IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 15, 2018, accepted August 31, 2018, date of publication September 17, 2018, date of current version October 8, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2869780

A Feasibility of Respiration Prediction Based on Deep Bi-LSTM for Real-Time Tumor Tracking

RAN WANG[1,2], XIAOKUN LIANG[1,2], XUANYU ZHU[1], AND YAOQIN XIE[1]

[1]Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China
[2]University of Chinese Academy of Sciences or Shenzhen College of Advanced Technology, University of Chinese Academy of Sciences, Shenzhen 518055, China

Corresponding author: Yaoqin Xie (yq.xie@siat.ac.cn)

This work was supported in part by the National Key Research and Development Program of China under Grant 2016YFC0105102, in part by the Leading Talent of Special Support Project in Guangdong under Grant Y77504, in part by the Shenzhen Key Technical Research Project under Grant JSGG20160229203812944, in part by the National Science Foundation of Guangdong under Grant 2014A030312006, in part by the Beijing Center for Mathematics and Information Interdisciplinary Sciences, and in part by the National Natural Science Foundation of China under Grant 81871433.

ABSTRACT In radiotherapy, the position of thoracic-abdominal tumor is changing due to respiratory motion. Real-time tracking of thoracic-abdominal tumors is of great significance in improving the treatment effect of radiotherapy. The accurate prediction of thoracic-abdominal tumor motion is required to compensate for system latency in image-guided adaptive radiotherapy systems. The purpose of this paper is to identify an optimal prediction model to improve the treatment effect of radiotherapy. A seven-layer bidirectional long short term memory (Deep Bi-LSTM) and one output layer deep neural network is proposed to predict respiration motion for a latency about 400 ms. 103 malignant lung tumor patients' respiratory motion data is used to train model. Mean absolute error (MAE), root mean square error (RMSE), and normalized mean square error are introduced to evaluate the performance of predictive results. Deep Bi-LSTM has great performance in the cases with relative long latency, average MAE of 0.074 mm, RMSE of 0.097 mm, and normalized root mean square error (nRMSE) of 0.081 with latency about 400 ms are obtained from predictive results of Deep Bi-LSTM. It demonstrates that the prediction accuracy of our proposed Deep Bi-LSTM is about five times better than traditional autoregressive integrated moving average model and about three times better than adaptive boosting and multi-layer perceptron neural network when the latency of 400 ms. The method can be applied to improve tracking accuracy and increase efficiency in thoracic-abdominal radiotherapy, which is practical and attractive for clinical application in the near future.

INDEX TERMS Radiotherapy, respiratory prediction, tumor tracking, bidirectional long short term memory.

I. INTRODUCTION

Compensation of respiratory motion can significantly improve the tumor treatment effect during radiation therapy. Radiotherapy aims to accurately irradiate the tumor target, while avoid damaging organ at risk (OAR) around the tumor. However, in the course of radiotherapy, the location of thoracic-abdominal organs and tissues are changing due to respiratory motion, which is likely to cause the actual tumor to exceed the target area, while OAR may also enter the target area of the plan, and suffer a certain dose of radiation, which will greatly influence the effect of radiotherapy. More seriously, it may cause complications for patients. Therefore, accurate determination of tumor position is of great importance to reduce the adverse factors of respiratory motion.

At present, many methods have been put forward to solve the problem of respiratory motion in radiotherapy, mainly including breath holding technology or abdominal compression technology [1], [2], respiratory gating technology [3] and real-time tracking technology [4]. The breath holding technique is active or passive control of the patient's breathing to reduce the tumor movement, which is simple and easy, but the tolerance is poor for some patients with limited radiotherapy precision. The respiratory gating technology is to synchronize the dose exposure in a specific phase of the respiratory cycle, which prolong the treatment time. Different from the two methods above, in real-time tracking technology, patients can breathe normally and the system can keep up with the tumor motion and adjust the radiation beam [5]–[7] or the position of treatment couch [8] to ensure relative position between radiation beam and tumor. Real-time tracking technology must overcome the system latency. Currently, there is about 75 to 100ms latency in radiotherapy

system, which including the latency of information acquisition and robotic system response [9]. Generally, the prediction algorithm will have to deal with a latency of at least 100 to 150 ms. So far, the most feasible and invasive way of real-time tracking is an indirect tumor tracking. The synchronous dynamic breathing tracking system has been used in clinic by CyberKnife system.

The Cyberknife is a robotic image guided system that delivers stereotactic body radiotherapy(SBRT), tracks tumors during respiration, and automatically adjusts treatment for any patient movement [10]. The CyberKnife is a frameless robotic radiosurgery system used for treating benign tumors, malignant tumors and other medical conditions [11]. In Cyberknife, the radiation source is mounted on a general purpose industrial robot and the system include an image guidance system [12]. The Cyberknife has been used to treat a broad range of tumors throughout the body, including prostate, lung, spine, liver, pancreas, kidney, and other tumors. Currently, there have been increasing numbers of successful reports of using SBRT against hepatocellular carcinoma(HCC) and other liver tumors [13].

Various prediction methods have been investigated for respiration prediction. Seregni *et al.* [14], Collobert *et al.* [15] and Sharp *et al.* [16] apply linear prediction model. The limitation of linear model is the poor robustness when the state change of linear system. Mnih and G. Hinton [17] made respiration prediction using Kalman filter, which can estimate the dynamic behavior in real-time. Khashei and Bijari [18], Babu and Reddy textitet al. [19] and Shirato *et al.* [20] use autoregressive integrated moving average(ARIMA) model for respiration prediction. ARIMA is one of the most commonly used models in time series analysis. It predicts the future time series through a linear combination of its past value, past error, other time series's current value and past value. Compared with other traditional time series analysis methods, the accuracy of ARIMA is relatively high. Azimi *et al*. [21] and others [22], [23] apply Artificial Neural network(ANN) to respiration prediction, ANN is expert in describing the nonlinear characteristics of various factors and is the most commonly used method in current respiratory prediction. Sunet *et al.* [24] proposed an adaptive boosting and multi-layer perceptron neural network(ADMLP-NN) to make respiration prediction. ADMLP-NN prediction performance is the results of multiple multi-layer perceptron(MLP) interactions. ADMLP-NN is also a kind of ANN method and achieve a relatively great performance in respiration prediction.

Although the methods mentioned above have been applied in prediction of respiration motion, the precision of prediction is not enough especially in the case with long latency. To improve the prediction accuracy of respiratory prediction, a deep bidirectional long short term memory(Deep Bi-LSTM) method is proposed in this paper. A 7-layer bidirectional LSTM and one output layer deep neural network is proposed to predict respiration motion for a latency about 400ms. 103 malignant lung tumor patients' respiratory motion data is used to train model. Mean absolute error(MAE), root mean square error(RMSE) and normalized mean square error(RMSE) are introduced to evaluate the performance of predictive results. Deep Bi-LSTM has great performance in the cases with relative long latency, average MAE of 0.074mm, RMSE of 0.097mm and nRMSE of 0.081 with latency about 400ms are obtained from predictive results of Deep Bi-LSTM. Besides, we compared Deep Bi-LSTM with other methods(ARIMA, ADMLP-NN) that achieve relatively great performance. Our method is about 5 times better than ARIMA model and about 3 times better than ADMPL-NN when the latency of 400ms.

II. MATERIALS AND METHODS

A. RESPIRATORY DATABASE

The database of respiratory motion used in this paper come from an open-access database during CyberKnife treatment at Georgetown University Hospital by courtesy of Dr. Kevin Cleary and Dr. Sonja Dieterich [9]. The database contains breathing recordings of 103 patients with the total of 306 respiratory motion traces. No private information is recorded. All patients had malignant tumor manifestations in the lung. Each patient's respiratory data are recorded with three fiducial markers, each of which is placed on the chest consecutively. The position of three fiducial marker are recorded using an optical tracking device with 26Hz sampling. The recording time for each trace is distributed from 25 minutes to 132 minutes.

B. WORKFLOW OF PREDICTION

Figure 1 shows the overall workflow of prediction algorithm of respiration motion. The raw respiratory motion signal s_i need to be preprocessed before training. Data preprocessing includes 4 steps as follows shown in Fig. 1: (*a*) data intercepting, (*b*) removal of abnormal values, (*c*) data smoothing, (*d*) data normalization. For the respiratory motion database, the amount of position points on respiration signals are relatively large for training our model, there are 306 traces in the database, and about 100 thousand position points on each trace. The data intercepting is to pick 154 traces(a patients at least one trace is picked) with better respiration pattern from the database. Take the data normalization into consideration, in order to allow the target data to be evenly distributed between 0 and 1, we need to remove abnormal data(individual maximum and minimum). For the data smoothing, Savitzky-Golay filter [25] is used for respiratory signals. Savitzky-Golay filter is widely used in data stream smoothing and denoising. It is a filtering method based on local polynomial least square fitting in time domain. The characteristics of the filter is to filter the noise and ensure the shape and width of the signal unchanged. The data normalization step is to normalize signals to [0,1], which can improve the speed of convergence and the accuracy of algorithms.

The black solid line box in Fig. 1 shows the Deep Bi-LSTM with 7-layer bidirectional LSTM and one output layer.

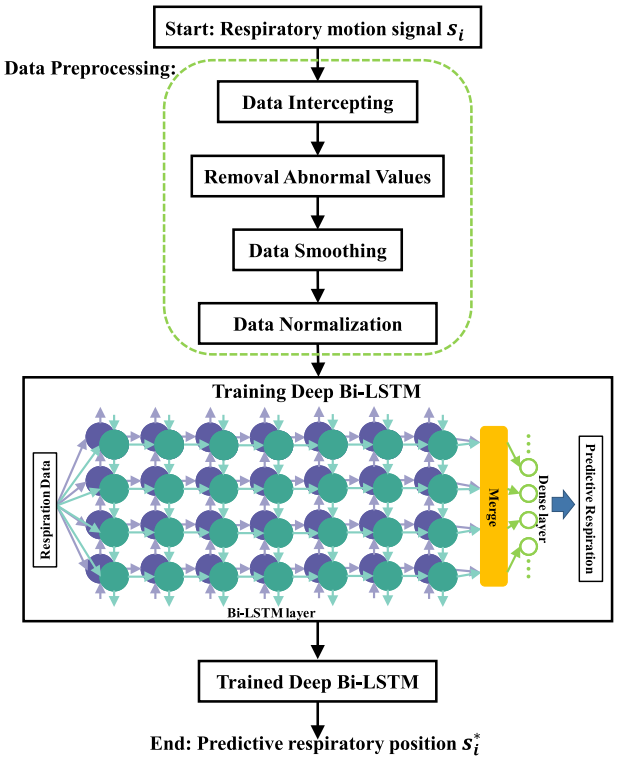


FIGURE 1. Workflow of respiration motion prediction. s_i is the original respiratory motion signal data and s_i^* is the predictive respiration motion signal. four steps of data preprocessing show in the green dashed box. The black solid line box indicating the structure of Deep Bi-LSTM.

The preprocessed data is used to train our model. Trained model would calculate predictive respiration curve s_i^*. In this paper, s_i^* is position of the 400ms after the current moment. Generally, the prediction algorithm will have to deal with a latency of at least 100 to 150 ms [9], latency of 400ms is fully able to deal with radiotherapy system latency.

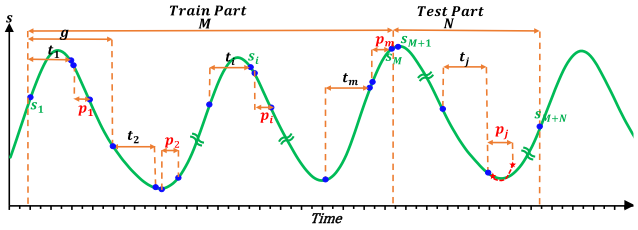


FIGURE 2. Respiration trace intercepted from the database. Every patient's respiration trace is split to training part(from s_1 to s_M) and testing part(from s_{M+1} to s_{M+N}). t_i and p_i is a section on the trace, and consist of a training sample for model's inputs and outputs, g represents the interval between two adjacent training samples, m and n is the number of training samples and testing samples on one respiration signal.

C. TRAINING DATA PROCESSING

The preprocessed data is divided into training part(from s_1 to s_M) and testing part(from s_{M+1} to s_{M+N}). In training part as shown in Fig. 2, hypothetically, s_i is the most recent observation at time i, a section on signals t_i is the input of model, t_i is the number of signal position points. The aim is to predict a section on the trace p_i which is also model's output based on t_i, p_i also corresponding to predictive latency, Owing to the frequency of respiration signal is about 26Hz, the latency of $p_i = 1, 5, 10$ is about 40ms, 200ms, 400ms, respectively. p_i and t_i consist of a training sample. The training deep Bi-LSTM model is based on training set X_{train} as Eq. 1 and label set Y_{train} as Eq. 2. g represents the interval between two adjacent training samples, M and N is the length of training part and test part on a respiration signal, m and n is the number of training samples and testing samples on a respiration signal. The training set and label set are:

$$X_{train} = [t_1, t_2, ..., t_m]^T \tag{1}$$
$$Y_{train} = [p_1, p_2, ..., p_m]^T \tag{2}$$

The testing set and label set are:

$$X_{test} = [t_{m+1}, t_{m+2}, ..., t_{m+n}]^T \tag{3}$$
$$Y_{test} = [p_{m+1}, p_{m+2}, ..., p_{m+n}]^T \tag{4}$$

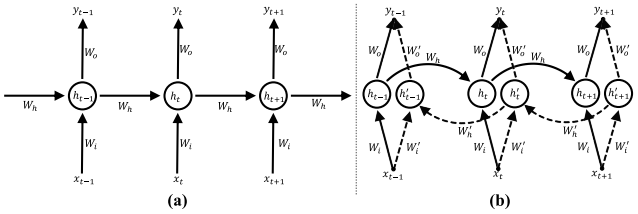


FIGURE 3. (a) Unfolding form of RNN; (b) unfolding form of Bi-RNN.

D. RESPIRATION PREDICTION ALGORITHM

Recurrent neural network(RNN) is specifically suitable for processing time series data and has been widely applied in many fields [17], [26], [27]. In traditional neural network models, two adjacent layers are fully connected and the nodes between each layer are connectionless, which would get a bad result when handle sequence data, but RNN is expert in dealing with such situations, the nodes between the same hidden layer are connected, and the input of hidden layer includes not only the output of input layer, but also the output of hidden layer at the last time showing as Fig. 3(a). When the network receives input at time t, the value of hidden layer is h_t, and the output value is y_t. The key point is that the value of h_t depends not only on the x_t, but also on the h_{t-1}. RNN is calculated as:

$$y_t = f_1(W_o h_t) \tag{5}$$
$$h_t = f_2(W_i x_t + W_h h_{t-1}) \tag{6}$$

Equation 5 calculates the output of each layer, and Eq. 6 calculates the result of hidden layer. W_i and W_o is the weight matrix of input and output layer, respectively. W_h is the weight matrix between adjacent RNN cells. W_h is the major difference between recurrent layer and fully connected layer. Back-propagation through time(BPTT) [28] is applied to train RNN.

In practice, RNN does not deal with long sequences well. A main reason is that RNN is prone to generate

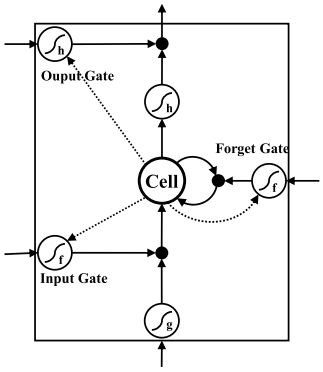


FIGURE 4. LSTM memory block with one cell.

gradient disappearance in training, which leads to gradient can't be transmitted in long sequences. To tackle this issue, LSTM [29], [30] is proposed. In LSTM network, the summation units in the hidden layer of standard RNN are replaced by memory blocks. The LSTM architecture consists of a set of recurrently connected subnets, known as memory blocks [31] shown in Fig. 4. Each block contains one or more self-connected memory cells and three multiplicative units (the input, output and forget gates) that provide continuous analogues of write, read and reset operations for the cells. The three gates are nonlinear summation units that collect activations from inside and outside the block, and control the activation of the cell via multiplications (small black circles). The multiplicative gates allow LSTM memory cells to store and access information over long periods of time, thereby mitigating the vanishing gradient problem.

Bidirectional RNN(Bi-RNN) is an upgraded version of RNN [32], Bi-RNN is made up of two RNN superimposition as shown in Fig. 3(b). Bi-RNN need to train two RNN network that input are forward and backward of input sequences,respectively, and two RNN connect to a same output layer, which means that the output at every moment time knows the complete information of the entire input sequence in Bi-RNN, not just the information before the current time. As in Fig. 3(b), the calculation of a hidden layer's output y_t depends not only on $\{x_1, ..., x_{t-1}\}$, but also on $\{x_{t+1}, ...\}$.

The hidden layer of Bi-RNN should save two values, h_t participates in forward calculation and the h'_t is involved in reverse calculation. The final output value y_t depends on the h_t and h'_t, calculated as:

$$y_t = f_1(W_o h_t + W_o h'_t) \tag{7}$$

h_t calculated as Eq. 6, h'_t calculated as:

$$h'_t = f_2(W'_i x_t + W'_h h_{t+1}) \tag{8}$$

The forward and backward RNN are separate in calculation, the weights are not shared, i.e., W_i and W'_i, W_o and W'_o, W_h and W'_h are different weight matrices, respectively. In forward calculation, the value of hidden layer h_t is related to h_{t-1}, while the value of hidden layer h'_t is related to h'_{t+1}. The final output depends on the sum of forward and backward computation as follows:

$$y_t = f_1(W_o h_t + W'_o h'_t) \tag{9}$$

E. EVALUATION

In order to evaluate the performance of proposed algorithm and compare with other methods, several evaluation metrics are introduced. The respiration prediction algorithm can use the following metrics to evaluate.

We define the error as e_i:

$$e_i = y_i - y_i^* \tag{10}$$

where y_i is the actual respiratory motion trace, y_i^* is predictive respiratory motion trace. Mean absolute error(MAE) is a measure of difference between two continuous variables. MAE is defined as [33]:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|e_i| \tag{11}$$

Where N is the number of investigated points. As the name suggests, the mean absolute error is an average of the absolute errors $|e_i|$. The mean absolute error uses the same scale as the data being measured. This is known as a scale-dependent accuracy measure and therefore cannot be used to make comparisons between series using different scales [34]. The mean absolute error is a common measure of forecast error in time series analysis [35]–[37].

Root mean square error (RMSE) is a frequently used measure of the differences between values (sample or population values) predicted by a model or an estimator and the values observed. The RMSE represents the sample standard deviation of the differences between predicted values and observed values. RMSE is a measure of accuracy, to compare forecasting errors of different models for a particular dataset and not between datasets, as it is scale-dependent [35]. RMSE is the square root of the average of squared errors. The effect of each error on RMSE is proportional to the size of the squared error; thus larger errors have a disproportionately large effect on RMSE. Consequently, RMSE is sensitive to outliers [38], [39]. RMSE is the most commonly used accuracy measure [40]–[42] and is defined as:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(e_i)^2} \tag{12}$$

Normalizing the RMSE facilitates the comparison between datasets or models with different scales. Though there is no consistent means of normalization in the literature, common choices are the mean or the range (defined as the maximum value minus the minimum value) of the measured data [43]. This value is commonly referred to as the normalized root-mean-square error(nRMSE), and often expressed as a percentage, where lower values indicate less residual variance. In many cases, especially for smaller samples, the sample

range is likely to be affected by the size of sample which would hamper comparisons. nRMSE is defined as:

$$nRMSE = \frac{RMSE}{\sigma} = \frac{RMSE}{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \bar{y})^2}} = \sqrt{\frac{\sum_{i=1}^{N}(e_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2}} \quad (13)$$

Where the $\bar{y}$ is the mean of the actual data points.

All the calculation of evaluation metrics are based on the following network parameter settings: 7 bidirectional LSTM, 100 epochs, 512 batch sizes, $T = 50$, $G = 50$, op =Adam, $lr = 0.001$ and predicting about 400 ms forward. The hardware experimental platform is a workstation (CPU E5-2637 v2 @ 3.50GHz×8, NVIDIA GeForce GTX 1080 8G, 32G RAM).

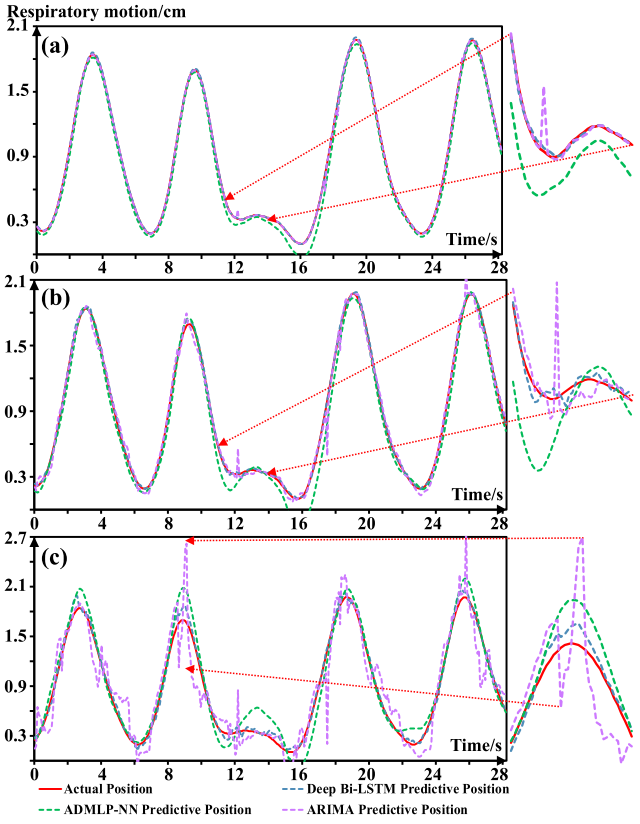


FIGURE 5. Actual respiration trace and Deep Bi-LSTM, ADMLP-NN and ARIMA model's predictive trace. Subgraph (*a*) – (*c*) latency is 40ms, 200ms, 400ms ($p_i = 1, 5, 10$), respectively.

III. RESULTS

Figure 5 illustrates the three methods(Deep Bi-LSTM, ADMLP-NN and ARIMA) actual performance when latency is 40ms, 200ms and 400ms ($p_i = 1$, $p_i = 5$ and $p_i = 10$), respectively. From the figure, we can learn that Deep Bi-LSTM remain accurate prediction performance even the latency of 400ms. ADMLP-NN and ARIMA maintain a sub-optimal performance when the latency of less 200ms.

TABLE 1. Performance of predictive algrithm based on Deep Bi-LSTM.

Latency(ms)	ave.MAE(mm)	ave.RMSE(mm)	ave.nRMSE(None)
40 ($p_i = 1$)	0.015904	0.019124	0.015981
200 ($p_i = 5$)	0.027445	0.034275	0.028606
400 ($p_i = 10$)	0.074523	0.096926	0.080646

The performance of ARIMA is poor when the latency of 400ms. Table 1 shows the final performance of Deep Bi-LSTM with parameters optimized. The results is calculated based on parameters: epochs(model training times) = 100, batch size = 1024, $t_i = 50$, op =Adam and $lr = 0.01$.

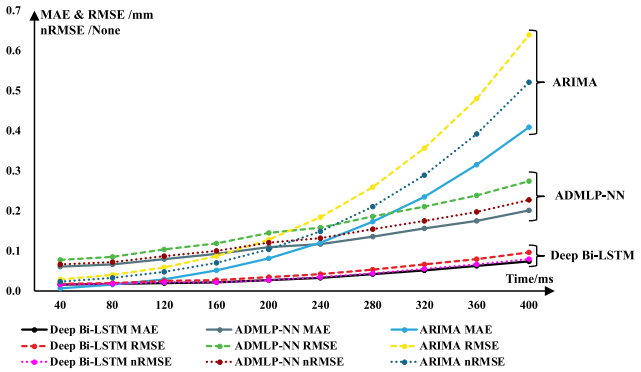


FIGURE 6. Performance of Deep Bi-LSTM, ADMLP-NN and ARIMA using different metrics, evaluation metrics are MAE, RMSE and nRMSE, respectively.

Figure 6 shows the comparison of the performance among three prediction methods. The performance is evaluated by MAE, RMSE and nRMSE, respectively. Obviously, the proposed Deep Bi-LSTM outperforms better than other two methods and the longer the latency, the better the performance of the proposed method as compared to ADMLP-NN and ARIMA.

When latency is 400ms ($p_i = 10$) in Fig. 6, The MAE, RMSE and nRMSE of Bi-LSTM are reduced by 62.9%, 64.6% and 64.6%(from 0.201mm to 0.0745mm, from 0.274mm to 0.0969mm and from 0.228 to 0.0806) as compared with ADMLP-NN, and 81.8%, 84.8% and 84.5%(from 0.409mm to 0.0745mm, from 0.639mm to 0.0969mm and from 0.521 to 0.0806) as compared with ARIMA. ADMLP-NN get a bad performance when the latency is less than 160ms compared with Deep Bi-LSTM and ARIMA. ARIMA can achieve the same performance as Deep Bi-LSTM when latency is less than 80 ms. The performance of ARIMA is very poor when latency is more than 300ms, which also shows in Fig. 5(c).

From the Fig. 5 and Fig. 6 we can demonstrate that Deep Bi-LSTM always has great performance in the cases with relative long latency. On the contrary, ARIMA is not suitable for long latency situation. It shows that the accuracy of proposed Deep Bi-LSTM method is about 5 times better than ARIMA and about 3 times better than ADMLP-NN when the latency is 400ms.

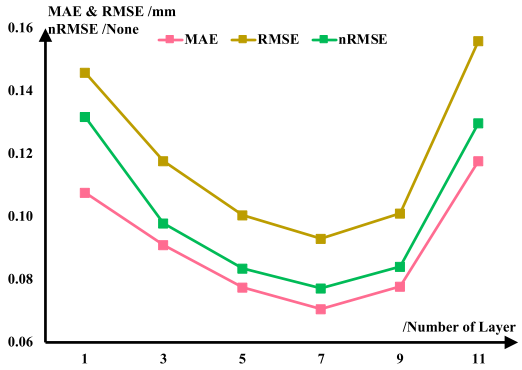


FIGURE 7. Performance of Deep Bi-LSTM correlation to number of network layers.

IV. DISCUSSION

The optimization of layer number is investigated. Fig. 7 shows the correlation between the performance of Deep Bi-LSTM and the number of network layers. Obviously, Deep Bi-LSTM with 7 layers would achieve better performance. It is demonstrated that the network is not as deep as better, deeper network may achieve poorer performance.

Parameter optimization is an important issue for the prediction accuracy. Table 2 shows three major parameters(t_i, op, lr) for our Deep Bi-LSTM.

TABLE 2. Influence of different parameters(t_i, op and lr) on Deep Bi-LSTM. The latency of folloeing results is 400ms.

Parameters	ave.MAE(mm)	ave.RMSE(mm)	ave.nRMSE(None)
t_i=30	0.30437	0.39311	0.32478
t_i=40	0.25240	0.33780	0.27628
t_i=50	0.14476	0.19320	0.15888
t_i=60	0.17593	0.24222	0.20380
op=SGD	2.63816	3.39950	1.33255
op=RMSprop	0.45057	0.56589	0.22182
op=Adam	0.14476	0.19320	0.15888
lr=0.01	0.20796	0.27373	0.10730
lr=0.001	0.14476	0.19320	0.15888
lr=0.0001	0.88687	1.10974	0.43500

A section on respiration signal t_i is model's input. The length of t_i represents the amount of information to make prediction, different length of t_i may cause different prediction results. op is the optimizer used in Deep Bi-LSTM. A good optimizer can speed up the training model process and achieve better performance. SGD is the most common optimizer, it can be said that there is no acceleration effect; RMSprop is an upgraded version of SGD and Adam is also an upgraded version of RMSprop. But it does not mean that the more advanced optimizer, the better the result. In order to choose a better optimizer for Deep Bi-LSTM, We compare these optimizers in Table 2. lr is learning rate of op. The optimizer will be affected by learning rate. If the learning rate a is too small, the training time will be very long; on the contrary, it may cross the local minimum and result in no convergence. These learning rates can usually be considered: 0.1, 0.01, 0.001, 0.0001, the different performance of different lr shown in Table 2.

All the parameters are based on 7 bidirectional LSTM layers, 30 epochs, 1024 batch sizes and latency is about 400ms($p_i = 10$). When one of the parameters is adjusted, other parameters remain the default values($t_i = 50$, $op =$ Adam and $lr = 0.01$). In addition, the results of each parameter would be calculated 10 times, the average results are filled in the Table 2. From Table 2, it is clear that $t_i = 50$ would be better with smallest MAE, RMSE and nRMSE. Optimizer op gets Adam is better, Adam optimizer has great performance for respiration prediction. Learning rate choose $lr = 0.01$.

From Table 2, Our final results is based on optimal parameters in Table 2 and 100 epochs, 1024 batch sizes. we finally use the parameters($t_i = 50$, $op =$Adam, $lr = 0.01$) to train our Deep Bi-LSTM.

V. CONCLUSION

Based on the current published literature, this paper proposed a Deep Bi-LSTM model intended for respiration motion prediction during respiratory. Deep Bi-LSTM has great performance in the cases with relative long latency, average MAE of 0.074mm, RMSE of 0.097mm and nRMSE of 0.081 with latency about 400ms are obtained from predictive results of Deep Bi-LSTM. It demonstrates that the prediction accuracy of our proposed Deep Bi-LSTM is about 5 times better than traditional ARIMA model and about 3 times better than ADMPL-NN when the latency of 400ms. Deep Bi-LSTM could potentially be used for prediction of respiration trajectory during treatment delivery, which is practical and attractive for clinical application in the near future.

REFERENCES

[1] J. W. Wong *et al.*, "The use of active breathing control (ABC) to reduce margin for breathing motion," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 44, no. 4, pp. 911–919, 1999.

[2] K. K. Herfarth *et al.*, "Extracranial stereotactic radiation therapy: Set-up accuracy of patients treated for liver metastases," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 46, pp. 329–335, Jan. 2000.

[3] H.-M. Lu *et al.*, "A respiratory-gated treatment system for proton therapy," *Med. Phys.*, vol. 34, no. 8, pp. 3273–3278, 2007.

[4] M. J. Murphy, "Tracking moving organs in real time," *Seminars Radiat. Oncol.*, vol. 14, pp. 91–100, Jan. 2004.

[5] W. Mao *et al.*, "Image-guided radiotherapy in near real time with intensity-modulated radiotherapy megavoltage treatment beam imaging," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 75, no. 2, pp. 603–610, 2009.

[6] Y. Ge, R. T. O'Brien, C.-C. Shieh, J. T. Booth, and P. J. Keall, "Toward the development of intrafraction tumor deformation tracking using a dynamic multi-leaf collimator," *Med. Phys.*, vol. 41, no. 6, p. 061703, 2014.

[7] T. Neicu, H. Shirato, Y. Seppenwoolde, and S. B. Jiang, "Synchronized moving aperture radiation therapy (SMART): Average tumour trajectory for lung patients," *Phys. Med. Biol.*, vol. 48, no. 5, pp. 587–598, 2003.

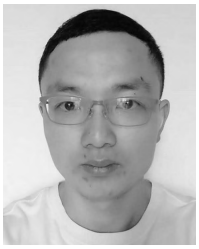
[8] I. Buzurovic, K. Huang, Y. Yu, and T. K. Podder, "A robotic approach to 4D real-time tumor tracking for radiotherapy," *Phys. Med. Biol.*, vol. 56, no. 5, pp. 1299–1318, 2011.

[9] F. Ernst, *Compensating for Quasi-Periodic Motion in Robotic Radiosurgery*. New York, NY, USA: Springer, 2012.

[10] H. Kato *et al.*, "Cyberknife treatment for advanced or terminal stage hepatocellular carcinoma," *World J. Gastroenterol.*, vol. 21, no. 46, p. 13101, 2015.

[11] È. Coste-Manière, D. Olender, W. Kilby, and R. A. Schulz, "Robotic whole body stereotactic radiosurgery: Clinical advantages of the CyberKnife integrated system," *Int. J. Med. Robot. Comput. Assist. Surg.*, vol. 1, no. 2, pp. 28–39, 2005.

[12] F. Colombo, L. Casentini, C. Cavedon, P. Scalchi, S. Cora, and P. Francescon, "Cyberknife radiosurgery for benign meningiomas: Short-term results in 199 patients," *Neurosurger*, vol. 64, no. 2, pp. A7–A13, 2009.
[13] R. A. Ibarra *et al.*, "Multicenter results of stereotactic body radiotherapy (SBRT) for non-resectable primary liver tumors," *Acta Oncol.*, vol. 51, no. 5, pp. 575–583, 2012.
[14] M. Seregni, A. Pella, M. Riboldi, R. Orecchia, P. Cerveri, and G. Baroni, "Real-time tumor tracking with an artificial neural networks-based method: A feasibility study," *Phys. Medica*, vol. 29, no. 1, pp. 48–59, 2013.
[15] R. Collobert, J. Weston, L. Bottou, M. Karlen, K. Kavukcuoglu, and P. Kuksa, "Natural language processing (almost) from scratch," *J. Mach. Learn. Res.*, vol. 12, pp. 2493–2537, Aug. 2011.
[16] G. C. Sharp, S. B. Jiang, S. Shimizu, and H. Shirato, "Prediction of respiratory tumour motion for real-time image-guided radiotherapy," *Phys. Med. Biol.*, vol. 49, no. 3, pp. 425–440, 2004.
[17] A. Mnih and G. Hinton, "Three new graphical models for statistical language modelling," in *Proc. Int. Conf. Mach. Learn.*, 2007, pp. 641–648.
[18] M. Khashei and M. Bijari, "A novel hybridization of artificial neural networks and ARIMA models for time series forecasting," *Appl. Soft Comput.*, vol. 11, no. 2, pp. 2664–2675, Mar. 2011.
[19] C. N. Babu and B. E. Reddy, "A moving-average filter based hybrid ARIMA–ANN model for forecasting time series data," *Appl. Soft Comput.*, vol. 23, pp. 27–38, Oct. 2014.
[20] H. Shirato *et al.*, "Physical aspects of a real-time tumor-tracking system for gated radiotherapy," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 48, no. 4, pp. 1187–1195, 2000.
[21] P. Azimi, H. R. Mohammadi, E. C. Benzel, S. Shahzadi, and S. Azhari, "Use of artificial neural networks to predict recurrent lumbar disk herniation," *Clin. Spine Surg.*, vol. 28, pp. E161–E165, Apr. 2015.
[22] J. H. Goodband, O. C. L. Haas, and J. A. Mills, "A comparison of neural network approaches for on-line prediction in IGRT," *Med. Phys.*, vol. 35, no. 3, pp. 1113–1122, 2008.
[23] J. Rottmann and R. Berbeco, "Using an external surrogate for predictor model training in real-time motion management of lung tumors," *Med. Phys.*, vol. 41, no. 12, p. 121706, 2014.
[24] W. Sun, F. F. Yin, M. Jiang, J. Dang, and T. You, "Respiratory signal prediction based on adaptive boosting and multilayer perceptron neural network," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 96, no. 2, p. E702, 2017.
[25] I. Zawisza, H. Yan, and F. Yin, "SU-C-BRF-07: A pattern fusion algorithm for multi-step ahead prediction of surrogate motion," *Med. Phys.*, vol. 41, no. 6, p. 98, 2014.
[26] D. Sussillo *et al.*, "A recurrent neural network for closed-loop intracortical brain–machine interface decoders," *J. Neural Eng.*, vol. 9, no. 2, p. 026027, 2012.
[27] T. Mikolov, M. Karafiát, L. Burget, J. Cernocký, and S. Khudanpur, "Recurrent neural network based language model," in *Proc. 11th Annu. Conf.*, 2010, pp. 1045–1048.
[28] M. Mozer, "A focused back-propagation algorithm for temporal pattern recognition," *Complex Syst.*, vol. 3, no. 4, pp. 349–381, 1989.
[29] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
[30] R. Jozefowicz, W. Zaremba, and I. Sutskever, "An empirical exploration of recurrent network architectures," in *Proc. Int. Conf. Mach. Learn.*, 2015, pp. 2342–2350.
[31] A. Graves, *Supervised Sequence Labelling With Recurrent Neural Networks*. Berlin, Germany: Springer, 2012, p. 385.
[32] A. Graves and J. Schmidhuber, "Framewise phoneme classification with bidirectional LSTM and other neural network architectures," *Neural Netw.*, vol. 18, pp. 602–610, Jul. 2005.
[33] C. J. Willmott and K. Matsuura, "Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance," *Climate Res.*, vol. 30, no. 1, pp. 79–82, 2005.
[34] Y. Shi, "Evaluating forecast accuracy," LAP LAMBERT Academic, Tech. Rep., 2017.
[35] R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy," *Int. J. Forecast.*, vol. 22, no. 4, pp. 679–688, 2006.
[36] Y. Sheng, "Fuzzy and hybrid prediction of position signal in synchrony respiratory tracking system," in *Proc. SIP*, 2007, pp. 459–464.
[37] I. Buzurovic, T. K. Podder, K. Huang, and Y. Yu, "Tumor motion prediction and tracking in adaptive radiotherapy," in *Proc. IEEE Int. Conf. Bioinf. Bioeng. (BIBE)*, May 2010, pp. 273–278.
[38] R. G. Pontius, O. Thontteh, and H. Chen, "Components of information for multiple resolution comparison between maps that share a real variable," *Environ. Ecol. Statist.*, vol. 15, no. 2, pp. 111–142, 2008.
[39] C. J. Willmott and K. Matsuura, "On the use of dimensioned measures of error to evaluate the performance of spatial interpolators," *Int. J. Geograph. Inf. Sci.*, vol. 20, no. 1, pp. 89–102, 2006.
[40] S. Choi, Y. Chang, N. Kim, S. H. Park, S. Y. Song, and H. S. Kang, "Performance enhancement of respiratory tumor motion prediction using adaptive support vector regression: Comparison with adaptive neural network method," *Int. J. Imag. Syst. Technol.*, vol. 24, pp. 8–15, Mar. 2014.
[41] K. Huang, I. Buzurovic, Y. Yu, and T. K. Podder, "A comparative study of a novel AE-nLMS filter and two traditional filters in predicting respiration induced motion of the tumor," in *Proc. IEEE Int. Conf. Bioinf. Bioeng. (BIBE)*, May 2010, pp. 281–282.
[42] D. Putra, O. C. L. Haas, J. A. Mills, and K. J. Burnham, "A multiple model approach to respiratory motion prediction for real-time IGRT," *Phys. Med. Biol.*, vol. 53, p. 1651, Feb. 2008.
[43] M. J. Murphy, M. Isaakson, and J. Jalden, "Adaptive filtering to predict lung tumor motion during free breathing," in *Computer Assisted Radiology and Surgery*. Berlin, Germany: Springer, 2002, pp. 539–544.



RAN WANG received the B.S. degree from Jiangnan University, Jiangsu, China, in 2016. He is currently pursuing the M.S. degree with the Shenzhen College of Advanced Technology, University of Chinese Academy of Sciences, Shenzhen, China. His research interests focus on medical image analysis, medical physics, and image guided radiotherapy.

XIAOKUN LIANG received the B.S. degree from Southern Medical University, Guangzhou, China, in 2013, and the M.S. degree from Guangdong Medical University, Dongguan, China, in 2016. He is currently pursuing the Ph.D. degree with the Shenzhen College of Advanced Technology, University of Chinese Academy of Sciences, Shenzhen, China. His research interests focus on medical image analysis, medical physics, and image guided radiotherapy.

XUANYU ZHU is currently pursuing the bachelor's degree from Northeastern University, Shenyang, China. He is currently participating in the Excellent Engineer Joint Cultivation Program with the Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China. His research interests include medical material analysis, medical image physics, and image guided radiation therapy.

YAOQIN XIE received the B.S., M.S., and Ph.D. degrees from Tsinghua University, Beijing, China, in 1995, 1998, and 2002, respectively. From 2002 to 2010, he was a Lecturer with the School of Physics, Peiking University, China. From 2006 to 2008, he was a Research Fellow with the Department of Radiation Oncology, Stanford Medical School, USA. Since 2010, he has been a Full Professor with the Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China. His research interests include medical image analysis, medical physics, and image-guided radiotherapy.