



Title	Application of motion prediction based on a long short-term memory network for imaging dose reduction in real-time tumor-tracking radiation therapy
Author(s)	Numakura, Kazuki; Takao, Seishin; Matsuura, Taeko et al.
Citation	Physica medica : European journal of medical physics, 125, 104507 https://doi.org/10.1016/j.ejmp.2024.104507
Issue Date	2024-09
Doc URL	https://hdl.handle.net/2115/97259
Rights	© <2024>. This manuscript version is made available under the CC-BY-NC-ND 4.0 license https://creativecommons.org/licenses/by-nc-nd/4.0/
Rights(URL)	https://creativecommons.org/licenses/by-nc-nd/4.0/
Type	journal article
File Information	LSTMpaper_R1.pdf



Title

Application of motion prediction based on a long short-term memory network for imaging dose reduction in real-time tumor-tracking radiation therapy

Authors

Kazuki Numakura^a, Seishin Takao^{b, c}, Taeko Matsuura^{b, c}, Kouhei Yokokawa^c, Ye Chen^{b, c}, Yusuke Uchinami^d, Hiroshi Taguchi^e, Norio Katoh^d, Hidefumi Aoyama^d, Satoshi Tomioka^b, Naoki Miyamoto^{b, c*}

^a Graduate School of Biomedical Science and Engineering, Hokkaido University, North 13, West 8, Kita-ku, Sapporo, Hokkaido, 060–8628, Japan

^b Faculty of Engineering, Hokkaido University, North 13, West 8, Kita-ku, Sapporo, Hokkaido, 060-8638, Japan

^c Department of Medical Physics, Hokkaido University Hospital, North 14, West 5, Kita-ku, Sapporo, Hokkaido, 060–8648, Japan

^d Faculty of Medicine, Hokkaido University, North 15, West 7, Kita-ku, Sapporo, Hokkaido, 060-8638, Japan

^e Department of Radiation Oncology, Hokkaido University Hospital, North 14, West 5, Kita-ku, Sapporo, Hokkaido 060–8648, Japan

*Corresponding author

Faculty of Engineering, Hokkaido University, North 13, West 8, Kita-ku, Sapporo, Hokkaido, 060-8638, Japan. Tel.: +81-11-706-6673

E-mail address: miya-nao@eng.hokudai.ac.jp (Naoki Miyamoto)

Abstract

Purpose

To demonstrate the possibility of using a lower imaging rate while maintaining acceptable accuracy by applying motion prediction to minimize the imaging dose in real-time image-guided radiation therapy.

Methods

Time-series of three-dimensional internal marker positions obtained from 98 patients in liver stereotactic body radiation therapy were used to train and test the long-short-term memory (LSTM) network. For real-time imaging, the root mean squared error (RMSE) of the prediction on three-dimensional marker position made by LSTM, the residual motion of the target under respiratory-gated irradiation, and irradiation efficiency were evaluated. In the evaluation of the residual motion, the system-specific latency was assumed to be 100 ms.

Results

Except for outliers in the superior–inferior (SI) direction, the median/maximum values of the RMSE for imaging rates of 7.5, 5.0, and 2.5 frames per second (fps) were 0.8/1.3, 0.9/1.6, and 1.2/2.4 mm, respectively. The median/maximum residual motion in the SI direction at an imaging rate of 15.0 fps without prediction of the marker position, which is a typical clinical setting, was 2.3/3.6 mm. For rates of 7.5, 5.0, and 2.5 fps with prediction, the corresponding values were 2.0/2.6, 2.2/3.3, and 2.4/3.9 mm, respectively. There was no significant difference between the irradiation efficiency with and that without prediction of the marker position. The geometrical accuracy at lower frame rates with prediction applied was superior or comparable to that at 15 fps without prediction. In comparison with the current clinical setting for real-time image-guided radiation therapy, which uses an imaging rate of 15.0 fps without prediction, it may be possible to reduce the imaging dose by half or more.

Conclusions

Motion prediction can effectively lower the frame rate and minimize the imaging dose in real-time image-guided radiation therapy.

Keywords: motion prediction, fiducial marker, real-time tumor-tracking radiation therapy, image-guided radiation therapy

1. Introduction

In stereotactic body radiation therapy (SBRT), the use of motion management techniques is important to ensure high dosimetric and geometric accuracy in the treatment of mobile tumors, such as those in the lung, liver, and pancreas. Real-time tumor-tracking radiation therapy (RTTRT) is one effective motion management technique [1]. RTTRT has been clinically applied in photon therapy [2-4] and in proton beam therapy [5-7]. In RTTRT, the three-dimensional positions of internal markers inserted near the tumor are determined in real time by orthogonal kilovoltage (kV) X-ray imaging using an acquisition rate of up to 30 frames per second (fps). The treatment beam is on only when the marker is located within an allowable three-dimensional region, which is defined as the gating window. In our institution, the normal frame rate used for X-ray imaging in RTTRT is 15.0 fps. In general, the target localization accuracy achieved using a combination of an internal surrogate and X-ray imaging is better than that obtained using an external surrogate [8]; however, additional dose due to X-ray imaging is accompanied. According to the AAPM Task Group 180 [9], the accumulated imaging dose in image-guided radiation therapy (IGRT) should be less than 5% of the dose prescribed for the target. The cumulative skin dose caused by X-ray imaging in typical RTTRT for liver cancer is less than 1 Gy [4]. However, it is clinically important to minimize the imaging dose as additional imaging techniques, such as cone beam computed tomography, may be used to verify the patient position. One way of reducing the imaging dose in real-time imaging is to lower the X-ray imaging frame rate. However, because the ratio of false beam-on time increases, the residual motion of the target during gated irradiation increases.

The delay due to the imaging interval and system latency can be compensated for by controlling the gate signal in advance. This can be done by predicting the positions of the markers one or a few frames ahead; in this way, the irradiation accuracy can be maintained even at lower frame rates. In the tracking irradiation using external surrogates, a method for predicting respiration signals to compensate for the system latency has been proposed [10-13]. In the case of respiratory-gated irradiation in MR-linac, there have been studies on the use of motion prediction to compensate for the system latency [14-16]. However, most previous studies have aimed to improve the accuracy of the tracking irradiation or gated irradiation at fixed frame rates, and few studies have focused on the application of prediction to minimize the imaging dose by using real-time tracking of the three-dimensional positions of internal surrogates.

In this study, we evaluated the accuracy of predictions of the three-dimensional position of an internal marker using a long short-term memory (LSTM) network [17]. The residual motion of the target under gated irradiation and the irradiation efficiency at different imaging frame rates were also evaluated. The aim was to demonstrate the possibility of using a lower imaging rate while maintaining acceptable irradiation accuracy by applying motion prediction, thus minimizing the imaging dose. The results demonstrate the potential for reducing the imaging dose compared with the current imaging dose used in RTTRT.

2. Materials and methods

2.1. Beam gating with prediction of the marker position

Figure 1 shows charts in which the actual target position, predicted target position, and beam-on/off signal are plotted against time. In this study, the latency $T_{\text{gate-on}}$ was defined as the time duration from the trigger of X-ray imaging until generation of beam gating signal. The latency $T_{\text{beam-on}}$ was defined as the time duration from beam gating signal to actual beam on. $T_{\text{total-on}}$ is sum of $T_{\text{gate-on}}$ and $T_{\text{beam-on}}$. In this study, the latencies are fixed values that are specific to the treatment system. Temporal factors such as the imaging interval and system-specific latency can affect the accuracy of the irradiation in respiratory-gated radiation therapy. Due to these temporal uncertainties, there will be a discrepancy between the ideal beam-on/off time and the actual beam-on/off time, as shown in Figure 1(a). In the case of treatment without motion prediction, as the imaging frame rate decreases, the imaging interval increases, and the ratio of false beam-on will increase. As shown in Figure 1(b), predictions of the marker position at T_{pred} ahead can be used to control the gate signal in advance to compensate for the latency so that the ideal beam-on/off state is approached. Here, T_{pred} is corresponding to prediction time window and is increment of the imaging interval T_{image} . In this study, the crossing time was considered to correspond to the intersection of the gate window with the line connecting the current and predicted marker positions, and the gate signal was controlled in advance to take account of the latency time. In this way, the beam-on/off timings could be made closer to those of the ideal situation.

In this study, the prediction performance was evaluated at imaging rates of 7.5, 5.0, and 2.5 fps. $T_{\text{gate-on}}$ and $T_{\text{beam-on}}$ were assumed to be 40 and 100 ms, which are equivalent to the latency of current working RTRT systems [18], respectively. The same values of latencies were assumed for beam-off. Because the crossing time is evaluated by simple linear interpolation between the current and predicted marker positions, timing discrepancy of beam gating will be increased for long T_{pred} even if the prediction accuracy is high. On the other hand, when T_{pred} is shorter than $T_{\text{total-on}}$ or $T_{\text{total-off}}$, the gate control based on prediction would not be in time. Therefore, T_{pred} for the frame rate of 7.5 fps was set to 266.6 ms (2 frames ahead). The position at 1 frame ahead was predicted for 5.0 and 2.5 fps datasets.

2.2. Datasets used for evaluation

Time-series of three-dimensional marker positions obtained during RTRT for liver SBRT were used to train and test an LSTM network. The RTRT system consisted of a combination of a TrueBeam (Varian Medical Systems, USA) and a SyncTraX real-time tumor-tracking system (Shimadzu, Japan) [4]. Respiratory-gated irradiation using the internal markers was conducted in free-breathing. The characteristics of the marker trajectory data are summarized in Table 1. The three-dimensional marker positions were recorded at imaging rates of 15.0 or 20.0 fps. If the test datasets include the data with small motion for which no motion management is needed, the prediction accuracy might be underestimate. Therefore, the test data were limited to those data for which the motion in the superior-inferior (SI) direction was greater than 4 mm.

Table 1: Characteristics of the training and test datasets of the internal marker motion. Values of the mean $\pm$ standard deviation are given together with the minimum and maximum values in brackets.

		Training	Test
Number of patients		68	30
Number of datasets		3064	1197
Duration of datasets [s]		60.8 ± 32.3 (30.0–339.1)	76.2 ± 62.1 (30.0–592.7)
Motion range [mm]	LR	2.2 ± 1.5 (0.4–20.9)	2.3 ± 1.5 (0.4–10.8)
	SI	12.6 ± 4.9 (1.1–33.8)	14.0 ± 4.4 (4.0–31.0)
	AP	4.7 ± 2.4 (0.4–16.9)	5.6 ± 2.7 (0.5–14.5)

2.3. Training the LSTM network

In this study, MATLAB (Mathworks, USA) was used for data processing and for training the LSTM network. The trajectory data obtained at 20.0 fps were downsampled to 15.0 fps to give a uniform sampling rate. The prediction performance was evaluated at imaging rates of 7.5, 5.0, and 2.5 fps. For the lower frame rates, the trajectory data were produced by reducing the number of sampling points in the 15.0-fps data. The number of channels in input layer and output layer were three which was corresponding to the three-dimensional coordinates of the internal marker. The optimum values of the hyperparameters, including the learning rate, batch size, number of LSTM layers, and number of units per LSTM layer, were determined as the combination that gave the lowest root mean squared error (RMSE) for the validation data (15% of the training data) in the training process with grid search. In this study, the parameters were optimized using the 5.0-fps data as representative. The optimum values were as follows: learning rate: 0.001, batch size: 250, number of units: 200, and number of LSTM layers: 4. The optimizer was fixed to the Adaptive moment estimation algorithm (Adam). In a previous study in which the hyperparameters used in LSTM were investigated [12], it was reported that the parameters that have a large impact on the prediction performance are the number of LSTM layers, learning rate, and optimizer. Our parameters were equivalent to the optimal parameters of the previous study. For all the frame rates that were evaluated in this study, the LSTM network was trained using these same hyperparameter values. For each frame rate, the metrics discussed in the next section were evaluated.

2.4. Evaluation

For each marker trajectory test dataset, the following metrics were evaluated. In each case, data from the initial 10 s were used for fine-tuning the LSTM network; the remaining data were used for the evaluation.

Prediction error

The prediction error was defined as the RMSE between the predicted marker position and the actual marker

position:

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

where y_i and $\hat{y}_i$ are the actual and predicted positions, respectively.

Residual motion

To determine the accuracy of respiratory-gated irradiation, the residual motion of the target during the beam-on period was evaluated. The residual motion was defined as the 95th percentile of the displacement of the recorded actual marker positions with reference to the center of the gating window during the beam-on period at an imaging rate of 15 fps.

Beam-on/off accuracy and beam efficiency

The beam efficiency was defined as the ratio of the beam-on time to the total imaging time. In addition, as an index of the temporal accuracy, the true beam-on (beam-off) rate was defined as the percentage of time that the beam was determined to be on (off) during the ideal beam-on (beam-off) period. The ideal beam-on/off times were determined according to the actual positions of the marker at an imaging rate of 15.0 fps.

3. Results

3.1. Prediction error

Figure 2 shows an example of the actual and predicted marker positions at frame rates of 7.5, 5.0, and 2.5 fps together with ideal beam-on/off and prediction-based beam-on/off times. The periods where the beam-on/off states are “true” or “false” are also shown. Without prediction, the times when the beam switches on or off are delayed compared to the ideal situation, especially at imaging rates of 5.0 and 2.5 fps, whereas, with prediction, these times closely match those for the ideal situation at all frame rates. The results for the true or false beam-on/off times confirm this. The typical examples of high and low prediction accuracy are shown in Figure 43.

The RMSE in the marker positions is shown in Figure 4. As a reference, RMSE of no-predictor model was evaluated by comparing each point and the point after T_{pred} in a similar way to the previous study [19]. The RMSE in the SI direction is larger than that in the anterior–posterior (AP) and left–right (LR) directions because the movement in the SI direction was greater than that in the other directions. As the imaging frame rate decreases, the RMSE increases. Except for the outliers in the SI direction, the median/maximum values of the RMSE for imaging rates of 7.5, 5.0, and 2.5 fps are 0.8/1.3, 0.9/1.6, and 1.2/2.4 mm, respectively. RMSE of no-predictor in 7.5 fps was larger than that in 5.0 fps because 266 ms ahead was predicted for 7.5 fps, and 200 ms ahead was predicted for 5.0 fps. However, the prediction performance in 7.5 fps was slightly better than 5.0 fps. One of the reasons could be the imaging interval T_{image} . Since the LSTM network state was updated frequently in 7.5 fps compared with 5.0 fps, the prediction performance was better in 7.5 fps although the prediction

window was greater than that of 5.0 fps.

3.2. *Residual motion of the target*

Figure 5 illustrates the residual motion of the target at each of the frame rates used. When prediction was applied, the amount of the residual motion was less than that without prediction. Except for the outliers in the SI direction, the median/maximum residual motion at an imaging rate of 15.0 fps without prediction, which is equivalent to currently clinical settings, is 2.3/3.6 mm. At 7.5, 5.0, and 2.5 fps with prediction, the corresponding values are 2.0/2.6, 2.2/3.3, and 2.4/3.9 mm, respectively. At an imaging rate of 7.5 fps with prediction, the residual motion is smaller than that at 15.0 fps without prediction. For imaging rates of 5.0 and 2.5 fps with prediction, the results are similar to those for a rate of 15.0 fps without prediction.

3.3. *Beam-on/off accuracy and beam efficiency*

The results for the beam efficiency, true beam-on rate, and true beam-off rate are shown in Figure 6. No significant difference in the irradiation efficiency was found between the cases where prediction was applied and where it was not applied. Similar to the results for the residual motion, the true beam-on/off rate was higher at an imaging rate of 7.5 fps with prediction than at 15.0 fps without prediction. At 5.0 and 2.5 fps with prediction, on average, the true beam-on/off rate was slightly lower than that at 15.0 fps without prediction.

The above results suggest that, using prediction, it should be possible to achieve an irradiation accuracy that is as good as, or better than, that achieved at an imaging rate of 15.0 fps without prediction.

4. Discussion

Here, we consider the possibility of reducing the imaging dose using prediction of the target position by comparing this to the situation where an imaging rate of 15.0 fps without prediction is used, the latter setting being that currently used for RTRT in clinical situations. It has been reported that, in particle beam therapy, the interplay effect is negligible for a residual target motion of 4 mm [20]. Except for outliers, imaging without prediction at rates below 7.5 fps includes cases in which the residual motion is greater than 4 mm. Hence, it has been considered that 15.0 fps is a clinically appropriate imaging rate if prediction is not applied. The results obtained in this study show that, compared with imaging at 15.0 fps without prediction, the residual motion at an imaging rate of 7.5 fps prediction was small and the irradiation efficiency was comparable. Therefore, by using 7.5-fps imaging with prediction, it should be possible to reduce the imaging dose by half while achieving a higher accuracy than that currently achieved in RTRT. It was also found that, except for outliers, the residual motion at 5.0 and 2.5 fps with prediction was less than 4 mm in all directions and that, again, the irradiation efficiency was similar to that achieved at 15.0 fps without prediction; this implies that even greater reductions in the imaging dose may be possible. For an imaging rate of 5.0 fps, assuming a total treatment time of approximately 1 h, the skin dose and the dose received at the isocenter would be approximately 0.2 and 0.1 Gy, respectively, in the two directions combined. Thus, in SBRT, for a prescribed dose of 40 Gy, the estimated

imaging dose could be expected to be lower than that recommended in the guidelines. The beam efficiency is varied for each patient as shown in Figure 6. If the respiratory motion is unstable, the beam efficiency will be decreased. As a result, the treatment time is prolonged, and the imaging dose is increased. Hence, we have to minimize the imaging dose while maintaining the acceptable irradiation accuracy in any cases. In addition to reducing the imaging rate, other effective methods of reducing the imaging dose include optimization of the X-ray conditions (the tube voltage and current and the pulse width used) [21] and image processing [22]. In addition, respiratory coaching and breath-holding techniques may assist in reducing the treatment time.

Since the accuracy of predictions of respiratory-induced motion should be comparable to the accuracy of the predictions made in this study, it is thought that predictions can be used to optimize the imaging frame rate for sites other than the liver. Prediction-based frame rate optimization could also be applied to other kind of IGRT techniques. Hunt *et al.* [23] reported on an internal marker detection technique based on 9.5-Hz kV and megavoltage (MV) X-ray imaging used in prostate treatment. Nguyen *et al.* [24] reported that marker locations in the prostate could be detected using kV X-ray imaging at 11 Hz. When these IGRT techniques are applied to SBRT, prediction should be effective at optimizing the imaging rate and minimizing the imaging dose. Motion prediction may also be applicable to markerless methods of evaluating tumor positions in kV images without internal markers. Hirai *et al.* [25] used 15.0-fps kV X-ray imaging for the validation of five lung and five liver cases; Terunuma *et al.* [26] used 30.0-fps imaging in their study. It would be possible to apply motion prediction to optimize the frame rate in these cases because the target position can be derived frame by frame.

Several limitations to this study remain. Although the ratio was small (approximately 3%), there were some individual large errors, as shown by the outliers in Figure 3. These outliers were due to such things as large baseline shifts that occurred during treatment and large variations in respiratory cycles and amplitude as shown in Figure 3. In clinical use, it is recommended that marker position data are obtained in rehearsal imaging before treatment and that the accuracy and applicability of the predictions are evaluated in advance. In the 1197 test datasets, the frame rate with the lowest RMSE was 7.5 fps in 862 cases and 5.0 fps in 335 cases. This result suggests the possibility of optimizing the frame rates on a per-patient or per-fraction basis. However, if the acceptable accuracy can not be obtained, it is recommended to use the 15-fps imaging without motion prediction. The other limitation was that the prediction accuracy may be varied on the algorithms. Algorithms for predicting respiratory motion other than LSTM have been reported. Samadi *et al.* compared the performance of several algorithms, including LSTM, Bi-LSTM, and gated recurring units (GRU) [11], and concluded that the GRU algorithm performed best. It has also been reported that there is not much difference between the prediction performance of various deep learning-based algorithms when the system latency was less than or equal 300 ms [13]. The optimal prediction algorithm for a particular situation may depend on the type of data used for prediction (e.g., one-dimensional external surrogate motion or three-dimensional internal surrogate motion). Although it was not focused to find the best algorithm in this study, the other technique might show better performance.

5. Conclusion

In this study, the prediction performance of an LSTM network and the residual motion of a target under gated irradiation at a range of imaging frame rates were evaluated to determine the optimal frame rate for real-time imaging used in RTTRT. No significant difference between the irradiation efficiency with and without prediction of the target position was found. The residual target motion using 7.5-fps imaging with prediction was smaller than that for 15.0-fps imaging without prediction. At 5.0 and 2.5 fps with prediction, the residual motion was approximately equal to that at 15.0 fps without prediction. Hence, there is a potential for reducing the imaging dose in RTTRT by applying motion prediction. In comparison with the current clinical setting for RTTRT, which uses an imaging rate of 15.0 fps without prediction, it may be possible to reduce the imaging dose by half or more.

Declaration of interest

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

Author Contribution

Kazuki Numakura Conceptualization, Methodology, Investigation, Formal analysis, Writing – original draft. Seishin Takao: Methodology, Investigation, Writing – review & editing. Taeko Matsuura: Writing – review & editing. Kouhei Yokokawa: Writing – review & editing. Ye Chen: Writing – review & editing. Yusuke Uchinami: Writing – review & editing. Hiroshi Taguchi: Writing – review & editing. Norio Kato: Writing – review & editing. Hidefumi Aoyama: Writing – review & editing. Satoshi Tomioka: Writing – review & editing. Naoki Miyamoto: Conceptualization, Methodology, Investigation, Formal analysis, Writing – original draft.

Funding

This research is supported by AMED under Grant Number JP23he2302001.

References

- [1] Shirato H, Shimizu S, Kunieda T, Kitamura K, van Herk M, Kagei K, et al. Physical aspects of a real-time tumor-tracking system for gated radiotherapy. *Int J Radiat Oncol Biol Phys*. 2000;48:1187-95.
- [2] Katoh N, Onimaru R, Sakuhara Y, Abo D, Shimizu S, Taguchi H, et al. Real-time tumor-tracking radiotherapy for adrenal tumors. *Radiother Oncol*. 2008;87:418-24.
- [3] Inoue T, Katoh N, Onimaru R, Shimizu S, Tsuchiya K, Suzuki R, et al. Stereotactic body radiotherapy using gated radiotherapy with real-time tumor-tracking for stage I non-small cell lung cancer. *Radiation oncology*. 2013;8:69.

- [4] Uchinami Y, Katoh N, Abo D, Taguchi H, Yasuda K, Nishioka K, et al. Treatment outcomes of stereotactic body radiation therapy using a real-time tumor-tracking radiotherapy system for hepatocellular carcinomas. *Hepatol Res.* 2021;51:870-9.
- [5] Shimizu S, Miyamoto N, Matsuura T, Fujii Y, Umezawa M, Umegaki K, et al. A proton beam therapy system dedicated to spot-scanning increases accuracy with moving tumors by real-time imaging and gating and reduces equipment size. *PloS one.* 2014;9:e94971.
- [6] Matsuura T, Miyamoto N, Shimizu S, Fujii Y, Umezawa M, Takao S, et al. Integration of a real-time tumor monitoring system into gated proton spot-scanning beam therapy: an initial phantom study using patient tumor trajectory data. *Med Phys.* 2013;40:071729.
- [7] Kanehira T, Matsuura T, Takao S, Matsuzaki Y, Fujii Y, Fujii T, et al. Impact of Real-Time Image Gating on Spot Scanning Proton Therapy for Lung Tumors: A Simulation Study. *Int J Radiat Oncol Biol Phys.* 2017;97:173-81.
- [8] Ionascu D, Jiang SB, Nishioka S, Shirato H, Berbeco RI. Internal-external correlation investigations of respiratory induced motion of lung tumors. *Med Phys.* 2007;34:3893-903.
- [9] Ding GX, Alaei P, Curran B, Flynn R, Gossman M, Mackie TR, et al. Image guidance doses delivered during radiotherapy: Quantification, management, and reduction: Report of the AAPM Therapy Physics Committee Task Group 180. *Med Phys.* 2018;45:e84-e99.
- [10] Wang G, Li Z, Li G, Dai G, Xiao Q, Bai L, et al. Real-time liver tracking algorithm based on LSTM and SVR networks for use in surface-guided radiation therapy. *Radiation oncology.* 2021;16:13.
- [11] Samadi Miandoab P, Saramad S, Setayeshi S. Respiratory motion prediction based on deep artificial neural networks in CyberKnife system: A comparative study. *Journal of applied clinical medical physics / American College of Medical Physics.* 2022:e13854.
- [12] Lin H, Shi C, Wang B, Chan MF, Tang X, Ji W. Towards real-time respiratory motion prediction based on long short-term memory neural networks. *Phys Med Biol.* 2019;64:085010.
- [13] Jeong S, Cheon W, Cho S, Han Y. Clinical applicability of deep learning-based respiratory signal prediction models for four-dimensional radiation therapy. *PloS one.* 2022;17:e0275719.
- [14] Li Y, Li Z, Zhu J, Li B, Shu H, Ge D. Online prediction for respiratory movement compensation: a patient-specific gating control for MRI-guided radiotherapy. *Radiation oncology.* 2023;18:149.
- [15] Lombardo E, Liu PZY, Waddington DEJ, Grover J, Whelan B, Wong E, et al. Experimental comparison of linear regression and LSTM motion prediction models for MLC-tracking on an MRI-linac. *Med Phys.* 2023;50:7083-92.
- [16] Ehrbar S, Braga Kaser S, Chamberlain M, Kräyenbühl J, Wilke L, Mayinger M, et al. MR-guided beam gating: Residual motion, gating efficiency and dose reconstruction for stereotactic treatments of the liver and lung. *Radiother Oncol.* 2022;174:101-8.
- [17] Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Comput.* 1997;9:1735-80.

- [18] Shiinoki T, Kawamura S, Uehara T, Yuasa Y, Fujimoto K, Koike M, et al. Evaluation of a combined respiratory-gating system comprising the TrueBeam linear accelerator and a new real-time tumor-tracking radiotherapy system: a preliminary study. *Journal of applied clinical medical physics / American College of Medical Physics*. 2016;17:202-13.
- [19] Lombardo E, Rabe M, Xiong Y, Nierer L, Cusumano D, Placidi L, et al. Evaluation of real-time tumor contour prediction using LSTM networks for MR-guided radiotherapy. *Radiother Oncol*. 2023;182:109555.
- [20] Mori S, Asakura H, Kandatsu S, Kumagai M, Baba M, Endo M. Magnitude of residual internal anatomy motion on heavy charged particle dose distribution in respiratory gated lung therapy. *Int J Radiat Oncol Biol Phys*. 2008;71:587-94.
- [21] Miyamoto N, Ishikawa M, Bengua G, Sutherland K, Suzuki R, Kimura S, et al. Optimization of fluoroscopy parameters using pattern matching prediction in the real-time tumor-tracking radiotherapy system. *Phys Med Biol*. 2011;56:4803-13.
- [22] Miyamoto N, Ishikawa M, Sutherland K, Suzuki R, Matsuura T, Toramatsu C, et al. A motion-compensated image filter for low-dose fluoroscopy in a real-time tumor-tracking radiotherapy system. *J Radiat Res*. 2014.
- [23] Hunt MA, Sonnick M, Pham H, Regmi R, Xiong JP, Morf D, et al. Simultaneous MV-kV imaging for intrafractional motion management during volumetric-modulated arc therapy delivery. *Journal of applied clinical medical physics / American College of Medical Physics*. 2016;17:473-86.
- [24] Nguyen DT, O'Brien R, Kim JH, Huang CY, Wilton L, Greer P, et al. The first clinical implementation of a real-time six degree of freedom target tracking system during radiation therapy based on Kilovoltage Intrafraction Monitoring (KIM). *Radiother Oncol*. 2017;123:37-42.
- [25] Hirai R, Sakata Y, Tanizawa A, Mori S. Real-time tumor tracking using fluoroscopic imaging with deep neural network analysis. *Phys Med*. 2019;59:22-9.
- [26] Terunuma T, Sakae T, Hu Y, Takei H, Moriya S, Okumura T, et al. Explainability and controllability of patient-specific deep learning with attention-based augmentation for markerless image-guided radiotherapy. *Med Phys*. 2023;50:480-94.

Figure captions

Figure 1: Time-series representing the actual marker motion and measured marker position. Ideal beam-on/off and gate-on/off signals are also shown together with actual beam-on/off signals (a) without prediction and (b) with prediction.

Figure 2: Example of actual marker positions recorded at an imaging rate of 15.0 fps and predicted marker position in the SI direction at rates of (a) 7.5, (b) 5.0, and (c) 2.5 fps. Yellow colored area represents the gating window. The ideal beam-on/off patterns and the actual patterns without prediction and with prediction are also shown.

Figure 3: Example of actual marker positions and predicted marker positions in the SI direction at rate of 5.0 fps in high prediction accuracy (RMSE of 0.4 mm) and low prediction accuracy (RMSE of 2.3 mm).

Figure 4: RMSE in the marker positions for frame rates of 7.5, 5.0, and 2.5 fps in the (a) LR, (b) SI, and (c) AP directions

Figure 5: Residual motion of the target during gated beam delivery in the (a) LR, (b) SI, and (c) AP directions

Figure 6: Results for (a) the beam efficiency, (b) true beam-on rate, and (c) true beam-off rate

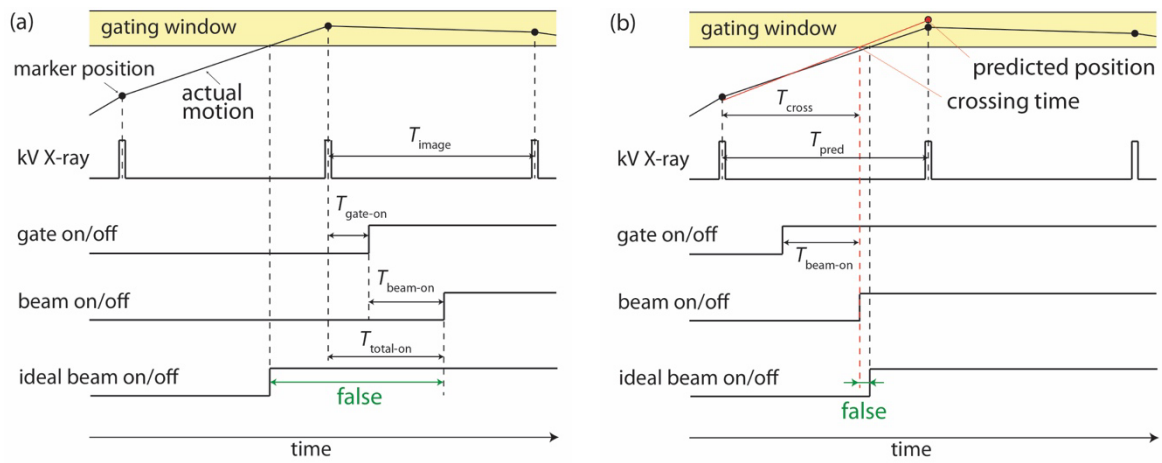


Figure 1

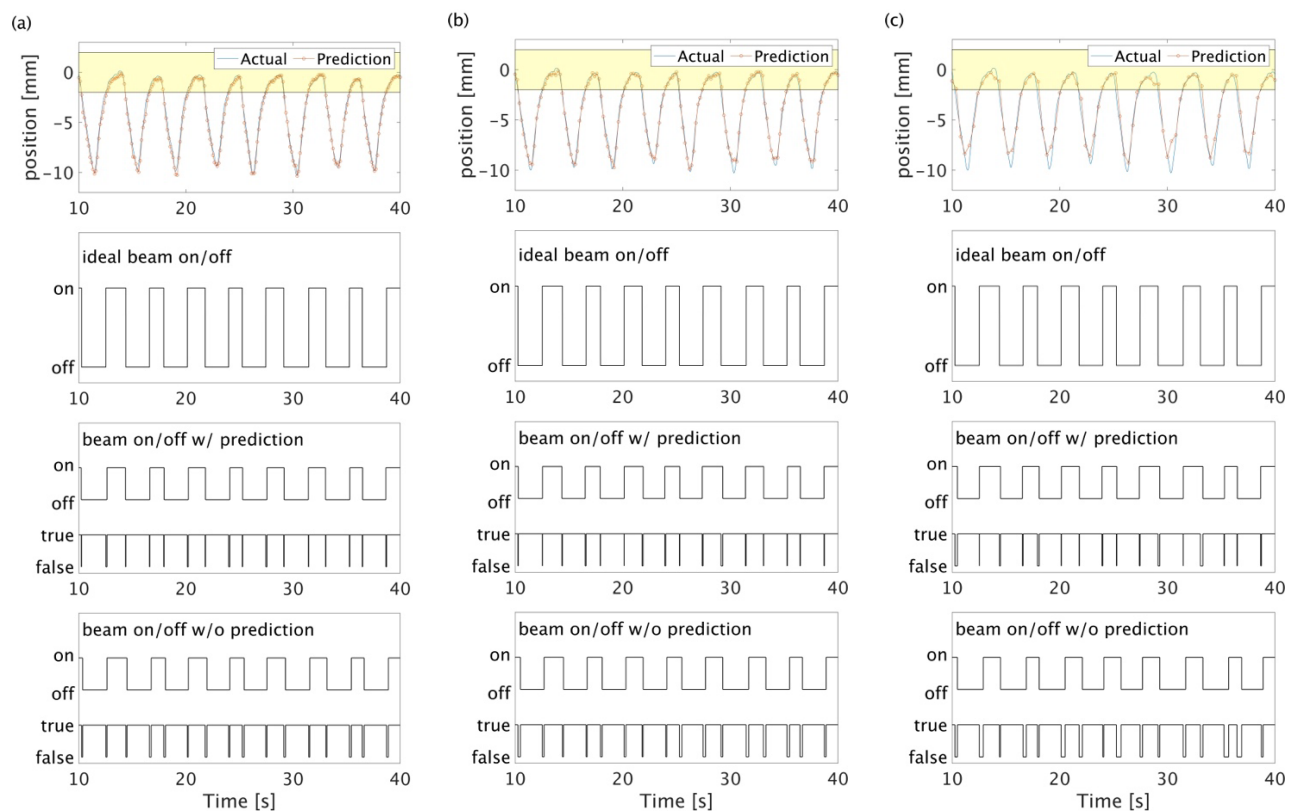


Figure 2

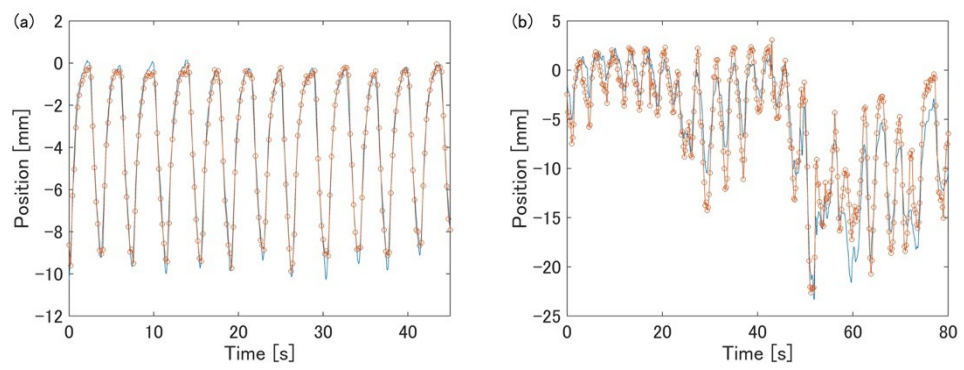


Figure 3

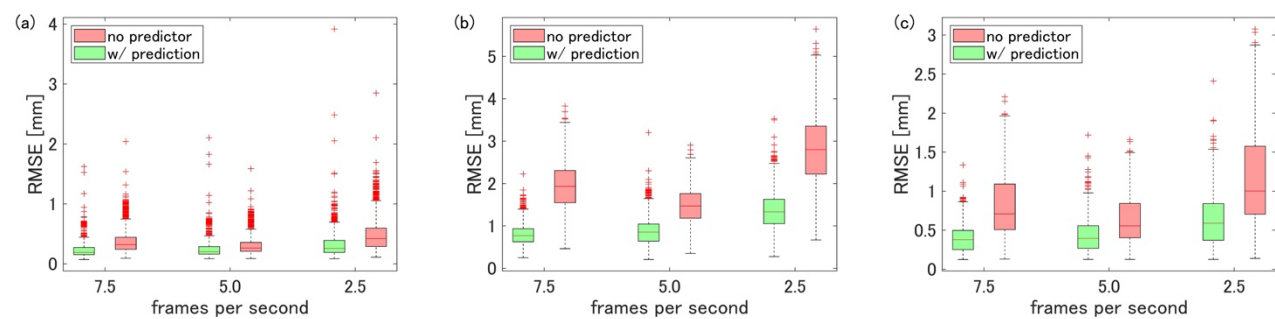


Figure 4

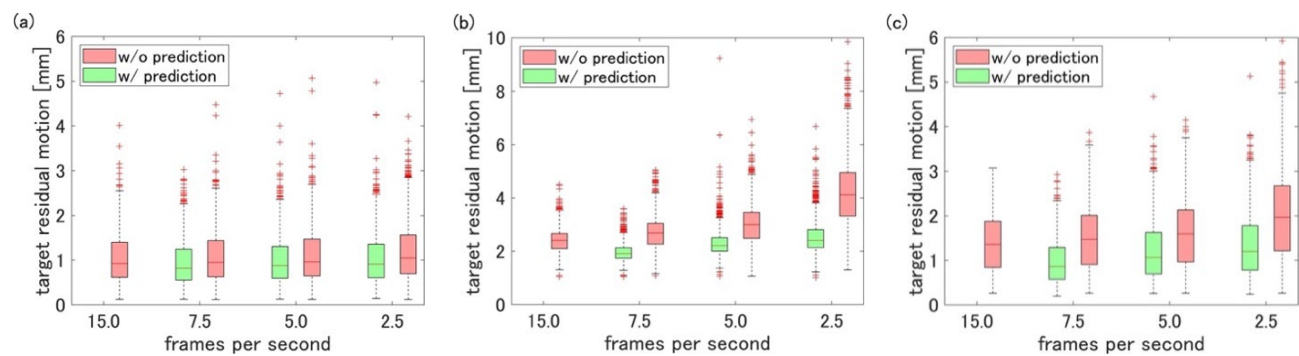


Figure 5

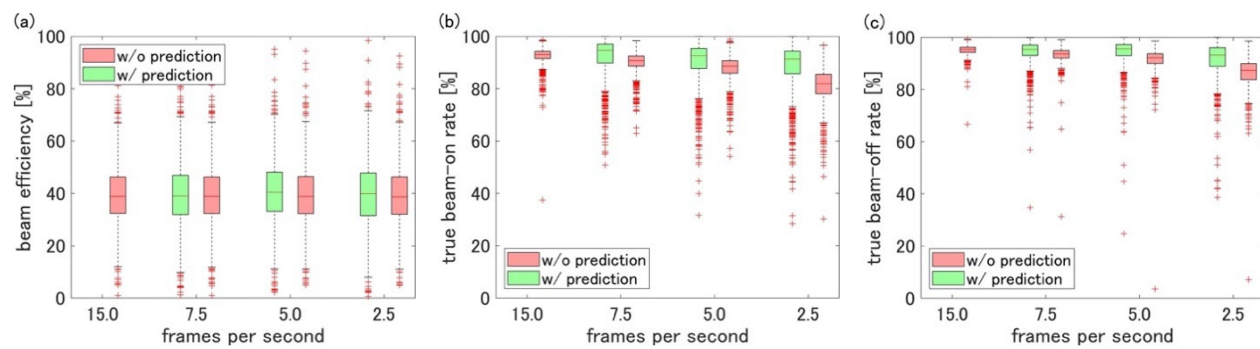


Figure 6