

Deep Learning based Respiratory Pattern Classification and Applications in PET/CT Motion Correction

Yin Guo, Nicha Dvornek, Yihuan Lu, Yu-jung Tsai, James Hamill, Michael Casey, Chi Liu

Abstract—Respiratory motion has to be corrected in PET/CT imaging for precise tumor detection and quantification. The optimal motion correction methods for regular breathers and irregular breathers could be different. In this study, we developed deep learning based methods to automatically classify patient breathing patterns and investigated the impact of breathing pattern variability on gating performance. We implemented a hybrid neural network consisting of convolutional (Conv) layers, recurrent layers (LSTM, long short-term memory) and a linear classifier to differentiate breathing patterns. 1295 respiratory traces collected using RPM (Real-time Position Management) system were used for training and testing, as well as additional traces acquired using the Anzai system. We optimized the deep neural network with respect to data preprocessing, augmentation, weighted loss function and generalization capability. The results showed that the proposed deep learning model has reached a high prediction accuracy, with a sensitivity of 92.0% and a specificity of 91.8%. Using phase gating approach, for regular breathers, end-expiration phase gating can effectively correct the respiratory motion. In contrast, for irregular breathers, larger amount of intra-gate motion was present in the gated PET/CT images and more sophisticated motion correction methods are required.

I. INTRODUCTION

Respiratory motion is one of the major causes of degradation of PET image quality. Image blurring and artifacts due to breathing are unavoidable, as the usual duration of PET acquisition is much longer than a respiratory cycle. Respiratory motion averages 7.5 mm in mean amplitude for long tumors, and 18.6 mm for abdominal tumors (liver, kidney, pancreas) [1]. These motions might have a substantial impact on tumor detection and quantification. Previous researches have demonstrated that, the blurred images lead to 10-75% underestimation of maximum standard uptake value, and up to 2-fold overestimation of tumor volume [2]. Moreover, in PET/CT that employs CT for attenuation correction, the mismatch between breath-hold CT, a snapshot of one respiratory location within one breathing cycle, and PET, a multi-minute average over many breathing cycles, can

Manuscript received Dec 12, 2019. This work was supported in part by NIH grant R01CA224140.

Yin Guo was with the Department of Biomedical Engineering, Yale University, New Haven, CT 06511 USA. He is now with the Department of Radiology, University of Washington, Seattle, WA 98109 USA (telephone: 203-390-0978, e-mail: yinguo@uw.edu).

Nicha Dvornek, Yihuan Lu, Yu-jung Tsai are with the Department of Radiology and Biomedical Imaging, Yale University, New Haven, CT 06511 USA.

James Hamill and Michael Casey are with the Siemens Medical Solutions, Knoxville, TN 37932 USA.

Chi Liu is with the Department of Biomedical Engineering and the Department of Radiology and Biomedical Imaging, Yale University, New Haven, CT 06511 USA (telephone: 203-785-4269, e-mail: chi.liu@yale.edu).

cause significant mis-localization, artifacts and quantitative errors through CT-based attenuation correction [1]. Therefore, respiratory motion has to be corrected.

Different breathing patterns might require different motion correction methods. A previous study [2] has defined regular breathers, whose respiratory traces have consistent amplitudes and spend more time during the end-expiration phase that is relatively motion free, and irregular breathers, whose traces suffer from significant amplitude variation and baseline change. The histogram of irregular traces does not have an end-expiration peak and typically spread arbitrarily over various displacements. Representative breathing patterns are demonstrated in Fig. 1. Skewness can serve as an intuitive metric to quantify the asymmetric of histogram shapes, but lacks sensitivity and specificity.

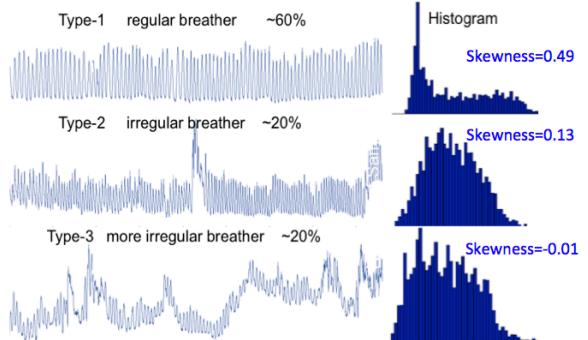


Figure 1. Breathing pattern variability [2].

Patients' breathing patterns often change over time, which brings even more complications. Changes often occur within a PET imaging session, or breathing patterns can be substantially different during PET and CT acquisitions. For approximately every 10 minutes, breathing patterns of patients are likely to change [3], which causes additional attenuation correction mismatch and tumor quantitative errors. The variability of respiratory patterns thus requires precise differentiation of regular from irregular breathers, and detection and tracking of changes in these patterns. Additionally, the impact of breathing patterns changes on existing motion correction methods have to be investigated.

Deep neural networks have served as a promising tool in an automatic fashion for supervised sequence classification tasks. Both convolutional and recurrent networks have seen many successes in processing time-series, such as arrhythmia detection [4], fMRI disease state diagnosis [5] and genetic sequencing [6]. Yet the problem of varying breathing patterns has not been addressed by this new technique.

In this study, we trained a neural network to differentiate irregular breathers from regular breathers with high accuracy, and explored the impact of irregular breathing patterns on respiratory gating. This classification can guide us to build an automated and personalized PET/CT motion correction framework, where for regular breathers, less complex algorithms (like phase gating and end-expiration gating) [7] are expected to provide equivalent accuracy while saving precious console computation time, and for irregular breathers, more sophisticated and time-consuming event-by-event motion correction is required to correct for the extra intra-gate motion [8].

II. METHODS

A. Deep learning network structure

For the RPM supervised sequence classification task, the neural network is a hybrid structure consisting of convolutional (Conv) layers, recurrent layers and a linear classifier. RPM traces are fed as input into the network.

The 1D Conv layers (Fig. 2) [9] serve as feature extractors from local input patches, like the 2D counterpart for image processing. The property of weight sharing and translate invariance in temporal dimension allows for efficient computation. With 1D Conv layers, longer sequences are convolved and pooled into shorter ones with rich Conv features.

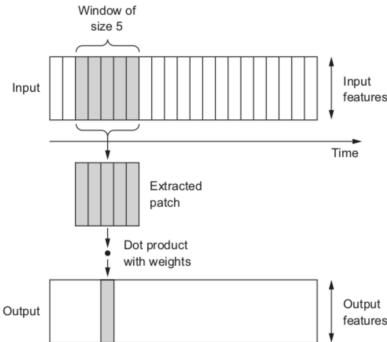


Figure 2. Diagram of 1D convolution [9].

However, these layers process input patches independently, and are not sensitive to the order of the timesteps, unlike the proceeding LSTM units (Fig. 3) [10]. LSTM is a popular RNN layer, explicitly designed for handling the long-term dependency problem. Three interactive gates work interactively to process the time-series, allowing features to propagate through long sequences [11].

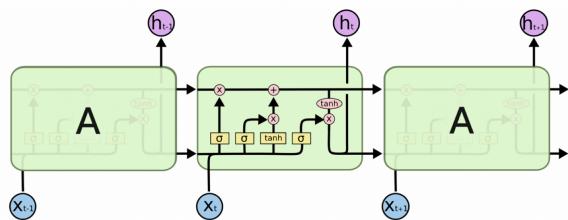


Figure 3. Diagram of LSTM [11].

Finally, binary breathing pattern ($y = 0$ as regular breather or $y = 1$ as irregular breather) is predicted by the linear classifier. The schematic diagram of the hybrid classification network is shown in Fig. 4.

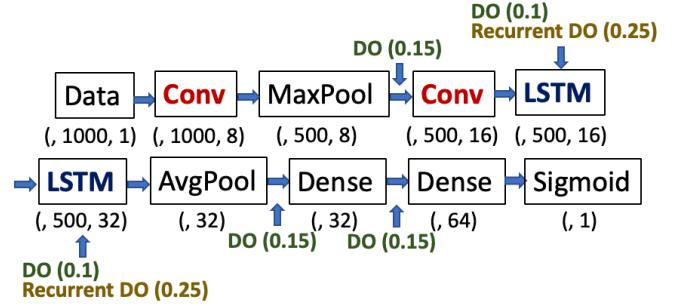


Figure 4. The hybrid structure of classification network, (batchsize, timesteps, features) denoting the dimension of tensor.

B. Dataset preparation

The training and validation dataset of respiratory traces for the network are collected from the RPM system, which monitored the chest motion during a 7-minute PET scan. This study recruits 1295 patients at the University of Washington, and each trace has been individually inspected visually and determined to be regular or irregular breather. Around 60% traces are from regular breathers and the remaining 40% otherwise.

Additional traces by the Anzai system and corresponding PET imaging data acquired at Yale PET Center have been used for further testing. Both single bed and continuous bed motion (CBM) acquisition protocols record breathing traces between 1 and 2 hours in length.

In order to fit traces collected from both motion tracking systems into a generalized framework, traces have been downsampled to 3 Hz, balancing the tradeoff between signal fidelity and computational expenses. Then we performed a retrospective calculation of amplitudes of each breathing cycle, so that all traces were normalized to a mean of zero and a mean amplitude of one across all cycles.

C. Training and testing setup

Random crops of traces were used for data augmentation. We used dropout (DO) and recurrent dropout (recurrent DO) units (Fig. 4) to prevent overfitting. The cross-entropy (CE) loss function is weighted as follows [12],

$$\text{Weighted CE } (p_t, y) = -(1 - p_t)^2 \log(p_t)$$

$$p_t = \begin{cases} p, & y = 1 \\ 1 - p, & y = 0 \end{cases}$$

where p is the calculated classification probability by the output layer of the sigmoid function. Compared with traditional cross-entropy loss $\text{CE}(p_t, y) = -\log(p_t)$, the term $(1 - p_t)^2$ can help balance the importance between training examples that the network finds it easy or hard to classify. Additionally, we determined the threshold of p by investigating the ROC curve.

All the training and evaluation are performed with Keras (TensorFlow backend) as the deep learning framework, using

a server with one Titan Xp GPU. We randomly select 20% subjects out of 1295 as the test set, while using the remaining for training. The prediction accuracy of neural networks consisting pure convolutional layers, pure LSTM layers and the hybrid structure has also been tested and compared.

The capability of model generalization was further tested with Anzai traces. Reconstructions of PET images, with both no motion correction and phase gating, were generated with Siemens e-7 tools. Attenuation correction was performed with the corresponding CT data.

III. RESULTS

A. Test of RPM traces

The neural network is able to properly differentiate irregular RPM traces from regular ones, achieving a high testing accuracy of 91.9%, with an ordinary decision boundary of $p = 0.5$, as reported in Table 1. The sensitivity and specificity are 92.0% and 91.8%, respectively, and the area under the ROC curve is 0.97.

TABLE I. RPM TEST RESULTS

	Predicted 'regular'	Predicted 'irregular'	Sum
True 'regular'	101	9	110
True 'irregular'	7	81	88
Sum	108	90	198

The comparison of prediction performance of three different neural network structures is summarized in Table 2. The hybrid structure outperforms the pure convolutional or recurrent ones, showing that to use 1D Conv layers as preprocessing steps before LSTMs is beneficial for further exploiting the temporal features of breathing patterns.

TABLE II. COMPARISON OF THREE NEURAL NETWORK STRUCTURES

	Accuracy	Sensitivity	Specificity
Pure convolutional	81.0%	81.4%	80.6%
Pure recurrent	86.5%	86.4%	86.6%
Hybrid structure	91.9%	92.0%	91.8%

The ROC curve was used for selection of the classification threshold. Since high true positive rate (correct labeling of irregular breathers) is desired, we set a decision threshold of $p = 0.439$, increasing sensitivity to 97.7% while specificity as a tradeoff is therefore 84.5%. Examples of correctly classified results of RPM traces are shown in Fig. 5.

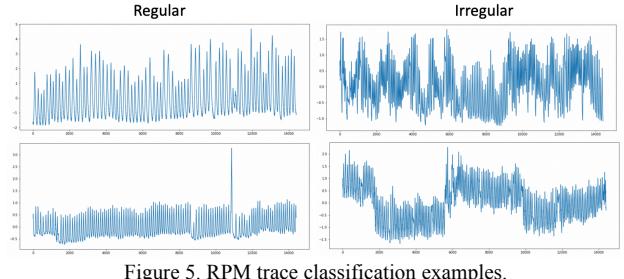


Figure 5. RPM trace classification examples.

B. Test of Anzai traces

Due to the trace normalization schemes, the network is equally applicable to Anzai tracers, as classified examples shown in Fig. 6. As the neural network is able to output prediction results given an arbitrary trace length, accurate classification results can be produced for fragments of Anzai traces (for example, each 5-minute CBM scan path).

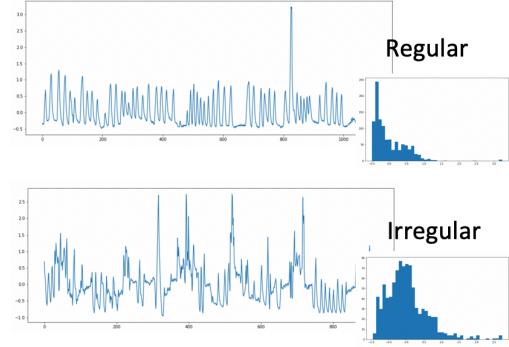


Figure 6. Anzai trace classification examples (CBM).

Furthermore, anomaly detection has been performed on the traces, using a sliding-window of 30 seconds to classify every few breathing cycles, so that we are able to detect when the breathing pattern changes. Examples in Fig. 7 showed that the network is able to detect changed breathing patterns within a scan. Breathing pattern changes can be tracked precisely, regardless of the overall breathing pattern of regular or irregular.

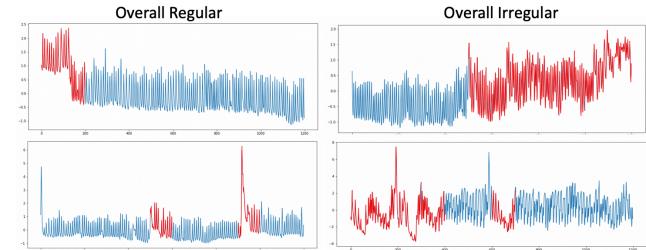


Figure 7. Anzai trace anomaly detection examples (single bed), blue as regular and red as irregular.

C. Effects of breathing on phase gating reconstruction

We have performed initial experiments to investigate the effects of breathing patterns, determined by our proposed deep learning based classification, on image reconstruction with phase gating.

As a sample patient shown in Fig. 8, for two consecutive 5-minute scans, the breathing pattern was regular for the first 5 minutes but changed to irregular pattern for the second 5 minutes. Such breathing pattern changes have been correctly identified by the proposed network. Without motion correction, the image of the second 5 minutes is more blurred, likely due to the additional motion caused by breathing irregularity. In this case the inspiration gating did not improve much due to intra-gate motion. The end expiration phase gating provided more effective motion reduction for the 0-5 min scan with regular breathing pattern, while additional blurring was observed in the 5-10 min scan with irregular breathing pattern, due to intra-gate motion.

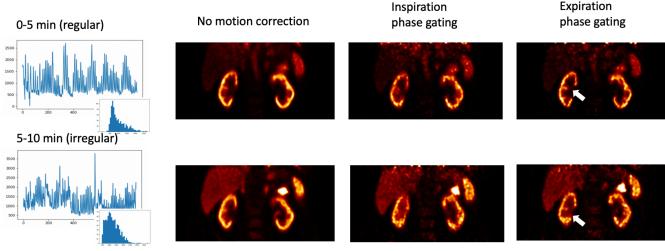


Figure 8. Sample results of the same patient at two consecutive 5 min scans with different breathing patterns and resulting reconstructed images with and without phase gating.

As another example shown in Fig. 9, after the detected shift of regular respiratory pattern to irregular occurs at around 50 minutes post injection, the intra-gated motion of end-expiration phase gating increased, causing additional blurring as compared to the 40-50 min scan, as shown in the coronal slices of abdominal structures.

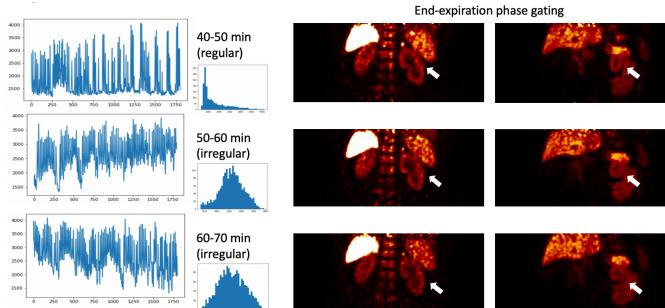


Figure 9. Sample results of the same patient at three consecutive 10 min scans of different breathing patterns and resulting reconstructed images with end-expiration.

IV. DISCUSSION

Due to the large amount of intra-gate motion caused by additional inter-cycle variability and inter-cycle variability, respiratory motion correction for irregular breathers are particularly challenging. Previous studies have shown that event-by-event listmode based motion correction utilizing the correlation between internal organs and external (INTEX) motion monitoring signal can effectively correct intra-gate motion, as the irregular breathing pattern information is maintained in the external motion signal and is taken into full

account in the INTEX motion correction framework [13]-[17]. However, such event-by-event correction is time-consuming.

On the other hand, several other efficient motion correction methods have been implemented on clinical scanners [7]-[8]. For regular breathers with smaller amount of intra-gate motion, the motion correction performance of existing methods is expected to be similar to that of the more sophisticated approach, such as INTEX.

In clinical practice, both image reconstruction and acquisition are typically performed on the same console computer. Therefore, precious console time prevents performing event-by-event motion correction for every patient. Using the proposed deep learning based breathing pattern classification method, one is able to determine if a patient is a regular or irregular breather. For regular breathers, existing motion correction methods on the scanner should provide sufficient motion correction. Only for the irregular breathers, event-by-event correction method will be performed to optimize the computing time of scanner console.

In terms of classification accuracy tradeoff between sensitivity and specificity of detecting irregular breathers, we chose to achieve high sensitivity at the expense of specificity. If a regular breather is mis-classified as an irregular breather, the consequence is that event-by-event motion correction will be performed thus more computational resources are used. In contrast, if an irregular breather is mis-classified as a regular breather, using existing motion correction methods lacking intra-gate motion correction will result in sub-optimal image quality, which can subsequently affect patient diagnosis and management. Therefore, we chose the threshold $p = 0.439$ in order to achieve 97.7% sensitivity but 84.5% specificity.

V. SUMMARY

In this study, we developed a neural network to identify the irregularity of the recorded breathing traces for PET/CT. We also showed that irregular breathing patterns would degrade the motion correction based on gating due to inter-cycle and intra-cycle motion variations. The developed classification network provides high sensitivity and has the potential of facilitating a workflow of automated and personalized PET/CT motion correction.

REFERENCES

- [1] Nehmeh, S. A., & Erdi, Y. E. (2008). Respiratory motion in positron emission tomography/computed tomography: a review. In *Seminars in nuclear medicine* (Vol. 38, No. 3, pp. 167-176). WB Saunders.
- [2] Liu, C., Pierce, L. A., II, A. M. A., & Kinahan, P. E. (2009). The impact of respiratory motion on tumor quantification and delineation in static PET/CT imaging. *Physics in medicine and biology*, 54(24), 7345.
- [3] Keall, P. J., Mageras, G. S., Balter, J. M., Emery, R. S., Forster, K. M., Jiang, S. B., ... & Ramsey, C. R. (2006). The management of respiratory motion in radiation oncology report of AAPM Task Group 76 a. *Medical physics*, 33(10), 3874-3900.
- [4] Rajpurkar, P., Hannun, A. Y., Haghpanahi, M., Bourn, C., & Ng, A. Y. (2017). Cardiologist-level arrhythmia detection with convolutional neural networks. *arXiv preprint arXiv:1707.01836*.
- [5] Dvornek, N. C., Ventola, P., Pelphrey, K. A., & Duncan, J. S. (2017). Identifying autism from resting-state fMRI using long short-term memory networks. In *International Workshop on Machine Learning in Medical Imaging* (pp. 362-370). Springer, Cham.

- [6] Quang, D., & Xie, X. (2016). DanQ: a hybrid convolutional and recurrent deep neural network for quantifying the function of DNA sequences. *Nucleic acids research*, 44(11), e107-e107.
- [7] Liu, C., Alessio, A., Pierce, L., Thielemans, K., Wollenweber, S., Ganin, A., & Kinahan, P. (2010). Quiescent period respiratory gating for PET/CT. *Medical physics*, 37(9), 5037-5043.
- [8] Meier, J. G., Wu, C. C., Cuellar, S. L. B., Truong, M. T., Erasmus, J. J., Einstein, S. A., & Mawlawi, O. R. (2019). Evaluation of a Novel Elastic Respiratory Motion Correction Algorithm on Quantification and Image Quality in Abdominothoracic PET/CT. *Journal of Nuclear Medicine*, 60(2), 279-284.
- [9] Chollet, F. (2018). Deep Learning with Python and Keras. Manning Publications.
- [10] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8), 1735-1780.
- [11] Olah, C. (2015). Understanding LSTM networks. Retrieved from <https://colah.github.io/posts/2015-08-Understanding-LSTMs>.
- [12] Lin, T. Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017). Focal loss for dense object detection. In *Proceedings of the IEEE international conference on computer vision* (pp. 2980-2988).
- [13] Liu, C., Alessio, A. M., & Kinahan, P. E. (2011). Respiratory motion correction for quantitative PET/CT using all detected events with internal-external motion correlation. *Medical physics*, 38(5), 2715-2723.
- [14] Chan, C., Jin, X., Fung, E. K., Naganawa, M., Mulnix, T., Carson, R. E., & Liu, C. (2013). Event-by-event respiratory motion correction for PET with 3D internal-1D external motion correlation. *Medical physics*, 40(11), 112507.
- [15] Yu, Y., Chan, C., Ma, T., Liu, Y., Gallezot, J. D., Naganawa, M., ... & Liu, C. (2016). Event-by-event continuous respiratory motion correction for dynamic PET imaging. *Journal of Nuclear Medicine*, 57(7), 1084-1090.
- [16] Chan, C., Onofrey, J., Jian, Y., Germino, M., Papademetris, X., Carson, R. E., & Liu, C. (2017). Non-rigid event-by-event continuous respiratory motion compensated list-mode reconstruction for PET. *IEEE transactions on medical imaging*, 37(2), 504-515.
- [17] Lu, Y., Fontaine, K., Mulnix, T., Onofrey, J. A., Ren, S., Panin, V., ... & Fulton, R. (2018). Respiratory motion compensation for PET/CT with motion information derived from matched attenuation-corrected gated PET data. *Journal of Nuclear Medicine*, 59(9), 1480-1486.