PROCEEDINGS OF SPIE

Automatic labeling of respiratory phases and detection of abnormal respiratory signals in free-breathing thoracic dynamic MR image acquisitions based on deep learning

Sun, Changjian, Udupa, Jayaram, Tong, Yubing, Wu, Caiyun, McDonough, Joseph, et al.

SPIE.

Automatic labeling of respiratory phases and detection of abnormal respiratory signals in free-breathing thoracic dynamic MR image acquisitions based on deep learning

Changjian Sun[1,2‡], Jayaram K. Udupa[2*], Yubing Tong[2‡], Caiyun Wu[2], Joseph M. McDonough[3], Catherine Qiu[3], Carina Lott[3], Jason B. Anari[3], Drew A. Torigian[2], Patrick J. Cahill[3]

[1] College of Electronic Science and Engineering, Jilin University, Changchun, China.

[2] Medical Image Processing Group, 602 Goddard building, 3710 Hamilton Walk, Department of Radiology, University of Pennsylvania, Philadelphia, PA 19104, United States.

[3] Wyss Campbell Center for Thoracic Insufficiency Syndrome, Children's Hospital of Philadelphia, Philadelphia, PA, 19104, United States

*Corresponding author; ‡ co-first authors.

ABSTRACT

4D thoracic images constructed from free-breathing 2D slice acquisitions based on dynamic magnetic resonance imaging (dMRI) provide clinicians the capability of examining the dynamic function of the left and right lungs, left and right hemi-diaphragms, and left and right chest wall separately for thoracic insufficiency syndrome (TIS) treatment [1]. There are two shortcomings of the existing 4D construction methods [2]: a) the respiratory phase corresponding to end expiration (EE) and end inspiration (EI) need to be manually identified in the dMRI sequence; b) abnormal breathing signals due to non-tidal breathing cannot be detected automatically which affects the construction process. Since the typical 2D dynamic MRI acquisition contains ~3000 slices per patient, handling these tasks manually is very labor intensive. In this study, we propose a deep-learning-based framework for addressing both problems via convolutional neural networks (CNNs) [3] and Long Short-Term Memory (LSTM) [4] models. A CNN is used to extract the motion characteristics from the respiratory dMRI sequences to automatically identify contiguous sequences of slices representing exhalation and inhalation processes. EE and EI annotations are subsequently completed by comparing the changes in the direction of motion of the diaphragm. A LSTM network is used for detecting abnormal respiratory signals by exploiting the non-uniform motion feature sequence of abnormal breathing motions. Experimental results show the mean error of labeling EE and EI is ~0.3 dMRI time point unit (much less than one time point). The accuracy of abnormal cycle detection reaches 80.0%. The proposed approach achieves results highly comparable to manual labeling in accuracy but with close to full automation of the whole process. The framework proposed here can be readily adapted to other modalities and dynamic imaging applications.

1. INTRODUCTION

4D constructed thoracic images can improve disease visualization and quantification and thus provide a basis for lung function evaluation and study of disease processes such as thoracic insufficiency syndrome (TIS) [5]. Recently, several MRI-based 4D construction methods have been proposed [2, 6]. In our ongoing TIS projects [1, 2, 6], thoracic 4D MRI plays a very important role. TIS is a complex condition that involves deformities of the spine, rib cage, sternum, and chest wall and which can seriously compromise lung function. In most cases, children with TIS are also born with congenital spinal disorders, such as scoliosis [5, 7]. In normal developing children, lung growth parallels chest and spine growth. In children with TIS, lung growth is limited by rib deformities and spinal curves. As a result, children may become dependent on nasal oxygen or ventilator support to breathe, which makes it difficult to acquire dMRI data sets and perform 4D construction by using external devices or gating techniques to obtain respiratory signals. We previously developed a 4D construction method that was based on free-breathing slice acquisitions and a graph-based global optimization technique [2], which completely eliminated the need for any signals for gating.

One of the shortcomings of the previous approach is that human effort is needed to manually label end expiration (EE) and end inspiration (EI) time points in the respiratory cycle of the dMRI sequences. For one patient, considering that typically, for each of 30-40 sagittal slice locations across the thorax, ~80 dMRI slice data are acquired under free-breathing conditions over several respiratory cycles, the total number of resulting slices is 2400 to 3200. Labeling all of these frames manually takes about 4 hours per patient and greatly reduces the throughput of the entire pipeline for deriving useful quantitative information from these data. Auto-labeling of EE and EI time points via an optical flow-based method was proposed recently to address this concern [6]. This method has one drawback, which prompted us to develop the proposed solution. During dMRI acquisition, often patients take a deep breath or perform shallow breathing. This affects quantification derived with the assumption of tidal breathing conditions. Our current approach is to manually detect these abnormal cycles when present and simply discard such data sets from further analysis completely. Even though the data sets are discarded, the manual detection process itself is time-consuming.

The main contribution of this paper is a combined method, a deep-learning-based tool, for automatic labeling of respiratory signals as well as for detecting abnormal respiratory cycles. With such a tool, we can perform labeling of breathing phases and the detection of abnormal respiratory cycles with very minimal user interaction. The framework is based on convolutional neural networks (CNNs) [3] and Long Short-Term Memory (LSTM) [4] models. A CNN is used to extract the motion characteristics from the respiratory dMRI sequences to automatically identify contiguous sequences of slices representing exhalation and inhalation processes. EE and EI annotations are subsequently completed by comparing the changes in the direction of motion of the diaphragm. Irregular (shallow and deep) breathing patterns cause the CNN to extract non-uniform motion feature sequences corresponding to each sagittal scanning position. These feature sequences can be used to distinguish between normal tidal breathing and abnormal breathing cycles. An LSTM model is used for detecting abnormal cycles based on these motion feature sequences of normal and abnormal breathing. The accuracies of labeling EE and EI time points and detecting abnormal breathing cycles are evaluated by comparing with the manual method. The experimental results show that the proposed method outperforms previous methods in terms of accuracy and almost fully automates the process.

2. MATERIALS AND METHODS

Image data

Our experimental data were retrospectively collected from the Children's Hospital of Philadelphia (CHOP). This study was conducted following approval from the Institutional Review Board at CHOP along with a Health Insurance Portability and Accountability Act waiver. The slices in our acquisition are 224×256 pixels with a pixel size of 0.78×0.78 - 1.46×1.46 mm^2. A total of 48 dMRI data sets gathered from 45 subjects were utilized in our study. This ensemble included data sets from 20 normal children, 20 pediatric patients with TIS, and 5 normal adults, three of whom were scanned twice. Since our study is sagittal slice location specific and not patient-specific, the total number of such locations in our data sets is ~1800 and the number of locations with abnormal cycles is 112.

Main idea of auto-labeling and abnormal breathing signal detection

In our set up, for each of 35-40 sagittal slice locations across the chest, 80 dMRI slices are acquired rapidly (in 200-300 ms per slice) while the patient is undergoing free breathing. Since all processing is done identically in this paper on the sequence of slices acquired for each sagittal location, we will confine our description to one fixed sagittal location and represent the sequence of 80 slices acquired for a sagittal location by $A = \left\{ f_{T_1}, f_{T_2}, ..., f_{T_M} \right\}$. This constitutes a time sequence of slices. Our goal for the combined CNN-LSTM strategy is to discard any subsequences within A that constitute abnormal breathing cycles, and for the normal cycles, to mark each f_{T_i} in A with one of the three labels: EE, EI, and none.

The main idea of our method is to use the CNN model to extract the motion information associated with respiratory phase successively from two adjacent frames. Subsequently, we derive the motion information (feature sequences) of the time series through CNN. Meanwhile, the normal EE and EI phases are labeled according to the characteristic sequence. Finally, LSTM performs screening and culling of abnormal respiratory signals by employing the feature sequences. Prior to training,

the optical flow method is applied to the images for preliminary extraction of motion characteristics of adjacent frames. Figure 1 is a schematic overview of the framework.
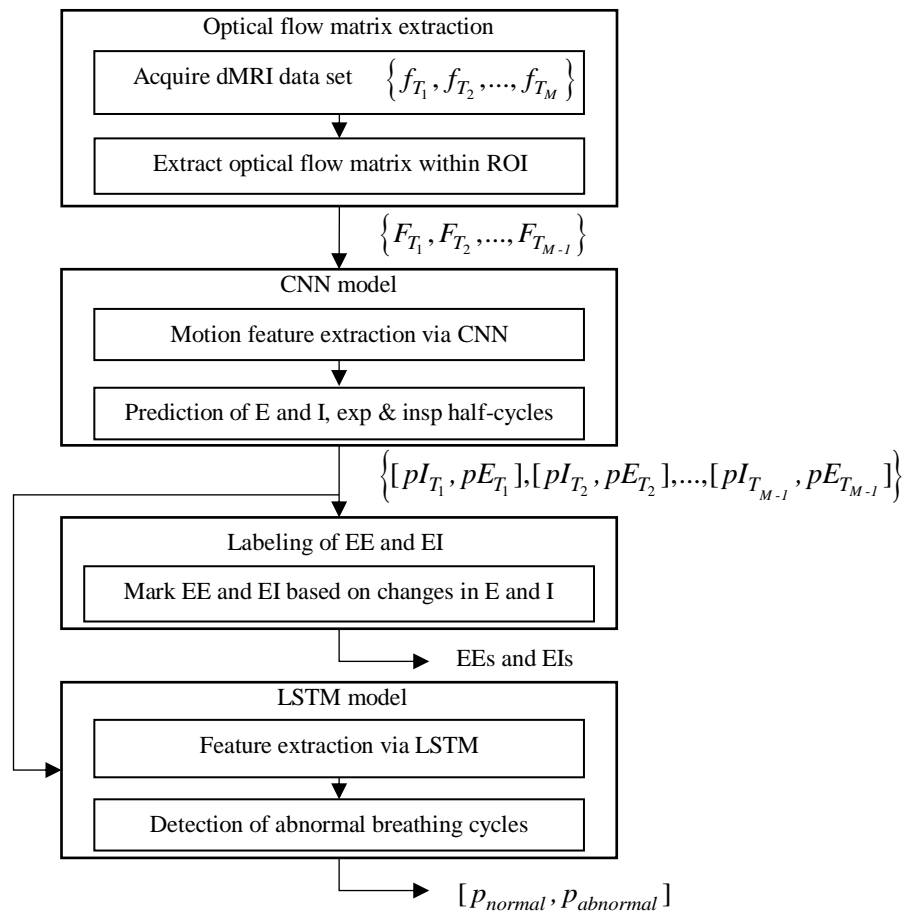

Figure 1. A schematic representation of the framework for automatic labeling of respiratory phases and detection of abnormal respiratory signals.

Motion feature extraction based on optical flow and CNN

We initially specify manually on one slice a region of interest (ROI) covering the right hemi-diaphragm. This specification is required on one sagittal location only per patient dMRI scan. For the left hemi-diaphragm and all other sagittal locations, the ROI is determined automatically based on this manually provided information. The reason for the ROI is to confine all processing to the ROI only since the hemi-diaphragms are the most indicative of the nature of the motion that takes place. First, we compute optical flow within the ROI. The optical flow image contains target motion information estimated from two adjacent time slices, which is used to train the CNN model to get the features of the expiratory-inspiratory motion. The process of selecting ROI and the architecture of CNN are shown in Figure 2.

After obtaining a sequence of 79 ROI optical flow images $F = \left\{ F_{T_1}, F_{T_2}, ..., F_{T_{M-1}} \right\}$ from the given time sequence of 80 slices $A = \left\{ f_{T_1}, f_{T_2}, ..., f_{T_M} \right\}$ at each sagittal location, the slice images in F are divided into two categories: E denoting the expiration phase, and I denoting the inspiration phase. The CNN is trained to classify the images in F into these two classes. Our network has 4 convolutional layers, 3 pooling layers, and 3 fully-connected layers. After each convolutional layer, a ReLU activation is used. We use a two-component softmax output layer to predict the motion of the slice under consideration. The output of the CNN is a series of two-dimensional vectors

$$P=\left\{[pI_{T_1},pE_{T_1}],[pI_{T_2},pE_{T_2}],...,[pI_{T_{M-1}},pE_{T_{M-1}}]\right\},\quad (1)$$

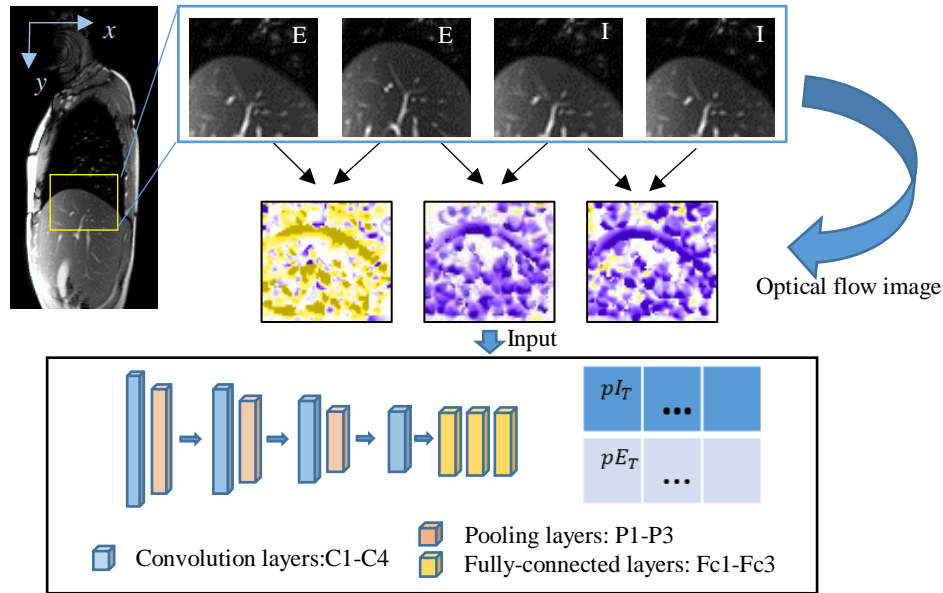characterizing the class probabilities for E and I such that $pI_{T_i} + pE_{T_i} = 1$.

Figure 2. Choosing an ROI is the only operation that requires human interaction. Thereafter, each optical flow image is calculated from every pair of two time slices that are adjacent in the time sequence as an input to the CNN.

Labeling of EE and EI phases by using motion features

From the sequence of probabilities *P*, we can detect EE and EI phases by using the following rules:

$$\text{EI: } pI_{T_i} > pE_{T_i} \text{ AND } pI_{T_{i+1}} < pE_{T_{i+1}} \qquad (2)$$

$$\text{EE: } pI_{T_i} < pE_{T_i} \text{ AND } pI_{T_{i+1}} > pE_{T_{i+1}}. \qquad (3)$$

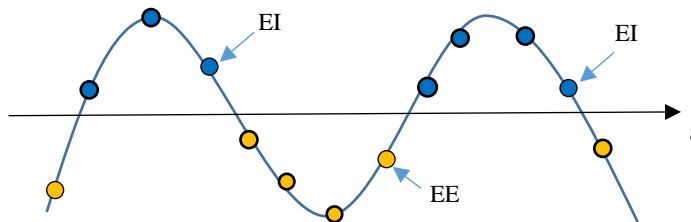Figure 3 illustrates this principle.

Figure 3. The EE and EI annotations for one breathing cycle. The blue and orange dots denote E phase points and I phase points, respectively.

Abnormal breathing signal detection based on LSTM

LSTM [4] is a network architecture for analyzing signals and is designed for sequence processing. Our idea is to use the probability sequence, which indirectly contains timing information, to train an LSTM network to predict those acquisitions

A that may contain an abnormal respiratory cycle. Another characteristic of LSTM networks is that they require far fewer training samples than other deep networks. To analyze the inter-time-slice information, we apply a sequential learning strategy by inputting a series of features extracted from the time-distributed CNN into our LSTM layer. Each series *P* of probabilities associated with each (sagittal) z-position is treated as one training sample for the LSTM network. Our LSTM layer has 64 units and the final output is two types of probabilities indicating the likelihood of the input slice sequence *A* containing/ not containing an abnormal cycle.

3. EXPERIMENTS AND RESULTS

By calculating the deviation in the time instance of each auto-labeling result from the closest ground truth marking of time points, the average of all results is used to describe the error in auto-labeling. We also compare the performance of the labeling method based only on optical flow (abbreviated as OF) [6] with the method of combined OF and CNN (abbreviated as CNN-OF). The combined results for labeling EE and EI are summarized in Table 1. The CNN-OF method is closer to the result of manual labeling than OF.

In the experiment of abnormal signal detection, we collected abnormal signals from 112 sagittal locations from 20 subjects. Similarly, we also collected normal signals from 112 sagittal locations. Based on these 224-position data sets, we performed 8-fold cross-validation. The mean accuracy of detecting abnormal cycles is 80.0%. The true positive rate and true negative rate are 79.5% and 80.3%, respectively. Although CNN-OF performs slightly better than direct OF method, it is an integrated strategy which together with LSTM performs also abnormal signal detection.

Table 1. Auto-labeling (EE+EI) error (mean±SD) in terms of number of time points of deviation from manual labeling.			
	Right lung	Left lung	Both
OF	0.23±0.16	0.35±0.28	0.29±0.19
CNN-OF	0.23±0.11	0.32±0.17	0.27±0.12

4. CONCLUSIONS

To improve the efficiency of 4D thoracic image construction and to identify slices that are not suitable for the construction process or that would mislead construction, we proposed a deep-learning-based method for automatic labeling of respiratory signals and detection of abnormal respiratory signals from dynamic thoracic MRI datasets. With this framework, we can perform labeling of breathing phases and the detection of abnormal respiratory signals with only minimal user interaction in the same single system. In our experimental results compared to a previously reported labeling method, the proposed method has improved accuracy considerably for slices in the vicinity of the heart. For the task of detecting abnormal signals, the true positive rate and true negative rate reached 79.5% and 80.3%, respectively. They are expected to improve after obtaining more data sets in the future. The proposed approach is useful for building a 4D image construction system and can be adapted to any other dynamic modality for automatic labeling of signals for 4D image construction purposes.

ACKNOWLEDGEMENT

REFERENCES

[1] Tong Y, Udupa JK, Wileyto PE, et al. Quantitative dynamic Lung MRI (QdMRI) volumetric study of pediatric patients with thoracic insufficiency syndrome, Radiology, 293 (7): 206-213, 2019.

[2] Tong Y, Udupa J K, Ciesielski K C, et al. Retrospective 4D MR image construction from free-breathing slice Acquisitions: A novel graph-based approach. Medical image analysis, 2017, 35: 345-359.

[3] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems, 2012: 1097-1105.

[4] Gers F A, Schmidhuber J, Cummins F. Learning to forget: Continual prediction with LSTM. 1999.

[5] Campbell RM Jr, M. D., Mayes, T. C., Mangos, J. A., et al. The characteristics of thoracic insufficiency syndrome associated with fused ribs and congenital scoliosis. Journal of Bone & Joint Surgery American Volume, 2003, 85-A(3): 399.

[6] Sun C, Udupa J K, Tong Y, et al. Auto-labeling of respiratory time points in free-breathing thoracic dynamic MR image acquisitions for 4D image construction. Medical Imaging 2019: Biomedical Applications in Molecular, Structural, and Functional Imaging. International Society for Optics and Photonics, 2019, 10953: 109531B.

[7] Campbell Jr R M, Smith M D. Thoracic insufficiency syndrome and exotic scoliosis. JBJS, 2007, 89: 108-122.

[8] Barron J L, Fleet D J, Beauchemin S. Performance of optical flow techniques. International journal of computer vision, 1994, 12(1): 43-77.