

Unobtrusive and Automatic Classification of Multiple People's Abnormal Respiratory Patterns in Real Time using Deep Neural Network and Depth Camera

Yunlu Wang, Menghan Hu, Yuwen Zhou, Qingli Li, *Senior Member, IEEE*, Nan Yao, Guangtao Zhai, *Senior Member, IEEE*, Xiao-Ping Zhang, *Senior Member, IEEE*, Xiaokang Yang, *Fellow, IEEE*

Abstract—Respiratory pattern is a representation of human breathing activity, which can reflect people's physical and psychological condition. Capturing the unexpected abnormal respiratory pattern unobtrusively of the patient or the potential patient has great significance. In the current work, we attempt to capitalize on depth camera and deep learning architecture to achieve the accurate and unobtrusive measurement of abnormal respiratory patterns, and the whole system can classify multiple people's respiratory patterns in a real-time manner. The challenges in this task are threefold: the real-time online system means that the ROI (region of interest) needs to be located and tracked automatically; the amount of real-world data is not enough for training to obtain the robust deep neural network; and the intra-class variation is large and the outer-class variation is small. Consequently, human joints tracking is applied to determine the location of subjects' shoulder and chest. Based on the characteristics of actual respiratory signals, a novel and efficient Respiratory Simulation Model (RSM) is proposed to generate abundant and high-quality training data. Finally, we apply a GRU neural network with bidirectional and attentional mechanisms (BI-AT-GRU) to classify 6 clinically significant respiratory patterns (Eupnea, Tachypnea, Bradypnea, Biots, Cheyne-Stokes and Central-Apnea). The performance of the obtained BI-AT-GRU is tested by the data that is actually measured by depth camera. Experimental results demonstrate that the proposed model can classify 6 different respiratory patterns with the accuracy, precision, recall and F1 of 94.5%, 94.4%, 95.1% and 94.8% respectively. In comparative experiments, the obtained BI-AT-GRU specific to respiratory pattern classification outperforms the existing state-of-the-art viz. BI-AT-LSTM, GRU and LSTM.

This work is sponsored by the Shanghai Sailing Program (No. 19YF1414100), the National Natural Science Foundation of China (No. 61831015, No. 61901172, No. 61975056), the STCSM (No.18DZ2270700), the Science and Technology Commission of Shanghai Municipality (No. 18511102500, No. 19511120100), and the foundation of Key Laboratory of Artificial Intelligence, Ministry of Education (No. AI2019002).

Y. Wang, M. Hu and Q. Li are with the Shanghai Key Laboratory of Multidimensional Information Processing, East China Normal University, Shanghai 200062, China.

Y. Wang, M. Hu, G. Zhai and X. Yang are also with the Key Laboratory of Artificial Intelligence, Ministry of Education, Shanghai, 200240, China.

Y. Zhou is with the SEU-Monash Joint Graduate School, Southeast University, Suzhou, 215123, China.

N. Yao is with the Shanghai Jianglai Data Technology Co., Ltd, Shanghai, 200241, China.

X.-P. Zhang is with the Department of Electrical, Computer and Biomedical Engineering, Ryerson University, Canada.

G. Zhai and X. Yang are also with the Institute of Image Communication and Information Processing, Shanghai Jiao Tong University, Shanghai 200240, China.

Corresponding authors: Menghan Hu (mhhu@ce.ecnu.edu.cn), Guangtao Zhai (zhaiguangtao@sjtu.edu.cn).

BI-AT-GRU. Moreover, other experimental results indicate that the proposed online measuring system, deep neural network and the modeling ideas have the potential to be extended to the large scale applications such as public places, sleep scenario, and office environment. The demo videos of the proposed system are available at: <https://doi.org/10.6084/m9.figshare.11493666.v1>

Index Terms—Remote monitoring, Breathing pattern, Physiological signal measurement, Deep learning architecture, Recurrent neural network, Attentional mechanism

I. INTRODUCTION

RESPIRATION is a core physiological process for all living creatures on Earth. For humans, respiration is a cyclic process of inhalation and exhalation: after inhaling oxygen-containing air, the respiratory system repeatedly performs gas exchanges across the alveolar-capillary membrane, with oxygen entering the bloodstream and carbon dioxide exiting through the mouth and nostrils [1]. We breathe every second. A person's physiological state [2]–[4] as well as mood and stress [5] may be reflected by representation of some respiratory parameters. Respiratory training is an indispensable daily care for patients with respiratory diseases [6]–[8]. Therefore, we should pay attention to our common and vital physiological process — respiration.

Respiratory rate, along with oxygen saturation, heart rate and blood pressure, is regarded as one of the four major physiological indicators [9]. Respiratory rate is an important respiratory sign, but it is not the only one. Other respiratory signs, namely, respiratory depth and rhythm and changes in chest movements, are also of great importance but often neglected [10], [11]. The above respiratory signs being comprehensively considered, the assessment of the respiratory activity continuing for a period of time (60 or more seconds) is actually equivalent to an examination on the respiratory pattern. 12 to 20 breaths per minute is normal for a healthy adult during quiet awake periods [10]. The normal respiratory pattern, called Eupnea, has normal respiratory depth, rhythm and chest movements [12], [13]. There are also a few abnormal respiratory patterns. It is noteworthy that these abnormal patterns are not unlikely to have more than just unusual respiratory rates. Biots' respiration is seen with periods of rapid respirations of nearly equal depth, which is regularly followed by periods of apnea where breathing terminates temporarily

for 15 seconds or longer [10]. Cheyne-Stokes is characterized by periods of respirations during which each breath gets progressively deeper and then decreasing, a decrease resulting in a period of apnea [14]. Central Apnea is an abnormal pattern that starts with normal breathing for a while and then enters a period of apnea [15]. Integrated various respiratory signs, respiratory patterns are able to more comprehensively reflect the conditions of respiratory activity. Many clinical literatures suggested that abnormal respiratory patterns are able to predict a few specific diseases, providing relatively detailed clues for clinical treatments [15]–[17]. For example, during the night, Cheyne-Stokes is an efficient indicator to judge the prognosis of Congestive Heart Failure (CHF) patients. The Cheyne-Stokes pattern, if detected in a CHF patient during the night, will provide a valuable reference for the patient's rehabilitation [14]. The respiratory patterns and their common symptoms are elaborated in **Table I**.

Unfortunately, these abnormal respiratory patterns occur in a way difficult for people to notice by themselves. For example, it is not easy for people to notice and evaluate their abnormal respiratory patterns, particularly during sleep. Unless the patient (or the potential patient) has been in a sleep monitoring room of a hospital, his or her abnormal respiratory pattern will not be easily detected in the progress of respiration. Thus, if we can develop a system capable of remotely and unobtrusively detecting these unnoticeable abnormal breathing patterns under various scenarios (hospitals, workplaces, homes, etc.), people who have the diseases may be diagnosed as early as possible!

Contact measurement devices constitute the majority of what is traditionally used for the assessment of respiratory activity [1]. Though having the virtue of accuracy, these devices are heavy and expensive, and in most cases need to be used in health facilities, causing pain as well as inconvenience to patients during their use [18]. For example, connecting and winding tubes and wires may be cumbersome for sleep monitoring. Besides, the detection of breathing patterns may be affected due to the prior and subjective knowledge, which can lead to a sense of "precognition". However, the best time to evaluate a patient's breathing pattern is when they are not aware that they are being examined [19].

From the above we can conclude that non-contact measurement methods are more suitable for detecting abnormal respiratory patterns. According to the principle of measurement, non-contact detection of breathing can be divided into three categories viz. radar doppler effect based method, thermal imaging technology and motion detection. The continual use of radar has the potential risk of the explosion to release radiation, and the detection of the respiratory patterns may be affected if a band adjacent to the band of the radar being used occurs nearby [20], [21]. Thermal imaging cameras are susceptible to ambient heat, and its ROI is forbidden to be covered by tubes, blankets, masks, etc [22]. Besides, the pixels of thermal imaging are inclined to saturate during deep breathing, in which case the measurement of respiratory depth will be affected, making it difficult to distinguish respiratory patterns with different respiratory amplitudes, such as Biots and Cheyne-Stokes [23]. The method based on motion detection is able to

solve the above problems, because its measurement is based on the variation in displacement and not on the changes in pixel intensity. Wiesner and Yaniv capitalized on the visible camera to track color markers placed on the subjects' abdomens, completing the estimation of respiratory rates [24]. Another idea to obtain respiratory signals based on motion detection is to take advantage of depth cameras because depth information of a scene can now be acquired by inexpensive cameras with the progress of sensing technology [25], [26]. Soleimani et al. constructed a three-dimensional model of human chest using the depth signals measured by the depth camera, and successfully extracted respiratory signals which they applied to pulmonary function tests [27]. Wijenayake and Park extracted human respiratory signals with the depth camera, and proposed a method to remove the spatial and temporal noise in deep data using PCA [28]. Bae et al. tracked the movement of markers placed on human chest and abdomen to acquire respiratory signals using Kinect camera [29]. Al-naji and Chahl exploited the Kinect's nature to be able to work in a dark environment for monitoring respiratory activity of patients who fell asleep, and realized the detection of cardiopulmonary abnormalities during sleep [30]. Being non-contact, unobtrusive, portable and low-cost, the respiratory measurement technology based on depth camera is capable of meeting the need to conduct respiratory measurements in various environments (working environment, sleeping environment and sports environment, etc.). Nonetheless, we found that most of the relevant work focused on improving the accuracy of the extraction of respiratory signals as well as restraining the noise in movement [31]. Some paid attention to the detection of abnormal respiratory patterns, but did not classify them specifically [32]. Thus, the available clinical clues were limited.

Real-time measurement and the automatic choice of suitable ROI are two significant challenges in non-contact measurement methods. Real-time measurement means that the data needs to be captured, processed and displayed frame by frame. The real-time respiratory monitoring system can catch abnormal respiratory patterns, and subsequently provide medical clues the first time. In non-contact measurement methods, the automatic choice of suitable ROI directly affects the quality of respiratory signals. In previous studies, the measurement methods based on thermal camera mostly take face, nose or mouth as ROI [23] [33] while the measurement methods based on motion detection mostly take thoracic and abdominal area as ROI. In addition, some studies pay attention to the automatic selection of ROI. For instance, Pereira et al. [34] divided thermal images into several boxes and made adaptive selection according to the signal quality of each box. Automatic ROI selection is an important prerequisite for the real-time online measurement system, although most studies based on motion detection ignore this.

The model for the classification of respiratory patterns based on contact measurement techniques performs relatively well because of the more waveform stability and less noise found in respiratory signals. De Chazal et al. put forward the automatic detection method based on electrocardiogram (ECG) signals capable of accurately detecting obstructive apnea [35]. Fekr et al. utilized a support vector machine to model the position

signals extracted by the acceleration sensor, completing the classification of the five abnormal breathing modes (Bradypnea, Tachypnea, Kussmaul, Biots, cheyne-strokes) [36]. As mentioned above, the non-contact measurement method is considered more suitable for detecting respiratory patterns. Its acquired respiratory signals are, nevertheless, rather inclined to be influenced by the environment and individual differences, leading to the possible inability of extending the features extracted by traditional methods to the complete data set. In addition, handicraft features often require careful engineering design and abundant professional knowledge. With few required manual operations, the methods based on deep learning are more applicable to complex and abstract sets of data when compared to those traditional methods [37]. Deep learning has thus been extensively used in fields like agricultural engineering [38], clinical medicine [39], [40], and, internet of things [41]. In terms of respiratory pattern detection, Cho et al. took advantage of the thermal imaging camera to extract the respiratory signals which would be converted into two-dimensional spectral sequences. Besides, they capitalized on the convolutional neural networks (CNN) to achieve the identification of deep breathing which represents the level of mental stress [42]. Kim et al. used 1D CNN to classify the respiratory signals estimated by radars into four categories (Eupnea, Bradypnea, Tachypnea and Apnea, respectively), producing good accuracy [43]. Therefore, the classification of respiratory signals extracted by the non-contact measurement system with the aid of deep learning is a study worth trying and of much significance. In the above study, all the data sets required for model construction were obtained, though, by assessing the respiratory activities of the test subjects. This approach for capturing different types of respiratory patterns yields a limited set of data. Were there no large amount of training data to support, the neural network might not be able to give full play to its advantages. Furthermore, researchers often present studies that the classification models generally adopt the general network architecture in the field of deep learning (such as CNN) without specific designs for respiratory pattern classification. At the same time, the network is not optimized according to the characteristics of the collected data.

In this paper, we capitalize on an improved GRU neural network that introduces bidirectional and attentional mechanisms (BI-AT-GRU) for respiratory pattern classification. With regard to the training data, we pioneer a simple but effective respiratory signal model for generating the simulated respiratory signals that are used to train the neural network. We utilized Microsoft Kinect depth camera to measure 6 respiratory patterns of 20 subjects (Eupnea, Bradypnea, Tachypnea, Biots, Cheyne-Stokes and Central-Apnea) to verify the feasibility of the network and the effectiveness of the training with simulation data. A total of 605 sets of data are collected to verify the performance of BI-AT-GRU, and the results are compared with the performance of general LSTM and general GRU network. In addition, the end-to-end model is used for a real-time system which can classify multiple people's respiratory patterns via engineering means.

The contributions of this work are as follows: 1) achieve the automatic ROI selection method based on human joint track-

ing; 2) the proposed RSM (Respiratory Simulation Model) can handle the gap between the large amount of training data and scarce real-world data; 3) the deep learning architecture we used viz. BI-AT-GRU performs outstandingly in two tasks (identification and classification) and two window lengths (30 seconds and 60 seconds); and 4) the proposed method can display multi-people's respiratory waveforms in real time, and subsequently identify and classify abnormal breathing patterns automatically, which has a broad application prospect.

II. PIPELINES OF PROPOSED METHOD

Previous researches tend to set up an offline system to process the collected respiratory data, and they seldom pay attention to the situation of multiple people. One of the most difficult challenges for multi-person and real-time measurement is to implement automated ROI selection. We propose an online system by combining the deep learning architecture and the engineering method to identify and classify multiple people's respiratory patterns in a real-time manner. The core function of the deep learning architecture is to allow the online system to accurately classify respiratory patterns, and the engineering method is to mainly achieve the low-latency communication among modules.

The pipelines of the proposed method are shown in Fig. 1. We obtain depth image sequences using Microsoft Kinect, and each frame of the depth video is transmitted to the server through the USB 3.0 cable. In the server, three key points of chest and shoulder area are tracked, and the ROI of each person is subsequently drawn with these vertexes. The average depth value in the ROI is calculated, and the value of all frames constitutes the 1D respiratory signal. This signal is displayed in real time by a Python plotting library viz. Matplotlib [44]. Finally, the classification results can be obtained by inputting these depth values into the carefully prepared end-to-end model. The respiratory waveform is displayed in real time and the classification results are given every 300 or every 600 points.

The whole system can be divided into three parts: acquisition of depth images, the processing of data and the determination of results. These modules can work independently on different devices: the acquisition of depth images can be completed by a depth camera; the processing of depth images can be completed in PC or server with an end-to-end model trained by stimulated data; and the respiratory waveform and the diagnosis of each person can be displayed on the terminal.

Demo videos of this method working in situations of one subject and two subjects can be downloaded online (<https://doi.org/10.6084/m9.figshare.11493666.v1>). In these two videos, we record depth images with Microsoft Kinect, and the calculation and visualization module both work in the same PC.

In this paper, section III will introduce how we measure respiration in a real-time manner by the depth camera; section IV will introduce the proposed respiratory pattern classification model; section V will introduce experimental process and show the results; section VI is discussion; and section VII is conclusion.

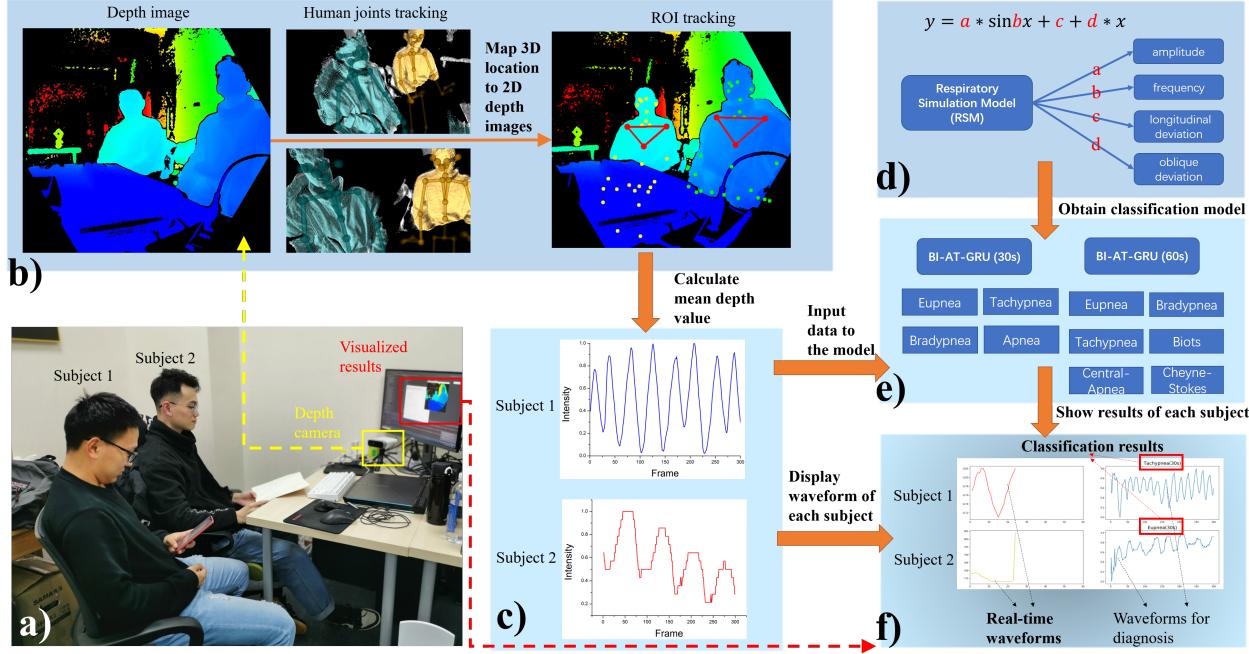


Fig. 1: Pipelines of the proposed method. a) The depth camera (yellow box) records depth images, and the respiratory waveforms and patterns of each subject are shown on the screen (red box). b) In PC or server, we detect and track human joints and map these point cloud data to 2D depth images to track ROIs (red triangles). c) The respiratory waveforms are obtained by calculating the mean depth value of each ROI, and they are displayed and processed in a real-time manner. When there are 300 or 600 data points, the well-trained BI-AT-GRU will seize them to export classification results. d) We utilize the proposed RSM to generate abundant and reliable training data. The RSM takes various features of the real-world respiratory waveform into account, and the training process is finished offline and the obtained classifier is connected to the online system. e) The end-to-end model trained by data generated by RSM is used to detect and classify abnormal respiratory patterns. The lengths of detection window have 30s and 60s in this paper. BI-AT-GRU model of 30s detection window will classify signal into 4 patterns while 60s will classify them into 6 patterns. f) In visualized results module, the respiratory waveforms are displayed in real time and the diagnosis of each subject is shown every 30 seconds and 60 seconds.

III. REAL-TIME MEASUREMENT OF HUMAN RESPIRATION BY DEPTH CAMERA

In this paper, we first acquire the raw depth image sequences via the depth camera, and then the ROI is located using automatic and manual methods. The automatic ROI location method is developed based on the human joint tracking, and this is an important guarantee for real-time measurement system. The actually measured respiratory signals are afterwards extracted to test the performance of the obtained BI-AT-GRU model. In order to analyze the impact of different ROIs on the classification results, we also made manual ROI selection in the experimental stage.

A. Acquisition of depth image sequences

Kinect v2 (Microsoft released in 2014) and Azure Kinect (Microsoft released in 2019) were used to record depth images of subjects when they breathed. They were both equipped with a depth camera based on Time of Flight (TOF) technology. A total of 20 participants were asked to sit in a chair and learn to imitate 6 respiratory patterns. Spirometer of a sleep monitor (GY-6620, HeNan HuaNan Medical Science and Technology Co., LTD.) was used to check their respiratory situations to

make sure they had learned the pattern (in practice, some subjects confused certain respiratory patterns). We excluded the actually measured respiratory data that was inconsistent with these obtained by the gold standard method for the subsequent modeling procedure. Subjects were at 0.5m-3m from the depth camera and were asked to breathe in a specific respiratory pattern for one minute at a time.

B. ROI selection and tracking

Breathing habits vary from person to person: when they breathe, the chest of some people changes more; the abdomen of some changes more and the shoulders of some shake more. Therefore, in depth images, we selected three ROIs, namely chest, abdomen and shoulder. **Fig. 2** shows actually measured Central-Apnea waveforms of one subject specific to these three ROIs. We use manually selected ROI for the modeling to ensure the accuracy and validity of the respiratory data, and we analyze these data from different ROIs to eliminate the impact of automated selection.

However, the manually selected ROI has two defects in the online system: 1) manual ROI selection is required before each use, which causes inconvenience to users; and 2) the original ROI may not cover the effective areas when the user moves.

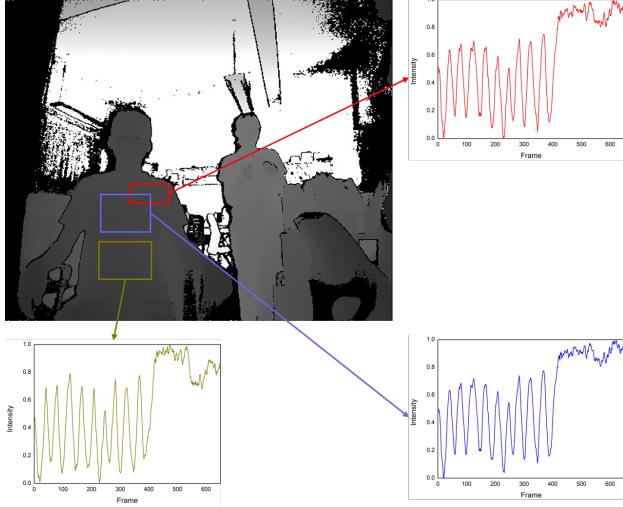


Fig. 2: Actually measured Central-Apnea waveforms of one subject specific to three ROIs. They are extracted from the manually selected ROIs. X axis is frame and Y axis is intensity.

Therefore, manually selected ROI is not appropriate for a real-time respiratory measurement system in practical applications.

We tested respiratory pattern classification model on three different manually selected ROIs, and results indicated that chest and shoulder performed relatively well (shown in section V). If these two areas can be located automatically, ROI can be obtained.

Microsoft Azure Kinect provides API (Application Programming Interface) for body tracking, which returns every subject's 3D position of each joint. Afterwards, we map 3D position of required joints viz. shoulders (left and right) and chest to the 2D depth image. This step can be represented by the following equations:

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \begin{bmatrix} R & T \\ 0^T & 1 \end{bmatrix} * \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (1)$$

$$Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{dx} & 0 & u_0 \\ 0 & \frac{1}{dy} & v_0 \\ 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} fX_c \\ fY_c \\ Z_c \end{bmatrix} \quad (2)$$

where X_w , Y_w and Z_w represent 3D position in the real world; R (3×3) and T (3×1) represent the camera calibration parameter; X_c , Y_c and Z_c represent the camera coordinate; f represents focal length; dx and dy represent the length (mm) per row and column in depth images, u_0 and v_0 represent the difference between the principal point and the camera coordinate. We map u and v to the depth images to obtain 2D position of required joints.

The triangle drawn by these three vertices is target area viz. ROI. When the subject is 0.5m-3m away from the camera, ROI can be tracked automatically, regardless of subjects' movements. Fig. 3 shows that all human joints can be accurately tracked in real time under different situations. It illustrates that ROI can be automatically selected regardless of the posture of

a single subject or multiple subjects (even if part of the body is concealed).

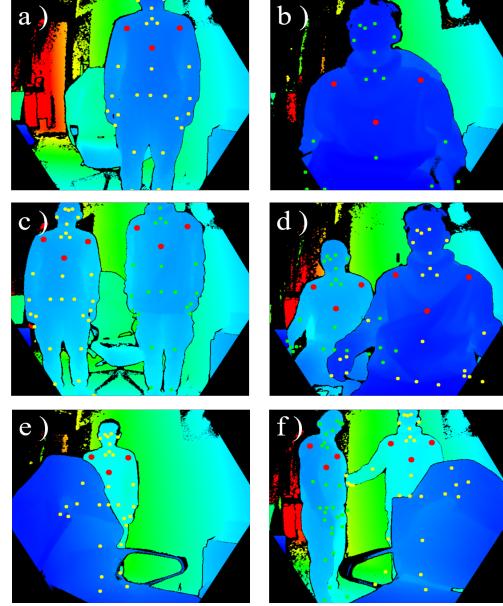


Fig. 3: Human joints detection and ROIs tracking under different situations. Red points are three vertices of ROI. a) One subject stands. b) One subject sits. c) Two subjects stand. d) Two subjects sit. e) One subject's body is partially covered by a chair. f) One subject's body is partially covered by a chair and another subject stands sideways.

C. Depth data processing

By calculating the average value of depth data in certain ROI under each frame, we preliminarily extracted respiratory signals. This process can be completed through the following equation:

$$\bar{D}(k) = \frac{1}{n} \sum_{i,j \in N} D(i,j,k) \quad (3)$$

where $D(i,j,k)$ represents the depth data at the pixel (i,j) in frame k ; N represents the pixel coordinate vector of each ROI and n represents the number of pixels in a certain ROI. All frames of $\bar{D}(k)$ were smoothed by a moving average filter with a data span of 5 to eliminate sudden changes of waveform.

Finally, min-max normalization was carried out on respiratory signals:

$$x^*(k) = \frac{x(k) - \min[x(k)]}{\max[x(k)] - \min[x(k)]} \quad (4)$$

IV. RESPIRATORY PATTERN CLASSIFICATION MODEL

The proposed respiratory pattern classification model consists of five parts: 1) develop respiratory simulation model and generate simulated respiratory pattern data; 2) acquire the actually measured respiratory pattern data using depth camera; 3) establish BI-AT-GRU model via simulated respiratory pattern data and validate it using the actually measured data; 4)

TABLE I: Respiratory situation and common symptom of 6 respiratory patterns.

Respiratory Pattern	Respiratory Situation	Common Symptom
Eupnea	Normal breathing	Normal
Bradypnea	Slow breathing	Over fatigue
Tachypnea	Rapid breathing	Acute respiratory distress
Biots	Deep and fast breathing, then choking	Meningitis
Cheyne-Stokes	Shallow-deep-shallow breathing, then choking	Congestive heart failure
Central-Apnea	Normal breathing, then choking	Primary hypoventilation syndrome

conduct the comparative experiments; and 5) develop a real-time respiratory pattern detection device by combining our model with the imaging system.

Deep learning is usually supported by a large amount of data. However, respiratory pattern data is highly heterogeneous and no relevant data sets have been released by researchers, so it is extremely difficult, or even impossible to utilize the transfer learning approach. Thus, based on plentiful preliminary research work, a simple but innovative Respiratory Simulation Model (RSM) is proposed by us to generate simulated respiratory pattern data for training the end-to-end model. Considering characteristics of waveforms belonging to different respiratory patterns, we add bidirectional and attentional mechanisms into GRU model. We make corresponding modifications to BI-AT-GRU which is adopted in the area of natural language processing and first apply it to classify respiratory patterns. The actually measured respiratory pattern data measured by depth camera was used for testing the performance of BI-AT-GRU.

Section IV will introduce respiratory patterns, generation of simulated data and BI-AT-GRU architecture.

A. Respiratory patterns

Respiratory pattern is defined by a series of signs of human respiratory process, such as respiratory frequency, respiratory depth, respiratory rhythm and changes in chest movements. Respiratory patterns can be divided into normal breathing (Eupnea) and abnormal breathing. People with Eupnea take about 12-20 breaths a minute with normal respiratory depth and rhythm as well as chest movement. Abnormal breathing is manifested as the abnormality of one or more of the above-mentioned indicators. In general, abnormal respiratory patterns indicate abnormalities in physiology, mood and stress. It can be used as clues and references for clinical diagnosis. Respiratory patterns researched in this paper include Eupnea, Bradypnea, Tachypnea, Biots, Cheyne-Stokes and Central-Apnea. The respiratory situation and common symptoms of these six respiratory patterns are summarized in **Table I** and their standard waveforms are shown in **Fig. 4**.

B. Respiratory Simulation Model (RSM)

Respiration is a cyclic process of inhalation and exhalation, which is reflected in the rise and fall of the waveform. Thus, respiratory signals measured by non-contact method can be approximated by sine wave. Actually measured respiratory signals, especially those measured by non-contact method, are

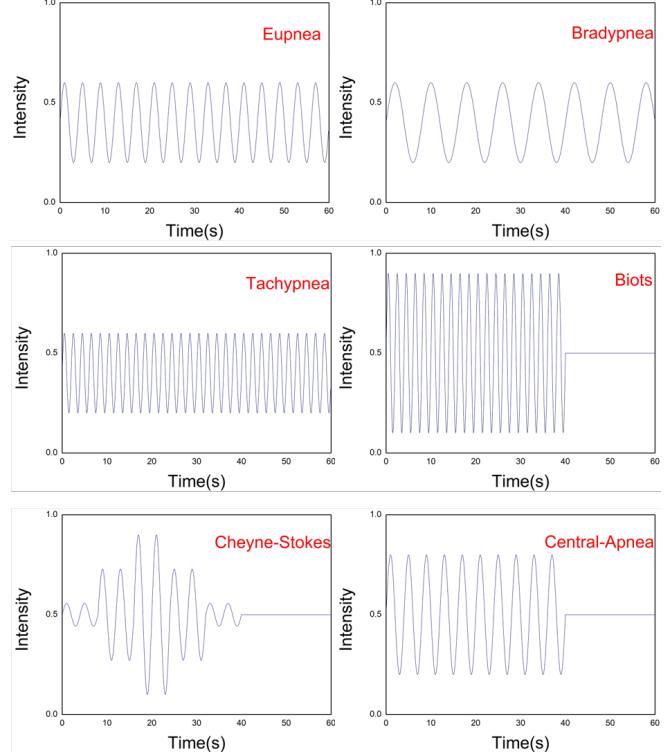


Fig. 4: Standard waveforms of 6 respiratory patterns. X axis is time and Y axis is signal intensity from 0 to 1.0. Notice that the time window is 60 seconds.

prone to deviation due to environmental changes, leading to fluctuations in respiratory depth and frequency within a certain range. Affected by body movements in measurements, signals also tend to have longitudinal and oblique deviation. Considering the possible deviation mentioned above, the actually measured respiratory signals can be defined by the following equation:

$$y_{i \in \omega} = a_i \sin(b_i x) + c_i + d_i x \quad (5)$$

where ω is a constant time window, indicating the cycle of examining a respiratory pattern; i is a variable time, which is used to represent changes of signals belonging to the same respiratory pattern; the point dividing each period of i can be called a ‘breakpoint’; a_i is respiratory depth; b_i is respiratory rate; c_i and d_i are longitudinal deviation degree and oblique deviation degree of respiratory signal, respectively.

Gaussian white noise with SNR (Signal Noise Ratio) is added to the simulated respiratory signal to render it closer to

TABLE II: Parameter setting of each respiratory pattern in RSM.

Respiratory Pattern	Number of breakpoints	a_i	b_i	c_i	d_i	Apnea time (s)	SNR
Eupnea	10~15	0.2~0.6	1.37~1.77	-0.1~0.1	-0.1~0.1	0	20
Bradypnea	5~10	0.3~0.6	0.79~1.15	-0.1~0.1	-0.1~0.1	0	20
Tachypnea	10~15	0.3~0.6	2.09~3.14	-0.2~0.2	-0.1~0.1	0	20
Biots	5~10	0.6~0.9	2.25~4.00	-0.05~0.05	-0.1~0.1	15~20	20
Cheyne-Stokes	5~10	0.1~0.3	1.37~1.57	-0.05~0.05	-0.1~0.1	15~20	20
Central-Apnea	5~10	0.3~0.5	1.37~1.57	-0.05~0.05	-0.1~0.1	15~20	20

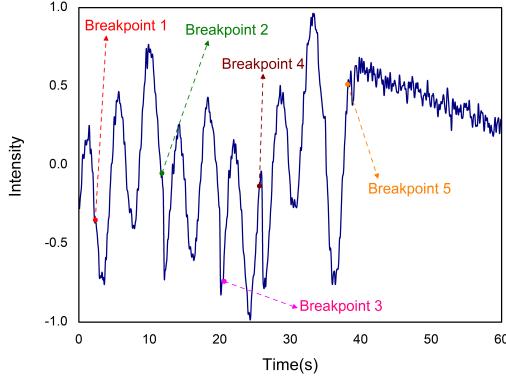


Fig. 5: Process of generating a Central-Apnea waveform via RSM. This waveform is randomly generated by proposed simulation model. The breathing phase before Breakpoint 5 is normal breathing, and after is choking for about 20 seconds.

the actually measured respiratory signal. The *awgn* (Additive White Gaussian Noise) function in MATLAB (The Mathworks, Inc., Natick, MA, USA) is used:

$$y' = awgn(y, SNR, 'measured') \quad (6)$$

Different features in the process of respiration can be simulated such as speed, depth and time of apnea by changing values of a_i , b_i , c_i and d_i in (5). A specific partial waveform can be generated by randomizing the parameters of each waveform in a pre-set range. We combine every partial waveform through breakpoints then specific respiratory patterns can be obtained. To illustrate the process of generating data and the effectiveness of RSM, we select an abnormal respiratory pattern to demonstrate. **Fig. 5** shows the process of generating a Central-Apnea waveform.

The random range of parameters was set for the 6 respiratory patterns by reference to standard waveforms (**Fig. 4**) and waveforms measured by depth camera. **Table II** shows the parameter settings of RSM. We release the code for researchers interested in this issue so they can reproduce this process (<https://doi.org/10.6084/m9.figshare.9978833.v1>).

C. BI-AT-GRU for respiratory patterns classification

Bidirectional and attentional mechanisms were added to the traditional GRU, thus building BI-AT-GRU framework according to characteristics of our task. We first applied BI-AT-GRU which is traditionally adopted in natural language

processing to classify respiratory patterns. In this paper, BI-AT-GRU was trained by simulated data generated by RSM and tested by the data measured by the depth camera.

Both the simulated respiratory data and the actually measured respiratory data can be regarded as time series data, a kind of sequential data. Recurrent Neural Network (RNN) is a Neural Network model proposed by Elman [45] in 1990, which is very suitable for sequential data modeling. RNN model can be represented by the following equations:

$$h_m = f(Wx_m + US_{m-1}) \quad (7)$$

$$o_m = softmax(VS_m) \quad (8)$$

where m represents m th step; h_m represents hidden layer; x_m represents input; o_m represents output and S_m represents the state of m th step. f is a non-linear activation function and W, U, V are parameters obtained by training procedure.

Long-Short Term Memory (LSTM) is an important variant network of RNN. It was proposed by Hochreiter and Schmidhuber [46] in 1997 to solve the problem of long training time of RNN and long-term memory loss in long sequences. Gated Recurrent Unit (GRU) proposed by Cho et al. [47] in 2014 is a simplified variant of LSTM. GRU merges three doors of LSTM (forget, input and output) into two doors (update and reset). The structure of GRU can be expressed by the following equations:

$$h_t = (1 - z_t) * h_{t-1} + z_t h'_t \quad (9)$$

$$h'_t = \tanh[r_t * (Uh_{t-1}) + Wh_t + b_h] \quad (10)$$

where h_t is the hidden layer of t step; h_{t-1} is the hidden layer of $t-1$ step; h'_t is current new state of t step which can be regarded as a summary of the above information h_{t-1} and the input information of t step x_t . U, W, b are the parameters obtained by training procedure.

GRU uses update gate and reset gate to control the retention, forgetting and updating of sequence information. z_t is the update gate which determines how much past information is forgotten and how much new information is added. r_t is the reset gate which determines how much of the above information is passed to the new state. z_t and r_t can be expressed by the following equations:

$$z_t = \sigma(W_z x_t + U_z h_{t-1} + b_z) \quad (11)$$

$$r_t = \sigma(W_r x_t + U_r h_{t-1} + b_r) \quad (12)$$

where x_t represents the input at t step and U, W, b are parameters obtained by training procedure. Both LSTM and

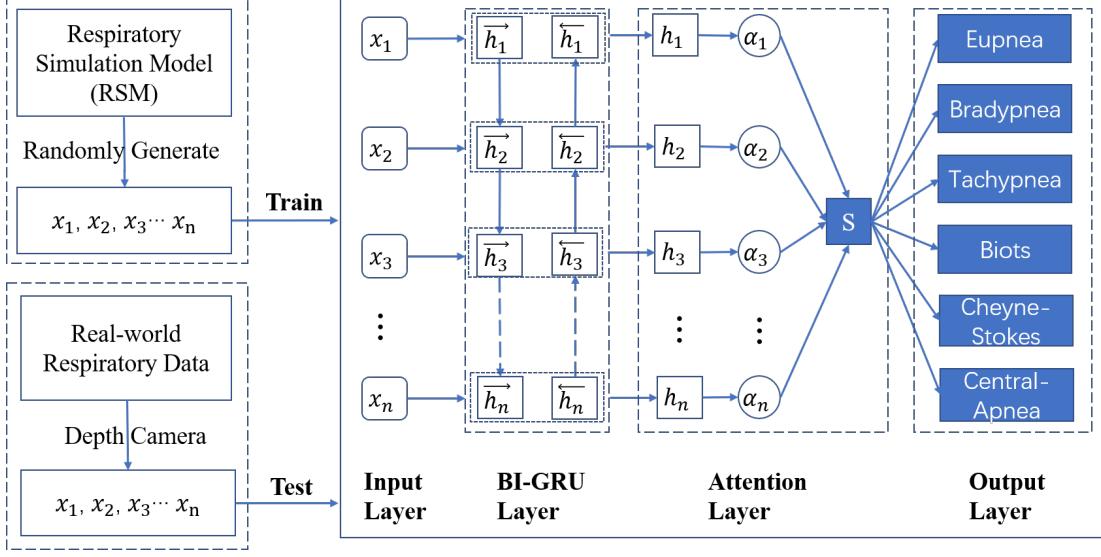


Fig. 6: BI-AT-GRU for respiratory pattern classification. This network consists of four layers: Input Layer, BI-GRU Layer, Attention Layer and Output Layer. It is trained by simulated data randomly generated by RSM and tested by the data measured by the depth camera.

GRU can solve the problem of gradient disappearance of RNN. Thus, they have been widely applied in sequence data modeling. Compared with LSTM, GRU has fewer parameters and simpler structure, so it is easier to converge when the amount of data is small.

We utilized an improved GRU viz. BI-AT-GRU for respiratory pattern classification by adding bidirectional and attentional mechanisms to GRU network. The network architecture is shown in Fig. 6. The whole network is divided into input layer, BI-GRU layer, attention layer and output layer. The input layer is used to input simulated data (training stage) or depth data (testing stage) $x_1, x_2, x_3, \dots, x_n$ at each point of respiratory waveform.

By observing and studying different respiratory pattern waveforms measured by the depth camera, we found two particular characteristics: 1) compared with left to right observation of respiratory waveform, we can acquire more information when we observe the signals in reverse chronological order (right to left observation); and 2) there are key “turning points” in some combined respiratory patterns such as Central-Apnea which consists of normal breathing and apnea. Those turning points can provide reliable evidence for judging respiratory patterns. Based on these two findings, we added bidirection and attention into GRU, respectively corresponding to BI-GRU layer and attention layer in BI-AT-GRU.

Bidirection: Bidirectional RNN (BI-RNN) was proposed by Bahdanau et al. [48] and used in text translation. Inspired by this work, we added a bidirectional mechanism to GRU to obtain forward sequence information \vec{h}_t from beginning to end of breathing process and backward sequence information \overleftarrow{h}_t from end to beginning. This can be expressed by following equations:

$$\vec{h}_t = \overrightarrow{GRU}(x_t), \quad t \in [1, n] \quad (13)$$

$$\overleftarrow{h}_t = \overleftarrow{GRU}(x_t), \quad t \in [n, 1] \quad (14)$$

$$h_t = [\vec{h}_t, \overleftarrow{h}_t] \quad (15)$$

where x_t is input simulated data generated by RSM (in training stage) or input depth data (in testing stage) at time t ; n represents length of the time series data, which is the period of a respiratory pattern.

Attention: Attention RNN has been widely applied in document classification [49], speech recognition [50] and relation classification [51]. In above studies, the attention mechanism is used to express the importance of words in sentences or sentences in documents. In our task—classification of respiratory patterns, each point in the respiratory waveform is of different importance to determining certain respiratory pattern. Therefore, we added attentional mechanism into GRU, which can be expressed by following equations:

$$u_t = \tanh(W_a h_t + b_a) \quad (16)$$

$$\alpha_t = \text{softmax}(V_a u_t) \quad (17)$$

$$S = \sum_t \alpha_t h_t \quad (18)$$

where h_t is the state of t step; W_a and b_a are the parameters obtained by training phase. First, a \tanh function is used to obtain the hidden representation u_t . Then normalized importance weight α_t of each point in the waveform is obtained through a softmax function. V_a is also a parameter obtained by training processing and it can be understood as the importance of data point to certain respiration pattern. Finally, the sum of product of α_t and h_t at each point is calculated to obtain the output of the attention layer, which is denoted as S . S here is high-level representation of certain waveform for classification.

Output layer uses the output S of the attention layer to classify respiratory patterns:

$$p = \text{softmax}(W_c S + b_c) \quad (19)$$

$$\hat{y} = \underset{y}{\operatorname{argmax}}(p) \quad (20)$$

where S represents the output of attention layer; y represents the label of each respiratory pattern and \hat{y} represents the predicted label by BI-AT-GRU.

In training stage, we encoded the label in one-hot format and cross-entropy was adopted as loss function:

$$L = - \sum_i y^i \log(p^i) \quad (21)$$

where y^i is the i th value of the label, and p^i is the i th normalized value output by softmax function.

V. EXPERIMENTS AND RESULTS

A. Experimental settings

The size of original training set was 120,000 including 6 respiratory patterns. Each pattern had 20,000 sets of data which were generated randomly by the proposed RSM. Hidden layers, attention hidden layers and batch size were 128, 16, 128, respectively. To get the actually measured data of 6 respiratory patterns, we measured the respiratory signal of 20 subjects (12 females and 8 males). They are students from East China Normal University. Subjects sat in a relaxed manner and imitated every respiratory patterns for one minute, with a one-minute break between each pattern. The depth camera was at 0.5m-3.0m from the subjects and each subject breathed 1-3 groups. The size of the original test data was 605. Among them, there were 108 groups of Eupnea, Bradypnea and Tachypnea; 97 groups of Cheyne-strokes and Central-Apnea and 87 groups of Biots. Some wrong respiratory data caused by subject's misunderstanding of respiratory patterns were eliminated for modelling.

B. Testing on BI-AT-GRU and other models

We trained BI-AT-GRU, BI-AT-LSTM, GRU and LSTM with the same training set (120,000 samples). Accuracy and loss curve of BI-AT-GRU in training phase is demonstrated in **Fig. 7**.

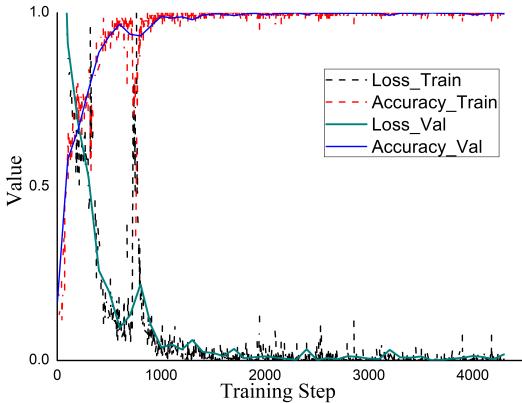


Fig. 7: Accuracy and loss curve in training phase of BI-AT-GRU. X axis is training step and Y axis is value. Notice that the validation sets were also generated by Respiratory Simulation Model (RSM).

Subsequently, we tested the performance of the above models on the same test set (605 samples) with four evaluation metrics viz. Accuracy, Precision, Recall and F1:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (22)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (23)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (24)$$

$$F1 = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (25)$$

where TP represents True Positive; TN represents True Negative; FP represents False Positive and FN represents False Negative.

Results are demonstrated in **Table III**. It can be seen from **Table III** that: 1) the four metrics of BI-AT-GRU are higher than other networks, and networks with bidirectional and attentional mechanisms perform better than their basic networks, which indicates that our modifications on GRU are successful; 2) GRU based networks perform slightly better than LSTM based in our task; and 3) the classification accuracy of all networks maintains a high level, which proves the feasibility of training the network with RSM data.

TABLE III: Test results on the actually measured data.

Model	Accuracy	Precision	Recall	F1
BI-AT-GRU	94.5%	94.4%	95.1%	94.8%
BI-AT-LSTM	90.1%	90.1%	91.9%	91.0%
GRU	89.6%	89.2%	91.1%	90.1%
LSTM	88.1%	87.8%	91.3%	89.5%

To figure out the performance of the developed classification models for each respiratory pattern, we plotted the confusion matrix of each model (**Fig. 8**). In terms of four models, the classification error mainly came from prediction of Cheyne-Stokes to be Central-Apnea. The possible reason is that these patterns are only different in the depth of breathing, which reflected in amplitude of the waveform. While in normalization processing, amplitude of respiratory signal is sensitive to the time window. In addition, the movement of subject's body can also lead to mutations of amplitude, thus increasing the error rate. It can be seen that BI-AT-GRU has the lowest error rate on this issue, which is an important reason why it performs best. Moreover, the performance of GRU and LSTM with bidirection and attention is better than their basic networks on this issue, which also indicates the contribution of these two modules.

C. Performance of BI-AT-GRU specific to different ROIs

To figure out how differently ROIs affect classification results, we divided the original test set into three groups (194 sets for abdomen, chest and shoulder respectively) and tested their performance on BI-AT-GRU (**Table IV**).

We find that performance of chest and shoulder is better than abdomen. Furthermore, when subject sits, the abdomen is

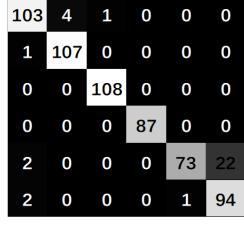
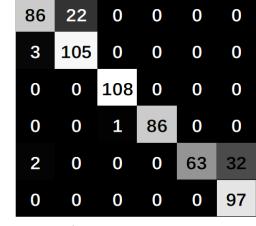
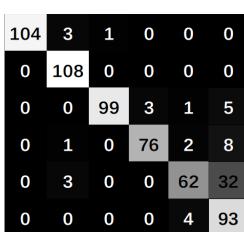
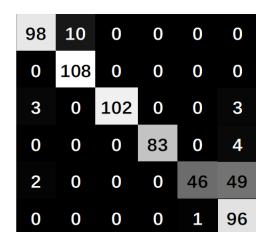
	
a) BI-AT-GRU	b) BI-AT-LSTM
	
c) GRU	d) LSTM

Fig. 8: Confusion matrix of 4 models. Each row is the number of real labels and each column is the number of predicted labels. From left to right or from top to bottom is: Eupnea, Bradypnea, Tachypnea, Biots, Cheyne-Stokes and Central-Apnea.

TABLE IV: Performance of BI-AT-GRU specific to three different ROIs.

ROI	Accuracy	Precision	Recall	F1
Abdomen	92.3%	91.6%	93.1%	92.3%
Chest	96.4%	96.4%	96.7%	96.5%
Shoulder	96.9%	96.9%	97.2%	97.1%

sometimes covered. Therefore, we combine chest and shoulder into one ROI and track it automatically using human joints recognition in our real-time system.

D. Testing on different Window lengths

Different lengths of time windows also affect performance of classifiers. Typically, it takes more than one minute to evaluate breathing patterns because some patterns are defined as a combination of two respiratory patterns. Biots, for example, is a combination of Tachypnea and Apnea.

TABLE V: Test results of BI-AT-GRU under two window lengths. Notice that BI-AT-GRU of 30s detection window has four patterns while 60s BI-AT-GRU has six patterns.

Window Length (s)	Accuracy	Precision	Recall	F1
30	96.5%	96.4%	96.3%	96.4%
60	94.5%	94.4%	95.1%	94.8%

To verify whether a shorter window can classify some simple respiratory patterns, we retrained BI-AT-GRU for four patterns classification (Eupnea, Bradypnea, Tachypnea and Apnea). The training set and test set are obtained respectively by transforming the original data set, and the operation is as follows: Eupnea, Bradypnea and Tachypnea waveforms of 60s

were divided into the front and back sections as two samples; and the last 30s of Biots, Cheyne-Stokes and Central-Apnea waveforms (asphyxia part) were extracted as Apnea. Since time of Apnea may be less than 30s in the actually measured data of above-mentioned respiratory patterns, the extracted waveform may contain some interference. Test results are demonstrated in **Table V**.

Table V illustrates that 30s BI-AT-GRU also performs well in task of four patterns classification. Therefore, combining 30s BI-AT-GRU for a preliminary judgment and 60s BI-AT-GRU for a detailed classification is a feasible scheme in our real-time system.

E. Identification and classification under long window length

In above experiments, we focus on respiratory patterns of short time window e.g. 30s or 60s. However, abnormal respiratory patterns usually occur only once in a few minutes. Therefore, identifying and classifying abnormal respiratory patterns during long periods of normal breathing is closer to real-world applications.

We set the time window length as 10 minutes. In normal breathing group, we randomly mixed 10 normal breathing waveforms viz. Eupnea from the original test data; and in abnormal breathing group, we randomly mixed 9 Eupnea and one abnormal respiratory pattern from the original test data, and the order is also random. The test set size was 50 for each group.

We applied a 60s window to slide on these long-time waveforms to extract data and the stride was 30s. Subsequently, all data from the same long-time waveforms were input to the end-to-end model to get a set of results. For the identification task, if there is one or more predicted abnormal pattern in long time window, we define it as abnormal respiratory pattern; and for the classification task, we take the predicted label (except Eupnea, if there exist only Eupnea labels, obviously it is normal breathing) that appears the most as the classification result.

Test results of these two tasks in long-time window are shown in **Table VI**. In identification task, GRU outperforms other models and F1 score reaches 87.0%. In addition, all models yield high recall, which suggests that our models have a high sensitivity to abnormal breathing patterns. However, low precision means our models tend to mistake normal breathing for abnormal breathing. In classification task, BI-AT-GRU outperforms other models and F1 score reaches 72.5%. Performance declines in all models because it is more difficult to classify than to identify abnormal respiratory patterns.

F. Analysis of results under long windows

The proposed method performs outstandingly in short window length situation. However, when it comes to the synthetic long window waveform, the evaluation indexes drop obviously. There are two possible reasons: 1) In the real-time system, it takes certain time to locate the human joints of each frame depth image and automatically select ROI (especially in the situation of multiple people), and this brings down the sampling rate. We set the frame rate of depth camera as 15

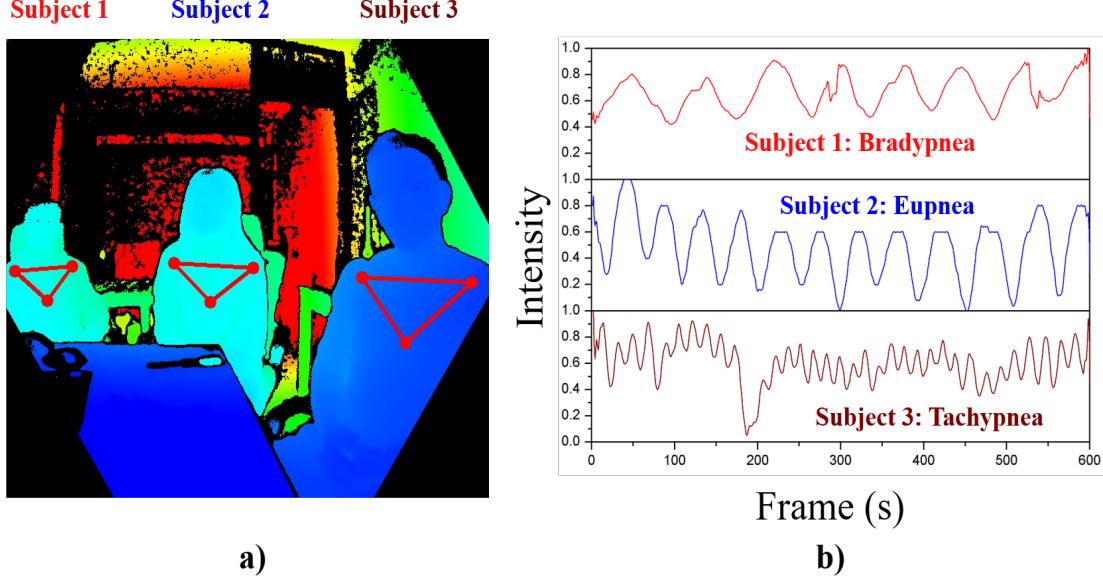


Fig. 9: Example waveforms and classification results from three subjects breathing at the same time. a) Three subjects breathe at the same time and their ROIs can be tracked at the same time. b) Example waveforms and their classification results of three subjects in one minute. 30s BI-AT-GRU and 60s BI-AT-GRU give the same results in this example because subjects' respiratory patterns do not change in one minute. Notice that we resample the signal based on linear interpolation because the frame rate declines in the situation of multiple people.

TABLE VI: Performance of two tasks in 10 minutes. Notice that the test set was spliced from single actually measured data.

Task	Model	Accuracy	Precision	Recall	F1
Identify	BI-AT-GRU	79.0%	70.4%	100.0%	82.6%
	BI-AT-LSTM	71.0%	63.7%	98.0%	77.2%
	GRU	85.0%	77.0%	100%	87.0%
	LSTM	67.0%	60.2%	100%	75.2%
Classify	BI-AT-GRU	65.0%	70.0%	75.5%	72.5%
	BI-AT-LSTM	50.0%	54.0%	75.3%	62.9%
	GRU	65.0%	61.7%	76.0%	68.1%
	LSTM	43.0%	49.0%	76.4%	59.7%

fps. However, compared with the spirometer, the frame rate of respiratory signals acquired by the depth camera fluctuated in the experiment. The respiratory rate is an important indicator for evaluating respiratory patterns, which is influenced by the fluctuating frame rate, and this fluctuation finally affects the experimental results to some extent. 2) The long window waveform synthesized by short waveforms may have relatively large mutations at the splicing point. In the normal breathing waveform, this synthetic mutation is captured by the sliding window and input into the neural network model. Therefore, it tends to misdiagnose normal breathing as abnormal breathing.

VI. DISCUSSION

A. Discussion about multiple people situation

In the experiments, we use the data of one person's situation to validate the effectiveness and robustness of RSM and BI-AT-GRU. Furthermore, the proposed system can classify multiple people's breathing patterns in real time. **Fig. 9** shows the

waveforms and the classification results from three subjects breathing at the same time measured by our system.

Two factors need to be considered when it comes to multiple people. The first factor is the frame rate. We observe that the frame rate decreases with the increasing number of subjects because it takes time to automatically calculate all subjects' joints. Thus, we resample the signal in the multi-person application scenarios. The second factor is the hardware limitations of the depth camera. The depth camera cannot work at a too close or too far distance, and its maximum field of vision is limited. Therefore, the number of people recorded in a single depth image is limited. We can increase the number of depth cameras to expand the field of vision when there are a large number of people to be measured, such as respiratory infections screening in public places.

B. Discussion about the test data

In the experiments, the original test data is obtained from 20 subjects who are coached to imitate 6 respiratory patterns. Based on the results on test set, it is feasible to utilize well-designed simulated data to drive the deep neural network in the task of identifying and classifying abnormal breathing patterns. There are three shortcomings of our test data: 1) the test data comes from 20 subjects, and it can only cover part of the breathing patterns rather than the entire space of possible breathing patterns in the real world; 2) the test data is not collected from patients who are actually suffering from these abnormal patterns; and 3) the "gold standard" waveforms for classification are computer-derived sinusoidal waveforms instead of "real-world" patient data.

Our work can be extended to real application scenarios. In the future, the following additional work may be required: 1) consult experts to determine the specific respiratory patterns to be classified and their gold standards; and 2) collect large amounts of the real patients' respiratory data to fine-tune the pre-trained model and validate it.

VII. CONCLUSION

In this paper, we propose an unobtrusive and automatic measurement method based on human joints tracking and depth camera for determining multiple people's respiratory patterns in a real-time manner. Subsequently, we make minor modifications to BI-AT-GRU, which is widely utilized in natural language processing, and first apply it for classifying respiratory patterns. BI-AT-GRU is trained by the simulated data randomly generated by the proposed Respiratory Simulation Model, and it is tested by the data measured by the depth camera. In validation experiments, the obtained BI-AT-GRU yields the excellent performance with the accuracy, precision, recall, F1 of 94.5%, 94.4%, 95.1%, 94.8%, respectively. In comparative experiments, BI-AT-GRU specific to respiratory pattern classification outperforms the existing state-of-the-art viz. BI-AT-LSTM, GRU and LSTM. The results indicate that the proposed method is accurate and robust for remotely and unobtrusively classifying abnormal respiratory patterns.

Compared with previous researches, our method has the following pros: it can automatically select ROI and can be applied to multiple people situation in a real-time manner. The proposed method can be used to monitor, identify and classify abnormal respiratory patterns of multiple people. For clinical purposes, it can offer a pre-examine to potential patients, providing clues to treatment and assess the patient's prognosis; for public places, it can be applied to detect respiratory infections efficiently in the crowd; and for family, it can monitor family members' respiratory 24 hours a day and provide timely warning. In future research, we plan to fuse the other non-contact measurement techniques into the system and explore the more practical deep neural network architecture, thus making the system applicable to more challenging situations.

REFERENCES

- [1] F. Q. AL-Khalidi, R. Saatchi, D. Burke, H. Elphick, and S. Tan, "Respiration rate monitoring methods: A review," *Pediatric pulmonology*, vol. 46, no. 6, pp. 523–529, 2011.
- [2] J. F. Fieselmann, M. S. Hendryx, C. M. Helms, and D. S. Wakefield, "Respiratory rate predicts cardiopulmonary arrest for internal medicine inpatients," *Journal of general internal medicine*, vol. 8, no. 7, pp. 354–360, 1993.
- [3] M. Cretikos, J. Chen, K. Hillman, R. Bellomo, S. Finfer, A. Flabouris, M. S. Investigators *et al.*, "The objective medical emergency team activation criteria: a case-control study," *Resuscitation*, vol. 73, no. 1, pp. 62–72, 2007.
- [4] M. J. Griffiths, D. F. McAuley, G. D. Perkins, N. Barrett, B. Blackwood, A. Boyle, N. Chee, B. Connolly, P. Dark, S. Finney *et al.*, "Guidelines on the management of acute respiratory distress syndrome," *BMJ Open Respiratory Research*, vol. 6, no. 1, p. e000420, 2019.
- [5] H. W. Timm, "Analyzing deception from respiration patterns," *Journal of Police Science & Administration*, 1982.
- [6] F. Benetazzo, A. Freddi, A. Monterù, and S. Longhi, "Respiratory rate detection algorithm based on rgb-d camera: theoretical background and experimental results," *Healthcare technology letters*, vol. 1, no. 3, pp. 81–86, 2014.
- [7] M. J. Johnson, M. Kanaan, G. Richardson, S. Nabb, D. Torgerson, A. English, R. Barton, and S. Booth, "A randomised controlled trial of three or one breathing technique training sessions for breathlessness in people with malignant lung disease," *BMC medicine*, vol. 13, no. 1, p. 213, 2015.
- [8] C. Jones, B. Sangthong, O. Pachirat, and D. Jones, "Slow breathing training reduces resting blood pressure and the pressure responses to exercise," *Physiological research*, vol. 64, no. 5, 2015.
- [9] G. Zamzmi, D. Goldgof, R. Kasturi, Y. Sun, and T. Ashmeade, "Machine-based multimodal pain assessment tool for infants: a review," *arXiv preprint arXiv:1607.00331*, 2016.
- [10] J. C. Paz and M. P. West, *Acute Care Handbook for Physical Therapists-E-Book*. Elsevier Health Sciences, 2013.
- [11] G. Yuan, N. A. Drost, and R. A. McIvor, "Respiratory rate and breathing pattern," *McMaster University Medical Journal*, vol. 10, no. 1, pp. 23–25, 2013.
- [12] I. A. Rybak, J. F. Paton, R. Rogers, and W. S. John, "Generation of the respiratory rhythm: state-dependency and switching," *Neurocomputing*, vol. 44, pp. 605–614, 2002.
- [13] W.-J. St and J. Paton, "Defining eupnea," *Respiratory physiology & neurobiology*, vol. 139, no. 1, pp. 97–103, 2003.
- [14] P. A. Lanfranchi, A. Braghieri, E. Bosimini, G. Mazzuero, R. Colombo, C. F. Donner, and P. Giannuzzi, "Prognostic value of nocturnal cheyne-stokes respiration in chronic heart failure," *Circulation*, vol. 99, no. 11, pp. 1435–1440, 1999.
- [15] J. A. Dempsey, S. C. Veasey, B. J. Morgan, and C. P. O'Donnell, "Pathophysiology of sleep apnea," *Physiological reviews*, vol. 90, no. 1, pp. 47–112, 2010.
- [16] V. Mohsenin, "Sleep-related breathing disorders and risk of stroke," *STROKE-DALLAS-*, vol. 32, no. 6, pp. 1271–1276, 2001.
- [17] F. Roux, C. D'Ambrosio, and V. Mohsenin, "Sleep-related breathing disorders and cardiovascular disease," *The American journal of medicine*, vol. 108, no. 5, pp. 396–402, 2000.
- [18] M. J. Tobin, "Respiratory monitoring in the intensive care unit 1 2," *Am Rev Respir Dis*, vol. 138, pp. 1625–1642, 1988.
- [19] M. J. Mador and M. J. Tobin, "Effect of alterations in mental activity on the breathing pattern in healthy subjects," *American Review of Respiratory Disease*, vol. 144, no. 3_pt_1, pp. 481–487, 1991.
- [20] A. D. Droitcour, T. B. Seto, B.-K. Park, S. Yamada, A. Vergara, C. El Hourani, T. Shing, A. Yuen, V. M. Lubecke, and O. Boric-Lubecke, "Non-contact respiratory rate measurement validation for hospitalized patients," in *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2009, pp. 4812–4815.
- [21] J. Kranjec, S. Beguš, J. Drnovšek, and G. Geršak, "Novel methods for noncontact heart rate measurement: A feasibility study," *IEEE transactions on instrumentation and measurement*, vol. 63, no. 4, pp. 838–847, 2013.
- [22] A. Al-Naji, K. Gibson, S.-H. Lee, and J. Chahl, "Monitoring of cardiorespiratory signal: Principles of remote measurements and review of methods," *IEEE Access*, vol. 5, pp. 15 776–15 790, 2017.
- [23] J. Rumiński, "Analysis of the parameters of respiration patterns extracted from thermal image sequences," *Biocybernetics and Biomedical Engineering*, vol. 36, no. 4, pp. 731–741, 2016.
- [24] S. Wiesner and Z. Yaniv, "Monitoring patient respiration using a single optical camera," in *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2007, pp. 2740–2743.
- [25] X. Liu, D. Zhai, R. Chen, X. Ji, D. Zhao, and W. Gao, "Depth super-resolution via joint color-guided internal and external regularizations," *IEEE Transactions on Image Processing*, vol. 28, no. 4, pp. 1636–1645, 2018.
- [26] ———, "Depth restoration from rgb-d data via joint adaptive regularization and thresholding on manifolds," *IEEE Transactions on Image Processing*, vol. 28, no. 3, pp. 1068–1079, 2018.
- [27] V. Soleimani, M. Mirmehdi, D. Damen, J. Dodd, S. Hannuna, C. Sharp, M. Camplani, and J. Viner, "Remote, depth-based lung function assessment," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 8, pp. 1943–1958, 2016.
- [28] U. Wijenayake and S.-Y. Park, "Real-time external respiratory motion measuring technique using an rgb-d camera and principal component analysis," *Sensors*, vol. 17, no. 8, p. 1840, 2017.
- [29] M. Bae, S. Lee, and N. Kim, "Development of a robust and cost-effective 3d respiratory motion monitoring system using the kinect device: Accuracy comparison with the conventional stereovision navigation system," *Computer methods and programs in biomedicine*, vol. 160, pp. 25–32, 2018.

- [30] A. Al-Naji and J. Chahl, "Detection of cardiopulmonary activity and related abnormal events using microsoft kinect sensor," *Sensors*, vol. 18, no. 3, p. 920, 2018.
- [31] F. Ernst and P. Saß, "Respiratory motion tracking using microsoft's kinect v2 camera," *Current Directions in Biomedical Engineering*, vol. 1, no. 1, pp. 192–195, 2015.
- [32] A. Al-Naji, K. Gibson, S.-H. Lee, and J. Chahl, "Real time apnoea monitoring of children using the microsoft kinect sensor: a pilot study," *Sensors*, vol. 17, no. 2, p. 286, 2017.
- [33] M.-H. Hu, G.-T. Zhai, D. Li, Y.-Z. Fan, X.-H. Chen, and X.-K. Yang, "Synergetic use of thermal and visible imaging techniques for contactless and unobtrusive breathing measurement," *Journal of biomedical optics*, vol. 22, no. 3, p. 036006, 2017.
- [34] C. B. Pereira, X. Yu, T. Goos, I. Reiss, T. Orlikowsky, K. Heimann, B. Venema, V. Blazek, S. Leonhardt, and D. Teichmann, "Noncontact monitoring of respiratory rate in newborn infants using thermal imaging," *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 4, pp. 1105–1114, 2018.
- [35] P. De Chazal, C. Heneghan, E. Sheridan, R. Reilly, P. Nolan, and M. O'Malley, "Automated processing of the single-lead electrocardiogram for the detection of obstructive sleep apnoea," *IEEE Transactions on Biomedical Engineering*, vol. 50, no. 6, pp. 686–696, 2003.
- [36] A. Fekr, M. Janidarmian, K. Radecka, and Z. Zilic, "A medical cloud-based platform for respiration rate measurement and hierarchical classification of breath disorders," *Sensors*, vol. 14, no. 6, pp. 11204–11224, 2014.
- [37] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, p. 436, 2015.
- [38] A. Kamilaris and F. X. Prenafeta-Boldú, "Deep learning in agriculture: A survey," *Computers and electronics in agriculture*, vol. 147, pp. 70–90, 2018.
- [39] J. De Fauw, J. R. Ledsam, B. Romera-Paredes, S. Nikolov, N. Tomasev, S. Blackwell, H. Askham, X. Glorot, B. O'Donoghue, D. Visentin *et al.*, "Clinically applicable deep learning for diagnosis and referral in retinal disease," *Nature medicine*, vol. 24, no. 9, p. 1342, 2018.
- [40] R. Poplin, A. V. Varadarajan, K. Blumer, Y. Liu, M. V. McConnell, G. S. Corrado, L. Peng, and D. R. Webster, "Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning," *Nature Biomedical Engineering*, vol. 2, no. 3, p. 158, 2018.
- [41] H. Li, K. Ota, and M. Dong, "Learning iot in edge: Deep learning for the internet of things with edge computing," *IEEE Network*, vol. 32, no. 1, pp. 96–101, 2018.
- [42] Y. Cho, N. Bianchi-Berthouze, and S. J. Julier, "Deepbreath: Deep learning of breathing patterns for automatic stress recognition using low-cost thermal imaging in unconstrained settings," in *2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII)*. IEEE, 2017, pp. 456–463.
- [43] S.-H. Kim and G.-T. Han, "1d cnn based human respiration pattern recognition using ultra wideband radar," in *2019 International Conference on Artificial Intelligence in Information and Communication (ICAIC)*. IEEE, 2019, pp. 411–414.
- [44] J. D. Hunter, "Matplotlib: A 2d graphics environment," *Computing in science & engineering*, vol. 9, no. 3, p. 90, 2007.
- [45] J. L. Elman, "Finding structure in time," *Cognitive science*, vol. 14, no. 2, pp. 179–211, 1990.
- [46] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [47] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using rnn encoder-decoder for statistical machine translation," *arXiv preprint arXiv:1406.1078*, 2014.
- [48] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," *arXiv preprint arXiv:1409.0473*, 2014.
- [49] Z. Yang, D. Yang, C. Dyer, X. He, A. Smola, and E. Hovy, "Hierarchical attention networks for document classification," in *Proceedings of the 2016 conference of the North American chapter of the association for computational linguistics: human language technologies*, 2016, pp. 1480–1489.
- [50] J. K. Chorowski, D. Bahdanau, D. Serdyuk, K. Cho, and Y. Bengio, "Attention-based models for speech recognition," in *Advances in neural information processing systems*, 2015, pp. 577–585.
- [51] P. Zhou, W. Shi, J. Tian, Z. Qi, B. Li, H. Hao, and B. Xu, "Attention-based bidirectional long short-term memory networks for relation classification," in *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, 2016, pp. 207–212.