Identification of Respiratory Sounds Collected from Microphones Embedded in Mobile Phones

AUTHOR(S):

Fukuyama, Keita; Sugiyama, Osamu; Chin, Kazuo; Satou, Susumu; Matsumoto, Shigemi; Muto, Manabu

CITATION:

Fukuyama, Keita ...[et al]. Identification of Respiratory Sounds Collected from Microphones Embedded in Mobile Phones. Advanced Biomedical Engineering 2022, 11: 58-67

ISSUE DATE:

2022

URL:

http://hdl.handle.net/2433/282086

Advanced Biomedical Engineering
11: 58–67, 2022.

Original Paper

DOI:10.14326/abe.11.58

Identification of Respiratory Sounds Collected from Microphones Embedded in Mobile Phones

Keita Fukuyama,[*] Osamu Sugiyama,[*] Kazuo Chin,[**] Susumu Satou,[***] Shigemi Matsumoto,[*] Manabu Muto[†, #]

Abstract Sudden deterioration of condition in patients with various diseases, such as cardiopulmonary arrest, may result in poor outcome even after resuscitation. Early detection of deterioration is important in medical and long-term care settings, regardless of the acute or chronic phase of disease. Early detection and appropriate interventions are essential before resuscitating measures are required. Among the vital signs that indicate the general condition of a patient, respiratory rate has a greater ability to predict serious events such as thromboembolism and sepsis than heart rate and blood pressure, even in early stages. Despite its importance, however, respiratory rate is frequently overlooked and not measured, making it a neglected vital sign. To facilitate the measurement of respiratory rate, a non-invasive method of detecting respiratory sounds was developed based on deep learning technology, using a built-in microphone in a smartphone. Smartphones attached to the bed headboards of 20 participants undergoing polysomnography (PSG) at Kyoto University Hospital recorded respiratory sounds. Sound data were synchronized with overnight respiratory information. After excluding periods of abnormal breathing on the PSG report, sound data were processed for each 1-minute period. Expiration sound was determined using the pressure flow sensor signal on PSG. Finally, a model to identify the expiration section from the sound information was created using a deep learning algorithm from the convolutional Long Short Term Memory network. The accuracy of the learning model in identifying the expiratory section was 0.791, indicating that respiratory rate can be determined using the microphone in a smartphone. By collecting data from more patients and improving the accuracy of this method, respiratory rates could be more easily monitored in all situations, both inside and outside the hospital.

Keywords: respiratory sound identification, non-touch vital sign monitoring, smartphone.

Adv Biomed Eng. **11**: pp. 58–67, 2022.

1. Introduction

Sudden deterioration of condition in patients with various diseases, such as cardiopulmonary arrest, may result in poor outcome even after resuscitation. These patients may therefore require additional medical resources for treatment. Detecting the deterioration of patients at an early stage is essential in medical and long-term care settings, regardless of the acute or chronic phase, as it enables appropriate interventions before resuscitating measures are required [1, 2].

One vital sign that indicates the general condition of a patient is respiratory rate, which can better predict severe events such as thromboembolism and sepsis than heart rate and blood pressure from the early stages of the underlying condition [3, 4]. Respiratory rate, however, is frequently not measured, making it a neglected vital sign [5].

The major mechanism for continuous monitoring of respiratory rate requires sensors in contact with the sur-

This study was presented at the Symposium on Biomedical Engineering 2021, September, 2021.
Received on July 30, 2021; revised on October 29, 2021 and January 1, 2022; accepted on January 11, 2022.

[*] Department of Real World Data Research and Development, Graduate School of Medicine, Kyoto University, Japan.
[**] Division of Sleep Medicine, Department of Internal Medicine, Department of Sleep Medicine and Respiratory Care, Nihon University of Medicine, Japan.
[***] Department of Respiratory Care and Sleep Control Medicine, Graduate School of Medicine, Kyoto University, Japan.
[†] Department of Clinical Oncology, Kyoto University Hospital, Japan.
[#] 54 Kawara-cho, Shogoin, Sakyo-ku, Kyoto city, Kyoto 606–8507, Japan.
E-mail: mmuto@ kuhp.kyoto-u.ac.jp

face of the body [6]. A few systems, however, are capable of monitoring respiratory rate without contact between sensors and the body. Respiratory rate monitoring using wearable or specialized non-contact devices is not sufficiently widespread to cover the entire population, and is not suitable for long-term health monitoring for all patients at risk.

Smartphones are among the most popular electronic devices used worldwide, with several sensors for general purposes. In 2021, 3.8 billion people, or 48.33% of the world population, were estimated to own smartphones [7]. The ability of smartphones to monitor respiratory rate could allow widespread and continuous monitoring for many patients at risk. Methods of monitoring respiratory rate with smartphones include placing smartphones on the face and placing smartphones that emit ultrasonic waves in front of the patient to detect movement of the thorax. These methods are not suitable for continuous monitoring of patients' respiratory rates in hospitals and long-term care settings [8, 9].

To facilitate the measurement of respiratory rate, we developed a non-invasive method of detecting respiratory sounds using a microphone built into a smartphone and placing the latter at the headboard of the patient's bed. It was difficult to correctly identify breathing sounds from the sound data recorded by a smartphone. Respiratory sounds, excluding snoring, are very faint and easily masked by sounds of body movements, bruxism, and footsteps. Breathing itself is a periodic phenomenon, but the cycle is not constant and the sounds produced by breathing are not uniform among patients. To accurately determine the subtle sounds of breathing recorded by a smartphone, information from other synchronized sensors that can accurately identify breathing is required. We therefore record sounds simultaneously in patients undergoing polysomnography (PSG) examination to obtain highly accurate respiratory information, and synchronize the audio data with this respiratory information. This has led to the development of a model that identifies breath sounds using the recorded sound data and respiratory information. By improving the method of identifying respiratory sounds recorded by smartphones, it should be possible to continuously monitor respiratory rates in various settings including hospitalized patients, clients of care facilities, and outpatients.

2. Methods

The methods used to obtain sound data synchronized with respiration information, to create a model to identify respiration, and to evaluate the performance of the model are described below. The workflow of the procedure is shown in **Fig. 1**.

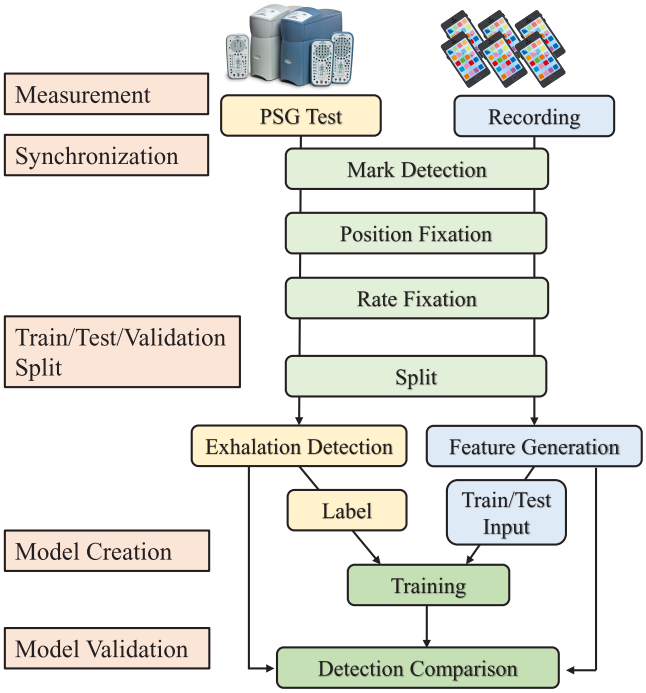


Fig. 1 Outline of the experiment.

2.1 Measurements

To obtain sound data from a microphone built into a smartphone and synchronize the sounds with respiratory information, we recorded sounds on smartphones in 20 patients with sleep apnea undergoing PSG examination using Alice 6 (Philips) in a single room at Kyoto University Hospital. The measurement is outlined in **Fig. 2**. To verify whether respiratory rate can be measured via smartphone in various individuals, the 20 patients included both men and women, who varied widely in age and severity of sleep apnea.

This study was reviewed and approved by the Ethics Committee of Kyoto University Hospital (R2614). Written informed consent was obtained from each adult participant and parents of children.

Equipment needed for PSG examinations included a barometric pressure flow sensor, a temperature sensor worn under the nose, and a snoring sensor, along with other sensors. The barometric pressure flow sensor (P-Flow), which corresponds to the nasal pressure sensor in the American Academy of Sleep Medicine guidelines for scoring of sleep and associated events [10], detects pressure changes due to inspiration and expiration through the cannula of the nostrils by guiding it to the pressure transducer. The temperature sensor (T-Flow), which corresponds to the airflow sensor in the guidelines, detects the temperature changes around the nose using a thermistor. The snoring sensor (SNOR), which corresponds to the snoring sound sensor in the guidelines, uses the piezo effect to detect vibrations at the skin

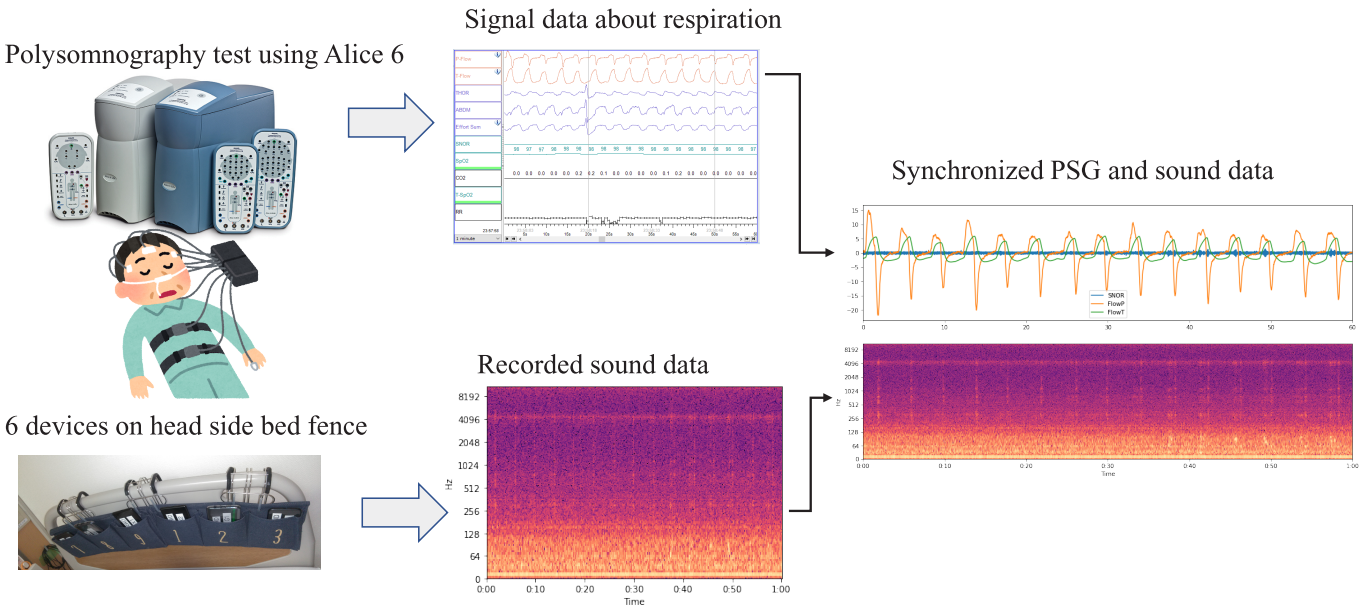


Fig. 2 Outline of measurements.
SNOR (blue line): signal from the piezo snoring sensor that detects vibrations due to snoring.
FlowP (orange line): signal from the pressure flow sensor that detects changes in barometric pressure under the patient's nostrils.
FlowT (green line): signal from the temperature sensor that detects changes in temperature under the patient's nose.

due to snoring. Along with signal information over time, information on the period during which abnormal respiration occurs can be obtained as a report.

During this measurement, six smartphones (two iPhones, Galaxy phones, and Xperia phones) were lined up on the headboard, and data were recorded overnight. The six smartphones were inserted into the fixtures placed on the head side of the bed fence in a fixed order (Fig. 2), with the microphones facing upward and the displays facing away from the patient. The distance between the microphone of each smartphone and the patient's nose was within approximately 50 cm at the beginning of PSG.

Six smartphones were used to avoid data loss from measurement error and to avoid creating a model optimized for one smartphone model and OS. Sound was recorded by iPhones using the "Voice memo" app and by Android phones using the "Easy Voice Recorder (DIGIPOM)" app. PSG data were interpreted by consultation with a pulmonologist.

2.2 Data synchronization

We generated a mark using the snoring sensor of the PSG system and an artificial sound output from a Bluetooth speaker at the beginning of data collection to synchronize breathing information obtained by PSG with the data recorded by the smartphones. The marking and synchronization p is outlined in **Fig. 3**. Any time discrepancy between these data was corrected by the rate for PSG.

2.3 Data extraction

The data used for learning were extracted from the recorded and PSG data. PSG data were collected from the time the ward light was switched off at night to the time it was turned on in the morning. The period that included sounds from televisions and radios was excluded, as were periods of abnormal breathing such as PSG-defined obstructive apnea, central apnea, and respiratory effort-related arousal responses. This process reduces the data usage in cases of severe sleep apnea.

Data extracted for learning were divided into 1-minute periods, with 60% of the data used for training, 20% for testing, and 20% for validation.

Based on the barometric pressure flow sensor signal, which sensitively detects exhaled breath, a period of 1.5 seconds starting 0.75 seconds before and ending 0.75 seconds after the local minimum barometric pressure was defined as one exhaled breath. Periods that were not apnea but could not be detected by the barometric pressure flow sensor, such as mouth breathing, were excluded from the training and test data.

2.4 Feature generation

The following processing was performed to input each 60-second period of 44.1 khz monaural sound data labeled with expiration span into the machine learning model. The sound data were transformed by a short-time Fourier series to generate the power spectrum. After reducing noise with a Gaussian filter of kernel size 25×25,

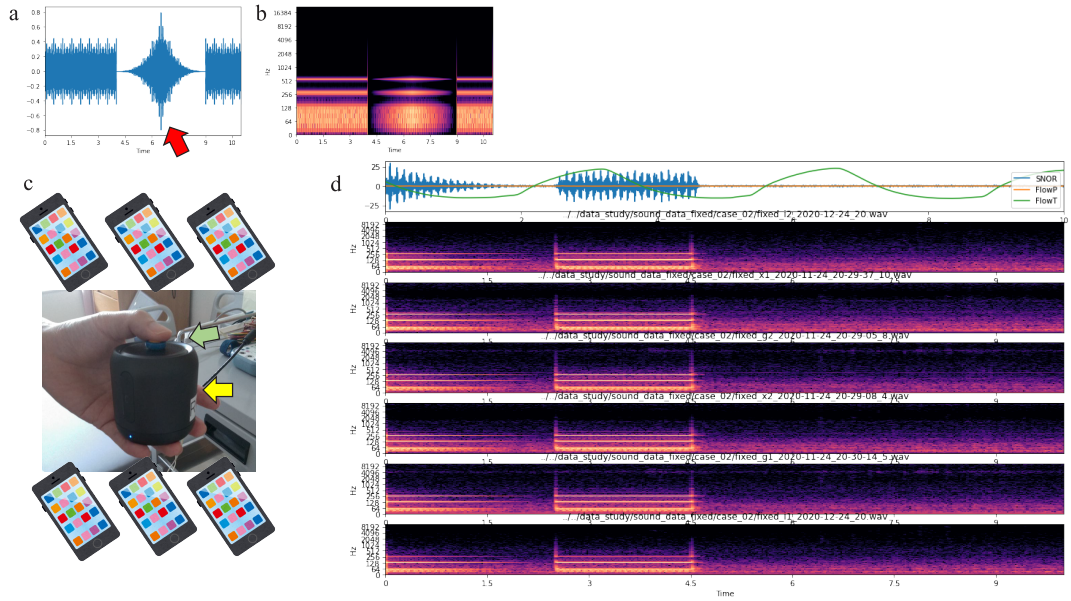


Fig. 3 Schema of data synchronization.
a: The artificial sound signals for synchronization. Red arrow shows the peak to detect.
b: The power spectrum of the artificial sound. The sound was created by mixing sinusoidal waves of multiple frequencies.
c: Schema of marking. Green arrow: The piezo snoring sensor of PSG system. Yellow arrow: Bluetooth speaker to output the sound.
d: The head of synchronized PSG data and sound.

the average of each frequency component was calculated, and this parameter was subtracted for each frequency.

Because coarse noise and snoring frequently appear as high-frequency components, the part of the subtracted spectrum matrix at frequencies ≥ 16537.5 kHz was filtered. A two-dimensional array of 769 frequency components and 5168 time components was used as input. The flowchart is shown in **Fig. 4**.

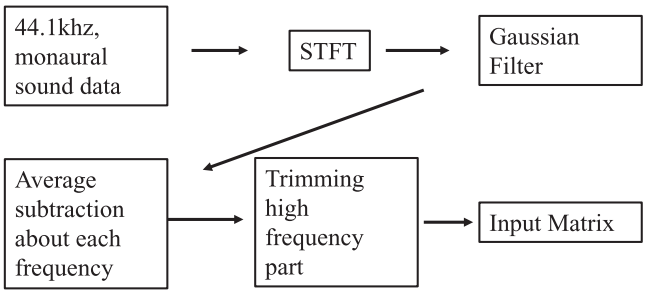


Fig. 4 Flowchart of processing input data.

2.5 Convolutional LSTM network

Signal segmentation was performed using a Convolutional Long Short-Term Memory (ConvLSTM) neural network as a machine learning model to detect exhaled sounds within each 1-minute sound recording. The structure of the model was determined with reference to the model used in a study on the segmentation of an electrocardiogram waveform [11].

ConvLSTM is a model commonly used to illustrate sequence information by showing imaging features including sound [12–14]. The convolution layer extracts image-like features of a two-dimensional matrix derived from the power spectrum for the frequency axis and compresses the features into a dense vector from 769 to 128. The LSTM layer captures changes in the features sent from the convolution layer over time. The Dense and Softmax layer without width for time series processes sequentially 5168 inputs and outputs 5168 pieces of classification data. The outline of the model is shown in **Fig. 5** and **Supplemental Document S1**. Detailed training conditions are shown in **Table 1**.

2.6 Validation

To evaluate the performance of the proposed method, the ability to detect breath sounds was evaluated by comparing with the pressure flow sensor. The exhaled breath detection status was verified based on the respiratory information for validation. Participants with a breath detection frequency ≤ 9 per minute despite the PSG chart showing normal respiration were regarded as detection failures.

Accuracy of respiration detection using the sound data was evaluated by comparing with the exhaled air detected by the pressure air flow sensor. Data in which

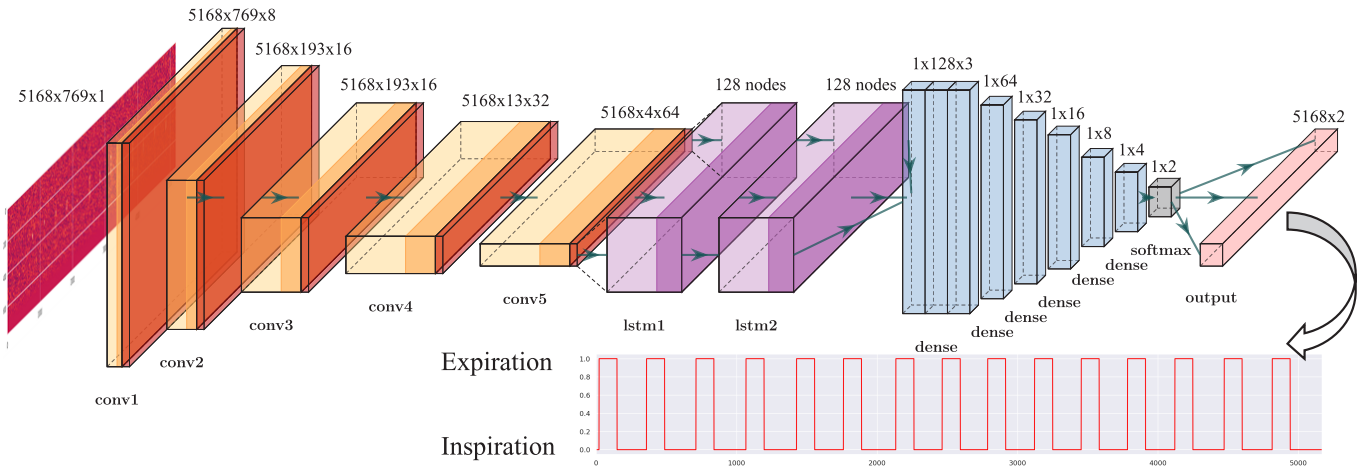


Fig. 5 Schema of the ConvLSTM model.
The input matrix is processed in the order from left to right. Locally, two-dimensional features are compressed by the convolution layer, and time-series information is determined by the two-stage LSTM layer. The Dense and Softmax layers sequentially process the time-series information output by the LSTM layer and output classification information for each event.

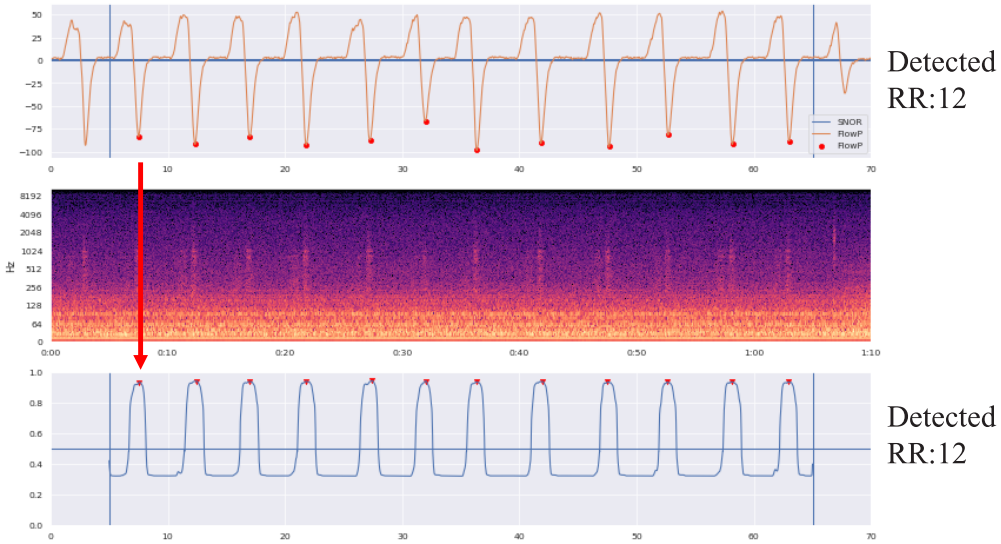


Fig. 6 Schema of model validation.
Upper: Pressure flow sensor signals and SNOR signals from PSG. Red circles indicate local minimum signals from pressure flow sensor with numbers corresponding to the respiratory rate (the number is 12 in this case). Middle: Recorded sound data. Lower: Model output of likelihood of expiration. Red triangles indicate peak of likelihood of expiration with numbers corresponding to the respiratory rate obtained from sound data (the number is 12 in this case). The red arrow shows the relation of accuracy calculation. If the peak of likelihood is within 1 second of the local minimum pressure flow sensor signal, the output is assumed to be accurate.

both the pressure flow sensor and the model failed to detect respiration were regarded as non-evaluable and excluded from the accuracy assessment. Because the length of expiratory sound was about 1 second, the range to detect the expiration peak from sound data was set from 0.5 seconds before to 0.5 seconds after the minimum value of the pressure flow sensor suggesting that expiration was observed. Accuracy was calculated by assuming that a correct result was obtained when this peak of expiration-like output was within the range. The outline of the model validation is shown in **Fig. 6**. Accuracy for case i, validation segment j, device k, and expiration index from PFlow l when detection times for expiration from PFlow and sound were represented as x and y, re-

Table 1 Training condition.

Convolution size in each layer:		(3,5)
Size of the output tensor of each layer:		In document S1
Activation function:		
	ReLu (excluding LSTM)	
Hyperparameters of the LSTM:		
	Direction:	Bidirectional
	Number of hidden nodes:	64 for each direction
	Activation:	tanh
	Recurrent activation:	hard sigmoid
	Dropout:	0
	Recurrent dropout:	0
	Return Sequence:	TRUE
Initialization method of the model:		
	Kernel initializer:	Glorot uniform
	Bias initializer:	Zeros
Optimization algorithms:		
	ADAM (with differential learning rate)	
Loss value:	Binary cross entropy	
Batch size:		12
Learning rate:		
	LSTM:	1.00E-04
	Dense:	1.00E-08
	Conv2d:	1.00E-08
	Batch normalization:	1.00E-08
Termination judgment conditions:		
	Stop if improvement of loss value is less than 0.001 during 10 epochs.	
Scheduled number of learning epochs:		500
Learning curve:		**Fig. S2**
Data size:		
	Train:	13800 samples, 439 G
	Test:	4600 samples, 147.5 G

spectively, was mathematically defined as:

$$Acc_{i,j_i,k_{i,j},l_{i,j_i}} = \begin{cases} 1 \; for \; \exists m \in \mathbb{N}(|x_{l_{i,j_i}} - y_{m_{i,j_i,k_{i,j}}}| \leq 0.5) \\ 0 \; for \; \nexists m \in \mathbb{N}(|x_{l_{i,j_i}} - y_{m_{i,j_i,k_{i,j}}}| \leq 0.5) \end{cases}$$

The accuracy of the entire validation was defined as:

$$Acc_{whole} = \sum_{i,j,k,l} Acc_{i,j_i,k_{i,j},l_{i,j_i}} / \sum_{i,j,k,l} 1$$

The accuracy of case i was defined as:

$$Acc_i = \sum_{j,k,l} Acc_{i,j_i,k,l_{i,j_i}} / \sum_{j,k,l} 1$$

The detailed definition of accuracy is described in **Supplemental Document S1**.

The relationship between the number of exhaled breaths per minute detected from the pressure air flow sensor and voice data was visualized.

Because of concerns that the data available for learning would decrease depending on the severity of sleep apnea syndrome (SAS), the relations among severity, data availability, and the accuracy of the model were evaluated.

3. Results

3.1 Patient characteristics

Twenty patients, comprising 13 males and 7 females with median age 55 years (range, 15–77 years) were included in this study. Their demographic and clinical characteristics are shown in **Table 2**.

Of the 20 patients, six had severe sleep apnea syndrome (SAS). As SAS became more severe, abnormal breathing and apnea increased. The amount and percentage of data after filtering in each process are summarized in **Table 3**. The median length of respiratory information available for training, testing, and validation was 6.8 hours.

3.2 Results of learning

A total of 7,946 minutes of respiratory information and recorded data from six smartphones were obtained from the 20 patients. These data were divided into a training set of 4754 minutes, a test set of 1598 minutes, and a validation set of 1594 minutes. Three samples of training and test data were created from 1-minute measurements. **Fig. S2** shows the learning curve. The final accuracy of the test set was 0.777.

3.3 Detection of respiration

Table 4 shows the rates of detection of respiration by the pressure-flow sensor and from sound. The pressure-flow sensor was able to detect > 90% of respiration, whereas sound could detect only approximately 51%.

The accuracy for breath detection by the machine learning model was 79.1% when respiration was detected by both the pressure flow sensor and sound. **Figure 7** shows the distribution of the estimated respiratory rate when respiratory was successfully detected.

3.4 Relations with accuracy and SAS

As SAS became more severe, the data available for learning decreased due to apnea and abnormal breathing. We evaluated whether the ability of the model to detect exhaled breath was optimized for people with mild SAS and those without SAS. The accuracy in each case was not clearly associated with SAS severity. **Supplemental Figure S3** shows the distribution of accuracy relative to

Table 3 Data usage during processing.

	hours(median[IQR])	%(median[IQR])
whole examinations	10.49[10.24–10.71]	100.0[100.0–100.0]
from light on/off	8.6[8.21–9.04]	82.7[79.5–84.6]
without TV or radio sound	8.58[8.21–8.98]	82.1[79.5–83.6]
without abnormal breath	7.79[7.31–8.15]	74.1[70.3–78.1]
train/test/validation	6.79[5.96–7.33]	64.2[58.2–70.1]
train	4.06[3.56–4.38]	38.5[34.8–42.0]
test	1.37[1.19–1.47]	12.9[11.7–14.0]
validation	1.37[1.2–1.47]	12.9[11.7–14.0]

Table 2 Clinical characteristics.

Number	20
Age (median [IQR])	55 [36, 68]
Gender = female/male (%)	7/13 (35.0/65.0)
BW (kg) (median [IQR])	64.15 [53.45, 69.82]
BH (cm) (median [IQR])	166.70 [161.38, 169.62]
BMI (median [IQR])	23.14 [20.35, 25.43]
AHI (median [IQR])	14.90 [6.20, 31.75]
SAS Severity	
normal (AHI < 5)	3 (15.0)
mild (5 <= AHI < 15)	7 (35.0)
moderate (15 <= AHI < 30)	4 (20.0)
severe (30 <= AHI)	6 (30.0)

Abbreviations: BW, body weight; BH, body height; BMI, body mass index; AHI, apnea hypopnea index; SAS, sleep apnea syndrome

Table 4 Detection rate (%) of respiration.

	detected by sound	not by sound	Total
detected by pressure sensor	48.8	43.2	92.0
not by pressure sensor	1.9	6.1	8.0
total	50.7	49.3	100

the severity of SAS.

4. Discussion

4.1 Respiratory rate monitoring by smartphone

Placement of a smartphone with a built-in microphone under the nose could be used to determine the respiratory rate based on voice information [9]. Although the patients did not report discomfort, this method is not suited for long-term continuous monitoring.

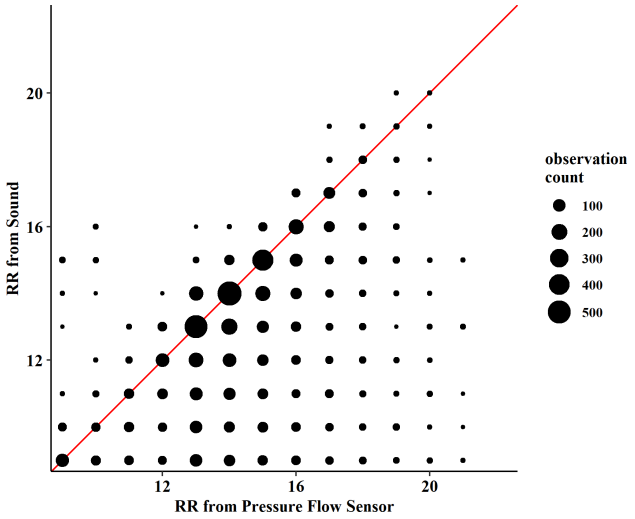


Fig. 7 Estimation of respiratory rate.

Respiratory rate can also be determined by monitoring the movement of the thorax following emission of ultrasonic waves by a microphone built into a smartphone [8, 15]. Because these methods involve ultrasonic irradiation, the smartphone must be placed in front of the chest, making it unsuitable for patients on a bed covered with a thick blanket.

Respiratory rate has also been determined from image information by placing a finger on the built-in camera of a smartphone [16], and thoracic movement detected by a standard smartphone held on a person's chest [17], but these methods are not suitable for long-term monitoring. Earphones have been used as a microphone to detect breath sounds [18] and are expected to improve performance when breath sounds are minute and difficult to determine, but this method would require additional equipment and evaluation of the performance of each earphone.

We propose a method for non-contact monitoring of

vital signs by a smartphone placed on the headboard of a bed. This method can be applied to various patients, including outpatients.

4.2 Relationship with snoring sounds

Smartphone-recorded snoring and PSG data have been used to verify the severity of sleep apnea [19]. Snoring is a very loud and characteristic waveform of breathing-related sounds recorded during sleep. Snoring usually occurs during inspiration, except in patients with severe SAS. Thus, to avoid designing a model that recognizes snoring dominantly, we designed a model that detects exhaled breath sound.

4.3 Limitations

In this study, the pressure airflow sensor was able to detect > 90% of normal respirations, whereas recorded sound could detect only 51%. The main reason for the low detection rate using sound was that breath sounds were tranquil and were masked by white noise or could not be discriminated from other coarse sound sources. Further improvements should include the removal of background noise to detect minute breath sounds more accurately.

One of the causes of the low accuracy of detecting expiration is that the sounds are weak and only a part of them can be detected. Another cause is the presence of gross noise; the weak expiration sound is further attenuated as a result of the process of subtracting the signal intensity of time-averaged frequency, and only a part of the expiration can be detected. In order to reduce the influence of the coarse spike-like noise, a method such as subtracting the moving average of a certain length instead of the frequency averaged for 1 minute can be considered.

Because this study targeted sleeping patients with no critical or systemic diseases who were undergoing PSG examinations, it was not possible to record the breath sounds of severely ill patients with tachypnea. Breath sounds due to airflow are likely to increase in patients with tachypnea due to the increase in ventilation volume. However, learning and identification of breathing sounds without PSG that provides detailed respiratory information is required for patients with tachypnea and severe sickness. Further studies are needed to develop a learning model that can differentiate tachypnea from sounds.

4.4 Significance

This study showed that breath sounds could be detected in data recorded by smartphones not in direct contact with patients. This suggests that important vital signs indicative of patient deterioration could be monitored by general purpose devices in many patients without increasing the burden on medical staff.

In the validation set, the model detected respiratory sounds in only one-half of the recorded data, suggesting that this method is inappropriate for use in intensive care units, where accurate monitoring is required. This method, however, has the potential to detect 50% of respiration sounds during sleep, thus allowing continuous monitoring of large numbers of patients at risk without the need for any specialized devices.

Dyspnea and tachypnea are clinical indicators of critically ill patients with COVID-19 infection [20, 21]. Objective detection of the deterioration of respiratory conditions may be a marker of early-stage tachypnea in outpatients and inpatients, allowing appropriate early interventions.

5. Conclusion

We have proposed a method for detecting breathing sounds using a microphone embedded in a mobile phone. By simultaneous recording of sounds in patients undergoing PSG examinations, we succeeded to create a model capable of signal segmentation by deep learning. The created model was able to detect breath sounds in one-half of the data in the validation set, with accuracy of expiration sound identification of 79.1%, when correct detection was defined as the output from the model matching the breath detection by the pressure air flow sensor of the PSG system.

Conflicts of interest

Professor Chin was employed in a donation course funded by Philips-Respironics, Fukuda Denshi, Fukuda Life Tech Tokyo, and ResMed. Dr. Sato has received grants from Philips-Respironics, Fukuda Denshi, Fukuda Lifetec Keiji, and ResMed.

Acknowledgment

This study was supported by industry-academia joint research fund (150200700012) from Department of Real World Data Research and Development.

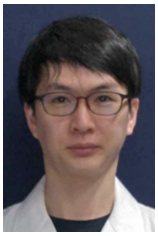
We thank the members of the PSG Laboratory of Kyoto University Hospital.

References

1. DeVita MA, Braithwaite RS, Mahidhara R, Stuart S, Foraida M, Simmons RL: Use of medical emergency team responses to reduce hospital cardiopulmonary arrests. Qual Saf Health Care. **13**(4), 251–254, 2004.
2. Sandroni C, Nolan J, Cavallaro F, Antonelli M: In-hospital cardiac arrest: incidence, prognosis and possible measures to improve survival. Intensive Care Med. **33**(2), 237–245, 2007.
3. Fieselmann JF, Hendryx MS, Helms CM, Wakefield DS: Respiratory rate predicts cardiopulmonary arrest for internal medicine inpatients. J Gen Intern Med. **8**(7), 354–360, 1993.

4. Schein RMH, Hazday N, Pena M, Ruben BH, Sprung CL: Clinical antecedents to in-hospital cardiopulmonary arrest. Chest. **98**(6), 1388–1392, 1990.
5. Cretikos MA, Bellomo R, Hillman K, Chen J, Finfer S, Flabouris A: Respiratory rate: the neglected vital sign. Med J Aust. **188**(11), 657–659, 2008.
6. Massaroni C, Nicolò A, Lo Presti D, Sacchetti M, Silvestri S, Schena E: Contact-based methods for measuring respiratory rate. Sensors (Basel). **19**(4), 908, 2019.
7. BankMyCell.com. HOW MANY SMARTPHONES ARE IN THE WORLD? <https://www.bankmycell.com/blog/how-many-phones-are-in-the-world#1579705085743-b3697bdb-9a8f>.[accessed on July 20, 2021]
8. Ge L, Zhang J, Wei J: Single-frequency ultrasound-based respiration rate estimation with smartphones. Comput Math Methods Med. 2018, 3675974, 2018.
9. Nam Y, Reyes BA, Chon KH: Estimation of respiratory rates using the built-in microphone of a smartphone or headset. IEEE J Biomed Health Inform. **20**(6), 1493–1501, 2016.
10. Berry. RB, Quan. SF, Abreu. AR, Bibbs. ML, Harding. SM, Mao. M-M, et al.: The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications. American Academy of Sleep Medicine; 2020.
11. Malali A, Hiriyannaiah S, Siddesh GM, Krishnarajanagar GS, Sanjay NT: Supervised ECG wave segmentation using convolutional LSTM. ICT Express. **6**(3), 166–169, 2020.
12. Jung S, Park J, Lee S, editors: Polyphonic sound event detection using convolutional bidirectional LSTM and synthetic data-based transfer learning. ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP); 2019 12–17 May 2019.
13. Komatsu A, Nishimura R, Kitaoka N: Environmental sounds recognition with convolutional-LSTM. 2019 IEEE 8th Global Conference on Consumer Electronics (GCCE). 691–693, 2019.
14. Wakamoto R, Mabu S, Kido S, Kuremoto T: Lung sound classification using deep neural networks with pre-training -comparison of the performance between CNN, LSTM and convolutional LSTM. Trans Inst Electr Eng Jpn C, A Publ Electron, Inf Syst Soc. **140**(12), 1402–1409, 2020.
15. Wang X, Huang R, Mao S, editors: Sonarbeat: sonar phase for breathing beat monitoring with smartphones. 2017 26th International Conference on Computer Communication and Networks (ICCCN); 2017 31 July-3 Aug. 2017.
16. Nam Y, Lee J, Chon KH: Respiratory rate estimation from the built-in cameras of smartphones and tablets. Ann Biomed Eng. **42**(4), 885–898, 2014.
17. Aly H, Youssef M, editors: Zephyr demo: ubiquitous accurate multi-sensor fusion-based respiratory rate estimation using smartphones. 2016 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS); 2016 10–14 April 2016.
18. Ren Y, Wang C, Yang J, Chen Y, editors: Fine-grained sleep monitoring: hearing your breathing with smartphones. 2015 IEEE Conference on Computer Communications (INFOCOM); 2015 26 April-1 May 2015.
19. Nakano H, Hirayama K, Sadamitsu Y, Toshimitsu A, Fujita H, Shin S, et al.: Monitoring sound to quantify snoring and sleep apnea severity using a smartphone: proof of concept. J Clin Sleep Med. **10**(1), 73–78, 2014.
20. Berlin DA, Gulick RM, Martinez FJ: Severe Covid-19. N Engl J Med. **383**(25), 2451–2460, 2020.
21. Wu Z, McGoogan JM: Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72 314 cases from the Chinese Center for Disease Control and Prevention. JAMA. **323**(13), 1239–1242, 2020.

Keita Fukuyama



Keita Fukuyama is currently a program-Specific Researcher in the Department of Real World Data Research and Development, Graduate school of Medicine, Kyoto University. He works at Kyoto university hospital as a bioinformatician supporting oncologists. He received his PhD form Graduate School of Medicine, Kyoto University in 2021. His research fields are medical information, health economics, and optimization of clinical procedures. He is a member of JSMO, JSS and JSGS.

Osamu Sugiyama

Osamu Sugiyama received his PhD from School of Science for Open and Environmental Systems, Keio University in 2009. He is currently a Program-Specific Associate Professor in the Department of Real World Data Research and Development, Graduate School of Medicine, Kyoto University. His research fields are Medical Informatics, Communication Robot, Robot Audition. His research interests are in machine learning utilizing medical information and health promotion with personal health records. He is affiliated with the Robotic Society of Japan, the Japanese Society for Artificial Intelligence and the Architectural Institute of Japan.

Kazuo Chin

Kazuo Chin, MD, PhD graduated from the Kyoto University School of Medicine in 1981 and received his doctoral degree in internal medicine from Kyoto University Graduate School of Medicine in 1990. He was a program Specific Professor in the Department of Respiratory Care and Sleep Control Medicine, Graduate School of Medicine, Kyoto University from 2008 to 2021. He is currently a Professor in the Department of Sleep Medicine and Respiratory Care, Division of Sleep Medicine, Nihon University of Medicine. At the same time, he works part-time in the Department of Human Disease Genomics, Center for Genomic Medicine, Graduate School of Medicine, Kyoto University. His research interests are respiratory care and sleep medicine including sleep disordered breathings.

Susumu Satou

Susumu Satou is currently an Associate Professor of the Department of Respiratory Care and Sleep Control Medicine, Graduate School of Medicine, Kyoto University. He is a head of a section of sleep laboratory of Kyoto University Hospital. His subspeciality is Respiratory Medicine, and he received his PhD in 2004. His current research fields are Respiratory Medicine, Respiratory Mechanics, Biomedical Engineering, and Image Analysis, including Outcome research. He is a member of ATS, ERS, APSR, AARC, JRS, JSSR, JSRCM, JSRCR, JARM, JSA, and JSIM. He is a deputy head of Assembly of Respiratory Structure and Function of APSR (Asia-Pacific Society of Respirology).

Shigemi Matsumoto

Shigemi Matsumoto is currently a program- Specific Professor in the Department of Real World Data Research and Development, Graduate school of Medicine, Kyoto University. He works at Kyoto university hospital as one of head of medical oncologist. His subspecialty is gastrointestinal, hepatobiliary and rare cancer. He received his PhD form Graduate School of Medicine, Kyoto University in 2009. His research fields are medical information and health utilizing Real World Data (RWD) in Oncology. He is a member of JSMO, JSCO, ASCO, ESMO and JAMI.

Manabu Muto

Dr. Manabu Muto is Professor of Department of Therapeutic Oncology, Graduate School of Medicine at Kyoto University form April 2013, and also a Head of Department of Clinical Oncology at Kyoto University Hospital Cancer Center, a Head of Hospital Cancer Division, Department of Cancer Center Kyoto University Hospital, a Department of Cancer Genome Medicine Kyoto University Hospital, a Director of Center of Clinical Bio Resource Center Kyoto University Hospital, a Director of Center of Kyoto Innovation Center for New Generation Clinical Trials and iPS Cell Therapy Kyoto University Hospital, and also Advisor to a Director of The Hospital, Kyoto University Hospital.