

E-021

# Automatic Cough Detection with Using Throat Microphone for Health Monitoring

Bhuiyan Rakib<sup>1</sup>, Kazuhiro Yoshimura<sup>1</sup>, Shingo Kuroiwa<sup>2</sup>, Masafumi Nishida<sup>1</sup>, Masafumi Nishimura<sup>1</sup>

<sup>1</sup>Graduate School of Integrated Science and Technology  
Shizuoka University, Japan

<sup>2</sup>Graduate School of Science and Engineering  
Chiba University, Japan

**Abstract**— Cough is a symptom of many respiratory and non-respiratory disorders. Cough sound has been the subject of several investigations. We simulated real-life data in this work, which included not just coughing but also speaking, eating-related sound events such as chewing and swallowing, and the subject's movement-related sound. Furthermore, cough data was induced using the nebulizer, despite the fact that most prior investigations relied on pretended cough data. In a real-world context, we aimed to show the usefulness of automatic cough detection and tried to detect coughing in dysphagia patients triggered by a mealtime mis-swallow. As a result of cough detection in real environment, we obtained 0.82 in F-measure by using multi-layer LSTM and a throat microphone. In the case of a dysphagia patient, we obtained 0.62 in the F-measure using the same method.

**Keywords**—long short-term memory (LSTM), nebulizer, throat microphone, dysphagia

## I. INTRODUCTION

The most common condition for which individuals seek medical care is coughing. This is motivated by the urgent need to assess the severity of cough and therapy efficacy in patients with respiratory issues, as well as the well-documented inaccuracy of cough evaluation based on patient self-reports [1]. Many investigations have been made in case of cough detection. A study employed using volunteering coughs [2] to identify cough using a CNN-based technique. They have recorded data on a number of mobile phones, which has raised privacy concerns. Because it can capture both conversations and coughs. Furthermore, it is self-evident that the mobile phone's setting and position have a direct impact on the audio data quality. Another study [3] was conducted on recognizing cough, including a comparison of CNN and RNN in the instance of cough detection. They used cough data that was generated by feigning it, and each participant read the identical phonetically balanced voice cues, which is unnatural. Because a person's communication in the actual world is spontaneous and unpredictable. Another concerning issue dysphagia which is a condition caused by a lack of nerve or muscle control. Dysphagia can make it difficult to eat and drink, and it can lead to aspiration pneumonia. It causes difficulty in swallowing, which leads to intense coughing. Fujita et al. [4] sought to identify a dysphagia patient's cough using throat microphone and SVM. But in the F-measure, he achieved a score of 0.21. However, in this work, we attempted

to identify cough in a real-world setting and to increase the detecting accuracy of cough detection in dysphagia patients with using multi-layer LSTM and throat microphone assuring the privacy issue.

## II. DATA ACQUISITION

Coughs were captured in a genuine circumstance utilizing 1ch throat microphone, as shown in Fig.1 The connection of the throat microphone to the subject's throat is shown in red circle. We chose the dining episode to create a realistic situation since the individuals ate (our subject ate different types of food including rice, cookie, chocolate etc.), drank, talked (with another person), and moved about freely. In the meantime, our subjects have used nebulizer and 1% sodium citrate shown in Fig.2 for inducing the real coughs. At the time of eating, we gathered acoustic data from 18 participants (16 males and 2 females in their early twenties). Each recording included one complete eating episode covered around 30-35 minutes. Our dataset has a length of around 571 minutes and 1000 cough events. We also obtained 419 cough incidences from a total of 55 people who were suffering from various diseases (cerebrovascular, congenital cerebrovascular, Dementia, Encephalitis, amyotrophic lateral sclerosis) using a throat microphone in the hospital environment.

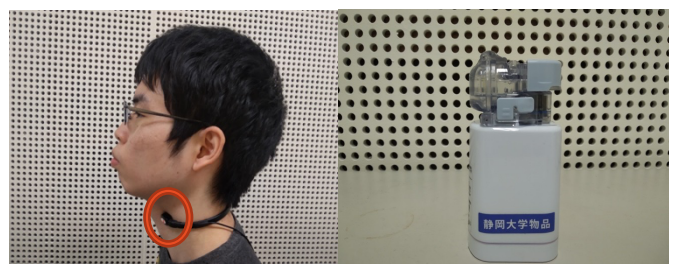


Fig. 1

Fig. 2

Fig.(1, 2). Gathering data using throat microphone and with using nebulizer and 1% sodium citrate

Using a throat microphone and a tablet in the patient's home setting [4], we acquired around 760 cough data from one dysphagia patient. Recording was performed during breakfast (20-30 minutes) around 1 month. The data collection is limited to breakfast time every day, because breakfast time is less affected by fatigue due to daily speech and behavior than other food intaking time. We [5] also have shown the effectiveness of throat microphone in case of cough detection in real environment. Other important fact is that the throat

microphone only records the target person's speech. The subject basically agrees to this recording; thus, it means that we will not invade other person's privacy besides the target person.

### III. CLASSIFICATION MODEL

In our research, we tried to identify cough in a real-world scenario as well as cough in dysphagia patients caused by a mis-swallow during mealtime. For our purpose, we used multi-layer LSTM. We created a realistic environment that mimicked a human eating episode that includes eating, drinking, talking, coughing, and normal movement of the human. We gathered cough data from a dysphagia patient in his home environment during normal eating, speaking, and moving at his own comfort. We prepared a 4-layer LSTM model under the condition, 39-dimensional MFCC as feature extraction technique, frame-based approach and frame size was 25ms with overlapping of 15ms and performed a three class-classification of cough sounds, other sound events and silence. The Adam optimizer was used to minimize the cross-entropy loss in case of LSTM. The LSTM model was trained through Mini-Batch with a batch size of 200 and the epoch size was 150.

### IV. EVALUATION

A 3-class (cough, other sound events (chewing, swallowing, speech, movement-related sound of the subjects) and silence) classification was conducted. We performed cross validation using the cough data of 18 subjects for multi-layer LSTM model. For training the models we used data of 17 subjects that having total length around 505 minutes includes around 950 cough sounds (around 14 minutes in total length), around 163 minutes of speech and 134 minutes of eating-related sound including chewing and swallowing. We have tested our model using the rest of the subject having length around 30-35 minutes including 50-60 cough sounds (around 45s in length) and 7-10 minutes of speech and around 6-10 minutes of eating-related sound. Using the same technique, we attempted to recognize dysphagia patient cough events using healthy person cough episodes (about 1000 cough events) as train data. Unfortunately, our F-measure was 0.47. Then We utilized 419 cough episodes collected from healthy persons and another 419 cough episodes collected from different patients as our train data for dysphagia cough identification. Table 1 demonstrates the results of cough detection evaluation in a real-world setting of healthy people and for a dysphagia patient utilizing a throat microphone and a multi-layer LSTM.

Table 1. Frame-based classification result of detecting cough events by using multi-layer LSTM in case of real world environment and dysphagia patient

Microphone	Healthy Control			Dysphagia Patient		
	Recall	Precision	F1-Score	Recall	Precision	F1-Score
Throat	0.85	0.79	0.82	0.67	0.57	0.62

In the case of cough detection in a real-world setting, we achieved a pretty good result. In the instance of dysphagia patient cough detection, the result is not that much good for trailing data, but it has improved a long way comparing to the prior research [4].

### V. CONCLUSION

Our objective is to realize privacy-protected health-monitoring system which can detect speaking, eating and coughing. We have also shown the feasibility study of automatic cough detection in real environment and attempted to detect cough of dysphagia patient with using throat microphone and deep neural network. Throat microphone also ensures the privacy because it can only record target subject's related sound. This characteristic helps to prevent privacy invasion of people around the subject.

### ACKNOWLEDGEMENT

This work was partially supported by JSPS KAKENHI Grant Number 18H03260.

### REFERENCES

- [1] S. Birring, T. Fleming, S. Matos, A. Raj, D. Evans, and I. Pavord, "The leicester cough monitor: preliminary validation of an automated cough detection system in chronic cough," *European Journal*, vol. 31, no. 5, pp 1013-1018, 2008.
- [2] Filipe Barata, Kevin kipfer, Mauric Weber, Peter Tinschert Elgar Fleisch, Tobias Kowatsch "Towards Device-Agnostic Mobile Cough Detection with Convolutional Neural Network". *IEEE International Conference on Health Care Informatics*, 2019.
- [3] Justice Amoh, "Deep Neural Network for Identifying cough Sound". *IEEE Transaction on Biomedical Circuits and Systems*, vol 10, pp 1003-1011, 2016.
- [4] 藤田 祥太, 古川 大輔, 村西 幸代, 黒岩 眞吾, 西田 昌史, 西村 雅史, "嚥下機能低下の早期発見に向けた食事時の咳嗽自動検出", *IPSJ 82<sup>nd</sup> National Convention*, 5ZE-08,2020.
- [5] Bhuiyan Rakib, Masafumi Nishimura, "Automatic cough detection using Deep Neural Network and a Throat Microphone", *IIAI-AAI 2021* (to be published)