

Deep learning-based cough recognition model helps detect the location of coughing sounds in real time

14 August 2020

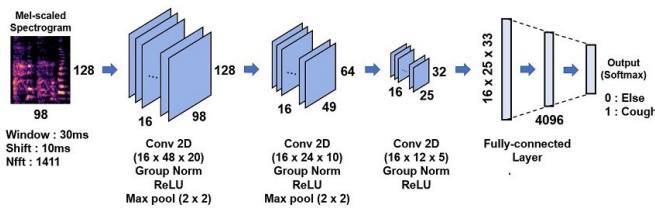


Figure 1. Architecture of the cough recognition model based on CNN. Credit: The Korea Advanced Institute of Science and Technology (KAIST)

The Center for Noise and Vibration Control at KAIST announced that their coughing detection camera recognizes where coughing happens, visualizing the locations. The resulting cough recognition camera can track and record information about the person who coughed, their location, and the number of coughs on a real-time basis.

Professor Yong-Hwa Park from the Department of Mechanical Engineering developed a [deep learning](#)-based cough recognition model to classify a coughing [sound](#) in real time. The coughing event classification model is combined with a [sound camera](#) that visualizes their locations in public places. The research team said they achieved a best test accuracy of 87.4%.

Professor Park said that it will be useful medical equipment during epidemics in public places such as schools, offices, and restaurants, and to constantly monitor patients' conditions in a hospital room.

Fever and coughing are the most relevant respiratory disease symptoms, among which fever can be recognized remotely using thermal

cameras. This new technology is expected to be very helpful for detecting epidemic transmissions in a non-contact way. The cough event classification model is combined with a sound camera that visualizes the cough event and indicates the location in the video image.

To develop a cough recognition model, a supervised learning was conducted with a convolutional neural network (CNN). The model performs binary classification with an input of a one-second sound profile feature, generating output to be either a cough event or something else.

In the training and evaluation, various datasets were collected from Audioset, DEMAND, ETSI, and TIMIT. Coughing and others sounds were extracted from Audioset, and the rest of the datasets were used as background noises for data augmentation so that this model could be generalized for various background noises in public places.

The dataset was augmented by mixing coughing sounds and other sounds from Audioset and background noises with the ratio of 0.15 to 0.75, then the overall volume was adjusted to 0.25 to 1.0 times to generalize the model for various distances.

The training and evaluation datasets were constructed by dividing the augmented dataset by 9:1, and the test dataset was recorded separately in a real office environment.

In the optimization procedure of the network model, training was conducted with various combinations of five acoustic features including spectrogram, Mel-scaled spectrogram and Mel-frequency cepstrum coefficients with seven optimizers. The performance of each combination was compared with the test dataset. The best test accuracy of 87.4% was achieved with Mel-scaled Spectrogram

as the acoustic feature and ASGD as the optimizer.

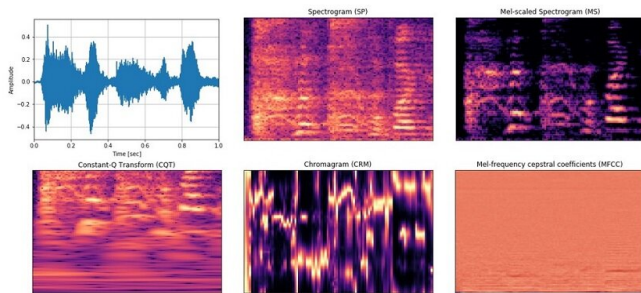


Figure 2. Examples of sound features used to train the cough recognition model. Credit: The Korea Advanced Institute of Science and Technology (KAIST)

distinguishes cough events and other events even in a noisy environment. In addition, it can track the location of the person who coughed and count the number of coughs in real time. The performance will be improved further with additional training data obtained from other real environments such as hospitals and classrooms.

Professor Park said, "In a pandemic situation like we are experiencing with COVID-19, a [cough](#) detection camera can contribute to the prevention and early detection of epidemics in public places. Especially when applied to a hospital room, the patient's condition can be tracked 24 hours a day and support more accurate diagnoses while reducing the effort of the medical staff."

Provided by The Korea Advanced Institute of Science and Technology (KAIST)

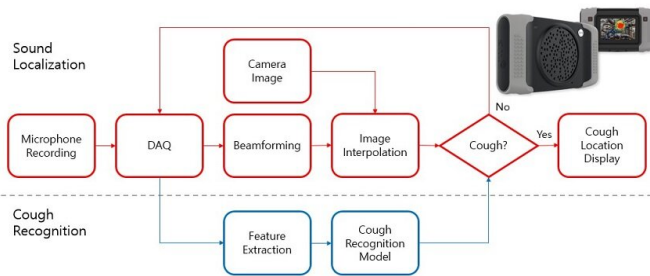


Figure 3. Cough detection camera and its signal processing block diagram. Credit: The Korea Advanced Institute of Science and Technology (KAIST)

The trained cough recognition model was combined with a sound camera. The sound camera is composed of a microphone array and a camera module. A beamforming process is applied to a collected set of acoustic data to find out the direction of incoming sound source. The integrated cough recognition [model](#) determines whether the sound is cough or not. If it is, the location of cough is visualized as a contour image with a 'cough' label at the location of the coughing sound source in a video image.

A pilot test of the cough recognition camera in an office environment shows that it successfully

APA citation: Deep learning-based cough recognition model helps detect the location of coughing sounds in real time (2020, August 14) retrieved 25 October 2020 from <https://techxplore.com/news/2020-08-deep-learning-based-recognition-real.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.