

# 호흡기 이상 조기 감지를 위한 사운드 분석 기술개발

홍혁기<sup>1\*</sup>, 정찬희<sup>1</sup>, 박진호<sup>1</sup>, 이효기<sup>2</sup>, 조에담<sup>2</sup>, 이용훈<sup>2</sup>

한국전자기술연구원<sup>1</sup>, (주)에이앤티솔루션<sup>2</sup>

## Development of Sound Analysis Technology for the Early Detection of Respiratory Abnormalities

Hyuckki Hong<sup>1\*</sup>, Chanhee Jeong<sup>1</sup>, Jinho Park<sup>1</sup>, Hyoki Lee<sup>2</sup>, Aedam Cho<sup>2</sup>, Yonghoon Lee<sup>2</sup>  
Medical IT Convergence Research Center, Korea Electronics Technology Institute, Korea<sup>1</sup>  
AnT Solution Corporation<sup>2</sup>

\*hkhong@keti.re.kr

### Abstract

Respiratory abnormalities manifested through cough can be broadly divided into upper airway (UA) and lower airway (LA) disorders. In addition, based on the duration of cough, they may be classified as acute cough—often caused by viral infections or allergies—subacute cough, such as post-infectious cough, and chronic cough, which is commonly associated with conditions including chronic obstructive pulmonary disease (COPD) and asthma. While previous research on cough sound analysis has mainly focused on the frequency of cough events, this study proposes a deep learning-based approach that not only quantifies the number of cough sounds but also classifies their duration. Using cough recordings from patients with viral infections, the model is developed and subsequently evaluated on sound data from patients with various respiratory disorders. The results suggest the potential of cough sound analysis as a non-invasive tool for the early detection and monitoring of respiratory abnormalities.

### 1. 연구 배경

호흡기 이상에 따른 기침 사운드는 폐와 기도를 보호하는데 도움이 되는 중요한 보호 반사 작용이다.[1] 기침은 주로 급성( $\leq 3$ 주), 아급성(3–8주), 만성( $> 8$ 주)으로 분류하는 임상적 기준이 있으며 바이러스나 세균에 의한 감염 상태에서의 기침은 급성으로, 이러한 질병에 의해서 잔류하는 기침을 아급성으로 분류한다.[2] 감염에 의한 급성 기침의 병태생리적 특징은 기도 점막의 염증 반응이 일어나고, 이로 인한 신경 수용기 자극이 기침 반사를 유도한다. 또한 점액 분비 증가 또는 점액의 변화, 점액 정체에 있을 경우 기침은 점액을 제거하는 역할을 하게 된다.[3]

본 연구에 중점은 감염성 기침 사운드의 연속성을 중심으로 임상적 중증도가 달라진다는 본 연구에서는 정상인의 기침 사운드와 감염성 상기도염(UA: upper airway) 및 하기도염(LA: lower airway)의 기침 사운드를 이용하여 기침의 회수와 정도에 따른 연속성의 차이를 분류하는 머신러닝 기술을 개발한다. 또한 실제 환경에서 발생할 수 있는 다양한 생활환경 노이즈가 있는 상태에서의 탐지를 비교한다.

### 2. 연구 방법

개발에 사용된 기침 사운드는 오픈데이터 세트인 'CougVid'의 연령별, 성별, 지리적 위치, COVID-19 상태를 나타내는 25,000개 이상의 크라우드 소싱 사운드 데이터 중 4명의 숙련된 의사가 분류한 2,800개 사운드 데이터와 google에서 공개한 데이터 셋을 레이블한 fluSense project Audio set[4]을 이용하였다. 사용된 데이터에는 기침, 재채기, 콧물, 목 가다듬기, 대화, 다양한 환경 노이즈가 포함되어 있다.

개발에 적용된 사운드는 입력 사운드 비교를 위하여 PCM mono 16 bit와 PCM mono 8 bit 포맷의 사운드를 모든

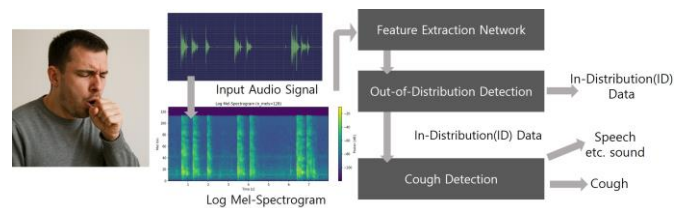


그림 1. 기침 검출을 위한 구성 흐름도

사운드에 적용하여 비교하였다. 이는 추후 실시간 저사양 인공지능 분류를 위한 성능비교를 위하여 수행하였다.

신호처리의 과정은 오디오 신호를 로그 멜 스펙트로그램(Log Mel-Spectrogram)으로 변환 후 특징을 추출하는 네트워크를 구축하였다. 멜 스펙트로그램을 이용한 변환의 경우 비선형적 특성을 보이거나 1KHz 까지는 선형에 가까우며, 1KHz를 넘어서면 스케일이 압축되는 특성을 나타낸다. 이러한 특성을 이용하여 기준 값을 1KHz(임계 값보다 40dB 높은 주파수)로 설정하였다. 그 다음 OOD(Out-of-Distribution) 검출 기를 사용하여 ID 및 OOD 데이터를 인식하였다. ID 데이터는 기침과 음성으로 분류되고 OOD 탐지와 기침 탐지에 동일하게 사용되었다. FENet, ResNet 및 VGG는 이미지와 신호에서 특징을 추출하는 데 사용되는 대표적인 신경망 프레임이다. 개발에 사용된 시스템을 설계하기 위해 세 가지 모델이 내장된 시스템을 각각 사용하여 성능을 비교하였다. FENet은 다양한 주파수 특성을 추출할 수 있는 CNN 기반 모델로 오디오 데이터 분석에 자주 사용되고 ResNet은 훈련 효율성을 향상시키고 기울기 소실(gradient vanishing)문제를 최소화 하기 위해 잔여 블록을 추가하는 CNN 기반 모델이기도 하다. VGG Net에는 피라미드 모양의 후속 컨볼루션 레이어(convolutional layers)도 포함되어 있어 컴퓨터 비전 작업에 서 좋은 결과를 얻을 수 있다.

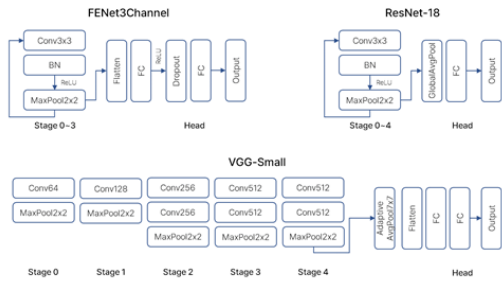


그림 2. 모델 아키텍처

### 3. 연구 결과

입력은 16 kHz 파형을 50% overlap 윈도우로 분할하고, STFT(25ms/10ms/FFT512)-Mel(64bins)-로그압축(40dB클리핑)-정규화를 적용하였다. 분류 대상(ID)은 기침 vs 비기침(또는 기침 vs 음성)이며, 재채기·목 가다듬기·대화·환경소음 등은 OOD로 정의하였다. 학습은 동일한 파라미터(AdamW, cosine LR, label smoothing 0.1)로 수행하였다.

ResNet-18이 소폭의 최고 F1/Acc를 보였으나, FENet-S는 파라미터 16배 감소, 연산량 3배 감소, 지연 38% 감소로 효율 대비 손실이 작다. VGG-Small은 견고한 기준선이나 지연·연산 비용이 크다.

표 1. 모델 별 분류 성능

Model	FENet-S	ResNet-18	VGG-Small
Acc [%]	92.6	<b>93.4</b>	91.8
F1_cough [%]	92.1	<b>93.0</b>	91.2
ROC-AUC	0.967	<b>0.972</b>	0.960
Prec/Rec [%]	92.8 / 91.4	<b>93.7 / 92.3</b>	91.9 / 90.5
Params [M]	<b>0.7</b>	11.7	19.0
MACs [G]	<b>0.20</b>	0.60	2.50
Latency [ms] (FP32, BS=1)	<b>7.8</b>	12.5	19.6

### 4. 결론 및 향후 과제

본 연구는 공개 데이터와 표준 멜 스펙트로그램 전처리를 기반으로 FENet-S/ResNet-18/VGG-Small을 비교하였다. FENet-S는 지연·연산 효율을 크게 확보하면서 ResNet-18과 경쟁적인 성능을 달성하였다.

향후 연구로는 (1) 도메인/장치 일반화를 위한 도메인 적응·테스트타임 적응, (2) 오디오 자기지도 사전학습(SSL) 기반 파인튜닝, (3) 라벨 합치도 보고 및 불확실성 정량화, (4) 고도화된 OOD, (5) 지표 표준화, (6) 온디바이스 최적화(QAT + distillation, structured pruning, 그래프 최적화)가 요구된다.

### 5. Acknowledgements

이 연구는 2025년도 산업통산자원부 및 한국산업기술기획평가원(KEIT) 연구비 지원에 의한 연구임 (RS-2024-004206 11)

### 6.참고 문헌

- [1] Kian Fan Chung, Stuart B Mazzonec, Lorcan McGarveyd, Woo-Jung. “SongChronic cough as a disease: implications for practice, research, and health care.” *The lancet respiratory Medicine*, Vol 13, Issue 2, p.110-112, 2025
- [2] Richard S. Irwin, Cynthia L, Anne B. Chang, Kenneth W. Altman. “Classification of Cough as a Symptom in Adults

and Management Algorithms. “ *Chest journal*, Vol 153, Issue 1, p.196-209, 2018

[3] Richard S. Irwin, Michael H. Baumann, Donald. Bolser, “Diagnosis and Management of Cough Executive Summary” *Chest journal*, Vol 129, Issue 1, 1S-23S, 2006

[4] Yuhan Chen, Pankaj Attri, Jeffrey Barahona, Michelle L. Hernandez, Delesha Carpenter, Alper Bozkurt. “Robust Cough Detection with Out-of-Distribution Detection.” *IEEE Journal of Biomedical and Health Informatics*, Vol 27, Issue 7, 3210-3221, 2023