

# 소리와 가속도 데이터를 이용한 멀티모달 기침 감지 모델

강재식<sup>\*,</sup>, 백문기, 최형탁, 윤승원, 이규철<sup>+</sup>  
 충남대학교 컴퓨터공학과

e-mail : {steelroadk, znmzment, poo4967, yoonenoch11, kclee}@cnu.ac.kr

## Multimodal Cough Detection Model Using Audio and Acceleration Data

Jae-Sik Kang<sup>\*</sup>, Moon-Ki Back, Hyung-Tak Choi, Yoon-Seung Won, Kyu-Chul Lee<sup>+</sup>  
 Department of Computer Engineering, Chungnam National University

### 요 약

전 세계적으로 인플루엔자에 의해 매년 29~64만의 사망자가 발생하며 사회, 경제적 피해를 일으키고 있다. 기침에 의해 생성된 비말은 인플루엔자의 주요 전파 방법으로, 기침 감지 기술을 통해 확산 방지가 가능하다. 이전의 기침 감지에 대한 연구는 기침 소리와 전통적인 기계학습기법을 사용하였다. 본 논문은 기침 소리와 더불어 기침 시 발생하는 신체의 움직임 정보를 동시에 학습하는 멀티모달 딥러닝 기반의 기침 감지 모델을 제안한다. 도출된 모델과 기존의 모델과의 성능 비교를 통해 제안한 모델이 이전의 기침 감지 모델보다 정확한 기침 인식이 가능함을 보였다. 본 논문이 제안하는 모델은 스마트 워치와 같은 웨어러블 기기에 적용되면 인플루엔자의 확산 방지에 크게 기여할 수 있을 것이다.

### 1. 서론

기침은 인플루엔자와 같은 호흡기 병원균의 주요 전파 방법으로 현대의 밀집된 주거형태는 인플루엔자 전파 방지에 취약할 수밖에 없다. 인플루엔자는 에어로졸 물질인 비말 핵을 통해 전파되는데, 이는 기침이나 대화뿐만 아니라 단순 호흡을 통해서도 쉽게 생성되기 때문에 전파 방지가 대단히 어렵다. 인플루엔자는 매년 전 세계적으로 크고 작은 유행을 일으키는데, 유행이 시작되면 통상 2~3주 이내에 인구의 10~20%가 감염[1]되고, 전 세계적으로 매년 29만에서 64만을 죽음에 이르게 한다[2]. 이에 각국은 자국의 상황에 맞는 인플루엔자 대유행 지침을 만들고 이에 대비하고 있지만, 완벽히 예방하지 못하고 있다[3].

이전의 기침 감지를 위한 연구들은 HMM(Hidden Markov Model)과 SVM(Support Vector Machine)와 같은 전통적인 기계학습을 사용하였다[4, 5]. 반면 연구가 이루어지고 있는 딥러닝은 상대적으로 데이터 특징 추출 및 분석에 필요한 인력을 많이 필요로 하는 전통적인 기계학습 방식을 대체하려는 방법으로 모델에서 데이터에 대한 특징을 분석하고 학습한다. 기침 감지에서도 딥러닝을 이용한 연구들이 진행되고 있다[6-9]. 통상 기침이 발생하면 기침에 따른 소리와 신체의 움직임이 나타나지만, 이전의 연구들은 기침으로 발생하는 소리 데이터만을 사용하고 있다. 신체의 움직임은 가속도 센서를 통해 감지가 가능하

며, 가속도 센서를 이용하여 신체의 움직임을 예측하는 연구도 진행되었다[10]. 딥러닝 분야에서 서로 다른 모달리티(Modality)를 가진 여러 데이터를 하나의 모델로 학습하는 멀티모달(Multimodal) 방식은 이미 제시된 방법[11]으로 기침 감지에 기침 소리 데이터와 가속도 데이터를 동시에 이용하는 모델 설계가 가능하며 더욱 높은 기침 감지 성능을 도출할 수 있다.

본 논문의 구성은 2장에서 관련 연구에 대하여 소개하고 3장에서는 멀티모달 기침 감지 모델에 대해 다룬다. 4장에서 모델의 성능을 비교하고 5장에서 향후 연구를 제시한다.

### 2. 관련 연구

<표 1> 딥러닝 기반의 기침 감지 연구

	Feature Extraction	Classifier	Sensitivity	Specificity	Accuracy
[6]	MFCC	NN	82%	96%	-
[7]	MFCC	DNN	77%	97%	95%
[8]	STFT	RNN	84.2%	75.2%	79.9%
[9]	STFT	CNN	94.0%	91.7%	-

<표 1>은 딥러닝 기법이 적용된 기침 감지모델의 성능을 정리한 표이다. 해당 연구들은 분류기(Classifier)의 기침과 비기침 분류 성능을 향상하기 위해 기침 소리 데이터로부터 특징 데이터를 추출하는 방식을 사용하였다. 소리 데이터에서 특징 추출을 위해 STFT(Short-Time Fourier Transform)와 MFCC(Mel-Frequency Cepstrum Coefficients) 알고리즘이 사용되었고 분류기는 DNN(Deep Neural Network), NN(Neural Network), RNN(Recurrent Neural

\* 본 연구는 과학기술정보통신부 및 정보통신기술진흥센터의 SW중심대학지원 사업의 연구결과로 수행되었음(2015-0-00930)

+ 교신저자

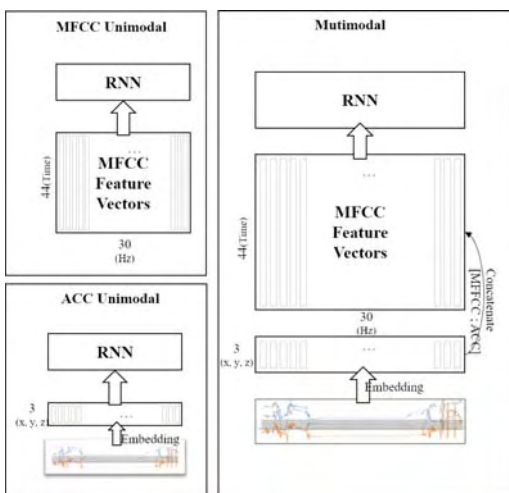
Network), CNN(Convolution Neural Network) 방식들을 사용하였다. MFCC와 STFT는 기본적으로 시간-주파수의 특징을 추출하는 공통점이 존재하지만 MFCC는 인간의 귀가 가지는 비선형적인 주파수 특성을 고려한 Mel Filter Bank를 적용한 알고리즘이기 때문에 높은 주파수 대역에 대한 특징 추출은 STFT 보다 약하다. 본 논문에서 수행한 MFCC 단일모달 모델의 도출결과를 통해 기침 데이터의 특징 추출은 MFCC가 STFT보다 좋은 성능을 발휘함을 보였으며 분류기 자체의 성능은 CNN이 RNN보다 좋은 성능을 보였다.

### 3. 기침 감지 모델

#### 3.1 기침 데이터 수집 및 구성

인플루엔자 감염자의 기침 소리 데이터와 가속도 데이터를 수집하기 위해 마이크로폰과 가속도 센서가 탑재된 삼성전자의 스마트 워치 기어 S3를 사용하였다. 해당 기기를 인플루엔자에 감염된 18명(남성 12명, 여성 6명)에게 착용하여 약 7일간의 기침 소리와 가속도 데이터를 수집하였다. 수집된 소리 데이터에서 기침을 1초 단위로 추출했으며 이를 30차 MFCC를 통해 특징 벡터로 추출하였다. 가속도 데이터는 기침 전후 1초씩 총 3초에 해당하는 데이터를 추출하였고, 임베딩(Embedding)과정을 거쳐 3×30 행렬로 구성하였다. 구성된 학습데이터 세트(기침 소리와 해당 시점의 가속도 값이 쌍을 이루는 데이터)는 총 43,610개이며 기침 데이터 세트의 개수 4,361개, 비기침 데이터 세트 39,249개로 1:9의 비율을 가진다.

#### 3.2 모델 설계



(그림 1) 기침 감지를 위한 단일모달(좌)과 멀티모달(우) 방식의 기침 감지 모델 구성도

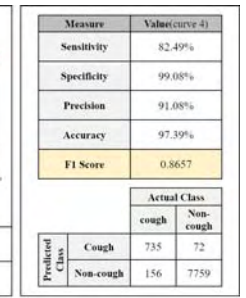
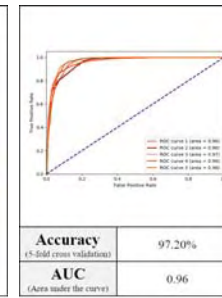
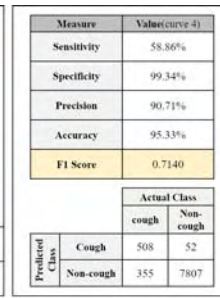
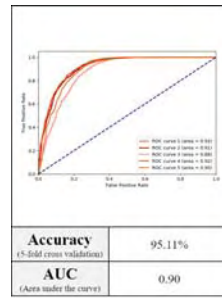
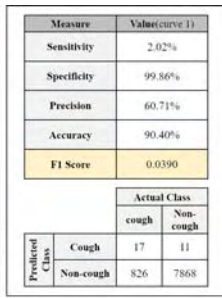
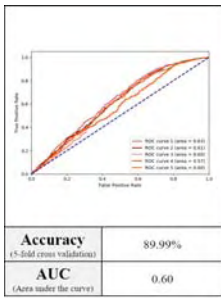
멀티모달 방식 적용 여부에 따른 성능의 변화 및 성능을 평가하기 위해 (그림 1)과 같은 모델들을 설계하였다. MFCC 단일모달은 30차 MFCC를 통해 특징이 추출된 기침 소리 데이터만을 이용한 단일모달 기침 감지 모델이며, ACC 단일모달은 기침에 의해 발생한 속도 데이터만을 이

용한 단일모달 기침 감지 모델이다. 멀티모달은 30차 MFCC를 통해 추출된 기침 소리 특징 데이터(44×30 행렬)에 가속도 데이터(3×30 행렬)를 결합(Concatenate)한 결합 데이터(47×30 행렬)를 이용한 멀티모달 방식의 기침 감지 모델이다. 분류기는 RNN의 일종인 LSTM(Long Short-Term Memory)을 이용했으며, 해당 LSTM을 96개의 히든 유닛(Hidden Unit)으로 이루어진 3개 레이어(Layer)와 2개의 출력(Output)으로 구성하였다. 2개의 출력은 다시 소프트맥스(Softmax)를 거쳐 하나의 바이너리 값을 가진다. 분류기는 세부적으로 기침 감지 모델(MFCC 단일모달, ACC 단일모달, 멀티모달)에 사용되는 학습 데이터 구조에 따라 서로 다른 입력 크기와 시간단계(Timestep)을 가진다. 각각의 모델에 구성된 입력과 시간단계의 경우 MFCC 단일모달 모델은 44 입력 크기에 30 시간단계로 구성하였으며, ACC 단일모달 모델은 3 입력 크기에 30 시간단계로 구성하였다. 멀티모달은 47 입력 크기에 30 시간단계로 구성하였다.

#### 4. 도출된 모델 성능 비교

각각의 모델은 5-폴드 교차 검증(5-fold Cross Validation)을 적용하여 학습 및 검증을 수행하였으며, 성능 평가에 평가지표 민감도(Recall), 정밀도(Precision), F1-점수(F1-Score)를 사용하였다.

(그림 2)는 기침 가속도 데이터만을 이용한 단일모달의 성능지표이고, (그림 3)는 MFCC로 추출한 기침 소리 데이터만을 이용한 단일모달 학습 모델의 성능지표이다. (그림 2)의 기침 가속도 데이터만을 학습한 모델의 경우는 평균 정확도는 89.99%, AUC는 0.60 기록하였다. 민감도, 정밀도로 봤을 때 민감도는 2.02% 정밀도는 60.71%로 해당 모델이 기침, 비기침에 대한 분류에서 좋은 성능을 발휘하지 못했다. (그림 3)의 기침 소리만을 학습한 모델의 경우 평균 정확도는 95.33%, AUC는 0.90을 기록했으며, 민감도, 정밀도로 봤을 때 민감도는 58.86% 정밀도는 90.71%의 결과 값을 보였다. 해당 수치는 모델이 기침에 대한 분류보다 비기침에 대한 분류 성능이 더 우수함을 나타낸다. 또한 기침 소리 데이터가 기침 감지에 있어서 중요한 요인으로 사용 가능함을 확인가능 하였다. (그림 4)는 본 논문이 도출한 기침 소리 데이터와 가속도 데이터를 결합하여 학습한 멀티모달 방식의 모델의 성능지표이다. 모델의 평균 정확도는 97.20%, AUC는 0.96으로 기침 소리 데이터만으로 학습한 모델보다 향상된 성능 결과를 얻었다. 해당 모델의 민감도 82.49%, 정밀도 91.08%, F1-점수 0.8657이며, 해당 수치를 기침 소리를 단독으로 학습한 모델과 비교 했을 때, 기침과 비기침 분류 모두에서 성능이 단일모달을 이용한 모델보다 향상되었다. F1-점수는 0.8657로 단일모달 모델에 비해 보다 안정적인 기침 감지 성능을 발휘하는 모델임을 증명하였으며, 또한 단일모달로 구성하였을 때 성능이 낮게 나타난 기침 가속도 데이터가 본 논문이 사용한 결합 기법으로 결합 시 전체



(그림 2) ACC 단일모달 성능지표

(그림 3) MFCC 단일모달 성능지표

(그림 4) 멀티모달 성능지표

적인 분류 성능과 모델의 안정성이 향상됨을 확인하였다. <표 1>의 이전에 연구된 다른 논문과의 비교에서 이전 연구 논문에서 정확한 성능지표가 공개되지 않아 본 논문에서 사용된 평가 지표를 모두 비교할 수 없었으나, 공개된 성능지표들만을 이용하여 성능을 비교하였다. 그 결과 [6-8] 논문의 경우와 비교하였을 때, 본 논문이 제시하는 멀티모달 방식의 기침 감지 모델이 해당 논문에서 사용된 단일모달 방식보다 기침과 비기침을 분류하는 성능이 더 우수함을 확인 가능하였다. 또한 공개된 성능지표로만 보았을 경우 가장 우수한 성능을 보이는 [9] 논문과도 본 논문이 제시하는 방식이 성능상으로 동등하거나보다 우수한 것으로 판단되었다.

본 논문이 제시한 자체 성능 평가와 타 논문과의 비교분석을 통해 제시하는 멀티모달 방식의 기침 감지 모델이 보다 안정적인 모델 구성이 가능하며, 기침과 비기침 감지 성능 또한 우수함을 보였다. 또한 단일모달 방식의 기침 감지 모델에서 낮은 성능을 보인 가속도 데이터가 본 논문이 제시하는 소리 데이터와 결합 기법을 이용하여 멀티모달 모델을 구성했을 때 기침 소리 또는 가속도 데이터만을 학습한 단일모달 모델보다 높은 성능을 발휘함을 본 논문에서 보였다.

**5. 결론 및 향후 연구**

본 논문에서는 실제 데이터를 이용한 인플루엔자 감염자의 기침에서 발생하는 소리 데이터와 가속도 센서 데이터를 이용한 위한 멀티모달 방식의 기침 감지 모델을 제안하였다. 특히 본 모델은 현실에 실현이 가능한 기침의 식별을 위해 실험환경의 데이터가 아닌 실제 데이터를 수집 이용하는 모델을 설계하였다. 제안된 모델은 기침 소리 데이터에서 MFCC 특징 데이터를 추출한 후, 이를 가속도 데이터와 결합하여 멀티모달 방식을 이용했으며, RNN 기반의 LSTM 모델 학습을 통해 기침과 비기침 중 하나로 식별하였다. 또한 실험을 통해 실제 특징 데이터로써 효용성이 낮은 가속도 데이터가 소리 데이터와 상관관계가 존재함을 확인하였다.

향후 연구과제로 일반적인 감기 또는 흡연 등으로 발생하는 습성 기침(Wet Cough)과 바이러스 등으로 발생하는 건성기침(Dry Cough)을 구분할 수 있는 후속 연구들이 요구되며 보다 높은 성능을 보일 수 있는 멀티모달 기법

에 대한 후속연구 또한 진행할 계획이다.

**참고문헌**

[1] 질병관리본부 국립보건연구원, "감염병 실험실진단", 2005

[2] CDC(Centers for Disease Control and Prevention), "Seasonal flu death estimate increases worldwide", 2017

[3] Jong Koo Lee, et al. "Government Measures against Pandemic Influenza", Korean Medical Association, 2010

[4] S. Matos, et al. "Detection of cough signals in continuous audio recordings using hidden Markov models", IEEE Transactions on Biomedical Engineering, 2006

[5] Thomas Drugman, et al. "Assessment Of Audio Features For Automatic Cough Detection", European Signal Processing Conference, 2011

[6] Martinek J, et al. "Cough Sound Analysis", Acta Medica Martiniana 2013

[7] Marcel Młynczak, et al. "Automatic cough episode detection using a vibroacoustic sensor", International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2015

[8] Justice Amoh, et al. "Deep Neural Networks for Identifying Cough Sounds", IEEE Transactions on Biomedical Circuits and Systems, 2016

[9] Justice Amoh, et al. "DeepCough: A Deep Convolutional Neural Network in A Wearable Cough Detection System", IEEE Biomedical Circuits and Systems Conference (BioCAS), 2015

[10] Nishkam Ravi, et al. "Activity Recognition from Accelerometer Data", Aaai, 2005

[11] Jiquan Ngiam, et al. "Multimodal Deep Learning", ICML, 2011