

이상 호흡음 탐지를 위한 딥러닝 활용

변규린¹, 양희규¹, 추현승^{1,2*}

¹성균관대학교 AI시스템공학과

²성균관대학교 전자전기컴퓨터공학과

{byungyurin21, huigyuu, choo}@skku.edu

Harnessing Deep Learning for Abnormal Respiratory Sound Detection

Gyurin Byun¹, Huigyu Yang¹, Hyunseung Choo^{1,2}

¹Dept. of AI Systems Engineering, Sungkyunkwan University

²Dept. of Electrical and Computer Engineering, Sungkyunkwan University

요약

Deep Learning(DL)을 사용한 호흡음의 자동 분석은 폐 질환의 조기 진단에 중추적인 역할을 한다. 그러나 현재의 DL 방법은 종종 호흡음의 공간적 및 시간적 특성을 분리하여 검사하기 때문에 한계가 있다. 본 연구는 컨볼루션 연산을 통해 공간적 특징을 캡처하고 시간 컨볼루션 네트워크를 사용하여 이러한 특징의 공간적-시간적 상관 관계를 활용하는 새로운 DL 프레임워크를 제안한다. 제안된 프레임워크는 앙상블 학습 접근법 내에 컨볼루션 네트워크를 통합하여 폐음 녹음에서 호흡 이상 및 질병을 검출하는 정확도를 크게 향상시킨다. 잘 알려진 ICBHI 2017 챌린지 데이터 세트에 대한 실험은 제안된 프레임워크가 호흡 이상 및 질병 검출을 위한 4-Class 작업에서 비교모델 성능보다 우수함을 보여준다. 특히 민감도와 특이도를 나타내는 점수 메트릭 측면에서 최대 45.91%와 14.1%의 개선이 이진 및 다중 클래스 호흡 이상 감지 작업에서 각각 보여준다. 이러한 결과는 기존 기술보다 우리 방법의 두드러진 이점을 강조하여 호흡기 의료 기술의 미래 혁신을 주도할 수 있는 잠재력을 보여준다.

1. 서론

호흡음은 폐 질환에 대한 필수 진단 정보를 제공하는 중요한 의학적 특성을 제공한다 [1]. 특히 호흡음들은 주파수, 피치 및 에너지와 같은 특성에 따라 정상 호흡음과 이상 호흡음 [2]로 구분된다. 이상 호흡음은 폐 질환과 관련된 정보를 전달하는 Wheeze와 Crackle로 구성된다. 기도 폐쇄로 인한 고음인 Wheeze는 천식, 만성폐쇄성 폐질환(COPD)과 같은 만성 질환에 널리 퍼져 있다. Crackle는 60-2000Hz 주파수 범위에 걸쳐 기도가 터지는 불연속적인 소리를 포함하며 일반적으로 비만성 질환과 관련이 있다.

폐 이상 및 만성 및 비만성 질환을 식별하는 일반적인 방법은 디지털 청진기를 사용하여 호흡음을 녹음하는 것이다 [3], [4]. 그러나 이 방법은 호흡음을 해석하는 전문가가 필요하다. 따라서 자동화된 호흡음 분석은 의료 전문가가 폐 상태의 조기 진단을 도울 수 있는 유망한 솔루션 [5]-[7]이다. 비록 자동화된 호흡 이상 및 질병 감지를 위한 초기 연구는 Machine Learning (ML)에 중점을 두었지만, 최근에는 DL기법이 호흡음을 분류하는 데 우수한 성과를 보여주고 있다. 종종 호흡음의 시공간적 특성을 분리하기 때문에 특징추출에 제한이 있다. 본 논문에서 시공간적 특징을 모두 추출할 수 있는 모델을 제안한다.

2. 관련 연구

ICBHI [13] 공공데이터 셋이 공개됨에 따라 호흡음 분류에 관한 ML과 DL 방법론들의 많은 연구가 활발히 이루어졌다. 로지스틱 회귀, 가우시안 혼합 모델, 랜덤 포레스트 및 히든 마코프 모델을 포함하여 비정상적인 호흡 소리를 식별하기 위해 많은 ML이 개발되었다 [7]-[11]. 공개적으로 액세스할 수 있는 데이터 세트가 부족하기에 제안된 특징추출 방법 및 알고리즘의 효과를 평가하는 것은 어렵다 [12].

DL에 기반한 최근 프레임워크, 특히 컨볼루션 신경망(CNN)은 데이터에서 공간 표현을 직접 추출하고 결합하는 방법을 학습할 수 있다. [14]는 변형 컨볼루션 자동 인코더를 사용하여 데이터 세트의 균형을 맞추기 위해 데이터 확대를 위한 CNN 모델을 도입했다. 달성된 결과는 3-Class 및 6 Class 폐 질환에 대한 99.3%와 99.0%의 정확도를 보인다.

본 연구에서는 최적의 이상 호흡음을 분류하기 위해 간단하면서도 효과적인 모델을 제안한다. 공공데이터를 이용하여 효과적인 데이터 특징추출 기법인 Mel Spectrogram, MFCC 방법을 통해 폐음의 특징적인 정보를 추출하고 딥러닝을 통해 이상 호흡음을 분류하고자 한다. TCN 기반의 병렬 구조와 스택 앙상블 모델을 조합하여 분류 성능의 우수함을 보여주었다.

3. 제안 모델

3.1 특징 추출

멜 스펙트로그램과 MFCC 추출과정은 다음과 같은 단계를 따른다. 먼저 호흡음 신호를 작은 크기의 프레임으로 자르며 시간 도메인을 주파수 성분으로 변환하는 고속 푸리에 변환을 적용하여 스펙트로그램을 얻는다. 그리고 저주파수 영역대에 주로 분포하는 호흡음의 특성을 고려하여 고주파수 대비 저주파수 영역을 더 세밀히 분석하는 멜 필터 बैं크를 통해 스펙트로그램을 멜 스펙트로그램으로 변환한다. 두 번째 음성 특징 벡터인 MFCC는 추출된 로그 멜 스펙트로그램에 이산 코사인 변환을 통하여 추출한다. 다음으로, ImageNet 데이터셋에서 사전 훈련된 가중치를 가진 VGG19 모델을 기본 모델로 사용한다. 이 모델은 입력 이미지에서 높은 수준의 특징을 추출한다.

3.2 제안모델

추출된 특징은 여러 TCN 블록을 통과한다. 잔차 연결 및 배치 정규화를 가진 컨볼루션 계층으로 구성된 이 블록은 데이터의 시간 의존성을 캡처하기 위해 서로 다른 확장 속도를 가진 TCN 계층을 활용하여 특징의 융합된 표현을 생성한다. 이러한 융합된 특징은 공간 차원을 줄이고 컴팩트한 표현을 얻기 위해 전역 평균 풀링을 거친다.

순환 구조인 LSTM 및 GRU는 개선된 메모리 유지력 때문에 다양한 시퀀스 모델링 작업[15], [16]에서 TCN보다 성능이 떨어진다. LSTM 및 이와 유사한 구조는 고정된 크기의 메모리를 유지하는 반면, TCN은 시간적 관계를 추출할 때 전체 시퀀스를 참조할 수 있다. 또한, 순환 구조를 제거함으로써 계산 비용이 감소한다.

3.3 Stacking Ensemble

본 연구에서는 단일모델의 단점을 보완할 수 있는 스택킹 기법을 사용한다. 이는 여러 모형에서 예측된 값들을 다시 학습 데이터로 사용하고 최종 예측모형을 만드는 방법으로 각 알고리즘의 장단점을 보완하여 일반적으로 단일모델보다 성능이 좋은 것으로 알려져 있다[17]. 스택킹 앙상블기법의 효율성을 높이고 모델들의 단점을 서로 보완할 수 있도록 K-Nearest Neighbors, DT, SVM, Gaussian Naive Bayes, RF, Gradient Boosting 분류 알고리즘을 조합하여 사용하였다. 각 모델의 결과를 결합하는 메타 모델은 Simple Linear Model에 해당하는 Logistic Regression을 사용한다.

4. 실험

본 연구 실험을 위해 손실함수는 Cross Entropy를 사용하였고, Optimizer 함수는 Adam으로 설정하였다. Adam Optimizer에 설정된 Learning rate은 $1 \cdot e^{-5}$ 로 초기 설정을 하였고, 모델 학습이 일정 에폭 이상 개선이 없을 시 ReduceLROnplateau 함수를 통하여 학습률을

$1 \cdot e^{-6}$ 까지 줄여 학습을 안정적으로 수렴하게 하였다[16]. 훈련단계에서 오버피팅을 방지하기 위하여 Early Stopping을 적용하였다.

4.1 Dataset

본 연구에서는 ICBHI 데이터 세트를 사용하여 이상 및 질병을 감지한다. 도전 과제의 데이터 세트인 ICBHI는 6898개의 호흡 주기를 포함하는 5.5시간의 녹음으로 구성된다. 3642개 Normal, 1864개 Crackle, 886개 Wheeze 나머지 506개는 126명 피험자의 Crackle와 Wheeze를 모두 포함한다. ICBHI 데이터 세트에서 호흡 주기의 길이는 0.2초에서 16.2초 사이로 다양한 920개의 주석이 달린 녹음이 포함되어 있다.

4.2 실험 결과

ICBHI [13]에 따라 동일한 평가 측정 기준을 채택한다. 데이터의 60%는 학습용으로, 40%는 테스트용으로 사용한다. 검증 데이터로는 학습 세트의 10%를 사용했다. 실험은 4-class (Normal, Crackle, Wheeze, Crackle and Wheeze)에 대해 진행한다. S_c 로 표시된 평가 점수는 민감도 S_c 및 특이도 S_p 점수의 산술 평균으로 결정된다. 이 접근 방식은 실제 양과 실제 음의 비율을 모두 고려하여 모델의 성능을 종합적으로 평가한다. 표1은 학습한 모델들의 Test 결과 평가지표를 정리한 것이다.

우리는 Bi-GRU, CNN-GRU 및 CNN-Bi GRU와 비교 분석을 수행하여 이상 호흡음 분류에 대한 효과를 측정한다. 비교모델과 결과를 분석하면 모든 평가지표에서 제안된 모델이 순환 및 CNN 변형 모델을 큰 차이로 능가할 수 있음을 보여준다. 기준 모델인 Bi-GRU(552,769), CNN-GRU(2,605,249) 및 CNN-Bi GRU(5,240,513)의 학습 가능한 매개변수 수에 비해 제안된 방법은 276,225개의 학습 가능한 매개변수만 포함한다.

〈표1〉 Evaluation results of the ICBHI

Methods	S_c	S_e	S_p
Bi-GRU	54.8	50.1	68.1
CNN-GRU	56.5	47.7	74.3
CNN-Bi GRU	51.6	58.8	70.4
Our	74.7	69.6	77.7

5. 결론 및 향후 연구 계획

본 논문에서는 제한된 데이터 세트여도 정확한 호흡 이상감지를 위해 설계된 강력하고 가벼운 프레임워크를 다룬다. 특히 ICBHI 데이터에 적합한 전처리 방식과 특징추출을 통하여 이상 호흡음을 효과적으로 분류할 수 있는 TCN 기반의 병렬 합성곱 신경망 모델을 제안하였다. 다양한 딥러닝 모델에서의 결과를 비교하였으며, 모든 성능지표에서 가장 우수한 결과를 나타내는 것을 확인할 수 있었다. 본 연구는 수집된 환자들의 호흡음 수의 부족과 세분화된 호흡질환

분류의 한계점이 있으므로, 추가로 데이터 증강기법 적용과 호흡 관련 질환 분류 모델을 위한 연구가 필요하다. 본 연구를 통해 향후 AI를 접목한 기술을 통해 폐질환 이상 여부를 조기에 스크리닝하여 신속한 치료가 가능해질 수 있으며 사회적 취약계층들에게 의료 접근 방식 개선 및 의료 서비스의 질 향상에 도움이 될 것으로 기대된다.

ACKNOWLEDGEMENT

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 ICT명품인재양성사업 (IITP-2023-2020-0-01821), 인공지능대학원 (성균관대학교, No.2019-0-00421), 인공지능 혁신 허브 연구 개발 (No.2021-0-02068)의 지원을 받아 수행된 연구임.

참고문헌

- [1] J. A. Dar, K. K. Srivastava, and S. A. Lone, "Spectral features and optimal hierarchical attention networks for pulmonary abnormality detection from the respiratory sound signals," *Biomedical Signal Processing and Control*, 2022.
- [2] Y. Choi and H. Lee, "Interpretation of lung disease classification with light attention connected module," *Biomedical Signal Processing and Control*, 2023.
- [3] L. Pham, H. Phan, R. Palaniappan, A. Mertins, and I. McLoughlin, "Cnn-moe based framework for classification of respiratory anomalies and lung disease detection," *IEEE journal of biomedical and health informatics*, 2021.
- [4] S. Gairola, F. Tom, N. Kwatra, and M. Jain, "Respirenet: A deep neural network for accurately detecting abnormal lung sounds in limited data setting," *IEEE EMBC*, 2021.
- [5] B. M. Rocha, D. Filos, L. Mendes, G. Serbes, S. Ulukaya, Y. P. Kahya, N. Jakovljevic, T. L. Turukalo, I. M. Vogiatzis, E. Perantoni et al., "An open access database for the evaluation of respiratory sound classification algorithms," *Physiological measurement*, 2019.
- [6] S. B. Shuvo, S. N. Ali, S. I. Swapnil, T. Hasan, and M. I. H. Bhuiyan, "A lightweight cnn model for detecting respiratory diseases from lung auscultation sounds using emd-cwt-based hybrid scalogram," *IEEE Journal of Biomedical and Health Informatics*, 2020.
- [7] Z. Zhao, Z. Gong, M. Niu, J. Ma, H. Wang, Z. Zhang, and Y. Li, "Automatic respiratory sound classification via multi-branch temporal convolutional network," *IEEE ICASSP*, 2022.
- [8] P. Bokov, B. Mahut, P. Flaud, and C. Delclaux, "Wheezing recognition algorithm using recordings of respiratory sounds at the mouth in pediatric population," *Computer in biology and medicine*, 2016.
- [9] I. Sen, M. Saraclar, and Y. P. Kahya, "A comparison of svm and gmm-based classifier configurations for diagnostic classification of pulmonary sounds," *IEEE Transactions on Biomedical Engineering*, 2015.
- [10] J. Acharya, A. Basu, and W. Ser, "Feature extraction techniques for low-power ambulatory wheeze detection wearables," in *2017 39th Annual International Conference of the IEEE EMBC*, 2017.
- [11] N. Jakovljević and T. Lončar-Turukalo, "Hidden markov model based respiratory sound classification," in *Precision Medicine Powered by pHealth and Connected Health: ICBHI 2017*, Thessaloniki, Greece, Springer, 2018.
- [12] R. X. A. Pramono, S. Bowyer, and E. Rodriguez-Villegas, "Automatic adventitious respiratory sound analysis: A systematic review," 2017.
- [13] B. Rocha et al., "A respiratory sound database for the development of automated classification," in *Proc. Int. Conf. Biomed. Health Informat.*, pp. 33-37, 2017.
- [14] M. T. García-Ordás, J. A. Benítez-Andrades, I. García-Rodríguez, C. Benavides, and H. Alaiz-Moretón, "Detecting respiratory pathologies using convolutional neural networks and variational autoencoders for unbalancing data," *Sensors*, 2020.
- [14] C. Lea, M. D. Flynn, R. Vidal, A. Reiter, and G. D. Hager, "Temporal convolutional networks for action segmentation and detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, pp. 1003-1012, 2017.
- [15] E. MatthewDavies and S. Böck, "Temporal convolutional networks for musical audio beat tracking," in *Proc. 27th Eur. Signal Process. Conf.*, pp. 1-5, 2019.
- [16] L. Breiman, "Bagging predictors," *Mach. Learn.*, vol. 24, no.2, pp. 123-140, 1996.
- [17] Y. Xiao, J. Wu, Z. Lin, and X. Zhao, "A deep learning-based multi-model ensemble method for cancer prediction," *Comput. Methods Programs Biomed.*, vol. 153, pp. 1-9, 2018.