심전도 자동 해석 모델에 연령과 성별 특징을 추가한 심혈관질환 다중분류

정환규 [1], 박진주 [1], 이명은 [1], 김유리 [2], 양형정 [1*]
[1] 전남대학교 인공지능융합학과
[2] 전남대학교 병원 순환기내과

98ghksrb98@gmail.com, pjj7367@naver.com, myungeun07@gmail.com, geniusyul@gmail.com, hyungjeong@gmail.com

Multi-Label Classification of Cardiovascular Diseases Using an Automated ECG Interpretation Model with Age and Gender Features

Hwangyu Jung[1], Jinju Park[1], Myungeun Lee[1], Yoo Ri Kim[2], Hyung-Jeong Yang[1*]
[1]Dept. of Artificial Intelligence Convergence, Chonnam National University
[2]Division of Cardiology, Department of Internal Medicine, College of Medicine, Chonnam National University

요 약

임상 환경에서 심장 전문의는 표준 12 리드 심전도 및 연령, 성별과 같은 임상적 특징들을 기반으로 심혈관질환을 진단한다. 그러나 기존의 딥러닝 기반 심전도 자동 해석 모델들은 임상적 특징을 포함하지 않고 심전도만으로 심혈관질환을 분류한다. 따라서, 본 연구에서는 심전도, 연령 및 성별 정보를 활용한 심혈관질환 다중분류 모델을 제안한다. 제안된 모델은 연령과 성별 정보를 전처리된 심전도 데이터와 결합하여, 임상적 특징에 따른 심전도 파형의 구조적 특징을 효과적으로 분석하였고 심혈관질환 자동분류 정확성을 향상시켰다. 실험 결과 연령과 성별 정보를 추가했을 때 심전도만을 사용한 기존 모델대비 F1-score 평가지표에서 더 우수한 성능을 보였으며, 이를 통해 연령과 성별 정보가 심혈관질환 분류에 있어 중요한 역할을 하고 있음을 확인하였다.

1. 서론

심전도(Electrocardiogram, ECG)는 심장의 전기 신호를 기록하여 심장의 리듬과 전기적 활동을 분석하는 검사다. 이러한 심전도는 심혈관질환과 밀접한 관계가 있으며, 여러 심장 질환을 진단하는 중요한 도구로 사용된다. 또한, 심혈관질환은 생존율이 낮아 조기 발견이 중요하지만, 심전도에 대한 전문가의 해석이 필요하여 많은 시간과 비용이 소요된다[1]. 이에 따라 심혈관질환 자동화 분류 연구가 많이 수행되고 있다[2].

임상 환경에서 심장 전문의는 표준 12 리드 심전도 및 연령, 성별과 같은 임상적 특징들을 기반으로 심혈관질환을 진단한다. 최근 Ahmadi Pooria, et al. [3] 연구에서 심전도 값과 이상 소견 해석 시 나이와 성별이 중요한 고려 요소임을 보여준 연구를 발표했다. 즉, 휴식기 심박수는 여성에서 더 높지만, P 파, RR 간격, QRS 지속시간 등 다른 심전도 값은 남성에서 더 높았고, 두 성별 모두에서 나이가 증가할수록 주요 심전도 이상의 위험이 증가한다고 했다. 그러나 기존의 딥러닝 모델을 활용한 심혈관질환 자동 분류 연구들은 심전도 데이터만을 활용하여 심혈관질환을 분류하며 이는 연령과 성별 등의 임상적 특징에 대한 정보를 반영하지 않고 있다.

따라서, 본 연구에서는 심전도와 연령 및 성별 정보를 활용한 심혈관질환 다중분류 모델을 제안하고자 한다. 2 장에서는 기존의 심혈관질환 분류 연구를 소개하고 3 장에서는 연령과 성별 정보를 심전도 데이터에 결합한 제안 모델을 소개한다. 4 장에서는 실험 및 결과를 5 장에서는 결론 및 향후 연구를 기술하고자 한다.

2. 관련연구

심전도를 활용한 심혈관질환 분류 방법으로 Asgari et al.[4]는 심방세동(Atrial Fibrillation; AF)을 분류하기 위해 정상 웨이블릿 변환(Stationary Wavelet Transform)을 사용하여 특징을 추출하고 서포트 벡터

머신(Support Vector Machine; SVM)을 사용하였다. Chumachenko et. al. [5]는 심박 변동성(Heart Rate Variability; HRV)의 시간 및 주파수 영역 특징을 사용하고, K최근접 이웃(K-nearest neighbor; KNN)과 랜덤 포레스트 등을 사용하여 심근경색을 분류하였다. 시계열 클러스터링을 사용하는 비지도학습은 Annam et al. [6]에서 제안되었다. 기계학습 모델은 계산 효율성이 높고 예측 속도가 빠른 장점을 가지지만, 시퀀스 데이터를 1차원으로 변환하는 과정에서 시간적으로 연속적인 정보를 충분히 고려하지 못할 수 있다. 이로 인해 중요한 시간적 특징이 손실되어 예측의 정확도가 감소할 수 있다.

최근 많은 연구[7,8]에서 심전도를 활용한 부정맥 진단을 위해 합성곱 신경망 (Convolutional Neural Network; CNN)이 적용되었다. 심전도 시계열 데이터에 1D-CNN을 적용하거나, 심전도 신호를 이미지로 변환하여 2D-CNN을 적용하였다. Sandra et. al. [7]는 합성곱 신경망과, 엔트로피 기반으로 SincNet 등을 활용하였다. 이 아키텍처는 CNN의 특징 학습 기능을 활용하여 멀티 리드 심전도의 특징을 학습한다. 또한, 시계열을 고려해 특징을 추출하는 TCN을 활용한 연구로는 Prabhakararao et al. [8] 연구가 있다. 연구에서는 각 리드별로 TCN 모듈을 통해 단일 레이블을 각각의 모델에 학습하는 앙상블학습을 이용해 이진분류를 수행하였다.

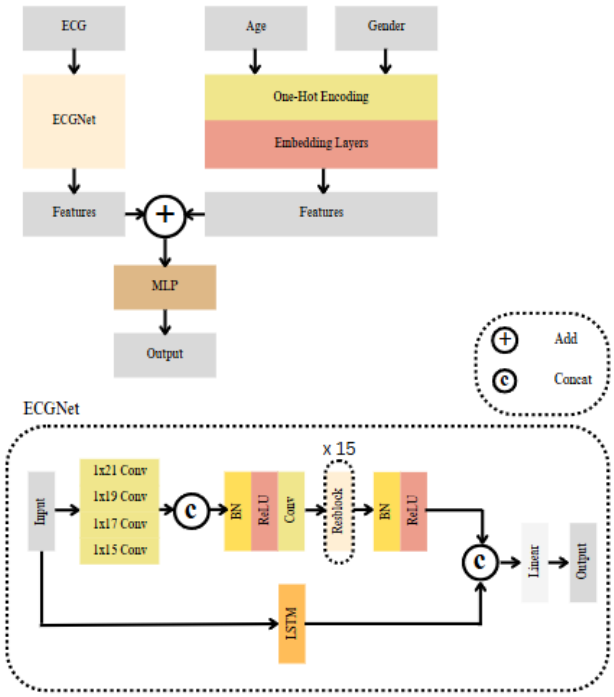
이와 같이, 기존의 많은 연구들이 심전도 데이터만을 이용해 심혈관질환 분류를 수행하였다. 현재 심전도 데이터와 임상적 데이터를 특징으로 사용한 논문은 많지 않다.

3. 제안 방법

본 연구에서는 전처리된 심전도 데이터를 모델의 입력으로 사용하고 연령, 성별에 따른 심전도 파형의 구조적 특징을 활용하기 위해 연령, 성별 정보에 선형 변환 기법[9]을 적용하여 이후에 결합하는 Late Fusion 모델을 제안한다.

Late Fusion 모델의 입력은 전처리된 심전도 데이터이며, 길이가 1000인 12 리드 심전도(32, 1000, 12) 구조를 가진다. 입력 데이터는 Backbone 모델인 ECGNet[10]을 통해 다양한 크기의 커널로 병렬 처리되어 특징이 추출되고 모두 연결되어 하나의 특징 벡터로 변환된다. 이후, 15개의 Residual block을 거치며 특징이 추출되고 LSTM을 통해 얻은 특징과 더해져 최종 심전도 특징이 추출된다.

연령, 성별 정보의 경우 연령 정보를 5차원으로, 성별 정보를 2차원으로 원핫인코딩한 후, 모델을 통해 나온 심전도 특징과 길이가 같도록 선형 변환이 진행된다. 선형 변환된 연령, 성별 특징과 심전도 특징을 결합하고



(그림 1) 제안 모델 아키텍처

최종 출력층으로 전달되어 연령, 성별을 고려한 심혈관질환 다중분류가 수행된다. Late Fusion의 아키텍처는 그림 1과 같다.

4. 실험 및 결과

4.1 데이터셋 및 전처리

LargeECG 데이터셋[11]은 10초 길이, 500Hz의 45,152 개의 심전도로 구성되어 있다. PTB-XL 데이터셋[12]은 10초 길이, 500Hz의 21,799 개의 심전도로 구성되어 있으며 상세 정보는 표 1과 같다. 본 연구에서는 빈도수가 10개 미만인 데이터를 제거 후 LargeECG 46개, PTB-XL 44개의 레이블로 실험을 진행하였다.

<표 1> 데이터셋 분포

Dataset	Sample	Label	Age range	Gender (male:female)
LargeECG	45,152	51	0 ~ 89	56:44
PTB-XL	21,799	44	0 ~ 95	52:48

데이터 전처리의 경우 심전도 데이터는 심장의 주요 주파수 성분을 보존하고, 전자 기록 장치의 노이즈 제거를 위해 주파수 0.5hz~45hz의 범위로 필터링하였다. 이후 모든 데이터가 유사한 스케일을 갖도록 Z-정규화를 수행하였다.

<표 2> 제안 모델 성능 평가

Dataset	Features	ACC	Precision	Recall	F1-score
LargeECG	ECG-only	0.5975	0.8630	0.7533	0.7815
	w/ age	0.6090	0.8700	0.7618	0.7888
	w/ gender	0.6071	0.8652	0.7566	0.7842
	w/ age, gender	**0.6158**	**0.8751**	**0.7644**	**0.7923**
PTB-XL	ECG-only	0.4503	0.6154	0.5594	0.5723
	w/ age	0.4730	0.6261	0.5684	0.5829
	w/ gender	0.4749	0.6389	0.5776	0.5924
	w/ age, gender	**0.4858**	**0.6433**	**0.5784**	**0.5962**

연령데이터는 Null 값은 Missing values로, 0세~4세는 유아(Early childhood), 5세~14세는 어린이(Childhood), 15세~22세는 청소년 (Adolescence), 23세~41세는 청년(Youth), 42세~56세는 성인(Maturity), 57~68세는 중년(Old age), 69세 이상은 노년(Late old)으로 그룹화를 진행한 후, 원핫인코딩을 적용하였다. 이 중 표본이 적은 14세 이하의 데이터를 제거하여 5개의 그룹으로 실험을 진행하였다.

4.2 실험 결과

심혈관질환 다중분류 성능 평가 지표로는 정확도(Accuracy), 정밀도(Precision), 재현율(Recall), F1-score를 사용하였고 연구의 일반화 성능을 확보하기 위해 5겹 교차 검증(5-fold cross-validation)을 적용하였다.

연령과 성별 정보가 성능 평가에 중요한 역할을 하는지 판단하기 위해 심전도 데이터만을 사용했을 때와 심전도 데이터에 연령, 성별 정보를 결합하였을 때를 구성하여 성능 평가를 진행하였다. 모든 결과는 샘플별로 계산한 값의 평균을 나타내며 정확도의 경우 샘플 당 모델이 예측한 값이 정답과 정확히 일치할 때만 1의 값을 나타낸다.

표 2는 제안 모델을 두 가지 데이터셋을 통해 평가한 결과를 나타내며 표 3은 모델별 성능 평가 결과를 나타낸다. 심전도 데이터만을 사용했을 때보다 연령과 성별 정보를 추가했을 때 전반적으로 성능이 향상됨을 보였다. 이는 모델이 임상적 맥락에서 중요한 비시계열적 특징과 심전도의 시계열 특징을 동시에 고려할 수 있게 되어 보다 정밀하게 심혈관질환을 분류한 것으로 보인다.

실험 결과를 통해 모델이 심전도 데이터의 복잡한 시계열 정보와 임상적 특징을 효과적으로 분석할 수 있음을 확인하였다. 또한, 심혈관질환 분류에서 연령과 성별이 모델의 성능을 향상시키는 데 중요한 요소임을 확인할 수 있었다.

<표 3> 모델별 성능 평가

Dataset	Model	Features	ACC	F1-score
Large ECG	ResNet101 [13]	ECG-only	0.5644	0.7482
		w/ age, gender	0.5979	0.7773
	BiLSTM [14]	ECG-only	0.5701	0.7485
		w/ age, gender	0.5789	0.7595
	Ours	ECG-only	0.5975	0.7815
		w/ age, gender	**0.6158**	**0.7923**
PTB-XL	ResNet101	ECG-only	0.4317	0.5594
		w/ age, gender	0.4627	0.5719
	BiLSTM	ECG-only	0.4451	0.5605
		w/ age, gender	0.4641	0.5722
	Ours	ECG-only	0.4503	0.5723
		w/ age, gender	**0.4858**	**0.5962**

5. 결론

본 연구에서는 심전도와 연령, 성별 정보를 활용한 심혈관질환 다중분류 모델을 제안하였다. 12 리드 심전도

데이터만을 사용했을 때와 연령, 성별의 임상적 정보를 결합했을 때의 분류 성능을 비교했으며, 연령과 성별 정보를 포함한 모델이 심전도 데이터만을 사용한 모델보다 정확도 및 F1-score가 향상됨을 확인하였다.

이를 통해 연령과 성별 정보가 심혈관질환의 세부 분류에 있어 비교적 중요한 역할을 하고 있음을 확인하였다.

향후 연구에서는, 연령과 성별 외에도 과거력, BMI, 흡연 여부 등 다양한 임상적 정보들을 반영한 연구를 진행할 예정이다.

사사

이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(RS-2023-00208397).
본 연구는 과학기술정보통신부 및 정보통신기획평가원의 인공지능융합혁신인재양성사업 연구 결과로 수행되었음(IITP-2023-RS-2023-00256629).
이 논문은 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원-지역지능화혁신인재양성사업의 지원을 받아 수행된 연구임(IITP-2025-RS-2022-00156287).
본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학 ICT 연구센터사업의 연구결과로 수행되었음(IITP-2025-RS-2024-00437718).

참고문헌

[1] Roth Gregory A., et al. "Global burden of cardiovascular diseases and risk factors, 1990–2019: update from the GBD 2019 study." Journal of the American college of cardiology 76.25: 2982-3021(2020).

[2] Ran Shaolin, et al. "Label correlation embedding guided network for multi-label ECG arrhythmia diagnosis." Knowledge-Based Systems 270: 110545(2023).

[3] Ahmadi Pooria, et al. "Age and gender differences of basic electrocardiographic values and abnormalities in the general adult population; Tehran Cohort Study." BMC Cardiovascular Disorders 23.1: 303(2023).

[4] Shadnaz Asgari, Alireza Mehrnia, and Maryam Moussavi. “Automatic detection of atrial fibrillation using stationary wavelet transform and support vector machine”. In: Computers in biology and medicine 60, pp. 132–142(2015).

[5] Chumachenko, Dmytro, et al. "Machine learning methods in predicting patients with suspected myocardial infarction based on short-time HRV data." Sensors 22.18: 7033(2022).

[6] Jagadeeswara Rao Annam and Bapi Raju Surampudi. “AAMI Based ECG Heart-Beat Time-Series Clusterling Using Unsupervised ELM and Decision Rule”. In: ICIT. IEEE, pp. 137–141, 2016.

[7] Śmigiel Sandra, Krzysztof Pałczyński, and Damian Ledziński. "ECG signal classification using deep learning techniques based on the PTB-XL dataset." Entropy 23.9: 1121(2021).

[8] Prabhakararao Eedara, and Samarendra Dandapt. "Multi-Label ECG Classification using Temporal Convolutional Neural Network." arXiv preprint arXiv:2306.03844 (2023).

[9] Xiao Ran, et al. "Integrating multimodal information in machine learning for classifying acute myocardial infarction." Physiological Measurement 44.4: 044002 (2023).

[10] Murugesan Balamurali, et al. "Ecgnet: Deep network for arrhythmia classification." 2018 IEEE International Symposium on Medical Measurements and Applications (MeMeA). IEEE, 2018.

[11] Zheng, J., H. Guo, and H. Chu. "A large scale 12-lead electrocardiogram database for arrhythmia study (version 1.0. 0)." PhysioNet 2022Available online httpphysionet orgcontentecg arrhythmia10 0accessed on 23 (2022).

[12] Wagner, P., N. Strodthoff, et al. "PTB-XL, a large publicly available electrocardiography dataset." PhysioNet httpphysionet.orgcontentptb-xl1.0.3 (2022).

[13] He, Kaiming, et al. "Deep residual learning for image recognition." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.

[14] Zhang, Shu, et al. “Bidirectional long short-term memory networks for relation classification.” Proceedings of the 29th Pacific Asia conference on language, information and computation. 2015.