

Machine-learning-based out-of-hospital cardiac arrest (OHCA) detection in emergency calls using speech recognition*

Jong In Kim^{1,**} · Joo Young Lee² · Jio Chung³ · Dae Jin Shin⁴ · Dong Hyun Choi^{5,6} ·

Ki Hong Kim^{6,7,8} · Ki Jeong Hong^{6,7,8,9} · Sunhee Kim¹⁰ · Minhwa Chung^{1,2,***}

¹*Interdisciplinary Program in Cognitive Science, Seoul National University, Seoul, Korea*

²*Department of Linguistics, Seoul National University, Seoul, Korea*

³*V/S Works, Seoul, Korea*

⁴*SoundMind, Seoul, Korea*

⁵*Department of Biomedical Engineering, Seoul National University College of Medicine, Seoul, Korea*

⁶*Laboratory of Emergency Medical Services, Seoul National University Hospital Biomedical Research Institute, Seoul, Korea*

⁷*Department of Emergency Medicine, Seoul National University Hospital, Seoul, Korea*

⁸*Department of Emergency Medicine, Seoul National University College of Medicine, Seoul, Korea*

⁹*Disaster Medicine Research Center, Seoul National University Medical Research Center, Seoul, Korea*

¹⁰*Department of French Language Education, Seoul National University, Seoul, Korea*

Abstract

Cardiac arrest is a critical medical emergency where immediate response is essential for patient survival. This is especially true for Out-of-Hospital Cardiac Arrest (OHCA), for which the actions of emergency medical services in the early stages significantly impact outcomes. However, in Korea, a challenge arises due to a shortage of dispatcher who handle a large volume of emergency calls. In such situations, the implementation of a machine learning-based OHCA detection program can assist responders and improve patient survival rates. In this study, we address this challenge by developing a machine learning-based OHCA detection program. This program analyzes transcripts of conversations between responders and callers to identify instances of cardiac arrest. The proposed model includes an automatic transcription module for these conversations, a text-based cardiac arrest detection model, and the necessary server and client components for program deployment. Importantly, The experimental results demonstrate the model's effectiveness, achieving a performance score of 79.49% based on the F1 metric and reducing the time needed for cardiac arrest detection by 15 seconds compared to dispatcher. Despite working with a limited dataset, this research highlights the potential of a cardiac arrest detection program as a valuable tool for responders, ultimately enhancing cardiac arrest survival rates.

Keywords: cardiac arrest, out-of-hospital cardiac arrest (OHCA), machine learning-based OHCA detection

* This study was conducted with the support of the National Fire Agency's research and development project, "Information Standardization and Data Utilization of 911 Emergency Calls" The 119 emergency call data was collected through cooperation with the Sejong Fire Headquarters, Jeju Fire Headquarters, and Seoul Fire Headquarters for the purpose of building audio data.

** prows12@gmail.com, Corresponding author

*** mchung@snu.ac.kr, Corresponding author

Received 24 November 2023; Revised 11 December 2023; Accepted 11 December 2023

© Copyright 2023 Korean Society of Speech Sciences. This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. 서론

심정지는 심장의 혈액공급이 원활히 이루어지지 않아, 결과적으로 세포 내 혈액공급이 제대로 이루어지지 않은 상태를 말한다. 질병관리청에 따르면 급성심장정지 생존율은 7.5%이며, 2020년도 급성심장정지 발생률은 인구 10만 명당 61.66명으로, 총 32,000명에 달하였다(Kim et al., 2022). 심정지는 일반적으로 병원 내 심정지(in-hospital cardiac arrest), 병원 전 심정지(out-of-hospital cardiac arrest)로 분류된다(Park, 2014). 병원내심정지(IHCA)와 달리 병원밖심정지(out-of-hospital cardiac arrest, OHCA)에서는 119 구급대의 반응시간이 중요한 것으로 알려져 있다.

심정지의 국내 발생현황 추세를 살펴보면, 2018년 기준 39.7명으로, 2006년부터 꾸준한 증가 추세에 있다(Yoon, 2020). 심정지는 한국인의 주요 사망원인 중 하나이며, 중요한 보건 문제로 인식되어 왔다. 이는 심정지 발생이 개인에게 경제적, 정신적 부담을 가중시키기 때문이다. 심정지의 생존에는 심정지 발생 시 초기 대응 및 대처가 중요한 요소로 알려져 있다(O'Keeffe et al., 2011; Vukmir & Sodium Bicarbonate Study Group, 2004). 심정지는 적절한 초기 대처를 시행하지 않은 경우에는 3-5분 내에 뇌가 손상되며 사망에 이를 수 있다(Safar, 1986).

심정지의 생존을 높이는 요소로는 병원도착 전에 심정지 발생을 목격했는지 여부, 일반인의 심폐소생술 혹은 제세동 시행 여부, 병원에 도착하기 전 심폐소생술 여부, 신고에서 현장 도착까지 걸린 시간 등이 있다(Kim & Chun, 2020). 특히 심정지 발생 시에 소방전문 인력의 정확하고 신속한 응급처치 여부가 중요한 요소로 간주되어 오고 있다(Berg et al., 2020; Eisenberg et al., 1990). 이는 병원 전 단계(구급 단계)에서 심정지 발생 시 초기 대응 및 치료가 응급 의료 체계 내에서 적절하게 이루어졌는지에 대한 여부가 심정지의 생존률을 높이는 데 중요한 역할을 하고 있음을 의미한다.

응급 의료체계는 병원 전 단계(pre-hospital phase)와 병원 단계(in-hospital phase)로 구성된다(Park, 2014). 병원 전 단계에서는 환자 발생 시 신고 접수와 수보요원(dispatcher)의 응급처치 지도, 구급차 출동, 응급 구조사 및 구급대원의 환자에 대한 응급조치로 이루어진다(Chung, 2015). 병원 단계(in-hospital phase)에서는 환자에 대한 진단 및 수술이 이루어진다. 수보요원의 역할은 신고자가 최초 신고한 전화를 접수하며, 적합한 응급처치를 지도하고, 구급대를 출동시키며, 구급대원에게 현장상황을 공유하여 적절한 조치를 시행할 수 있도록 지도한다. 이때 수보요원의 반응 시간을 줄이는 것은 병원밖심정지(OHCA)의 생존율에 큰 영향을 미친다.

119 종합상황실(Oh, 2022)에 따르면 구급상황실에 접수된 신고 발생 건수는 2021년 기준 1,207만 5,804건이다. 수보요원은 전국에 838명이며, 1명당 14,410건을 처리했다. 이는 소수의 수보요원이 다량의 신고접수에 대응해야 하는 상황을 의미한다. 즉 수보요원의 부담을 경감하기 위해서, 특정 질환에 대해 자동으로 판단해주는 시스템이 있을 경우, 보조적인 수단으로서 유용하게 쓰일 수 있을 것이다.

또한 표 1에 따르면(New Jersey Government, 2020; Phoenix Government, 2016) 심정지 수보 지침은 일반적으로 신고자의 위치와 전화번호, 환자의 의식 여부, 환자의 호흡 여부, 환자의 나이를 확인한다. 수보요원은 미국, 일본, 한국이 공통적으로 환자의 위치와 전화번호를 확인한다. 환자의 의식과 호흡은 심정지를 판단하는 기준이 되는 지표가 된다. 나이는 미국과 한국의 경우 의식과 호흡을 질문하기 후 확인하나, 일본의 경우 의식과 호흡 전에 질문한다. 이러한 심정지 수보지침의 특성을 머신러닝 심정지 탐지 시스템에 반영한다면 심정지를 판단하는데 유용하게 쓰일 수 있을 것이다.

응급의료체계 내에서 머신러닝을 활용한 연구들이 진행되어 왔다. Lee et al.(2016)에서는 응급의료시스템 내에서 음성인식기를 활용하여 응급 환자 정보를 추출하는 시스템을 개발하였다. 자동음성인식기는 40만 단어에 해당하는 한국어 연속 음성인식기를 활용하였고, 응급환자 정보 추출은 템플릿을 기반으로 레빈슈타인 거리를 활용하였다.

Kim et al.(2022)은 119 응급 서비스를 통해 이송되었던 환자들의 병원 내의 심정지를 예측하는 머신러닝 모델을 개발하였다. 이 연구에서는 심정지를 예측하기 위해서 병원 내(in-hospital) 요인을 활용하였다. 머신러닝 모델은 gradient boosting model을 활용하였다. 실험 결과에 따르면 산소 공급량, 수축기 혈압, 맥박수, 나이, 산소 포화도 등이 심정지 생존의 중요한 요소였다. Park(2022)은 병원밖심정지(OHCA) 환자를 대상으로 구급현장에서 자발적으로 순환을 회복할 수 있는지 여부를 예측하는 모델을 개발하였다. 훈련 시 학습데이터는 5,240명, 평가데이터는 1,709명을 활용하였으며, 개발 모델 가운데 사건-시간 시점 분석 모델이 가장 좋은 성능을 보였다.

응급 의료체계 내에서 신고자와 수보자의 통화 내역을 분석하여 심정지를 예측하는 연구로는 Blomberg et al.(2019)이 있는데, 이는 덴마크 코펜하겐의 병원밖심정지(OHCA)와 관련된 통화를 수집하여 심정지 여부를 예측하는 모델을 개발한 것이다. 학습 데이터로는 918건이 병원밖심정지(OHCA)를 활용하였다. 이 실험에서 머신러닝 모델이 수보요원보다 심정지 예측정확도는 84.1%, 심정지 인지시간은 10초 단축되었다. 이를 통해 기계학습 모델이 수보요원의 심정지 판단을 돕는데 주요한 역할이 될 수 있음을 보여주었다. 또한 Byrsell et al.(2021)은 심정지 탐지를 1분 안에 수보요원보다 빠르게 탐지할 수 있는 기계학습 모델을 제시하였다. 그리고 병원밖심정지(OHCA)탐지의 영향을 주는 통화의 특성을 조사하였는데, 학습데이터는 851건을 활용하였다. 제시한 기계학습 모델의 심정지 정확도는 84%이며, 심정지 인지시간 단축은 수보요원보다 22초 빨랐다. 이를 통해 심정지 탐지에 있어서 머신러닝을 활용한 모델이 좋은 보조도구가 될 수 있음을 보였다.

119 구급 상황관리센터의 심정지 인지시간이 단축되면 심정지의 생존율이 향상될 수 있다. Gnesin et al.(2021) 연구에서는 신속한 구급출동이 병원밖심정지(OHCA)의 생존율을 향상시킬 수 있음을 보였다. 수보요원의 심정지 인지 단축 시간은 신속한 구급출동을 가능하게 하며, 이를 통해 심정지 환자의 생존

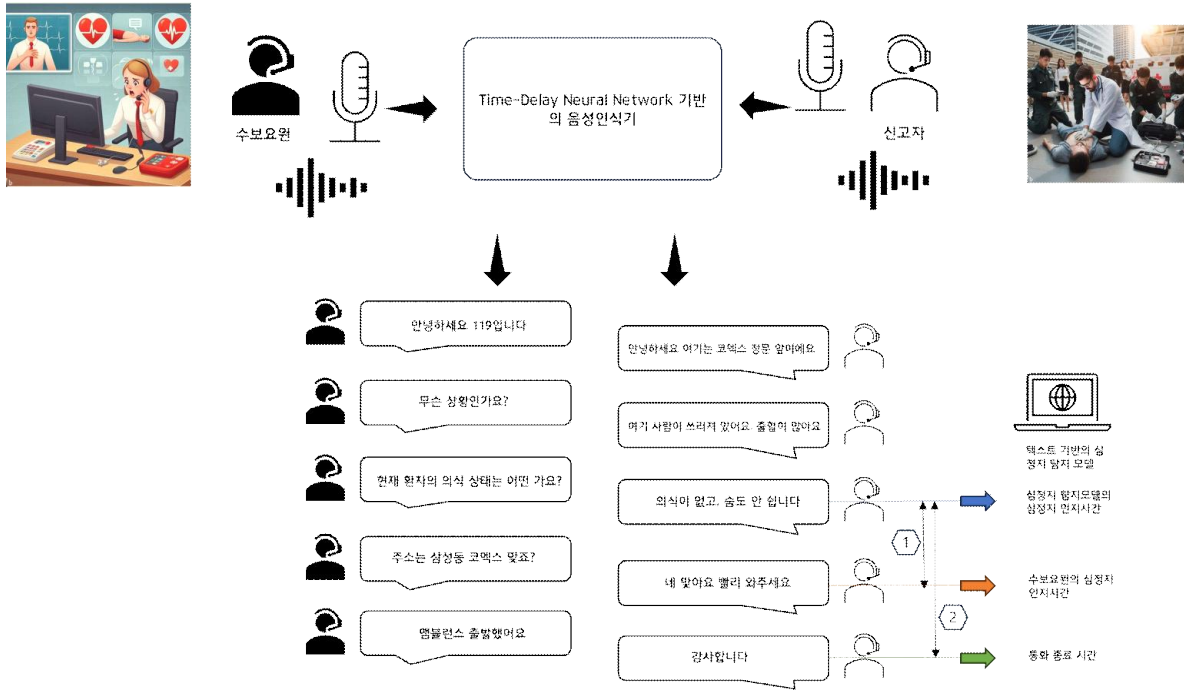


그림 1. 음성인식 기반의 심정지 탐지 프로그램의 개념도

Figure 1. Conceptual framework of speech recognition-based out-of-cardiac arrest (OHCA) detection program

율을 높일 수 있다.

즉, 심정지 자동 탐지 시스템의 개발은 과중되어 있는 소방본부 인력을 보조하는 보조수단으로서의 역할과, 초기 응급 대처 및 대응이 중요한 심정지 질환에 신속하게 대응하는데 그 필요성이 있다. 그러나 심정지 탐지 프로그램은 개인정보 보호 및 심정지 녹취 데이터 수집의 어려움 그리고 기술적인 문제 때문에 활발히 연구되지는 못하였다. 본 연구는 이러한 문제를 해결하는 방안으로 수보요원과 신고자의 통화 녹취록을 분석하여 심정지를 자동 탐지하는 프로그램을 제안하는 것을 목적으로 한다.

표 1. 심정지 수보 지침

Table 1. OHCA-related emergency call guideline

항목	미국(피닉스)	한국	일본
위치, 전화번호	초기 질문	초기 질문	초기 질문
의식 여부	의식 여부 확인	의식 여부 확인	대화 가능 여부로 의식 확인
호흡 여부	의식과 호흡이 모두 비정상이면 심정지	의식 호흡이 모두 비정상이면 심정지	호흡을 우선적으로 질문
나이	의식과 호흡 질문 이후	의식과 호흡 질문 이후	의식과 호흡 질문 전

2. 머신 러닝 기반 심정지 탐지 프로그램

머신 러닝 기반 심정지 탐지 프로그램은 수보요원과 신고자의 전화통화를 분석하여 심정지 여부를 판단하는 프로그램으로(그림 1) 3가지 모듈로 이루어져 있다. 첫째는 수보요원과 신고자의 대화내용을 텍스트로 전사하는 자동전사 모델과, 전사

된 텍스트를 기반으로 심정지 여부를 탐지하는 모델, 서버를 기반으로 하여 각 모델들을 유기적으로 연결하기 위한 UI 및 서버 개발로 구성되어 있다.

2.1. Kaldi 기반의 Chain 모델 기반 신고자와 수보요원의 통화 자동전사 모델

신고자와 수보요원의 통화 녹취록을 자동 전사하기 위해서, Kaldi 기반의 Chain 모델을 활용하였다. Kaldi 기반의 Chain 모델은 DNN-HMM 기반 모델의 일종이다(Povey et al., 2011).

Kaldi 기반의 Chain 모델은 다음의 훈련 과정을 거친다. 1단계에서는 각 음성에 대해서 음성 특징을 추출하여 벡터의 형태로 저장한다. 2단계에서는 monophone training을 통해 음향 모델을 학습한 뒤 alignment를 진행한다. 3단계에서는 트라이폰 학습을 통해 음향 모델을 학습하여, 적절한 파라미터를 찾는다. 4단계에서는 ivector를 추출하여 화자 정보를 저장한 후, 5단계에서 chain 기반의 음향 모델을 학습한다. 그 후에 6단계에서 decoding 과정을 수행한다.

Chain 모델은 뉴럴 네트워크의 아웃풋의 프레임 레이트를 3배로 줄여 디코딩 시 연산량을 줄였다. 이는 심정지 탐지 프로그램에서 real-time decoding에 적합한 형태라고 할 수 있다. Chain 모델은 일반적인 DNN-HMM 모델보다 디코딩 시 3배 정도 빠른 것으로 알려져 있다.

2.2. Fasttext 를 이용한 텍스트 기반의 심정지 탐지 모델

Fasttext는 Facebook에서 개발한 텍스트 분류에 적합한 툴킷으로, Linear Bag-of-words 알고리즘을 통해 단어의 의미에 대해

워드 임베딩 형태로 학습을 진행한다(Bojanowski et al., 2017). Fasttext 알고리즘은 일반적으로 연산량이 적어 속도가 빠르며, 가볍고, 속도 대비 정확도가 높은 편이다.

Linear Bag-of-words는 속도가 빠르고 정확도가 높기 때문에 자동전사된 텍스트의 심정지 여부를 빠르게 판단할 수 있다. 심정지 여부를 빠르게 판단할 수 있다는 것은 프로그램을 구성하는데 핵심적인 요소이다.

텍스트 기반의 심정지 탐지 알고리즘은 증상 표현 여부를 판단하는 모델, 병원밖심정지(OHCA)를 판단하는 모델, “숨”, “호흡”이 포함된 문장에서 호흡과 의식 여부를 판단하는 모델의 3개의 텍스트 세부 분류모델로 구성되어있다. 그림 2와 같이 자동 전사모델에서 전사된 N번째 텍스트는 텍스트 기반의 심정지 모델의 3가지 세부 모델을 통해 심정지 여부를 판별한다.

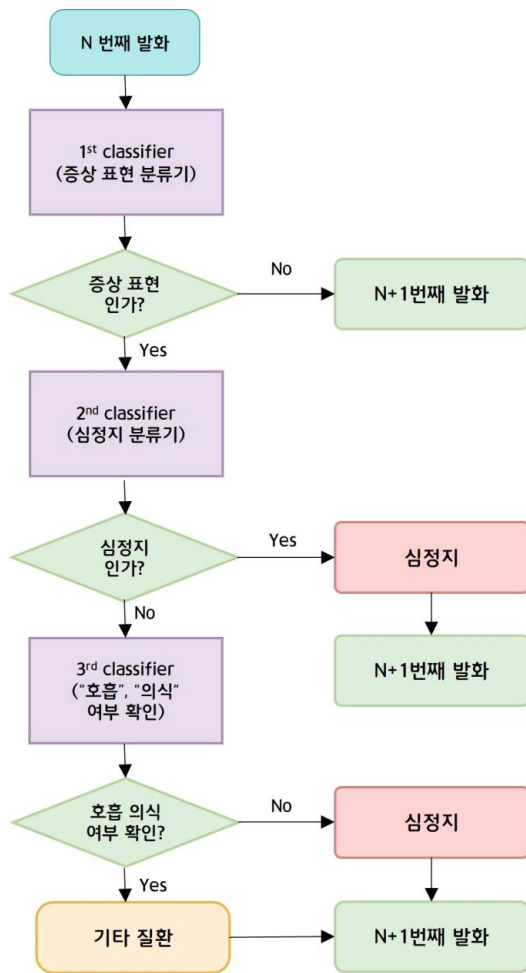


그림 2. 텍스트 기반 심정지 탐지 모델의 알고리즘
Figure 2. The algorithm for text-based OHCA detections

증상 표현 모델은 통화 내용 중 신고자가 발화한 텍스트에서 증상표현에 대해서 발화했는지 여부를 판단하기 위한 모델이다. 예시로 신고자가 수보요원에게 “환자가 가슴 통증이 있습니다” 라고 발화한다면 “증상 표현”이라고 판단한다. 다른 예로 신고자가 수보요원에게 “여보세요 119요?”라고 질문한다면

“증상 표현이 아님”이라고 판단한다.

심정지 분류 모델은 자동전사된 텍스트를 바탕으로 병원 밖 심정지(OHCA)인지 아닌지를 이진 분류하는 모델이다. 본 분류 모델에서 핵심적인 모델이며, 신고자와 수보요원의 대화를 분석하여 심정지 여부를 판단하는 모델이다. 심정지 분류 모델의 학습은 수보요원과 신고자의 전체 발화를 대상으로 학습시켰다. 즉 수보요원과 신고자의 전사 텍스트 전체를 분석하여 심정지 여부를 판단한다.

숨, 호흡, 의식 기반의 심정지 분류 모델은 통화 중 “숨”, “호흡”, “의식” 이란 단어가 있을 경우에 작동되며, 호흡과 의식 여부를 통해 심정지 여부를 판단하는 모델이다. 이 모델은 호흡이나 의식이 없을 경우 심정지라고 판단한다. 이는 수보지침에서 심정지 판단에 “숨”과 “호흡” 이 중요한 역할을 하고 있음을 반영하기 위한 모델이며, 두 번째 심정지 분류 모델을 보완하기 위한 모델이다.

제안한 모델의 심정지 정확도는 통화 종료 시점을 기준으로 측정한다. 3차 심정지 시스템은 의식과 호흡 여부를 판단하기 위함이며, 2차 심정지에서 놓친 케이스에 대해서, 다시 한번 검토하는 목적으로 설계하였다.

2.3. 심정지 탐지 프로그램을 위한 서버 및 클라이언트 개발

심정지 탐지 프로그램의 서버는 실시간 음성인식이 가능한 음성인식 서버와 심정지 증상을 판단하는 증상분류 API로 이루어진다.

음성인식 서버는 c++ 기반의 소켓 통신을 이용한 서버이며, Kaldi 기반의 음성인식기를 활용할 수 있는 형태로 개발되었다. 서버는 다채널로 구성되었고, 음향 모델과 언어모델을 통합하여 여러 채널을 활용하도록 구성되어 있다. 심정지 분류 API의 개발을 위해 FastAPI를 활용한 RestfulAPI 서버를 개발하였다.

심정지 탐지 프로그램의 클라이언트는 C#의 윈도우폼 기반으로 개발되었다. 클라이언트 프로그램은 소켓 방식으로 자동 전사 모델과 통신하고, 결과를 바탕으로 텍스트 기반의 심정지 API와 http통신을 통해 심정지 여부를 탐지하는 프로그램이다.

클라이언트 프로그램의 GUI(graphical user interface)는 그림 3과 같이 구성되어 있다. 상단에는 수보요원과 신고자의 대화 내용에 따른 표시하고, 좌측 하단에는 심정지와 기타질환을 구분할 수 있는 증상 분석 결과가 표시된다. 우측에는 분석로그를 통해 모델이 심정지를 판단한 시간 정보를 제시한다. 좌측 최하단에는 가장 마지막에 말한 내용을 표시하며, 우측 최하단에는 마이크를 누를 수 있는 버튼을 표시한다.

2.4. 통합 심정지 탐지 프로그램

본 연구에서는 앞서 제시된 개별 모듈을 통합하여 심정지 시스템을 구현하였다. 구현 방식은 다음과 같다. 신고자와 수보요원이 마이크에 음성을 발화하면 발화된 음성은 칼디 기반의 음성인식 모듈을 통해 텍스트로 자동전사된다. 이 때 클라이언트에서는 칼디 기반의 음성인식 서버에 전사 정보를 요청하고, 전사된 결과 텍스트는 그림 3과 같이 UI(user interface)를 통해 표

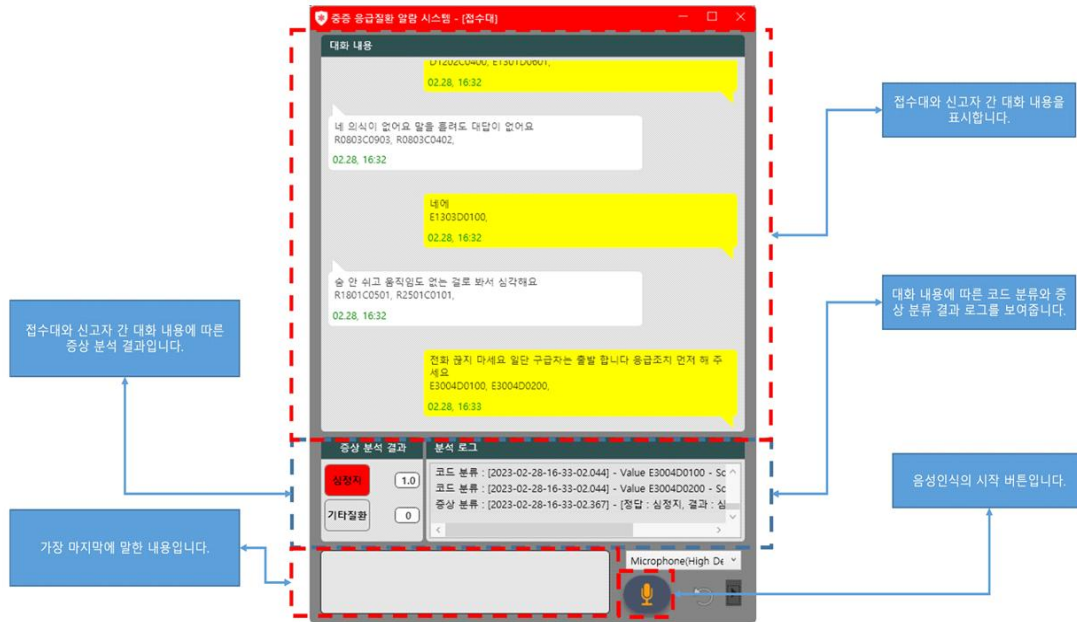


그림 3. 심정지 탐지 프로그램의 UI(User Interface)
 Figure 3. The user interface (UI) design for cardiac arrest detection software

시된다. UI에서는 신고자와 수보자의 텍스트 정보를 구분하여 표시한다. 이후 전사된 텍스트 정보는 텍스트 기반의 심정지 시스템을 통해 심정지 여부를 판단한다.

본 프로그램은 통화내용 분석을 통해 심정지 탐지 정확도 측정과 심정지 최초 인지시간 측정을 수행한다.

심정지 탐지 정확도 측정은 다음과 같다. 통화 전체 시간을 미리 예측할 수 없기 때문에 심정지 수보요원과 신고자의 통화 내용을 일정 단위로 끊어서 분석한다. 음성인식을 통해 연속적으로 통화내역의 일부분을 연속적으로 전사하며, 심정지 여부를 판단한다. 이때 시스템은 심정지와 기타질환 중 하나로 판단을 내린다. 이러한 방식으로 시스템에서 N번째 발화에 대한 심정지 여부의 판단이 끝난 이후, N+1번째 신고자와 수보요원의 음성을 자동전사한다. 또한 텍스트 기반의 심정지 판단 시, N에서 N+1번째로 진행될수록, 자동 전사텍스트를 누적한다. 즉 현재 음성인식에서 인식한 발화뿐 아니라 이전에 음성인식에서 전사한 텍스트도 누적적으로 포함하여 심정지 여부를 판단한다. 이 과정은 신고자와 수보요원의 통화가 끝날 때까지 반복된다. 이는 통화 종료 시점까지 진행된다. 그 후, 심정지로 판단된 경우와 기타질환으로 판단된 경우의 총합을 비교하여 심정지 여부를 판단한다.

제안한 모델의 심정지 인지시간 측정은 모델이 심정지라고 최초로 인지한 시점을 기준으로 측정하였다. 제안한 모델은 통화 내용 분석을 통해 심정지 여부를 판단하게 되는데, 이 때 최초 심정지라고 판단한 시간을 측정한다. 수보요원의 심정지 인지시간 측정은 서울소방본부에서 제공받은 수보요원의 심정지 인지시간 DB를 활용하였다. 심정지 인지시간 단축은 모델이 최초로 심정지라고 인지한 시점과 서울소방본부에서 제공한 수보요원의 실제 인지시간과의 차이를 이용하여 측정하였다.

3. 실험

3.1. 데이터셋

본 연구에서 사용된 음성 데이터셋과 텍스트 데이터셋은 서울재난소방본부와 서울종합방재센터에서 제공받았다. 서울 소방의 음성 데이터셋과 실험은 본부 내에서 실험을 진행하는 것만이 허용되었다. 서울 소방의 텍스트 데이터셋은 외부 반출은 가능했으나, 개인정보에 대해서는 비식별화를 진행하였다.

3.1.1. 신고자와 수보요원의 자동전사 모델 학습을 위한 데이터셋

신고자와 수보요원의 자동전사의 음향 모델 학습을 위한 데이터셋은 9,700시간의 음성 데이터셋을 활용하였다. 표 2와 같이 speed perturbation을 통해, 29,100시간의 음향 데이터셋을 학습시켰다. Speed perturbation은 전체 데이터셋을 기준으로 오디오의 속도를 0.9배와 1.1배로 변환하는 방식으로 진행하였다. 언어 모델 학습을 위해서는 총 289,448,559개의 발화를 이용하여 학습시켰으며, 제주 소방에서 제공받은 2,571콜(81시간)도 포함되어있다.

표 2. 음향 모델 학습을 위한 음성데이터셋(trainset)
 Table 2. Speech dataset for speech-to-text acoustic model training (trainset)

음성데이터	시간
연구실 기보유 음성데이터셋	3,619시간
제주 소방 데이터셋	81시간
AIHUB 저음질 전화망 음성 데이터	6,000시간
총 음향 데이터셋	9,700시간
Speed perturbation 적용 후 음향 데이터셋(×3)	29,100시간

본 실험에서 학습된 자동전사 모델을 평가하기 위한 테스트셋으로는 서울 소방본부에서 제공받은 수보요원과 신고자의 통화 녹취 음성 데이터셋 중 100건을 이용하였다. 소방본부 실험 데이터를 활용하기 위해, 100건의 데이터에 대한 태깅 작업을 시행하였고, 이를 위해 수보요원 파트와 신고자 파트를 분류해서 세그먼트를 진행하였다.

3.1.2. 텍스트 기반의 심정지 탐지 모델 학습을 위한 데이터셋

텍스트 기반의 심정지 탐지 모델을 학습하기 위해서 서울소방본부와 제주 소방본부에서 제공받은 텍스트 기반의 통화녹취 데이터셋을 이용하였다. 텍스트 기반의 심정지 탐지 모델에서는 제주 소방 데이터셋, 서울 소방데이터셋과 Aihub 보건민원콜 데이터셋을 포함하여 약 29,340개의 데이터셋을 이용하여 모델링하였다. 이외에도 성능을 향상시키기 위해 일부 데이터를 보강하였다(표 3).

표 3. 텍스트 기반의 심정지 탐지 모델 학습을 위한 데이터셋
Table 3. Dataset for text-based cardiac arrest detection model (Trainset)

	제주 소방(콜)	서울 소방(콜)
전체 콜	2,571	6,122
심정지 콜	1,773	2,145
기타질환 콜	798	3,977

3.1.3. 통합 심정지 탐지 시스템 관련 평가 테스트셋

통합 심정지 탐지 시스템을 평가하기 위하여 테스트셋 1,198건을 구축하였다. 테스트셋 구축은 서울 소방본부 내에서 진행되었으며, 심정지 799건, 기타질환 399건의 비율로 구성하였다. 본 테스트셋을 활용하여 자동전사 모델과 텍스트 기반의 심정지 모델을 통하여 모델의 정확도를 측정하였다.

3.1.4. 비식별화 이슈

서울 소방본부에서는 통화녹취록 반출을 텍스트DB에 한하여 허가하였다. 그러나 반출을 하기위해서는 비식별화 작업을 진행하였다. 특히 주소, 위치, 전화번호, 이름 등의 개인정보에 대해서는 [주소], [위치], [전화번호], [이름]으로 변환하는 방식으로 비식별화를 진행하였다.

3.2. 메트릭

3.2.1. 통화 자동 전사 모델의 성능에 관한 메트릭

통화 자동 전사 모델의 성능에 관한 메트릭은 WER(word error rate)와 CER(character error rate)를 이용하였다. WER과 CER은 음성인식에서 주로 사용되는 메트릭으로, 삽입(insertion), 대치(substitution), 삭제(deletion)을 통해, 자동 전사의 정확도를 판단하는 척도이다.

3.2.2. 심정지 탐지 모델의 성능

심정지 탐지 모델의 성능을 측정하기 위한 메트릭은 정확도

에 관한 메트릭 즉 F1 score, accuracy 등을 이용하였다. 또한 모델의 예측 정도를 보기 위해 confusion matrix를 이용하였다.

3.2.3. 심정지 인지시간 단축을 측정하는 메트릭

심정지 인지시간 단축을 측정하는 척도는 초(s)이며, 심정지라고 판단한 589건 중 구급에서 심폐소생술을 시행한 231건에 대해서만 평가를 수행하였다. 심정지 인지 단축시간은 두 가지 기준에 의해서 측정하였다. 기준 1은 모델이 심정지라고 예측한 시간과 수보요원이 심정지라고 예측한 시간의 차이이다. 기준 2는 모델이 심정지라고 예측한 시간과 통화가 끝난 시점의 시간의 차이이다.

4. 실험 결과

4.1. 신고자와 수보요원의 통화 자동전사 성능

신고자와 수보요원의 통화 녹취 자동전사 성능은 다음과 같다. 서울소방으로부터 제공받은 테스트셋 100건에 대해 WER 기준으로 12.5%이며 CER 기준으로 6.7%이다.

4.2. 머신러닝 기반의 심정지 탐지 시스템의 성능

본 실험에서 제안한 머신러닝 기반의 심정지 탐지 시스템의 성능은 테스트셋 1,198건에 대하여, F1 score 기반 79.49%이며, 정확도는 74.62%이다(표 4).

표 4. 머신러닝 기반의 심정지 탐지 모델의 성능
Table 4. Assessment of the performance of a machine learning-based cardiac arrest detection model

평가 척도	값(%)
Sensitivity	73.72
Specificity	76.44
Precision	86.44
False positive rate	23.56
False negative rate	26.28
Accuracy	74.62
F1 score	79.49

Confusion matrix를 살펴보면 심정지를 심정지라 판단한 경우는 589건, 심정지가 아닌 것을 심정지가 아니라고 판단한 경우는 305건이었다(표 5).

표 5. 음향 모델 트레이닝셋
Table 5. Confusion matrix for OHCA detection results

	True positive	True negative
Predicted positive	589	94
Predicted negative	210	305

4.3. 머신러닝 기반의 심정지 탐지의 인지시간

인지시간 단축에 관한 기준은 다음과 같다. 실제 심정지를 심정지라 판단한 231건에 관해서 최초 심정지 인지 시간과 수보요원의 인지시간 차이를 기준(기준 1)으로 하면 15초 단축되었

다. 실제 심정지를 심정지라 판단한 231건의 경우 최초 심정지 인지시간과, 통화 종료 시점 간의 인지시간 차이를 기준(기준 2)으로 하면 21초 단축하였다(표 6).

표 6. 심정지 인지시간 단축
Table 6. Reduced OHCA detection time compared to human call-taker in 119

	기준 1	기준2
Case	231	231
Predicted positive	15초	21초

4.4. 수보요원과 신고자 통화 내용 분석

실험 분석 결과는 다음과 같다. 우선 전체 녹취록에서는 주소와 위치에 관한 내용이 가장 고빈도로 나타났다. 그리고 호흡여부와 의식 여부 그리고 전화번호에 대한 단어가 많이 나타났다. 수보요원이나 신고자에 대해서도 비슷한 경향성을 보였다. 주소와 위치에 대한 질문과 답변, 호흡과 의식에 대한 답변 등이 고빈도로 빈출되었다(표 7).

표 7. 심정지 관련 통화 내용 최다 빈출 단어(Top 10)
Table 7. Top 10 most frequent words in OHCA-related call conversations

	전체		수보요원		신고자	
1	[주소]	10,685	[주소]	3,242	[주소]	7,443
2	[위치]	3,698	[위치]	970	[위치]	2,726
3	주소	1,051	여보세요	473	전화	779
4	[주소]요	995	전화	420	[주소]요	694
5	숨	873	주소	405	주소	645
6	숨을	783	잠시만요	375	숨을	599
7	의식	768	[주소]요	301	숨	577
8	구급차	693	숨	296	의식	514
9	의식이	664	의식	256	의식이	491
10	[전화번호]	621	구급차	234	구급차	456

5. 토론

본 연구에서는 음성인식 기반의 심정지 탐지 프로그램을 제안하였다. 심정지 탐지 프로그램은 개인 정보 보호 이슈 및 데이터 수집의 어려움에 따라 활발히 연구되어 오지 못했고, 대부분의 응급의료 시스템에서의 머신러닝 연구는 텍스트 기반에 머물러 있었다. 본 연구는 음성 기반의 시스템으로 신고자와 수보자의 음성 통화를 활용하여 심정지를 판단하는 시스템을 개발했다는 데 의의가 있다.

또한 실험 결과를 바탕으로 논의해 볼 점은 다음과 같다.

첫째, 신고자와 수보요원의 자동전사 모델의 정확도는 100건에 대해 단어오류율(WER) 12.5%이며, 음절오류율(CER) 기준으로 6.7%이다. 음성인식의 모델의 정확도는 잡음에 의해 크게 영향을 받는 경우가 많다. 기존 연구와 비교했을 때, 자동전사의 모델은 상대적으로 좋은 성능을 보여줬다. 그러나 초기 훈련 모델에서는 특정 지역명, 상호명과 같이 고유명사에 대해서는 상대적으로 낮은 정확도를 보여줬다. 본 연구에서는 이를 극복하기 위해 언어 모델에 주소 DB를 보충하여 성능을 보완하고자

하였다. 실제 수보요원과 신고자의 녹취록을 보면, 주소를 확인하는 부분이 대화의 대부분을 차지하는 경우가 많았다. 즉 자동전사 모델에서는 신고자의 주소와 위치 정보를 정확하게 전사해줄 수 있는지 여부가 심정지 통화 자동 전사에서 중요한 요소로 작용할 것이다.

둘째, 음성 기반의 심정지 모델의 성능은 정확도 기준 74.62%, F1 score 기준으로 79.49%, sensitivity 73.72%를 기록하였다. 해외 사례와 다르게 국내에는 음성 기반의 심정지 모델의 연구가 거의 이루어지지 않았다. 대부분의 선행 연구들은 응급 상황에서의 텍스트 기반의 분류 시스템을 구축하는데 초점이 맞춰져 있다. 그럼에도 본 연구에서 제안한 심정지 시스템은 높은 성능을 보여주지는 못했다. 이에 대한 원인은 데이터셋의 부족과 세부적인 라벨링의 부족 때문이다. 실제로 텍스트 기반의 심정지 판단 모델은 콜 전체 대화내용을 기반으로 학습되었다. 이는 수보요원이 통화 녹취록에서 심정지인지를 언제 판단했는지에 대한 라벨링이 없었기 때문이다. 즉 녹취록 전체에 해당하는 심정지 라벨링이 아닌 수보요원과 신고자가 발화한 개별 문장에 대한 라벨링이 가능하면 성능을 향상시킬 수 있을 것으로 본다.

셋째, 심정지 기반의 탐지 모델의 인지시간 단축시간은 15초이다. 국내의 선행 연구에서는 주로 응급 상황에서의 텍스트 기반의 머신러닝 시스템을 제안하였기에 직접적으로 비교하기는 힘들다. 해외의 몇몇 사례와 비교하여 데이터셋이 동일하지 않기 때문에 동일선상에 성능을 비교하기에는 어려움이 있다. 하지만 실제 실용화 단계에 이르기 위해서는 더 많은 데이터셋을 통한 다양한 상황에서의 녹취록의 학습이 필요하다. 만약 심정지 녹취록 데이터셋을 보강하여 강건하게 모델링 할 수 있다면, 짧은 대화로도 심정지 여부를 예측 할 수 있는 시스템을 개발할 수 있을 것이다. 또한 데이터셋이 보강된다면 소규모의 데이터에 적합한 머신러닝 알고리즘뿐 아니라, 대규모의 데이터셋에 적합한 딥러닝 알고리즘을 활용하여 성능을 향상시킬 수 있을 것이다. 그리고 이는 심정지 탐지 모델이 수보요원의 보조적인 역할을 수행할 수 있을 것이다.

넷째, 수보요원과 신고자 통화내용 분석을 통해서 심정지 수보지침의 사항들이 통화 내용에 반영되었는지를 살펴보았다. 실험 결과 심정지를 탐지하는데 있어서, 호흡과 의식의 여부를 확인하는 것은 중요하다는 점이 나타났다. 또한 시스템이 구급차 출동을 위하여 주소나 위치 전화번호 등의 사항에 대해서 명확하게 인지하는 것이 특히 중요하다는 점을 보여줬다.

6. 결론

본 연구는 수보요원과 신고자 사이의 통화내용을 분석하여 심정지 여부를 판단하는 소프트웨어로 구현한 연구이다. 본 연구에서는 딥러닝 방식의 음성인식 모델과 텍스트 기반 심정지 분류 모델을 활용하여 심정지 탐지 프로그램을 구현하였다.

본 연구의 한계점은 다음과 같다.

첫째, 수보요원과 신고자의 통화녹취 데이터셋의 부족이다. 일반적으로 딥러닝 모델을 트레이닝하기 위해서는 일정 정도

수준 이상의 데이터셋이 필요하다. 본 실험은 서울종합방재센터에서 특정 기간 동안 제한된 환경에서 수행되었기 때문에 소규모의 데이터로 실험을 진행하였다.

둘째, 심정지 여부를 세부적으로 판단하기 위해, 문장 단위의 수보 코드가 필요하다. 서울소방본부에서 제공받은 데이터셋은 전체 콜에 대해서 심정지 여부인지만을 태깅해놓았다. 이러한 태깅 정보의 부족을 극복하기 위해 “증상 표현 모듈”을 도입하였다. 그러나 이 문제를 근본적으로 해결하기 위해서는 개별 발화에 대한 수보 코드가 필요하다. 이러한 수보 코드를 활용하여 모델링 한다면, 심정지 판단의 정확도와 신뢰도를 높일 수 있을 것이다.

향후 연구에서는 문장 단위의 심정지 수보 코드를 정의하고, 개별 발화에 대한 모델링을 실행하는 방식으로 모델의 성능을 높일 예정이다.

감사의 글

본 연구는 소방청 연구개발사업인 “119 구급 신고 정보 표준화 및 자료 활용 방안 연구” 과제의 지원을 받아 진행되었습니다. 119 구급신고 데이터는 세종 소방본부, 제주 소방본부, 서울 소방본부의 협조를 받아 녹취 데이터 구축을 하였습니다.

References

- Berg, K. M., Cheng, A., Panchal, A. R., Topjian, A. A., Aziz, K., Bhanji, F., Bigham, B. L., ... Lavonas, E. J. (2020). Part 7: Systems of Care: 2020 American Heart Association guidelines for cardiopulmonary resuscitation and emergency cardiovascular care. *Circulation*, *142*(16_suppl_2), S580-S604.
- Blomberg, S. N., Folke, F., Ersbøll, A. K., Christensen, H. C., Torp-Pedersen, C., Sayre, M. R., Counts, C. R., ... Lippert, F. K. (2019). Machine learning as a supportive tool to recognize cardiac arrest in emergency calls. *Resuscitation*, *138*, 322-329.
- Bojanowski, P., Grave, E., Joulin, A., & Mikolov, T. (2017). Enriching word vectors with subword information. *Transactions of the Association for Computational Linguistics*, *5*, 135-146.
- Byrsell, F., Claesson, A., Ringh, M., Svensson, L., Jonsson, M., Nordberg, P., Forsberg, S., ... Nord, A. (2021). Machine learning can support dispatchers to better and faster recognize out-of-hospital cardiac arrest during emergency calls: A retrospective study. *Resuscitation*, *162*, 218-226.
- Chung, B. H., Shin, J. H., & Cho, U. H. (2015). Development of 119 integrated emergency management system training simulator system. *The Journal of the Korea Institute of Electronic Communication Sciences*, *10*(4), 461-467.
- Eisenberg, M. S., Horwood, B. T., Cummins, R. O., Reynolds-Haertle, R., & Hearne, T. R. (1990). Cardiac arrest and resuscitation: A tale of 29 cities. *Annals of Emergency Medicine*, *19*(2), 179-186.
- Gnesin, F., Møller, A. L., Mills, E. H. A., Zyliftari, N., Jensen, B., Bøggild, H., Ringgren, K. B., ... Torp-Pedersen, C. (2021). Rapid dispatch for out-of-hospital cardiac arrest is associated with improved survival. *Resuscitation*, *163*, 176-183.
- Kim, H. S., & Chun, J. H. (2020). Determinants on survival outcomes of sudden out-of-hospital cardiac arrest: A multilevel analysis. *The Korean Journal of Emergency Medical Services*, *24*(2), 7-26.
- Kim, J., Jeong J., & Kweon, S. (2022). Incidences of out-of-hospital sudden cardiac arrest in the Republic of Korea, 2021. *Public Health Weekly Report*, *15*(51), 3007-3020.
- Kim, J. H., Choi, A., Kim, M. J., Hyun, H., Kim, S., & Chang, H. J. (2022). Development of a machine-learning algorithm to predict in-hospital cardiac arrest for emergency department patients using a nationwide database. *Scientific Reports*, *12*(1), 21797.
- Lee, K., Chung, J., Shin, D., Chung, M., Kang, K., Jang, Y., & Jang, K. (2016). Emergency dispatching based on automatic speech recognition. *Phonetics and Speech Sciences*, *8*(2), 31-39.
- New Jersey Government. (2020). Emergency medical dispatch guidecards. Retrieved from <https://www.nj.gov/911/home/highlights/EMD%20Guidecards%202020%20Final.pdf>
- O’Keeffe, C., Nicholl, J., Turner, J., & Goodacre, S. (2011). Role of ambulance response times in the survival of patients with out-of-hospital cardiac arrest. *Emergency Medicine Journal*, *28*(8), 703-706.
- Oh, S. Y. (2020). 119 reports in 2021, up 7.1% (nearly 800,000) year-over-year. Retrieved from <http://m.safetyin.co.kr/news/articleView.html?idxno=15029>
- Park, I., Kim, Y., & Kang, S. (2014). A study on the step-by-step survival factors of prehospital cardiac arrest patients. *Health and Social Welfare Review*, *34*(2), 484-513.
- Park, J. H. (2022). *Use of machine learning based time-to-event analysis for prediction of field return of spontaneous circulation after out-of-hospital cardiac arrest* (Doctoral dissertation). Seoul National University, Seoul, Korea.
- Phoenix Government. (2016). Fire emergency dispatcher: 911 operator. Retrieved from <https://www.phoenix.gov/hrsite/documents/fireguide.pdf>
- Povey, D., Ghoshal, A., Boulianne, G., Burget, L., Glembek, O., Goel, N., Hannemann, M., ..., Vesely, K. (2011). The Kaldi speech recognition toolkit. *Proceeding of the IEEE 2011 Workshop on Automatic Speech Recognition and Understanding (No. CONF)*. Waikoloa, HI.
- Safar, P. (1986). Cerebral resuscitation after cardiac arrest: a review. *Circulation*, *74*(6 Pt 2), IV138-53.
- Yoon, H., Park, S., Ahn, J., & Lee, Y. (2020). Major findings of the 2006-2018 survey on acute myocardial infarction. *Weekly Health and Disease*, *13*(1), 77-90.

Vukmir, R. B., & Sodium Bicarbonate Study Group. (2004). Witnessed arrest, but not delayed bystander cardiopulmonary resuscitation improves prehospital cardiac arrest survival. *Emergency Medicine Journal*, 21(3), 370-373.

- **김종인 (Jong In Kim)** 교신저자
서울대학교 인지과학협동과정 박사과정
서울시 관악구 관악로 1 서울대학교 인문대학 3-301
Tel: 02-880-6164
Email: prowsl2@gmail.com
관심분야: 머신러닝, 음성인식, 음성분류
- **이주영 (Joo Young Lee)**
서울대학교 언어학과 박사과정
서울시 관악구 관악로 1 서울대학교 인문대학 3-301
Tel: 02-880-6164
Email: excalibur12@snu.ac.kr
관심분야: 음성학, 음성인식, 음성모델링
- **정지오 (Jio Chung)**
브이에스웍스 대표
서울특별시 강서구 염창동 240-21 1205호
Tel: 02-880-6164
Email: jiochung@vsworks.co.kr
관심분야: 음성처리, 음성인식
- **신대진 (Dae Jin Shin)**
사운드마인드 대표
서울특별시 강서구 염창동 240-21 1205호
Tel: 070-4643-8843
Email: djshin@eidware.com
관심분야: 음성처리, 음성인식
- **최동현 (Dong Hyun Choi)**
서울대학교 의과대학 (의공학) 연구원
응급의학전문의, 의학석박사과정(의공학)
서울특별시 종로구 대학로 101(연건동 28) 서울대학교병원
Tel: 02-2072-4052
Email: donghyun369@naver.com
관심분야: 응급의학
- **김기홍 (Ki Hong Kim)**
서울대학교병원 응급의학과 임상조교수
응급의학전문의, 의학석사(응급의학)
서울특별시 종로구 대학로 101(연건동 28) 서울대학교병원
Tel: 02-2072-4052
Email: emphysiciankkh@gmail.com

관심분야: 응급의학

- **홍기정 (Ki Jeong Hong)**
서울대학교병원 응급의학과
서울대학교 의과대학 응급의학교실 임상부교수
서울특별시 종로구 대학로 101(연건동 28) 서울대학교병원
Tel: 02-2072-4052
Email: emkjhong@gmail.com
관심분야: 응급의학
- **김선희 (Sunhee Kim)**
서울대학교 불어교육과 전임교수
서울시 관악구 관악로 1 서울대학교
Tel: 02-880-7693
Email: sunhkim@snu.ac.kr
관심분야: 음성학, 음운론, 음성언어처리, 발음교육
- **정민화 (Minhwa Chung)** 교신저자
서울대학교 언어학과 교수
서울시 관악구 관악로 1 서울대학교
Tel: 02-880-9195
Email: mchung@snu.ac.kr
관심분야: 음성학, 음성공학

119 응급신고에서 수보요원과 신고자의 통화분석을 활용한 머신 러닝 기반의 심정지 탐지 모델*

김 중 인¹ · 이 주 영² · 정 지 오³ · 신 대 진⁴ · 최 동 현^{5,6} · 김 기 흥^{6,7,8} · 홍 기 정^{6,7,8,9} · 김 선 희¹⁰ · 정 민 화^{1,2}

¹서울대학교 인문대학 인지과학 협동과정, ²서울대학교 인문대학 언어학과, ³브이에스웍스,
⁴사운드마인드, ⁵서울대학교 의과대학 의공학교실, ⁶서울대학교병원 의생명연구원 응급의료연구실,
⁷서울대학교병원 응급의학과, ⁸서울대학교 의과대학 응급의학교실,
⁹서울대학교 의학연구원 재난의학연구센터, ¹⁰서울대학교 사범대학 불어교육과

국문초록

심정지는 초기 대응에 따라 생존율과 예후에 영향을 미치는 중요한 응급 상황이다. 특히 병원밖심정지(out-of-hospital cardiac arrest, OHCA)의 경우, 119 구조대의 초기 조치가 심정지 환자의 생존율을 높이는 데 결정적인 역할을 한다. 그러나 국내에서는 수보요원의 수가 제한적이지만 다량의 신고 전화에 응대해야 하는 현실이다. 이런 상황에서 머신러닝 기반의 OHCA 탐지 프로그램은 수보요원의 보조 역할로 심정지 환자의 생존률을 높일 수 있다. 본 연구에서는 이러한 문제를 해결하기 위해 머신러닝 기반의 심정지(OHCA) 탐지 프로그램을 개발하였다. 이 프로그램은 수보요원과 신고자의 통화 녹취록을 분석하여 심정지 여부를 판단한다. 제안한 모델은 수보요원 및 신고자와의 통화를 자동으로 전사하는 모델, 텍스트 기반의 심정지 탐지 모델, 그리고 프로그램 개발을 위한 서버와 클라이언트로 구성되어 있다. 실험 결과, 본 연구에서 제안한 모델은 F1 점수 기준으로 79.49%의 성능을 보였으며, 수보요원과 비교하여 심정지 감지 시간을 15초 단축하였다. 이 연구는 소규모 데이터셋을 사용하였음에도 불구하고, 심정지 기반의 탐지 프로그램이 수보요원의 보조 역할로 심정지 생존률에 기여할 수 있음을 입증하였다.

핵심어: 심정지, 병원밖심정지(out-of-hospital cardiac arrest, OHCA), 심정지 탐지 프로그램

참고문헌

- 김지수, 권상희, 정지현. (2022). 2021년 급성심장정지 발생 현황. *주간 건강과 질병*, 15(51), 3007-3020.
- 김효실, 전진호. (2020). 병원 밖에서 발생한 급성 심장정지 환자의 생존결과에 영향을 미치는 요인: 다수준 분석. *한국응급구조학회지*, 24(2), 7-26.
- 박정호. (2022). *기계학습 기반 사건 시점 분석을 활용한 병원밖 심정지 환자의 구급 현장 자발순환회복 예측 모델 개발*. 서울대학교 박사학위논문.
- 박일수, 김유미, & 강성홍. (2014). 병원전 심정지 환자의 단계별 생존요인에 관한 연구. *보건사회연구*, 34(2), 484-513.
- 윤하정, 박수정, 안주연, & 이연경. (2020). 2006~2018년 급성심장정지조사 주요 결과. *주간 건강과 질병*, 13(1), 77-90.
- 오세용. (2022). 2021년 119신고, 전년 대비 7.1%(약 80 만건) 증가. Retrieved from <http://m.safetyin.co.kr/news/articleView.html?id=xno=15029>
- 이규환, 정지오, 신대진, 정민화, 강경희, 장윤희, & 장경호. (2016). 음성인식 기반 응급상황관제. *말소리와 음성과학*, 8(2), 31-39.

* 본 연구는 소방청 연구개발사업인 "119 구급 신고 정보 표준화 및 자료 활용 방안 연구" 과제의 지원을 받아 진행되었습니다. 119 구급신고 데이터는 세종소방본부, 제주소방본부, 서울소방본부의 협조를 받아 녹취 데이터 구축을 하였습니다.