

청각장애인을 위한 위험 소리 감지 및 알람 시스템

조유겸, 황태욱, 배무성, 임태훈, 김정창

국립한국해양대학교

bamppy12@g.kmou.ac.kr, xodnr1130@g.kmou.ac.kr, baemuseong@g.kmou.ac.kr,
lth3097@g.kmou.ac.kr, jchkim@kmou.ac.kr

Hazardous Sound Detection and Alert System for Hearing Impaired People

Yu-gyeom Cho, Tae-uk Hwang, Museong Bae, Taehun Lim, Jeongchang Kim

Korea Maritime and Ocean University (KMOU)

요약

본 논문에서는 청각장애인이 위급 상황 소리를 감지하기 어려운 문제를 해결하기 위해 스마트폰의 내장 마이크로부터 소리를 수집하여 CNN 기반 모델로 실시간 분석 후, 스마트폰에서 진동/시각 알람으로 사용자에게 위험 상황을 즉각적으로 전달하는 모바일 애플리케이션을 개발하였다.

I. 서론

2013 년 인천에서는 환경미화원으로 근무하던 청각장애인이 뒤에서 돌진해오는 차량의 소리를 인지하지 못해 교통사고를 당하여 사망한 일이 발생하였다[1]. 이처럼 위급 상황에서 소리 정보를 인지하기 어려운 청각장애인은 교통사고 및 위급 상황에서 대피 지연 등 다양한 위험에 노출될 수 있다[2]. 이에 본 논문에서는 스마트폰을 활용하여 청각장애인을 위한 모바일 기반 위험 소리 감지 및 알람 시스템을 설계하고 구현한다. 소리 분류를 위해 CNN(Convolutional Neural Network)을 채택하였으며, 오디오 데이터를 Log-Mel Spectrogram 으로 변환하여 입력하였다. 제안 시스템은 환경 소리를 실시간으로 수집하고, 딥러닝 모델을 통해 분석하여, 위험이 감지되면 사용자에게 진동과 시각적 알람을 제공하는 기능을 갖춘다.

II. 본론

본 논문에서 제안한 소리 감지 및 알람 시스템은 스마트폰의 내장 마이크를 이용하여 2 초 동안

오디오를 수집하여 Log-Mel Spectrogram 으로 변환한다. 이는 오디오 신호를 시간-주파수 영역의 2 차원 이미지 형태로 표현함으로써 CNN 모델이 각 클래스의 패턴을 효과적으로 학습할 수 있도록 돕는다[3]. 변환된 스펙트로그램은 딥러닝 기반 분류 모델에 학습데이터 및 입력데이터로 사용되며, 모델은 siren, car_horn, noise 의 3 가지 클래스를 판별한다. 이 중 siren 또는 car_horn 클래스의 예측 확률이 0.6 이상일 경우 해당 소리를 위험으로 간주하며, noise 클래스가 감지된 경우에는 위험 소리로 판단하지 않는다. 또한, 동일한 위험 소리가 3 회 예측 주기 동안 연속으로 감지되었을 때, 푸시 알람이 발생하도록 하여 중복 알람을 방지한다.

본 시스템에서는 경량화를 위해 학습된 모델을 TensorFlow Lite 형식으로 변환하였으며[4], 실시간 동작이 가능하도록 구현하였다. 그림 1 은 각 클래스의 Log-Mel Spectrogram 예시를 보여준다. 3 가지 클래스 중 siren 은 중간 주파수 대역의 변조된 패턴, car_horn 은 저역 중심의 강한 수평 패턴, noise 는 전 주파수에 걸쳐 고르게 분포된 불규칙적인 에너지

특성을 보인다. 이러한 스펙트럼 차이로 인하여 소리 종류의 분류가 가능하다.

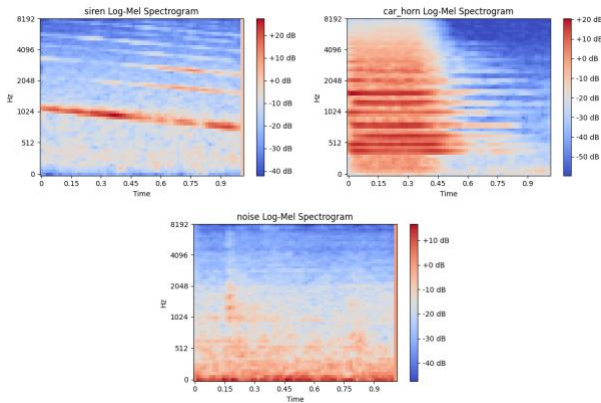


그림 1. 각 상황별 Log-Mel Spectrogram

사용자는 감지 결과를 애플리케이션 화면을 통해 실시간으로 확인할 수 있다. 그림 2 는 위험 소리가 탐지되었을 때의 앱 실행 화면을 보여주며, 상단에는 감지된 소리에 대한 상태 정보가, 하단에는 사용자가 직접 감지를 제어할 수 있는 ‘감지 중지’ 버튼이 제공된다. 본 연구에서는 아이폰에서 사용 가능한 앱을 구현하였다.

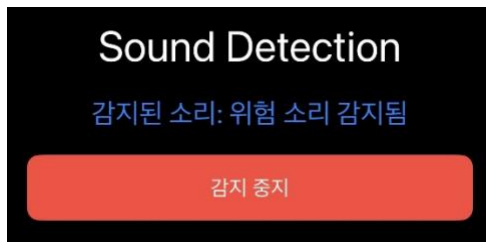


그림 2. 위험 소리 감지 시 애플리케이션 실행 화면.

모델의 성능 평가는 학습에 사용되지 않은 테스트셋 300 개(각 클래스당 100 개)를 기반으로 진행되었으며, 그 결과는 아래의 그림 3 을 통해 확인할 수 있다.

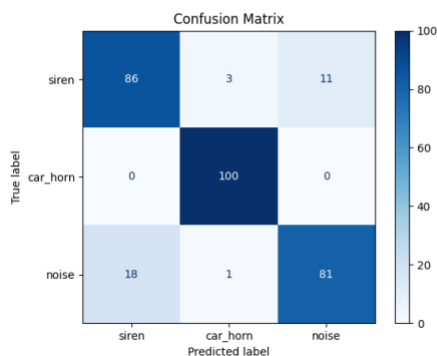


그림 3. 각 클래스 분류 결과의 Confusion Matrix

전체 평균 정밀도(Precision)는 89.0%로 측정되었으며, 클래스별 F1-score 는 car_horn 98.1%, noise 84.4%, siren 84.3%로 나타났다. 특히 car_horn 클래스는 정밀도(Precision) 96.2%, 재현율(Recall) 100.0%로 매우 우수한 성능을 보였다. 반면 siren 클래스는 noise 클래스와의 혼동으로 인해 정밀도(82.7%)와 F1-score 가 상대적으로 낮게 측정되었으며, 이는 향후 다양한 사이렌 데이터를 통한 학습 개선이 필요함을 보여준다.

III. 결론

본 논문에서는 청각장애인을 위한 실시간 위험 소리 감지 및 알림 시스템을 제안하였다. 제안한 모델은 Log-Mel Spectrogram 과 CNN 기반 분류기를 활용하여 위험 소리 발생 시 스마트폰 진동과 시각적 알림으로 실시간 경과를 제공한다. 그리고 각 클래스들을 전반적으로 높은 정밀도와 F1-score 로 구분하였으며, Confusion Matrix 분석을 통해 클래스 간 경계도 확인할 수 있었다. 향후에는 다양한 환경과 소리 유형에 대한 추가 학습과 함께 Apple Watch 등 웨어러블 기기와 연동을 통해 접근성과 활용도를 더욱 높일 수 있을 것으로 기대된다.

참 고 문 헌

- [1] 에이블뉴스. “청각장애인 위험감지·경고시스템 특성화 필요성-①.” 2019 년 1 월 7 일.

<https://www.ablenews.co.kr/news/articleView.html?idxno=86087>

- [2] 한국교통안전공단. 『2022 년 보행자 사고 통계 보고서』. 세종: 한국교통안전공단, 2022.

<https://www.koroad.or.kr/>, pp. 3-5

- [3] McFee, Brian, et al. “librosa: Audio and Music Signal Analysis in Python.” SciPy 2015 Proceedings, 2015. https://brianmcfee.net/papers/scipy2015_librosa.pdf, pp. 1-3.

- [4] TensorFlow. “TensorFlow Lite Documentation.”

<https://www.tensorflow.org/lite>