여성안심 귀갓길 비명 인식 인공지능 시스템

오윤선[1], 이다온[2], 이솔[3], 김형훈[4], 최재준[5]
[1 2 3 5]한양여자대학교 스마트 IT 과
[4](주) 삼성전자

ysun0906@naver.com, rainbow16457@naver.com, leesol7531@naver.com,
pastelom@gmail.com, jchoi@hywoman.ac.kr

Scream Recognition AI System for Women's Safe Way Home

Yoon-Sun Oh[1], Da-On Lee[2], Sol Lee[3], Hyeong-Hoon Kim[4], Jae-Jun Choi[5]
[1 2 3 5]Dept. of Smart IT Hanyang Women University
[4]Samsung Electronics Co. Ltd

요 약

최근의 범죄 동향을 살펴보면 묻지마 범죄와 여성 혐오 범죄가 계속해서 늘어나는 추세이다. 낮과 밤을 구분하지 않고 일어나는 사건·사고들에 특히나 범죄에 노출되기 쉬운 여성들은 일상생활에서조차 안전을 보장받지 못하고 있다. 이를 위해 생겨난 여성안심 귀갓길이라고 지정된 보호구역은 보호라는 말이 무색하게 제대로 사용할 수 있는 방범 시스템이 제대로 마련되어 있지 않은 실정이다. 이와 같은 상황을 개선하기 위해 기존의 시각적 방범과는 달리 탐지 범위가 넓고 상대적으로 제한이 적은 음향으로 접근한 새로운 방범 시스템 개발을 통해 여성들의 안전을 보장할 수 있도록 제안한다.

1. 서론

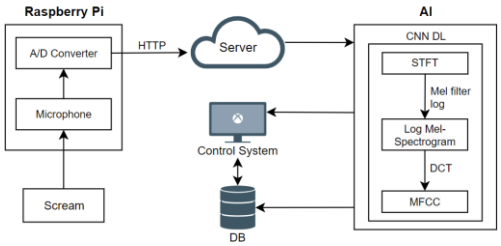
일반적으로 '감시 시스템'이라고 하면 CCTV 와 같은 이미지 및 비디오 신호 기반 모니터링이 주류를 이룬다. CCTV 는 눈, 비, 안개 및 밤과 같이 영상 촬영이 어려운 환경에서는 객체를 인식하는 데 문제가 있으며, 객체 간의 간섭이 존재할 때에도 목표 객체 인식이 상당히 어렵다.[1]

이러한 문제점을 개선하기 위해 본 논문을 통해 상대적으로 환경의 제약에 강하고 탐지 범위가 넓은 음향 인식을 활용하여 상황 판단 및 즉각적인 신고 기능이 탑재된 인공지능 모델을 구축한다.

관제 시스템을 통해 주변 관할서로 신고 대처가 이뤄지고 범죄 현황에 대한 정보 제공을 통해 여성들의 불안감을 해소하는 것이 본 논문의 목표이다.

2. 제안하는 시스템

2-1. 시스템 구조



(그림 1) 시스템 구조

라즈베리파이의 마이크를 통해 음향 데이터를 수집하여 인공지능 모델로 전송한다. 전달받은 소리를 인공지능이 분석하고, 위급상황으로 판명되었을 경우 관제 시스템으로 신고 알림을 전송한다.

2-2. 시스템 동작 절차

라즈베리파이에 연결한 마이크를 통해 주변 음향 데이터를 수집한 후, 마이크에 부착된 A/D 컨버터로 신호를 변환한다.

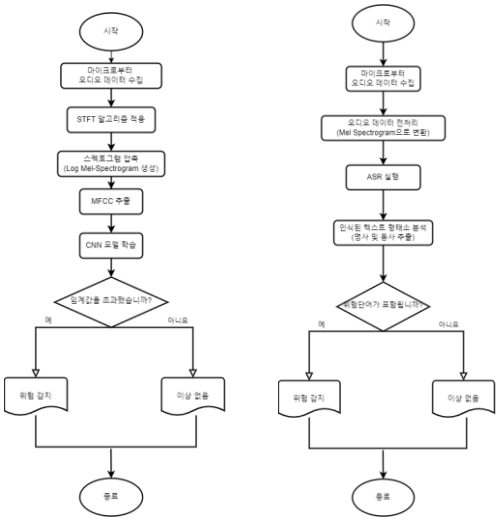
음향 데이터에 STFT 알고리즘 적용 후 스펙트로그램을 추출한다. 효율성과 성능 향상을 위해 스펙트로그램의 주파수 성분을 Mel 필터에 따라 압축한다.

$$M(f) = 2595 \cdot \log_{10}\left(1 + \frac{f}{700}\right)$$

진폭 축에 Log 를 적용해 한 번 더 압축하여 최종적으로 Log Mel-Spectrogram 을 생성한다. 이를 통해 소리들의 공통되는 특징 샘플링 데이터를 파악한다. 여기서 한 단계 더 나아가 저음 특성이 부각되도록 Log Mel-Spectrogram 에 DCT 를 적용해 압축한 MFCC 알고리즘을 적용한다.

음성인식을 위하여 OpenAI 인 whisper 를 사용하였다. 이를 통해 인식된 음성의 문장과 신고가 될 음성의

문장을 설정한 후 비교 분석하여 두 문장이 겹칠 때 신고 기능이 작동하게 한다. 공백이나 특수문자 등의 예외를 줄이기 위해 komoran 라이브러리를 사용하여 형태소 분석으로 이 중 동사와 명사만을 추출하고 제거 및 추가할 단어를 정의해 비교하였다.



(그림 2) 음향분류 모델 순서도
(그림 3) 음성인식 모델 순서도

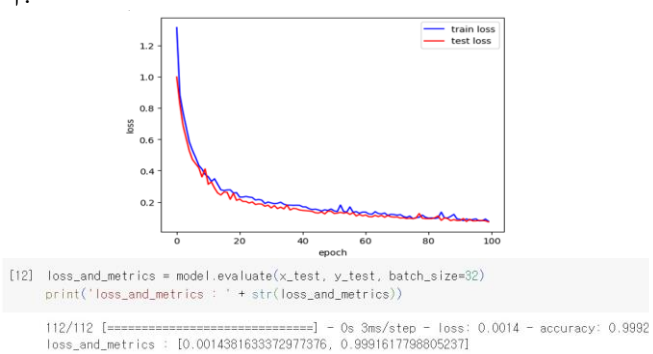
위 과정을 거쳐 변환된 그래프 이미지 데이터에 CNN 모델을 적용하여 음향과 음성으로 상황을 판단하는 인공지능 모델을 구축한다.[2][3]

인공지능이 위급상황으로 판단된 결과를 관제 시스템으로 전달하고 신고한다. 신고로 이어진 데이터들은 DB 에 축적되고, 이를 바탕으로 범죄 현황(발생 시각, 장소 등)에 대한 통계를 도출할 수 있다.

3. 학습 결과

본 연구에서는 위급 상황에서의 음성과 비명, 일반 환경 소리를 직접 녹음하고 추가로 AI-Hub 에서 수집하여 데이터셋을 정제하였다. 총 8,200 개의 파일에서 데이터 증강 기법을 활용하여 20,000 개의 데이터셋으로 학습하였고, 위급상황과 일반 상황을 분류하기 위해 각 클래스의 확률을 추정하는 softmax 함수를 마지막 층에 사용하였다.[2]

학습 결과 그림 4 와 같이 99%의 높은 정확도를 보였다.



(그림 4) 학습 결과 그래프 및 로그

4. 결론

4-1. 결론 요약

본 논문에서는 여성안심 귀갓길에 기존 방범 기기들과는 차별성이 있는 새로운 방범 시스템을 제안하였다.

피해자의 신고가 들어왔을 때는 범죄 행위가 이미 종료된 후가 대부분이지만 본 시스템은 인공지능이 사고 발생 당시에 범죄를 인식함과 동시에 신고가 이루어진다. 즉, 피해자가 조치를 취할 수 없더라도 실시간으로 본인의 상황을 알릴 수 있다는 것이다.

비명 인식으로 즉각적인 신고가 들어가는 시스템이 있다는 것을 인지하고 있는 것만으로도 어두운 귀갓길을 지나가는 여성들의 두려움을 완화시켜줄 수 있다. 기존 방범 시스템에서 65%의 여성들이 귀갓길 불안감을 호소한 반면, 비명 인식기 시연 결과 42%로 23% 감소하였음을 알 수 있었다.

관제 시스템에 축적된 빅데이터를 바탕으로 범죄 발생 현황을 대시보드를 통해 시각적으로 나타내어 해당 구역은 지속적인 순찰과 경비 시스템으로 보호하고 있음을 보여준다. 이를 통해 보행자의 불안감을 해소하고, 범죄자에게는 심리적 위축 효과를 줄 수 있다.

4-2. 향후 연구과제

3 장에서의 검증 결과를 통해 높은 정확도를 보였다는 것을 확인할 수 있었다. 이번 연구에서는 여성안심 귀갓길로 축소해 진행하였지만, 더 다양한 분야로 확장하여 연구할 것이다.

기존 방범 기기인 CCTV 와 결합하여 비명이 발생한 방향으로 카메라를 돌려 한층 더 정확히 판별하는 시스템 구축에 활용할 수 있다. 반대로 영상 관제가 불가능한 제한적인 장소(공중 화장실, 탈의실 등)에도 활용될 수 있을 것으로 기대된다.

더 나아가 특정 장소에 한정되지 않고, 음향 분석 및 실시간 상황 전달 기술을 바탕으로 도로 사고 발생 신고, 조난자 구조 등 다양한 상황에 활용될 수 있을 것으로 기대된다.

참고문헌

[1] 유광현, "딥러닝 기반의 비명소리 인식 및 구현", 공학석사학위논문 전남대학교 대학원, pp.7, 2018

[2] 이소연 외 3 명, "음향 데이터 기반 지능형 상황 인식 시스템" 대한전자공학회 하계학술대회 논문집, pp.2731, 2022

[3] 이소연 외 2 명, “음향 데이터를 활용한 딥러닝 기반 긴급차량 우선 신호 시스템”, Jounal of Platform Technology, 제 9 권, 제 3 호, pp.47-49, 2021

※ 본 프로젝트는 과학기술정보통신부 정보통신창의 인재양성사업의 지원을 통해 수행한 ICT 멘토링 프로젝트 결과물입니다.