

인공지능 기반 다중 소리 감지 스마트허브 제작

이태민⁰, 성병준*, 이창현*, 김성수*, 김병수*, 한찬우*, 박준호(교신저자)*

⁰경운대학교 항공소프트웨어공학과,

*경운대학교 항공소프트웨어공학과

e-mail: xoals4354@naver.com⁰, {wnsdl700440, amckdgis, tjdtm1031, todaud0032, hhgg7884}@naver.com*, jhpark@ikw.ac.kr*

Artificial intelligence-based multi-sound recognition smart hub production

Tae-min Lee⁰, Byung-jun Sung*, Chang-heon Lee*, Seong-soo Kim*, Byeong-su Kim*, Chan-woo Han*, Joon-ho Park(Corresponding Author)*

⁰Department of Aeronautical Software Engineering, Kyungwoon University,

*Department of Aeronautical Software Engineering, Kyungwoon University

● 요약 ●

본 논문에서는 딥러닝 소리 인식을 이용하여 실내에서 발생할 수 있는 다양한 소리를 시각적인 정보로 제공하는 스마트허브 시스템을 제안한다. 인공지능 모델은 2D-CNN 구조를 활용하여 학습을 진행하였고, 스마트허브 하드웨어는 라즈베리파이를 이용하여 구현하였다. 제안된 시스템은 청각장애인을 위해 설계된 다양한 청각 정보를 시각 정보로 전달하는 다양한 제품을 하나로 대체할 수 있을 뿐만 아니라, 설치 및 운반이 간편하여 누구나 사용하기 쉬워서 활용도가 높을 것으로 기대된다.

키워드: 딥러닝(Deep Learning), 소리 인식(Sound Recognition), 청각장애인(deaf person)

I. Introduction

청각장애인은 대부분의 판단을 시각을 통해 이뤄지기 때문에 일반인보다 수용할 수 있는 정보량이 제한적이다. 정보의 접근성이 낮아서, 재난이 발생할 때 정보를 늦게 획득하거나, 정보를 전혀 인지 못하는 일이 있다[1]. 경보기의 경우 소리뿐만 아니라, 시각적으로 위험을 알려줄 수 있는 시각 경보기의 의무설치가 되어 있으나 제약사항이 존재한다. 본 논문에서는 딥러닝 소리 인식을 이용하여 실내에서 발생할 수 있는 다양한 소리를 시각적인 정보로 제공하는 스마트허브 시스템을 제안하였다.

정보를 분류할 수 있다.

Table 1. 학습 데이터

단위(개)

아기 울음	경적	개	고양이	유리 깨짐	노크	비명	경보	도난	화재	응급
400	400	400	320	400	300	358	400	277	300	300

II. Preliminaries

본 논문에서는 높은 정확도를 보장하는 CNN 구조를 이용하여 학습을 진행하였다. 소리 인식 인공지능은 총 4개로 다중 소리 분류 인공지능은 기본 감지, 사이렌 감지로 구성되어 있고, 이중 소리 분류 인공지능은 아기 울음소리 감지, 비명 감지로 구성되어 있다. 기본 감지 기능의 소리 분류는 아기 울음소리, 경적, 개 짖는 소리, 유리 깨지는 소리, 노크 소리, 고양이 울음소리, 비명, 사이렌 소리를 분류할 수 있고 사이렌 감지 기능에는 도난 경보, 화재 경보, 응급

III. Design and Development

1. Development Environment

프로그램 개발은 Window10 환경에서 진행되었다. 사용된 라이브러리는 다음과 같다.

Table 2. Library version

Library	PyQt5	librosa	keras	numpy
version	5.15.7	0.9.2	2.11.0	1.23.5

2. Data preprocessing

인공지능 학습을 위해 각기 다르게 포맷된 소리 데이터를 정규화시켜주고 특징추출기법을 활용하여 MFCC (Mel-Frequency-Cepstral-Coefficient)와 Mel-Spectrogram으로 변환해주었다.

학습할 소리 데이터를 모노 타입으로 통일하고 각기 다른 Sampling Rate를 Re-Sampling 하여 22,050Hz로 통일시켜주고, Bit Depth를 정규화시켜주는 과정을 거쳤다. 그리고 소리 데이터에서 특징을 추출하기 위해 일반적으로 많이 활용하는 MFCC와 Mel-Spectrogram으로 변환해주었다. MFCC 기법은 Mel-Spectrum에서 배운 구조를 분석하여 추출된 값으로 역푸리에 변환을 적용하여 주파수 정보의 상관관계가 높은 문제를 해결한 기법이며, Mel-Spectrogram 기법은 인간의 청각 특징을 반영한 mel scale이 적용되고 주파수 특성이 시간에 따라 달라지는 오디오를 분석이 가능한 기법이다[2].

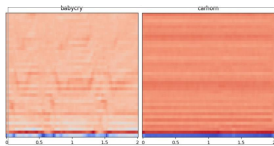


Fig. 1. MFCC

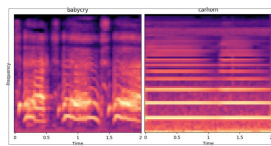


Fig. 2. Mel-Spectrogram

3. Model Structure

모델은 Conv2D 4-layer, Max_pooling2D 3-layer, Global average pooling2D 1-layer, Dense 3-layer로 구성하여 MFCC와 Mel-Spectrogram으로 추출된 데이터로 각각 학습을 진행하였다.

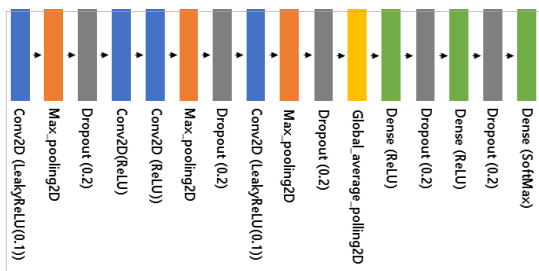


Fig. 3. Model Structure

4. Learning results

학습 결과 MFCC 기법을 이용한 모델은 Test Accuracy가 95.73171%, Mel-Spectrogram 기법을 이용한 모델은 Test Accuracy가 93.29268%를 기록하였고, MFCC 기법을 활용한 모델로 시스템을 구성하였다.

Table 3. Learning results

Model	Training	Validation	Test
MFCC	99.95748%	97.651%	95.73171%
Mel-Spectrogram	98.2568%	96.30873%	93.29268%

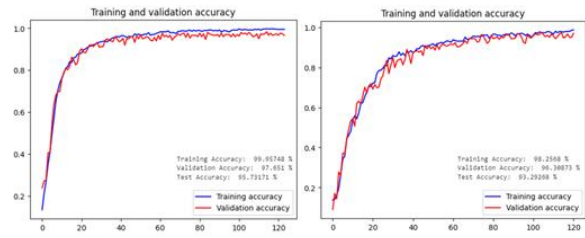


Fig. 4. MFCC Accuracy(left), Mel-Spectrogram Accuracy(right)

5. System Process

실시간 소리 감지가 실행되면 소리 인식 부분이 Loop문 안에서 계속 반복되며 작동한다. 해당 부분은 GUI의 원활한 작동을 위해 멀티스레드 형태로 구현하였다.

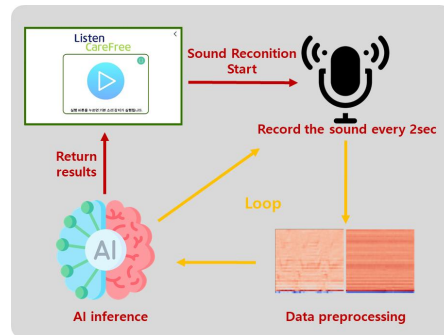


Fig. 5. AI inference process

IV. Conclusions

본 논문에서는 인공지능 기반 다중 소리 감지 스마트허브를 제시하였다. 기존의 청각장애인은 일상생활 속 소리를 잘 들을 수 있도록 제안된 스마트허브를 활용하여 들려오는 일상생활 속 소리를 감지해 청각 장애인에게 시각적 정보로 변환하여 알려준다.

본 시스템은 스마트허브에서 시각적인 정보로 변환해주기에 스마트허브를 주기적으로 확인해야 하는 제약이 있다. 더 많은 소리 들을 학습시켜 범용성을 늘릴 필요가 있고 스마트허브를 매번 확인하기에는 어려움이 있어 IoT 기술을 도입하여 스마트허브를 확인하지 않고도 정보를 전달받는 기능은 앞으로의 개선안으로 남겨둔다.

REFERENCES

[1] 이현아 and 이준우. (2017). 청각장애인의 위험 판단을 위한 소리 정보, 인지 조건과 행동 반응에 관한 연구. 한국장애인복지학, 38(38), 149-180.
 [2] 정기원. (2021). Introduction to Analysis for Sound Data, Korea University.