2025-2 AI 프로젝트IV HCI 시스템 설계

위험소리 감지 알림 UX 서비스 기획 (3)

인공지능학과 22619027 지정원

Contents

00 Previous Research

01 FlowChart

02 Dataset

03 Experiments

- LightWeighted CNN
- RCNN

00. Previous research

소리 분류 모델을 이용한 골목길에서의 차량-보행자 충돌 위험
방지 시스템 (2024.12)

그림 1. 앱 초기 화면(좌), 차량 인식 화면(오)

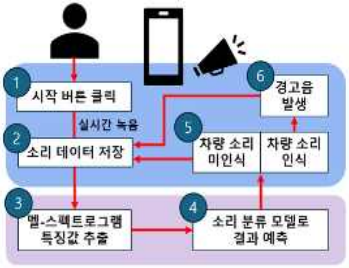


그림 2. 앱 동작 구조

표 1. LSTM, CNN 기반의 소리 분류 모델 성능 비교

	LSTM	CNN
Accuracy	93.3%	96.2%
Precision	93.5%	96.4%
Recall	93.3	96.2%
F1-score	93.3	96.2%

00. Previous research

Smart headphones warn of nearby cars- [Journal Watch](Jeremy Hsu, *IEEE Spectrum*, Vol. 57, No. 2, Feb. 2020, pp.8)

Designing Smart Headphones That Warn Pedestrians of Imminent Dangers | Columbia Science Commits

Designing Smart Headphones That Warn Pedestrians of Imminent Dangers

December 9, 2019

Florida, R. (2019, December 9)*Designing smart headphones that warn pedestrians of imminent dangers.*Columbia Science Commitment.

JOURNAL WATCH

Smart Headphones Warn of Nearby Cars

HOW CAN PEDESTRIANS safely tune out the world? Perhaps with a pair of intelligent headphones that alert them to oncoming vehicles.

The number of pedestrians killed on U.S. roads reached a three-decade high in 2018. Smart headphones are unlikely to prevent most pedestrian deaths—but a few seconds' warning might spare some lives.

Xiaofan Jiang, an assistant professor of electrical engineering at Columbia University, developed the Pedestrian Audio Wearable System (PAWS) with collaborators at the University of North Carolina at Chapel Hill and Barnard College. They published their work in the October 2019 *IEEE Internet of Things Journal.*

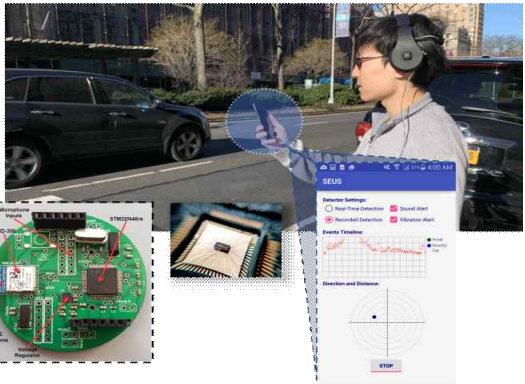
Many cars with collision warning systems rely on cameras, radar, or lidar to detect objects. But Jiang and his colleagues decided to use inexpensive microphones to serve as low-power sensors for their system.

The group placed four microphones in different

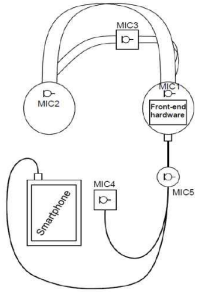
00. Previous research

S. Xia et al., "Improving Pedestrian Safety in Cities Using Intelligent Wearable Systems," in IEEE Internet of Things Journal, vol. 6, no. 5, pp. 7497-7514, Oct. 2019.

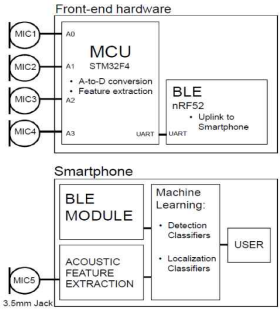
PAWS: Improving Pedestrian Safety – Columbia ICSL



System Implementation



PAWS three stage segmented architecture



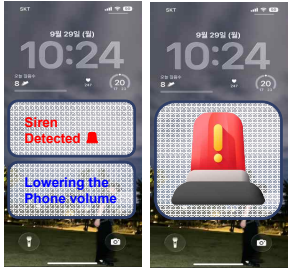
00. Previous research

- 선행 연구와의 차이점

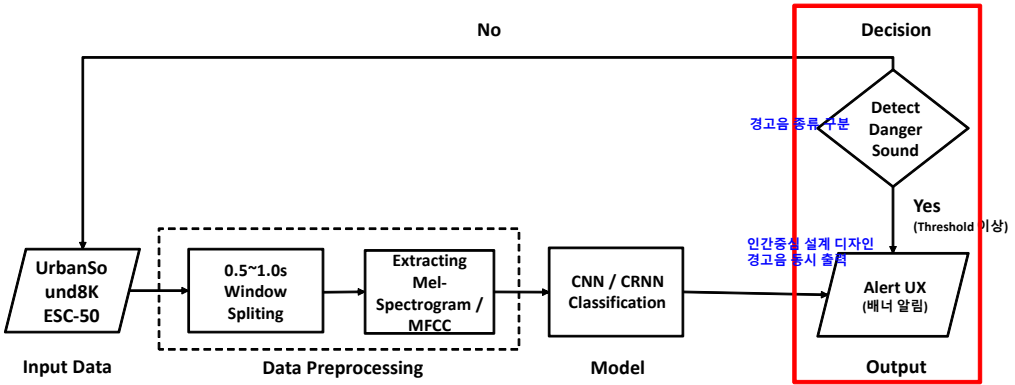
	선행 연구	진행 연구
① 연구 초점	하드웨어 기반 실시간 탐지 시스템 구현	AI 탐지 결과를 사용자에게 어떻게 전달할 것이냐? (UX 디자인)

	선행 연구	진행 연구
② 경고	이어폰에서 차 경적소리 감지하면 단순 “삡” 소리	어떤 위험음인지 “종류”까지 구분 “경고음” 삽입 볼륨 down

00. Previous research

	선행 연구	진행 연구
③ 디자인	 개발자 중심 설계 디자인 (DCD)	 사용자 중심 설계 디자인 (UCD)

01. FlowChart



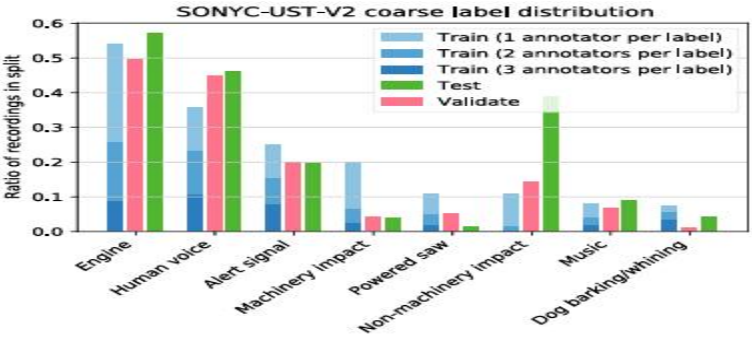
02. Dataset

fold별 car_horn(1)/siren(8) 분포:

classID	1	8
fold		
1	36	86
2	42	91
3	43	119
4	59	166
5	98	71
6	28	74
7	28	77
8	30	80
9	32	82
10	33	83

car horn	 72567-1-1-0.wav	 145577-1-0-0.wav
siren	 102871-8-0-0.wav	 159743-8-0-0.wav

02. Dataset


SONYC-UST-V2 coarse label distribution
Train (1 annotator per label)
Train (2 annotators per label)
Train (3 annotators per label)
Test
Validate
Ratio of recordings in split
0.0
0.1
0.2
0.3
0.4
0.5
0.6
Engine
Human voice
Alert signal
Machinery impact
Powered saw
Non-machinery impact
Music
Dog barking/whining

02. Dataset

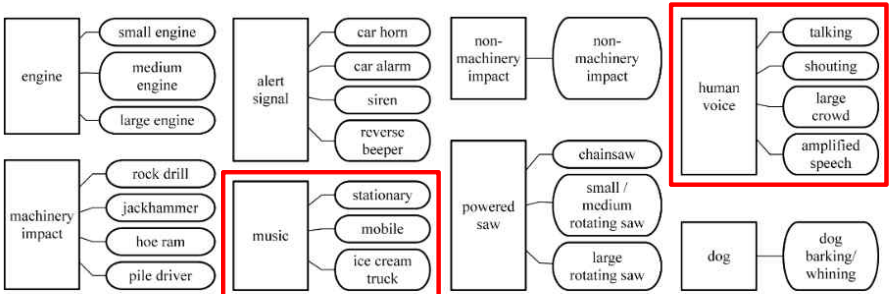
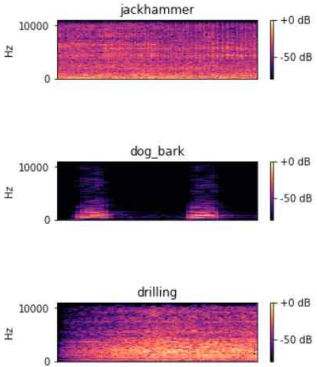


Figure 2. Hierarchical taxonomy of urban sound tags in the DCASE Urban Sound Tagging task. Rectangular and round boxes respectively denote coarse and fine tags.

02. Dataset

human voice	00_007540	00_009223
	00_000914	00_003399

02. Data Preprocessing



- 위험음과 동일한 전처리 시행

SR(Sampling Rate) = 16000Hz

N_MELS = 64

WIN_LEN(Window Length) = 1024

HOP_LEN(Hop Length) = 320

02. Data Preprocessing

UrbanSound8K

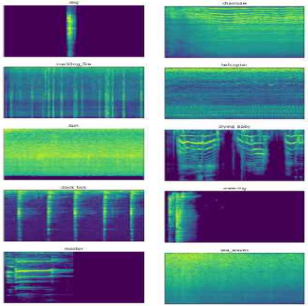
+

SONYC-UST-V2 (human voice)

+

music / another noise

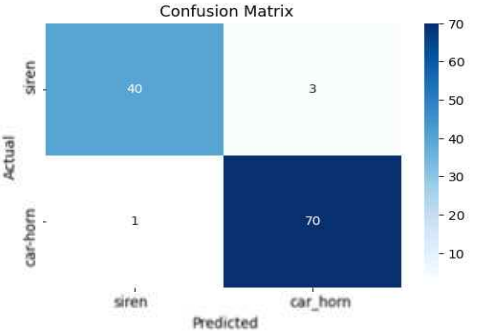
이어폰 착용 환경 흉내



Mel-Spectrogram

03. Experiments

- Original Danger Sound

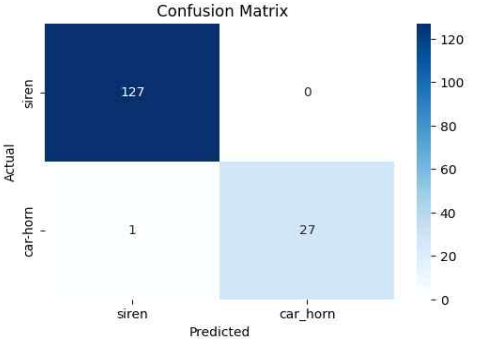


	precision	recall	f1-score
siren	0.99	1.00	1.00
car_horn	1.00	0.96	0.98
accuracy			**0.99**
macro avg	1.00	0.98	0.99
weighted avg	0.99	0.99	0.99

Lightweighted CNN

03. Experiments

- Original Danger Sound

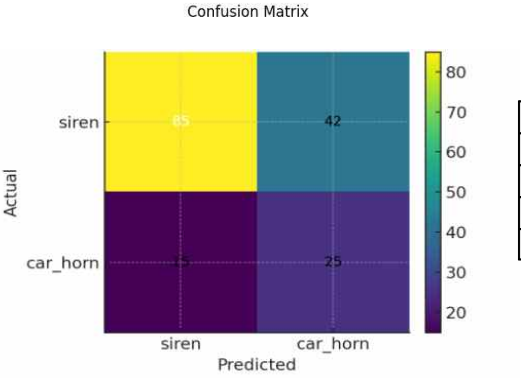


	precision	recall	f1-score
siren	0.99	1.00	1.00
car_horn	1.00	0.96	0.98
accuracy			**0.99**
macro avg	1.00	0.98	0.99
weighted avg	0.99	0.99	0.99

RCNN

03. Experiments

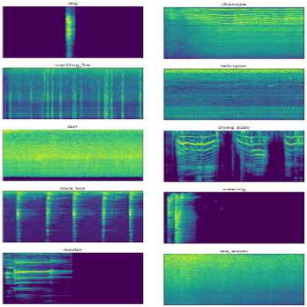
- After Synthesizing Noise Sound



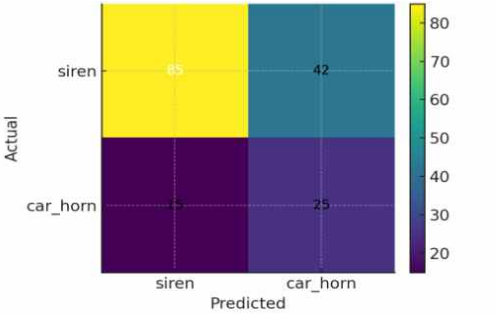
	precision	recall	f1-score
siren	0.85	0.67	0.75
car_horn	0.37	0.62	0.47
accuracy			**0.66**
macro avg	0.61	0.65	0.61
weighted avg	0.61	0.65	0.61

Lightweighted CNN

04. What to do next



① 재전처리 시도 및 Mel-Spectrogram 재추출



② precision, recall, f1-score 성능 향상

감사합니다 .