2025-2 AI 프로젝트IV HCI 시스템 설계

위험소리 감지 알림 UX 서비스 기획

인공지능학과 22619027 지정원

Contents

Problem Recognition


디지털타임스

사회 | 일반

[SNS, 그후] 손에 스마트폰, 귀엔 이어폰 낀 채 `휘청휘청`… 거리의 `시한폭탄`

박상길 기자 구독 + 입력 2024-03-06 14:49 수정 2024-03-06 18:59

경기일보 사회 | 사회일반

‘듣고 싶은 것만 들어요’… 교통사고 부른 노이즈캔슬링

승인 2024-09-30 09:00

한준호 기자 hjh1212@kyeonggi.com
기자페이지 >

무선 이어폰 착용 사고 위험↑… 도내 보행 ‘교통사고’ 주의보
최근 3년간 연평균 8천600건… 시민들 ‘안전의식’ 수반 중요

실제 도로교통공단의 교통사고분석시스템에 따르면 2020년부터 2022년까지 발생한 보행자 교통사고는 10만9877건으로 전체 교통사고의 18%를 차지했으며 이로 인한 사망자는 3044명에 달했다. 특히 고령 보행자가 차지하는 비율은 매년 증가했는데 고령 보행사망자의 전체 보행사망자 대비 비율은 2020년부터 2022년까지 2.3%p(포인트) 늘어났다.

이 같은 사망 사고의 가장 큰 원인은 '노이즈 캔슬링' 기능을 활성화한 무선 이어폰을 사용하기 때문이라는 지적이다. 도로교통공단에 따르면 노이즈 캔슬링 기능이 활성화한 이어폰을 끼고 다니는 개인형 이동장치 사고의 경우 2020년 897건에서 2022년 2386건으로 2.6배 급증했다.

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)

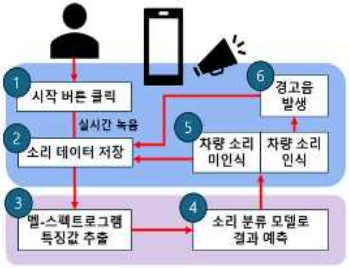
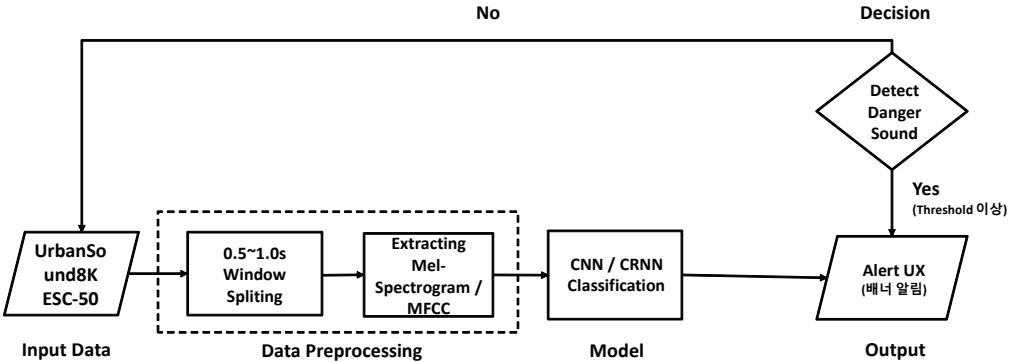


그림 2. 앱 동작 구조

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. FlowChart



01. Dataset

classID:

A numeric identifier of the sound class:

0 = air_conditioner

1 = car_horn

2 = children_playing

3 = dog_bark

4 = drilling

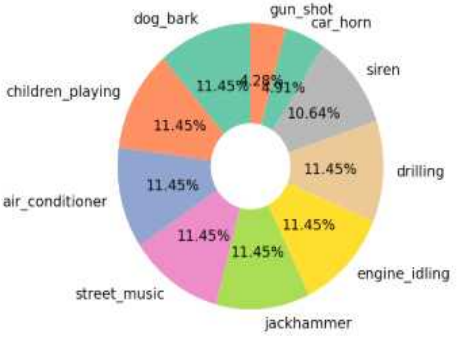
5 = engine_idling

6 = gun_shot

7 = jackhammer

8 = siren

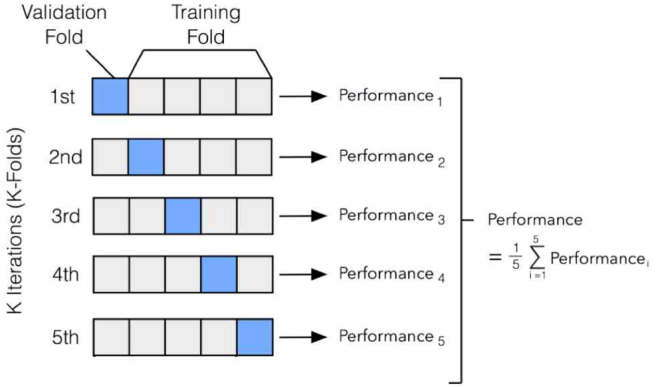
9 = street_music



01. Dataset

Input (7.09 GB)

- Data Sources
  - UrbanSound8K
    - fold1
    - fold10
    - fold2
    - fold3
    - fold4
    - fold5
    - fold6
    - fold7
    - fold8
    - fold9
    - UrbanSound8K.csv

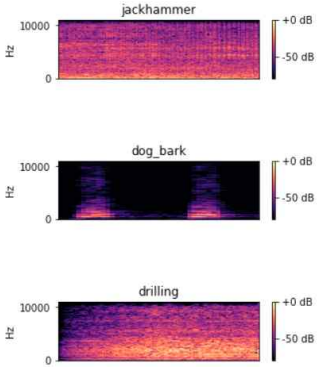


01. Dataset

```
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. Data Preprocessing



SR(Sampling Rate) = 16000Hz

N_MELS = 64

WIN_LEN(Window Length) = 1024

HOP_LEN(Hop Length) = 320

03. Model 1- Lightweighted CNN

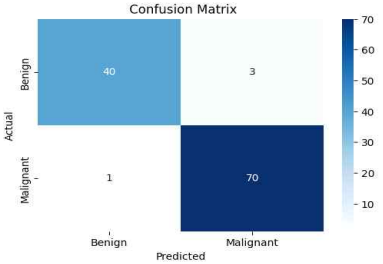
Layer (type)	Output Shape	Param #
input_layer (InputLayer)	(None, 64, 48, 1)	0
conv2d (Conv2D)	(None, 64, 48, 16)	160
batch_normalization (BatchNormalization)	(None, 64, 48, 16)	64
max_pooling2d (MaxPooling2D)	(None, 32, 24, 16)	0
conv2d_1 (Conv2D)	(None, 32, 24, 32)	4,640
batch_normalization_1 (BatchNormalization)	(None, 32, 24, 32)	128
max_pooling2d_1 (MaxPooling2D)	(None, 16, 12, 32)	0
conv2d_2 (Conv2D)	(None, 16, 12, 64)	18,496
batch_normalization_2 (BatchNormalization)	(None, 16, 12, 64)	256
global_average_pooling2d (GlobalAveragePooling2D)	(None, 64)	0
dropout (Dropout)	(None, 64)	0
dense (Dense)	(None, 2)	130

```
Total params: 23,874 (93.26 KB)
Trainable params: 23,650 (92.38 KB)
Non-trainable params: 224 (896.00 B)
```

epochs=30

batch_size=64

early_stopping(patience=5)



	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

03. Model 2- RCNN

Layer (type)	Output Shape	Param #
input_layer (InputLayer)	(None, 64, 48, 1)	0
conv2d (Conv2D)	(None, 64, 48, 16)	144
batch_normalization (BatchNormalization)	(None, 64, 48, 16)	64
activation (Activation)	(None, 64, 48, 16)	0
max_pooling2d (MaxPooling2D)	(None, 32, 24, 16)	0
conv2d_1 (Conv2D)	(None, 32, 24, 32)	4,608
batch_normalization_1 (BatchNormalization)	(None, 32, 24, 32)	128
activation_1 (Activation)	(None, 32, 24, 32)	0
max_pooling2d_1 (MaxPooling2D)	(None, 16, 12, 32)	0
conv2d_2 (Conv2D)	(None, 16, 12, 64)	18,432
batch_normalization_2 (BatchNormalization)	(None, 16, 12, 64)	256
activation_2 (Activation)	(None, 16, 12, 64)	0
permute_time_first (Permute)	(None, 12, 16, 64)	0
reshape_seq (Reshape)	(None, 12, 1024)	0
bidir_grv (Bidirectional)	(None, 128)	413,568
dropout (Dropout)	(None, 128)	0
dense (Dense)	(None, 2)	258

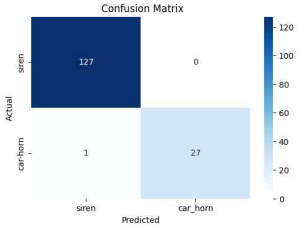
```
Total params: 442,450 (1.69 MB)
Trainable params: 442,226 (1.69 MB)
Non-trainable params: 224 (896.00 B)
```

epochs=30

batch_size=64

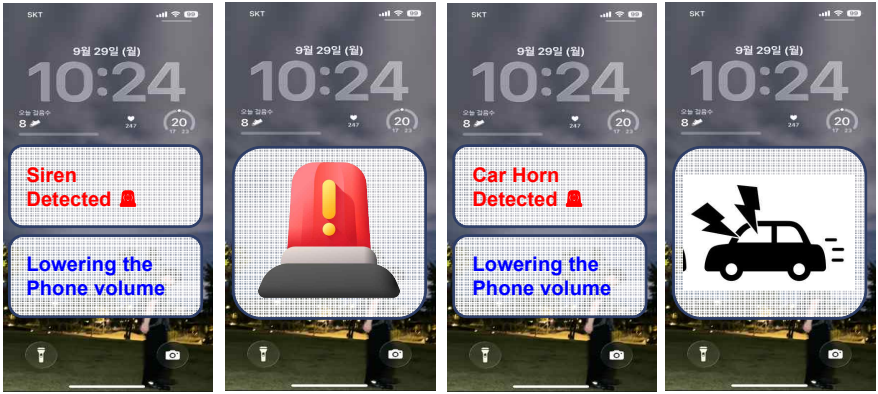
early_stopping(patience=5)

ReduceLROnPlateau
(patience=2, factor=0.5)

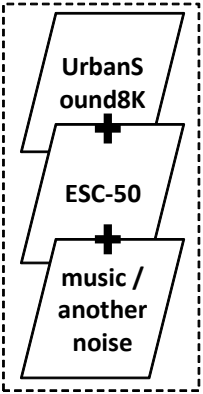


	precision	recall	f1-score
siren	0.99	1.00	1.00
car_horn	1.00	0.96	0.98
accuracy			0.99
macro avg	1.00	0.98	0.99
weighted avg	0.99	0.99	0.99

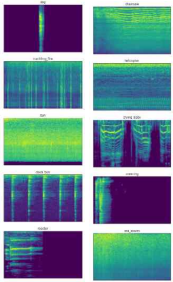
04. Design


SKT
9월 29일 (월)
10:24
오늘 걸음수
8
247
20
Siren
Detected
Lowering the
Phone volume
SKT
9월 29일 (월)
10:24
오늘 걸음수
8
247
20
SKT
9월 29일 (월)
10:24
오늘 걸음수
8
247
20
Car Horn
Detected
Lowering the
Phone volume
SKT
9월 29일 (월)
10:24
오늘 걸음수
8
247
20

05. What to do next



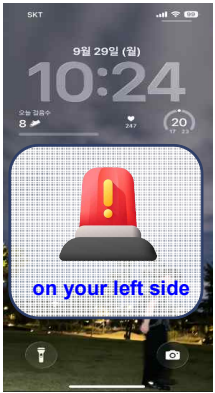
이어폰 착용 환경 흉내



Mel-Spectrogram

05. What to do Next

Detectioning a location of the sound



감사합니다 .