

방문자 알리미

목차

1. <서론> 주제선정 및 가치판단

- 선정한 주제
- 가치
- 범위
- 대상 인터뷰

2. 주제 구현 방법 및 보조자료

- 구현과정 이미지 및 Flow Chart
- AI Hub 화자인식 모델 적용 과정
- 제한된 환경(라즈베리파이5)에서의 화자인식 모델 적용
- 하드웨어 구성
- 음성 인식 모델 및 구현
- 음성 특징 등록
- 음성 인식 및 적외선 센서 감지
- 등록되지 않은 사용자 처리

3. 개발 계획 및 기자재

- 제품 구현 시 필요한 하드웨어 부품
- 제품스케치(설계) 외부/내부
- 팀원간 업무 내용
- 작품 제작 추진 계획 및 일정표
- 지원 경비 사용 계획

1. <서론> 주제선정 및 가치판단

화자인식을 활용한 현관 방문자 알리미

1-1 주제 개요

화자 인식을 통해 현관에 방문한 사람이 낯선 자임을 판별한다.
인지 능력이 성숙하지 못한 아이에게 낯선 자임을 알 수 있도록 주의를 준다.
보호자에게는 집에 낯선 자가 방문했음을 알린다.
(아이가 집에 혼자 있을 경우를 상정함)

1-2 주제를 선택한 이유 및 근거

요즘 현관 cctv 는 집의 보안과 사용자의 안전을 위해 수요가 점점 늘어나고 있습니다. 하지만 현관 cctv 는 법적으로 문제가 될 가능성이 있습니다. 개인정보 보호법 제 25 조에 따르면 '누구든지 다음 각 호의 경우를 제외하고는 공개된 장소에 영상정보처리기를 설치·운영하여서는 아니 된다.'라고 명시되어 있기 때문입니다. 이러한 문제점을 없애기 위해 이 주제를 선정하게 되었습니다

관련기사

[세대 현관 앞 CCTV 설치로 인한 이웃 간 분쟁 < 법률/유권해석/노동 < 전문관리 < 기사본문 - 아파트관리신문 \(aptn.co.kr\)](#)

1-3 문제 해결의 가치 및 범위

가치

-아이에게 현관 앞에 사람을 낯선 자임을 인식시킨다.

아이의 경우 인지 능력이 성숙하지 못하기 때문에 방문자가 친근하게 다가올 시 경계가 풀어질 수 있다. 하지만 장치를 통해 방문자가 위험하다는 사실을 오디오를 통해 주의를 준다면 문을 열어 일어날 수 있는 범죄를 예방할 수 있을 것이다.

-사용자에게 낯선 자가 방문한 기록을 남겨 보안을 강화하고
범죄 발생 시 대응을 용이하게 한다.

구현하려는 주제에는 적외선 센서가 부착되어 현관 앞에 사람이 나타날 시에
사용자에게 연락이 가게끔 되어있다. 이러한 장치를 통해 강도가 집에 침입한 경우
침입한 시간을 특정할 수 있기 때문에 범죄 초기에 대응을 빠르게 할 수 있다.

범위

현관 CCTV의 경우 범죄자의 인상착의를 알 수 있어 범죄를 수사하기에 더 용이할 수
있다. 하지만 앞서 언급했던 것과 같이 법적인 문제가 따를 수 있다는 단점이 존재한다.
그렇기 때문에 이 주제로 구현하려는 제품을 통해 법적인 문제를 회피하고 범죄가
일어난 시간을 특정하여 주변 CCTV와 연계해 범죄를 수사하는 방향으로 초점을 맞추는
것이 좋을 것 같다.

1-4 맞벌이 부부 대상으로 진행한 인터뷰

진행일자 2024.03.19(화)

Q1. 맞벌이 부부로서 가장 걱정되는 부분은 무엇인지

-맞벌이 부부이다 보니 아이를 위해 되도록 집에 일찍 들어오려고 노력을 하지만 직장
생활을 하다 보면 가끔 늦는 경우가 생길 수밖에 없고 먼저 퇴근한 사람이 퇴근하고
바로 집에 오긴 하지만 회사에서 일을 하는 시간 동안에 아이가 혼자 있기 때문에 사고
발생 등 여러 상황에 대한 걱정이 있다.

Q2. 아이를 집에 홀로 둘 때 사고예방 등 목적으로 사용중인 보안 제품이 있는지

-대부분 맞벌이 부부가 설치하는 홈캠을 현재 사용중이다.

Q3. 아이가 혼자 있을 때 가족 외에 집에 자주 방문할 만한 사람이 있는지

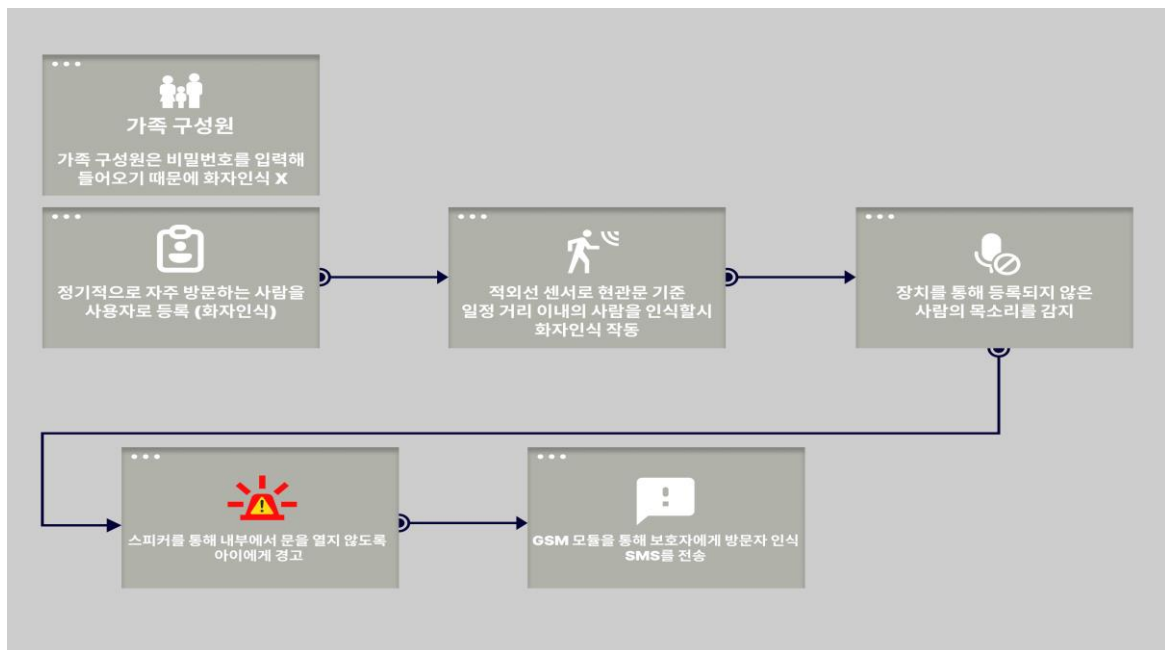
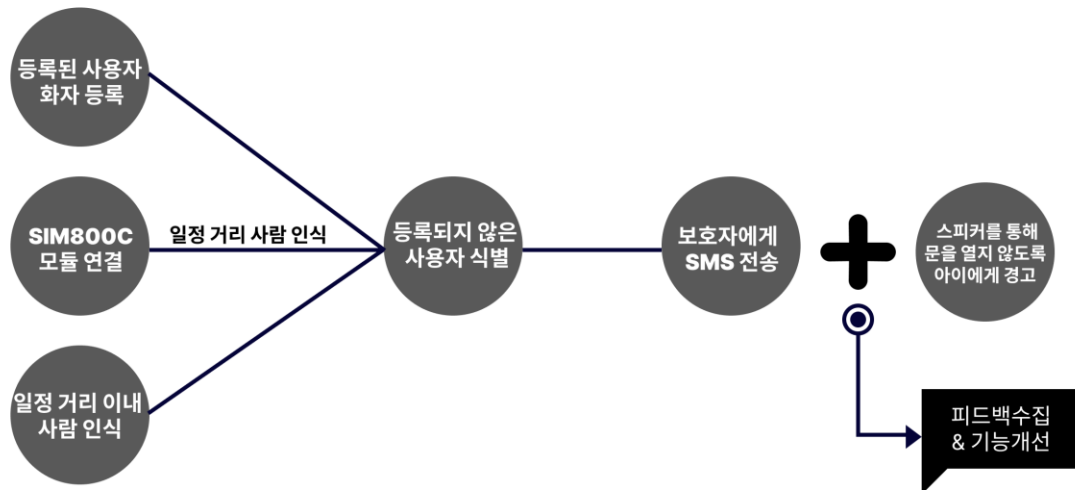
-자주는 아니지만 가끔씩 친정이나 시댁에서 아이를 돌봐주러 와주실 때가 있고,
아파트 유지보수를 해주는 사람들(가스점검,아파트 정기소독)이 벨을 누를 때가 있다.

Q4. 안전을 위해 추가적으로 현관문에 음성인식 장치를 달 의향이 있는지

-복도식 아파트라 같은층에 세대가 많아서 CCTV를 달지 못했다. 집 내부에는 홈캠이
있어 걱정이 덜하지만 집 외부에서 누가 문을 열어달라고 하거나 할 경우가 생길 수
있어 걱정이 되는 부분이 많다. 사전에 아이에게 모르는 사람은 문을 절대 열어주지
말라고 교육을 자주 시키지만 혹시 모를 상황에 대비해 보안 관련 장치가 있다면
추가적으로 설치할 의향이 있다.

2. 주제 구현방법 및 보조자료

<구현 과정에 대한 이미지 및 Flow Chart>



AIHub 화자인식 모델 적용 과정



1. 모델 활용 준비작업

- Install Kaldi toolkit <https://kaldi-asr.org/doc/install.html>
- path.sh 에서 KALDI_ROOT 수정
- 다음 디렉토리를 Kaldi 경로에서 실행할 폴더로 복사
 - \$KALDI_ROOT/egs/voxceleb/v1/local
 - \$KALDI_ROOT/egs/sre08/v1/sid
 - \$KALDI_ROOT/egs/wsj/s5/steps
 - \$KALDI_ROOT/egs/wsj/s5/utls

2. 테스트

1) 두 음성 파일간의 유사도 측정

- score_two_wavs.sh
- e.g., score_two_wav.sh A0003-0000F1012-10120100-00000100.wav C0002-0085F1111-100000_0-00021235.wav plda_score.txt

2) 모델 성능 계산(EER)

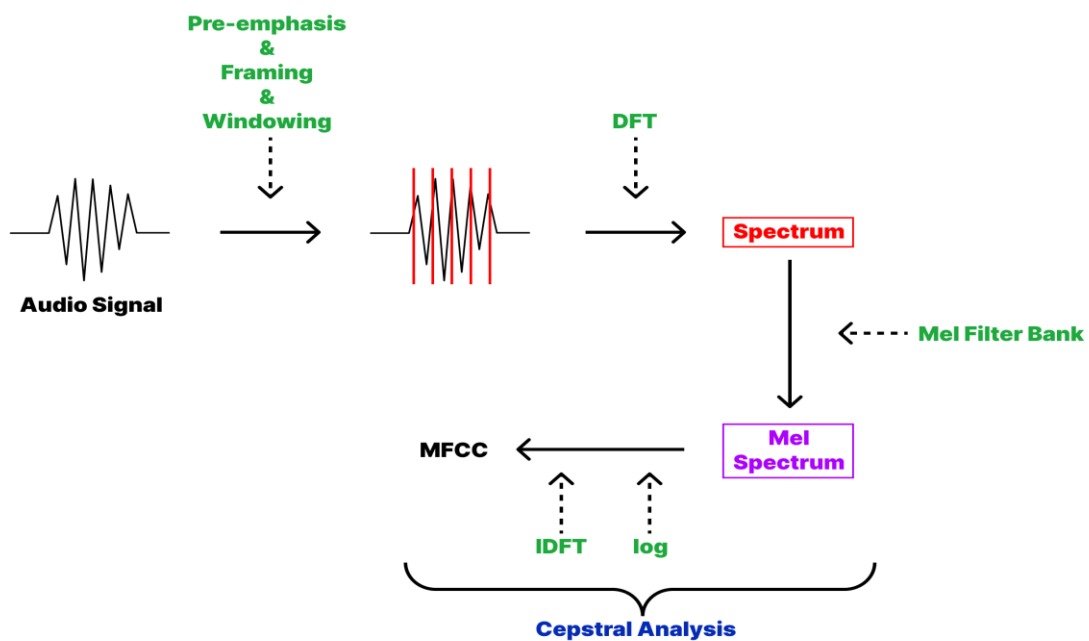
- 테스트 데이터의 wav.scp, utt2spk, trials 를 생성
 - 테스트에 사용할 wav 파일들을 특정 경로에 복사
 - gen_data_dir.sh

- ▣ e.g., `gen_data_dir.sh /home/nia/testdata ./input/testdata1`
- `compute_eer.sh`
 - ▣ e.g., `compute_eer.sh input/testdata1 scores/test1`
- 결과 저장 경로 내 test xvectors, eer, scores 파일 생성

데이터셋	활용 모델	오류 지표(EER)	적합성
AI-Hub (aihub.or.kr)	X-vectors	4.606%	○



파이썬에서 MFCC 를 통한 화자 인식을 구현하는 과정



1. ****데이터 수집****: 먼저, 화자 인식을 위한 음성 데이터를 수집합니다. 이 데이터는 여러 화자의 음성 샘플을 포함해야 하며, 가능한 많은 다양성을 가질수록 좋습니다.
2. ****MFCC 특징 추출****: 수집한 음성 데이터에서 MFCC(Mel-Frequency Cepstral Coefficients)를 추출합니다. 이는 각 화자의 음성을 대표하는 특징 벡터를 생성하는 과정입니다. Librosa 와 같은 라이브러리를 사용하여 MFCC 를 추출할 수 있습니다.
3. ****데이터 전처리****: 추출된 MFCC 특징 벡터를 사용하여 데이터를 전처리합니다. 이 단계에서는 주로 특징 스케일링이나 차원 축소 등의 기술을 사용하여 데이터를 준비합니다.
4. ****학습 데이터 분할****: 전처리된 데이터를 학습 및 테스트 세트로 분할합니다. 이를 위해 데이터를 랜덤하게 섞고, 일반적으로 학습 세트와 테스트 세트를 80:20 또는 70:30 등의 비율로 분할합니다.
5. ****모델 학습****: 분할된 학습 세트를 사용하여 화자 인식 모델을 학습합니다. SVM(Support Vector Machine), k-NN(k-Nearest Neighbors), 신경망 등 다양한 분류 알고리즘을 사용할 수 있습니다.
6. ****모델 평가****: 학습된 모델을 사용하여 테스트 세트에서 화자 인식을 수행하고, 예측 결과를 실제 레이블과 비교하여 모델의 성능을 평가합니다. 이를 통해 정확도, 정밀도, 재현율 등의 성능 지표를 계산할 수 있습니다.
7. ****모델 튜닝****: 모델의 하이퍼파라미터를 조정하고 다양한 실험을 통해 모델을 향상시킵니다. 이를 통해 최적의 화자 인식 모델을 찾을 수 있습니다.

8. ****실시간 화자 인식****: 최종적으로 개발된 모델을 사용하여 실시간으로 오디오 입력에서 화자를 인식할 수 있습니다. 이를 위해 오디오 입력을 실시간으로 읽고 MFCC 를 추출한 후 모델에 입력으로 전달하여 화자를 인식합니다.

구현 예시 : [음성/오디오 신호처리] Python 을 이용한 deep learning - 화자 인식 : 네이버 블로그 (naver.com)

제한된 환경 (라즈베리 파이 5)에서의 화자인식 모델 적용 방법

1. 사용될 음성 모델의 경량화

1.1. TensorFlow Lite 로의 경량화



원본 모델 준비:

AIHub 에서 다운로드한 모델을 사용 :

<https://aihub.or.kr/aihubdata/data/view.do?currMenu=&topMenu=&aihubDataSetSn=537>

TensorFlow Lite Micro 로의 경량화:

TensorFlow 에서 제공하는 tfLite 변환 도구를 사용하여 모델을 TensorFlow Lite 형식으로 변환합니다.

예시:

```
tensorflow_lite_converter --saved_model_dir=model --output_file=model.tflite
```

TensorFlow Lite 모델 통합:

TensorFlow Lite 런타임을 사용하여 모델을 통합하고, 이를 PC 프로그램이나 라즈베리 파이 피코에서 사용

하드웨어 구성:

라즈베리 파이 5 에 적외선 센서와 마이크를 연결합니다. 적외선 센서는 지정된 거리 내의 움직임을 감지하고, 마이크는 음성을 수집합니다. 또한 SIM800C 모듈을 라즈베리 파이와 UART 통신을 통해 연결합니다.

음성 인식 모델 구현:



AIHub 의 화자 인식 모델을 사용하여 음성 인식 기능을 구현합니다.
이 모델을 사용하여 지정자의 음성을 식별합니다.

음성 특징 등록:

지정자의 음성을 AIHub 의 화자 인식 모델을 통해 등록합니다.

음성 인식 및 적외선 센서 감지:

적외선 센서를 사용하여 현관문 근처에서 움직임을 감지합니다.
움직임이 감지되면 마이크를 통해 음성을 수집하고, 화자 인식
모델을 사용하여 음성을 분석합니다.



등록된 사용자 확인:

분석된 음성을 사용하여 AIHub 의 화자 인식 모델을 통해 등록된 사용자인지 확인합니다. 등록된 사용자의 경우 추가 조치를 취하지 않고 프로그램을 종료합니다.

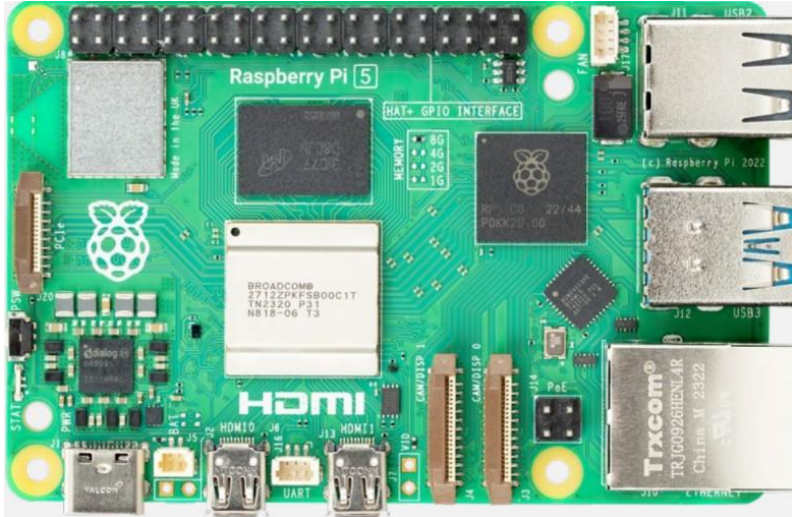
등록되지 않은 사용자 처리:

등록되지 않은 사용자의 경우, 적외선 센서를 사용하여 현관문 근처에서 움직임을 감지한 후, SIM800C 모듈을 사용하여 보호자에게 SMS 를 보냅니다.

SMS 보내기:

SIM800C 모듈을 사용하여 보호자에게 SMS 를 보냅니다. SMS 에는 현관문 근처에서 등록되지 않은 사용자의 음성이 감지되었음을 알리는 메시지가 포함됩니다

1. 하드웨어 구성



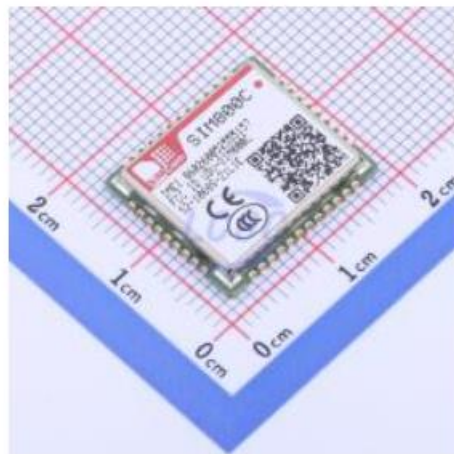
라즈베리 파이 5: 이 프로젝트의 중심이 되는 장치입니다. 음성 및 센서 데이터의 수집, 처리, 제어를 담당합니다.



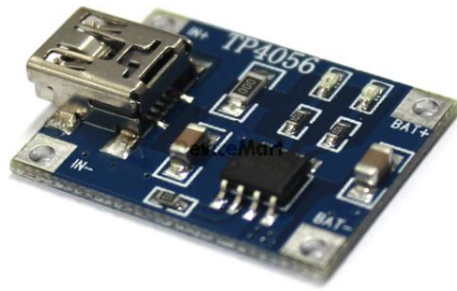
마이크: 음성 데이터를 수집하는데 사용됩니다. 라즈베리 파이에 연결하여 음성 신호를 디지털로 변환합니다.



적외선 센서: 현관문 근처의 움직임을 감지하는데 사용됩니다.
움직임이 감지되면 음성 데이터를 수집하고 처리하는 트리거로
사용됩니다.



SIM800C 모듈: SMS 를 보내는데 사용됩니다. GSM 네트워크를 통해
보호자에게 경고 메시지를 보냅니다.



전원 공급 장치: 라즈베리 파이 및 기타 하드웨어를 전원으로 공급합니다. 필요한 경우 배터리 또는 전원 어댑터를 사용할 수 있습니다.

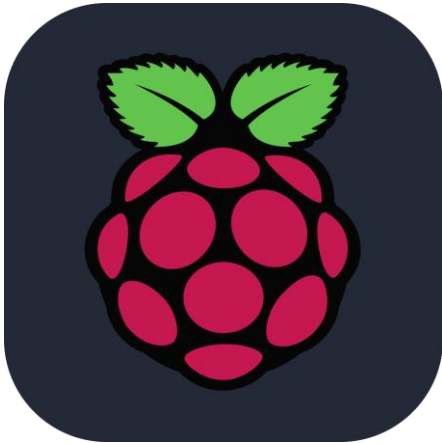
저항, 케이블 및 기타 부품: 회로를 연결하고 구성하는 데 필요한 부품들입니다. 저항은 전류를 제한하고, 케이블은 장치들을 연결하는데 사용됩니다.

외장 케이싱: 모든 하드웨어를 보호하고 외부 환경으로부터 보호하는데 사용됩니다.

2. 음성 인식 모델 구현

모델 다운로드 및 설정:

AIHub 에서 제공하는 화자 인식 모델을 다운로드합니다.
모델의 형식은 TensorFlow, PyTorch 등이 될 수 있습니다.
다운로드 후, 모델을 라즈베리 파이로 복사하고 필요한
라이브러리와 종속성을 설치합니다.



라즈베리 파이 설정:

라즈베리 파이에서 사용할 수 있도록 TensorFlow 또는
PyTorch 등의 라이브러리를 설치합니다. 필요한 경우 가속기를
설정하여 모델의 성능을 향상시킵니다.

모델 로드:

다운로드한 모델을 라즈베리 파이에서 로드합니다. 모델을 메모리에 로드하고 초기화합니다.

음성 데이터 수집 및 전처리:

마이크로부터 음성 데이터를 수집하고, 모델에 입력할 수 있는 형식으로 전처리합니다. 음성 데이터를 적절한 크기로 자르고, 필요한 경우 스펙트로그램 등의 특성을 추출합니다.

모델 추론:

전처리된 음성 데이터를 화자 인식 모델에 입력하여 추론합니다. 모델이 예측한 화자의 ID를 반환합니다.

등록된 사용자 확인:

모델이 예측한 화자의 ID를 등록된 사용자 목록과 비교하여 등록 여부를 확인합니다.

결과 처리:

등록된 사용자의 경우 추가 조치를 취하지 않고, 등록되지 않은 사용자의 경우 SMS 를 보내는 등의 액션을 수행합니다.

3. 음성 특징 등록

지정자의 음성 데이터 수집:

각 지정자의 음성을 수집합니다. 이때, 가능한 한 많은 다양한 음성 샘플을 수집하는 것이 좋습니다. 음성 샘플은 다양한 환경에서 수집되어야 하며, 다양한 발음과 억양을 포함해야 합니다.

음성 데이터 전처리:

수집된 음성 데이터를 전처리합니다. 이 단계에서는 음성 샘플의 크기를 조정하고, 필요에 따라 노이즈를 제거하거나 음성을 정규화할 수 있습니다.

화자 인식 모델 훈련:

전처리된 음성 데이터를 사용하여 화자 인식 모델을 훈련합니다. 훈련된 모델은 각 지정자의 음성 특징을 학습하고 저장합니다.

화자 특징 저장:

훈련된 모델에서 추출된 화자 특징을 저장합니다. 이는 각 지정자의 고유한 음성 특징을 나타냅니다.

등록된 사용자 목록 업데이트:

등록된 사용자 목록에 각 가족 구성원의 정보를 추가합니다. 각 사용자에게 대한 고유한 식별자와 함께 해당 사용자의 화자 특징을 저장합니다.

4. 음성 인식 및 적외선 센서 감지



적외선 센서 감지:

현관문 근처의 적외선 센서가 움직임을 감지합니다. 이 때,
적외선 센서가 움직임을 감지하면 음성 인식 프로세스를
시작합니다.

음성 데이터 수집:

움직임이 감지되면 마이크를 통해 음성 데이터를 수집합니다.
이 음성 데이터는 후속 처리를 위해 저장되거나 실시간으로
처리될 수 있습니다.

음성 데이터 전처리:

수집된 음성 데이터를 전처리합니다. 이 과정에는 음성 신호를
디지털 형식으로 변환하고, 필요한 경우 노이즈를 제거하거나
음성을 정규화하는 등의 작업이 포함될 수 있습니다.

음성 인식 모델 적용:

전처리된 음성 데이터에 대해 화자 인식 모델을 적용하여
등록된 사용자를 식별합니다. 모델은 등록된 사용자와
일치하는지 확인하기 위해 화자의 특징을 분석합니다.

등록된 사용자 확인:

모델이 예측한 화자가 등록된 사용자 목록에 있는지
확인합니다. 등록된 사용자의 경우에는 추가적인 조치를
취하지 않고 프로세스를 종료합니다.

등록되지 않은 사용자 처리:

모델이 예측한 화자가 등록되지 않은 사용자인 경우에는 적외선 센서가 감지한 움직임이 무시되지 않도록 추가 조치를 취해야 합니다. 예를 들어, 등록되지 않은 사용자의 음성이 감지된 경우, 이에 대한 경고 메시지를 보내거나 보호자에게 알릴 수 있습니다.

5. 등록되지 않은 사용자 처리

SMS 전송: 등록되지 않은 사용자가 감지되면 SIM800C 모듈을 사용하여 보호자에게 SMS 를 보냅니다. 이 SMS 에는 현관문 근처에서 등록되지 않은 사용자의 음성이 감지되었음을 알리는 메시지를 포함합니다. 이를 통해 보호자는 즉시 알림을 받아야 합니다.

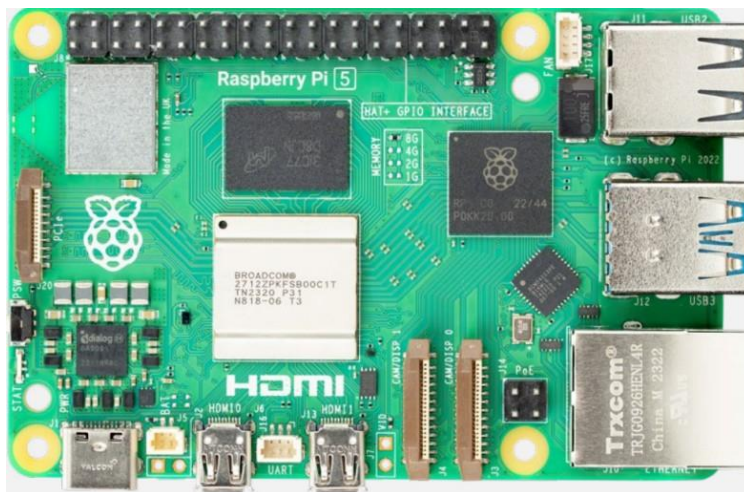
경고음 발생: 등록되지 않은 사용자가 감지되었을 때, 경고음을 발생시켜 현관문 주변에 있는 사람들에게 경고합니다. 이는 가족 구성원에게 경고를 줄 수 있습니다.

보호자에게 전화 알림: SMS 뿐만 아니라 등록되지 않은 사용자가 감지되면 보호자에게 전화 알림을 보내는 기능을 추가할 수도 있습니다. 이렇게 하면 보다 신속한 대응이 가능합니다.

3. 개발 계획 및 기자재

-제품 구현 시 필요한 하드웨어 부품

(1) 라즈베리파이 보드



라즈베리파이 5

148,170 원

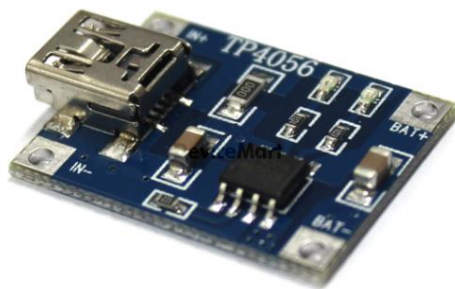
(2) 적외선 센서



[YwRobot] 아두이노 적외선 PIR 인체감지센서 모듈 V2

4,800 원

(3) 배터리 충전 모듈



[OEM] 리튬배터리 1A 충전모듈 (Mini-5P)

900 원

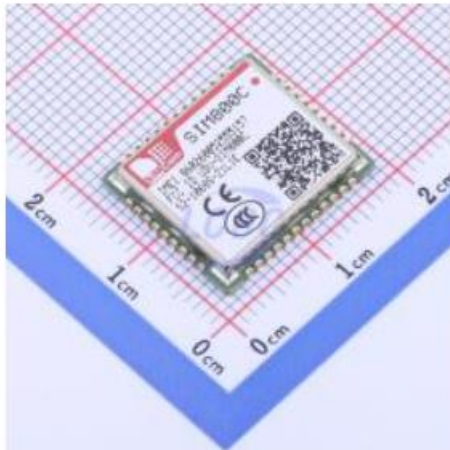
(4) 마이크로폰 센서



[Seeed] ReSpeaker 2-Mics Pi HAT for Raspberry Pi

17,500 원

(5) GSM 통신 모듈

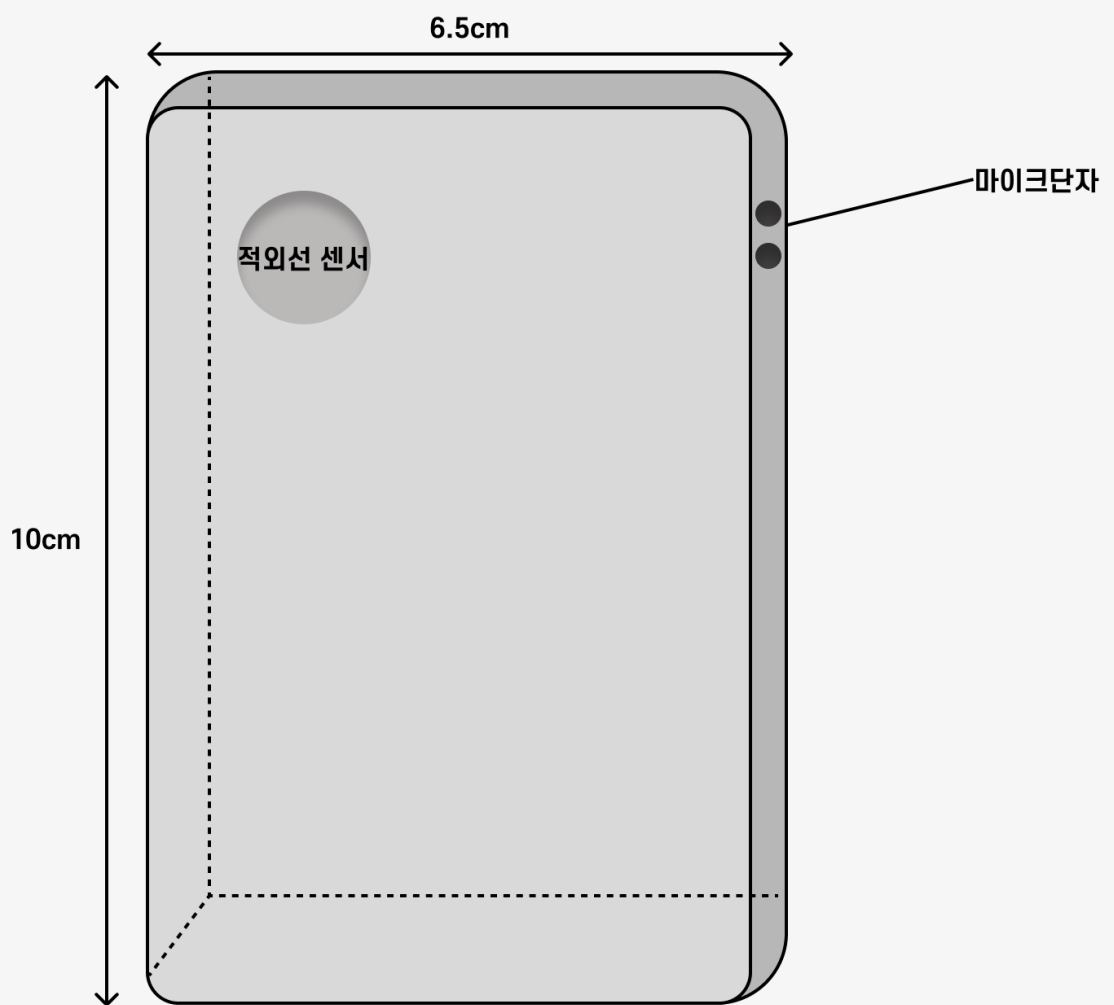


[SIMCom Wireless Solutions] SIM800C

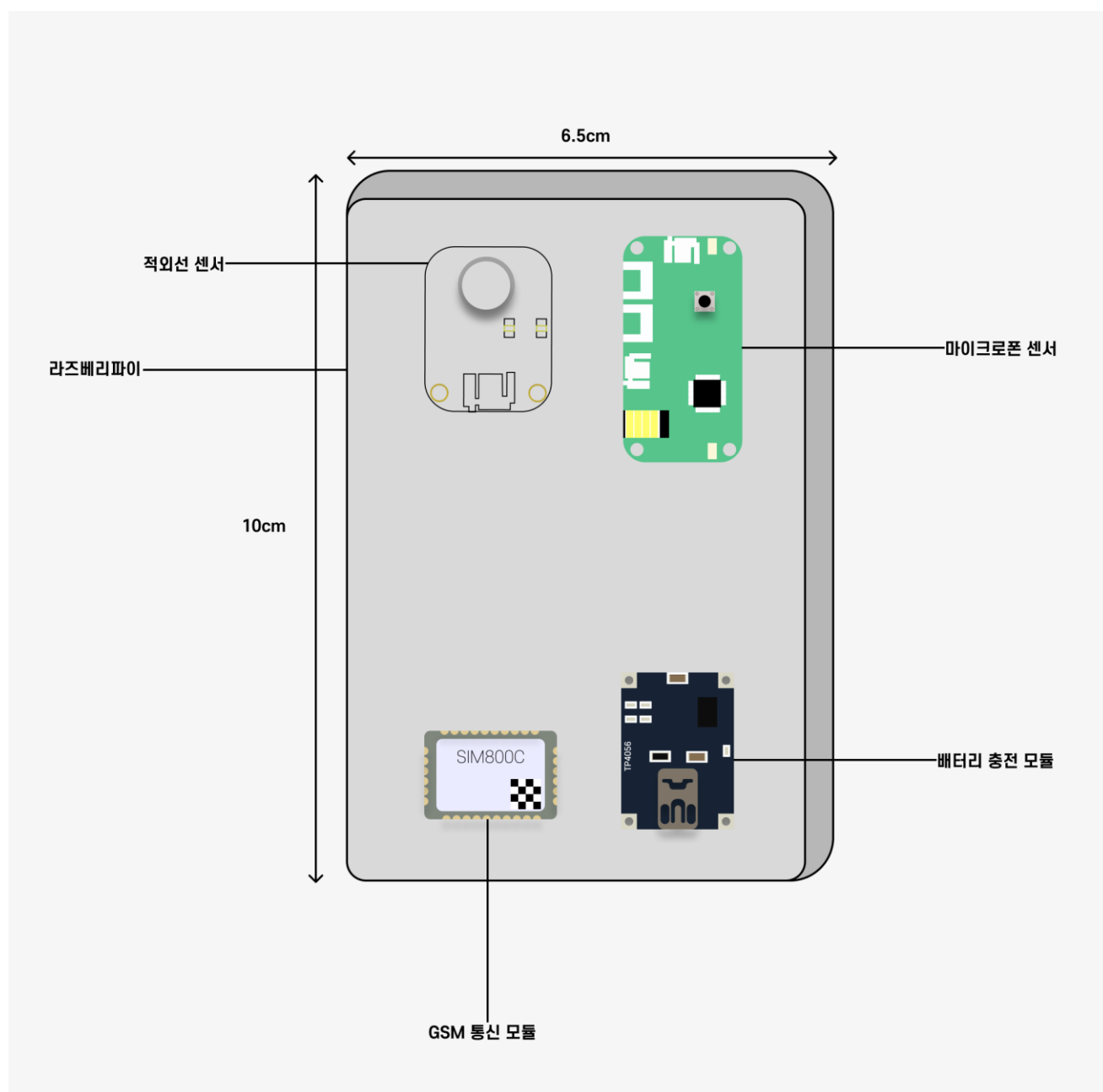
11,740 원

* 제품 스케치(설계)

<외부>



<내부>



-팀원간 업무 내용

길진성

회의록 작성
자재 구입
제품 기획
음성 데이터 수집
음성 데이터 처리 및 분석 소프트웨어 개발

김현욱

제품 기획
음성데이터 수집
SIM800C 모듈을 활용한 통신 기술 구현
회로 구현 및 시나리오 동작 테스트

김태성

제품 기획
음성 데이터 처리 및 분석 소프트웨어 개발
음성데이터 수집
회로 구현 및 시나리오 동작 테스트

정선진

회로 구현 및 시나리오 동작 테스트
SIM800C 모듈을 활용한 통신 기술 구현
음성데이터 수집
제품 기획

- 작업 제작 추진 계획 및 일정

	1 주	2 주	3 주	4 주	5 주	6 주	7 주	8 주
	물품 구입 및 자세한 계획 수립 & 구조 장치의 하드웨어적 설계도(디자인)제작		음성 데이터 처리 및 분석 소프트웨어 개발			음성데이터 수집		중간 고사 기간

	9 주	10 주	11 주	12 주	13 주	14 주	15 주	16 주
	음성 데이터 전처리 & SIM800C 모듈을 활용한 통신 기술 구현				다양한 시나리오에서의 작동 여부 테스트 & 오류 수정			

- 지원 경비 사용 계획

NO.	물품 이름	금액
1	<i>라즈베리파이 5</i>	148,170 원
2	<i>[YwRobot] 아두이노 적외선 PIR 인체감지센서 모듈 V2</i>	4,800 원
3	<i>[OEM] 리튬배터리 1A 충전모듈 (Mini-5P)</i>	900 원
4	<i>[Seeed] ReSpeaker 2-Mics Pi HAT for Raspberry Pi</i>	17,500 원
5	<i>[SIMCom Wireless Solutions] SIM800C</i>	11,740 원

6		22,000
7		
8		
9		
	합계	183,110 원