



고양이 울음소리 기반 모니터링 시스템 “묘(猫)묘안”

오픈소스기초설계 나반 10팀 20212222 최다은, 20213607 손혜정, 20221940 서지민

Aim & Backgrounds : (1) 배경설명과 사례 분석



<고양이 울음소리의 의미와 중요성>

고양이 울음소리 = 보호자와의 핵심 상호작용 수단

수의사가 구분한 고양이 울음소리

- ❑ 어미에게 필요한 것을 요청
- ❑ 하악질(위협을 느낄 시 경고의 울음)
- ❑ 불안감 표현 울음소리
- ❑ 낮시간 동안 소모 못한 에너지 표출

<고양이 모니터링에 이용하는 펫캠 기능 조사>

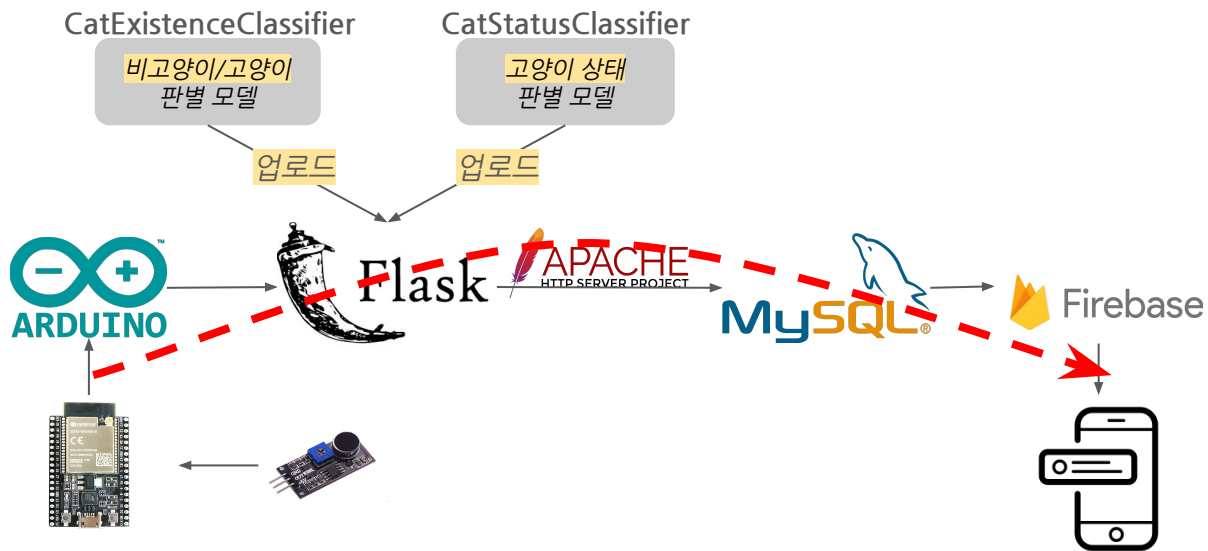
	쿠팡 판매량 1위: Tapo C210	쿠팡 판매량 2위: 샤오미 홈캠 C300	쿠팡 판매량 3위: Tapo C200
모니터링 방식	모션 감지 기능을 통한 자동 모니터링		
소리 활용 여부	양방향 실시간 소리 통신 제공함 하지만 <u>소리 데이터를 모니터링에 활용하지는 않음</u>		

>> 청각 정보는 고양이 모니터링에 중요한 정보 <<
>> 기존 모니터링 수단인 펫캠은 이 정보를 활용하지 않고 있음 <<

Aim & Backgrounds : (2) 펫캠의 문제와 극복방안, 기대효과

문제 정의	극복 방안	기대 효과
카메라 시야에 제한된 모니터링	>> 청각 정보로 시야 외부도 모니터링	>> 효과적인 자동 모니터링 방식으로 펫캠 사용자의 불편함 감소
사용자의 수동 모니터링	>> 고양이 울음소리 로 스트레스 상황 자동 모니터링	>> 고양이 스트레스 상황을 빠르게 파악하고 대처해 고양이의 건강 수준 증진
사생활 유출 가능성	>> 카메라 꺼도 모니터링 가능	
>> 고양이 울음소리로 모니터링이 가능한 시스템을 개발하자 <<		

Contents : (1) 시스템 구조



<시스템 동작 순서와 사용한 오픈소스 프레임워크들>

1. 아두이노: 녹음 데이터 생성

- ESP32 보드와 LM393 센서로 녹음
- **와이파이 통신**을 통해 서버로 녹음한 오디오 데이터 전송

2. Apache + Flask 서버:

녹음 데이터 가공 및 분류

- 녹음한 오디오 전처리와 **MFCC 변환**
- 미리 학습시켜 업로드한 두 모델을 녹음한 오디오에 적용해 **분류**

3. MySQL DB: 데이터 관리

- 서버에서 분류된 데이터의 입출력

4. Firebase: 사용자 알림

- MySQL의 데이터에 접근해 사용자에게 모바일 알림 전달

Contents : (2) 구현 방법

Step 1

: 오디오 인식 구현

① 아날로그 오디오 데이터 수집

ESP보드 + 소리 센서 ->
10초 간격 소리 수집

② 아날로그 데이터 서버 전송

와이파이 통신 +
실시간 스트리밍

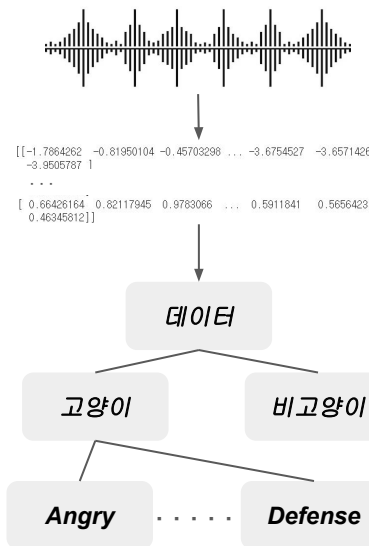
③ MFCC 데이터 저장

- 1) 데이터 wav 변환
- 2) MFCC 추출
- 3) txt파일로 저장

④ MFCC 데이터를 MySQL 데이터베이스에 저장

Step 2

: 모델 학습 구현



① 오디오를 MFCC 특징벡터로 변환

② 전처리 과정 (길이 조정, 스케일링)

③ 고양이/비고양이 판별 CNN 모델 설계

④ 고양이 상태 판별 CNN 모델 설계

Step 3

: 모바일 알림 구현

① MySQL을 FCM에 연결

② MySQL의 데이터와 FCM 연동

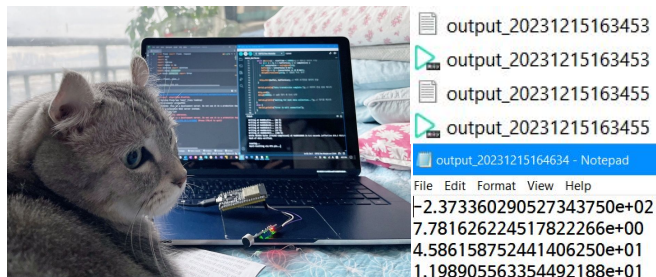
③ 울음소리 발생시 알림 전송

- 안드로이드 스튜디오 이용
- 울음소리별 다른 알림 전송

알림 예시)
00시 00분 고양이가 'Angry'
울음소리를 발생했습니다.

Contents : (3) 주요 결과 및 토의

구현 결과 및 예상 결과



<센서와 아두이노 연결 후 MFCC 데이터 파일 생성>

```
[13]
# 오디오 파일 경로
audio_file_path = DATA_DIR + 'data/cat-sound/sample/LM393_cat.mp3'

# 예측 수행
make_prediction(model, audio_file_path)

1/1 [=====] - 0s 28ms/step
Your cat is in Angry status.
Prediction Confidence: 0.78337276
```

< 고양이 상태 구분 모델 sample test >



< 학습 완료한 고양이/비고양이 모델과
고양이 상태 판별 모델 >



< fcm 모바일 알림 예상 이미지 >

추후 개발 방향과 토의

1. 현재 주요 3단계 모두 로컬에서 작동하도록 구현 완료함
2. 하지만 Apache, Flask 서버와 연결하지 못해 직접 테스트한 적 없음
3. 추후 서버와 각 단계를 연결하는 과정을 진행할 예정
4. CatExistenceModel의 경우 acc가 너무 낮아 비고양이(소음) 데이터와 모델 조정이 필요함

Appendix : (1) 고양이 울음소리 특징

"... meowing is mainly a human-directed vocalization and in principle represents a useful tool for cats to communicate emotional states to their owners."

→ 논문¹⁾에 따르면 야옹 소리는 고양이가 감정 상태를 전달하는 유용한 도구

수의사가 구분한 고양이 울음소리

- ❑ 어미에게 필요한 것을 요청
- ❑ 하악질(위협을 느낄 시 경고의 울음)
- ❑ 발정기 메이팅 콜
- ❑ 채터링(사냥감 발견)
- ❑ 불안감 표현 울음소리
- ❑ 낮시간 동안 소모 못한 에너지 표출

>> 고양이 울음소리는 고양이 상태를 판별하는 중요 정보임<<

Appendix : (2) MFCC 알고리즘의 이해와 적용

❑ MFCC 알고리즘

: 소리의 고유한 주파수 특징을 추출해 수치적으로 나타내는 알고리즘

❑ MFCC 추출 단계

① STFT로 주어진 음성 신호를 작은 프레임 단위로 나누어 주파수 영역 데이터로 변환

② Mel Filter Bank로 인간의 청각 특성을 고려한 척도인 멜 스펙트럼을 데이터에서 추출

③ 데이터를 0~1사이의 값으로 스케일링하고 로그값을 씌운 후 DCT를 수행해 특징 행렬을 압축 표현

④ 해당 프레임의 특징(=MFCC 계수) 추출

❑ MFCC Hyperparameter

- A. 고양이 울음소리가 사람의 말소리와 비슷한 주파수 대역대에서 발생한다고 가정하고 진행함.
- B. 사람의 말소리는 최대 16000Hz로 설정함 (sr = 16000로 설정)

```
frame_size = 0.025 # frame_size : 프레임 길이
frame_step = 0.01 # frame_step : 프레임 간의 거리
n_fft=512 # n_fft: FFT를 수행할 때 사용하는 샘플 수
hop_length=n_fft//3 # hop_length : 프레임 간 겹치는 길이
```

```
for mp3_file in mp3_files:
    try:
        mp3, sr = librosa.load(os.path.join(class_folder_path, mp3_file), sr=16000)
        mfcc = librosa.feature.mfcc(y=mp3, sr=16000, n_mfcc=20, n_mels=40,
                                    win_length=n_fft, hop_length=hop_length, lifter=22)

        mfcc = sklearn.preprocessing.scale(mfcc, axis=1)
        # 전체 데이터 평균이 약 3.83초로 350으로 padding
        padded_mfcc = pad2d(mfcc, 350)
```

< mfcc 특징 벡터 변환 code >

Appendix : (3) 학습 데이터 특징과 분석

A. 학습에 사용한 데이터의 특징

➤ 논문¹⁾의 저자에게 공유받은 데이터로

총 10개의 class를 가지며

개별로 약 600개의 mp3 파일 보유

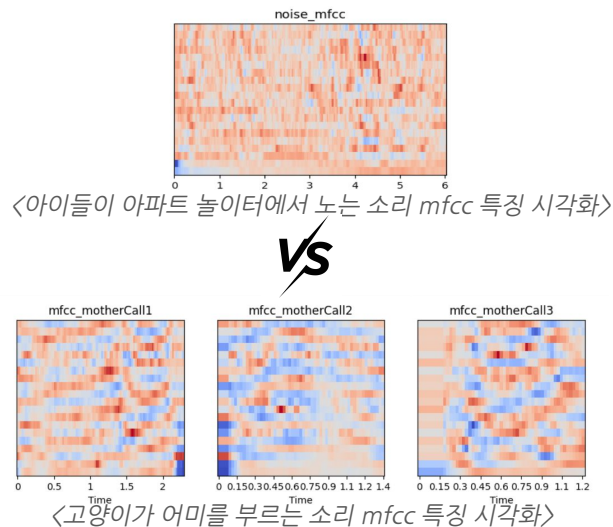
➤ Folder 'Defence' has 582 file(s).
Folder 'Fighting' has 600 file(s).
Folder 'HuntingMind' has 578 file(s).
Folder 'Mating' has 602 file(s).
Folder 'MotherCall' has 592 file(s).
Folder 'Paining' has 582 file(s).
Folder 'Resting' has 592 file(s).
Folder 'Warning' has 600 file(s).
Folder 'Happy' has 595 file(s).
Folder 'Angry' has 600 file(s).
Sum : '5923'

```
label_dict={'Defence':0, 'Fighting':1, 'HuntingMind':2, 'Mating':3,  
            'MotherCall':4, 'Paining':5,  
            'Resting':6, 'Warning':7, 'Happy':8, 'Angry':9}
```

〈클래스별 파일 개수와 클래스 종류〉

B. 학습에 사용한 데이터 분석

➤ mfcc 특징벡터 시각화 결과, 고양이 울음소리와
소음 간의 주파수 패턴 차이 관찰 가능



Appendix : (4) 모델 설계와 학습 결과

* 자세한 코드는 아래 깃허브 참고

https://github.com/thisischeese/opensrc_final-project.git

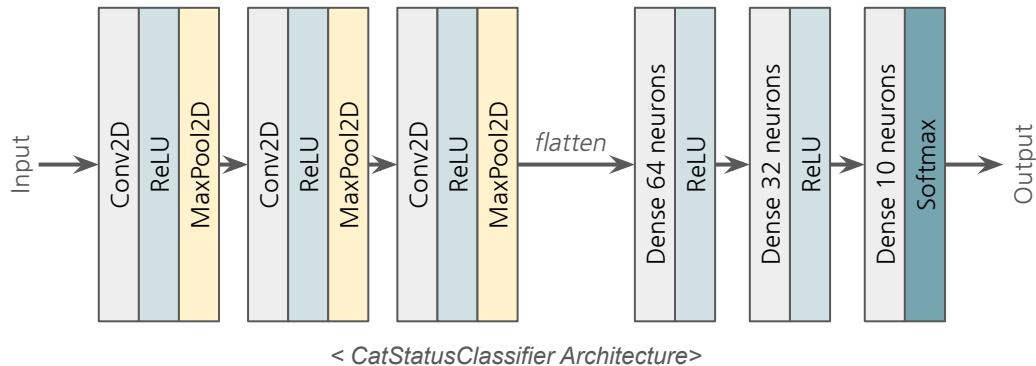
□ 모델 설명

: Conv2D와 MaxPooling2D를 각각 3개씩
쌓고 한개의 Flatten층, 3개의 Dense 층을 쌓은
CNN 모델

Model: "model_1"

Layer (type)	Output Shape	Param #
input_3 (InputLayer)	[(None, 20, 350, 1)]	0
conv2d_8 (Conv2D)	(None, 17, 347, 96)	1632
max_pooling2d_8 (MaxPoolin g2D)	(None, 4, 86, 96)	0
flatten_2 (Flatten)	(None, 33024)	0
dense_6 (Dense)	(None, 64)	2113600
dense_7 (Dense)	(None, 32)	2080
dense_8 (Dense)	(None, 10)	330
=====		
Total params: 2117642 (8.08 MB)		
Trainable params: 2117642 (8.08 MB)		
Non-trainable params: 0 (0.00 Byte)		

< 고양이 상태(10개 class) 분류 모델 : CatStatusClassifier summary >



□ Result :

	CatStatusClassifier	CatExistenseClassifier
val_acc	0.97	0.26(추후 모델 조정 필요)

Reference

1. 논문링크
<https://www.kci.go.kr/kciportal/ci/sereArticleSearch/ciSereArtiView.kci?sereArticleSearchBean.artid=ART002158668>
2. 기사 <https://m.hankookilbo.com/News/Read/A2023101115480001460>
3. Yagya Raj Pandeya, Dongwhoon Kim and Joonwhoan Lee, Domestic Cat Sound Classification Using Learned Features from Deep Neural Nets (<https://www.mdpi.com/2076-3417/8/10/1949>)
4. Yagya Raj Pandeya and Joonwhoan Lee, Domestic Cat Sound Classification Using Transfer Learning (http://www.ijfis.org/journal/download_pdf.php?doi=10.5391/IJFIS.2018.18.2.154)
5. “1) MFCC 개요”, 위키독스, 2023년 4월 9일 수정, 2023년 12월 14일 접속, <https://wikidocs.net/193361>
6. ”MFCC (Mel-Frequency Cepstral Coefficient)”, 네이버 블로그, 2019년 9월 28일 수정, 2023년 12월 3일 접속, <https://m.blog.naver.com/sooftware/221661644808>
7. https://github.com/FilipTirnanic96/mfcc_extraction
8. https://github.com/EwhaSpeakUP/SpeakUP_ML
9. <https://m.hankookilbo.com/News/Read/A2023101115480001460>
10. Prato-Previde E, Cannas S, Palestini C, Ingrassia S, Battini M, Ludovico LA, Ntalampiras S, Presti G, Mattiello S. What's in a Meow? A Study on Human Classification and Interpretation of Domestic Cat Vocalizations. *Animals (Basel)*. 2020 Dec 14;10(12):2390. doi: 10.3390/ani10122390. PMID: 33327613; PMCID: PMC7765146.