

반려묘 먹이 제공 전 발성 분류에서 개체 간 전이 성능과 임계값 개인화 효과 분석

호다예*¹, 김희재*², 정주호*³, 정조운**

Inter-Individual Transfer Performance and Threshold Personalization Effects in Pre-Feeding Vocalization Classification of Domestic Cats

Daye Ho*¹, HeeJae Kim*², Joo Ho Jeong*³, and Jo Woon Chong**

This work was supported by the BK21 FOUR Project. This work was also supported by the National Research Foundation of Korea (NRF), funded by the Ministry of Science and Information and Communication Technology of Korea (grant number 2026OK9045).

요 약

본 연구는 CatMeows 데이터셋을 이용하여 반려묘의 먹이 제공 전 발성과 낯선 환경 격리 발성을 구분하는 이진 분류 문제를 구성하고, 개체 간 전이 성능과 임계값 개인화 효과를 분석하였다. 전처리 후 17마리 고양이의 273개 음원에서 발성 길이, 평균 로그 에너지, 평균 영교차율, 왜도, 첨도의 시간 영역 특징을 추출하였다. 개인화, 전이, 임계값 개인화 기반 전이 조건을 비교한 결과, XGBoost는 임계값 개인화 기반 전이 조건에서 Accuracy 0.881, Food Recall 0.783, PR-AUC 0.934로 가장 높은 평균 성능을 보였다. 특히 전이 조건에서 낮은 Food Recall은 test cat별 임계값 조정 후 증가하였다. 이는 새로운 고양이에 모델을 적용할 때 개체별 임계값 조정의 필요성을 보여준다. 또한 향후 머신러닝 기반 자동 급식기에서 급식 전 발성 신호를 활용한 개체별 상태 판단에 활용될 수 있을 것으로 기대된다.

Abstract

This study used the CatMeows dataset to classify waiting-for-food vocalizations and vocalizations during isolation in an unfamiliar environment in cats, and analyzed inter-individual transfer performance with threshold personalization. Five time-domain features were extracted from 273 audio samples from 17 cats after preprocessing. Among the three evaluation conditions, XGBoost achieved the best average performance under transfer with threshold personalization, with an Accuracy of 0.881, Food Recall of 0.783, and PR-AUC of 0.934. Test cat-specific threshold adjustment improved Food Recall under the transfer condition. These results suggest that individual-specific threshold adjustment is needed when applying the model to new cats and may support machine learning-based automatic feeding systems.

Keywords

cat vocalizations, signal processing, machine learning, feature selection, personalization

* 성균관대학교 전자전기컴퓨터공학과 석사과정
- ORCID¹: <https://orcid.org/0009-0000-3262-6663>
- ORCID²: <https://orcid.org/0009-0005-1536-563X>
- ORCID³: <https://orcid.org/0009-0000-4803-5286>
** 성균관대학교 전자전기공학부 부교수(교신저자)
- ORCID: <https://orcid.org/0000-0002-3161-2742>

• Received: Jan. 16, 2026, Revised: Jun. 29, 2026, Accepted: Jul. 02, 2026
• Corresponding Author: Jo Woon Chong
School of Electronic and Electrical Engineering, Sungkyunkwan University,
Suwon-si, Republic of Korea
Tel.: +82-31-299-4604, Email: jwchong@skku.edu

I. 서 론

반려묘의 비만과 체중 관리는 수의학 및 반려동물 관리 분야에서 지속적으로 다루어져 온 문제이며[1], 이를 위해서는 칼로리 섭취를 조절하는 급식 관리가 필요하다[2]. 또한 급식 방식과 급식 환경은 고양이의 정서 및 행동적 건강과도 관련되므로[3], 급식 관리는 사료 제공량뿐 아니라 섭식 행동 변화와 함께 이해될 필요가 있다. 하지만 보호자 또는 관리자가 반려묘의 급식 상태와 섭식 행동을 연속적으로 관찰하기 위해서는 지속적인 관찰 시간이 요구되며, 시각적 관찰에 의존할 경우 드물게 나타나는 행동을 놓칠 수 있다[4]. 이에 따라 자동 급식 시스템은 정해진 급식량과 급식 빈도에 따라 사료를 제공하거나, 특정 개체의 먹이 접근을 제한하는 방식으로 활용되어왔다[5].

다만 기존 자동 급식 시스템은 주로 급식량, 급식 빈도, 먹이 접근 여부와 같은 급식 관리 정보를 다루기 때문에[5], 발성이나 먹이 요구 행동과 같이 급식 상황에서 나타나는 요구 신호를 충분히 반영하는 데 한계가 있다. 이러한 먹이 요구 행동은 발성 신호로도 나타날 수 있으며, 실제로 먹이 요구 상황에서 고양이가 특정 음향 신호를 사용할 수 있음이 보고된 바 있다[6]. 이러한 관점에서 급식 관련 상황을 포함한 고양이 발성 데이터셋을 이용한 자동 분류 연구가 수행되어왔다[7]-[10]. Ntalampiras 등은 먹이 제공 전 대기 상황(Waiting for food), 낯선 환경에서의 격리 상황(Isolation in an unfamiliar environment), 빗질 상황(Brushing)에서 수집된 고양이 발성을 자동으로 분류하였다[7]. 이후 해당 발성 데이터는 CatMeows 데이터셋으로 공개되었고[8], 발성 상황 분류와 성별, 품종 등 개체 특성 분류 연구에 활용되었다[7]-[10].

기존 CatMeows 기반 분류 연구는 주로 전체 데이터셋 수준에서 발성 상황 또는 개체 특성이 분류 가능한지를 확인하는 데 초점을 두었다[7]-[10]. 이러한 접근은 데이터셋 전체에서의 분류 가능성을 확인하는 데에는 유용하지만, 학습에 포함되지 않은 새로운 고양이에 대한 성능을 직접 보여주지는 못한다. 또한 동물 발성에는 개체별 음향 특성이 포함

될 수 있으므로[11], 새로운 고양이에 대한 성능은 학습 개체와 평가 개체를 분리한 조건에서 검증할 필요가 있다.

본 연구는 CatMeows 데이터셋의 먹이 제공 전 대기 상황과 낯선 환경에서의 격리 상황에서 수집된 발성을 구분하는 머신러닝 기반 이진 분류 과제를 구성하고, 동일 고양이 내부 개인화 조건, 개체 간 전이 조건, 임계값 개인화 기반 전이 조건에서의 분류 성능을 비교하였다[7][8]. 또한 실제 반려묘 모니터링 환경에서의 적용 가능성을 고려하여 입력 특징을 구성하였다. 기존 CatMeows 기반 연구에서는 MFCC(Mel-Frequency Cepstral Coefficients)와 시간 변조 특징(Temporal modulation features)을 사용하였으며[7], 이들 특징과 시간 영역 특징은 모두 프레임 단위 신호 처리를 기반으로 한다. 다만 MFCC와 시간 변조 특징은 Fourier transform 계열의 추가 변환 절차를 포함하는 반면[12][13], 시간 영역 특징은 원시 시계열 또는 프레임 단위 신호에서 직접 계산할 수 있고[14], 상대적으로 낮은 계산 복잡도를 갖는 특징군으로 활용될 수 있다[15]. 이에 본 연구에서는 발성 길이(Vocalization length)[16], 로그 에너지(Log-energy)[17], 영교차율(ZCR, Zero-crossing rate), 왜도(Skewness), 첨도(Kurtosis)의 5개 특징을 시간 영역 특징으로 사용하였다[18]. 이후 동일 고양이 내부 평가를 통해 개체 내 발성 구분 가능성을 확인하고, 여러 고양이로 학습한 모델을 학습에 포함되지 않은 고양이에 적용하여 개체 간 전이 성능을 평가하였다. 마지막으로 소량의 대상 고양이 데이터를 이용한 임계값 조정을 통해 임계값 개인화가 분류 성능에 미치는 영향을 분석하였다.

II. 관련 연구

2.1 반려묘 자동 모니터링 연구

반려묘의 상태를 자동으로 파악하기 위한 연구에서는 웨어러블 센서와 영상 분석 기반 접근이 활용되어 왔다. Smit et al.[4]은 반려묘에 3축 가속도 센서를 부착한 뒤, 수집된 센서 데이터에 기계 학습 알고리즘을 적용하여 행동 상태를 분류하였

다. 같은 연구진은 이후 실제 가정 환경에서 얻은 가속도 데이터를 이용해 활동, 섭식, 그루밍, 배변, 휴식 등의 일상 행동을 정량화하였다[19]. 이들 연구는 보호자나 관리자의 장시간 관찰 없이도 센서 데이터를 통해 반려묘의 행동 변화를 추적할 수 있음을 보여준다.

영상 기반 방법도 반려묘 행동 모니터링에 사용되어왔다. Chen et al.[20]은 Raspberry Pi와 YOLOv3-Tiny 기반 객체 탐지 모델을 이용하여 가정 내 반려묘 행동을 실시간으로 감지하는 시스템을 제안하였다. 카메라 기반 방법은 반려묘의 위치, 자세, 움직임을 직접 확인할 수 있다는 장점이 있으며, 웨어러블 센서와 함께 반려묘 행동 정보를 자동으로 기록하는 방법으로 활용될 수 있다.

다만 이러한 연구들은 주로 움직임, 자세, 활동량과 같이 외부에서 관찰 가능한 행동 지표를 다룬다. 반면 발성은 비접촉 방식으로 수집할 수 있는 신호이므로, 반려묘 모니터링에서 기존 센서 및 영상 기반 방법을 보완하는 입력으로 활용될 수 있다.

2.2 고양이 발성 분류

고양이 발성에 관한 초기 연구에서는 발성의 음성학적 특성과 기능적 의미를 분석하려는 시도가 이루어졌다. Moelk[21]는 고양이 발성을 발성 방식에 따라 여러 유형으로 구분하였고, Schötz et al.[22]은 집 고양이 울음소리(Meow)의 f_0 , duration, f_0 contour가 발성 상황과 정서 상태에 따라 달라질 수 있음을 보고하였다. 또한 인간 청취자를 대상으로 한 연구에서도 고양이 발성에 상황 정보가 포함될 가능성이 검토되었다[23][24]. Nicastro와 Owren[23]은 인간 청취자가 고양이 발성 상황을 무작위 선택보다 높은 정확도로 분류할 수 있음을 보고하였고, Prato-Previde et al.[24]도 인간이 일부 고양이 발성의 맥락을 구분할 수 있음을 보였다. 이러한 결과는 고양이 발성에 상황 정보를 반영하는 음향적 단서가 포함될 가능성을 보여준다.

이후 기계학습 기반 고양이 발성 상황 분류 연구도 수행되었다. Ntalampiras et al.[7]은 Maine Coon과 E-european Shorthair 고양이의 발성을 먹이 제공 전

대기 상황, 낮선 환경에서의 격리 상황, 빗질 상황의 세 가지 상황에서 수집하고, MFCC와 temporal modulation features 및 HMM(Hidden Markov Model) 기반 분류 구조를 이용하여 발성 상황을 자동 분류하였다. 실험 결과, DAG-HMM(Directed Acyclic Graph Hidd-en Markov Model) 모델은 세 가지 발성 상황 분류에서 95.94%의 인식률을 보였다[7]. 이는 CatMeows 데이터셋 전체에서 발성 상황이 높은 정확도로 분류될 수 있음을 보여준다. 이후 해당 데이터는 CatMeows 데이터셋으로 공개되어[8], 고양이 성별 및 품종 분류와 같은 개체 특성 분류 연구에도 활용되었다[9][10].

그러나 기존 CatMeows 기반 연구는 주로 전체 데이터셋 수준에서 발성 상황 또는 개체 특성이 분류 가능한지를 확인하는 데 초점을 두었다[7]-[10]. 반면 학습에 포함되지 않은 새로운 고양이에 대한 개체 간 전이 성능과, 새로운 개체의 발성 분포 차이를 반영하기 위한 분류 기준 조정 효과는 상대적으로 충분히 분석되지 않았다[7][11][25]. 또한 기존 연구에서 사용된 MFCC와 시간 변조 특징은 프레임 단위의 주파수 분석 및 변환 절차를 포함하므로, 실제 센서 기반 모니터링 환경에서는 특징 추출 과정의 계산 부담을 함께 고려할 필요가 있다[12][13]. 이에 본 연구는 먹이 제공 전 대기 상황 발성의 자동 식별을 위한 이진 분류 과제를 구성하고, FFT 기반 주파수 분석을 수행하지 않는 시간 영역 특징을 사용하여 동일 고양이 내부 개인화 성능, 새로운 고양이에 대한 개체 간 전이 성능, 임계값 기반 개인화 효과를 함께 평가하였다.

III. 방법론

3.1 데이터셋 및 특징 추출

3.1.1 데이터셋 및 이진 분류 과제

본 연구에서는 공개 고양이 발성 데이터셋인 CatMeows를 사용하였다[8]. CatMeows 데이터셋은 Maine Coon과 European Shorthair 두 품종에 속하는 총 21마리의 고양이로부터 수집된 440개의 wav 파

일로 구성되어 있으며, 상황별로는 빗질 상황 127개, 낮선 환경에서의 격리 상황 221개, 먹이 제공 전 대기 상황 92개의 음원 파일을 포함한다. 각 음원 파일은 단일 발성을 포함하며, 평균 길이는 약 1.83초이고, 발성 시작과 종료 구간에 각각 약 0.5초의 무음 구간을 포함하고 있다[8].

세 가지 발성 상황 중 빗질 상황은 보호자가 고양이를 직접 빗질하는 과정에서 수집된 발성으로, 보호자와의 지속적인 신체 접촉이 포함된 상호작용 상황에 해당한다. 반면 먹이 제공 전 대기 상황과 낮선 환경에서의 격리 상황은 각각 음식 제공 지연 상황과 낮선 환경에서의 고립 상황에서 수집된 발성으로, 보호자 또는 관리자의 대응이 필요한 상태와 관련될 수 있다. 이에 본 연구에서는 빗질 상황을 제외하고, 먹이 제공 전 대기 상황과 낮선 환경에서의 격리 상황에서의 발성을 각각 food와 isolation으로 표기하여 2-클래스 분류 과제를 구성하였다. 이때 food 클래스는 음식 제공이 지연되는 상황에서의 발성을, isolation 클래스는 낮선 환경에서 혼자 남겨진 상황에서의 발성을 의미한다.

3.1.2 음원 전처리

본 연구에서는 무음 구간이 음향 특징 추출 결과에 미치는 영향을 줄이기 위해 음원의 시작 및 종료 부분의 0.5초 무음 구간을 제거하였다. 최대 절대 진폭 $A_{\max}$ 는 다음과 같이 정의하였다.

$$A_{\max} = \max_{1 \leq n \leq N} |x[n]| \quad (1)$$

식 (1)에서 $x[n]$ 는 무음 구간 제거 후 남은 파형의 n 번째 샘플값을, N 은 해당 파형의 전체 샘플 수를 의미한다. 선행연구의 저진폭 신호 제외 절차를 참고하여[26], $A_{\max} \leq 0.05$ 인 음원을 유효한 고양이 발성 신호가 충분하지 않은 음원으로 판단하여 분석 대상에서 제외하였다. 최종적으로 17마리 고양이의 273개 음원이 특징 추출에 사용되었으며, 이 중 food 클래스는 63개, isolation 클래스는 210개였다.

3.1.3 특징 추출

전처리된 파형으로부터 시간 영역 특징을 추출하였다. 시간 영역 특징은 음원 신호의 시간적 길이와 파형 진폭 분포의 특성을 반영하기 위해 추출하였다[18][27][28]. 본 연구에서는 발성 길이(Vocalization length)[16], 로그 에너지(Log-energy)[17], 영교차율(ZCR), 왜도(Skewness), 첨도(Kurtosis)[18]의 5개 특징을 시간 영역 특징으로 사용하였다.

Vocalization length는 발성 구간의 시간적 길이를 나타내는 특징으로 사용된다[16]. Vocalization length는 전처리 후 남은 유효 발성 구간의 시간적 길이로 정의하였다. 전처리된 파형 $x[n]$ 의 전체 샘플 수를 N , 샘플링 주파수를 f_s 라 할 때, vocalization length L_{voc} 는 $L_{voc} = N/f_s$ 로 계산하였다.

Log-energy와 ZCR는 전처리된 파형을 일정 길이의 프레임으로 분할한 뒤, 프레임 단위로 계산하였다. 이후 각 음원에 대해 프레임별 값의 평균을 최종 특징으로 사용하였다. 이때, m 번째 프레임의 i 번째 샘플값을 $x_m[i]$, m 번째 프레임에 포함된 샘플 수를 N_m , 한 음원에서 분할된 전체 프레임 수를 M 이라 하였다.

Log-energy는 프레임 단위로 파형 샘플값의 제곱합에 로그를 취해 계산한 시간 영역 에너지 특징이다[17]. m 번째 프레임의 log-energy E_m 은 식 (2)와 같이 계산하였다.

$$E_m = \log \left(\sum_{i=1}^{N_m} x_m[i]^2 \right) \quad (2)$$

이후 한 음원에 포함된 전체 M 개 프레임에 대해 E_m 의 평균을 계산하여 해당 음원의 mean log-energy로 사용하였다.

ZCR은 파형이 시간 영역에서 0을 기준으로 부호가 변화하는 빈도를 나타내는 특징으로, 신호의 진동 특성 및 고주파 성분의 상대적 변화를 반영한다. m 번째 프레임의 ZCR을 Z_m 이라 하면, Z_m 은 식 (3)과 같이 나타낼 수 있다.

$$Z_m = \frac{1}{2(N_m - 1)} \sum_{i=2}^{N_m} |sgn x_m[i] - sgn(x_m[i-1])| \quad (3)$$

여기서 $sgn(y)$ 는 y 의 부호를 나타내는 함수로, $y \geq 0$ 인 경우 1, $y < 0$ 인 경우 -1이다. 전체 M 개 프레임의 ZCR을 평균하여 해당 음원의 mean zero-crossing rate을 최종 특징으로 사용하였다.

Skewness와 kurtosis는 파형 진폭 분포의 형태를 나타내기 위해 사용하였다. Skewness는 진폭 분포의 비대칭성을 나타내며, kurtosis는 진폭 분포의 첨도를 나타낸다. 본 연구에서 kurtosis로 표기한 특징은 SciPy 라이브러리의 기본 설정인 Fisher 정의에 따라 계산된 excess kurtosis를 의미한다[29]. 일반적인 kurtosis는 정규분포에서 3의 값을 가지며, excess kurtosis는 일반적인 kurtosis에서 3을 뺀 값으로 정규분포에서 0의 값을 가진다[29]. 이에 따라 본 연구에서는 정규분포를 중립 기준으로 두고 파형 진폭 분포의 첨도 차이를 해석하기 위해 excess kurtosis를 사용하였다.

전처리된 전체 음원 파형의 n 번째 샘플값을 $x[n]$, 파형 샘플값의 평균을 $\bar{x}$, 전체 샘플 수를 N , γ 차 중심 모멘트를 μ_γ 이라 할 때, $\bar{x}$ 와 μ_γ 각각 식 (4)와 식 (5)와 같이 정의하였다.

$$\bar{x} = \frac{1}{N} \sum_{n=1}^N x[n] \quad (4)$$

$$\mu_\gamma = \frac{1}{N} \sum_{n=1}^N (x[n] - \bar{x})^\gamma \quad (5)$$

이를 이용하여 skewness S 와 excess kurtosis K 를 각각 식 (6)과 식 (7)과 같이 계산하였다.

$$S = \frac{\mu_3}{\mu_2^{3/2}} \quad (6)$$

$$K = \frac{\mu_4}{\mu_2^2} - 3 \quad (7)$$

3.2 분류 성능 평가 조건

3.2.1 평가 조건 개요

고양이 발성에는 개체별 음향 특성이 포함될 수 있으므로, food와 isolation을 구분하는 음향적 단서가 모든 고양이에서 동일하게 나타난다고 가정하기 어렵다[11]. 이에 본 연구에서는 먼저 동일 고양이 내부에서 food와 isolation 발성을 구분할 수 있는지를 확인하는 개인화 조건을 설정하였다. 기계학습에서 개인화는 특정 개인을 대상으로 성능을 최적화하는 접근으로 설명되며[30], 본 연구에서는 고양이별 특징 부분집합을 사용하는 방식으로 개인화 조건을 구성하였다.

또한 음성 기반 분류 연구에서는 평가 대상 화자가 학습 데이터에 포함되는 경우와 포함되지 않는 경우를 구분하여 성능을 비교한 바 있다[31]. 이는 동일 발성 주체 내부에서의 성능과 새로운 발성 주체에 대한 성능이 서로 다를 수 있기 때문이다. 본 연구에서는 이러한 평가 관점을 고양이 개체 단위로 적용하여, 학습 고양이와 평가 고양이를 분리한 전이 조건을 설정하였다.

마지막으로, 전이 모델을 새로운 고양이에 적용할 때 동일한 판정 임계값이 모든 고양이에게 적합하지 않을 수 있다. 이진 분류에서는 예측 확률을 클래스 라벨로 변환하기 위한 임계값이 필요하며, 해당 임계값은 특정 성능 기준에 맞게 조정될 수 있다[32]. 이에 본 연구에서는 test cat의 enrollment set을 이용해 임계값을 조정된 뒤 evaluation set에서 성능을 평가하는 임계값 개인화 기반 전이 조건을 설정하였다.

본 연구의 세 가지 조건은 개인화 조건(A), 전이 조건(B), 임계값 개인화 기반 전이 조건(C)으로 정의하였다. 각 조건의 학습 데이터, 개인화 방식, 평가데이터는 표 1에 정리하였다. 앞서 설명한 음원 전처리 이후 17마리 고양이의 273개 음원이 최종 특징 추출 대상에 포함되었다. 이 중 food와 isolation 클래스 모두에서 최소 4개 이상의 음원이 남아 있어 개체별 평가와 enrollment/evaluation 분할이 가능한 6마리를 test cat 후보로 선정하였다.

표 1. 실험 조건별 학습 및 평가 설정

Table 1. Training and evaluation settings for each experimental condition

Condition	Training data	Personalization method	Evaluation data
A. Personalization	Train folds within the same cat	Cat-specific feature selection	Test folds from the same cat
B. Transfer	Vocalization samples from the 16 cats excluding the test cat	Not applied	Vocalization samples from the excluded test cat
C. Transfer with threshold personalization	Vocalization samples from the 16 cats excluding the test cat	Test cat-specific threshold adjustment	Evaluation set of the test cat

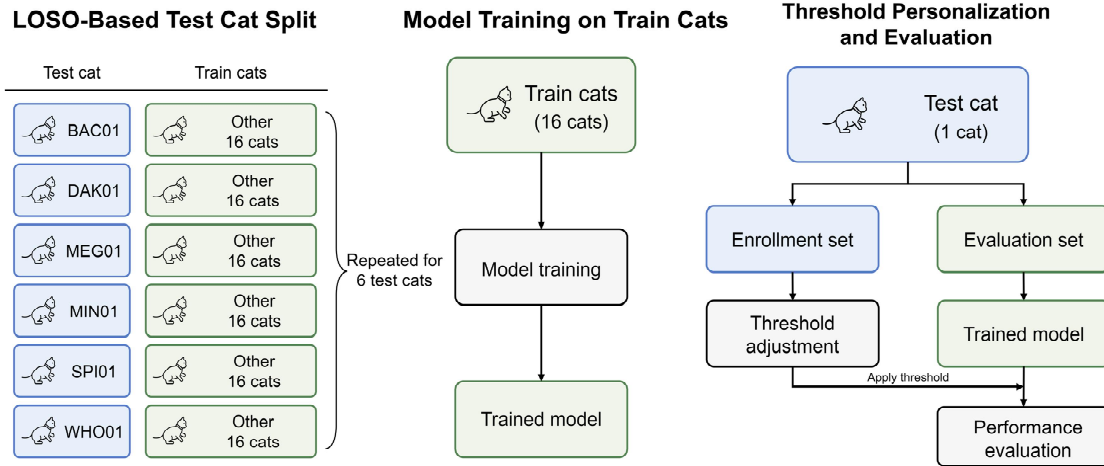


그림 1. 임계값 개인화 기반 전이 조건(C 조건)의 평가 구조

Fig. 1. Evaluation framework for transfer with threshold personalization

전이 조건과 임계값 개인화 기반 조건에서는 개체 단위 평가를 LOSO(Leave-One-Subject-Out) 방식으로 수행하였다. 즉, 6마리의 test cat 후보 중 한 마리를 평가 대상 고양이로 지정하고, 해당 고양이를 제외한 16마리의 음원 파일로 모델을 학습하였다. 이 과정을 6마리의 test cat 후보에 대해 각각 반복하여 학습에 포함되지 않은 새로운 고양이에 대한 성능을 평가하였다. 이 중 임계값 개인화 기반 전이 조건(C)의 평가 구조는 그림 1과 같다.

데이터 분할 및 모델 학습 과정에서 발생할 수 있는 난수성의 영향을 줄이고 결과의 안정성을 확인하기 위해 각 실험은 seed 0부터 9까지 총 10회 반복하였다. A 조건에서는 동일 고양이 내부 음원 파일의 cross-validation 분할을 seed마다 다르게 구성하였고, B 조건에서는 동일한 LOSO 분할을 유지한 상태에서 모델 학습 과정의 난수성을 반영하였다. C 조건에서는 test cat의 enrollment set과 evaluation set 분할을 seed마다 새로 구성하였다.

3.2.2 개인화 조건

개인화 조건에서는 각 고양이의 음원 파일만을 사용하여 개체별 분류 모델을 학습하고 평가하였다. 다른 고양이의 음원 파일은 학습에 포함하지 않았으며, 동일 고양이 내부에서 train set과 test set을 나누었다. 이를 통해 특정 고양이 내부에서 food와 isolation 발성이 구분 가능한지를 평가하였다.

또한 개인화 조건에서는 모든 고양이에게 동일한 입력 특징을 적용하지 않고, 고양이별로 food와 isolation 구분에 상대적으로 유용한 특징을 선택하였다. 이를 위해 각 고양이에 대해 5개 후보 특징별로 food 클래스와 isolation 클래스 간 차이를 평가하였다. 두 클래스의 표본 수와 분산이 동일하지 않을 수 있으므로 Welch's two-sample t-test를 적용하였다 [33]. p-value는 클래스 간 평균 차이가 통계적으로 유의한지를 판단하는 데 사용될 수 있으나, 해당 차이의 실제 크기를 직접적으로 나타내지는 않는다. 따라서 고양이별 특징 선택에서는 클래스 간 차이의 유의성뿐 아니라 차이의 크기도 함께 고려하기 위해 각 후보 특징에 대해 p-value와 He-dges'g 기반 effect size를 함께 산출하였다[34].

Hedges'g는 food 클래스와 isolation 클래스 간 평균 차이의 크기를 표준화하여 정량화하기 위해 사용하였다[34]. 이에 따라 본 연구에서는 통계적으로 유의한 차이를 보이면서도 클래스 간 차이의 크기가 충분한 특징을 선택하기 위해, $p < 0.05$ 를 만족하고 Hedges'g의 절댓값이 0.8 이상인 특징을 해당 고양이의 개인화 모델 입력 특징으로 선택하였다[34].

이렇게 정의한 개인화 특징 부분집합은 해당 고양이의 모든 교차검증 fold에 동일하게 적용하였다. 예를 들어, 특정 고양이에서 mean log-energy와 vocalization length가 선택된 경우, 해당 고양이의 모든 fold에서 두 특징만을 입력으로 사용하였다.

각 고양이 내부 학습 및 평가는 stratified k-fold cross-validation으로 수행하였다[35]. Fold 수는 해당 고양이에서 두 클래스 중 더 적은 음원 파일 수에 맞추어 설정하였으며, food와 isolation 음원 수가 각각 n_{food} , $n_{isolation}$ 일 때 $k = \min(n_{food}, n_{isolation})$ 로 정의하였다. 예를 들어, food 음원이 5개, isolation 음원이 12개인 고양이의 경우 $k = 5$ 로 설정하였다. 이와 같이 fold 수를 소수 클래스의 음원 수 이하로 제한함으로써, 각 test fold에 food와 isolation 클래스가 모두 포함되도록 하였다. 최종 성능은 각 fold의 평가 결과를 종합하여 고양이별로 산출하였으며, 이를 바탕으로 6마리 고양이에 대한 개체별 성능을 계산하였다.

3.2.3 전이 조건

전이 조건에서는 여러 고양이로부터 학습한 분류 모델이 학습에 사용되지 않은 새로운 고양이에게 적용될 수 있는지를 평가하였다. 본 연구에서 전이는 별도의 pre-training/fine-tuning 절차가 아니라, 여러 고양이로 학습한 모델을 학습에 포함되지 않은 새로운 고양이에게 적용하는 개체 간 전이 설정을 의미한다. 이를 위해 LOSO 방식을 사용하였다. 각 반복에서 6마리의 test cat 후보 중 한 마리를 t-test cat으로 지정하고, 해당 고양이의 음원 파일은 학습 과정에서 완전히 제외하였다. 모델은 test cat을 제외한 16마리의 음원 파일로 학습되었으며, 학습된 분류 모델을 test cat의 음원 파일에 적용하여 성능을 평가하였다. 이 조건에서는 test cat에 대한 특징

선택이나 임계값 조정을 적용하지 않고, vocalization length, mean log-energy, mean zero-crossing rate, skewness, kurtosis의 5개 음향 특징 모두 분류 모델의 입력으로 사용하였다.

3.2.4 임계값 개인화 기반 전이 조건

임계값 개인화 기반 전이 조건에서는 전이 조건과 동일한 LOSO 구조를 사용하되, test cat의 일부 라벨 데이터를 이용하여 판정 임계값을 조정하였다. 각 반복에서 한 마리의 고양이를 test cat으로 지정하고, test cat을 제외한 16마리의 음원 파일로 분류 모델을 학습하였다. 입력 특징은 전이 조건과 동일하게 vocalization length, mean log-energy, mean zero-crossing rate, skewness, kurtosis의 5개 음향 특징 전체를 사용하였다.

이후 학습에 사용되지 않은 test cat의 음원 파일을 stratified split으로 enrollment set과 evaluation set으로 나누었다. Test cat 후보는 두 클래스 모두에서 최소 4개 이상의 음원이 남아 있는 고양이로 제한하였으므로, 50:50 분할을 통해 enrollment set과 evaluation set 모두에 양 클래스 샘플이 포함되도록 하였다.

Enrollment set은 모델 파라미터를 재학습하는 데 사용하지 않고, 해당 test cat의 판정 임계값을 선택하는 데에만 사용하였다. Evaluation set은 임계값 선택 과정에서 제외하고, 선택된 임계값을 적용한 최종 성능 평가에만 사용하였다[36]. 임계값 조정은 enrollment set에서 산출된 food 클래스 예측 확률을 기준으로 수행하였다. 먼저 예측 확률을 정렬한 뒤, 인접한 확률값 사이의 중간값을 후보 임계값으로 설정하였다[37]. 각 후보 임계값을 적용했을 때 enrollment set의 accuracy가 가장 높은 값을 해당 test cat의 개인화 임계값으로 선택하였고[32], 선택된 임계값을 evaluation set에 적용하여 최종 성능을 평가하였다. 이 조건에서 개인화는 모델 구조나 입력 특징을 새로 학습하는 방식이 아니라, test cat의 소량 라벨 데이터를 이용해 임계값을 조정하는 방식으로 수행되었다.

3.3 분류 모델 및 평가 방법

본 연구에서는 전이 및 개인화 전략의 효과가 분류 모델 종류에 따라 달라지는지를 확인하기 위해 총 다섯 가지 분류 모델을 비교하였다. 후보 모델은 SVM(Support Vector Machine), R-F(Random Forest), XGBoost[38], DT(Decision Tree)[39], CNN(Convolutional Neural Network)[40]으로 구성하였다. 모든 모델은 동일한 음향 특징 세트를 입력으로 사용하였다.

SVM은 특징 공간에서의 선형 분리를 기반으로 하는 linear kernel을 적용하였으며, decision tree, random forest, XGBoost는 특징 간 비선형 관계와 상호작용을 반영할 수 있는 트리 기반 모델로 포함하였다[38][39]. CNN의 경우 5개 음향 특징 벡터를 1D 입력으로 구성하여 신경망 기반 분류 모델로 사용하였다[40].

본 연구의 데이터는 food 클래스와 isolation 클래스의 음원 수가 불균형하므로, 단순 Accuracy뿐 아니라 food 클래스 탐지 성능을 나타내는 Food Recall과 불균형 데이터에서 양성 클래스 탐지 성능을 반영하는 PR-AUC를 함께 고려하였다. 또한 isolation 클래스에 대한 탐지 성능을 확인하기 위해 Isolation Recall을 산출하였고, 여러 임계값에서의 전반적인 분류 가능성을 평가하기 위해 ROC-AUC도 계산하였다. 성능 지표는 각 고양이에 대해 10개 seed 반복 결과를 평균한 뒤, 6마리 고양이의 개체별 평균 성능을 기준으로 평균±표준편차로 산출하였다.

IV. 결 과

4.1 평가 조건별 특징 적용 결과

본 연구의 평가 조건은 개인화 조건(A), 전이 조건(B), 임계값 개인화 기반 전이 조건(C)으로 구성하였다. 개인화 조건에서는 고양이별로 선택된 특징 부분집합을 입력으로 사용하였고, 전이 조건과 임계값 개인화 기반 전이 조건에서는 별도의 특징 선택 없이 5개 후보 특징 전체를 입력으로 사용하였다. 그림 2는 개인화 조건에서의 특징 선택 결과를 해석하기 위해, 6마리 평가 대상 고양이에 대해 food 클래스와 isolation 클래스 간 5개 후보 특징의 분포를 box plot으로 나타낸 것이다. 고양이별 특징 선택은 Welch's two-sample t-test 결과를 기준으로 수행하

였다. 그림 2에서 5개 후보 특징의 클래스별 분포는 고양이별로 서로 다른 양상을 보였다. 일부 고양이에서는 mean log-energy 또는 vocalization length에서 food와 isolation 클래스 간 중앙값 차이가 관찰되었고, mean ZCR은 여러 고양이에서 food 클래스가 isolation 클래스보다 높은 값의 분포를 보였다. 반면 skewness와 kurtosis는 고양이에 따라 클래스 간 분포의 겹침 정도와 차이의 방향이 다르게 나타났다. 이는 두 클래스 간 분포 차이가 모든 고양이에서 동일한 특징으로 나타나지 않음을 보여준다.

개인화 조건 특징 선택 결과, 고양이별로 서로 다른 특징 조합이 선택되었다. MEG01에서는 mean log-energy만 선택된 반면, WHO01에서는 vocalization length, mean ZCR, mean log-energy의 세 가지 특징이 선택되었다. BAC01에서는 mean ZCR과 kurtosis, DAK01과 MIN01에서는 vocalization length와 mean log-energy, SPI01에서는 vocalization length와 mean ZCR이 선택되었다. 이는 food와 isolation 클래스 간 차이를 설명하는 음향적 단서가 개체마다 다를 수 있음을 의미한다.

4.2 평가 조건별 분류 모델 성능 비교

본 절에서는 먼저 개인화, 전이, 임계값 개인화 기반 전이 각 조건에서 분류 모델별 성능을 비교하였다. 비교 대상 모델은 XGBoost, SVM, RF(Random Forest), DT, CNN으로 구성하였다. 표에서 Acc는 accuracy를, Food Rec와 Iso Rec는 각각 food 클래스와 isolation 클래스에 대한 recall을 의미한다.

4.2.1 개인화 조건에서 분류 모델 성능 비교

표 2에서 확인할 수 있듯이, 개인화 조건에서는 평가 지표에 따라 가장 높은 평균값을 보인 분류 모델이 다르게 나타났다. RF는 Accuracy와 ROC-AUC에서 각각 0.858 ± 0.124 , 0.888 ± 0.111 로 가장 높은 평균값을 보였고, CNN은 Food Recall, Isolation Recall, PR-AUC에서 각각 0.854 ± 0.151 , 0.849 ± 0.063 , 0.847 ± 0.159 로 가장 높은 평균값을 보였다. SVM은 PR-AUC가 0.816 ± 0.193 으로 나타났으나, Food Recall은 0.661 ± 0.277 로 다섯 모델 중 가장 낮았다.

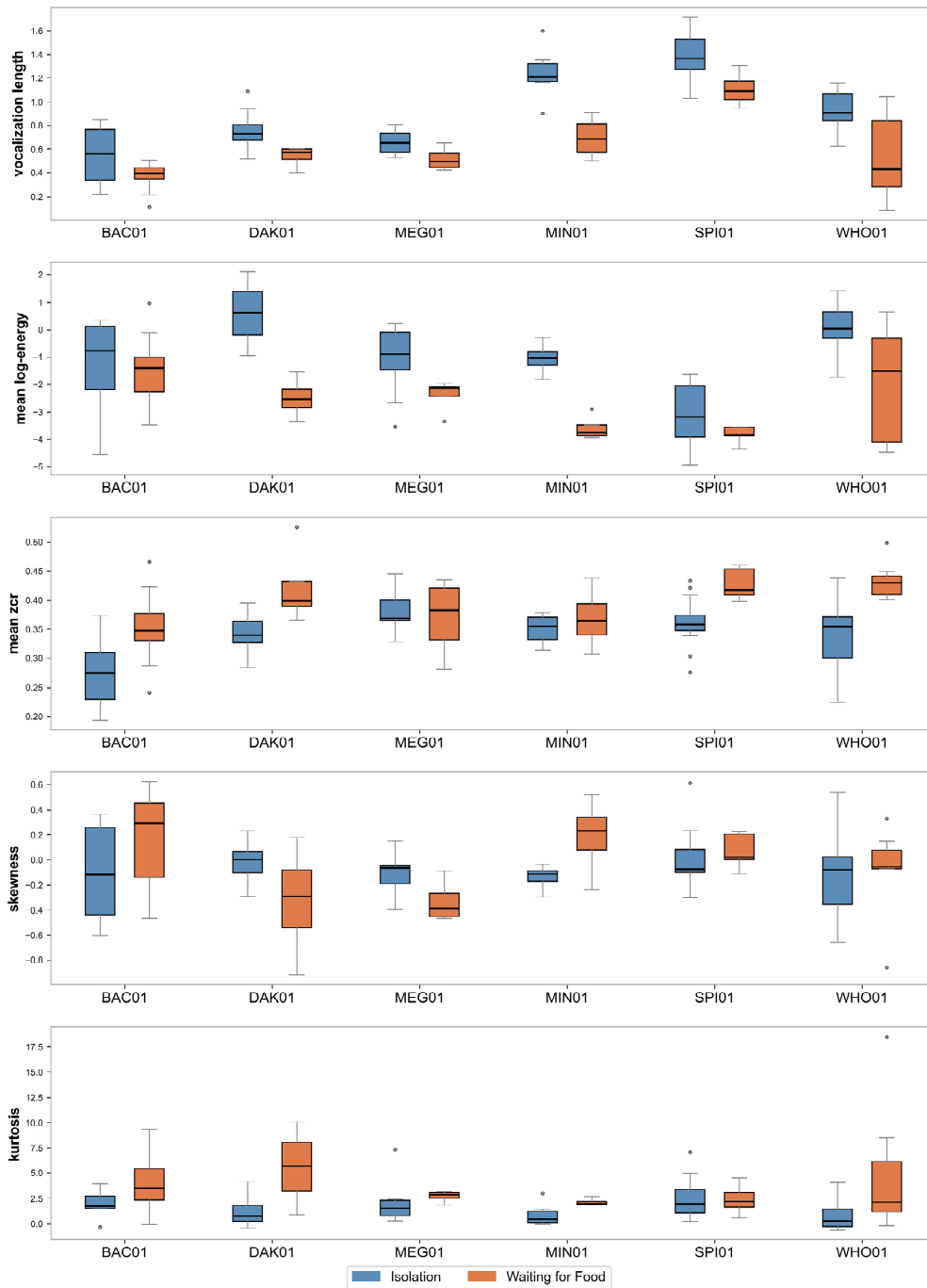


그림 2. 개인화 조건에서의 고양이별 후보 특징 분포

Fig. 2. Cat-specific distributions of candidate features for the personalization condition

표 2. 개인화 조건에서 분류 모델별 성능 비교

Table 2. Performance comparison of classification models under the personalization condition

Model	Acc	Food Rec	Iso Rec	ROC-AUC	PR-AUC
XGBoost	0.795±0.090	0.773±0.176	0.820±0.103	0.866±0.065	0.796±0.141
SVM	0.817±0.135	0.661±0.277	0.810±0.272	0.847±0.156	0.816±0.193
RF	0.858±0.124	0.805±0.098	0.833±0.224	0.888±0.111	0.836±0.145
DT	0.823±0.123	0.773±0.075	0.814±0.179	0.793±0.119	0.687±0.091
CNN	0.834±0.079	0.854±0.151	0.849±0.063	0.886±0.094	0.847±0.159

DT는 Accuracy 0.823 ± 0.123 , Food Recall 0.773 ± 0.075 , Isolation Recall 0.814 ± 0.179 로 나타났으나, PR-AUC는 0.687 ± 0.091 로 다섯 모델 중 가장 낮았다. 따라서 개인화 조건에서는 특정 모델이 모든 평가 지표에서 일관되게 가장 높은 평균값을 보이지 않았다. 모델별 성능 양상은 Accuracy, Food Recall, Isolation Recall, ROC-AUC, PR-AUC에 따라 다르게 나타났다.

4.2.2 전이 조건에서 분류 모델 성능 비교

표 3에 제시된 결과와 같이, 전이 조건에서는 XGBoost가 Accuracy, ROC-AUC, PR-AUC에서 각각 0.835 ± 0.122 , 0.926 ± 0.071 , 0.924 ± 0.070 으로 가장 높은 평균값을 보였다. RF 역시 ROC-AUC와 PR-AUC에서 XGBoost와 유사한 수준의 성능을 보였으나, Food Recall은 0.494 ± 0.313 으로 낮게 나타났다.

전이 조건에서는 XGBoost가 Accuracy, ROC-AUC, PR-AUC에서 가장 높은 평균값을 보였다. 다만 XGBoost의 Food Recall은 0.644 ± 0.417 로 Isolation Recall보다 낮았고, 고양이별 편차도 크게 나타났다. SVM은 class 간 recall 차이가 가장 뚜렷한 모델로, Isolation Recall은 전체 모델 중 가장 높았으나 Food Recall은 0.300 ± 0.286 으로 가장 낮았다. RF 역시 ROC-AUC와 PR-AUC는 XGBoost와 유사한 수준

이었지만, Food Recall은 상대적으로 낮게 나타났다. 반면 CNN은 Food Recall이 XGBoost보다 높았으나, Accuracy와 Isolation Recall은 낮았다.

4.2.3 임계값 개인화 기반 전이 조건에서 분류 모델 성능 비교

표 4에서 확인할 수 있듯이, 임계값 개인화 기반 전이 조건에서는 XGBoost가 모든 평가 지표에서 가장 높은 평균값을 보였다. 전이 조건과 비교했을 때, XGBoost의 Food Recall은 0.644 ± 0.417 에서 0.783 ± 0.175 로 증가하였고, Accuracy도 함께 증가하였다. 반면 PR-AUC는 전이 조건과 유사한 수준을 유지하였다. 모델별 비교에서도 XGBoost의 우위가 유지되었으며, SVM과 RF 모두 전이 조건보다 Food Recall 평균값은 증가했지만 전체 평가 지표의 평균값은 XGBoost보다 낮았다. RF는 ROC-AUC와 PR-AUC에서 각각 0.932 ± 0.061 , 0.924 ± 0.050 으로 XGBoost와 유사한 수준을 보였으나, Food Recall은 0.718 ± 0.076 으로 상대적으로 낮았다. SVM은 Food Recall이 0.751 ± 0.135 로 비교적 높았지만, ROC-AUC와 PR-AUC는 XGBoost와 RF보다 낮았다. DT와 CNN은 전반적으로 낮은 성능을 보였으며, 특히 DT는 Food Recall과 PR-AUC에서 가장 낮은 값을 보였다.

표 3. 전이 조건에서 분류 모델별 성능 비교

Table 3. Performance comparison of classification models under the transfer condition

Model	Acc	Food Rec	Iso Rec	ROC-AUC	PR-AUC
XGBoost	0.835 ± 0.122	0.644 ± 0.417	0.938 ± 0.087	0.926 ± 0.071	0.924 ± 0.070
SVM	0.719 ± 0.196	0.300 ± 0.286	0.983 ± 0.041	0.883 ± 0.112	0.873 ± 0.110
RF	0.792 ± 0.130	0.494 ± 0.313	0.953 ± 0.073	0.925 ± 0.061	0.908 ± 0.051
DT	0.770 ± 0.073	0.463 ± 0.284	0.923 ± 0.113	0.693 ± 0.100	0.525 ± 0.186
CNN	0.685 ± 0.125	0.673 ± 0.380	0.650 ± 0.392	0.792 ± 0.192	0.764 ± 0.160

표 4. 임계값 개인화 기반 전이 조건에서 분류 모델별 성능 비교

Table 4. Performance comparison of classification models under the transfer with threshold personalization condition

Model	Acc	Food Rec	Iso Rec	ROC-AUC	PR-AUC
XGBoost	0.881 ± 0.070	0.783 ± 0.175	0.910 ± 0.121	0.934 ± 0.063	0.934 ± 0.061
SVM	0.825 ± 0.101	0.751 ± 0.135	0.822 ± 0.200	0.888 ± 0.108	0.896 ± 0.084
RF	0.840 ± 0.088	0.718 ± 0.076	0.863 ± 0.173	0.932 ± 0.061	0.924 ± 0.050
DT	0.757 ± 0.075	0.475 ± 0.322	0.854 ± 0.190	0.703 ± 0.076	0.578 ± 0.154
CNN	0.790 ± 0.070	0.731 ± 0.229	0.742 ± 0.098	0.796 ± 0.096	0.799 ± 0.114

반면 XG-Boost는 Accuracy, Food Recall, Isolation Recall, ROC-AUC, PR-AUC의 모든 평가 지표에서 가장 높은 평균값을 보였다. 특히 Food Recall과 Isolation Recall이 각각 0.783 ± 0.175 , 0.910 ± 0.121 로 함께 높게 나타났고, ROC-AUC와 PR-AUC도 각각 0.934 ± 0.063 , 0.934 ± 0.061 로 가장 높은 값을 보였다. 따라서 임계값 개인화 기반 전이 조건에서는 XGB-oost가 food 클래스와 isolation 클래스 관련 지표에서 모두에서 가장 높은 평균값을 보였다. 이에 따라 다음 절에서는 XGBoost를 기준 모델로 사용하여 개인화, 전이, 임계값 개인화 기반 전이 조건 간 성능 변화를 비교하였다.

4.3 XGBoost 기준 실험 조건별 성능 비교

앞 절의 분류 모델 비교에서 XGBoost가 임계값 개인화 기반 전이 조건의 모든 평가 지표에서 가장 높은 평균 성능을 보였으므로, 본 절에서는 XGBoost를 기준 모델로 설정하고 개인화, 전이, 임계값 개인화 기반 전이 조건의 성능을 표 5에 제시하여 비교하였다.

표 5. XGBoost 기준 실험 조건별 주요 성능 비교
Table 5. Comparison of key performance metrics across experimental conditions based on XGBoost

Condition	Accuracy	Food Recall	PR-AUC
Personalization	0.795 ± 0.090	0.773 ± 0.176	0.796 ± 0.141
Transfer	0.835 ± 0.122	0.644 ± 0.417	0.924 ± 0.070
Transfer with threshold personalization	0.881 ± 0.070	0.783 ± 0.175	0.934 ± 0.061

XGBoost 기준 조건 비교에서 가장 두드러진 변화는 전이 조건에서 PR-AUC와 Food Recall이 서로 다른 방향으로 변한 점이다. 전이 조건의 PR-AUC는 개인화 조건보다 높았으나, Food Recall은 낮았다. 또한 Food Recall의 표준편차는 개인화 조건보다 크게 나타나, test cat에 따른 차이가 크게 나타났다.

임계값 개인화 기반 전이 조건에서는 Accuracy, Food Recall, PR-AUC가 세 조건 중 가장 높은 평균값을 보였다. 특히 Food Recall은 전이 조건보다 높

아졌고, 표준편차는 감소하였다. 반면 PR-AUC는 전이 조건과 유사한 수준을 유지하였다. 따라서 XGB-oost 기준 조건 비교에서는 전이 조건에서 Food Recall의 감소와 변동성 증가가 관찰되었고, 임계값 개인화 기반 전이 조건에서는 Food Recall의 평균값 증가와 변동성 감소가 함께 확인되었다.

V. 논 의

본 연구에서 가장 중요한 결과는 전이 조건에서 PR-AUC와 Food Recall이 서로 다른 양상을 보였다는 점이다. XGBoost 기준으로 전이 조건은 개인화 조건보다 높은 PR-AUC를 보였지만, 기본 임계값을 적용했을 때의 Food Recall은 낮게 나타났다.

이는 여러 고양이의 데이터를 이용해 학습한 모델이 food 클래스에 해당하는 발성을 비교적 높은 확률 순위에 배치할 수 있더라도, 기본 임계값을 그대로 적용하면 최종 분류 결과에서는 food 클래스로 충분히 판정하지 못할 수 있음을 의미한다.

임계값 개인화 기반 전이 조건에서는 모델을 다시 학습하지 않고 test cat의 enrollment set만을 이용해 임계값을 조정하였다. 그 결과 Food Recall은 전이 조건보다 증가하였고, PR-AUC는 전이 조건과 유사한 수준을 유지하였다. PR-AUC가 여러 임계값에서의 precision-recall 관계를 반영하는 지표라는 점을 고려하면, 이러한 결과는 모델의 확률 기반 순위화 성능 자체가 크게 변했다기보다는, test cat에 맞는 판정 기준을 적용하면서 최종 분류 결과 기준의 food 클래스 판정 성능이 달라진 것으로 해석할 수 있다. 따라서 새로운 고양이에 모델을 적용할 때에는 모델 성능뿐 아니라, 예측 확률을 최종 클래스로 변환하는 임계값 설정도 함께 고려할 필요가 있다.

개인화 조건에서 선택된 특징이 고양이마다 다르게 나타난 점도 주목할 만하다. 일부 고양이에서는 mean log-energy가 food 클래스와 isolation 클래스를 구분하는 데 사용되었지만, 다른 고양이에서는 mean zero-crossing rate나 vocalization length가 선택되었다. 이는 food 클래스 발성과 isolation 클래스 발성을 구분하는 음향적 단서가 모든 고양이에서 동일하게 나타나지 않을 수 있음을 시사한다. 다만 본 연구에서 사용한 특징은 평균적인 시간 영역 특성에 중심

을 두고 있으므로, 발성 내부의 시간적 변화나 프레 임 간 변동성을 충분히 반영하지는 못한다.

본 연구의 결과를 해석할 때에는 데이터 규모와 클래스 불균형을 함께 고려해야 한다. 품질 필터링 이후 최종 분석에 포함된 음원은 17마리 고양이의 273개였으며, 이 중 개체별 평가가 가능했던 test cat은 6마리였다. 또한 food 클래스의 음원 수가 isolation 클래스보다 적어, Food Recall의 변화가 일부 test cat의 데이터 분포나 분할 방식에 영향을 받을 가능성이 있다. 따라서 본 연구의 결과는 전체 반려묘 집단에 대한 일반적인 결론이라기보다는, 제한된 공개 데이터셋 조건에서 관찰된 전이 성능과 임계값 조정 효과로 해석하는 것이 적절하다.

또한 본 연구는 공개 데이터셋을 기반으로 수행되었기 때문에 실제 환경에서 발생하는 입력 조건을 충분히 반영하지 못했다. 실제 가정이나 보호소에서는 배경소음, 마이크 위치 변화, 녹음 장치 차이, 여러 고양이의 동시 발성 등이 함께 나타날 수 있다. 또한 실제 기기에 적용하기 위해서는 연속 녹음 데이터에서 발성 구간을 먼저 검출하고, 이후 해당 발성이 실제 급식 요구 상황에서 나타나는 발성인지 판단하는 과정이 필요하다. 따라서 후속 연구에서는 실제 급식 전후 환경에서 수집한 반려묘 발성 데이터를 이용하여, 발성 구간 검출과 급식 요구 발성의 자동 식별을 함께 평가할 필요가 있다.

그럼에도 본 연구는 CatMeows 데이터셋을 이용해 새로운 고양이에 대한 개체 간 전이 성능과 test cat별 임계값 조정 효과를 분리하여 분석했다는 점에서 의의가 있다. 특히 모델을 재학습하지 않더라도 소량의 test cat 라벨 데이터를 이용한 임계값 조정만으로 먹이 제공 전 대기 상황 발성의 최종 판정 성능이 달라질 수 있음을 확인하였다. 따라서 실제 급식 요구 발성 자동 식별 시스템에 새로운 개체를 적용할 때에는, 소량의 개체별 데이터를 이용한 임계값 보정 절차를 함께 고려할 필요가 있다.

VI. 결 론

본 연구는 CatMeows 데이터셋을 이용하여 food 클래스와 isolation 클래스를 구분하는 이진 분류 문

제를 구성하고, 개인화 조건, 전이 조건, 임계값 개인화 기반 전이 조건의 성능을 비교하였다. 실험 결과, XGBoost 기준으로 임계값 개인화 기반 전이 조건에서 Accuracy, Food Recall, PR-AUC의 평균값이 가장 높게 나타났다. 특히 전이 조건에서는 PR-AUC가 높게 나타났음에도 기본 임계값에서의 Food Recall이 낮았으나, test cat의 enrollment set을 이용해 임계값을 조정한 후에는 Food Recall이 증가하고 PR-AUC는 유사한 수준을 유지하였다. 따라서 본 연구는 food 클래스와 isolation 클래스를 구분하는 이진 분류 과제에서, 새로운 고양이에 대한 전이 성능을 평가할 때 임계값 조정 효과를 함께 고려할 필요가 있음을 보였다는 점에서 의의가 있다.

References

- [1] C. Saavedra, C. Pérez, C. Oyarzún, and Á. Torres-Arévalo, "Overweight and obesity in domestic cats: epidemiological risk factors and associated pathologies", *Journal of Feline Medicine and Surgery*, Vol. 26, No. 11, pp. 1-11, Nov. 2024. <https://doi.org/10.1177/1098612X241285519>.
- [2] S. Ligout, X. Si, H. Vlaeminck, and S. Lyn, "Cats reorganise their feeding behaviours when moving from ad libitum to restricted feeding", *Journal of Feline Medicine and Surgery*, Vol. 22, No. 10, pp. 953-958, Oct. 2020. <https://doi.org/10.1177/1098612X19900387>.
- [3] M. Delgado and L. M. S. Dantas, "Feeding Cats for Optimal Mental and Behavioral Well-Being", *Veterinary Clinics of North America: Small Animal Practice*, Vol. 50, No. 5, pp. 939-953, Sep. 2020. <https://doi.org/10.1016/j.cvs.2020.05.003>.
- [4] M. Smit, S. J. Ikuir, R. A. Corner-Thomas, C. J. Andrews, I. Draganova, and D. G. Thomas, "The Use of Triaxial Accelerometers and Machine Learning Algorithms for Behavioural Identification in Domestic Cats (*Felis catus*): A Validation Study", *Sensors*, Vol. 23, No. 16, Art. no. 7165, Aug. 2023. <https://doi.org/10.3390/s23167165>.

- [5] A. Witzel-Rollins, M. Murphy, C. M. Springer, T. D. Moyers, and J. D. Albright, "Evaluation of a pet-separating automatic feeder and high-frequency meal feeding for weight loss in multi-cat households", *Journal of Feline Medicine and Surgery*, Vol. 24, No. 8, pp. e281-e288, Aug. 2022. <https://doi.org/10.1177/1098612X221105046>.
- [6] K. McComb, A. M. Taylor, C. Wilson, and B. D. Charlton, "The cry embedded within the purr", *Current Biology*, Vol. 19, No. 13, pp. R507-R508, Jul. 2009. <https://doi.org/10.1016/j.cub.2009.05.033>.
- [7] S. Ntalampiras, L. A. Ludovico, G. Presti, E. Prato Previde, M. Battini, S. Cannas, C. Palestini, and S. Mattiello, "Automatic Classification of Cat Vocalizations Emitted in Different Contexts", *Animals*, Vol. 9, No. 8, Art. no. 543, Aug. 2019. <https://doi.org/10.3390/ani9080543>.
- [8] L. A. Ludovico, S. Ntalampiras, G. Presti, S. Cannas, M. Battini, and S. Mattiello, "CatMeows: A Publicly-Available Dataset of Cat Vocalizations", *27th International Conference on MultiMedia Modeling (MMM 2021)*, Prague, Czech Republic, pp. 230-243, Jun. 2021. https://doi.org/10.1007/978-3-030-67835-7_20.
- [9] M. Kukushkin and S. Ntalampiras, "Automatic Acoustic Classification of Feline Sex", *16th International Audio Mostly Conference*, Trento, Italy, pp. 156-160, Sep. 2021. <https://doi.org/10.1145/3478384.3478385>.
- [10] W. Raccagni and S. Ntalampiras, "Acoustic Classification of Cat Breed Based on Time and Frequency Domain Features", *2021 30th Conference of Open Innovations Association (FRUCT)*, Oulu, Finland, pp. 184-189, Oct. 2021. <https://doi.org/10.23919/FRUCT53335.2021.9599975>.
- [11] D. Stowell, T. Petrusková, M. Šálek, and P. Linhart, "Automatic acoustic identification of individuals in multiple species: improving identification across recording conditions", *Journal of the Royal Society Interface*, Vol. 16, No. 153, Art. no. 20180940, Apr. 2019. <https://doi.org/10.1098/rsif.2018.0940>.
- [12] C. Choo, Y.-U. Chang, and I.-Y. Moon, "FPGA-Based Hardware Accelerator for Feature Extraction in Automatic Speech Recognition", *Journal of Information and Communication Convergence Engineering*, Vol. 13, No. 3, pp. 145-151, Sep. 2015. <https://doi.org/10.6109/jicce.2015.13.3.145>.
- [13] H. K. Maganti and M. Matassoni, "An auditory based modulation spectral feature for reverberant speech recognition", *Interspeech*, pp. 570-573, Sep. 2010. <https://doi.org/10.21437/Interspeech.2010-225>.
- [14] C. Constantinescu and R. Brad, "An Overview on Sound Features in Time and Frequency Domain", *International Journal of Advanced Statistics and IT&C for Economics and Life Sciences*, Vol. 13, No. 1, pp. 45-58, 2023. <https://doi.org/10.2478/ijasitels-2023-0006>.
- [15] L. Czúni and P. Z. Varga, "Time Domain Audio Features for Chainsaw Noise Detection Using WSNs", *IEEE Sensors Journal*, Vol. 17, No. 9, pp. 2917-2924, May 2017. <https://doi.org/10.1109/JSEN.2017.2670232>.
- [16] L. Perrier, A. de W. Thézy, A. Pradeau, C. Schradin, M. D. Greenfield, N. Mathevon, and F. Levréro, "A cooperatively breeding mouse shows flexible use of its vocal repertoire according to social context", *Behavioural Brain Research*, Vol. 486, Art. no. 115575, May 2025. <https://doi.org/10.1016/j.bbr.2025.115575>.
- [17] M. Rotaru, S. Ciocină, and F. Albu, "An efficient GSC VSS-APA beamformer with integrated log-energy based VAD for noise reduction in speech reinforcement systems", *International Symposium on Signals, Circuits and Systems (ISSCS)*, Iasi, Romania, pp. 1-4, Jul. 2013. <https://doi.org/10.1109/ISSCS.2013.6651240>.
- [18] H. Ankişhan, "Classification of acoustic signals with new feature: Fibonacci space (FSp)", *Biomedical Signal Processing and Control*, Vol. 48, pp. 221-233, Feb. 2019. <https://doi.org/10.1016/j.bsp.2019.02.006>.

- j.bspc.2018.08.037.
- [19] M. Smit, R. A. Corner-Thomas, I. Draganova, C. J. Andrews, and D. G. Thomas, "How Lazy Are Pet Cats Really? Using Machine Learning and Accelerometry to Get a Glimpse into the Behaviour of Privately Owned Cats in Different Households", *Sensors*, Vol. 24, No. 8, Art. no. 2623, Apr. 2024. <https://doi.org/10.3390/s24082623>.
- [20] R.-C. Chen, V. S. Saravanarajan, and H.-T. Hung, "Monitoring the behaviours of pet cat based on YOLO model and raspberry Pi", *International Journal of Applied Science and Engineering*, Vol. 18, No. 5, Art. no. 2021137, pp. 1-12, Sep. 2021. [https://doi.org/10.6703/IJASE.202109_18\(5\).016](https://doi.org/10.6703/IJASE.202109_18(5).016).
- [21] M. Moelk, "Vocalizing in the House-Cat; A Phonetic and Functional Study", *American Journal of Psychology*, Vol. 57, No. 2, pp. 184-205, Apr. 1944. <https://doi.org/10.2307/1416947>.
- [22] S. Schötz, J. van de Weijer, and R. Eklund, "Melody Matters: An Acoustic Study of Domestic Cat Meows in Six Contexts and Four Mental States", *VIHAR 2019*, London, UK, pp. 29-34, Aug. 2019. <https://doi.org/10.7287/peerj.preprints.27926v1>.
- [23] N. Nicastro and M. J. Owren, "Classification of Domestic Cat (*Felis catus*) Vocalizations by Naive and Experienced Human Listeners", *Journal of Comparative Psychology*, Vol. 117, No. 1, pp. 44-52, Mar. 2003. <https://doi.org/10.1037/0735-7036.117.1.44>.
- [24] E. Prato-Previde, S. Cannas, C. Palestrini, S. Ingrassia, M. Battini, L. A. Ludovico, S. Ntalampiras, G. Presti, and S. Mattiello, "What's in a Meow? A Study on Human Classification and Interpretation of Domestic Cat Vocalizations", *Animals*, Vol. 10, No. 12, Art. no. 2390, Dec. 2020. <https://doi.org/10.3390/ani10122390>.
- [25] I. Nolasco, et al., "Learning to detect an animal sound from five examples", *Ecological Informatics*, Vol. 77, Art. no. 102258, Aug. 2023. <https://doi.org/10.1016/j.ecoinf.2023.102258>.
- [26] S. Kim, J. Choi, B.-K. Kim, H. Kim, D. Kang, J. W. Choi, Y. G. Yoon, and S. Cho, "Reconstruction Modeling and Validation of Brown Croaker (*Micthys miiuy*) Vocalizations Using Wavelet-Based Inversion and Deep Learning", *Sensors*, Vol. 25, No. 19, Art. no. 6178, Oct. 2025. <https://doi.org/10.3390/s25196178>.
- [27] S. Gunasekaran and K. Revathy, "Automatic Recognition and Retrieval of Wild Animal Vocalizations", *International Journal of Computer Theory and Engineering*, Vol. 3, No. 1, pp. 136-140, Feb. 2011. <https://doi.org/10.7763/IJCTE.2011.V3.295>.
- [28] S. Gunasekaran and K. Revathy, "Content-Based Classification and Retrieval of Wild Animal Sounds Using Feature Selection Algorithm", *2010 Second International Conference on Machine Learning and Computing*, Bangalore, India, pp. 272-275, Feb. 2010. <https://doi.org/10.1109/ICMLC.2010.11>.
- [29] SciPy Developers, "scipy.stats.kurtosis", *SciPy Documentation*, <https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.kurtosis.html>. [accessed: Jun. 07, 2026]
- [30] J. Schneider and M. Vlachos, "Personalization of Deep Learning", *Data Science – Analytics and Applications: Proc. of the 3rd International Data Science Conference – iDSC2020*, Online, pp. 89-96, 2021. https://doi.org/10.1007/978-3-658-32182-6_14.
- [31] J. Rybka and A. Janicki, "Comparison of speaker dependent and speaker independent emotion recognition", *International Journal of Applied Mathematics and Computer Science*, Vol. 23, No. 4, pp. 797-808, Dec. 2013. <https://doi.org/10.2478/amcs-2013-0060>.
- [32] C. Thiele and G. Hirschfeld, "cutpointr: Improved Estimation and Validation of Optimal Cutpoints in R", *Journal of Statistical Software*, Vol. 98, No. 11, pp. 1-27, Jun. 2021. <https://doi.org/10.18637/jss.v098.i11>.
- [33] B. L. Welch, "The generalization of 'Student's'

problem when several different population variances are involved", *Biometrika*, Vol. 34, No. 1-2, pp. 28-35, 1947. <https://doi.org/10.1093/biomet/34.1-2.28>.

- [34] D. Lakens, "Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs", *Frontiers in Psychology*, Vol. 4, Art. no. 863, Nov. 2013. <https://doi.org/10.3389/fpsyg.2013.00863>.
- [35] R. Kohavi, "A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection", 14th International Joint Conference on Artificial Intelligence, Montreal, Quebec, Canada, pp. 1137-1145, Aug. 1995.
- [36] G. C. Cawley and N. L. C. Talbot, "On over-fitting in model selection and subsequent selection bias in performance evaluation", *Journal of Machine Learning Research*, Vol. 11, No. 70, pp. 2079-2107, Jul. 2010.
- [37] C. Lee, J. Lee, and D. W. Kim, "A Hellinger-based discretization method for numeric attributes in classification learning", *Knowledge-Based Systems*, Vol. 20, No. 4, pp. 419-425, Jun. 2007. <https://doi.org/10.1016/j.knosys.2006.06.005>.
- [38] Y. Jin, "Sales Prediction Based on Machine Learning Approach: Support Vector Machine, Xgboost and Random Forest", *Highlights in Science, Engineering and Technology*, Vol. 88, pp. 104-109, Mar. 2024. <https://doi.org/10.54097/6dxq4p06>.
- [39] J. R. Quinlan, "Induction of Decision Trees", *Machine Learning*, Vol. 1, pp. 81-106, Mar. 1986. <https://doi.org/10.1023/A:1022643204877>.
- [40] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-Based Learning Applied to Document Recognition", *Proc. of the IEEE*, Vol. 86, No. 11, pp. 2278-2324, Nov. 1998. <https://doi.org/10.1109/5.726791>.

저자소개

호 다 예 (Daye Ho)



2019년 3월 : 부산대학교
분자생물학과(이학사)
2025년 9월 ~ 현재 :
성균관대학교
전자전기컴퓨터공학과
석사과정
관심분야 : 인공지능, 생체신호

김 희 재 (HeeJae Kim)



2025년 2월 : 울산대학교
전기공학부 의공학전공(공학사)
2026년 3월 ~ 현재 :
성균관대학교
전자전기컴퓨터공학과 석사과정
관심분야 : 인공지능, 영상처리

정 주 호 (Joo Ho Jeong)



2024년 6월 : BS degree, Dept. of
Santa Barbara Applied
Mathematics, California
University
2026년 3월 ~ 현재 :
성균관대학교
전자전기컴퓨터공학과
석사과정
관심분야 : 인공지능, 기하학적 딥러닝

정 조 운 (Jo Woon Chong)



2009년 2월 : KAIST 전기 및
전자공학과(공학박사)
2024년 3월 ~ 현재 :
성균관대학교
전자전기공학부 부교수
관심분야 : Biomedical Signal
Processing, Medical
Instrumentation, On-Device AI