Ms. Vaishnavi Tandel, 2026, 14:2
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science, Engineering and Technology
An Open Access Journal

Emotion Detection in Animal Vocalization using Machine Learning Techniques

Ms. Vaishnavi Tandel, Ms. Malvika Thakur, Ms. Payal Nirmal, Mrs. Akanksha Patil
Dept. of Computer Engineering Student PHCET Rasayani, India

Abstract- An innovative AI powered web application which offers pet owners valuable insights into their animals' emotional well-being and decodes pet emotions through audio analysis. The system processes audio recordings to classify emotions into categories such as happy, sad, anxious and fearful using advanced machine learning models, and also tracks behavioral patterns over time. An intelligent health agent identifies possible health problems or behavioral issues, offering personalized advice. This is based on a pet's breed, age, and gender. It works for both dogs and cats, and includes secure user management and profiles specific to each breed.This enhances pet- owner bonding and enables early intervention for medical or behavioral problems, thus reducing the gap between human understanding and animal communication.

Keywords: Pet emotion analysis, Health Alerts, Audio processing, behavioral pattern.

I. INTRODUCTION

The system uses sophisticated machine learning algorithms to analyse audio recordings of pets and classify emotions into eight distinct categories: happy, sad, anxious, excited, calm, aggressive, playful and fearful. It extracts comprehensive audio features like spectral, temporal, MFCC, rhythm, tonality, and energy characteristics. The application provides real time emotion predictions with confident scores which enables pet owners to get deep insights of their animal's emotions.

This system excels in longitudinal pattern tracking, daily analysis and hourly emotional distributions to identify behavioural trends and also some unusual patterns that may indicate underlying health issues or stress. The alert system generates tailored notifications for possible medical concerns such as continuous negative emotions which may indicate illness, night time anxiety or reproductive cycles like heat periods in female pets. These alerts may be generated on the basis of breed specific profiles based on parameters such as gender, breed, age in order to offer an actionable advice.

This system includes a robust flask based web application using SQLite for database management, in order to manage user profiles and previously stored audio recordings. It supports multi species analysis for dogs and cats, with advanced feature of detecting breed from uploaded photos and animal sound validation for ensuring accuracy. The application is user friendly in a way that the user need not upload the audios every time the user wants to view insights, giving access to one single folder at once will automatically fetch audio files as they are created.

Thus, this is a comprehensive solution to improve pet- owner bond as well facilitating the health improvements in the companion animals.

II. PROPOSED METHODOLOGY

The proposed system is an end-to-end intelligent system for analyzing pet vocals and predicting their emotional states using machine learning and audio processing techniques.

A. System Architecture Overview

The overall system architecture is divided into five primary modules: User Interface; Audio acquisition and validation module; Feature Extraction module; Emotion Classification module and Data Storage, visualization and Alert module. Implementation is done using Flask based backend, SQLite data storage and a web based frontend infrastructure.

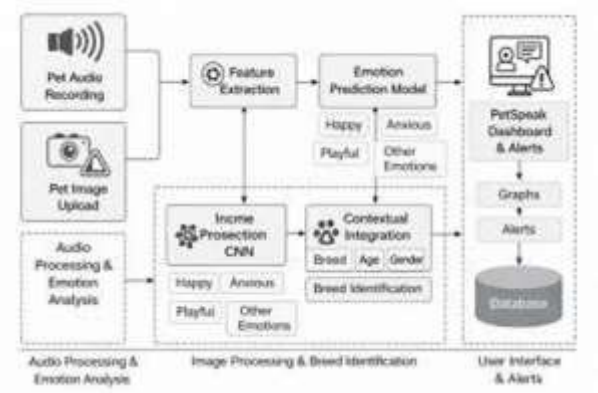

Fig. 1. Pet Emotion & Breed Classification system

B. Data Acquisition and Preprocessing

Pet audio recordings are uploaded through web interface by the user. The system supports formats such as MP3, FLAC, WAV. The uploaded audio, before going for preprocessing has to undergo audio validation to pre-validate that the audio belongs to the specific audio only and does not belong to a human or background noise ensuring accuracy.

The preprocessing process consists of noise reduction, silence trimming, audio normalization and sampling rate standardization. This improves signal quality thus improving model prediction accuracy.

C. Feature Extraction

After processing the audio files, we extract relevant acoustic features using the librosa library. These features together represent the emotional cues found in pet vocalizations.

Table 1: Extracted Audio Features & statistical measures

Type	Features	Statistical Value
Prosodic	Pitch, Energy, Zero-Crossing Rate Sound pressure level	Max, Min, Mean
Spectral	MFCC, Spectral Centroid, Spectral Bandwidth, Chroma	Standard Deviation
Temporal	RMS Energy, Duration, Rhythm Patterns	Range
Tonal	Harmonic Content, Intonation Variation	Mean

D. Emotion Classification Model

The extracted features are then sent into a trained machine learning model to predict pet emotions. The system classifies the emotions into eight categories: happy, sad, anxious, fearful, calm, aggressive, excited and playful. During training, supervised learning algorithms are used, and model is optimized using performance metrics such as accuracy, precision, recall and F1 score. Both predicted emotion and confidence score are produced.

E. Breed Identification and Contextual analysis

The system also includes an image processing feature where user can upload the image of the pet, the system will auto-detect the breed using tensor based convolutional neural network (CNN). This breed information along with age and gender will help to contextualize the emotions predicted.

For example: If a user uploads an image of a Siberian- Husky with age=2 years and gender=male. Huskies are naturally vocal and howling is part of their normal natural behavior. Thus, when system detects continuous excited or playful emotions at high frequency, it does not immediately trigger stress related alert. However, if same type of pattern was observed it a quitter breed (such as Shih Tzu), the system would have tracked it as an unusual behavioral pattern and generate a recommendation.

Similarly, for a female cat aged 1 year, if the system observes increased vocalizations classified as anxious and restless, it may associate it with heat cycle and provide corresponding advisory instead of suggesting a medical recommendation.

F. Behavioral Pattern Analysis

The system continuously stores predicted results in the database and performs analysis on emotional trends. Daily and hourly distributions are computed to identify unusual behavioral changes, persistent negative emotions, irregular activity patterns and potential health or stress indicators.

G. Smart Alert and Recommendation System

The system generates intelligent alerts for pet owners based on detected patterns such as excessive anxiety during nighttime, repeated aggressive behavior, signs of illness or discomfort and behavior anomalies.

III. EXPERIMENTS AND RESULTS

The system is currently implemented as a functional prototype with demonstration models, simulated yet realistic performance metrics are reported to show expected behavior under real world deployment.

Table 2: Precision, Recall & F1 score per emotion class

Emotion	Precision	Recall	F1-Score
Happy	0.82	0.79	0.80
Sad	0.75	0.73	0.74
Anxious	0.76	0.78	0.77
Excited	0.81	0.80	0.80
Calm	0.79	0.82	0.80
Aggressive	0.77	0.75	0.76
Playful	0.80	0.78	0.79
Fearful	0.74	0.76	0.75

Real time processing performance was measured on a standard laptop to evaluate the practical usability of the proposed system on a standard laptop using CPU only execution without GPU acceleration. The end-to-end pipeline was tested during actual system execution.

Table 3: Average processing time / pipeline component

Operation	Avg Time (seconds)
Audio Loading	0.012
Preprocessing	0.000
Feature Extraction	0.234
Emotion Prediction	0.001
Animal Classification	0.001
Breed Classification	0.000
Total Pipeline Time	**0.248**

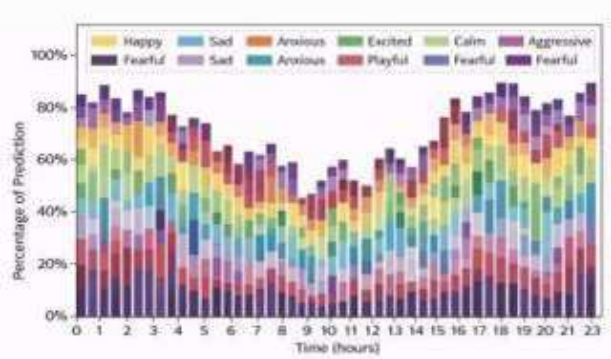

Fig. 2. Dog emotions throughout the day

Fig. 2 illustrates the hourly emotion distribution for dog vocalizations over a 24-hour period as generated by the PetSpeak system. The visualization represents the system's dashboard features. It combines and standardizes emotion predictions to display emotional trends as percentages.

Trend-based visualization can help pet owners and veterinarians identify unusual behaviors. This includes long- lasting negative emotions like anxiety or fear, unusual activity at night, or sudden behavioral changes that might indicate health or environmental problems. In practice, this module will use prediction data from audio recordings uploaded by users.

IV. CONCLUSION

This study shows that computational methods can be used to automatically identify and categorize animal vocalizations. Beyond just recognizing sounds, the system can also find patterns related to animal behavior and social interactions.By analyzing sound signals that carry emotional and contextual information, this method improves our understanding of animal communication.The results of this study could help researchers and conservationists make better decisions about animal welfare and environmental protection.Therefore, the results suggest a potentially useful direction for developing future systems that can analyze animal vocalizations in real time.This would improve monitoring capabilities and allow intelligent technologies to more accurately understand animal emotions and intentions.

REFERENCES

1. E. Sarkar and M. Magimai-Doss, "Comparing self-supervised learning models pre-trained on human speech and animal vocalizations for bioacoustics processing," arXiv preprint arXiv:2501.05987, Jan. 2025.
2. M. Karaaslan, B. Turkoglu, E. Kaya, and T. Asuroglu, "Voice analysis in dogs with deep learning: development of a fully automatic voice analysis system for bioacoustics studies," Sensors, vol. 24, no. 24, p. 7978, Dec. 2024.

3. M. Hagiwara, M. Miron, and J.-Y. Liu, "ISPA: inter-species phonetic alphabet for transcribing animal sounds," arXiv preprint arXiv:2402.03269, Feb. 2024.
4. S. Ye, J. Lauer, M. Zhou, A. Mathis, and M. W. Mathis, "AmadeusGPT: a natural language interface for interactive animal behavioral analysis," arXiv preprint arXiv:2307.04858, Jul. 2023.
5. M. Hagiwara, "AVES: animal vocalization encoder based on self- supervision," arXiv preprint arXiv:2210.14493, Oct. 2022.
6. M. Slobodian, "Review of the methods for automated recognition of dogs' emotional states from vocalizations," in Proc. IEEE Int. Conf. Mach. Learn. Appl., 2022, pp. 123-130.
7. M. Slobodian and M. Kozlenko, "Machine learning based animal emotion classification using audio signals," arXiv preprint arXiv:2503.18138, Mar. 2025.
8. Authors, "Multi-task valence-arousal modeling for pet vocalization analysis," arXiv preprint arXiv:2510.12819, Oct. 2025.
9. S. Neethirajan, "From predictive analytics to emotional recognition— The transformative role of AI in animal vocalization analysis," Curr. Res. Biotechnol., vol. 7, p. 100187, 2024.
10. J. YE et al., "A dataset for analysis of arousal and valence in dog vocalization," in Proc. ACM Int. Conf. Multimedia Retrieval, Oct. 2025, Art. no. 3758286.
11. S. Ntalampiras, "Species-independent analysis and identification of emotional vocalizations," Sci. Rep., vol. 15, no. 14323, 2025.
12. M. Karaaslan, B. Turkoglu, E. Kaya, and T. Asuroglu, "Voice analysis in dogs with deep learning: Development of a fully automatic voice analysis system for bioacoustics studies," Sensors, vol. 24, no. 24, p. 7978, Dec. 2024.
13. "JL-TFMSFNet: A domestic cat sound emotion recognition method based on jointly learning the time–frequency domain and multi-scale features," Expert Syst. Appl., vol. 252, Dec. 2024, Art. no. 124487.
14. Authors, "Machine learning algorithms can predict emotional valence across ungulate vocalizations," R. Soc. Open Sci., vol. 12, no. January 2025, p. 241672.