### Abstract

Audio recognition, Convolutional Neural Networks, Mel Frequency Cepstral Coefficients

### I. Introduction

**- Background information on the project**

In recent years, with the rapid development of deep learning, sound recognition has also been widely used in vary or fields.

**- Research objectives and goals**

The objective of the project is to investigate the sounds that occur preceding a dog’s barking in a local environment. By conducting this research, we aim to develop an automated system for dispensing treats or snacks to the dog before it barks, utilizing the noises as a tool to encourage positive behaviour. The high- level approach is to use two classifiers for classification and recognition purposes: one to identify dog barking and the other to recognize other environmental sounds. By detecting the sound of the dog’s bark, it is possible to determine which noises triggered the barking, and further can be employed to provide positive reinforcement for pet dogs.

**- Research questions**

Unfamiliar and loud noises in the living environment or situations can cause anxiety and restlessness in dogs, resulting in excessive barking. Moreover, each dog may respond differently to various sounds. Consequently, dog owners naturally seek to train their pets to minimize behaviors that cause them discomfort. The objective of this research project is to utilize two classifiers to detect the sounds that induce dogs to bark. The first classifier aims to identify and categorize distracting noises, while the second classifier employs the results of the first classifier to identify the factors that trigger dog barking in an audio clip. The ultimate goal is to predict the behavior of a dog barking and intervene when a particular distraction noise is detected. Unlike the previously mentioned studies, this research not only focuses on accurately classifying environmental sounds but also predicts whether or not a dog will bark after hearing a certain audio clip.

**- Significance of the study**

By researching the relevant work the challenges for both sound classifications would be in data processing, suitable feature extraction, and ap- propriate model selection. Further research will explore the strengths and limitations of each method to identify the optimal performing model. The ultimate goal is to establish a model capable of predicting dog barking and automating dog behaviour training, and this research could make significant contributions towards enhancing the welfare of pet dogs and improving the quality of life for their owners.

**- Scope of the study**

### II. Literature Review

**- Overview of environmental sound classification**

**- Techniques for sound classification**

Numerous studies have focused on the classification of environmental sounds. In its early stages, sound recognition relied primarily on basic machine learning recognition methods including Hidden Markov Models (HMM), Gaussian Mixture Models(GMM), Artificial Neural Networks (ANN) and Support Vector Machines (SVMs) to perform sound classification on environmental sounds[3][8]. Due to the advancements in deep learning in recent years, neural networks have demonstrated exceptional suitability for recognition tasks in effectiveness and accuracy, and convolutional neural networks(CNN) have become the predominant approach for voice recognition[7]. For sound feature extraction, common methods include MFCC, Gammatone, and LBP-HOG features.[9] There is also extensive published work on the binary classification of sounds, with popular models including Support Vector Machine (SVM)[2] and the CNN as mentioned above. In addition, numerous studies on biological characteristics classification have cited the effectiveness of random forest trees[6] and logistic regression[5] in similar classification tasks.

- Binary classification of sounds

- Biological characteristics classification

- Previous studies on dog behavior and barking

- Existing datasets for environmental sound classification

### III. Methodology

Environmental sound classification (ESC) is a vast research area, and numerous scholars have conducted extensive research in this field, and several related datasets are available. This project of classifying environmental sounds will use a dataset consisting of a collection of urban sounds, which includes 8,732 common sounds heard in a city that contains dog bark[4]. Waveform signals will be extracted from all sound files for feature extraction. One disadvantage of this dataset is that the classification of environmental noise in real-life situations is not limited to those included in the dataset. Therefore, it is necessary to expand the classification of environmental noise and include an ’other’ category in the classification. A multi-class classifier will be developed using a CNN-based neural network model. The emphasis of this model will be on using appropriate convolution kernels to efficiently extract waveform signal features. To achieve the best training results for this classification task, various convolutional neural network structures, data inputs, number of pooling layers, and learning rate parameters will be explored [3].

The second classification task is a binary classification for audio, aimed at predicting whether a dog will bark. As each dog may behave differently, the training data will consist of combined sounds based on UrbanSound8k. Individual noises and barking sounds will be merged into longer audio files using Audio Clips. Since each audio file is lengthy, feature extraction becomes challenging due to the vast amount of data involved. To tackle this issue, our initial approach is to segment each audio file into smaller pieces and mark the class as ’N’ for the segments that the dogs didn’t react with. This technique will help reduce the size of each data piece while increasing the amount of data to facilitate the extraction of waveform feature graphics. Several models will be built and compared for this classification task to select the best-performing one. The models to be tested include SVM, Random Forest, Logistic Regression, and CNN. Furthermore, techniques such as data augmentation, regularization, and hyperparameter tuning [1] will be utilized to enhance the performance of the model. The incorporation of unsupervised learning methods may also prove effective for this classification task.

- Data collection and preprocessing

- Feature extraction methods

- Classification models (CNN, SVM, Random Forest, Logistic Regression)

- Model evaluation and comparison

- Data augmentation and regularization techniques

- Hyperparameter tuning

- Unsupervised learning methods

- Implementation details and software tools used

### IV. Results and Discussion

- Presentation of results for each classification task

- Evaluation of model performance

- Comparison of different models and techniques used

- Analysis of findings and discussion of insights

- Limitations and future research directions

V. Conclusion and Recommendations

- Summary of the study

- Implications of the research for dog behavior training

- Recommendations for future research

- Conclusion and final remarks

### VI. References

- List of all sources cited in the report

### VII. Appendix

- Code samples, tables, figures, and other supplementary materials