### 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, including automatic speech recognition (ASR)\cite{kabir2021survey} and environmental sound classification (ESC) \cite{bansal2022environmental}, and has been widely applied in smart home, medical, security monitoring fields. The emergence of audio recognition based on the development of machine learning and neural networks has made it possible to handle complex audio classification tasks. In addition to the areas mentioned, there has been quite a wide range of applications for animal voice recognition. The convolutional neural network (CNN) models outstanding ability to distinguish differences in audio of has made it possible perform breed identification and environmental quality issue based on animal sound analysis\cite{wang2019cough}.

Large numbers of automation of decisioning and classification of sound-based processing tasks have been raised and discussed. In this paper, a method to predict the cause of dog barking based on environmental sound recognition will be demonstrated. Due to the large variety and amount of noise in the city, these noises would cause anxiety in dogs and leads to excessive barking. Each individual dog is sensitive to different noises and lead to different reactions, as a result each dog will have their different noises that trigger barking depending on personality and environment. It is therefore natural for dog owners to seek to train their pets to minimize the behaviours that cause discomfort.

This prediction task of dog barking the following difficulties: (1) the nature of large amount and variety of urban noises around households therefore for general classification of noised a dataset contains large range of different noised is required; (2) difficulty to collect a large amount of data of each individual as the barking trigger sound is different, and there is almost no audio datasets about dog trigger barking is available, hence audio data augmentation technique is important; (3) Based on breed, characteristic, growing environment etc. each dog would has different behavioural patterns, a general strategy to cover all to predict barking behaviour for a specific dog.

This paper discusses a machine learning-based approach to automate the prediction and barking behaviour of domesticated dogs to achieve behaviours that can intervene and prevent barking triggered by environmental noise and could further positively motivate the correct behaviour of dogs. **Section 1…Section 5.**

By conducting this research, an automated system for dispensing treats or snacks to the dog before it barks and using the model as a tool to encourage positive behaviour is aimed to be achieved.

### 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\cite {7324337} \cite{uzkent2012non}. 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\cite{su2019environment}. For sound feature extraction, common methods include MFCC, Gammatone, and LBP-HOG features.\cite {8096153}

**- Binary classification of sounds**

There is also extensive published work on the binary classification of sounds, with popular models including Support Vector Machine (SVM)\cite{muhammad2014pathological} and the CNN as mentioned above. In addition, numerous studies on biological characteristics classification have cited the effectiveness of random forest trees\cite{statnikov2008comprehensive} and logistic regression\cite{springer2015logistic} in similar classification tasks.

**- Biological characteristics classification**

**- Previous studies on dog behaviour and barking**

In existing research of dog behaviour, dog barking is not only a reaction to express anger, but also for many reasons. Studies have shown that dogs will bark when faced with Fear, Noise and Separation anxiety, and female, neutered and more fearful dogs significantly responding to noise more strongly. In addition, dog barking has been shown to have a role in communication. When dogs bark in communication and other behaviours, dog barking will be different from other dogs, and it will be affected by human preferences and social environments.

**- 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. 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.

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 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

- 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