KARNATAKA STATE COUNCIL FOR SCIENCE AND TECHNOLOGY

Indian Institute of Science campus, Bengaluru

Telephone: 080 -23600978, 23341652 || Email: spp@kscst.org.in
Website: www.kscst.iisc.ernet.in/spp.html or www.kscst.org.in/spp.html

1.	**Project Reference Number: 46S_BE_5281** **Name of the College: Siddaganga Institute of Technology**
2.	**Title of the project: Noise Removal, Feature Extraction & Classification of Environmental Sound**
3.	**Name of the College: Siddaganga Institute of Technology** **Department: Computer Science & Engineering**
4.	**Name of project guide:** **Name: Dr. Shobha K** **Email id:** shobhak@sit.ac.in **Contact No:** 7760434039 **Names of the students:** **Name: Chandana P** **Email id:** 1si19cs032@sit.ac.in **Contact No:** 9686046129 **Name: Ganya M** **Email Id:** 1si19cs041@sit.ac.in **Contact No:** 8971777650 **Name: Jyothishree** **Email id:** 1si19cs056@sit.ac.in **Contact No:** 9480506814
5.	**Keywords:** Convolutional Neural Network Denoising algorithm Mel-frequency cepstral coefficients (MFCCs) Feature Extraction

6.	**Objectives:** Noise reduction is important to preserve the details of the desired sound. Considering the importance of this the goal of this project is to design an accurate model that can classify environmental sounds which involves following objectives: 1. To develop a web app which reduce and classifies various environmental sounds during online meetings to reduce background noise. 2. To improve speakers voice quality during online meetings. 3. Valuable insights and information about various environmental sounds will be provided to the user.
7.	**Introduction:** Sound classification is a widely researched area with many practical applications, such as speech recognition, audio indexing, and audio content analysis. By automatically categorizing sounds into predefined classes, sound classification can improve the efficiency and performance of various sound processing systems and lead to new insights and understanding of the sound signals. Feature extraction for noise removal is a process in signal processing where relevant features are extracted from noisy signals to improve the quality of the signals. The goal of feature extraction for noise removal is to identify the most important and discriminating features of the signals that are not affected by the noise, and then use these features to remove the noise. Constant exposure to background noise can be mentally taxing and increase stress levels during meetings. Removing background noise creates a more pleasant and comfortable listening experience. When participants can clearly hear and understand each other, they are more likely to actively engage in the meeting and contribute their ideas. It creates a conducive environment for meaningful discussions, enhances productivity, and reflects a higher level of professionalism. Sound classification is the process of assigning a class label to a given sound signal based on its properties and characteristics. The goal of sound classification is to automatically categorize sounds into predefined classes, such as speech, music, or noise. There are various methods for sound classification, including traditional machine learning algorithms, such as decision trees and support vector machines, and deep learning algorithms, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs). Another popular method is feature extraction, where relevant features are extracted from the sound signals and then used as input to a machine learning algorithm.

8. **Methodology:**

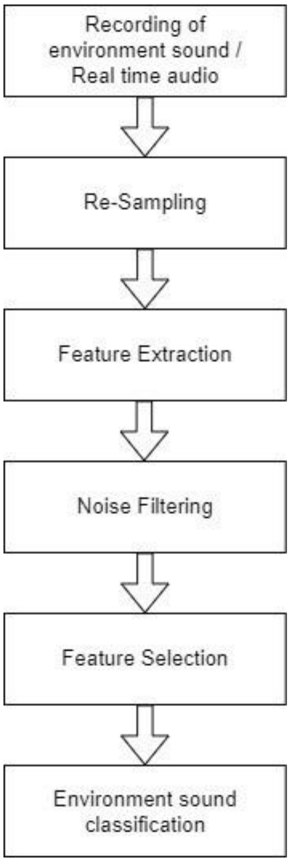

Fig 1.1. Sound Classification

Resampling: The audio signal is resampled to a desired sampling rate. Resampling involves changing the number of samples per second in the audio signal. Increasing the sampling rate can provide more detail but also requires more computational resources. Decreasing the sampling rate can reduce computational requirements but may result in a loss of detail. The appropriate sampling rate depends on the specific application and the characteristics of the audio data.

Feature Extraction: Once the audio signal has been resampled, features are extracted to represent the sound and discriminate between the signal of interest and background noise. Commonly used features for audio classification include Mel-frequency cepstral coefficients (MFCCs), spectrograms, or other time-frequency representations. These features capture different aspects of the audio signal such as pitch, timbre, and energy distribution.

Noise Filtering: Enable these options in your conferencing software settings to help filter out background noise during the meeting.

Feature selection is a crucial step in sound classification, as it determines which features of the sound signals will be used to train the classifier and make predictions. The goal of feature selection is to identify the most relevant and informative features of the sound signals that can effectively discriminate between different sound classes. In sound classification, feature selection is particularly important for improving the performance and accuracy of the classifier.

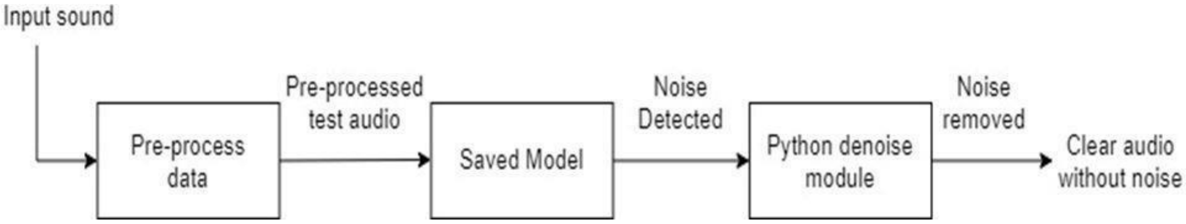

Fig 1.2. Noise Removal

Figure 1.2 depicts the detailed schematic representation of the activities performed in the process of noise removal. Preprocessing has been done and audio is saved in saved model.If any noise detected, python denoise module is used to remove noise. After that clearaudio without noise is produced.

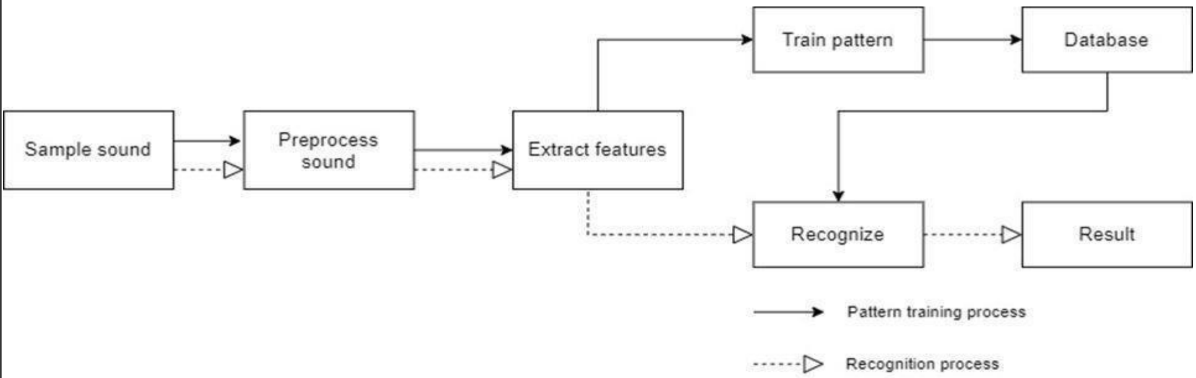

Fig 1.3. Feature Extraction

Figure 1.3 depicts the detailed schematic representation of the activities performed in the process of feature extraction. When the sample sound is given from the dataset it preprocessand extract the relevant features , train the pattern to store it in database. In the recognitionprocess , it will fetch the train pattern from the database and recognize to give correct result. The extracted features are fed into a classification algorithm, such as neural network or asupport vector machine (SVM), which learns to distinguish between different sound classes.The classifier is trained on labeled data, where each sound sample is associated with a knownclass label. The training process enables the classifier to learn the discriminative patterns inthe features that correspond to different sound categories.

9.	**Results and Conclusion:**

The classification of sounds during online meetings would be a system that can automatically identify and distinguish between different types of sounds present in the background of audio recordings during online meetings, such as children playing, siren sounds, dog barking, doorbell ringing, car horn and so on.

This system would be to provide a more immersive and distraction-free online meeting experience, by reducing the impact of background noise on the audio quality and allowing participants to focus on the conversation.

Graph of Clean audio v/s Noisy audio

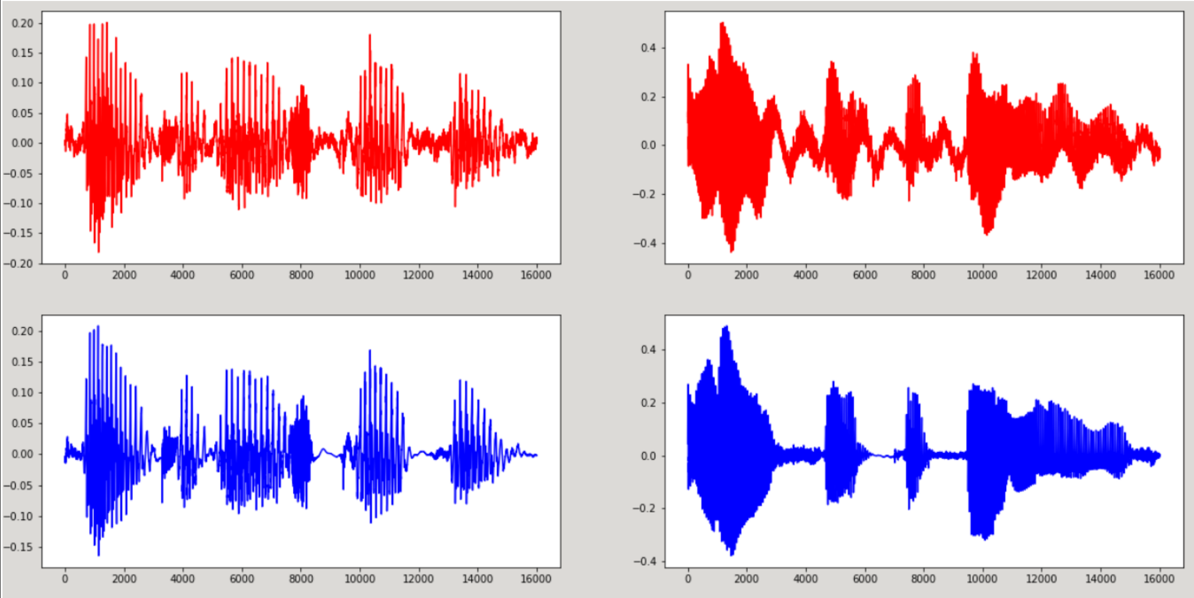

Fig 2.1

In fig 2.1, the first subplot in the first row contains a plot of the 10th signal from the noisy_train dataset and the 10th signal from the clean_train dataset. The second subplot in the first row contains a plot of the 100th signal from the noisy_train dataset and the 100th signal from the clean_train dataset. The third subplot in the second row contains a plot of the 1000th signal from the noisy_train dataset and the 1000th signal from the clean_train dataset. The fourth subplot in the second row contains a plot of the 10000th signal from the noisy_train dataset and the 10000th signal from the clean_train dataset.

Graph of Learning rate v/s Loss

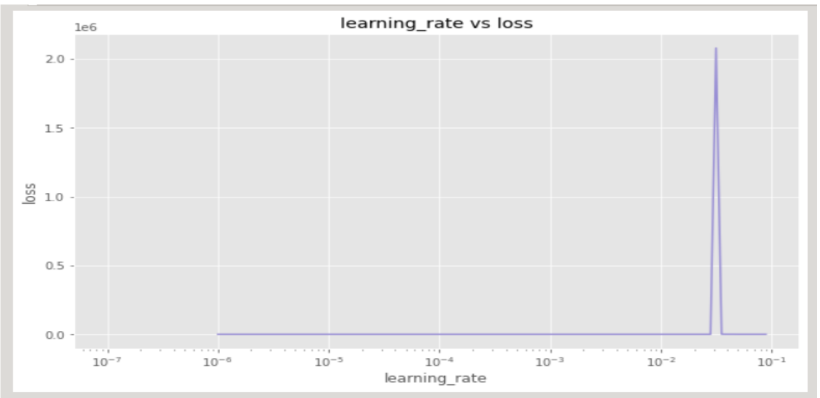

Fig 2.2

In fig 2.2, the graph shows that loss of audio signal is minimum in every stage of learning rate and increases at a particular high learning rate.

When compared with RNN it provides an accuracy of about 70% while CNN provides approximately 60% accuracy. But RNN supports noise removal effectively and CNN supports Classification of sound. Overall after pre-processing the dataset the comparision of CNN and RNN shows that RNN has a high impact in correctly classifying or removing noise from the audio samples.

The aim of this project is to make environmental sound classification in the most efficient way. It is aimed to find the combination of the most successful feature extraction for the most efficient environmental sounds. In this, various learning techniques are applied to the features which are obtained after time-frequency measured feature extraction techniques andfor using different feature extraction techniques and learning techniques together various combinations are obtained to find best sound classification system. Comparisons are obtained from the test data result and in this way, the best combinations are determined. Further elaboration on these techniques can be achieved by obtaining better classification rates. The field of deep learning, which is constantly evolving, enables the testing of new types of feature extraction and classification techniques, and paves the way for environmental sound classification studies.

10.	**Scope and future work:** The project could evolve to offer personalized audio settings based on individual preferences. By analyzing user's audio patterns and preferences, the system could automatically adjust audio parameters like volume, equalization, or background noise levels to provide an optimized audio experience for each participant. This could enhance user satisfaction and reduce the need for manual adjustments. Future advancements in the project could focus on developing intelligent noise suppression techniques. These techniques could automatically detect and suppress specific types ofnoises, such as keyboard typing, mouse clicks, sirens, or construction sounds. By selectively removing only the unwanted noise sources while preserving the speaker's voice, the system could significantly improve audio quality. Another future scope for the project is integrating the noise removal and sound classification technology directly into existing online meeting platforms and communication tools. This would eliminate the need for separate noise removal software and enable seamless integration with popular platforms like Zoom, Microsoft Teams, or Google Meet. Integrationcould be achieved through APIs or plugins, making the technology easily accessible to a wide range of users.