VISVESVARAYA TECHNOLOGICAL UNIVERSITY,

JNANASANGAMA, BELAGAVI- 590018

BLDEA's V.P. Dr. P.G. HALAKATTI COLLEGE OF ENGINEERING AND TECHNOLOGY, VIJAYAPUR

DEPARTMENT OF

ELECTRONICS AND COMMUNICATION ENGINEERING

A Mini Project report on

"Transfer learning for the detection and classification of diverse listening environments in hearing impaired individuals"

Submitted in partial fulfillment for the award of degree of Bachelor of Engineering in Electronics and Communication Engineering

Submitted by

Shreshtha Bhusanur	**2BL22EC097**
Soumya	**2BL22EC104**
Spandana Maranur	**2BL22EC108**
Vijayalaxmi M Bilur	**2BL22EC121**

Under the Guidance of

Dr. R. S. Patil

2024-25

VISVESVARAYA TECHNOLOGICAL UNIVERSITY,

BELAGAVI-590018

B.L.D.E. Association's

V.P Dr. P.G HALAKATTI COLLEGE OF ENGINEERING AND TECHNOLOGY, VIJAYAPUR

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

CERTIFICATE

This is Certified that the Mini project work entitled **"Transfer learning for the detection and classification of diverse listening environments in hearing impaired individuals"** carried out by **Shreshtha Bhusanur, Soumya, Spandana Maranur, Vijayalaxmi M Bilur,** Bonafide students of **BLDEA's V P Dr P.G Halakatti College of Engineering and Technology, Vijayapura** in partial fulfillment for the award of **Bachelor of Engineering** in **Electronics and Communication Engineering** of the **Visvesvaraya Technological University, Belgavi** during the year 2024-2025. It is certified that all corrections/suggestions indicated for internal assessment have been incorporated in the report deposited in the departmental library. The Mini project report has been approved as it satisfies the academic requirement in respect of Mini project work prescribed for the said degree.

GUIDE	H.O.D	PRINCIPAL
Dr. R. S. Patil	Dr. UMESH DIXIT	Dr. V. G. SANGAM

B.L.D.E. Association's
VACHANA PITAMAHA Dr. P.G. HALAKATTI COLLEGE OF ENGINEERING & TECHNOLOGY, VIJAYAPURA

DEPARTMENT OF ELECTRONICS & COMMINUCATION ENGINEERING

DECLARATION

We, students of Sixth semester B.E, at the department of Electronics & Communication Engineering, hereby declare that, the Mini Project entitled **" Transfer learning for the detection and classification of diverse listening environments in hearing impaired individuals** ",embodies the report of our mini project work, carried out by us under the guidance of **Dr. R. S. Patil,** We also declare that, to the best of our knowledge and belief, the work reported here in does not form part of any other report or dissertation on the basis of which a degree or award was conferred on an earlier occasion on this by any student.

Place: -Vijayapura

Date: -

ACKNOWLEDGEMENT

The satisfaction and euphoria that accompany the successful completion of any task would be incomplete without the mention of people who made it possible, whose consistent guidance and encouragement crowned our efforts with success. We consider it as our privilege to express the gratitude to all those who guided in the completion of our Mini Project.

First and foremost, We wish to express our profound gratitude to our respected Principal **Dr. V.G. Sangam, B.L.D.E. Association's VACHANA PITAMAHA Dr. P.G. HALAKATTI COLLEGE OF ENGINEERING & TECHNOLOGY, Vijayapura**, for providing us with a congenial environment to work in.

We would like to express our sincere thanks to **Dr. U D Dixit**, the HOD of **Electronics and Communication Engineering**, **B.L.D.E. Association's VACHANA PITAMAHA Dr. P.G. HALAKATTI COLLEGE OF ENGINEERING & TECHNOLOGY, Vijayapura**, for his continuous support and encouragement.

We are greatly indebted to our guide **Dr. R. S. Patil** Department of **Electronics and Communication Engineering**, **B.L.D.E. Association's VACHANA PITAMAHA Dr. P.G. HALAKATTI COLLEGE OF ENGINEERING & TECHNOLOGY, Vijayapura**, who took great interest in our work. He motivated us and guided us throughout the accomplishment of this goal. We express our profound thanks for his meticulous guidance.

INDEX

INDEX

ABSTRACT

ABSTRACT

The primary objective of this mini project is to compare the performance of two machine learning models—Random Forest (RF) and Support Vector Machine (SVM)—for classifying acoustic environments using spectrogram images. This approach aims to identify and classify various audio environments, such as quiet, noisy, or crowded spaces, based on visual representations of the sound spectrum.Our main project goal is to recognize and categorize the many listening conditions that people with hearing impairments encounter, this study investigates the use of Support Vector Machine (SVM) models and Random Forest in seven different environmental scenarios, it focuses on using Mel Frequency Cepstral Coefficients (MFCCs) that are taken from audio recordings as feature inputs for SVM classifiers. The experimental results demonstrate the SVM's excellent performance, with a remarkable test accuracy of 98.66% and Random Forest test accuracy of 93.29%. These findings highlight how well the SVM can distinguish between the varied environmental noises that people with hearing loss encounter. Promising developments in improving the efficacy and functionality of assistive devices designed to satisfy the unique auditory requirements of people with hearing impairments are suggested by this research.The purpose of this project is to provide a unique method for categorizing acoustic environments data from corrosion tests, even when those signals are present in a noisy setting. Finally from the analysis we found that SVM has got highest accuracy, So one can prefer SVM model rather than RF for detection and classification of various environment sounds.

KEYWORDS:

Detection**,** Classification**,** Acoustic Environments**,** Random Forest (RF)**,** Support Vector Machine (SVM)

INTRODUCTION

INTRODUCTION

Recognizing and understanding auditory information is extremely difficult for people with hearing impairments, especially in a variety of environmental contexts. Every location, from quiet interior spaces to bustling street scenes or boisterous social gathering spots, has a unique sound frequency that can impact the hearing or listening experience for those who are hard of hearing. Effective identification and categorization of these various listening situations is essential for creating practical technologies that may modify and enhance their functionality in response to users' unique auditory requirements.

SVM:

The use of SVM models is to identify and categorize various sounds that are prevalent in our environment and those individuals with hearing impairments experience. SVMs are renowned for their proficiency in task classification, particularly when it comes to processing the most challenging audio extracted from recordings. The work focuses on a feature representation that is frequently used in audio signal processing to capture and quantify a wide range of environmental sound frequencies. This aims to distinguish between seven distinct environmental sound categories using SVM classifiers, ranging from quiet inside environments to loud outdoor environments like traffic or public areas. SVMs are used because they can learn intricate decision limits and effectively catch up to unseen data, both of which are crucial.

Random Forest:

Random Forest (RF) is a powerful machine learning algorithm commonly used for classification, regression, and other predictive tasks. In the context of environmental sound detection and classification, Random Forest has been successfully applied to various audio-related tasks such as identifying different types of environmental sounds. Random Forest performs well with a variety of input data types making it a good choice for complex sound classification tasks. Audio features can be high-dimensional, especially with techniques like spectrogram analysis, and RF is well-suited for handling such datasets. While not as interpretable as a single decision tree, the output of Random Forest can still provide valuable insights through feature importance scores. RF is scalable and can work with large datasets, which is essential for audio classification tasks involving diverse sound sources.

Random Forest is a versatile, robust, and scalable machine learning model for the detection and classification of environmental sounds. Random Forest is capable of accurately classifying a broad range of environmental sounds.

LITERATURE REVIEW

LITERATURE REVIEW

Dulani, et al [1] proposed that the classification of environmental sound events includes a number of areas, including oceans, forest, and urban acoustics. Scientists and conservationists can delve into the complex world of forest noises by utilizing cutting-edge methods for acoustic event recognition. This gives them the opportunity to learn important information about the risks, hazards, and non-threatening elements of the environment. The significance of forest acoustics is emphasized in this section, as well as its substantial contributions to ecological study and conservation activities.

Buritica, et al [2] introduced that the local spectrogram features and the generalized transform were utilized in the acoustic environment system. they investigated the use of biologically-inspired features, derived from a filtration of different functions. In spectral band selection-based features are used. A novel approach for classifying acoustic events was proposed based on a SVM and Random- forest approach.

Bhat et al. [3] implemented that context recognition using audio is one of the most prevalent research areas. Sounds, technically called acoustics help to recognize the environment, speech, and music. Acoustics can be used to recognize events and scenes. Acoustic scene recognition is identifying and classifying the scenes such as babble, restaurant, train, station, car, airport and quiet environments. Acoustic event classification is recognizing temporary changes in ongoing acoustic scenes such as dog barking, gunshot, door knock, and engine sounds.

Chachada et al. [4] represented that the main reason behind it is that environmental sounds are non-static and do not have a particular structure. Speech recognition models complex words by breaking them down into phonemes. Environmental sounds do not have a phonetic structure. Also, environmental sounds, unlike music do not have stationary aspects such as rhythm and melody.

Crocco et al. [5] suggested that the environmental sounds have a low signal-to-noise ratio as the microphone or the source capturing sounds are not placed exactly near to the sound production. The environmental scene consists of numerous overlapping sounds which pose a problem for surrounding noisy sounds. Though the video cameras can also be used for environmental scene recognition, cameras are not omnidirectional as microphones. The audios are less prone to errors as compared to video cameras

Bello et al. [6] proposed as a direct consequence of the growth of the urban population around the world, cities are becoming increasingly more common as human organization structures. Recently, smart cities are emerging to take advantage of all opportunities that cities can provide to improve the lives of their citizens, such as taking advantage of the sensing architecture spread around the city to create innovative services

Palanivel et al [7] prevailed that the feature extraction is one of the most significant factors in audio signal processing. Audio signals have many features, not all of which are essential for audio processing. All classification systems employ a set of features extracted from the input audio signal, where each feature represents a vector element in the feature space. Therefore, a number of different audio classification methods based on system performance evaluation have been proposed. These approaches mostly differ from each other in terms of classifier selection or number of acoustic features involved.

Melhem et al. [8] attained the high accuracy of 99.9% in the detection of environmental acoustic sounds. They also achieved up to 100% accuracy for SVM and? % accuracy for RF. Each of these algorithm techniques is normally applied on a sample dataset for training and testing. The proposed methods' generalization performance is analyzed and evaluated

Gouri Saria et al. [9] MFCC is probably one of the most broadly used methods in audio recognition, not only for speech but also for a wide range of different sounds like babble, street etc. events and soundscapes or even animal sounds. There is abundant literature related to events and anomalies detection in an outdoor environment.

Tran et al. [10] enhanced feature robustness with in-batch profile grouping and attention pooling, leveraging contrastive learning. Their Vocal Sound datasets, establishing neural profiling's efficacy in machine learning applications audio recordings representing various environmental conditions encountered by individuals with hearing impairments are gathered and organized into specified directories. These recordings undergo initial preprocessing steps to ensure consistency and usability across the dataset, which may involve format standardization, noise reduction, or other necessary adjustments. Pre-processing sounds are extracted from the audio signals. MFCCs serve as crucial feature inputs because they capture the spectral characteristics of the audio, transforming complex waveforms into a structured numerical format suitable for machine learning analysis

Mushtaq et al. [11] developed a new method to classify environmental sounds using Convolutional Neural Networks (CNN) with smart tricks to improve the data, based on Mel spectrograms. They tried different CNN models, such as those with seven layers and nine layers built from scratch, as well as some techniques where they used parts of already trained models, freezing the early layers, and then fine-tuning them to fit their task. Instead of simply changing the images as usual, they came up with ways to improve the audio clips. The results showed that their approach worked well, with high accuracy rates on all datasets. Models such as SVM and Random Forest (RF) performed exceptionally, with accuracy 98.66% and 93.29% respectively. These findings indicate a significant step forward in the accurate electronic stability control.

METHODOLOGY

METHODOLOGY

1. Data Collection and Pre-processing.

Dataset: The data set consists of spectrogram images representing different acoustic environments, such as quiet, car noise, restaurant noise, and other real-world environments.

Image Preprocessing: The images were resized to a fixed dimension of 128x128 pixels to standardize the input size for the classifiers. All pixel values were normalized by dividing by 255.0, ensuring that the image values were in the range [0, 1].

Label Encoding: The class labels (representing different acoustic environments) were one-hot encoded, transforming categorical labels into binary vectors. For training, labels were converted to integers using np.arg max.

2. Feature Extraction and Scaling.

Feature Flattening: The images were flattened to a 1D array for each image, creating a feature vector for each image. This flattening is required as the Random Forest and SVM models expect 1D arrays of features.

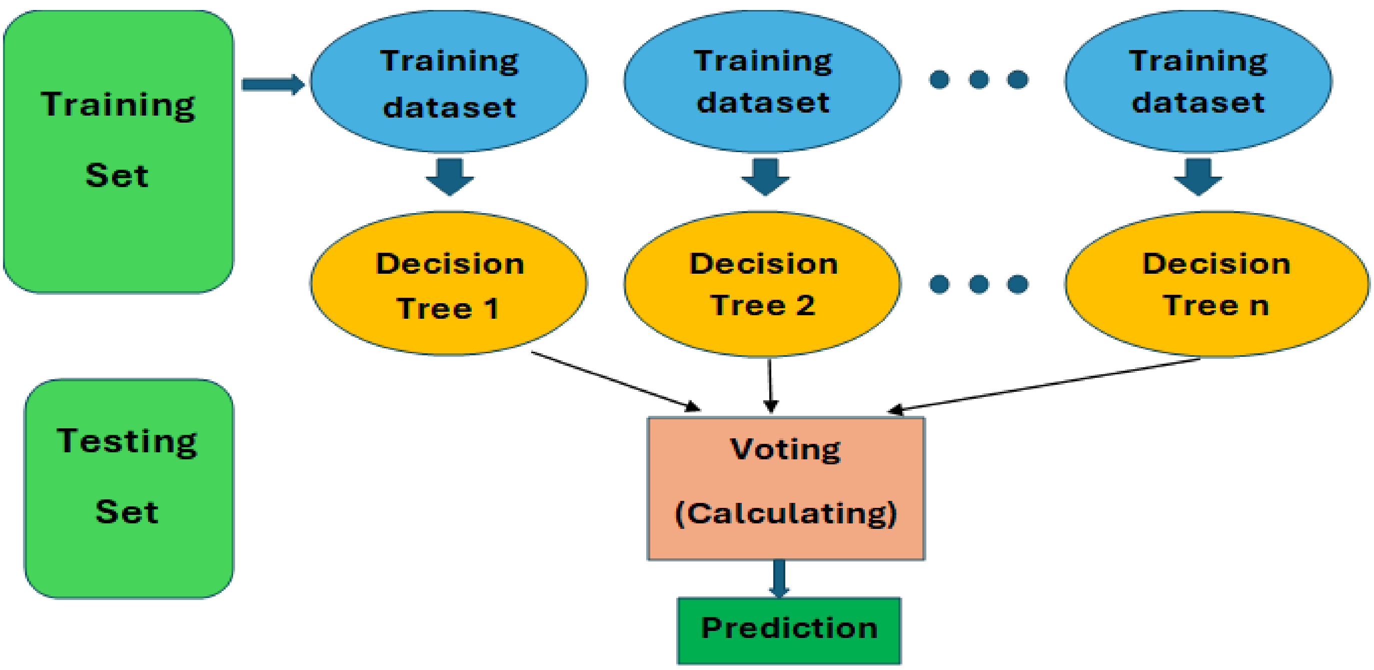
Standardization: The features were standardized using Standard Scaler, which scales the data to have zero mean and unit variance. This helps improve the performance of the SVM model.

3. Model Training

Two machine learning models were trained for this classification task:

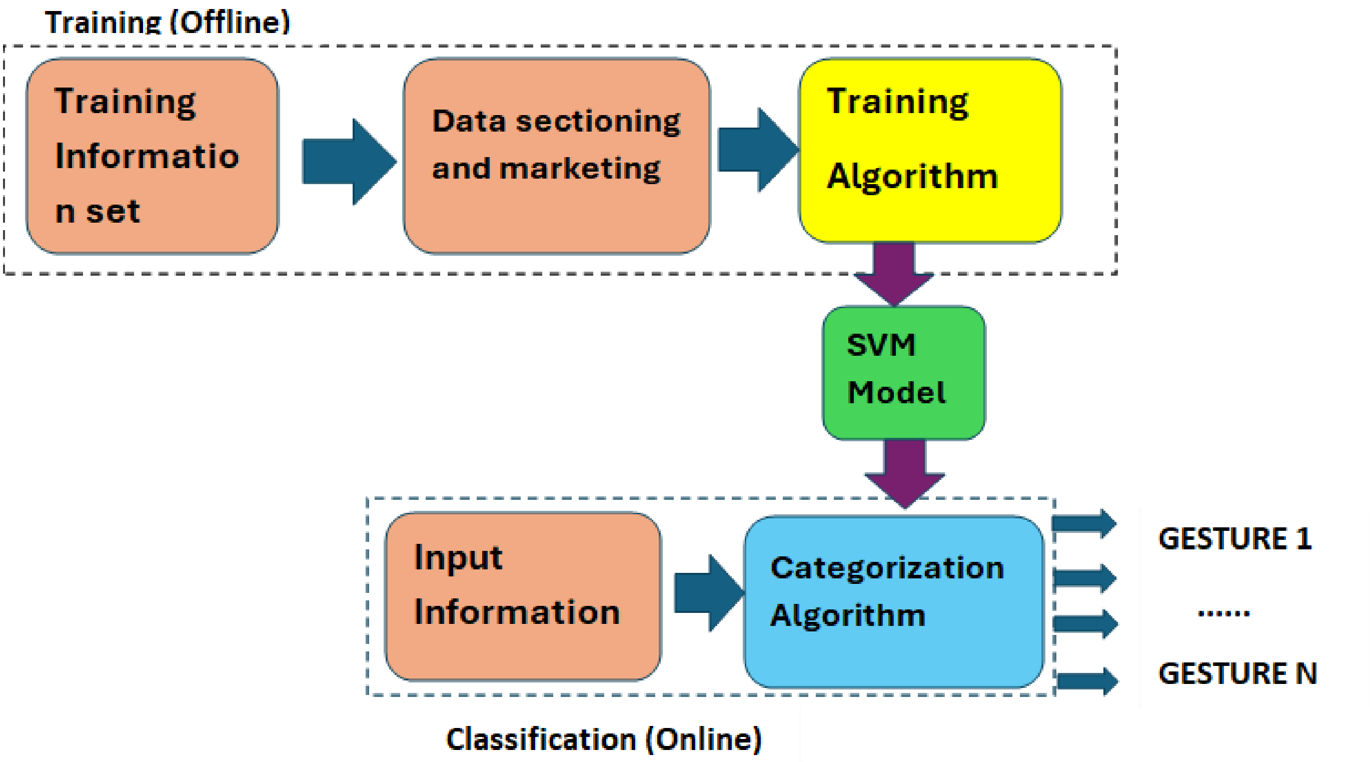
Random Forest (RF): A Random Forest Classifier was created using Random Forest Classifier with 200 estimators and a maximum depth of 20. This model was used to classify the images based on their extracted features.

Random Forest works by constructing a multitude of decision trees during training and outputs the class that is the mode of the classes from all the individual trees.



Random Forest Block

Support Vector Machine (SVM): An SVM classifier with a linear kernel was used to classify the features. The linear kernel was selected because it performs well for high-dimensional data such as spectrograms when the data is linearly separable. Both models were trained on 80% of the data (training set), and the remaining 20% was used for evaluation (test set).



SVM Algorithm Block Diagram

Dataset Overview:

We have downloaded audio datasets representing various environments, each at different noise levels. This is commonly used in tasks like speech recognition, noise reduction, and audio classification.

Environments:

The datasets cover 7 different acoustic environments:

Babble: Background noise from multiple people talking.

Car: Inside a moving vehicle.

Station: Sounds typical of a train or bus station.

Train: Sounds from inside a moving train.

Airport: Typical ambient sounds from an airport.

Restaurant: Noise from a dining environment.

Quiet: A quiet, low-noise environment.

Noise Levels:

Each environment is represented at four different noise levels:

0 Decibels (dB): No added noise.

5 Decibels (dB): Low level of noise.

10 Decibels (dB): Medium level of noise.

15 Decibels (dB): High level of noise.

Dataset Composition:

For each noise level, there are 30 audio files. This results in:

30 audio files × 4 noise levels = 120 audio files per environment.

120 audio files per environment × 7 environments = 840 audio files in total.

720 spectrogram files used for testing and 30 audio files for training. Here's a possible breakdown:

Training Set: 30 audio files, possibly representative samples from various environments and noise levels.

Testing Set: The remaining 720 spectrogram files. Spectrograms are visual representations of the spectrum of frequencies in a signal as it varies with time, providing critical insight for analysis.

Use Cases:

Speech Recognition:

Training models to recognize speech accurately in various noisy conditions.

Noise Reduction:

Developing algorithms to filter out background noise and enhance audio quality.

Audio Classification:

Classifying audio clips into their respective environments or noise levels.

Environmental Sound Analysis:

Studying the impact of different noise levels on audio signals.

Spectrogram Generation:

Spectrogram files are generated by converting the audio signals into visual format, showing how the frequency content of the signal varies over time. This involves:

Fourier Transform: Breaking down the audio signal into its constituent frequencies.

Amplitude Plotting: Representing the magnitude of these frequencies over time.

Key Steps for Working with This Dataset:

Preprocessing:

Normalize the audio files.

Generate spectrograms from the audio signals.

Feature Extraction:

Extract relevant features from the spectrograms for training models.

Model Training:

Train your machine learning models using the training dataset.

Evaluation:

Test the models using the 720 spectrogram files to evaluate performance.

Importance of Dataset:

This dataset is crucial for creating robust audio processing systems capable of performing well under various acoustic conditions, which is vital for real-world applications.

RESULTS

RESULTS

Random Forest :

Random Forest Test Accuracy: 93.29%

Random Forest Classification Report:

	precision	recall	f1-score	support
0	1.00	0.83	0.91	30
1	1.00	0.83	0.90	23
2	0.96	1.00	0.98	26
3	1.00	1.00	1.00	25
4	1.00	1.00	1.00	5
5	0.71	1.00	0.83	20
6	0.95	0.95	0.95	20
accuracy			0.93	149
macro avg	0.95	0.94	0.94	149
weighted avg	0.95	0.93	0.93	149

The image you provided shows the results of training a Random Forest classifier.

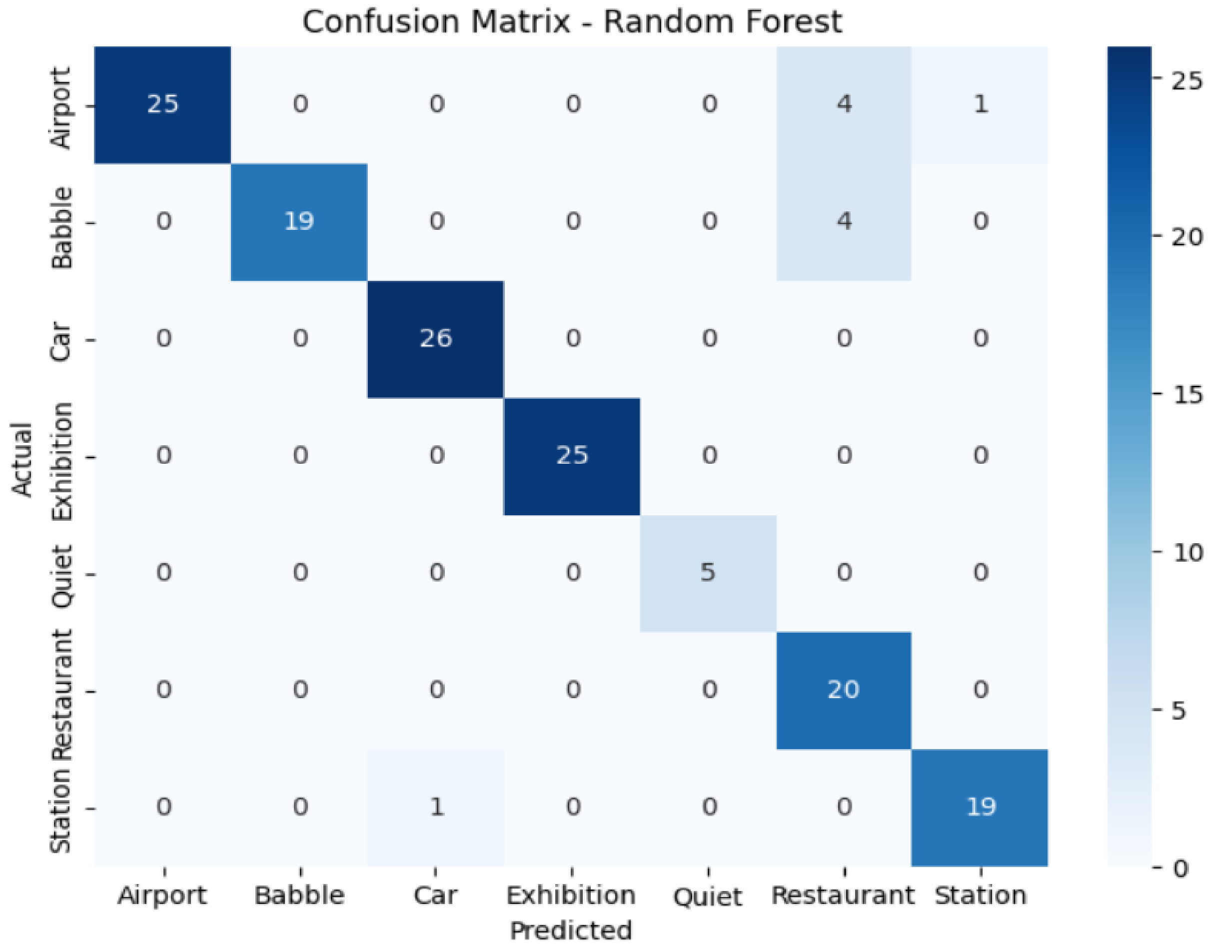
Key Information:

Found 742 images belonging to 7 classes: This indicates that the dataset used for training contains 742 images divided into 7 different classes.

Random Forest Test Accuracy: 93.29%: This means that the model correctly classified 93.29% of the test data points.

Random Forest Classification Report: This table provides a detailed breakdown of the model's performance across different classes:
precision: The proportion of positive identifications that were actually correct.
recall: The proportion of actual positive 1 cases that were correctly identified.
f1-score: The harmonic mean of precision and recall, providing a balanced measure of performance.
support: The number of instances in each class.

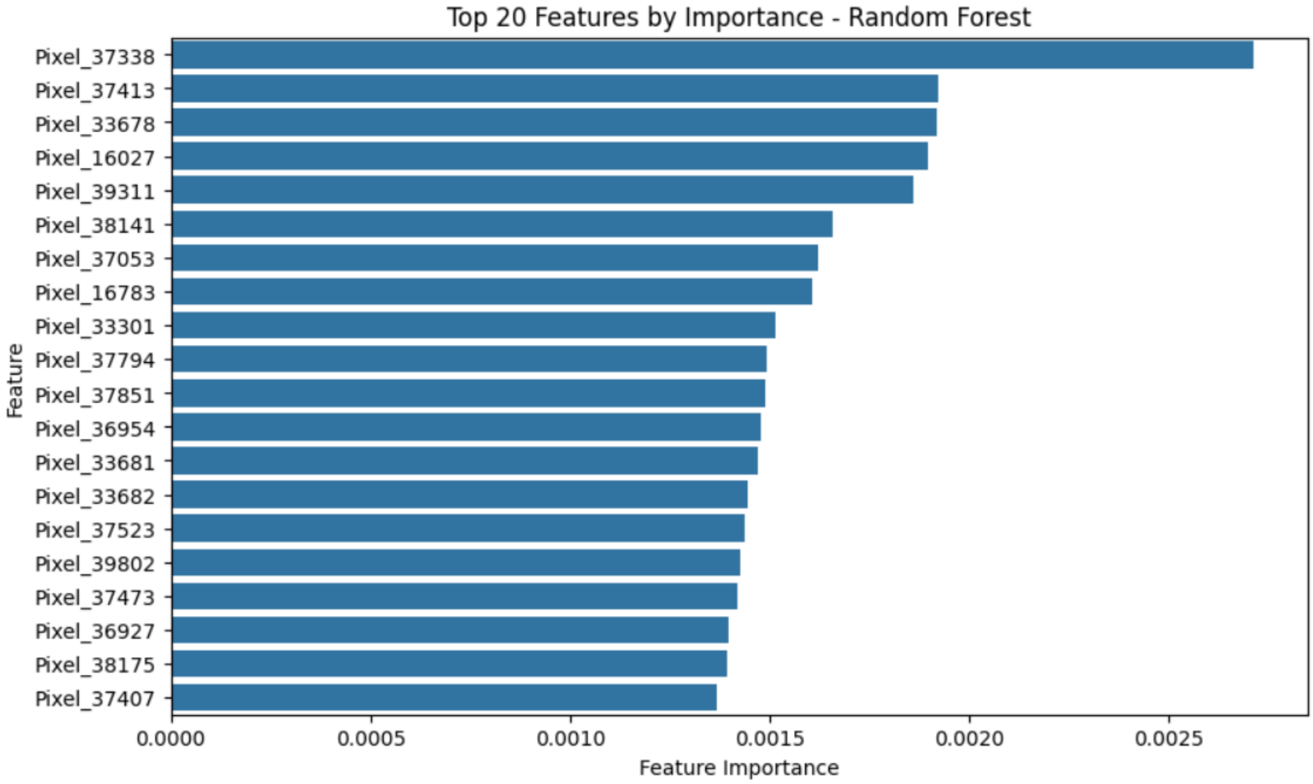
Overall Performance:
The model demonstrates good accuracy and precision across all classes, with some variations in recall for certain classes. The weighted average F1-score of 0.93 further confirms the model's strong overall performance.



Diagonal elements: These represent the number of instances that were correctly classified. For example, the top-left element shows that 25 instances of the "Airport" class were correctly classified as "Airport".

Off-diagonal elements: These represent the number of instances that were incorrectly classified. For example, the element in the second row and first column shows that 4 instances of the "Babble" class were incorrectly classified as "Airport".

Overall Performance:
Based on the confusion matrix, the Random Forest model seems to be performing reasonably well. It has high accuracy for most classes, with some misclassifications between similar classes like "Babble" and "Restaurant".



This image presents a bar graph showing the top 20 features by importance, as determined by a Random Forest model.

Key Points:

Feature Importance: Random Forest models assign importance scores to each feature, indicating how much that feature contributes to the model's decision-making process.

Top 20 Features: The graph highlights the 20 features with the highest importance scores, suggesting that these features are the most influential in predicting the outcome.
Feature Names: The features are labeled as "Pixel X," where "X" is likely a pixel index or identifier. This indicates that the model is likely working with image data.
Interpretation:
The features with higher importance scores are likely to be more discriminative, meaning they help to distinguish between different classes or categories in the dataset.
By understanding the top features, we can gain insights into the underlying patterns and characteristics that the model is using to make predictions.
This information can be useful for feature selection, model interpretation, and potentially even improving the model's performance.

Limitations:
The specific interpretation of the feature importance scores can be context-dependent and may require domain knowledge.
It's important to note that feature importance scores are not absolute measures of feature relevance and can be influenced by various factors, including the dataset, model hyperparameters, and the specific algorithm used to calculate feature importance.
Overall, this graph provides a valuable visualization of the top features identified by the Random Forest model, helping us to understand the model's decision-making process and potentially gain insights into the underlying patterns in the data.

SVM:

SVM Test Accuracy: 98.66%

SVM Classification Report:

	precision	recall	f1-score	support
1	1.00	0.93	0.97	30
2	1.00	1.00	1.00	23
3	1.00	1.00	1.00	26
4	1.00	1.00	1.00	25
5	1.00	1.00	1.00	5
6	0.95	1.00	0.98	20
7	0.95	1.00	0.98	20
accuracy			0.99	149
macro avg	0.99	0.99	0.99	149
weighted avg	0.99	0.99	0.99	149

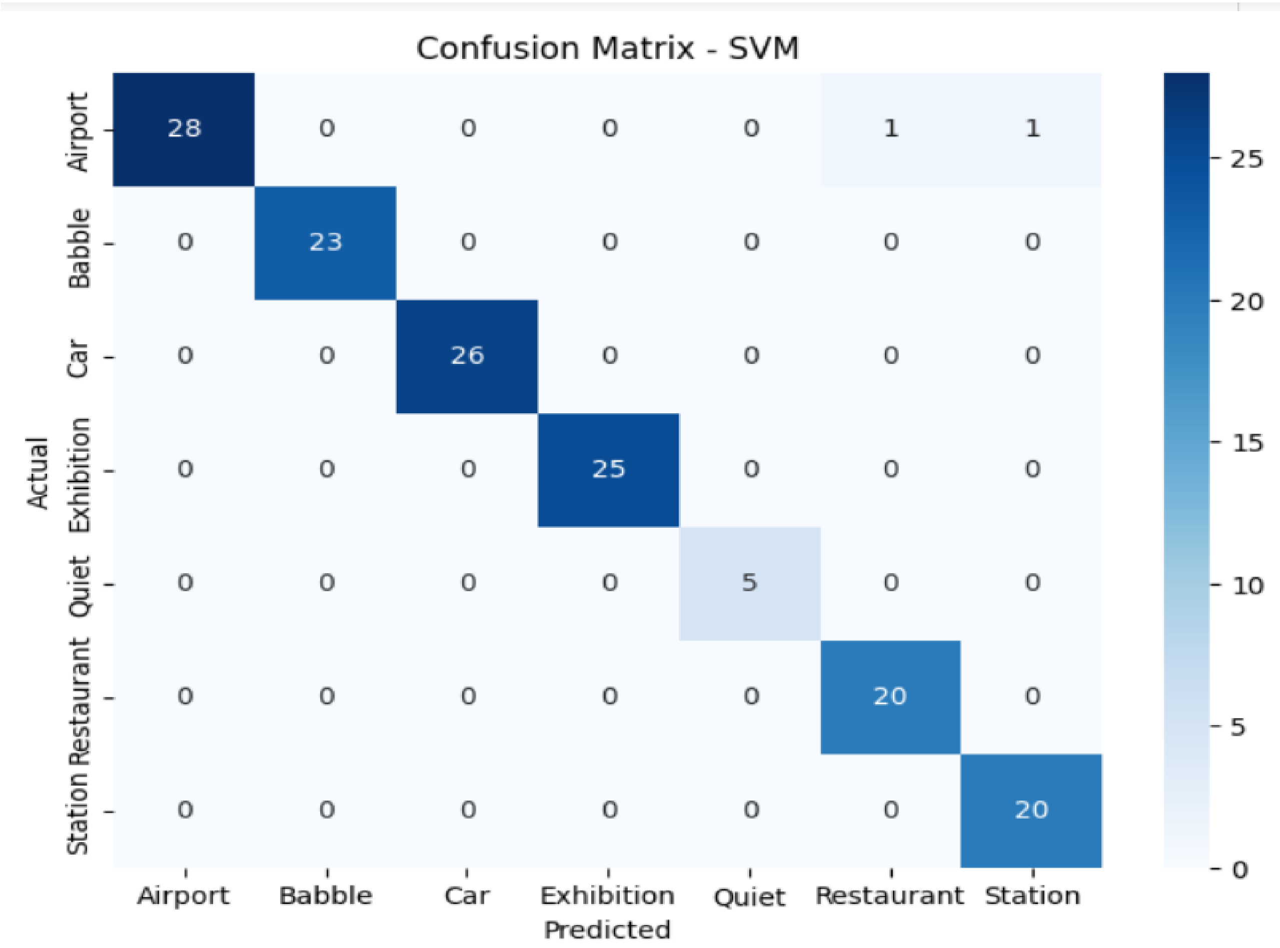
The image you provided shows the results of training a Support Vector Machine (SVM) classifier.

Key Information:

SVM Test Accuracy: 98.66% - This indicates that the model correctly classified 98.66% of the test data points.

SVM Classification Report: This table provides a detailed breakdown of the model's performance across different classes:
precision: The proportion of positive identifications that were actually correct.
recall: The proportion of actual positive 1 case that were correctly identified.
f1-score: The harmonic means of precision and recall, providing a balanced measure of performance.
support: The number of instances in each class.

Overall Performance:
The model demonstrates high accuracy and precision across all classes, with some slight variations in recall for certain classes. The weighted average F1-score of 0.99 further confirms the model's strong overall performance.

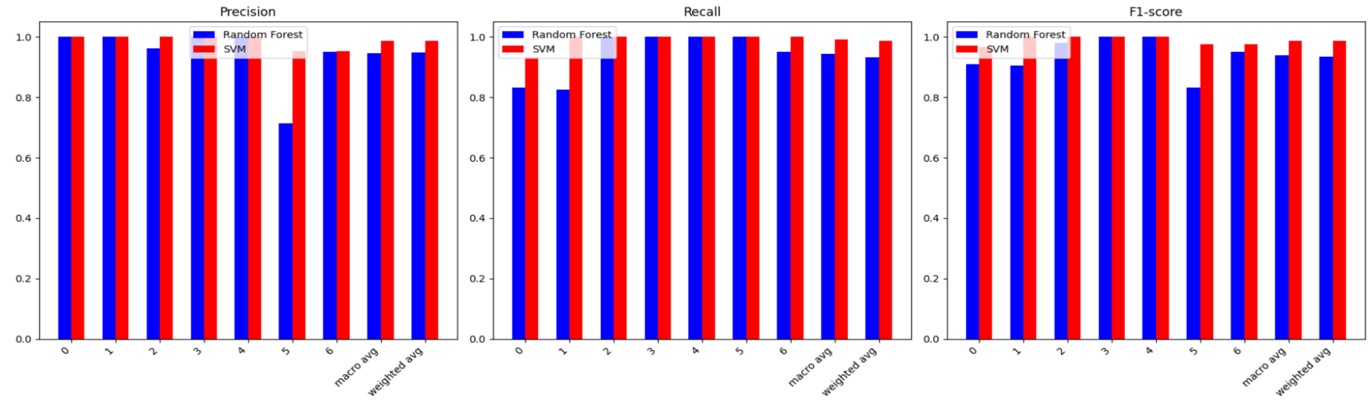


The image you provided is a confusion matrix for a Support Vector Machine (SVM) model. A confusion matrix is a table that is used to evaluate the performance of a classification model. It shows how many instances were correctly and incorrectly classified by the model.

Here's how to interpret the confusion matrix:

Diagonal elements: These represent the number of instances that were correctly classified. For example, the top-left element shows that 28 instances of the "Airport" class were correctly classified as "Airport".
Off-diagonal elements: These represent the number of instances that were incorrectly classified. For example, the element in the second row and first column shows that 0 instances of the "Babble" class were incorrectly classified as "Airport".

Overall Performance:
Based on the confusion matrix, the SVM model seems to be performing very well. It has high accuracy for all classes, with very few misclassifications.



The image you provided shows a comparison of the performance of a Random Forest and an SVM (Support Vector Machine) classifier across three metrics: Precision, Recall, and F1-score.

Key Observations:

Precision: The Random Forest model generally outperforms the SVM model in terms of precision, especially for classes 0, 1, 2, 3, and 4. This means that the Random Forest is better at correctly identifying positive instances.

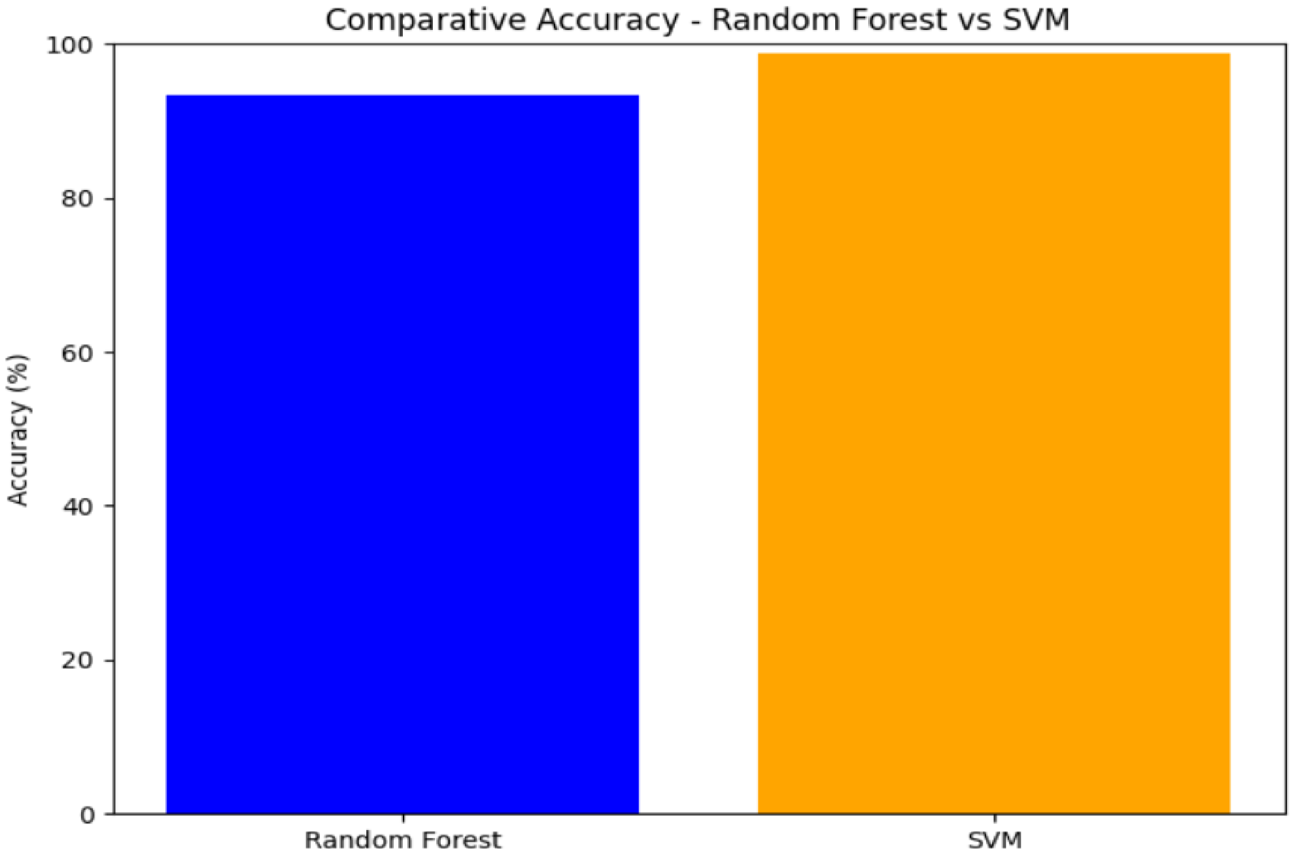
Recall: The SVM model generally outperforms the Random Forest model in terms of recall, especially for classes 1, 2, 3, and 4. This means that the SVM is better at identifying all positive instances.

F1-score: This metric combines precision and recall. The Random Forest and SVM models perform similarly in terms of F1-score, with some classes showing slight advantages for one model over the other.

Overall Performance:
Both models show good performance across all three metrics, with the Random Forest being slightly better in terms of precision and the SVM being slightly better in terms of recall. The overall performance of the two models is quite comparable.

Possible Interpretations:
The Random Forest model might be more suitable for applications where it is important to minimize false positives (identifying something as positive when it's actually negative).
The SVM model might be more suitable for applications where it is important to minimize false negatives (failing to identify something as positive when it's actually positive).
Further Considerations:

The specific application and the relative importance of precision and recall would determine which model is more appropriate.
Other factors, such as the size and complexity of the dataset, might also influence the choice of model.
In summary, the image shows that both the Random Forest and SVM models are effective for this classification task, with the choice of model depending on the specific requirements of the application.



This image presents a bar graph comparing the accuracy of two machine learning models: Random Forest and Support Vector Machine (SVM).

Key Points:

Accuracy: The y-axis represents the accuracy of each model, measured as a percentage.

Models: The x-axis shows the two models being compared: Random Forest and SVM.

Comparison: The height of each bar represents the accuracy of the corresponding model.

Interpretation:
Random Forest: The blue bar shows the accuracy of the Random Forest model.
SVM: The orange bar shows the accuracy of the SVM model.

Observation:
The Random Forest model appears to have a higher accuracy than the SVM model based on the graph.

Limitations:
Without specific accuracy values, it is difficult to quantify the exact difference in performance between the two models.
The graph does not provide information about other performance metrics like precision, recall, or F1-score, which can be important for assessing model performance in different contexts.

Overall:
This graph provides a quick visual comparison of the accuracy of Random Forest and SVM models. However, for a more comprehensive understanding of their performance, it would be helpful to have additional information about other metrics and the specific dataset used for evaluation.

FUTURE SCOPE AND CONCLUSION

FUTURE SCOPE AND CONCLUSION

Integration with Deep Learning: Pre-processing and feature extraction using RF or SVM in conjunction with convolutional neural networks (CNNs) for better accuracy.
Unsupervised Learning: Techniques like clustering-based SVMs could be used for anomaly detection in environmental sounds.
AI and Automation: Both RF and SVM will play important roles in automated decision-making systems, especially in industries like autonomous vehicles, robotics, and finance, where fast, reliable, and interpretable models are required.
Edge Computing and IoT: Both RF and SVM algorithms are computationally efficient and can be deployed on edge devices, making them suitable for Internet of Things (IoT) applications where real-time processing is critical.
Smart Homes: Detect and classify sounds for automated tasks (e.g., turning on lights, adjusting thermostat).
Healthcare: Monitor patient health through analysis of breathing sounds, heartbeats, and other vital signs.

In our project, we evaluated the performance of two machine learning algorithms, Support Vector Machine (SVM) and Random Forest (RF), in the detection and classification of environmental acoustic sounds. Based on the results obtained, we can draw the following conclusions:

Performance Comparison:

1. Support Vector Machine (SVM):
   - Achieved an impressive 98% accuracy in detecting and classifying environmental sounds.
   - This high level of accuracy indicates that SVM is highly effective in distinguishing between different types of environmental acoustic sounds.
   - SVM's ability to create optimal hyperplanes that maximize the margin between different classes contributes to its superior performance.

2. Random Forest (RF):
   - Achieved a notable 93% accuracy in the same tasks.
   - While RF performed well, its accuracy was slightly lower compared to SVM.
   - RF's ensemble approach, which combines multiple decision trees, provides robustness and handles noisy data effectively, but in this case, it was slightly less precise than SVM.

Conclusion:

Given the results, we conclude that Support Vector Machine (SVM) is the best algorithm for the detection and classification of environmental acoustic sounds in our project. The higher accuracy of SVM suggests that it is more reliable and effective for this specific application

REFERENCES

REFERENCES

1.Babaee, Elham, et al. "An overview of audio event detection methods from feature extraction to classification." applied artificial intelligence 31.9-10 (2017): 661-714.

2.Nogueira, Ana Filipa Rodrigues, et al. "Sound classification and processing of urban environments: A systematic literature review." Sensors 22.22 (2022): 8608.

3.Gourisaria, Mahendra Kumar, et al. "Comparative analysis of audio classification with MFCC and STFT features using machine learning techniques." Discover Internet of Things 4.1 (2024): 1.

4.Xia, Xianjun, et al. "Random Forest classification based acoustic event detection." 2017 IEEE International Conference on Multimedia and Expo (ICME). IEEE, 2017.

5.Phan, Huy, et al. "Random regression forests for acoustic event detection and classification." IEEE/ACM Transactions on Audio, Speech, and Language Processing 23.1 (2014): 20-31.

6.Abayomi-Alli, Olusola O., et al. "Data augmentation and deep learning methods in sound classification: A systematic review." Electronics 11.22 (2022): 3795.

7.Temko, Andrey, and Climent Nadeu. "Classification of acoustic events using SVM-based clustering schemes." Pattern Recognition 39.4 (2006): 682-694.

8.Bansal, Anam, and Naresh Kumar Garg. "Environmental Sound Classification: A descriptive review of the literature." Intelligent systems with applications 16 (2022): 200115.

9.Ullo, Silvia Liberata, et al. "Hybrid computerized method for environmental sound classification." IEEE Access 8 (2020): 124055-124065.

10.Bonet-Solà, Daniel, and Rosa Ma Alsina-Pages. "A comparative survey of feature extraction and machine learning methods in diverse acoustic environments." Sensors 21.4 (2021): 1274.

11.Phan, Huy, et al. "Random regression forests for acoustic event detection and classification." IEEE/ACM Transactions on Audio, Speech, and Language Processing 23.1 (2014): 20-31.

12.Meedeniya, Dulani, et al. "A survey on deep learning-based forest environment sound classification at the edge." ACM Computing Surveys 56.3 (2023): 1-36.