Maintenance and Preservation of Mridangam Using Machine Learning

Shrikrishna R, Chanmalla Lohtih, V S Navvin, Pinni Suraj, Dr. Senthilselvi A

Student, Student, Student, Student, Professor
Department of Computer Science and Engineering,
SRM Institute of Science and Technology, Chennai, India
rshrikrishna84@gmail.com， lohithkumar3001@gmail.com， vsnavvin2005@gmail.com, pinnisuraj@gmail.com,
senthila3@srmist.edu.in

Abstract— **Mridangam is a southern Indian percussion instrument. The external factors and the continual use of the instrument call for its preservation and maintenance. Traditional maintenance methods rely heavily on the skills of the individual craftspeople and these methods may be subjective and differ from one person to another. This research utilizes machine learning methods to further enhance the maintenance and preservation of Mridangam. With the use of predictive modeling, the present study proposes a scheme for rational maintenance recommendations. The focus lies on increasing the lifespan of the instrument, enhancing its tonal quality, and maintaining an easy-to-access maintenance method for the musicians and artisans.**

Index Terms— **Mridangam, SVM, MFCC, Bandwidth, Spectral centroid, flatness.**

I. Introduction

Mridangam in below fig.1 is an instrument which is used in South Indian classical forms of music, deeply appreciated for its rich metallic tonal quality and variations in rhythm. The Mridangam is a double-sided drum made of jackfruit wood with animal hide. The percussion instrument is essential not only for Carnatic music performances but also forms another cultural and historical factor in the musical heritage of India. Mridangam maintenance is a tough job requiring the hand of skilled artisans, who work tirelessly to keep its tone and sustained strength.

Fig.1 Mridangam

However, the care of Mridangam is ignored or insufficiently done despite its cultural significance. The drumheads are very responsive to the environment such as temperature and humidity changes, resulting in changes in the amount of tension and quality of sound, and even physical damage. Gradually, the drumheads and shell materials deteriorate with use and need to be retightened, or even replaced. Traditionally, all these maintenance tasks are done through manual inspections and subjective reading, which are time-consuming and inconsistent.

It's a modern world, and that includes technological inventions. In fact, the greatest research development innovation in recent years is machine learning, and most of the recent advancements open up new opportunities for solving even the most sophisticated maintenance problems of the conventional tools. Machine learning has taken up most applications, including instrument classification, raga identification, and acoustic features analysis in music. Such realizations indicate the merit of machine learning concerning its precision, efficiency, and scalability. Through machine learning, prediction modeling and diagnostics can be tailored towards the specific maintenance requirements of the Mridangam.

Machine learning integrated into the maintenance of the Mridangam can change how the instrument will be cared for. The ML models will be able to analyze acoustic signals, monitor environmental conditions, and track usage patterns, thus predicting wear and giving valuable action-oriented maintenance tips. This can ensure high sound quality and instrument longevity. The other aspect would include ML-based tools that can enable this traditional skill to be distributed among musicians across the globe and thereby democratize Mridangam maintenance.

This research utilizes machine learning methodologies to plan maintenance and conservation of the Mridangam. The research would indeed have to solve real-life problems such as wear detection, environmental condition monitoring, and automated tuning assistance. Thus, this research would try to make the heritage of Mridangam more sustainable and efficient by filling the gap between open traditional knowledge and contemporary technology.

II. LITERATURE SURVEY

Chhabra et al., 2020 [1], discuss about the drum instrument classification by a machine learning technique using self-recorded .wav audio files. The authors extract various features from these sound recordings and implement machine learning models to classify, then they achieve the classification of the fundamental components of drums-like bass drums, snares, toms, and cymbals. Their work speaks for effective feature extraction and the usage of ML techniques for automating the auditory tasks.

Srishti Vashishtha et al., 2024 [2] delves into the classification of musical instrument sounds using k-Nearest Neighbor and Convolutional Neural Network models. While the kNN model has a higher accuracy (96%) as compared to CNN (88%), the models still indicate the application of machine learning techniques in the area of audio classification. Applications hence include music production, transcription, and education.

Revathy & Pillai, 2022 [3] carries out the investigation of emotional classification in songs with the use of Russell's Circumplex model. The research implements Support Vector Machine ,Logistic Regression, and Artificial Neural Networks for the classification of emotions based on valence and energy features. The findings showed an almost perfect accuracy (99.9%) using Logistic Regression. The paper concludes on the potential of machine learning in music recommender systems and emotional analysis.

Shashidhar et al., 2023 [4] proposes the use of 1D Convolutional Neural Network for music emotion recognition, utilizing the temporal structures of music signals. The model effectively classifies emotions with high accuracy and outperformed all traditional ML methods. It finds remarkable applications in music therapy, recommendation systems, and education.

Harryanto et al., 2022 [5] uses the transformer architecture has been evaluated for the purpose of music genre classification and across GTZAN dataset. It relies on Mel-Frequency Cepstral Coefficients as features for classification transformer resulted in better accuracy 75.1% when compared to traditional neural networks, which strongly projects the future of advanced architectures in classifying audio related tasks.

This paper by Huaysrijan & Pongpinigpinyo, 2021 [6] studies the Thai classical music instrument recognition by the CNN. It compares the use of MFCC and Mel-Spectrogram as feature extraction methods. It can be observed that the CNN model with MFCC features achieved remarkable accuracy at 99.44%, thus recognizing the importance of conserving cultural heritage through automated recognition via deep learning.

This research by Julian & Mukund, 2024 [7] offers an approach and framework for a Music-to-score Conversion (MSC) approach using pitch estimation and Fourier Transform as the major signal processing techniques. It also exposes the need to explore what potential MSC stands for in music education, production, and accessibility improvement, especially for people with hearing impairments in the conversion of auditory to visual.

In this study by KunWang & Zhao Lv, 2021 [8], exceptionally anomalous sounds in operational conditions of machines are detected with an unsupervised strategy, namely Memory-Amended Deep Autoencoder (MemAE). This model proves superior to classic autoencoder methods in that it reduces missed detections of anomalies and produces a marked performance improvement over different types of machines such as toy conveyor systems and valves.

This research paper by Sungeetha et al., 2024 [9] describes the application of deep learning to the classification of environmental sounds with emphasis on CNN and feature representations such as mel spectrograms. Improvement in the context of noise awareness and surveillance for smart city applications solves the challenges posed by unstructured ambient noise.

This work by Onisha et al., 2024 [10] investigated multi-label sound classification by comparing CNN and LSTM-GRU models for MFCC feature extraction. The CNN model demonstrated formidable accuracy of 97% and temporal dependency capturing would prove inferior for the LSTM-GRU. The advances in audio processing will find applications in automatic organization and sound detection systems.

III. PROPSED METHODOLGY

This study is to build a machine-learning driven model to estimate Mridangam condition just on the basis of the sound generated out of it. It utilizes the most advanced audio analysis and classification techniques to capture a recording of the Mridangam and demonstrates whether it is actually good or needs minor or major repairs or maintenance. In short, the design analyzes the entire process of digital signal processing and supervised machine learning for the instrument to retain quality and duration**.**

Figure.2 which is below shows the flow of evaluating the condition of a mridangam through machine learning. The steps include recording the sound of the mridangam, preprocessing, and feature extraction. These features are used to train the machine learning model and are further tested and validated. Once the model is made robust, it helps to evaluate the instrument condition. Accordingly, the instrument is classified as Good Condition, Needs Tuning, or Needs Maintenance.

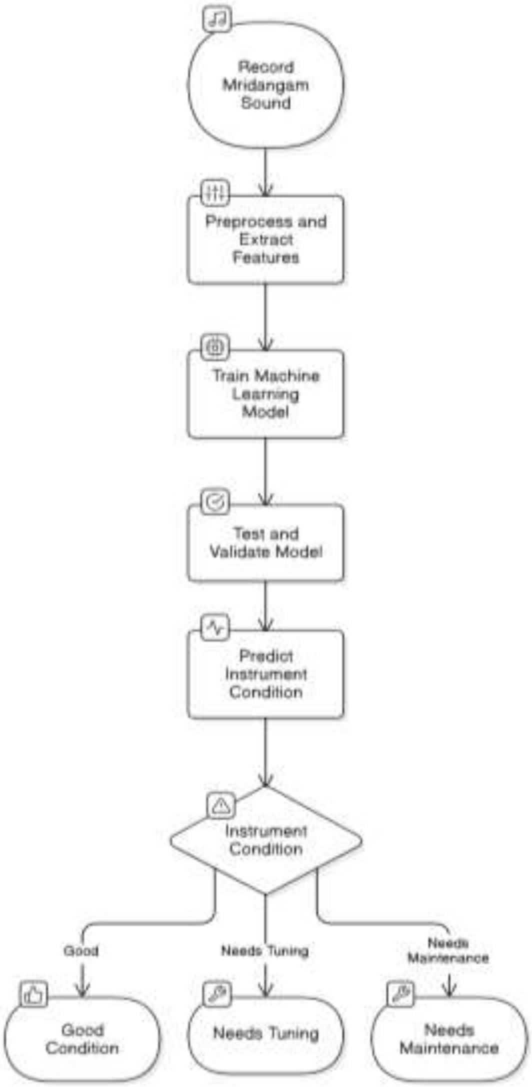

Fig.2 Architecture Diagram

1. Audio Data Collection

Audio data collection is the first step of the process. It entails an organized collection of sound recordings from various Mridangams under diverse situations:

- **Healthy:** Well-tuned condition, no damage or looseness
- **Needs Tuning:** Slightly off-pitch, with inconsistency in tonal response
- **Needs Maintenance:** Damaged skin, weak bass, loose straps, or faded stroke clarity

Recordings comprise strokes such as tha, dhin, nam, chappu, and gumki, which are recorded by experts for the sake of authenticity. Each audio file is tagged with the specific type of condition of the Mridangam, as judged by the expert opinion, at the time of its recording.

2. Preprocessing and Feature Extraction

Before any recordings of the Mridangam get into the machine learning model, we extract salient audio features, which separate tonal and rhythmic elements.

- **Mel-Frequency Cepstral Coefficients (MFCC):** Take account of the sound perceptual properties.
- **Spectral centroid, bandwidth, and flatness:** Measure brightening and balance of that sound.
- **Root Mean Square (RMS):** Represent the sharpness and energy, respectively, of the percussive attack.

The audio sample is segmented into frames, and from each segment, audio features are extracted to construct a detailed profile. The features are then standardized and compiled to constitute a dataset for model training.

3. Training the Machine Learning Models

Classification models are trained on the labeled feature dataset to predict the state of the Mridangam.
The model uses here is:
Support Vector Machine (SVM): It is known to be good when dealing with high-dimensional audio features.

The models are trained on cross-validation procedures such that overfitting is avoided, and the rigor is maintained. The dataset is split into training, validation, and test sets, and the hyperparameters of these models are tuned accordingly to obtain the best performance.

4. Testing and Validation

Training is complete then; testing is performed on a separate dataset consisting of Mridangam recordings to check for accuracy in a real-world setting. Consequently, the performance of the model is assessed based on:

- **Accuracy:** The overall correctness of the model's predictions.
- **Precision and Recall:** This is important for distinguishing borderline or critical cases.
- **Confusion Matrix:** To assess false positives and false negatives for each of the condition classes.

In addition, a blind test is performed, where the system analyzes unknown samples and compares the results with expert evaluations. This step is helpful for testing the practicality and acceptance of conventional knowledge.

5. Deployment and Applications

The final model is packaged into a software or API, where musicians or technicians go through the following steps: Record a stroke or sequence using a phone or microphone. Upload that sound to the system. Get instant feedback as to whether the particular instrument is well-tuned or requires attention. The system reduces dependence on manual tuning checks, preserves the tonal heritage of the Mridangam, and enables a casual user to maintain the instrument fairly well.

IV. RESULTS AND DISCUSSIONS

1. Model Performance

After the dataset is refined and training process of the SVM using the dataset achieves the best performance in the final model. Final model achieves accuracy of 77%, precision of 0.76, recall of 0.65 could be expected from the model in terms of performance metrics, thus demonstrating a reasonable model capability to determine Mridangam condition from audio features.

Table.1 The Comparison of the precision and recall for the audio quality

Audio quality	Precision	Recall
Good	0.81	0.96
Needs_tuning	0.80	0.32
Needs_maintenance	0.66	0.67

In the table.1, the model achieved the highest performance in the 'good' class-with precision values of 0.81 and a very high recall of 0.96-which is an indicator that good quality instruments are identified with high certainty. In contrast, the results for the needs_tuning class were poor, as seen with a very low recall of 0.32, indicating that many such samples were classified wrongly into other classes. The needs_maintenance category was successfully balanced for both metrics-with precision and recall values at 0.66 and 0.67, respectively-as this suggests an intermediate degree of success in classification.

2. Confusion Matrix

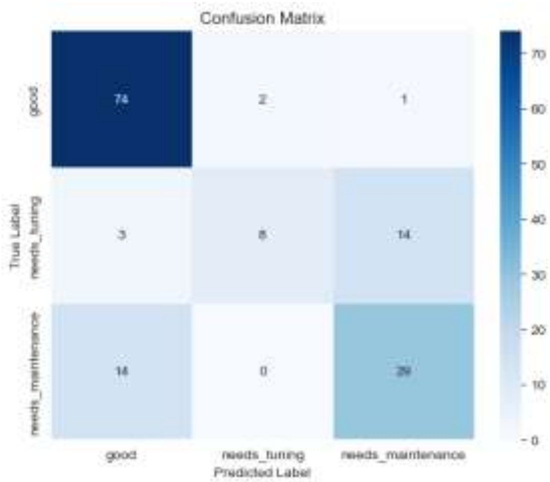

Fig.3 Confusion matrix

Figure.3 shows the detailed confusion matrix with respect to the classifications it obtained.

- A huge proportion of samples classified as good turned out to be actually classified as such (74 out of 77) except for few acoustic exceptions/wrong misclassifications.
- An overlap of acoustic characteristics may cause needs_tuning confused with needs_maintenance (14 of 25).
- The model slightly fails to separate between needs_maintenance and good, and 14 count as falsely classified.

These numbers show how difficult it is to discriminate between small changes in sound signifying that an instrument needs tuning and major maintenance.

3. Real World Application

The results demonstrate that machine learning, specifically SVM-based classifiers, can be effectively used to classify the acoustic condition of traditional Indian percussion instruments like the Mridangam. Such technology could assist artists, students, and instrument makers in assessing instrument condition non-invasively, potentially leading to smartphone applications or studio tools for automated diagnostics.

V. CONCLUSION AND FUTURE RESEARCH

Examine the evaluation condition of Mridangam sounds through machine learning methods using audio features that have been extracted from recorded sounds. The Mel-frequency cepstral coefficients (MFCCs) were applied to train Support Vector

Machines (SVMs), where analysis and classification were performed on three sample classes of Mridangam: good, needing tuning and maintenance. The model achieved an accuracy of 77% and exceptionally developed precision in well-conditioned instruments.

While the model was generally very successful, it showed difficulties in differentiating between more-tuning instruments and the other category of needing-maintenance instruments, as shown by the confusion matrix and recall values. This indicates further enhancing the system's ability to differentiate between subtle acoustic differences in these conditions; however, these findings demonstrate that sound-based analysis can be a potential reliable tool for assessing conditions in instruments.

Future work could include an increase in more heterogeneous samples and clearer playing conditions to enhance model generalization. Combining more audio features for augmenting classification performance or using modern deep learning techniques such as CNNs may also help. This system can eventually undergo development towards real-time mobile or desktop applications usable by musicians and technicians for fast, non-invasive evaluations of traditional percussion instruments.

REFERENCES

[1] Chhabra, A., Veer Singh, A., Srivastava, R., & Mittal, V. (2020). Drum Instrument Classification Using Machine Learning. *Proceedings - IEEE 2020 2nd International Conference on Advances in Computing, Communication Control and Networking, ICACCCN 2020*, 217–221. https://doi.org/10.1109/ICACCCN51052.2020.9362963

[2] Srishti Vashishtha, Rachna Narula, & Poonam Chaudhary. (2024). Classification of Musical Instruments' Sound using kNN and CNN. *2024 11th International Conference on Computing for Sustainable Global Development (INDIACom)*. https://doi.org/978-93-80544-51-9/24

[3] Revathy, V. R., & Pillai, A. S. (2022). Multi-class classification of song emotions using Machine learning. *2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering, ICACITE 2022*, 2317–2322. https://doi.org/10.1109/ICACITE53722.2022.9823535

[4] Shashidhar, R., Balivada, D., Shalini, D. N., Hirebasur Krishnappa, K., & Roopa, M. (2023). Music Emotion Recognition using Convolutional Neural Networks for Regional Languages. *IEEE 1st International Conference on Ambient Intelligence, Knowledge Informatics and Industrial Electronics, AIKIIE 2023*. https://doi.org/10.1109/AIKIIE60097.2023.10390450

[5] Harryanto, A. A. A., Gunawan, K., Nagano, R., & Sutoyo, R. (2022). Music Classification Model Development Based on Audio Recognition using Transformer Model. 2022 3rd International Conference on Artificial Intelligence and Data Sciences: Championing Innovations in Artificial Intelligence and Data Sciences for Sustainable Future, AiDAS 2022 - Proceedings, 258–263. https://doi.org/10.1109/AiDAS56890.2022.9918787

[6] Huaysrijan, A., & Pongpinigpinyo, S. (2021). Deep Convolution Neural Network for Thai Classical Music Instruments Sound Recognition. *ICSEC 2021 - 25th International Computer Science and Engineering Conference*, 283–288. https://doi.org/10.1109/ICSEC53205.2021.9684611

[7] Julian, A., & Mukund, V. (2024). Music to Score Conversion using Machine Learning. *IEEE International Conference on Data Engineering and Communication Systems, ICDECS 2024*. https://doi.org/10.1109/ICDECS59733.2023.10502423

[8] KunWang, & Zhao Lv. (2021). Unsupervised Anomalous Sound Detection For Machine Condition Using Memory-augmented Deep Autoencoder. 2021 17th International Conference on Computational Intelligence and Security (CIS). https://doi.org/10.1109/CIS54983.2021.00061

[9] Sungeetha, D., Nanthini, S., Sivabalan, R., Soundari, A. G., Legapriyadharshini, N., & Rajaram, G. (2024). Deep Learning based Detection and Delineation Sound in Smart City System. 8th International Conference on Electronics, Communication and Aerospace Technology, ICECA 2024 - Proceedings, 870–873. https://doi.org/10.1109/ICECA63461.2024.10801135

[10] Onisha, T. A., Kim, J., & Seol, J. (2024). Multi Label Sound Classification using Deep Learning Models. *2024 IEEE/ACIS 22nd International Conference on Software Engineering Research, Management and Applications, SERA 2024 - Proceedings*, 129–134. https://doi.org/10.1109/SERA61261.2024.10685563