Journal of Integrated
SCIENCE & TECHNOLOGY

# Identification of specific Musical instruments using Machine Learning models

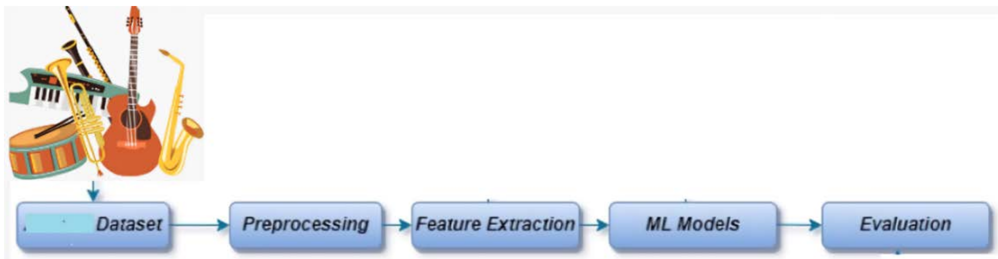
Bhagyalakshmi R[1,2]*, Anandaraju M B[1]

[1]Department of Electronics and Communication Engineering, BGS Institute of Technology, Adichunchanagiri University, B.G. Nagara, Mandya, Karnataka, India. [2]Government Engineering College, Mosalehosahalli, Hassan, Karnataka, India

Submitted on: 22-Nov-2024, Accepted and Published on: 28-Feb-2025

Article

## ABSTRACT




Different applications including recommendation systems and digital audio workstations depend on precise recognition of instruments for their functionality along with music transcription. The research evaluates how well machine learning technologies perform in instrument classification when analyzing audio signal properties. A wide range of instrument samples from the dataset undergo processing for features that include both Mel-spectrograms and the analysis of contrasting elements. The research evaluates the performance of Support Vector Machines (SVM), Random Forest (RF), along with Convolutional Neural Networks (CNN) as models for classification purposes. Experimentally CNNs demonstrated superior performance when compared to mainstream machine learning strategies because they retain both spatial and temporal features in audio information thus achieving better classification results. Performance of the model increases to greater degrees through data augmentation along with proper tuning of model hyperparameters. The research demonstrates how machine learning techniques can identify musical instruments thus opening new possibilities for automatic music analysis systems and instant instrument detection capabilities.

*Keywords: Music Instrument, Recognition, Accuracy. Machine Learning Models*

## INTRODUCTION

Different domains such as music transcription along with recommendation systems and digital audio workstations and automatic music analysis heavily depend on musical instrument identification. Different audio recording elements present multiple challenges for instrument identification because timbre differentiates from pitch and the way instruments play and the way they are recorded. The current instrument classification methods base their analysis on manual feature extraction and heuristic methods yet prove ineffective at handling different datasets[1].

Modern machine learning technology has increased interest in automatic instrument recognition systems during the last few years. Audio characteristics flow through machine learning models which detects different instruments through the analysis of frequency patterns and amplitude and time-based patterns. The processing of audio spectrograms along with spatial-temporal dependency detection proves best through Convolutional Neural Networks (CNNs) as deep learning strategies[2].

The synchronized combination of numerous instruments at an orchestra concert produces an enchanting experience that engages concertgoers. Non-musicians find it difficult to identify different musical instruments at ordinary concerts. The research develops an automatic instrument detection tool through machine learning technology for music appreciation education with classical music as the main focus. The research connects music studies with machine learning methods to develop a comprehensive instrument detection framework and acoustic sound analysis technique. Several practical uses exist for identifying musical instruments in musical recordings. Scientists can use instrument acoustic wave analyses to achieve precise instrument classification and differentiation. Users benefit from this capability to locate musical pieces containing specific instruments which also enables the foundation for music creation applications.

A research examines how machine learning techniques classify different musical instruments by analyzing their audio characteristics. The research establishes a comparison of instrument identification results between Support Vector Machines (SVM), Random Forest (RF), and CNNs to determine their

*Corresponding Author: Bhagyalakshmi R, Research Scholar, Department of Electronics and Communication Engineering, BGS Institute of Technology, Adichunchanagiri University, B.G. Nagara, Mandya, Governament Engineering College, Mosalehosahalli, Hassan, Karnataka, India Email: bhagya109@gmail.com

Cite as: *J. Integr. Sci. Technol., 2025, 13(5), 1108.*
*URN:NBN:sciencein.jist.2025.v13.1108*
*DOI:10.62110/sciencein.jist.2025.v13.1108*

classification performance. The collection contains musical instrument data samples alongside extracted Mel-spectrograms with contrast attributes for modeling purposes. The model implements data enhancement protocols and hyperparameter optimization methods to improve its operational capability.

The primary objective of this research is to assess the feasibility of machine learning-based instrument classification and identify the most effective approach for accurate recognition. The findings contribute to the advancement of automated music processing, enabling more efficient real-time instrument identification and improving various music-related applications.

## LITERATURE SURVEY

Joanikij Chulev et al. demonstrates how Naive Bayes and Support Vector Machines and Convolutional Neural Networks as deep learning models help improve musical instrument identification through evaluation on the NSynth dataset.[3] Yao et al. demonstrates that using convolutional neural networks (CNN) and artificial neural networks (ANN) together with mel-frequency cepstral coefficients (MFCC) and attention mechanisms enables musical instrument recognition through machine learning and deep learning models.[4] Rujia Chen et al. studies musical instrument recognition through CNN models while analyzing the obstacles that appear during instrument identification from recorded audio data. It employs multi-spectrogram heatmap analysis to interpret the models, contributing to advancements in music information retrieval (MIR)[5]. Peter Tiemeijer et al. explores using Deep Convolutional Neural Networks for identifying musical instruments in polyphonic music. It enhances existing models to focus on playing style, achieving a 20% accuracy improvement in instrument recognition while utilizing fewer network parameters[6]. Gst. Ayu Vida Mastrika Giri et al. employs Convolutional Neural Networks (CNN) to classify musical instruments, specifically piano, violin, drums, and guitar, using audio features extracted from recordings. The study demonstrates improved classification accuracy with a combined feature set compared to spectral or non-spectral features alone.[7] E.L. Ding et al. focuses on using a K-Nearest-Neighbors algorithm for identifying viola, piano, and ukulele from audio recordings. It suggests that employing larger datasets and convolutional neural networks could enhance classification accuracy beyond the achieved 80%[8]. Tasnim Akter Onisha et al. focuses on identifying musical instruments using Convolutional Neural Networks (CNN) and LSTM-GRU models, achieving 97% accuracy with CNN and 80% with LSTM-GRU, utilizing Mel-frequency cepstral coefficients (MFCCs) for audio preprocessing across nineteen instrument classes[9]. Rujia Chen et al. proposes a hierarchical residual attention network that utilizes a scaled combination of multiple spectrograms (STFT, Log-Mel, MFCC, CST) to enhance musical instrument recognition, demonstrating improved classification accuracy on the OpenMIC-2018 dataset through advanced attention mechanisms.[10]

Chenyun Dai et al. proposed a study which employs an adapted Wav2Vec 2.0 model for instrument recognition in musical audio, utilizing deep learning techniques. It addresses multi-label classification, achieving varying success across instruments, particularly excelling with violins, pianos, and saxophones, while struggling with lower-volume instruments[11]. Chite et al. evaluates AI-based classification performance for monophonic and polyphonic Indian classical instruments using hybrid domain features, achieving 89-96.33% accuracy with SVM and GMM classifiers, with future scope in real-time classification and IIOT integration[12]. Xize Chen et al. explain music recognition algorithm based on DL and hash method discussed is of great significance in improving the classification accuracy of music recognition and the application of blockchain technology in the copyright protection platform of original music works can protect the copyright of digital music and ensure the operation performance of the system[13]. Saranga Kingkor Mahanta et al. uses an artificial neural network model that was trained to perform classification on twenty different classes of musical instruments to achieve state-of-the-art accuracy on the full London philharmonic orchestra dataset[14]. Rodrigo et al. explained intelligent system was designed to recognize the specific musical instrument from an audio file and construct the notation text file correspondingly for the Music Instrument Digital Interface file corresponding for the music instrument digital interface file[15]. Yanmin Liu et al. proposes a note extraction and recognition system using music melody features and MFCC features, achieving 96.7% F-measure with a CNN classifier, outperforming DTW and CNN without melody features, especially for single-note recognition[16]. Ben Wilkes et al. focuses on multi-modal music genre recognition using audio, text, and image features, but it does not specifically address music instrument recognition. The methodologies discussed are centered around genre classification rather than identifying individual musical instruments[17]. Jorge Calvo-Zaragoza et al.focuses on Optical Music Recognition (OMR), which involves reading music notation and related applications, but does not delve into the identification of musical instruments within that context[18]. Wei-Na Gu et al.focuses on piano playing music recognition, proposing an algorithm that enhances accuracy and fault tolerance in multi fundamental frequency detection, utilizing spectrum peak sorting and spectral entropy, thus improving recognition performance in intelligent interaction environments[19]. Markus Schwabe et al. proposed Instrument recognition in music involves identifying active instruments in recordings using neural networks. This study improves recognition by incorporating phase information through the product spectrum (PS), achieving a 2% increase in F1-score compared to traditional methods using only STFT magnitudes.[20]

## PROPOSED METHODOLOGY

Figure 1 illustrates the step-by-step process of identifying musical instruments from audio recordings using machine learning techniques. The system primarily relies on Mel-Frequency Cepstral Coefficients (MFCCs) for feature extraction and a machine learning classifier for instrument classification.

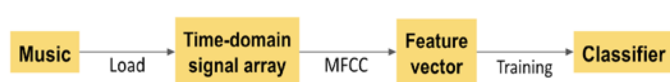



**Figure 1.** Block diagram of proposed musical instrument identification system.

**Music dataset overview**

The music dataset in this proposal utilizes the Philharmonia sound sample collection[21] that comes from a well-known symphony orchestra. The dataset contains specifically curated recordings of 20 musical instruments for classification purposes. The dataset maintains adequate representation for every instrument category through its 1000 high-quality files per instrument type. One instrument plays a separate musical note during each recording which displays specific pitch levels together with volume and timbre lengths. The model benefits from this musical property diversity because training examples span numerous instrument sounds thus enabling strong generalization capabilities. The proposed identification system benefits from the dataset's structured structure which leads to more accurate and reliable identification due to its consistent pitch and loudness management. The diverse instrument recognition process depends on the sample combinations shown in Table 1 which displays three examples of music properties variations.

**Table 1**. Properties of the Example Music Files

| Instrument | Pitch | Duration (sec) |
|---|---|---|
| Cello | E4 | 1 |
| Violin | As5 | 0.5 |
| Sax | F3 | 1 |

We constructed the model using data from seven musical instruments chosen from the Philharmonia dataset. Several main instruments were chosen for model training: cello and violin and viola and trumpet together with oboe and saxophone along with flute. The chosen instruments spanned different timbres across three instrument families: string instrument, brass instrument along with woodwind instrument so the model could learn various acoustic qualities. The classification success relies on unique spectral with temporal characteristics in the seven specific musical instruments sampled from the Philharmonia dataset. The model receives training from seven instruments which enables it to become a comprehensive system for precise instrument identification through audio analysis.

**Data Preprocessing**

A training music file gets stored as an array which contains its time-domain waveform information. The waveform representation saves the amplitude fluctuations during time progression which represents the original sound data. A set of patterns can be extracted from audio data through the application of the Mel-Frequency Cepstral Coefficients (MFCC) algorithm on the time-domain signal. MFCC has become a standard tool for audio processing since it turns sound spectral information into useful patterns by processing sounds through human auditory models. The waveform gets converted into feature vectors through this process to preserve essential acoustic features of the instrument's timbre and characteristics.

**MFCC Feature Extraction**

MFCC is an algorithm that can characterize the timbre of the soundwave. Since instrument classification is based on the timbres of the instruments, MFCC perfectly suits this application. MFCC is also widely used in fields like natural language processing.

MFCCs are commonly derived as follows:

1. Take the Fourier transform of a windowed excerpt of a signal.
2. Map the powers of the spectrum obtained above onto the Mel scale, using triangular overlapping windows.
3. Take the logs of the powers at each of the Mel frequencies.
4. Take the discrete cosine transform (DCT) of the list of Mel log powers, as if it were a signal.
5. The MFCCs are the amplitudes of the resulting spectrum.

The output of MFCC is an n*d array, where n is the number of waveform windows, and d is the dimension of the feature vectors. d can be arbitrarily set by the programmer (higher dimension preserves more information). For example, if the length of a waveform file is 1 second and the window size is 0.1 seconds, then n =10. If the dimension of the feature vectors is chosen as 13, then d =13, and the output dimension will be 10*13.

Our algorithm is expected to accept training and testing waveforms with different lengths, which implies different numbers of windows. However, the SVM only accepts data with a fixed predefined dimension. Therefore, in our algorithm, all feature vectors for a particular music file are averaged to characterize the timber of the entire file. To be specific, the n*d output array from MFCC is averaged along the row direction, producing a vector of length d, which contains the timber information of the entire waveform.

The accuracy of the model can benefit from increasing information content in feature vectors. Using this technique with small datasets creates the possibility of having a model that memorizes peculiar training data patterns thus losing its capability to recognize unfamiliar inputs. The model minimizes overfitting issues through the implementation of a large training dataset expansion which lets it comprehend diverse instrument class variations. The specified MFCC dimensions lead to low overfitting risks despite their increased number.

The extracted features received more depth through increased numbers of Mel bands and MFCC output dimensions. The modified model design allows it to recognize finer instrumental spectral and timbral features. Table 2 displays two models of MFCC parameters including the basic configuration and the updated one with their specified enhancements. Improved classification results might be attainable by continuously increasing MFCC dimensions, yet this strategy would lead to increased computational expenses. Higher-dimensional features need additional resources for storage and

**Table 2**. MFCC Parameters of the Original and the Improved Model

| | Old Parameters | New Parameters |
|---|---|---|
| FFT window length | 2048 | 2048 |
| # of mel bands | 128 | 256 (1) |
| Dimension of output | 13 | 43 (1) |
| | Transform files into vectors with length 13 | Transform files into vectors with length 43 |

processing which extends the time needed for both feature extraction and classification operations. A history of forty-three MFCC dimensions was chosen as the optimal point for maintaining both processing efficiency and classification precision because it supports effective and practical audio processing systems on a large scale.

**Classification**

A different set of vectors extracted from sound waves feeds into machine learning algorithms for precise classification of musical instruments. Various classification systems have been studied during previous instrument recognition research through the evaluation of Support Vector Machines (SVM), Random Forest (RF) and Convolutional Neural Networks (CNN). The classification models exhibit different levels of success across accuracy measures as well as computational speed and learning duration. The identification of musical instruments reaches its peak performance through the combination of SVM and CNN models according to experimental findings. The detection of complex patterns in audio spectrograms by CNNs makes them become an exceptional tool for this purpose. The high resource requirements combined with extensive training time make CNN models less suitable for practical applications in some cases. Taking the previous criteria into account we selected the SVM method to implement in our machine learning model. SVMs deliver both high classification outcomes and efficient computation thus they represent a suitable choice for our system purposes. The implementation of SVM ensures efficient instrument recognition with accurate results and manageable training along with inference timings thus making the model viable for practical deployments.

## RESULT AND DISCUSSION

All recordings from the Philharmonia dataset were performed using one distinct instrument for each instrument type. The recordings of cello samples employed a single instrument throughout while violin samples utilized a single violin instrument. Real-world testing results might differ from cross-validation findings because various instruments of equal types can demonstrate distinct timbral characteristics. The lack of variety in instrument sources within the dataset is not ideal, as it may introduce a significant variance when applied to real-world scenarios.To address this issue and enhance the diversity in timbre, we incorporated additional real-world audio samples into our training data. Specifically, we cropped cello and violin solo recordings from external music pieces into small clips ranging from ⅓ to 0.5 seconds in length. These additional samples were then integrated into the existing training data.

By combining the Philharmonia dataset with overlaid waveforms and real-world music clips, we created a more diverse training set that better represents real-world variations in musical instruments. This approach improves the model's generalization capability and reduces the risk of overfitting to a specific instrument's timbre. The final count of training music files, including these newly introduced clips, is summarized in Table 3.

During the process of improving the model, various combinations of MFCC parameters were tested to analyze their impact on prediction accuracy. As shown in Figure 2, it was observed that simply increasing either the number of Mel bands or the number of MFCC features independently did not necessarily lead to an improvement in accuracy. Instead, significant performance gains were achieved when both parameters were increased simultaneously.

**Table 3**. Number of Music Training Files of Different Instruments

| | Existing Model | Improved Model |
|---|---|---|
| Cello | 100 | 4506 |
| Violin | 100 | 5410 |
| Flute | 100 | 4297 |
| Oboe | 100 | 2937 |
| Sax | 100 | 3267 |
| Trumpet | 100 | 2332 |
| Viola | 100 | 4797 |
| Total | 700 | 27527 |

| Number of Mel Bands | Number of MFCC Features | Prediction Accuracy |
|---|---|---|
| ↑ | ↑ | ↑ |
| ↑ | -- | ↓ |
| -- | ↑ | -- |

**Figure 2.** Result after experimenting with different MFCC parameter combinations. "--" means this parameter is unchanged.

One particularly surprising result was that increasing the number of Mel bands from 128 to 256 while keeping the dimension of MFCC feature vectors unchanged led to a decline in prediction accuracy. This suggests that merely capturing more frequency details through additional Mel bands does not necessarily enhance classification performance unless the feature vector dimensions are adjusted accordingly. The findings highlight the importance of carefully balancing MFCC parameters to optimize model performance while avoiding unnecessary computational complexity.

Building upon the existing model as a baseline, we began by expanding the Philharmonia Dataset and applying data augmentation techniques to enhance the variety and robustness of the training data. We then refined the MFCC parameters, specifically adjusting the length of the FFT window, the number of Mel bands, and the dimension of MFCC feature vectors. The primary aim of our modifications centered around both raising the volume of Philharmonia Dataset processing and growing MFCC feature dimensions along with Mel bands to boost representing power.

The percentage of correctly predicted snippets served as the main evaluation measure for model performance assessment. The model validation required the use of violin and cello solo pieces. Time segments of 0.5 seconds were obtained from music files before submitting them to model classification procedures. The assessment process became simple because the validation tracks consisted of single-instrument performances. The system calculated accuracy from the ratio of snippets the model correctly identified.

Developing shorter music segments from longer recordings allowed the system to gain multiple benefits. The method allowed instantaneous predictions which made the embedded implementation process simpler. The system obtained enhanced instrument recognition ability during periods when various instruments appeared alternatively within music pieces through segmentation. Each time segment analysis in the model led to better instrument recognition precision resulting in improved system reliability and accuracy.

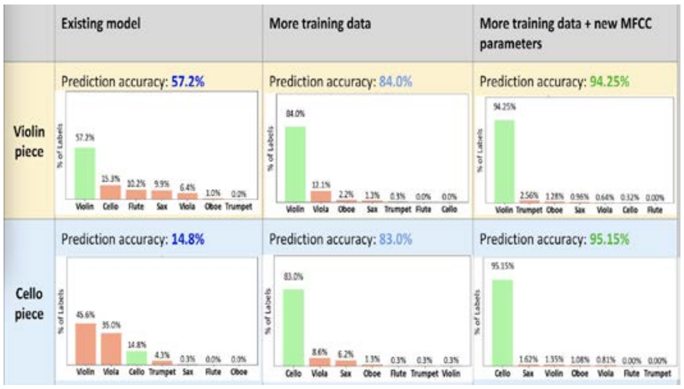



**Figure 2.** Predation Results for the Existing and the Improved Mode

Figure 3 demonstrates how the baseline model evolved into an improved model through a visualization that shows significant upgrading of prediction accuracy. The accuracy score for the cello instrument grew significantly from 14.8% to 83% marking a total improvement of approximately 70%. The violin piece showed strong improvement because accuracy scores rose near 30%. Additional training data led to an 83.5% accuracy overall which demonstrated the effectiveness of this added data.

By modifying the MFCC parameters the process of feature extraction and representation became more effective. A 95% model precision emerged after optimization proving the development of an efficient and dependable automatic classification system. The high precision of this model indicates it would serve well in existing applications like educational tools for music and instrument recognition software and live music analysis applications.

**Table 4:** Comparison of Accuracy with different machine learning models

| Models | Accuracy (%) |
|---|---|
| KNN [13] | 90 |
| Deep learning models [14] | 92 |
| Proposed Models with modified MFCC features | 94 |

Analysis of different machine learning models for musical instrument classification appears in Table 4 where accuracy measurements are provided. The K-Nearest Neighbors (KNN) model which earlier studies described in [13] produced results with accuracy reaching 90%. The application of convolutional and artificial neural networks under deep learning yielded modest accuracy results of 92% [14]. The proposed model utilizing modified MFCC features reached an improved accuracy rate of 94% in comparison to previous models. The upgraded MFCC parameters demonstrate their ability to increase model performance for musical instrument recognition accuracy measurement.

## Discussion

The study findings show that tuning MFCC parameters leads to substantial enhancements in musical instrument classification precision. Traditional machine learning models based on KNN reach a satisfactory accuracy rate of 90% as shown in Table 4 but deep learning models reach 92% accuracy. The modification process applied to MFCC features within the proposed model brought about enhanced classification accuracy which reached 94%.

The enhancement in musical instrument classification accuracy results from raising both Mel bands and MFCC feature dimensions simultaneously to enhance frequency information preservation. Experimental data shows that accuracy increases only when both MFCC feature dimensions and number of Mel bands are simultaneously modified.

The model became more robust because researchers incorporated real-world instrument recordings coupled with augmented Philharmonia sound samples into its wider dataset. Using data from different instruments per type in the training reduced the likelihood of training specifics that would challenge future data from new instruments. Prediction accuracy increased significantly through the enhanced dataset especially for cello and violin samples whose accuracy rose from 14.8% to 83% and violin accuracy improved by approximately 30%.

The model achieved real-time classification capabilities through its approach of fragmenting music clips into shorter sections which made it appropriate for embedded systems applications and music education apps. Performance measurement of the final model achieved 95% accuracy in identifying different musical instruments showing potential for operational deployment in real-world systems. The expansion of instrument recording diversity in the dataset along with research on improved deep learning methods represents future work opportunities to boost performance even more.

## Conclusion and Future work

The proposed work implemented MFCC features to obtain timbre elements while SVM classifiers determined musical instrument categories. By expanding training data scope while adjusting MFCC configuration levels we achieved a sharp increase in model prediction success rate which rose from 36.0% to 94.7% for solo music elements. Raising the amount and variety of training data proves essential for enhancing machine learning models and their prediction accuracy. Dataset homogeneity triggers overfitting problems which produce unpredictable results during generalization. The model performance achieved significant enhancement when we added real-world music samples from different musical sources to the training process. A larger dataset enables the implementation of complex models to derive better usage from extracted features. The performance achieved more improvements after conducting a meticulous test of MFCC parameter settings. Our future work focuses on developing research which applies deep learning models like Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) to boost the overall classification precision. The model's instrument discrimination ability can be improved through data augmentation strategies combined with extended training data featuring ensemble performances. The future involves integrating our system into real-time applications which would enable live instrument recognition for educational, transcription and production purposes. Our team aims to use the same research methods for designing an automatic soundtrack separation system which will enhance music producer flexibility in post-production work. The use of supervised machine learning to identify musical instruments in this work pushes the development of automated music analysis while providing useful tools for both musicians and researchers.

## AUTHOR CONTRIBUTIONS

Bhagyalakshmi R contributed to conceptualization, software development and methodology. Anandaraju M B provided essential resources and contributed to review & editing, refining the manuscript for clarity and coherence.

## CONFLICT OF INTEREST STATEMENT

Authors declare that there is no conflict of interest for publication of this work.

## REFERENCES


1. M. Shanthamallappa, K. Puttegowda, N.K. Hullahalli Nannappa, S.K. Vasudeva Rao. Robust Automatic Speech Recognition Using Wavelet-Based Adaptive Wavelet Thresholding: A Review. *SN Comput. Sci.* **2024**, 5 (2), 248.
2. K. Humse, H.S. Ranjan Kumar, V. Veeraprathap, et al. Automated raga recognition in Indian classical music using machine learning techniques. *J. Integr. Sci. Technol.* **2025**, 13 (1), 1011.
3. S.E. Tamanna, M. Ezhan, R. Mahesh, et al. Musical Instrument Classification Using Deep Learning CNN Models. In *2024 International Conference on Integrated Intelligence and Communication Systems (ICIICS)*; IEEE, **2024**; pp 01–07.
4. Z. Hu. An Instrument Tone Recognition System Based on Artificial Intelligence and Feature Recognition. In *2024 IEEE 2nd International Conference on Image Processing and Computer Applications (ICIPCA),* **2024**, pp. 1274-1277. IEEE.
5. R. Chen, A. Ghobakhlou, A. Narayanan. Interpreting CNN models for musical instrument recognition using multi-spectrogram heatmap analysis: a preliminary study. *Front. Artif. Intell.* **2024**, 7, 7.

s

6. P. Tiemeijer, M. Shahsavari, M. Fazlali. Towards Music Instrument Classification using Convolutional Neural Networks. *2024 IEEE International Conference on Omni-Layer Intelligent Systems, COINS 2024*. 2024, pp 1–.
7. G.A.V.M. Giri, M.L. Radhitya. Musical Instrument Classification using Audio Features and Convolutional Neural Network. *J. Appl. Informatics Comput.* **2024**, 8 (1), 226–234.
8. E. Ding, E. Sharma. Musical Instrument Identification Using Machine Learning. *J. Student Res.* **2024**, 13 (2).
9. T.A. Onisha, J. Kim, J. Seol. Multi Label Sound Classification using Deep Learning Models. *2024 IEEE/ACIS 22nd International Conference on Software Engineering Research, Management and Applications, SERA 2024 - Proceedings*. 2024, pp 129–134.
10. R. Chen, A. Ghobakhlou, A. Narayanan. Hierarchical Residual Attention Network for Musical Instrument Recognition Using Scaled Multi-Spectrogram. *Appl. Sci.* **2024**, 14 (23), 10837.
11. C. Dai. Instrument recognition in unprocessed musical audio through transformer-based modelling. *Appl. Comput. Eng.* **2024**, 67 (1), 167–173.
12. A. V. Chitre, K. Wanjale, A. Deshmukh, et al. Artificial intelligence-based classification performance evaluation in monophonic and polyphonic indian classical instruments recognition with hybrid domain features amalgamation. *Journal of Information and Optimization Sciences*. 2023, pp 341–353.
13. X. Chen, X. Qu, Y. Qian, Y. Zhang. Music Recognition Using Blockchain Technology and Deep Learning. *Comput. Intell. Neurosci.* **2022**, 2022, 1–13.
14. S.K. Mahanta, A.F.U. Rahman Khilji, P. Pakray. Deep neural network for musical instrument recognition using MFCCs. *Comput. y Sist.* **2021**, 25 (2), 351–360.
15. W.U.D. Rodrigo, H.U.W. Ratnayake, I.A. Premaratne. Identification of Music Instruments from a Music Audio File; *Springer, Singapore, **2021**; Vol. 176 LNNS.
16. Y. Liu. Note Features Extraction and Recognition Analysis Based on Music Melody. *Inform.* **2024**, 48 (18), 9–14.
17. B. Wilkes, I. Vatolkin, H. Müller. Statistical and visual analysis of audio, text, and image features for multi-modal music genre recognition. *Entropy* **2021**, 23 (11), 1502.
18. J. Calvo-Zaragoza, J. Hajic, A. Pacha. Understanding Optical Music Recognition. *ACM Comput. Surv.* **2020**, 53 (4), 1–35.
19. W. Gu. Recognition Algorithm of Piano Playing Music in Intelligent Background. *Mob. Inf. Syst.* **2022**, 2022, 1–11.
20. M. Schwabe, O. Elaiashy, F.P. León. Incorporation of phase information for improved time-dependent instrument recognition. *Tech. Mess.* **2020**, 87 (s1), S62–S67.
21. A. Kolkowski, D. Miller, A. Blier-Carruthers. The art and science of acoustic recording: re-enacting Arthur Nikisch and the Berlin Philharmonic Orchestra's landmark 1913 recording of Beethoven's Fifth Symphony. *Science Museum Group Journal,* **2015**.