International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

Classification of Classical Indian Music Tabla Taals using Deep Learning

Dr. S. T. Patil[1], Pardeshi Palak Abhay[2], Lambture Suvarnalaxmi Prakash[3], Kotangle Mallaika Arun[4], Pagare Aniket Arun[5]
[1, 2, 3, 4, 5] Department of Computer Engineering, Vishwakarma Institute of Technology, Pune, Maharashtra, India.
[1] patil.st@vit.edu, [2] palak.pardeshi22@vit.edu, [3] suvarnalaxmi.lambture22@vit.edu, [4] mallaika.kotangle22@vit.edu,
[5] aniket.pagare22@vit.edu

Abstract— In the research that we are bringing to light, we profoundly explore the categorization of Classical Indian Music Tabla Taals. This emphasizes widely recognized taals such as Addhatrital, Ektal, Rupak, Dadra, Deepchandi, Jhaptal, Trital, and Bhajani. To push the boundaries of our understanding, we implement a mixed-methods approach tethering both Feedforward Neural Networks (FNN) and Convolutional Neural Networks (CNN). These state-of-the-art technologies enable us to dissect and categorize tabla taals efficiently. In essence, the hallmark of Classical Indian music is its complex and multifaceted rhythms brought to life by the primal percussive instrument - the tabla. The conception and reproduction of these nuanced taals require technical finesse. Thus, accompanying the digital revolution and the eclectic musical preferences, it becomes essential for advanced methodologies to pinpoint and classify tabla taals. The hardcover of our research opens up to the magnificent crafting of an unmatched model employing both FNN and CNN. This blend enables us to recognize diverse features unique to tabla taals like Addhatrital, Ektal, Rupak, Dadra, Deepchandi, Jhaptal, Trital, and Bhajani. The model obtained its bosom knowledge during training from an assortment of Classical Indian music recordings showcasing these invigorating taals. This fosters a broader understanding regarding the array of minute differences brimming within each rhythmic inheritance. To bring user interaction to life, we have embedded a Graphical User Interface (GUI). This empowers users to introduce an audio file filled with table music from the taals listed and receive on-the-spot recognition. refining their connection and knowledge of the taal in question. Our research findings procure paramount importance in the scape of music analysis, especially framed within the heart of Classical Indian Music. We propose a system that would serve as a tool for amateur table players to learn the skill well and master their art. Instructors could also utilize it for training purposes. It opens a new window of possibilities providing an advanced model for intuitive, swift, and accurate automated identification of tabla taals.

Keywords— Addhatrital, Bhajani, Classical Indian Music, CNN, Dadra, Deepchandi, Ektal, FNN, hybrid model, Jhaptal, music classification, Rupak, Tabla, Taals, Trital.

I. INTRODUCTION

For years, music, which is understood in different dialects, was one of the most important means of the expression of people. It has created a deep emotional effect in our hearts, reducing our stress levels and thereby ensuring joyful living. The complex musical landscape has emerged due to cultural and historical influences. Various changes in music have spanned a period of decades, revealing the complex web of various cultural inputs. It is therefore important that within this large musical spectrum, classification of these musical records for easy organization and comprehension is a key element.

In the North Indian Classical Music system, the tabla is generally employed as percussion for vocal music in polyphonic compositions. The auditory system perceptually groups musical elements thus the tabla component is easily filtered, and prominent rhythmic features like tala and tempo can be decoded from a polyphonic composition. [1] It is a traditional Indian percussion instrument that has a great reputation for its complicated and rhythmic patterns. It is composed of two hand-played drums, the smaller drum known as 'Dayan' is played using the dominant hand which is made of wood and the larger drum known as 'Bayan' is played with the non-dominant hand and is usually constructed from metal, clay, or a mixture Differentiation of tabla taals is a unique challenge that arises from dynamic nature of the instrument. Various hand techniques, strokes, as well as intricate patterns employed by the tabla players do not allow accurate isolation and classification of the individual taals. The use of subtle variations in strokes, tempo changes, and improvisational nuances also make it difficult to identify tabla taals through automated classification, and that made our research a difficult one.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

Vibhags are the sections where the talas are fixed patterns of some length composed of beats or matras. The whole complete cycle is called 'avart' but the whole tala will begin and terminate with the first beat called 'sam', meaning that the first beat, will always add up or should be considered as adding up at the end of avart. Tala has two gestures which are called the clap(tali) and wave(khali) which follow the vibhag on each talas. [2] In addition to that they mark accented or unaccented parts and they give vital indications to the soloist, when the tala is too difficult and when the soloist might be confused.

Our study takes into consideration a specific slice of the entire pie of the music world, i.e., Classical Indian Music Tabla Taals. It also includes the complex rhythms of tabla in the context of classical Indian music. In this project, Feedforward Neural Networks are applied together with Convolutional Neural Networks to distinguish between different tabla taals within specific taals and some popular ones.

II. RELATED WORK

Several studies in the field of music genre classification have explored the application of deep learning techniques to categorize and analyze diverse musical genres. Gowriprasad. R and K. Sri Rama Murty, in the paper titled 'Onset Detection of Tabla Strokes using LP Analysis' propose a pre-processing method to improve onset detection of intricate strokes. The document discusses the problems of tabla stroke resonance and its influence on the energy-based and spectral flux-based onset detector when applied to raw signals. However, the proposed algorithm uses LP analysis and HE together to overcome this situation. It results in emphasizing the onset times, thus, increasing the accuracy of stroke onset detection. The document provides the results of the evaluations based on the spectrogram illustrations and the datasets. The presented signal processing approach results in significant improvement in F-measure. It is suggested to fine-tune the detection of tabla strokes. [3]

Uttam Kumar Roy authored a study titled 'Composing Recorded Tabla Sound to Accompany Musicians'. This research reveals that the tabla, a main percussion instrument in Asian music, lacks sufficient trained performers. Therefore, synthetic tabla sounds may be necessary. Tabla syllables are recorded to develop a scheme for synthesizing them. The scheme stresses a particular strategy used in tabla to achieve harmonic overtones. It explores the complexities of rhythmic structures and serialization of tabla syllables, emphasizing the importance of tuning pitch and increasing sound frequencies. The paper also presents the efficient re-sampling techniques for the tabla sound, giving references to related studies and research on the Indian musical drums. [4]

George Tzanetakis and his co-authors, Ajay Kapur and Richard I. McWalter discuss a system for rhythm-based music information retrieval, with emphasis on drum transcription, in their paper titled 'Subband-based Drum Transcription for Audio'. This is the comparison of wavelet and filter algorithms using F-Measure and detailed experimental results. It comprises adaptive thresholding and peak picking on the frequency bands' envelope amplitude. It discusses the systems, outlines the experimental setup, and gives results for evaluation. [5]

Prakash Persad, Jorrel Bisnath, and Ruel Ellis examine the problem in conventional learning environments, especially in classical Indian drumming, which remained the same for over 2000 years in the work 'Investigating the Use of a Robot with Tabla Education'. The tabla is very difficult to play and requires one to exercise mastery in musculature and playing techniques given the complex nature and the unique tonal quality. Additionally, tabla music is not standardized, making learning for both students and tutors even more difficult. The authors recommend that a robotic tabla tutor should be used together with a human tutor to improve teaching and practice to overcome the challenges. Experiments with a prototype robot tabla player show promise as a teaching aide. The paper stresses the need to develop metrics for measuring the viability and effectiveness of the idea of robot-tabla-tutor. [6] All these studies helped us identify the limitations and research gap in the field and led to the creation of our problem statement.

III. DATASET

Searching for a suitable dataset for our Classical Indian Music Tabla Taals-focused research turned out to be quite challenging because there were no ready tools that satisfied our needs at our disposal. Therefore, we embarked on the process of rigorous compilation of our raw data from the onset. This entailed carefully listening to each song and placing it in its respective genre folder. Utilizing YouTube and Spotify as key sources of diverse musical content, we assembled our archive.

We found it difficult to collect substantial data for each taal, resulting in an unbalanced dataset. Deep learning model training becomes cumbersome in the case of an imbalanced dataset comprising a huge number of instances for different classes. However, our case had a few taals of data that were not enough, thus leading to an uneven distribution throughout the dataset, making the process of data extraction difficult.

Article Received: *25 July 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

For our project, we created a dataset of 424 .wav files of Tabla tunes and split it as a training (80%) and test (20%) set. It provides a balanced mix for a diversified representation which in turn enabled our model to learn and generalize effectively. Our model attains an impressive test set accuracy of 92% with a training configuration of 30 epochs and a batch size of 512 samples. This indicates the strength of our model to correctly categorize the tabla tune under various taals, which suggests its applicability to the real-world music genre classification problem. The system also throws an error if a file other than a .wav format file is submitted to be analyzed for table taals.

Although it was not easy to produce the dataset, our rich collection provides a solid basis for detailed research into classifying tabla blows in several taals, including Addhatrital, Ektal, Rupak, Dadra, Deepchandi, Jhaptal, Ttital, and Bhajani. This pushes us therefore to use careful strategies for model training to ensure we comprehend the diverse rhythmic patterns in Classical Indian Music.

The following table provides an overview of the dataset and the number of instances used in each of the classes, which in turn made up our entire dataset of 424 .wav audio files:

TABLE 1: NUMBER OF INSTANCES

Sr. No	Taal	Count
1	Addhatrital	77
2	Trital	72
3	Ektal	72
4	Rupak	60
5	Dadra	16
6	Deepchandi	19
7	Jhaptal	60
8	Bhajani	48
	Total	**424**

IV. METHODOLOGY

A. Data Collection:

Our dataset comprises 424 audio samples of Classical Indian Music table taals in .wav format, covering taals such as Addhatrital, Trital, Ektal, Rupak, Jhaptal, Bhajani, Dadra, and Deepchandi. Ensuring diversity in taals, this dataset formed the foundation for robust model training and thus helped us achieve greater accuracy within a limited number of epochs.

B. Data Preprocessing:

Librosa Feature Extraction:

Leveraging the Librosa library, the following essential audio features are extracted:

1. Chroma_stft
2. Root Mean Square (RMS)
3. Spectral Centroid
4. Spectral Bandwidth
5. Spectral Rolloff
6. Zero-Crossing Rate (ZCR)
7. Mel-frequency Cepstral Coefficients (MFCCs)

These features provide a comprehensive representation of the tabla taals, capturing both temporal and spectral characteristics. The average of these features helps in the generation of a spectrogram.

Spectrogram Generation:

The audio files are transformed into spectrograms using the Librosa library for further analysis of the audio file and detection of taal.

Parameters for Spectrogram:

1. Window Size: 60 milliseconds
2. Hop Length: 16 milliseconds
3. FFT (Fast Fourier Transform) Size: 72

Spectrograms offer a visual representation of the frequency content over time, aiding in the identification of tabla taals.

C. Model Architecture:

Hybrid Model (FNN + CNN):

A hybrid model is designed to incorporate both Feedforward Neural Network (FNN) and Convolutional Neural Network (CNN) layers. The layers focus on capturing spatial patterns within the spectrograms, enhancing the model's ability to recognize intricate features specific to tabla taals.

The following table provides a compact summary of the architecture, with an FNN and a CNN layer. The layers are ordered, specifying the network type, the layer name, the output size, and other data like filter shapes or the number of neurons. Under 'Units', it shows the number of units in Fully Connected layers when applicable:

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

TABLE 2: ARCHITECTURE SUMMARY OF CNN AND FNN LAYERS

Layer (Type)	**Output Size**	**Filter Shape/ Neurons**	**Units**
Input (FNN)	W x H x C	-	-
Convolution 1 (CNN)	W/2 x H/2 x 32	3x3xCx32	-
Max Pooling 1 (CNN)	W/4 x H/4 x 32	2x2	-
Convolution 2 (CNN)	W/4 x H/4 x 64	3x3x32x64	-
Max Pooling 2 (CNN)	W/8 x H/8 x 64	2x2	-
Convolution 3 (CNN)	W/8 x H/8 x 128	3x3x64x128	-
Max Pooling 3 (CNN)	W/16 x H/16 x 128	2x2	-
Flatten (FNN)	1 x (W/16)x(H/16) x128	-	-
Fully Connected 1 (FNN)	1 x 512	-	512
Fully Connected 2 (FNN)	1 x 256	-	256
Fully Connected 3 (FNN)	1 x 128	-	128
Fully Connected 4 (FNN)	1 x 64	-	64
Output (FNN)	1 x Output	-	Output

D. Model Training:

Optimization and Loss Functions:

The model is trained using the Adam optimizer with a learning rate of 0.001. Categorical Cross-Entropy is employed as the loss function to guide the model towards accurate classification.

The following figure shows a graphical representation of the model loss:

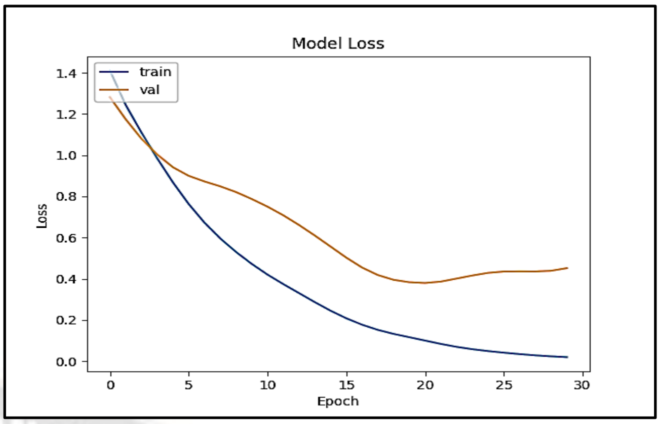

Fig 1: Model Loss

Training Process:

The training process spans 30 epochs with a batch size of 512, enabling the model to learn complex patterns inherent in tabla taals.

E. Model Evaluation:

Dataset Split:

The dataset is divided into training (80%) and testing (20%) sets to assess model performance comprehensively.

Key Metrics:

Accuracy is computed to measure the model's overall classification performance.

A confusion matrix is generated to visualize the distribution of predicted classes against actual classes, offering insights into classification nuances.

The following figure shows a graphical representation of the model accuracy:

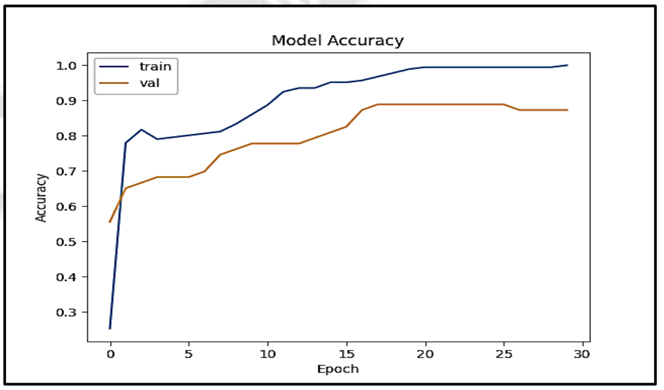

Fig 2: Model Accuracy

F. GUI Integration:

Flask Web Application:

A Flask-based Graphical User Interface (GUI) is developed to facilitate user interaction.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

Users can upload .wav audio files containing tabla taals, and the model detects the corresponding taal present in the file, providing an intuitive, easy-to-use, and accessible user experience.

V. RESULTS AND DISCUSSION

Valuable insights were gained from the commendable performance of the model that was implemented. The following presents a summary of the key outcomes:

A. Distinguishing between various tabla taals was a task at which the classification model excelled, with a high accuracy rate of 92% on the test set. This model was highly proficient in achieving this feat.

B. Employing an exhaustive data processing approach, the model utilized the Librosa library to extract crucial audio characteristics. The method involved using chroma short-time Fourier transform, spectral centroid, spectral roll-off, root mean square, spectral bandwidth, zero-crossing rate, and Mel-frequency cepstral coefficients (MFCCs). By doing so, the model's capacity to identify distinct traits of varied tabla taals was vastly amplified.

C. Strategically designed is the neural network architecture, which ingeniously merged Convolutional Neural Networks (CNN) alongside Feedforward Neural Networks (FNN). Its composition comprised fully connected layers, max-pooling layers, convolutional layers, and a flattened layer, and the intricate nature of this structure effortlessly extracted patterns from the audio data to increase the fine-tuned identification of genres.

The following image represents the GUI of the application:

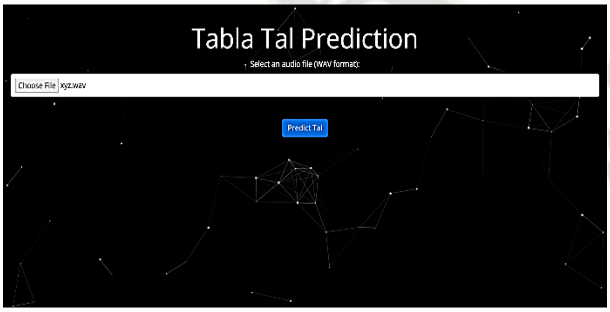

Fig 3: Tabla Taal Predictor

VI. FUTURE WORK

Extending the scope of the project, we can elevate its user interface. Additional features of computational musicology can be added for better results through statistical analysis and other mathematical algorithms. [7] We could integrate tools for live recording and monitoring, which could grant us the ability to scrutinize tabla taals as they happen, allowing for instantaneous assessment. Audio fingerprinting using the sum of a multi-band spectrum can be used to pitch the jitters encountered in the audio. [8] To produce a comprehensive dataset to enhance future models, a data logging structure to house categorized taals can be implemented. To facilitate even more interaction and education on our platform, it would be wise to offer users a mechanism for feedback and tips for executing top-performing taals. A variety of sensor techniques can be used to capture body gestures and train a student performing the Tabla. [9] The Object Oriented feature Polymorphism can be used in designing a generic interface that embraces a set of related activities to classify complex taals in sub-classes and super-classes for better accuracy. [10] All these features would make this project a great asset for tabla enthusiasts.

The following image shows the current scope of work of the application:

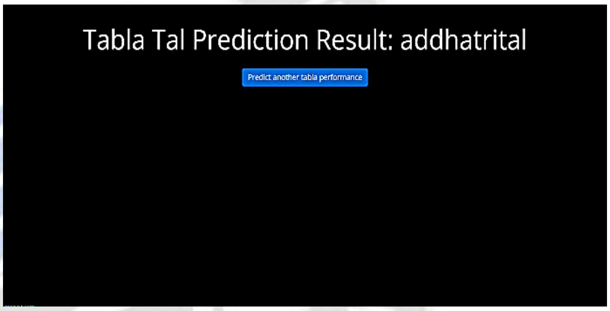

Fig 4: Predicted Taal

VII. CONCLUSION

Our rigorous exploration of the complexities of Tabla Taals in Classical Indian Music has yielded a highly customized and effective classification system. We adopted a unique method, involving a combination of Convolutional Neural Networks (CNNs) and Feedforward Neural Networks (FNNs), to achieve this feat. Our comprehensive training has enabled our model to accurately identify a range of tabla taals, including the likes of Bhajani, Ektal, Addhatrital, Jhaptal, Rupak, Trital, Dadra, and Deepchandi. This powerful model achieves remarkable levels of precision, providing a much deeper insight into the diverse and intricate rhythmic structures of Classical Indian Music.

Not only have we incorporated technical sophistication into our model, but we have also acknowledged the distinctive traits and intricacies characteristic of Classical Indian Music Tabla Taals. Our dataset creation and implementation of advanced deep-learning techniques have resulted in a finely tuned model capable of capturing the nuances of this traditional form of music.

Additionally, the integration of a GUI in our model makes it more accessible and practical. This allows users to easily input audio files that have tabla taals for immediate genre recognition which brings together advanced deep learning models and user-friendly applications. The combination of FNN and CNN reflects our dedication to incorporating cutting-edge methodologies to achieve exact classification.

In summary, our research work is not only at the forefront of music analysis and classification but also an expression of Classical Indian Music Tabla Taals, as well as an effort to preserve and appreciate its cultural heritage. This is a living example that embodies the confluence of traditionalism and technology while giving us deeper knowledge into the intricacies of rhythm making up the essence of Classical Indian Music.

VIII. REFERENCES

[1] Susmita Bhaduri, Anirban Bhaduri, Dipak Ghosh, 'Detecting tala Computationally in Polyphonic Context - A Novel Approach', arXiv:1611.05182, American Journal of Computer Science and Information Technology(2018), https://doi.org/10.48550/arXiv.1611.05182

[2] M. Clayton, Time in Indian Music: Rhythm, Metre and Form in North Indian Rag Performance. Oxford University Press, Google Books, 2000.

[3] R. Gowriprasad. and K. S. R. Murty, "Onset detection of Tabla Strokes using LP Analysis," 2020 International Conference on Signal Processing and Communications (SPCOM), Bangalore, India, 2020, pp. 1-5, doi: 10.1109/SPCOM50965.2020.9179587.

[4] U. K. Roy, "Composing recorded Tabla sound to accompany musicians," 2015 Annual IEEE India Conference (INDICON), New Delhi, India, 2015, pp. 1-6, doi: 10.1109/INDICON.2015.7443114.

[5] G. Tzanetakis, A. Kapur and R. I. McWalter, "Subband-based Drum Transcription for Audio Signals," 2005 IEEE 7th Workshop on Multimedia Signal Processing, Shanghai, China, 2005, pp. 1-4, doi: 10.1109/MMSP.2005.248626.

[6] P. Persad, J. Bisnath and R. Ellis, "Investigating the Use of a Robot with Tabla Education," 2008 Second IEEE International Conference on Digital Game and Intelligent Toy Enhanced Learning, Banff, AB, Canada, 2008, pp. 103-107, doi: 10.1109/DIGITEL.2008.45.

[7] Soubhik Chakraborty, Guerino Mazzola, Swarima Tewari, Moujhuri Patra, Computational Musicology in Hindustani Music, Springer Cham, 2014, https://doi.org/10.1007/978-3-319-11472-9

[8] S. S. Ranjani, K. Karthik, P. K. Bora, and V. Abdulkareem, "Multi-band sum of spectrogram-based audio fingerprinting of Indian film songs for multi-lingual song retrieval," 2015 International Conference on Advances in Computing, Communications, and Informatics (ICACCI), Kochi, India, 2015, pp. 1414-1420, doi: 10.1109/ICACCI.2015.7275811.

[9] M. S. Benning, A. Kapur, B. Till, G. Tzanetakis and P. Driessen, "A Comparative Study on Wearable Sensors for Signal Processing on the North Indian Tabla," 2007 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing, Victoria, BC, Canada, 2007, pp. 621-624, doi: 10.1109/PACRIM.2007.4313313.

[10] D. De and S. Roy, "Polymorphism in Indian Classical Music: A pattern recognition approach," 2012 International Conference on Communications, Devices and Intelligent Systems (CODIS), Kolkata, India, 2012, pp. 612-615, doi: 10.1109/CODIS.2012.64222