International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7
DOI: https://doi.org/10.17762/ijritcc.v11i7.7844
***Article Received**: 29 April 2023 **Revised**: 24 June 2023 **Accepted**: 02 July 2023*

Acoustic Feature Identification to Recognize Rag Present in Borgit

Spandan Kumar Barthakur[1], Parismita Sarma[2*], Gaurav Narayan Baruah[3]
[1]Department of Information Technology, Gauhati University, Guwahati, Assam, India
e-mail: spandankumarb@gmail.com
[2*]Department of Information Technology, Gauhati University, Guwahati, Assam, India
*e-mail: pari@gauhati.ac.in
[3]Department of Information Technology, Gauhati University, Guwahati, Assam, India
e-mail: gauravbaruah20@gmail.com

Abstract— In the world of Indian classical music, raga recognition is a crucial undertaking. Due to its particular sound qualities, the traditional wind instrument known as the borgit presents special difficulties for automatic raga recognition. In this research, we investigate the use of auditory feature identification methods to create a reliable raga recognition system for Borgit performances. Each of the Borgits, the devotional song of Assam is enriched with rag and each rag has unique melodious tune. This paper has carried out few experiments on the audio samples of rags and a few Borgits sung with those rugs. In this manuscript three mostly used rags and a few Borgits with these rags are considered for the experiment. Acoustic features considred here are FFT (Fast Fourier Transform), ZCR (Zero Crossing Rates), Mean and Standard deviation of pitch contour and RMS(Root Mean Square). After evaluation and analysis it is seen that FFT and ZCR are two noteworthy acoustic features that helps to identify the rag present in Borgits. At last K-means clustering was applied on the FFT and ZCR values of the Borgits and were able to find correct grouping according to rags present there. This research validates FFT and ZCR as most precise acoustic parameters for rag identification in Borgit. Here researchers had observed roles of Standard deviation of pitch contour and RMS values of the audio samples in rag identification.

Keywords- Borgit, Rag, Fast Fourier Transform, Zero Crossing Rates, Root Mean Square and Standard Deviation.

I. INTRODUCTION

The Borgits are versatile song being introduced by two holy souls of Assam Srimanta Sankardeva and his disciple Srimanta Madhavdeva in 15^{th} to 16^{th} century in Assamese. The scripts of Borgits are in Brajawali language [1,2]. Borgit consists of one basic component known as Rag. Borgits are played in different sattras with two specific instruments i.e. Khol and Taal [3]. No much research on Borgit are going on till now and no experimental evidence is found which can recognize unique acoustic features of rag and Borgit. Understanding the acoustic measures experimentally and automatic identification of rag in a Borgit are main aims of this research. The authors of this paper had tried to identify the most distinguished acoustic features of Borgit and hence identify the rag of that borgit from those feature values. After going through different digital observations on the acoustics components of the Borgit, it is seen that all the Borgits have at least one of the mostly used rag of Indian classical music[4]. Rag recognition uses some methods which describes and categorizes different notes from a musical audio file[5]. By digitally indentifying a rag that exist in Borgit, musician as well as researchers will be benefitted and will be able to preserve their heritage and culture in digital code form for future generation. Till now no fruitful research outcome on Borgit is found which may characterize a Borgit in terms of probable acoustic parameters. Each Borgit is allotted with a specific time space for singing during the day and which is termed as 'Prahar' of singing. Generally in Borgit, we get two kinds of rag i.e. Bandha(close)-rag and Mela(open)-rag[6]. The main features of Rags are based on musical techniques, with the following seven characteristics [7]. Notes (swaras), Aaroh and Avroh, Vadi and Samvad , Gamakas, Pakad, Tala, Thaat. Mostly found rags in the Borgits composed by Shankara and Madhava are Ahir bhairava, Ashowaree, Kalyana, Kou etc [8]. Apart from a couple of rags like 'Kau', the rest of the rags are used in Indian classical music also.

Outline: This paper puts light on different acoustic parametric characteristics of Borgit, the distinguished acoustic attributes with help of which the rag in the Borgit can be identified. The introduction part of this manuscript consistes of general introduction to Borgits, objective of the research, purpose and novality. The literature review is included in introduction chapter itself. This literarture survey section mentions different related research papers not specifically on Borgit but reaserach work done on other Indian classical songs. Section 2 explains the material and method used in this paper. It describes the methods and software tools used to compute different musical

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7
DOI: https://doi.org/10.17762/ijritcc.v11i7.7844
***Article Received**: 29 April 2023 **Revised**: 24 June 2023 **Accepted**: 02 July 2023*

attributes. Section 3 shows all the result we achieved and make an analysis, while section 4 is on conclusion and future work.

Purpose: Main purpose of this research is to experimentally determine most contributing acoustic parameters exhibiting creditable differences among the rags while rolling inside the Borgits. Results obtained through this experiment will help the researchers in the interdisciplinary research domains of Computer Music, Mathematical Music Theory and Music Information Retrieval. Moreover automatic rag recognition from Borgit may be integrated as a module with any music identification tool.

Novelity: Many contemporary research related to rag and their automatic identification in Indian Classical Music were studied but no specific works on Borgits in the light of mathematical music theory or in music information retrieval system are found. So this is a unique reaserch work performed on audio samples of Borgit and authors have initiated it as foundation of a broad study of musical arrangement of Borgits. After reading many of the research articles in the domain of audio identification, we are getting that there is no previous work done on Assamese borgit rag identification with extracted feature of this experiment i.e, ZCR, FFT , pitch.

Few researchers work on Hindustani classical music while few work on ICM raga like Yaman and Bhairavi, but there is no work found of the raga we are taken that are Asaowari, Dhanashree and AhirBhairav. So we can state that our work is of unsolved problems.

In this context we can say that we will peroform audio analysis on Assamese borgit while there is no previous work on Assamese borgit rag identification. Identification of rag in borgit and extraction of different features like Pitch contour, zero crossing rate which deals with the rate at which a signal changes it sign from positive to negative and vice versa and Fast fourier transform where the fourier values measured magnitude and phase of frequency components. After getting numerical representation in terms of those parameters we have applied K means clustering algorithms on FFT and ZCR values.

A. Objectives

Understanding the acoustic measures and automatic identification of rag in a Borgit are main aims of this research. Most of Indian classical music is based on rag, a melodic construction, which cherish the song with its unique tune. The main objectives of this research are as follows:

i. Make awareness on musical arrangement of the Borgits.
ii. Studying and digitally indentifying rag that exist in Borgits, musician and researchers can preserve their rich heritage and culture for future generation.
iii. This experimental research on acoustic attributes will contribute to musicological research of Borgit. Feature identification of rag and Borgits may be a field of research in broad musical landscape with interdisciplinary research in Mathematical Music Theory, Music Information Retrieval and Music Informatics.

II. LITERATURE SURVEY

A review of the relevant literature, their major contributions and technique used are given below.

A thorough study project was suggested by Ekta and Savita [9] that explores the crucial topic of rag detection methods. Their study's major goal is to provide efficient techniques for identifying and categorizing distinct kinds of rags in a variety of situations. In this context, the term "rags" refers to irregular or unstructured items that may exist in visual data or photographs. The researchers used a wide variety of classifiers, each with certain advantages and traits, to accomplish this goal. In their paper, they examine a variety of classifiers, including the Bayesian net, a probabilistic graphical model that facilitates effective inference, the naive Bayes classifier, a straightforward but potent probabilistic classifier, the Support Vector Machine (SVM), a well-liked supervised learning algorithm known for its capacity to handle high-dimensional data, the decision table, which supports rule-based decision-making, the random forest ensemble learning technique, and the multi-layer support vector Ekta and Savita assessed the effectiveness of various classifiers when used to conduct rag identification tasks through diligent experimentation and thorough review. To thoroughly evaluate the performance of each classifier, they took into account a number of performance parameters, including accuracy, precision, recall, and F1 score. The results of their study offer insightful information about the advantages and disadvantages of each classifier for rag detection. This information advances computer vision and image processing, creating new opportunities for robotics, industrial automation, and environmental monitoring, among other applications. Overall, the work of Ekta and Savita makes a substantial contribution to the field of pattern recognition and highlights the significance of choosing the right classifiers for certain tasks like rag identification. Their research lays a strong foundation for future studies and spurs new advancements in the field of object identification and classification.

A fascinating experiment was carried out by Joshi D. et al. [10] with a focus on Hindustani classical music. Their research focuses on the raga classification, notably on Yaman and Bhairavi, two important ragas. The researchers used the K Nearest Neighbor (KNN) and Support Vector Machine (SVM), two popular classifiers, to achieve their study goals. These

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7
DOI: https://doi.org/10.17762/ijritcc.v11i7.7844
Article Received: 29 April 2023 Revised: 24 June 2023 Accepted: 02 July 2023

classifiers were used on the rigorously collected and ready-for-experiment rag dataset of Yaman and Bhairavi. Surprisingly, the outcomes of their experiment showed that both KNN and SVM classifiers had excellent accuracy. The SVM classifier demonstrated its effectiveness in various classification tasks by achieving a remarkable 95% accuracy when tested utilizing an 80/20 Train/Test ratio. Notably, the KNN classifier also produced results that were very promising. The researchers found that some Neighbor values surpassed SVM in terms of accuracy by experimenting with different Neighbor values in the 80/20 Train/Test ratio arrangement. As a result, KNN with the ideal Neighbor value produced the maximum accuracy and outperformed SVM in this particular situation. The results of Joshi D. et al.'s experiment have important ramifications for the study of raga classification and Hindustani classical music. The KNN and SVM classifiers' excellent accuracy highlights the promise of machine learning methods to improve our comprehension and analysis of classical music. Furthermore, the discovery that KNN is a more accurate classifier in some circumstances offers insightful information for future study and motivates additional investigation of other classifier setups for specific classification tasks. Overall, Joshi D. et al.'s research makes a substantial contribution to the field of music and machine learning, opening the door for innovative uses in the fields of music analysis, recommendation systems, and cultural heritage preservation. Their efforts serve as a testament to the ability of artificial intelligence to decipher the subtleties of Hindustani classical music and deepen our understanding of this priceless form of art. A ground-breaking digital music library with a wide selection of songs has been unveiled by Chakrabarty et al. [11] Each song is linked to a particular rag name and its related healing properties in the context of music therapy. This cutting-edge library intends to investigate the healing potential of music and its effects on mental and emotional health. The basis of their research is the creation of a sophisticated algorithm intended to ascertain the degree of similarity between various songs found in the collection. The method can accurately gauge how similar two songs are by using the statistical idea of the correlation of coefficient. Chakrabarty et al. can find patterns and connections between musical aspects including rhythm, melody, and tonality using the Correlation of Coefficient, a commonly used statistical measure. With the help of this innovative method, they can group songs together in the digital music library based on their commonalities. The field of music therapy can benefit greatly from this novel computational strategy. Therapists can more effectively customize music interventions for people and increase the efficacy of music-based healing treatments by combining songs with comparable traits and healing properties. The rich connections between music and emotions can also be explored by researchers, musicians, and therapists with the help of this digital music library, opening the door to more specialized and focused therapeutic applications. The study of Chakrabarty et al. represents a significant advancement in the fusion of music and technology for the enhancement of human wellbeing. By utilizing the power of music as a potent instrument for emotional healing and general wellness, their study has provided a profound understanding of the healing potential of music that has the potential to alter the way we approach mental health and therapeutic interventions.

Dodia *et al.* [12] proposed a work that extracts audio features with the help of a Machine Learning algorithm and a classifier such as SVM, KNN and CNN. They selected two rags from Indian classical music, named Bhimpalasi and Yaman. Average accuracy found for K-NN classifier is 92 per cent and for SVM is 91 per cent.

An extensive survey was carried out by Kalyani et al. [13] with a focus on the crucial area of rag detection methods in Indian Classical Music. Their study's main goal was to investigate the qualities and traits of Indian classical music that are useful for accurate rag identification.

The researchers carefully considered numerous methods for rag identification in their study and noted the essential characteristics of Indian Classical Music that are essential to this process. These characteristics cover a variety of musical aspects, such as pitch values, note patterns, and particular traits particular to each raga.

The results of Kalyani et al.'s survey shed important information on the success of several rag identification methods, highlighting the most promising strategies for this difficult endeavor. These findings have important ramifications for the development of computational musicology and music information retrieval, ultimately leading to a better comprehension and appreciation of the nuances of Indian classical music.

Overall, Kalyani et al.'s research is a useful tool for the music research community because it provides an in-depth analysis of rag detection methods and their relative accuracy. Their work not only exemplifies the capability of machine learning for raga analysis and recognition, but also opens up possibilities for further investigation and improvement of these approaches to improve comprehension and preserve this important cultural heritage.

Santosh and Satya [14] proposed a research for the identification of rag. According to them the most notable barriers are music pitch and mood, extra tones, data translation, and the rag pace.

A cutting-edge method for rag recognition in Indian classical music was introduced by Devansh et al. [15] by fusing the strength of deep learning and signal processing techniques. Their innovative approach is based on using unprocessed spectrograms taken directly from audio sources as the main data format.

The researchers developed a series of preprocessing procedures to improve the raw spectrograms and maximize their compatibility for deep learning models in order to achieve accurate and effective rag recognition. To extract pertinent patterns and characteristics from the spectrograms, these preprocessing processes most likely involved activities like noise reduction, normalization, and feature extraction.

Their research exemplifies the impressive potential of deep learning models in handling rag recognition tasks and shows how artificial neural networks are capable of learning intricate patterns and representations from unprocessed audio input. The suggested method achieves a remarkable testing accuracy of 98.98% on a subset of 10 ragas derived from the CompMusic dataset by utilizing the expressive potential of deep learning.

The remarkable accuracy attained by Devansh et al.'s method is a significant development in the field of rag recognition technology and has great promise for practical uses. The identification of ragas must be accurate at such high levels in order to support many fields like music information retrieval, music recommendation systems, and musicological study.

Their study makes a significant contribution to the study of computational musicology and highlights how machine learning can be used to preserve and study Indian classical music. Devansh et al. present a cutting-edge method that is anticipated to have an impact on future advancements in the field of computational music studies and music recognition by fusing deep learning approaches with expertise in signal processing. Additionally, their findings emphasize the value of utilizing cutting-edge technologies to deepen our comprehension of cultural heritage and foster a love of classical music traditions.

MFF-SAug (Multi-Feature Fusion with Speech Augmentation), a ground-breaking research strategy developed by Jothimani and Premalatha [16], aims to improve emotion prediction from speech. To increase the precision of emotion recognition, their methodology employed a number of signal processing techniques and feature extraction techniques.

In order to improve the quality of voice signals, the researchers used three essential speech augmentation processes in their work. First, noise removal was used to remove unnecessary noise from the speech data, improving the clarity of the audio's emotional cues. Second, White Noise Injection was used to diversity the dataset and add new variations that help the model generalize to various noise situations more effectively. The pitch of speech signals was also modified using Pitch Tuning, which helped provide a more accurate representation of emotional expressions.

An innovative categorization method using a Deep Recurrent Neural Network (DRNN) was presented by Michele et al. [17] for the analysis of audio signals gathered from construction sites. Their study's goal was to accurately categorize and distinguish different pieces of construction machinery and equipment using the acoustic data they create. The popular Recurrent Neural Network (RNN) variation known as the DRNN was chosen because of its propensity to efficiently capture temporal interdependence and sequential patterns within audio inputs. Because of this, it excels at jobs involving time-series data, such as audio. The researchers gathered a wide range of audio recordings from numerous construction sites for their investigation, capturing a variety of tools and machinery in use. These audio samples were used to train the DRNN model, which was then adjusted to correctly categorize the construction noises. Amazingly, during evaluation on the test set, the proposed DRNN technique displayed outstanding performance. The model's overall accuracy was outstanding, coming in at 97%. This exceptional accuracy demonstrates the DRNN's ability in successfully differentiating between various audio signals connected to construction, outperforming previous state-of-the-art methods in this field. The research of Michele et al. has important applications for the construction sector and beyond. At construction sites, accurate audio classification can improve security, productivity, and monitoring. This method may provide real-time monitoring and early detection of possible difficulties by automating the recognition of different construction machine noises, which will improve site management and safety procedures. Their effective implementation of DRNN for audio classification also has larger implications for other audio-related activities, including environmental monitoring, sound event detection, and smart city applications. It is a valuable contribution to the field of audio signal processing and machine learning to be able to correctly categorize audio signals from difficult contexts like construction sites. In conclusion, Michele et al.'s research demonstrates the effectiveness of Deep Recurrent Neural Networks in audio classification tasks and offers a potential method for accurately classifying audio signals connected to construction. Their DRNN model's excellent accuracy highlights its potential to transform several industrial and smart city applications, establishing new benchmarks for audio-based categorization methods.

A innovative algorithm designed exclusively to assess the degree of similarity between various songs was presented by Sheikh et al. [18]. The GTZAN database, Emotive music dataset, and Music Audio Benchmark Dataset (MABD) were

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7
DOI: https://doi.org/10.17762/ijritcc.v11i7.7844
Article Received: 29 April 2023 Revised: 24 June 2023 Accepted: 02 July 2023

used in their study to test the algorithm's performance across various musical genres and styles.

The algorithm was specifically designed to examine and contrast the numerous musical qualities and aspects found in the songs. The method can quantify the similarity of songs by extracting pertinent data from audio signals, such as spectral properties, rhythm patterns, and melodic structures.

Sheikh et al. examined their suggested strategy on two different models to gauge its correctness. The accuracy of the first model in determining song similarity was 83.79%, while the accuracy of the second model was even better at 86.5%. Regardless of the many datasets used for evaluation, these remarkable accuracy rates demonstrate how well the system captures musical commonalities.

The method developed by Sheikh et al. has numerous real-world uses. Numerous domains, including music recommendation systems, music information retrieval, and content-based music search engines, might greatly benefit from the capacity to reliably identify song similarities. By recommending related songs based on user preferences, it can improve users' musical discovery experiences and promote a deeper understanding of musical trends and patterns across many genres.

Furthermore, the algorithm's application to various datasets shows that it is resilient and flexible enough to accommodate various music collections. Since music datasets can differ greatly in terms of genre distribution, audio quality, and cultural influences, this adaptability is essential in real-world applications.

The study by Sheikh et al. makes an important contribution to computational musicology and music information retrieval. Their work has the potential to increase the state of the art in music-related research and applications, ultimately boosting our entire musical experience and knowledge by creating an accurate and adaptable algorithm for song similarity analysis.

In order to handle a specific use case where users can play musical notes or audio pieces on an instrument, Preeth et al. [19] devised a model. The main goal of their research was to create a system that could precisely identify the traits of various musical keys. The researchers used the LibROSA Python module, a potent tool for music and audio analysis, to accomplish this goal. Key detection tasks are ideally suited for LibROSA since it offers a wide variety of functionalities for processing and extracting important information from audio signals. The Preeth et al. model probably goes through numerous levels of processing and analysis. First, the musical instrument's audio input is preprocessed with LibROSA to remove crucial traits and qualities. To prepare the audio for further analysis, this step entails activities including pitch extraction, time-domain analysis, and feature engineering. The model is then trained using machine learning techniques to discover the distinctions and patterns connected to various musical keys. This entails feeding the model labeled data including audio samples in various keys so that it can recognize and learn to distinguish between key-specific properties. To gauge the model's effectiveness in key detection, new, previously unheard audio samples are used for testing. The model's ability to correctly identify musical notes is evidence of how well it can distinguish between various audio sources. The research of Preeth et al. has important applications. Real-time accompaniment production, automatic transcription of musical works, and intelligent music tutoring systems can all benefit from the model's capacity to recognize musical keys. It offers a useful tool that helps musicians, students, and music lovers study and comprehend the structure of musical compositions. Preeth et al.'s research shows the promise of open-source tools in enhancing music analysis and computational musicology by utilizing the LibROSA Python package. Machine learning algorithms and strong libraries like LibROSA work together to promote innovation in music technology, opening up new avenues for musical discovery and artistic expression. In conclusion, Preeth et al.'s approach, notably in key detection, marks a substantial advancement in the field of autonomous music analysis. Their work advances the creation of intelligent music systems and offers insightful knowledge about the peculiarities of various musical keys, ultimately enhancing our appreciation and comprehension of music.

All the above-mentioned papers are using different machine learning algorithms and deep learning techniques to recognize the type of the songs with specific accuracy. All of them focused their research on different Indian classical songs and a few rags but no research on Borgit or identification of rag present in Borgit is carried out. For example in [6] they have used Hindustani musical dataset, while we use Borgit audio files for extraction of feature. They have used KNN and SVM method. In our work we used k-means clustering for grouping the Borgit which are similar in terms of FFT and ZCR as they are more specific to rags in Borgit. ZCR, FFT feature value helped us in better identification of ragas. In [16] Jothimani, S. and Premalatha, K., has worked on pre-processed speech signals while we work on Assamese Borgit audio files with Pitch and its significance of mean, FFT and ZCR features for rag identification. Thus by taking help from the above mentioned papers and experimenting on some other acoustic parameters of our own we have come to a few edge cutting conclusions which are mentioned at Result and Discussion section of this manuscript.

III. PROPOSED METHOD

Authors had gone through a number of techniques and analyzed different parameters used to recognize rag in Borgit. The

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7
DOI: https://doi.org/10.17762/ijritcc.v11i7.7844
Article Received: 29 April 2023 Revised: 24 June 2023 Accepted: 02 July 2023

proposed problem definition, dataset used and different methods used to extract features of the rags and the Borgits are discussed below.

A. Problem definition

In general, it becomes challenging for common people to identify the rag used in Assamese folk music. Most of the time ascending and descending scales are utilized together in an artistic way, for instance, many rags may share the same notes and even identical distinctive phrases[20]. Determining the characteristics of each rag in terms of audiotary features and recognize the particular rag present in a Borgit is main focus of this research.

B. Dataset preparation

The authors of this paper has collected around 200 (Two hundred) Borgits from different sources which are original and in recorded format. Three base rags “Ahir Bhairav’’, “Ashowaree” and ”Dhanshree” are selected for this experiment. Initially four to nine Borgits for each of these rags and their audio clips were selected and later they were collected from different confidential sources. Three Indian rags which we are going to search in the Borgits are also original in nature(without mixing with other rags). They were collected as mp3 file and were converted to .wav format thus making them easy to process on digital platform. The rags considered for our experiment is approximately 30 sec duration (as the same tune may repeat many times) and the experimented Borgits are with their full length audio.

C. Feature extraction

Feature extraction from audio file means extraction of similar patterns (tune that is repeated in that audio signal). These patterns will be coded using different software and extracted as features of the song. Two rags can be distinguished with the help of these features and the features can be built up from the attributes of the audio signal and will be used to identify the Borgit.

The group of rags is called “thaat”. There are approximately 4 lakh rags in Indian classical music[21].

We have selected four acoustic parameters based on our literature survey and could measure their values in each rag and Borgit using software tools and statistical functions or by python programming. Audio features used for the experiment are:

1) Pitch Contour
2) Zero Crossing Rate (ZCR)
3) Root-Mean Square (RMS)
4) Fast Fourier transform (FFT)

The acoustic features we are considering here are Pitch contour, Zero Crossing Rate(ZCR), Fast Fourier Transform (FFT) and Root Mean Square (RMS). Feature evaluation was carried out on a few frequently sung popular Borgits with the rags “Ashowaree’’, “Ahir Bhairav” and “Dhanshree”. Table I shows our experimented audio samples for rag and Borgits along with their evaluated acoustic values.

D. Pitch extraction

The frequency of a sound wave directly affects its pitch. A higher pitch derives from a greater frequency, whereas a lower frequency produces a lower pitch [22]. As a result, a sound wave's frequency plays a key role in determining how pitch is perceived.

The scientists used a statistical technique based on the correlation coefficient to create an automatic system to recognize rag in a Borgit. This process probably entailed comparing and analyzing the features or frequency patterns of the audio waves relating to various ragas.

The writers' research led them to the melodic element known as "shadja," which is distinctive and present in every rag. A "shadja" is a certain swara (note or pitch) that functions as the foundational note or base note for a particular raga. This shadja note is a distinguishing feature of each raga's melodic structure, making its existence helpful in differentiating and identifying various ragas.

The authors' automated rag identification approach, which makes use of the correlation coefficient and the identification of the shadja note, is predicted to have a big impact on Indian classical music. The system can efficiently classify and identify ragas from audio inputs by identifying the fundamental swara associated with each raga, supporting a variety of applications like music recommendation, content-based music retrieval, and musicological study.

E. Pre-processing for Pitch contour extraction

A statistical measure that is frequently used to quantify the link between two sets of data points is the correlation coefficient. The Correlation Coefficient is used to assess the similarity between two series of fundamental frequencies of several ragas in the context of rag identification in Borgit structures.

The Correlation Coefficient can be used to determine whether or not the fundamental frequencies of several ragas played on a borgit reflect a consistent pattern. The fundamental frequencies of the ragas may exhibit a similar pattern or trend if the correlation coefficient is high, indicating a significant positive link. A low correlation coefficient value, on the other hand, denotes a weak or negligible correlation and indicates that the fundamental frequencies of the ragas are significantly different.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7
DOI: https://doi.org/10.17762/ijritcc.v11i7.7844
***Article Received**: 29 April 2023 **Revised**: 24 June 2023 **Accepted**: 02 July 2023*

The ability for researchers and musicians to recognize similar patterns or melodic structures that define a certain raga is made possible by the use of correlation coefficient to compare fundamental pitches. The Correlation Coefficient sheds light on the fundamental musical traits that set one raga apart from another by assessing the similarity or dissimilarity between the fundamental frequencies of various ragas.

In general, the use of the correlation coefficient to identify rags utilizing Borgit structures improves our comprehension of Indian classical music and makes it easier to create automatic systems for raga analysis and recognition. We can better understand the connections between musical parts and deepen our understanding of the intricate melodies and nuanced ragas by utilizing statistical techniques like the correlation coefficient.

This experiment was carried out with all the Borgits. By evaluating pitch for each of the Borgit we studied and analyzed variation of pitch on each rag and Borgit.

A screenshot of our experiment on rag “Ahir Bhairav”, its waveform, spectrogram and pitch contour is shown in figure 1.

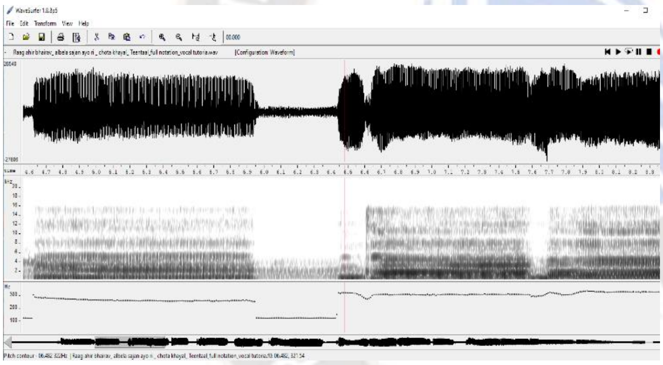

Figure 1. Waveform, Spectrogram and Pitch Contour of Rag “Ahir Bhairav”

The Pitch values of all our sample Borgits with rag “Ahir Bhairav’’, “Ashowaree” and “Dhanshree” are obtained. A lower standard deviation denotes more stable pitch contour whereas higher value means greater variability or diffusion in pitch values.

Significance of Mean and Standard Deviation

Mean and Standard Deviation are two important metrics that can extract features of audio signal [29]. Mean value gives loudness of the audio signal. A higher mean value increases the dynamic range and a lower value decreases dynamic range. For an audio signal standard deviation gives variation of the signal around the mean value. It plays crucial role in removing noises from audio segment. It is also used as a parameter to access audio signal quality[30]. Figure 2 shows a screenshot that evaluates the Mean and Standard deviation with pitch contour, waveform and spectrogram of the Borgit “Utho Utho bapu Chanda bayan”.

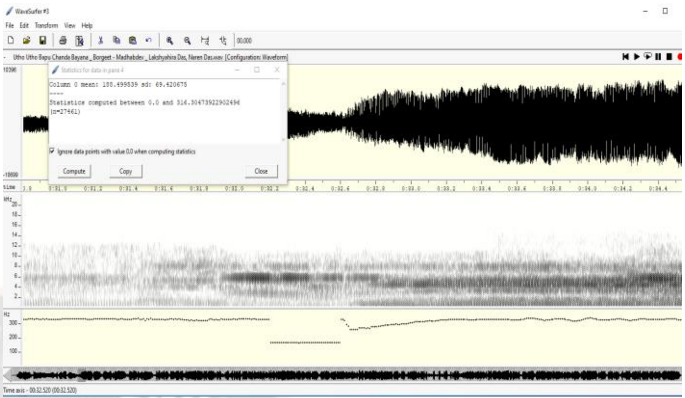

Figure. 2. Mean and Standard deviation with pitch contour, waveform and spectrogram of Borgit “Utho Utho bapu Chanda bayan”

F. Preprocessing for FFT, ZCR and RMS evaluation

To get the values for FFT, ZCR and RMS initially we load the audio file in Anaconda jupyter environment. First we loaded the base rag in jupyter(python) and analyze it in python. Then all the Borgits are loaded in same environment of jupyter python for the evaluation of FFT, ZCR and RMS values.

FFT Extraction

The Discrete Fourier Transform (DFT) of a data series is computed using the Fast Fourier Transform (FFT), a strong and effective method [27]. A mathematical procedure called the Fourier transform enables us to examine a signal's frequency constituents. We may create the signal's power spectrum by taking the square of the magnitude spectrum that the Fourier transform produced [28]. The frequency composition and dispersion of a signal can be learned a lot from the power spectrum.

The researchers used FFT to extract frequency domain data from the audio samples in the context of evaluating Borgit music. They used Python to carry out the FFT computation using the Jupyter platform, which is a well-liked setting for data analysis and visualization.

In Table 1, along with other pertinent factors, the results of the FFT method are shown. These numbers probably comprise the frequency components and their respective magnitudes, enabling the researchers to examine the spectral properties of the Borgit composition "Shayama komal lochan" and its accompanying raga "Ahir Bhaiarav."

The magnitude spectrum graph for the Borgit composition under investigation is shown in Figure 3. The magnitude spectrum graph shows the audio signal's various frequency components' amplitudes. Understanding the music piece's spectral properties requires having a visual depiction of how energy is divided among its many frequencies.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7
DOI: https://doi.org/10.17762/ijritcc.v11i7.7844
***Article Received**: 29 April 2023 **Revised**: 24 June 2023 **Accepted**: 02 July 2023*

The magnitude spectrum graph can be used to analyze the major frequencies, harmonics, and other spectral features that are unique to the ragas "Ahir Bhaiarav" and "Shayama komal lochan." This study is essential for recognizing and differentiating various ragas and comprehending their distinctive musical features.

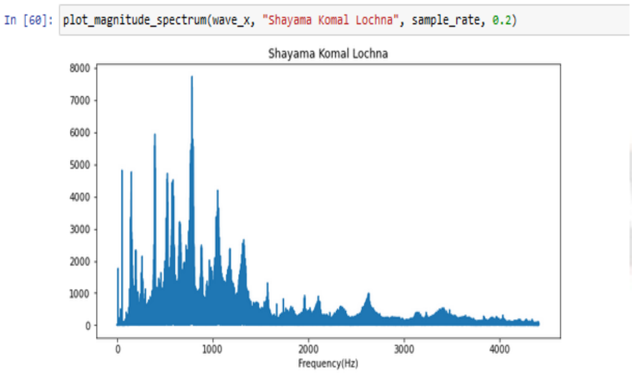

Figure 3. Snapshot of magnitude Spectrum graph of Borgit “Shayama komal lochan” with rag “Ahir Bhaiarav”

ZCR Extraction

Zero crossing rate (ZCR) is a useful feature commonly used in digital signal processing and audio analysis. It is defined as the number of times that a signal crosses the horizontal axis per unit time. The zero crossing rate has various applications in the field of audio processing including speech recognition, music genre classification and speaker identification [23]. ZCR can be calculated using the equation [24]

$$ZCR = 0.5 \sum_{i} |sgn(si) - sgn(si+1)| \quad (1)$$

In equation (1), i takes values from t∗S to (t+1)∗(S−1), sgn is the Signum function, s_i is the amplitude of the ith frame and S is the number of frames in the audio signal.

The extracted ZCR values of the three rags and the Borgits are shown in table 1.

ZCR evaluation is performed here only with 100 samples in the range 15000 to 15100, taken from total 798411 wave samples. So based on n0 and n1 values, if we count in the graph we get how many zero crossing number in that graph.

ZCR values and graphs are shown in figure 4 for Borgits “Shayama komal lochana” with this rag “Ahir Bhairav”

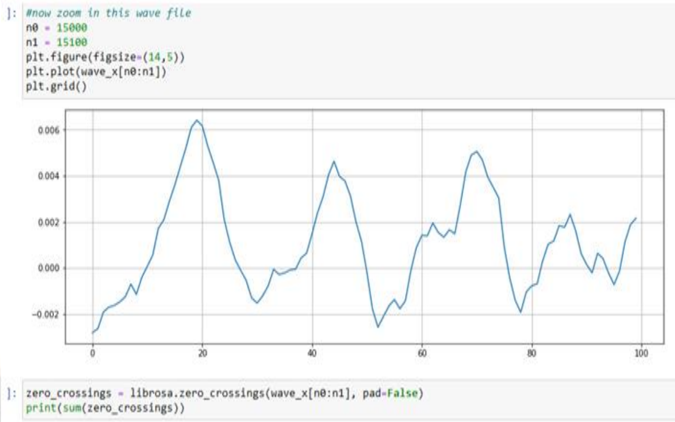

Figure 4. Graph and ZCR value of Borgit “Shayama komal lochana”

RMS extraction

Root Mean Square (RMS) is a mathematical concept used to determine the average magnitude of a set of values [25]. It is calculated by taking the square root of the mean of the squared values. Dynamic range of a sound can be measured with RMS value. Dynamic range of an audio signal gives difference between loudness and softness of the sound. For a given signal, x = {x1, x2….xn}, the RMS value, X_{rms}, is shown in equation (2) and (3). Equation (3) is decomposed signal of original signal x in equation (2).

$$X_{rms} = \sqrt{\frac{x^2}{n}} \quad (2)$$

$$X_{rms} = \sqrt{\frac{1}{n}(x1^2 + x2^2 + \cdots + xn^2)} \quad (3)$$

In this experiment RMS values are extracted separately for both the rags and their corresponding set of Borgit files.

IV. RESULT AND DISCUSSIONS

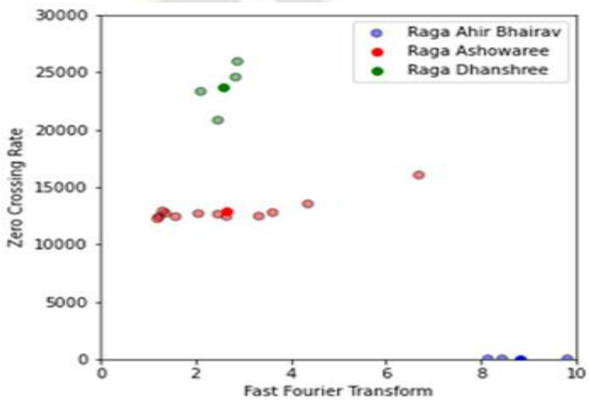

Figure 5. K – Means clustering on ZCR and FFT values of rags and Borgits

In this paper we are going to explain our experiment to determine whether the above mentioned four acoustic features of speech signal are also useful to identify the rag implanted in a Borgit or not.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7
DOI: https://doi.org/10.17762/ijritcc.v11i7.7844
***Article Received**: 29 April 2023 **Revised**: 24 June 2023 **Accepted**: 02 July 2023*

Table 1 shows experimental results performed on each of the three rags and the Borgits. As mentioned earlier the feature values evaluated with standard deviation, Fast

Fourier Transform(FFT), Zero Crossing Rate(ZCR) and Root Mean Square (RMS) are listed in this table. After

performing analysis and observations on these values researchers have noticed that some feature values of rag and Borgits with that rag shows in close proximity. This is apparent for FFT and ZCR values for all the rags and Borgits. Standard deviation values of pitch contour for rags and Borgits always do not show identical behavior. It means two Borgits with two different rags do not show much differences in their standard deviation values on pitch contour. So there is scope for more research on this feature in future. Table 1 below shows all the feature values the researchers have evaluated for the songs and rags.

Table 1. Experimental Feature set with values

Sl no	Rag Name	Borgit Name (duration)	Pitch Contour	Fast Fourier Transform(Hz)	Zero Crossing Rate (Hz)[Sum of zcr]	Root Mean Square (RMS)
1	Ahir Bhairav(30sec)	→	66.892	8.12059	11	(1,1560)
		Shayama Kamal Lochana (Borgit)	60.074	9.81026	11	(1,1560)
		Utho Utho Bapu (Borgit)	69.420	8.142235	10	(1,1208)
		Chintohu Govinda (Borgit)	64.951	8.4450	11	(1,1748)
2	Ashowaree (30sec)	→	71.911	2.8354	23335	(1,1107)
		Jaya Jaya Jadava (Borgit)	72.987	2.1188	23335	(1,1513)
		Anande Govinda bai (Borgit)	79.840	6.6921	16043	(1.1370)
		Rama Gosain karoho gohari (Borgit)	72.227	2.4714	20831	(1,1181)
		Ramka bani japahu mon (Borgit)	76.65	2.8822	25946	(1, 1411)
		Suna Suna Re Sina Bairi Pramana(Borgit)	71.6475	2.8444	24587	(1. 1512)
3	Dhanashri(30 sec)	→	52.9977	1.6683	12905	(1, 1543)
		Haripada pankaja(Borgit)	52.9979	3.6200	12765	(1, 1433)
		Narayana kahe bhakti (Borgit)	53.5313	3.3384	12458	(1, 1543)
		Narayana kahe bhakti (Borgit)	53.5313	3.3384	12458	(1, 1543)
		Naryana ke guna janaba(Borgit)	57.7032	1.5824	12410	(1, 1544)
		Pamaru Monai Kamane Hariro(Borgit)	58.6589	1.233	12422	(1, 1655)
		Hari Guna Koise(Borgit)	54.7246	1.374	12700	(1, 1544)
		Balahuram (Borgit)	55.3425	2.06	12683	(1, 1765)
		Hari guna gayata(Borgit)	53.766	2.65	12430	(1, 1743)
		Mana meri rama charana(Borgit)	56.876	2,47	12631	(1, 1654)
		Parabhate bihare chale(Borgit)	57.644	1.20	12907	(1, 1545)
		Rama meri hridaya(Borgit)	55.765	1.31	12226	(1, 1544)
		Sajere sakhi nanduka(Borgit)	58.988	4.31	13526	(1, 1632)

A. *Apply K-Means clustering*

We have used k-means clustering algorithm to discover which are the most influential parameters achieved in our experiment to recognize rag of the Borgit. K-Means clustering is an unsupervised clustering technique. After observing numeric distribution of the parameters (Mean value, FFT, ZCR and RMS), we determined that FFT and ZCR are the most trivial parameters to recognize rags. So k-means clustering was applied on FFT and ZCR values and it gave correct results and matched with which rag it really belongs to. So K-Means

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7
DOI: https://doi.org/10.17762/ijritcc.v11i7.7844
***Article Received**: 29 April 2023 **Revised**: 24 June 2023 **Accepted**: 02 July 2023*

clustering proves that FFT and ZCR are the unique features for identification of rags in the Borgits. Figure.5. showsa the K – Means clustering on ZCR and FFT values of rags and Borgits **Critical Discussion**: Acoustic feature identification plays a crucial role in recognizing and categorizing rags, contributing to the development of automated systems for rag recognition. However, it is important to critically examine the effectiveness and limitations of this approach.

One of the key challenges in acoustic feature identification is the inherent complexity and variability of rags themselves. Rags are characterized by a combination of melodic patterns, tonal nuances, and expressive elements, making their accurate identification a complex task. Acoustic features such as pitch, timbre, and spectral content are commonly engaged to capture the distinctive characteristics of rags. While these features provide

valuable information, they may not fully summarize the nuanced and subjective aspects of rag performance, which are often carried through subtle variations in intonation, ornamentation, and phrasing.

Another critical aspect is to consider the diversity and context-dependent nature of rag. Rags can be interpreted in a variety of ways by different musical cultures, geographical areas, and even individual musicians. Methods for identifying acoustic features must take these variances into consideration and be flexible enough to accommodate various stylistic nuances. Failure to do so could result in inaccurate or insufficient rag recognition.

After analyzing all the factors above the following conclusions are achieved:

i. Borgits are supplemented with rag and they are sung with melodious tune of rag. In Indian culture most of the classical music are enriched with rag, this experiment proves that four measurable acoustic parameters FFT, ZCR, Standard Deviation of Pitch Contour and RMS can be extracted and used as features of rag. In case of RMS evaluation we observed they (RMS feature values) provide only range of the total samples, not a specific value. So we have not noticed important characteristics of RMS in rag identification.

ii. We have evaluated values of all the above mentioned parameters from three selected rags “Ahir Bhairav’’, “Ashowaree” and ”Dhanshree” and nineteen Borgits sung with tune of those rags. After performing some preliminary observations we came to conclusion that mean or Standard deviation of pitch contour as well as RMS are not contributing much in rag identification. On the other hand FFT and ZCR actively participate in recognition of Rag so the Borgits sung with those rag.

iii. We have used K-Means clustering algorithm and noticed that FFT and ZCR can be considered for clustering with three rags as centroid or initial seed of cluster, the algorithm was able to cluster all the Borgits correctly.

iv. We tried to prove whether there is any significance of MFCC in rag and Borgits identifications, we evaluated MFCC values but observed no significant patterns hence discarded MFCC in identifying rag from Borgit.

V. CONCLUSION

In this paper, Borgits of Assamese culture and rags in Borgits are studied and analyzed. Here only three rags and nineteen Borgits with these rags were considered for this basic experiment. The suggested method makes use of a number of significant elements that were collected from the audio signals for the analysis and identification of rags and Borgits. These characteristics include Standard Deviation of Pitch Contour, Root Mean Square, Fast Fourier Transform, and Zero Crossing Rate. The approach seeks to identify commonalities between various rags and Borgits based on their auditory patterns by examining these aspects. When comparing the similarities between rags and Borgits, standard deviation—a measurement of the spread or variability of pitch contour—is thought to be a key factor. However, in this particular experiment, its efficacy is constrained by the values of two or more rags that overlap. Due to the data points' overlap, it may be difficult to distinguish between several rags based just on their standard deviation values. Three more similarity metrics, FFT, ZCR, and RMS, have been investigated for audio analysis in this method in addition to Standard Deviation. The frequency content of the audio signals may be examined using FFT, while the rate of signal change can be determined using ZCR and the total energy of the audio can be determined using RMS. ZCR and FFT have shown to be very efficient in clustering and detecting similarities between rags and experimental Borgit audio samples among all the acoustic parameters. The method's capacity to differentiate various audio patterns and classify them into meaningful clusters based on their similarity is substantially influenced by these properties. Diverse audio elements and strategies for Borgit recognition are continually being explored in the broad domains of Music Information Retrieval (MIR). The development of content-based music retrieval, music recommendation engines, intelligent music systems, and other applications that improve human comprehension and relationship with music depend heavily on MIR. The proposed approach, in the context of Indian Classical Music, highlights the significance of utilizing different audio aspects like FFT, ZCR, RMS, and Standard Deviation for

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7
DOI: https://doi.org/10.17762/ijritcc.v11i7.7844
***Article Received**: 29 April 2023 **Revised**: 24 June 2023 **Accepted**: 02 July 2023*

locating and examining rags and Borgits. Multiple acoustic factors can be used to produce more reliable and precise results, advancing music information retrieval and enhancing our understanding of the rich musical history of rag and bogit.

***Research limitations**:* Three rags and nineteen borgits are considered for all the experiments here but there are more Borgits and in future we will work with all of them.

***Social implication**:* Use of Artificial technologies similar to Automatic rag recognition using AI technology may face challenges from social acceptance due to fear of losing jobs, unfamiliarity with technology etc. Affected individuals and adequate support to them will help to adopt a technology aided society in future.

In conclusion, it will be essential to incorporate expert knowledge and take into account the perceptual components of rag performance if automated rag recognition using acoustic data is to advance.

Future work

- Several experiments can be carried out to measure various rags with corresponding Borgits and identify its similarities.
- A larger database can be considered and also a machine learning model can be designed to classify and analysis rags and Borgit.
- Analysis and comparison of the different extraction and classification process in Python and Matlab can be done.
- An automatic rag recognition tool for Borgit can be designed.

REFERENCES

[1]. B. Bhattacharyya, 2020. Sankardeva: Some Reflections on Religious Language.
[2]. S. Kakoty, 2018. DIMENSIONS OFTHE MOBILE THEATRE OFASSAM (Doctoral dissertation).
[3]. Insights IAS "INSIGHTSIAS "(available at: Sattriya - INSIGHTSIAS (insightsonindia.com) (Accessed 1st may, 2023)
[4]. N. Saikia, 2023. Indian Music: Ancient Traditions, Contemporary Rights. Contemporary Rights (April 23, 2023).
[5]. A.V.Bang,R.G.Purandare,A.K.Ratnaparkhi,(2023) An algorithm for enhancement of audio content classification, Bulletin of Electrical Engineering and Informatics, 12 (4), pp. 2262-2268. DOI: 10.11591/eei.v12i4.3663.
[6]. K.C. Baral, ed., 2023. Cultural Forms and Practices in Northeast India. Springer Nature.
[7]. Joshi, D. et al., 2021, February. Indian Classical Rag Identification using Machine Learning. In ISIC (pp. 259-260).
[8]. S.K. Baruah, (2022), Mahapurush Sri SriShankaradeva and Mahapurush Sri Sri Madhavadeva Borgeet, 7th Edition, Jyoti Publication.
[9]. E. Patel, and S. Chauhan, 2017. Raag detection in music using supervised machine learning approach. International Journal of Advanced Technology and Engineering Exploration, 4(29), p.58.
[10]. Joshi, D. et al., 2021, February. Indian Classical Rag Identification using Machine Learning. In ISIC (pp. 259-263).
[11]. S. Chakrabarty,et al.,2022. An Approach to Discover Similar Musical Patterns. IEEE Access, 10, pp.47322-47339.
[12]. Dodia, H. et al., 2020. Identification of Rag by Machine Learning with Chromagram. International Research Journal of Engineering and Technology (IRJET), 7(05).
[13]. K.C. Waghmare and B.A. Sonkamble, 2017. Rag identification techniques for classifying Indian classical music: A survey. International Journal of Signal Processing Systems, 5(4), pp.1-6.
[14]. Ipseeta Nanda, Monika SIngh, Lizina Khatua. (2023). Automated Irrigation System Using IoT Cloud Computing. International Journal of Intelligent Systems and Applications in Engineering, 11(2s), 360–365. Retrieved from https://ijisae.org/index.php/IJISAE/article/view/2728
[15]. Kumar, M.S. and Devi, M.S., 2020, Rag Recognition Using Machine Learning. In 2020 (0950-0707).
[16]. D.P. Shah, et al., 2021. Rag Recognition in Indian Classical Music Using Deep Learning. In Artificial Intelligence in Music, Sound, Art and Design: 10th International Conference, EvoMUSART 2021, Held as Part of EvoStar 2021, Virtual Event, April 7–9, 2021, Proceedings 10 (pp. 248-263). Springer International Publishing.
[17]. S. Jothimani, and K. Premalatha, 2022. MFF-SAug: Multi feature fusion with spectrogram augmentation of speech emotion recognition using convolution neural network. Chaos, Solitons & Fractals, 162, p.112512.
[18]. M. Scarpiniti, et al., 2021, January. Deep recurrent neural networks for audio classification in construction sites. In 2020 28th European Signal Processing Conference (EUSIPCO) (pp. 810-814). IEEE.
[19]. M. Sheikh Fathollahi and F. Razzazi, 2021. Music similarity measurement and recommendation system using convolutional neural networks. International Journal of Multimedia Information Retrieval, 10, pp.43-53.
[20]. P. Raguraman, et al., 2019, March. Librosa based assessment tool for music information retrieval systems. In 2019 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR) (pp. 109-114). IEEE.
[21]. T.M. Krishna, and V. Ishwar, 2012. Carnatic music: Svara, gamaka, motif and rag identity. In Serra X, Rao P, Murthy H, Bozkurt B, editors. Proceedings of the 2nd CompMusic Workshop; 2012 Jul 12-13; Istanbul, Turkey. Barcelona: Universitat Pompeu Fabra; 2012.. Universitat Pompeu Fabra.
[22]. M.Altayeb,A.Al-Ghraibah, (2022) Classification of three pathological voices based on specific features groups using support vector machine, International Journal of Electrical and Computer Engineering, 12 (1), pp. 946-956. DOI: 10.11591/ijece.v12i1.pp946-956.
[23]. G.Agarwal,S.Gupta,S.Agarwal,A.K.Rai, (2023) Emotion classification for musical data using deep learning techniques, International Journal of Reconfigurable and Embedded

Systems, 12 (2), pp. 240-247. DOI: 10.11591/ijres.v12.i2.pp240-247.

[24]. P.R.K.Rao,S.Ravi,T.Haritha,(2021),Purging of silence for robust speaker identification in colossal database, International Journal of Electrical and Computer Engineering, 11 (4), pp. 3084-3092. DOI: 10.11591/ijece.v11i4.pp3084-3092

[25]. Dr. V. Arthi. (2020). A Novel Channel Estimation Technique in MIMO-OFDM Mobile Communication Systems. International Journal of New Practices in Management and Engineering, 9(02), 08 - 14. https://doi.org/10.17762/ijnpme.v9i02.84

[26]. M.J.AlDujaili,A.Ebrahimi-Moghadam,A.Fatlawi, (2021),Speech emotion recognition based on SVM and KNN classifications fusion, International Journal of Electrical and Computer Engineering, 11 (2), pp. 1259-1264. DOI: 10.11591/ijece.v11i2.pp1259-1264

[27]. N.Benayad,Z.Soumaya,B.D.Taoufiq,A.Abdelkrim, (2022) Features selection by genetic algorithm optimization with k-nearest neighbour and learning ensemble to predict Parkinson disease International Journal of Electrical and Computer Engineering, 12 (2), pp. 1982-1989. DOI: 10.11591/ijece.v12i2.pp1982-1989.

[28]. Liu, Wei, et al., 2020. On-line estimation of road profile in semi-active suspension based on unsprung mass acceleration. Mechanical Systems and Signal Processing, 135, p.106370.

[29]. Paul Garcia, Ian Martin, Laura López, Sigurðsson Ólafur, Matti Virtanen. Personalized Learning Paths Using Machine Learning Algorithms. Kuwait Journal of Machine Learning, 2(1). Retrieved from http://kuwaitjournals.com/index.php/kjml/article/view/166

[30]. Mohapatra, B.N. and Mohapatra, R.K., 2017, February. FFT and sparse FFT techniques and applications. In 2017 Fourteenth International Conference on Wireless and Optical Communications Networks (WOCN) (pp. 1-5). IEEE.

[31]. Z. Yang, et al., 2022. A reconstruction method for graph signals based on the power spectral density estimation. Digital Signal Processing, 122, p.103347.

[32]. M.A.A. Albadr, et al., 2021. Mel-frequency cepstral coefficient features based on standard deviation and principal component analysis for language identification systems. Cognitive Computation, 13, pp.1136-1153.

[33]. H.A. Abdulmohsin, 2021. A new proposed statistical feature extraction method in speech emotion recognition. Computers & Electrical Engineering, 93, p.107172.