Journal of Information & Optimization Sciences
ISSN 0252-2667 (Print), ISSN 2169-0103 (Online)
Vol. 44 (2023), No. 3, pp. 341–353
DOI : 10.47974/JIOS-1345

Artificial intelligence-based classification performance evaluation in monophonic and polyphonic indian classical instruments recognition with hybrid domain features amalgamation

Abhijit V. Chitre *
Department of Electronics and Telecommunication
Vishwakarma Institute of Information Technology
Pune
Maharashtra
India

Kirti Wanjale †
Aradhanaa Deshmukh §
Department of Computer Engineering
Vishwakarma Institute of Information Technology
Pune
Maharashtra
India

Shyamsunder P. Kosbatwar ‡
Department Computer Science and Engineering
Dnyanshree Institute of Engineering and Technology
Satara
Maharashtra
India

Anup Ingle @
Department of Electronics and Telecommunication
Vishwakarma Institute of Information Technology
Pune
Maharashtra
India

* *E-mail:* abhijit.chitre@viit.ac.in (Corresponding Author)
† *E-mail:* kirti.wanjale@viit.ac.in
§ *E-mail:* araddhanadeshmukh@mmcoe.edu.in
‡ *E-mail:* shyamkosbatwar@gmail.com
@ *E-mail:* anup.ingle@viit.ac.in

Sheela Hundekari [#]
MIT-ADT University
Pune
Maharashtra
India

Abstract

In computer music, instrument recognition is a critical part of sound modeling. Pitch, timbre, loudness, duration, and spatialization are all components of musical sounds. All of these components play a significant part in determining the quality of the tonal sound. It is possible to alter the first four parameters, but timbre always poses a challenge [6]. It was inevitable that timbre would take center stage. Musical instruments are distinguished from one other by their distinct sound quality, independent of their pitch or volume. To distinguish between monophonic and polyphonic music recordings, this method might be used.

In Musical Information Retrieval, classification plays one of the critical role. Monophonic instrument classification can be found in literature with quiet a substantial combinations of features and classifiers. Polyphonic instrument classification witnessed less references in the literature and is still an area to be explored specifically when it comes to Indian Classical domain. The present paper exactly focusses on this experimentation.

Several Indian instruments were used to produce training data sets for the proposed approach's evaluation purposes. Among the instruments utilized are the flute, harmonium, and sitar. Statistical and spectral factors are used to classify Indian musical instruments along with the Artificial Intelligence-based methods. Hybrid features from multiple domains that extract essential musical properties are extracted. Accuracy is demonstrated through an Indian Musical Instrument SVM and GMM classification. With monophonic sounds, SVM and Polyphonic produce an average accuracy of 89% and 91%. GMM outperforms SVM in monophonic recordings by a factor of 96.33 and polyphonic recordings by a factor of 93.33, according to the results of the studies. The future scope of this recognition framework can be an Artificial Intelligence System with a system linked with the Industrial Internet of Things (IIOT) framework to develop a standalone system or application which can be used for real-time classification of instruments.

Subject Classification: *68T01 General topics in artificial intelligence,68T01 General topics in artificial intelligence.*

Keywords: *Indian classical instruments, Support vector machine (SVM), Gaussian mixture modelling, K-nearest neighbor classifier (KNN).*

1. Introduction

Musical data mining can benefit from the identification of musical instruments. We can categorize or search in a musical audio database without having to include descriptive data in the audio stream, by virtue of

[#] *E-mail:* **sheela.hundekari@mituniversity.edu.in**

to this identification technique. The identification of musical instruments is an important part of the music information retrieval process. The knowledge we may uncover by understanding the instruments used in a particular piece of music is priceless. Structure coding, automated signal annotation, and a musician's tool are all examples of musical content analysis applications. There are several additional fields of study that have ties to musical instrument identification systems. Artificial Intelligence-based systems are becoming the Multi-domain feature categorization and the accuracy of classifiers is presented in this study. Figure 1 depicts the overall system of analysis and classification. The samples of music pieces from instruments are recorded. In the second stage features extraction of these samples is carried out and in the third stage, the classification with KNN, SVM, and GMM is carried out to check the combination of features and classifier producing the highest accuracy. This system has experimented with monophonic as well as of polyphonic recorded musical notes.

Music from India is one of the world's oldest forms of expression. Indian music relies heavily on the use of musical instruments. The vast majority of Indian musical instruments were created over time. Several sorts of instruments, including string, wind-blown, and percussion instruments, are used in Indian music.[2] The flute, harmonium, and sitar are three of the most often utilized Indian classical instruments. Instrument classification and automatic indexing have been used in music information retrieval systems for a long time. [1].

This research aims to carry out the study prominently on Indian classical instruments. Also, the most critical part of the research presented

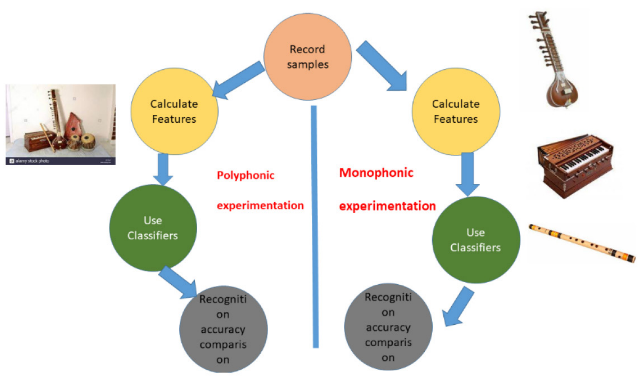

Figure 1

Block Diagram of the Proposed System

in the paper is to work with polyphonic musical pieces and achieve the classification accuracy to an acceptable value. Work is scarce in the literature focusing on a polyphonic musical instrument. This makes this research even more prominent.

The future scope of the research work could be an extension of these feature extraction mechanisms based on Artificial Intelligence systems and deployment of this entire mechanism as a system to the Internet of Things (IoT) enabled devices, where the application can be utilized for classification Figure 1 shows block diagram of the system.

2. Literature Survey

Significant research has been conducted in the literature to investigate various strategies for automatic musical instrument recognition. The majority of the research is on Western musical instruments such as the piano, clarinet, trumpet, guitar, and violin.

Daniel Mintz's study [5] offers a ground-breaking technique for timbre-based sound synthesis. The proposed method for determining timbre quality is heavily influenced by the spectral shape, amplitude envelope, and Spectro-temporal fluctuation, and timber separation converts these parameter values to direct envelopes. Juan José Burred et al. [1] discussed a system for encoding timbre models, which was necessary for classifying instrument samples. The proposed work comprises a spectro-temporal envelope with PCA for unknown samples and a GMM-based classifier. When applied with a five-class database, this approach claims to have a classification accuracy of 94.9 percent. The technique enables precise timbre modeling via the use of many descriptors. We extract spectral, temporal, spectro-temporal, and statistical characteristics. Swe Zin Kalayar Khine et al. [4] presented a system for evaluating the timbre quality of a performer. The proposed work comprises a voice detection algorithm that makes use of vocal parameter extraction from each song and a cepstral coefficient, which are both crucial for recognizing vocalist timbre. This parameter is classified using GMM. According to the classification, singer detection had a typical mistake rate of 12.2 percent. Poli [5] developed a method based on signal processing for determining the spectral envelope of MFCC data that can be utilized to identify singers. The algorithm's accuracy was determined by comparing sound samples detected manually and those detected automatically.

In the recent literature also the approaches to music classification can be observed. S.Prabhavathy et.al [10] demonstrated the classification with SVM and KNN, where the authors prominently focus on the MFCC coefficient. The accuracy of 95% is claimed by the authors. This study prominently is focused on identifying the musical genre. Lare Haidar[11] presented an accuracy of around 70% in identifying three instruments piano, drums, and flute. The deep learning approach is presented to carry out this experimentation. Convolutional Neural Network along with Mel spectrogram as input is described in the paper. Accuracy of around 90% is claimed in the paper by S.Prabhavathy et. al[18] by using the CNN model to recognize the category of instruments. Paper by Utsav Shukla et.al [13] demonstrates an effort to identify 4 similar String instruments (Acoustic Guitar, Cello, Violin, and Electric Guitar) in the audio recordings. Authors have used 1D MLPs and 2D CNN architectures for classifying the sounds and compared the performance of different audio features. Our experiments also show that using image-based transfer learning models like Inception and VGG gives better results among the aforementioned architectures**.**

S.Prabhvathy et.al [14] demonstrated the results of musical instrument classification using the pre-trained deep learning model Alexnet. Accuracy of around 90% is obtained. The work is carried out on 16 instruments.

Anubhav Chhabra et.al [16] presented the Drum instrument classification with machine learning. The work is on four classes of drum instruments. Spectral features are calculated and using Random Forest and Naïve Baye's classifier the classification is carried out. Accuracy of around 93% is claimed with Random Forest and 95% accuracy is claimed with Naïve Baye's classifier.

From the literature review, it is evident that some experiments were carried out on Western instruments. Also, many instruments in terms of different classes can be observed but there is hardly any evidence on the classification of polyphonic instruments where the input audio file itself has many instruments which we need to classify s which distinct instruments can be recognized by the proposed system. According to the literature review, polyphonic and monophonic hybrid feature-based classification can be studied as a novel paradigm for developing a classifier and feature pair that achieves the highest recognition accuracy[17-19].

3. Preliminaries

Raw data or input data have a random set of values. These statistics give a constant value which helps in characterizing the population i.e. data. When a set of values tend to cluster around some particular values, then this is useful for characterizing the major central tendency. Some statistics help in knowing the graphical shape of the population. The Statistical features considered for our work are Mean, Variance, Standard deviation, Skewness, Kurtosis, Average power, and Average Magnitude.

3.1 *Gaussian Mixture Model (GMM)*

A Bayesian classifier must be able to handle as many data dimensions, classes, and Gaussian components as is practical. Often referred to as the Gaussian components of a mixed model, a weighted sum of Gaussian probability density functions characterizes a Gaussian mixture model [4]. Supervised learning is used to train a classifier [13]. Classification using Bayesian methods is based on the idea that the class with the lowest probability should be chosen.

Assume a feature vector is to be classified into one of nine classifications [3]. In this case, let the feature vector be defined as the sum of the dimensions of the vectors be $f = [f_1, f_2, \ldots f_D]^T$ where D is the dimension of the vector. Posteriori probabilities can be computed with the Bayes formula as:

$$P(w_k \mid f) = \frac{p(f|w_k)P(w_k)}{p(f)} \tag{1}$$

Where, $P(w_k|f)$ is the pdf of class $w_k$in the feature space and $P(w_k)$ is the a priori probability, which tells the probability of the class before measuring any features. If it is unknown, then it is estimated from the proportion of class in the training set $p(f)$ emulates a scaling factor ensuring the true probabilistic nature of the posterior probability set.

Choosing a class with high posterior probability gives minimum error probability. Mean μ and variance σ^2 characterize the Gaussian probability density function. It can be defined as

$$\mathcal{N}(x;\mu,\Sigma) = \frac{1}{(2\pi)^{\frac{D}{2}} |\Sigma|^{1/2}} \exp\left[-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right] \tag{2}$$

where μ is the mean vector and Σ its covariance matrix.

GMM is the combination of several Gaussian distributions so different subclass inside one class can be represented by using the GMM model. PDF is defined as given below [8]

$$p(x,;\theta) = \sum_{c=1}^{C} \propto_c \mathcal{N}(x, \mu_c, \textstyle\sum_c) \tag{3}$$

where, $\propto_c$ is weight of component c, $0 < \propto_c < 1$ for all components. $\sum_{c=1}^{C} \propto_c = 1$.

4. Experimentation and Results

This study focuses on the identification of musical instruments. The most commonly utilized classical Indian instruments, such as the flute, harmonium, and sitar, are employed in this study. Initially, the sounds of these three instruments are gathered from standard recordings of them. These sounds have been preprocessed. An audio file of 3 seconds is obtained from the time frame in which the instrument is playing alone (monophonic) and converted to the.wav file format utilizing a file converter during preprocessing. The sampling frequency selected is 44.1 kHz. Audacity software is used to convert samples into 5 sec. samples. Some of the samples are prerecorded and some samples are recorded with the instrument being played. For this study, we have considered 300 musical samples i.e. 100 samples each form flute, harmonium, and sitar. We have divided this database into two i.e. training database and a testing database. The training database consists of 70 musical samples each form flute, harmonium, and sitar i.e. a total of 210 samples. The testing database consists of 30 musical samples each from flute, harmonium, and sitar i.e. a total of 90 samples. For polyphonic experimentation, 300 samples are used which have a combination of three and four instruments. 210 samples are used for training and 90 samples for testing.

4.1 *Classifier Result*

NN is also a multi-class classifier and accurately classifies the Indian Musical Instrument into three classes with an accuracy of 64.44% in the Spectral domain & 72.92% in the statistical domain. Figure 5 shows the results of classification using the KNN classifier for Monophonic music of Indian Musical Instruments with varying values of the training-testing ratio. KNN seems to give consistent classification accuracy over a varying range of training-testing ratios. We used Spectral & Statistical Features to recognition of Indian Musical Instruments. Figure 2 KNN Comparative result

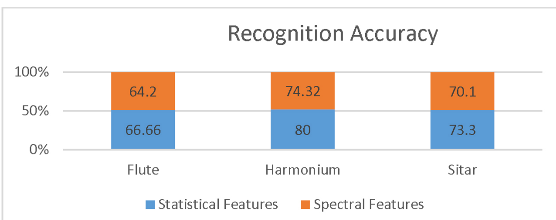

Figure 2

KNN Comparative Result

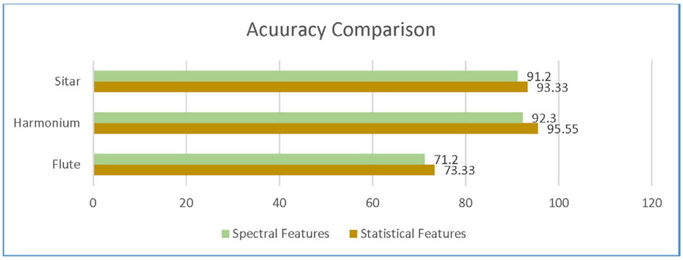

Figure 3

SVM Comparative Result

The above graphical representation describes the individual accuracy of each type of classifier. As depicted in Figure 6 and given in Table 1, the SVM binary classifier classifies the input sound sample of a musical instrument into a woodwind and string-based class with high accuracy of Harmonium 95.55% Figure 3 SVM Comparative result.

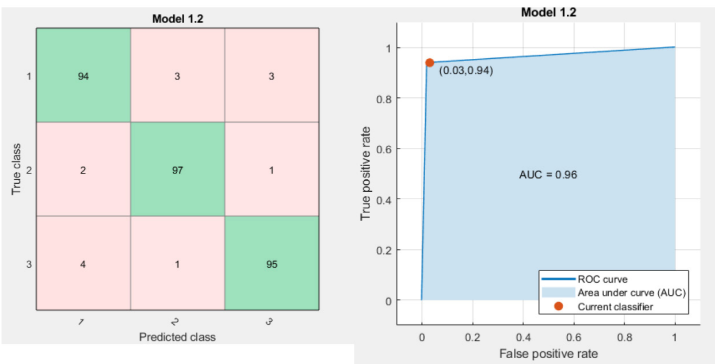

Figure 4

Confusion Matrix and ROC of SVM Classifier

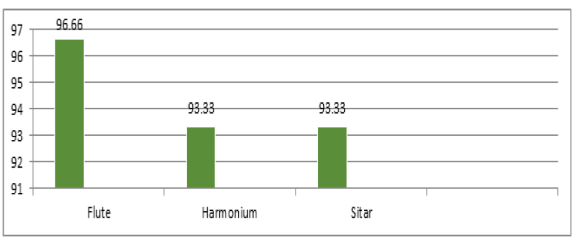

Figure 5

GMM Comparative Result

The above graphical representation in Figure 5, mentions the individual accuracy of each type of classifier. The GMM classifier classifies the input sound sample of a musical instrument into three classes with high accuracy of Flute 96.55%.

4.3 *Experimentation and Results (Polyphonic)*

The purpose of this stage of analysis was to count and classify the instruments contained in the input files. This stage is distinct from the previous phases of analysis in that the sample files utilized for feature extraction were larger and contained many musical instruments. While the accuracy attained was slightly lower than in earlier phases of research, this final level of experimentation enabled exponential outcomes in the classification and identification of Indian musical instruments. Figure 4 talks abou Confusion Matrix and ROC of SVM Classifier.

KNN classifier accurately classifies the Indian Musical Instrument into three classes with an accuracy of 80% in the Spectral domain with two instruments and 75% in 75% with three instruments. Figure 7 shows the results of classification using the KNN classifier for Polyphonic music of

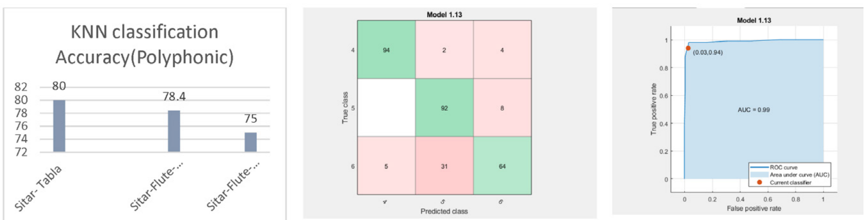

Figure 6

KNN Classification Accuracy for Polyphonic with Confusion Matrix and ROC

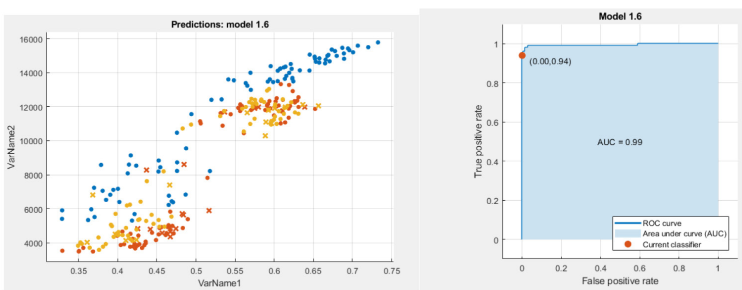

Figure 7

SVM Classification Accuracy for Polyphonic with feature scatter plot and ROC

Table 1

SVM Accuracy

Polyphonic		
Class	**Statistical Parameter**	**Spectral Parameter**
Sitar-Tabla	93.33	92.1
Sitar-Flute-Tabla	87.55	88.8
Sitar-Flute-Santoor-Tabla	88.89	80

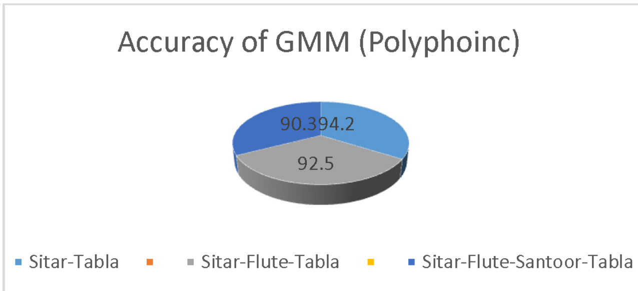

Figure 8

GMM Classification Accuracy for Polyphonic

Indian Musical Instruments. While from table 1 it can be observed that the accuracy of SVM is substantially high for two instruments up to 93.33% and even for five instruments simultaneously it reaches 88.88%.

The graphical representation in Figure 8 describes the accuracy of GMM. The GMM classifier classifies the input sound sample of two musical instruments with remarkable accuracy of 94.2 % and with four instruments it classifies with the accuracy of 90.3%.

5. Discussion and Conclusions

The overall experimentation can be summarized with Monophonic and Polyphonic music samples and two sets of features (Statistical and Spectral) with the combination of classifiers. SVM, KNN, and GMM are implemented as classifiers with the combination of the above features.

The results can be summarized as follows

For monophonic musical samples, KNN produces more accuracy in the spectral domain than in the statistical domain which reflects the fact that statistical domain parameters are more critical in instrument identification. Spectral parameters do contribute to deciding the quality of tone but as per as the class of instrument is concerned the statistical parameters play move vital role. This fact is also revealed by SVM which also demonstrates greater accuracy with statistical parameters. Being more effective SVM reaches an accuracy of identification up to 95.5% specifically for Harmonium. For flute, the behavior produced by the air column generates a structure that resembles the memory generating system which is more complex than the harmonium. This is observed in the results.GMM proves to be the best classifier for monophonic instrument identification. As GMM is a mixture of multiple Gaussian distributions and the behavior of the instrument note also resembles the Gaussian distribution when observed in the probabilistic domain enhances the result. Multiple Gaussian exhibits the most effective identification. Indian Classical Instruments exhibit the most natural sound which can be modeled using Gaussian distribution which models the natural phenomenon. So the density function of sound from Indian instruments can be modeled with Gaussian distribution. This gets reflected in very high accuracy in recognition with GMM even when worked out for polyphonic combination. Polyphonic SVM proves to be the best classifier with an accuracy of 94% for two instruments as the best performance of the SVM is reflected in binary classification. However, for more than two classes combined GMM supersedes SVM and gives the maximum accuracy of 88% for four classes combined. So up to two classes combined the SVM proves to be the better classifier but for more than two classes GMM proves to be the best classifier. Figure 6 graphically represents the comparison of classification accuracy for each instrument. As can be concluded that, GMM outstands the other two classifiers and proves its efficiency by providing the highest accuracy.

Future Scope:

The work presented in the paper can be extended to western instruments as well. It is an AI-based generic framework and methodology which can be adapted to any type and class of instrument. Moreover, the polyphonic recognition accuracy cab further is improved if the feature ranking mechanism is established to select the most significant features contributing to classification.

References

[1] J. J. Burred, A. Robel and T. Sikora, "Dynamic Spectral Envelope Modeling for Timbre Analysis of Musical Instrument Sounds", *IEEE*

Transactions on Audio, Speech, and Language Processing, vol. 18, no. 3 pp. 663-674 (2010).

[2] Sang Hyun Park, "Musical Instrument Extraction through Timbre Classification", NVIDIA Corporation Santa Clara, CA 95050.

[3] Marques and P. J. Moreno, "A Study of Musical Instrument Classification using Gaussian Mixtures Models and Support Vector Machines", Compaq Cambridge Research Laboratory, vol. Tech. Report 99/4 (1999).

[4] Swe Zin Kalayar Khine, Tin Lay New, Haizhou Li, "On Timbre Based Perceptual Feature For Singer Identification", Institute for Infocomm Research 21 Heng Mui Keng Terrace, Singapore.

[5] Giovanni De Poli and Paolo Prandoni, "Sonological Models for Timbre Characterization", *Journal of New Music Research*, Vol. 26:2 pp. 170-197 (1997).

[6] Christianini, N., and J. Shawe-Taylor, "An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods", Cambridge University Press, Cambridge, UK (2000).

[7] B. Kostek, "Musical instrument classification and duet analysis employing music information retrieval techniques", *Proceedings of the IEEE*, vol. 92, no. 4, pp. 712-729 (2004).

[8] Chih-Wei Hsu and Chih-Jen Lin, "A comparison of methods for multiclass support vector machines", *IEEE Transactions on Neural Networks*, vol. 13, no. 2, pp. 415-425 (2002).

[9] Luca Turchet, Carlo Fischione, Georg Essl, DamiaN Kellar,Mathieu Barthet,"Internet of Musical things:Vision and Challenges", IEE Access,Vol 6, pp. 61994-62017 2018.

[10] S. Prabavathy, V. Rathikarani, P. Dhanalakshmi," Classification of Musical Instruments using SVM and KNN", *International Journal of Innovative Technology and Exploring Engineering* (IJITEE) ISSN: 2278-3075 (Online), Volume-9, Issue-7 (May 2020).

[11] Lara Haidar-Ahmad," Music and Instrument Classification using Deep Learning Technics", CS230: Deep Learning, Winter 2018, Stanford University, CA.

[12] S. Prabavathy, V. Rathikarani, P. Dhanalakshmi," An Enhanced Musical Instrument Classification using Deep Convolutional Neural Network", *International Journal of Recent Technology and Engineering* (IJRTE) ISSN: 2277-3878, Volume-8 Issue-4 (November 2019).

[13] Utsav Shukla, Utkarsh Tiwari, Vaibhav Chawla , Shailendra Tiwari,” Instrument Classification using image based Transfer Learning”, 2020 5th International Conference on Computing, Communication and Security (ICCCS).

[14] S. Prabavathy, V. Rathikarani, P. Dhanalakshmi,” Musical Instruments Classification using Pre-Trained Model”, *International Research Journal of Engineering and Technology*”, Volume: 07 Issue: 05 (May 2020).

[15] Karthikeya Racharla, Vineet Kumar, Chaudhari Bhushan Jayant, Ankit Khairkar, Paturu Harish,” Predominant Musical Instrument Classification based on Spectral Features

[16] Anubhav Chhabra; Aryan Veer Singh; Ritesh Srivastava; Veena Mittal,” Drum Instrument Classification Using Machine Learning”, 2020 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN)

[17] Amit Kumar Gupta, Pushpa Gothwal, Dinesh Goyal & Carlos M. Travieso-Gonzalez (2022) IoT-Galvanized pandemic special E-Toilet for generation of sanitized environment, Journal of Discrete Mathematical Sciences and Cryptography, DOI: 10.1080/09720529.2022.2068607.

[18] Amit Kumar Gupta, Vijander Singh, Priya Mathur & Carlos M. Travieso-Gonzalez. Prediction of COVID-19 pandemic measuring criteria using support vector machine, prophet and linear regression models in Indian scenario, *Journal of Interdisciplinary Mathematics*, (2020). DOI: https://doi.org/10.1080/09720502.2020.1833458.

[19] Jitendra Singh Yadav, Amit Kumar Gupta & Arjit Saxena. A review on gender identification using machine learning based on fingerprints, *Journal of Information and Optimization Sciences*, 40:5, 1121-1129 (2019), DOI: https://doi.org/10.1080/02522667.2019.1638002.