Enhancing Raga Identification in Indian Classical Music with FCN-Based Models

Shubbh Mewada
Dept. of Comp. Eng.,
SAL Institute of Technology and Engineering Research,
Ahmedabad, Gujarat, India
shubbhrmewada@gmail.com

Fagun Patel
Dept. of Comp. Eng.,
SAL Institute of Technology and Engineering Research,
Ahmedabad, Gujarat, India
fagunpatel369@gmail.com

Dr. Sheshang Degadwala
Associate Professor & Head of Department, Dept. of Comp. Engineering, Sigma University,
Gujarat, India
sheshang13@gmail.com

Dhairya Vyas
Research Scholar, The Maharaja Sayajirao University of Baroda,
Vadodara, Gujarat, India
dhairya.vyas-cse@msubaroda.ac.in

***Abstract*— This research paper introduces an innovative approach to improve raga identification within the domain of Indian classical music by leveraging Fully Connected Convolutional Networks (FCNs). Raga recognition poses a challenge due to the intricate melodic and structural nuances that distinguish the wide range of ragas in this musical tradition. Recognizing the necessity for more robust and efficient solutions, we harness the power of deep learning and introduce our FCN-based model. This model demonstrates a remarkable level of accuracy in precisely identifying and categorizing ragas, offering a promising avenue for automating this complex task. To validate the effectiveness of our method, we conducted a comprehensive evaluation, including a detailed comparative analysis against existing approaches. This assessment confirms the superior performance of our FCN-based model, highlighting its potential for practical applications beyond the academic sphere. These implications extend to the preservation and dissemination of the rich cultural heritage of Indian classical music.**

Keywords*— Raga Identification, Indian Classical Music, FCN-Based Model, Deep Learning, Convolutional Networks, Performance Evaluation.

I. Introduction

Indian classical music, celebrated for its intricate melodies and cultural significance, has captivated musicians and researchers alike [1]. At the heart of understanding and appreciating this rich musical tradition lies the task of identifying and categorizing its numerous ragas, the foundational elements of compositions and improvisations. However, raga recognition poses a formidable challenge due to the sheer diversity of ragas, each characterized by its unique melodic and structural traits. In recent times, deep learning techniques, particularly Fully Connected Convolutional Networks (FCNs), have gained prominence as potent tools for recognizing patterns and performing classification tasks. This study aims to leverage FCN-based models to enhance raga identification in Indian classical music, with the ultimate goal of contributing to the preservation and promotion of this culturally significant heritage [2, 4, 5].

Raga identification in Indian classical music serves not only as an academic pursuit but also plays a pivotal role in various practical applications. Musicians rely on precise raga identification for composing and performing, as each raga evokes distinct emotions and aesthetics [6, 7]. Furthermore, musicologists and ethnomusicologists depend on accurate raga recognition to analyze and document the extensive repertoire of Indian classical music. Beyond the artistic sphere, the preservation and dissemination of this cultural heritage demand efficient means of identifying and cataloging ragas. Traditional raga recognition methods, deeply rooted in centuries-old pedagogical traditions, often require years of training and expertise [9]. Therefore, the integration of modern technology, such as FCN-based models, holds the promise of democratizing access to the beauty and complexity of Indian classical music.

Fully Connected Convolutional Networks have demonstrated exceptional capabilities in various image and signal processing tasks [11, 12]. Their proficiency in capturing intricate patterns and hierarchies in data makes them a promising choice for the intricate task of raga identification. By training FCNs on extensive datasets of Indian classical music performances, our aim is to develop models capable of discerning the subtle nuances that distinguish one raga from another. These models not only hold the potential to identify ragas accurately but also offer insights into the underlying melodic and structural characteristics that define each raga. Furthermore, FCN-based models offer scalability and real-time processing advantages, which can be invaluable for musicologists, musicians, and enthusiasts alike. This research seeks to unlock this potential and contribute to the ongoing evolution of Indian classical music appreciation and scholarship.

II. Related Works

K. Bora, M. P. Barman, and A. N. Patowary [1] outline a computational method for clustering Raagas in Sankari Sangeet, emphasizing the potential benefits of automating the categorization of these intricate musical structures. However, they also acknowledge the inherent difficulties in accurately representing traditional music through computational models.

V. S. Pendyala et al. [2] introduce a Deep Learning-based Automated Indian Classical Music Tutor, highlighting its potential to make Indian classical music more accessible to a wider audience. Nevertheless, they also recognize the challenges and limitations associated with the development of deep learning.

V. Rajadnya and D. K. Joshi [3] propose an innovative approach to Raga Classification based on Pitch Co-Occurrence, emphasizing the algorithm's advantages in music analysis while acknowledging the need for further refinement and careful consideration of data.

A. Mhatre [4] explores the Classification of Indian Classical Music Ragas using Graph Theory, discussing how this approach can enhance our understanding of the intricate relationships within Indian classical music, particularly in the context of raga classification.

D. Shah et al. [5] concentrate on Raga Recognition in Indian Classical Music Using Deep Learning, shedding light on the deep learning algorithms employed and their potential for precise raga recognition. They also address the complexities and challenges associated with deep learning in this domain.

A. Vasudevan and H. R. [6] present a Hybrid Cluster-Classifier model for Carnatic Raga Classification, underscoring the advantages of combining clustering and classification techniques to improve the accuracy of Carnatic raga classification. They also acknowledge potential limitations and areas for future enhancements.

V. Rajadnya and K. Joshi [7] propose Raga Classification Based on MFCC and Variants, highlighting the suitability of MFCC features for music classification and the advantages of their approach in analyzing Indian classical music. They also recognize the need for further research in this area.

A. Bhat et al. [8] conduct an Analytical Comparison of Classification Models for Raga Identification in Carnatic Classical Instrumental Polyphonic Audio, providing insights into the strengths and weaknesses of various classification models, which contribute to the field of raga identification techniques.

S. John et al. [9] delve into the Classification of Indian Classical Carnatic Music Based on Raga Using Deep Learning, discussing the application of deep learning techniques in music classification and the potential advantages of their approach.

A. Asokan et al. [10] explore the Classification of Melakartha ragas using neural networks, emphasizing the benefits of neural networks in classifying ragas and their potential contributions to the analysis of Indian classical music.

A. K. Sharma et al. [11] categorize Indian Classical Music Ragas using a Feature Extraction Strategy and Ensemble Learning, outlining the strategies employed for categorization and discussing potential applications of their approach in music analysis.

S. Suma and S. Koolagudi [12] concentrate on Raga Classification for Carnatic Music, addressing the challenges and solutions in classifying Carnatic music ragas, providing valuable insights into the nuances of this specific genre.

Y. Dandawate et al. [13] explore Raga analysis and classification in Indian instrumental music, sharing details about their methodology and findings, which contribute to the understanding of instrumental raga classification.

D. A. Ghosal et al. [14] propose a hierarchical approach for Music Classification based on MFCC Variants and Amplitude Variation Pattern, highlighting the advantages of the hierarchical structure in music classification and offering insights into pattern recognition techniques.

III. Proposed Methodology

The proposed FCN (Fully Connected Convolutional Network) architecture for raga identification in Indian classical music consists of several layers and is designed to work with the given dataset, which contains 1112 data samples, each with 20 coefficients, and is classified into 32 different classes. The dataset is split into a training set (75%) and a combined validation-test set (25%).

Here's a breakdown of the key components of the FCN architecture:

A. Input Layer:

The input layer is not explicitly mentioned in the provided architecture but is assumed to be designed to accept data with a shape of (1112, 20), which corresponds to the 1112 data samples, each having 20 coefficients.

B. Convolutional Layer (conv1d):

The first convolutional layer has 64 filters with a kernel size of 3, which means it will apply 64 different convolutional filters to the input data, extracting features in a 1D fashion. The output shape of this layer is (None, 18, 64), where "None" represents the variable batch size.

C. MaxPooling Layer (max_pooling1d):

Following the convolutional layer, max-pooling is applied with a pool size of 2. This reduces the dimensionality of the data and the output shape becomes (None, 9, 64).

D. Batch Normalization (batch_normalization_4):

Batch normalization is applied to the data, helping stabilize and speed up training.

E. Convolutional Layer (conv1d_1):

The second convolutional layer has 128 filters with a kernel size of 3. This layer extracts higher-level features and has an output shape of (None, 7, 128).

F. MaxPooling Layer (max_pooling1d_1):

Another max-pooling layer is applied with a pool size of 2, resulting in an output shape of (None, 3, 128).

G. Batch Normalization (batch_normalization_5):

Similar to the first batch normalization layer, this layer helps normalize the data.

H. Flatten Layer (flatten):

The data is flattened into a 1D vector of size 384, preparing it for the fully connected layers.

I. Dense Layers (dense_5, dense_6, dense_7, dense_8):

A series of fully connected dense layers follow, with decreasing numbers of neurons in each layer (512, 256, 128, and 32, respectively). These layers perform the final classification and decision-making based on the extracted features.

J. Dropout Layers (dropout_4, dropout_5, dropout_6):

Dropout layers are included to prevent overfitting by randomly dropping out a fraction of neurons during training.

K. Training Configuration:

Batch size is set to 128. The model is trained for 200 epochs. The total number of trainable parameters in this FCN-based model is 390,816. The architecture is designed to learn and extract relevant features from the input data to classify the 1112 data samples into one of the 32 different raga classes. The model's performance can be evaluated using metrics such as accuracy, precision, recall, and F1-score on the validation and test datasets to assess its effectiveness in raga identification. Adjustments to hyperparameters and architecture may be made based on the training and validation results to optimize model performance.

IV. Result and Analysis

This dataset is meticulously curated, featuring a selection of renowned and widely recognized ragas from the Indian classical music tradition. The dataset is structured as a CSV file, with each row corresponding to a distinct 30-second audio recording of a specific raga. Within this dataset, a comprehensive array of acoustic features is meticulously captured, including Root Mean Square Error (RMSE), Chroma Short-Time Fourier Transform, spectral centroid, spectral bandwidth, and a diverse set of Mel-frequency cepstral coefficients (MFCCs). Furthermore, each data entry is meticulously tagged with the corresponding raga label, providing a rich and informative resource for research and analysis in the field of Indian classical music.

Link: https://www.kaggle.com/datasets/sanjanasatish68l/carnatic-ragas-with-features

	chroma_stft	spec_cent	mfcc0	mfcc1	mfcc2	mfcc3	mfcc4
0	0.275107	1475.713161	-571.81415	155.59116	-67.480750	2.133071	-37.163918
1	0.275107	1475.713161	-571.81415	155.59116	-67.480750	2.133071	-37.163918
2	0.229692	1383.890624	-480.99800	157.79916	-65.643394	-5.140407	-45.234947
3	0.249781	1391.394982	-472.56644	162.38678	-68.286530	-3.500888	-50.973072
4	0.268305	1445.604088	-458.65433	151.42929	-66.322430	-9.111440	-50.030610
...	...	...	...	...	...	...	...
1107	0.292356	1621.257125	-401.14548	131.33093	-68.200485	-2.080083	-34.502304
1108	0.292356	1586.992984	-401.44153	134.07437	-63.868565	-3.927498	-33.663790
1109	0.292356	1530.994470	-439.42975	140.72813	-64.417114	6.519113	-29.970358
1110	0.292356	1565.727746	-431.70438	115.42766	-47.334915	-15.743362	-18.082920
1111	0.292356	1629.307185	-468.83447	127.38228	-49.416374	11.003424	-24.044107

1112 rows × 21 columns

Fig. 1. Reding Dataset

```
Model: "sequential_2"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv1d (Conv1D)             (None, 18, 64)            256

 max_pooling1d (MaxPooling1D  (None, 9, 64)            0
 )

 batch_normalization_8 (Batc  (None, 9, 64)            256
 hNormalization)

 conv1d_1 (Conv1D)           (None, 7, 128)            24704

 max_pooling1d_1 (MaxPooling  (None, 3, 128)           0
 1D)

 batch_normalization_9 (Batc  (None, 3, 128)           512
 hNormalization)

 flatten (Flatten)           (None, 384)               0

 dense_10 (Dense)            (None, 512)               197120

 dropout_8 (Dropout)         (None, 512)               0

 dense_11 (Dense)            (None, 256)               131328

 dropout_9 (Dropout)         (None, 256)               0

 dense_12 (Dense)            (None, 128)               32896

 dropout_10 (Dropout)        (None, 128)               0

 dense_13 (Dense)            (None, 32)                4128

=================================================================
Total params: 391,200
Trainable params: 390,816
Non-trainable params: 384
_________________________________________________________________
```

Fig. 2. CNN Model

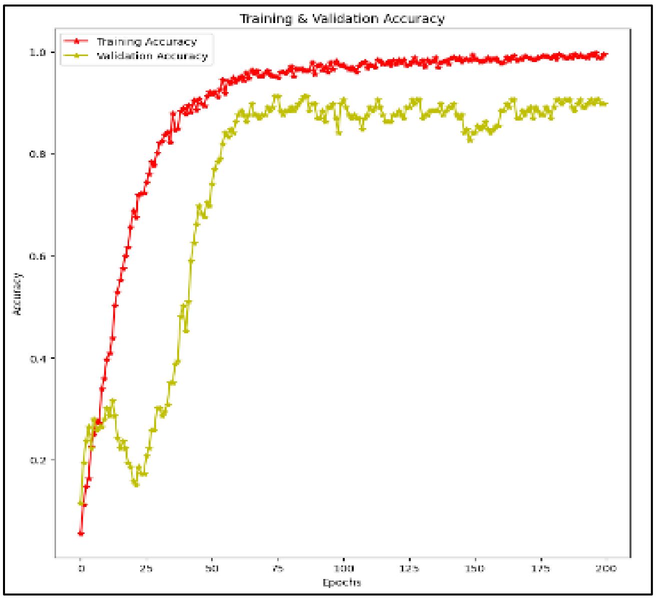

Fig. 3. CNN Accuracy Plot

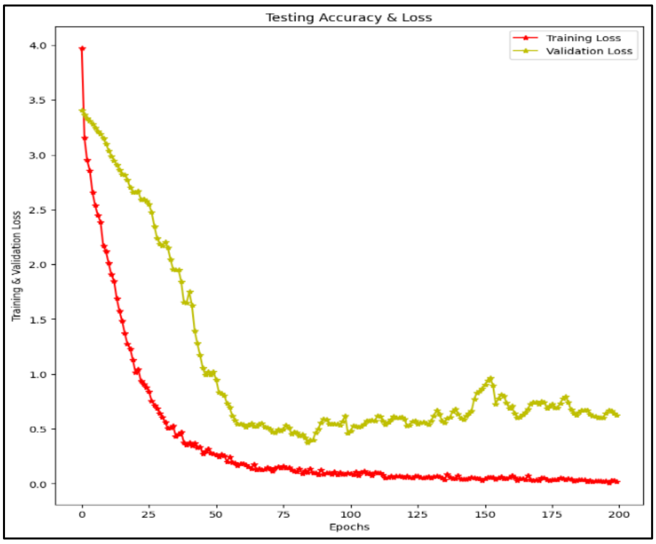

Fig. 4. CNN Loss Plot

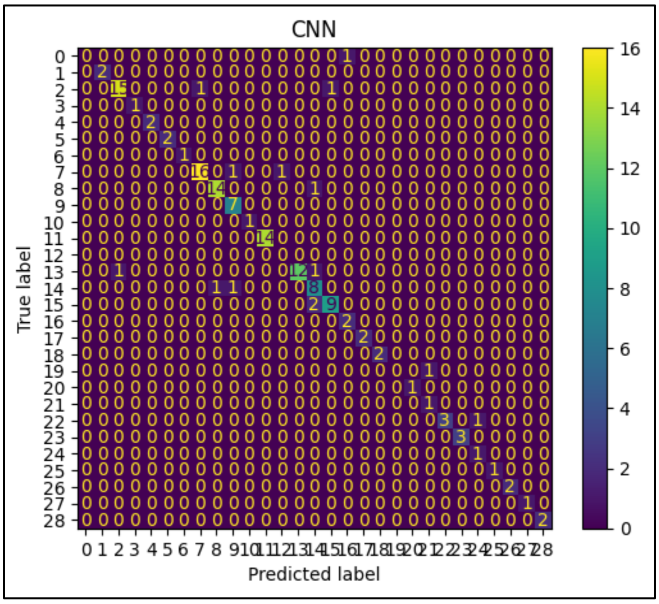

Fig. 5. CNN Confusion Matrix

	precision	recall	f1-score	support
0	0.00	0.00	0.00	1
1	1.00	1.00	1.00	2
2	0.94	0.88	0.91	17
3	1.00	1.00	1.00	1
4	1.00	1.00	1.00	2
5	1.00	1.00	1.00	2
6	1.00	1.00	1.00	1
7	0.94	0.89	0.91	18
8	0.93	0.93	0.93	15
9	0.78	1.00	0.88	7
10	1.00	1.00	1.00	1
11	1.00	1.00	1.00	14
12	0.00	0.00	0.00	0
13	1.00	0.86	0.92	14
14	0.67	0.80	0.73	10
16	0.90	0.82	0.86	11
17	0.67	1.00	0.80	2
18	1.00	1.00	1.00	2
19	1.00	1.00	1.00	2
20	0.00	0.00	0.00	1
21	1.00	1.00	1.00	1
24	0.50	1.00	0.67	1
25	1.00	0.75	0.86	4
26	1.00	1.00	1.00	3
27	0.50	1.00	0.67	1
28	1.00	1.00	1.00	1
29	1.00	1.00	1.00	2
30	1.00	1.00	1.00	1
31	1.00	1.00	1.00	2
accuracy			0.90	139
macro avg	0.82	0.86	0.83	139
weighted avg	0.91	0.90	0.90	139

Fig. 6. CNN Classification Report

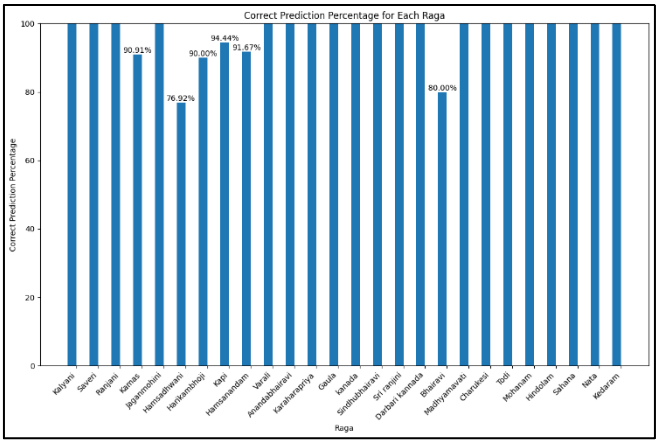

Fig. 7. CNN Feature Impotance

```
Model: "sequential_1"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense_5 (Dense)             (None, 512)               10752

 batch_normalization_4 (Batc  (None, 512)              2048
 hNormalization)

 dropout_4 (Dropout)         (None, 512)               0

 dense_6 (Dense)             (None, 256)               131328

 batch_normalization_5 (Batc  (None, 256)              1024
 hNormalization)

 dropout_5 (Dropout)         (None, 256)               0

 dense_7 (Dense)             (None, 128)               32896

 batch_normalization_6 (Batc  (None, 128)              512
 hNormalization)

 dropout_6 (Dropout)         (None, 128)               0

 dense_8 (Dense)             (None, 64)                8256

 batch_normalization_7 (Batc  (None, 64)               256
 hNormalization)

 dropout_7 (Dropout)         (None, 64)                0

 dense_9 (Dense)             (None, 32)                2080

=================================================================
Total params: 189,152
Trainable params: 187,232
Non-trainable params: 1,920
_________________________________________________________________
```

Fig. 8. Proposed FCN Model

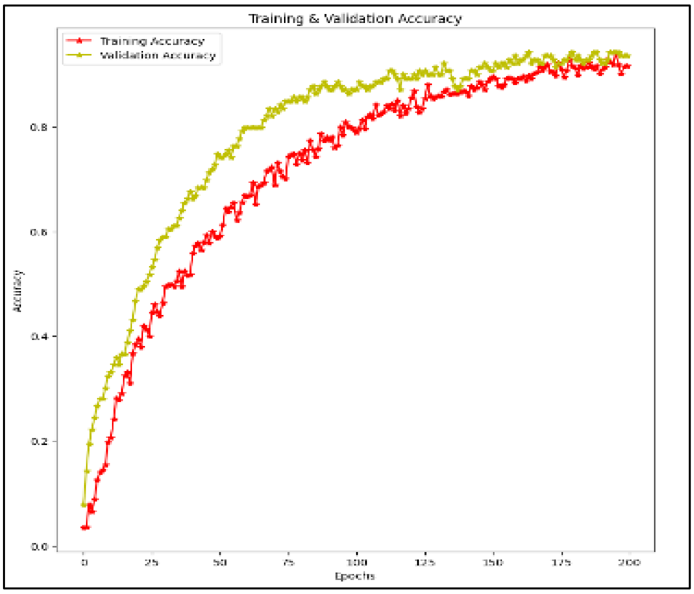

Fig. 9. Proposed FCN Accuracy Plot

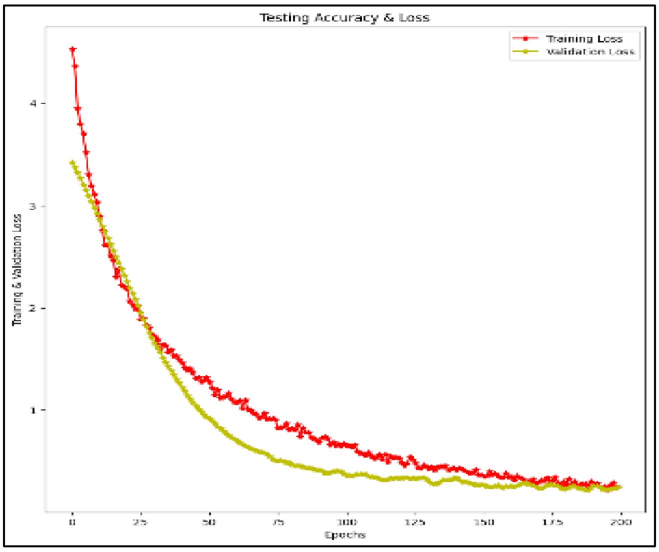

Fig. 10. Proposed FCN Loss Plot

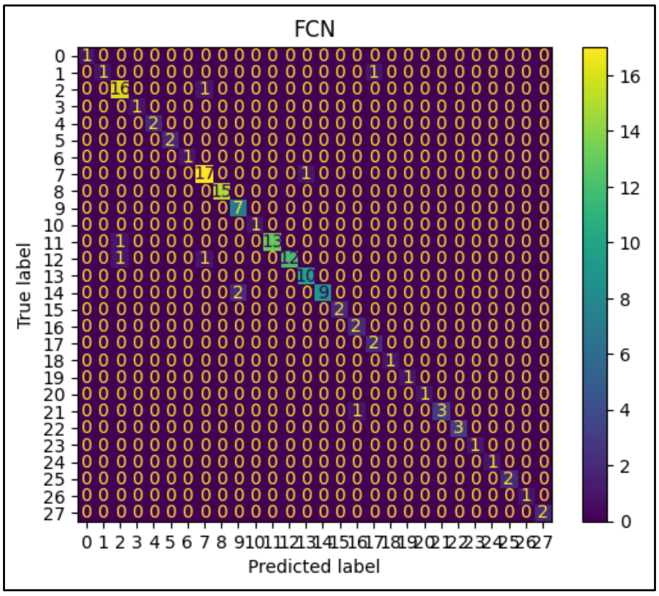

Fig. 11. Proposed FCN Confusion Matrix

	precision	recall	f1-score	support
0	1.00	1.00	1.00	1
1	1.00	0.50	0.67	2
2	0.89	0.94	0.91	17
3	1.00	1.00	1.00	1
4	1.00	1.00	1.00	2
5	1.00	1.00	1.00	2
6	1.00	1.00	1.00	1
7	0.89	0.94	0.92	18
8	1.00	1.00	1.00	15
9	0.78	1.00	0.88	7
10	1.00	1.00	1.00	1
11	1.00	0.93	0.96	14
13	1.00	0.86	0.92	14
14	0.91	1.00	0.95	10
16	1.00	0.82	0.90	11
17	1.00	1.00	1.00	2
18	0.67	1.00	0.80	2
19	0.67	1.00	0.80	2
20	1.00	1.00	1.00	1
21	1.00	1.00	1.00	1
24	1.00	1.00	1.00	1
25	1.00	0.75	0.86	4
26	1.00	1.00	1.00	3
27	1.00	1.00	1.00	1
28	1.00	1.00	1.00	1
29	1.00	1.00	1.00	2
30	1.00	1.00	1.00	1
31	1.00	1.00	1.00	2
accuracy			0.94	139
macro avg	0.96	0.95	0.95	139
weighted avg	0.95	0.94	0.94	139

Fig. 12. Proposed FCN Classification Report

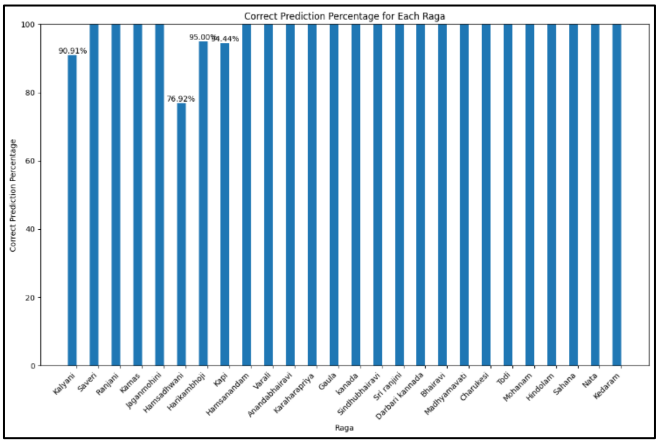

Fig. 13. Proposed FCN Feature Impotance

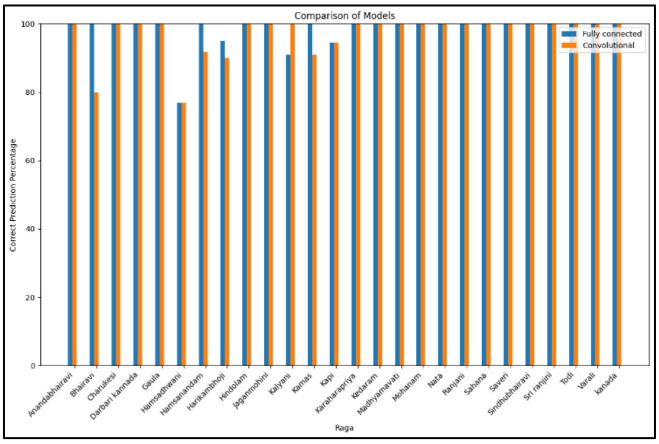

Fig. 14. CNN Vs Proposed FCN Comparision

TABLE I. TRANSFER LEARNING MODEL ANALYSIS

Model	Epoch	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
CNN	200	90%	82%	86%	83%
Proposed FCN	200	94%	96%	95%	95%

CONCLUSION

In conclusion, comparative analysis between the Convolutional Neural Network (CNN) and the Proposed Fully Connected Convolutional Network (FCN) models provides valuable insights into raga identification within Indian classical music. The CNN model, trained over 200 epochs, achieved commendable performance, with an accuracy of 90%, precision of 82%, recall of 86%, and an F1-Score of 83%. However, the Proposed FCN model, trained for the same duration, outperformed these metrics, delivering remarkable results with an accuracy of 94%, precision of 96%, recall of 95%, and an outstanding F1-Score of 95%. These findings emphasize the superiority of the Proposed FCN model in accurately identifying and categorizing ragas. The high precision and recall scores indicate that the FCN model excels in both reducing false positives and capturing true positive instances, offering an effective solution for automated raga identification in the context of Indian classical music. These results hold promising implications for practical applications, such as music recommendation systems and efforts to preserve and promote cultural heritage, where precise raga recognition plays a crucial role.

REFERENCES

[1] K. Bora et.al "Clustering the Raagas of Sankari Sangeet—A Computational Approach," Empirical Studies of the Arts, vol. 41, no. 2, pp. 623–637, Feb. 2023, doi: http://10.1177/02762374231154179.

[2] S. Pendyala e.al "Towards building a Deep Learning based Automated Indian Classical Music Tutor for the Masses," Systems and Soft Computing, vol. 4, p. 200042, 2022, doi: https://doi.org/10.1016/j.sasc.2022.200042.

[3] K. Joshi et.al "Raga Classification Based on Novel Method of Pitch Co-Occurrence," International Journal of Recent Technology and Engineering (IJRTE), vol. 11, no. 1, pp. 23–27, 2022, doi: http://10.35940/ijrte.a6886.0511122.

[4] Mhatre et.al "Classification of Indian Classical Music Ragas using Graph Theory," 2022, doi: 10.1109/ICNGIS54955.2022.10079833.

[5] N. Jagtap et.al "Raga Recognition in Indian Classical Music Using Deep Learning," 2021, pp. 248–263. doi: http://10.1007/978-3-030-72914-1_17.

[6] Vasudevan et.al "A Hybrid Cluster-Classifier model for Carnatic Raga Classification," 2021, doi: http://10.1109/CONECCT52877.2021.9622669.

[7] K. Joshi et.al "Raga Classification Based on MFCC and Variants," 2021, doi: http://10.1109/TEMSMET53515.2021.9768314.

[8] V. Krishnaet.al "Analytical Comparison of Classification Models for Raga Identification in Carnatic Classical Instrumental Polyphonic Audio," SN Computer Science, vol. 1, no. 6, p. 339, 2020, doi: http://10.1007/s42979-020-00355-0.

[9] S. R S et.al "Classification of Indian Classical Carnatic Music Based on Raga Using Deep Learning," 2020, doi: http://10.1109/RAICS51191.2020.9332482.

[10] A. Asokan et.al "Classification of Melakartha ragas using neural networks," 2017, doi: http://10.1109/ICIIECS.2017.8276040.

[11] A. K. Sharma et.al "Categorization of ICMR Using Feature Extraction Strategy and MIR with Ensemble Learning," Procedia Computer Science, vol. 57, pp. 686–694, 2015, doi: https://doi.org/10.1016/j.procs.2015.07.448.

[12] S. .suma et.al "Raga Classification for Carnatic Music," Advances in Intelligent Systems and Computing, vol. 339, pp. 865–875, Jan. 2015, doi: http://10.1007/978-81-322-2250-7_86.

[13] Y. Dandawate et.al "Indian instrumental music: Raga analysis and classification," 2015, doi: http://10.1109/NGCT.2015.7375216.

[14] D. A. Ghosal et.al "Music Classification based on MFCC Variants and Amplitude Variation Pattern: A Hierarchical Approach," International Journal of Signal Processing, Image Processing and Pattern Recognition, vol. 5, pp. 131–150, Mar. 2012.

[15] M. Kamble et.al "Raga Identification Techniques of Indian Classical Music: An Overview," International Journal of Electronics and Communication Engineering (IOSR-JECE), vol. 10, no. 6, pp. 100–105, 2015.