Carnatic Music Raga Identification using Music Information Retrieval & Deep Learning

Fall 2025

INFO 556 Information Retrieval & Web Search

Final Project

Mentored by:
Dr. Xiao Hu
Presented by:
Sundar Ram Subramanian
(SID: 23946913)

Flow of the session

1. Context Setting
2. Objectives
3. Methodology
4. Dataset Overview
5. Code Architecture
6. Code Walkthrough
7. Results & Evaluation
8. Challenges & Future work

About Carnatic Music

Carnatic music is one of the oldest and most historic music systems in the world. Raga is the single most crucial melodic framework of this music art form that provides the rules and notes for a Carnatic composition. It contains a primordial sound (*nāda*), tonal system (*swara*), pitch (*śruti*), scale, ornaments (*gamaka*), and important tones. There are several number of Ragas. The 72 Melakarta Ragas (parent ragas) are formed by certain rules and constraints to the 7 tones & 5 semitones. All other Ragas called Janya Ragas are derived from melakartha ragas.

Project Motivation & Objectives

- Learning to identify a song’s raga is a core competency developed during an Indian Carnatic music education. Therefore, Raga recognition is an important step in computational musicology & MIR as far as Indian Carnatic music is considered.
- This Project aims to leverage MIR techniques to extract and analyze Carnatic audio for 8 melakarta ragas, followed by building a fusion deep learning model (CNN+LSTM) for Raga classification.

Methodology

Dataset collection & preparation → Audio Preprocessing → Feature Extraction → Model Building → Model Evaluation

Dataset Overview

- Indian Art Music Raga Recognition datasets
  - A Computational Music Dataset maintained in UPF repository.
  - Huge dataset with audio in .mp3 format and proper labelling of Ragas in hierarchical directories.

```
RagaDataset/
├── Carnatic
│   ├── audio
│   ├── features
│   └── _info_
├── Hindustani
│   ├── audio
│   ├── features
│   └── _info_
```

  - shortlisted 8 Melakarta Ragas, namely hanumathodi, harikaambhoji, kaamavardhini, kalyani, kharaharapriya, maayamaalavagowlai, shankaraabharanam & shanmugapriya.
  - Minimum of 5 songs from each of the shortlisted 8 ragas were selected based on their audio size and a separate sample dataset having only audios were created

Code Architecture

Sruthi Standardization

Sruthi - base pitch of the performer

- Get Audio
- Extract Sruthi (f_0)
- Standardize f_0 (C3)
- Clip audio to 30s

Raga Feature Extraction

Raga - spine of Carnatic Music

- Pitch Contour
- Dynamic Melody
- Gamakas
- Mel-Spectral
- Cepstral (MFCC)
- Chroma

Deep Learning Model

Fusion Model - CNN + LSTM

- CNN ~ Spectral& Chroma embeds
- LSTM ~ Pitch & Gamakas embeds
- Fusion CNN + LSTM
- Softmax Output

THE UNIVERSITY OF ARIZONA
College of
Information Science

Code Walkthrough

Visual Studio Code:

- Helpers.py
- Pipelines.py
- Feature extraction code.ipynb

Collab

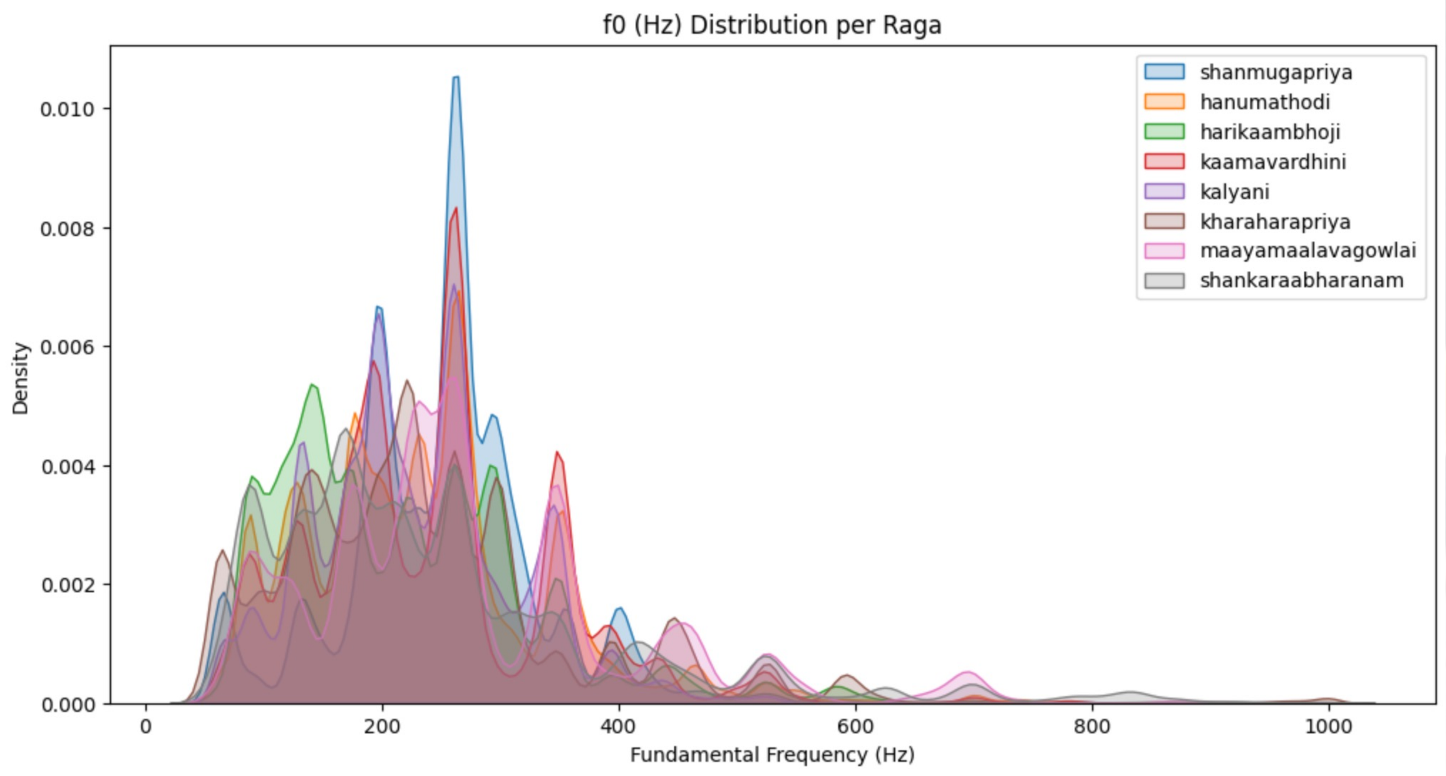
- EDA
- Model Building

GitHub

- https://github.com/sundarram1608/Carnatic-Music-Raga-Identification-using-MIR-and-Deep-Learning.git

THE UNIVERSITY OF ARIZONA
College of
Information Science

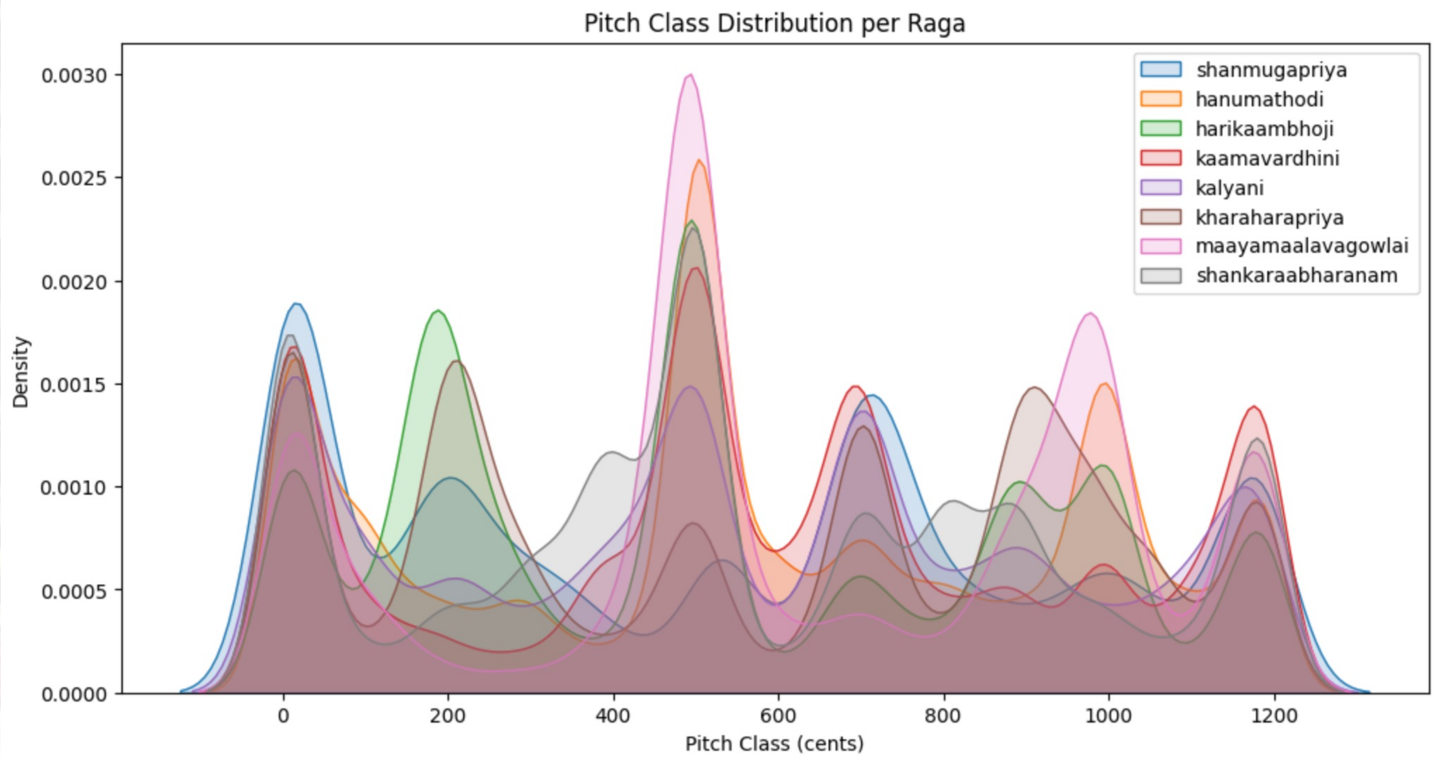
Distribution of f0 in Hz across Ragas



- It can be seen that the Sruthi across audio clips are standardized.
- All ragas sit in a common pitch range.
- Each raga has unique pitch density “fingerprints”
- Even with massive overlap, subtle peaks differ.

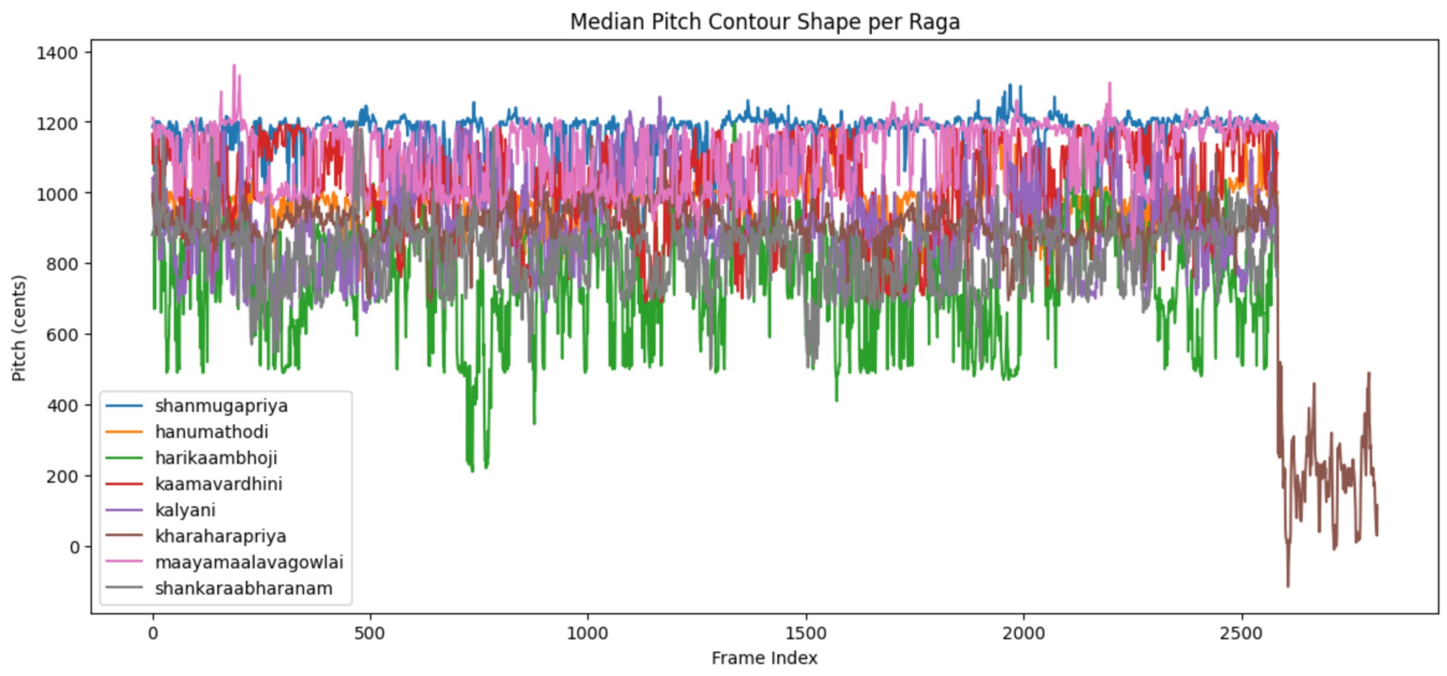
THE UNIVERSITY OF ARIZONA
College of Information Science

Pitch class distribution



- All ragas share strong peaks at Sa (0 cents), Pa (700 cents), and upper Sa (1200 cents) due to tonic fixation and common panchama.
- The discriminative power lies in Ri, Ga, Ma, Dha, Ni peaks
- Each raga shows different prominence, sharpness, and peak shifting.
- **Ragams are separable based on pitch-class distributions despite sharing the same tonic.**
- Ga, Ma, Ni positions and widths offer maximum discriminative power.

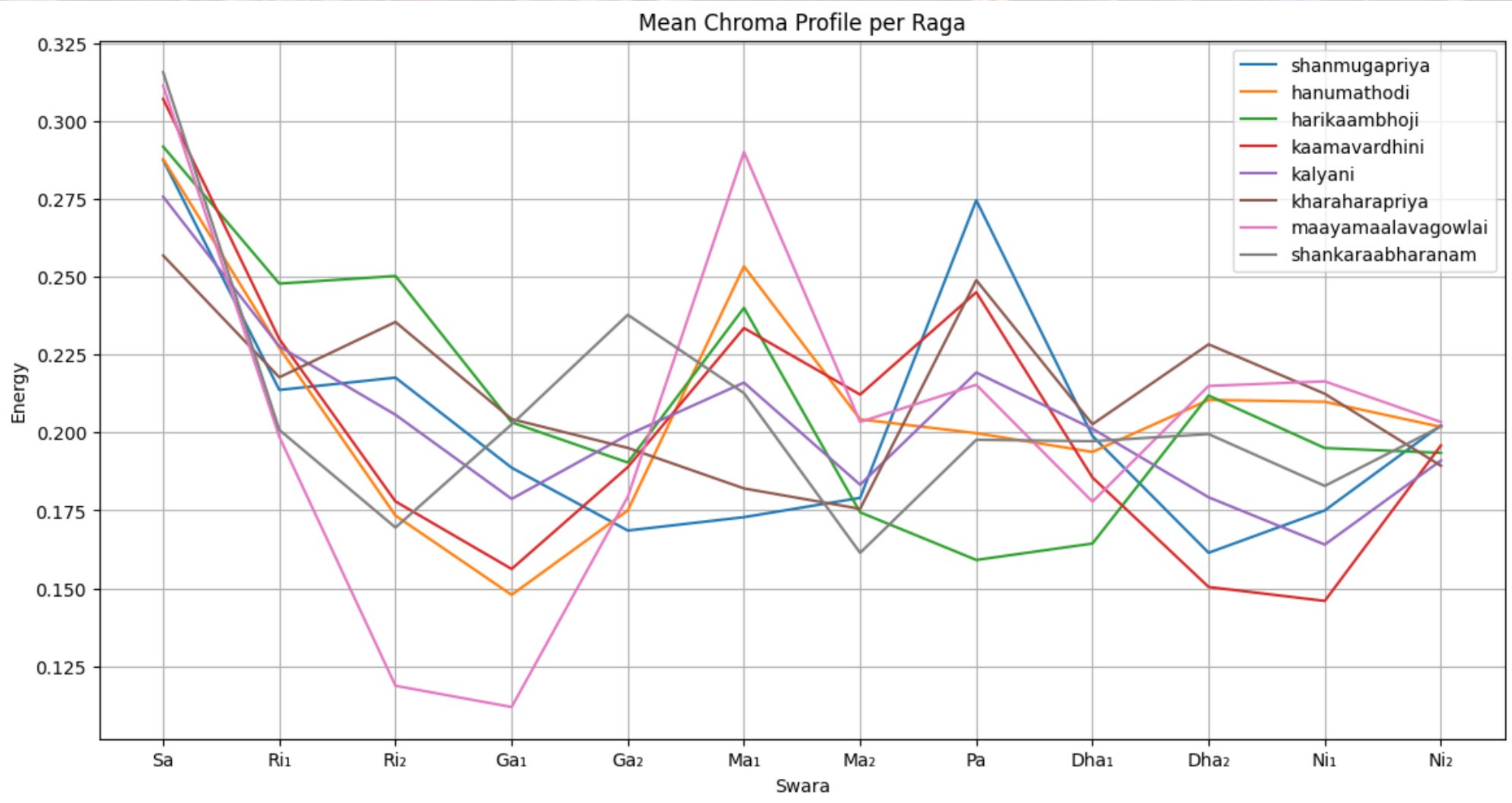
Median Pitch Contour Shape



- Captures the “average melodic fingerprint” of each raga across all clips.
- Shanmukhapriya shows the most stable, plateau-like contour.
- Harikambhoji shows strong downward fluctuations — a key identifying trait.
- Maayamaalavagowla displays large oscillations representing intense kampita.
- Shankarabharanam shows mid-range stability with lower sthayi dips toward clip endings.

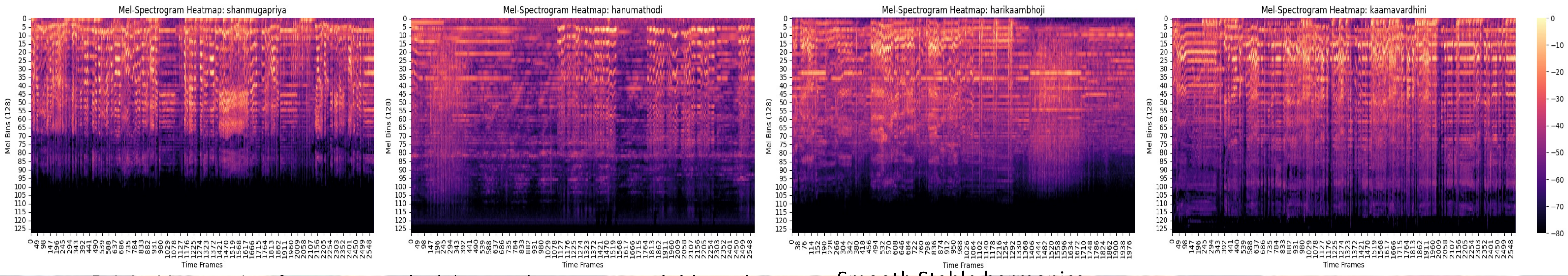
THE UNIVERSITY OF ARIZONA
College of Information Science

Mean Chroma Profile



- Each Line represents average energy across swaras
- This indicates how strongly each swara is emphasized in a raga
- Different ragas show distinct swara emphasis patterns.
- Sa and Pa remain stable anchors across all ragas.
- Discriminative differences appear mainly in other swaras.
- Gamaka-heavy ragas show more energy variation, while scale-based ragas show smooth energy curves.

Mel Spectrogram Heatmap

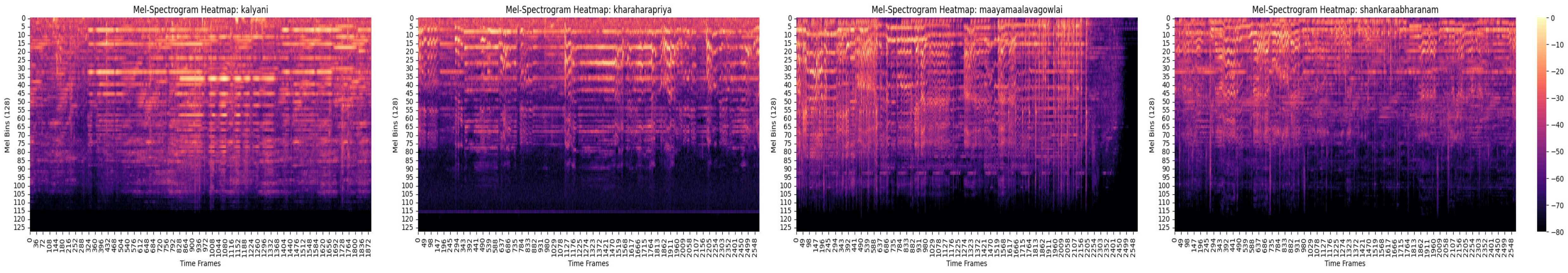


Bright Harmonics & Uniform oscillations

thick low-mid presence with blurred oscillations

Smooth Stable harmonics

High Brilliance with rapid oscillations



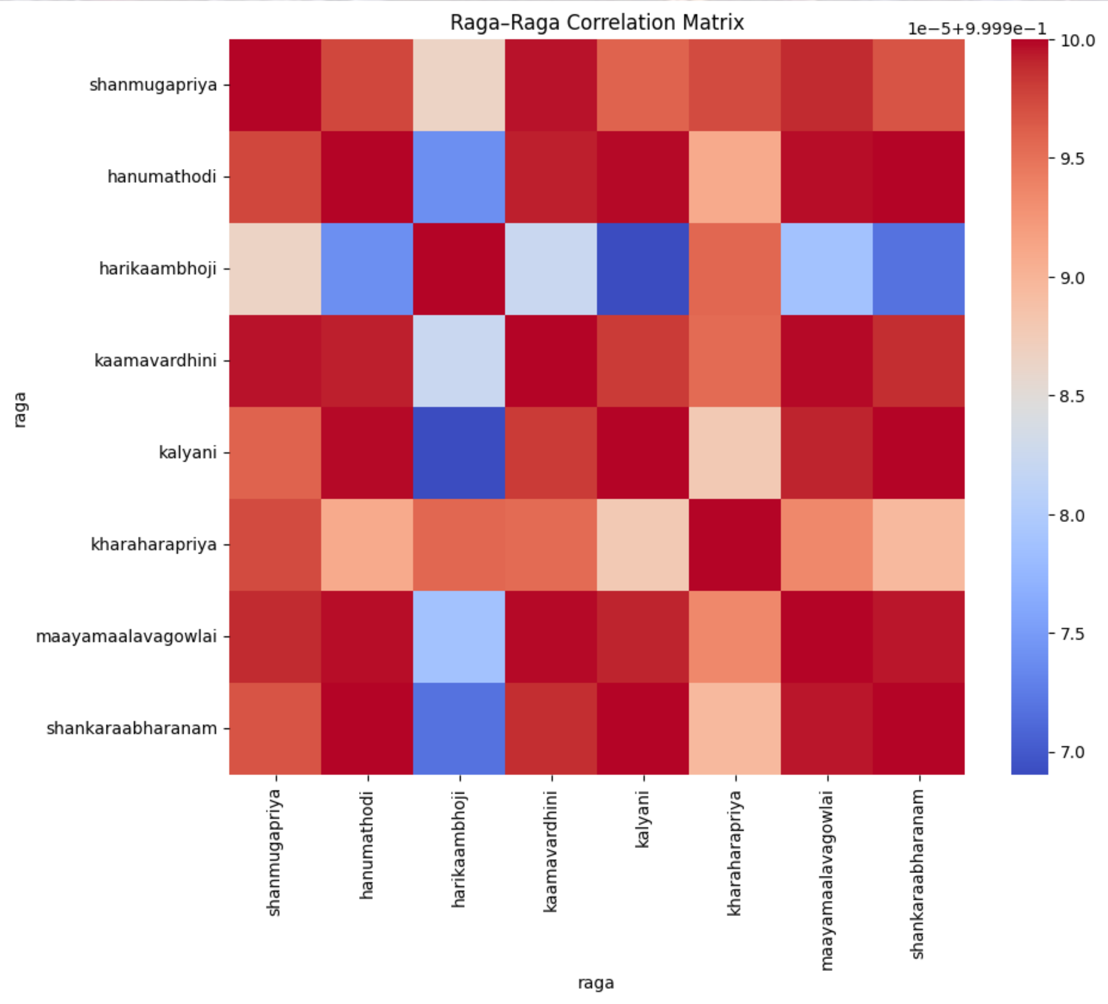
Smooth & wide oscillations

Soft Spectral Profile

Dense High frequency brightness

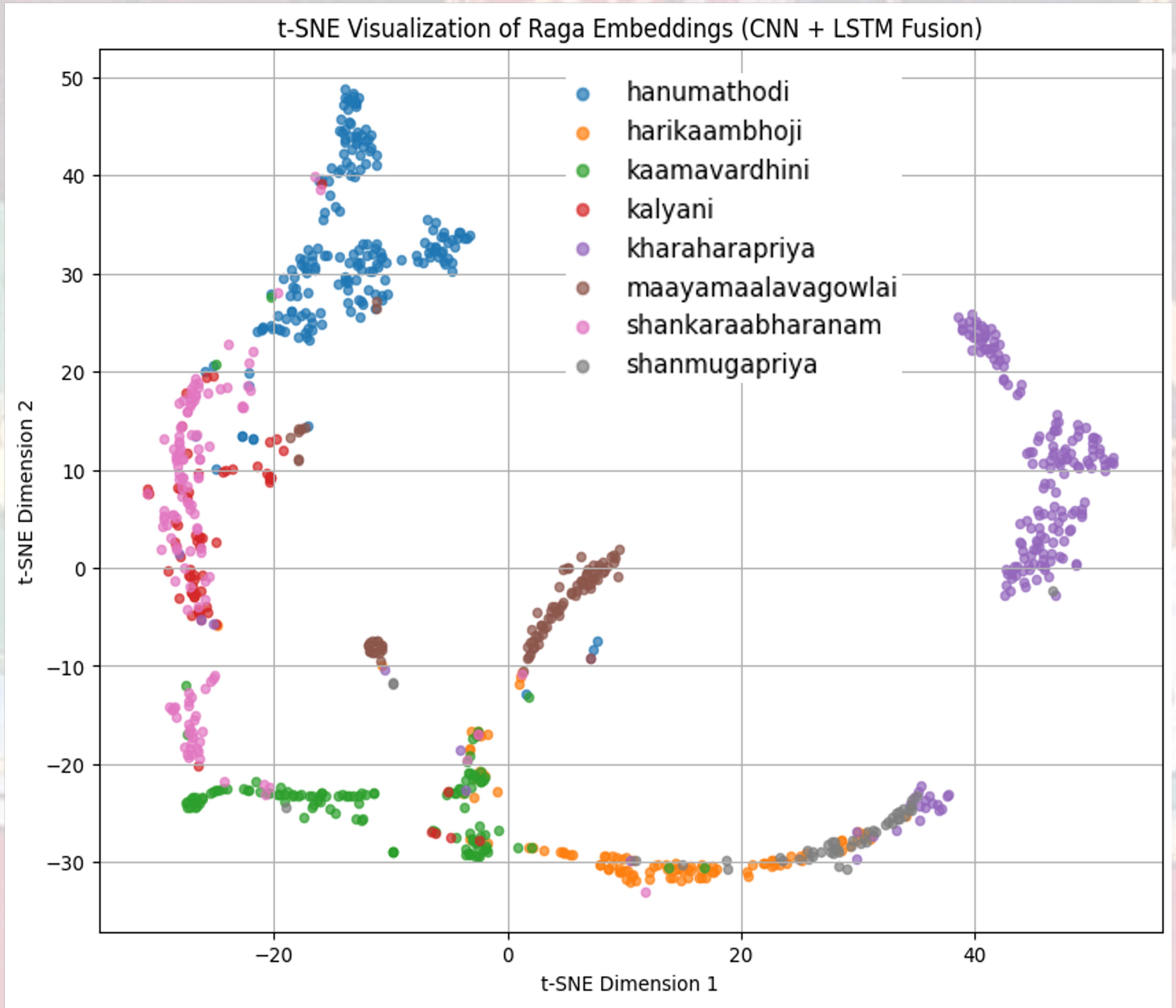
Balanced & stable spectral

Raga-Raga Correlation



- It can be seen that most of the Ragas are correlated and hence a lot of overlap.
- This indicates the major challenge in Raga Identification

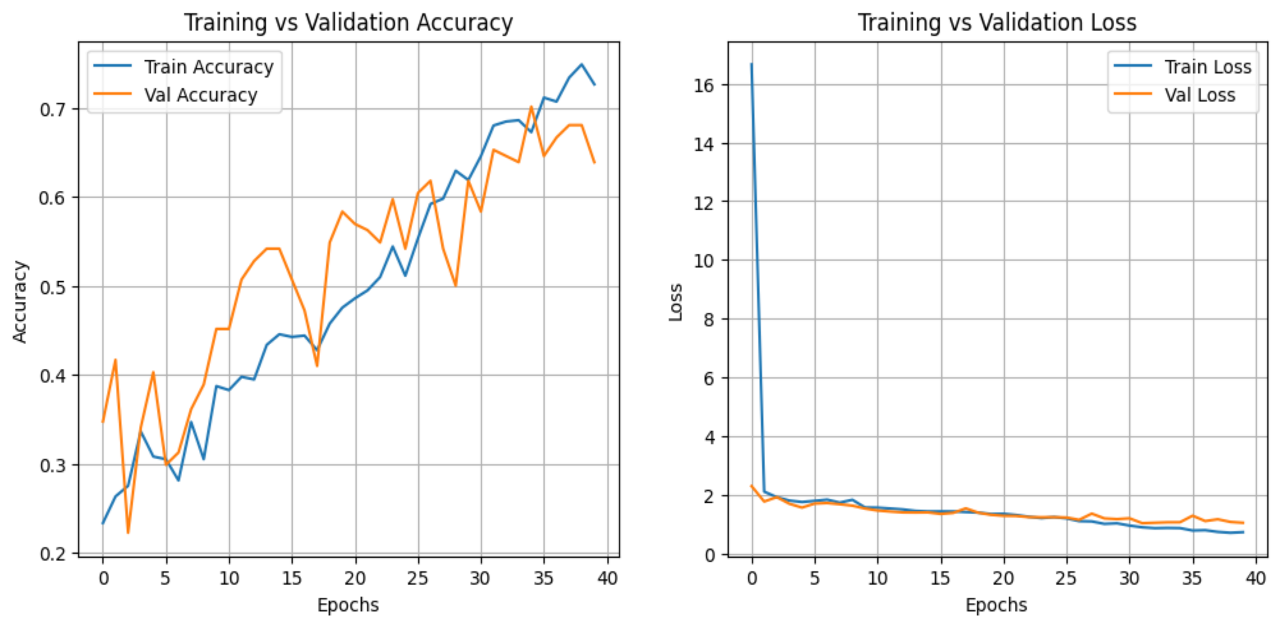
Embedding Visualizations



THE UNIVERSITY OF ARIZONA
College of Information Science

Results & Model Evaluation

```
==================================
 Test Loss : 1.0408
 Test Acc  : 64.58 %
==================================
```



CLASSIFICATION REPORT:

	precision	recall	f1-score	support
hanumathodi	0.70	0.79	0.74	29
harikaambhoji	0.48	0.67	0.56	15
kaamavardhini	0.61	0.61	0.61	18
kalyani	0.00	0.00	0.00	11
kharaharapriya	0.87	0.90	0.89	30
maayamaalavagowlai	0.78	0.50	0.61	14
shankaraabharanam	0.52	0.75	0.61	20
shanmugapriya	0.00	0.00	0.00	7

THE UNIVERSITY OF ARIZONA
College of Information Science

Challenges

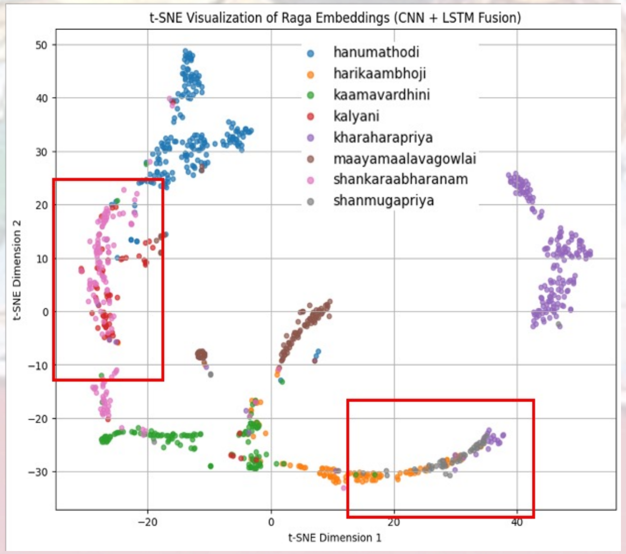
Class Imbalance due to sampling:

Class	Frequency
hanumathodi	194
harikaambhoji	102
kaamavardhini	125
kalyani	69
kharaharapriya	195
maayamaalavagowlai	93
shankarabharanam	130
shanmugapriya	49

The number of vocal recordings per raga vary greatly. Some ragas have abundant data while others have very few examples.

Dataset can be increased by leveraging the full dataset instead of sampling.

Overlapping raga embeddings:



Several Carnatic ragas share the same note sets
(e.g., Shankarabharanam and Kalyani).

Difference in phrasing, note transitions, and gamakas are very difficult to identify & extract

Future Work

- ❑ Expand Dataset (More examples)
- ❑ Improve Class Balance (Adding more samples of same raga)
- ❑ Create Raga Datasets based on induced emotions and review how characteristic features in digital signals vary.
- ❑ Review and Research on what other user centric characteristic features can be retrieved from digital audio information for better separation of Raga embeddings.
- ❑ Leverage SOTA Transformer models for Model Building & Classification
- ❑ Building a user-centric Shazam like app for Carnatic music leveraging the built models

THE UNIVERSITY OF ARIZONA
College of
Information Science

References

[1] Serrà, J., Ganguli, K. K., Sentürk, S., Serra, X., & Gulati, S. (2016). Indian Art Music Raga Recognition Dataset (audio) (1.0) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.7278511

[2] Sridhar, Rajeswari & Geetha, T.V. (2009). Raga Identification of Carnatic music for music Information Retrieval. SHORT PAPER International Journal of Recent Trends in Engineering. 1. Research Gate

[3] Shah, Devansh & Jagtap, Nikhil & Talekar, Prathmesh & Gawande, Kiran. (2021). Raga Recognition in Indian Classical Music Using Deep Learning. 10.1007/978-3-030-72914-1_17. Research Gate

[4] Ravikoti, Sridhar. (2020, September). Identifying Ragas in Carnatic Music with Machine Learning LinkedIn

[5] Great Learning Snippets. (2020, July). Deep Learning Based Raga Classification in Carnatic Music Medium

[6] Srinivasan, Shriya. (2021, August). Carnatic Raga Recognition Medium

[7] Divan, Shreyas. (2021, November). Raga identification using ML and DL Medium