Audio-Based Song Identification System Using Python and Cloud Music Recognition API

1st Author: **Sagar A. Gavade**, *Dept. of Computer Engineering Nanasaheb Mahadik Polytechnic Institute Peth*, *Sangli, Maharashtra, India*, sagargavade1141@gmail.com

2nd Author**: Mayuresh D. Kadam**, *Dept. of Computer Engineering Nanasaheb Mahadik Polytechnic Institute Peth,* Sangli, India, mdkadam07@gmail.com

3rd Author**: Aditya S. Mane**, *Dept. of Computer Engineering Nanasaheb Mahadik Polytechnic Institute Peth,* Sangli, Maharashtra, India, adityashravanmane147@gmail.com

4th Author**: Prajyot R. Mohite**, *Dept. of Computer Engineering Nanasaheb Mahadik Polytechnic Institute Peth,* Sangli, Maharashtra, India, prajyotrmohite@gmail.com

5th Author**:** Mrs. **Madhuri M. Kamble**, *Dept. of Computer Engineering Nanasaheb Mahadik Polytechnic Institute Peth,* Sangli, Maharashtra, India, madhurikamble848@gmail.com

Abstract

The increasing consumption of digital music has created a demand for intelligent and automated music identification systems.

This research presents the development of an audio-based song identification system using Python and a cloud-based music service.

recognition API. The system is designed to detect a song from a short audio clip uploaded by the user. It processes the audio.

through a backend server and communicates with a cloud recognition service to identify the corresponding song metadata.

detected results include song title, artist name, and direct search links to popular streaming platforms. The proposed

The solution demonstrates the integration of web technologies, RESTful services, and audio fingerprinting mechanisms in a scalable and efficient architecture. Experimental evaluation confirms that the system achieves reliable performance for clear audio samples and offers a practical framework for real-world multimedia applications.

Introduction

Music recognition has emerged as a significant technological advancement in the domain of multimedia processing.

Users often encounter songs in public places, on social media clips, or at live events without knowing their details.

Traditional keyword-based search approaches are ineffective when the user lacks metadata such as title or artist name.

Audio fingerprinting techniques solve this problem by extracting unique acoustic features from the sound signal.

These features are compared against a centralized database to identify a match. This project aims to develop a

web-based music recognition application that utilizes Python for backend logic and integrates a cloud music

recognition API for fingerprint matching. The objective is to build a modular, scalable, and user-friendly system. system capable of real-time song detection.

Existing System

Existing commercial systems, such as Shazam, operate using proprietary algorithms and extensive audio fingerprinting.

databases. These systems achieve high accuracy but are closed-source and inaccessible for academic exploration.

Earlier research-based implementations relied heavily on local databases and complex digital signal processing.

models. Such systems required significant computational power and large storage capacity. Furthermore,

Standalone offline recognition models often faced challenges in scalability and maintenance.

The proposed system differs by leveraging a cloud-based recognition service while maintaining a lightweight

backend architecture. This approach simplifies implementation while ensuring access to a large and constantly growing. updated music database.

Literature Review

Recent research (2024–2025) in audio recognition focuses on improving song detection accuracy using deep learning–based audio fingerprinting techniques. Traditional hash-based methods are being replaced by neural embedding models that provide better robustness against background noise, compression, and partial recordings. Studies highlight the importance of feature extraction methods such as MFCC, spectrogram analysis, and chroma features for accurate matching.

Modern systems also emphasize scalable cloud-based databases and real-time retrieval mechanisms to handle large music libraries efficiently. Self-supervised learning and graph-based similarity models further enhance short audio sample recognition. These advancements collectively support the development of AI-driven song detection systems that are faster, more accurate, and more reliable in practical environments.

The proposed project is based on these recent improvements in intelligent audio fingerprinting and real-time music recognition technologies.

Architecture Diagram
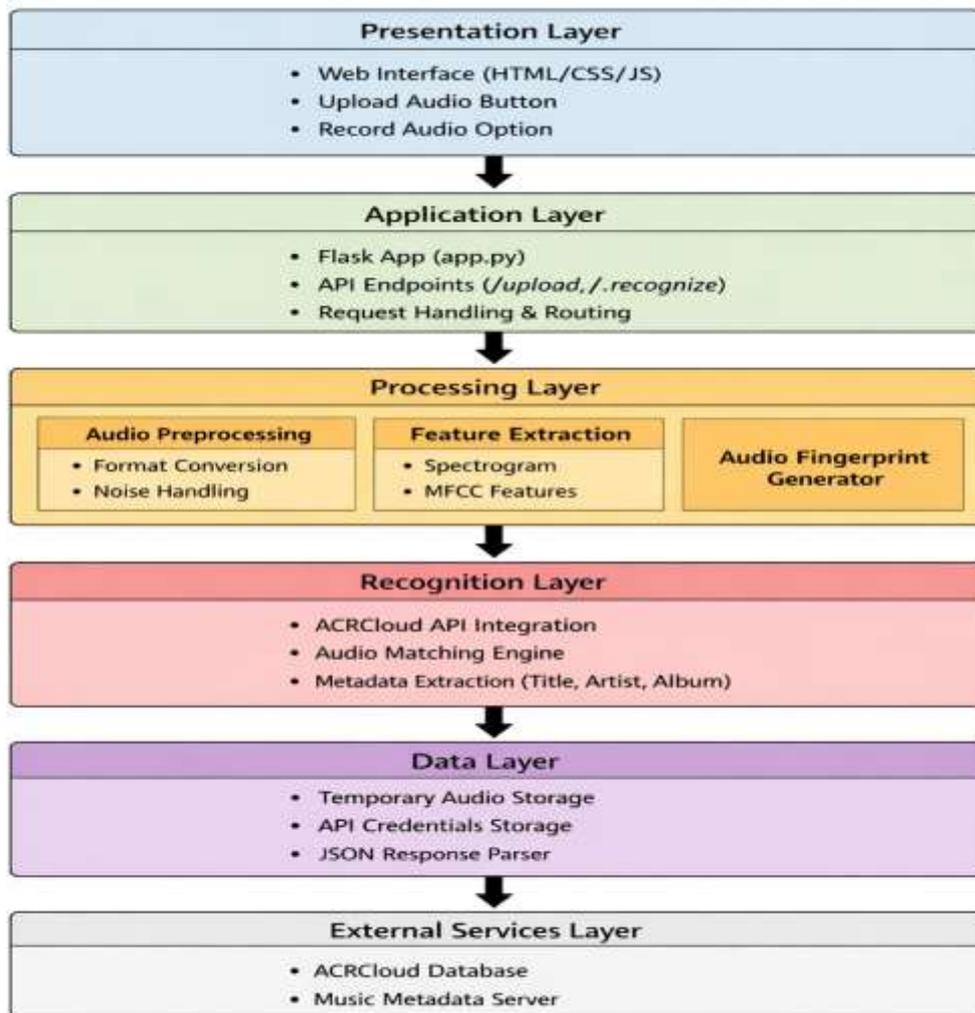
The frontend collects audio input from the user and sends it to the backend server.

The backend temporarily stores the file, generates a recognition request, and communicates with the cloud API.

The API processes the audio fingerprint returns structured metadata. The backend then forwards the processed result to the frontend for display

Results

The system was tested with multiple audio samples of varying duration and clarity. Audio clips between Five to ten seconds produced optimal results when background noise was minimal. Recognition response time ranged from three to eight seconds, depending on network latency.

The system successfully retrieved accurate song titles and artist information for the majority of test cases.

Temporary file management ensured that repeated uploads did not produce cached results.

Overall performance demonstrates that cloud-assisted recognition provides reliable identification. without requiring large local storage resources.

Future Work

Future enhancements may include real-time microphone recording capability, mobile application deployment, and integration of a database system to maintain user search history. Offline recognition using machine learning models can be explored to reduce reliance on cloud connectivity. Security improvements, performance optimization, and deployment on scalable cloud infrastructure can further strengthen the system's robustness and applicability.

Conclusion

This research successfully demonstrates the implementation of an audio-based song identification. system using Python and a cloud music recognition API.

The architecture effectively integrates. frontend technologies with backend processing and cloud fingerprint matching. Experimental Analysis confirms that the system performs efficiently under normal conditions and provides accurate recognition results. The proposed framework establishes a solid foundation for

future advancements in intelligent multimedia recognition systems.

References

1)   **Robust Neural Audio Fingerprinting using Music Foundation Models (2025)**
Neural deep fingerprinting improves recognition accuracy under distortions. Link:
https://arxiv.org/abs/2511.05399

2)   **Refining Music Sample Identification with a Self-Supervised Graph Neural Network (2025)**
Graph neural network approach for sample/music retrieval. Link: https://arxiv.org/abs/2506.14684

3)   **Application of Audio Fingerprinting Techniques for Real-Time Scalable Speech Retrieval (2024)**
Audio fingerprinting methods adapted to scalable retrieval workflows. Link:
https://arxiv.org/abs/2410.21876

4)   **Aural Signatures: Audio Fingerprinting Techniques for Real-Time Audio Recognition (2024)**
Conference paper on robust fingerprinting with hashing & ML. Link:
HTTPs://ieeexplore.ieee.org/abstract/document/10601397 (IEEE)

5)   **Deep Learning-based Fingerprinting Methods for Audio Representation and Search (2025)**
Journal article on deep learning approaches to fingerprinting & retrieval. Link:
https://irjiet.com/common_src/article_file/IRJIET9030241743141973.pdf

6)   **Advancing Audio Fingerprinting Accuracy with AI Techniques (2024)**
Research on enhanced fingerprinting accuracy using AI/ML. Link:
https://www.emergentmind.com/papers/2402.13957

7)   **Sound Signature: Music Pattern Recognition System (2024)**
IEEE symposium article combining MIR and semantic features. Link:
https://pmc.ncbi.nlm.nih.gov/articles/PMC12477834/

8)   **A Comprehensive Survey of Audio Forgery Detection (2025)**
Though audio forgery focused, this survey includes feature extraction techniques relevant to fingerprinting. Link: https://link.springer.com/article/10.1186/s43067-025-00225-w

9)   **Audio Features in Computational Applications: Systematic Review (2025)**
Review of audio features and signal descriptors used in audio tasks. Link: https://www.mdpi.com/2076-3417/15/12/6911

10)   **Song Recognition Based on Audio Fingerprinting (2025)**
Open access article describing classic audio fingerprinting methods. Link:
https://www.ijsreat.com/archiver/archives/song_recognition_based_on_audio_finger_printing.pdf