

Audio Feature Extraction Methods for Machine Learning and Deep Learning-Driven Sound Classification: A Scientific Literature-Based Analysis

Sandhya Rani K.M^{1*}, Dr. S D Khamitkar², Dr. Parag U Balachandra³

***Corresponding Author:** ¹ Research Scholar, School of Computational Sciences, SRTM University, Nanded

sandhya.kiran.rao@gmail.com



ORCID: <https://orcid.org/0009-0009-2027-8513>

²School of Computational Sciences, SRTM University, Nanded

s.khamitkar@gmail.com

³School of Computational Sciences, SRTM University, Nanded

srtmun.parag@gmail.com



ORCID: <https://orcid.org/0000-0001-5422-9423>

ABSTRACT

Sound classification has become an important task in computational audio analysis because audio signals can provide meaningful information regarding the environment, speech, music, machines, animals, health conditions, and forensic evidence. Although recent studies often emphasize machine learning and deep learning architectures, the representation of audio signals before classification remains a central methodological consideration. This study presents a literature-based analysis of audio feature extraction techniques employed in learning-driven sound classification. By synthesizing recent research, it identifies key feature categories, examines the relationships between features and model architectures, explores domain-specific feature preferences, and addresses reproducibility issues. The findings highlight that commonly used features include MFCCs, Mel spectrograms, STFT/spectrograms, chroma, RMS energy, zero-crossing rate, spectral descriptors, pitch and rhythm features, and multi-feature fusion. Spectrogram-like inputs are predominantly utilized with CNN models, whereas classical machine learning approaches typically rely on concise handcrafted features. This study emphasizes that feature extraction should be considered a domain- and model-dependent design decision rather than a routine preprocessing step.

KEYWORDS: Audio feature extraction; sound classification; machine learning; deep learning; MFCC; Mel spectrogram; feature fusion; Librosa.