MUSIC GENRE CLASSIFICATION

Introduction

“Where words fail, music speaks.”
Music is a universal form of expression, and genre classification plays an important role in organizing, retrieving, and recommending songs. Traditional methods relied on handcrafted features such as pitch, tempo, and rhythm, but recent advancements in deep learning allow automatic extraction of complex audio features through spectrogram analysis.
This project implements a CNN-based approach for music genre classification using spectrogram images derived from the GTZAN dataset. A GUI is also provided to enable easy audio upload and real-time genre prediction.

Abstract

This project focuses on the automatic classification of music into predefined genres using deep learning techniques. The GTZAN dataset, which contains 1,000 audio tracks across 10 genres, was used for training and evaluation. Spectrograms of the audio clips were generated and processed as input to a Convolutional Neural Network (CNN). The trained model predicts the music genre of unseen audio files with reasonable accuracy, and a simple Graphical User Interface (GUI) was developed for user interaction.

Tools Used

Python: Programming Language
Librosa: Audio processing and feature extraction
Matplotlib: Spectrogram visualization
TensorFlow/Keras: Deep Learning framework for building CNN
Tkinter: GUI framework for user interaction
GTZAN Dataset: Dataset containing audios and spectrograms of 10 music genres.

Steps Involved in building Project

1) Data Preparation
    - GTZAN dataset containing 10 genres (blues, classical, country, disco, hiphop, jazz, metal, pop, reggae, rock).
    - Converted each audio file into spectrogram images (128×128 pixels).
    - Normalized data and organized it into training and testing sets.
2) Feature Extraction
    - Mel-spectrograms were generated using Librosa.
    - Spectrograms were resized and scaled to prepare inputs for the CNN.

3) Model Building (CNN)
    - A CNN was designed with multiple convolutional, pooling, and dropout layers.
    - Input shape: 128×128×3 (RGB spectrogram images).
4) Model Training
    - Data was split into training and validation sets.
    - Model trained for multiple epochs(25).
5) Evaluation & Testing
    - Predictions were tested on both training spectrograms and new audio files.
    - GUI implemented for real-time testing: user uploads an audio file → spectrogram generated → CNN predicts genre with confidence score.
6) GUI Integration
    - Built using Tkinter.
    - Features: Upload audio button, prediction popup with genre and confidence

Conclusion

The project successfully demonstrates music genre classification using a CNN model trained on spectrogram images. While initial versions of the model faced challenges such as label mismatches and low confidence predictions, refinements in preprocessing and training improved accuracy. A GUI was also developed to provide end-users with an interactive way to classify audio files in real time.

Future improvements could include using more advanced architectures (e.g., CNN-LSTM, ResNet), larger datasets, and real-time streaming audio classification for deployment in music recommendation systems.