

ABSTRAK

Musik memiliki peran penting dalam kehidupan manusia sebagai bentuk ekspresi budaya, emosi, dan kreativitas. Genre musik menjadi salah satu aspek penting dalam membedakan karakteristik musik berdasarkan gaya, ritme, instrumen, dan latar belakang budayanya. Meskipun pendengar sering dapat mengenali genre musik secara umum, perbedaan antar-genre sering kali sulit dijelaskan secara rinci oleh masyarakat awam tanpa pengalaman mendalam. Dengan meningkatnya jumlah musik digital yang tersedia, pendengar menghadapi tantangan dalam menemukan musik sesuai preferensi secara manual. Oleh karena itu, diperlukan solusi berbasis teknologi untuk mengklasifikasikan genre musik secara otomatis.

Penelitian ini menggunakan metode *Convolutional Neural Network* (CNN) dengan arsitektur ResNet101 untuk mengklasifikasikan 10 genre musik dari GTZAN Dataset, yang mencakup genre *blues, classical, country, disco, hiphop, jazz, metal, pop, reggae, dan rock*. Data *preprocessing* dilakukan untuk meningkatkan kualitas data melalui *cleaning, splitting data, augmentasi, dan ekstraksi fitur* menggunakan Mel Frequency Cepstral Coefficient (MFCC. Augmentasi melibatkan teknik seperti *time stretching, pitch negative, pitch positive, noise injection, dan time shifting* yang bertujuan meningkatkan keragaman data sehingga performa model menjadi lebih optimal.

ResNet101 dipilih karena arsitektur ini mampu meningkatkan akurasi melalui mekanisme *residual learning* dan *skip connections*, yang memfasilitasi pelatihan jaringan dalam tanpa mengurangi stabilitas gradien. Selain itu, penerapan *dropout* memberikan regularisasi tambahan untuk mengurangi risiko *overfitting*. Penelitian terdahulu menunjukkan efektivitas ResNet101 dalam berbagai aplikasi klasifikasi, termasuk pengenalan pola pada data citra maupun sinyal audio.

Hasil pengujian menunjukkan bahwa penerapan metode *Convolutional Neural Network* (CNN) ResNet101 dengan ekstraksi *fitur Mel Frequency Cepstral Coefficient* (MFCC) untuk klasifikasi genre musik berhasil menghasilkan akurasi terbaik sebesar 58% pada 10 kelas genre musik menggunakan parameter optimal seperti *optimizer SGDprop, dropout 0,125-0,25, batch size 32, dan learning rate 0,001*. Dengan adanya hasil penelitian ini diharapkan dapat menjadi acuan untuk mengembangkan sistem yang dapat mendukung pencinta musik dalam mengeksplorasi genre secara lebih efisien.

Kata kunci: Klasifikasi, GTZAN, Genre Musik, Augmentasi, *Mel Frequency Cepstral Coefficient, Convolutional Neural Network, ResNet101, Confusion Matrix*

ABSTRACT

Music plays a significant role in human life as a form of cultural, emotional, and creative expression. Music genres are one of the key aspects in distinguishing the characteristics of music based on style, rhythm, instruments, and cultural background. While listeners can often recognize music genres generally, the differences between genres are often difficult for the general public to explain in detail without in-depth experience. With the increasing amount of digital music available, listeners face challenges in manually finding music that matches their preferences. Therefore, technology-based solutions are needed to automatically classify music genres.

This research using the Convolutional Neural Network (CNN) method with the ResNet101 architecture to classify 10 music genres from the GTZAN Dataset, which includes blues, classical, country, disco, hip-hop, jazz, metal, pop, reggae, and rock genres. Data preprocessing is performed to improve data quality through cleaning, data splitting, augmentation, and feature extraction using Mel Frequency Cepstral Coefficient (MFCC). Augmentation techniques such as time stretching, pitch negative, pitch positive, noise injection, and time shifting are applied to enhance data diversity, thereby optimizing model performance.

ResNet101 was chosen because its architecture can enhance accuracy through residual learning and skip connections, facilitating deep network training without reducing gradient stability. Additionally, the application of dropout provides extra regularization to reduce the risk of overfitting. Previous studies have demonstrated the effectiveness of ResNet101 in various classification applications, including pattern recognition in image and audio signal data.

The results show that applying the Convolutional Neural Network (CNN) method with the ResNet101 architecture and Mel Frequency Cepstral Coefficient (MFCC) feature extraction for music genre classification achieved the best accuracy of 58% across 10 music genre classes, using optimal parameters such as the SGDprop optimizer, dropout rates of 0.125–0.25, a batch size of 32, and a learning rate of 0.001. These findings are expected to serve as a reference for developing systems that support music enthusiasts in exploring genres more efficiently.

Keywords: *Classification, GTZAN, Music Genre, Augmentation, Mel Frequency Cepstral Coefficient, Convolutional Neural Network, ResNet101, Confusion Matrix*