ABSTRAK

Kemajuan teknologi digital telah mempermudah manipulasi audio, termasuk konversi format yang dapat menyamarkan keaslian file audio WAV, sehingga menimbulkan tantangan serius dalam bidang forensik digital, validasi hak cipta, dan autentikasi suara. File WAV asli, yang dikenal karena kualitasnya yang tidak terkompresi, sering kali sulit dibedakan dari file hasil konversi secara kasat mata atau pendengaran, terutama akibat proses kompresi-dekompresi yang dapat mengaburkan karakteristik spektral. Hal ini berpotensi memunculkan risiko penyalahgunaan data audio, seperti pemalsuan dalam konteks hukum atau pelanggaran hak cipta. Oleh karena itu, diperlukan metode yang andal untuk mengidentifikasi keaslian audio WAV dengan akurasi tinggi guna mendukung kebutuhan keamanan digital. Penelitian ini berfokus pada pengembangan model klasifikasi untuk membedakan file audio WAV asli dan hasil konversi, menjawab kebutuhan akan teknik autentikasi audio yang efektif dan akurat.

Penelitian ini menggunakan pendekatan machine learning dengan Mel-Frequency Cepstral Coefficients (MFCC) untuk mengekstrak fitur spektral audio dan Extreme Gradient Boosting (XGBoost) sebagai algoritma klasifikasi. Data audio WAV asli dikumpulkan dari komunitas audio untuk membentuk dataset penelitian. Proses ekstraksi fitur MFCC mencakup tahapan pre-emphasis, framing, windowing, transformasi Fourier, Mel filter bank, dan Discrete Cosine Transform, yang menghasilkan fitur numerik untuk merepresentasikan karakteristik spektral audio. Fitur-fitur ini kemudian digunakan sebagai input untuk melatih model XGBoost, yang dipilih karena keunggulannya dalam menangani data non-linear, efisiensi memori, dan kemampuan regularisasi untuk mencegah overfitting. Dataset dibagi menjadi 80% data latih dan 20% data uji untuk memastikan evaluasi model yang optimal, dengan parameter seperti log loss digunakan sebagai metrik evaluasi selama pelatihan.

Hasil pengujian menunjukkan bahwa model mencapai akurasi 83,42%, presisi 83,88%, recall 83,42%, dan F1-score 83,39%, mengindikasikan kemampuan konsisten untuk membedakan audio asli dan hasil konversi dengan tingkat kebenaran sekitar 83-84%. Pengujian tambahan pada 20 file audio baru menghasilkan akurasi 90%, meskipun terjadi dua kesalahan klasifikasi pada file hasil konversi akibat kemiripan karakteristik spektral. Penelitian ini berhasil menjawab permasalahan dengan menghasilkan model klasifikasi yang efektif untuk autentikasi audio, diharapkan dapat mendukung aplikasi forensik digital, perlindungan hak cipta, dan verifikasi suara. Kontribusi penelitian ini terletak pada pengembangan metodologi machine learning untuk pemrosesan sinyal audio non-stasioner, membuka peluang untuk pengembangan lebih lanjut dalam analisis keaslian audio di era digital.

Kata Kunci: Audio WAV, MFCC, XGBoost, Klasifikasi Audio, Autentikasi Digital.

ABSTRACT

The advancement of digital technology has facilitated audio manipulation, including format conversion that can obscure the authenticity of WAV audio files, posing significant challenges in digital forensics, copyright validation, and voice authentication. Original WAV files, known for their uncompressed quality, are often indistinguishable from converted files through auditory or visual inspection, particularly due to compression-decompression processes that obscure spectral characteristics. This raises the risk of audio data misuse, such as forgery in legal contexts or copyright infringement. Therefore, a reliable method is needed to accurately identify the authenticity of WAV audio files to support digital security needs. This research focuses on developing a classification model to distinguish original WAV audio files from those converted, addressing the need for effective and accurate audio authentication techniques.

This study employs a machine learning approach using Mel-Frequency Cepstral Coefficients (MFCC) for spectral feature extraction and Extreme Gradient Boosting (XGBoost) as the classification algorithm. The dataset, consisting of original WAV audio files, was collected from an audio community to form the research data. The MFCC feature extraction process involves pre-emphasis, framing, windowing, Fourier transform, Mel filter bank, and Discrete Cosine Transform, producing numerical features that represent the audio's spectral characteristics. These features serve as input for training the XGBoost model, selected for its ability to handle non-linear data, memory efficiency, and regularization to prevent overfitting. The dataset was split into 80% training and 20% testing data to ensure optimal model evaluation, with parameters such as log loss used as the evaluation metric during training.

The test results demonstrate that the model achieves an accuracy of 83.42%, precision of 83.88%, recall of 83.42%, and F1-score of 83.39%, indicating consistent performance in distinguishing original and converted audio with an accuracy rate of approximately 83-84%. Additional testing on 20 new audio files yielded a 90% accuracy, though two converted files were misclassified due to spectral similarities. This research successfully addresses the problem by developing an effective classification model for audio authentication, expected to support applications in digital forensics, copyright protection, and voice verification. Its contribution lies in advancing machine learning methodologies for processing non-stationary audio signals, paving the way for further development in audio authenticity analysis in the digital era.

Keywords: *WAV Audio, MFCC, XGBoost, Audio Classification, Digital Authentication.*