

ABSTRACT

This research develops a fake stereo audio detection system using Stereo Panning Features (SPF) and classifiers such as Support Vector Machine (SVM) and eXtreme Gradient Boosting (XGBoost). The system aims to distinguish between real stereo audio and fake stereo audio generated by traditional methods (such as the Haas effect) and advanced anti-forensic techniques like Generative Adversarial Networks (GANs). Experiments were conducted comparing two main features: MFCC and SPF, for detecting fake stereo audio. The results show that the SPF-based models outperform the MFCC models in terms of accuracy, precision, recall, and F1-score. The SPF + XGBoost model achieved the highest performance with 99.902% accuracy, 99.806% precision, and a very low false acceptance rate (FAR) of 0.195%. These findings indicate that SPF is an effective feature for detecting stereo audio manipulations, particularly GAN-generated fake audio. The system developed can be utilized to detect audio manipulations more accurately and efficiently, showing significant potential for applications in audio forensics, particularly for detecting deepfake audio. The study highlights the potential of using spatial features in improving the accuracy of audio manipulation detection.

Keywords: Fake Audio, Audio Detection, Stereo Panning Features, GAN, SVM, XGBoost, Audio Forensics.