Abstract:

Persons can recognize a speaker by listening to their voice. The topic of speaker recognition research is fascinating since there are still a lot of unanswered questions and gaps in the literature that need further research. Many techniques utilizing deep learning (DL) and machine learning (ML) have been utilized to address issues with speaker recognition, particularly in research using large volumes of voice data. At the moment, the volume of data is growing at an incredibly rapid rate. Around the world, there will inevitably be a data explosion. Numerous sources produce hundreds of petabytes of data, such as social media, mobile devices, f inancial market data, astronomy, personal archives, health data, and cameras. As a result, finding the right technique for converting huge amounts of data into information that improves people's lives is a difficult field of studies. In this work, a convolution neural network (CNN) based deep learning model for speaker identification is proposed. The suggested CNN-based methodology employs the standard Mel Frequency Cepstra1 coefficients (MFCCs)-based feature extraction method, which is the most widely used feature selection method for audio and voice signals. The speaker identification system is presented in brief in this research work, after which the general architecture of the system utilizing the CNN model is discussed. Comparing the results of this study to others is challenging due to variations in methodologies and implementation contexts, which influence accuracy rates and other evaluation metrics. Despite this, it is remains necessary to observe the variations in a brief summary to understand the distinctive features of each strategy when comparison to the proposed one.