A Real-Time Convolutional Approach To Speech Emotion Recognition
J. Engelbart

Abstract

In the task of Speech Emotion Recognition, the speakers' emotion is classified based on audio recordings of their voice. This research explores the use of machine learning models to recognize emotions in speech in real-time. For this purpose, this research presents the DeepEmoNet architecture. DeepEmoNet uses a CNN based architecture to classify emotion on a raw audio input. The experiments show that this architecture achieves a score that can compete with the current state-of-the-art approaches. Furthermore, the experiments show that the model is able to achieve that score in real-time.