Diversity of Voices: Audio Visual Continuous Speech Recognition with Traditional Approach

Authors : Partha Protim Majumder, Sajeeb Das, Sharun Akter Khushbu

Abstract : Bengali is widely spoken in the world, but Bengali speech recognition has not received much attention. Here, we are conducting the toughest task because it must be performed in a noisy place in our study. Another challenge we overcome is dealing with speeches and collecting data on third genders, and our approach is to recognize the gender in speeches. All of the Bangla speech samples used in this study were short and were taken from real-life situations. We employed the male, female, and third-gender categories of speech. In this study, we derive the feature from the spoken word. We used MFCC(1-20), ZCR,rolloff,spec_cen, RMSE, and chroma_stft. Here, we used the algorithms Gboost, Random Forest, K-Nearest Neighbors (KNN), Decision Tree, Naive Bayes, and Logistic Regression (LR) to assess the performance of recognition metrics, and we got the highest performance from random forest in recognizing the gender of the speeches.