Speech Emotion Recognition for Autism Spectrum Disorder Using Deep Learning

MD Gunathilake[1#] and GAI Uwanthika[1]

[1]Faculty of Computing, General Sir John Kotelawala Defence University, Sri Lanka

#36-cs-0002@kdu.ac.lk

Children who belong to autism spectrum disorder have difficulty in identifying emotions and expressing their emotions because it is hard to identify the emotions like anger, disgust, fear, happiness, sadness and surprise in other people and in themselves. This can be even more severe when it could not be found at the beginning and may lead to impairment of social communication of the child. Through the proposed systematic methodology, children can identify their basic emotions and try to express them. This evolved methodology was developed using python language. For emotion recognition, a deep machine learning model like Recurrent Neural Network (RNN) using Keras with a TensorFlow backend was used. RNN consists of four layers with two long short-term memory (LSTM) layers. To optimize the performance of the model Adam optimizer was used. For the training and testing of the model online available data were used. For the classification of the emotion's valuable features of the audio signal like Zero Crossing Rate (ZCR), Chroma STFT, Mel-Frequency Cepstral Coefficient (MFCC), Root Mean Square (RMS) value, and Mel spectrogram were extracted using the python libROSA library. Due to the lack of the data amount and GPU requirements model's performance can decrease. This model performed well with the TESS data corpus with 91% test accuracy.

Keywords: *speech emotion recognition, autism spectrum disorder, ZCR, Chroma STFT, MFCC, RMS, MEL spectrogram*