OpenSmile

- Fake Speech detection
- Extracts 88 dimensions vector from Speech
- Test was made with session, gender, domain, and synthesizer variability.
- Except for domain variability model worked well for others.
- Extracted features like signal energy , pitch, loudness, MFCC, PLPCC etc
- The degradation in the domain variability case needs to be addressed

Modulation Spectrogram

- Fake speech detection
- Test was made with session, gender, domain, and synthesizer variability.
- Speech -> FIR Filter Bank -> some processing -> Modulation Spectrogram Image
- Resnet -34 architecture : 34 layer CNN with Relu and Softmax
- The degradation in the performance due to domain variability has to be addressed

Vowel Onset Point Analysis

- Speaker Verification
- Aimed at identifying Vowel like regions
- Less number of frames from vowel-like regions sufficient to model in clear conditions
- Gaussian Mixture Modeling (GMM) is employed for speaker modelling
- Future focus on better resolution and accuracy with algorithms to to separate vowel region from other regions of speech

Rhythm Formant Analysis

- Classifies depressed vs non depressed speech
- Features : Amplitude Modulation and Frequency Modulation and their combinations.
- FM RF as a feature worked out well
- Experiment with DT , LR and Random Forest.
- Used top 6 FM RF features
- DT gave less biased results
- Larger and balanced datasets , and some more relevant features could help the performance.

Srihari K G
IIT Dharwad