Natural and Synthetic Speech Classification

Vikram Bhatt[1,*]

[1]Department of Computational and Data Sciences
[*]vikrambhatt@iisc.ac.in

ABSTRACT

Classify speech as natural (genuine) or synthetic (machine generated). we train a DNN on the spoofing challenge training data to discriminate between human and spoofed speech signals. Delta filterbank spectra (DFB), delta plus double delta Mel-frequency cepstral coefficients (DMCC), delta plus double delta linear prediction cepstral coefficients (DLPCC) and product spectrum-based cepstral coefficients (DPSCC) features are used as inputs to the DNN. For each feature, posteriors and bottleneck features (BNF) are then generated for all the spoofing challenge data using the trained DNN. The DNN posteriors are directly used to decide if a test recording is spoofed or human. For spoofing detection with the acoustic level features and the bottleneck features we build a standard Gaussian Mixture Model (GMM) classifier.