ABSTRACT

An automatic speaker verification (ASV) is one of the challenging problem in speech processing since there are so many models of machine learnings those capable of synthesizing a fake speech from a given text. This paper discusses the impact of SpecAugment to state of the art methods such as Gaussian Mixture Models (GMM) and Deep Neural Networks (DNNs). Some experiments on a speech dataset sampled from the ASVSpoof2019, which is specially made to tackle the threat of spoofing, show that GMM produces an Equal Error Rate (EER) of 19.0% that is better than the DNNs system with EER of 24.0%. However, after combining with a traditional augmentation technique, the DNN gives a better EER of 15.3% than GMM with EER of 15.7%.