World Academy of Science, Engineering and Technology
International Journal of Mathematical and Computational Sciences
Vol:14, No:12, 2020

USE-Net: SE-Block Enhanced U-Net Architecture for Robust Speaker Identification

Authors : Kilari Nikhil, Ankur Tibrewal, Srinivas Kruthiventi S. S.

Abstract : Conventional speaker identification systems often fall short of capturing the diverse variations present in speech data due to fixed-scale architectures. In this research, we propose a CNN-based architecture, USENet, designed to overcome these limitations. Leveraging two key techniques, our approach achieves superior performance on the VoxCeleb 1 Dataset without any pre-training. Firstly, we adopt a U-net-inspired design to extract features at multiple scales, empowering our model to capture speech characteristics effectively. Secondly, we introduce the squeeze and excitation block to enhance spatial feature learning. The proposed architecture showcases significant advancements in speaker identification, outperforming existing methods, and holds promise for future research in this domain.

Keywords : multi-scale feature extraction, squeeze and excitation, VoxCeleb1 speaker identification, mel-spectrograms, USENet
Conference Title : ICSRD 2020 : International Conference on Scientific Research and Development
Conference Location : Chicago, United States
Conference Dates : December 12-13, 2020