

## REFERENCES

- [1] Y. Jung, Y. Choi, and H. Kim, "Self-Adaptive Soft Voice Activity Detection using Deep Neural Networks for Robust Speaker Verification," 2019.
- [2] W. Ping *et al.*, "Deep Voice 3: Scaling Text-to-Speech with Convolutional Sequence Learning," pp. 1–16, 2017.
- [3] S. Arik, J. Chen, K. Peng, W. Ping, and Y. Zhou, "Neural voice cloning with a few samples," *Adv. Neural Inf. Process. Syst.*, vol. 2018-Decem, no. Nips, pp. 10019–10029, 2018.
- [4] T. Nakamura, Y. Saito, S. Takamichi, Y. Ijima, and H. Saruwatari, "V2S attack: building DNN-based voice conversion from automatic speaker verification," 2019.
- [5] Z. Ren, Z. Chen, and S. Xu, "Triplet Based Embedding Distance and Similarity Learning for Text-independent Speaker Verification," 2019.
- [6] J. S. Chung, A. Nagrani, and A. Zisserman, "VoxceleB2: Deep speaker recognition," *Proc. Annu. Conf. Int. Speech Commun. Assoc. INTERSPEECH*, vol. 2018-Sept, no. ii, pp. 1086–1090, 2018.
- [7] D. Wang, L. Li, Z. Tang, and T. F. Zheng, "Deep speaker verification: Do we need end to end?," *Proc. - 9th Asia-Pacific Signal Inf. Process. Assoc. Annu. Summit Conf. APSIPA ASC 2017*, vol. 2018-Febru, pp. 177–181, 2018.
- [8] H. Hussain, S. H. Salleh, C. M. Ting, A. K. Ariff, I. Kamarulafizam, and R. A. Suraya, "Speaker verification using Gaussian mixture model (GMM)," *IFMBE Proc.*, vol. 35 IFMBE, pp. 560–564, 2011.
- [9] A. Nagraniy, J. S. Chungy, and A. Zisserman, "VoxCeleb: A large-scale speaker identification dataset," *Proc. Annu. Conf. Int. Speech Commun. Assoc. INTERSPEECH*, vol. 2017-Augus, pp. 2616–2620, 2017.
- [10] H. Salehghaffari, "Speaker Verification using Convolutional Neural Networks," 2018.
- [11] S. Pascual, A. Bonafonte, and J. Serra, "SEGAN: Speech enhancement generative adversarial network," *Proc. Annu. Conf. Int. Speech Commun. Assoc. INTERSPEECH*, vol. 2017-Augus, no. D, pp. 3642–3646, 2017.
- [12] E. Variani, X. Lei, E. McDermott, I. L. Moreno, and J. Gonzalez-Dominguez, "Deep neural networks for small footprint text-dependent speaker verification," *ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc.*, pp. 4052–4056, 2014.
- [13] A. Torfi, J. Dawson, and N. M. Nasrabadi, "Text-Independent Speaker Verification Using 3D Convolutional Neural Networks," *Proc. - IEEE Int. Conf. Multimed. Expo*, vol. 2018-July, 2018.
- [14] Q. Wang, C. Downey, L. Wan, P. A. Mansfield, and I. L. Moreno, "Speaker diarization with LSTM," *ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc.*, vol. 2018-April, pp. 5239–5243, 2018.
- [15] A. Sriram, H. Jun, Y. Gaur, and S. Satheesh, "Robust speech recognition using generative adversarial networks," *ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc.*, vol. 2018-April, pp. 5639–5643, 2018.
- [16] K. Wang, J. Zhang, S. Sun, Y. Wang, F. Xiang, and L. Xie, "Investigating generative adversarial networks based speech dereverberation for robust speech recognition," *Proc. Annu. Conf. Int. Speech Commun. Assoc. INTERSPEECH*, vol. 2018-Sept, pp. 1581–1585, 2018.
- [17] D. S. Park *et al.*, "SpecAugment: A Simple Data Augmentation Method for Automatic Speech Recognition," pp. 2–6, 2019.
- [18] M. Todisco *et al.*, "ASVspoof 2019: Future Horizons in Spoofed and Fake Audio Detection," pp. 3–7, 2019.