

## Documents

Gunawan, T.S.<sup>a</sup>, Sarif, M.R.M.<sup>a</sup>, Kartiwi, M.<sup>a</sup>, Ahmad, Y.A.<sup>b</sup>

**Development of U-Net Architecture for Audio Super Resolution**

(2023) *Proceedings of the 9th International Conference on Computer and Communication Engineering, ICCCE 2023*, pp. 132-137.

DOI: 10.1109/ICCCE58854.2023.10246115

<sup>a</sup> International Islamic University Malaysia, Electrical and Computer Engineering Department, Kuala Lumpur, 53100, Malaysia

<sup>b</sup> International Islamic University Malaysia, Information Systems Department, Kuala Lumpur, 53100, Malaysia

**Abstract**

Audio processing is used in a wide range of applications, including telecommunications and music streaming. Audio quality degradation during transmission and processing is a common issue in these fields, often caused by bandwidth constraints and the use of subpar equipment. This problem is exacerbated when the task requires converting low-quality audio input to high-resolution output, which is difficult for deep neural networks to do. To improve audio superresolution, this paper proposes a novel solution to this problem by embedding a U-Net-based architecture model within deep neural networks. Over 100 iterations, the U-Net architecture was trained, with loss values and Mean Squared Error (MSE) monitored at each epoch. A diverse dataset of audio signals with varying Signal-to-Noise Ratio (SNR) values ranging from 1 dB to 30 dB was used. The model's average SNR of 17.29 dB exceeds thresholds where listener detection of enhancements becomes difficult, demonstrating its ability to preserve subtle audio details. Furthermore, the Log-Spectral Distance (LSD) values revealed a mere 1.41 dB difference between the actual and reconstructed spectrograms, indicating that the model can recover lost information during upsampling. This research suggests a promising method for improving audio quality, particularly when bandwidth constraints or insufficient equipment prevent high-resolution audio transmission and processing. © 2023 IEEE.

**Author Keywords**

audio quality enhancement; audio super-resolution; Low-quality audio; U-Net architecture

**Index Keywords**

Audio acoustics, Bandwidth, Deep neural networks, Mean square error, Network architecture; Audio processing, Audio quality, Audio quality enhancement, Audio super-resolution, Low qualities, Low-quality audio, NET architecture, Quality enhancement, Superresolution, U-net architecture; Signal to noise ratio

**Funding details**

Universiti Malaysia PahangUMP

International Islamic University MalaysiaIUMIUMP-SRCG22-001-0001

ACKNOWLEDGMENT This research is funded by the University Malaysia Pahang (UMP) and International Islamic University Malaysia (IIUM) sustainable research collaboration 2022 grant scheme no. IUMP-SRCG22-001-0001.

**References**

- Mira, R., Vougioukas, K., Ma, P., Petridis, S., Schuller, B.W., Pantic, M.  
**End-to-end video-to-speech synthesis using generative adversarial networks**  
(2022) *Ieee Transactions on Cybernetics*,
- El-Hajjar, M., Hanzo, L.  
**A survey of digital television broadcast transmission techniques**  
(2013) *Ieee Communications Surveys & Tutorials*, 15 (4), pp. 1924-1949.
- Gunawan, T.S., Zain, M.K.M., Muin, F.A., Kartiwi, M.  
**Investigation of lossless audio compression using ieee 1857.2 advanced audio coding**  
(2017) *Indonesian Journal of Electrical Engineering and Computer Science*, 6 (2), pp. 422-430.
- Wang, H., Wang, D.  
**Towards robust speech super-resolution**

(2021) *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 29, pp. 2058-2066.

- Deng, J.  
**Exploiting time-frequency patterns with LSTM-RNNs for low-bitrate audio restoration**  
(2020) *Neural Computing and Applications*, 32 (4), pp. 1095-1107.
- Rakotonirina, N.C.  
**Self-attention for audio super-resolution**  
(2021) *2021 IEEE 31st International Workshop on Machine Learning for Signal Processing (MLSP)*, pp. 1-6.
- Lim, T.Y., Yeh, R.A., Xu, Y., Do, M.N., Hasegawa-Johnson, M.  
**Time-frequency networks for audio super-resolution**  
(2018) *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 646-650.
- Tong, H., Yang, Z., Wang, S., Hu, Y., Saad, W., Yin, C.  
**Federated learning based audio semantic communication over wireless networks**  
(2021) *2021 IEEE Global Communications Conference (GLOBECOM)*, pp. 1-6.
- Saito, Y., Takamichi, S., Saruwatari, H.  
**Statistical parametric speech synthesis incorporating generative adversarial networks**  
(2017) *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 26 (1), pp. 84-96.
- Gui, J., Sun, Z., Wen, Y., Tao, D., Ye, J.  
**A review on generative adversarial networks: Algorithms, theory, and applications**  
(2021) *IEEE Transactions on Knowledge and Data Engineering*, 35 (4), pp. 3313-3332.
- Su, J., Wang, Y., Finkelstein, A., Jin, Z.  
**Bandwidth extension is all you need**  
(2021) *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 696-700.

**Correspondence Address**

Gunawan T.S.; International Islamic University Malaysia, Malaysia; email: tsgunawan@iiu.edu.my

**Publisher:** Institute of Electrical and Electronics Engineers Inc.

**Conference name:** 9th International Conference on Computer and Communication Engineering, ICCCE 2023

**Conference date:** 15 August 2023 through 16 August 2023

**Conference code:** 192690

**ISBN:** 9798350325218

**Language of Original Document:** English

**Abbreviated Source Title:** Proc. Int. Conf. Comput. Commun. Eng., ICCCE 2-s2.0-85173656002

**Document Type:** Conference Paper

**Publication Stage:** Final

**Source:** Scopus