

## REFERENCE

- Ashwin, J. S., & Manoharan, N. (2018). Audio denoising based on short time Fourier transform. *Indonesian Journal of Electrical Engineering and Computer Science*, 9(1), 89–92. <https://doi.org/10.11591/ijeecs.v9.i1.pp89-92>
- Dang, F., Chen, H., & Zhang, P. (2022). DPT-FSNet: Dual-path transformer based full-band and sub-band fusion network for speech enhancement. *ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 6857-6861.
- Fink, D. M. (2023). A new definition of noise. *The Hearing Journal*, 76(11), 6–7. <https://doi.org/10.1097/01.HJ.0000995260.15519.9e>
- Hossain, M. N., Basir, S., Hosen, M. S., Asaduzzaman, A. O., Islam, M. M., Hossain, M. A., & Islam, M. S. (2022). Supervised single channel speech enhancement method using UNET. *Electronics*, 12(14), 3052. <https://doi.org/10.3390/electronics12143052>
- Kanazawa, L., Tokumura, T., Nakagawa, T., Ozeki, Y., & Suzuki, K. (2025). Impact of office acoustic environment on workplace productivity: Effects of background sounds investigated using speech. *The Journal of the Acoustical Society of Japan*, 81(9), 600-603.
- Kang, Z., Huang, Z., & Lu, C. (2022). Speech enhancement using U-Net with compressed sensing. *Applied Sciences*, 12(9), 4161. <https://doi.org/10.3390/app12094161>
- Ke, Y., Li, A., Zheng, C., Peng, R., & Li, X. (2021). Low-complexity artificial noise suppression methods for deep learning-based speech enhancement algorithms. *EURASIP Journal on Audio, Speech, and Music Processing*, 2021(1), 17. <https://doi.org/10.1186/s13636-021-00204-9>
- Kim, C., & Stern, R. (2008). Robust signal-to-noise ratio estimation based on waveform amplitude distribution analysis. *Proceedings of Interspeech 2008*, 2598–2601. <https://doi.org/10.21437/Interspeech.2008-644>
- Lee, G. W., & Kim, H. K. (2020). Multi-task learning U-Net for single-channel speech enhancement and mask-based voice activity detection. *Applied Sciences*, 10(9), 3230. <https://doi.org/10.3390/app10093230>
- Loizou, P. C. (2013). *Speech enhancement: Theory and practice* (2nd ed.). CRC Press. <https://doi.org/10.1201/b14529>

- Nustede, E. J., & Anemüller, J. (2023). *Single-channel speech enhancement with deep complex U-Networks and probabilistic latent space models*. arXiv. <https://arxiv.org/abs/2309.01535>
- Pandey, A., & Wang, D. (2022). Self-attending RNN for speech enhancement to improve cross-corpus generalization. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 30, 1374-1385.
- Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. *International Conference on Medical Image Computing and Computer-assisted Intervention*, 234-241. Springer.
- Schiller, I., Aspöck, L., & Schlittmeier, S. (2023). The impact of a speaker's voice quality on auditory perception and cognition: A behavioral and subjective approach. *Frontiers in Psychology*, 14, 1243249. <https://doi.org/10.3389/fpsyg.2023.1243249>
- Shetty, S. S., & Reeja. (2015). *Audio noise removal – The state of the art*.
- Taal, C. H., Hendriks, R. C., Heusdens, R., & Jensen, J. (2010). A short-time objective intelligibility measure for time-frequency weighted noisy speech. *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, 4214-4217. <https://doi.org/10.1109/ICASSP.2010.5495701>
- Tan, K., & Wang, D. (2018). A convolutional recurrent neural network for real-time speech enhancement. *Proceedings of Interspeech 2018*, 3229–3233. <https://doi.org/10.21437/Interspeech.2018-1405>
- Upadhyay, N., & Jaiswal, R. K. (2016). Single channel speech enhancement: Using Wiener filtering with recursive noise estimation. *Procedia Computer Science*, 84, 22–30. <https://doi.org/10.1016/j.procs.2016.04.061>
- Upadhyay, N., & Karmakar, A. (2015). Speech enhancement using spectral subtraction-type algorithms: A comparison and simulation study. *Procedia Computer Science*, 54, 574-584. <https://doi.org/10.1016/j.procs.2015.06.066>
- Wang, Z.-Q., Wang, P., & Wang, D. (2020). Complex spectral mapping for single- and multi-channel speech enhancement and robust ASR. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 28, 1778–1787. <https://doi.org/10.1109/TASLP.2020.2998279>
- Wang, Z.-Q., Wichern, G., Watanabe, S., & Le Roux, J. (2023). STFT-domain neural speech enhancement with very low algorithmic latency. *IEEE/ACM*

*Transactions on Audio, Speech, and Language Processing*, 31, 397–410.  
<https://doi.org/10.1109/TASLP.2022.3224285>

- Wu, J., Li, Q., Yang, G., Li, L., Senhadji, L., & Shu, H. (2023). Self-supervised speech denoising using only noisy audio signals. *Speech Communication*, 149, 63–73.
- Xu, R., Wu, R., Ishiwaka, Y., Vondrick, C., & Zheng, C. (2020). Listening to sounds of silence for speech denoising. *Advances in Neural Information Processing Systems (NeurIPS)*, 33, 11912–11923.  
<https://proceedings.neurips.cc/paper/2020/hash/6d7d394c9d0c886e9247542e06ebb705-Abstract.html>
- Yan, X., Yang, Z., Wang, T., & Guo, H. (2020). *An iterative graph spectral subtraction method for speech enhancement*. arXiv.  
<https://arxiv.org/abs/2006.08497>
- Zhao, H., Zarar, S., Tashev, I., & Lee, C.-H. (2018). *Convolutional-recurrent neural networks for speech enhancement*. arXiv.  
<https://arxiv.org/abs/1805.00579>
- Zhou, H., Molesworth, B. R. C., Burgess, M., & Hatfield, J. (2023). The effect of moderate broadband noise on cognitive performance: A systematic review. *Cognition, Technology & Work*, 26(1), 1–36.