

Cooking With MSG: Enhancing the Output of Source Separation Using Generative Adversarial Networks

Boaz Cogan^{*1}, Noah Schaffer^{†1}, Ethan Manilow¹, and Bryan Pardo¹

Arxiv DOI: [10.21105/joss.01667](https://arxiv.org/abs/10.21105/joss.01667)

¹ Interactive Audio Lab, Northwestern University.

Abstract

Music source separation is the task of isolating individual components, or sources, in a musical mixture, e.g., isolating a singer from a full band. In general, the output of source separation systems, including recent state-of-the-art methods, exhibits two common issues: unwanted noise not present in the original source (“artifacts”), and missing content from the original source. These two issues greatly affect the perceptual quality of the output. We propose to use an additional post-processing step applied to the output of source separation systems, designed to reduce artifacts and restore missing content to the initial source estimates. Taking inspiration from recent advances in speech processing, we use a Generative Adversarial Network (GAN) as a post-processor. We call our system the Make it Sound Good (MSG) post-processor. Illustrative results on the widely-used MUSDB18 dataset show that MSG is able to reduce the presence of artifacts and restore missing frequencies from an upstream source separation system.

License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)).

In partnership with



Figures

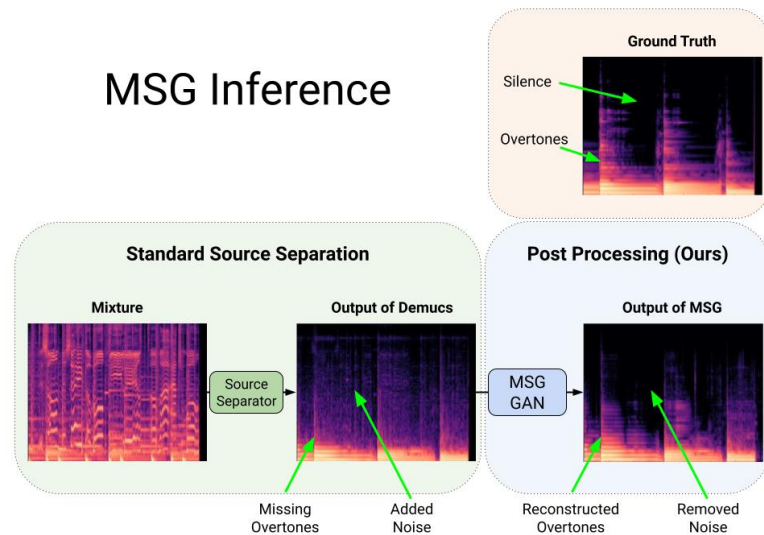


Figure 1: MSG Post-Processor

References

^{*}co-first author

[†]co-first author



- Défossez, A., Usunier, N., Bottou, L., & Bach, F. (2019). Music source separation in the waveform domain. *arXiv Preprint arXiv:1911.13254*.
- Germain, F. G., Chen, Q., & Koltun, V. (2018). Speech denoising with deep feature losses. *arXiv Preprint arXiv:1806.10522*.
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2020). Generative adversarial networks. *Communications of the ACM*, 63(11), 139–144.
- Kumar, K., Kumar, R., Boissiere, T. de, Gestin, L., Teoh, W. Z., Sotelo, J., Brébisson, A. de, Bengio, Y., & Courville, A. (2019). Melgan: Generative adversarial networks for conditional waveform synthesis. *arXiv Preprint arXiv:1910.06711*.
- Le Roux, J., Wisdom, S., Erdogan, H., & Hershey, J. R. (2019). SDR–half-baked or well done? *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 626–630.
- Su, J., Jin, Z., & Finkelstein, A. (2020). HiFi-GAN: High-fidelity denoising and dereverberation based on speech deep features in adversarial networks. *arXiv Preprint arXiv:2006.05694*.