



Self-refining of Pseudo Labels for Music Source Separation with Noisy Labeled Data

Junghyun Koo *¹, Yunkee Chae †², Chang-Bin Jeon¹, and Kyogu Lee ^{1,2,3}¶

¹ Department of Intelligence and Information ² Interdisciplinary Program in Artificial Intelligence ³ Artificial Intelligence Institute, Seoul National University ¶ Corresponding author

Published: 04 November 2023

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/)).

Abstract

Music source separation (MSS) faces challenges due to the limited availability of correctly-labeled individual instrument stems. This paper proposes an automated approach for refining mislabeled instrument tracks in a partially noisy-labeled dataset. Similar, yet different from self-training (Xie et al., 2020), our approach learns directly from noisy labeled data and re-labels the training data. We train the classifier to perform multi-label instrument recognition with mixtures that are synthesized by randomly selecting each stem from the noisy labeled dataset. After this training procedure, we refine the original noisy dataset with the trained classifier and use this new dataset to train the final classifier Ψ , which we refer to as self-refining training. Our self-refining technique results in only a 1% accuracy degradation for multi-label instrument recognition compared to a classifier trained with a clean-labeled dataset. Using the refined multi-labeled dataset, we train the MSS models by randomly selecting stems from the dataset. As a result, we've enhanced the MSS performance compared to the baseline, which was trained with an original noisy labeled dataset, and achieved 9th place in the leaderboard A of the SDX challenge 2023 (Fabbro et al., 2023).

Fabbro, G., Uhlich, S., Lai, C., Choi, W., Martinez-Ramirez, M., Liao, W., Gadelha, I., Ramos, G., Hsu, E., Rodrigues, H., & others. (2023). The sound demixing challenge 2023—music demixing track. *Under Review*.

Xie, Q., Luong, M.-T., Hovy, E., & Le, Q. V. (2020). Self-training with noisy student improves imagenet classification. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 10687–10698.

*co-first author

†co-first author