SDX
Sound Demixing Workshop 2023

BS-RoFormer: The SAMI-ByteDance Music Source Separation System for Sound Demixing Challenge 2023

Ju-Chiang Wang *[1], **Wei-Tsung Lu**†[1], **Qiuqiang Kong** [1], **and Yun-Ning Hung** [1]

1 SAMI, ByteDance Inc.

Abstract

Recently, multi-band frequency-domain approaches such as Band-Split Recurrent Neural Networks (BSRNN) (Luo & Yu, 2023) have been explored and achieved very promising results. In this abstract, we introduce a novel approach based on Band-Split RoPE Transformer (termed as BS-RoFormer) (Lu et al., 2023). Similar to BSRNN, BS-RoFormer relies on a band-split module to project the input complex spectrogram into subband-level representations. Then, instead of RNNs, we arrange a stack of hierarchical Transformers to model the inner-band as well as inter-band sequences for multi-band mask estimation. To improve the training efficacy, we use the Rotary Position Embedding (RoPE) (Su et al., 2021). The BS-RoFormer system trained on MUSDB18HQ and 500 extra songs ranked the first place in the standard MSS track (Leaderboard C) of Sound Demixing Challenge (SDX'23) (Fabbro et al., 2023). It outperformed the second best by a large margin in SDR. Benchmarking our SDX'23 system on MUSDB18HQ shows state-of-the-art results, with an average SDR of 11.99 dB. In ablation study (Luo & Yu, 2023), we demonstrate that a smaller version of BS-RoFormer without extra training data is also competitive, achieving an average SDR of 9.92 dB on MUSDB18HQ.

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. *arXiv Preprint arXiv:2308.06979*.

Lu, W.-T., Wang, J.-C., Kong, Q., & Hung, Y.-N. (2023). Music source separation with band-split RoPE transformer. *arXiv Preprint arXiv:2309.02612*.

Luo, Y., & Yu, J. (2023). Music Source Separation With Band-Split RNN. *IEEE/ACM Trans. Audio Speech Lang. Process.*, *31*, 1893–1901.

Su, J., Lu, Y., Pan, S., Murtadha, A., Wen, B., & Liu, Y. (2021). Roformer: Enhanced transformer with rotary position embedding. *arXiv Preprint arXiv:2104.09864*.

---

*co-first author
†co-first author