



StemGMD: A Large-Scale Multi-Kit Audio Dataset for Deep Drums Demixing

Alessandro Ilic Mezza ¹, Riccardo Giampiccolo ¹, Alberto Bernardini ¹, and Augusto Sarti ¹

¹ Image and Sound Processing Lab, Politecnico di Milano 

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

In music demixing, drums is conventionally treated as a single stem despite being an ensemble of percussive instruments in itself. At the same time, the availability of publicly-available datasets specifically tailored for drums demixing has been severely lacking. To address this gap, we present StemGMD ([Alessandro Ilic Mezza et al., 2023](#)), the first large-scale multi-kit audio dataset of isolated single-instrument drum stems. Each audio clip in StemGMD is synthesized from MIDI recordings of expressive drums performances taken from Magenta’s Groove Midi Dataset (GMD) using ten real-sounding sample libraries from Logic Pro X. Totalling over 1224 hours, StemGMD is the largest dataset of drums to date and the first to comprise isolated audio clips for every instrument in a canonical nine-piece drum kit, i.e., kick drum, snare, high tom, mid-low tom, floor tom, open hi-hat, closed hi-hat, ride cymbal, crash cymbal. StemGMD also contains single hits for each drum piece at varying velocities, and is inherently aligned with the MIDI content and metadata found in GMD, allowing for a broad range of applications such as Automatic Drum Transcription. Furthermore, having access to isolated audio stems enables a vast array of diverse data augmentation methods that draw inspiration from common music production practices. Alongside the dataset, we release a reference model built upon a bank of parallel U-Nets that separates five stems (kick drum, snare, tom-toms, hi-hat, cymbals) from a stereo drum mixture through spectro-temporal soft masking ([A. I. Mezza et al., 2023](#)). Such model is meant to serve as a baseline for future research and might complement existing music demixing models.

Mezza, A. I., Giampiccolo, R., Bernardini, A., & Sarti, A. (2023). *Toward deep drum source separation*.

Mezza, Alessandro Ilic, Giampiccolo, R., Bernardini, A., & Sarti, A. (2023). *StemGMD: A large-scale audio dataset of isolated drum stems for deep drums demixing* (Version 1.0.0) [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.7860223>