

Over the past few decades, neuroscience has seen major advances in understanding the mind thanks to functional magnetic resonance imaging (fMRI), which produces three-dimensional snapshots of brain activity. Recently, large multi-subject datasets have become widely available, leading to the development of new methods that can leverage this data. A challenge of working with multi-subject datasets is that subjects have distinct anatomical and functional structure, making direct aggregation of subject data unfeasible. Standard anatomical alignment methods provide a partial solution to this problem as functional brain topographies do not necessarily align. Lately, functional alignment techniques have been developed to overcome these limitations. The Shared Response Model (SRM) is the state-of-the-art method that maps functional topographies from every subject data to a low-dimensional shared response. This unsupervised machine learning method learns a mapping for each subject and a low-dimensional shared response, which can be applied to map one subject to another, reduce noise, discriminate between groups, and more. Despite its great predictive performance, SRM is hard to compute over a few tenths of subjects. We perform algorithmic optimizations that reduce the memory and runtime complexity and allow us to estimate the model for hundreds of subjects. We further develop a code-optimized distributed version of SRM. This distributed algorithm exhibits promising results of strong scaling results of up to 5x with 20 nodes on real datasets, and successful weak scaling by aligning a synthetic dataset of 1024 subjects in 512 nodes in 51 seconds.