Decentralized Parallel Independent Component Analysis for Multimodal, Multisite Data

Chan Aek Panichvatana[1]
Jiayu Chen, Bradley Baker, Bishal Thapaliya, Vince D Calhoun and Jingyu liu
[1] Georgia State University

Large amounts of neuroimaging and omics data have been generated for studies of mental health. Collaborations among research groups that share data have shown increased power for new discoveries of brain abnormalities, genetic mutations, and associations among genetics, neuroimaging and behavior. However, sharing raw data can be challenging for various reasons. A federated data analysis allowing for collaboration without exposing the raw dataset of each site becomes ideal. Following this strategy, a decentralized parallel independent component analysis (dpICA) is proposed in this study which is an extension of the state-of-art Parallel ICA (pICA). pICA is an effective method to analyze two data modalities simultaneously by jointly extracting independent components of each modality and maximizing connections between modalities. We evaluated the dpICA algorithm using neuroimage and genetic data from patients with schizophrenia and health controls, and compared its performances under various conditions with the centralized pICA. The results showed dpICA is robust to sample distribution across sites as long as numbers of samples in each site are sufficient. It can produce the same imaging and genetic components and the same connections between those components as the centralized pICA. Thus our study supports dpICA is an accurate and effective decentralized algorithm to extract connections from two data modalities.