NiPyPe: A flexible, lightweight and extensible neuroimaging data processing framework

Over the past twenty years, advances in non-invasive *in vivo* neuroimaging have resulted in an explosion of studies investigating human cognition in health and disease. Current imaging studies acquire multi-modal image data (e.g., structural, diffusion, functional) and combine with non-imaging behavioural data, patient and/or treatment history and demographic and genetic information. Several sophisticated software packages (e.g., AFNI, BrainVoyager, FSL, FreeSurfer, NiPy, R, SPM) are used to process and analyze such extensive data. Current neuroimaging software offer users an incredible opportunity to analyze their data in different ways, with different underlying assumptions. In a typical analysis, algorithms from these packages, each with its own set of parameters, process the raw data. However, data collected for a single study can be diverse (highly multi-dimensional) and large, and algorithms suited for one dataset may not be optimal for another. This complicates analysis methods and makes data exploration and inference challenging, and comparative analysis of new algorithms difficult.

This heterogeneous collection of specialized applications creates several technical, practical and social issues that hinder replicable, efficient and optimal use of neuroimaging analysis approaches: 1) No uniform access to neuroimaging analysis software and usage information; 2) No framework for comparative algorithm development and dissemination; 3) Personnel turnover in laboratories often limit methodological continuity and training new personnel takes time; 4) Neuroimaging software packages do not address computational efficiency; 5) Methods sections in published articles are inadequate for reproducing results; and 6) In-depth knowledge of neuroimaging analysis algorithms is limited to few individuals.

To address these issues, we present NiPyPe (Neuroimaging in Python: Pipelines and Interfaces), an open source, community-developed, Python-based software package that easily interfaces with existing software for efficient analysis of neuroimaging data and rapid comparative development of algorithms. NiPyPe, solves the issues by providing a uniform interface to existing neuroimaging software and by facilitating interaction between these packages using workflows. NiPyPe provides an environment that encourages interactive exploration of algorithms from different packages (e.g., SPM, FSL), eases the design of workflows within and between packages, and reduces the learning curve necessary to use different packages. NiPyPe is addressing limitations of existing pipeline systems and creating a collaborative platform for neuroimaging software development. Processing modules and their inputs and outputs are described in an object-oriented manner providing the flexibility to interface with any type of software. The workflow execution engine has a plug-in architecture and supports both local execution on multicore machines and remote execution on clusters. Python has already been embraced by the neuroscientific community and is rapidly gaining popularity. NiPyPe, based on Python, has immediate access to an extensive community and its software, technological resources and support structure. NiPyPe is distributed with a BSD License allowing anyone to make changes and redistribute it. Development is done openly with collaborators from many different labs, allowing rapid adaptation to the varied needs of the evolving neuroimaging community.