



HAL
open science

New longitudinal and deep learning pipelines in the Clinica software platform

Alexandre Routier, Arnaud Marcoux, Mauricio Diaz Melo, Jorge Samper-González, Adam Wild, Alexis Guyot, Junhao Wen, Elina Thibeau-Sutre, Simona Bottani, Stanley Durrleman, et al.

► **To cite this version:**

Alexandre Routier, Arnaud Marcoux, Mauricio Diaz Melo, Jorge Samper-González, Adam Wild, et al.. New longitudinal and deep learning pipelines in the Clinica software platform. OHBM 2020 - Annual meeting of the Organization for Human Brain Mapping, Jun 2020, Montreal / Virtual, Canada. hal-02549242

HAL Id: hal-02549242

<https://inria.hal.science/hal-02549242>

Submitted on 21 Apr 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

New longitudinal and deep learning pipelines in the Clinica software platform

Alexandre Routier¹, Arnaud Marcoux¹, Mauricio Diaz Melo², Jorge Samper-González¹, Adam Wild¹, Alexis Guyot¹, Junhao Wen¹, Elina Thibeau-Sutre¹, Simona Bottani¹, Stanley Durrleman¹, Ninon Burgos¹, Olivier Colliot^{1,8}

¹ ARAMIS Lab, ICM, Inserm U1127, CNRS UMR 7225, Sorbonne University, Inria, Paris, France

² Inria Paris, SED, 75013 Paris, France

Introduction

We present new advances made to Clinica (www.clinica.run), an open source software platform for clinical neuroscience studies. Neuroimaging studies are challenging since they involve several data analysis steps such as image preprocessing, extraction of image-derived features or statistical analysis. The development of machine learning methods for neuroimaging also involves most of these steps. The objective of Clinica is to automate the processing and statistical analysis of neuroimaging data and ease the development of machine learning approaches.

New functionalities have been integrated to Clinica to enable the longitudinal analysis of T1w MRI and PET data, and the development of deep learning classification approaches. Other advances aim to consolidate the platform.

Methods

The core of Clinica is written in Python and mainly relies on Nipype (Gorgolewski et al., 2011) to create pipelines. These pipelines involve the combination of different software packages widely used in the neuroimaging community: FreeSurfer (Fischl, 2012), FSL (Jenkinson et al., 2012), SPM (Frackowiak et al., 1997), ANTs (Avants et al., 2014), MRtrix3 (Tournier et al., 2012), and PETPVC (Thomas et al., 2016). Features extracted with the different pipelines can be used as inputs to statistical or machine learning analysis. The input neuroimaging data are expected to follow the BIDS data structure (Gorgolewski et al., 2016).

New functionalities of Clinica can be divided into three main parts.

1. Clinica now comprises pipelines that are specifically dedicated to the processing of longitudinal data sets: `t1-freesurfer-longitudinal` and `pet-surface-surface`. The longitudinal FreeSurfer stream (Reuter et al., 2012) processes a series of T1w MR images acquired at different time points for the same subject to increase the accuracy of volume and thickness estimates. The resulting pial and white surfaces can then be used for the projection of the PET signal using the methodology described in (Marcoux et al., 2018). Finally, longitudinal surface-based data can be plugged to the `statistics-surface` pipeline, which relies on SurfStat (Worsley et al., 2009).
2. Previous releases provided integration between outputs of Clinica and machine learning algorithms from scikit-learn (Pedregosa et al., 2011). We now provide pipelines that preprocess neuroimaging data for deep learning classification: the `t1-linear` mainly performs affine registration of T1w images to a standard space using ANTs (Avants et al., 2014) and `dl-prepare-dl` prepares raw or processed data for integration with PyTorch (e.g. extraction of patches/slices).
3. Other functionalities were also developed or enhanced:
 - creation of Docker and Singularity images to automate the installation of software dependencies and ease the execution of pipelines;
 - update of the `adni-2-bids` converter that now includes ADNI3 and new modalities (fMRI, new PET tracers);
 - development of `nifd-2-bids` that converts the NIFD dataset (<http://4rtni-ftldni.ini.usc.edu>), which contains data of patients with frontotemporal lobar degeneration, to BIDS;

- update of the `dwi-preprocessing` pipeline that now uses the FSL eddy tool (Andersson et al., 2016a 2016b);
- development of the `statistics-volume` pipeline that enables statistical analysis on voxel-based features using the general linear model through the wrapping of SPM functions (Friston et al, 2007).

Pipelines available in Clinica are listed in **Figure 1**.

Results

The new functionalities of Clinica aim to answer the needs of its target audience. Neuroscientists and clinicians conducting clinical neuroscience studies will benefit from longitudinal pipelines, and researchers developing advanced machine learning algorithms will benefit from new pipelines that ease the application of deep learning approaches to neuroimaging data.

Conclusions

Clinica (www.clinica.run) is an open source software platform that aims to make clinical research studies easier and to pursue the community effort of reproducibility.

References

- Andersson, J. L. R., et al. (2016a). Incorporating outlier detection and replacement into a non-parametric framework for movement and distortion correction of diffusion MR images. *NeuroImage* 141, 556572.
- Andersson, J. L. R., and Sotiropoulos, S. N. (2016b). An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *Neuroimage* 125, 10631078.
- Avants, B. B., et al. (2014). The Insight ToolKit image registration framework. *Front Neuroinform* 8.
- Fischl, B. (2012). FreeSurfer. *NeuroImage* 62, 774781.
- Frackowiak, R. S. J., et al. (1997). *Human Brain Function*. Academic Press USA
- Friston, K., et al. (2007). *Statistical Parametric Mapping*. Elsevier
- Gorgolewski, K., et al. (2011). Nipype: A Flexible, Lightweight and Extensible Neuroimaging Data Processing Framework in Python. *Front. Neuroinform.* 5.
- Gorgolewski, K., et al. (2016). The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Scientific Data* 3, 160044.
- Jenkinson, M., et al. (2012). FSL. *Neuroimage* 62, 782790. doi:10.1016/j.neuroimage.2011.09.015.
- Pedregosa, F., et al. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* 12, 28252830.
- Marcoux, A., et al. (2018). An Automated Pipeline for the Analysis of PET Data on the Cortical Surface. *Frontiers in Neuroinformatics* 12.
- Reuter, M., et al. (2012). Within-subject template estimation for unbiased longitudinal image analysis. *NeuroImage* 61, 14021418.
- Thomas, B. A., et al. (2016). PETPVC: a toolbox for performing partial volume correction techniques in positron emission tomography. *Physics in Medicine and Biology* 61, 79757993.
- Tournier, J.-D., et al. (2012). MRtrix: Diffusion tractography in crossing fiber regions. *Int. J. Imaging Syst. Technol.* 22, 5366.
- Worsley, K., et al. (2009). SurfStat: A Matlab toolbox for the statistical analysis of univariate and multivariate surface and volumetric data using linear mixed effects models and random field theory. *NeuroImage* 47, S102.

Figures

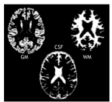

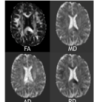









	Voxel-based pipelines	Surface-based pipelines		Other pipelines	
Anatomical MRI	<p>t1-volume Tissue segmentation (GM, WM, CSF), inter-subject registration using Dartzel, spatial normalization to standard space (MNI) Dependencies: SPM, CAT12</p> <ul style="list-style-type: none"> • Voxel-based features (GM, WM, CSF) • Regional features (average GM) using atlases (AAL2, AICHA, Hammers, LPBA40, Neuromorphometrics) 	<p>t1-freesurfer t1-freesurfer-Longitudinal Cortical surface extraction, segmentation of subcortical structures, cortical thickness estimation, spatial normalization to standard space (FSAverage) Dependencies: FreeSurfer</p> <ul style="list-style-type: none"> • Surface-based features (cortical thickness) • Regional features (average cortical thickness) using atlases (Desikan, Destrieux) 	Diffusion MRI	<p>dwi-preprocessing: dwi-dti Correction of raw DWI data, extraction of DTI-based measures, normalization to standard space (MNI) Dependencies: FSL, ANTs, MRtrix3</p> <ul style="list-style-type: none"> • Voxel-based features (FA, MD, AD, RD) • Regional features (average FA, MD, AD, RD) using atlases (JHU DTI81, JHUTracts) 	<p>dwi-preprocessing: dwi-connectome Tractography & connectome Dependencies: FreeSurfer, FSL, MRtrix</p> <ul style="list-style-type: none"> • Probabilistic tractography • Structural connectome using atlases (Desikan, Destrieux) 
PET	<p>pet-volume Registration to T1 MRI, partial volume correction, spatial normalization to standard space (MNI), intensity normalization Dependencies: SPM, PETPVC, CAT12</p> <ul style="list-style-type: none"> • Voxel-based features (FDG uptake, amyloid uptake) • Regional features (average FDG, amyloid uptake) using atlases (AAL2, AICHA, Hammers, LPBA40, Neuromorphometrics) 	<p>pet-surface pet-surface-Longitudinal Registration to T1 MRI, intensity normalization, partial volume correction, projection to cortical surface, spatial normalization to standard space (FSAverage) Dependencies: FreeSurfer, FSL, SPM, PETPVC</p> <ul style="list-style-type: none"> • Surface-based features (FDG uptake, amyloid uptake) • Regional features (average cortical thickness) using atlases (Desikan, Destrieux) 	Machine Learning	<p>machinelearning-prepare-spatial-svm Preparation of T1 MRI and PET data for spatially regularized SVM Dependencies: None</p> <ul style="list-style-type: none"> • Regularization that accounts for the spatial and anatomical structure of neuroimaging data leading to a more regular and anatomically interpretable decision function. • Used as input for machine learning classification 	<p>(No command line interface) Classification based on machine learning Dependencies: scikit-learn</p> <ul style="list-style-type: none"> • Voxel-based, surface-based or regional features • Classifications (SVM, ϵ_2 logistic regression, random forest) using cross-validations (K-fold, repeated K-fold, repeated hold-out) 
Statistics	<p>statistics-volume Voxel-based mass-univariate analysis with SPM Dependencies: SPM, Matlab</p> <ul style="list-style-type: none"> • Voxel-based features from t1-volume or pet-volume pipelines • Group comparison or correlations analysis using GLM 	<p>statistics-surface Surface-based mass-univariate analysis with SurfStat Dependencies: Matlab</p> <ul style="list-style-type: none"> • Surface-based features from t1-freesurfer-[longitudinal] or pet-surface-[longitudinal] pipelines • Group comparison or correlations analysis using GLM 	Deep Learning	<p>t1-linear Bias field correction, affine registration and cropping Dependencies: ANTs</p> <ul style="list-style-type: none"> • T1w on ICBM 2009c nonlinear symmetric template • Used as input for di-prepare-di pipeline 	<p>di-prepare-di Convert features extracted by Clinica to PyTorch tensors Dependencies: None</p> <ul style="list-style-type: none"> • 3D Volumes, slices or patches from t1-linear, t1-volume or pet-volume pipelines • Tensors for PyTorch 

Figure 1 List of the pipelines currently available in Clinica with their dependencies and outputs. Older pipelines are indicated in blue while new or updated pipelines are indicated in purple. Explanations regarding the atlases can be found in the Wiki of Clinica.