

ClinicaDL: an open-source deep learning software for reproducible neuroimaging processing

Elina Thibeau-Sutre, Mauricio Diaz, Ravi Hassanaly, Alexandre M Routier,

Didier Dormont, Olivier Colliot, Ninon Burgos

► To cite this version:

Elina Thibeau-Sutre, Mauricio Diaz, Ravi Hassanaly, Alexandre M Routier, Didier Dormont, et al.. ClinicaDL: an open-source deep learning software for reproducible neuroimaging processing. 3IA Doctoral Workshop, Nov 2021, Toulouse, France. hal-03423072v1

HAL Id: hal-03423072 https://hal.science/hal-03423072v1

Submitted on 9 Nov 2021 (v1), last revised 24 Nov 2021 (v2)

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.

ClinicaDL: an open-source deep learning soft-ware for reproducible neuroimaging processing



AIRIE

<u>Elina Thibeau-Sutre¹, Mauricio Diaz¹, Ravi Hassanaly¹, Alexandre Routier¹,</u> Didier Dormont^{1,2}, Olivier Colliot¹, Ninon Burgos¹

¹Sorbonne Université, Institut du Cerveau, Inserm, CNRS, AP-HP Pitié-Salpêtrière, Inria Équipe-projet ARAMIS, PaRis Artificial Intelligence Research InstitutE Paris, France

²AP-HP Pitié-Salpêtrière, Département de Neuroradiologie, Paris, France

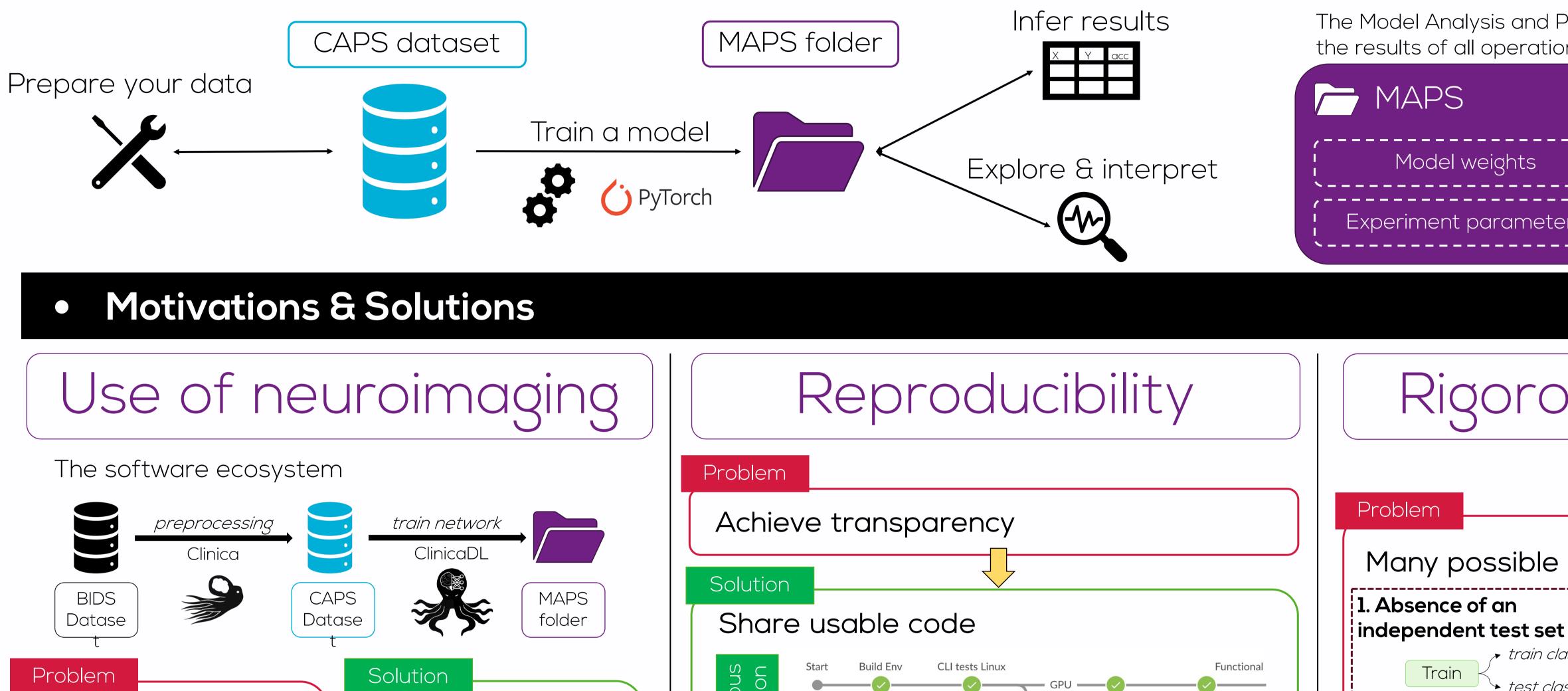
elina.thibeausutre@icm-institute.org

Deep learning has become one of the Moreover, the whole deep learning most used data analysis technique for community faces a reproducibility crisis medical image analysis. Unfortunately, that discrédits its results. Hence there is this recent massive use of deep learning an urgent need in publishing open-source has also been associated with software, data sets and scripts that allow methodological flaws in many studies reproducing the methodologies described which results are contaminated by data in deep learning studies. leakage.

@AramisLabParis

Finally, deep learning users who are not neuroimaging specialists have difficulty in accessing properly formatted and pre-processed data sets. This issue has been partly tackled by a data set format established by the community: the Brain Imaging Data Structure (BIDS).

Software overview



MAPS

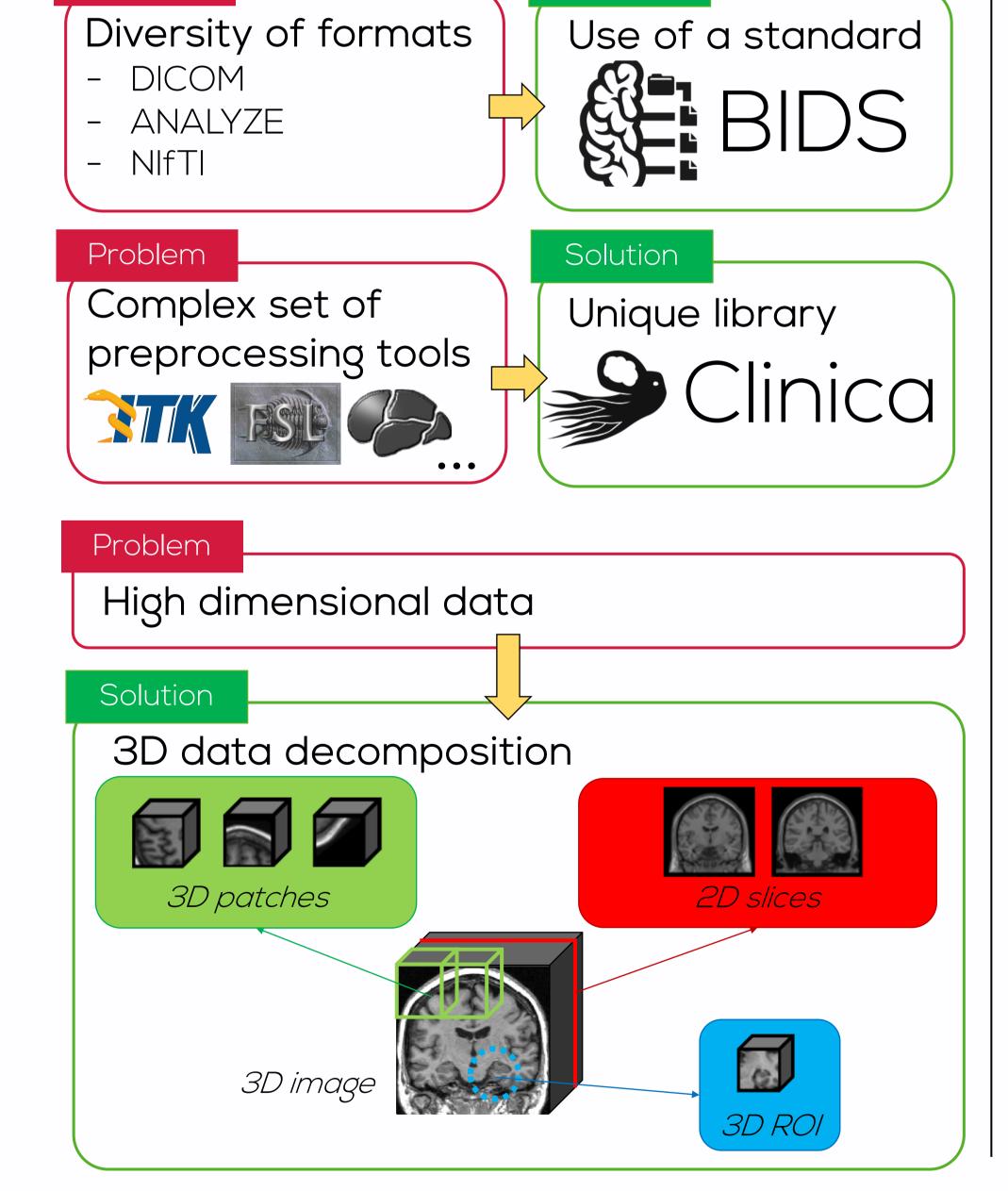
The Model Analysis and Processing Structure (MAPS) contains the results of all operations performed on a model.

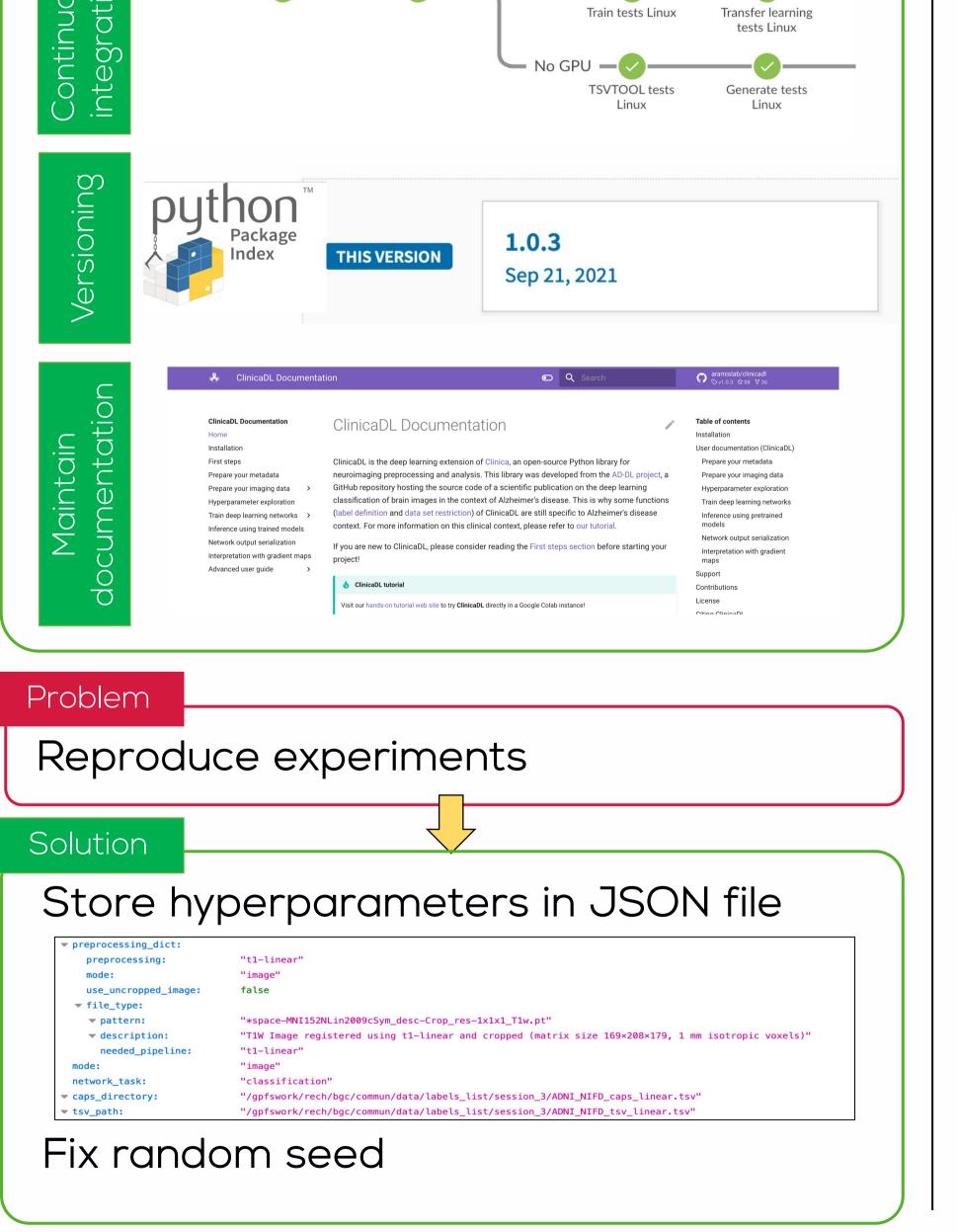
MAPS	Training logs
Model weights	Evaluation on test sets
Experiment parameters	Model interpretation

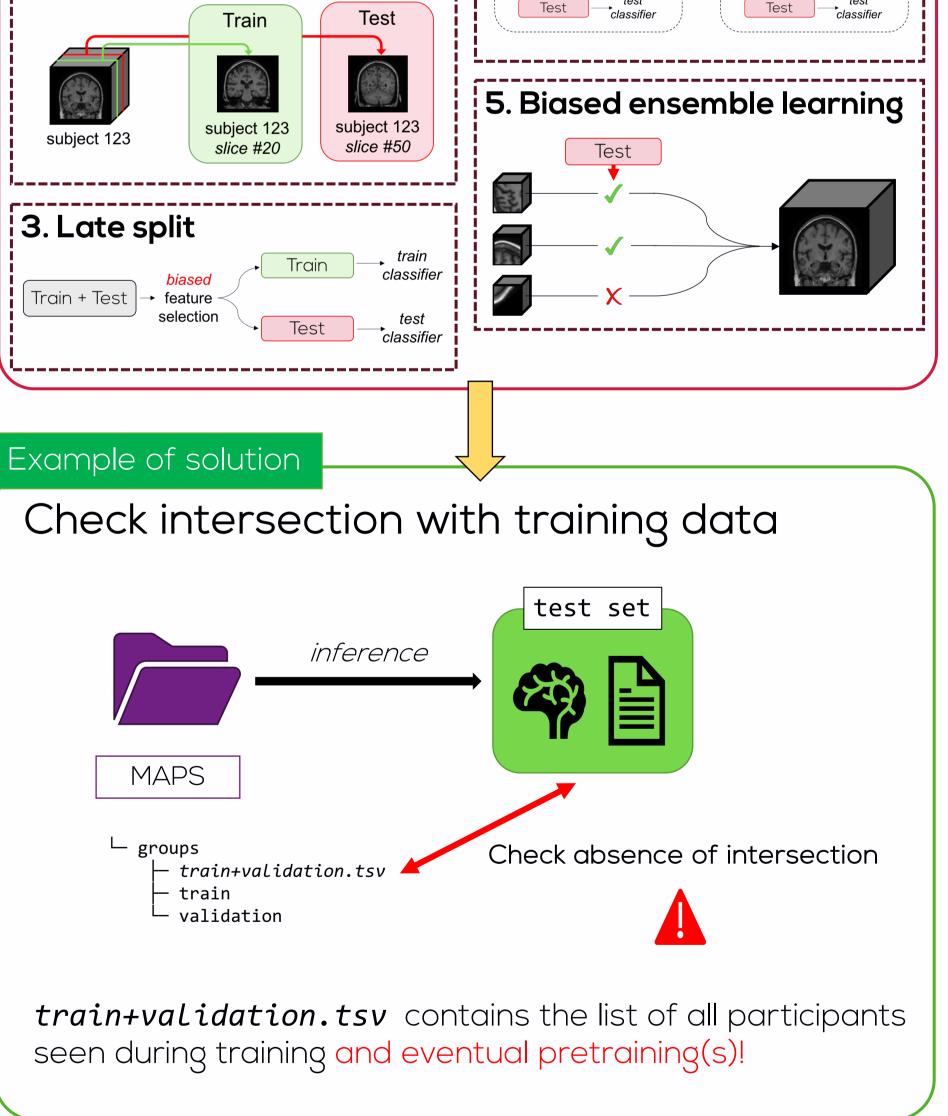
Rigorous validation

Many possible scenarios of data leakage 1. Absence of an









Conclusion

ClinicaDL is an open-source software for deep learning processing on neuroimaging data. With this software, we solve the three main issues encountered by deep learning users who are not specialist of the neuroimaging domain:

> (1) the data management and preprocessing of neuroimaging data sets, (2) the contamination of results by data leakage, (3) the lack of reproducibility of deep learning experiments.

Useful links



aramis-lab/clinicaDL



clinicadl.readthedocs.io

