ARAMIS LAB
BRAIN DATA SCIENCE

www.aramislab.fr
www.clinica.run

ninon.burgos@icm-institute.org

3e colloque sur l'imagerie médicale à l'heure de l'intelligence artificielle
14 octobre 2020

Computer-aided diagnosis from neuroimages:

A framework for objective & reproducible experiments

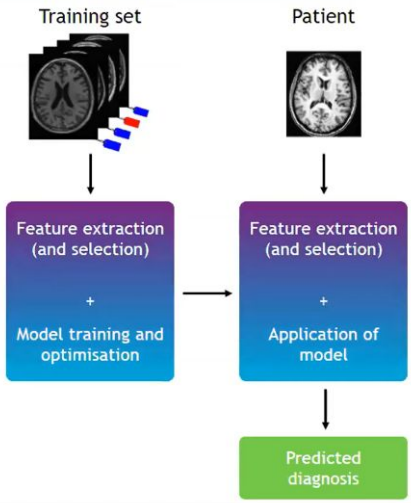
Ninon Burgos, CNRS Researcher
Aramis Lab, Paris Brain Institute, France

YouTube FR
Rechercher

Deep learning for computer-aided diagnosis

Basic elements of a deep learning classification pipeline

- Training data set
- Feature extraction from raw data and dimensionality reduction
- Model training and optimization
- Application to test data

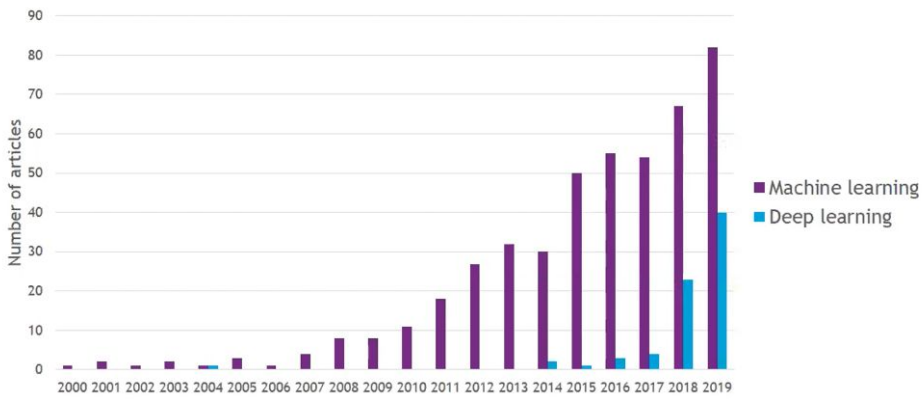


Ninon Burgos
425 vues • 4 déc. 2020
2 JE N'AIME PAS PARTAGER ENREGISTRER ...
3e colloque sur l'imagerie médicale à l'heure de
Institut du Cerveau - Paris Brain Institute - 1/11

YouTube FR
Rechercher

ML/DL for Alzheimer's diagnosis & prognosis

A very active field of research



Ninon Burgos
425 vues • 4 déc. 2020
2 JE N'AIME PAS PARTAGER ENREGISTRER ...
3e colloque sur l'imagerie médicale à l'heure de
Institut du Cerveau - Paris Brain Institute - 1/11

ML for Alzheimer's diagnosis & prognosis

A very active field of research, to which we contribute!

DL for Alzheimer's diagnosis & prognosis

Literature review

- Alzheimer classification from structural MRI using deep learning

Study	Performance					Approach	Data leakage	Number of citations
	AD vs CN	sMCI vs pMCI	MCI vs CN	AD vs MCI	Multi-class			
(Aderghal et al., 2017b)	ACC=0.84	--	ACC=0.65	ACC=0.67†	--	ROI-based	None detected	16
(Aderghal et al., 2018)	BA=0.90	--	BA=0.73	BA=0.83	--	ROI-based	None detected	9
(Bäckström et al., 2018) [4]	ACC=0.90	--	--	--	--	3D subject-level	None detected	20
(Cheng et al., 2017)	ACC=0.87	--	--	--	--	3D patch-level	None detected	12
(Cheng and Liu, 2017)	ACC=0.88	--	--	--	--	3D subject-level	None detected	8
(Islam and Zhang, 2018)								23
(Korolev et al., 2017)								72
(Li et al., 2017)								12
(Li et al., 2018)								7
(Lian et al., 2018)	ACC=0.90						None detected	30
(Mingxia Liu et al., 2018a)	ACC=0.91	ACC=0.78†	--	--	--	3D patch-level	None detected	59
(Mingxia Liu et al., 2018c)	ACC=0.91	--	--	--	--	3D patch-level	None detected	26
(Qiu et al., 2018)	--	ACC=0.82†	--	--	--	2D slice-level	None detected	8
(Senanayake et al., 2018)	ACC=0.76	--	ACC=0.75	ACC=0.76	--	3D subject-level	None detected	3
(Shmulev et al., 2018)	--	ACC=0.62	--	--	--	3D subject-level	None detected	5
(Vaillant and Soni, 2017)	ACC=0.81	--	--	--	ACC=0.87†	2D slice-level	None detected	8

Average accuracy (AD vs CN): 86.4%

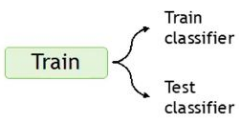
Study	Performance					Approach	Data leakage (type)	Number of citations
	AD vs CN	sMCI vs pMCI	MCI vs CN	AD vs MCI	Multi-class			
(Aderghal et al., 2017a)	ACC=0.91	--	ACC=0.66	ACC=0.70	--	ROI-based	Unclear (b,c)	13
(Basaia et al., 2019)	BA=0.99	BA=0.75	--	--	--	3D subject-level	Unclear (b)	25
(Hon and Khan, 2017)	ACC=0.96	--	--	--	--	2D slice-level	Unclear (a,c)	32
(Hosseini Asl et al., 2018)	ACC=0.99	--	ACC=0.94	ACC=1.00	ACC=0.95†	3D subject-level	Unclear (a)	107
(Islam and Zhang, 2017)	--	--				2D slice-level	Unclear (b,c)	23
(Lin et al., 2018)							Unclear (b)	22
(Manhua Liu et al., 2018)								39
(Taqi et al., 2018)								16
(Vu et al., 2017)								20
(S.-H. Wang et al., 2018)	ACC=0.98						Unclear (b)	49
(Bäckström et al., 2018) [4]	ACC=0.99	--	--	--	--	3D subject-level	Clear (a)	20
(Farooq et al., 2017)	--	--	--	--	ACC=0.99†	2D slice-level	Clear (a,c)	31
(Gunawardena et al., 2017)	--	--	--	--	ACC=0.96†	3D subject-level	Clear (a,b)	8
(Yu et al., 2018)	ACC=0.86	--	ACC=0.86	ACC=0.77	ACC=0.80†	3D subject-level	Clear (a,c)	8
(Wang et al., 2017)	--	--	ACC=0.91	--	--	2D slice-level	Clear (a,c)	11
(Wang et al., 2019)	ACC=0.99	--	ACC=0.93	ACC=0.94	ACC=0.97†	3D subject-level	Clear (b)	17
(Wu et al., 2018)	--	--	--	--	0.95††	2D slice-level	Clear (a,b)	7

Average accuracy (AD vs CN): 93.8%

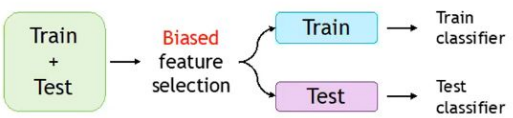
Wen, Thibeau-Sutre et al., Medical Image Analysis, 2020

Main causes of data leakage in DL scenarios

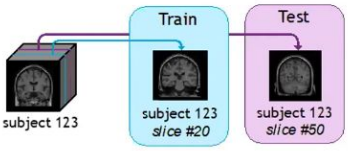
- No independent test set



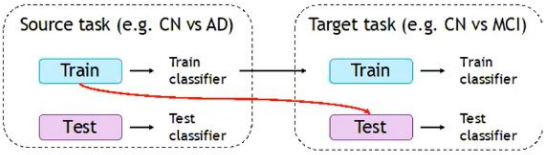
- Late split



- Biased within-subject split



- Biased transfer learning



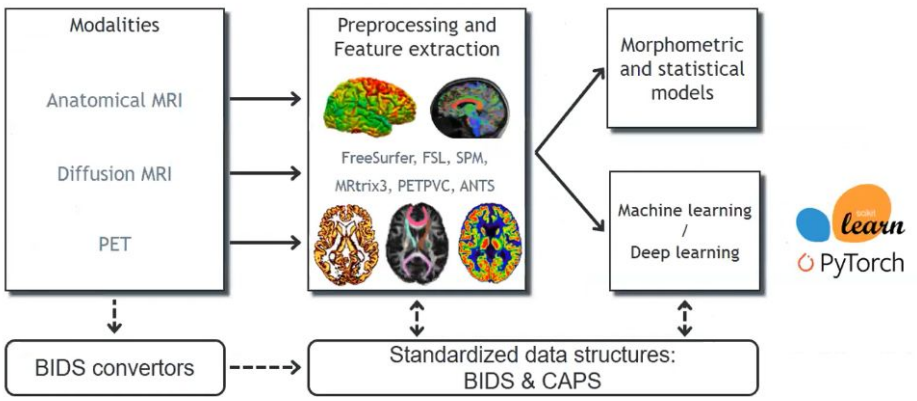
Wen, Thibeau-Sutre et al., Medical Image Analysis, 2020

DL for Alzheimer's diagnosis & prognosis

- **Assessment**: flaws in the validation and lack of details
  - Impossible to compare approaches and draw conclusions
- **Objective**: identify which factors (preprocessing, network architecture, etc.) have an influence on the classification accuracy
- **Solution proposed**: framework to rigorously compare new approaches
  - ClinicaDL
  - Experiments

Clinica

Software platform for clinical neuroimaging studies



ClinicaDL

Framework for the objective & reproducible classification of Alzheimer's disease using deep learning

Evaluation :

- **More reproducible**
  - Storing of data using community standards
  - Fully automatic data manipulation
  - Code sharing
- **More objective**
  - Baseline approaches against which new methods can easily be compared
  - Rigorous validation

Samper-González et al., NeuroImage, 2018; Wen, Thibeau-Sutre et al., Medical Image Analysis, 2020

https://www.youtube.com/watch?v=3P6xpAt6E2Y&list=PLpupy7iBJC30PoGe9jCJozw1DvuH3sWDB&index=1