

Normalize MRI for robust radiomics in oncology using Deep Learning

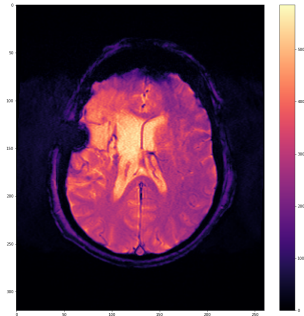
Sacha Dedeken

supervised by John Klein, Jérémie Boulanger from CRISAL and David Pasquier,
Alexandre Escande from Oscar Lambret

March 8, 2022

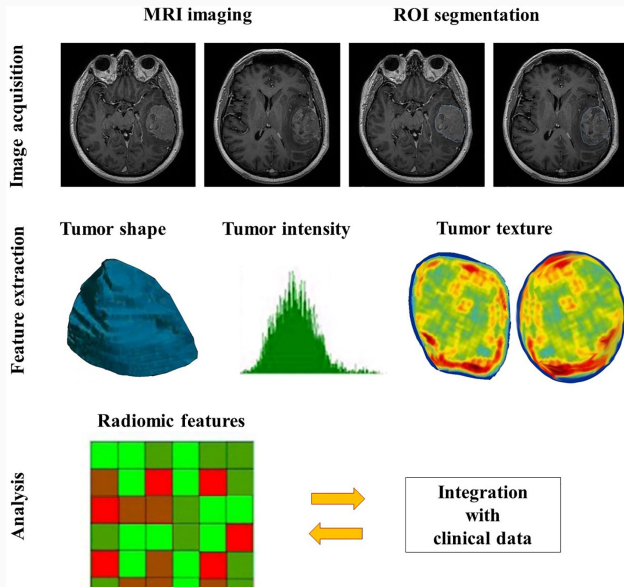
Master 2 Data Science, University of Lille / Centrale Lille / IMT Nord-Europe

Human workflow of MRI



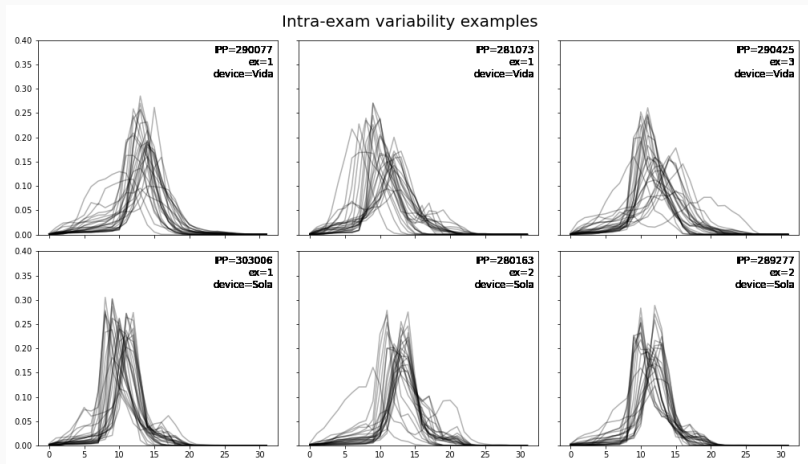
Common examination for diagnosis, prognosis and treatment decision for cancer

The radiomics revolution



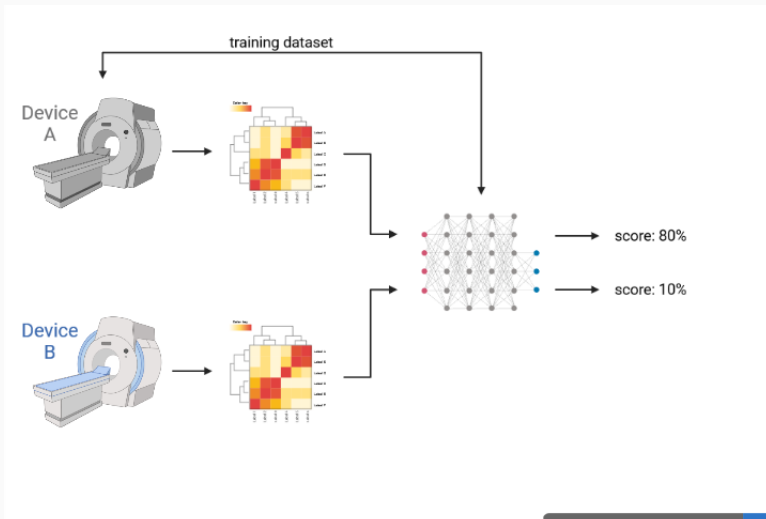
The robustness issue

Three sources of variability : Electromagnetic field, Intra exam variability and Inter exam variability



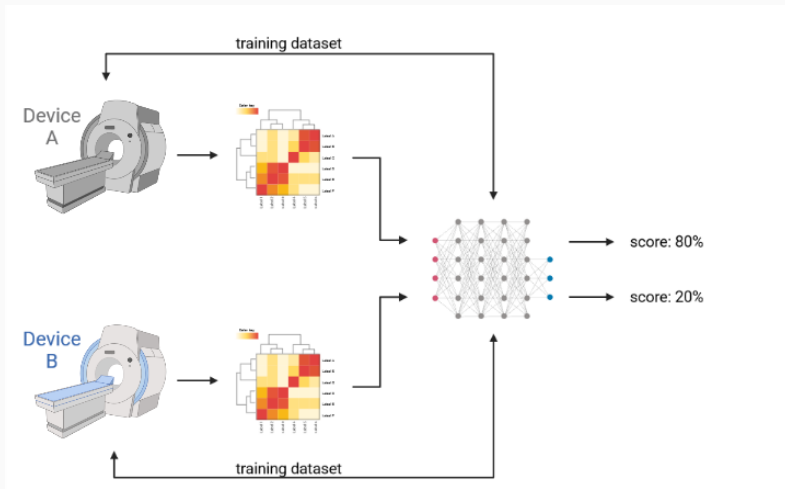
Objectives of the project

How to reduce inter-device variability?



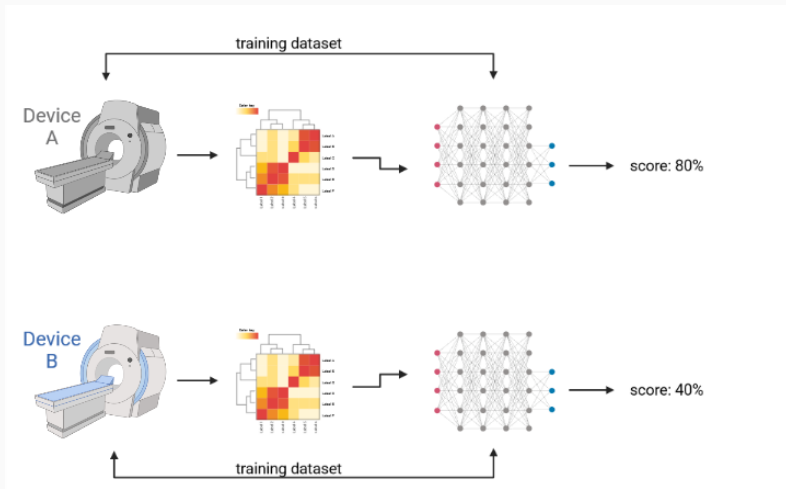
Objectives of the project

False good idea #1 : Just merge the two dataset!



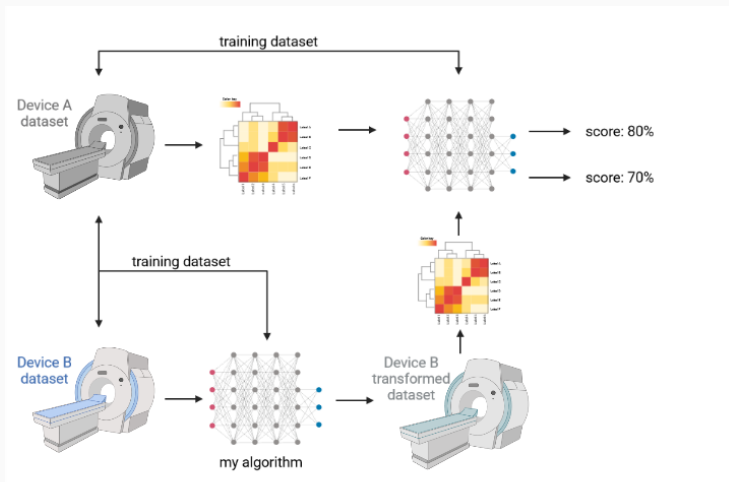
Objectives of the project

False good idea #2 : Just train another algorithm!



Objectives of the project

New approach: Train deep learning to do histogram transport to harmonize the whole dataset



The most popular approach: ComBat

- Adapted from Johnson et al. (2007) in genomics
- This method works directly on radiomics that are described as:

$$y_{ij} = \alpha + X_{ij}\beta + \gamma_i + \delta_i\epsilon_{ij}$$

- It creates corrected values using maximum likelihood approach:

$$y_{ij}^{\text{ComBat}} = \frac{y_{ij} - \hat{\alpha} - X_{ij}\hat{\beta} - \hat{\gamma}_i}{\hat{\delta}_i} + \hat{\alpha}$$

- Already validated for PET, CT and MRI with limited amelioration

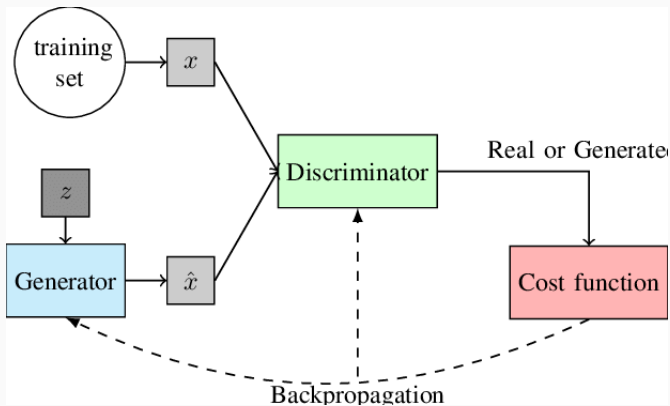
The good pre-processing (technical details)

- Create a mask to consider only brain voxels
- Extract histogram from Region Of Interest (ROI)
- Standardization (- mean and / by standard deviation)
- Normalization over 32 values + discretization

ID	exam	slice	device	h0	...	h31
291755	2	18	1	0.02	...	0.09

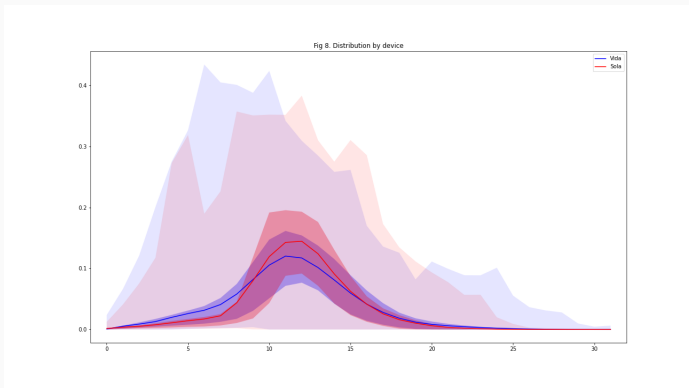
How to use GAN for histogram transport?

Generative Adversarial Network are perfect for this type of task!



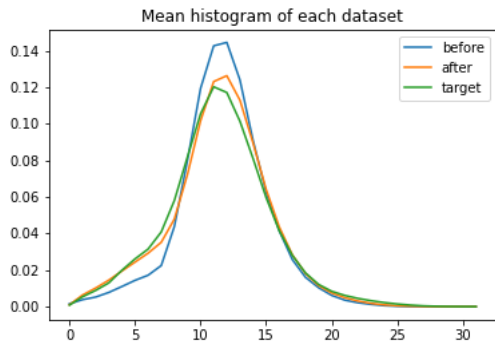
How to evaluate GAN performances

Visual assessment comparing to the whole distribution is very limited: it's hard to look after a correct distribution movement!



How to evaluate GAN performances

Moments convergence (mean and diagonal of covariance matrix)



$$score = \frac{\chi^2(h_s, h_t) - \chi^2(h_g, h_t)}{\chi^2(h_s, h_t)}$$

- +60% improvement on mean test score, but...
- not sensitive to individual transformations
- more a sanity check than a satisfying metric

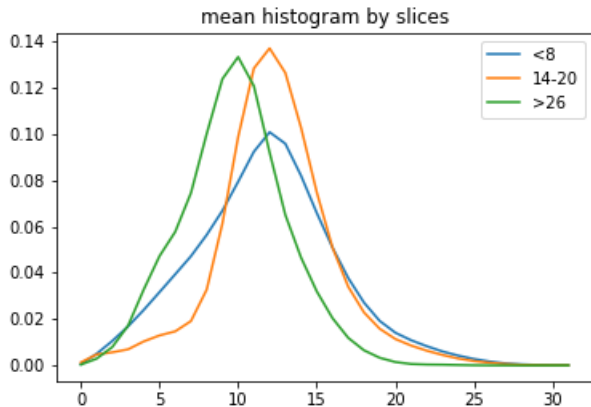
We can do better: an 8-steps experimental protocol

1. **Dataset analysis** on size, modality, body region, prediction possibility
2. **Pre-processing**
3. **Train-test split** both devices
4. **Algorithm training**
Gen: $Vida(Vida_{train} + Vida_{test})$
Dis: $Skyra_{train}$
5. **Histogram generation**
Gen: $Vida_{train}, Vida_{test}$
 $\implies Vida_{train}^G, Vida_{test}^G$
6. **Radiomics creation**
7. **Radiomics-based predictor** evaluated on 6 conditions
8. **Results analysis** with comparison to ComBat

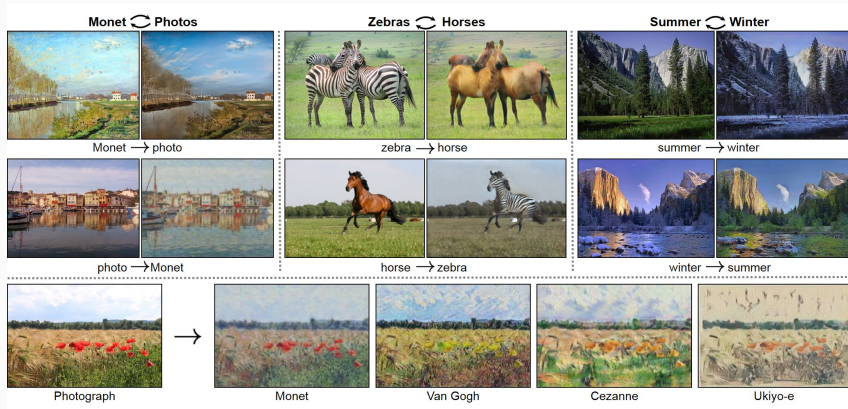
Personal view

- Very happy to have pursued this project in M2
- Finding correct medical data and clean them takes time
- This literature review, experimental protocol and python toolbox pave the road to future projects
- Now fully convinced that I want to work in data science and machine learning applied to the medical field
- Found a directly linked internship in the Radiomics Research Team of SOPHiA Genetics near Bordeaux
- I consider doing a thesis in deep learning for radiomics

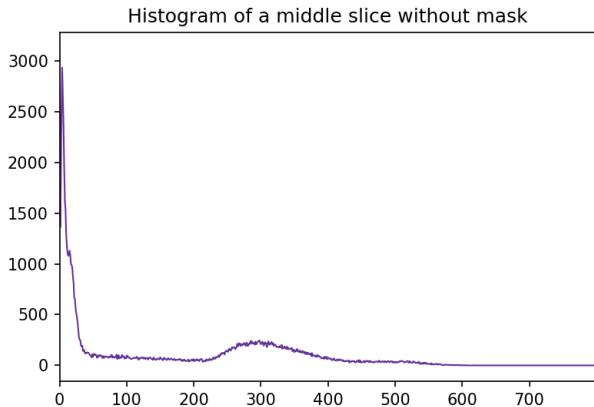
Bonus #1: ConditionalGAN



Bonus #2: CycleGAN



Bonus #3: Brain Extraction Tool



Bonus #4: An excellent paper very similar to our approach

comBat versus cycleGAN for multi-center MR images harmonization

Stenzel Cackowski¹

STENZEL.CACKOWSKI@UNIV-GRENOBLE-ALPES.FR

Emmanuel L. Barbier¹

EMMANUEL.BARBIER@UNIV-GRENOBLE-ALPES.FR

Michel Dojat¹

MICHEL.DOJAT@UNIV-GRENOBLE-ALPES.FR

Thomas Christen¹

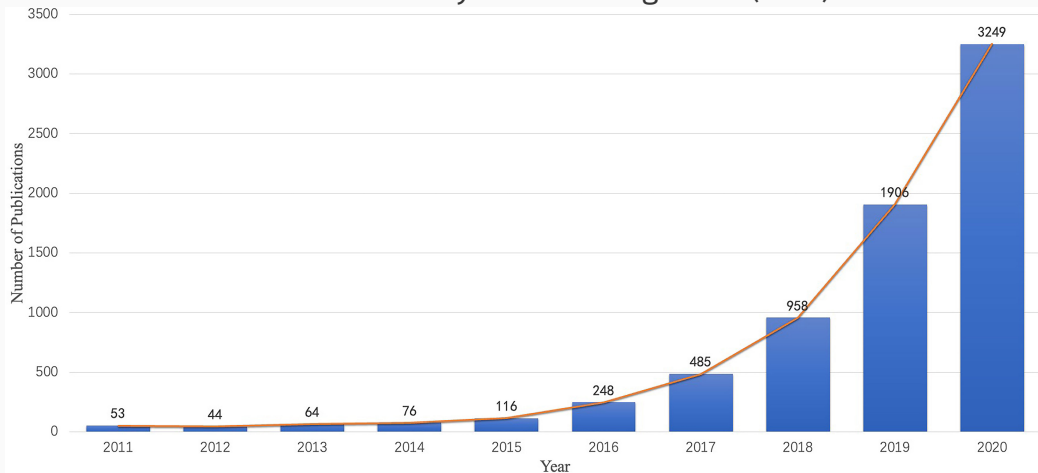
THOMAS.CHRISTEN@UNIV-GRENOBLE-ALPES.FR

¹ *Université Grenoble Alpes, Inserm U1216, Grenoble Institut for Neurosciences*

Editors: Under Review for MIDL 2021

Bonus #5: Radiomics between revolution and buzz word

Bibliometric analysis from Ding et al. (2021):



Bonus #6: Evaluation on 6 subgroups

Train your supervised predictor on $Skyra_{train}$ then evaluate it on 6 conditions to answer the following question "Does the classifier obtain a good score..." :

- $Skyra_{train}$: ...on its training data?
- $Skyra_{test}$: ...on new data from the same device?
- $Vida_{train} + Vida_{test}$: ...on new data from another device without harmonization?
- $Vida_{train}^G$: ...on new but harmonized data, coming from the GAN training set?
- $Vida_{test}^G$: ...on new but harmonized data, unseen by the GAN?
- $Vida_{combat}$: ...on new but harmonized data via the ComBat method?