Artificial
Intelligence in
Medicine

Collection of the dataset for Radiomic and Dosiomic analysis in RT

Part 1

03/02/2020 - Stefano Piffer

Purpose

Apply a retrospective exploratory MR-CT-based radiomics and dosiomic analysis based on ML, to investigate imaging biomarkers of clinical outcomes in paediatric patients affected by medulloblastoma.

Features from MR-CT scans and dose distribution (both PTV and OAR) will be associated with overall survival, recurrence-free survival, and loco-regional recurrence-free survival after IMRT.

1° year milestone → complete the database

To date almost completed

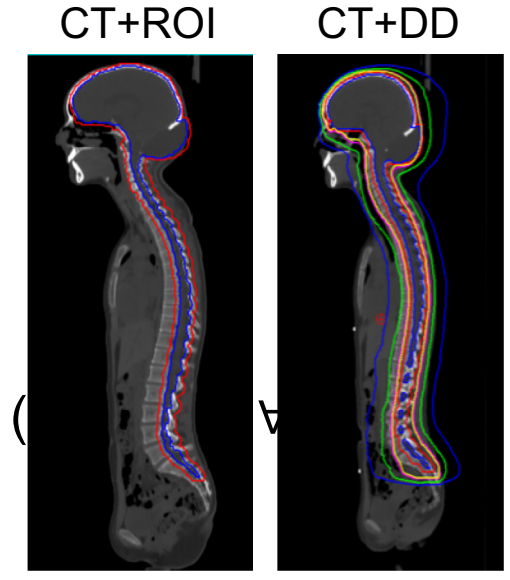
Treatment scheme for medulloblastoma

Multimodality approach:

Surgery → Histology (medulloblastoma)

Chemotherapy → Protocols & Drugs

Radiotherapy → CT, ROI,
Dose Distributions (DD)



Hospital centers involved for database build

AOU Meyer for clinical data and MRI
(1 MR scanner)

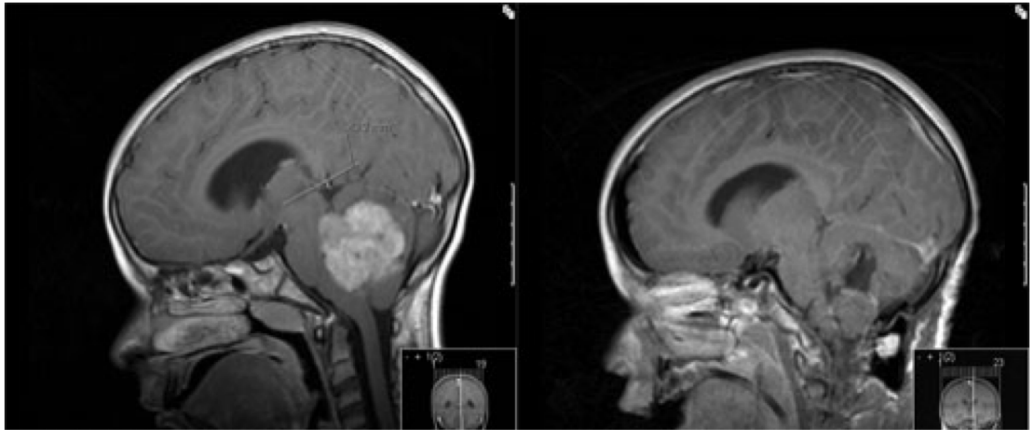
AOU Careggi for radiotherapy
(1 CT scanner & 1 machine for RT)

Difficulties encountered:

- Patient from 2011 to 2018
- Dearchiving and Anonimizing is a slow and meticulous process
- Several archives to access (1 for CT, 1 for Radiotherapy, 1 for MRI, 1 for Careggi clinical data, 1 for Meyer clinical data)
- Smooth all these data to make the database homogeneous
- Interconnect different professional figures (Radiologist, Radiotherapist, Medical Physicist, Physicist)

Magnetic Resonance Images

- MRI pre - and post - surgery and every three months until the end of the multimodal treatment.



- Time lapse: form 8 months to 8 years
- For each resonance exam at least 3 sequences that contain information: T1 - mdc, T2, FLAIR (Fluid Attenuated Inversion Recovery)
  (~30 GB ∀ patient)

Study Population

- Up to now 50 patients

- For each of them:
  - Imaging data (MRI & CT)
  - Radiotherapy data (dose distribution; prescription dose; fractionations; boost; machine parameters; structures)
  - Clinical data (histology; risk class; drugs; motor, cognitive and sensory deficits; possible radio-induced toxicity; metastasis; relapse; end state)

- A lot of data to interpret and analyze, to handle and storage (overall more than 1.5 TB)

Preliminary test

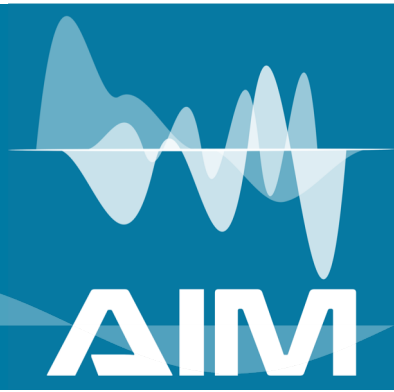
- Thanks to a preliminary test on a subset of patients (presented later), the features “Radiological evidence of neuro damage” was not so robust, it can be misinterpreted.

- We posed the problem to the radiologists of Meyer who is:
  - reviewing each patient's medical history
  - redefining clinical outputs
  - redefining the volume where to calculate the features (more local than the whole cranio-spinal axis)

Noteworthy

- At the moment the database is of modest size but it is possible to expand it with the data of patients from other hospitals.
- There are open-source online databases even more numerous than this but they are far more heterogeneous and incomplete.
- Only one MR, only one CT, only one RT machine: added value for the calculation of radiomic and dosiomic features extracted from the images (features are affected by the inter-scanner variability).
- Oral presentation at ViMaBi2019 (Salerno, Italy) - “Workshop on Visual Computing and Machine Learning for Biomedical Applications”.
- Oral presentation at RAD2020 (Herceg Novi, Montenegro) - “Eighth International Conference on Radiation in Various Fields of Research”
- Publisced on “Communications in Computer and Information Science” https://doi.org/10.1007/978-3-030-29930-9_6

thank you



Artificial
Intelligence in
Medicine

Challenges in ML analysis implementation on RT data

Part 2

03/02/2020 - Leonardo Ubaldi

Tools

- Python:
  - SimpleITK
  - Pandas
  - Matplotlib
  - Scikit-learn
  - PyRadiomics

} Analysis

- Advanced Normalization Tools (ANTs) ⟹ Registration (CT on ADC)

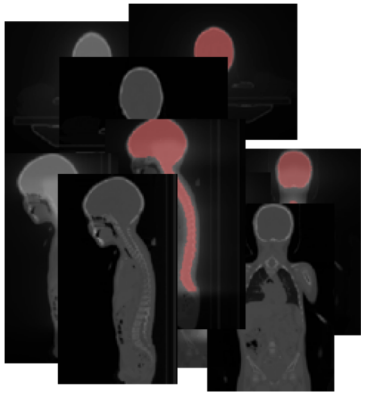
Pyradiomics: radiomic features

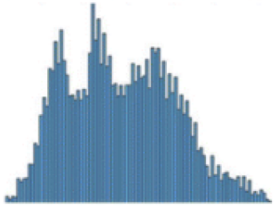
- First Order Statistics
- Shape based
- Gray Level Co-occurrence Matrix Features
- Gray Level Run Length Matrix Features
- Gray Level Size Zone Matrix Features
- Neighbouring Gray Tone Difference Matrix Features
- Gray Level Dependence Matrix Features

94 features

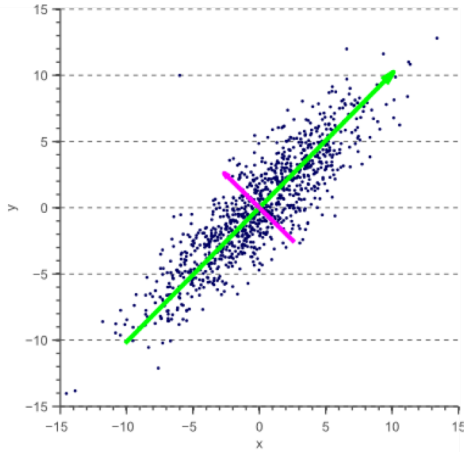
Workflow

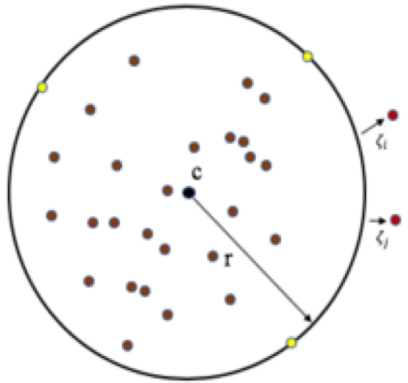
Database (CT, Dose, Structures) → Feature extraction → Feature selection → Classification





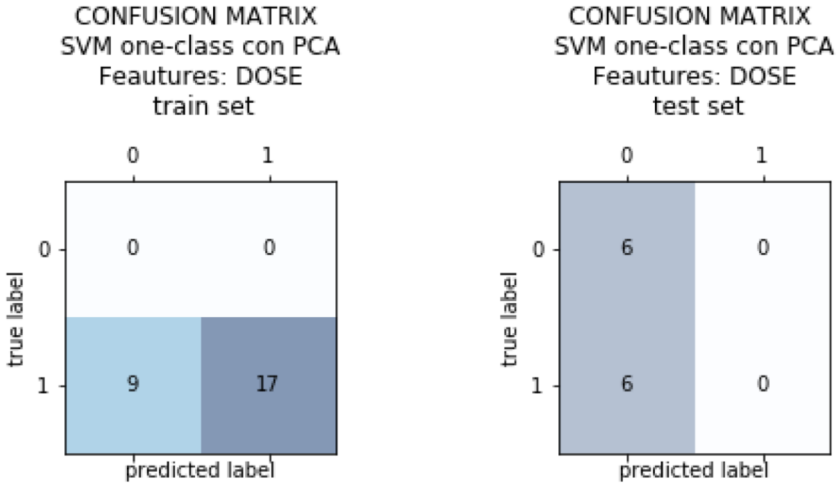
pyradiomics
python
+
RADIOMICS





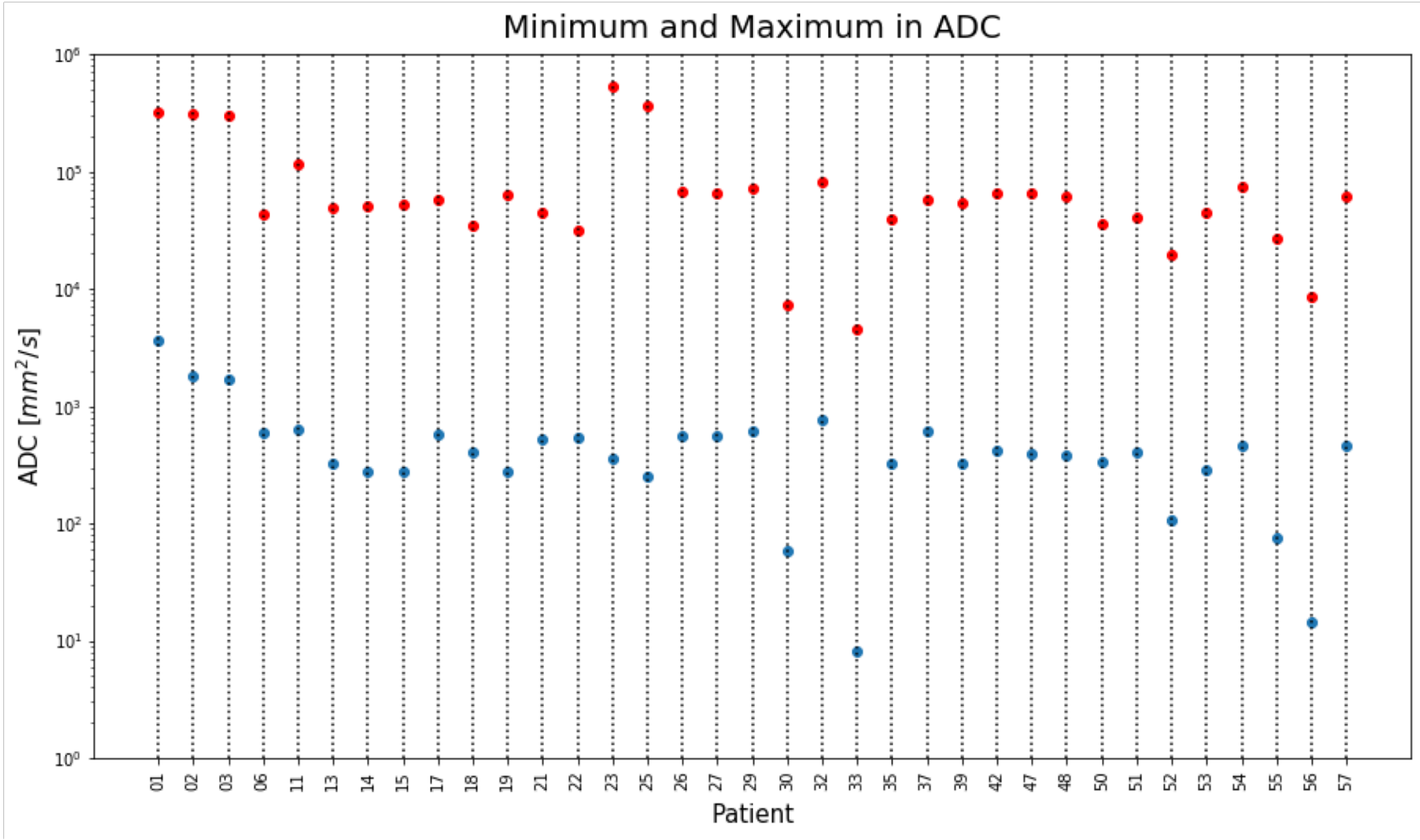
Preliminary results from dose distribution

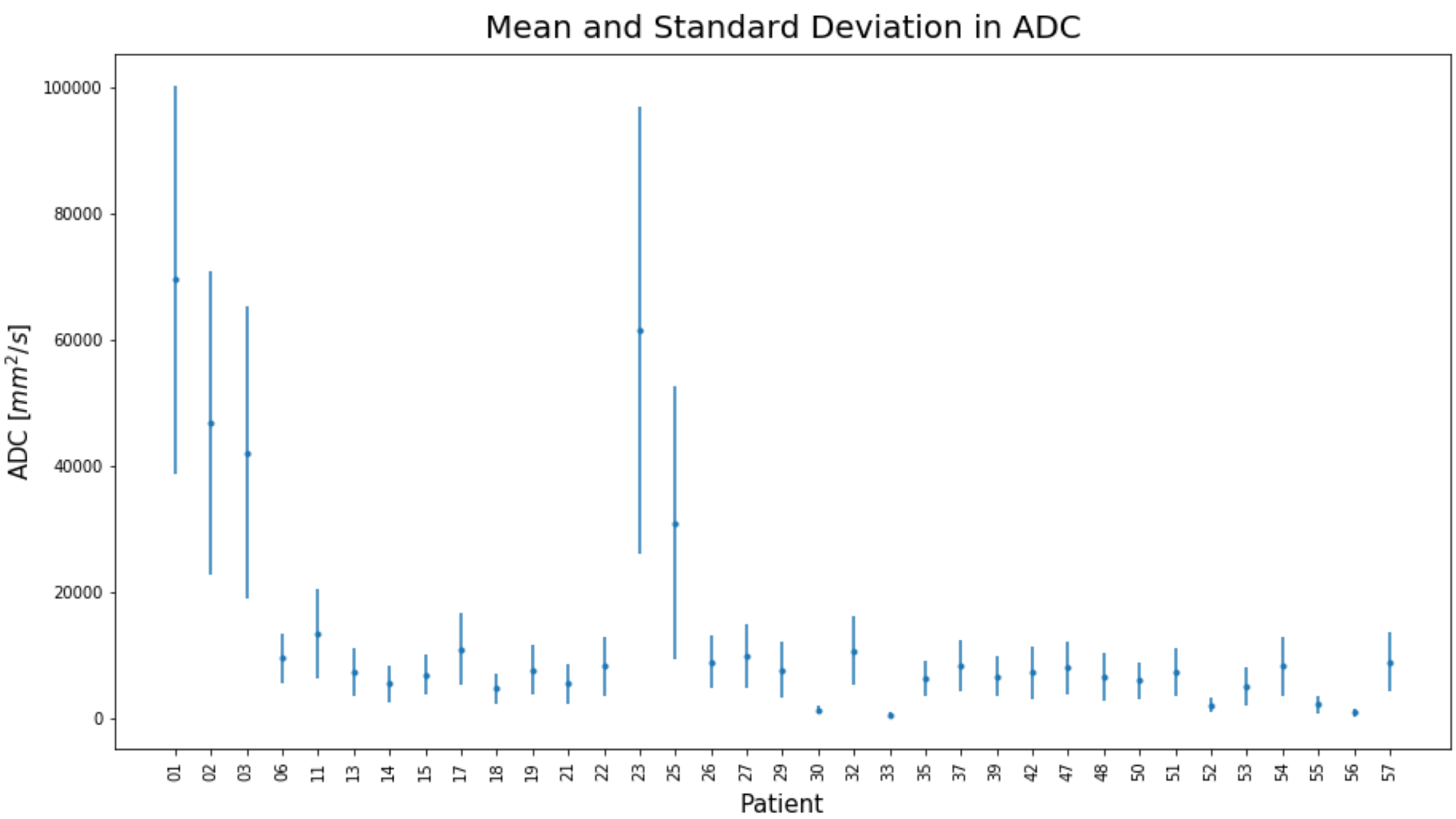
- ‘Radiological evidence of neuro damage’ → 38 subject
  - Outcome 0 → 6 subject
  - Outcome 1 → 32 subject
- Confusion matrix obtained with One-Class SVM using features extracted from dose distribution

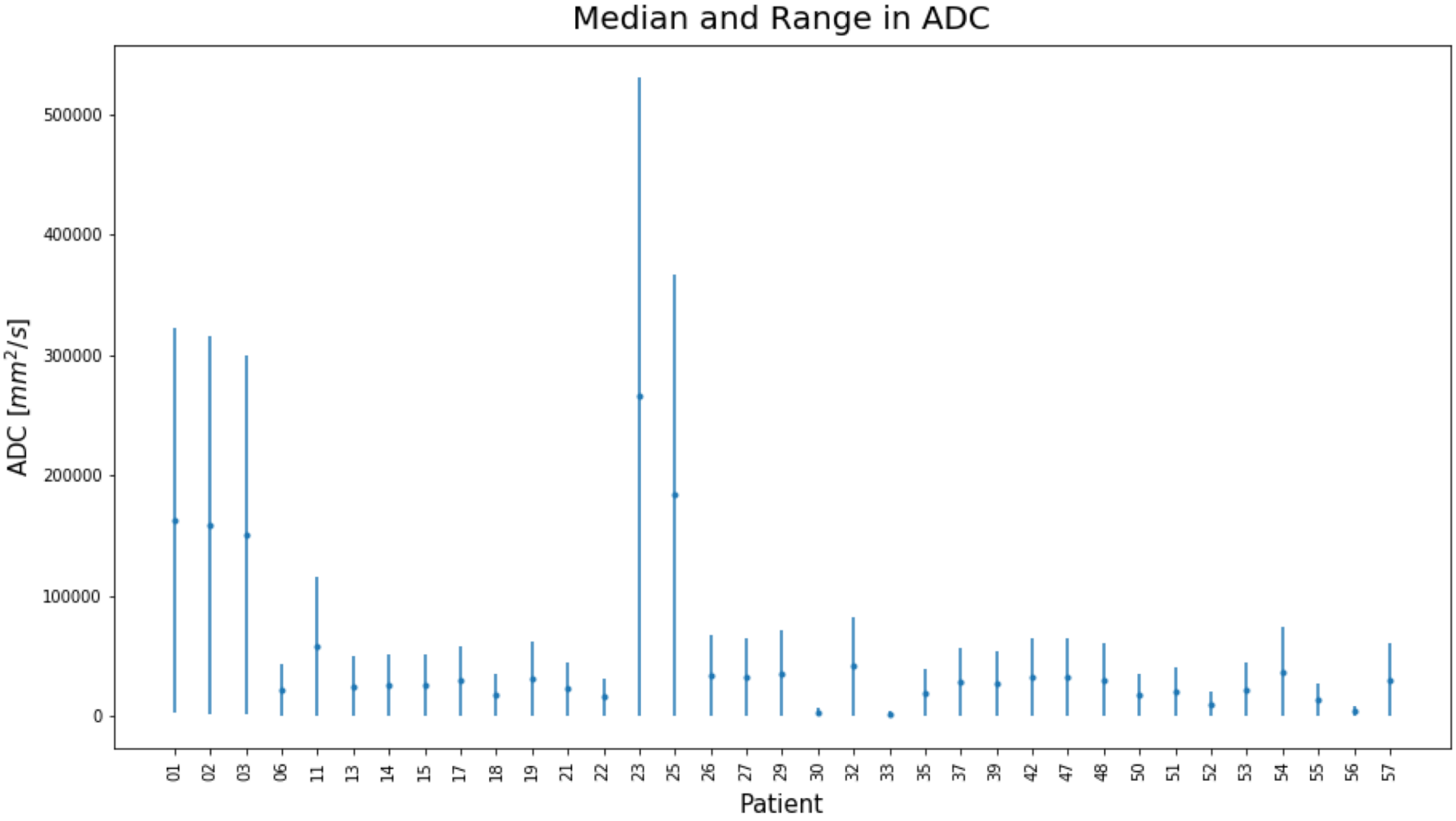


Registration and features extraction from ADC

- ADC :
  - Apparent diffusion coefficient is a measure of the magnitude of diffusion (of water molecules) within tissue, and is commonly clinically calculated using MRI with diffusion-weighted imaging (DWI).
  - 35 subjects
- Registration:
  - CT on ADC
  - Registration with rigid transformation with ANTs
- Features extraction:
  - From ADC using ROI brain
  - Pyradiomics


Minimum and Maximum in ADC
ADC [mm^2/s]
10^0
10^1
10^2
10^3
10^4
10^5
10^6
01 02 03 06 11 13 14 15 17 18 19 21 22 23 25 26 27 29 30 32 33 35 37 39 42 47 48 50 51 52 53 54 55 56 57
Patient

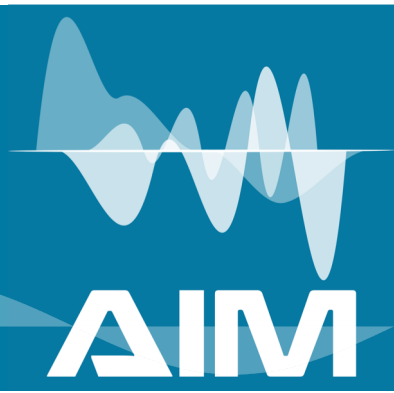

Mean and Standard Deviation in ADC
ADC [mm^2/s]
100000
80000
60000
40000
20000
0
01
02
03
06
11
13
14
15
17
18
19
21
22
23
25
26
27
29
30
32
33
35
37
39
42
47
48
50
51
52
53
54
55
56
57
Patient


Median and Range in ADC
ADC [mm^2/s]
0
100000
200000
300000
400000
500000
01
02
03
06
11
13
14
15
17
18
19
21
22
23
25
26
27
29
30
32
33
35
37
39
42
47
48
50
51
52
53
54
55
56
57
Patient

Problems

- Small database: 38 subjects
  - 32 subjects → outcome 1
  - 6 subjects → outcome 0
- Different protocol in MRI
- Clinical outputs

thank you



First Order Statistics

- Energy: sommatoria dei quadrati dei valori dei voxel
- Entropy: misura la quantità media di informazioni necessaria per codificare i valori dell'immagine
- Skewness: misura l'asimmetria di una distribuzione rispetto al valore medio
- Kurtosis: misura quanto una distribuzione è piccata rispetto alla distribuzione normale
- Uniformity: misura di omogeneità

Gray Level Co-occurrence Matrix

- Viene costruita la GLCM
  - E' una matrice che descrive le frequenze con cui ogni livello di grigio appare accanto ad altri livelli di grigio ad una distanza definita (offset)
- Vengono estratte features dalla matrice GLCM

Gray Level Size Zone Matrix

- Viene costruita la matrice GLSZM
  - Descrive il numero di voxel adiacenti che contengono lo stesso valore di intensità di grigio, nelle tre dimensioni
- Vengono estratte features dalla matrice GLSZM

Gray Level Run Length Matrix

- Viene costruita la matrice GLRLM
  - descrive con quale frequenza nell'immagine è possibile osservare file di elementi con una certa lunghezza e con una certa intensità
- Vengono estratte features dalla matrice GLRLM

Neighbouring Gray Tone Difference in Matrix

- Viene costruita la matrice NGTDM
  - Descrive le differenze esistenti fra ogni elemento dell'immagine e quelli immediatamente confinanti
- Vengono estratte features dalla matrice NGTDM

Recap database

- Distribuzione di dose → 57 soggetti
- Strutture → 55 soggetti
- MRI post-operatorie e pre-trattamento → 50 soggetti
  - T1 (3D) → ~ metà con MDC
  - DWI (ADC) → 50 soggetti
- 'Radiological evidence of neuro damage' → 40 soggetti
  - Outcome 0 → 7 soggetti
  - Outcome 1 → 33 soggetti
- Massima intersezione utilizzabile (utilizzando le ADC) → 35 soggetti
  - Outcome 0 → 6 soggetti
  - Outcome 1 → 29 soggetti