Radiomics and ML-segmentation on Facio-Scapulo-Humeral dystrophy (FSHD) and lung tumor

Francesca Brero
Raffaella Cabini

next_AIM kickoff meeting
WP3: Applications to real data
18/02/2022

TABLE OF CONTENTS

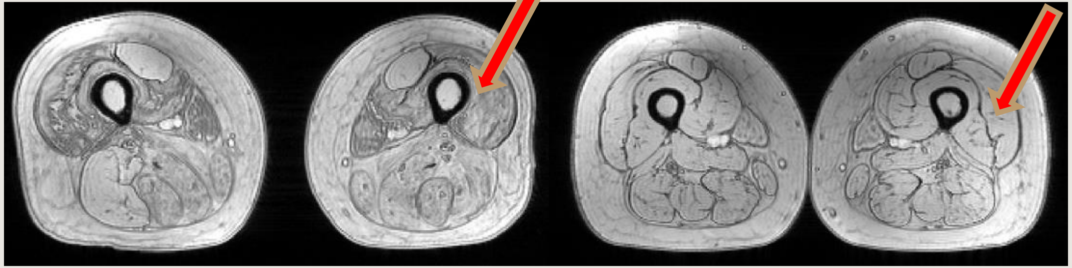
01

Facio-Scapulo-Humeral dystrophy (FSHD)

FSHD disorder: disease progression assessment

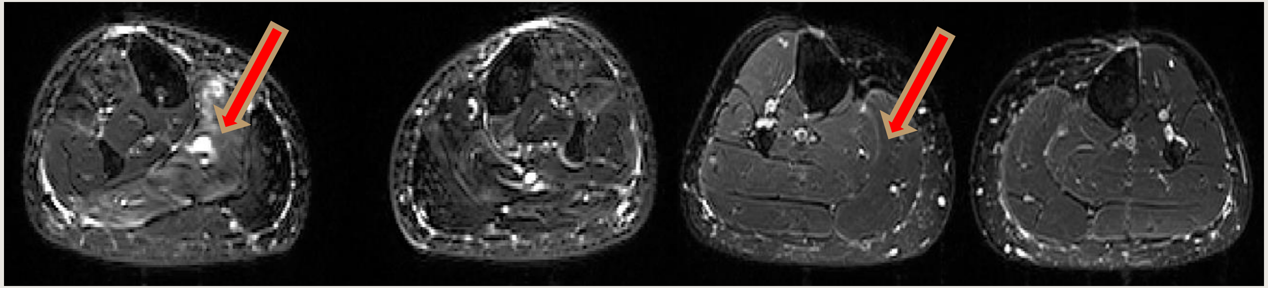
Quantitative Radiological Biomarkers

patient healthy control



- Fat Fraction (FF) by Fatty Riot algorithm (1)

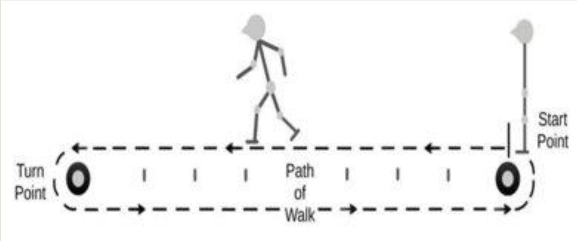
patient healthy control



- Edema (wT2) by EPG signal simulation (2)

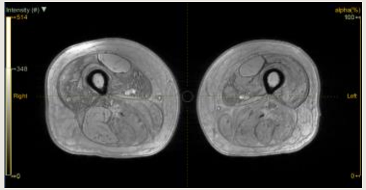
Clinical tests

- Clinical Severity Score (CSS)
- Six Minutes Walking Test (6MWT)

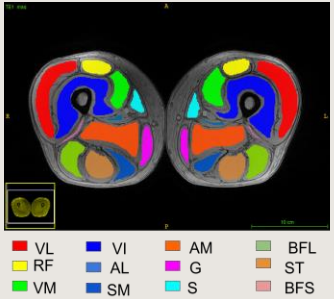


(1) E.B. Welch et al. Proc. 21st Sci. Meet. Int. Soc. Magn. Reson. Med. (2013) - (2) doi.org/10.1002/nbm.3459 & doi.org/10.1002/jmri.24619

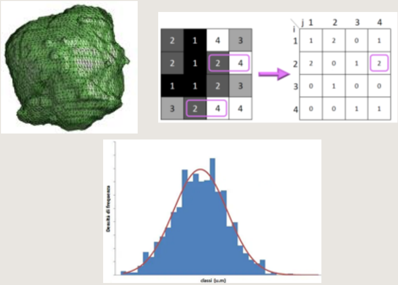
Idea



1. MRI acquisition of patients' thighs

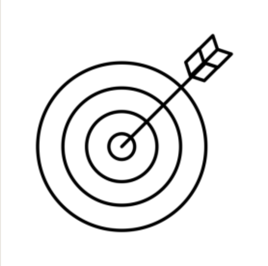


2. Pre-processing and segmentation



3. Radiomics

4. ML algorithms



5. Quantitative Fat Fraction and Water-T2 prediction

(Small) Dataset

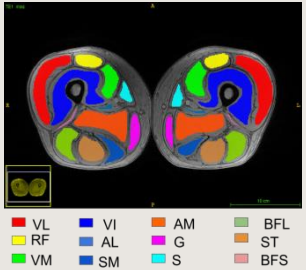
24 patients @ 3T MRI scanner
(Istituto Neurologico Nazionale IRCCS, Pavia)

- 3D 6-point multi-echo gradient echo (MEGE)
- multi-echo spin echo with 17 echoes (MESE)

Settings	MEGE	MESE
Number of slices	52	7
Slice Thickness (mm)	5	10
Distance Factor (%)	20	300
Resolution	1x1x5 mm^3	1.2x1.2x10 mm^3
Repetition Time TR (ms)	3.5	4100
Echo Time TE (ms)	1.7 - 9.2	10.9 - 185.3
Scan Time (min)	15	5.13

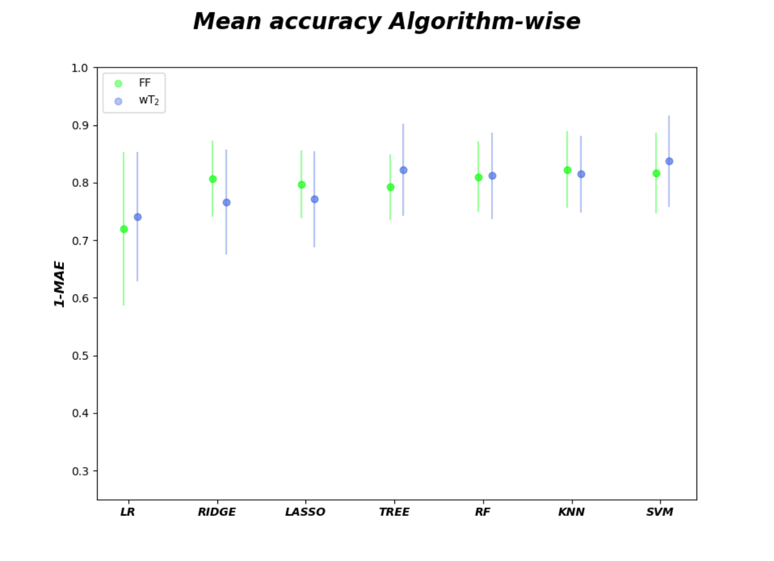
Features extraction and pre-processing

- 12 ROIs for thigh muscles
- Pre-processing (denoise of the images by a Gaussian Filter, N4 Bias Field Correction and Histogram Matching)
- Texture analysis was performed on MEGE images (LIFEx) - 42 radiomic features were obtained with 2D extraction from each region of interest



Muscle-wise Fat Fraction (FF) and Water-T2 (wT2) prediction

- Supervised regression problem
- Fatty Riot algorithm (MEGE sequence) and by EPG signal simulation (MESE sequence): ML **ground truth for FF and wT2**
- Parametric linear, ridge and lasso regression and the non-parametric KNN, SVM, tree, and RF algorithms.
- **FF mean algorithms accuracy floats around 80%** and alike considerations applied to the mean wT2 accuracy.
- **Lower prediction performance** both in FF and wT2 **for BFS and AL muscles** compared to the other muscles. This could be traced back to quick BFS and AL muscle degradation in neuromuscular disease that result in segmentation struggling → possible alteration of the texture analysis



Mean Absolute Error

$$MAE_j = \frac{\sum_{i=1}^{N} |y_i - \bar{y}_i|}{N}$$

j runs over the different thigh muscles and *i* over the observations associated to each muscle

Which is our next AIM ?

Removal of the necessity of texture analysis -> deep learning methods

02

Harmonization of features Extraction on stage III unresectable NSCLC

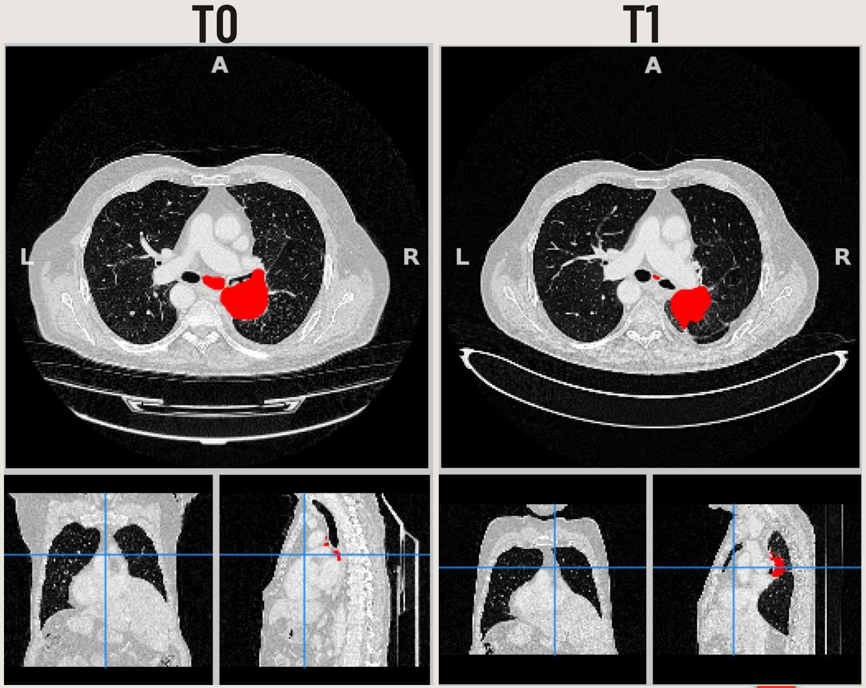
(Small) dataset

23 patients with inoperable stage III lung adenocarcinoma undergoing radio-chemo-immuno therapy.

2 CT images acquired at 2 time points:

- Baseline: at the diagnosis
- After the chemo-radio-immuno therapy

Homogeneus dataset of clinical features



Aim

Predict:

- Probability of Metastasis
- Probability of Relapses
- Response to the therapy

(Small) dataset

Images from 10 different centers:

Image Parameters:

Pixel
From (0.62 x 0.62) mm^2 to (0.98 x 0.98) mm^2

Slice thickness
From 0.3 mm to 3.0 mm

Reconstruction Parameter: Convolutional Kernel
11 types

Acquisition Parameters:

Scanner
4 different vendors

Current
From 56 mA to 581 mA

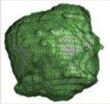
Contrast Agent
2 types

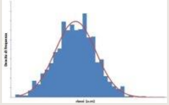
kV-peak
100 kVp, 120 kVp, 130 kVp, 140 kVp

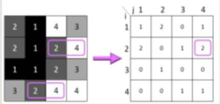
Exposure Time
From 350 s to 1000 s

Features Extraction

42 features computed by Lifex (IBSI standard compliant), including six categories of features:







4 Size- and Shape-based features

6 First-order Statistics features

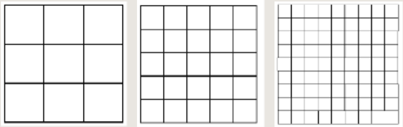
32 Higher order Statistics features:
7 grey-level co-occurency matrix features, 11 grey-level run length matrix features, 3 neighboring grey-level difference matrix features, 11 grey-evel zone length matrix features

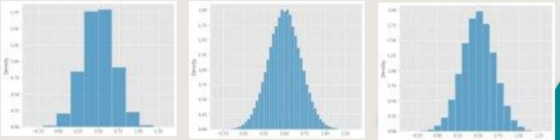
LIFEx soft
Local Image Feature Extraction

Images Parameters Harmonization

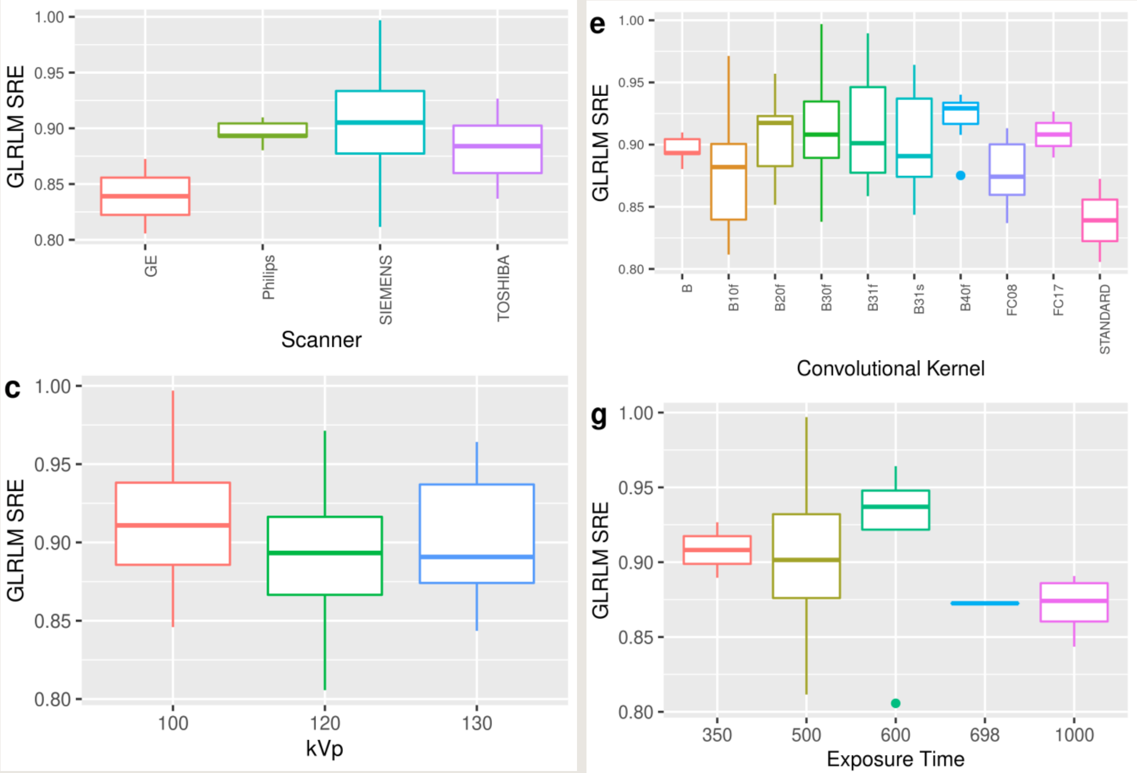
To harmonize images (before the extraction):

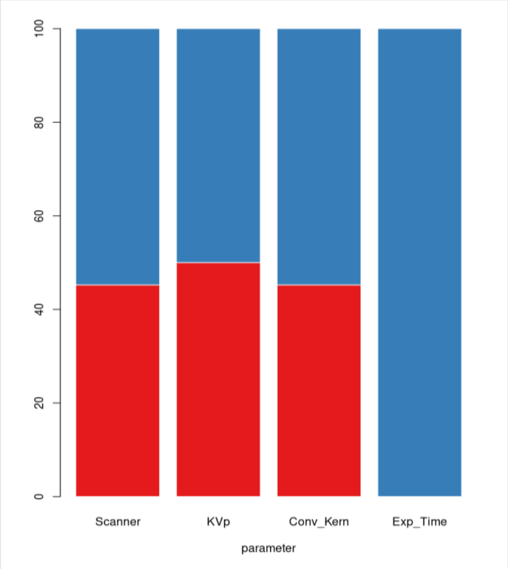
1. **Spatial Resampling:** 1mm x 1mm x 1mm
2. **Intensity Rescaling:** -1000 HU, 3000HU
3. **Intensity Discretization:** 400 bins of size 10 HU





Acquisition Parameters Harmonization





Percentage of features dependent on the parameter

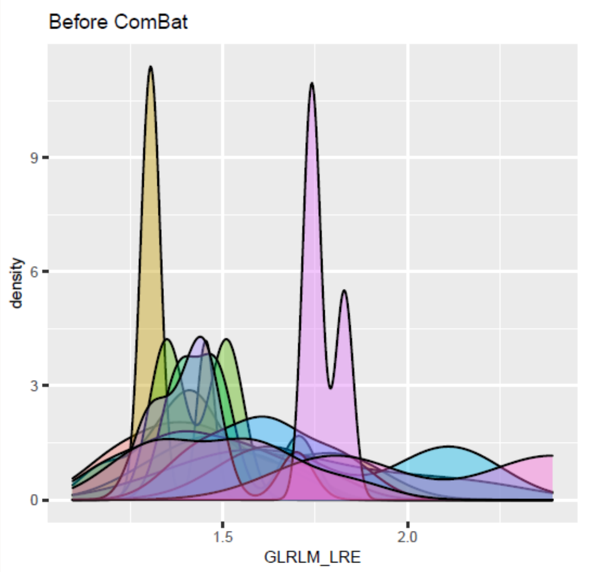
Percentage of features not dependent on the parameter

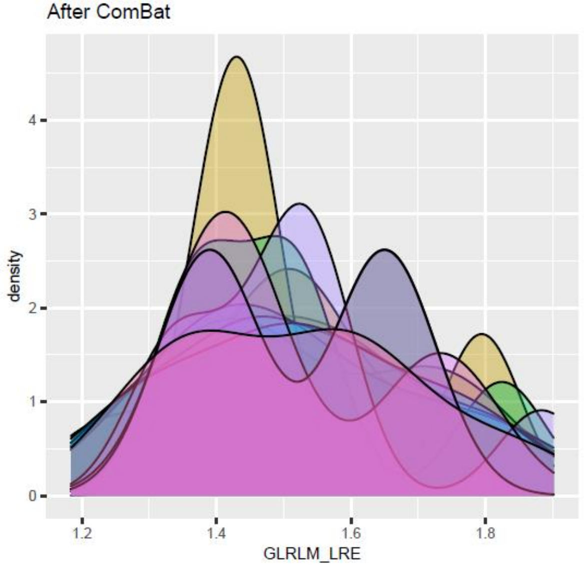
Scanner vendor, kVp and Convolutional Kernel have a statistically significant impact on features distributions

Acquisition Parameters Harmonization

ComBat: realigns features as a function of median and variance.
Long-ComBat: ComBat algorithm adapted to longitudinal data.
Batch: numeric value associated with each combination of parameters.

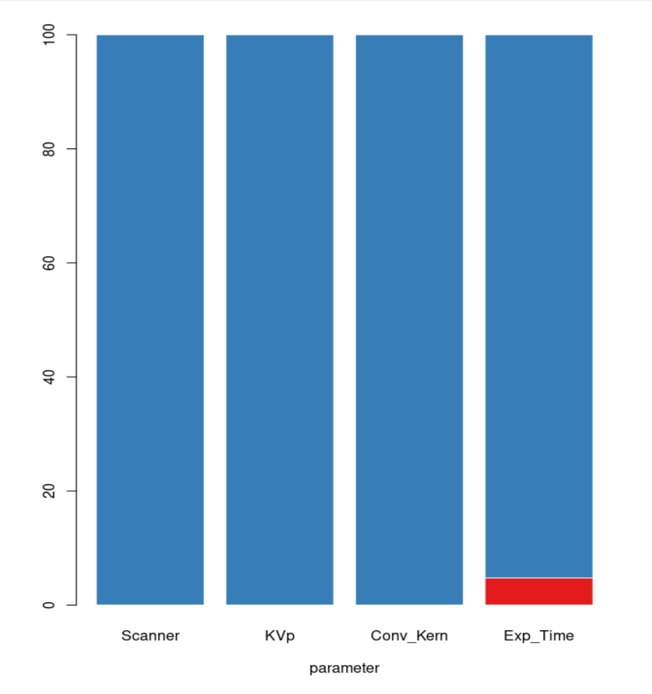
kV peak - convolution kernel combinations define **14 batches**





Acquisition Parameters Harmonization

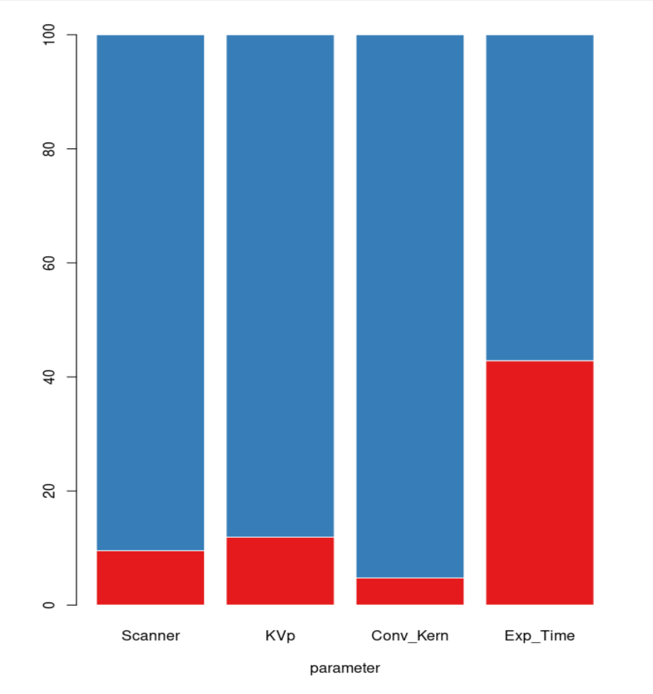
Combat



Percentage of features dependent on the parameter

Percentage of features not dependent on the parameter

Long-Combat

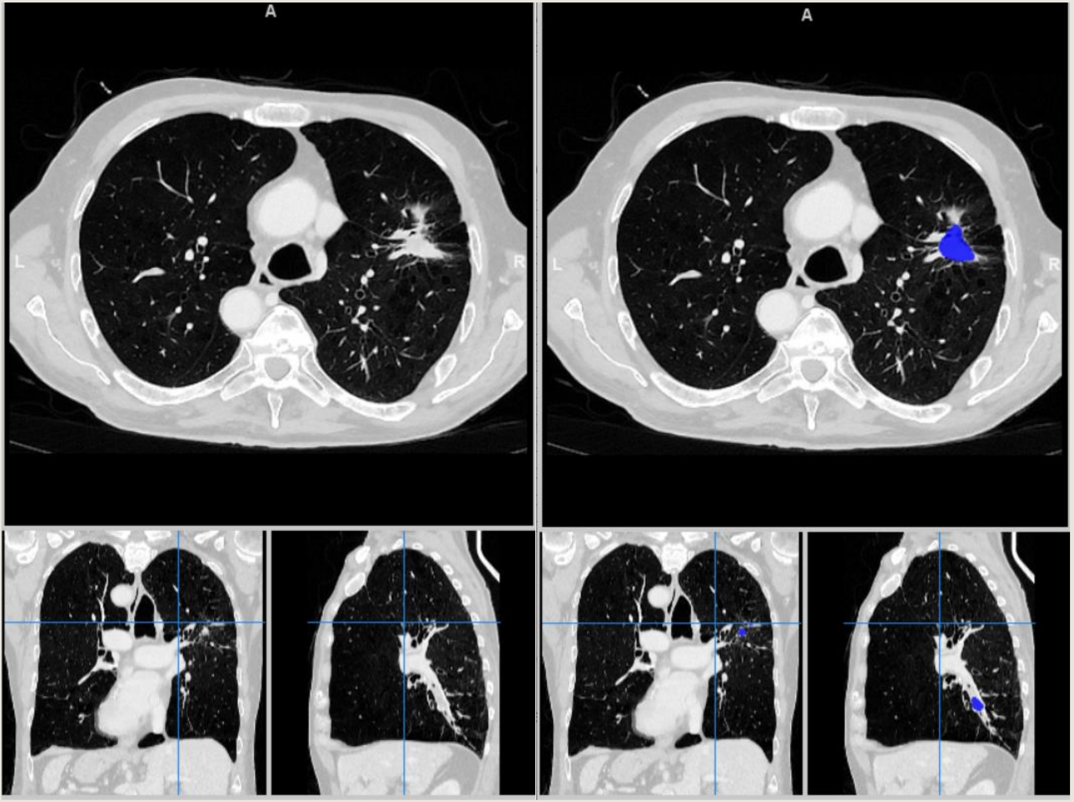


Percentage of features dependent on the parameter

Percentage of features not dependent on the parameter

nextAIM: study how harmonization processes influence predictive models

nn-Unet for tumor segmentation



Automatic segmentation to prevent human segmentation variability

Dataset:

- "Lung1" images dataset from The Cancer Imaging archive (https://wiki.cancerimagingarchive.net/display/Public/NSCLC-Radiomics) - 422 images
- Italian Lung Cancer database (from Policlinico San Matteo) - 178 images

Conclusions

1. Possibility of applying ML for different tasks such as quantitative biomarkers prediction in muscular dystrophy and automatic segmentation
2. Importance of the development of robust analysis pipeline for small-datasets
3. Images, features and segmentation harmonization steps are necessary with small-datasets
4. Importance of availability of public datasets useful for neural network training

THANKS!

Leonardo Barzaghi, Elena Ballante, Francesca Brero, Raffaella Cabini,
Giulia Colelli, Ian Postuma, Sara Panebianco, Ilaria Villa