Artificial Intelligence in Medicine

next AIM

Modality-independent explainable anomaly detection tool for neuroimaging

F. Sensi, A. Lanza, A. Chincarini

next_AIM workshop on XAI techniques for medical data analysis
16-18/10/2024, Bari

GENERAL CONFERENCE FAIR Future Artificial Intelligence Research 23|24SET2024 NAPOLI

INFN

Funded by the European Union NextGenerationEU | Ministero dell'Università e della Ricerca | Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA | FAIR Future Artificial Intelligence Research

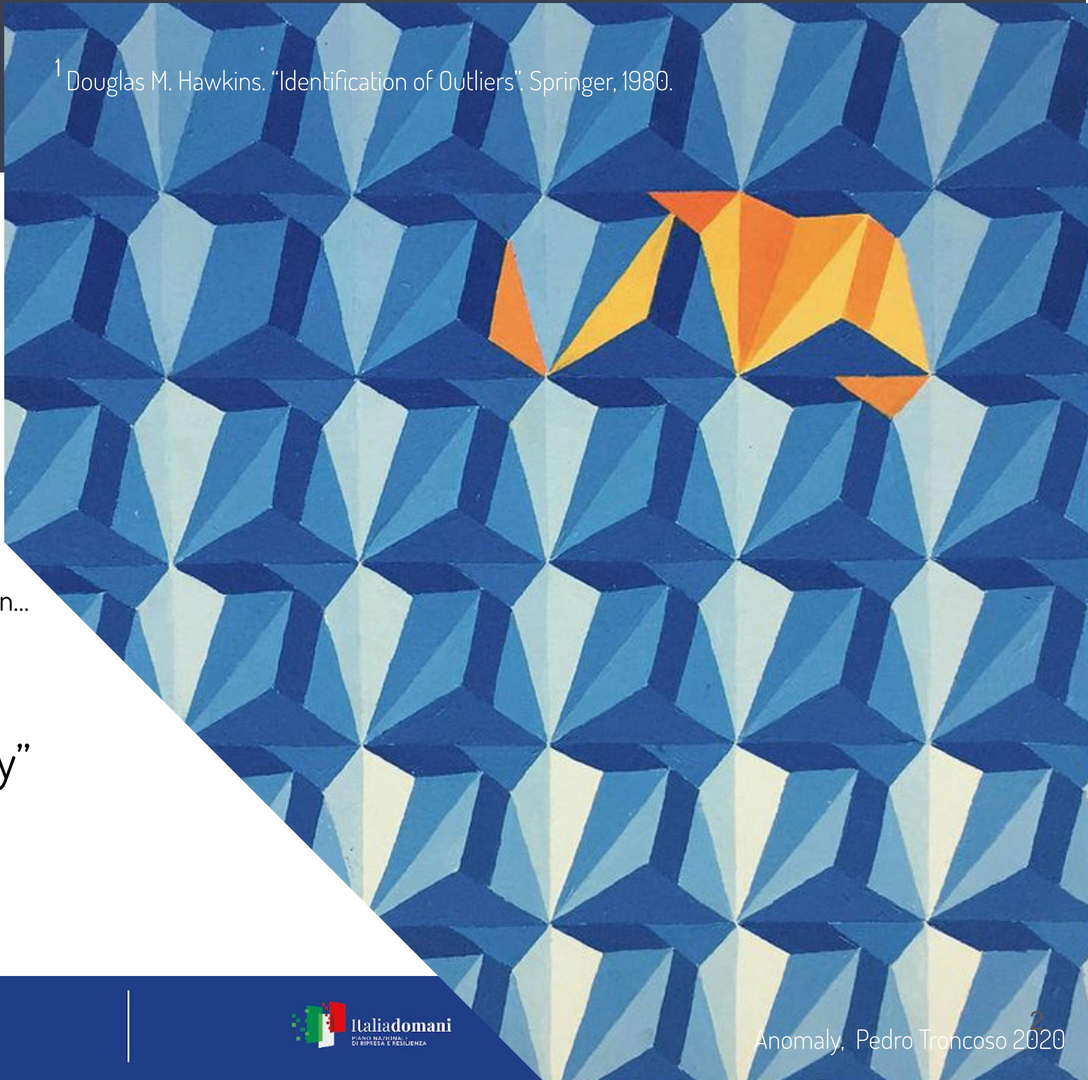
ANOMALY DETECTION

The science of identifying outliers

“...observations that deviate so much from other observations as to arouse suspicion that they were generated by a different mechanism” [1]

Most diverse domains: card fraud detection, industrial damage detection...

- Pathology as deviation from “normality”



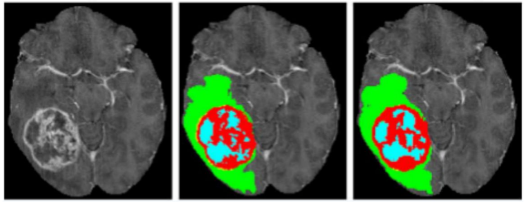
Anomaly, Pedro Troncoso 2020

[1] Douglas M. Hawkins. “Identification of Outliers”. Springer, 1980.

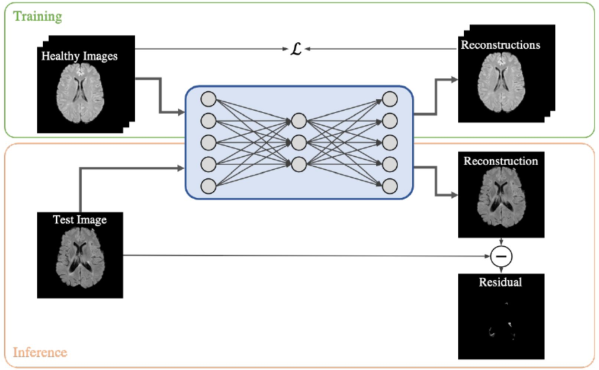
Funded by the European Union NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA

STATE OF THE ART

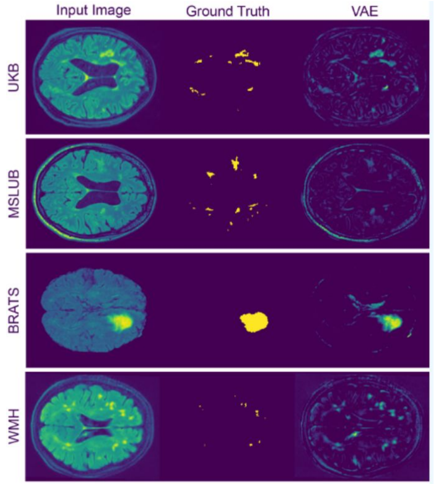
Unsupervised anomaly detection (clustering, Markov Random-fields, Dictionary Learning, ..)
Supervised anomaly detection (Convolutional neural networks)
Semi-supervised anomaly detection (Auto-encoders)
Try emulate the capability of the human eye to exploit the prior knowledge of how healthy brains should appear, in order to reach anomaly detection performance comparable to the ones of neuroradiologists



Pereira et al. “Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images”, 2016



Auto Encoder + GAN



Pinaya et al, 2022

PROPOSED METHOD

A framework for general purpose & modality agnostic Anomaly Detection.
The main outputs of our machine-learning algorithm are anomaly score and probability maps

- **Explainable:** merely practical motivation: might need to justify the decision to someone (patient, ..)
- **Generalizable** might not know a priori what is to be found
- **Applicable to small data sets** it can be “localized” on centers data

Recipe:

1) Choose metrics to measure distance from a normative dataset (1-to-many, voxel-per-voxel)
2) Define “normality” boundaries
3) Assign anomaly score o probability

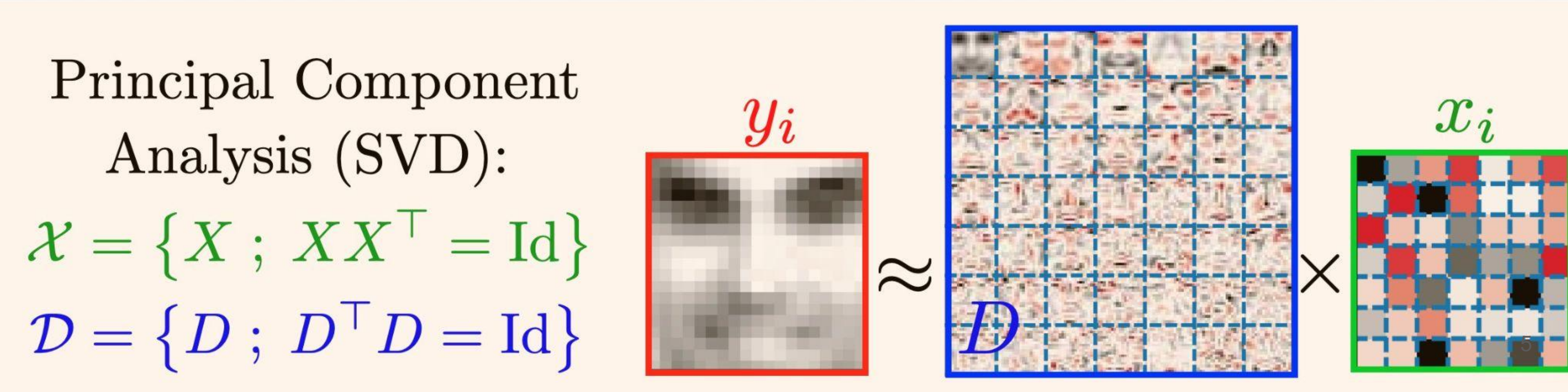
Funded by the European Union NextGenerationEU | Ministero dell'Università e della Ricerca | Italiadomani Piano Nazionale di Ripresa e Resilienza | FAIR Future Artificial Intelligence Research

DISTANCE METRICS > PCA-based

Find normative set eigenvectors → Project target image → Measure recon. error $(I_{orig} - I_{rec})/\sigma_{normative}$

Linear version of an auto-encoder
Good sensibility to small intensity differences
Decent sensibility to texture differences

Typically we have: $N_{normative} \times d \Rightarrow d \times (N_{normative} - 1)$ if $N_{normative} < d$



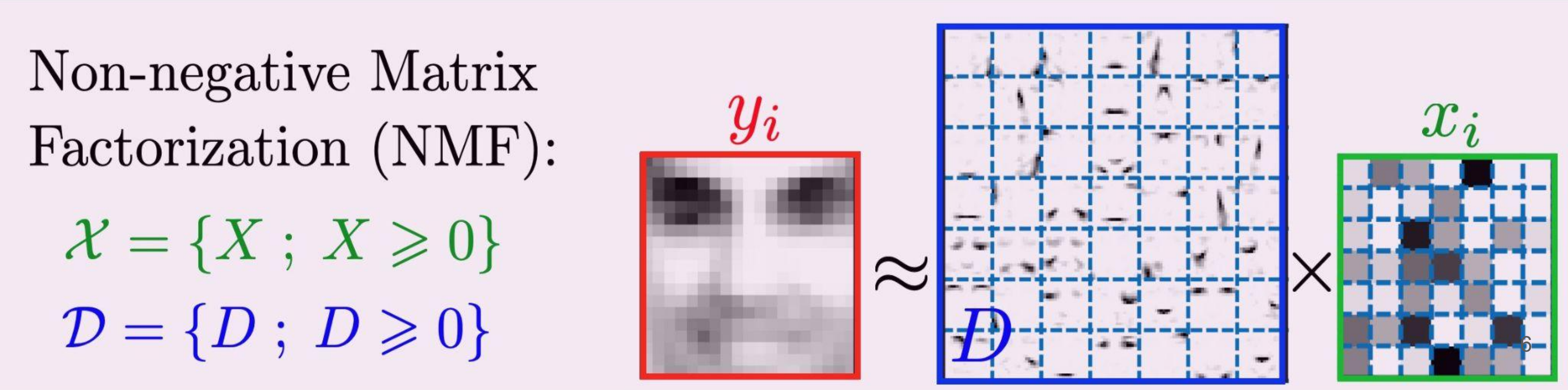
DISTANCE METRICS > NMF-based

Find normative set basis → Project target image → Measure recon. error $(I_{orig} - I_{rec})/\sigma_{normative}$

Unlike PCA, the elements of the base are non-negative and sparse: they represent the individual parts of the data

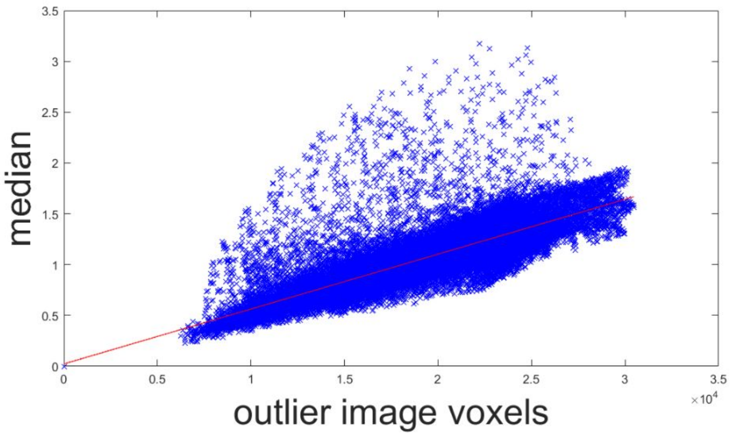
Poor sensibility
No mirror artefacts

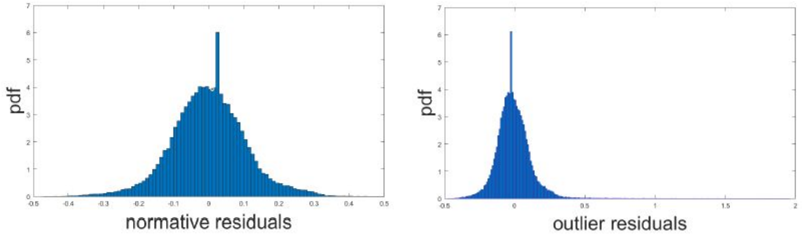
We chose: $\sqrt{N_{normative}}$ basis images



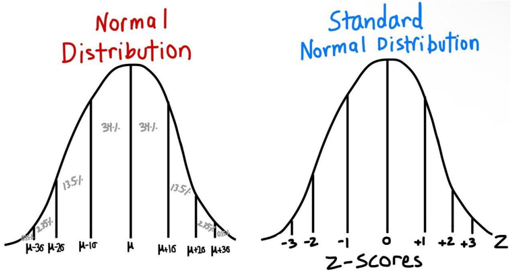
DISTANCE METRICS > Res, Z-score

Standardized residuals → (vertical distance between the point and the fitting line)/$\sigma_{\text{normative}}$





Example: residuals of normative images are distributed as a Gaussian, while the pdf of the outlier image residuals has a pronounced right tail



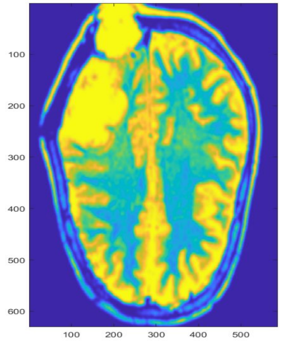
Z-score → distance from normative set mean in σ units

DISTANCE METRICS > STD, H

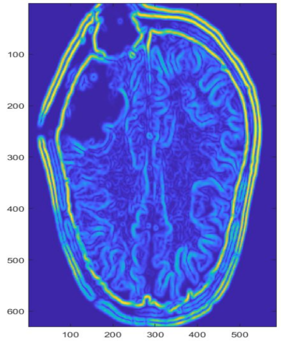
“Texture” metrics: capture pattern-based anomalies. Defined on a 5x5x5 voxel volume.

Standard deviation metric $\frac{STD_A(i,j,k) - M(i,j,k)}{S(i,j,k)}$

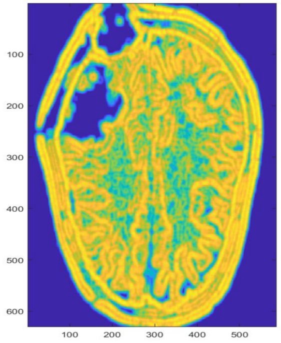
Entropy metric $\frac{H_A(i,j,k) - M(i,j,k)}{S(i,j,k)}$



orignal MRI image (brain tumor)



Entropy image

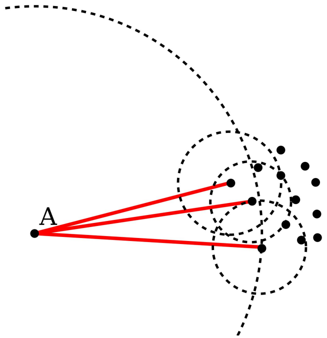


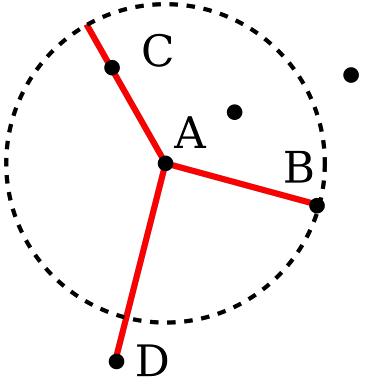
STD image

Funded by the European Union NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani Piano Nazionale di Ripresa e Resilienza
Future Artificial Intelligence Research

LOCAL OUTLIER FACTOR

Local Outlier Factor: comparing the local density of a point with the densities of its k-nearest neighbors. A point that has a much lower density than its neighbors is considered outlier.





K-distance (D) is the distance of a point to its *k_th* neighbor

Reachability distance (RD) is the distance need to travel from particular point to its neighbor point

max(*k*-distance(B), distance(A, B))

Reachability distance (LRD) is the inverse of the average RD of its neighbors

$$LRD_k(x) = 1 / \left(\frac{\sum_{o \in N_k(x)} d_k(x, o)}{|N_k(x)|} \right)$$

LOF score is the is the average LRD of the neighbors divided by object's own LRD

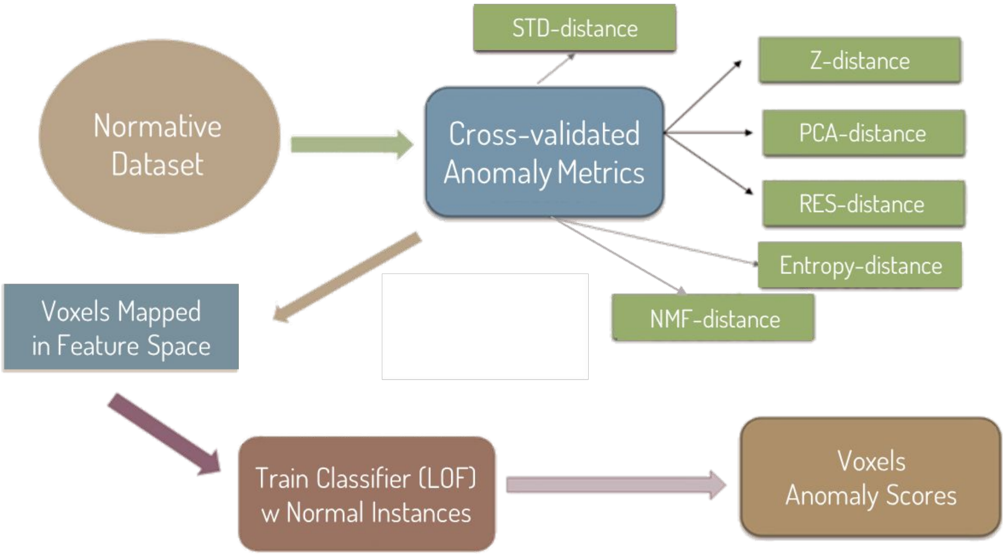
$$LOF(x) = \frac{\sum_{o \in N_k(x)} \frac{LRD_k(o)}{LRD_k(x)}}{|N_k(x)|}$$

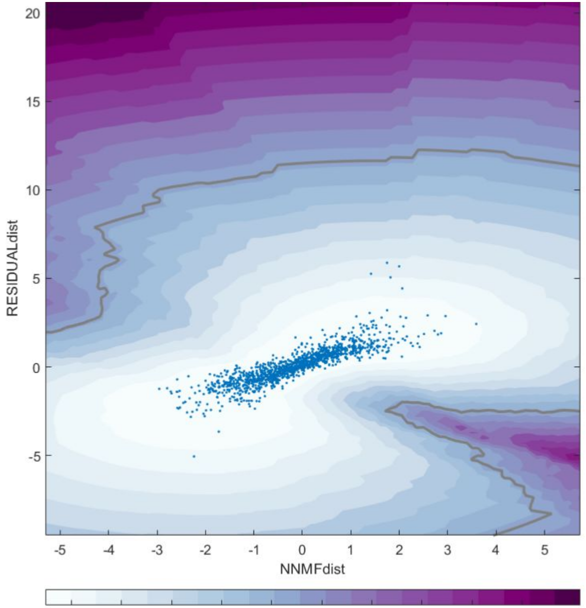
TRAINING

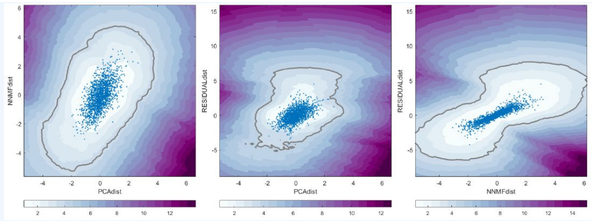
Distance images are created from normative instances:

LOF operates in 2D feature subspaces (e.g., NMF-distance vs RES-distance)

The system assigns an anomaly score to each pixel, and those above
a certain threshold (determined by the algorithm) are classified as anomalies



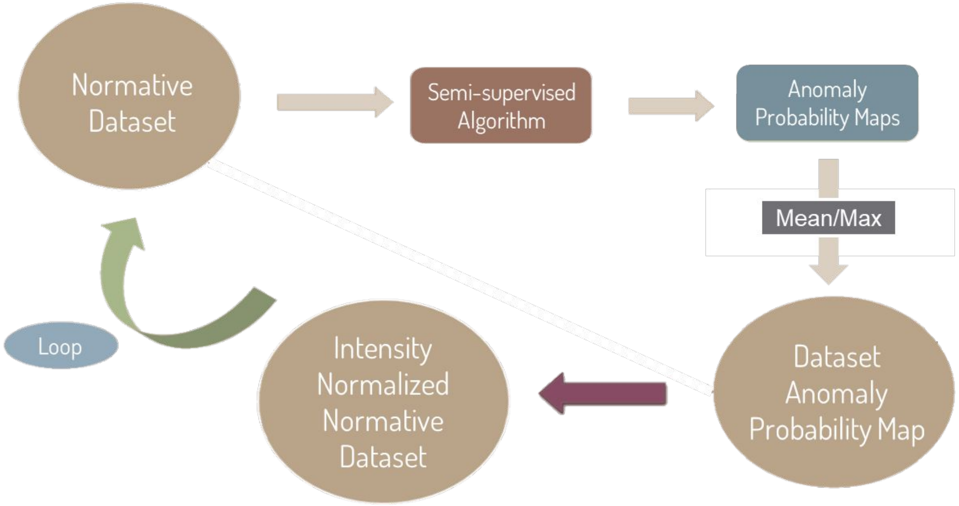




INTENSITY NORMALIZATION

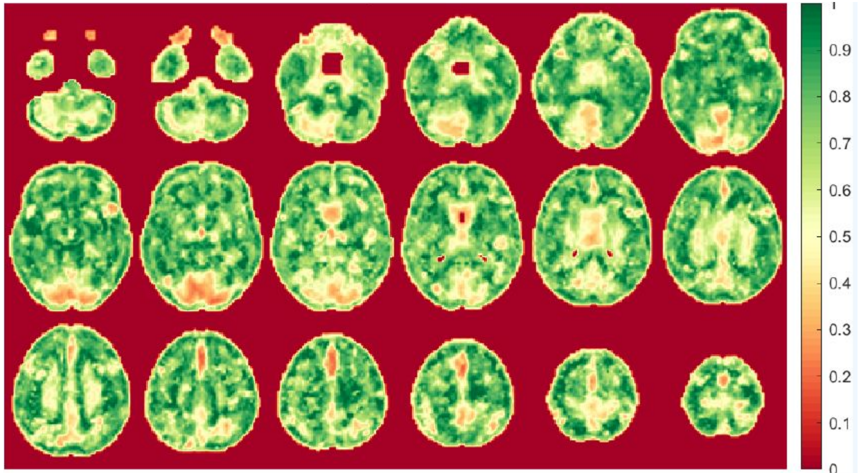
Intensity normalization is typically region-based (e.g., segmentation, non-affected region), whole-brain or limited to normal subjects.
Risks: pathology dependent, sensitive to outliers..

To remain general, we chose to stick with total counts in data-driven masks (auto-calibration)



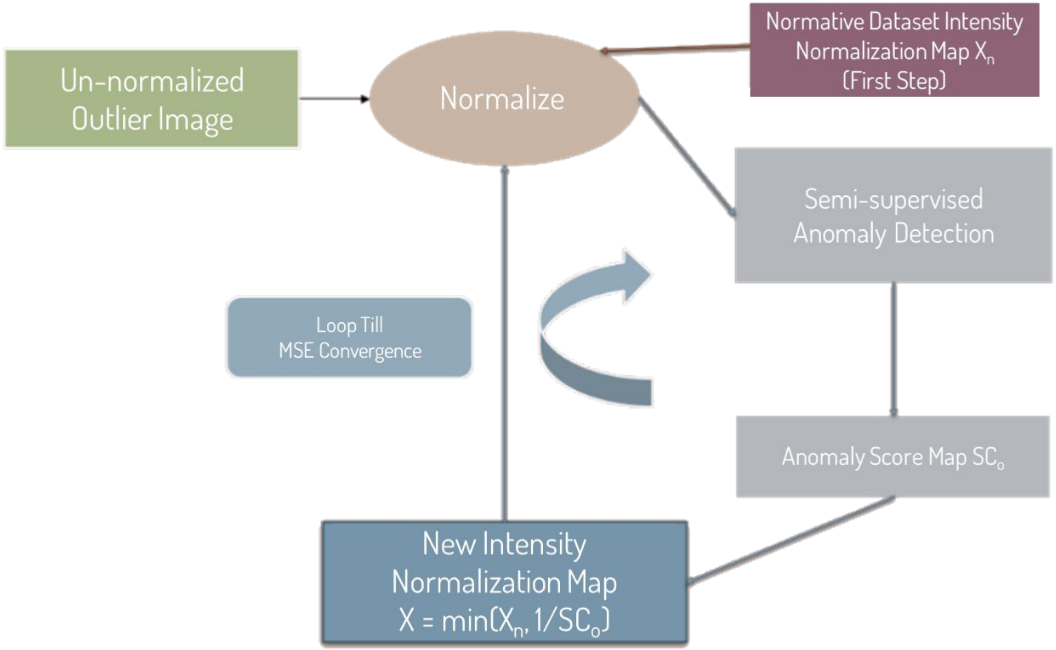
Auto-calibration for normative dataset (a *faster* version is also available..)

STEP 3



Xn map for ^{18}FDG-PET normative dataset

TESTING



Auto-calibration for test image

Test images undergo a similar procedure

Intensity normalization is so that normative dataset intensity normalization map X_n is combined to with S_{Co} to obtain an intensity normalization where anomalous regions have a lower weights

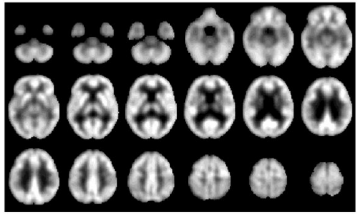
$$x(i) = \min\left(X_n(i), \frac{1}{SC_o(i)}\right)$$

Funded by the European Union NextGenerationEU | Ministero dell'Università e della Ricerca | Italiadomani Piano Nazionale di Ripresa e Resilienza | FAIR Future Artificial Intelligence Research

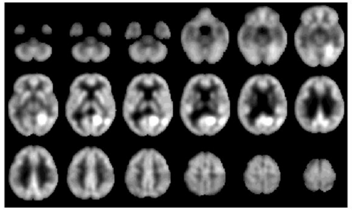
SYNTH FDG-PET DATA

The anomaly images are created by adding pseudo-Gaussian intensity anomalies onto images of the normative dataset

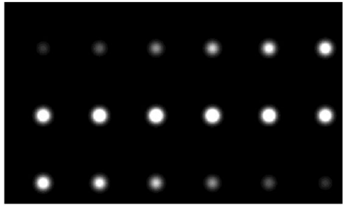
$$H(x,y,z) = a\exp\left(-(\frac{x-x_0}{\sigma})^s\right)\exp\left(-(\frac{y-y_0}{\sigma})^s\right)\exp\left(-(\frac{z-z_0}{\sigma})^s\right)$$



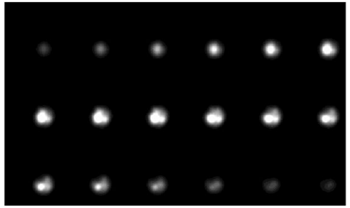
(a) ORIGINAL



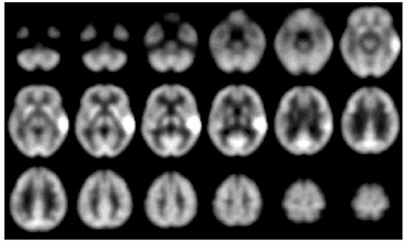
(b) synthA



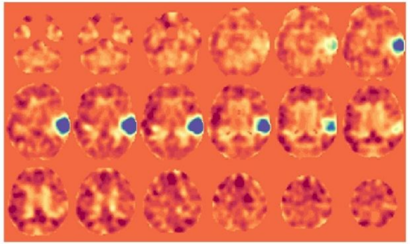
(a) synthH



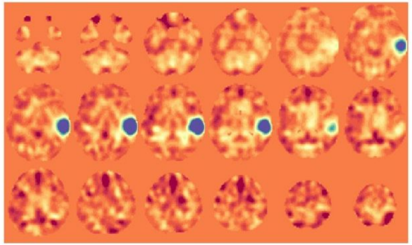
(b) delta



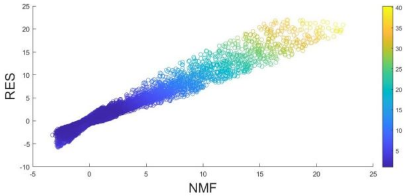
(a) SynthA



(b) NMF-distance

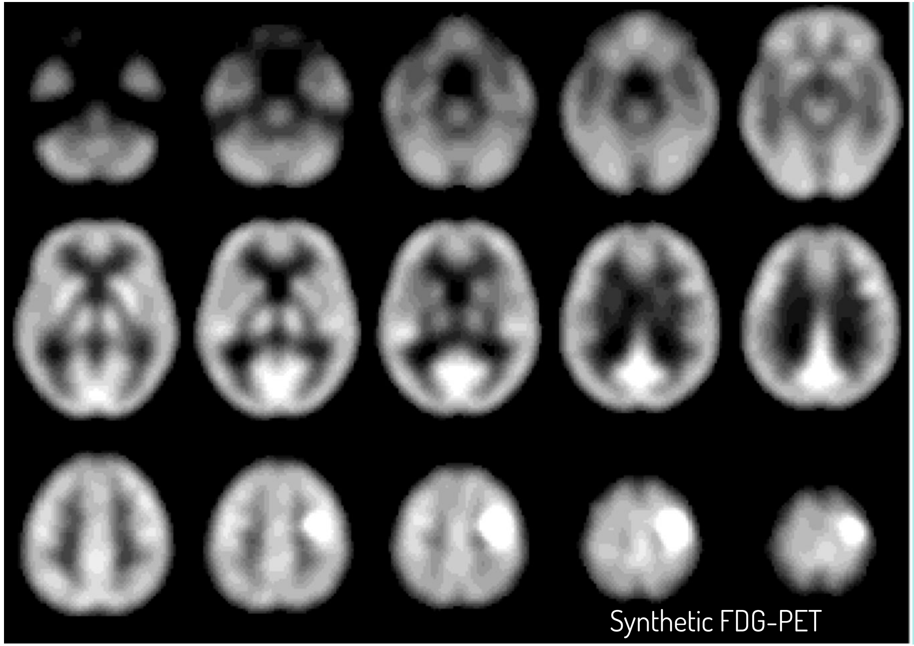


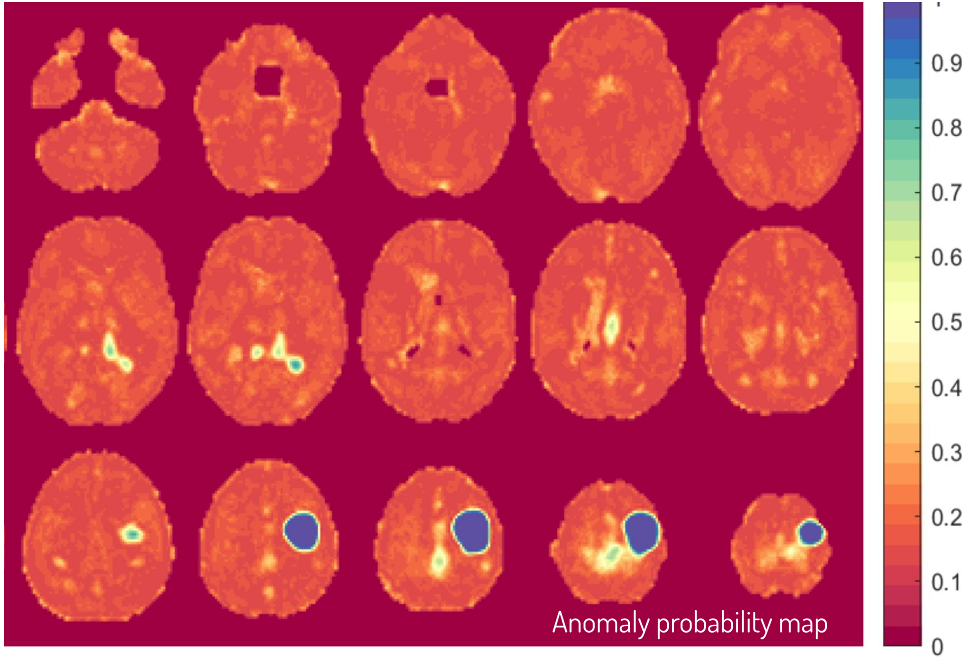
(a) RES-distance



(b) LOF scores

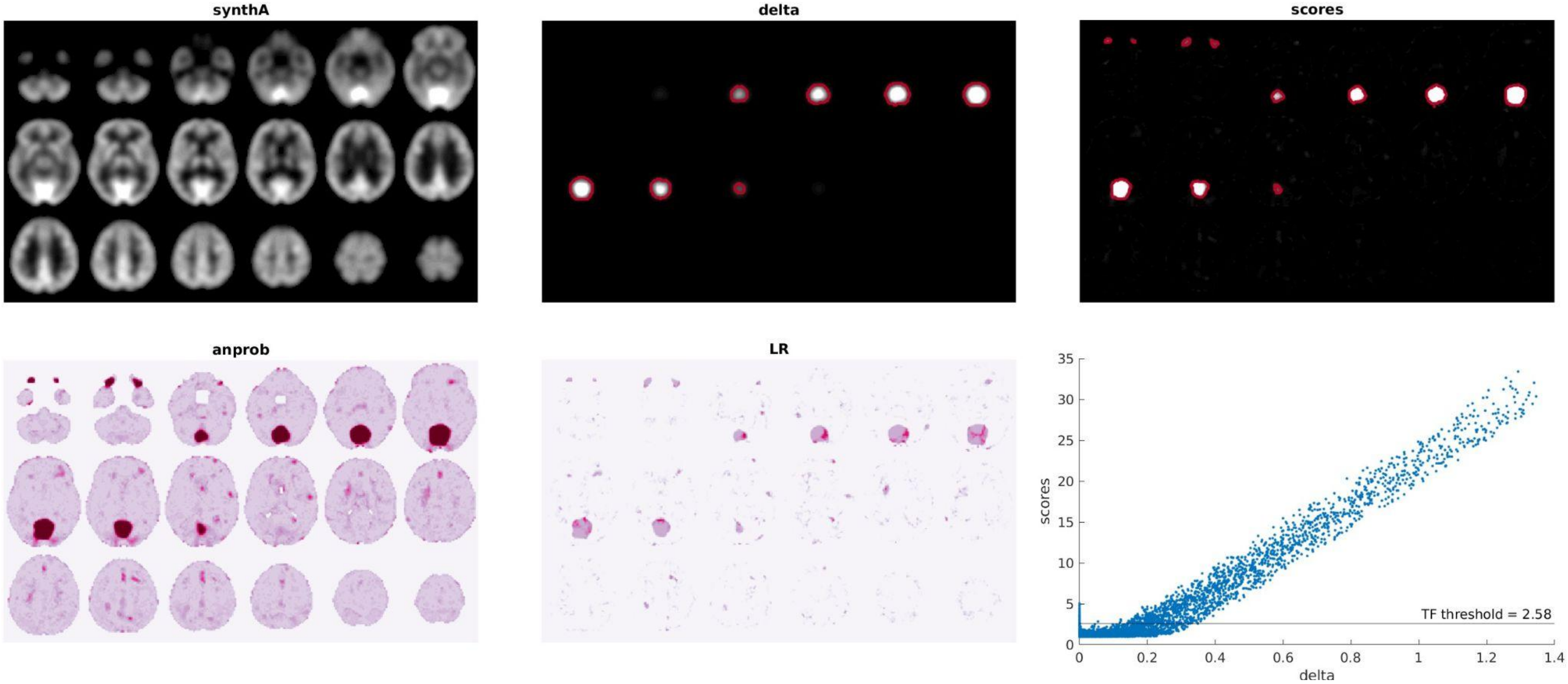
SOME RESULTS





100 synthetic FDG-PET images created from 125 real FDG-PET of healthy subjects (San Martino Hospital, Genoa).
Example shows a high accuracy in hyper-metabolism detection.

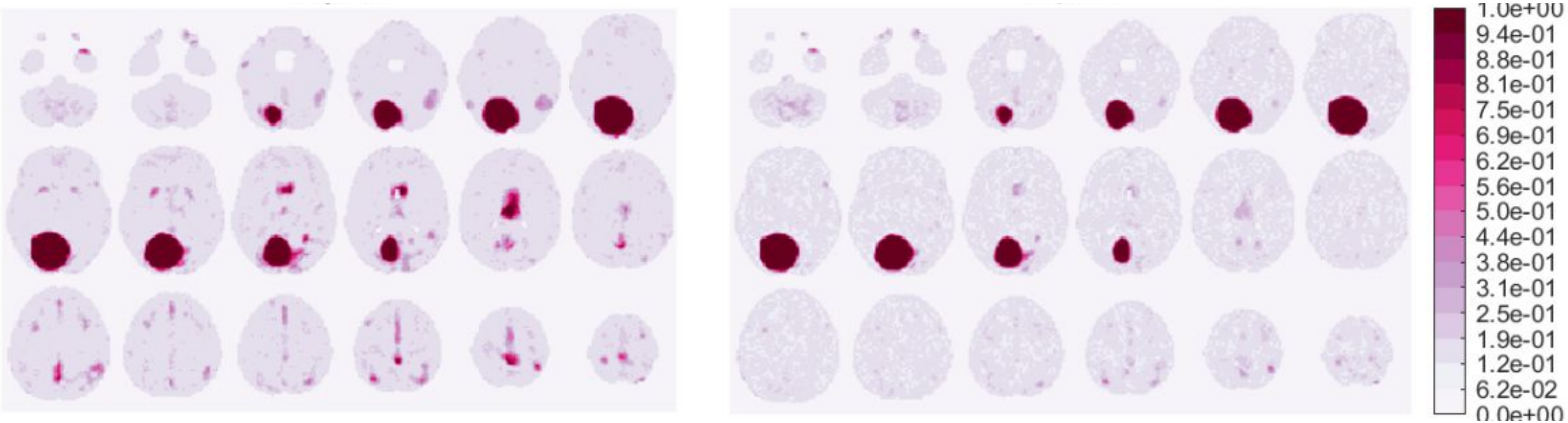
SOME RESULTS



SOME RESULTS

Auto-calibration proves to be best suited in this context: anomaly probability maps come out cleaner

Whole-brain normalized (NMF vs RES) - Auto-calibrated

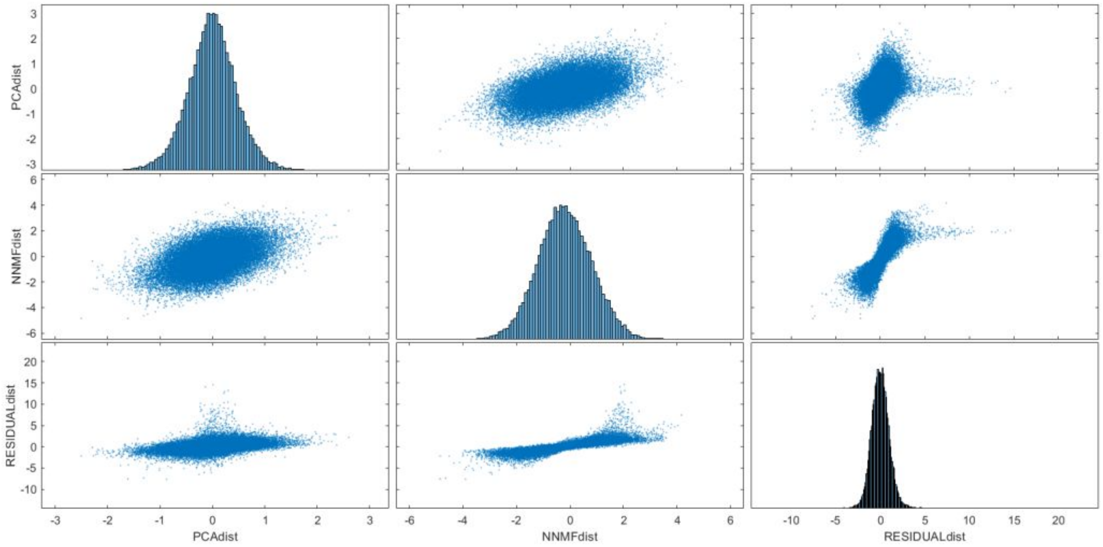


Anomaly probability

Funded by the European Union NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA
FAIR Future Artificial Intelligence Research

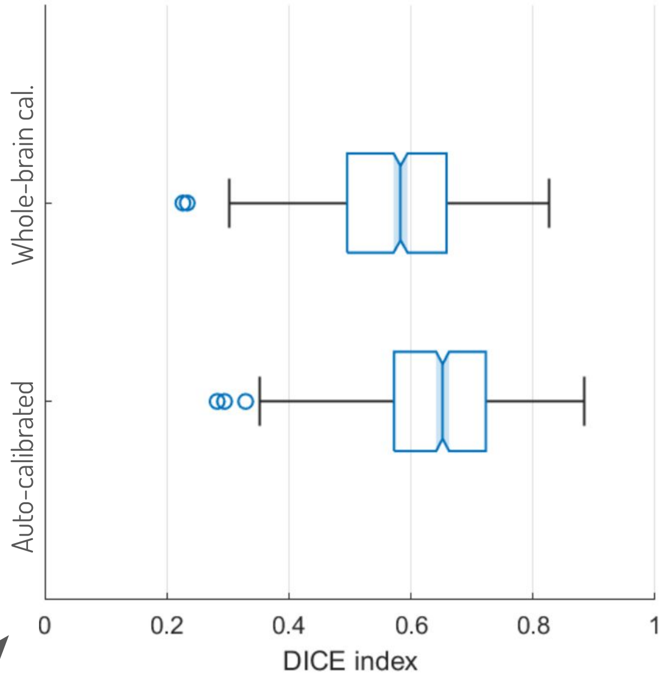
SOME RESULTS

Histogram shapes show discrepancies in metrics comparison



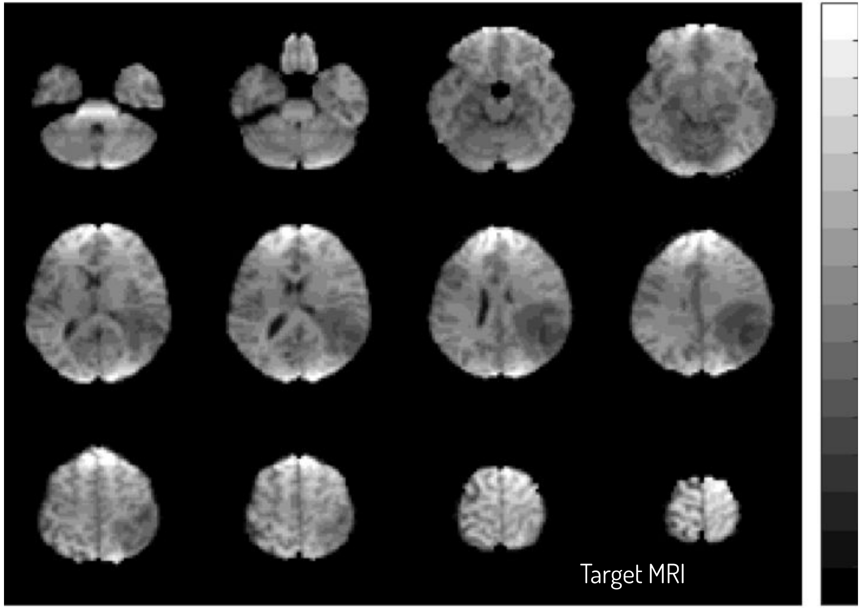
$$cDICE = 1 - \exp \frac{\sum x(y<0)}{\sum x(y>0)\log(0.5)}$$

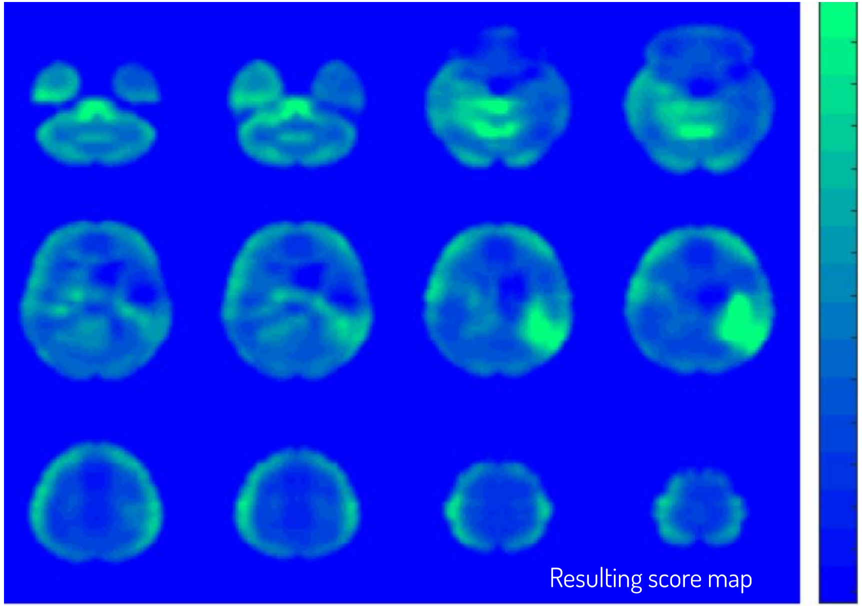
Importance of intensity normalization



Funded by the European Union NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA
FAIR Future Artificial Intelligence Research

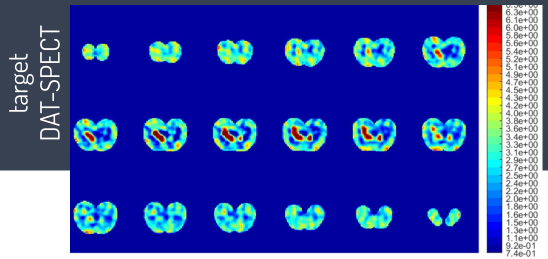
MRI DATA





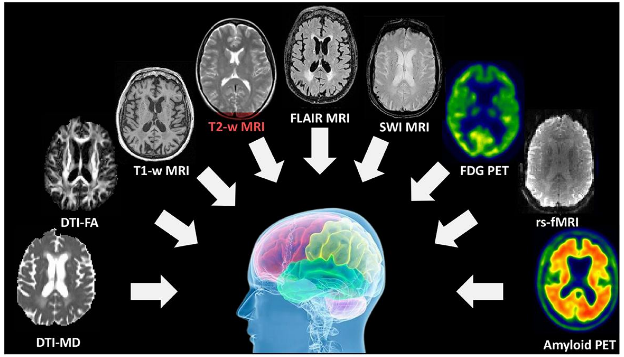
20 MRI of glioma affected patients; system trained on 75 healthy subjects (San Martino Hospital, Genoa).
Example shows the accuracy in detection. The task was complicated by the fact that anomaly is not ipo nor hyper-intense

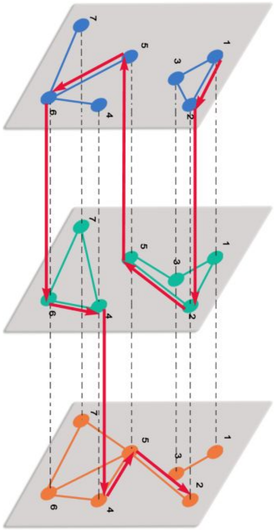
PERSPECTIVES

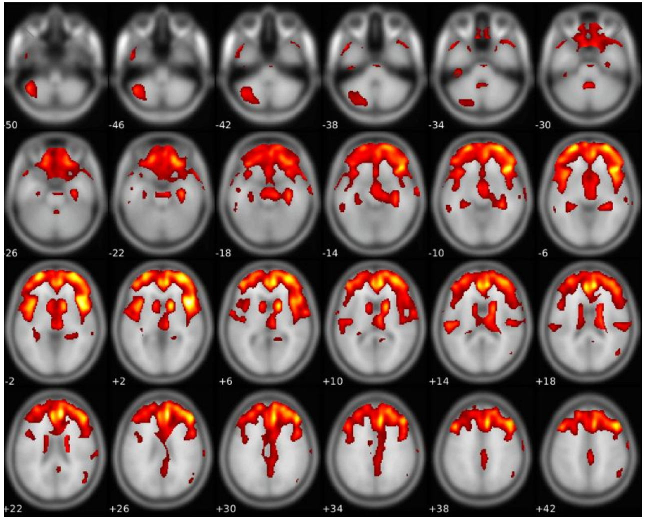


To test: Different metrics, alternatives to Local Outlier Factor
To test: Compare with well-known voxel-based morphometry software (FDG-PET)

To study: Score maps from different modalities composed into multi-layered matrix to study pathology models?







Thank you for your attention

GENERAL CONFERENCE
FAIR Future Artificial Intelligence Research
23|24SET2024
NAPOLI

Funded by the European Union NextGenerationEU | Ministero dell'Università e della Ricerca | Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA | FAIR Future Artificial Intelligence Research