Synthetic PET images generation from small datasets

Lorenzo Di Meco, Andrea Chincarini, Daniel Remondini

nextAIM Meeting - Milano 13-15 Febbraio 2023

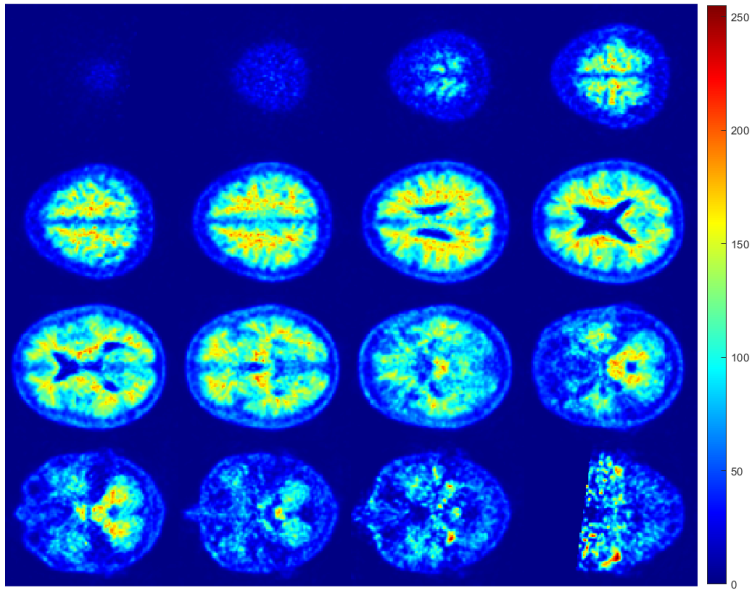
Aim

Generation of synthetic brain PET images to overcome privacy issues about management of sensitive data.
Nonlinear Dimensionality Reduction
Data Manifold mapping inversion

Dataset

1001 brain PET images from 21 European research centres
457 negative, 540 positive Alzheimer's disease diagnosis and 4 unknown patients

PET images: 97-by-115-by-97 matrices, reshaped into 1082035-dimensional vectors
Gray levels standarized at single-patient level

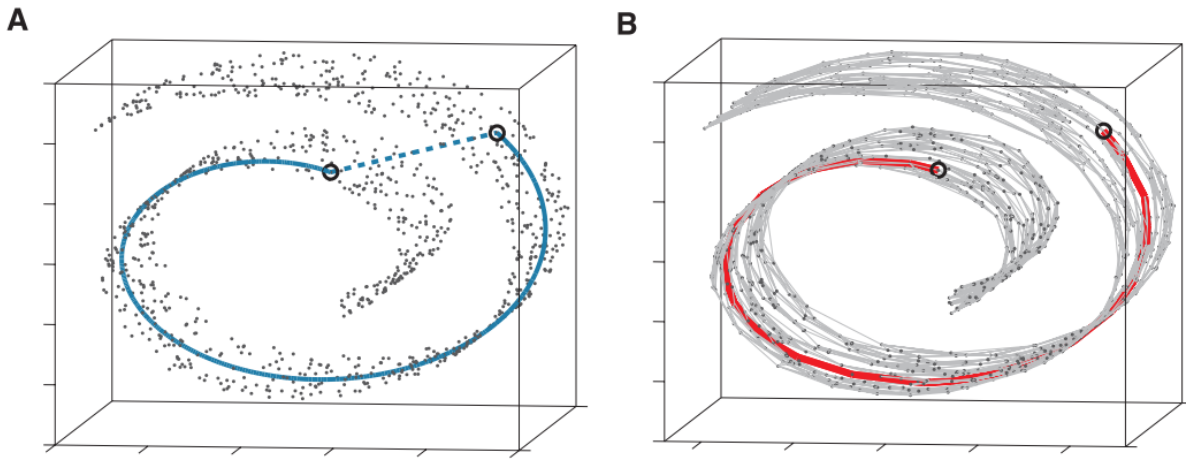


Dimensionality reduction: PCA and Isomap

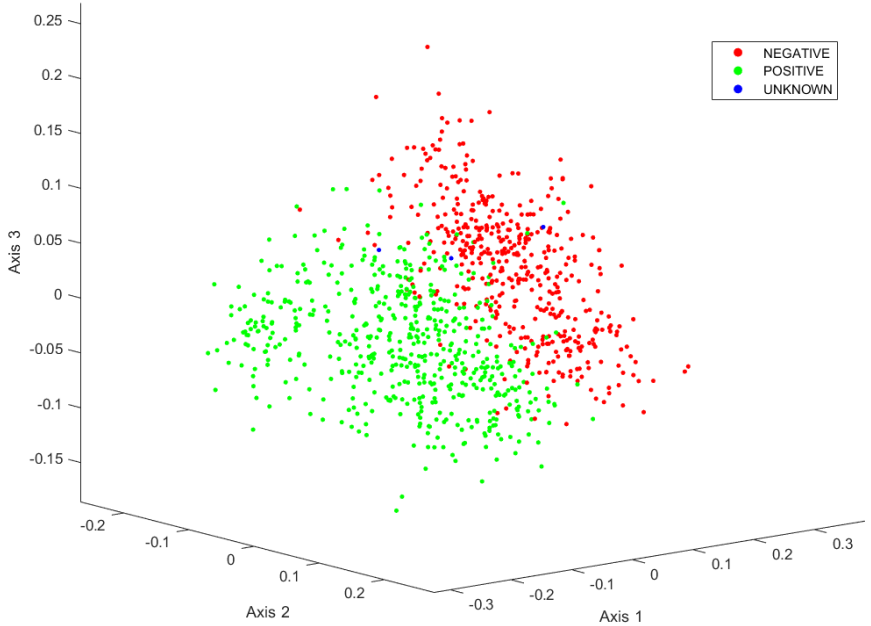
PET image $\vec{x} \in \mathbb{R}^{1082035} \rightarrow \vec{z} = \mathbf{\Phi}(\vec{x}) \in \mathbb{R}^{10}$

1) PCA: covariance matrix diagonalization.

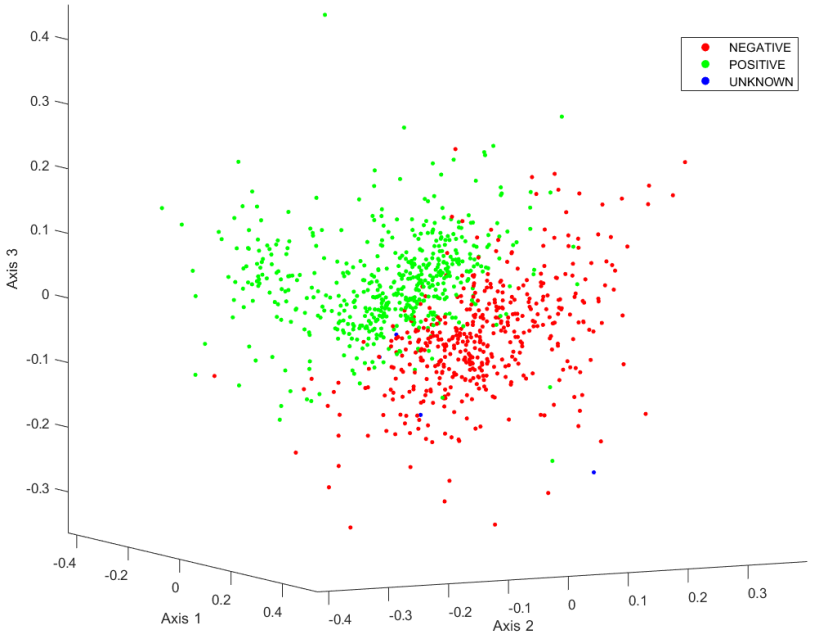
2) Isomap: non-invertible generalisation of MDS algorithm.



PCA and Isomap dimensionality reduction



PCA



Isomap

Back to original image space

1) PCA: analytical backprojection onto original space
2) Isomap: cubic RBF interpolant

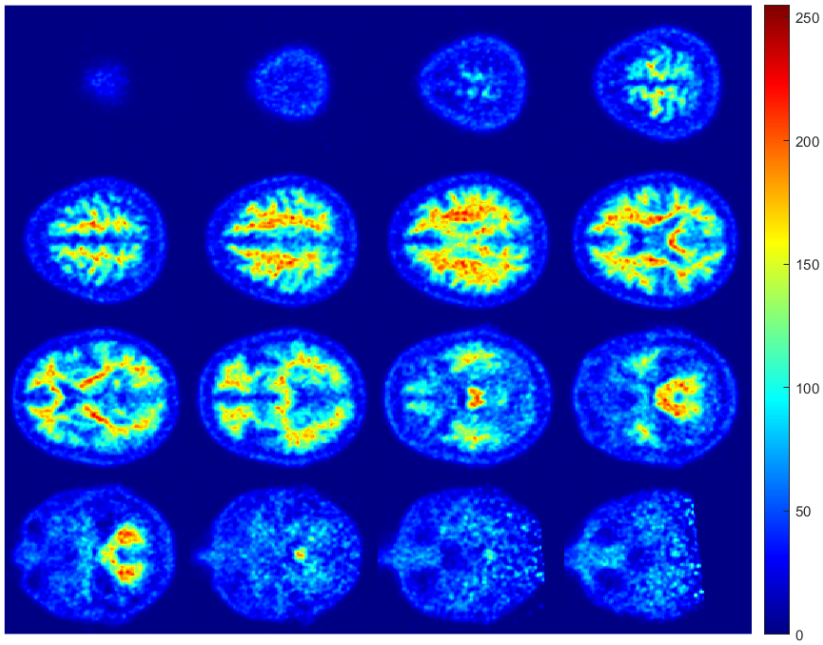
Comparison of reconstruction:

LOOCV Cross-validation of inverse mapping algorithm.

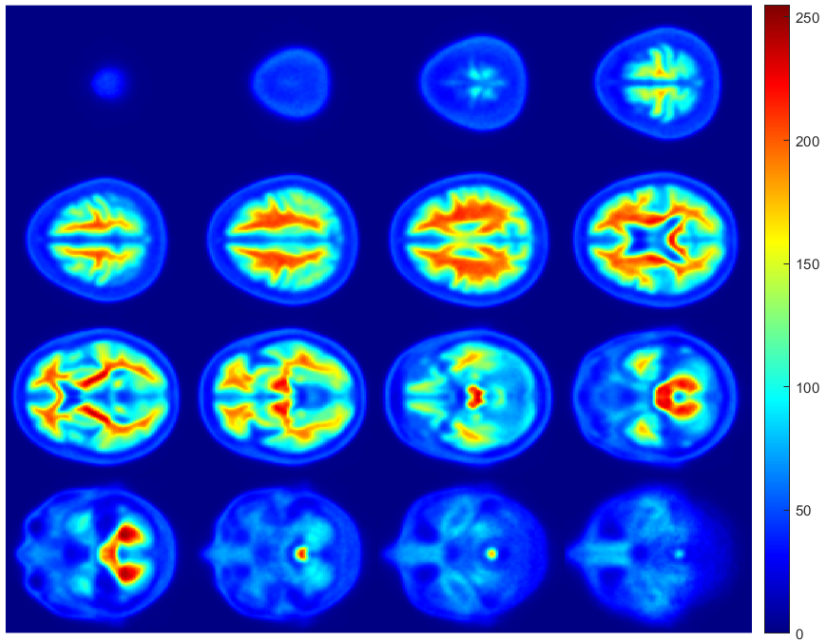
Reconstructed and original images comparison:

- Euclidean distance reconstruction error RMSE
- Structural Similarity Index SSI

Image reconstruction with PCA

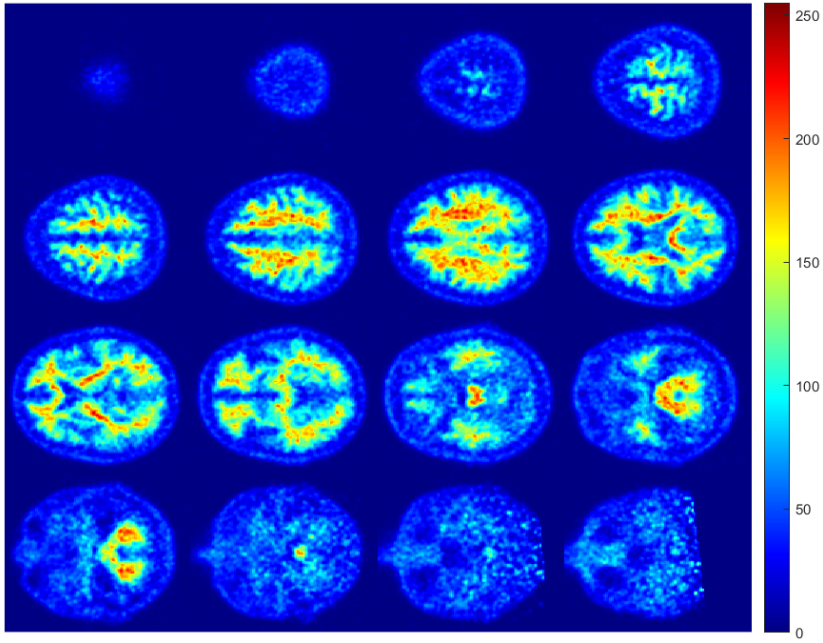


Original PET image

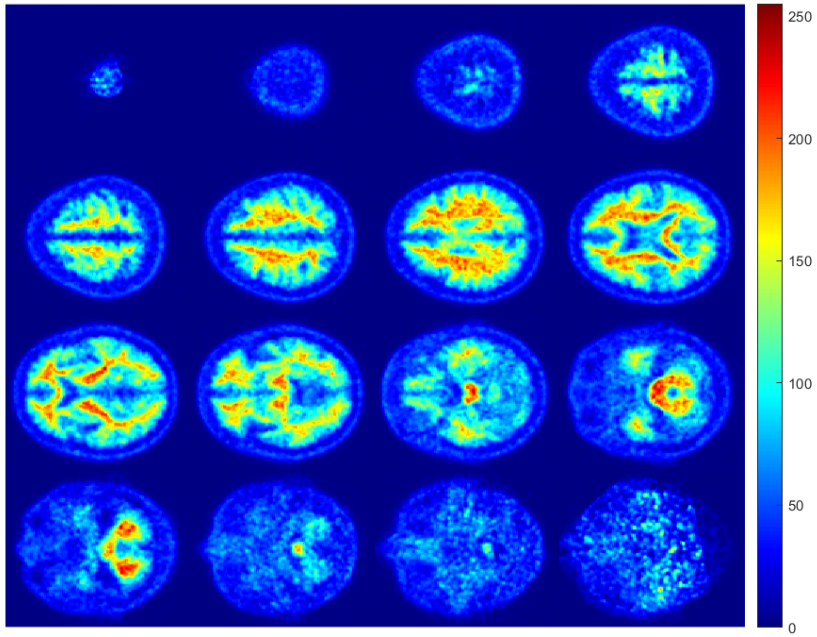


PCA reconstruction
SSIM ≈ 0.83

Image reconstruction with cubic RBF interpolant

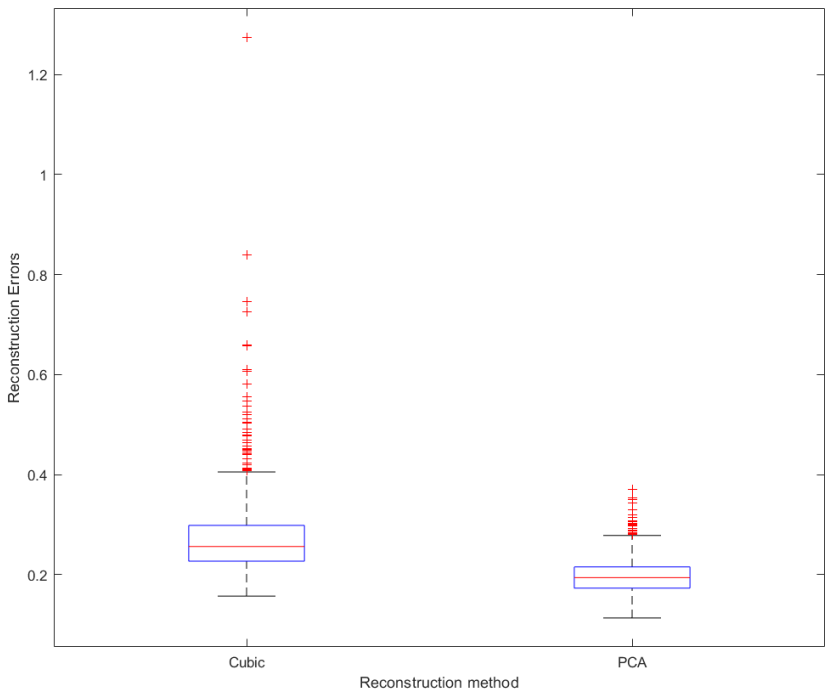


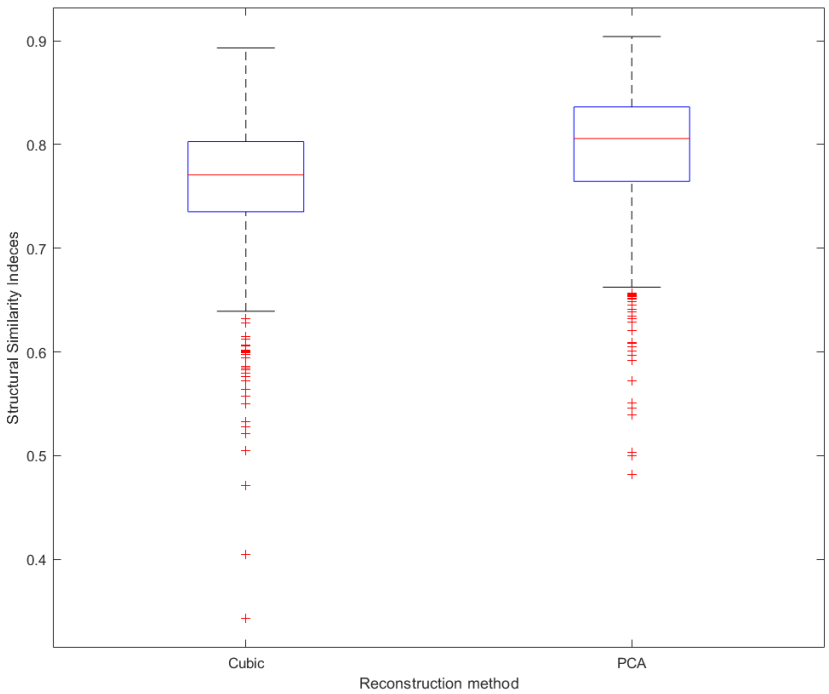
Original PET image



Cubic RBF reconstruction
SSIM ≈ 0.78

Dataset statistics: PCA is better!





$$E_{\text{cubic}} = 0.24 \pm 0.03$$
$$\text{E}_{\text{PCA}} = 0.18 \pm 0.02$$

$$\text{SSIM}_{\text{cubic}} = 0.76 \pm 0.06$$
$$\text{SSIM}_{\text{PCA}} = 0.79 \pm 0.06$$

Synthetic brain PET images

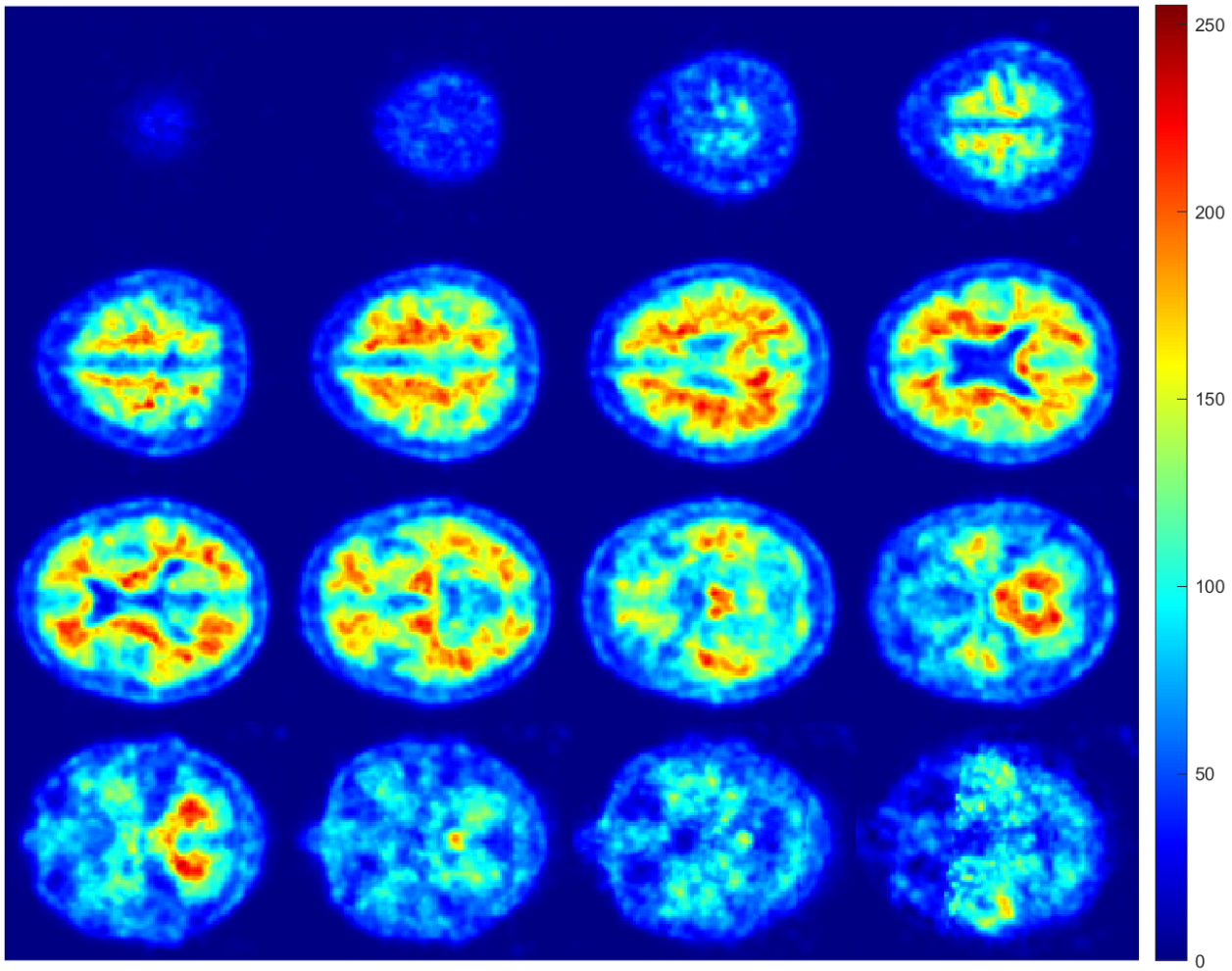
- Random choice of 2 neighbouring data points $\vec{z}_1$ and $\vec{z}_2$ from same patient class on 10-d Isomap space;
- interpolation of a random point $\vec{z}$:

$$\vec{z} = \vec{z}_1 + r \cdot (\vec{z}_2 - \vec{z}_1) \quad \in \quad \mathbb{R}^{10} \qquad r \in (0, 1)$$

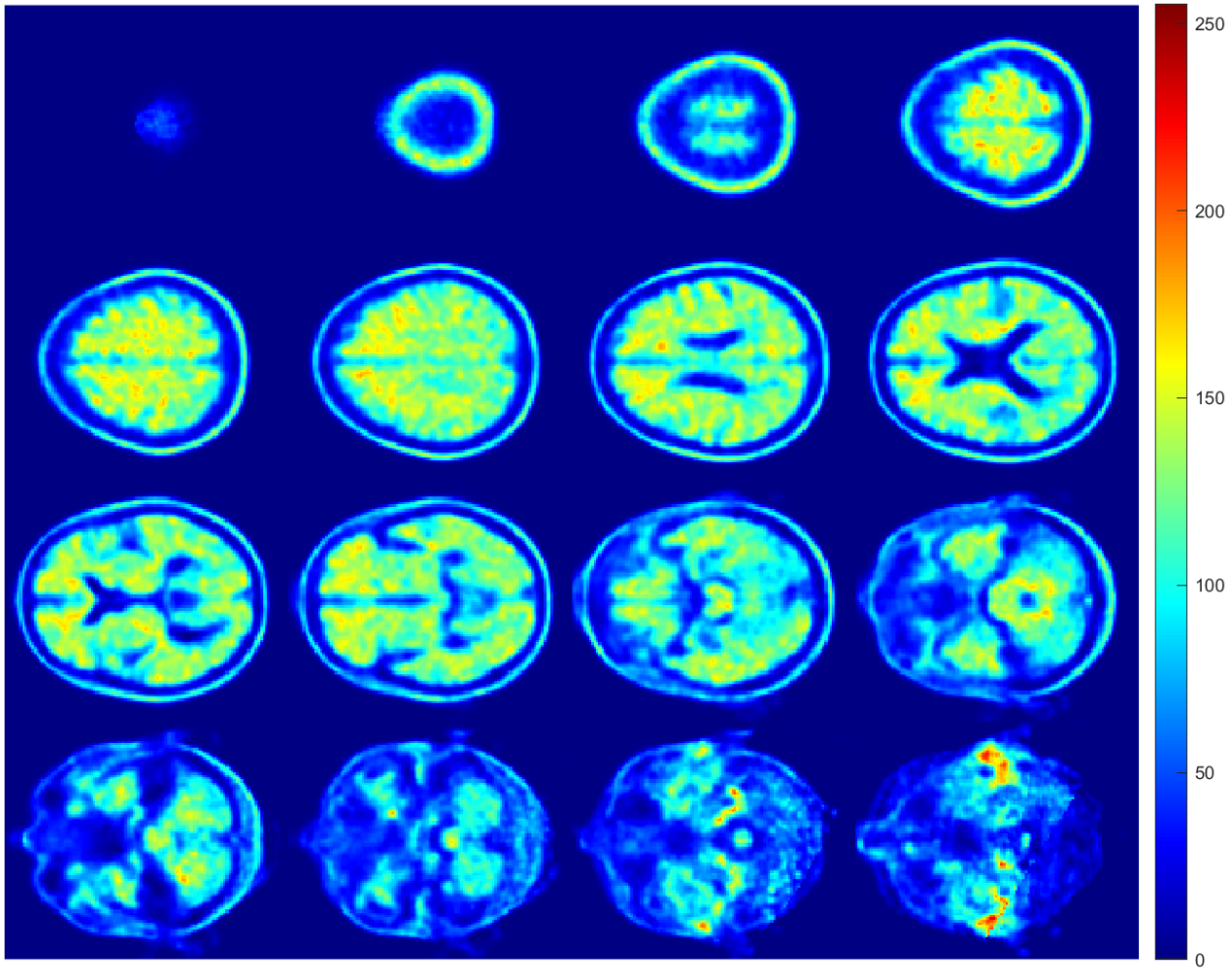
- back-reconstruction of synthetic PET via inverse mapping:

$$\vec{x} = \mathbf{\Phi}^{\dagger}(\vec{z}) \quad \in \quad \mathbb{R}^{1082035}$$

Synthetic negative PET image



Synthetic positive PET image



Manual validation by panel of experts

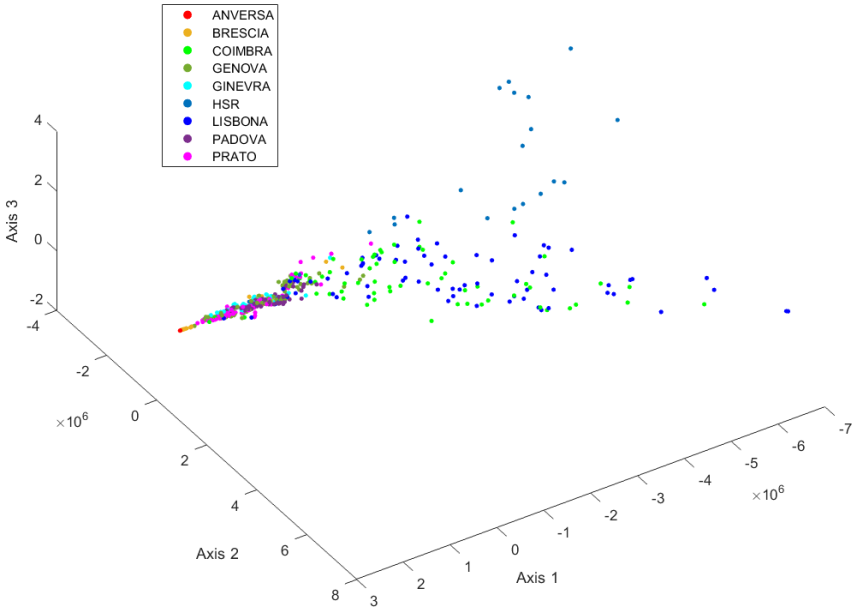
- 10 synthetic PET images: 5 negative and 5 positive;
- 10 real images: 5 negative and 5 positive;
- visual assessment by 4 experienced clinicians.

	Real	Synthetic	**p-value**
Clinician 1	2/10	6/10	0.63
Clinician 2	4/10	4/10	0.66
Clinician 3	7/10	8/10	0.07
Clinician 4	5/10	5/10	1

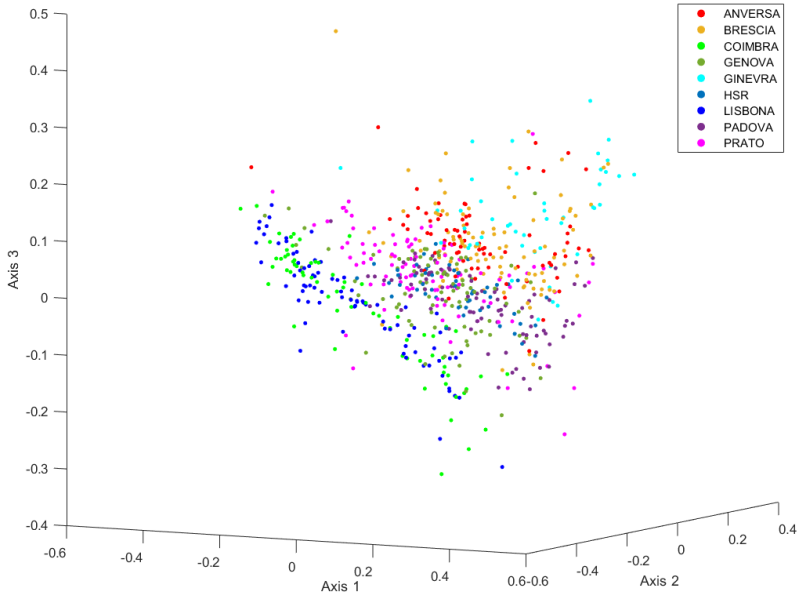
Conclusions

Generation of synthetic PET samples through low-dimensional mapping inversion
Manual validation by experts suggests good reconstruction quality
RMSE SSI quality indexes less reliable than visual inspection

Effect of data normalization on multi-center acquisition batches

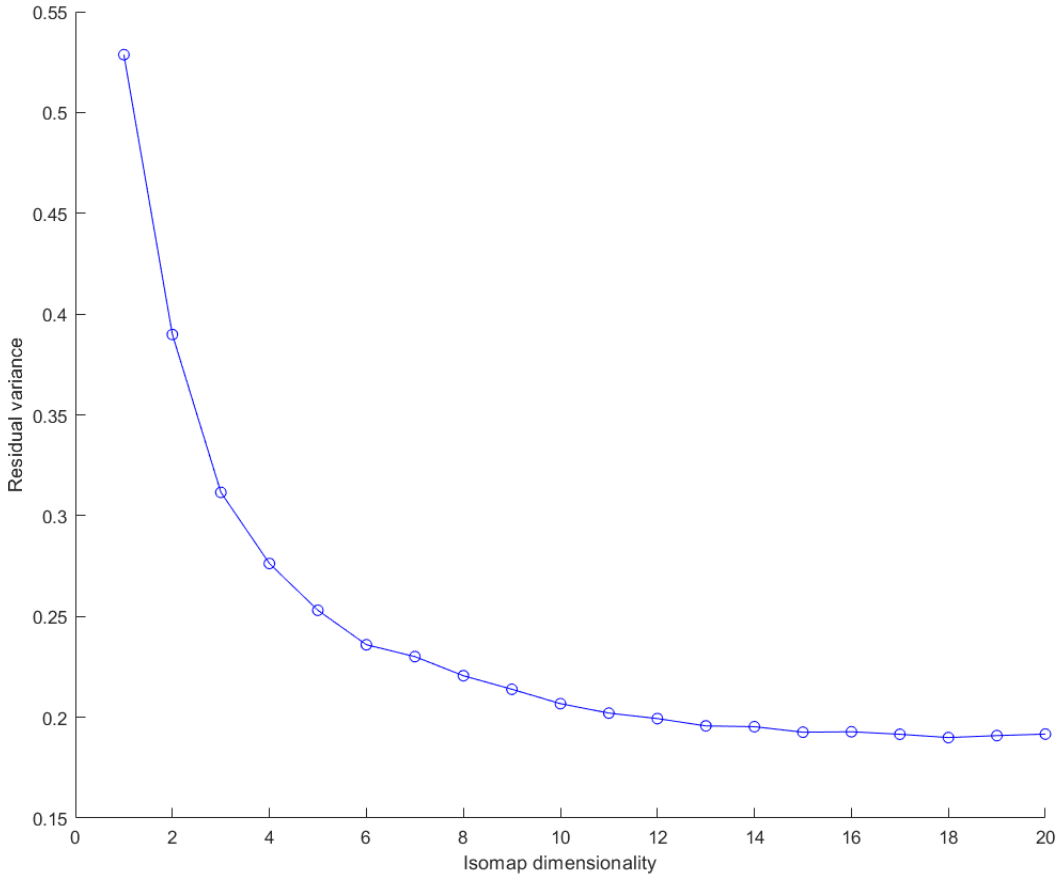


Before normalization



After normalization

Residual Variance by Isomap: flattening around 10



Reconstruction performance measures

Euclidean distance reconstruction Error RMSE:

$$E = \|\boldsymbol{x}^{(i)} - \boldsymbol{\Phi}^{\dagger}(\boldsymbol{z}^{(i)})\|$$

Structural Similarity Index Measure:

$$\mathrm{SSIM}(\boldsymbol{a}, \boldsymbol{b}) = \frac{(2\mu_a\mu_b + C_1)\,(2\sigma_{ab} + C_2)}{(\mu_a^2 + \mu_b^2 + C_1)\,(\sigma_a^2 + \sigma_b^2 + C_2)}$$

μ_a, μ_b: voxel sample means

σ_a, σ_b: standard deviations

σ_{ab}: cross-covariance of local windows