

Revealing Patterns of Symptomatology in Parkinson's Disease: A Latent Space Analysis with 3D Convolutional Autoencoders

E. Delgado de las Heras, F.J. Martinez-Murcia, I.A. Illán, C. Jiménez-Mesa, D. Castillo-Barnes, J. Ramírez, J.M. Górriz
Dpt. Signal Theory, Networking and Communications. University of Granada (Spain)*

1. INTRODUCTION

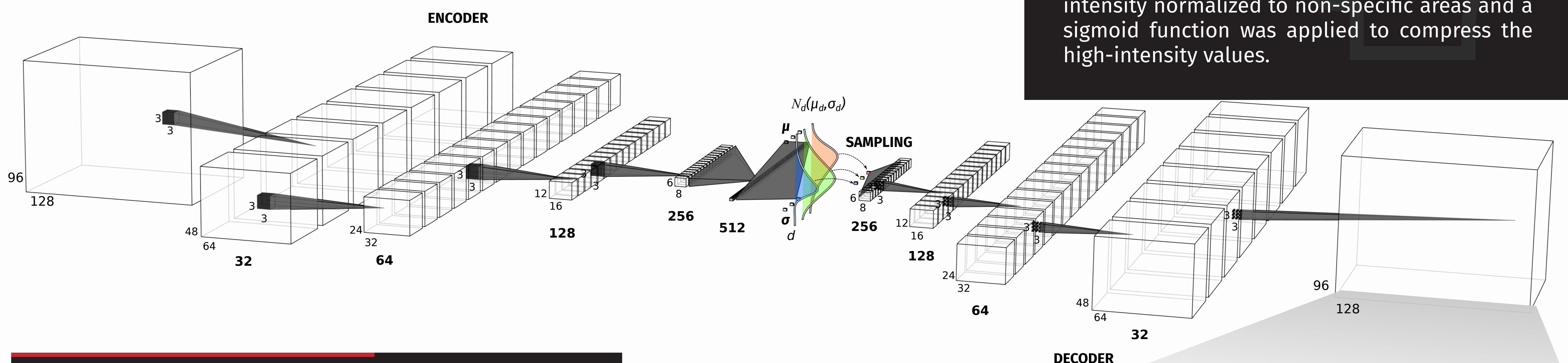
Parkinson's disease (PD) affects more than 6.2 million people worldwide. It is characterized by a loss of dopamine-producing neurons in the brain, causing tremors, rigidity, and cognitive decline among other symptoms. Ioflupane I-123 binds to the presynaptic dopamine transporters (DaTs), allowing to visualize and quantify DaT concentration at the striata,

which is key to characterize PD in SPECT imaging. Early detection and monitoring of PD progression could be tackled by modelling a self-supervised, low-dimensional representation of a FPCIT dataset, which enables us to longitudinally compare images and identify patterns of change that are indicative of neurodegeneration.

2. METHODOLOGY

This work proposes a novel approach to detect and quantify subtle changes in DaT concentration and distribution in the brain using 3D convolutional variational AEs (CVAEs). The composite variables of the latent space are then modelled using Decision Trees and XGBoost regression, and the spaces are interpreted through visualization and SHAP.

2.2 3D CONVOLUTIONAL VAE

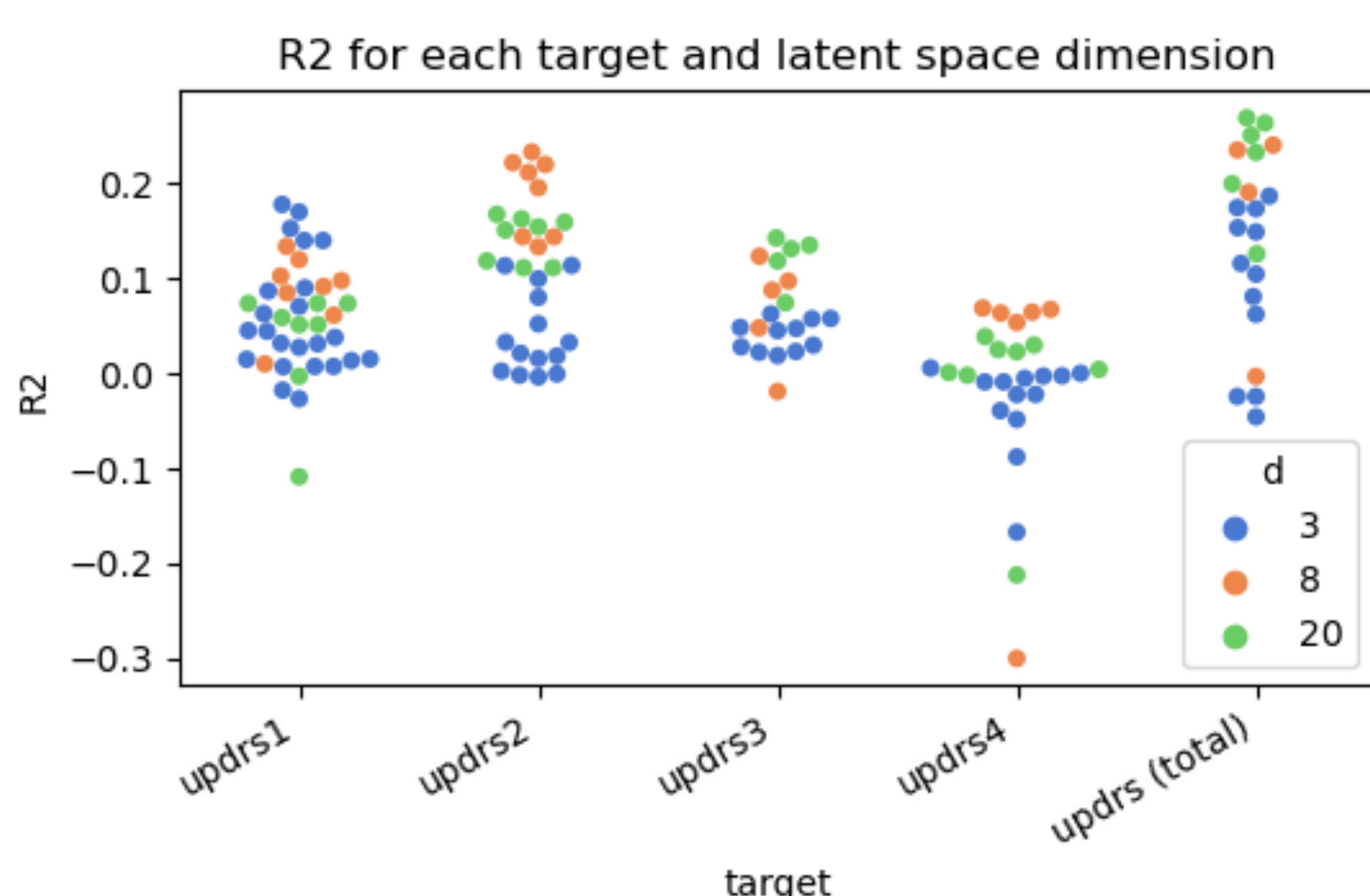


2.3. EVALUATION

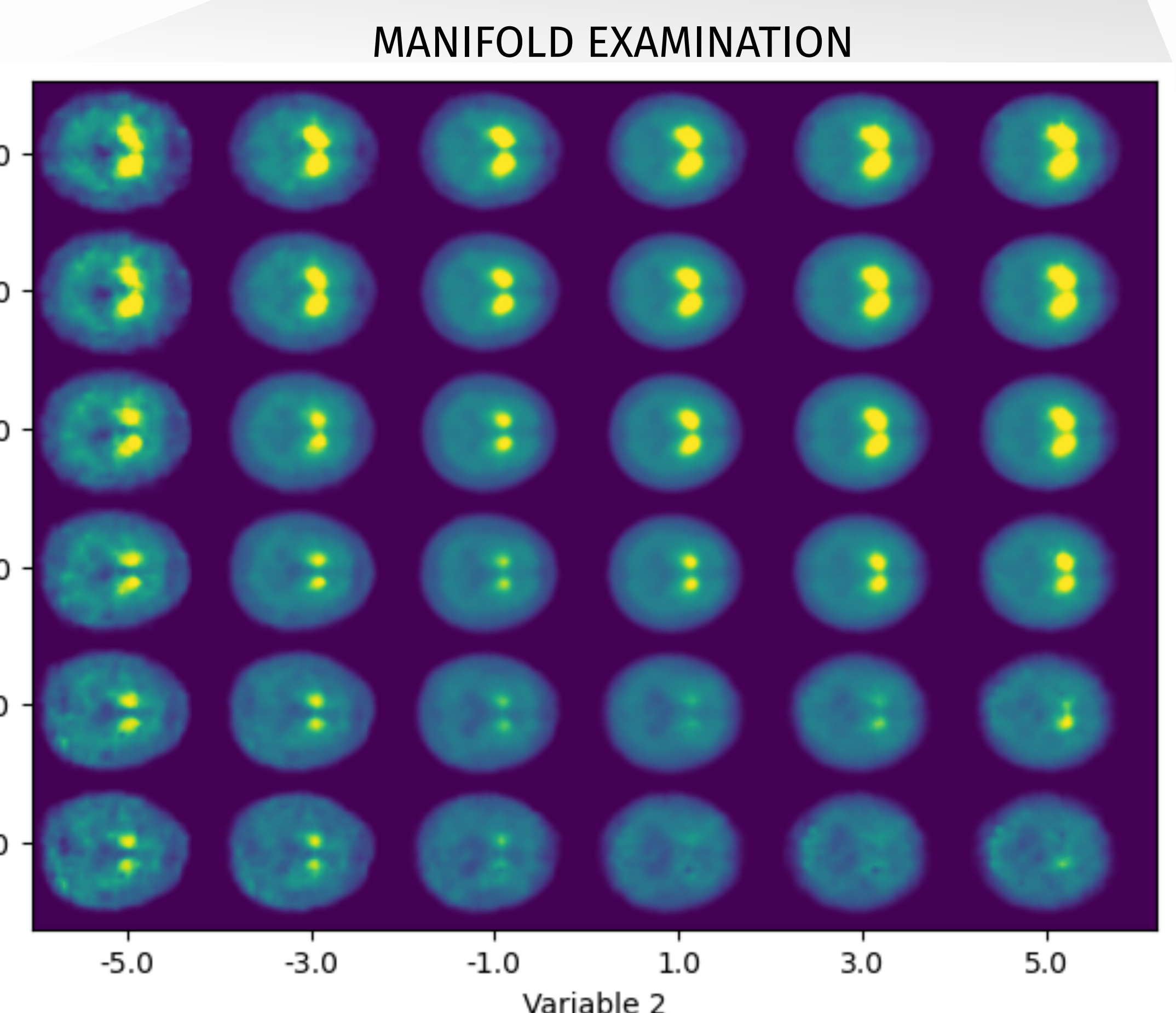
- Use the μ parameters of the latent distribution $N_d(\mu_d, \sigma_d)$ as features.
- **K-Means features** (KMF) are generated to capture important data characteristics.
- Regression using **Decision Trees** (DT) and **XGBoost**.
- Performance metrics such as **Mean Absolute Error** (MAE), **Root Mean Squared Error** (RMSE), and coefficient of determination **R2** are computed using 10-fold cross-validation.
- SHapley Additive exPlanations (SHAP) is applied to the outputs of each system.

- **R2>0** for almost all target UPDRS categories and d-dimensional latent spaces.
- Low latent dimensionality (3, 4) is generally for individual **perceived symptomatology** (UPDRS 1, 2 and 4).
- **UPDRS - total** is better accounted for by the 20-D model (MAE = 12.21, R2=0.26), pointing at higher complexity of the composite and stronger relationship with FPCIT patterns.
- **UPDRS 3**, more dependent on measured motor symptoms, benefits from the 20-D space.

3. RESULTS



Target	D	Model	MAE	RMSE	R2
updrs1	3	DT	3.53	4.5	0.15
	8	DT (KMF)	3.58	4.55	0.13
	20	DT (KMF)	3.69	4.71	0.07
updrs2	3	XGB (KMF)	3.96	5.23	0.11
	8	XGB (KMF)	3.68	4.86	0.23
	20	XGB (KMF)	3.74	5.08	0.16
updrs3	3	XGB (KMF)	10.78	13.11	0.06
	8	XGB (KMF)	10.21	12.93	0.09
	20	XGB (KMF)	10.03	12.59	0.14
updrs4	3	DT	0.7	1.44	0
	8	XGB	0.62	1.39	0.07
	20	DT (KMF)	0.71	1.43	0.02
updrs (total)	3	XGB (KMF)	13.15	16.42	0.17
	8	XGB	12.72	15.75	0.24
	20	XGB	12.21	15.51	0.26



For top performing 20-D XGB model of UPDRS (total), SHAP reveals:

- top-3 features that contribute to the output of the algorithm are latent variables 2, 10 and 0.
- Approx. linear dependency between importance to the algorithm output and values of the latent variables.
- Composition of variables 2 and 10 accounts for relevant characteristics of Ioflupane SPECT imaging: the **general intensity of the striata**, the **separation between them** and the **uptake ratio between its anterior and posterior parts**.
- Anterior-posterior uptake ratio more related to progression in the first symptoms. Average intensity for the automatic diagnosis of PD.

5. CONCLUSIONS

- The latent features of trained CVAEs are related to different aspects of the MDS-UPDRS scale with $R^2 > 0.20$.
- Best performance for UPDRS (total) with latent variables 2 and 10.
- Anterior/Posterior ratio are more related to progression of the first symptoms. Average striatal intensity for the diagnosis of PD.

This work paves the way for exploratory analysis of links between neuroimaging patterns and neurological disorders in an hypothesis-free environment.

Full paper link:

