

Anomaly Detection in the Euclid Q1 Data Release using Foundation Model Embeddings

Maxime RONCERAY¹, Marc HUERTAS-COMPANY²

Malgorzata SIUDEK², Alexandre CHANSON¹, Astroinfo2025
1 : Université de Tours - 2 : Instituto de Astrofísica de Canarias

1 - MOTIVATIONS

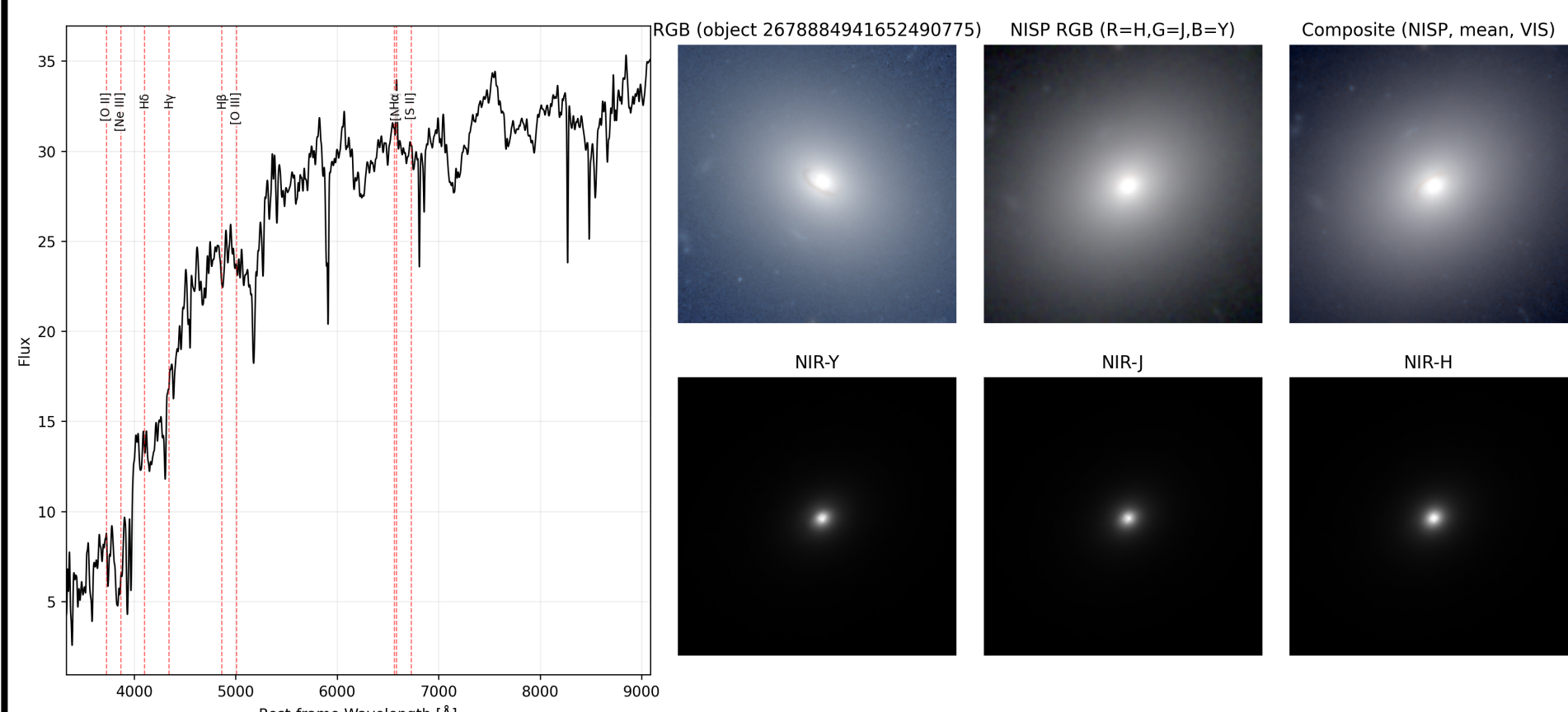
The Euclid Q1 survey provides photometric data for over 30 million sources, making automated methods essential for identifying anomalies such as strong lenses, rare morphologies, and instrumental artifacts. This project evaluates the use of foundation model embeddings (AstroPT, AstroCLIP, AION)[1,2,3] to detect anomalies in Euclid images, supported by spectral information provided by the Dark Energy Spectroscopic Instrument (DESI)[4].

OBJECTIVES :

- Characterize the latent representations by assigning anomaly scores and isolating atypical objects.
- Compare the three multimodal models to determine whether they converge toward a consistent and scientifically meaningful latent space.

EUCLID & EXPECTED ANOMALIES

- Q1 imaging: 30M sources across 63 deg²
- Tracing the last half of the cosmic history
- 40k with images + spectra (multimodal subset)



TOP : Euclid galaxy mosaics in the visible and near-infrared, highlighting morphological diversity and survey depth from the Q1 release.

LEFT : example of a typical data sample, containing one optical (VIS) image and 3 near-infrared (NISIP) images with the corresponding spectrum.

STRONG LENSES

Massive foreground galaxies bend the light of distant sources, producing arcs or multiple images. These rare systems let us measure total mass (including dark matter) and probe the expansion of the Universe.

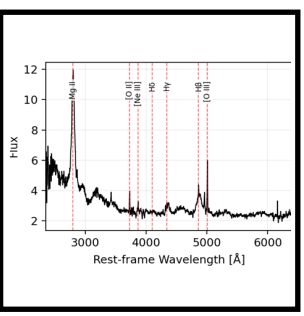
INSTRUMENTAL ANOMALIES

Some outliers come from the instrument itself, cosmic rays, diffraction spikes, saturation. They are not astrophysical, but identifying them is essential for clean data and for calibrating anomaly-detection methods.

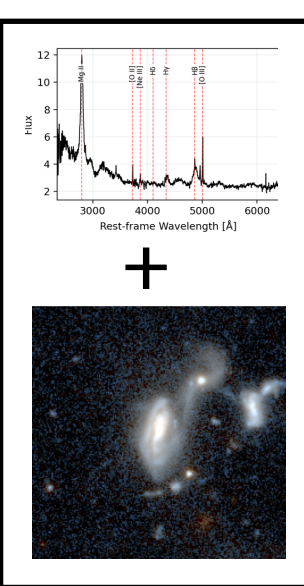


GALAXY MERGERS

when two systems interact and begin to coalesce, producing tidal tails, distortions, and other unusual morphologies.



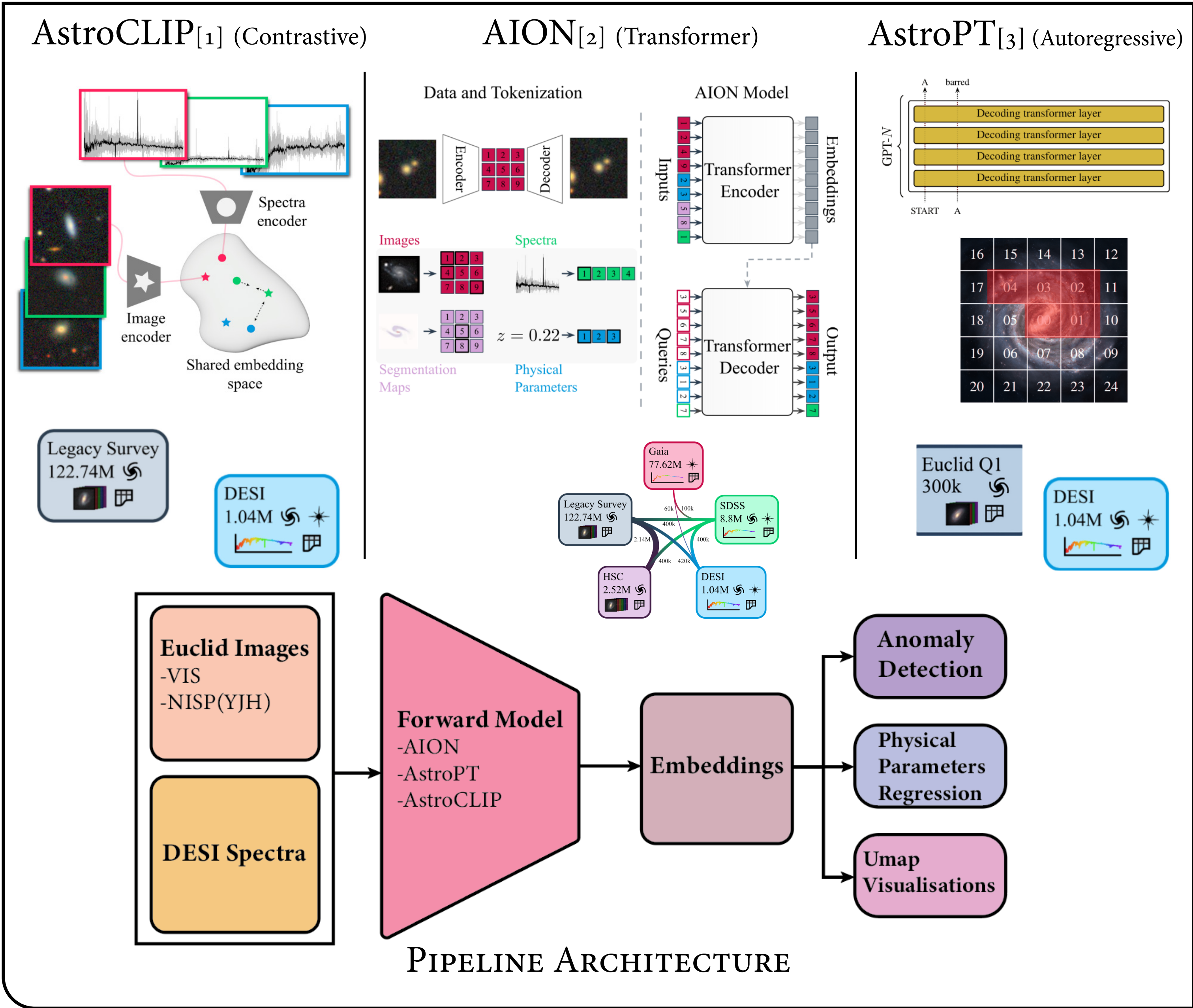
ACTIVE GALACTIC NUCLEI (AGN) galaxies whose central black hole is actively accreting matter, making them extremely bright. They signal periods of intense black-hole growth and energetic feedback.



DUAL AGN

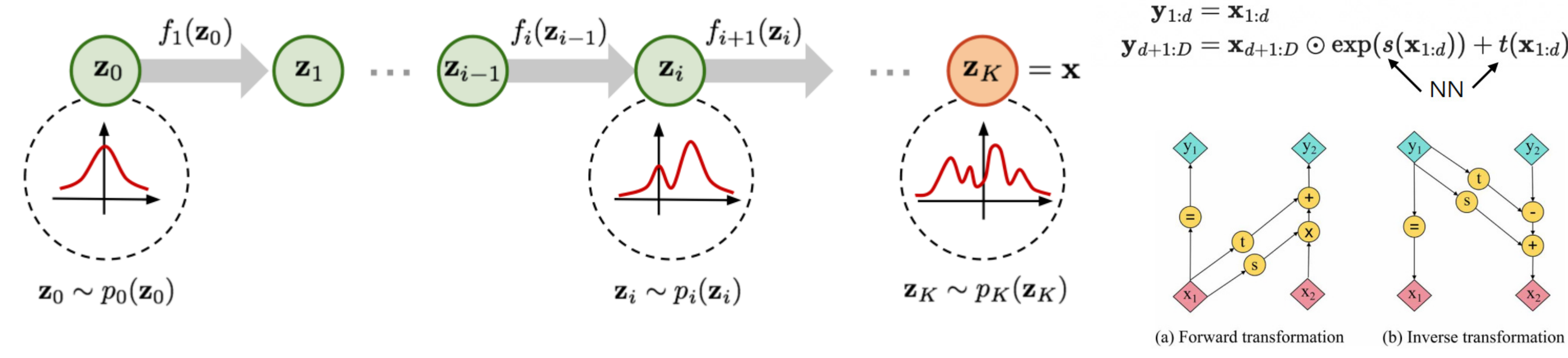
both galaxies in a merging pair host an active black hole. They are rare and scientifically valuable because they trace the early stages of black-hole binary formation and joint growth during mergers.

2 - FOUNDATION MODELS



3 - OUR METHOD

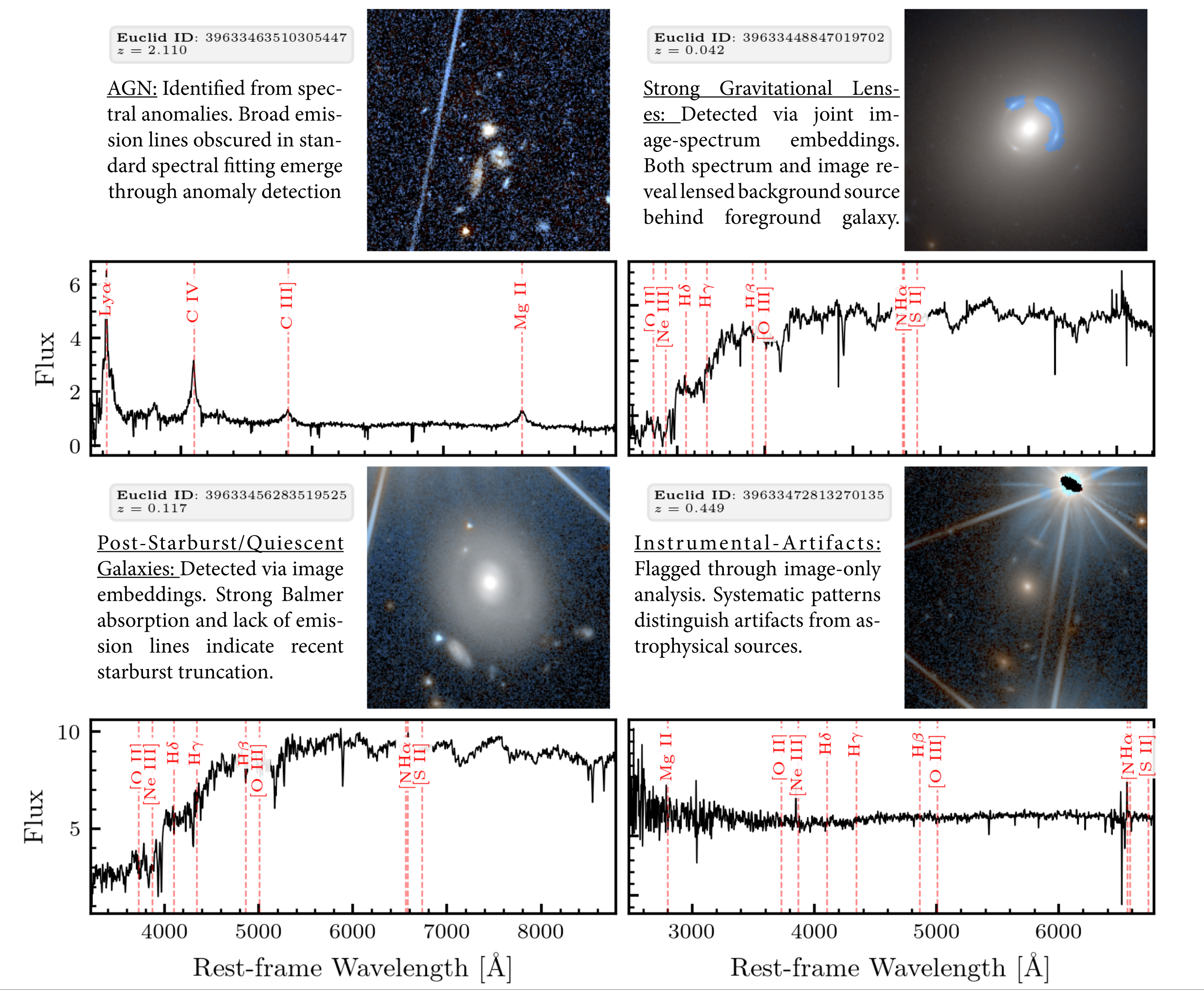
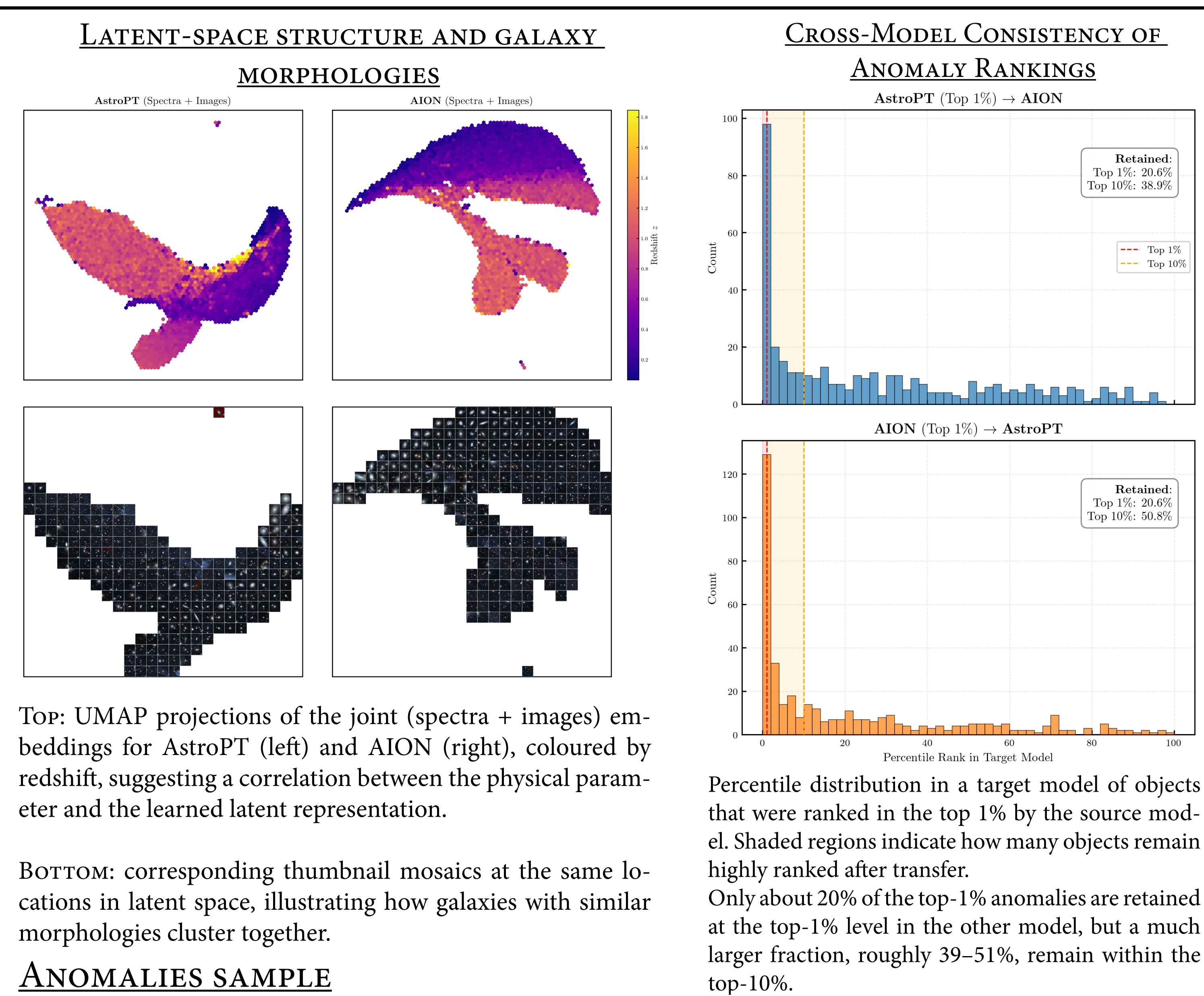
Anomaly ranking based on Normalizing Flows (RealNVP)



ANOMALY SCORE: the standardized negative log-likelihood; objects with the lowest density under the flow are ranked and flagged as anomalies.

$$a(\mathbf{x}) = - \left[\log p_0(f(\mathbf{x})) + \sum_{k=1}^K \log |\det J_{f_k}(\mathbf{x})| \right] s_{\sigma}(\mathbf{x}) = \frac{a(\mathbf{x}) - \mu}{\sigma}$$

4 - PRELIMINARY RESULTS



REFERENCES

[1] - AstroCLIP: A Cross-Modal Foundation Model for Galaxies, L. Parker, F. Lanusse et al. arXiv:2310.034
[2] - Euclid Collaboration et al. Euclid Quick Data Release (Q1) Exploring galaxy properties with a multi-modal foundation model. 2025. arXiv: 2503.15312
[3] - Liam Parker et al. AION-1: Omnimodal Foundation Model for Astronomical Sciences. 2025. arXiv: 2510.17960
[4] - DESI Collaboration et al. Data Release 1 of the Dark Energy Spectroscopic Instrument. 2025. ArXiv : 2503.14745
Misc : Part of this work was made during the Astroinfo2025 collaboration