Euclid Quick Data Release (Q1)

Exploring galaxy properties with a multi-modal foundation model

Euclid Collaboration: M. Siudek[*,1,2,3], M. Huertas-Company[3,1,4,5], M. Smith[6,7], G. Martinez-Solaeche[8], F. Lanusse[9], S. Ho[10,11,12], E. Angeloudi[3,13], P. A. C. Cunha[14,15], H. Domínguez Sánchez[16], M. Dunn[3,1,13], Y. Fu[17,18], P. Iglesias-Navarro[3,13], J. Junais[3,13], J. H. Knapen[3,13], B. Laloux[19,20], M. Mezcua[2,21], W. Roster[20], G. Stevens[22], J. Vega-Ferrero[23], N. Aghanim[24], B. Altieri[25], A. Amara[26], S. Andreon[27], N. Auricchio[28], H. Aussel[9], C. Baccigalupi[29,30,31,32], M. Baldi[33,28,34], S. Bardelli[28], P. Battaglia[28], A. Biviano[30,29], A. Bonchi[35], E. Branchini[36,37,27], M. Brescia[38,19], J. Brinchmann[15,14], S. Camera[39,40,41], G. Cañas-Herrera[42,43,17], V. Capobianco[41], C. Carbone[44], J. Carretero[45,46], S. Casas[47], F. J. Castander[2,21], M. Castellano[48], G. Castignani[28], S. Cavuoti[19,49], K. C. Chambers[50], A. Cimatti[51], C. Colodro-Conde[3], G. Congedo[52], C. J. Conselice[53], L. Conversi[54,25], Y. Copin[55], F. Courbin[56,57], H. M. Courtois[58], M. Cropper[59], A. Da Silva[60,61], H. Degaudenzi[62], G. De Lucia[30], A. M. Di Giorgio[63], J. Dinis[60,61], C. Dolding[59], H. Dole[24], F. Dubath[62], C. A. J. Duncan[53], X. Dupac[25], S. Dusini[64], S. Escoffier[65], M. Farina[63], R. Farinelli[28], F. Faustini[35,48], S. Ferriol[55], F. Finelli[28,66], S. Fotopoulou[22], M. Frailis[30], E. Franceschi[28], S. Galeotta[30], K. George[67], B. Gillis[52], C. Giocoli[28,34], J. Gracia-Carpio[20], B. R. Granett[27], A. Grazian[68], F. Grupp[20,67], S. Gwyn[69], S. V. H. Haugan[70], W. Holmes[71], I. M. Hook[72], F. Hormuth[73], A. Hornstrup[74,75], K. Jahnke[76], M. Jhabvala[77], E. Keihänen[78], S. Kermiche[65], A. Kiessling[71], B. Kubik[55], M. Kümmel[67], M. Kunz[79], H. Kurki-Suonio[80,81], Q. Le Boulc'h[82], A. M. C. Le Brun[83], D. Le Mignant[84], S. Ligori[41], P. B. Lilje[70], V. Lindholm[80,81], I. Lloro[85], G. Mainetti[82], D. Maino[86,44,87], E. Maiorano[28], O. Mansutti[30], S. Marcin[88], O. Marggraf[89], M. Martinelli[48,90], N. Martinet[84], F. Marulli[91,28,34], R. Massey[92], S. Maurogordato[93], H. J. McCracken[94], E. Medinaceli[28], S. Mei[95,96], M. Melchior[97], Y. Mellier[98,94], M. Meneghetti[28,34], E. Merlin[48], G. Meylan[99], A. Mora[100], M. Moresco[91,28], L. Moscardini[91,28,34], R. Nakajima[89], C. Neissner[101,46], S.-M. Niemi[42], J. W. Nightingale[102], C. Padilla[101], S. Paltani[62], F. Pasian[30], K. Pedersen[103], W. J. Percival[104,105,106], V. Pettorino[42], S. Pires[9], G. Polenta[35], M. Poncet[107], L. A. Popa[108], L. Pozzetti[28], F. Raison[20], A. Renzi[109,64], J. Rhodes[71], G. Riccio[19], E. Romelli[30], M. Roncarelli[28], R. Saglia[67,20], Z. Sakr[110,111,112], A. G. Sánchez[20], D. Sapone[113], B. Sartoris[67,30], J. A. Schewtschenko[52], P. Schneider[89], T. Schrabback[114], M. Scodeggio[44], A. Secroun[65], G. Seidel[76], M. Seiffert[71], S. Serrano[21,115,2], P. Simon[89], C. Sirignano[109,64], G. Sirri[34], L. Stanco[64], J. Steinwagner[20], P. Tallada-Crespí[45,46], A. N. Taylor[52], I. Tereno[60,116], S. Toft[117,118], R. Toledo-Moreo[119], F. Torradeflot[46,45], I. Tutusaus[111], L. Valenziano[28,66], J. Valiviita[80,81], T. Vassallo[67,30], G. Verdoes Kleijn[18], A. Veropalumbo[27,37,36], Y. Wang[120], J. Weller[67,20], A. Zacchei[30,29], G. Zamorani[28], F. M. Zerbi[27], I. A. Zinchenko[67], E. Zucca[28], V. Allevato[19], M. Ballardini[121,122,28], M. Bolzonella[28], E. Bozzo[62], C. Burigana[123,66], R. Cabanac[111], A. Cappi[28,93], D. Di Ferdinando[34], J. A. Escartin Vigo[20], L. Gabarra[124], J. Martín-Fleitas[100], S. Matthew[52], N. Mauri[51,34], R. B. Metcalf[91,28], A. Pezzotta[125,20], M. Pöntinen[80], C. Porciani[89], I. Risso[126], V. Scottez[98,127], M. Sereno[28,34], M. Tenti[34], M. Viel[29,30,32,31,128], M. Wiesmann[70], Y. Akrami[129,130], I. T. Andika[131,132], S. Anselmi[64,109,133], M. Archidiacono[86,87], F. Atrio-Barandela[134], C. Benoist[93], K. Benson[59], D. Bertacca[109,68,64], M. Bethermin[135], L. Bisigello[68], A. Blanchard[111], L. Blot[136,133], M. L. Brown[53], S. Bruton[137], A. Calabro[48], B. Camacho Quevedo[21,2], F. Caro[48], C. S. Carvalho[116], T. Castro[30,31,29,128], Y. Charles[84], F. Cogato[91,28], A. R. Cooray[138], O. Cucciati[28], S. Davini[37], F. De Paolis[139,140,141], G. Desprez[18], A. Díaz-Sánchez[142], J. J. Diaz[3], S. Di Domizio[36,37], J. M. Diego[16], P.-A. Duc[135], A. Enia[33,28], Y. Fang[67], A. G. Ferrari[34], P. G. Ferreira[124], A. Finoguenov[80], A. Fontana[48], A. Franco[140,139,141], K. Ganga[95], J. García-Bellido[129], T. Gasparetto[30], V. Gautard[143], E. Gaztanaga[2,21,144], F. Giacomini[34], F. Gianotti[28], G. Gozaliasl[145,80], M. Guidi[33,28], C. M. Gutierrez[146], A. Hall[52], W. G. Hartley[62], S. Hemmati[147], C. Hernández-Monteagudo[13,3], H. Hildebrandt[148], J. Hjorth[103], J. J. E. Kajava[149,150], Y. Kang[62], V. Kansal[151,152], D. Karagiannis[121,153], K. Kiiveri[78], C. C. Kirkpatrick[78], S. Kruk[25], J. Le Graet[65], L. Legrand[154,155], M. Lembo[121,122], F. Lepori[156], G. Leroy[157,92], G. F. Lesci[91,28], J. Lesgourgues[47], L. Leuzzi[91,28], T. I. Liaudat[158], A. Loureiro[159,160], J. Macias-Perez[161], G. Maggio[30], M. Magliocchetti[63], E. A. Magnier[50], F. Mannucci[162], R. Maoli[163,48], C. J. A. P. Martins[164,15], L. Maurin[24], M. Miluzio[25,165], P. Monaco[166,30,31,29], C. Moretti[32,128,30,29,31], G. Morgante[28], C. Murray[95], K. Naidoo[144], A. Navarro-Alsina[89], S. Nesseris[129], F. Passalacqua[109,64], K. Paterson[76], L. Patrizii[34], A. Pisani[65,167], D. Potter[156], S. Quai[91,28], M. Radovich[68], S. Sacquegna[139,140,141], M. Sahlén[168], D. B. Sanders[50], E. Sarpa[32,128,31], A. Schneider[156], D. Sciotti[48,90], D. Scognamiglio[71], E. Sellentin[169,17],

L. C. Smith[170], K. Tanidis[124], G. Testera[37], R. Teyssier[167], S. Tosi[36, 126], A. Troja[109, 64], M. Tucci[62], C. Valieri[34], A. Venhola[171], D. Vergani[28], G. Verza[10], P. Vielzeuf[65], N. A. Walton[170], and J. G. Sorce[172, 24]

(Affiliations can be found after the references)

March 17, 2025

ABSTRACT

Modern astronomical surveys, such as the *Euclid* mission, produce high-dimensional, multi-modal data sets that include imaging and spectroscopic information for millions of galaxies. These data serve as an ideal benchmark for large, pre-trained multi-modal models, which can leverage vast amounts of unlabelled data. In this work, we present the first exploration of *Euclid* data with `AstroPT`, an autoregressive multi-modal foundation model trained on approximately 300 000 optical and infrared *Euclid* images and spectral energy distributions (SEDs) from the first *Euclid* Quick Data Release. We compare self-supervised pre-training with baseline fully supervised training across several tasks: galaxy morphology classification; redshift estimation; similarity searches; and outlier detection. Our results show that: (a) `AstroPT` embeddings are highly informative, correlating with morphology and effectively isolating outliers; (b) including infrared data helps to isolate stars, but degrades the identification of edge-on galaxies, which are better captured by optical images; (c) simple fine-tuning of these embeddings for photometric redshift and stellar mass estimation outperforms a fully supervised approach, even when using only 1% of the training labels; and (d) incorporating SED data into `AstroPT` via a straightforward multi-modal token-chaining method improves photo-z predictions, and allow us to identify potentially more interesting anomalies (such as ringed or interacting galaxies) compared to a model pre-trained solely on imaging data.

Key words. Methods: data analysis – Surveys – Galaxies: general

1. Introduction

Astronomy's abundance of data requires appropriately scalable, automated frameworks to mine it effectively. This demand has already driven the adoption of advanced machine-learning and deep-learning techniques for a variety of purposes (see for example Domínguez Sánchez et al. 2018, Cheng et al. 2020, Cheng et al. 2021, Vega-Ferrero et al. 2021, Walmsley et al. 2022a, Walmsley et al. 2023, Huertas-Company et al. 2024, or see Huertas-Company & Lanusse 2023, Smith & Geach 2023 for general reviews). However – in most prior work – a small subset of labelled examples are used to train supervised models. This means that only a small fraction of available data is used, and that the trained models are very sensitive to 'domain drift' (Domínguez Sánchez et al. 2019) thus limiting their full deployment for scientific analysis (Huertas-Company & Lanusse 2023).

Recent works have demonstrated the transformative impact of models pre-trained on large amounts of unlabelled data (that is, 'foundation models'), which can then be fine-tuned to perform a variety of tasks. This is particularly evident for language processing, where the best-performing 'frontier' models now contain many billions of parameters and require 'web-scale' data of the order of terabytes to petabytes to be sufficiently pre-trained (e.g., Grattafiori et al. 2024). Furthermore, the recent rediscovery of neural scaling laws – empirical relationships that demonstrate how model performance improves predictably with increases in model size, data set size, or computational resources – has led to great growth in the size of deep-learning models (Cortes et al. 1993; Kaplan et al. 2020; Henighan et al. 2020).

Astronomical data sets are also growing rapidly in size and complexity, creating the need for models that can scale with data set size and perform multiple tasks effectively (Smith et al. 2024; Pan et al. 2024; Walmsley et al. 2024; Merz et al. 2024; The Multimodal Universe Collaboration 2024). The advent of the next generation of large-scale sky surveys such as *Euclid* (Laureijs et al. 2011), the Dark Energy Spectroscopic Instrument Legacy Survey (DESI; DESI Collaboration et al. 2022, 2024a), and the Vera C. Rubin Legacy Surveys of Space and Time (LSST; Ivezić et al. 2019) opens new opportunities for developing a new generation of astronomical foundation models. These data sets offer multi-wavelength imaging and spectroscopy of millions of galaxies, with relatively few labelled data, and thus provide an ideal sandbox for developing the next generation of machine-learning models. *Euclid*'s mission is especially valuable for its combination of high-resolution imaging with spectroscopy over a wide field of view, which creates a unique opportunity for advancing the next generation of large-scale machine-learning models. This synergy between cutting-edge surveys and artificial intelligence-driven methods will be instrumental in accelerating adoption of machine learning in astrophysics (Euclid Collaboration: Humphrey et al. 2023; Euclid Collaboration: Leuzzi et al. 2024; Euclid Collaboration: Aussel et al. 2024; Euclid Collaboration: Enia et al. 2024; Euclid Collaboration: Mellier et al. 2024).

Several works have started exploring large pre-trained models in astronomy (e.g., Vega-Ferrero et al. 2024). Hayat et al. (2021) and Sarmiento et al. (2021) first demonstrated the quality of the embedding spaces on Sloan Digital Sky Survey (SDSS; Gunn et al. 1998, 2006) and Mapping Nearby Galaxies at APO (MaNGA; Bundy et al. 2015) data using contrastive learning, a self-supervised technique where a model is trained to bring similar samples closer in an embedding space while pushing dissimilar ones apart (Chen et al. 2020; Radford et al. 2021). Smith et al. (2024) showed that autoregressive galaxy imaging models follow similar scaling laws to those reported for text. Complementing these imaging methods, `AstroCLIP` (Parker et al. 2024) pioneered a multi-modal strategy by aligning spectra and imaging via contrastive learning. Following a similar approach, the `PAPERCLIP` framework (Mishra-Sharma et al. 2024) explores the intersection of astronomical imaging and natural language by fine-tuning a contrastive language–image pretraining (CLIP; Radford et al. 2021) CLIP model to align observations (e.g., galaxy images) and textual descriptions. This allows users to search for images based on descriptive language, and vice versa. Zhang et al. (2024) and Rizhko & Bloom (2024) present the first attempts to incorporate time-series data with CLIP, constructing an astronomical multi-modal data set from time-series photometry data, spectra, and astrophysical metadata. These advancements illustrate the expanding frontier of multi-modal approaches, which combine data modalities to provide a more com-

---

prehensive understanding of astronomical phenomena. As just one example of a downstream use-case, researchers can use these models to gain insight into both the structural properties and the dynamical evolution of galaxies, reaching state-of-the-art performance on classification and redshift estimation without any explicit optimisations towards these tasks.

Following these developments, in this work we pre-train and use a multi-modal autoregressive foundation model (AstroPT) to explore the optical and infrared imaging components, and the spectral energy distributions (SEDs) of *Euclid* Q1 galaxies (Euclid Quick Release Q1 2025). Multi-modal models are capable of integrating and learning from different types of data (such as images, spectra, and text) simultaneously, while autoregressive models predict data sequentially by learning dependencies across inputs, such as pixel sequences in images or wavelength features in spectra or SEDs. Furthermore, autoregressive models can operate on the union of available observations without being restricted solely to cross-matched data sets, which typically represent only a small fraction of the available data. These properties make models like AstroPT well-suited to analysing the diverse, high-dimensional data produced by modern surveys. In Fig. 1 we illustrate the rapid growth of foundation models in astronomy, demonstrating an increasing trend in model size as astronomical data sets expand. With each new data release – assuming that galaxy scaling laws hold in the same way as previously found in textual models – models are projected to scale closer to the size of GPT-2 and GPT-3, large-scale transformer-based language models developed by OpenAI, representing significant advancements in natural language processing (Brown et al. 2020). Additionally, incorporating spectral data into multi-modal models significantly enhances their complexity, potentially requiring a further increase in model size.
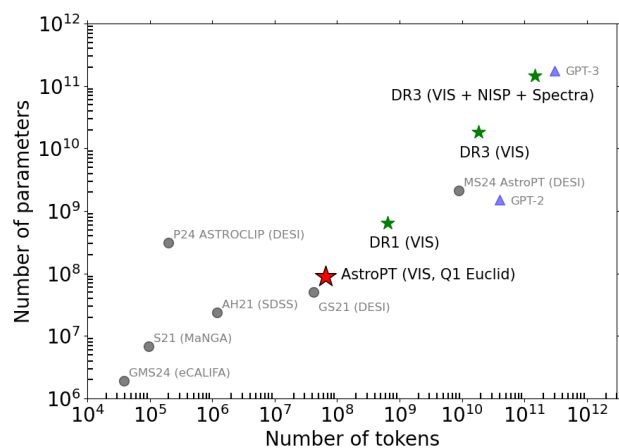

Fig. 1: Scaling laws for foundation models in astronomy compared to the GPT series. The number of tokens is defined as subword units in text (GPT models); pairs of augmented images (contrastive learning models); or a number of patches (AstroPT model). The projected scaling of future models incorporating *Euclid* data in upcoming data releases are marked with green stars. Our current model, trained on Q1 *Euclid* data, is represented by a red star. The models from the literature are referenced as follows: AH21 (Hayat et al. 2021); S21 (Sarmiento et al. 2021), GS21 (Stein et al. 2021); GMS24 (Martínez-Solaeche et al. 2024); P24 (Parker et al. 2024); and MS24 (Smith et al. 2024).

We evaluate the performance of an AstroPT model on a variety of downstream tasks – specific, practical applications of a pre-trained model – such as galaxy morphology classification, redshift prediction, or stellar mass estimations. By comparing these results to those achieved by a baseline fully supervised model, we assess the advantages of pre-training and multi-modal learning for astronomical data analysis. Unsupervised and self-supervised learning methods are particularly appealing for downstream tasks because they can generate semantically meaningful representations without requiring extensive labelled data sets. Such methods were successfully applied in astronomy for galaxy morphology studies (e.g., Hocking et al. 2018; Slijepcevic et al. 2024), estimations of photometric redshift (photo-z; e.g., Merz et al. 2024; Pattnaik et al. 2025), and physical properties (e.g., Euclid Collaboration: Bisigello et al. 2023; Chu et al. 2024; Euclid Collaboration: Enia et al. 2024; Gai et al. 2024), as well as for anomaly detection (see Fotopoulou 2024, for a review), and similarity searches (e.g., Stein et al. 2021; Walmsley et al. 2022b; Parker et al. 2024). Recent advancements in self-supervised learning have significantly closed the performance gap between unsupervised and supervised learning (e.g., Cheng et al. 2020; Mohale & Lochner 2024), making these methods viable alternatives for downstream tasks.

This paper is structured as follows. In Sect. 2, we introduce the *Euclid* mission and the Q1 sample selection. The AstroPT framework is presented in Sect. 3, together with other approaches to predict galaxy properties. In Sect. 4 we discuss the performance of AstroPT on *Euclid* imaging data, in comparison to other methods. Section 5 summarises the main results. In App. A the hyperparameters of the model are given, App. B presents the embeddings for VIS, and NISP imaging. The star–galaxy separation is discussed in App. C, and physical properties in App. D. Finally, the accuracy of morphological classification, and predictions of stellar masses and photo-z is detailed in App. E.

2. Data

The *Euclid* mission, launched in 2023, is a mission from the European Space Agency (ESA), dedicated to investigating the nature of dark energy, and dark matter by measuring the accelerated expansion of the Universe. *Euclid* primarily focuses on gathering high-resolution imaging and spectroscopic data of around 10^9 galaxies out to $z \approx 3$, and millions of galaxies out to $z \approx 6$ over $14\,000\ \mathrm{deg}^2$ (Euclid Collaboration: Mellier et al. 2024). *Euclid* provides broadband optical imaging through the visible imaging channel (VIS; Euclid Collaboration: Mc-Cracken et al. 2025), and multi-band near-infrared spectroscopy and imaging with the near-infrared spectrometer and photometer (NISP; Euclid Collaboration: Jahnke et al. 2024). The VIS instrument captures optical data in the visible range corresponding to r, i, and z filters (I_E covers $\lambda = 0.53\text{–}0.92\,\mu\mathrm{m}$; Euclid Collaboration: Cropper et al. 2024), with a spatial resolution of approximately $0.''18$ (Euclid Collaboration: Cropper et al. 2024). This high-resolution imaging provides an excellent foundation for morphological studies, allowing the extraction of detailed structural properties of galaxies. The NISP captures infrared data (Y_E, J_E, and H_E cover $\lambda = 0.95\text{–}2.02\,\mu\mathrm{m}$; Euclid Collaboration: Schirmer et al. 2022) and is designed for photometric redshift (photo-z) determination and galaxy clustering measurements, which complement the VIS data by providing additional depth and accuracy for cosmic surveys.

2.1. Q1 sample

As part of the *Euclid* mission, Q1 represents the first major public data release. Q1 provides a snapshot of early observations taken during the initial phase of the mission, covering ∼29.7 million sources over 63 deg² of the Euclid Deep Fields (EDFs). These fields are the Euclid Deep Field North, (EDF-N), the Euclid Deep Field (EDF-S), and Euclid Deep Field Fornax (EDF-F). The Q1 release includes raw and processed images (calibrated VIS, NISP frames, and background-subtracted mosaics, and segmentation maps), along with the MER photometric catalogue (Euclid Collaboration: Romelli et al. 2025). A detailed description of the Q1 data release is presented in Euclid Collaboration: Aussel et al. (2025), Euclid Collaboration: McCracken et al. (2025), Euclid Collaboration: Polenta et al. (2025), and Euclid Collaboration: Romelli et al. (2025).

The MER catalogue includes multi-wavelength photometry, combining data from *Euclid* VIS and NISP photometry with external photometric surveys spanning the ultraviolet to the near-infrared wavelengths. In particular, the catalogue includes fluxes in the I_E, H_E, J_E, and Y_E *Euclid* bands, complemented by optical bands (u, g, r, i, and z) from the Dark Energy Survey (Dark Energy Survey Collaboration et al. 2016), and the ultraviolet near-infrared optical northern survey (UNIONS; Gwyn et al. in prep.). UNIONS imaging includes CFHT u, r, HSC g, z, and Pan-STARRS i, z, and is an independent collaboration aimed at enhancing the scientific yield of large and deep surveys of the northern sky. To ensure robust cross-ground-based survey comparisons, the photometry is measured using `A-PHOT` (Merlin et al. 2019), employing fixed circular apertures with diameters of 2 full width at half maximum (FWHM) of the worst-seeing band (typically around 2″), computed individually for each source based on PSF-matched images. The bands used to construct SEDs (i.e., the collection of photometric measurements), and their corresponding fields are listed in Table 1.
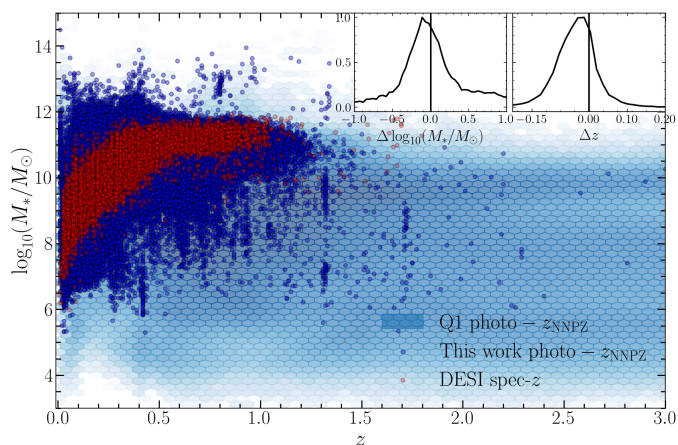

Fig. 2: Stellar mass-redshift relation for the Q1 sample (in gray), the sample used in this work (in blue), and our sample with DESI spectroscopic information (in red). The inset plots show the residuals of redshift and stellar mass differences between the *Euclid* pipeline and DESI, respectively.

This data set is further complemented by a {star, galaxy, quasar (QSO)} trinary classification, the photo-z, and galaxy physical properties estimated with the k-nearest-neighbour algorithm (NNPZ; Tanaka et al. 2018; Euclid Collaboration: Desprez et al. 2020; Euclid Collaboration: Enia et al. 2024, 2025; Euclid Collaboration: Tucci et al. 2025), among others. While these estimates are useful for certain downstream tasks (see Sect. 4.2), spectroscopic redshifts (DESI Collaboration et al. 2024b) and stellar masses (Siudek et al. 2024) from DESI are used for validation to ensure reliability. The DESI sample constitutes only a small fraction of the entire Q1 data set, and does not cover the full panoply of *Euclid* Q1 galaxies (see Fig. 2).

In this work, we select sources with a segmentation area greater than 800 pixels in the VIS data, which corresponds to galaxies that are sufficiently large to have accurate morphological classifications (Euclid Collaboration: Walmsley et al. 2025a). The *Euclid* survey operates with a pixel scale of approximately 0.″1 per pixel, corresponding to a physical resolution of approximately 0.5 kpc at $z \approx 0.5$. This criterion ensures that the sample includes only well-resolved objects, allowing for accurate morphological characterisation. We additionally restrict our sample to galaxies with $H_E < 22.5$. This magnitude limit ensures the inclusion of objects that are bright enough to be detectable in the near-infrared bands, which will be crucial for obtaining their spectra. Figure 3 shows the distribution of *Euclid* magnitudes for sources in our sample, broken down by the I_E, Y_E, J_E, and H_E bands. For the selected sample the images are complemented with the SEDs of sources covering the optical through to infrared fluxes. An example SED constructed from *Euclid* and auxiliary survey data is shown in Figure 4. This work solely builds on the imaging and photometric data, while we leave the addition of the spectra for future work.
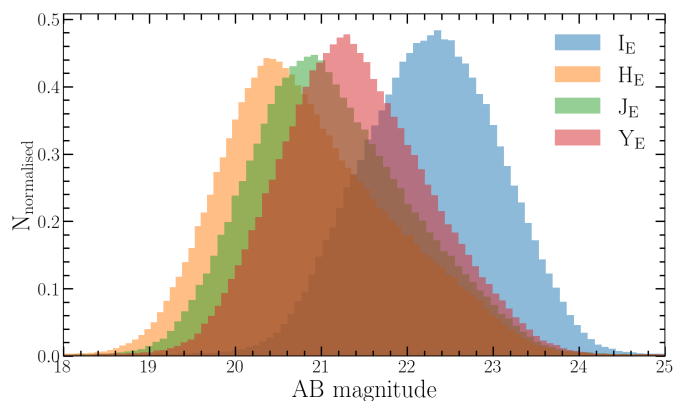
Fig. 3: Distribution of *Euclid* photometry for the sample used in this paper.

Our final set comprises approximately 330 000 galaxies spread across the Q1 fields, shown in blue in Fig. 2. We split our galaxy data set using an 80/10/10 training-test-validation split. This selection introduces constraints in both redshift and stellar mass, which may influence the representation of our data and the performance of downstream tasks. In particular, our sample is mostly limited to $z \lesssim 1$ and $\log_{10}(M_*/M_\odot) \gtrsim 7$. These cuts ensure higher-quality images for our analysis but could introduce biases, particularly in studies of low-mass and high-redshift galaxies. The spectroscopic subset from DESI further refines the selection, potentially impacting our ability to characterise the most massive systems with precise redshift information.

3. Methodology

In this section, we outline the machine learning methodologies employed to analyse the galaxy data set, starting with `AstroPT` in Sect. 3.1. In Sect. 3.2 we introduce three separate models used

Table 1: Photometric bands, instruments, observed depths, and fields included in the *Euclid* photometric catalogue used to build the SEDs of Q1 sources.

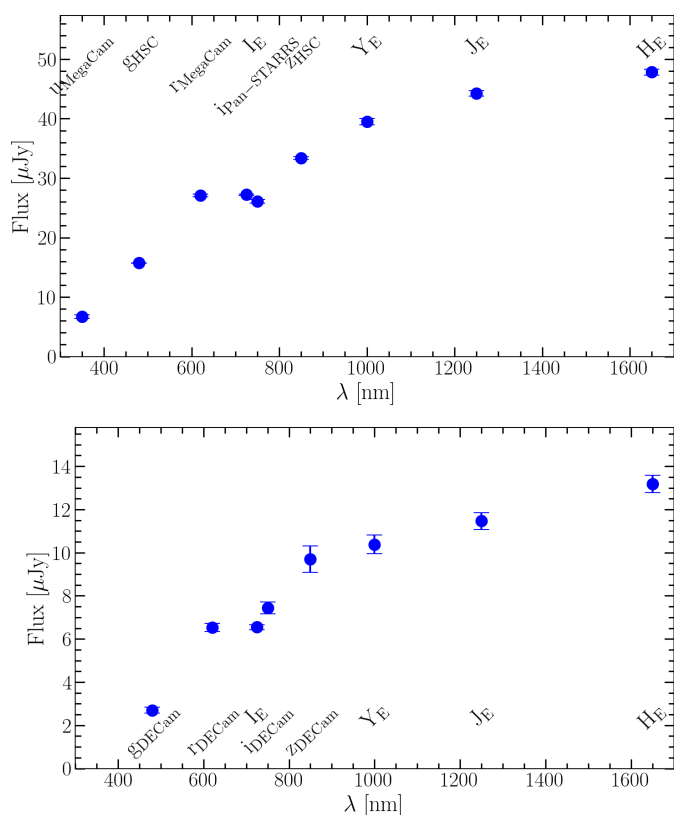| Band	Instrument	λ [nm]	Depth [mag]	Field
u	CFHT/MegaCam	372	23.5	EDF-N
g	HSC	480	25.3	EDF-N
r	CFHT/MegaCam	640	24.1	EDF-N
i	Pan-STARRS	755	23.3	EDF-N
z	HSC	891	23.5	EDF-N
g	DECam	473	24.6	EDF-S, EDF-F
r	DECam	642	24.3	EDF-S, EDF-F
i	DECam	784	23.7	EDF-S, EDF-F
z	DECam	926	22.9	EDF-S, EDF-F
I_{E}	VIS	715	24.7	EDF-N, EDF-S, EDF-F
Y_{E}	NISP	1085	23.1	EDF-N, EDF-S, EDF-F
J_{E}	NISP	1375	23.2	EDF-N, EDF-S, EDF-F
H_{E}	NISP	1773	23.2	EDF-N, EDF-S, EDF-F

Fig. 4: Representative SEDs of galaxies from the Q1 EDF-N (top) and EDF-S fields (bottom). The SED includes multi-wavelength photometric data spanning the ultraviolet to the near-infrared listed in Table 1.

to evaluate the performance of different techniques in downstream tasks (Sect. 4.2).

3.1. AstroPT

`AstroPT`[1] is an autoregressive transformer designed to extract scientifically useful embeddings from high-resolution galaxy images. A visualisation of the `AstroPT` architecture is shown

---

[1] https://github.com/Smith42/astroPT

in the right panel in Fig. 5. Due to space constraints, we do not go into detail about the transformer architecture in this paper. We instead recommend Sect. 4(d) of Smith & Geach (2023) for an astronomy-friendly introduction of the transformer model, and Sects. 7 and 9 of the same publication for a further discussion of the ideas behind `AstroPT` and other foundation models.

`AstroPT` is initially trained on the task of predicting the next segment of a stream of astronomical data. This training process incentivises the model to learn a compact representation of the data, which captures the essential patterns and relationships inherent in the astronomical observations it is pre-trained on. Our data set consists of 224×224 pixel FITS cutouts in the VIS and NISP bands fixed to the same angular size from mosaics, i.e., MER background-subtracted calibrated frames. The cutout size is an arbitrary choice, corresponding to a commonly used input size for deep-learning models due to their compatibility with standard convolutional neural network architectures and `AstroCLIP` (Parker et al. 2024). The fixed size allows for efficient training on GPUs while maintaining a balance between spatial resolution and computational cost. However, this choice has implications for the representation of galaxy morphology: (i) smaller galaxies may be dominated by background noise or multiple objects might be in the frame; and (ii) larger galaxies might not fit entirely within the cutout, leading to truncation of their outer regions. These effects introduce potential biases into the embeddings, as their descriptive power may depend on the relative size of the galaxy within the cutout. For images, the training process involves optimising a Huber loss (Huber 1964) objective, with the model attempting to predict the next 16×16 pixel patch in a sequence of galaxy image patches (as illustrated in the left panel of Fig. 5). In this study, we use the same model architecture as the 90 million parameter model described in Smith et al. (2024). To ensure consistency across inputs, each image patch is normalised by taking its z-score, defined as

$$z = \frac{x - \mu}{\sigma}, \tag{1}$$

where x is the pixel value, μ, is the mean and σ is the standard deviation of the pixel values across the patch. This normalisation centres the data around zero and scales it to have unit variance, improving stability during pre-training (e.g., El-Nouby et al. 2024). However, we acknowledge that global normalisation may have an impact on information, like a contrast reduction in faint structures, particularly in the outskirts of galaxies. The verification of alternative normalisation strategies, such as local

normalisation or 'min-max' scaling, is left to future work. To incorporate SEDs, we use a simple token-chaining mechanism. This process involves concatenating the SED data as a token sequence, to be fed in after the imaging tokens, thus allowing the model to process both modalities (imaging, and SED information) simultaneously. The SEDs are divided into patches of a length of one, and the SEDs are normalised using the z-score over the entire SED sequence. While the z-score normalisation preserves the relative shape of SEDs, it could, in principle, reduce the distinctiveness of extreme SEDs (e.g., very blue or very red galaxies). Testing alternative approaches, such as converting fluxes to magnitudes before normalisation, is left for future work. We use separate encoder and decoder layers for the images and SEDs. The encoders take as input the raw image or SED patch data, and process these data so that they produce embedding patches that are of a fixed size. These embedding patches are then passed into the model 'trunk', shown on the right of Fig. 5. After the model trunk processes these patches, the output is then projected back into a prediction of the next image or SED patch by the decoder head, which consists of the final layers of the model.

In our case, `AstroPT` was pre-trained three times: firstly on solely optical (VIS) images; secondly, on optical and NISP (VIS+NISP) images; and thirdly on a chained modality stream of optical and NISP images and pre-calculated SEDs (VIS+NISP+SED). A 90-million parameter model was used in all cases due to its efficiency, as larger architectures did not show significant performance improvements. We state our full hyper-parameters in Appendix A.

3.2. Models and evaluation metrics

We investigate three methodologies for predicting galaxy properties, utilising supervised and self-supervised techniques. These approaches leverage the same data set, and are applied to both regression tasks (such as photo-z and stellar mass estimation) and classification tasks (such as morphological classification of late/early-type galaxies).

1. Supervised convolutional neural network (CNN): This approach involves training a CNN (Fukushima 1980) using the same cutouts as used for `AstroPT`, ensuring consistency in the input data. The implementation of SEDs is beyond the scope of this paper, and so is left for future work. Our CNN consists of convolutional layers followed by dense layers, which are trained end-to-end using the mean squared error loss (MSE) function for regression tasks, and the binary cross-entropy (BCE) loss function for classification tasks. This method is a fully supervised approach where the network relies entirely on labelled data sets to learn the mapping from input features to output predictions.

2. Linear model: Here we evaluate the quality of our representations derived from the `AstroPT` framework via a simple linear or logistic regression model. This is a minimalistic and interpretable approach, where we train our regressor directly on the embeddings generated by the `AstroPT` framework. This approach leverages the representations learned by the self-supervised model and evaluates their utility without additional modifications to the embeddings. For regression tasks, the linear regression model outputs continuous values, while for classification tasks, the logistic regression model outputs class probabilities.

3. MLP model: This method also uses self-supervised embeddings from the `AstroPT` model as fixed input features.

A multi-layer perceptron (MLP; Rosenblatt 1958) is then trained on labelled data for either regression or classification tasks. The MLP consists of four fully connected layers. To ensure stable, efficient training, batch normalisation is applied after each layer and dropout regularisation (with a 30% probability) is used to prevent overfitting. We use the leaky rectified linear unit (ReLU; Nair & Hinton 2010; Xu et al. 2015) as our activation function. The model is trained using the AdamW optimiser (Loshchilov & Hutter 2017) with weight decay, and the Huber(Huber 1964) or BCE loss is used as the loss function for regression or classification tasks, respectively.

All source code and full implementation details for our downstream tasks are available in the `AstroPT` GitHub repository[2], which provides further insight into the underlying methods. Table 2 summarises the key details about the methodologies used for downstream tasks.

The morphological labels for all our sample come from the morphology catalogue(GZ; Euclid Collaboration: Walmsley et al. 2025b), for $\approx$ 5% of our sample the stellar mass labels, and the spectroscopic redshift are taken from from DESI Early Data Release (DESI Collaboration et al. 2024b), and associated Value Added Catalogue of physical properties of DESI galaxies (Siudek et al. 2024).

To assess the robustness of the models, a Monte Carlo simulation is conducted with 50 (reduced to 10 for supervised CNN due to time limitations) runs for different percentage of data used for training, each time evaluating key metrics given below. For evaluating our models' performance, we use different sets of metrics depending on whether the tasks are regression or classification.

For predicting continuous values, such as photo-z, or stellar mass estimates, we rely on the following metrics commonly used in regression tasks

– Bias (Δx): Indicates systematic offsets in predictions. We apply the normalisation to account for the scale dependence of redshift errors that scale with $1 + x_{\text{true}}$, we keep it as well for stellar mass estimates:

$$\Delta x = \frac{x_{\text{pred}} - x_{\text{true}}}{1 + x_{\text{true}}}, \qquad (2)$$

where x_{true} and x_{pred} correspond to the true and predicted values, respectively.

– Normalised median absolute deviation (NMAD): A robust statistic to measure scatter that is less sensitive to outliers. It is defined as

$$\text{NMAD} = 1.4826 \times \text{median}\left(|\Delta x - \text{median}(\Delta x)|\right). \qquad (3)$$

– Outlier fraction: The fraction of predictions with $|\Delta x|$ exceeding 0.15, and 0.25 for photo-z, and stellar mass, respectively, quantifying catastrophic prediction errors:

$$\text{Outlier fraction} = \frac{\text{Number of galaxies with } |\Delta x| > 0.15(0.25)}{N}. \qquad (4)$$

The choice of the threshold of 0.15 for photo-z is commonly used to define catastrophic outliers. For stellar mass, while no strict convention exists, a threshold of 0.25 corresponds to standard stellar mass error (Siudek et al. 2024).

---

[2] https://github.com/Smith42/astroPT/tree/euclid_scripts/scripts/euclid
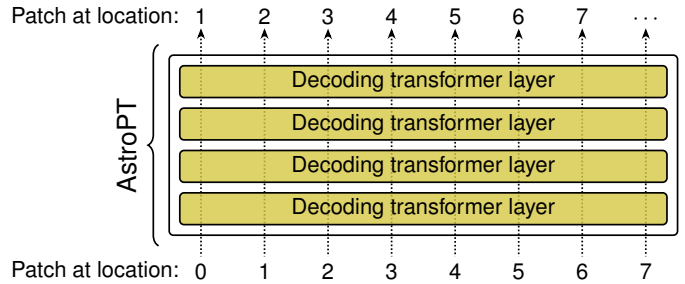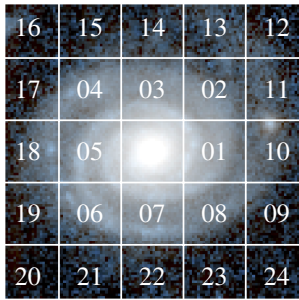
(a) The sequence order of galaxy image patches used in this work.

(b) A schematic of `AstroPT`'s architecture.

Fig. 5: `AstroPT's` training objective is predicting the next patch in a sequence of galaxy image (and other modalities') patches. On the left we show the order of galaxy patches used in this work, and on the right we show how each patch is processed by `AstroPT`.

Table 2: Comparison of different methodologies for regression and classification downstream tasks.

Model	Loss function	Key features
Supervised CNN	MSE / BCE	Non-linear, deep feature extraction; trained on labelled data.
Linear Model	MSE / Log Loss	Simple regression, assumes linear relationships, trained on embeddings.
MLP Model	Huber / BCE	Deep architecture, trained on embeddings.

For binary classification tasks such as morphological classification (such as early-type versus late-type galaxies), we use the following metrics.

– Accuracy: The proportion of correct predictions (both true positives and true negatives) among all predictions:

$$\text{Accuracy} = \frac{N_{\text{TP}} + N_{\text{TN}}}{N_{\text{TP}} + N_{\text{FP}} + N_{\text{TN}} + N_{\text{FN}}}, \tag{5}$$

where N_{TP}, N_{TN}, N_{FP}, and N_{FN} are the number of true positives, true negatives, false positives, and false negatives, respectively.

– Precision: The proportion of true positives among all instances predicted as positive:

$$\text{Precision} = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FP}}}. \tag{6}$$

– Recall (true positive rate, TPR): The proportion of true positives among all actual positive instances:

$$\text{Recall} = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FN}}}. \tag{7}$$

– False positive rate (FPR): The proportion of false positives among all actual negatives:

$$\text{FPR} = \frac{N_{\text{FP}}}{N_{\text{FP}} + N_{\text{TN}}}. \tag{8}$$

– F1 score: The harmonic mean of precision and recall, balancing the trade-off between them:

$$F1 = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}. \tag{9}$$

4. Results

In this work, we explore the latent representations of galaxies derived from our pre-trained `AstroPT` models applied to three distinct data sets. The complementary nature of VIS, VIS+NISP, and VIS+NISP+SED data facilitates a systematic evaluation of the information content in the latent spaces derived from our models. This comparison also allows us to assess how well each wavelength constrains a galaxy's physical parameters and to understand the limitations of relying solely on optical imaging.

4.1. `AstroPT`: Quality of self-supervised embeddings

In this section, we assess the quality of self-supervised embeddings obtained from the *Euclid* VIS, VIS+NISP, and VIS+NISP+SED data via `AstroPT`. We do this by analysing their correlation with source properties. In all downstream tasks (including anomaly and similarity search, see Sect. 4.2), we use the raw 768-dimensional neural network embeddings as input. For visualisation purposes, we project this high-dimensional embedding space onto two dimensions using uniform manifold approximation and projection (UMAP; McInnes et al. 2018). We initialise UMAP with a consistent set of hyperparameters for all visualisations (`n_neighbours = 15`, `min_dist = 0.1`, `n_components = 2`, `random_state = 42`) to ensure consistency across tasks. We validate different UMAP parametrisations, to: (i) determine whether the randomness in UMAP initialisation impacts the visualisations; (ii) to emphasize local structures that can potentially create tighter clusters; and (iii) to highlight the global relationships. While the shapes of the projections varied slightly across these settings, the trends related to galaxy morphological and physical properties remained consistent. Thus, we stick to the default values that provide a good trade-off between the local, and global structures. The UMAP projection facilitates an intuitive exploration of clustering patterns and structures within the data. UMAP is particularly effective in preserving local and global structures, often outperforming methods such as principal component analysis (PCA) in re-

vealing underlying relationships in the data (e.g., McInnes et al. 2018). However, we note that the dimensionality reduction algorithm ultimately influences the patterns and separations observed in these visualisations. While the UMAP projections suggest that relevant morphological and physical information is encoded in the `AstroPT` embeddings, the degree of separation achieved depends on the algorithm's ability to preserve the global structure.

To more directly evaluate the utility of these embeddings, we also examine their performance in downstream tasks through approaches like the linear and MLP models (see Sect. 4.2). These analyses provide a complementary and quantitative assessment of the embeddings' capacity to capture source properties, independent of the visualisation technique used.

A visualisation of the self-supervised embeddings trained on the VIS+NISP+SEDs images with example cutouts, and SEDs is shown in Fig. 6. The figure is created by binning the embedding space into a 50×50 grid and randomly selecting a sample from each cell, plotting its corresponding colour-mapped galaxy image at that location. The images are colour-mapped based on the VIS and NISP band images. The RGB image is generated only for visualisation purposes by averaging the $Y_{\rm E}$-, $J_{\rm E}$-, and $H_{\rm E}$-band images to form an NISP channel, combining the VIS channel with this NISP channel to form a composite, and then stacking these three channels: VIS, NISP, and the composite. To preserve luminance, the luminosity channel of the resulting RGB image is replaced with the VIS channel, maintaining the brightness information while allowing for distinct colour contrasts. Zoom-in panels around the edges show that galaxies with similar morphologies tend to be clustered in different areas allowing for a separation of edge-on (left zoom-in panel), elliptical (top zoom-in panel), and spiral (bottom zoom-in panel) galaxies, while stars and low-quality images are separated in distinct cluster (right zoom-in panel). The zoom-in panels of spiral, and elliptical galaxies are selected within a photo-z range of 0.5 to 0.7, ensuring a fair comparison of their SEDs. While no striking differences are observed, the SEDs appear to broadly reflect morphological trends. Specifically, edge-on disc galaxies and ellipticals tend to be redder. However, this distinction is subtle and subject to interpretation. Additionally, the SEDs of stars and sources with lower data quality show spurious features. A similar plot, but generated using the `AstroPT` embeddings trained on VIS, and VIS+NISP data, is provided in Appendix B.

The visual similarity of sources located close to each other in the representative space is also confirmed by their properties. The outlier regions are clearly separated from the main representations and are mostly attributed to bad-quality imaging and stars. Figure 7 shows the relationship between the embeddings trained on VIS images only, on VIS+NISP images, and VIS+NISP+SED, and the detection quality flag. The detection quality flag (`DET_QUALITY_FLAG`) is a flag indicating specific issues with sources, such as contamination, blending, or proximity to bright stars (Merlin et al. 2015, 2016, 2019; Kümmel et al. 2022). The UMAP projections reveal significant structure within the embedding space, reflecting the quality of the data: sources with high-quality images (`DET_QUALITY_FLAG < 4`) cluster tightly, whereas sources with contamination (e.g., blending or proximity to bright stars) form distinct outlier clusters. Including NISP data enhances the separation of poorly flagged sources, suggesting that NIR data provides complementary morphological information compared to VIS data. This improvement may arise because NIR observations are less affected by dust extinction and better trace the underlying stellar mass distribution, which can help distinguish extended or blended sources that are challenging to classify in VIS alone. In particular, embed-

dings trained on VIS+NISP images reveal two primary branches: one dominated by sources flagged as unknown (i.e. not fitting galaxy, star, or QSO templates); and another branch associated with sources of lower quality. However, the separation between galaxies, stars, and QSOs remains less defined (see details in Appendix C). When SED information is incorporated into the training (VIS+NISP+SED), the embeddings show a bimodal structure, which is driven by the spatial distribution: the EDF-N and EDF-S are targeted by different telescopes (see Sect. 2). Besides these two main branches, stars are also more distinctly separated, benefiting from the additional spectral information provided by SEDs. Furthermore, the unknown sources form a more coherent and distinguishable cluster, indicating that the combination of imaging and SED data aids in isolating these objects from other classes. This highlights the potential of SED data to complement imaging in disentangling complex or ambiguous cases, such as blended sources or objects with uncertain classifications.

We also examined the correlation of the self-supervised embeddings with morphological parameters derived from the GZ catalogue (Euclid Collaboration: Walmsley et al. 2025b). Early-type galaxies typically exhibit smooth profiles, while late-type galaxies tend to show distinct features such as spiral arms or a bulging core, contributing to a more complex, less smooth morphology. Although not identical, the GZ classification between smooth and featured galaxies provided with the *Euclid* Q1 release (Euclid Collaboration: Walmsley et al. 2025b) closely approximates an early-type/late-type separation. The UMAP visualisations reveal distinct clustering patterns that correspond to an early-type/late-type separation and the presence of spiral arms as shown in Fig. 8. Across all embeddings – whether `AstroPT` is trained on VIS, VIS+NISP, or VIS+NISP+SED – specific regions of the latent space are consistently associated with early-type galaxies, which also exhibit a low probability of hosting spiral arms. Conversely, other areas correspond to late-type galaxies with pronounced spiral structures. While the gradients within VIS+NISP+SED embeddings are clear, the separations between morphological types become harder to identify. This suggests that the embeddings prioritise spectral diversity over morphological clarity when SED data are included, highlighting a trade-off between the abstraction of spectral information and the preservation of morphological distinctions. This highlights the ability of automatic separation of elliptical and star-forming galaxy populations.

Interestingly, edge-on galaxies form a clearly identifiable region in the embeddings trained on VIS images alone, highlighting the capability of `AstroPT` to capture these specific morphological features (see the left panel in Fig. 9). This morphological distinctiveness aligns with their SED properties: a rising flux toward longer wavelengths, indicative of more dust, supports expectations for edge-on star-forming disk galaxies. However, as additional data from the NISP bands, and SEDs are incorporated, this distinct separation becomes less pronounced in the VIS+NISP and VIS+NISP+SED embeddings, as seen in the middle, and right panels of Fig. 9. While the inclusion of NISP and SED data enriches the latent space by providing additional spectral information, it does not necessarily enhance the clarity of morphological separations. Instead, the embeddings appear to encode a broader range of physical and morphological characteristics, becoming more abstract as additional data are integrated. The reduction in morphological separation, particularly for edge-on galaxies, may also be attributed to the lower resolution of NISP images relative to VIS, which limits the preservation of finer structural details. This does not imply that `AstroPT`'s performance degrades with lower-resolution
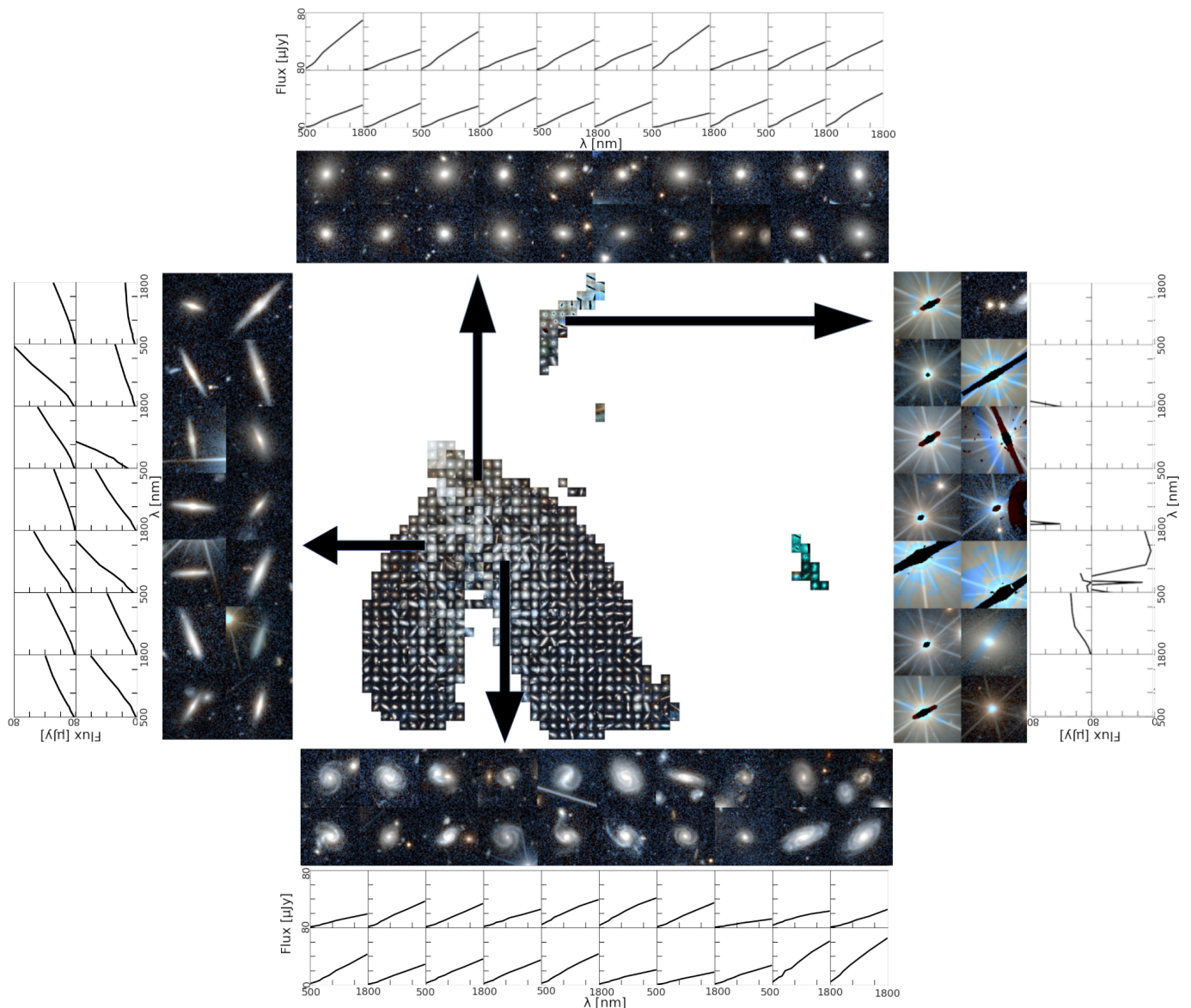
Fig. 6: UMAP visualisations of the embeddings from `AstroPT` trained on VIS+NISP+SEDs with example cutouts and SEDs, see Sect. 4.1 for more details.

data but instead highlights the trade-offs in representation fidelity when combining data sets of varying resolutions. These findings demonstrate the adaptability of `AstroPT` to encode rich, multidimensional information, albeit at the cost of losing some morphological specificity in certain contexts.

Similarly, we find that physical properties are correlated with the latent space and morphological parameters. Namely, the region associated with spiral galaxies is characterised by lower stellar masses, while the space occupied by elliptical galaxies tends to have higher masses. There is though no trend with the *Euclid* spectroscopic redshift and only a subtle trend of higher redshift for more massive galaxies using photo-zs. More details are given in Appendix D.

These visualisations confirm that the self-supervised embeddings capture meaningful physical and morphological properties, with improved performance when both VIS and NISP data are utilised, except for the shape elongation (i.e., edge-on, ellipticity) which are caught by VIS embeddings but blended when using VIS+NISP imaging. The latent space's ability to distinguish between low- and high-quality data, as well as between source types, underscores the robustness of the `AstroPT` framework for image analysis.

4.2. Downstream tasks

In our work, we evaluate the utility of self-supervised representations for different downstream tasks: morphology classification (see Sect. 4.2.1), photo-z estimations (see Sect. 4.2.2), and stellar mass predictions (see Sect. 4.2.3).

4.2.1. Morphological properties

The performance of three different methods – a supervised CNN, a linear model, and an MLP model (see Sect. 3.2)–is evaluated for their ability to classify galaxy as early-type or late-type based on the morphological catalogue (see Euclid Collaboration: Walmsley et al. 2025b). We define galaxies as early-type if their smooth probability exceeds 0.5 and as late-type otherwise. How-
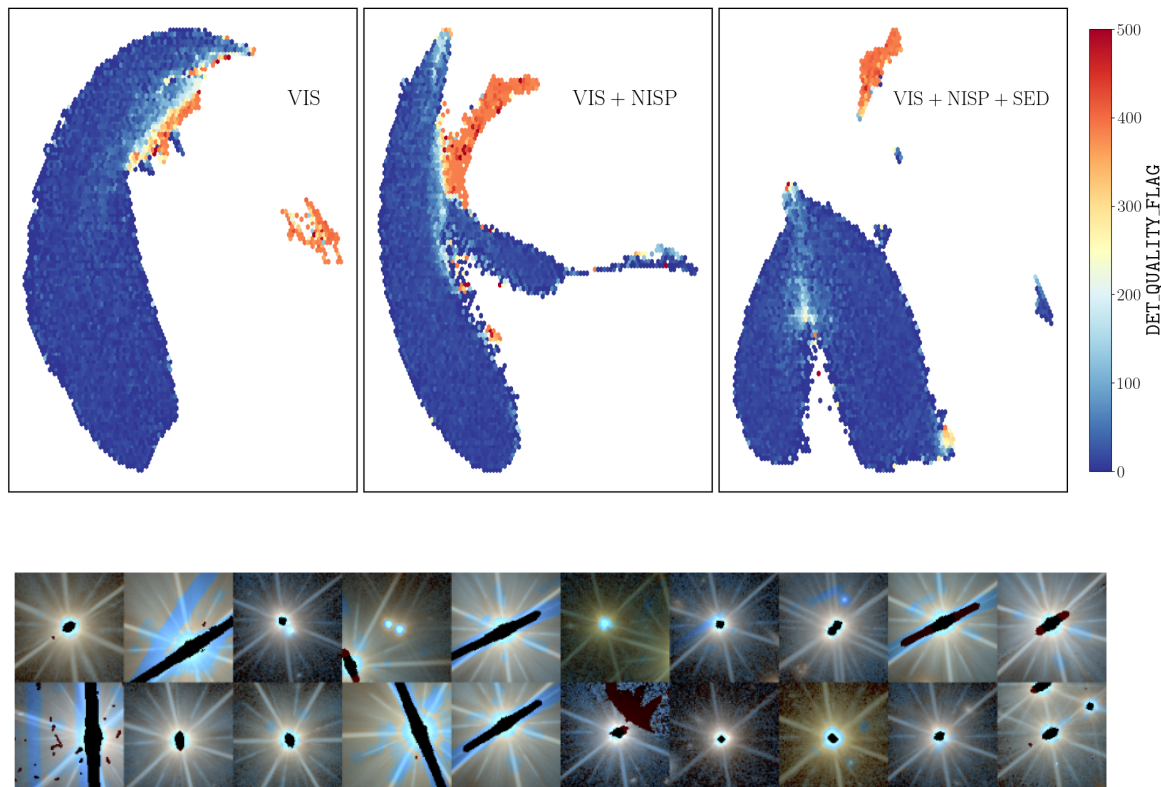
Fig. 7: *Top*: UMAP visualisations of the self-supervised embeddings from `AstroPT` trained on VIS (left), VIS+NISP (centre), and VIS+NISP+SED (right) images coloured by the detection quality flag. This flag allows to identify sources with contaminants, e.g., proximity star (`DET_QUALITY_FLAG` ≥ 4). *Bottom*: Example VIS+NISP cutouts of the outliers in the representative space.

ever, this threshold-based classification may introduce inaccuracies, as it relies on predicted continuous probabilities rather than an inherent binary distinction. The accuracy, precision, and recall as a function of the percentage of labelled data are shown in Fig. 10. Table 3 shows a statistical summary for models trained on 100% of the labels, while Table E.1 extends this comparison across all percentages of labelled data. As expected, the results suggest that performance varies with the percentage of the labelled data used for training, but this dependence is neither uniform nor consistent across methods. In particular, the linear and MLP models present relatively stable performance, with only minor fluctuations, suggesting that even modest amounts of labelled data are sufficient for robust classification. However, the supervised CNN shows a more pronounced dependence on labelled data, with a clear decline in accuracy, precision, and recall as the labelled data set size decreases. Interestingly, incorporating NISP imaging does not enhance classification performance; in some cases, it even slightly lowers accuracy, precision, and recall, possibly due to the NISP spatial resolution (see also discussion about edge-on galaxies in Sect. 4.1) or due to z-score normalisation (see Sect. 3). Similarly, adding SEDs does not improve classification, likely because the SEDs contain little direct morphological information. These findings suggest that while increasing the number of labelled examples provides some benefits, a self-supervised pre-training dramatically reduces the need for labelled samples. The self-supervised embeddings, trained on a relatively small labelled sample, already provide a robust foundation for morphological classification. Notably, the MLP model

consistently achieves the highest performance across all metrics, particularly in distinguishing between early-type and late-type galaxies.

4.2.2. Photometric redshifts

We verify the performance of different methods in predicting photo-z using the metrics described in Sect. 3.2. The spec-zs from DESI, available for about 5% of our sample, are assumed to be the true redshifts. The accuracy of *Euclid* photo-z derived with `NNPZ` as part of the Q1 release is shown in the top panel in Fig. 11. Figure 12 shows the bias, NMAD, and outlier fraction as a function of the percentage of labels used for training for photo-z obtained through these methods. A statistical summary is provided in Table 4 for a 100% of labels used for training, for the full table see Table E.2. The results highlight the impact of adding NISP imaging data, and SED information, and the effectiveness of different training approaches. The inclusion of NISP imaging improves photo-z accuracy, reducing metrics to levels comparable to those achieved by the official *Euclid* pipeline coming from the NNPZ template fitting algorithm (Euclid Collaboration: Tucci et al. 2025). These improvements are consistent with previous findings, such as those by Salvato et al. (2019), who demonstrated the importance of including NIR bands to constrain galaxy redshifts, particularly at higher redshifts. The performance is further improved when the SED information is included in the training. The bottom panel in Fig. 11 illustrates the ability of the MLP model to predict photo-zs, showing that
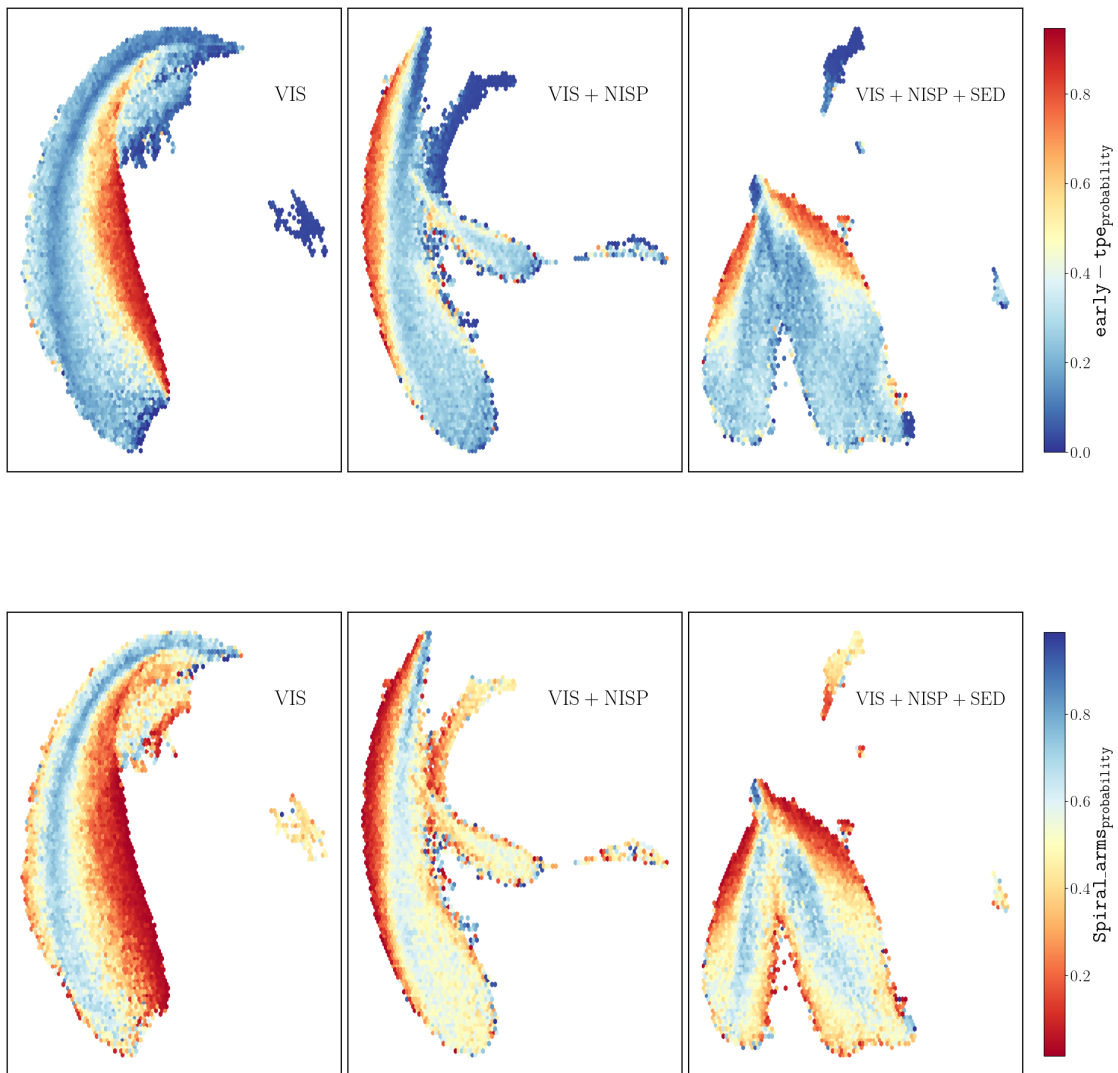
Fig. 8: UMAP visualisations of the self-supervised embeddings from `AstroPT` trained on VIS (left), VIS+NISP (centre), and VIS+NISP+SED (right) coloured by morphological parameters derived from the GZ catalogue (Euclid Collaboration: Walmsley et al. 2025b). Representations show clear clustering of smooth, non-spiral, and non-smooth, spiral galaxies.

the VIS+NISP+SED configuration achieves tighter clustering of predictions around the ideal 1:1 line, while the `NNPZ` photo-*z*s exhibit a more significant offset. Notably, the MLP model that catches the non-linearity tends to outperform other methods across all metrics, particularly for low percentages of labelled data. This approach achieves lower NMAD, and outlier fractions demonstrating its ability to adaptively leverage the added NISP, and SED information. For instance, with 100% labelled data, the MLP model achieves an NMAD = 0.048 ± 0.002 and an outlier fraction of $2.13\% \pm 0.38\%$, for training on VIS+NISP+SEDs approaching the performance of the official *Euclid* photo-*z* pipeline (NMAD = 0.037 and an outlier fraction of 2.21%). Linear models, while simpler, deliver competitive results for higher percentages of labelled data, where they approach the performance of *Euclid* pipeline, and more complex models. We note, that we observe a non-monotonic behaviour in model performance as a function of training sample size. Specifically, the NMAD and

outlier fraction peak at 5% training data, while performance improves at both smaller (1%) and larger (20%) sample sizes. This behaviour suggests the combination of the presence of the double descent effect (e.g., Schaeffer et al. 2023) in our regression model, and sample size balance. The double descent phenomenon occurs when test error initially decreases with increasing data size (classical regime), then spikes at an interpolation threshold, and finally decreases again as overparameterisation allows for better generalisation. In our case, with larger sample sizes (> 10%), the model moves into the overparameterised regime, where it can better capture meaningful patterns. On the other hand, when training on extremely limited data, the model may underfit, but given the low parameter count in linear regression this underfitting is less pronounced compared to the interpolation threshold. The fact that 1% performs comparably to 10% suggests that at extremely low sample sizes, the model is unable to overfit noise significantly, leading to stable results. In com-
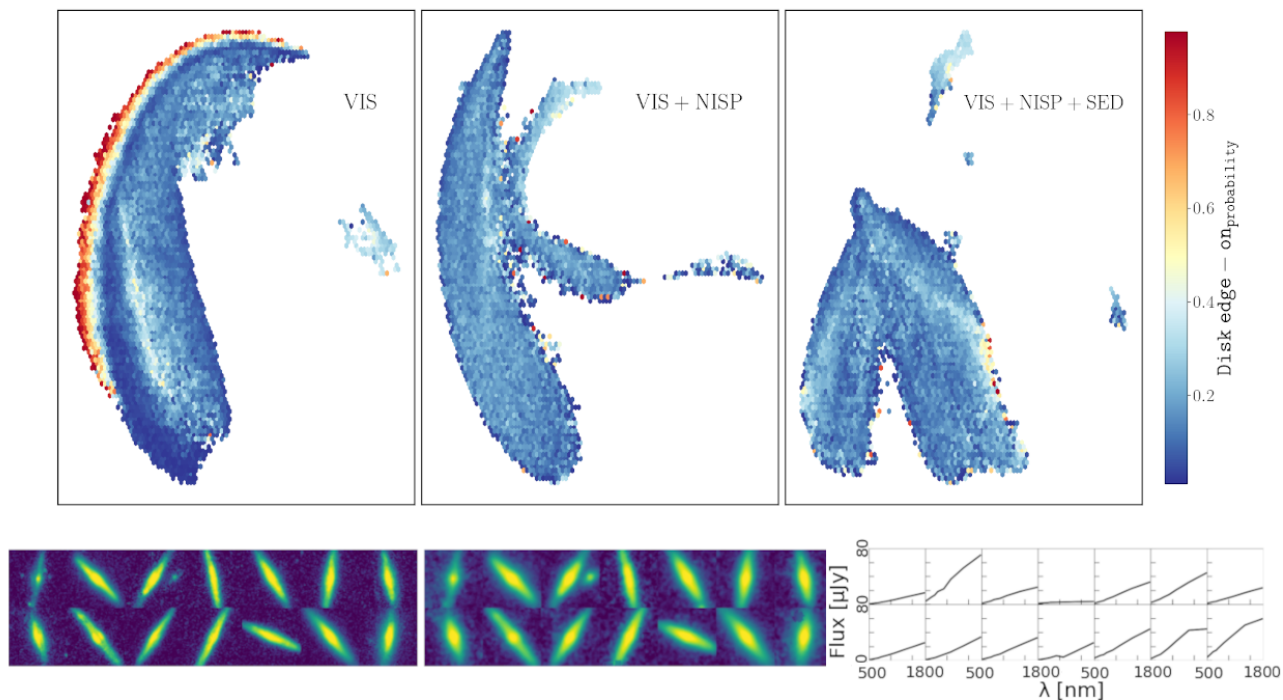
Fig. 9: UMAP visualisations of the self-supervised embeddings from `AstroPt` coloured by morphological parameters derived from the GZ catalogue (Euclid Collaboration: Walmsley et al. 2025b). *Left*: Representations from VIS images alone, showing a distinct region for edge-on galaxies. *Middle*: Representations incorporating both VIS and NISP data, where the separation of edge-on galaxies becomes less distinct. *Right*: Representations incorporating VIS, NISP, and SED data. *Bottom*: Example cutouts and SEDs of the edge-on galaxies.

Table 3: Comparison of early-type/late-type galaxy classification performance for different approaches. The model with the best (i.e., the highest) performance is bolded.

Model	Data	Accuracy	Precision	Recall	F1 score	FPR
Supervised CNN	VIS	0.84 ± 0.01	0.75 ± 0.02	0.71 ± 0.03	0.73 ± 0.01	0.11 ± 0.02
Linear model	VIS	0.86 ± 0.00	0.81 ± 0.01	0.73 ± 0.01	0.77 ± 0.00	0.08 ± 0.00
MLP model	**VIS**	$\mathbf{0.88 \pm 0.00}$	$\mathbf{0.80 \pm 0.01}$	$\mathbf{0.80 \pm 0.01}$	$\mathbf{0.80 \pm 0.00}$	$\mathbf{0.09 \pm 0.00}$
Supervised CNN	VIS+NISP	0.83 ± 0.01	0.74 ± 0.02	0.72 ± 0.02	0.73 ± 0.00	0.12 ± 0.02
Linear model	VIS+NISP	0.84 ± 0.00	0.79 ± 0.00	0.67 ± 0.01	0.72 ± 0.00	0.08 ± 0.00
MLP model	VIS+NISP	0.85 ± 0.00	0.77 ± 0.00	0.74 ± 0.01	0.76 ± 0.00	0.10 ± 0.00
Linear model	VIS+NISP+SED	0.85 ± 0.01	0.78 ± 0.01	0.73 ± 0.01	0.75 ± 0.01	0.10 ± 0.01
MLP model	VIS+NISP+SED	0.86 ± 0.01	0.78 ± 0.01	0.78 ± 0.01	0.78 ± 0.01	0.10 ± 0.01

parison, supervised CNN models exhibit modest improvements with increased labelled data, but generally underperform compared to other methods.

Overall, an MLP model trained with SED (VIS+NISP+SED data), not only achieves the best overall metrics, with improvements in bias and outlier fraction values in comparison to photo-zs from `NNPZ` for higher labelled data percentages, but also maintains superior results under low-data scenarios, demonstrating its robustness and adaptability for photo-z estimation tasks requiring only the SED information. These results underscore the effectiveness of leveraging advanced training strategies for accurate photometric redshift estimation.

4.2.3. Stellar masses

In this Section, we evaluate how `AstroPT` model is able to predict the stellar masses in comparison to more traditional approaches. As the ground truth, we assume the stellar masses of

the DESI Value-Added Catalogue (Siudek et al. 2024), which provides a robust data set derived from the SED fitting of optical and infrared photometry. First, we compare the stellar masses from *Euclid* official pipeline (Euclid Collaboration: Tucci et al. 2025) with the ones from DESI (see Fig. 13). The stellar masses are in agreement, as shown by the following 1:1 relation, but there is also a wide spread.

Similarly, as for the morphological parameters and photo-zs, we predict stellar mass using different methods and metrics introduced in Sect. 3.2. We compare this property for both VIS and VIS+NISP models, but do not compare for our VIS+NISP+SED model as the z-score normalisation across the SEDs removes the additional information needed to predict this property. The bias, NMAD, and outlier fraction as a function of the percentage of labelled data for the stellar mass are shown in Fig. 14. A statistical summary is given in Table 5. The MLP model, and linear model consistently achieve the lowest NMAD across, but only MLP is independent on the number of labels used for training, demon-
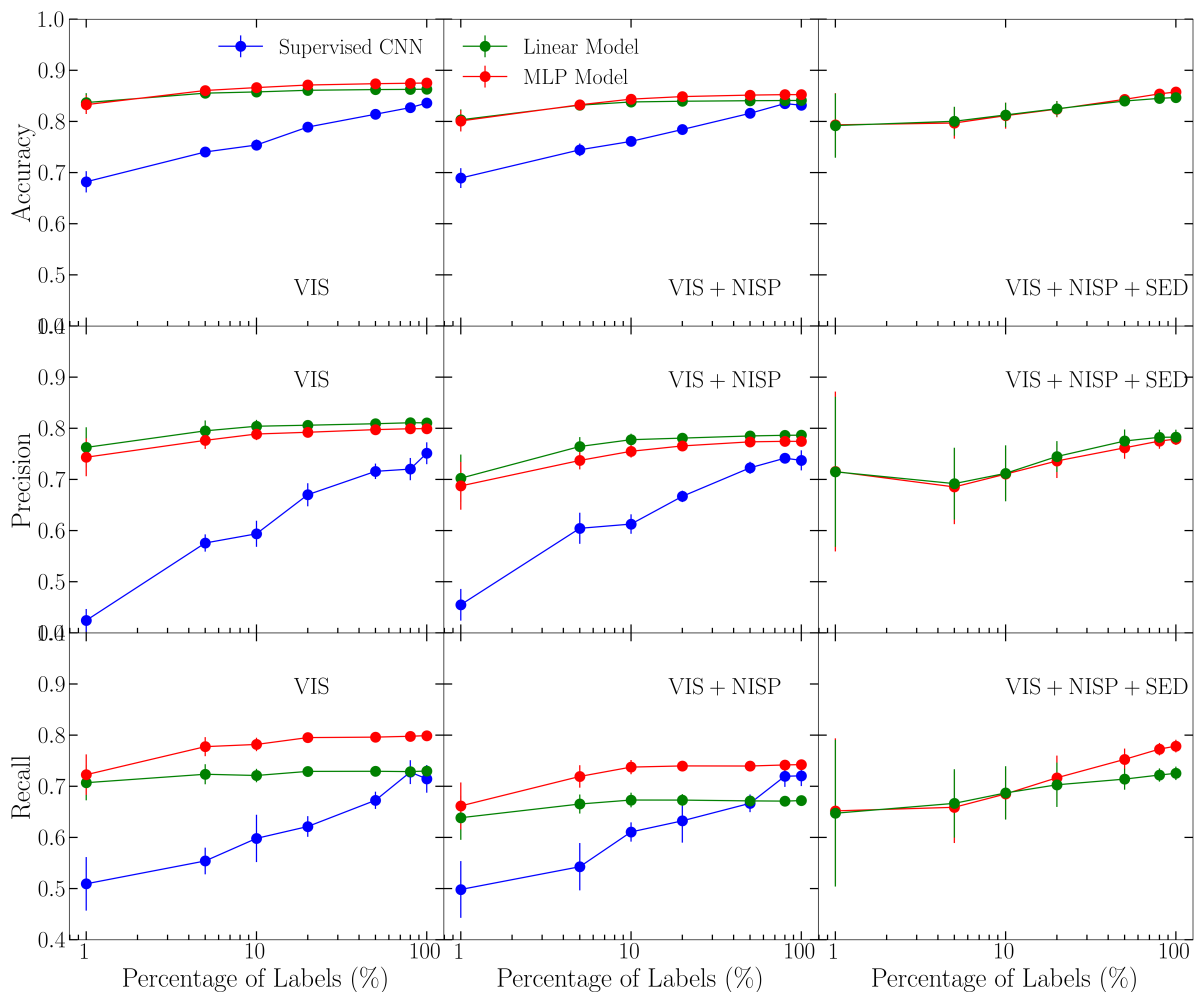
Fig. 10: Comparison of the accuracy, precision, and recall as functions of the percentage of labels used for training, for early-type/late-type classification. Results are shown for the VIS (left), VIS+NISP (centre), and VIS+NISP+SED configurations.

Table 4: Comparison of photo-z prediction performance for different approaches. The model with the best (i.e., the lowest) performance is bolded.

Model	Data	Bias	NMAD	Outlier fraction (%)
Supervised CNN	VIS	0.026 ± 0.004	0.118 ± 0.002	25.1 ± 0.7
Linear model	VIS	0.011 ± 0.003	0.100 ± 0.002	15.1 ± 0.6
MLP model	VIS	-0.002 ± 0.022	0.089 ± 0.004	12.7 ± 1.7
Supervised CNN	VIS+NISP	0.009 ± 0.006	0.103 ± 0.003	19.9 ± 0.3
Linear model	VIS+NISP	0.006 ± 0.002	0.063 ± 0.001	4.8 ± 0.4
MLP Model	VIS+NISP	-0.014 ± 0.012	0.057 ± 0.004	3.6 ± 0.5
Linear model	VIS+NISP+SED	0.004 ± 0.001	0.052 ± 0.001	2.7 ± 0.3
MLP model	**VIS+NISP+SED**	$\mathbf{-0.009 \pm 0.012}$	$\mathbf{0.048 \pm 0.002}$	$\mathbf{2.1 \pm 0.4}$
Euclid photo-z		-0.027	0.037	2.21

strating its robustness even with limited labelled data. The linear method, similarly as for the photo-z estimates shows the non-homogenous behaviour due to sample size, and double descent effect (see Sect. 4.2.2 for details). Nevertheless, for all labels used in training, both linear and MLP methods outperform estimates from *Euclid* official pipeline when the `AstroPT` is trained on both VIS and NISP data. In particular, stellar mass predictions with the linear model are statistically even closer to DESI stellar masses than official NNPZ estimates (with NMAD 0.023 ± 0.001, and 0.027, and outlier fraction $0.3 \pm 0.1\%$, 2.34% for MLP, and NNPZ, respectively. The performance degrades when solely VIS images are used for training. This highlights the added value of NISP's complementary wavelength coverage in constraining stellar mass. The distributions of the MLP stellar mass predictions is shown in the bottom panel in Fig. 13.

Table 5: Comparison of stellar mass prediction performance for different approaches. The model with the best (i.e., the lowest) performance is bolded.

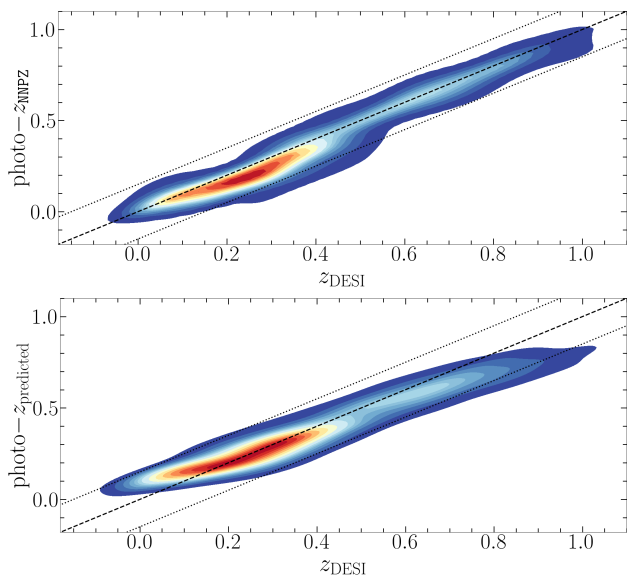| Model	Data	Bias	NMAD	Outlier fraction (%)
Supervised CNN	VIS	0.004 ± 0.017	0.063 ± 0.001	1.2 ± 0.4
Linear model	VIS	0.004 ± 0.001	0.050 ± 0.001	3.0 ± 0.3
MLP model	VIS	0.007 ± 0.002	0.052 ± 0.002	3.8 ± 0.4
Supervised CNN	VIS+NISP	0.011 ± 0.007	0.044 ± 0.001	0.5 ± 0.1
Linear model	**VIS+NISP**	**0.001 ± 0.001**	**0.023 ± 0.001**	**0.3 ± 0.1**
MLP model	VIS+NISP	0.002 ± 0.001	0.027 ± 0.003	0.6 ± 0.3
Euclid $\log_{10}(M_*/M_\odot)$		0.0011	0.0299	2.34

Fig. 11: Comparison of photo-z from *Euclid* official `NNPZ` pipeline (top) and predictions from the MLP model trained on `AstroPT` embeddings (bottom) versus DESI spec-z. A solid black line represents the 1:1 relation, while dotted lines indicate the outlier thresholds of 0.15.

4.2.4. Anomaly search

To identify peculiar objects within the learned feature space, we apply two unsupervised anomaly detection algorithms – isolation Forest (IF) and local outlier factor (LOF) – to the `AstroPT` embeddings. Both methods are implemented using the `Scikit-learn` Python package (Pedregosa et al. 2011). Before running the algorithms, we filter out spurious detections by removing sources with `DET_QUALITY_FLAG > 0`. To ensure robustness against sensitivity to parameter choices, we execute each algorithm multiple times with varying hyperparameters. Specifically, IF is tested with different numbers of base estimators, ranging from 50 to 200 in increments of 50, and the maximum number of features for splitting is varied between 1 and 20 in steps of 5. Additionally, we set the contamination parameter to an estimated fraction of outliers in the data set and do not use bootstrap sampling. For LOF, we explore the effect of different neighbourhood sizes by varying the number of neighbours from 5 to 50 in steps of 5. We use the 'ball-tree' (Dolatshah et al. 2015) algorithm for nearest neighbour searches and employ the cosine distance metric, as it has been shown to better capture similarities in high-dimensional spaces. The contamination parameter is set to reflect the expected proportion of outliers.

In the embeddings trained solely with VIS images, LOF predominantly detects anomalies often associated with spurious features, such as cosmic rays, bright stars in close proximity, or ghost artefacts (see upper panel in Fig. 15). We note that these objects were not flagged by standard quality-control metrics, indicating LOF's ability to identify subtle irregularities that might otherwise be overlooked. Additionally, LOF identifies extended objects in the outskirts of the cutout region, sometimes galaxies located far from image centres. These galaxies tend to be brighter and more extended than the central source, though such cases are relatively rare since most relevant features are concentrated near the image centre. In contrast, IF predominantly identifies large galaxies with rich morphological structures, including those with prominent rings indicative of recent merger events (see lower panel in Fig. 15). A key distinction between the two methods emerges in the spatial distribution of detected anomalies within the embedding space: LOF-identified outliers are scattered throughout, whereas IF-detected anomalies cluster closely together in the latent space (see Fig.16). These differences highlight the sensitivity of each method to different types of anomalies, with LOF favouring local density variations, and IF is adept at detecting sources with complex morphological structures.

When incorporating both VIS and NISP embeddings, the nature of the detected anomalies shifts. Both with IF and LOF algorithms, the VIS+NISP embedding highlights objects with multiple sources appearing within the field of view, potentially identifying interacting galaxies (see Fig. B.2 in Appendix B). However, this embedding is also prone to contamination by non-interacting systems, as the algorithms do not inherently account for redshift differences and instead focus on structural features within the images. Including redshift information as an additional constraint within the anomaly detection process could help disentangle interacting systems from unrelated projections of nearby and distant galaxies. Alternatively, utilising smaller image cutouts centred on each source might reduce confusion from neighbouring objects, improving both the accuracy and physical interpretability of detected anomalies. These refinements would enable more reliable identification of astrophysically meaningful outliers within the embedding space.

When SED information is included, the nature of the detected anomalies changes once again. This exemplifies how complementary information can alter the structure of the embedding space and, consequently, the anomaly detection results. As more information is incorporated, anomalies become potentially more interesting. When using the IF algorithm, we detect objects with significant dust presence, predominantly in the $H_{\rm E}$ band (see Fig. 17). The inclusion of SED data provides a measure of the relative importance of emission across different wavelengths, encoding valuable physical information. Interestingly, although the LOF algorithm still tends to detect anomalies driven by spurious
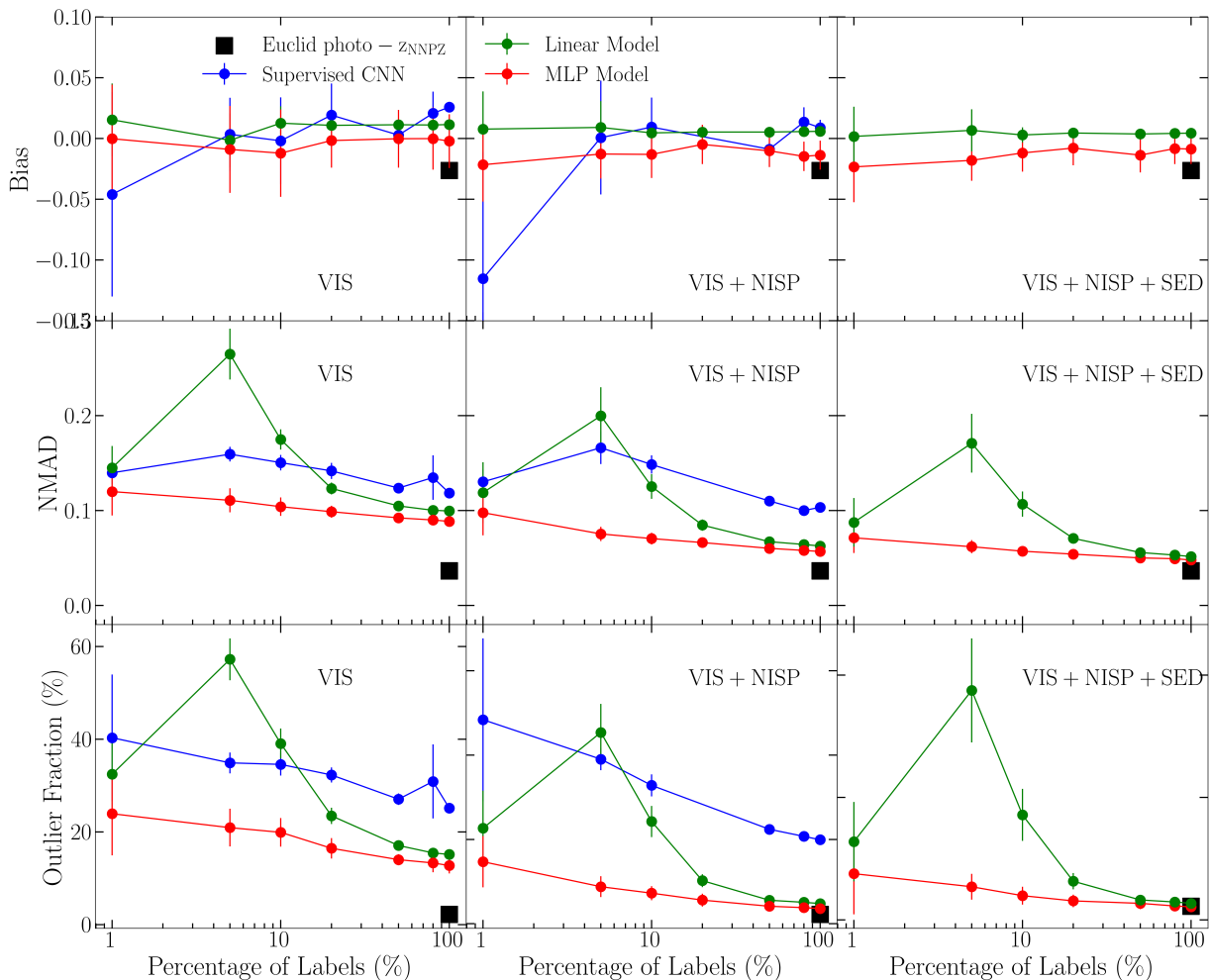
Fig. 12: Comparison of the bias, NMAD, and outlier fraction as functions of the percentage of labels used for training, for true redshift (DESI spec-*z*) versus *Euclid* photo-*z* predictions. Results are shown for the VIS (left), VIS+NISP (middle), and VIS+NISP+SED (right) configurations.

4.2.5. Similarity search

`AstroPT` projects objects with comparable features into the same regions of the embedding space, facilitating efficient similarity searches to identify objects of interest. This approach has shown great promise in various applications, particularly in identifying interacting galaxies. By using an interacting galaxy as a query, one can probe the embedding space to assess how many of its neighbouring galaxies also exhibit signs of interaction. This method allows for the rapid identification of potential interacting galaxies, contributing to large-scale surveys and studies of galaxy dynamics. In this work, we explore the potential of similarity search for identifying interacting galaxies and investigate how this approach can be adapted to astronomical data sets. As shown in Fig. 18, a significant fraction of the neighbouring galaxies to a query galaxy – chosen as an interacting galaxy – hereafter referred to as 'similar objects' – are indeed interacting. In particular, the distribution of the `merging_merger` parameter, which quantifies the probability of galaxies being in a merger process (Euclid Collaboration: Walmsley et al. 2025a), shows a higher fraction of merger candidates among the similar objects compared to the general galaxy population. This suggests that the similarity search effectively identifies galaxies involved in interactions. However, we also observe that the distribution of similar objects remains elongated, extending towards non-interacting galaxies as indicated by `merging_merger` ≈ 0, similar to the general distribution. This implies that, while the method is generally effective, it is not without contamination. In particular, the similarity search identifies some galaxies that are not truly interacting, highlighting the need for further refinement of the approach to reduce such false positives.

The presence of non-interacting neighbours, or 'false positives', in the identified clusters can be explained by the way `AstroPT` was trained and operates. During the training process, a fixed image size was chosen to ensure that the largest galaxies were fully captured within the image frame. This approach works well for galaxies at lower redshifts, where the central regions of galaxies can be clearly separated. However, for higher-redshift galaxies, the fixed image size may include multiple ob-

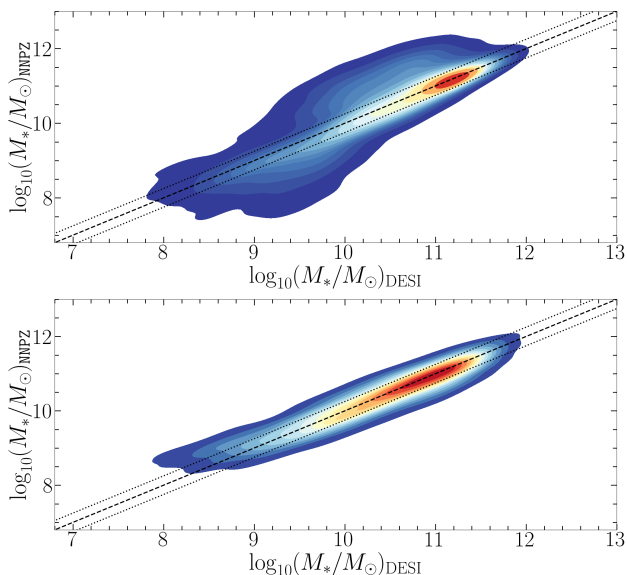features, it is nonetheless able to identify the two most prominent anomalies found with IF.

Fig. 13: Comparison of stellar mass from *Euclid* official NNPZ pipeline (top) and predictions from MLP model trained on AstroPT embeddings (bottom) versus DESI stellar masses from Siudek et al. (2024). A solid black line represents the 1:1 relation, while dotted lines indicate the outlier thresholds of 0.25 dex.

jects or extended outskirts of galaxies within the same frame. As a result, the algorithm may erroneously prioritise similarities in the outskirts of these galaxies, where there may be multiple objects or extended features, rather than focusing solely on the central regions where interactions are more likely to be detected. This can lead to the identification of objects as interacting when they are not, especially if the objects in question lie at different redshifts or are not truly interacting.

Despite these limitations, the approach shows great potential in identifying interacting galaxies. The contamination in the results, while non-negligible, offers insight into the nuances of the similarity search method and highlights the need for further refinement, particularly in distinguishing objects that are at different redshifts or in disentangling the central regions of interacting galaxies from their outskirts. These findings contribute to the broader understanding of similarity search applications in astrophysics and suggest areas for improvement in future models to enhance the accuracy of interaction detection.

5. Summary

In this paper, we have explored AstroPT, an autoregressive foundation model pre-trained on approximately $300\,000$ optical and infrared Q1 imaging and spectral energy distributions for deriving galaxy properties. By learning a unified and informative representation from multi-modal data, AstroPT provides a scalable solution for analysing large, unlabelled data sets, reducing dependence on human-labelled training samples while maintaining high performance for a variety of downstream tasks (such as photometric redshift estimation, morphology classification, anomaly detection).

The paper's highlights are as follows.

– Advancing foundation models in astronomy: AstroPT showcases the potential of multi-modal foundation models for a scalable and efficient resolution of many crucial astrophysical tasks, such as morphology classification, galaxy characterisation, and the search for peculiar sources. Unlike traditional task-specific models, AstroPT provides a generalisable feature space applicable across multiple domains, demonstrating the feasibility of self-supervised learning for astronomy at scale. Our results provide critical insight into the synergy between optical, infrared, and SED data, guiding future strategies for galaxy characterisation and cosmological analysis with *Euclid*.

– Multi-modal synergies: The integration of optical, and NIR photometry highlights the advantages of multi-modal approaches. Pre-trained embeddings capture the inherent diversity of galaxy morphologies (see Figs. 6 and 8), clearly isolating galaxy classes hardly distinguishable with standard methods, such as edge-on galaxies (see Fig. 9) and outliers, such as spurious objects and stars missed by standard techniques (see Figs. 7 and C.1).

– Unified representations for multiple tasks: Self-supervised embeddings provide a consistent representation for various downstream tasks, including photo-z and stellar mass estimates, anomalies, and similarity searches. This cross-task applicability underscores the flexibility of AstroPT's learned representations, which remain effective across different astrophysical domains. By leveraging the task-agnostic nature of self-supervised embeddings, we reduce dependence on large labelled data sets while maintaining high performance. This scalability is crucial for addressing the challenges of analysing large multi-modal data sets in future surveys.

– Morphological classification: Using solely optical images, our models achieve robust performance in predicting galaxy types. Surprisingly, incorporating NIR data and SEDs does not improve classification results, indicating that the additional spectral coverage does not introduce significant morphological information. The MLP and linear models demonstrate consistent performance across different data set sizes, maintaining higher robustness compared to traditional supervised approaches, even with a limited number of labels (see Fig. 10 and Table 3). This suggests that deep-learning-based embeddings effectively capture morphological characteristics using only optical imaging, potentially simplifying future survey strategies.

– Photo-z estimations: By integrating NIR, optical images, and SEDs, our models achieve competitive accuracy in photo-z estimates compared to the standard *Euclid* pipeline relying on traditional methods. Notably, the MLP model shows reduced bias and a lower fraction of catastrophic outliers, demonstrating that self-supervised embeddings can enhance photometric redshift predictions while being less reliant on large labelled training samples (see Fig. 12, and Table 4). This highlights the potential of deep-learning models adapted for standard tasks such as photo-z estimations, and the importance of multi-band photometry for capturing diverse galaxy properties and achieving higher precision.

– Stellar mass estimations: We explore the ability of AstroPT to reconstruct fundamental galaxy properties, such as stellar mass, showing that MLP-based models achieve high accuracy even in low-label regimes. The linear model performs comparably well, showcasing the flexibility of the learned representations (see Fig. 14 and Table 5). This result highlights the robustness of self-supervised learning in deriving astrophysical parameters, even when traditional template-fitting approaches require extensive labelled training sets.
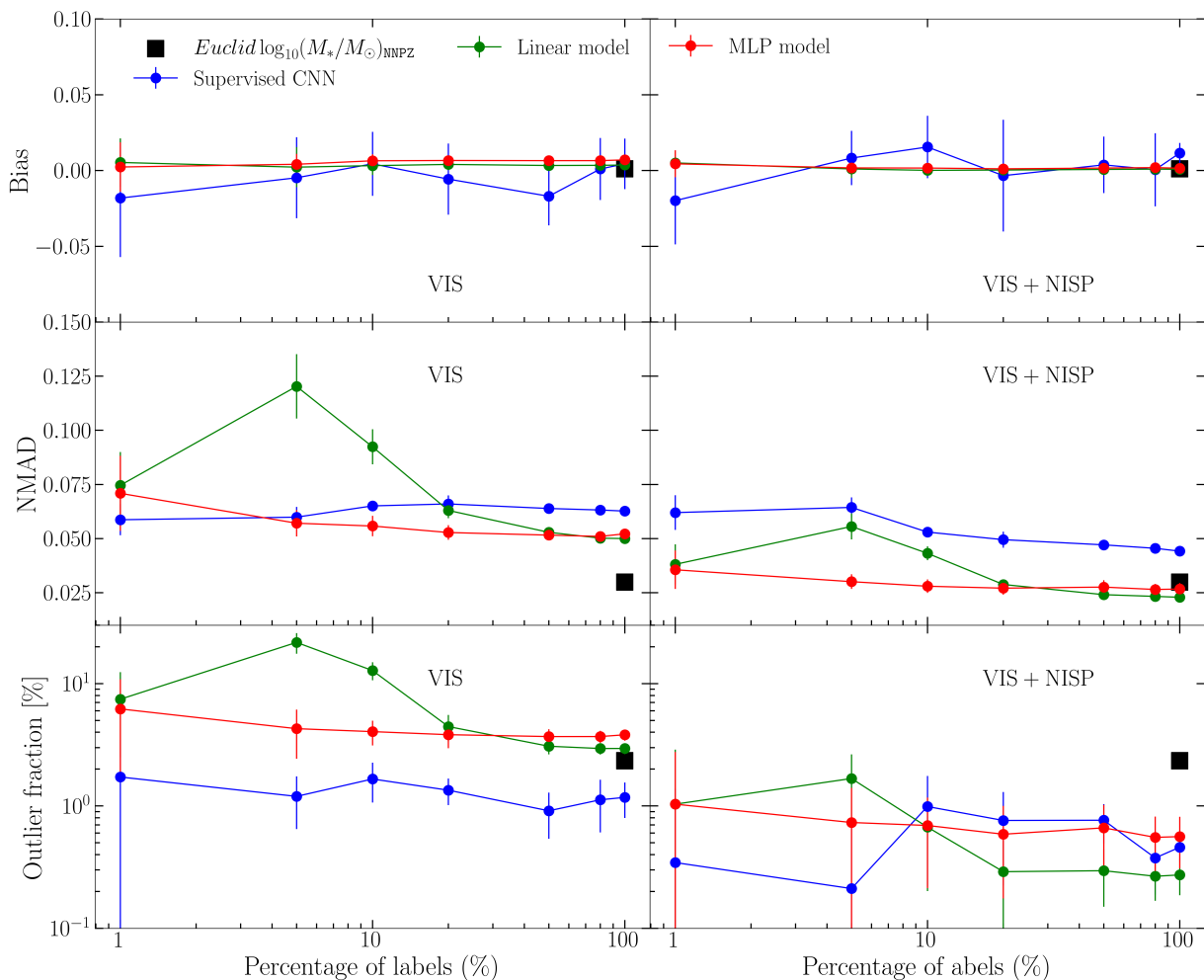
Fig. 14: Comparison of the NMAD and outlier fraction as functions of the percentage of labels used for training, for DESI stellar masses from Siudek et al. (2024) versus *Euclid* predictions. Results are shown for the VIS (left) and VIS+NISP (right) configurations.
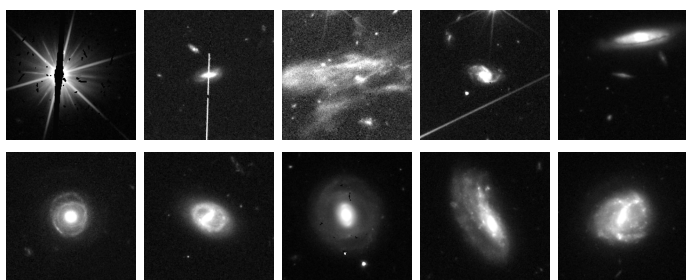

Fig. 15: Example VIS cutouts of anomaly detections found with LOF (upper panel) and IF (bottom panel) on VIS embeddings.

– Outlier and similarity search: `AstroPT`'s anomaly and similarity search capabilities allow the identification of rare astrophysical objects, including galaxies with rings, interacting systems, and other unusual morphologies. Including SEDs in training, we can also identify very dusty sources. This demonstrates the model's utility in detecting previously overlooked populations in large-scale galaxy surveys, paving the way for new discoveries in *Euclid* and future astronomical data sets.

– Limitations, and future perspectives: While `AstroPT` demonstrates strong performance across multiple tasks, some limitations remain. The reliance on pre-training with Q1 data means that domain adaptation might be required for different survey depths or observational conditions. Additionally, the current classification thresholds, and sample selection, could introduce selection biases, necessitating further refinement. Future work should explore strategies to mitigate these biases and improve generalisation to different data sets.

This work underscores the transformative potential of combining multi-modal data sets, self-supervised learning, and foundation models for tackling large-scale challenges in galaxy analysis.
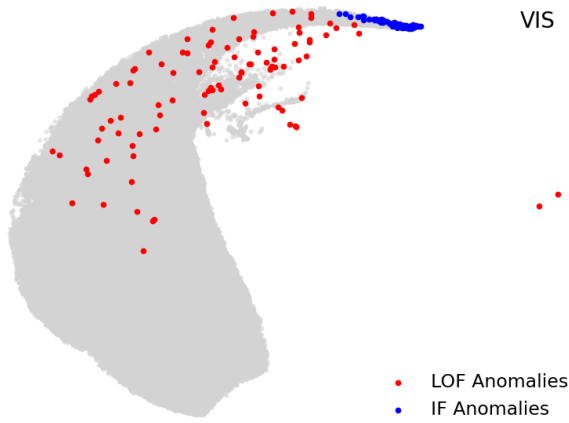
Fig. 16: UMAP visualisations of the self-supervised embeddings trained on VIS images with anomalies detected with LOF, and IF algorithms.
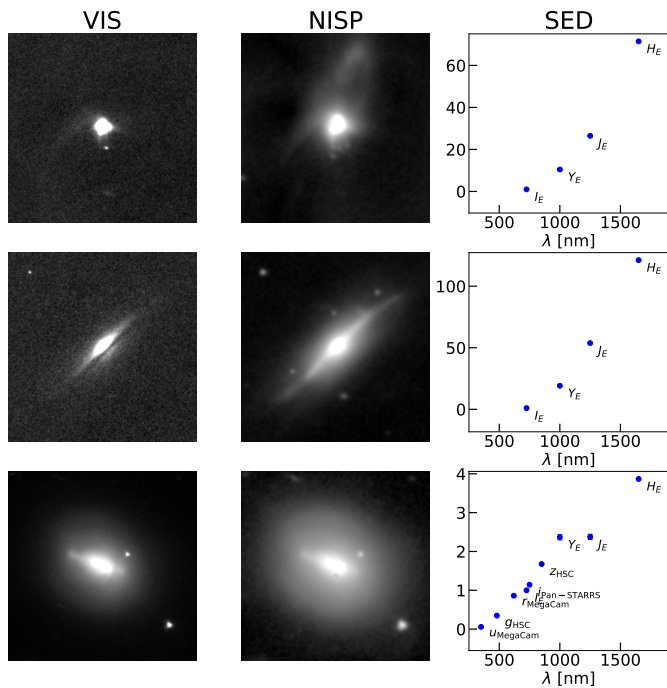

Fig. 17: Anomalies detected with the IF algorithm in the VIS+NISP+SED embedding. For each galaxy we show VIS and NSIP images and the SED. Fluxes are normalised to the *Euclid* VIS value.
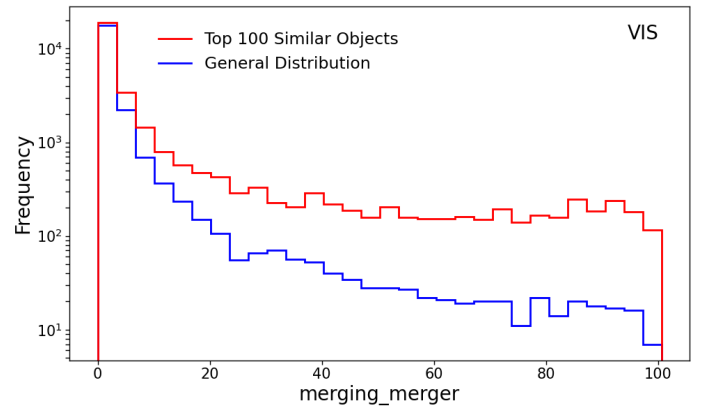

Fig. 18: Distribution of `merging_merger` parameter for similar objects computed with VIS embedding and general galaxy distribution.

References

Brown, T. B., Mann, B., Ryder, N., et al. 2020 [`arXiv:2005.14165`]
Bundy, K., Bershady, M. A., Law, D. R., et al. 2015, ApJ, 798, 7
Chen, T., Kornblith, S., Norouzi, M., & Hinton, G. 2020, arXiv e-prints, arXiv:2002.05709
Cheng, T.-Y., Conselice, C. J., Aragón-Salamanca, A., et al. 2020, MNRAS, 493, 4209
Cheng, T.-Y., Huertas-Company, M., Conselice, C. J., et al. 2021, MNRAS, 503, 4446
Chu, J., Tang, H., Xu, D., Lu, S., & Long, R. 2024, MNRAS, 528, 6354
Cortes, C., Jackel, L. D., Solla, S. A., Vapnik, V., & Denker, J. S. 1993, in Guide Proceedings (San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.), 327–334
Dark Energy Survey Collaboration, Abbott, T., Abdalla, F. B., et al. 2016, MNRAS, 460, 1270
DESI Collaboration, Abareshi, B., Aguilar, J., et al. 2022, AJ, 164, 207
DESI Collaboration, Adame, A. G., Aguilar, J., et al. 2024a, AJ, 167, 62
DESI Collaboration, Adame, A. G., Aguilar, J., et al. 2024b, AJ, 168, 58
Dolatshah, M., Hadian, A., & Minaei-Bidgoli, B. 2015, arXiv e-prints, arXiv:1511.00628
Domínguez Sánchez, H., Huertas-Company, M., Bernardi, M., et al. 2019, MNRAS, 484, 93
Domínguez Sánchez, H., Huertas-Company, M., Bernardi, M., Tuccillo, D., & Fischer, J. L. 2018, MNRAS, 476, 3661
El-Nouby, A., Klein, M., Zhai, S., et al. 2024, arXiv e-prints, arXiv:2401.08541
Euclid Collaboration: Aussel, B., Kruk, S., Walmsley, M., et al. 2024, A&A, 689, A274

Euclid Collaboration: Aussel, H., Tereno, I., Schirmer, M., et al. 2025, A&A, submitted

Euclid Collaboration: Bisigello, L., Conselice, C. J., Baes, M., et al. 2023, MNRAS, 520, 3529

Euclid Collaboration: Cropper, M., Al Bahlawan, A., Amiaux, J., et al. 2024, A&A, accepted, arXiv:2405.13492

Euclid Collaboration: Desprez, G., Paltani, S., Coupon, J., et al. 2020, A&A, 644, A31

Euclid Collaboration: Enia, A., Bolzonella, M., Pozzetti, L., et al. 2024, A&A, 691, A175

Euclid Collaboration: Enia, A., Pozzetti, L., Bolzonella, M., et al. 2025, A&A, submitted

Euclid Collaboration: Humphrey, A., Bisigello, L., Cunha, P. A. C., et al. 2023, A&A, 671, A99

Euclid Collaboration: Jahnke, K., Gillard, W., Schirmer, M., et al. 2024, A&A, accepted, arXiv:2405.13493

Euclid Collaboration: Leuzzi, L., Meneghetti, M., Angora, G., et al. 2024, A&A, 681, A68

Euclid Collaboration: McCracken, H., Benson, K., et al. 2025, A&A, submitted

Euclid Collaboration: Mellier, Y., Abdurro'uf, Acevedo Barroso, J., et al. 2024, A&A, accepted, arXiv:2405.13491

Euclid Collaboration: Polenta, G., Frailis, M., Alavi, A., et al. 2025, A&A, submitted

Euclid Collaboration: Romelli, E., Kümmel, M., Dole, H., et al. 2025, A&A, submitted

Euclid Collaboration: Schirmer, M., Jahnke, K., Seidel, G., et al. 2022, A&A, 662, A92

Euclid Collaboration: Tucci, M., Paltani, S., Hartley, W., et al. 2025, A&A, submitted

Euclid Collaboration: Walmsley, M., Holloway, P., Lines, N., et al. 2025a, A&A, submitted

Euclid Collaboration: Walmsley, M., Huertas-Company, M., Quilley, L., et al. 2025b, A&A, submitted

Euclid Quick Release Q1. 2025, https://doi.org/10.57780/esa-2853f3b

Fotopoulou, S. 2024, Astronomy and Computing, 48, 100851

Fukushima, K. 1980, Biological Cybernetics, 36, 193

Gai, M., Bove, M., Bonetta, G., Zago, D., & Cancelliere, R. 2024, MNRAS, 532, 1391

Grattafiori, A., Dubey, A., Jauhri, A., et al. 2024, ArXiv e-prints [2407.21783]

Gunn, J. E., Carr, M., Rockosi, C., et al. 1998, AJ, 116, 3040

Gunn, J. E., Siegmund, W. A., Mannery, E. J., et al. 2006, AJ, 131, 2332

Hayat, M. A., Stein, G., Harrington, P., Lukić, Z., & Mustafa, M. 2021, ApJ, 911, L33

Henighan, T., Kaplan, J., Katz, M., et al. 2020, ArXiv e-prints [2010.14701]

Hocking, A., Geach, J. E., Sun, Y., & Davey, N. 2018, MNRAS, 473, 1108

Huber, P. J. 1964, The Annals of Mathematical Statistics, 35, 73

Huertas-Company, M., Iyer, K. G., Angeloudi, E., et al. 2024, A&A, 685, A48

Huertas-Company, M. & Lanusse, F. 2023, PASA, 40, e001

Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, ApJ, 873, 111

Kaplan, J., McCandlish, S., Henighan, T., et al. 2020, ArXiv e-prints [2001.08361]

Kümmel, M., Álvarez-Ayllón, A., Bertin, E., et al. 2022, arXiv e-prints, arXiv:2212.02428

Laureijs, R., Amiaux, J., Arduini, S., et al. 2011, ESA/SRE(2011)12, arXiv:1110.3193

Loshchilov, I. & Hutter, F. 2017, arXiv e-prints, arXiv:1711.05101

Mampaso, A., García de la Rosa, I., Arias, P., et al. 2022, Study on environmental sustainability at the Instituto de Astrofísica de Canarias and proposed actions, Technical report, Instituto de Astrofísica de Canarias

Martínez-Solaeche, G., García-Benito, R., González Delgado, R. M., et al. 2024, A&A, 688, A160

McInnes, L., Healy, J., & Melville, J. 2018, arXiv e-prints, arXiv:1802.03426

Merlin, E., Amorín, R., Castellano, M., et al. 2016, A&A, 590, A30

Merlin, E., Fontana, A., Ferguson, H. C., et al. 2015, A&A, 582, A15

Merlin, E., Pilo, S., Fontana, A., et al. 2019, A&A, 622, A169

Merz, G., Liu, X., Schmidt, S., et al. 2024, arXiv e-prints, arXiv:2411.18769

Mishra-Sharma, S., Song, Y., & Thaler, J. 2024, arXiv e-prints, arXiv:2403.08851

Mohale, K. & Lochner, M. 2024, MNRAS, 530, 1274

Nair, V. & Hinton, G. E. 2010, in Proceedings of the 27th International Conference on International Conference on Machine Learning, ICML'10 (Madison, WI, USA: Omnipress), 807–814

Pan, J.-S., Ting, Y.-S., Huang, Y., Yu, J., & Liu, J.-F. 2024, ArXiv e-prints [2405.17156]

Parker, L., Lanusse, F., Golkar, S., et al. 2024, MNRAS, 531, 4990

Pattnaik, R., Kartaltepe, J. S., & Binu, C. 2025, arXiv e-prints, arXiv:2501.01070

Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, Journal of machine learning research, 12, 2825

Radford, A., Kim, J. W., Hallacy, C., et al. 2021, in International Conference on Machine Learning (PMLR), 8748–8763

Rizhko, M. & Bloom, J. S. 2024, arXiv e-prints, arXiv:2411.08842

Rosenblatt, F. 1958, Psychological Review, 65, 386

Salvato, M., Ilbert, O., & Hoyle, B. 2019, Nature, 3, 212

Sarmiento, R., Huertas-Company, M., Knapen, J. H., et al. 2021, ApJ, 921, 177

Schaeffer, R., Khona, M., Robertson, Z., et al. 2023, arXiv e-prints, arXiv:2303.14151

Siudek, M., Pucha, R., Mezcua, M., et al. 2024, A&A, 691, A308

Slijepcevic, I. V., Scaife, A. M. M., Walmsley, M., et al. 2024, RAS Techniques and Instruments, 3, 19

Smith, M. J. & Geach, J. E. 2023, Royal Society Open Science, 10, 221454

Smith, M. J., Roberts, R. J., Angeloudi, E., & Huertas-Company, M. 2024, arXiv e-prints, arXiv:2405.14930

Stein, G., Harrington, P., Blaum, J., Medan, T., & Lukic, Z. 2021, arXiv e-prints, arXiv:2110.13151

Strubell, E., Ganesh, A., & Mccallum, A. 2019, ACL Anthology, 3645

Tanaka, M., Coupon, J., Hsieh, B.-C., et al. 2018, PASJ, 70, S9

The Multimodal Universe Collaboration. 2024, ArXiv e-prints [2412.02527]

Vega-Ferrero, J., Domínguez Sánchez, H., Bernardi, M., et al. 2021, MNRAS, 506, 1927

Vega-Ferrero, J., Huertas-Company, M., Costantin, L., et al. 2024, ApJ, 961, 51

Walmsley, M., Bowles, M., Scaife, A. M. M., et al. 2024, ArXiv e-prints [2404.02973]

Walmsley, M., Géron, T., Kruk, S., et al. 2023, MNRAS, 526, 4768

Walmsley, M., Lintott, C., Géron, T., et al. 2022a, MNRAS, 509, 3966

Walmsley, M., Scaife, A. M. M., Lintott, C., et al. 2022b, MNRAS, 513, 1581

Xu, B., Wang, N., Chen, T., & Li, M. 2015, arXiv e-prints, arXiv:1505.00853

Zhang, G., Helfer, T., Gagliano, A. T., Mishra-Sharma, S., & Ashley Villar, V. 2024, Machine Learning: Science and Technology, 5, 045069

1. Instituto de Astrofísica de Canarias (IAC); Departamento de Astrofísica, Universidad de La Laguna (ULL), 38200, La Laguna, Tenerife, Spain
2. Institute of Space Sciences (ICE, CSIC), Campus UAB, Carrer de Can Magrans, s/n, 08193 Barcelona, Spain
3. Instituto de Astrofísica de Canarias, Vía Láctea, 38205 La Laguna, Tenerife, Spain
4. Université PSL, Observatoire de Paris, Sorbonne Université, CNRS, LERMA, 75014, Paris, France
5. Université Paris-Cité, 5 Rue Thomas Mann, 75013, Paris, France
6. School of Physics, Astronomy and Mathematics, University of Hertfordshire, College Lane, Hatfield AL10 9AB, UK
7. Aspia Space, Falmouth, TR10 9TA, UK
8. Instituto de Astrofísica de Andalucía, CSIC, Glorieta de la Astronomía, 18080, Granada, Spain
9. Université Paris-Saclay, Université Paris Cité, CEA, CNRS, AIM, 91191, Gif-sur-Yvette, France
10. Center for Computational Astrophysics, Flatiron Institute, 162 5th Avenue, 10010, New York, NY, USA
11. Center for Cosmology and Particle Physics, Department of Physics, New York University, New York, NY 10003, USA
12. Princeton University, Princeton, NJ 08540, USA
13. Universidad de La Laguna, Departamento de Astrofísica, 38206 La Laguna, Tenerife, Spain
14. Faculdade de Ciências da Universidade do Porto, Rua do Campo de Alegre, 4150-007 Porto, Portugal
15. Instituto de Astrofísica e Ciências do Espaço, Universidade do Porto, CAUP, Rua das Estrelas, PT4150-762 Porto, Portugal
16. Instituto de Física de Cantabria, Edificio Juan Jordá, Avenida de los Castros, 39005 Santander, Spain
17. Leiden Observatory, Leiden University, Einsteinweg 55, 2333 CC Leiden, The Netherlands
18. Kapteyn Astronomical Institute, University of Groningen, PO Box 800, 9700 AV Groningen, The Netherlands
19. INAF-Osservatorio Astronomico di Capodimonte, Via Moiariello 16, 80131 Napoli, Italy
20. Max Planck Institute for Extraterrestrial Physics, Giessenbachstr. 1, 85748 Garching, Germany
21. Institut d'Estudis Espacials de Catalunya (IEEC), Edifici RDIT, Campus UPC, 08860 Castelldefels, Barcelona, Spain
22. School of Physics, HH Wills Physics Laboratory, University of Bristol, Tyndall Avenue, Bristol, BS8 1TL, UK

23 Centro de Estudios de Física del Cosmos de Aragón (CEFCA), Plaza San Juan, 1, planta 2, 44001, Teruel, Spain

24 Université Paris-Saclay, CNRS, Institut d'astrophysique spatiale, 91405, Orsay, France

25 ESAC/ESA, Camino Bajo del Castillo, s/n., Urb. Villafranca del Castillo, 28692 Villanueva de la Cañada, Madrid, Spain

26 School of Mathematics and Physics, University of Surrey, Guildford, Surrey, GU2 7XH, UK

27 INAF-Osservatorio Astronomico di Brera, Via Brera 28, 20122 Milano, Italy

28 INAF-Osservatorio di Astrofisica e Scienza dello Spazio di Bologna, Via Piero Gobetti 93/3, 40129 Bologna, Italy

29 IFPU, Institute for Fundamental Physics of the Universe, via Beirut 2, 34151 Trieste, Italy

30 INAF-Osservatorio Astronomico di Trieste, Via G. B. Tiepolo 11, 34143 Trieste, Italy

31 INFN, Sezione di Trieste, Via Valerio 2, 34127 Trieste TS, Italy

32 SISSA, International School for Advanced Studies, Via Bonomea 265, 34136 Trieste TS, Italy

33 Dipartimento di Fisica e Astronomia, Università di Bologna, Via Gobetti 93/2, 40129 Bologna, Italy

34 INFN-Sezione di Bologna, Viale Berti Pichat 6/2, 40127 Bologna, Italy

35 Space Science Data Center, Italian Space Agency, via del Politecnico snc, 00133 Roma, Italy

36 Dipartimento di Fisica, Università di Genova, Via Dodecaneso 33, 16146, Genova, Italy

37 INFN-Sezione di Genova, Via Dodecaneso 33, 16146, Genova, Italy

38 Department of Physics "E. Pancini", University Federico II, Via Cinthia 6, 80126, Napoli, Italy

39 Dipartimento di Fisica, Università degli Studi di Torino, Via P. Giuria 1, 10125 Torino, Italy

40 INFN-Sezione di Torino, Via P. Giuria 1, 10125 Torino, Italy

41 INAF-Osservatorio Astrofisico di Torino, Via Osservatorio 20, 10025 Pino Torinese (TO), Italy

42 European Space Agency/ESTEC, Keplerlaan 1, 2201 AZ Noordwijk, The Netherlands

43 Institute Lorentz, Leiden University, Niels Bohrweg 2, 2333 CA Leiden, The Netherlands

44 INAF-IASF Milano, Via Alfonso Corti 12, 20133 Milano, Italy

45 Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Avenida Complutense 40, 28040 Madrid, Spain

46 Port d'Informació Científica, Campus UAB, C. Albareda s/n, 08193 Bellaterra (Barcelona), Spain

47 Institute for Theoretical Particle Physics and Cosmology (TTK), RWTH Aachen University, 52056 Aachen, Germany

48 INAF-Osservatorio Astronomico di Roma, Via Frascati 33, 00078 Monteporzio Catone, Italy

49 INFN section of Naples, Via Cinthia 6, 80126, Napoli, Italy

50 Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, HI 96822, USA

51 Dipartimento di Fisica e Astronomia "Augusto Righi" - Alma Mater Studiorum Università di Bologna, Viale Berti Pichat 6/2, 40127 Bologna, Italy

52 Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK

53 Jodrell Bank Centre for Astrophysics, Department of Physics and Astronomy, University of Manchester, Oxford Road, Manchester M13 9PL, UK

54 European Space Agency/ESRIN, Largo Galileo Galilei 1, 00044 Frascati, Roma, Italy

55 Université Claude Bernard Lyon 1, CNRS/IN2P3, IP2I Lyon, UMR 5822, Villeurbanne, F-69100, France

56 Institut de Ciències del Cosmos (ICCUB), Universitat de Barcelona (IEEC-UB), Martí i Franquès 1, 08028 Barcelona, Spain

57 Institució Catalana de Recerca i Estudis Avançats (ICREA), Passeig de Lluís Companys 23, 08010 Barcelona, Spain

58 UCB Lyon 1, CNRS/IN2P3, IUF, IP2I Lyon, 4 rue Enrico Fermi, 69622 Villeurbanne, France

59 Mullard Space Science Laboratory, University College London, Holmbury St Mary, Dorking, Surrey RH5 6NT, UK

60 Departamento de Física, Faculdade de Ciências, Universidade de Lisboa, Edifício C8, Campo Grande, PT1749-016 Lisboa, Portugal

61 Instituto de Astrofísica e Ciências do Espaço, Faculdade de Ciências, Universidade de Lisboa, Campo Grande, 1749-016 Lisboa, Portugal

62 Department of Astronomy, University of Geneva, ch. d'Ecogia 16, 1290 Versoix, Switzerland

63 INAF-Istituto di Astrofisica e Planetologia Spaziali, via del Fosso del Cavaliere, 100, 00100 Roma, Italy

64 INFN-Padova, Via Marzolo 8, 35131 Padova, Italy

65 Aix-Marseille Université, CNRS/IN2P3, CPPM, Marseille, France

66 INFN-Bologna, Via Irnerio 46, 40126 Bologna, Italy

67 Universitäts-Sternwarte München, Fakultät für Physik, Ludwig-Maximilians-Universität München, Scheinerstrasse 1, 81679 München, Germany

68 INAF-Osservatorio Astronomico di Padova, Via dell'Osservatorio 5, 35122 Padova, Italy

69 NRC Herzberg, 5071 West Saanich Rd, Victoria, BC V9E 2E7, Canada

70 Institute of Theoretical Astrophysics, University of Oslo, P.O. Box 1029 Blindern, 0315 Oslo, Norway

71 Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA, 91109, USA

72 Department of Physics, Lancaster University, Lancaster, LA1 4YB, UK

73 Felix Hormuth Engineering, Goethestr. 17, 69181 Leimen, Germany

74 Technical University of Denmark, Elektrovej 327, 2800 Kgs. Lyngby, Denmark

75 Cosmic Dawn Center (DAWN), Denmark

76 Max-Planck-Institut für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany

77 NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA

78 Department of Physics and Helsinki Institute of Physics, Gustaf Hällströmin katu 2, 00014 University of Helsinki, Finland

79 Université de Genève, Département de Physique Théorique and Centre for Astroparticle Physics, 24 quai Ernest-Ansermet, CH-1211 Genève 4, Switzerland

80 Department of Physics, P.O. Box 64, 00014 University of Helsinki, Finland

81 Helsinki Institute of Physics, Gustaf Hällströmin katu 2, University of Helsinki, Helsinki, Finland

82 Centre de Calcul de l'IN2P3/CNRS, 21 avenue Pierre de Coubertin 69627 Villeurbanne Cedex, France

83 Laboratoire d'etude de l'Univers et des phenomenes eXtremes, Observatoire de Paris, Université PSL, Sorbonne Université, CNRS, 92190 Meudon, France

84 Aix-Marseille Université, CNRS, CNES, LAM, Marseille, France

85 SKA Observatory, Jodrell Bank, Lower Withington, Macclesfield, Cheshire SK11 9FT, UK

86 Dipartimento di Fisica "Aldo Pontremoli", Università degli Studi di Milano, Via Celoria 16, 20133 Milano, Italy

87 INFN-Sezione di Milano, Via Celoria 16, 20133 Milano, Italy

88 University of Applied Sciences and Arts of Northwestern Switzerland, School of Computer Science, 5210 Windisch, Switzerland

89 Universität Bonn, Argelander-Institut für Astronomie, Auf dem Hügel 71, 53121 Bonn, Germany

90 INFN-Sezione di Roma, Piazzale Aldo Moro, 2 - c/o Dipartimento di Fisica, Edificio G. Marconi, 00185 Roma, Italy

91 Dipartimento di Fisica e Astronomia "Augusto Righi" - Alma Mater Studiorum Università di Bologna, via Piero Gobetti 93/2, 40129 Bologna, Italy

92 Department of Physics, Institute for Computational Cosmology, Durham University, South Road, Durham, DH1 3LE, UK

[93] Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, Laboratoire Lagrange, Bd de l'Observatoire, CS 34229, 06304 Nice cedex 4, France

[94] Institut d'Astrophysique de Paris, UMR 7095, CNRS, and Sorbonne Université, 98 bis boulevard Arago, 75014 Paris, France

[95] Université Paris Cité, CNRS, Astroparticule et Cosmologie, 75013 Paris, France

[96] CNRS-UCB International Research Laboratory, Centre Pierre Binetruy, IRL2007, CPB-IN2P3, Berkeley, USA

[97] University of Applied Sciences and Arts of Northwestern Switzerland, School of Engineering, 5210 Windisch, Switzerland

[98] Institut d'Astrophysique de Paris, 98bis Boulevard Arago, 75014, Paris, France

[99] Institute of Physics, Laboratory of Astrophysics, Ecole Polytechnique Fédérale de Lausanne (EPFL), Observatoire de Sauverny, 1290 Versoix, Switzerland

[100] Aurora Technology for European Space Agency (ESA), Camino bajo del Castillo, s/n, Urbanizacion Villafranca del Castillo, Villanueva de la Cañada, 28692 Madrid, Spain

[101] Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Campus UAB, 08193 Bellaterra (Barcelona), Spain

[102] School of Mathematics, Statistics and Physics, Newcastle University, Herschel Building, Newcastle-upon-Tyne, NE1 7RU, UK

[103] DARK, Niels Bohr Institute, University of Copenhagen, Jagtvej 155, 2200 Copenhagen, Denmark

[104] Waterloo Centre for Astrophysics, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada

[105] Department of Physics and Astronomy, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada

[106] Perimeter Institute for Theoretical Physics, Waterloo, Ontario N2L 2Y5, Canada

[107] Centre National d'Etudes Spatiales – Centre spatial de Toulouse, 18 avenue Edouard Belin, 31401 Toulouse Cedex 9, France

[108] Institute of Space Science, Str. Atomistilor, nr. 409 Măgurele, Ilfov, 077125, Romania

[109] Dipartimento di Fisica e Astronomia "G. Galilei", Università di Padova, Via Marzolo 8, 35131 Padova, Italy

[110] Institut für Theoretische Physik, University of Heidelberg, Philosophenweg 16, 69120 Heidelberg, Germany

[111] Institut de Recherche en Astrophysique et Planétologie (IRAP), Université de Toulouse, CNRS, UPS, CNES, 14 Av. Edouard Belin, 31400 Toulouse, France

[112] Université St Joseph; Faculty of Sciences, Beirut, Lebanon

[113] Departamento de Física, FCFM, Universidad de Chile, Blanco Encalada 2008, Santiago, Chile

[114] Universität Innsbruck, Institut für Astro- und Teilchenphysik, Technikerstr. 25/8, 6020 Innsbruck, Austria

[115] Satlantis, University Science Park, Sede Bld 48940, Leioa-Bilbao, Spain

[116] Instituto de Astrofísica e Ciências do Espaço, Faculdade de Ciências, Universidade de Lisboa, Tapada da Ajuda, 1349-018 Lisboa, Portugal

[117] Cosmic Dawn Center (DAWN)

[118] Niels Bohr Institute, University of Copenhagen, Jagtvej 128, 2200 Copenhagen, Denmark

[119] Universidad Politécnica de Cartagena, Departamento de Electrónica y Tecnología de Computadoras, Plaza del Hospital 1, 30202 Cartagena, Spain

[120] Infrared Processing and Analysis Center, California Institute of Technology, Pasadena, CA 91125, USA

[121] Dipartimento di Fisica e Scienze della Terra, Università degli Studi di Ferrara, Via Giuseppe Saragat 1, 44122 Ferrara, Italy

[122] Istituto Nazionale di Fisica Nucleare, Sezione di Ferrara, Via Giuseppe Saragat 1, 44122 Ferrara, Italy

[123] INAF, Istituto di Radioastronomia, Via Piero Gobetti 101, 40129 Bologna, Italy

[124] Department of Physics, Oxford University, Keble Road, Oxford OX1 3RH, UK

[125] INAF - Osservatorio Astronomico di Brera, via Emilio Bianchi 46, 23807 Merate, Italy

[126] INAF-Osservatorio Astronomico di Brera, Via Brera 28, 20122 Milano, Italy, and INFN-Sezione di Genova, Via Dodecaneso 33, 16146, Genova, Italy

[127] ICL, Junia, Université Catholique de Lille, LITL, 59000 Lille, France

[128] ICSC - Centro Nazionale di Ricerca in High Performance Computing, Big Data e Quantum Computing, Via Magnanelli 2, Bologna, Italy

[129] Instituto de Física Teórica UAM-CSIC, Campus de Cantoblanco, 28049 Madrid, Spain

[130] CERCA/ISO, Department of Physics, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, OH 44106, USA

[131] Technical University of Munich, TUM School of Natural Sciences, Physics Department, James-Franck-Str. 1, 85748 Garching, Germany

[132] Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str. 1, 85748 Garching, Germany

[133] Laboratoire Univers et Théorie, Observatoire de Paris, Université PSL, Université Paris Cité, CNRS, 92190 Meudon, France

[134] Departamento de Física Fundamental. Universidad de Salamanca. Plaza de la Merced s/n. 37008 Salamanca, Spain

[135] Université de Strasbourg, CNRS, Observatoire astronomique de Strasbourg, UMR 7550, 67000 Strasbourg, France

[136] Center for Data-Driven Discovery, Kavli IPMU (WPI), UTIAS, The University of Tokyo, Kashiwa, Chiba 277-8583, Japan

[137] California Institute of Technology, 1200 E California Blvd, Pasadena, CA 91125, USA

[138] Department of Physics & Astronomy, University of California Irvine, Irvine CA 92697, USA

[139] Department of Mathematics and Physics E. De Giorgi, University of Salento, Via per Arnesano, CP-I93, 73100, Lecce, Italy

[140] INFN, Sezione di Lecce, Via per Arnesano, CP-193, 73100, Lecce, Italy

[141] INAF-Sezione di Lecce, c/o Dipartimento Matematica e Fisica, Via per Arnesano, 73100, Lecce, Italy

[142] Departamento Física Aplicada, Universidad Politécnica de Cartagena, Campus Muralla del Mar, 30202 Cartagena, Murcia, Spain

[143] CEA Saclay, DFR/IRFU, Service d'Astrophysique, Bat. 709, 91191 Gif-sur-Yvette, France

[144] Institute of Cosmology and Gravitation, University of Portsmouth, Portsmouth PO1 3FX, UK

[145] Department of Computer Science, Aalto University, PO Box 15400, Espoo, FI-00 076, Finland

[146] Instituto de Astrofísica de Canarias, c/ Via Lactea s/n, La Laguna 38200, Spain. Departamento de Astrofísica de la Universidad de La Laguna, Avda. Francisco Sanchez, La Laguna, 38200, Spain

[147] Caltech/IPAC, 1200 E. California Blvd., Pasadena, CA 91125, USA

[148] Ruhr University Bochum, Faculty of Physics and Astronomy, Astronomical Institute (AIRUB), German Centre for Cosmological Lensing (GCCL), 44780 Bochum, Germany

[149] Department of Physics and Astronomy, Vesilinnantie 5, 20014 University of Turku, Finland

[150] Serco for European Space Agency (ESA), Camino bajo del Castillo, s/n, Urbanizacion Villafranca del Castillo, Villanueva de la Cañada, 28692 Madrid, Spain

[151] ARC Centre of Excellence for Dark Matter Particle Physics, Melbourne, Australia

[152] Centre for Astrophysics & Supercomputing, Swinburne University of Technology, Hawthorn, Victoria 3122, Australia

[153] Department of Physics and Astronomy, University of the Western Cape, Bellville, Cape Town, 7535, South Africa

[154] DAMTP, Centre for Mathematical Sciences, Wilberforce Road, Cambridge CB3 0WA, UK

[155] Kavli Institute for Cosmology Cambridge, Madingley Road, Cambridge, CB3 0HA, UK

[156] Department of Astrophysics, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland

[157] Department of Physics, Centre for Extragalactic Astronomy, Durham University, South Road, Durham, DH1 3LE, UK

[158] IRFU, CEA, Université Paris-Saclay 91191 Gif-sur-Yvette Cedex, France

[159] Oskar Klein Centre for Cosmoparticle Physics, Department of Physics, Stockholm University, Stockholm, SE-106 91, Sweden

[160] Astrophysics Group, Blackett Laboratory, Imperial College London, London SW7 2AZ, UK

[161] Univ. Grenoble Alpes, CNRS, Grenoble INP, LPSC-IN2P3, 53, Avenue des Martyrs, 38000, Grenoble, France

[162] INAF-Osservatorio Astrofisico di Arcetri, Largo E. Fermi 5, 50125, Firenze, Italy

[163] Dipartimento di Fisica, Sapienza Università di Roma, Piazzale Aldo Moro 2, 00185 Roma, Italy

[164] Centro de Astrofísica da Universidade do Porto, Rua das Estrelas, 4150-762 Porto, Portugal

[165] HE Space for European Space Agency (ESA), Camino bajo del Castillo, s/n, Urbanizacion Villafranca del Castillo, Villanueva de la Cañada, 28692 Madrid, Spain

[166] Dipartimento di Fisica - Sezione di Astronomia, Università di Trieste, Via Tiepolo 11, 34131 Trieste, Italy

[167] Department of Astrophysical Sciences, Peyton Hall, Princeton University, Princeton, NJ 08544, USA

[168] Theoretical astrophysics, Department of Physics and Astronomy, Uppsala University, Box 515, 751 20 Uppsala, Sweden

[169] Mathematical Institute, University of Leiden, Einsteinweg 55, 2333 CA Leiden, The Netherlands

[170] Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK

[171] Space physics and astronomy research unit, University of Oulu, Pentti Kaiteran katu 1, FI-90014 Oulu, Finland

[172] Univ. Lille, CNRS, Centrale Lille, UMR 9189 CRIStAL, 59000 Lille, France

Appendix A: Hyperparameters, compute, and carbon emissions

Table A.1: Hyperparameters used in pre-training `AstroPT-90M`.

Parameter	`AstroPT` pre-training
Model architecture	
Active parameters	90 millions
Number of layers	12
Model dimension	768
Number of heads	12
Training configuration	
Batch size	125 440
Learning rate	0.0006
Weight decay	0.1
Warmup steps	2000
Training steps	5000
Optimiser	AdamW

The training of deep-learning models requires considerable energy, which can contribute to carbon emissions. To counteract further emission from unnecessary retraining we follow the recommendations of (Strubell et al. 2019) and make available our model weights and inference codein the Hugging Face repository[3]. While this study made use of the 100% renewable energy-powered ITER Teide HPC cluster (Mampaso et al. 2022) for pre-training, for the sake of transparency we estimate and show our energy usage during the pre-training of `AstroPT` here: `AstroPT-90M` takes 2 hours to train on 4xA100-40G NVIDIA GPUs, which corresponds to 2 kWh of energy use according to the Machine Learning CO2 Impact Calculator (`mlco2.github.io`).

Appendix B: `AstroPT` embeddings for VIS and VIS+NISP imaging

Figure B.1 presents a visualisation of the `AstroPT` embeddings trained on the VIS (left) and combined VIS and NISP data (right). This plot demonstrates how the inclusion of both optical and infrared wavelengths influences the organisation and clustering of different galaxy types and stars in the embedding space.
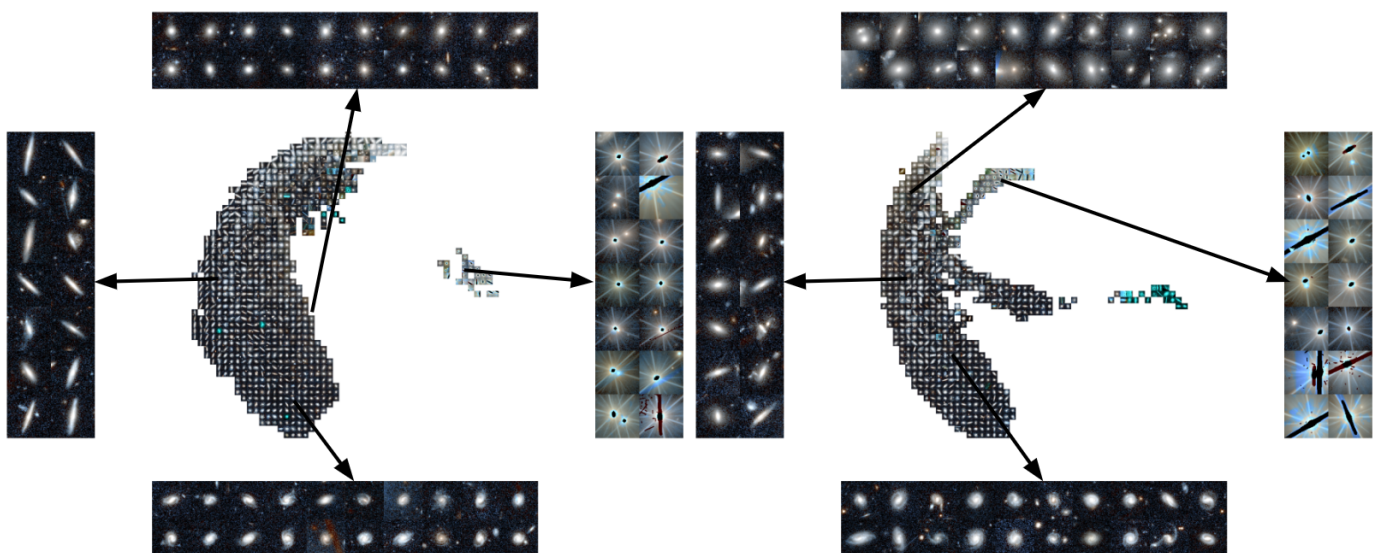
Fig. B.1: UMAP visualisations of the embeddings from `AstroPT` trained on VIS (left) and VIS+NISP (right) images with example cutouts, see text for more details.

In Fig. B.2, we show typical anomalies found with LOF (upper panel) and IF (bottom panel) for the VIS+NISP embedding. As can be observed, anomalies consist of images that include multiple objects within the cutouts.
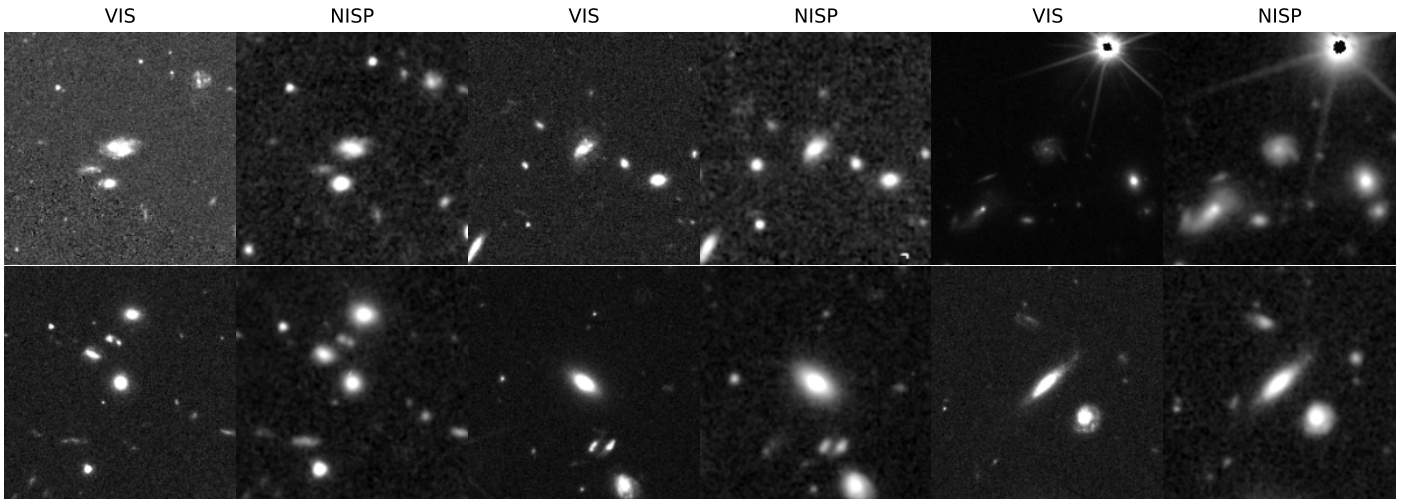
Fig. B.2: Example VIS and NISP cutouts of anomaly detections found on VIS+NISP embeddings with IF (upper panel) and LOF (bottom panel).
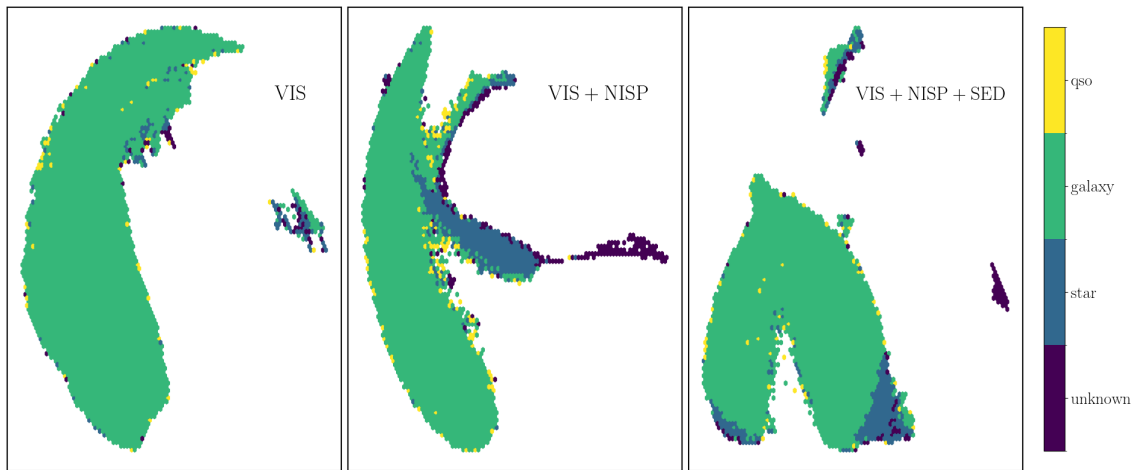
Fig. C.1: UMAP visualisations of the self-supervised embeddings from `AstroPT` coloured by the source type (QSO, galaxy, star, or unknown). The left panel corresponds to VIS-only embeddings, the centre panel includes both VIS and NISP data, and the right panel includes VIS, NISP, and SED data.

Appendix C: Star-QSO-galaxy separation

Figure C.1 shows the correlation of the embeddings with the source type (unknown, galaxy, star, QSO) derived from the spectroscopic pipeline. Clearly, the outlier regions of embeddings are occupied by stars and unknown sources, but stars are also found over all regions (see Fig. C.1). Galaxies, QSOs, and stars tend to occupy a broad distribution. Adding NISP data sharpens the separation of unknown objects where a long tail gathers the invalid NISP data. Incorporating SEDs further enhance the separation of unknown sources from the main space.

Appendix D: Quality of the embeddings: Physical properties

Figures D.1, and D.2 highlight the relationship between the latent space and physical properties, such as stellar mass, and redshift. There is a clear connection between morphological trends (see Fig. 8) and stellar masses. Early-type galaxies, which are typically more massive, occupy regions of the UMAP space associated with higher stellar masses. In contrast, spiral galaxies, known for their lower stellar masses, cluster in distinct, separate regions. This differentiation is especially pronounced when embeddings are trained with VIS+NISP data, suggesting that the additional NISP features help to better constrain stellar mass distributions. This

---

[3] https://huggingface.co/collections/msiudek/

is further enhanced when the SEDs are included in the training. The *Euclid* photometric redshift shows no strong correlation with the embeddings, although there is a subtle trend of higher stellar masses being linked to slightly higher redshifts, especially for VIS+NIS+SED, consistent with the evolutionary connection between galaxy mass and formation epoch. The trend with the redshift is clearer when the spectroscopic redshift from DESI is used. These results underscore the utility of self-supervised representations in capturing key physical galaxy properties, particularly with multi-band data.

Appendix E: Accuracy of predictions

Table E.1, E.2, and E.3 provide a detailed summary of the performance metrics for morphology classification, photo-z, and stellar mass predictions, respectively, across different models and data configurations (see Sect. 3.2). Metrics are reported for various percentages of labelled data used during training. These tables complements Figs. 10, 12, 14 by offering a more detailed view of the performance differences between the models. MLP model consistently achieves the best overall metrics, particularly when leveraging self-supervised embeddings with even a minimal set of labels. The linear model also demonstrates competitive performance, and serves as a robust baseline for comparison.
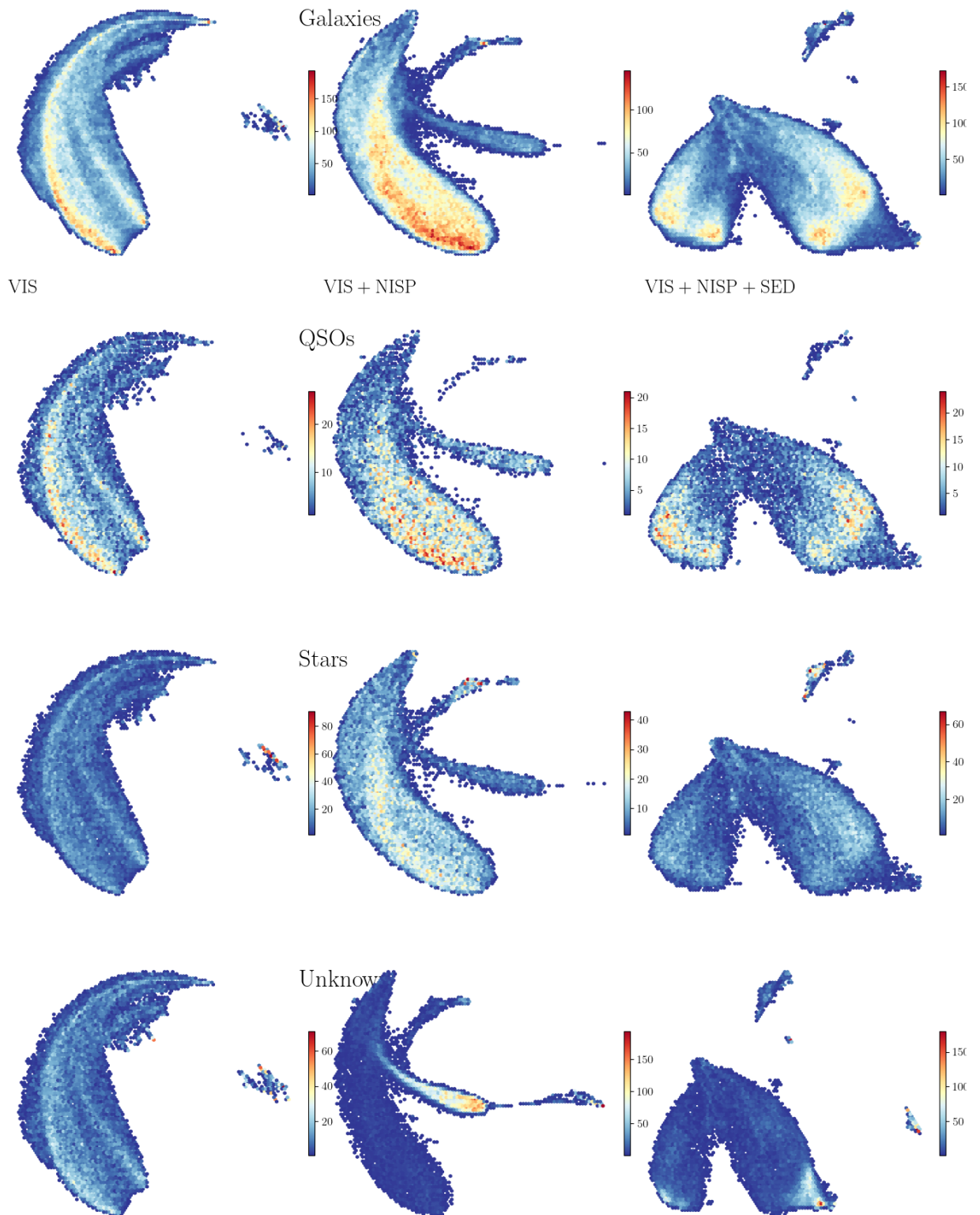
Fig. C.2: UMAP visualisations of the self-supervised embeddings from `AstroPT`. *Top*: Coloured by the stars. *Bottom*: Colorued by QSO. Left panels correspond to VIS-only embeddings, centre panels include both VIS, and NISP data, and the right panels include VIS, NISP, and SED data.
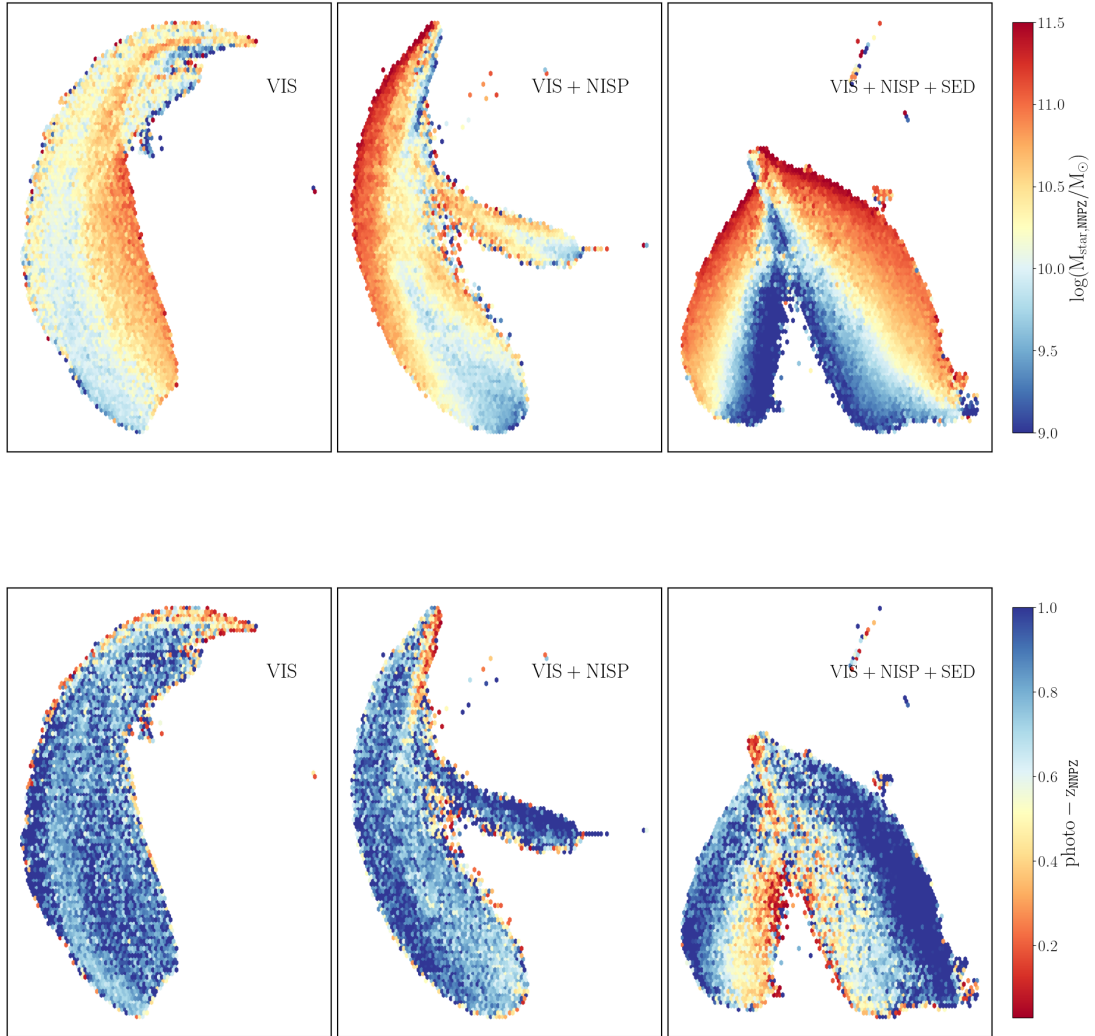
Fig. D.1: UMAP visualisation of self-supervised embeddings, highlighting correlations with stellar mass and photometric redshift. The left panels show the distribution of galaxies when `AstroPT` is trained on VIS images, while the centre panels represent the training with VIS+NISP images, and the right panels correspond to training on VIS, NISP images, and SEDs. The top UMAP visualisations are colour-coded by stellar mass from *Euclid* official pipeline (top). The bottom UMAP visualisations are colour-coded by *Euclid* official photometric redshifts.
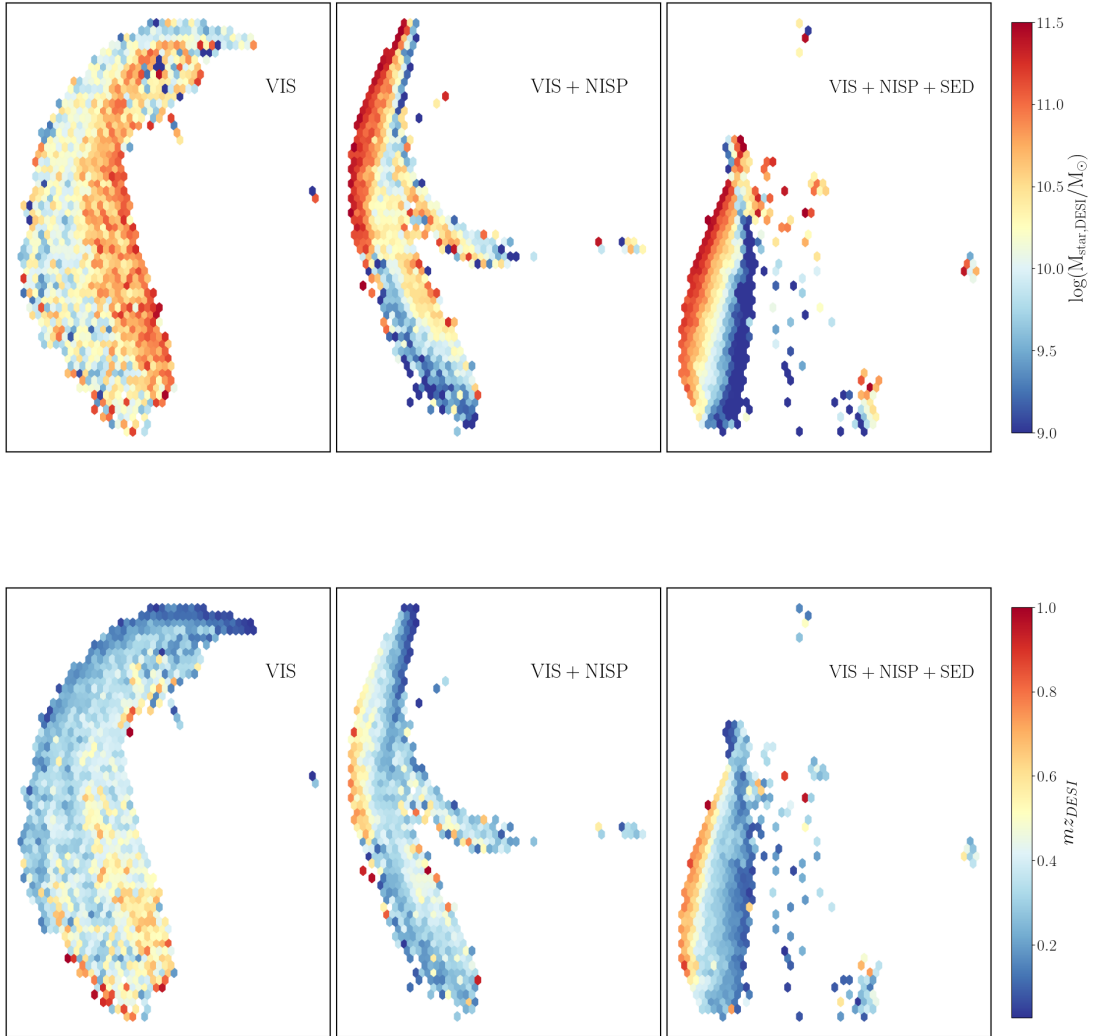
Fig. D.2: UMAP visualisation of self-supervised embeddings, highlighting correlations with stellar mass and photometric redshift. The left panels show the distribution of galaxies when `AstroPT` is trained on VIS images, while the centre panels represent the training with VIS+NISP images, and the right panels correspond to training on VIS, NISP images, and SEDs. The left UMAP visualisations are colour-coded by stellar mass from DESI VAC (Siudek et al. 2024). The right UMAP visualisations are colour-coded by the DESI spectroscopic redshift (DESI Collaboration et al. 2024b) available for about 5% of the sample.

Table E.1: Comparison of early-type/late-type galaxy classification performance for different approaches. Results are reported for different percentages of labelled data used for training (1%, 5%, 10%, 50%, 80%, and 100%).

Model	Data	Percentage	Accuracy	Precision	Recall	F1 score	FPR
Supervised CNN	VIS	1	0.68 ± 0.02	0.42 ± 0.02	0.51 ± 0.05	0.46 ± 0.02	0.25 ± 0.04
Supervised CNN	VIS	5	0.74 ± 0.01	0.58 ± 0.02	0.55 ± 0.03	0.56 ± 0.01	0.18 ± 0.02
Supervised CNN	VIS	10	0.75 ± 0.01	0.59 ± 0.03	0.60 ± 0.05	0.59 ± 0.01	0.18 ± 0.03
Supervised CNN	VIS	20	0.79 ± 0.01	0.67 ± 0.02	0.62 ± 0.02	0.64 ± 0.01	0.14 ± 0.02
Supervised CNN	VIS	50	0.81 ± 0.01	0.72 ± 0.01	0.67 ± 0.02	0.69 ± 0.01	0.12 ± 0.01
Supervised CNN	VIS	80	0.83 ± 0.01	0.72 ± 0.02	0.73 ± 0.02	0.72 ± 0.00	0.13 ± 0.02
Supervised CNN	VIS	100	0.84 ± 0.01	0.75 ± 0.02	0.71 ± 0.03	0.73 ± 0.01	0.11 ± 0.02
MLP model	VIS	1	0.83 ± 0.02	0.74 ± 0.04	0.72 ± 0.04	0.73 ± 0.03	0.12 ± 0.02
MLP model	VIS	5	0.86 ± 0.01	0.78 ± 0.02	0.78 ± 0.02	0.78 ± 0.01	0.10 ± 0.01
MLP model	VIS	10	0.87 ± 0.01	0.79 ± 0.01	0.78 ± 0.01	0.79 ± 0.01	0.10 ± 0.01
MLP model	VIS	20	0.87 ± 0.00	0.79 ± 0.01	0.80 ± 0.01	0.79 ± 0.01	0.09 ± 0.00
MLP model	VIS	50	0.87 ± 0.00	0.80 ± 0.01	0.80 ± 0.01	0.80 ± 0.00	0.09 ± 0.00
MLP model	VIS	80	0.87 ± 0.00	0.80 ± 0.01	0.80 ± 0.01	0.80 ± 0.00	0.09 ± 0.00
MLP model	VIS	100	0.88 ± 0.00	0.80 ± 0.01	0.80 ± 0.01	0.80 ± 0.00	0.09 ± 0.00
Linear model	VIS	1	0.84 ± 0.02	0.76 ± 0.04	0.71 ± 0.03	0.73 ± 0.03	0.10 ± 0.02
Linear model	VIS	5	0.86 ± 0.01	0.80 ± 0.02	0.72 ± 0.02	0.76 ± 0.01	0.08 ± 0.01
Linear model	VIS	10	0.86 ± 0.01	0.80 ± 0.01	0.72 ± 0.01	0.76 ± 0.01	0.08 ± 0.01
Linear model	VIS	20	0.86 ± 0.00	0.81 ± 0.01	0.73 ± 0.01	0.77 ± 0.01	0.08 ± 0.00
Linear model	VIS	50	0.86 ± 0.00	0.81 ± 0.00	0.73 ± 0.01	0.77 ± 0.00	0.08 ± 0.00
Linear model	VIS	80	0.86 ± 0.00	0.81 ± 0.01	0.73 ± 0.01	0.77 ± 0.00	0.08 ± 0.00
Linear model	VIS	100	0.86 ± 0.00	0.81 ± 0.01	0.73 ± 0.01	0.77 ± 0.00	0.08 ± 0.00
Supervised CNN	VIS+NISP	1	0.69 ± 0.02	0.46 ± 0.03	0.50 ± 0.06	0.47 ± 0.04	0.24 ± 0.03
Supervised CNN	VIS+NISP	5	0.74 ± 0.01	0.60 ± 0.03	0.54 ± 0.05	0.57 ± 0.02	0.16 ± 0.03
Supervised CNN	VIS+NISP	10	0.76 ± 0.01	0.61 ± 0.02	0.61 ± 0.02	0.61 ± 0.01	0.17 ± 0.02
Supervised CNN	VIS+NISP	20	0.78 ± 0.01	0.67 ± 0.01	0.63 ± 0.04	0.65 ± 0.02	0.15 ± 0.02
Supervised CNN	VIS+NISP	50	0.82 ± 0.01	0.72 ± 0.01	0.67 ± 0.02	0.69 ± 0.01	0.12 ± 0.01
Supervised CNN	VIS+NISP	80	0.83 ± 0.00	0.74 ± 0.01	0.72 ± 0.02	0.73 ± 0.01	0.11 ± 0.01
Supervised CNN	VIS+NISP	100	0.83 ± 0.01	0.74 ± 0.02	0.72 ± 0.02	0.73 ± 0.00	0.12 ± 0.02
MLP model	VIS+NISP	1	0.80 ± 0.02	0.69 ± 0.05	0.66 ± 0.05	0.67 ± 0.04	0.14 ± 0.02
MLP model	VIS+NISP	5	0.83 ± 0.01	0.74 ± 0.02	0.72 ± 0.02	0.73 ± 0.02	0.12 ± 0.01
MLP model	VIS+NISP	10	0.84 ± 0.01	0.76 ± 0.01	0.74 ± 0.01	0.75 ± 0.01	0.11 ± 0.01
MLP model	VIS+NISP	20	0.85 ± 0.00	0.77 ± 0.01	0.74 ± 0.01	0.75 ± 0.01	0.10 ± 0.00
MLP model	VIS+NISP	50	0.85 ± 0.00	0.77 ± 0.01	0.74 ± 0.01	0.76 ± 0.00	0.10 ± 0.00
MLP model	VIS+NISP	80	0.85 ± 0.00	0.77 ± 0.01	0.74 ± 0.01	0.76 ± 0.00	0.10 ± 0.00
MLP model	VIS+NISP	100	0.85 ± 0.00	0.77 ± 0.00	0.74 ± 0.01	0.76 ± 0.00	0.10 ± 0.00
Linear model	VIS+NISP	1	0.80 ± 0.02	0.70 ± 0.05	0.64 ± 0.04	0.67 ± 0.04	0.12 ± 0.02
Linear model	VIS+NISP	5	0.83 ± 0.01	0.76 ± 0.02	0.67 ± 0.02	0.71 ± 0.01	0.09 ± 0.01
Linear model	VIS+NISP	10	0.84 ± 0.01	0.78 ± 0.01	0.67 ± 0.01	0.72 ± 0.01	0.09 ± 0.01
Linear model	VIS+NISP	20	0.84 ± 0.00	0.78 ± 0.01	0.67 ± 0.01	0.72 ± 0.01	0.09 ± 0.01
Linear model	VIS+NISP	50	0.84 ± 0.00	0.79 ± 0.01	0.67 ± 0.01	0.72 ± 0.00	0.08 ± 0.00
Linear model	VIS+NISP	80	0.84 ± 0.00	0.79 ± 0.01	0.67 ± 0.01	0.72 ± 0.00	0.08 ± 0.00
Linear model	VIS+NISP	100	0.84 ± 0.00	0.79 ± 0.00	0.67 ± 0.01	0.72 ± 0.00	0.08 ± 0.00
MLP model	VIS+NISP+SED	1	0.79 ± 0.06	0.72 ± 0.16	0.65 ± 0.14	0.67 ± 0.11	0.13 ± 0.08
MLP model	VIS+NISP+SED	5	0.80 ± 0.03	0.69 ± 0.07	0.66 ± 0.07	0.67 ± 0.06	0.14 ± 0.03
MLP model	VIS+NISP+SED	10	0.81 ± 0.03	0.71 ± 0.05	0.69 ± 0.05	0.70 ± 0.04	0.13 ± 0.03
MLP model	VIS+NISP+SED	20	0.82 ± 0.02	0.74 ± 0.03	0.72 ± 0.04	0.73 ± 0.03	0.12 ± 0.02
MLP model	VIS+NISP+SED	50	0.84 ± 0.01	0.76 ± 0.02	0.75 ± 0.02	0.76 ± 0.02	0.11 ± 0.01
MLP Model	VIS+NISP+SED	80	0.85 ± 0.01	0.78 ± 0.01	0.77 ± 0.01	0.77 ± 0.01	0.11 ± 0.01
MLP model	VIS+NISP+SED	100	0.86 ± 0.01	0.78 ± 0.01	0.78 ± 0.01	0.78 ± 0.01	0.10 ± 0.01
Linear model	VIS+NISP+SED	1	0.79 ± 0.06	0.71 ± 0.15	0.65 ± 0.14	0.66 ± 0.11	0.13 ± 0.07
Linear model	VIS+NISP+SED	5	0.80 ± 0.03	0.69 ± 0.07	0.67 ± 0.07	0.68 ± 0.05	0.14 ± 0.04
Linear model	VIS+NISP+SED	10	0.81 ± 0.02	0.71 ± 0.05	0.69 ± 0.05	0.70 ± 0.04	0.13 ± 0.03
Linear model	VIS+NISP+SED	20	0.82 ± 0.01	0.74 ± 0.03	0.70 ± 0.04	0.72 ± 0.03	0.12 ± 0.02
Linear model	VIS+NISP+SED	50	0.84 ± 0.01	0.78 ± 0.02	0.71 ± 0.02	0.74 ± 0.02	0.10 ± 0.01
Linear model	VIS+NISP+SED	80	0.85 ± 0.01	0.78 ± 0.01	0.72 ± 0.01	0.75 ± 0.01	0.10 ± 0.01
Linear model	VIS+NISP+SED	100	0.85 ± 0.01	0.78 ± 0.01	0.73 ± 0.01	0.75 ± 0.01	0.10 ± 0.01

Table E.2: Comparison of photo-*z* prediction performance for different approaches. Results are reported for different percentages of labelled data used for training (1%, 5%, 10%, 50%, 80%, and 100%).

Model	Data	Percentage	Bias	NMAD	Outlier fraction (%)
Supervised CNN	VIS	1	-0.046 ± 0.084	0.140 ± 0.012	40.3 ± 13.6
Supervised CNN	VIS	5	0.003 ± 0.030	0.159 ± 0.008	34.9 ± 2.2
Supervised CNN	VIS	10	-0.002 ± 0.036	0.150 ± 0.008	34.6 ± 2.4
Supervised CNN	VIS	20	0.019 ± 0.026	0.142 ± 0.009	32.3 ± 1.6
Supervised CNN	VIS	50	0.003 ± 0.009	0.124 ± 0.004	27.1 ± 1.2
Supervised CNN	VIS	80	0.021 ± 0.018	0.135 ± 0.024	30.9 ± 8.0
Supervised CNN	VIS	100	0.026 ± 0.004	0.118 ± 0.002	25.1 ± 0.7
MLP model	VIS	1	-0.000 ± 0.044	0.120 ± 0.025	23.9 ± 9.0
MLP model	VIS	5	-0.009 ± 0.036	0.111 ± 0.013	20.9 ± 4.1
MLP model	VIS	10	-0.012 ± 0.036	0.104 ± 0.010	19.9 ± 3.1
MLP model	VIS	20	-0.002 ± 0.022	0.099 ± 0.006	16.5 ± 2.2
MLP model	VIS	50	-0.000 ± 0.024	0.092 ± 0.005	14.0 ± 1.2
MLP model	VIS	80	-0.000 ± 0.026	0.090 ± 0.006	13.3 ± 2.0
MLP model	VIS	100	-0.002 ± 0.022	0.089 ± 0.004	12.7 ± 1.7
Linear model	VIS	1	0.015 ± 0.030	0.145 ± 0.023	32.5 ± 7.1
Linear model	VIS	5	-0.002 ± 0.022	0.265 ± 0.027	57.2 ± 4.5
Linear model	VIS	10	0.013 ± 0.014	0.175 ± 0.011	39.1 ± 3.2
Linear model	VIS	20	0.011 ± 0.007	0.123 ± 0.006	23.5 ± 1.7
Linear model	VIS	50	0.011 ± 0.003	0.105 ± 0.003	17.1 ± 0.9
Linear model	VIS	80	0.011 ± 0.003	0.100 ± 0.002	15.5 ± 0.7
Linear model	VIS	100	0.011 ± 0.003	0.100 ± 0.002	15.1 ± 0.6
Supervised CNN	VIS+NISP	1	-0.115 ± 0.104	0.130 ± 0.013	48.4 ± 19.3
Supervised CNN	VIS+NISP	5	0.001 ± 0.047	0.166 ± 0.017	39.0 ± 2.6
Supervised CNN	VIS+NISP	10	0.009 ± 0.024	0.148 ± 0.010	32.8 ± 2.6
Supervised CNN	VIS+NISP	50	-0.009 ± 0.005	0.110 ± 0.004	22.4 ± 0.6
Supervised CNN	VIS+NISP	80	0.013 ± 0.012	0.100 ± 0.001	20.7 ± 0.9
Supervised CNN	VIS+NISP	100	0.009 ± 0.006	0.103 ± 0.003	19.9 ± 0.3
MLP Model CNN	VIS+NISP	1	-0.022 ± 0.030	0.098 ± 0.024	14.7 ± 6.1
MLP model	VIS+NISP	1	-0.022 ± 0.030	0.098 ± 0.024	14.7 ± 6.1
MLP model	VIS+NISP	5	-0.013 ± 0.020	0.075 ± 0.008	8.8 ± 2.5
MLP model	VIS+NISP	10	-0.013 ± 0.019	0.070 ± 0.006	7.3 ± 1.7
MLP model	VIS+NISP	20	-0.005 ± 0.016	0.066 ± 0.005	5.6 ± 1.5
MLP model	VIS+NISP	50	-0.010 ± 0.013	0.060 ± 0.004	4.2 ± 0.7
MLP model	VIS+NISP	80	-0.015 ± 0.012	0.058 ± 0.004	3.8 ± 0.6
MLP model	VIS+NISP	100	-0.014 ± 0.012	0.057 ± 0.004	3.6 ± 0.5
Linear model	VIS+NISP	1	0.008 ± 0.031	0.119 ± 0.032	22.6 ± 8.9
Linear model	VIS+NISP	5	0.009 ± 0.022	0.200 ± 0.030	45.4 ± 6.8
Linear model	VIS+NISP	10	0.005 ± 0.008	0.125 ± 0.013	24.2 ± 3.7
Linear model	VIS+NISP	20	0.005 ± 0.004	0.085 ± 0.006	10.2 ± 1.5
Linear model	VIS+NISP	50	0.005 ± 0.002	0.067 ± 0.002	5.6 ± 0.6
Linear model	VIS+NISP	80	0.006 ± 0.002	0.064 ± 0.002	5.1 ± 0.4
Linear model	VIS+NISP	100	0.006 ± 0.002	0.063 ± 0.001	4.8 ± 0.4
MLP model	VIS+NISP+SED	1	-0.023 ± 0.029	0.071 ± 0.016	7.5 ± 6.7
MLP model	VIS+NISP+SED	5	-0.018 ± 0.017	0.062 ± 0.007	5.4 ± 2.1
MLP model	VIS+NISP+SED	10	-0.012 ± 0.015	0.057 ± 0.005	4.0 ± 1.5
MLP model	VIS+NISP+SED	20	-0.008 ± 0.014	0.054 ± 0.004	3.1 ± 1.0
MLP model	VIS+NISP+SED	50	-0.014 ± 0.014	0.050 ± 0.003	2.7 ± 0.9
MLP model	VIS+NISP+SED	80	-0.008 ± 0.013	0.049 ± 0.003	2.3 ± 0.4
MLP model	VIS+NISP+SED	100	-0.009 ± 0.012	0.048 ± 0.002	2.1 ± 0.4
Linear model	VIS+NISP+SED	1	0.002 ± 0.025	0.088 ± 0.026	12.8 ± 6.5
Linear model	VIS+NISP+SED	5	0.007 ± 0.017	0.171 ± 0.031	37.5 ± 8.5
Linear model	VIS+NISP+SED	10	0.003 ± 0.006	0.107 ± 0.013	17.2 ± 4.2
Linear model	VIS+NISP+SED	20	0.004 ± 0.004	0.071 ± 0.005	6.3 ± 1.3
Linear model	VIS+NISP+SED	50	0.004 ± 0.002	0.056 ± 0.002	3.3 ± 0.5
Linear model	VIS+NISP+SED	80	0.004 ± 0.002	0.053 ± 0.001	2.9 ± 0.3
Linear model	VIS+NISP+SED	100	0.004 ± 0.001	0.052 ± 0.001	2.7 ± 0.3
Euclid photo-*z*		100	-0.027	0.037	2.21

Table E.3: Comparison of stellar mass prediction performance for different approaches. Results are reported for different percentages of labelled data used for training (1%, 10%, 50%, 80%, and 100%).

Model	Data	Percentage	Bias	NMAD	Outlier fraction (%)
Supervised CNN	VIS	1	-0.018 ± 0.039	0.059 ± 0.007	1.7 ± 1.7
Supervised CNN	VIS	5	-0.005 ± 0.027	0.060 ± 0.005	1.2 ± 0.6
Supervised CNN	VIS	10	0.004 ± 0.021	0.065 ± 0.002	1.7 ± 0.6
Supervised CNN	VIS	20	-0.006 ± 0.024	0.066 ± 0.004	1.3 ± 0.3
Supervised CNN	VIS	50	-0.017 ± 0.019	0.064 ± 0.002	0.9 ± 0.4
Supervised CNN	VIS	80	0.001 ± 0.021	0.063 ± 0.001	1.1 ± 0.5
Supervised CNN	VIS	100	0.004 ± 0.017	0.063 ± 0.001	1.2 ± 0.4
MLP model	VIS	1	0.002 ± 0.016	0.071 ± 0.017	6.2 ± 4.6
MLP model	VIS	5	0.004 ± 0.007	0.057 ± 0.006	4.3 ± 1.8
MLP model	VIS	10	0.006 ± 0.004	0.056 ± 0.005	4.0 ± 0.9
MLP model	VIS	20	0.007 ± 0.003	0.053 ± 0.003	3.8 ± 0.9
MLP model	VIS	50	0.006 ± 0.002	0.052 ± 0.002	3.7 ± 0.6
MLP model	VIS	80	0.006 ± 0.002	0.051 ± 0.002	3.7 ± 0.4
MLP model	VIS	100	0.007 ± 0.002	0.052 ± 0.002	3.8 ± 0.4
Linear model	VIS	1	0.005 ± 0.016	0.075 ± 0.015	7.4 ± 4.9
Linear model	VIS	5	0.002 ± 0.013	0.120 ± 0.015	21.7 ± 4.2
Linear model	VIS	10	0.003 ± 0.006	0.092 ± 0.008	12.8 ± 2.1
Linear model	VIS	20	0.004 ± 0.004	0.063 ± 0.004	4.5 ± 1.1
Linear model	VIS	50	0.003 ± 0.002	0.053 ± 0.002	3.1 ± 0.4
Linear model	VIS	80	0.003 ± 0.001	0.050 ± 0.001	2.9 ± 0.3
Linear model	VIS	100	0.004 ± 0.001	0.050 ± 0.001	3.0 ± 0.3
Supervised CNN	VIS+NISP	1	-0.020 ± 0.029	0.062 ± 0.008	0.3 ± 1.0
Supervised CNN	VIS+NISP	5	0.008 ± 0.018	0.064 ± 0.005	0.2 ± 0.5
Supervised CNN	VIS+NISP	10	0.015 ± 0.021	0.053 ± 0.002	1.0 ± 0.8
Supervised CNN	VIS+NISP	20	-0.003 ± 0.037	0.050 ± 0.004	0.8 ± 0.5
Supervised CNN	VIS+NISP	50	0.004 ± 0.019	0.047 ± 0.002	0.8 ± 0.3
Supervised CNN	VIS+NISP	80	0.001 ± 0.024	0.045 ± 0.001	0.4 ± 0.2
Supervised CNN	VIS+NISP	100	0.011 ± 0.007	0.044 ± 0.001	0.5 ± 0.1
MLP model	VIS+NISP	1	0.004 ± 0.009	0.036 ± 0.009	1.0 ± 1.7
MLP model	VIS+NISP	5	0.002 ± 0.003	0.030 ± 0.003	0.7 ± 0.7
MLP model	VIS+NISP	10	0.002 ± 0.002	0.028 ± 0.003	0.7 ± 0.5
MLP model	VIS+NISP	20	0.001 ± 0.002	0.027 ± 0.003	0.6 ± 0.4
MLP model	VIS+NISP	50	0.002 ± 0.002	0.028 ± 0.003	0.7 ± 0.4
MLP model	VIS+NISP	80	0.002 ± 0.001	0.026 ± 0.003	0.6 ± 0.3
MLP model	VIS+NISP	100	0.002 ± 0.001	0.027 ± 0.003	0.6 ± 0.3
Linear model	VIS+NISP	1	0.005 ± 0.008	0.038 ± 0.009	1.0 ± 1.9
Linear model	VIS+NISP	5	0.001 ± 0.006	0.056 ± 0.006	1.7 ± 1.0
Linear model	VIS+NISP	10	0.000 ± 0.003	0.043 ± 0.003	0.7 ± 0.5
Linear model	VIS+NISP	20	0.000 ± 0.002	0.029 ± 0.002	0.3 ± 0.2
Linear model	VIS+NISP	50	0.001 ± 0.001	0.024 ± 0.001	0.3 ± 0.1
Linear model	VIS+NISP	80	0.001 ± 0.001	0.023 ± 0.001	0.3 ± 0.1
Linear model	VIS+NISP	100	0.001 ± 0.001	0.023 ± 0.001	0.3 ± 0.1
Euclid $\log_{10}(M_*/M_\odot)$		100	0.0011	0.0299	2.34