Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **528,** 4188–4208 (2024) https://doi.org/10.1093/mnras/stad3956
Advance Access publication 2023 December 22

An extended catalogue of galaxy morphology using deep learning in southern photometric local universe survey data release 3

C. R. Bom [1,2★] A. Cortesi,[3] U. Ribeiro,[1] L. O. Dias,[1] K. Kelkar,[4] A. V. Smith Castelli,[5,6] L. Santana-Silva,[7] V. Lopes-Silva,[3] T. S. Gonçalves,[3] L. R. Abramo,[8] E. V. R. Lima,[9] F. Almeida-Fernandes,[9] L. Espinosa,[9] L. Li,[9] M. L. Buzzo [10] C. Mendes de Oliveira,[9] L. Sodré, Jr [9] F. Ferrari [11] A. Alvarez-Candal [12] M. Grossi [3] E. Telles,[12] S. Torres-Flores,[13] S. V. Werner [14] A. Kanaan,[15] T. Ribeiro[16] and W. Schoenell[17]

[1]*Centro Brasileiro de Pesquisas Físicas, Rua Dr Xavier Sigaud 150, CEP 22290-180, Rio de Janeiro, RJ, Brazil*
[2]*Centro Federal de Educação Tecnológica Celso Suckow da Fonseca, Rodovia Mário Covas, lote J2, quadra J, CEP 23810-000, Itaguaí, RJ, Brazil*
[3]*Valongo Observatory, Federal University of Rio de Janeiro, Ladeira Pedro Antonio 43, Saude Rio de Janeiro, RJ 20080-090, Brazil*
[4]*Universidad Técnica Federico Santa María, Av. España 1680, 2340000, Valparaíso, Chile*
[5]*Instituto de Astrofśica de La Plata, CONICET–UNLP, Paseo del Bosque s/n, B1900FWA La Plata, Argentina*
[6]*Facultad de Ciencias Astrónomicas y Geofísicas, UNLP, Paseo del Bosque s/n, B1900FWA La Plata, Argentina*
[7]*NAT-Universidade Cruzeiro do Sul / Universidade Cidade de São Paulo, Rua Galvão Bueno, 868, 01506-000 São Paulo, SP, Brazil*
[8]*Departamento de Física Matemática, Instituto de Física, Universidade de São Paulo, R. do Matão 1371, 05508-090 São Paulo, SP, Brazil*
[9]*IAG, Universidade de São Paulo, Rua do Matão 1225, São Paulo, SP, Brazil*
[10]*Centre for Astrophysics and Supercomputing, Swinburne University, John Street, Hawthorn, VIC 3122, Australia*
[11]*Instituto de Matemática, Estatística e Física, Universidade Federal do Rio Grande, 8 km Itália Av., 96203-900, Rio Grande, RS, Brazil*
[12]*Observatório Nacional R. Gen. José Cristino, 77, 20921-400 Rio de Janeiro, Brazil*
[13]*Departamento de Astronomia, Universidad de La Serena, Av. Cisternas 1200, La Serena, Chile*
[14]*School of Physics and Astronomy, University of Nottingham, Nottingham NG7 2RD, UK*
[15]*Departamento de Física, Universidade Federal de Santa Catarina, Florianópolis, SC 88040-900, Brazil*
[16]*Departamento de Astronomia, Instituto de Física, Universidade Federal do Rio Grande do Sul (UFRGS), Av. Bento Goncalves 9500, 91501-970, Porto Alegre, Brazil*
[17]*NOAO, P.O. Box 26732, Tucson, AZ 85726, USA*

Accepted 2023 December 18. Received 2023 December 18; in original form 2023 June 8

ABSTRACT

The morphological diversity of galaxies is a relevant probe of galaxy evolution and cosmological structure formation. However, in large sky surveys, even the morphological classification of galaxies into two classes, like late-type (LT) and early-type (ET), still represents a significant challenge. In this work, we present a Deep Learning (DL) based morphological catalogue built from images obtained by the Southern Photometric Local Universe Survey (S-PLUS) Data Release 3 (DR3). Our DL method achieves a purity rate of 98.5 per cent in accurately distinguishing between spiral, as part of the larger category of LT galaxies, and elliptical, belonging to ET galaxies. Additionally, we have implemented a secondary classifier that evaluates the quality of each galaxy stamp, which allows to select only high-quality images when studying properties of galaxies on the basis of their DL morphology. From our LT/ET catalogue of galaxies, we recover the expected colour–magnitude diagram in which LT galaxies display bluer colours than ET ones. Furthermore, we also investigate the clustering of galaxies based on their morphology, along with their relationship to the surrounding environment. As a result, we deliver a full morphological catalogue with 164 314 objects complete up to $r_{\rm petro} < 18$, covering $\sim$1800 deg^2, from which $\sim$55 000 are classified as high reliability, including a significant area of the Southern hemisphere that was not covered by previous morphology catalogues.

Key words: techniques: image processing – catalogues – galaxies: fundamental parameters – galaxies: structure.

1 INTRODUCTION

Galaxy structure was one of the first properties of galaxies that was ever directly observed and studied. Initially thought to be 'nebulae', it soon became evident that these objects showed distinct structural features like spiral arms or a smooth elliptical envelope (Herschel 1864; Zwicky 1940; Vaucouleurs 1959; van den Bergh 1998). Decades of studying galaxy shapes and structures thus resulted in several classification schemes, among which the 'Hubble tuning fork' system of classifying galaxies based on their observed characteristics has been widely used. Collectively known as galaxy 'morphologies',

★ E-mail: debom@cbpf.br

galaxies can broadly be divided into two main categories: early- and late-type galaxies (ETG and LTG). LTG are formed by spiral (S) and irregular/peculiar (Irr) galaxies. The spiral galaxies' branch, bifurcates into barred and un-barred systems. The ETG are composed of elliptical and lenticular galaxies. Elliptical galaxies display an increasing ellipticity, from round (E0) to flat (E7) systems. Lenticular galaxies lie at the apex of the Hubble tuning fork due to their hybrid structure, presenting a bulge and a disc, as spiral galaxies, but without spiral arms.

Morphological diversity often reflects the presence of different and composite stellar populations (Sánchez et al. 2007) and kinematics (Edelen 1969; Wang et al. 2020). For example, S galaxies are characterized by a star-forming disc with blue spiral arms, which indicate rotationally supported stellar kinematics. E galaxies have, in general, smoother features result from less morphological diversity due to a lack of star formation.. E galaxies present a range of kinematic profiles, being the E0 pressure-supported systems or slow rotators, while intermediate elliptical galaxies (E1/E7) present an increasing contribution of rotation to the total kinematic budjet (Cappellari et al. 2011; Bernardi et al. 2019).

Furthermore, galaxy morphologies are found to be tightly correlated to the colour bimodality observed in galaxy populations, thereby resulting in the existence of the younger blue star-forming galaxies with late-type (S) morphologies, and the older red passively evolving galaxies with early-type (E/S0) morphologies (Baldry et al. 2004). However, we are increasingly discovering that several sub-populations of galaxies do not neatly follow this dichotomy, i.e. red spirals and blue ellipticals exist (Bamford et al. 2009), and likely arise from a variety of physical processes, some of which may be environmentally driven (e.g. Vulcani et al. 2015).

Thus, the evolution of galaxy morphology has always been in tandem with the growth of galaxies' large-scale environment and their masses over cosmic time (Desai et al. 2007; Calvi et al. 2012; Crossett et al. 2014; Sarkar & Pandey 2020; Wu 2020). Indeed, using a dichotomous 'bulge/disc' definition for the Hubble-type morphologies, the redshift range $1 < z < 2$ is found to be abundant with bulge + disc systems (e.g. Huertas-Company et al. 2015; Margalef-Bentabol et al. 2016), while massive galaxies in the local universe are majorly bulge-dominated (Buitrago et al. 2013).

Furthermore, higher redshift galaxies pre-dominantly show peculiar/disturbed/irregular morphologies deviant from the classical morphologies observed at the Local Universe (e.g. Mortlock et al. 2013), suggesting that galaxies have undergone remarkable structural transformation over cosmic time (see also review by Conselice 2014). Undeniably, galaxy morphology is a crucial evolutionary key in tracing and understanding galaxy evolution throughout cosmic times (e.g. Shao et al. 2015).

Ample opportunities are now being presented to investigate galaxy morphologies through multiband sky surveys, giving us hundreds of thousands of galaxies while exploring large volumes of the sky at the same time (e.g. SDSS; York et al. 2000). The diverse methods employed by such sky surveys vary from human classification of specialists (Nair & Abraham 2010; Ann, Seo & Ha 2015), to citizen science (Lintott et al. 2008, 2010; Willett et al. 2013; Simmons et al. 2017), or from numerically estimating morphology from galaxy properties (Spiekermann 1992; Storrie-Lombardi et al. 1992; Walmsley et al. 2020) to novel techniques like Principal Component Analysis (PCA; Kelly & McKay 2004; Wjeisinghe et al. 2010), most of which heavily rely on image quality either due to resolution and/or sensitivity of the observations (e.g. Pović et al. 2015). However, migrating to automated methods of classifying galaxies is now necessary to deal with the huge data volumes resulting from such current and upcoming surveys e.g. the Legacy Survey of Space and Time (LSST; Tyson 2002; Axelrod 2006) by the Vera C. Rubin Observatory & sky surveys with the Nancy Grace Roman Space Telescope (Gehrels & Spergel 2015).

Machine Learning (ML) is a powerful automated tool for extracting useful information from complex and varied imaging data sets, and assist in decision-making processes such as classification trees. The use of machine learning is not only limited to galaxy morphologies (Tohill, Bamford & Conselice 2023) but also extends to the detection of gravitational lenses, the analysis of interacting galaxies, and the classification of quasars. (Freeman et al. 2013; Shamir, Holincheck & Wallin 2013; Holincheck et al. 2016; Bom et al. 2017; Ostrovski et al. 2017; Ma et al. 2019; Knabel et al. 2020; Zaborowski et al. 2022). These applications highlight the wide-ranging capabilities of ML in astrophysical research, enabling researchers to explore and understand diverse phenomena in the cosmos. In the last decade, a sub-field of ML known as Deep Learning (DL) has emerged as the main technique for computer vision applications (Abdel-Hamid et al. 2014; Lu, Wang & Zhou 2017; Vecchiotti et al. 2018), music classification (Choi et al. 2017), and medical prognostics and diagnostics (Li, Ding & Sun 2018; Hannun et al. 2019).

DL is applied in model development for processing complex, minimally reduced (or even raw) data from different sources, and extract relevant features that can then be effectively linked to other properties of interest. In particular, Deep Neural Networks (DNNs) are high-performance data-driven models that are capable of exceeding humans in classification tasks (Metcalf et al. 2019). In astronomy, several recent works have exploited this to show that DNNs can indeed be successfully used to identify not only the morphological features in raw images with minimal human intervention (Glazebrook et al. 2017; Petrillo et al. 2017, 2019a, b; Domínguez Sánchez et al. 2018; Lanusse et al. 2018; Cheng et al. 2019; Jacobs et al. 2019; Madireddy et al. 2019; Farias et al. 2020; Hausen & Robertson 2020; Bom et al. 2022), but also outliers in astronomical images (Margalef-Bentabol et al. 2020).

In this paper, we present the morphological classification of galaxies into LT and ET, using the new Southern Photometric Local Universe Survey DR3 (S-PLUS; Mendes de Oliveira et al. 2019). As a follow-up to Bom et al. (2021) hereafter BOM21, our main aim is to apply a high-performance DL algorithm to the imaging data, to obtain a novel and reliable morphological catalogue in the Southern hemisphere, with a complementary coverage to other morphological catalogues. Furthermore, we also develop the first Deep Network to evaluate the quality of the stamps and clean spurious detections. Finally, we take advantage of the high precision photometric redshifts derived using the 12 bands in S-PLUS to explore the dependence of morphology on the environment and colour, used as a proxy for the galaxy stellar population properties. We compare the classification presented in this work with Vega-Ferrero et al. (2021), and we discuss the implications arising by studying differently classified objects on the current understanding of galaxy morphological categories.

This paper is organized as follows, in Section 2, we describe the data from DR3 used in this work, the sample selection, and auxiliary data used, such as the photometric redshift. In Section 3, we present the Deep Learning method used for galaxy morphology classification, and the novel-ties in its implementation since S-PLUS DR1 morphology paper (BOM21, Bom et al. 2021). In Section 4, we present the results of the model, including deep learning performance. We also show the relation between environmental density and morphology, and we analyse the distribution of the different morphological classes in a (g–r) colour versus M_r absolute

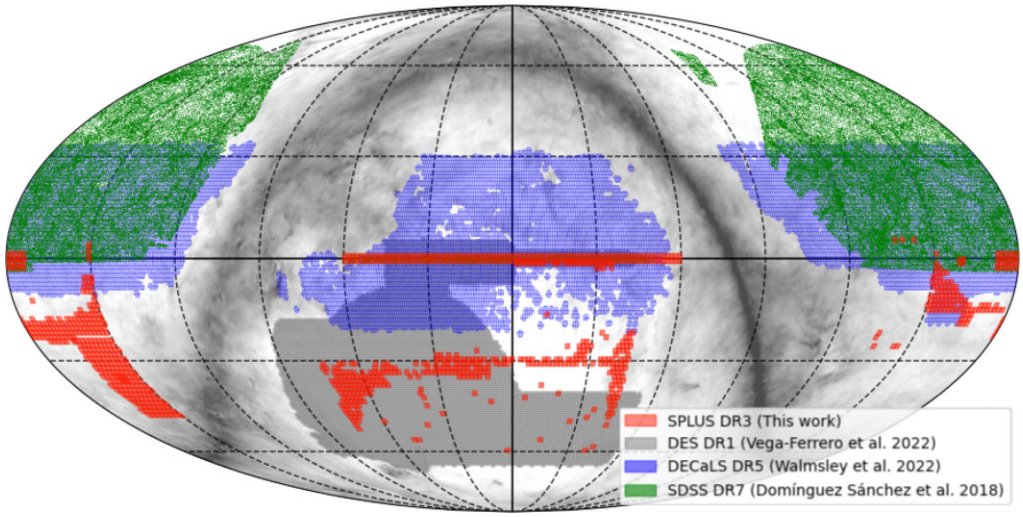

Figure 1. S-PLUS DR3 footprint (used in this work) and the footprint of some recent galaxies morphology catalogues available in the literature.

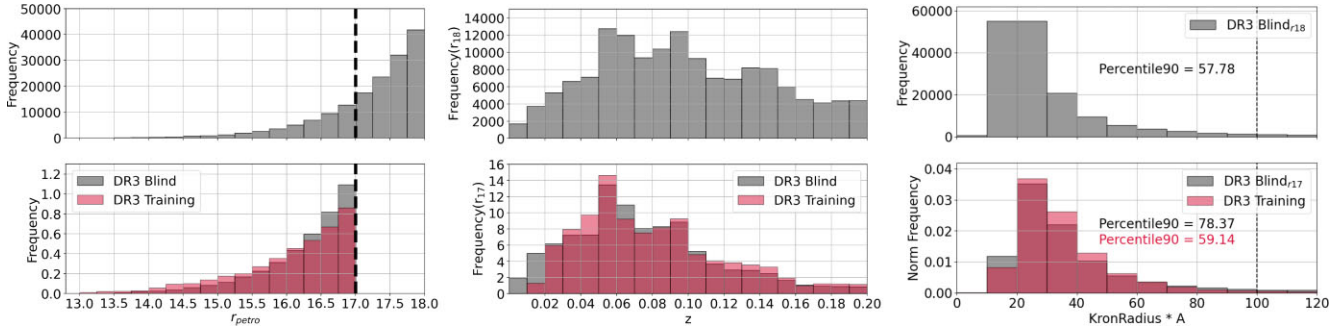

Figure 2. Samples characteristics. *Left.* Distribution of the *r*-petrosian magnitudes ($r_{\rm petro}$) for *II* (Blind set), i.e. up to magnitude $r_{\rm petro} \leq 18$ (*top*) and a normalized histogram of the magnitude distribution of *I* (training/validation set) and *II* (blind set) up to $r_{\rm petro} \leq 17$ *bottom*, i.e. the limiting magnitude of *I* – the training sample. *Middle.* As in the *left*-hand panel, but for the distribution of the photometric redshift for the training and blind sample. *Right.* As in the *left-hand* panel, but for the distribution of the Kron radius for the training and blind sample. The dashed vertical line represents the radius of a stamp of 200 px size, and we indicate the 90th percentile of the distribution for the different samples.

magnitude diagram. In Section 5, we present our summary and discuss the results.

2 DATA

2.1 Southern Photometric Local Universe Survey

The S-PLUS is performed with a robotic 86-cm telescope located at the Cerro Tololo Interamerican Observatory to cover ∼9300 deg^2 of the sky in 12 optical bands. S-PLUS uses a wide field optical camera with a field of view of 2 deg^2 and a plate scale of 0.55 arcsec pixel^{-1}. The optical filters (the so-called Javalambre filter system, with 5 SDSS-like bands and 7 narrow bands; Cenarro et al. 2019) are quite unique for the Southern hemisphere and are optimal for source classification, given its better definition of the spectral energy distribution of the observed objects, than the usual 4 or 5-band systems. The narrow bands are designed to be centred on important stellar features, for instance, the O II line, Ca H + K, H δ, and H α. The survey reaches a typical limiting magnitude of $r < 21$ AB mag for the broad bands and $r < 20$ AB mag for the narrow bands (Mendes de Oliveira et al. 2019).

The third public data release of S-PLUS (DR3) covers ∼2000 deg^2 over the Southern Sky. It includes the areas covered in the previous Data Releases such as the Stripe 82. However, the images were reprocessed, with a new reduction and calibration of the data being done from DR2 to DR3, as described in Almeida-Fernandes et al. (2022). In Fig. 1, we present the area covered by DR3 in comparison with other surveys with available morphological catalogues. The Stripe 82 area (at the equator) overlaps with numerous surveys in optical and other wavelengths, and it has been used as a benchmark for checking the data reduction and calibration procedures. Other important area covered by the DR3 is the Hydra supercluster (the long vertical red rectangle at the far left of Fig. 1).

2.1.1 *Sample selection*

We use the full DR3 catalogue containing ∼50 millions of sources. During the DR1 morphological classification, we selected the objects only by Petrosian magnitude in *r* band ($r_{\rm petro}$) < 17 AB mag and probability of being a galaxy prob$_{\rm gal} \geq 0.6$ (for further information see Nakazono et al. 2021a). However, we had a visual inspection phase to remove undesired spurious detection (see BOM21). The current catalogue covers an area of 1800 deg^2, which makes the visual inspection unfeasible in a reasonable time-scale with limited human resources. Therefore, we define more stringent cuts and include four extra constraints compared to BOM21. Additionally, we added an

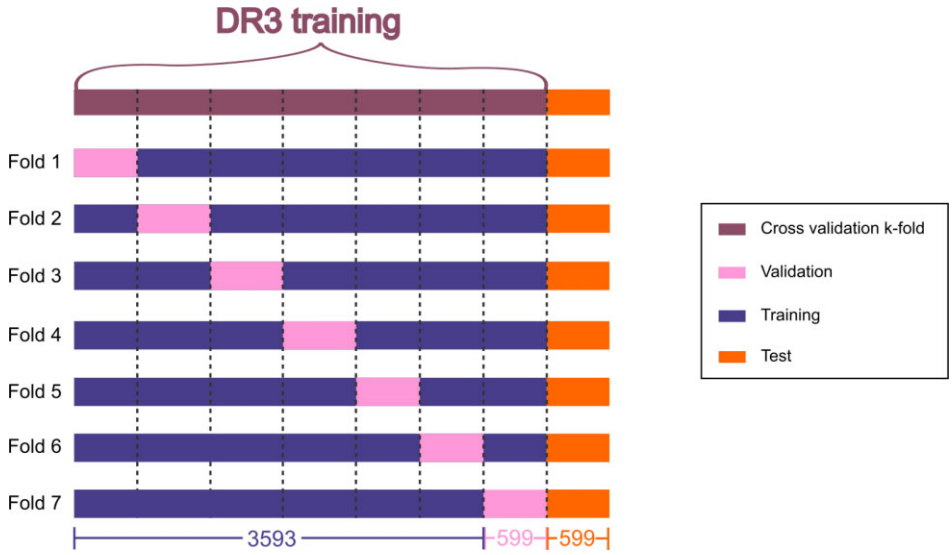

Figure 3. In this figure, we show the cross validation k-fold method applied to the training of the LTG/ETG model. Each fold is separated into training, validation and test. That process is made in a way that there is no match between each validation. Additionally, the training for each fold is slightly different, which reduces a possible bias concerning the selection of the objects that composes it. Considering that the technique will define an architecture with certain weights for each fold, the metrics in each training stage can be used to evaluate which fold has the best training set configuration. The numbers at the bottom indicate the size of the training, validation, and test sets in each fold.

Table 1. Sample description of the samples used in this work.

Sample	Subsample	Number of objects	Description
I	*DR3*-training	4192	ETG and LTG galaxies splitted between training and validation
I	*DR3*-test	599	ETG and LTG galaxies for performance test
II	*DR3*-blind	46 763	galaxies for blind classification with $r_{\rm petro} \leq 17$
II	*DR3*-extended	161 635	galaxies with $r_{\rm petro} \leq 18$ for blind classification

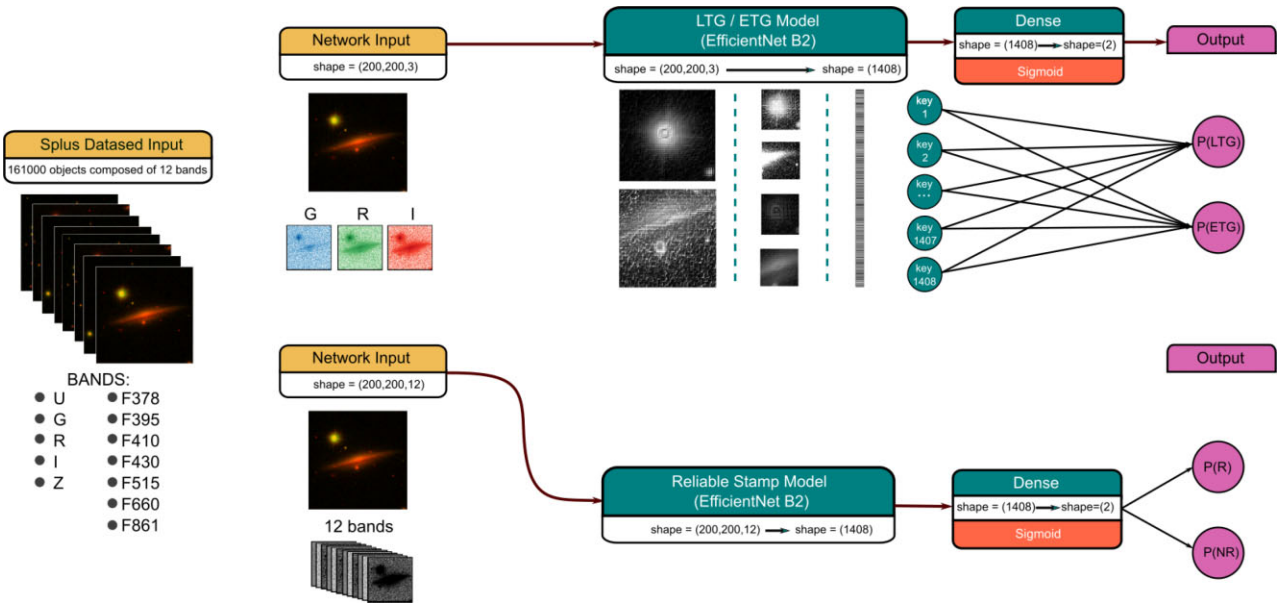

Figure 4. Workflow from the stamps taken from S-PLUS data while they passes through the model. Both architectures works in the same way, with the difference that the first one uses only the *G*, *R*, and *I* bands available in S-PLUS as the network input. The LTG/ETG model as well as the Reliable Stamp Model consists of some convolutional layers in the beginning responsible to compact and recognize patterns in the stamp, then, in the end, all that information passes through a dense layer that compacts it into a list containing 1408 keys represented by the bar code in the figure. Both models works with binary classification, then one more dense layer is needed to calculate the probability of each classification given by a sigmoid activation function.

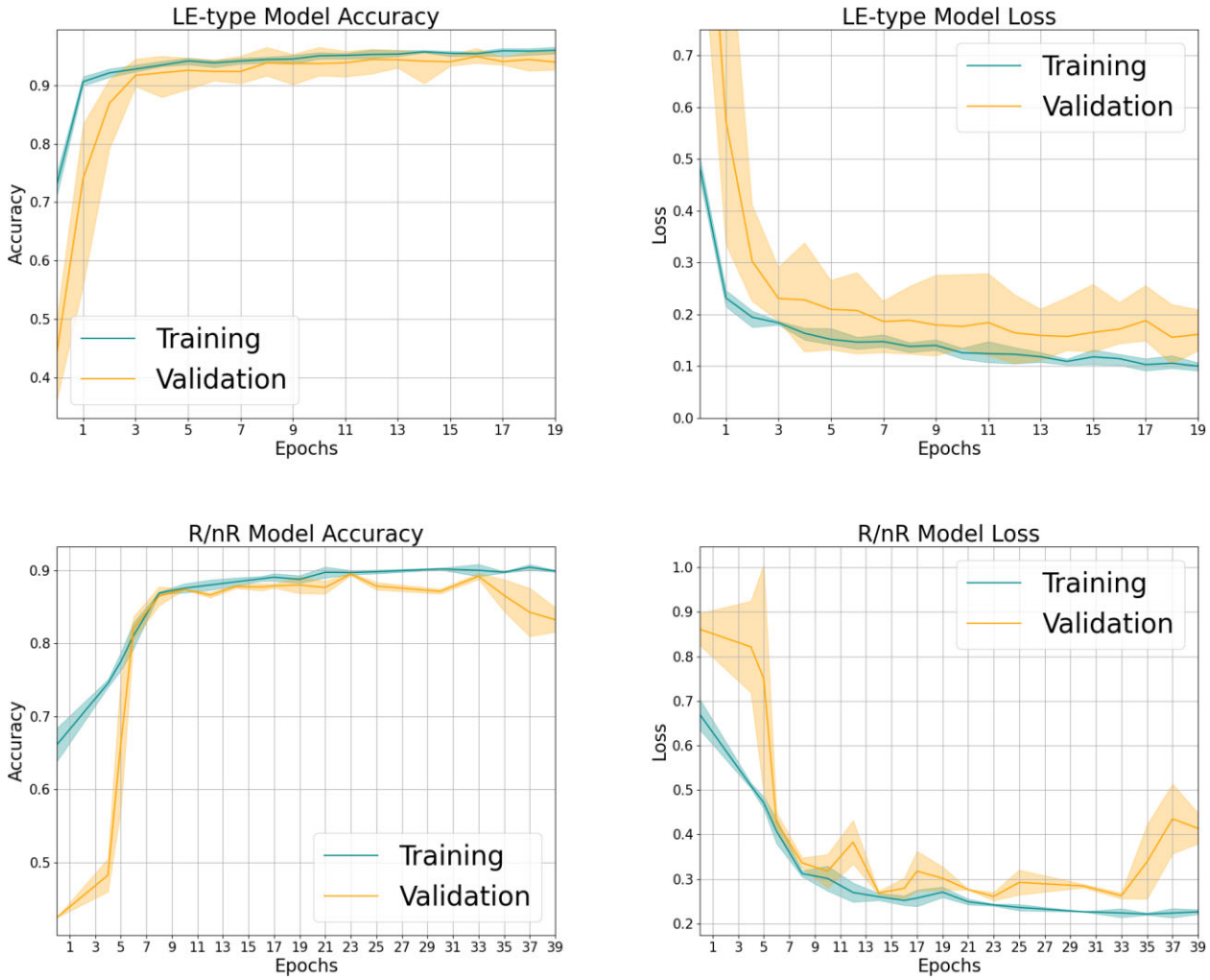

Figure 5. Accuracy and loss in the training of the late-/early-type (*Top*) model and Reliable Stamp model (*bottom*) as a function of epoch considering all folds. In blue we present these metrics for the training set and in orange the metrics for the validation set. The line in the middle represents the mean value between all 7 folds used in the cross validation k-fold method.

automated selection phase by Neural Network that is detailed in Section 3. Thus, we apply the following selection criteria to define our galaxy sample from the full catalogue of the S-PLUS DR3:

$$r_{petro} < 18 \text{ AB mag} \quad (1)$$

$$prob_{gal} \geq 0.7 \quad (2)$$

$$0 <= photoflag_r <= 3 \quad (3)$$

$$BrightStarFlag = 0 \quad (4)$$

$$R_{Kron} >= 3\ pixels \quad (5)$$

$$FWHM_n >= 1.5\ pixels \quad (6)$$

where $photoflag_r$ r is a photometry quality flag from *SExtractor* (Bertin & Arnouts 1996), R_{Kron} is the Kron radius, i.e. the first moment of the surface brightness light profile, $FWHM_n$ is the full width at half maximum of the object divided by the median FWHM of all bright non-saturated stellar objects of the field. All those features are available and described in the SPLUS catalogue. The probability of being a galaxy, $prob_{gal}$ (Nakazono et al. 2021b), and the flag indicating a presence of a bright star nearby, BrightStarFlag, are listed in the 'star-galaxy-quasar' and 'masks' Value Added Catalogues (VAC; see SPLUS.cloud for further details[1]). Specifically, $0 <= photoflag_r\ r <= 3$ ensure the goodness of *Sextractor* fit in most of the cases of interest. The *BrightStarFlag* parameter is very effective in removing bright stars and allows to cleaning the few stars that are erroneously assigned a probability higher than 0.7 of being galaxies by the star–galaxy–quasar classification (Nakazono et al. 2021b). The conditions $Kron_{Radius} >= 3$ and $FWHM_n >= 1.5$ are included to select resolved objects (the average $FWHM_{seeing} \simeq 1.2$ arcsec).

Following these selection criteria, we obtained a final catalogue of 164 314 objects, for which we created image stamps in the 12 bands, with a size of 200 × 200 pixels[2].[2] This choice of stamp size covers more than the 90th percentile of the galaxies in the sample, as shown in Fig. 2. The stamps are also normalized on each band independently before input the DL algorithm, we discuss the importance of colours in our catalogue later in Section 4.5.1. The

[1]https://splus.cloud/catalogtools

[2]The image cutout tasks can be found in this GitHub repository: https://github.com/lucatelli/splus-tools.

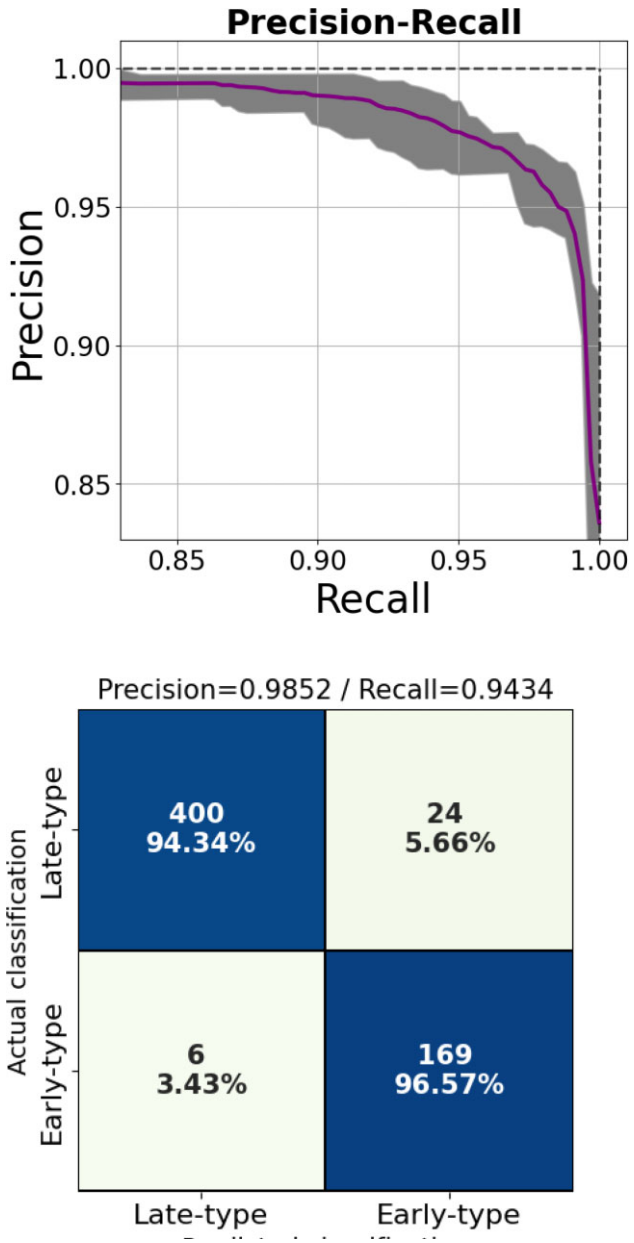

Figure 6. Performance in the *DR*3-Test sample for the ETG/LTG model. (Top) The Precision x Recall considering all folds. The purple line was made with the median value for every fold. (Bottom) The confusion matrix for the best fold.

final catalogue is mostly composed of reliable stamps, i.e. stamps centred on a galaxy, complete up to $r_{\rm petro} < 18$. Further improvement on the sample selection is described in Section 3.2.

2.1.2 *Samples definition*

The supervised DL assessment requires to be trained on a sample of objects with known classification, i.e. a *labelled set* (*Training/Validation and Test Set – I*), sharing as much as possible, the same properties of the sample where the algorithm will be applied in a second moment (*Blind Set – II*). In this section, we describe the characteristics of the two samples, but we refer to Section 3 for more details on the DL algorithm and its performance.

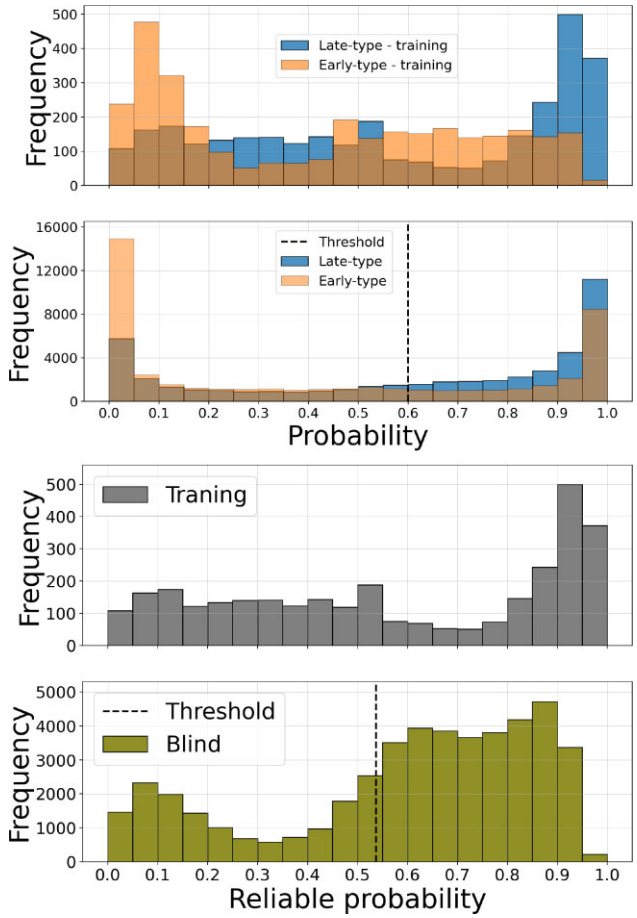

Figure 7. Probability distributions of the output classifications of our Deep Learning models for both test (second and fourth panel) and training sets (first and third panel). The top panels present the distribution for the late-type and early-type classification. On the bottom panels show the distribution of being a reliable stamp. In both cases, the dashed line represents the threshold used for the classification itself.

We used the same objects that were included in both the training and validation sets, following the test scheme used in BOM21. This scheme utilized unbiased morphological classification from Galaxy Zoo 1, distinguishing between elliptical and spiral galaxies. (Lintott et al. 2008; Bamford et al. 2009; Lintott et al. 2010) as true label. Such choice was possible since S-PLUS DR1 is included in S-PLUS DR3. It is important to note, though, that since the reduction pipeline has been improved between the two data releases, new stamps were created using the novel images, to ensure the homogeneity of the two data sets (I and II). Another relevant difference between the data from DR1 and DR3 is the new photometric calibration applied for the S-PLUS DR3. This calibration consists of fitting synthetic stellar templates to well-known data from other surveys, deriving precise zero-points and magnitudes that were tested on 170 STRIPE82 fields (see Almeida-Fernandes et al. 2022, for a detailed description of the method). We obtained the stamps for each object in the 12 bands from the DR3 data access, for both samples.

In total, there are 4232 objects in training sample *I*, while set *II* is composed of 164 314 objects. We refer to Table 1 for a description of the samples and subsamples used in this work. As presented in the top panel of Fig. 2, the training sample, i.e. sample *I*, is approximately

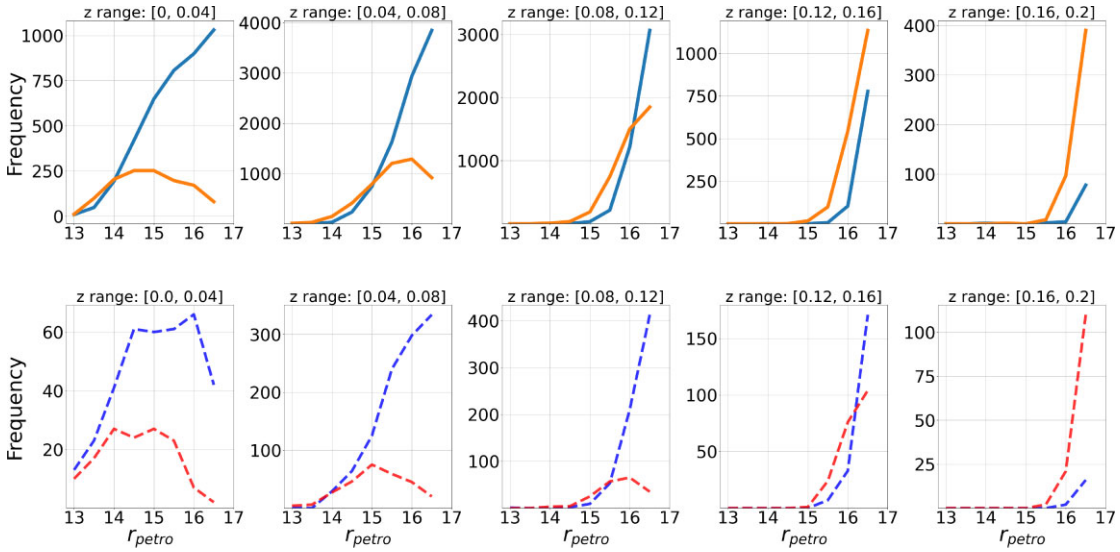

Figure 8. *r*-band apparent magnitude distribution for different photometric redshift bins for the blind sample (top panel – filled line), and for the training sample (bottom panel – dashed line), for elliptical (yellow/red) and spiral (blue/cyan) galaxies. Note that in the first magnitude bin the training sample has very low number of objects ($z < \leq 0.04$), since we used Galaxy Zoo data for the training, which is low in this mag bin (Lintott et al. 2008; Bamford et al. 2009; Lintott et al. 2010). On the other hand, in this first bin, in the blind sample the number of ETG is slightly higher than of LTG, for $r < 14$. This result reflects the fact that ETG are more massive (brighter) than spiral galaxies. At higher redshifts, this effect becomes visible also in the training sample, and it is due to the lower number of spiral galaxies being selected (see Section 2) due to their decreasing surface brightness profile for a given magnitude.

complete only up to $r_{\rm petro} < 17$.[3] As described in Section 2.1.1 in this work, we select objects up to $r_{\rm petro} < 18$. The implications of this choice in the DL performance are discussed in Section 4. Both the samples I and II show similar distribution of $R_{\rm petro}$ for magnitudes <17 (bottom panel of Fig. 2)

2.1.3 Photometric redshifts

The S-PLUS DR3 photometric redshift catalogue uses a DL model based on a Bayesian Mixture Density Network architecture. This specific configuration allows single-point estimates while also providing probability distribution functions (PDFs) for each galaxy. This network is trained on 12-band photometry from S-PLUS, cross-matched with the unWISE (Wide-field Infrared Survey Explorer; Lang 2014), GALEX (Galaxy Evolution Explorer; Niemack et al. 2009), and 2MASS (The Two Micron All Sky Survey; Skrutskie et al. 2006) catalogues (*W*1/*W*2, NUV/FUV, and *J*/*H*/*K* magnitudes, respectively). Spectroscopic redshift targets are compiled from various surveys, including SDSS DR16, 2dFRGS, 2dFLenS, 6dFGS, and others. A total of 262 521 objects are used for training/validation, and an independent test set.

Due to its unique filter system with a set of broad- and narrow-band photometry, the current model is capable of providing accurate photometric redshifts, while also maintaining low bias and negligible outlier fraction. In fact, within the magnitude range of interest of the present work, $r_{\rm petro} \in [14, 18]$, the median normalized bias stands ~ -0.0015, the scatter is ~ 0.015, and the outlier fraction is below 1 per cent. The catalogue not only includes single point estimates but also well-calibrated PDFs, enabling users to evaluate the uncertainties associated with each estimate. Further information regarding the methodology and resulting findings can be found in Lima et al. (2022).[4] Fig. 2 shows the distribution of the photometric redshift for samples *I* and *II*. Specifically, *II* is divided into the whole sample, with $r_{\rm petro} < 18$ and a sub-sample with $r_{\rm petro} < 17$, sharing the same magnitude limit as the training sample.

3 DEEP LEARNING CLASSIFICATION

3.1 Training, validation and test sample

We split the cross-match data between S-PLUS DR3 and Galaxy Zoo I STRIPE82, i.e. data set *I*, into training-validation-test sets. Data set *I* contains unbiased classification only (Lintott et al. 2008; Bamford et al. 2009; Lintott et al. 2010).The data includes a threshold of 80 per cent probability to identify true labels, with 29 per cent assigned to ETG and 71 per cent to LTG. This distribution reflects the proportion between the two classes, in the local Universe ($0 < z < 0.2$) as reported by Lintott et al. (2010).

We split the DR3-Training data set in 7 folds. These folds are subsamples of the training set used to perform a cross-validation procedure (Moreno-Torres, Sáez & Herrera 2012). We have evaluated other choices for a number of folds. However, with more folds the validation set is smaller, the validation loss starts to be more unstable, and we found a good trade-off with 7 folds. Thus, as shown in Fig. 3, we define 7 different training and validations sets, each containing $\sim$ 85 per cent and $\sim$ 15 per cent of the data, respectively. This separation is made so there is no match between the validation sample for every fold. Additionally, this method guarantees that each object will be used at least once in the test set. We use 599 objects as a test

[3]This magnitude limit is required in Galaxy Zoo 1 in order to perform the debiasing process, which requires spectroscopic redshifts, see Bamford et al. (2009) for more details

[4]The S-PLUS public data, including the photometric redshifts are also available in splus.cloud

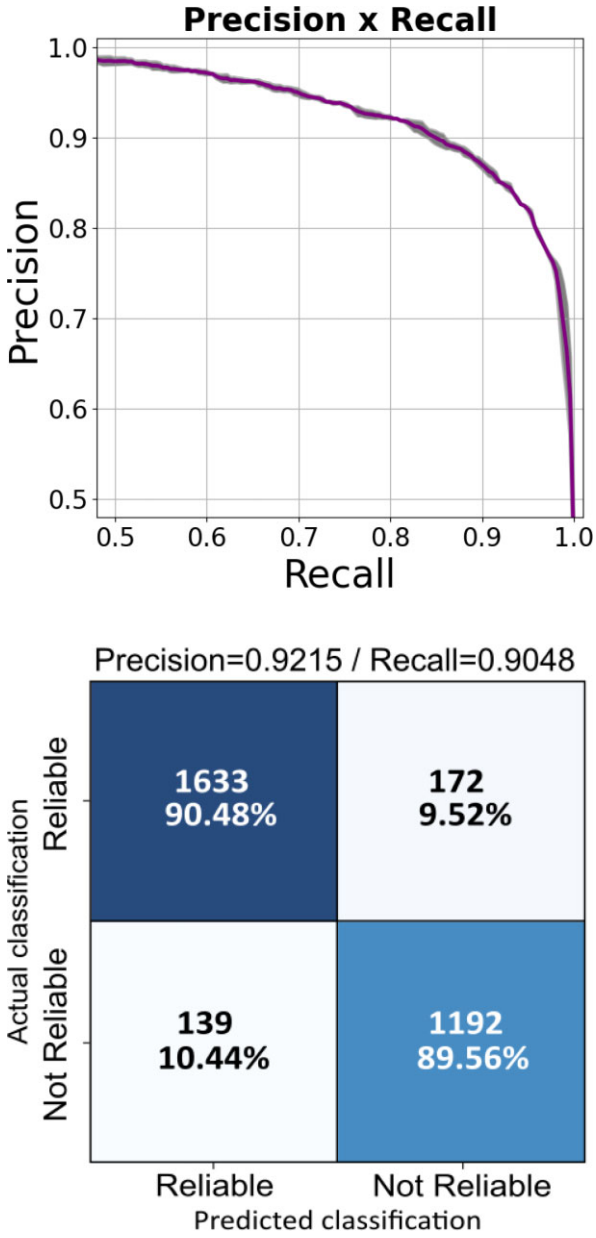

Figure 9. Performance in the *DR*3-R-Test concerning the reliable/non reliable model. (Top) The precision as a function of the recall plot considering all folds. The purple line was made with the mean value for every fold. (Bottom) The confusion matrix for the best fold.

set for performance evaluation, these are not used for training. As in BOM21, the training set based on debiased GZ1 contains 71 per cent of LTG and 29 per cent ETGs, and thus is an imbalanced data set. Therefore, in order to train the Neural Network to prevent our model of being biased towards the most abundant class, we adopt the same data treatment scheme presented in BOM21, applying weights to each class. For a set of N objects in the training set and if the number of objects in the class α is N_α, we define the weights as

$$w_\alpha = \frac{N}{mN_\alpha}, \tag{7}$$

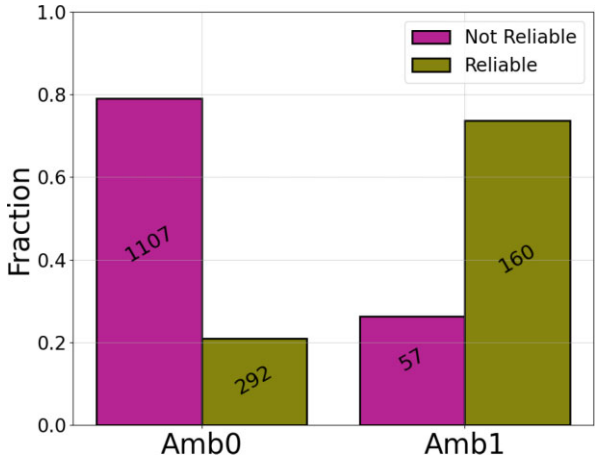

Figure 10. Normalized fraction of galaxies that belong to class Amb0 (left), i.e. galaxies that have a low probability of being ETG or LTG, and of galaxies belonging to the class Amb1 (right), i.e. galaxies with a high probability of being ETG and LTG, classified as no-reliable stamps (blue) and reliable stamps (orange).

where m is the total number of classes. This is a standard procedure in ML field.[5] The weights defined in equation (7) are then applied in the objective or loss function minimized during the training phase. This procedure enables each of the classes to have the same impact on the loss function.

3.2 Non reliable stamps

In the DR1 catalogue from BOM21, all the stamps were visually inspected prior to the analysis with the DL algorithm aiming to prevent biases in the classification process caused by spurious objects. This became no longer feasible due to the size of the current and future S-PLUS data releases. Therefore, in this study, in order to generate a more robust catalogue, we implemented a new DL model to separate *reliable* from *non-reliable* stamps. To this end, we use stamps that were excluded as *non-reliable* from the DR1 classification as a training sample for this new DL model. This approach is advantageous to avoid spurious classifications such as faint galaxies in the same field of nearby saturated stars whose spikes can affect the accuracy of the magnitude estimation of that galaxy. Given the considerable extent of the data set used in this study, other undesirable objects might include artefacts and problematic stamps in general. Therefore, it is essential to count with a robust method of distinguishing good images from low-quality images to use as an input for the main ETG/LTG DL model.

3.3 Deep learning model

Following a strategy similar to that in BOM21, in this work we also made use of EfficientNet algorithms (Tan & Le 2019), which are part of the Convolutional Neural Networks (CNN) family-models well-known for having high performance on visual pattern recognition problems in standard image data sets such as ImageNet (Deng et al. 2009).

EfficientNets are a set of convolutional neural networks crafted to deliver top-tier precision. These networks are uniquely optimized

[5]See e.g. https://www.tensorflow.org/tutorials/structured_data/imbalanced_data

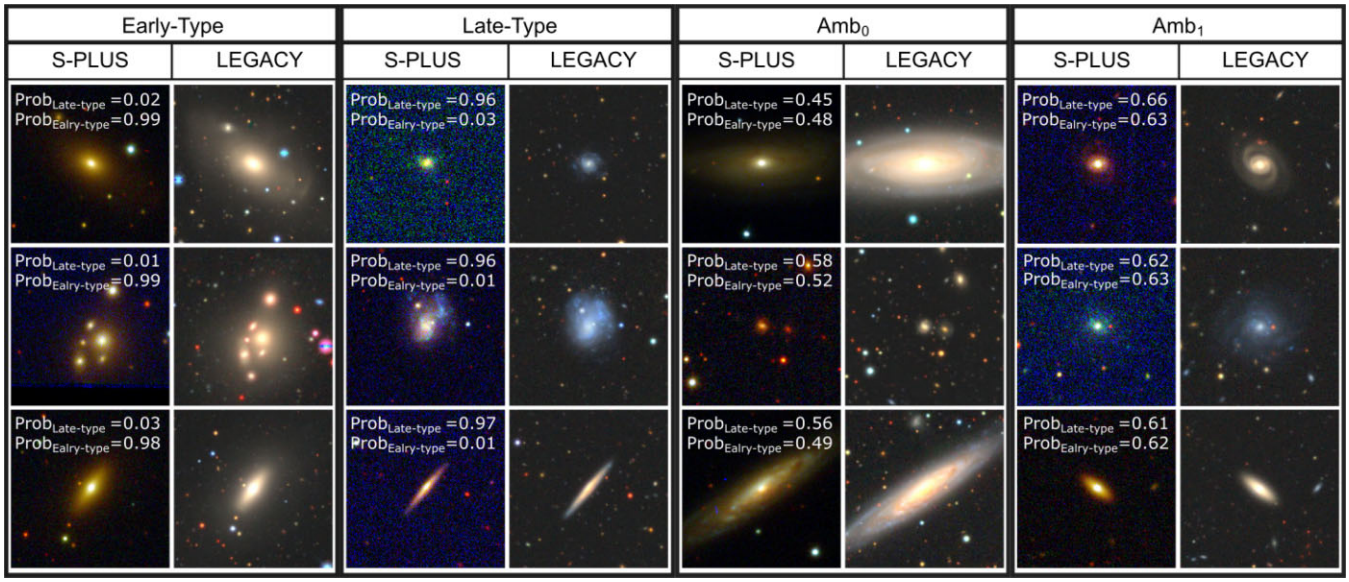

Figure 11. In the first two panels, we show some examples of **Reliable Stamps** that were classified as early-type (first panel) or late-type (second panel). In the last two panels, we have examples of **Reliable Stamps** that would fall in the Ambiguous classification. Amb_0 are those stamps that had a low probability of being ETG and also LTG according to the defined threshold ($\simeq$0.6). In the other hand, we have Amb_1 which are those objects that the model gave a high probability of belonging to both classes. Each panel is made with the same objects taken, respectively, from S-PLUS and LEGACY survey.

to provide for performance without compromising on computational efficiency, making them a prominent choice in current computer vision problems. This kind of Network is based on an initial model similar to a MobileNet (MnasNet; Tan et al. 2019) and can be also scalable by parametrizing the number of layers if needed to gain in performance by making a more complex network while constraining the number of Floating-point Operations Per Second (FLOPS).

The central feature of EfficientNets lies the concept of compound scaling. Rather than independently adjusting the network's width, depth, or resolution, EfficientNets modify all three dimensions simultaneously, using a consistent set of scaling factors. This integrated method ensures a harmonious balance among the network's width w, depth d, and input resolution r, culminating in enhanced performance and optimal efficiency. The scaling relations are:

$$\begin{aligned} d &= \alpha^{\phi} \\ w &= \beta^{\phi} \\ r &= \gamma^{\phi}, \end{aligned} \tag{8}$$

ϕ is an integer named compound coefficient. The coefficients α, β, and γ are optimized by a small grid search in the base MobileNet-like (Tan et al. 2019) model considering a fixed $\phi = 1$. The computational efficiency can be assessed by the number of FLOPS that scales with $d \cdot w^2 \cdot r^2$.

Therefore, each parameter choice defining a model and thus defining a family of models. Additionally, this kind of model can also be easily adapted to classify data sets with different resolutions (Bom et al. 2022). In this contribution, we made use of similar model based on EfficientNet B2 firstly described in Tan & Le (2019), with the minor adaptations detailed in BOM21. For a diagram presenting all the layers in this model please refer to fig. 5 (b) and (c) in Bom et al. (2022) paper.

None the less, we implemented several innovations compared to the workflow described in BOM21. First, we added a second EfficientNet B2 model to evaluate whether a stamp is reliable for morphological classification. The main goal of this NN is to identify spurious detections, such as crowded stamps where the central galaxy in the stamp is visually indistinguishable, stamps satured by close bright stars, and galaxies that are not completely contained in the stamps. We explore how this non reliable stamp model would be best defined in terms of inputs. After initial tests following BOM21 approach, we used all 12 bands as inputs in contrast to the ETG/LTG model that was shown to be best defined in terms of performance and stability of results by using g, r, i bands only. Although this choice is based on empirical results by adopting the same metrics presented in Bom21, the main difference here is likely due to the nature of patterns we are trying to characterize in the Reliable Stamp model. By visually inspecting the stamps, we find that some of the spurious detections presented large variability of shapes in different bands compared to reliable stamps, and thus are likely to be easily distinguishable by using more bands. For a full discussion of the band choice for finding ETG/LTG please refer to BOM21. A relevant difference related to the main ETG/LTG model developed for S-PLUS DR1 is that the probability assigned to a galaxy of being spiral or elliptical is no longer complementary, meaning that the sum of such probabilities is not equal to one, opening a space for a lot of interesting findings like the ones discussed in Section 4.3 and the possibility of pointing objects that do not fill in any category. This was implemented by changing the neural network activation function in the last layer from a *softmax* to a *sigmoid* and, therefore, change the algorithm to a multiclassification scheme. In Fig. 4, we present a scheme of both DL models, detailing the input bands and also presenting an example of a given stamp flowing towards some of the network convolutional filters.

4 RESULTS

4.1 Training

The training process was performed with a Rectified Adam (RADAM; Liu et al. 2019) optimizer and the loss function is a traditional cross-entropy (Goodfellow, Bengio & Courville 2016).

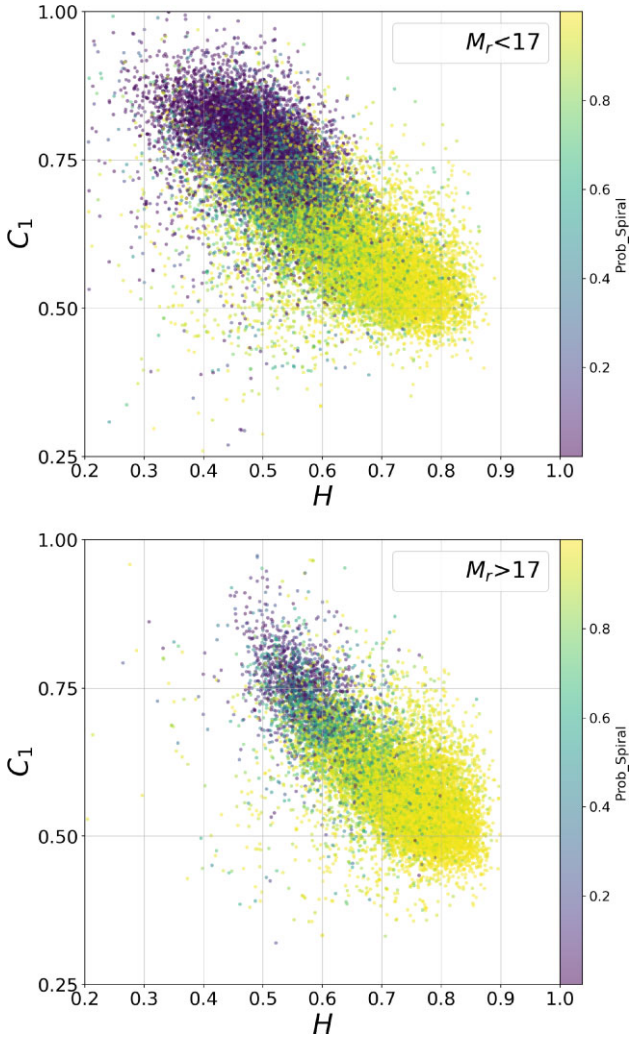

Figure 12. Concentration versus entropy parameters, colour coded according to their probability of being a spiral galaxy, for galaxies classified as 'reliable stamps', brighter than $r < 17$ (top) mag and with $17 < r < 18$ mag (bottom). ETG tend to have higher concentrations and lower entropy, while LTG, being disc dominated and star forming, are less concentrated and present higher entropy (see Ferrari, de Carvalho & Trevisan 2015). Interestingly, there is a lower number of galaxies with high concentration at $17 < r < 18$, probably caused by the decrease in the surface brightness.

In Fig. 5, we show the Loss and Accuracy achieved in the training procedure considering all 7 folds. The darkest line in the center corresponds to the mean value of those quantities for each epoch and the shaded area corresponds to the standard deviation between folds. In the top of Fig. 5, we present the results for ETG/LTG model using 3 broad-bands as input, similar to BOM21. The training converges fast, around the third-fifth epoch with high accuracy ∼0.9. The additional degree of freedom added compared to BOM21, i.e. the probability of being LTG or ETG set to be independent, does not seem to affect the performance significantly. Considering the errorbars, we did not find significant overfitting over the entire range. However, towards the end of the training, around the 15th epoch the figure suggests the beginning of slight overfitting. Furthermore, by evaluating the loss function, the reliable/non reliable model, that uses the 12 band set as the input, presents a more unstable behaviour: the convergence is slower, around the epoch 15. The validation presents some spikes that might be related to a regularization method present in the network. We also notice a tendency of overfitting from epoch ∼19 onwards. The validation accuracy does not reach 0.9 consistently. However, it is worth noticing that, differently from the 12 band model ETG/LTG presented in BOM21, the 12 band model for reliable/non reliable stamps has significant smaller errorbars suggesting that the model is robust, although the overall performance compared to ETG/LTG model as a classifier is expected to be lower. In the next section, we discuss the general performance metrics of our deep learning models considering the folds at their best epoch. We define the best epoch per fold as the one with lower loss calculated in the validation sample in the region where we are confident there is no overfitting within the errorbars. This criteria selects a single model per fold which best generalized for unseen data, i.e. lower loss in the validation sample. For final catalogue production, we use a similar criterion by selecting the lower validation loss among all best models per fold. This choice enables us to be efficient in running a single best-performing network for unseen validation data.

4.2 Performance

As outlined in the previous section, the cross-validation approach establishes a unique network configuration for each fold. Therefore, we may assess our model's performance on every individual fold. For both ETG/LTG or Reliable Stamp Model classification, we applied these individual folds to the test subsample. We evaluate the performance of our model by evaluating the trade-off between precision (purity) and recall (completeness). For the sample of galaxies of a class α we may define these quantities as:

$$\text{precision} = \frac{|\{\text{galaxies in } \alpha\} \cap \{\text{galaxies classified as } \alpha\}|}{|\{\text{galaxies classified as } \alpha\}|}. \quad (9)$$

$$\text{recall} = \frac{|\{\text{galaxies in } \alpha\} \cap \{\text{galaxies classified as } \alpha\}|}{|\{\text{galaxies in } \alpha\}|}. \quad (10)$$

For a given threshold t that defines which is our ETG if the predicted probability is higher than the t, precision or purity measures how many correct predictions were made out of all positive predictions, and recall or completeness presents how many true positives were found among all the actual positives.

4.2.1 ETG/LTG Model

In the upper of Fig. 6 we present the median precision-recall for t in the range [0,1] for all folds and its respective standard deviation. Later, we define the best theshold t_B as the t in the precision-recall curve closest to the point (1,1) which would represent a perfect classifier, i.e. with both purity and completeness equal to 1. This threshold t_B is set to ∼0.60. To understand the performance outcome with this choice we made use of a confusion matrix at the bottom of Fig. 6. This performance assessment shows the number of correct and incorrect predictions, grouped by each class and therefore presents model performance in a classification task by revealing where it gets confused and makes mistakes. The model demonstrates correct classification with over ∼ 94 per cent of both ETG and LTG classifications. It is worth mentioning that for this specific performance assessment, we had to assign each galaxy to one category unambiguously to obtain the confusion matrix and performance curves and therefore, for this specific performance analysis, we did not take advantage of the fact that the model assigns independent probabilities of ETG/LTG. We refer the reader to Section 4.3 for a detailed discussion about the advantages of having a multiclass evaluation, such as the possibility to select ambiguous objects. In Fig. 7 top we present the probability distribution of the

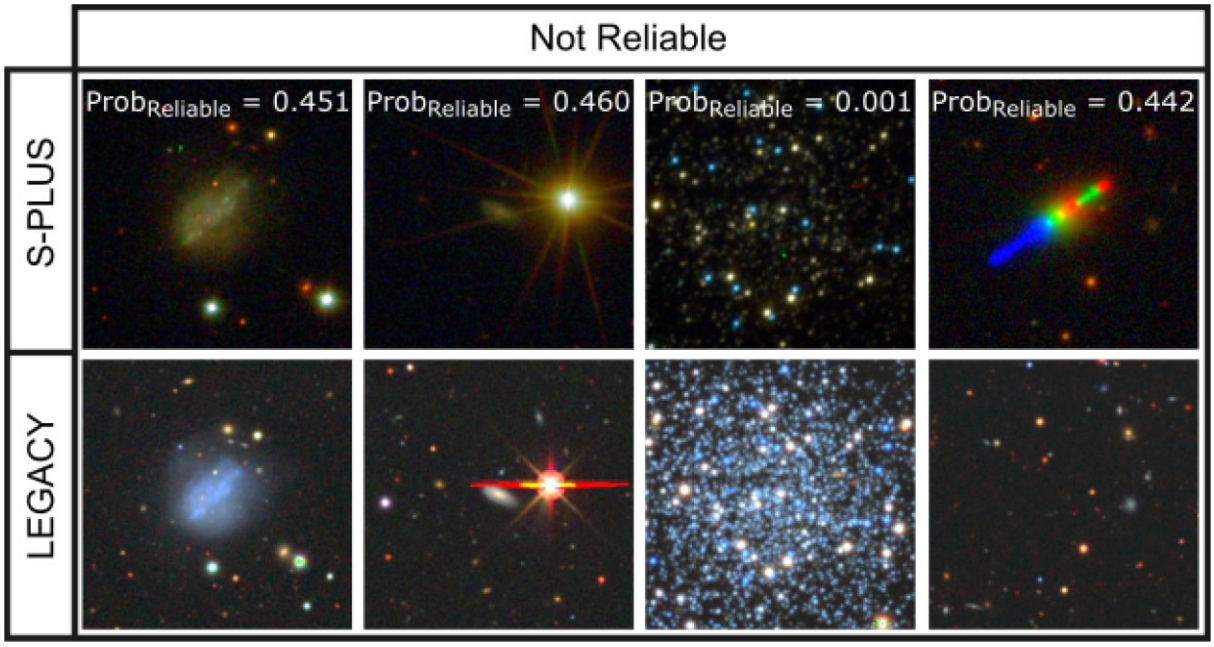

Figure 13. Example of not reliable stamps, from S-PLUS data (top) and LEGACY data (bottom). In the last column, it is visible an artefact, the third column present a crowded field, in the second column we find a saturated star compromising the galaxy image, and, finally, in the first column and irregular galaxy.

DR3-Blind set. We notice that the distribution for ETG and LTG classification are well separable with a strong peak around ∼0 and ∼1.0, as one should expect for a two-class classification.

We present in Fig. 8, a comparison of the distribution of the photometric redshifts (see Section 2.1.3) of ETG (orange/red) and LTG (cyan/blue) for the DR3-Blind/DR3-Training sample, respectively. It is noticeable that, against expectations, the number of ETG seems to be larger at higher redshifts than the number of LTG, both for the training and blind data sets. In fact, Buitrago et al. (2013) does not find any strong evolution between the fraction or density of spheroid and disc galaxies for $M_* > 11$ M$_\odot$ between $0 < z < 0.2$. We visually inspected galaxies classified as early type at $z > 0.15$ to verify whether the classification is affected by the lack of resolution of the spiral arms. We conclude that the classification is overall correct (see Sections 5.1 where we compare with the morphological classification performed by Cheng et al. (2020) and Vega-Ferrero et al. (2021) and that the lack of spiral galaxies at high redshift is related to the pre-selection of the stamps since high z spirals tend to have low surface brightness. The training sample used in this work is taken from The Galaxy Zoo 1 project (Lintott et al. 2008; Bamford et al. 2009; Lintott et al. 2010), which provides a debiased morphological classification Bamford et al. (2009) for galaxies in a redshift range between $z \geq 0.03$ and $z < 0.88$, where the lower limit is dictated by the incompleteness at low redshift, while the higher redshift is caused by the loss of objects fainter than $M_r < -20.25$.

4.2.2 *Reliable stamp classification*

We used the same analysis scheme to evaluate the Reliable Stamp model. The bottom part of Fig. 7 presents the Reliable stamp probability distribution. By comparing the probability distribution of both models we noticed that the reliable model presents wider peaks, which suggests the distribution is not as well separated as in the ETG/LTG model. This conclusion is also indicated by the loss optimization as discussed in Section 3.3.

In Fig. 9, we show the confusion matrix for the best fold and the precision versus recall plot considering all folds. The overall shape in the precision x recall curve is similar to the ETG/LTG model, however, the total area under the curve of the Reliable stamp model is smaller compared to the ETG/LTG model. The confusion matrix presents ∼ 90 per cent true positives, which is also interesting since there is a vast variability of what is a non reliable stamp. Additionally, by making a visual assessment over the objects classified as non reliable we can find some interesting objects that we believe are worth investigating. We discuss this with more detail in Section 4.4.

4.3 Early-type and late-type galaxies

Galaxies present a wide range of morphologies (e.g. van den Bergh 1998; Buta 2011), from almost spherical ellipticals to grand design spiral galaxies (Grosbøl & Dottori 2012), with the increasing importance of the disc component along the Hubble sequence. At the vertex of the Hubble tuning fork, lie the lenticular galaxies, which present bulge and disc components as spiral galaxies, but lack spiral arms and relevant star-forming regions. Moreover, the gallery of galaxy types also encompasses irregular galaxies. Elliptical and lenticular galaxies are classified as 'early-type', while spirals and irregulars are called 'late-type' galaxies (here ETG and LTG, respectively).

In a binary classification (ETG or LTG), though, we are forcing the galaxies into one of two classes, while the classification could be more gradual, reflecting the complexity of galaxy shapes, such as when using the numerical Hubble types. To account for this, the network architecture in this work was slightly changed when compared to the one used in BOM21, in order to make the probabilities of ETG or LTG not complementary, i.e. not necessarily summing to one. In fact, these probabilities are generated independently in a way that a galaxy can have a high probability (higher than the DL

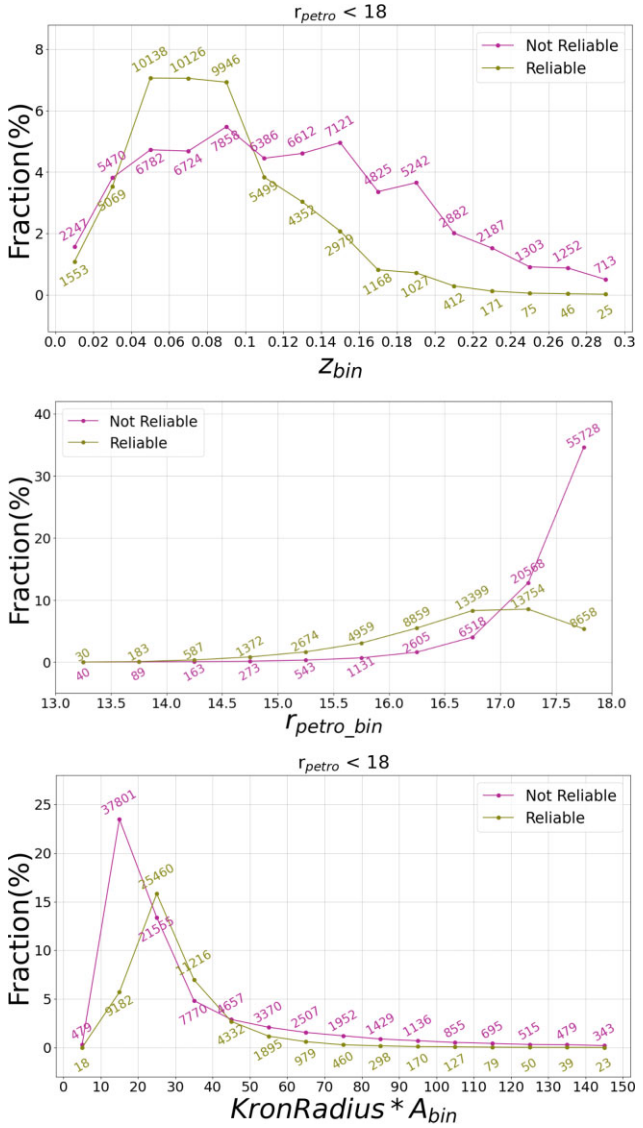

Figure 14. Number of not reliable and reliable stamps for increasing redshift bins (top), magnitude and Kron radius. The fraction was calculated in terms of the total number of objects in the catalogue.

threshold, see Fig. 7) of being both ETG and LTG. Galaxies that have a high probability of being both ETG and LTG are designated here as *Amb1*. This brings an interesting *ambiguity* to the model that can be explored to make the classification more gradual: a galaxy now can be classified as neither ETG nor LTG, and will be ascribed to class *Amb0*.

In our results, as shown in Fig. 10, we can see that most of the galaxies that had a low probability of being a ETG or LTG (*Amb0*) were also classified as non reliable stamps, while those with higher probability of being either ETG or LTG (*Amb1*) were also classified as reliable. We note here that the galaxies with high probability of being non reliable stamps and with a low probability of being ETG or LTG, are the highest in number (1107), while the majority of galaxies that have high probability of being ETG or LTG (160) are classified as reliable stamps.

Fig. 11 shows examples of reliable stamps, as defined using the 12 S-PLUS images, see Section 3.2, of galaxies belonging to the four different classes (ETG, LTG, Amb_0, Amb_1), from S-PLUS and Legacy surveys. The Legacy data are typically four magnitudes deeper than S-PLUS images and reveal faint outer features, so they can be used to understand the effects of depth and resolution in the ability of the DL method classify objects. In general, galaxies falling in the ETG class are ellipticals (left column, top and middle rows) or lenticulars (left column, bottom row). The LTG objects are either spiral or irregular galaxies (second column, first and middle rows), while the third row shows a disc-dominated lenticular galaxy. In Section 5.1, we compare the classification presented in this work with other works.

Galaxies are classified as *Amb0* or *Amb1* as the result of a combination of factors:

(i) faint/high redshift spiral galaxies can be misclassified as ETG, due to the pixel resolution and survey depth, which reflects in the difficulty of identifying the presence of spiral arms. In turn, they might present green dots of star formation, rendering them neither ETG nor LTG (see third column, middle panel of Fig. 11);

(ii) clumpy star-forming galaxies could also be assigned to neither class, due to their un-smooth appearance and the absence of clear spiral patterns, see third column top and bottom panel of Fig. 11;

(iii) bulge-dominated spirals (see last column, top and middle images), due to the low surface brightness of their spiral arms, clearly visible in the Legacy data, but close to the image noise in S-PLUS data, may have a high probability of being both ETG and LTG galaxies;

(iv) lenticular galaxies can also be found in the Amb_1 class, in particular lenticular galaxies with $B/T \simeq 0.5$ are associated to both classes, due to their hybrid nature. These results will be further discussed in Section 5.3.

(v) Galaxy size, as in the top and bottom panel of the third column (Amb_0), could compromise the performance of the DL classification, giving to the galaxy a low probability of being either late or early type.

(vi) The presence of companion galaxies, as in the middle row of the third and fourth columns, may also be a cause of ambiguity in the classification.

Points (i) and (iii) are mainly a consequence of the loss of resolution, which, given a fixed pixel scale, is directly connected to the objects brightness. In this work, we are using the training sample of BOM21 ($r < 17$) to classify object up to $r < 18$, see Section 2 for more details. Extrapolating the morphological classification for objects that are fainter, and therefore might present structural differences from the training sample, might lead to misclassification. To address this problem, we run MFMTK (Ferrari, de Carvalho & Trevisan 2015) to *i* band image stamps of the training and the blind samples, to obtain morphometric parameters, such as concentration (C) and entropy (H). In general, parametric and non-parametric quantities, as the Sersic index, *n*, (Peng et al. 2002; Barden et al. 2012) and extended CAS (Conselice 2014) estimates, respectively, correlates with galaxy morphologies. LTG have lower values of *C* and *n* indexes, reflecting the presence of the disc, and high values of H, which measures how clumpy is the image. In Fig. 12, we show the location of galaxies brighter, top, and fainter, bottom, of $r = 17$ mag, in the H–C plane, colour coded according to their probability of being spiral galaxies, estimated by the DL method. We can see that the transition between LTG and ETG is clearly depicted by this diagram, for both magnitude ranges ($r < 17$ and $r > 17$). We note that the number of objects in the bottom panel is lower than in the upper panel, since we selected only objects classified as 'reliable stamps'.

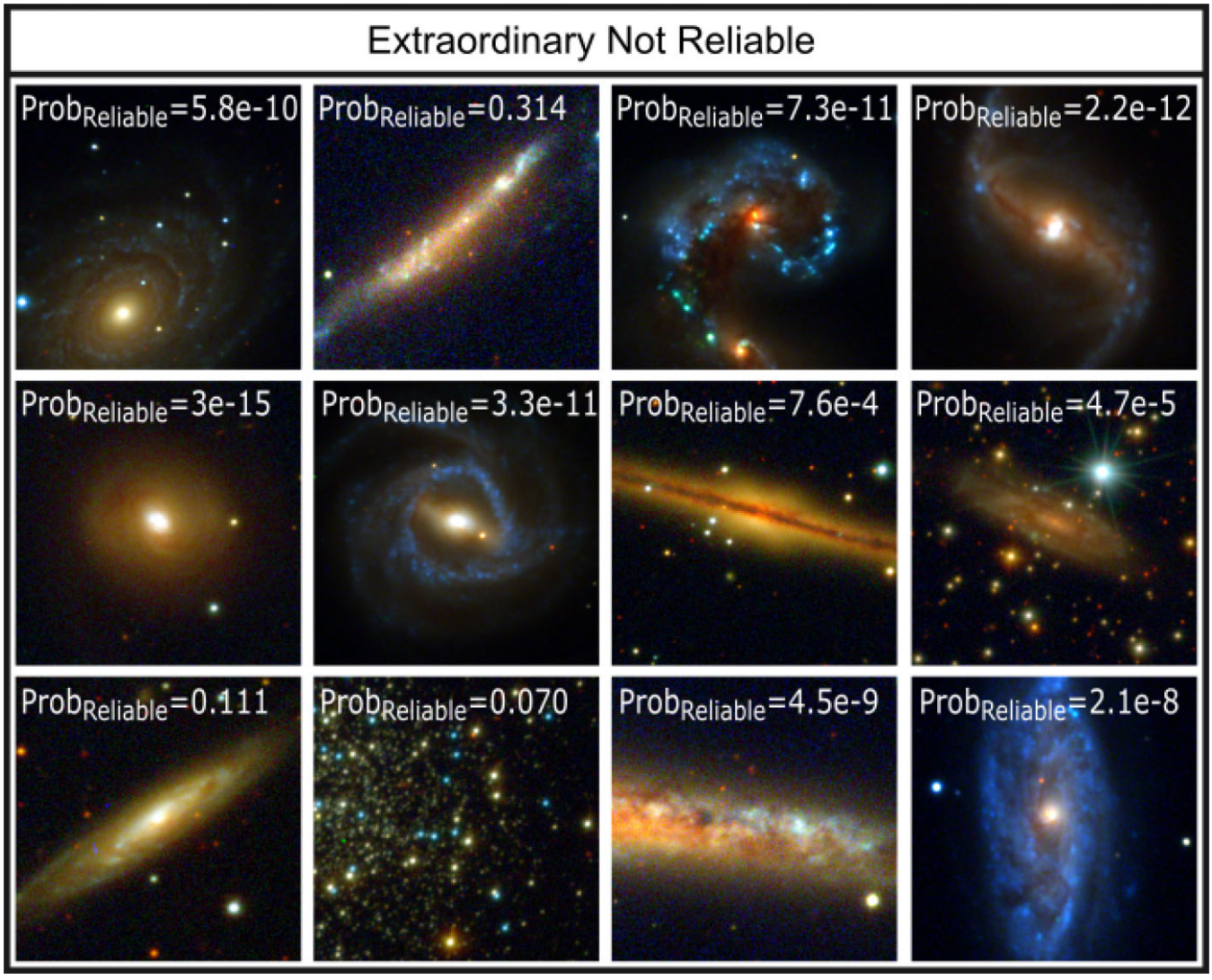

Figure 15. Example of extraordinary not reliable stamps. Large objects, whose projected radius is larger than the image stamp; star-forming galaxies, where more than one clump can be identified as an independent object during the catalogue extraction (Almeida-Fernandes et al. 2022); irregular galaxies as NGC 4038/NGC 4039; and dense regions of stars, maybe galactic clusters, can be found among the not reliable stamps. The probability of being a reliable stamp is given in the top left of each panel.

4.4 Extraordinary non-reliable stamps (NRS)

Fig. 13 shows some example of objects identified as NRS. Generally, they are objects nearby saturated stars or crowded fields. In fact, even if we select the sample of objects to be analyzed maximizing the probability of being galaxies, see Section 2.1.1, contaminants still appear in the sample and the deep learning code makes a great job in identifying spurious objects. The number of NRS is nearly constant with redshift, as shown in Fig. 14, while the number of Reliable stamps increases at $z = 0.04$ (where the number of objects in the training sample increases), remains constant till $z = 0.1$, and then decreases with increasing redshift. On the other hand, the number of RS is always higher than the number of NRS, for $r < 17$, where the number of NRS increases abruptly, as a combination of the decreasing in the image resolution and of the absence of objects in this magnitude range in the training sample. Finally, the number of RS is higher than the number of NRS only for intermediate radius, i.e. 20 < Kron Radius > 40 pixels, again reflecting the distribution of the training sample.

On the other hand, peculiar galaxies, especially if with clumpy star formation, or galaxies with a projected size larger than the stamp might fall in the category of NRS, as shown in Fig. 15. Somehow, the deep learning method not only allows us to identify unwanted objects, but it also helps in finding peculiar objects, of high interest/relevance.

4.5 Morphology as a probe of galaxy evolution and large-scale structure formation

Galaxies evolve through time via different mechanisms: major and minor mergers, secular evolution, harassment, stripping, and strangulation (Gunn & Gott 1972; Byrd & Valtonen 1990; Quilis, Moore & Bower 2000; Bournaud, Jog & Combes 2005; Aragón-Salamanca 2008; Kronberger et al. 2008). Many of these processes are environment-dependent, i.e. they can occur only in clusters of galaxies (strangulation), or they are more likely in the field or groups (e.g. mergers). In general, it is now believed that minor mergers are more common than major mergers and that they are the main

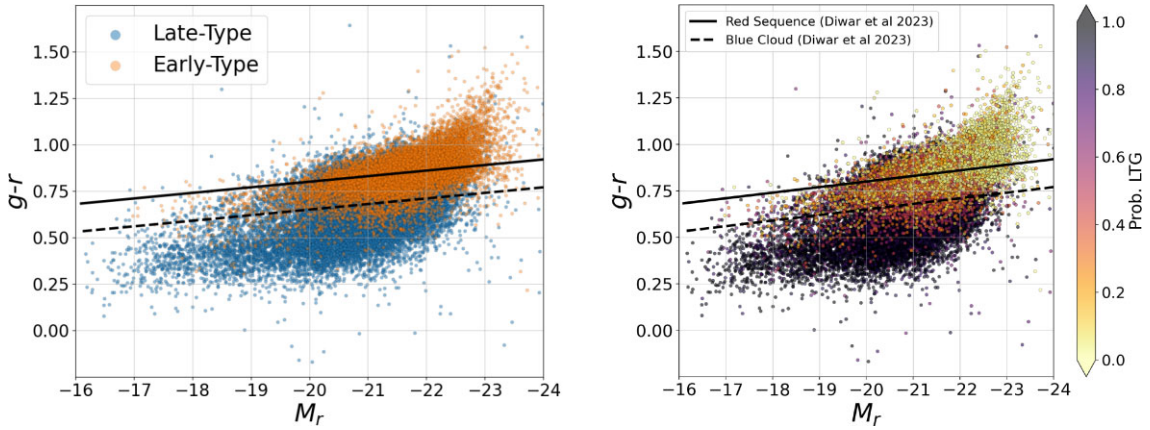

Figure 16. Colour–magnitude diagram of the objects classified as 'reliable stamps'. ($g - r$) Colour versus the absolute magnitude in r band, calculated using standard cosmological parameters and the luminosity distance (D_L) estimated from the photometric redshift. The *left*-hand panel shows the bin classification, while the *right*-hand panel, is colour coded according to the probability of being a LTG. The dashed and filled lines separate the red sequence, green valley and blue cloud (Dhiwar et al. 2022). The bulk of ETG lie within the red sequence and green valley area, but we can notice blue ETG and red spirals.

responsible for mass build-up in galaxies (Bournaud, Jog & Combes 2007). Different evolutionary scenarios leave specific imprints on the galaxy morphology; i.e. major mergers tend to disrupt the stellar orbits, resulting in pressure-dominated systems, characterized by an elliptical shape. On the other hand, secular processes, or environmentally driven mechanisms, as ram-pressure stripping, affect more the gaseous component, chasing the star formation. Moreover, a morphology–density relation had been proven in the last decades (Dressler 1980; Dressler et al. 1997; Cappellari et al. 2011; Buitrago et al. 2013), where ETG inhabit the densest regions of the Universe, while spiral and irregular galaxies are more common in the fields. Galaxy morphologies are a powerful proof of galaxy evolution as well as structure formation, as we will discuss in the next sections. In here we use only objects classified as reliable stamps with $photo_z$ $odds > 0.4$ and $r < 17$ mag. The magnitude have been corrected for galactic extinction using the Clayton, Cardelli and Mathis (Cardelli, Clayton & Mathis 1989) dust law.

4.5.1 *Correlation between morphology and colors*

Fig. 16 shows the colour–magnitude diagram ($g - r$ colour versus r-band absolute magnitude), colour coded according to the galaxies' morphologies. The left-hand panel presents the dual classification, where elliptical galaxies are shown in orange and spiral galaxies, in light-blue, while in the right-hand panel the colour scale shows the probability of being a spiral galaxy. It can be seen that elliptical (quiescent) galaxies inhabit the red sequence while spiral (star-forming) galaxies are mostly found in the 'blue cloud' as expected according to their dominant stellar populations (see e.g. Wong et al. 2012; Lima-Dias et al. 2021; Khanday et al. 2022). Interestingly, in the right panels it is possible to see that the probability of being spiral increases nearly from 0 to 1, going from the red cloud to the blue sequence, in a continuous manner. The intermediate region, where the probabilities range around 0.5, is known as Green Valley (see e.g. Zibetti et al. 2007) and it has been largely studied as a region of transition, where LTG could be quenching their star formation, turning into LTG, or ETG could be 'rejuvenating', due to some interaction with other galaxies or accretion of gas (Smethurst et al. 2015). The morphology seems to be reflecting this transformation since the quenching is 'removing' the spiral arms, decreasing the probability of being a spiral galaxy. On the other side, a sparkle of star formation in an early-type galaxy could create clumps, or star-forming regions, that would increase the probability of being a spiral galaxy. Finally, we can identify a small fraction of blue ETG and red LTG, highlighting that colour is not taken into account by the DL classification. Specifically, using the $g - r$ colour and the threshold defined in Dhiwar et al. (2022) to identify the red sequence and the blue cloud regions, we find that around 5 per cent of the blue galaxies were classified as early type and around 37 per cent of the red galaxies were classified as late type, see Section 3.3 for more details.

4.5.2 *Morphology–density relation*

There is a connection between the environment a galaxy live in and its morphology (Dressler 1980), but both the galaxy stellar population and environment evolve with time. While the galaxy stellar population is related to the galaxy mass (more massive galaxies are more metal rich at a given time, Leaman et al. (2013) and gas content, the morphology is more related to the environment (spiral galaxies tend to live in low density environment, ellipticals in the center of galaxy clusters). Yet, a merger, whose probability is dictated by the environment a galaxy lives in, would affect both the galaxy mass and stellar population. Note that 20 per cent of high mass ($M_* \geq 10^{9.5}$ $M_\odot$) galaxies have experienced a major merger since $z \simeq 6$ (Ventou et al. 2017) and minor mergers, accretions, fly-bys are very common in the history of the Universe.

We use a K-Nearest Neighbor method, with $k = 4, 5, 10$ (Baldry et al. 2010), to recover the projected density of the environment a galaxy inhabits, where $k = 4, 5$ refers to local environments, while $k = 10$ is related to larger scales. Specifically, the density (Σ_k) at any given k is:

$$\Sigma_k = \frac{k}{\pi D_k^2} \frac{1}{\psi(D))}, \tag{11}$$

where k is the k nearest neighbour, D is the comoving distance and $\psi(D)$ is a selection function to correct for the Malmquist bias (e.g. Santana-Silva et al. 2020). Fig. 17 presents the number density of LTG and ETG for increasing $k = 4$ density measures. The top panel of Fig. 17, presents all galaxies with magnitude $r \leq 17$, while the right-hand panel split them into magnitude bins (represented by different line shapes, see figure legend). ETG are identified by orange/red lines, while LTG are shown as cyan/blue lines. The morphological

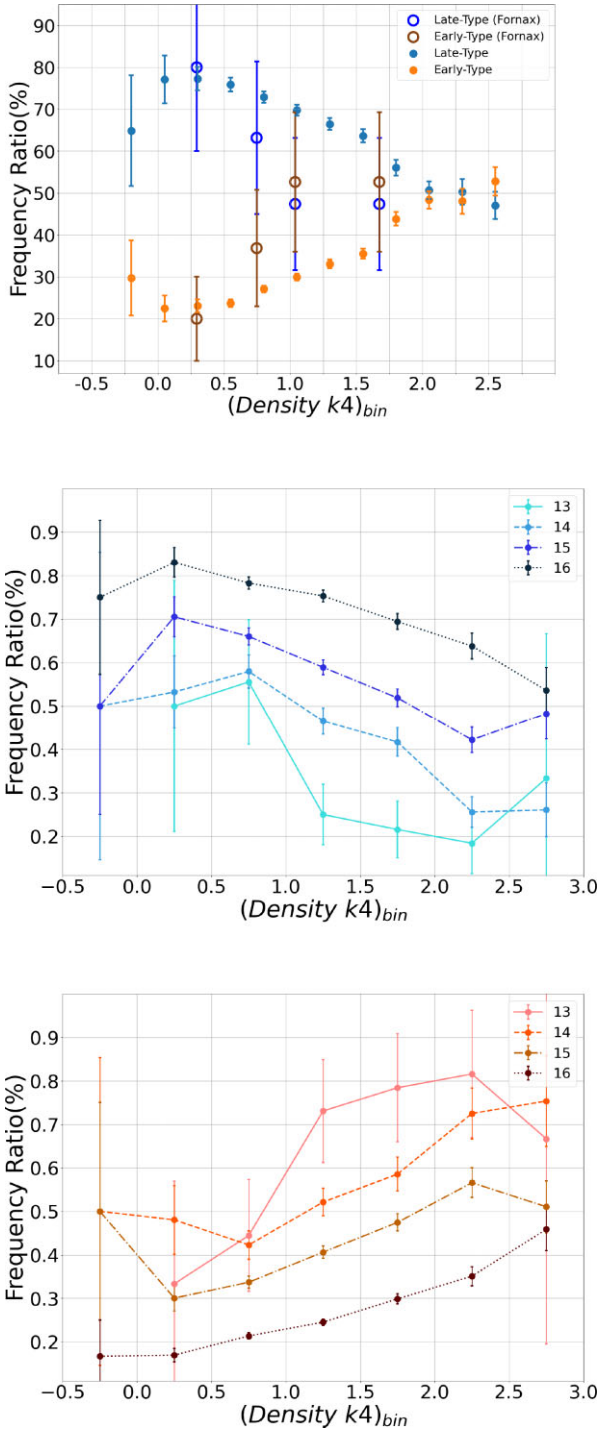

Figure 17. Morphology–density relation. Normalized fraction of LTG and ETG, for increasing density bin. The k4 estimator traces the local densities. The *top* panel shows the total distribution (for the general sample and galaxies in the Hydra cluster – see legend). The *middle* panel, present the same result, divided in different magnitudes bins for LTG, which are drawn in a scale of blue (considering the whole sample). In the *bottom* panel ETG are coloured in orange. For easier comparison, lines of equal magnitude have the same style (considering the whole sample). All galaxies used were classified as 'reliable stamps'.

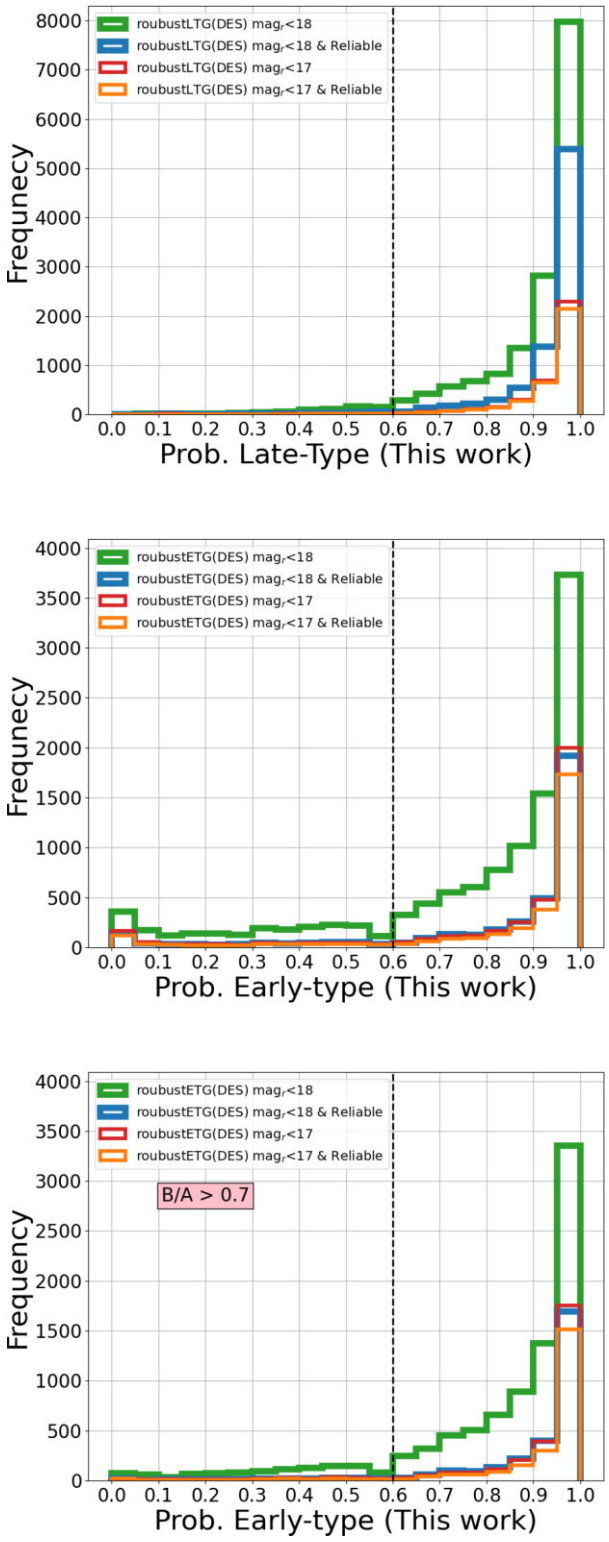

Figure 18. In this figure, we present a histogram showing the proportion of galaxies that were classified in accordance with our classification. The *top panel* represents the galaxies classified as roubust LTG by DES and the *middle* one represents those classified as roubust ETG by DES. The *bottom panel* is like the middle one, but for robust ETG with $b/a > 0.7$, see text. The histograms were made using the probability of belonging to the corresponding classes obtained in this work, with the dashed line being the threshold used in our classification, in other words, every galaxy that stands in the right side of this line was classified equally by both works.

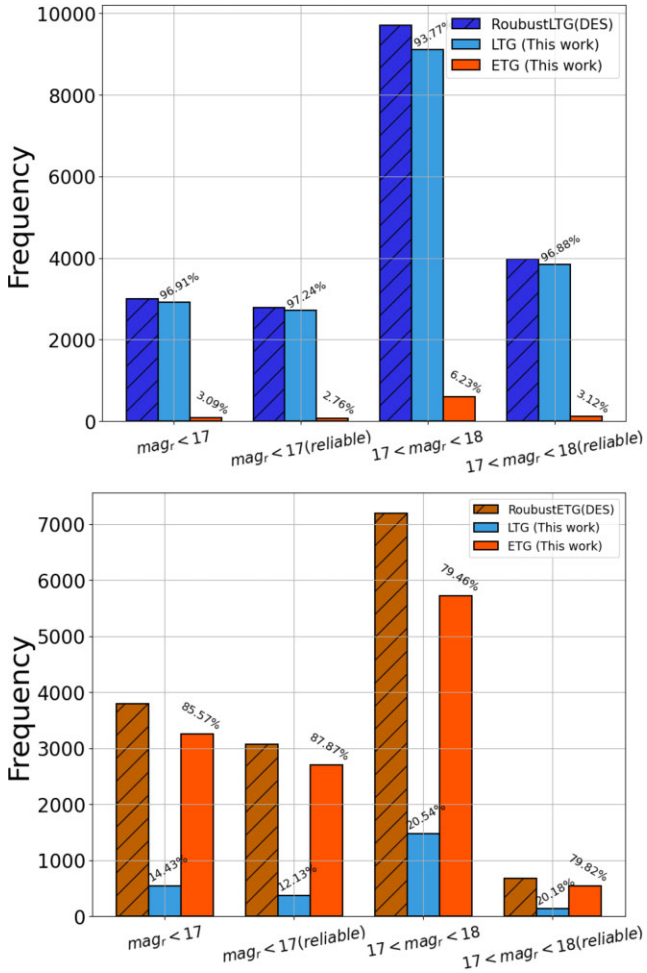

Figure 19. (Top) Number of objects classified as robust late types by Vega-Ferrero et al. (2021) for different magnitudes bins, dark blue histogram. The percentage at the top of each bin indicates the fraction of objects classified as late, cyan, and as early, orange, types in this work. (Bottom) Number of objects classified as robust early types by Vega-Ferrero et al. (2021) for different magnitudes bins, dashed orange histogram (note, we excluded objects classified as edge-on in Vega-Ferrero et al. (2021)). The cyan and orange histogram are as above. We note that we have an agreement of $\simeq 96$ per cent for reliable LTG in all magnitudes bin, which decreases to $\simeq 87$ per cent for reliable ETG brighter than $r < 17$, and $\simeq 80$ per cent for reliable ETG with $17 < r < 18$.

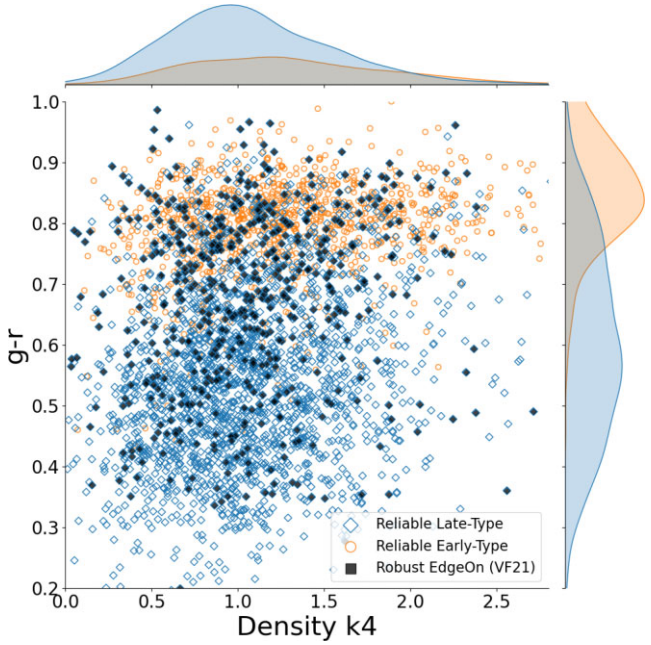

Figure 20. $(g - r)$ Colour versus density k4 distribution for RS galaxies. The filled yellow circles are classified as robust Edge-on in Vega-Ferrero et al. (2021). The histograms represent the colour and density distribution of each morphological class.

classification provided in this work clearly reflects the morphology density relation, with ETG occupying the densest regions, and LTG being the dominant population in the field/low-density regions, see left-hand panel in Fig. 17. In the top panel, we identify with open large circles LTG and ETG belonging to the Hydra cluster. We notice, that there are no galaxies associated with Hydra with $k_4 < 0.2$. Then, the overall trend for LTG and ETG with densities follow a similar trend for Hydra galaxies and galaxies belonging to the general sample, as expected since k_4 describes the local densities, and not the large-scale structure. We note that in the last bin for the Hydra cluster sample, the number of ETG and LTG overlap withing error bars, as a result of the difficulty of leading with crowded fields. A similar trend between morphology and environmental densities had been found by Thomas & Katgert (2006) for the Abell cluster survey, as well as by Pfeffer et al. (2022) in the EAGLE simulation.[6]

When looking at the magnitude dependence of the morphology–density relation, we see that it holds for different magnitudes bins, where the number density of ETG increases with increasing densities, while the opposite trend is found for LTG. Finally, we observe that the crossover density is lower for more luminous objects, indicating a correlation between lower densities and higher luminosities.

5 DISCUSSION AND CONCLUDING REMARKS

5.1 Comparison to other surveys

Vega-Ferrero et al. (2021) used DES galaxies with reliable morphological classification to assess whether CNNs are able to detect features that human eyes do not. To do that, they simulate the appearance that well morphologically classified DES galaxies would display at high redshifts, making them fainter and smaller. They find that, despite some of the features that distinguish ETGs from LTGs vanish after the simulation, the models are still able to correctly classify galaxies with an accuracy greater than 97 per cent. The main conclusion of that work is that it is possible to correctly classify galaxies from faint and small size images using CNNs models, satisfying the following conditions: final apparent magnitude below $m_r(z) < 22.5$, and the size of the final image larger than 32×32 pixels. DES data (DES-DR1; Abbott et al. 2018) has a median co-added catalogue depth of $m_r = 24.08$ at signal-to-noise ratio S/N = 10, with a pixel size of 0.2636 arcsec.

In comparison, S-PLUS has a scale of 0.55 arcmin pixel^{-1} and a depth in r band of $m_r = 19.6$ at signal-to-noise ratio S/N = 10

[6]In both works, the ET galaxies are subdivided into E and S0 galaxies.

(Almeida-Fernandes et al. 2022), resulting in lower resolution when compared with DES data, as clear from Fig. 11.

S-PLUS DR3 and DES DR1 overlap, see Fig. 1, resulting in a combined catalogue from Vega-Ferrero et al. (2021) and this work of 36 183 galaxies, brighter than $m_r < 18.0$ and with a mean redshift of $z_{ml} = 0.11847$. Comparing the classification presented in this work, considering the depth of the DES images, allows us to investigate the goodness of the classification and the advantages of combining the results of the two DL codes, i.e. studying the reliable early- and late-type classification. In Fig. 18, top panel, we present a histogram of the probability of being LTG obtained in this work, for galaxies classified as 'robust spirals' ($FLAG_{\rm LTG} == 5$) in Vega-Ferrero et al. (2021). The dashed line indicates the threshold used in this work, in other words, every galaxy that stands in the right side of this line is classified as a spiral in both works. The blue histogram shows the distribution of the probability of being a LTG for all 'robust spiral' galaxies and presents the larger discrepancy with Vega-Ferrero et al. (2021). The orange histogram shows the probability of being LTG for all 'robust spiral' stamps classified as reliable according to the second DL model, see Section 3.2. The green and red histograms represent all 'robust spiral' galaxies brighter than $r < 17$ mag, and among them all the ones classified as reliable stamps, respectively. The middle panel shows the same comparison for ETG. There is a non-negligible fraction of galaxies with zero probability of being ETG in this work, but classified as elliptical in Vega-Ferrero et al. (2021). In the bottom panel, we reproduce the same plot, now including only objects with $b/a > 0.7$. This choice drastically decreases the number of discrepant classifications. Similar results are obtained when performing the same comparison with Cheng et al. (2023).

In Fig. 19, top panel, we present the fraction of robust LTG (Vega-Ferrero et al. 2021) classified as late and early type in this work for different magnitudes bins, and in the bottom panel we present the same distribution for galaxies classified as robust early types in Vega-Ferrero et al. (2021). We exclude objects that are classified as edge-on by Vega-Ferrero et al. (2021). Figures available in the appendix,[7] it is possible to find examples of objects classified differently in the two papers. It is noticeable that in many cases of galaxies classified as early types in this work and late types in Vega-Ferrero et al. (2021), they are multiple object images, low surface brightness, bulge-dominated spiral galaxies, or faint/compact spiral galaxies, see Section 4.3. On the other hand, objects classified as late types in S-PLUS and early types in Vega-Ferrero et al. (2021) are often disc-dominated (edge-on) lenticular galaxies or merger/disturbed systems. In conclusion, the classification presented in this work is in agreement with Vega-Ferrero et al. (2021) with an average confidence level of $\simeq 88$ per cent, up to $r < 17$, for reliable ETG and $\simeq 97$ per cent, for reliable LTG, up to $r < 18$. The mismatch for ETG increases to 20 per cent for objects fainter than $r \simeq 17$ as a result of the fading of the spiral arms in the S-PLUS images. On the other hand, the mismatch for LTG is mostly caused by the association of disc-dominated lenticular galaxies or edge-on red spirals (Sodré, Ribeiro da Silva & Santos 2013) to this class in this work, while there is a perfect agreement between the two classifications when considering only edge-on objects (Vega-Ferrero et al. 2020) and $r < 17$. Implications from these results are further discusses in Section 5.3. Moreover, a visual inspection of the differently classified objects, see the panel figures in appendix A, reveals interesting objects resulting from a different structure of the DL codes and image depth and resolution, highlighting the importance of a diverse, open and collaborative scientific environment.

[7]The appendix is presented as an online supplementary material.

5.2 Combining morphology and precise photometric redshifts: where do galaxies live?

The relation between galaxies' morphology, their mass, stellar population properties, and the environment they live in has been studied in great detail and in a wealth of works (Paulino-Afonso et al. 2019; Coccato et al. 2020), as well as its redshift evolution (González Delgado et al. 2015). Recent works show that the bulge growth, measured as bulge over total light ratio, is directly connected with the quenching of the star formation (Paulino-Afonso et al. 2019; Dimauro et al. 2022; Werner et al. 2022). Group pre-processing is also found to play an important role in galaxies' star formation quenching and morphological evolution (González Delgado et al. 2022; Brambila et al. 2023).

S-PLUS photometric system allows the retrieval of reliable photometric redshifts with a scatter of 0.023 (Lima et al. 2022) and the recovery of reliable density estimates (Lopes da Silva et al. in preparation). In Fig. 20, we show the $(g-r)$ colour versus k4-density measures for reliable stamps with $Prob_{\rm Spiral} > 0.9$ (blue open dots), with $Prob_{\rm Elliptical} > 0.9$ (red open dots), and galaxies classified as edge-on by Vega-Ferrero et al. (2021) as filled yellow circles. The galaxy environment is more correlated to its morphology, than its colour, see Figs 17 and 20. The bulk of the ETGs have $(g-r) > 0.7$. LTG have colours in the range $0.1 \leq (g-r) \leq 1.0$, with the majority of the redder LTG, i.e. $(g-r) > 0.7$, being classified as edge-on in Vega-Ferrero et al. (2021). As shown in Fig. 11 disc-dominated lenticular galaxies can be associated to the late-type class, explaining the red colour of the discy-late-type galaxies. Moreover, edge-on star-forming spiral galaxies might suffer reddening, due to the presence of dust clouds surrounding the disc (Bamford et al. 2009; Sodré, Ribeiro da Silva & Santos 2013). Fig. 20 points out that ETG have red colours, $(g-r) > 0.7$ and are more common in denser environments. If colour is a proxy for galaxy stellar population, these findings would suggest that both the quenching of the stellar population and the environment are connected with the early-type morphology. On the other hand, late-type galaxy span a range of colour and their number seems to be more connected to the environment they live in, see Fig. 17. Similar results are found by Bamford et al. (2009) who show that nearly 30 per cent of the galaxies in their study, using GalaxyZoo 1 data release, are consistent with being red spiral, while only 10 per cent are blue elliptical galaxies. They also find that blue ETG live in low density environments, while red spiral inhabit higher density environment.

Galaxy morphology are also important to optimize halo occupation models and increase the accuracy of DM distribution models. Infact, galaxies trace the large-scale structure of the Universe, yet they account only for $\simeq 20$ per cent of the total matter (e.g. Planck Collaboration VI 2020), and their physics is affected by non-gravitational mechanisms such as baryonic effects, radiation pressure, feedback, etc. A simple but powerful tool that can bridge the gap between galaxies and the DM distribution is the halo model (Cooray & Sheth 2002). One of the consequences of that description is that galaxy abundances and their properties (such as stellar mass, colour, morphology, and star formation rate) can be traced back to the DM haloes and sub-haloes, as well as their properties (such as mass, age, concentration, and spin) Wechsler & Tinker (2018). From a large-scale structure perspective, the correlation function of the DM haloes is related to the correlation function of the DM particles by the halo abundance, bias, and halo density profile. To a

good approximation, more massive haloes are less abundant and are more highly biased with respect to the DM field, but other halo properties such as concentration, age, and even spin (angular momentum) also play an important role (Montero-Dorta et al. 2020). Galaxies that populate haloes and their sub-haloes inherit those properties, including their bias – but they can also bring additional information that is not manifested in the halo properties, and which are indicative of the baryonic mechanisms such as ram-pressure stripping. Galaxy morphology is, therefore, one of the additional indicators that can help distinguish between different types of haloes and their environments, leading to a more accurate and precise description of the correlation functions of those tracers.

5.3 S0 galaxies: considerations arising from the association to the early- and late-type classes

Lenticular galaxies are characterized by a hybrid morphology, with a bulge and a disc as spiral galaxies, but without spiral arms, as elliptical galaxies. van den Bergh (1990) suggest that the 'S0 classification type comprises a number of physically quite distinct types of objects that exhibit only superficial morphological similarities'. Recent studies based on observations (Fraser-McKelvie et al. 2018; Coccato et al. 2020; Domínguez Sánchez et al. 2020; Tous, Solanes & Perea 2020) and on simulations (Deeley et al. 2021) showed that this class of objects is, indeed, composed by two or more subgroups, formed via different physical mechanisms that lead to a similar morphology. Specifically, stripped spiral galaxies could be the progenitors of disc-dominated lenticular galaxies, if the gas and arms of the spiral galaxies would be removed by interactions with the cluster environment, or by harassment in a group environment or generally pestering in all environments (Cortesi et al. 2013; Jaffé et al. 2015; Johnston et al. 2021). Another group of lenticular galaxies could be the result of major or minor mergers and multiple accretions (Tapia et al. 2017). Others, low mass S0s, could be pristine galaxies, formed at redshift $z \simeq 2$ from mergers of the galaxy stellar/gaseous clumps (Saha & Cortesi 2018), or the result of secular evolution (Mishra, Wadadekar & Barway 2018).

The discrepancy between the probability of being LTG predicted in this work and in Vega-Ferrero et al. (2021) decreases when only objects with $B/A > 0.7$ are considered, see Fig. 18. Moreover, there is an extended population of LTG with red colours and high probability of being edge-on systems (Vega-Ferrero et al. 2021), see Fig. 20.

Specifically, there are 126 objects classified as spiral galaxies in this work and as robust ellipticals inVega-Ferrero et al. (2021), with $B/A \leq 0.7$ and $r_{\rm petro} < 17$ mag. At a visual inspection they are all consistent with being disc-dominated S0 galaxies (Coccato et al. 2020), or edge-one reddened spiral galaxies (Bamford et al. 2009; Sodré, Ribeiro da Silva & Santos 2013) and their average colour is $(g - r) \simeq 0.85$.

On the other hand, as discussed in Section 4.3, a fraction of galaxies classified as early-type galaxies in this work is comprised by bulge-dominated lenticular galaxies.

The multiple origin of the S0-like isophotal profile seems to be depicted by the DL algorithm used in this work. Specifically, lenticular galaxies which originates from (multiple) mergers, would be associated to the early-type galaxy class, due to their boxy isophotal profile, while stripped spiral galaxies (i.e. S0 galaxies resulting by ram pressure stripping of spiral galaxies) would be associated with the late-type galaxy class, since their isophotal profile is discy. Clearly, the inclination of the galaxy also plays an important role in the classification, especially for edge-on objects or face-on objects. In the first case, the presence of spiral arms is hard to be detected, rendering the edge-on S0s similar to edge-on late-type galaxies. In the second case, the projection of the disc on the plane of the sky could resemble an early type galaxy. We note here, that we do not directly classify S0 galaxies in this work, but we infer their presence in the ambiguous class or by studying objects associated to a different class in this work and in Vega-Ferrero et al. (2021). This topic will be further studied in a folow-up work, where the DL classification will be correlated with galaxies' bulge-to-total light profiles.

5.4 Summary

In this study, we employ a Deep Learning architecture among the top-ranking techniques for image classification of ETG/LTG while also introducing a model to predict the stamps that contain reliable information to be classified. Our method presented several innovations compared to the Bom21 model, including the possibility of objects that are not classified either as ETG or LTG.

Furthermore, we also make use of the precise photometric redshifts derived with 12 bands present in the S-PLUS. We recover the colour diagrams for the morphological types and examine the local environment and density of ETG/LTG. Additionally, we assess large-scale structure traced by morphology. As a result, we provide a novel Value Added Catalogue (VAC) of galaxy morphologies, in the full footprint of S-PLUS DR3 which includes areas never explored for other galaxy morphology catalogues. The catalogue is composed by the results of two DL methods, which, for every stamp, recover the probability of having a given morphology, and of being a reliable stamp, as detailed below.

5.4.1 A novel valued added morphology classification catalogue for the Southern hemisphere

In order to mediate between the variety of galaxies morphologies and the binary classification applied in this work, we allow for an independent classification into early- and late-type galaxies, i.e. the sum of the probability of belonging to each class does not sum up to one, see Section 4.2.1. As a consequence of this choice, some objects can be classified as belonging to either group (using the binary classification) or to any group, see Section 4.3. The study of these two peculiar types of objects allows us to identify bulge-dominated lenticular or spiral galaxies (Amb_1), as well as compact, flocculent, star-forming galaxies, see Fig. 11. Finally, this catalogue of galaxy morphologies covers areas of the Southern Sky for which there is no release of morphological catalogues (see Fig. 1).

5.4.2 A novel parameter to assign a probability of being a reliable stamp

An interesting correlation is found when comparing the number of galaxies with low probability of being LTG or ETG (Amb_0), with the probability of being a reliable stamp, see Section 3.2. In fact, the majority of objects not classified as early- nor late-type galaxies (see previous Section 5.4.1) have a low probability of being reliable stamps, see Fig. 10. Fig. 18 reveals that selecting only reliable stamps decreases the discrepancy with Vega-Ferrero et al. (2021) classification into ETG and LTG, especially for faint objects ($m_r >$ 17). Moreover, as shown in Fig. 15, among the non reliable stamps, there are extraordinary objects, as the Antenna Galaxy, which will be identified and studied in a follow-up work.

ACKNOWLEDGEMENTS

The S-PLUS project, including the T80-South robotic telescope and the S-PLUS scientific survey, was founded as a partnership between the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), the Observatório Nacional (ON), the Federal University of Sergipe (UFS), and the Federal University of Santa Catarina (UFSC), with important financial and practical contributions from other collaborating institutes in Brazil, Chile (Universidad de La Serena), and Spain (Centro de Estudios de Física del Cosmos de Aragón, CEFCA). We further acknowledge financial support from the São Paulo Research Foundation (FAPESP), the Brazilian National Research Council (CNPq), the Coordination for the Improvement of Higher Education Personnel (CAPES), the Carlos Chagas Filho Rio de Janeiro State Research Foundation (FAPERJ), and the Brazilian Innovation Agency (FINEP).

The authors who are members of the S-PLUS collaboration are grateful for the contributions from CTIO staff in helping in the construction, commissioning, and maintenance of the T80-South telescope and camera. We are also indebted to Rene Laporte and INPE, as well as Keith Taylor, for their important contributions to the project. From CEFCA, we particularly would like to thank Antonio Marín-Franch for his invaluable contributions in the early phases of the project, David Cristóbal-Hornillos and his team for their help with the installation of the data reduction package JYPE version 0.9.9, César Íñiguez for providing 2D measurements of the filter transmissions, and all other staff members for their support with various aspects of the project.

CMdO and LSJ acknowledge funding for this work from FAPESP grants 2019/26492-3, 2019/11910-4, 2019/10923-5 and 2009/54202-8. GS, CMdO, and LS acknowledge support, respectively, from CNPq grants 309209/2019-6, 115795/2020-0 and 304819/201794. NM acknowledges the University of São Paulo PUB grant 83–1 of 2020. AC acknowledge the financial support provided by FAPERJ grants E-26/200.607 e 210.371/2022(270993).

CRB acknowledges the financial support from CNPq (316072/2021-4) and from FAPERJ (grants 201.456/2022 and 210.330/2022) and the FINEP contract 01.22.0505.00 (ref. 1891/22). KK acknowledges full financial support from ANID through FONDECYT Postdoctrorado Project 3200139, Chile.

The authors made use of multi GPU Sci-Mind machines developed and tested for Artificial Intelligence and would like to thank P. Russano and M. Portes de Albuquerque for all the support in infrastructure matters.

The authors made use and acknowledge TOPCAT[8] tool to analyse the data and astrotools (Cardoso 2022) to visualize the objects. For complementary visual inspection and some panels, the authors made use of small cut outs images from the Legacy Survey. The Legacy Surveys consist of three individual and complementary projects: the Dark Energy Camera Legacy Survey (DECaLS; Proposal ID #2014B-0404; PIs: David Schlegel and Arjun Dey), the Beijing-Arizona Sky Survey (BASS; NOAO Prop. ID #2015A-0801; PIs: Zhou Xu and Xiaohui Fan), and the Mayall z-band Legacy Survey (MzLS; Prop. ID #2016A-0453; PI: Arjun Dey). DECaLS, BASS, and MzLS together include data obtained, respectively, at the Blanco telescope, Cerro Tololo Inter-American Observatory, NSF's NOIRLab; the Bok telescope, Steward Observatory, University of Arizona; and the Mayall telescope, Kitt Peak National Observatory, NOIRLab. The Legacy Surveys project is honored to be permitted to conduct astronomical research on Iolkam Du'ag (Kitt Peak), a mountain with particular significance to the Tohono O'odham Nation. The authors thanks J. Crosset for inspiring discussions and suggestions.

[8]http://www.starlink.ac.uk/topcat/ (TOPCAT)

DATA AVAILABILITY

We publicly release our VAC in the S-PLUS data base (splus.cloud).

REFERENCES

Abbott T. M. C. et al., 2018, ApJS, 239, 18
Abdel-Hamid O., Mohamed A.-r., Jiang H., Deng L., Penn G., Yu D., 2014, IEEE/ACM Trans. Audio Speech Lang. Process., 22, 1533
Almeida-Fernandes F. et al., 2022, MNRAS, 511, 4590
Ann H. B., Seo M., Ha D. K., 2015, ApJS, 217, 27
Aragón-Salamanca A., 2008, in Bureau M., Athanassoula E., Barbuy B.ed., Proc. IAU Symp. 245, Formation and Evolution of Galaxy Bulges. Kluwer, Dordrecht, p. 285
Axelrod T. S., 2006, in Gabriel C., Arviset C., Ponz D., Enrique S.eds, ASP Conf. Ser. Vol. 351, Astronomical Data Analysis Software and Systems XV. Astron. Soc. Pac., San Francisco, p. 103
Baldry I. K., Glazebrook K., Brinkmann J., Ivezić Ž., Lupton R. H., Nichol R. C., Szalay A. S., 2004, ApJ, 600, 681
Baldry I. K. et al., 2010, MNRAS, 404, 86
Bamford S. P. et al., 2009, MNRAS, 393, 1324
Barden M., Häußler B., Peng C. Y., McIntosh D. H., Guo Y., 2012, MNRAS, 422, 449
Bernardi M., Domínguez Sánchez H., Brownstein J. R., Drory N., Sheth R. K., 2019, MNRAS, 489, 5633
Bertin E., Arnouts S., 1996, A&AS, 117, 393
Bom C. R., Makler M., Albuquerque M. P., Brand t C. H., 2017, A&A, 597, A135
Bom C. R. et al., 2021, MNRAS, 507, 1937 (BOM21)
Bom C. et al., 2022, MNRAS, 515, 5121
Bournaud F., Jog C. J., Combes F., 2005, A&A, 437, 69
Bournaud F., Jog C. J., Combes F., 2007, Astron. Astrophys., 476, 1179
Brambila D., Lopes P. A. A., Ribeiro A. L. B., Cortesi A., 2023, MNRAS
Buitrago F., Trujillo I., Conselice C. J., Häußler B., 2013, MNRAS, 428, 1460
Buta R. J., 2013, Galaxy Morphology Terry D. Oswalt William C. Keel Springer Dordrecht 1–89 preprint(arXiv:1102.0550)978-94-007-5608-3
Byrd G., Valtonen M., 1990, ApJ, 350, 89
Calvi R., Poggianti B. M., Fasano G., Vulcani B., 2012, MNRAS, 419, L14
Cappellari M. et al., 2011, MNRAS, 416, 1680
Cardelli J. A., Clayton G. C., Mathis J. S., 1989, ApJ, 345, 245
Cardoso N. M., 2022, nmcardoso/astrotools: v0.1.0. Zenodo. Available at: https://doi.org/10.5281/zenodo.7268505
Cenarro A. J. et al., 2019, A&A, 622, A176
Cheng T.-Y., Li N., Conselice C. J., Aragón-Salamanca A., Dye S., Metcalf R. B., 2020, MNRAS, 494, 3750
Cheng T.-Y. et al., 2020, MNRAS, 493, 4209
Cheng T. Y. et al., 2023, MNRAS, 518, 2794
Choi K., Fazekas G., Sandler M., Cho K., 2017, IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, New Orleans, USA, p. 2392
Coccato L. et al., 2020, MNRAS, 492, 2955
Conselice C. J., 2014, Ann. Rev. Astron. Astrophys., 52, 291
Cooray A., Sheth R., 2002, Phys. Rep., 372, 1
Cortesi A. et al., 2013, MNRAS, 432, 1010
Crossett J. P., Pimbblet K. A., Stott J. P., Jones D. H., 2014, MNRAS, 437, 2521
Deeley S., Drinkwater M. J., Sweet S. M., Bekki K., Couch W. J., Forbes D. A., Dolfi A., 2021, MNRAS, 508, 895
Deng J., Dong W., Socher R., Li L.-J., Li K., Fei-Fei L., 2009, in Flynn P., ed., IEEE Conference on Computer Vision and Pattern Recognition. IEEE, Miami, FL, p. 248
Desai V. et al., 2007, ApJ, 660, 1151

Dhiwar S., Saha K., Dekel A., Paswan A., Pandey D., Cortesi A., Pandge M., 2022, MNRAS
Dimauro P. et al., 2022, MNRAS, 513, 256
Domínguez Sánchez H., Huertas-Company M., Bernardi M., Tuccillo D., Fischer J. L., 2018, MNRAS, 476, 3661
Domínguez Sánchez H., Bernardi M., Nikakhtar F., Margalef-Bentabol B., Sheth R. K., 2020, MNRAS, 495, 2894
Dressler A., 1980, ApJ, 236, 351
Dressler A. et al., 1997, ApJ, 490, 577
Edelen D. G. B., 1969, Ap&SS, 3, 56
Farias H., Ortiz D., Damke G., Jaque Arancibia M., Solar M., 2020, Astron. Comput., 33, 100420
Ferrari F., de Carvalho R. R., Trevisan M., 2015, ApJ, 814, 55
Fraser-McKelvie A., Aragón-Salamanca A., Merrifield M., Tabor M., Bernardi M., Drory N., Parikh T., Argudo-Fernández M., 2018, MNRAS, 481, 5580
Freeman P. E., Izbicki R., Lee A. B., Newman J. A., Conselice C. J., Koekemoer A. M., Lotz J. M., Mozena M., 2013, MNRAS, 434, 282
Gehrels N., Spergel D., 2015, J. Phys. Conf. Ser., 610, 012007
Glazebrook K., Jacobs C., Collett T., More A., McCarthy C., 2017, MNRAS, 471, 167
González Delgado R. M. et al., 2015, A&A, 581, A103
González Delgado R. M. et al., 2022, A&A, 666, A84
Goodfellow I., Bengio Y., Courville A., 2016, Deep Learning. MIT Press, Cambridge, Massachusetts
Grosbøl P., Dottori H., 2012, A&A, 542, A39
Gunn J. E., Gott III J. R., 1972, ApJ, 176, 1
Hannun A. Y., Rajpurkar P., Haghpanahi M., Tison G. H., Bourn C., Turakhia M. P., Ng A. Y., 2019, Nat. Med., 25, 65
Hausen R., Robertson B. E., 2020, ApJS, 248, 20
Herschel J. F. W., 1864, Phil. Trans. R. Soc. London Ser. I, 154, 1
Holincheck A. J. et al., 2016, MNRAS, 459, 720
Huertas-Company M. et al., 2015, ApJS, 221, 8
Jacobs C. et al., 2019, MNRAS, 484, 5330
Jaffé Y. L., Smith R., Candlish G. N., Poggianti B. M., Sheen Y.-K., Verheijen M. A. W., 2015, MNRAS, 448, 1715
Johnston E. J. et al., 2021, MNRAS, 500, 4193
Kelly B. C., McKay T. A., 2004, AJ, 127, 625
Khanday S. A., Saha K., Iqbal N., Dhiwar S., Pahwa I., 2022, MNRAS, 515, 5043
Knabel S. et al., 2020, AJ, 160, 223
Kronberger T., Kapferer W., Ferrari C., Unterguggenberger S., Schindler S., 2008, A&A, 481, 337
Lang D., 2014, AJ, 147, 108
Lanusse F., Ma Q., Li N., Collett T. E., Li C.-L., Ravanbakhsh S., Mandelbaum R., Póczos B., 2018, MNRAS, 473, 3895
Leaman R. et al., 2013, ApJ, 767, 131
Li X., Ding Q., Sun J.-Q., 2018, Reliab. Eng. Syst. Safety, 172, 1
Lima-Dias C. et al., 2021, MNRAS, 500, 1323
Lima E. V. R. et al., 2022, Astron. Comput., 38, 100510
Lintott C. J. et al., 2008, MNRAS, 389, 1179
Lintott C. et al., 2010, MNRAS, 410, 166
Liu L., Jiang H., He P., Chen W., Liu X., Gao J., Han J., 2020, OpenReview, ICLR 2020, preprint (arXiv:1908.03265), https://iclr.cc/virtual_2020/poster_rkgz2aEKDr.html
Lu J., Wang G., Zhou J., 2017, IEEE Trans. Image Process., 26, 4042
Ma Z. et al., 2019, ApJS, 240, 34
Madireddy S., Ramachandra N., Li N., Butler J., Balaprakash P., Habib S., Heitmann K., The LSST Dark Energy Science Collaboration, 2019, A Modular Deep Learning Pipeline for Galaxy-Scale Strong Gravitational Lens Detection and Modeling, preprint (arXiv:1911.03867)
Margalef-Bentabol B., Conselice C. J., Mortlock A., Hartley W., Duncan K., Ferguson H. C., Dekel A., Primack J. R., 2016, MNRAS, 461, 2728
Margalef-Bentabol B., Huertas-Company M., Charnock T., Margalef-Bentabol C., Bernardi M., Dubois Y., Storey-Fisher K., Zanisi L., 2020, MNRAS, 496, 2346
Mendes de Oliveira C. et al., 2019, MNRAS, 489, 241
Metcalf R. B. et al., 2019, A&A, 625, A119
Mishra P. K., Wadadekar Y., Barway S., 2018, MNRAS, 478, 351
Montero-Dorta A. D. et al., 2020, MNRAS, 496, 1182
Moreno-Torres J. G., Sáez J. A., Herrera F., 2012, IEEE Trans. Neural Netw. Learn. Syst., 23, 1304
Mortlock A. et al., 2013, MNRAS, 433, 1185
Nair P. B., Abraham R. G., 2010, ApJS, 186, 427
Nakazono L. et al., 2021a, MNRAS, 507, 5847
Nakazono L. et al., 2021b, MNRAS, 507, 5847
Niemack M. D., Jimenez R., Verde L., Menanteau F., Panter B., Spergel D., 2009, ApJ, 690, 89
Ostrovski F. et al., 2017, MNRAS, 465, 4325
Paulino-Afonso A. et al., 2019, A&A, 630, A57
Peng C. Y., Ho L. C., Impey C. D., Rix H.-W., 2002, AJ, 124, 266
Petrillo C. E. et al., 2017, MNRAS, 472, 1129
Petrillo C. E. et al., 2019a, MNRAS, 482, 807
Petrillo C. E. et al., 2019b, MNRAS, 484, 3879
Pfeffer J., Cavanagh M. K., Bekki K., Couch W. J., Drinkwater M. J., Forbes D. A., Koribalski B. S., 2022, MNRAS, 518, 5260
Planck Collaboration VI, 2020, A&A, 641, A6
Pović M. et al., 2015, MNRAS, 453, 1644
Quilis V., Moore B., Bower R., 2000, Science, 288, 1617
Saha K., Cortesi A., 2018, ApJ, 862, L12
Sánchez S. F., Cardiel N., Verheijen M. A. W., Pedraz S., Covone G., 2007, MNRAS, 376, 125
Santana-Silva L. et al., 2020, MNRAS, 498, 5183
Sarkar S., Pandey B., 2020, MNRAS, 497, 4077
Shamir L., Holincheck A., Wallin J., 2013, Astron. Comput., 2, 67
Shao X., Disseau K., Yang Y. B., Hammer F., Puech M., Rodrigues M., Liang Y. C., Deng L. C., 2015, A&A, 579, A57
Simmons B. D. et al., 2017, MNRAS, 464, 4420
Skrutskie M. F. et al., 2006, AJ, 131, 1163
Smethurst R. J. et al., 2015, MNRAS, 450, 435
Sodré L., Ribeiro da Silva A., Santos W. A., 2013, MNRAS, 434, 2503
Spiekermann G., 1992, AJ, 103, 2102
Storrie-Lombardi M. C., Lahav O., Sodre L., Storrie-Lombardi L. J., 1992, MNRAS, 259, 8
Tan M., Le Q., 2019, International Conference on Machine Learning, Association for Computing Machinery, New York, p. 6105
Tan M., Chen B., Pang R., Vasudevan V., Sandler M., Howard A., Le Q. V., 2019, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, Los Alamitos, CA. p. 2820
Tapia T., Eliche-Moral M. C., Aceves H., Rodríguez-Pérez C., Borlaff A., Querejeta M., 2017, A&A, 604, A105
Thomas T., Katgert P., 2006, A&A, 446, 31
Tohill C.-B., Bamford S., Conselice C., 2023, preprint (arXiv:2302.11482)
Tous J. L., Solanes J. M., Perea J. D., 2020, MNRAS, 495, 4135
Tyson J. A., 2002, in Tyson J. A., Wolff S.eds, Proc. SPIE Conf. Ser. Vol. 4836, Survey and Other Telescope Technologies and Discoveries. SPIE, Bellingham, p. 10
van den Bergh S., 1990, ApJ, 348, 57
van den Bergh S., 1998, Galaxy Morphology and Classification, 1st ed. Cambridge Univ. Press, Cambridge
Vaucouleurs G., 1959, Hand. Phys., 11, 275
Vecchiotti P., Vesperini F., Principi E., Squartini S., Piazza F., 2018, Multidisciplinary Approaches to Neural Computing, Springer, New York, p.161
Vega-Ferrero J. et al., 2021, MNRAS, 506,1927
Vega-Ferrero J. et al., 2021, MNRAS, 506, 1927
Ventou E. et al., 2017, A&A, 608, A9
Vulcani B., Poggianti B. M., Fritz J., Fasano G., Moretti A., Calvi R., Paccagnella A., 2015, ApJ, 798, 52
Walmsley M. et al., 2020, MNRAS, 491, 1554
Wang B., Cappellari M., Peng Y., Graham M., 2020, MNRAS, 495, 1958
Wechsler R. H., Tinker J. L., 2018, ARA&A, 56, 435
Werner S. V., Hatch N. A., Muzzin A., van der Burg R. F. J., Balogh M. L., Rudnick G., Wilson G., 2022, MNRAS, 510, 674

Willett K. W. et al., 2013, MNRAS, 435, 2835
Wjeisinghe D. B., Hopkins A. M., Kelly B. C., Welikala N., Connolly A. J., 2010, MNRAS, 404, 2077
Wong O. I. et al., 2012, MNRAS, 420, 1684
Wu J. F., 2020, ApJ, 900, 142
York D. G. et al., 2000, AJ, 120, 1579
Zaborowski E. et al., 2023, ApJ, 954, 68
Zibetti S., Ménard B., Nestor D. B., Quider A. M., Rao S. M., Turnshek D. A., 2007, ApJ, 658, 161
Zwicky F., 1940, Phys. Rev., 58, 478

SUPPORTING INFORMATION

Supplementary data are available at *MNRAS* online.

suppl_data

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

This paper has been typeset from a TEX/LATEX file prepared by the author.