



Publication Year	2024
Acceptance in OA	2025-02-05T12:54:28Z
Title	Simultaneous derivation of galaxy physical properties with multimodal deep learning
Authors	GAI, Mario, Bove, Mario, Bonetta, Giovanni, Zago, Davide, Cancelliere, Rossella
Publisher's version (DOI)	10.1093/mnras/stae1555
Handle	http://hdl.handle.net/20.500.12386/35810
Journal	MONTHLY NOTICES OF THE ROYAL ASTRONOMICAL SOCIETY
Volume	532

Simultaneous derivation of galaxy physical properties with multimodal deep learning

Mario Gai¹,¹★ Mario Bove,² Giovanni Bonetta,² Davide Zago² and Rossella Cancelliere²

¹*Istituto Nazionale di Astrofisica, Osservatorio Astrofisico di Torino, V. Osservatorio 20, I-10025 Pino Torinese (TO), Italy*

²*Department of Computer Sciences, University of Turin, Corso Svizzera 185, I-10149 Torino, Italy*

Accepted 2024 June 19. Received 2024 June 19; in original form 2023 June 27

ABSTRACT

Upcoming telescopes and surveys will revolutionize our understanding of the Universe by providing unprecedented amounts of observations on extragalactic objects, which will require new tools complementing traditional astronomy methods, in particular machine learning techniques, and above all, deep architectures. In this study, we apply deep learning methods to estimate three essential parameters of galaxy evolution, i.e. redshift, stellar mass, and star formation rate (SFR), from a data set recently analysed and tailored to the Euclid context, containing simulated *H*-band images and tabulated photometric values. Our approach involved the development of a novel architecture called the FusionNetwork, combining two components suited to the heterogeneous data, ResNet50 for images, and a Multilayer Perceptron (MLP) for tabular data, through an additional MLP providing the overall output. The key achievement of our deep learning approach is the simultaneous estimation of the three quantities, previously estimated separately. Our model outperforms state-of-the-art methods: overall, our best FusionNetwork improves the fraction of correct SFR estimates from ~ 70 to ~ 80 per cent, while providing comparable results on redshift and stellar mass.

Key words: galaxies: evolution – galaxies: general – galaxies: photometry – galaxies: star formation.

1 INTRODUCTION

The issue of galaxy evolution and its underlying physical processes is at the forefront of modern astrophysical research, and it is supported by a rapidly increasing wealth of multiwavelength data sets (images and spectra) provided by surveys and targeted observations. In particular, the recently launched *Euclid* mission (www.euclid-ec.org; Laureijs et al. 2012; Euclid Collaboration: Mellier et al. 2024) will provide an unprecedented uniform data set of visible and near IR images for about 1.5 billion objects in its Wide Survey (plus NIR spectra for >35 million sources), covering $\sim 1/3$ of the sky ($\sim 14\,000\text{ deg}^2$), which is expected to support crucial developments in the field.

Galaxies are supposed to evolve through cosmic time mainly thanks to star formation and mergers, e.g. Conselice et al. (2014); Madau & Dickinson (2014), evidencing a morphological change from irregular (and/or peculiar) galaxies at high redshift to more regular configurations at later epochs, e.g. Mortlock et al. (2013). Our understanding of the actual details of such general trend is expected to improve significantly in the next few years, thanks to new instruments and surveys.

The evolution of galaxies is often described through a few relevant physical properties, in particular the mass of their visible components (hereafter, stellar mass), and star formation rate (SFR) as a function of their redshift (i.e. distance/age). On one side, for fainter and fainter objects, it is difficult to get measurements sufficiently detailed to

achieve a significant individual description. On the other side, the scientific interest is anyway focused more on the statistical analysis of large sets of galaxies, aimed at discriminating general trends, rather than on individual variation, which may be dominated by peculiar events in the object history, or by their environment (Estrada et al. 2023).

Standard techniques (Bisigello et al. 2016, 2017; Ciesla, Elbaz & Fensch 2017; Iyer et al. 2018) to estimate galaxy physical properties matching templates or models with galaxy spectral energy distribution (SED) become more and more cumbersome with the increasing size of the samples made available by modern large surveys. In particular, SED fitting methods require huge processing power and extensive support of human experts in order to discriminate critical cases. Adoption of advanced tools for automated analysis, e.g. machine learning (ML) and/or data mining tools, is therefore a compelling need, as evidenced by the increasing interest in the scientific community, and its ongoing effort in building the specific expertise.

In this context, our work is focused on investigating possible improvements to the solution to a specific problem, i.e. estimation of redshift, stellar mass, and SFR on simulated *H*-band images and photometric magnitudes referred to the *Euclid* mission, by means of deep learning techniques. The science goals and overall framework are set by a recent publication (Euclid Collaboration: Bisigello et al. 2023), to which the interested reader is referred for further astrophysical insight. We work on the same data, by leave of that paper's Authors.

The photometric information is complemented with the morphological information associated to galaxy images. This is a

* E-mail: mario.gai@inaf.it

combination of differently structured data, since the image pixels are related to each other by the shape of the observed object (and the instrument response), whereas magnitudes are in principle independent scalar values. As a consequence, the tool extracting and joining the information from the two sides is hereafter labelled as *multimodal*.

In this context, the key aspects of our approach are:

(i) The simultaneous estimation of redshift, stellar mass, and SFR by a single network, rather than implementing three dedicated models. The rationale is that, for some classes of galaxies, we can expect correlations among their age, mass, and star formation history, dictated by their actual evolution. The deep learning machine may thus ‘grasp’ some of the underlying relationship among parameters, exploiting it to improve on their estimation from available data.

(ii) The new results we obtain for the three aforementioned quantities, in particular the improved values for SFR. Our model, a FusionNetwork made of a ResNet (He et al. 2016) combined with a Multilayer Perceptron (Rumelhart, Hinton & Williams 1986; Popescu et al. 2009), processes galaxy images and quantitative data at the same time, exploiting data correlations and leading to better performance.

The focus on collective diagnostics represents, in our opinion, the most significant difference with respect to Euclid Collaboration: Bisigello et al. (2023), demonstrating that significant margins towards better parameter estimation are still available. Besides, the merging of heterogeneous information (images and photometry) is performed in a more flexible way, explained in Section 3.3, allowing further tuning of the internal representation of the photometric inputs. The issue of collective, rather than independent, parameter estimation has been dealt with explicitly also by some of the authors in a simpler framework (Gai, Busonero & Cancelliere 2017).

Several recent works deal with various combinations of photometry and imaging, for different science goals and using a range of machine-learning approaches. Estimation of redshift, stellar mass, and SFR is also addressed in Humphrey et al. (2023), based on photometric data only, using semisupervised techniques, and focusing on training trends. The connections between galaxy stellar mass, SFR, and dark matter halo mass is investigated by Hausen et al. (2023), using a machine-learning method called Explainable Boosting Machines working on a range of simulated galaxy physical characteristics. Cabayol et al. (2021) propose LUMOS, a deep learning approach to measure photometry from galaxy images, for subsequent estimate of photometric redshift. Henghes et al. (2022) deduce photometric redshift from galaxy images in different filters using deep learning methods, based on convolutional neural networks (CNNs), on a $z \leq 1$ sample. A somewhat similar approach is handled in Syarifudin, Hakim & Arifyanto (2019). Li et al. (2022), and Treyer et al. (2024) exploit neural techniques using both photometric and galaxy images information in order to estimate only photometric redshift. Unsupervised ML by variational auto-encoders is proposed for galaxy morphology classification from *Hubble Space Telescope* (HST) and *JWST* images by Tohill et al. (2024). Besides, classification of galaxies, quasars, emission-line galaxies, and stars from visible and infrared photometric data using various ML methods is proposed by Zeraatgari et al. (2024). Simulation of Euclid images and photometry are being actively developed (Euclid Collaboration: Bretonnière et al. 2023; Euclid Collaboration: Merlin et al. 2023) in support to the forthcoming data reduction process, and possibly to the synergy with other modern projects (Liu et al. 2023).

In Section 2, we resume the main characteristics of the data set, of the pre-processing techniques, and of the evaluation criteria and

metrics used. In Section 3, the exploited neural architectures and our deep learning model are described. Section 4 is devoted to comparison of our results with those in the literature and a discussion of some of the implications. Finally, in Section 5, we draw our conclusions, with some hints on possible future developments.

2 DATA SET AND METRICS

The data set used in our work was derived from the COSMOS2015 multiwavelength public catalogue (Laigle et al. 2016), referring to the Cosmos Evolution Survey (COSMOS; Scoville et al. 2007) field, complemented by mock *Euclid* *H*-band images derived from the COSMOS-Drift And SHift (COSMOS-DASH; Mowla et al. 2019) survey with the *HST* Wide Field Camera 3 (WFC3).

The data are described in detail in Bisigello et al. (2020); Euclid Collaboration: Bisigello et al. (2023), in terms of data set structure and characteristics, and are used without modifications in our work. Hereafter, and in Section 2.1, we summarize the main data features for the reader’s convenience.

The photometric information (source magnitudes in different bands) is collected in tabular form; stars and X-ray sources (<1 per cent of the galaxy sample) have been removed from the original COSMOS2015 catalogue. The custom catalogue is inspired to the Euclid Wide Survey (Euclid Collaboration: Scaramella et al. 2022), including the four Euclid filters, i.e. I_E , Y_E , J_E , H_E , and extended with the addition of the u band from the Canada–France Imaging Survey (CFIS), and the Sloan Digital Sky Survey (SDSS; Gunn et al. 1998) magnitudes g , r , i , and z .

Observations comparable to those from the Euclid H_E band are available from the *HST*-WFC3 Imaging Survey in the COSMOS Field (COSMOS-DASH; Mowla et al. 2019), upon suitable transformation to *Euclid* resolution (matching its expected PSF rms size) and application of realistic noise. This also includes appropriate scaling of the source and background flux to achieve the desired S/N distribution.

In particular, the starting point are HST/F160W thumbnails of 51×51 pixels, centred on each galaxy, and the simulated H_E band images have 25×25 pixels. We remark that, for most of the expected sources, Euclid’s resolution is comparably poor, so that the available morphological information is limited.

By matching the COSMOS-DASH and COSMOS2015 catalogues, the simulated H_E images are linked to the set of mock Euclid magnitudes. The analysis is restricted to images with $S/N > 3$, further split by separation of the subset corresponding to $S/N > 10$, to identify respectively low and high flux regimes. In particular, the subsets include respectively 27 340 ($3 < S/N < 10$) and 9799 ($S/N > 10$) COSMOS-DASH galaxies, each associated with redshift, stellar mass, and SFR values.

The data set is built from highly imbalanced original data, due to the natural shortage of sufficiently bright galaxies ($S/N > 3(10)$) with low stellar mass ($\log_{10}(M_*/M_\odot) < 8$ (8.5)) and low SFR ($\log_{10}[\text{SFR}/(M_\odot \text{ yr}^{-1})] < -2.5$ (−3)), which are therefore under-represented. This is liable to turn out as one of the major issues with supervised ML tasks (Johnson & Khoshgoftaar 2019; Cheng et al. 2020), because learners will make decisions biased towards the most present values and, in extreme cases, they may completely ignore the less frequent values.

In order to mitigate this issue, the data from Euclid Collaboration: Bisigello et al. (2023) have been pre-processed by that paper’s authors, applying sample/image augmentation techniques on training and validation instances. In particular, magnitudes are perturbed according to S/N level (new instances having the nominal mean value

and random variations with standard deviation corresponding to the expected noise), and images are rotated by 10° steps, up to 35 times, to generate new data instances consistent with the starting catalogue information.

The data set built according to such prescriptions is considered to be the ground truth associated with the previously mentioned physical properties, within the limits associated with the astrophysical spread among source parameters, the noise of related observations, and estimate errors (also due to the underlying models). Future applications of our analysis and of the proposed approach will therefore benefit also from improvements to the data set itself.

For simplicity, we analyse only the data set using all nine photometric bands, i.e. the four Euclid plus the five ground-based ones, which is the case providing the best results in Euclid Collaboration: Bisigello et al. (2023).

2.1 Data pre-processing

For application of our neural approaches, the augmented data set from Euclid Collaboration: Bisigello et al. (2023) is split into training, validation, and test sets, respectively sized to 90, 5, and 5 per cent of the total, so that 24 605 instances are used for training, 1367 for validation, and 1367 for test. A necessary step is aimed at matching the original data format to an adequate internal representation by performing data normalization and removal of anomalous data, described below.

The normalization process helps in stabilizing training and avoiding biases on deep learning algorithms. Besides, non-normalized targets in regression problems can induce exploding gradients, resulting in the failure of the learning process (Sola & Sevilla 1997).

Because input magnitude values are spread over significantly different ranges, each band is rescaled to internal values in the $[0, 1]$ interval. Similarly, image pixels and targets, i.e. redshift, stellar mass, and SFR, are scaled to the same internal range. The transformation applied to each parameter x is described by the normalization equation (1):

$$y_i = \frac{x_i - \min(x)}{\max(x) - \min(x)}, \quad (1)$$

where x_i is the i^{th} component of the $\mathbf{x}$ vector made by all the parameter observations in the data set.

Anomalous data, i.e. data that, for any reason, has a numerical value far beyond the range of the others, can be expected to significantly degrade the performance of most machine learning tools (Khamis et al. 2001). Such values are often employed to identify lack of measurements rather than actual weird values of the physical quantity. In our case, just one data instance with value -99 has been suppressed in the SFR distribution, since we assume it was labelled by the providers as missing or bad data.

It is useful to separately deal also with a data set made only by instances describing galaxies for which low stellar masses and low SFR do not cause a shortage of their number. As done in Euclid Collaboration: Bisigello et al. (2023), galaxies with very low mass ($\log_{10}(M_*/M_\odot) < 8$) and very low SFR ($\log_{10} \text{SFR}/(M_\odot \text{ yr}^{-1}) < 0$) are removed from the full data set. We call this additional data set *balanced*.

2.2 Evaluation metrics

Several metrics are used to evaluate the performance of our implemented models, also following the guidelines used in the literature (Euclid Collaboration: Bisigello et al. 2023) and references therein.

The mean square error (MSE) is used as training loss function and to evaluate the overall model performance. Besides, in order to compare our results with the literature, the fraction of outliers (f_{out}), the bias (Δz), and the normalized median absolute deviation (NMAD) are calculated consistently with Euclid Collaboration: Bisigello et al. (2023).

The fraction of outliers f_{out} corresponds to over- or underestimated data relative to the size of the data set. The bias and NMAD are used to provide an indication of the statistical distribution of the estimates. In particular, it is expected that this distribution is approximated by a Gaussian, so that Δz indicates the redshift mean discrepancy (ideally zero), while the NMAD is related to its standard deviation. The calculation of these quantities differs depending on the target. In the following, the subscript ‘in’ refers to the target values, and the subscript ‘out’ refers to the estimated values (i.e. rescaled network outputs).

Redshift (z). Higher redshift values are expected to be more difficult to predict correctly. Thus a prediction is called an outlier if

$$|z_{\text{out}} - z_{\text{in}}| > 0.15(1 + z_{\text{in}}) \quad (2)$$

equations (3) and (4) define respectively the bias (Δz) and NMAD:

$$\Delta z = \text{median} \left[\frac{z_{\text{out}} - z_{\text{in}}}{1 + z_{\text{in}}} \right] \quad (3)$$

$$\text{NMAD} = 1.48 \text{ median} \left[\frac{|z_{\text{out}} - z_{\text{in}}|}{1 + z_{\text{in}}} \right]. \quad (4)$$

Stellar mass (M_*). As in Euclid Collaboration: Bisigello et al. (2023), we estimate for the entire sample the fraction of outliers, defined as galaxies for which the stellar mass is overestimated or underestimated by a factor two (~ 0.3 dex), so that a mass prediction is considered an outlier if:

$$\left| \log_{10} \left(\frac{M_{*,\text{out}}}{M_{*,\text{in}}} \right) \right| > 0.3 \quad (5)$$

Bias (ΔM_*) and NMAD are defined by equations (6) and (7):

$$\Delta M_* = \text{median} \left[\log_{10} \left(\frac{M_{*,\text{out}}}{M_{*,\text{in}}} \right) \right] \quad (6)$$

$$\text{NMAD} = 1.48 \text{ median} \left[\left| \log_{10} \left(\frac{M_{*,\text{out}}}{M_{*,\text{in}}} \right) \right| \right] \quad (7)$$

SFR. Similarly to the stellar mass case, outliers are defined as galaxies with SFR incorrect by, at least, a factor two (~ 0.3 dex):

$$\left| \log_{10} \left(\frac{\text{SFR}_{\text{out}}}{\text{SFR}_{\text{in}}} \right) \right| > 0.3 \quad (8)$$

Bias (ΔSFR) and NMAD are defined by equations (9) and (10):

$$\Delta \text{SFR} = \text{median} \left[\log_{10} \left(\frac{\text{SFR}_{\text{out}}}{\text{SFR}_{\text{in}}} \right) \right] \quad (9)$$

$$\text{NMAD} = 1.48 \text{ median} \left[\left| \log_{10} \left(\frac{\text{SFR}_{\text{out}}}{\text{SFR}_{\text{in}}} \right) \right| \right]. \quad (10)$$

3 DEEP LEARNING ARCHITECTURES

This section describes the details of the neural architectures used: the multilayer perceptron (MLP), the ResNet50 that is a particular CNN, and the fusion network, our model which combines them both.

3.1 Multilayer perceptron

MLPs (Rumelhart et al. 1986; Popescu et al. 2009) are the best known and widely used neural networks. An MLP is a type of

Table 1. Architecture of the MLP processing photometric data. The first layer has nine inputs, i.e. the magnitudes, while the last one has three output units.

Nr	Layer	N_{input}	N_{output}
1	fully connected	9	2000
2	fully connected	2000	1000
3	fully connected	1000	500
4	fully connected	500	3

artificial neural network that is composed of multiple layers of neurons interconnected to each other.

The structure of an MLP can be divided into three main parts: input layer, which contains one neuron for each sample feature, so the amount of neurons is equal to the dimensionality of the samples in the data set; one or more hidden layers, each consisting of a certain number of interconnected neurons where each neuron processes the outputs of the previous layer's neurons; output layer, which produces the final result, which is usually the solution of a regression or classification task.

To solve non-linear problems the MLP uses non-linear activation functions: a common one is the ReLU (rectified linear unit) which returns the input value if it is greater than 0, otherwise 0.

Learning in the MLP occurs by changing the connection weights after each input data is processed through backpropagation, based on the error in the output with respect to the expected result.

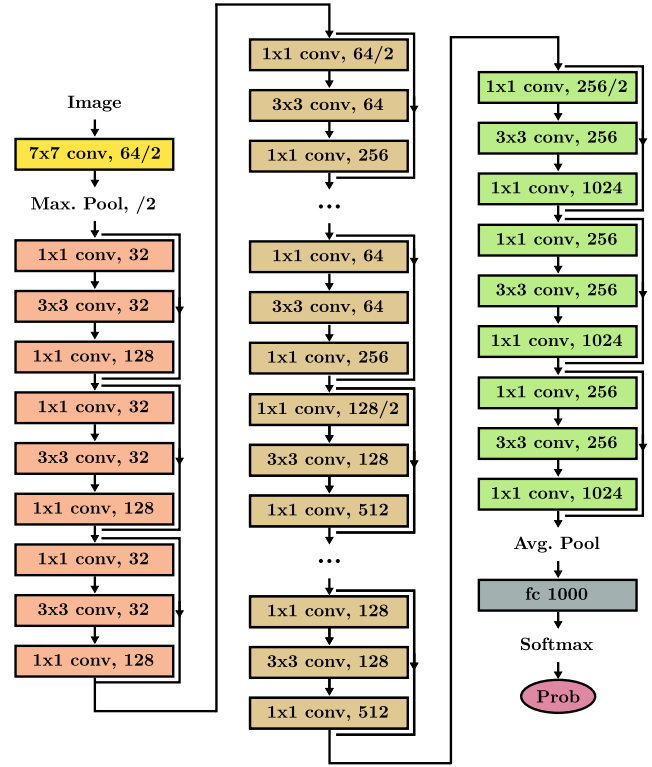
The architecture used in our tests is summarized in Table 1.

3.2 ResNet

CNNs (LeCun et al. 1989; O'Shea & Nash 2015) are neural networks that aim to efficiently process data that can be represented with a grid topology, such as images, and are commonly used for computer vision applications. CNNs are characterized by the *local connectivity* property and the *shared weight* property. The former means that neurons are not connected globally, but just to a few of the neurons from the previous layer, and the latter forces neurons to share their weights, processing the input the same way. The main types of layers are: the convolutional layer, whose main task is to extract features that best describe the input image; the pooling layer, which aims to gradually reduce the representation dimensionality, further reducing the number of parameters and the computational complexity of the model; finally, the fully connected layer (FC) is used to output the classification/regression values. As with the MLPs, activation functions are also used.

To get higher performance and accuracy it is common to make networks deeper by stacking layers. The rationale for adding multiple layers is the expectation of an increased capability to learn more complex patterns.

ResNet (Residual Networks; He et al. 2016) is a deep learning architecture that solves the problem of training deep neural networks through skip connections. These connections efficiently propagate gradients, enabling effective training even in very deep models. ResNet's innovation lies in the residual block, which learns residual functions and mitigates the issue of vanishing gradients, making it a highly influential neural network model in computer vision and other domains. A standard ResNet50 architecture is shown in Fig. 1.

**Figure 1.** The ResNet50 model. Skip connections and residual blocks are evidenced.

3.3 FusionNetwork

The heterogeneous structure of our data set suggests us an alternative model, the *FusionNetwork*, which exploits fusion, i.e. the process of joining data from multiple modalities, with the aim of extracting complementary and more complete information. Our proposed model combines an MLP, called hereafter I-MLP (most suited for tabulated photometric data), and a ResNet50 network (most suited for images). This fusion is obtained connecting the two parts through a final MLP, called F-MLP, and is represented in Fig. 2.

The two partial networks are expected to be able to ‘grasp’ some of the problem complexity and supply suitable representations to inputs, whereas the combining stage, with additional fine-tuning, may encode further correlations within the data not easily mapped by each of them separately. As a good starting point, each partial network is pre-trained on the data set. The architecture of the merging F-MLP is summarized in Table 2.

For the sake of comparison, we also run our own version of the CNN architecture from Euclid Collaboration: Bisigello et al. (2023); we called it CNN_s because the simultaneous estimation of redshift, stellar mass, and SFR is maintained. The CNN_s is obtained from our model using a classical CNN instead of the ResNet50, and directly feeding the photometric data into the network, without the processing performed by the three fully connected layers of our I-MLP (see Fig. 2).

4 RESULTS AND DISCUSSION

We experimented with two models: a standalone MLP and the FusionNetwork. Both these systems estimate redshift, stellar mass, and SFR at the same time, using a final layer with three outputs (simultaneous regression). The training loss is the average MSE of

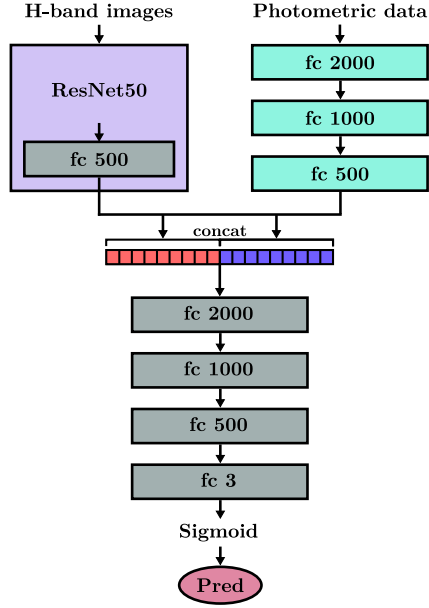


Figure 2. The FusionNetwork achieved by combination of an MLP (top right) and a Resnet50 architecture (top left).

Table 2. Architecture of the F-MLP used in the FusionNetwork to concatenate ResNet50 and the I-MLP.

Nr	Layer	N_{input}	N_{output}
1	fully connected	1000	2000
2	fully connected	2000	1000
3	fully connected	1000	500
4	fully connected	500	3

the three targets. The experiments are performed on a Linux-5.13.0 computer equipped with an Intel Core i7 8th Gen processor, 8 CPUs, and a NVIDIA TITAN RTX GPU (24gb). The software is written in PYTHON 3.9.12 with PYTORCH.

In all experiments, a batch of 612 data instances (with both images and numerical data) is used. At the end of each epoch, the model is tested on the validation set. This allows us to save the model obtaining the lowest loss function value on the validation set in order to obtain better generalization. For each experiment, the stop criterion is based on the lack of improvement of the loss function computed on the validation set, with respect to the best value, for 15 consecutive epochs. Moreover, a maximum limit of 1000 epochs is set. The tests are performed on the whole data set with $S/N > 3$.

4.1 MLP settings

The MLP architecture details are described in Table 1. We remind that the MLP only processes the nine numerical magnitude data used as input.

This system is trained using five different seeds. In this experiment, we consider the best run, i.e. the run achieving the lowest value of the validation loss function, and the average of the five runs, which provides an indication of the spread that might be expected from the results. The loss function on the training and validation, set for the best run, is shown in Fig. 3 (top). Training reaches convergence at epoch 673, and terminates at epoch 688 thanks to the stop criterion. The time required to train the network is 9 min and 19 s.

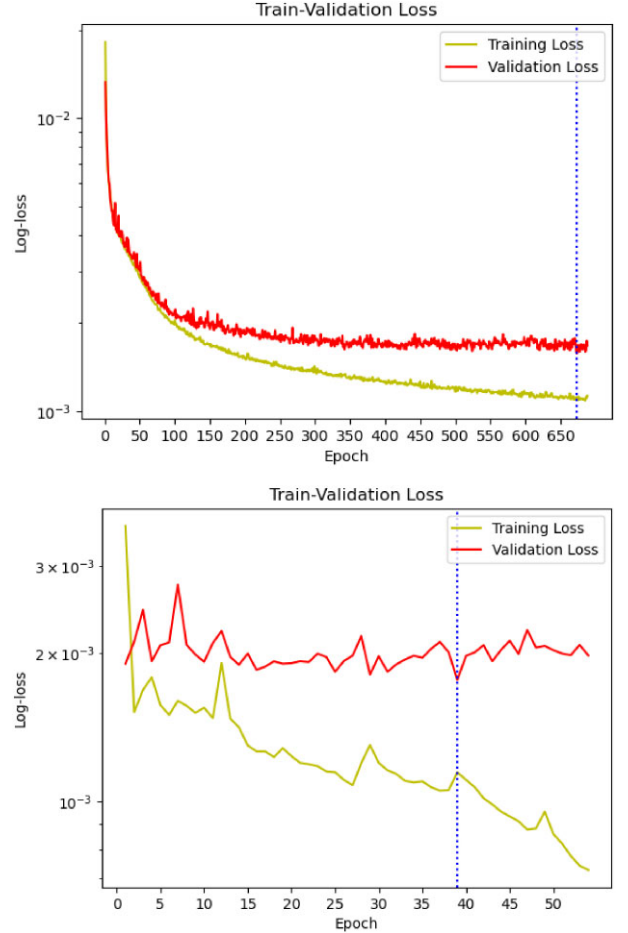


Figure 3. Trend of the loss function (in log units) over the training and validation set for simultaneous estimation of redshift, stellar mass, and SFR with MLP (top) and FusionNetwork (bottom). The vertical dotted line shows the lowest value achieved for the validation loss function.

4.2 FusionNetwork settings

A pre-training is used to initialize the weights of the two input components of the FusionNetwork (Section 3.3), i.e. I-MLP (Section 3.1) and ResNet50 (Section 3.2). The whole system is then trained using five different seeds for statistical relevance. The plot in Fig. 3 (bottom) shows the performance of the best model on the training and validation sets.

The model converges at the 39th epoch; the time required to train the best model is slightly above one hour.

Validation losses of the fusion model and MLP are comparable so that performances of both models will be compared with state-of-the-art approaches in the next section.

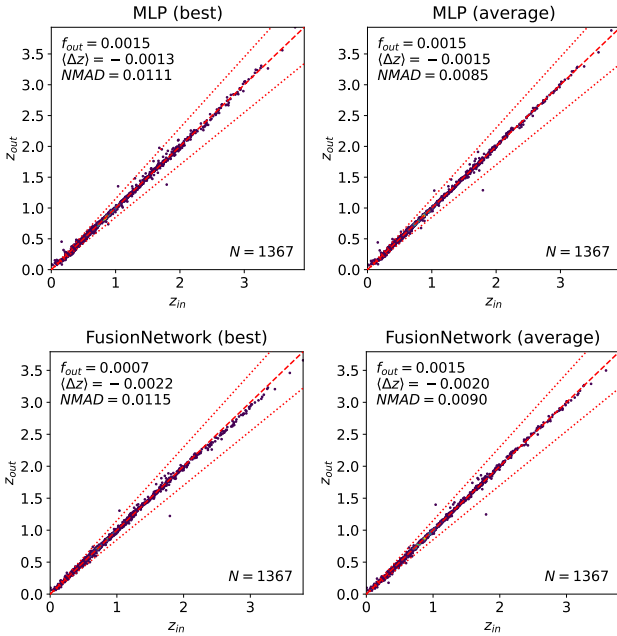
4.3 Result comparison

This section compares the results from different models in terms of outlier fraction, bias, and NMAD as described in Section 2.2.

We report the outputs from three different reference cases presented in Euclid Collaboration: Bisigello et al. (2023): the SED fitting model and their MLP and CNN implementations, respectively, called by the Authors DLNN and CNN. We also report the results from our implementation of the latter, labelled CNN_S: we anticipate here that it performs similarly to the CNN architecture.

Table 3. Results for redshift estimation. In bold, best values and closest competitors. SED, DLNN, and CNN are from Euclid Collaboration: Bisigello et al. (2023).

Algorithm	Model	f_{out}	Δz	NMAD
SED		0.127	-0.002	0.045
DLNN	best	0.001	-0.002	0.008
	average	0.002	-0.001	0.011
CNN	best	0.002	0.005	0.028
	average	0.003	-0.001	0.021
CNN _S	best	0.0037	-0.0043	0.0296
	average	0.0037	0.0011	0.0193
MLP	best	0.0015	-0.0013	0.0111
	average	0.0015	-0.0015	0.0085
FusionNetwork	best	0.0007	-0.0022	0.0115
	average	0.0015	-0.0020	0.0090

**Figure 4.** Comparison of the target value and estimated value of redshift. The dashed red line corresponds to perfect estimate, while the dotted red lines evidence the outlier threshold from equation (2). Top left of each panel: relevant values of f_{out} , bias, and NMAD. Bottom right: number of test set instances.

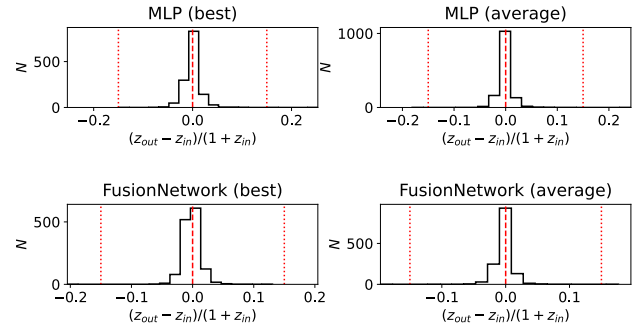
We recall that the MLP takes in input only tabular data (magnitudes), while the CNN is a multimodal system that also exploits images. These two neural models are trained for individual estimation of the physical properties, differently from our approach characterized by simultaneous estimation of the three parameters. In subsections 4.3.1, 4.3.2, and 4.3.3 we describe our findings for the three physical properties in details.

4.3.1 Redshift estimation

The main features of our redshift estimation are reported in Table 3 in statistical terms, and illustrated in Figs 4 and 5.

The DLNN and CNN already achieved significantly better results than the reference SED fitting method; in particular, the f_{out} is two orders of magnitude smaller.

Our MLP, working on photometric information only, is very robust in estimating the redshift, with f_{out} and bias in line with both DLNN

**Figure 5.** Error distribution on redshift estimation (difference between z_{in} and z_{out}). The red vertical dashed line (mean value) is consistent with a null difference and the red dotted lines correspond to the outlier threshold values (± 0.15) from equation (2).

and CNN, and NMAD comparable to DLNN and about two times better than CNN. Our CNN_S provides results quite similar to CNN, so that the same considerations hold.

Regarding the FusionNetwork, the best model achieves lower f_{out} than all others, but worse values of bias and NMAD w.r.t. MLP and DLNN models. The average results are comparable or better for f_{out} but worse for bias and comparable for NMAD.

From the comparison, it appears that the most robust model for redshift estimation only is the MLP. This might be justified by the noise introduced by the images, exceeding the information provided, and increasing the dispersion of the results, thus resulting in higher NMAD and bias values. On the other hand, the FusionNetwork may still be preferred in terms of pure outlier performance.

In Fig. 4, the target redshift is plotted against its estimates provided by the MLP (top) and the FusionNetwork (bottom), respectively showing the best (left) and average (right) models. At the top left of each panel we report the relevant values of f_{out} , bias and NMAD. At the bottom right, we report the number of instances in the test set.

Our models generate predictions close to the diagonal (desired 1:1 match) for the entire input range (i.e. $z \leq 3.5$). The spread of outliers is narrower at low redshift, increasing with z .

The plots in Fig. 5 show the distribution of redshift error (discrepancy between target and predicted value), respectively, for the MLP (top) and the FusionNetwork (bottom), in their best (left) and average (right) installations. The mean value (red dashed line) is very close to zero, with most instances located within the region delimited by the red dotted lines, associated with the minimum threshold for outliers from equation (2). Moreover, the distribution evidencing the lowest spread and offset corresponds to the average model of the FusionNetwork, demonstrating the effectiveness of a simple ensemble average approach.

4.3.2 Stellar mass estimation

The main features of the stellar mass estimate are reported in Tables 4 and 5 in statistical terms, and illustrated in Figs 6 and 7. As above, CNN and CNN_S behave quite similarly, and support the same considerations.

First of all, we evaluate the performance achieved over the full data set, summarized in Table 4.

The SED model provides the worst results among the tested cases. Our MLP shows, as in the redshift case, a clear improvement over the DLNN and CNN models on outliers (f_{out}). Slightly worse results are obtained on the bias, since the error distribution is slightly shifted

Table 4. Results for galaxy stellar mass estimate (full data set). In bold: best values and closest competitors. SED, DLNN, and CNN are from Euclid Collaboration: Bisigello et al. (2023).

Algorithm	Model	f_{out}	ΔM_*	NMAD
SED		0.135	0.002	0.121
DLNN	best	0.008	−0.017	0.054
	average	0.009	0.002	0.040
CNN	best	0.011	0.006	0.050
	average	0.010	0.001	0.044
CNN _s	best	0.0102	−0.0082	0.0654
	average	0.0110	0.0039	0.0454
MLP	best	0.0059	0.0055	0.0477
	average	0.0059	0.0076	0.0348
FusionNetwork	best	0.0044	−0.0096	0.0457
	average	0.0051	−0.0015	0.0358

Table 5. Results for galaxy stellar mass estimation (balanced data set). In bold: best values and closest competitors. SED, DLNN, and CNN are from Euclid Collaboration: Bisigello et al. (2023).

Algorithm	Model	f_{out}	ΔM_*	NMAD
SED		0.128	0.001	0.120
DLNN	best	0.005	−0.017	0.054
	average	0.006	0.002	0.040
CNN	best	0.006	0.006	0.050
	average	0.007	0.001	0.044
CNN _s	best	0.0075	−0.0083	0.0640
	average	0.0075	0.0039	0.0445
MLP	best	0.0045	0.0052	0.0472
	average	0.0030	0.0077	0.0343
FusionNetwork	best	0.0022	−0.0100	0.0453
	average	0.0022	−0.0020	0.0354

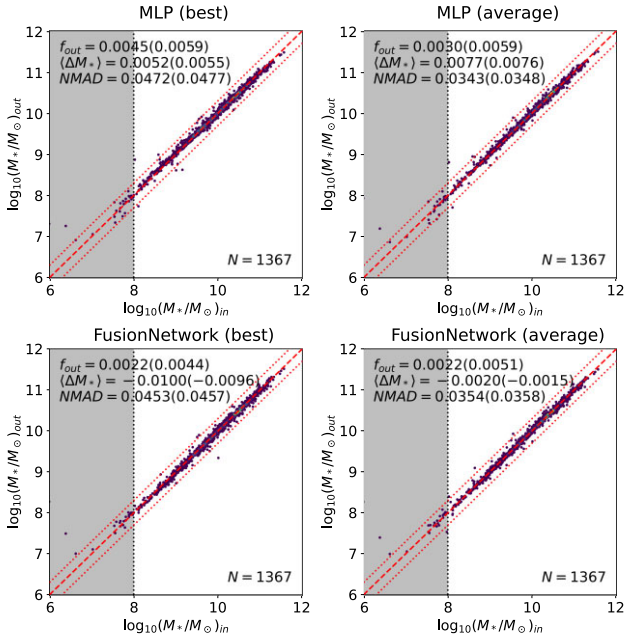


Figure 6. Comparison between the estimated stellar mass and target value. The red dashed line is the identity and the red dotted lines correspond to the outlier definition from equation (5). At the top left of each panel, we report the f_{out} , bias, and NMAD for galaxies in the whole data set, and (in brackets) for the balanced data set. On the bottom right we report the number of objects in the test set.

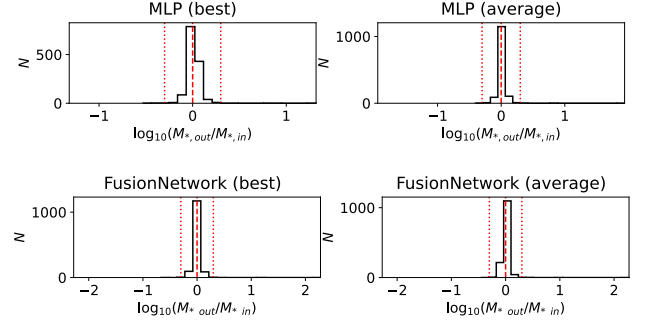


Figure 7. Error distribution on stellar mass estimation. The red vertical dashed line shows a null difference and the red dotted lines correspond to the outlier definition, equation (5).

towards positive values, while the NMAD is comparable to the DLNN and CNN.

The overall best result on f_{out} is achieved by the FusionNetwork best model, with a bias slightly shifted to negative values. Its NMAD is consistent with other methods. On the other hand, the FusionNetwork average model turns out to be quite robust in terms of error dispersion, obtaining the best bias value and quite competitive NMAD. Moreover, its f_{out} is second to the best, slightly higher than the best model, making it a very robust choice.

Furthermore, the performance assessment on stellar mass estimation is repeated over the balanced sample, deprived of underrepresented galaxy classes (see 2.1). Table 5 shows the new results achieved after filtering out these instances from the data set.

These new results show that outliers (f_{out}) are reduced by nearly 50 per cent with all models. Apparently, a large fraction of errors is in the underrepresented region of the data set.

The MLP and the FusionNetwork achieve a significantly lower level of outliers (f_{out}) than all other approaches. Moreover, the FusionNetwork average model and CNN biases are very close to zero. NMAD values of the best models are comparable, whereas relevant improvements are achieved by the average MLP and FusionNetwork models.

In Fig. 6, the target values of stellar mass (whole data set) are compared with our estimates, respectively for the MLP (top) and FusionNetwork (bottom), in their best (left) and average (right) versions. The grey area in the charts of the figure shows the underrepresented galaxies, corresponding to $\log_{10}(M_*/M_\odot) < 8$.

The points relative to the MLP and FusionNetwork models turn out to be very well distributed along the diagonal, evidencing the good performance obtained for stellar mass estimation.

Finally, the graphs in Fig. 7 show the histogram of discrepancy between the target M_* value and the estimates, in log units, for the MLP (top) and the FusionNetwork (bottom), in their best (left) and average (right) cases. Mean values (red dashed lines) are very close to zero, with most instances located within the red dotted lines, related to the outlier thresholds from equation (5).

In Fig. 8, we show the variation of the metrics with redshift (top) and I_B magnitude (bottom). The vertical lines highlight relevant values of the redshift distribution: the dashed line is the sample mean μ ; the densely dotted line evidences the range $\mu \pm \sigma$ (one standard deviation); the sparsely dotted line indicates the range $\mu \pm 2\sigma$. Fig. 8(a) shows that the FusionNetwork obtains the lowest outlier fraction (bottom panel) and NMAD (mid panel) across the majority of the redshift range. The bias (top panel) is comparable. In general, each architecture shows the lowest error at intermediate redshift values, where more samples are available. As the redshift

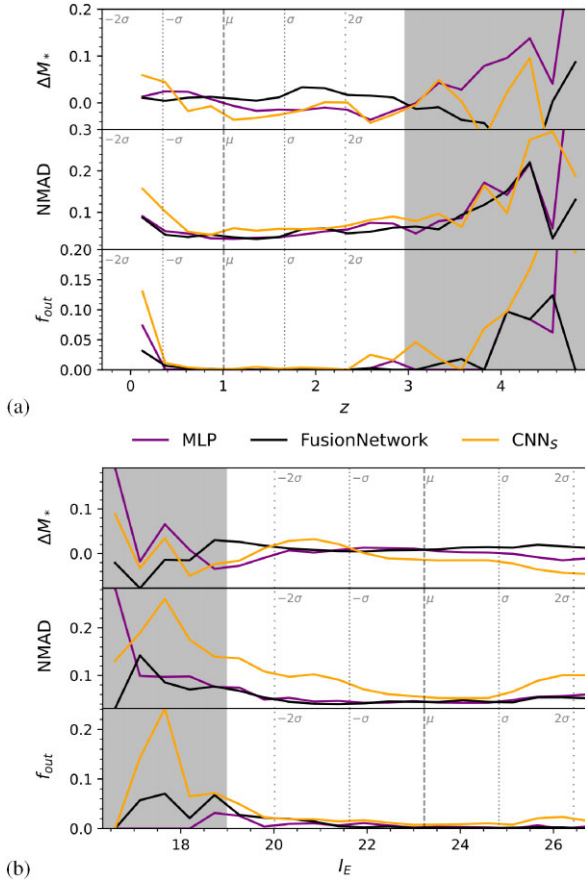


Figure 8. Variation of the bias (top), NMAD (centre), and fraction of outliers (bottom) of the recovered stellar mass with respect to the redshift (Fig. 8a) and to the I_E magnitude (Fig. 8b). The redshift (resp. I_E magnitude) range on the x-axis is divided into 20 bins, and the grey area evidences the bins with less than 2000 samples.

increases, outlier fraction and NMAD suffer a marginal rise, also due to the scarcity of samples with higher redshift. The metrics acquire less statistical significance in the redshift range > 3 (grey area), which is poorly represented in the data set. A similar trend can be seen by analysing the measures with respect to the I_E magnitude, shown in Fig. 8(b): slightly better performance of the FusionNetwork on outlier fraction (bottom panel) and NMAD (mid panel), with comparable bias (top panel), over most of the magnitude range, with large noise at the underrepresented bright end (grey area, $I_E < 20$).

4.3.3 SFR estimation

The results for the SFR estimation are reported in Tables 6 and 7, and shown in Figs 9 and 10. The considerations in the following hold for CNN_S as well as for CNN.

First of all, we evaluate the SFR performance achieved over the full data set, summarized in Table 6. Overall, the best f_{out} , ΔM_* , and NMAD are all achieved by the FusionNetwork average model. In terms of outlier performance, best and average MLP models are comparable to each other, and significantly better than DLNN and CNN. FusionNetwork best and average models are also comparable with each other, and significantly better than MLP. A similar progression can be seen on NMAD, with MLP better than DLNN and CNN, and further improvement achieved on FusionNetwork. The bias values ΔSFR are comparable, in the range of a few per cent, with

Table 6. Results for SFR estimation on the full data set. Bold: best values and closest competitors. SED, DLNN, and CNN are from Euclid Collaboration: Bisigello et al. (2023).

Algorithm	Model	f_{out}	ΔSFR	NMAD
SED		0.622	-0.065	0.637
DLNN	best	0.411	-0.021	0.350
	average	0.416	-0.045	0.359
CNN	best	0.440	0.016	0.383
	average	0.446	-0.063	0.390
CNN _S	best	0.3819	-0.0305	0.3097
	average	0.3614	0.0017	0.2949
MLP	best	0.3409	-0.0239	0.2719
	average	0.3424	-0.0093	0.2697
FusionNetwork	best	0.2970	0.0201	0.2484
	average	0.2963	0.0011	0.2394

Table 7. Results for SFR estimation in the balanced data set. Bold: best values and closest competitors. SED, DLNN, and CNN are from Euclid Collaboration: Bisigello et al. (2023).

Algorithm	Model	f_{out}	ΔSFR	NMAD
SED		0.560	-0.115	0.521
DLNN	best	0.310	-0.023	0.280
	average	0.315	-0.027	0.293
CNN	best	0.325	-0.046	0.293
	average	0.345	-0.055	0.306
CNN _S	best	0.2812	-0.0611	0.2359
	average	0.2600	-0.0282	0.2297
MLP	best	0.2500	-0.0285	0.2211
	average	0.2355	-0.0302	0.2149
FusionNetwork	best	0.2076	0.0145	0.1906
	average	0.2109	-0.0189	0.1937

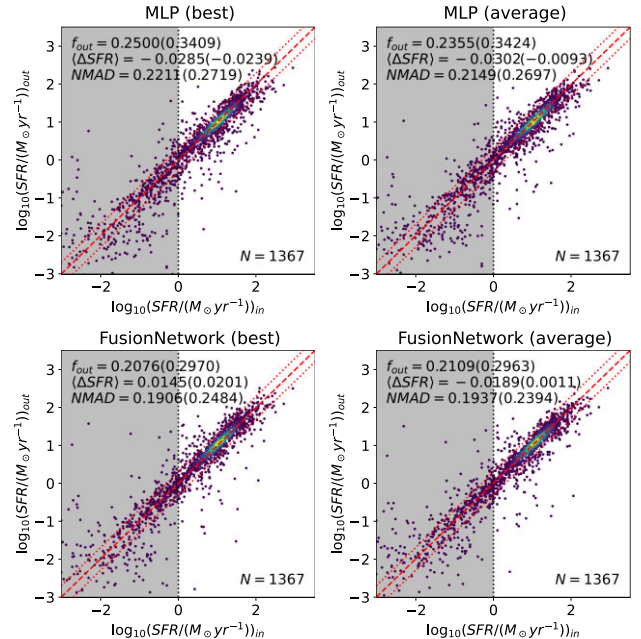


Figure 9. Comparison between the estimated SFR and the target value. The red dashed line is the identity, and the red dotted lines correspond to the outlier definition from equation (8). The grey area evidences the underrepresented galaxy classes, i.e. $SFR \leq 1 M_{\odot} \text{ yr}^{-1}$.

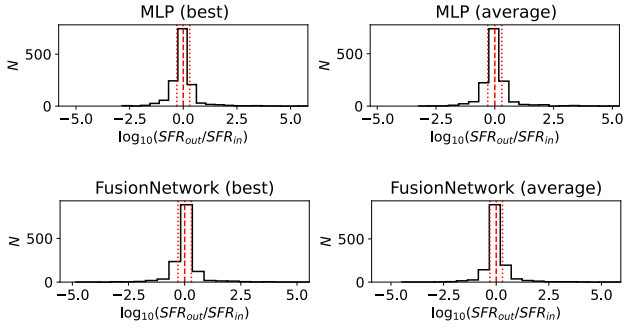


Figure 10. Error distribution on SFR estimation. The red vertical dashed line shows a bias close to zero, and the red dotted lines correspond to the outlier definition from equation (8).

the best result from the average CNN_s and FusionNetwork (<0.2 per cent).

Estimation of the SFR proves to be much more challenging than the other two parameters. In particular, for redshift and stellar mass predictions, most of the f_{out} results are in the order of 10^{-3} , while on SFR our best value of f_{out} is 0.2963 (30 per cent of outliers). None the less, our method leads to significant improvements over the SED fitting method (62 per cent), and performs significantly better than the DLNN (41 per cent) and CNN (44 per cent) models. This can be mainly explained, to our understanding, by the ability of the networks to simultaneously estimate all targets and by the adaptability of FusionNetwork.

In Table 7, we report the results achieved on the balanced data set. There is a noticeable improvement on f_{out} from both MLP (~ 24 per cent) and FusionNetwork (~ 21 per cent), with the best model of the latter being the best overall (20.76 per cent).

Bias values are comparable for most methods, within a factor two; the very small absolute value of ΔSFR on the whole data set, for the average CNN_s and FusionNetwork, may be just a happy coincidence. The NMAD on MLP (~ 22 per cent) improves significantly with respect to DLNN and CNN (~ 29 per cent), with smaller, but still appreciable, additional improvement (~ 19 per cent) on FusionNetwork.

In Fig. 9, the target values of SFR (whole data set) are compared with our estimates, respectively, for the MLP (top) and FusionNetwork (bottom), in their best (left) and average (right) versions. The grey area in the charts of the figure shows the underrepresented galaxy classes, i.e. $\log_{10}[\text{SFR}/(\text{M}_{\odot} \text{ yr}^{-1})] < 0$. It may be noted that, compared to previous plots (Figs 4 and 6), the point cloud is much scattered, with largest errors in the underrepresented area.

The graphs in Fig. 10 show the histogram of discrepancy between the target SFR value and the estimates, in log units, for the MLP (top) and the FusionNetwork (bottom), in their best (left) and average (right) cases. Mean values (red dashed lines) are still close to zero, with most instances within the red dotted lines, corresponding to the outlier thresholds from equation (8).

The variation of the recovered SFR metrics with redshift and I_E magnitude is shown in Fig. 11. SFR estimation in the low redshift range degrades as the number of samples decreases, similarly to the stellar mass prediction. Our experiments show decreasing outlier fraction and NMAD at increasing redshift, up to $z \approx 3$, where underrepresentation becomes dominant. As in stellar mass estimation, the lowest error is obtained for galaxies fainter than the bright end of the I_E band, partly because of their large number in the data set. Also, noisy ranges are reflected both in redshift and I_E magnitude distributions.

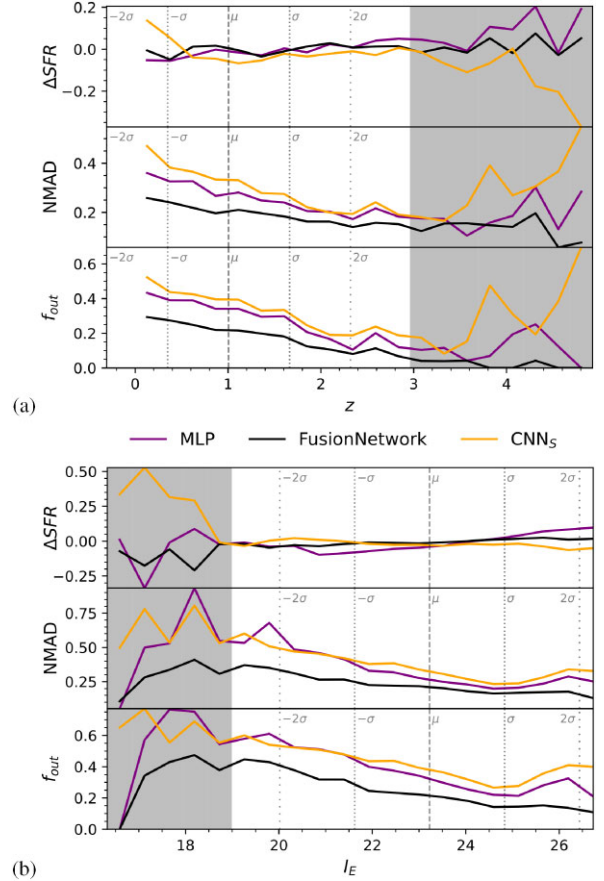


Figure 11. Variation of the bias (top), NMAD (centre), and fraction of outliers (bottom) of the recovered star formation rate with respect to the redshift (Fig. 11a) and to the I_E magnitude (Fig. 11b). The redshift (resp. I_E magnitude) range on the x-axis is divided into 20 bins, and the grey area evidences the bins with less than 2000 samples. Relevant redshift (resp. I_E magnitude) ranges are shown by the vertical lines: mean (dashed), one and two standard deviations.

4.4 Discussion

The results improvements achieved appear to be related to two different aspects. On one side, simultaneous estimation of physical properties, which may be expected to bear connections in classes of objects, gives our tools a chance to learn the ‘shape’ underlying our data distribution. Besides, images and photometric data have a different internal structure, since the former encode a bidimensional pixel-to-pixel relationship that is not present in a table of magnitudes, whose columns could, e.g., be easily swapped without any loss of information. This lead us to the choice of tools (MLP and Resnet50) dedicated to each kind of data, using an additional MLP layer to merge the intermediate results.

In general, the estimate of the SFR achieves the greatest improvement with respect to the other physical quantities. However, the SFR estimation still remains much more noisy than redshift and stellar mass. This is possibly implicit in the natural spread of SFR throughout the data set.

Also, we achieve a significant estimate improvement in data regions where there are more samples in the training set. This confirms that the networks fail to generalize correctly in underrepresented regions. Data availability depends on the distribution of objects in the Universe; however, given that more and more data may be

expected to become available with new instruments and forthcoming surveys, our capabilities to extract information, in particular by training deep learning tools, may be expected to improve as well. Future developments may therefore take into account the increasing data ensemble expected in coming years, so that, rather than freezing a specific architecture, an evolutive approach seems to be called for.

We recall that we analysed only the data set using all nine photometric bands, i.e. the four Euclid plus the five ground-based ones, which is the case providing the best results in Euclid Collaboration: Bisigello et al. (2023), neglecting the case restricted to the four Euclid bands. We may expect that more information is in general an asset, but the weight may be different depending on the associated noise level and, above all, on the astrophysical relevance of specific data subsets: e.g. the authors of that paper remarked that the shorter wavelength ground-based filters are especially useful to improve on SFR estimate, due to the higher sensitivity of the parameter to UV, rather than near-IR, radiation.

Diagnostics uncertainty suffers a natural increase in data starved regions, i.e. bright magnitude, high z , or low SFR populations. It may therefore be expected that future surveys, increasing the available training data set also in such regions, will allow for better parameter estimation, depending on the specific information content.

Comparing our performance with Euclid Collaboration: Bisigello et al. (2023) in the same conditions, corresponding to the balanced data set, the redshift estimate within a normalized error of 0.15 is approximately equivalent, i.e. achieved in 99.9 per cent of the instances; also, the stellar mass estimate within a factor two (~ 0.3 dex) is in both cases obtained for 99.5 per cent of the sample. Besides, the SFR estimate, within a factor two (~ 0.3 dex) improves from ~ 70 to ~ 80 per cent of the balanced sample.

In the former two cases, the tools used by Euclid Collaboration: Bisigello et al. (2023) already achieved a nearly perfect match between target and output, and it would be difficult to claim an effective improvement. The latter case is more challenging, and the performance gain is therefore both relevant in itself and appealing in the application perspective. In the selected context, the proposed approach of simultaneous estimation of redshift, stellar mass, and SFR appears to be fruitful, providing some improvements already in the comparison of the original CNN with our CNNs, which is a close enough tool apart from the capability of taking better advantage of parameter correlations hidden in the data.

The proposed FusionNetwork tool also appears to make a significant contribution to further improving the SFR estimate, arguably due to its higher learning flexibility associated with the separate pre-processing of images and photometric data, and in particular the exploitation of the ResNet capabilities on images. Therefore, FusionNetwork provides the best results on collective estimate of the three parameters ensemble, even if individual results are in some cases suboptimal. Besides, the MLP, providing intermediate performance with significantly reduced complexity (since it only works on photometric data) remains an interesting tool that may be conveniently retained for comparison and verification purposes.

Our future work may include further improvement on the proposed diagnostic tools as well as application to other science areas. The synergy between the Euclid Wide Survey and *Gaia* on the common ~ 40 per cent sky area is particularly of interest. Similar issues have been recently discussed by some of the authors (Gai et al. 2022), restricted mainly to the maintenance of the *Gaia* catalogue on the stellar population shared with China space station telescope (CSST). The investigation can be applied to the galaxy population at the faint end of *Gaia* and the bright end of *Euclid*, where complementary imaging and photometric measurements may be matched. *Gaia* will

provide an independent photometric determination in the visible (G band and BP/RP colours) for most sources in the common magnitude range, thus helping with respect to the general issue of better spectral definition of the *Euclid* targets. Synergy with other projects may be considered as well. Ground-based surveys will provide more accurate photometric and spectroscopic data on statistically significant subsets of galaxies, but not on the whole Euclid data set. Besides, the CSST (Liu et al. 2023) may provide a valuable additional contribution due to its photometry ranging from near UV to near IR, and a sky coverage quite comparable, although not totally superposed, to Euclid.

5 CONCLUSIONS

In this work, state-of-the-art deep learning architectures are used to estimate redshift, stellar mass, and SFR of galaxies over an heterogeneous data set including tabulated photometric data and simulated H -band images from *Euclid*.

One of the key aspects of our approach is the simultaneous estimation of the three physical properties, whereas machine learning techniques have been used in previous works to estimate them individually. The other main component consists in the hierarchical combination of tools, each optimized for processing a specific kind of data, in a multimodal architecture.

The best results are achieved on the balanced data set, excluding the underrepresented regions of the variables. The performance on the estimation of redshift and stellar mass is >99 per cent, comparable to that achieved in Euclid Collaboration: Bisigello et al. (2023). Conversely, significant improvements are reached on SFR estimates, with a considerable reduction of outliers from ~ 30 per cent in that paper to our ~ 20 per cent: simultaneous estimation and adoption of advanced ML tools appear to provide better capabilities in constraining the most uncertain quantity by consistency with the other parameters. Galaxy evolution studies and extragalactic science, in general, will undoubtedly benefit from such improved diagnostics performance.

We emphasize that our method is general and can be easily applied to other multimodal data consisting of images and tabular information. In particular, it can be easily extended to include additional photometric data, which would significantly improve the available astrophysical information, thus possibly improving the estimation of the desired parameters. Multimodal data combination is a potential asset towards exploitation of the synergy among several modern projects and surveys; specific options will be evaluated, also depending on the available resources and opportunities.

The *Euclid* space telescope recently entered its nominal operation phase, and will soon provide its first data release. Thus, our proposed approach could be effectively tested and used on real data within about a year. For application to real data, future developments on FusionNetwork (including tests on other modern ML tools) may be required in order to improve the match, e.g. to Euclid in-flight parameters (PSF shape, noise levels, and so on), and to extend the training data set with inclusion of up-to-date science data, e.g. new observations, in particular to improve the description of as yet poorly represented classes of galaxies. Recent advances on simulated Euclid photometry and imaging (Euclid Collaboration: Bretonnière et al. 2023; Euclid Collaboration: Merlin et al. 2023) may be conveniently included to this purpose.

ACKNOWLEDGEMENTS

We acknowledge usage of the data set from Euclid Collaboration: Bisigello et al. (2023), kindly provided by that paper's authors. The

activity has also been partially supported by the Gruppo Nazionale per il Calcolo Scientifico (GNCS) of the Italian Istituto Nazionale di Alta Matematica (INdAM). MG also acknowledges support by the Agenzia Spaziale Italiana (ASI) through contract 2018-24-HH.0 and its Addendum 2018-24-HH.1-2022. The paper clarity and readability benefits from the care and dedication of our anonymous referee, which were deeply appreciated.

DATA AVAILABILITY

Our input data are those from Euclid Collaboration: Bisigello et al. (2023), made available by the authors of that paper upon request. Our output data will be shared on reasonable request to the corresponding author. The code will be shared after publication.

REFERENCES

- Bisigello L. et al., 2016, *ApJS*, 227, 19
- Bisigello L. et al., 2017, *ApJS*, 231, 3
- Bisigello L. et al., 2020, *MNRAS*, 494, 2337
- Cabayol L. et al., 2021, *MNRAS*, 506, 4048
- Cheng T.-Y. et al., 2020, *MNRAS*, 493, 4209
- Ciesla L., Elbaz D., Fensch J., 2017, *A&A*, 608, A41
- Conselice C. J., Bluck A. F. L., Mortlock A., Palamara D., Benson A. J., 2014, *MNRAS*, 444, 1125
- Estrada N. et al., 2023, *A&A*, 671, A146
- Euclid Collaboration: R. Scaramella, J. Amiaux, Y. Mellier, C. Burigana, 2022, *A&A*, 662, A112
- Euclid Collaboration: L. Bisigello C. Conselice M. Baes et al. 2023, *MNRAS*, 520, 3529
- Euclid Collaboration: E. Merlin M. Castellano H. Bretonnière et al. 2023, *A&A*, 671, A101
- Euclid Collaboration: H. Bretonnière U. Kuchner M. Huertas-Company et al. 2023, *A&A*, 671, A102
- Euclid Collaboration: Y. Mellier Acevedo Abdurro'uf A. Achúcarro et al. 2024, preprint ([arXiv:2405.13491](https://arxiv.org/abs/2405.13491))
- Gai M., Busonero D., Cancelliere R., 2017, *PASP*, 129, 054502
- Gai M. et al., 2022, *Front. Astron. Space Sci.*, 9, 388
- Gunn J. E. et al., 1998, *AJ*, 116, 3040
- Hausen R., Robertson B. E., Zhu H., Gnedin N. Y., Madau P., Schneider E. E., Villaseñor B., Drakos N. E., 2023, *ApJ*, 945, 122
- He K., Zhang X., Ren S., Sun J., 2016, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). p. 770, USA
- Henghes B., Thiyaalingam J., Pettitt C., Hey T., Lahav O., 2022, *MNRAS*, 512, 1696
- Humphrey A., Cunha P. A. C., Paulino-Afonso A., Amarantidis S., Carvajal R., Gomes J. M., Matute I., Papaderos P., 2023, *MNRAS*, 520, 305
- Iyer K. et al., 2018, *ApJ*, 866, 120
- Johnson J., Khoshgoftaar T., 2019, *J. Big Data*, 6, 27
- Khamis A. B., Ismail Z., Haron K., Mohammed A. T., 2001, *J. Appl. Sci.*, 5, 1394
- Laigle C. et al., 2016, *ApJS*, 224, 24
- Laureijs R. et al., 2012, in Clampin M. C., Fazio G. G., MacEwen H. A., Oschmann Jacobus M. J., eds, Proc. SPIE Conf. Ser. Vol. 8442, Space Telescopes and Instrumentation 2012: Optical, Infrared, and Millimeter Wave. SPIE, Bellingham, p. 84420T
- LeCun Y., Boser B., Denker J. S., Henderson D., Howard R. E., Hubbard W., Jackel L. D., 1989, *Neural Comput.*, 1, 541
- Li R. et al., 2022, *A&A*, 666, A85
- Liu D. Z. et al., 2023, *A&A*, 669, A128
- Madau P., Dickinson M., 2014, *ARA&A*, 52, 415
- Mortlock A. et al., 2013, *MNRAS*, 433, 1185
- Mowla L. et al., 2019, *ApJ*, 880, 57
- O'Shea K., Nash R., 2015, arXiv e-prints, p. arXiv:1511.08458
- Popescu M.-C., Balas V., Perescu-Popescu L., Mastorakis N., 2009, WSEAS Trans. Circuits Syst., 8, 579
- Rumelhart D. E., Hinton G. E., Williams R. J., 1986, *Nature*, 323, 533
- Scoville N. et al., 2007, *ApJS*, 172, 1
- Sola J., Sevilla J., 1997, *IEEE Trans. Nucl. Sci.*, 44, 1464
- Syarifudin M. R. I., Hakim M. I., Arifyanto M. I., 2019, *J. Phys. Conf. Ser.*, 1231, 012013
- Tohill C., Bamford S. P., Conselice C. J., Ferreira L., Harvey T., Adams N., Austin D., 2024, *ApJ*, 962, 164
- Treyer M., Ait Ouahmed R., Pasquet J., Arnouts S., Bertin E., Fouchez D., 2024, *MNRAS*, 527, 651
- Zeraatgari F. Z., Hafezianzade F., Zhang Y., Mei L., Ayubinia A., Mosal-lanezhad A., Zhang J., 2024, *MNRAS*, 527, 4677

This paper has been typeset from a $\text{\LaTeX}$ file prepared by the author.