

Euclid: A machine-learning search for dual and lensed AGN at sub-arcsec separations[★]

L. Ulivi^{★1,2,3}, F. Mannucci³, M. Scialpi^{2,1,3}, C. Marconcini^{2,3}, G. Cresci³, A. Marconi^{2,3}, A. Feltre³, M. Ginolfi^{2,3}, F. Ricci^{4,5}, D. Sluse⁶, F. Belfiore^{3,7}, E. Bertola³, C. Bracci^{2,3}, E. Cataldi^{3,2}, M. Ceci^{2,3}, Q. D’Amato³, I. Lamperti^{2,3}, R. B. Metcalf^{8,9}, B. Moreschini^{2,3}, M. Perna¹⁰, G. Tozzi¹¹, G. Venturi^{12,3}, M. V. Zanchettin³, Y. Fu^{13,14}, M. Huertas-Company^{15,16,17,18}, N. E. P. Lines¹⁹, M. Mezcua^{20,21}, M. Pöntinen²², K. Rojas²³, V. Scottez^{24,25}, M. Siudek^{16,20}, H. Teimoorinia²⁶, I. T. Andika^{27,28}, J. A. Acevedo Barroso²⁹, B. Clément^{29,30}, F. Courbin^{31,32,33}, R. Gavazzi^{34,35}, L. R. Ecker^{36,11}, B. C. Nagam^{37,14}, R. Pearce-Casey³⁸, S. Schuldt^{39,40}, S. H. Vincken⁴¹, D. Stern⁴², A. Chakraborty³, S. Andreon⁴³, N. Auricchio⁹, C. Baccigalupi^{44,45,46,47}, M. Baldi^{48,9,49}, A. Balestra⁵⁰, S. Bardelli⁹, A. Biviano^{45,44}, E. Branchini^{51,52,43}, M. Brescia^{53,54}, S. Camera^{55,56,57}, G. Cañas-Herrera^{58,13}, V. Capobianco⁵⁷, C. Carbone⁴⁰, J. Carretero^{59,60}, M. Castellano⁵, G. Castignani⁹, S. Caviuoti^{54,61}, A. Cimatti⁶², C. Colodro-Conde¹⁵, G. Congedo⁶³, C. J. Conselice⁶⁴, L. Conversi^{65,66}, Y. Copin⁶⁷, H. M. Courtois⁶⁸, M. Cropper⁶⁹, A. Da Silva^{70,71}, H. Degaudenzi⁷², G. De Lucia⁴⁵, A. M. Di Giorgio⁷³, C. Dolding⁶⁹, H. Dole⁷⁴, F. Dubath⁷², C. A. J. Duncan⁶³, X. Dupac⁶⁶, S. Dusini⁷⁵, S. Escoffier⁷⁶, M. Fabricius^{11,36}, M. Farina⁷³, R. Farinelli⁹, F. Faustini^{5,77}, S. Ferriol⁶⁷, F. Finelli^{9,78}, M. Frailis⁴⁵, E. Franceschi⁹, M. Fumana⁴⁰, S. Galeotta⁴⁵, K. George⁷⁹, B. Gillis⁶³, C. Giocoli^{9,49}, J. Gracia-Carpio¹¹, A. Grazian⁵⁰, F. Grupp^{11,36}, S. V. H. Haugan⁸⁰, W. Holmes⁴², I. M. Hook⁸¹, F. Hormuth⁸², A. Hornstrup^{83,84}, K. Jahnke⁸⁵, M. Jhabvala⁸⁶, B. Joachimi⁸⁷, E. Keihänen⁸⁸, S. Kermiche⁷⁶, A. Kiessling⁴², B. Kubik⁶⁷, M. Kümmel³⁶, M. Kunz⁸⁹, H. Kurki-Suonio^{22,90}, A. M. C. Le Brun⁹¹, S. Ligori⁵⁷, P. B. Lilje⁸⁰, V. Lindholm^{22,90}, I. Lloro⁹², G. Mainetti⁹³, D. Maino^{39,40,94}, E. Maiorano⁹, O. Mansutti⁴⁵, O. Marggraf⁹⁵, M. Martinelli^{5,96}, N. Martinet³⁴, F. Marulli^{8,9,49}, R. J. Massey⁹⁷, E. Medinaceli⁹, S. Mei^{98,99}, M. Melchior⁴¹, Y. Mellier^{24,35}, M. Meneghetti^{9,49}, E. Merlin⁵, G. Meylan²⁹, A. Mora¹⁰⁰, M. Moresco^{8,9}, L. Moscardini^{8,9,49}, R. Nakajima⁹⁵, C. Neissner^{101,60}, R. C. Nichol¹⁰², S.-M. Niemi⁵⁸, C. Padilla¹⁰¹, S. Paltani⁷², F. Pasian⁴⁵, K. Pedersen¹⁰³, W. J. Percival^{104,105,106}, V. Pettorino⁵⁸, S. Pires¹⁰⁷, G. Polenta⁷⁷, M. Poncet¹⁰⁸, L. A. Popa¹⁰⁹, L. Pozzetti⁹, F. Raison¹¹, A. Renzi^{110,75}, J. Rhodes⁴², G. Riccio⁵⁴, E. Romelli⁴⁵, M. Roncarelli⁹, R. Saglia^{36,11}, Z. Sakr^{111,112,113}, D. Sapone¹¹⁴, B. Sartoris^{36,45}, M. Schirmer⁸⁵, P. Schneider⁹⁵, T. Schrabback¹¹⁵, A. Secroun⁷⁶, G. Seidel⁸⁵, S. Serrano^{21,116,20}, P. Simon⁹⁵, C. Sirignano^{110,75}, G. Sirri⁴⁹, L. Stanco⁷⁵, J. Steinwagner¹¹, P. Tallada-Crespí^{59,60}, D. Tavagnacco⁴⁵, A. N. Taylor⁶³, I. Tereno^{70,117}, N. Tessore^{87,69}, S. Toft^{118,119}, R. Toledo-Moreo¹²⁰, F. Torradeflot^{60,59}, I. Tutusaus¹¹², E. A. Valentijn¹⁴, L. Valenziano^{9,78}, J. Valiviita^{22,90}, T. Vassallo^{45,79}, G. Verdoes Kleijn¹⁴, A. Veropalumbo^{43,52,51}, Y. Wang¹²¹, J. Weller^{36,11}, G. Zamorani⁹, F. M. Zerbi⁴³, E. Zucca⁹, and A. A. Nucita^{122,123,124}

(Affiliations can be found after the references)

Received July 16, 2025; accepted —

ABSTRACT

Cosmological models of hierarchical structure formation predict the existence of a widespread population of dual accreting supermassive black holes (SMBHs) on kiloparsec-scale separations, corresponding to projected distances $< 0''.8$ at redshifts higher than 0.5. However, close companions to known active galactic nuclei (AGN) or quasars (QSOs) can also be multiple images of the object itself, strongly lensed by a foreground galaxy, as well as foreground stars in a chance superposition. Thanks to its large sky coverage, sensitivity, and high spatial resolution, *Euclid* offers a unique opportunity to obtain a large, homogeneous sample of dual/lensed AGN candidates with sub-arcsec projected separations. Here we present a machine learning approach, in particular a Convolutional Neural Network (CNN), to identify close companions to known QSOs down to separations of $\sim 0''.15$, comparable to the *Euclid* VIS point spread function (PSF). We investigated the performance of the CNN trained on a large sample of simulated single and dual AGN. We studied the effectiveness of the CNN in identifying dual AGN and demonstrated that it outperforms traditional techniques such as IRAFStarFinder. Applying our CNN to a sample of ~ 6000 QSOs from the Q1 *Euclid* data release (covering 63.1 deg^2), we find a fraction of about 0.25% dual AGN candidates with separation $\sim 0''.4$ (corresponding to $\sim 3 \text{ kpc}$ at $z=1$). Estimating the foreground contamination from stellar objects, we find that most of the pair candidates with separation higher than $0''.5$ are likely contaminants, while below this limit, contamination is expected to be less than 20%. For objects at higher separation ($> 0''.5$, i.e. 4 kpc at $z=1$), we performed PSF subtraction and used colour-colour diagrams to constrain their nature. We present a first set of dual/lensed AGN candidates detected in the Q1 *Euclid* data, providing a starting point for the analysis of future data releases.

Key words. quasars:general; Methods: data analysis – surveys

1. Introduction

The formation and growth of supermassive black holes (SMBHs) are closely linked to the evolution of their host galax-

Article number, page 1 of 20

[★] This paper is published on behalf of the Euclid Consortium.

^{★★} e-mail: lorenzo.ulivi@inaf.it

ies. In particular, galaxy mergers are thought to play a central role in driving gas inflows toward the central regions, potentially triggering both star formation and active galactic nuclei (AGN) activity (e.g. [Hernquist 1989](#); [Hopkins et al. 2008](#); [Blecha 2017](#)). When two galaxies merge, the SMBHs in their nuclei become part of the same post-merger galaxy, and eventually merge together, typically on timescales of the order of Gyr (e.g. [Khan et al. 2012](#)). In the final stages of a merger, both SMBHs may be simultaneously active, forming a dual AGN system (i.e. two accreting SMBHs within a single merged galaxy). These systems provide a unique laboratory for studying SMBH growth, AGN triggering, and the interplay between galaxy evolution processes and nuclear activity.

Simulations predict that dual AGN can form in both major and minor galaxy mergers (e.g., [Capelo et al. 2015, 2017](#); [Chen et al. 2023](#)), with longer observable dual-AGN lifetimes expected in major mergers. Conversely, the highest SMBH mass ratios are predicted to arise in minor mergers, which may dominate the observed population at lower luminosities ($< 10^{44} \text{ erg s}^{-1}$), especially in the regime accessible to facilities such as *Euclid*.

While theoretical models predict that dual AGN should be relatively common, especially at separations below several kpc (e.g. [Rosas-Guevara et al. 2019](#); [Volonteri et al. 2022](#); [Chen et al. 2023](#)), observational studies have found mixed results, with reported dual AGN fractions ranging from a few percent to about 25% depending on selection method and sample properties (e.g. [Koss et al. 2012](#); [Perna et al. 2025](#)). However, these estimates are possibly affected by several biases, including reliance on visual classification which can miss obscured or compact systems, as well as luminosity cuts, maximum separation, and the selection of the sample considered. Quantifying the frequency and properties of dual AGN as a function of separation, redshift, and host galaxy properties (e.g. extinction, star formation, black hole mass) is essential for constraining models of SMBH formation and co-evolution with galaxies. In particular, galactic scales ranging from tens of kiloparsecs to tens of parsecs represent a key regime, corresponding to the late stages of mergers when SMBHs are on the path to forming bound binaries.

On the observational side, there is still a lack of robust statistical samples of kpc-scale dual AGN systems at $z > 1$. Identifying such systems poses significant challenges, primarily due to the stringent requirements on spatial resolution and reliable AGN diagnostics. Detecting dual AGN at kpc separations requires sub-arcsec resolution, while their intrinsic rarity demands wide-area searches. To date, ground-based wide-field imaging and spectroscopic surveys have proven to be largely inefficient in uncovering sub-arcsec dual and lensed at high redshift.

Recently, the all-sky *Gaia* mission has enabled the identification of candidate dual and lensed AGN at sub-arcsec separations, extending to $z > 1$ (e.g. [Lemon et al. 2017](#); [Chen et al. 2022](#); [Mannucci et al. 2022, 2023](#); [Ciurlo et al. 2023](#); [Glikman et al. 2023](#); [Scialpi et al. 2024](#)). However, this approach is limited to optically bright quasars (QSOs) due to *Gaia*'s magnitude limit of $G \sim 21$, which corresponds to $L_{\text{bol}} \gtrsim 10^{46} \text{ erg s}^{-1}$ at $z \gtrsim 0.5$. As a result, current *Gaia*-selected samples probe only the most luminous AGN population, significantly biasing against more typical, moderate-luminosity AGN and limiting the completeness of such surveys.

Over the last ten years, astrophysics has undergone a significant shift in its approach, transitioning swiftly from dealing with relatively small datasets to operating with big data. Survey telescopes such as the ESA satellite *Euclid* ([Euclid Collaboration: Mellier et al. 2025](#)) generate vast amounts of data

that require automated tools leveraging advancements in high-performance computing, machine learning, data science, visualization and foundation models (e.g. [Euclid Collaboration: Rojas et al. 2025](#); [Acevedo Barroso et al. 2025](#); [Euclid Collaboration: Walmsley et al. 2025](#); [Euclid Collaboration: Lines et al. 2025](#); [Pearce-Casey et al. 2025](#); [Nagam et al. 2025](#); [Euclid Collaboration: Siudek et al. 2025](#)). Machine learning (ML) and, in particular, deep learning (DL) offer powerful data-driven approaches for an efficient analysis of large and complex datasets.

In this work, we explore the application of ML techniques to the identification of dual AGN systems in imaging data, with a particular focus on the first *Euclid* quick release, Q1 ([Euclid Quick Release Q1 2025](#)). As dual AGN are rare and often difficult to identify, especially at high redshift and small angular separations, advanced classification methods, such as convolutional neural networks (CNNs), offer a promising approach to improve both the completeness and reliability of detections.

The present paper is structured as follows. In Sect. 2 we describe the input catalogue of AGN, the data used in this work, which consists of the *Euclid* Q1 data, and the ML training set (Sect. 2.4) as well as the architecture of the CNN. In Sect. 3 we present the results obtained from simulations, compare our method with other standard techniques and discuss the identification of an optimal classification threshold. In Sect. 4 we apply our CNN to the Q1 data, we discuss foreground contamination, we extract the photometry exploiting the four *Euclid* bands and present a sample of dual AGN candidates. In Sect. 5 we compare our results with both simulation and observational works from the literature. Throughout this work, we assume $\Omega_m = 0.286$ and $H_0 = 69.9 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ([Bennett et al. 2014](#)).

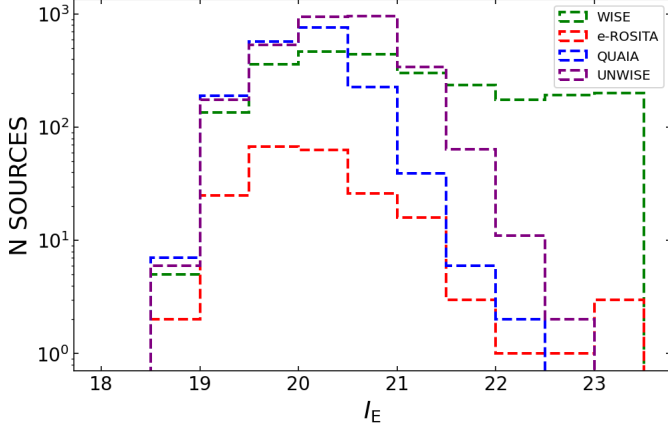
2. Data and methods

2.1. Euclid Data

This work is based on imaging data from the *Euclid* mission, which is conducting two major surveys: the *Euclid* Wide Survey (EWS) and the *Euclid* Deep Survey (EDS). The EWS ([Euclid Collaboration: Scaramella et al. 2022](#)) is designed to cover approximately one-third of the sky, amounting to $\sim 14\,000 \text{ deg}^2$, while avoiding the Galactic and ecliptic planes. The Deep Survey, in contrast, focuses on three carefully selected regions – the Euclid Deep Field (EDF) North, South, and Fornax – and achieves significantly greater depth through repeated exposures. In this study, we use data from the *Euclid* Q1 release ([Euclid Collaboration: Aussel et al. 2025](#)), which includes $\sim 63 \text{ deg}^2$ of imaging of the EDFs acquired to the depth of the Wide Survey. We use the `pdMerBksMosaic` data products ([Euclid Collaboration: Romelli et al. 2025](#)), which include background-subtracted mosaics and associated point spread function (PSF) catalogues (catalogue-PSF). For this analysis, we consider all Q1 fields which span approximately 350 tiles, each with a field of view of 0.57 deg^2 , corresponding to $0.75^\circ \times 0.75^\circ$. These observations reach a $\sim 10\sigma$ magnitude limit of 24.5 in the Visible Instrument (VIS) filter, I_E ([Euclid Collaboration: Cropper et al. 2025](#)), and a 5σ magnitude limit of 24.5 in each of the filters of the Near Infrared Spectrometer and Photometer (NISIP), namely Y_E , J_E , and H_E ([Euclid Collaboration: Jahnke et al. 2025](#)). The Q1 data give us a first glimpse into the possibilities of efficiently searching for dual AGN at sub-arcsec separations over the full *Euclid* survey of $\sim 14\,000 \text{ deg}^2$.

Table 1. List of QSO catalogues used to compile the search list.

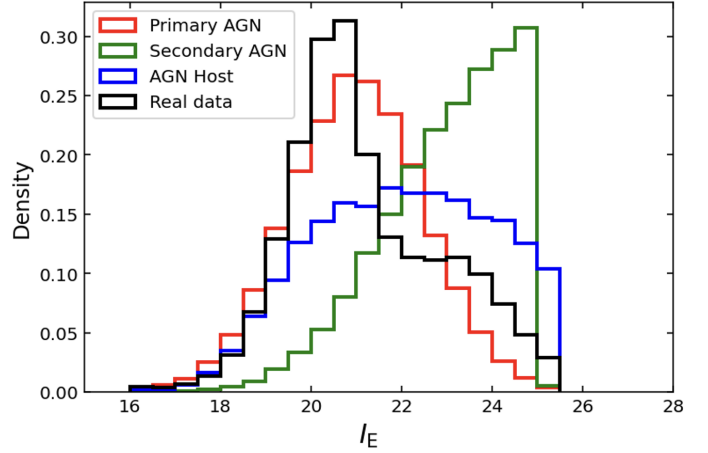
Catalogue	Properties	Number	Ref
Milliquas v8.0	Spectroscopy from literature	729 519	Flesch (2023)
Quaia G20.5	<i>Gaia</i> and <i>WISE</i> colours	878 145	Storey-Fisher et al. (2024)
eRosita	X-ray flux	177 971	Merloni et al. (2024)
<i>WISE</i>	<i>WISE</i> colours	2 275 316	Assef et al. (2018)
<i>Gaia</i> /unWISE	<i>Gaia</i> and <i>WISE</i> colours	1 228 393	Shu et al. (2019)
TOT (excluding duplications)		3 295 296	


Fig. 1. Number of Q1 AGN sample in $\Delta I_E = 0.5$ magnitude bins as a function of the selection method: *WISE* (green dashed line), e-ROSITA (red), QUAIA (blue), and UNWISE (purple).

2.2. Input QSO catalogue

The dual/lensed search strategy consists of looking for close companions near previously classified QSOs. We created the input QSO list by merging several external catalogues of both spectroscopically confirmed QSOs and robust photometric QSO candidates selected in various ways. Our aim is to privilege completeness over purity, while discarding the most unreliable candidates. More accurate assessments of the nature of candidate double systems could be later performed on an individual basis, e.g. with AO-assisted spectroscopy. For each catalogue we only selected the QSOs falling within the EWS footprint, i.e. excluding objects within 23° from the Galactic plane and 10° from the ecliptic plane. We also excluded QSOs close to the largest galaxies in the Local Group, e.g., Large Magellanic Cloud (LMC), Small Magellanic Cloud (SMC), M31, and M33. Table 1 reports the catalogues used and the number of objects obtained from each of them. In particular:

- Milliquas v8.0 ([Flesch 2023](#)): a compilation of all the known QSOs with a reliable redshift. Most of its entries are from the Sloan Digital Sky Survey (SDSS) QSO catalogue DR16Q ([Lyke et al. 2020](#)), the DESI EDR QSO catalogue ([Chaussidon et al. 2023](#)), the LAMOST survey ([Jin et al. 2023](#)), and the 2DF QSO survey ([Croom et al. 2004](#)).
- *WISE*: [Assef et al. \(2018\)](#) presented a catalogue of about 4.5 million QSO candidates based on *WISE* colours. To decrease the number of false positives, following [Stern et al. \(2012\)](#), we further select the objects with *WISE* colour $W1 - W2 > 0.8$ and having $W1 < 18.5$ and $W2 < 16.0$ (77% of the original sample).
- Quaia: [Storey-Fisher et al. \(2024\)](#) drew an all-sky QSO catalogue combining *Gaia* candidates ([Gaia Collaboration et al.](#)


Fig. 2. Distribution of I_E of the primary AGN (red), secondary AGN (green) and the associated host galaxy (blue) in simulations compared to the distribution of I_E of the considered sample of AGN in Q1 (black) in 0.5 magnitude bins.

2016) with unWISE infrared data (based on the *WISE* survey).

- eRosita: [Merloni et al. \(2024\)](#) present catalogues of both point-like and extended sources using the data acquired in the first six months of survey operations (eRASS1; completed in June 2020) over half of the sky, whose proprietary data rights lie with the German eROSITA Consortium. From their full catalogue, we selected all the sources that are clearly detected ($\text{DET_LIKE}_0 > 9$) and not extended (parameter $\text{EXT} = 0$).

We crossmatched the different catalogues using a $5''$ search radius to find duplicate sources. When a duplication is found, the coordinates derived from the *Gaia* and spectroscopic catalogues are preferred over those from eRosita and *WISE*. The final catalogue has about 3.29 million sources (see Table 1).

2.3. Euclid VIS stamps

Starting from the catalogue described in the previous section, we performed cutouts of the VIS images of $5'' \times 5''$ and $4' \times 4'$ around each AGN in the catalogue, ending up with 15 000 AGN images in the Q1 data. We used SExtractor ([Bertin & Arnouts 1996](#)) for object detection in both types of cutouts, employing specific parameters to optimise performance. In particular, the goal to use SExtractor in the small ($5'' \times 5''$) cutouts was to measure object properties and apply selection cuts in magnitude and size for our final input sample. The larger $4' \times 4'$ cutouts were processed with SExtractor to build a catalogue of host galaxies, which were then used in the simulation described in the following section. The SExtractor DETECT MINAREA was set to 2, the DETECT THRESH was set at 1.5 times the background

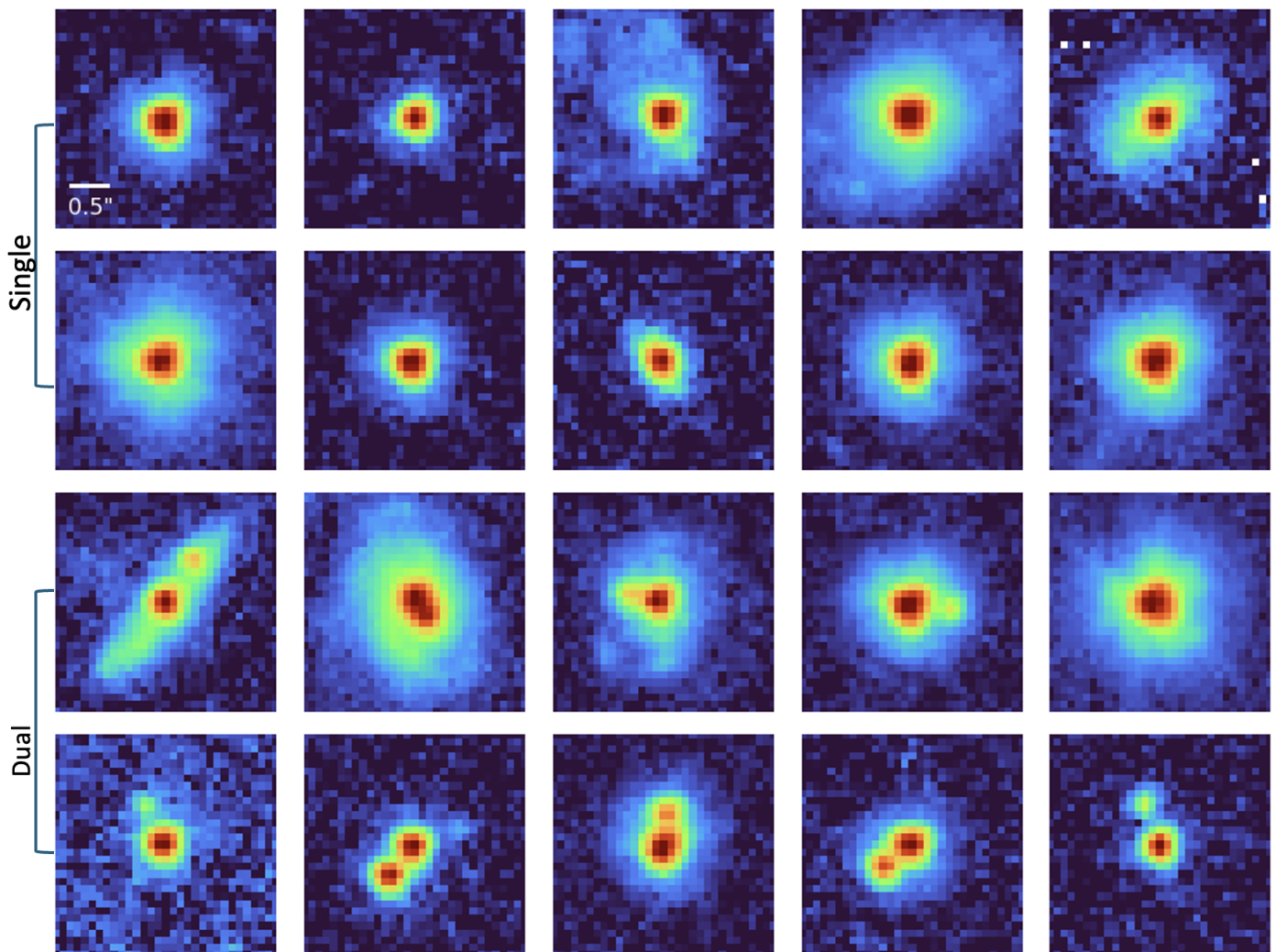


Fig. 3. Examples of ten simulations of a single AGN (first two rows) and of a dual AGN (second two rows) in I_E band. The size of each cutout is $3'' \times 3''$.

RMS noise, allowing us to detect faint objects. We applied a 4-pixel Gaussian filter (indicating that the sources are smoothed with a Gaussian kernel with $\sigma = 4$ pixel) and set the parameter that defines the minimum contrast between the two sources to consider them separated to 0.005. These parameter choices were tailored to improve detection sensitivity. MAG ZERO is provided in the header of each tile (Euclid Collaboration: Romelli et al. 2025). In this way, for each small cutout, we generated a source catalogue and assigned to the AGN candidate from the external catalogue the magnitude I_E of the brightest detected object within the cutout. We flagged those sources which were not centred in a square of $0''.3$ around the coordinates of the external catalogue. We centred the image in the pixel with the highest flux. For each AGN we keep track of the *Gaia* magnitude, the *WISE* colours and, the selection method.

In Fig. 1 we report the number of sources in 0.5 magnitude bins of I_E for different selection methods. We see that most of the QSOs in our catalogue are selected through (UN)WISE colour.

2.4. Simulations

The simulations described in this section were specifically designed for training and validating a neural network to detect single and dual AGN systems. Our strategy is based on adding syn-

thetic AGN sources, represented by PSFs, into real *Euclid* VIS galaxy images to create realistic examples of both single and dual AGN systems. We generated a dataset comprising 100 000 synthetic images, 50 000 containing a single AGN source and 50 000 containing a dual AGN system in the VIS band. Each AGN was modelled as a point source. When downloading the Q1 tiles, we also downloaded the grid of PSFs provided by the official pipeline (catalogue PSF). From each tile, we made cutouts of 200 randomly chosen PSFs throughout the field. The PSFs taken from the catalogue PSF are undersampled (19×19), so we used drizzle (Gonzaga et al. 2012) to oversample them using the ‘turbo’ kernel to keep the flux constant. We split the 100 000 simulations (50 000 single sources, 50 000 double sources) into 60% assigned to the training set and of the remaining 40%, 20% to the validation set and 80% to the test set.

The magnitude of each primary AGN was drawn from a Gaussian distribution centred at $I_E = 21$ with a standard deviation of $\sigma = 1.2$, chosen to closely match the observed magnitude distribution of the AGN from the external catalogue, as shown in Fig. 2. A priori, the relative contributions of the AGN and its host galaxy to the total observed luminosity are unknown. Previous studies have shown that this contribution can vary significantly with luminosity and redshift. For instance, Rakshit et al. (2020), analysing a sample of approximately 500 000 QSOs at

$z < 0.8$, found that the host galaxy contribution decreases with increasing total luminosity. At bolometric luminosities around $10^{45} \text{ erg s}^{-1}$, the host galaxy can contribute up to 80% of the total emission. Extending to higher redshifts ($z \sim 1.5$), [Schramm & Silverman \(2013\)](#) demonstrated that, particularly for AGN with lower luminosities, the host galaxy contamination can be even more significant, even exceeding the 90% of the total. Given this uncertainty, we developed simulations covering a broad parameter space to robustly explore the possible contributions of the AGN and its host galaxy. However, the main purpose of including host galaxies was to enhance the training of the neural network by introducing more realistic and challenging cases while preserving the underlying physics.

We associated to each point source a host galaxy with a flux between 0 and 10 times that of the point source and with $\text{CLASS_STAR} < 0.5$, where CLASS_STAR , taken from `SExtractor`, is defined as the probability to be a point source (1) or an extended object (0). In the case of host galaxies dimmer than $I_E = 27.5$, we just added random noise to the image as they would be below the detection limit of *Euclid*. To select the random noise, we computed for each tile the median and the median absolute deviation (MAD) and we made random cutouts of the FOV of the tile, accepting the cutout if the mean is less than the median and the $\sigma < 2 \text{ MAD}$. We designed 50 000 simulations with single objects and we ended up with 2% of simulated AGN painted without the host galaxy. Additionally, we added to each stamp a secondary source with a separation drawn from a uniform distribution between $0''.1$ and $0''.8$ from the primary AGN and whose magnitude was imposed to vary uniformly in magnitude between that of the primary and 25 mag, the limiting AB magnitudes (10σ for point-like sources) achieved in each footprint.

We explored several sets of simulations, each differing in some aspects to better capture the important features. The main differences between the simulation sets include: the distribution of magnitudes of the primary and secondary objects; the relationship between the AGN and host galaxy magnitudes; the distribution of projected separation between sources; the PSF used to model the AGN; and the level and standard deviation of the noise. These parameters are critical for the neural network because it has to be trained on data as close as possible to the *Euclid* dataset, while learning to recognize features. The main problem of using a neural network to identify dual AGN is inherent to their rareness, which poses a serious challenge to the training of the network. For this purpose, it is necessary to simulate many of such sources, which makes the distribution of simulated objects different from that of the *Euclid* data. This in itself is true for dual AGN, but it is also true for the simulated magnitude distributions of the two sources; in fact, the probability of having more faint primary and secondary objects will be larger than the probability of having bright objects, but this leads to imbalances in the simulations from which the network can learn biased features. However, in this work we focus on the simulation setup described above, which provided the best performance in training and validating our model with the Q1 data (see also Appendix A). Some examples of simulated single and double AGN are shown in Fig. 3.

2.5. Convolutional neural network classifier

The images were labelled with 2 if they contained dual sources with separations greater than $0''.25$, 1 if they contained dual sources with separation less than $0''.25$ or 0 if single sources. The $5'' \times 5''$ stamps are centred in the brightest pixels and cut into a

30×30 pixels ($3'' \times 3''$), which is first processed by a 2D convolution layer with a 4×4 filter size, then subsampled by a 2×2 max pooling layer. Two more identical units follow, with a growing dimensionality of the output space in the convolution, for a total of 2 convolutional layers and 2 max pooling layers. Each of these convolutional layers is followed by a ReLU activation layer. The output of these units is then processed through three single fully-connected layers. For the classification problem we used the cross-entropy cost function and found the weights using the ADAM ([Kingma & Ba 2014](#)) optimisation method. Note that in PyTorch ([Paszke et al. 2019](#)), no explicit softmax activation function is needed in the final layer, as it is handled internally by the cost function. The use of the ADAM optimizer improved the learning rate compared to tests with stochastic gradient descent (SGD). To prevent overfitting and enhance generalisation, we added dropout layers which randomly drop out a fraction (30%) of neurons during the training phase. We implemented our network as a multitask model, enabling it to predict, in addition to the nature of the system, the position of the secondary, the magnitude difference between the secondary and the primary, and the magnitude difference between the host and the primary. Multitask neural networks are powerful tools because they allow a model to learn several related tasks simultaneously, improving the performance while reducing the risk of overfitting, especially when the tasks are correlated. This approach often leads to more robust models and can capture structures in the data more effectively than training separate models for each task ([Ruder 2017](#); [Crawshaw 2020](#); [Ginolfi et al. 2025](#)).

We built another block in our CNN that predicts the coordinates x and y of the secondary AGN. To increase the effectiveness of the network, we made it so that the position of the secondary can vary between -1 and 1 , thus normalising to the maximum distance that the secondary can reach in our simulations. The activation function used for this task is a tanh function (so that it varies between -1 and 1). We assigned to the single sources the position of the secondary source at the value $x = 0$, $y = 0$. This can also give us an indication for detecting false positives, that is, single objects classified as dual; in fact when the network detects dual but predicts secondaries with coordinates very close to $(0,0)$, we can discard this prediction.

We also added two branches to the neural network to predict the magnitude difference between the primary and secondary sources ($\Delta I_{E,21}$) and, though more difficult, the difference in magnitude between the primary and its host galaxy ($\Delta I_{E,H1}$). For the single objects, we imposed as the true label of $\Delta I_{E,21}$ the maximum among those simulated in the double objects, so as to normalise all the sample between 0 and 1 and without restricting in a smaller parameter space the double objects. We did the same for the difference between the magnitude of the primary and that of the host galaxy. We used the mean squared error as the loss function for the regression tasks. For the multi-task learning framework, we computed the total loss as the sum of all individual task losses, with the classification loss contributing more strongly to the optimization process due to its relative scale compared to the other tasks. The final architecture of our model is illustrated in Fig. 4.

3. Simulation results

In this section, we present the results of our neural network applied to the simulated datasets. In particular, we study how the completeness and accuracy vary as a function of both the physical parameters of the systems and the quantities predicted by the network. We also compare the performance of our approach with

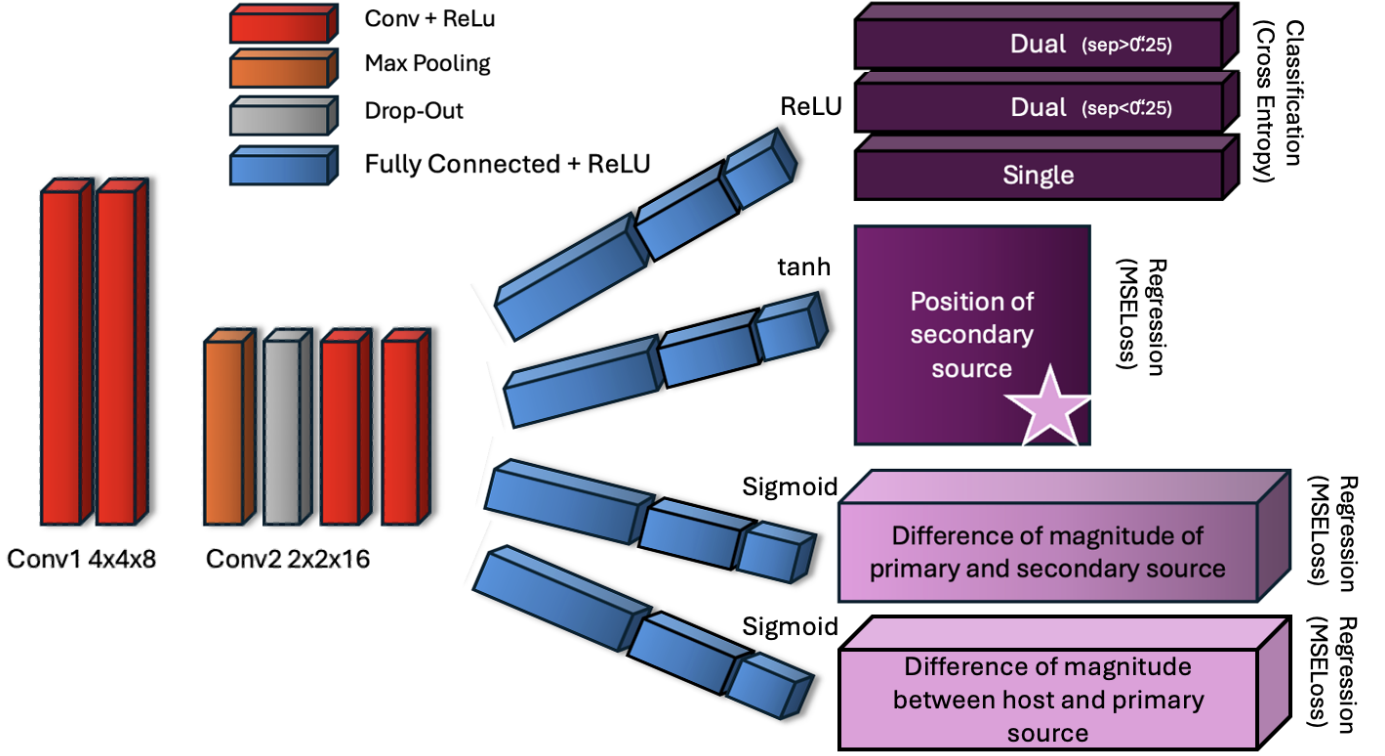


Fig. 4. Architecture of the convolutional neural network used for classification and feature extraction. The model consists of two convolutional layers (Conv1: 4×4×8, Conv2: 2×2×16) with ReLU activation, followed by max pooling and dropout layers. Fully connected layers (2304–80–20) process the features for output into multiple targets: the position of the secondary source, the difference in magnitude between primary and secondary sources, and the difference in magnitude between the host and primary source. Each output branch uses a different final activation function and loss: ReLU and cross-entropy loss function for the classification task, hyperbolic tangent for the position, and sigmoid for the magnitude differences. The loss functions are the mean squared error when applied to regression task.

standard methods, and select the most suitable network architecture for the application to the real data.

3.1. ROC curve and comparison with standard method

We evaluated the performance of the CNN by analysing its receiver operating characteristic (ROC; Bradley (1997)) curve. This curve is constructed by varying the threshold probability used to classify galaxies as dual or single, considering the number of True Positives (TPs), True Negatives (TNs), False Positives (FPs), and False Negatives (FNs). The TP rate, also called completeness, which measures the fraction of true mergers correctly identified, and the FP rate, which indicates the fraction of single AGN incorrectly classified as dual, are computed as follows:

$$\text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \quad (1)$$

$$\text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}}. \quad (2)$$

Figure 5 shows the ROC curves for the test set. We used a one-vs-rest (OvR) approach, defining classes 1 and 2 as the positive class and class 0 (single sources) as the negative class. The probability of the positive class was then computed as the sum of the predicted probabilities for classes 1 and 2. The perfect classifier reaches the top left corner, with the maximum TP rate and a FP rate of 0. Hence, we can use the area under the curve (AUC) as a metric of performance. The AUC curve in our CNN reaches

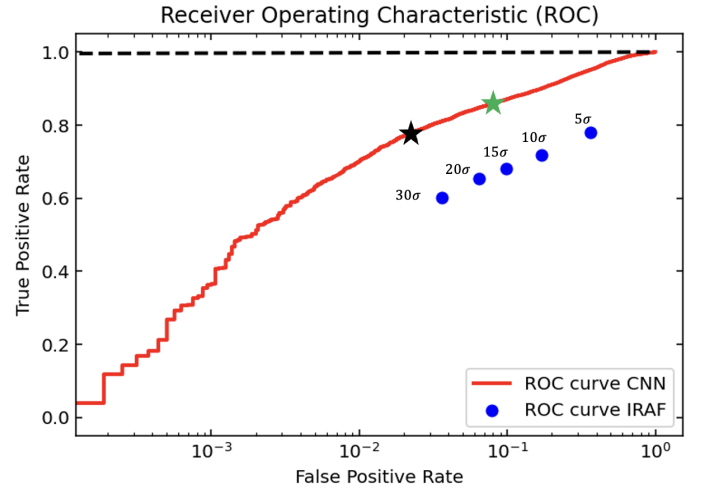


Fig. 5. ROC curve comparing the performance of the two detection methods in log scale. The red curve represents the ROC of the CNN, while the blue points correspond to the ROC obtained using the IRAFStarFinder algorithm with varying detection thresholds (5, 10, 15, 20, 30 σ). The green star marks the threshold closest to an ideal classifier, while the black star indicates the threshold that minimises the difference between the FPs and TPs assuming a real case scenario (see Sect. 3.4).

0.98. We want to compare the performance of our CNN with the

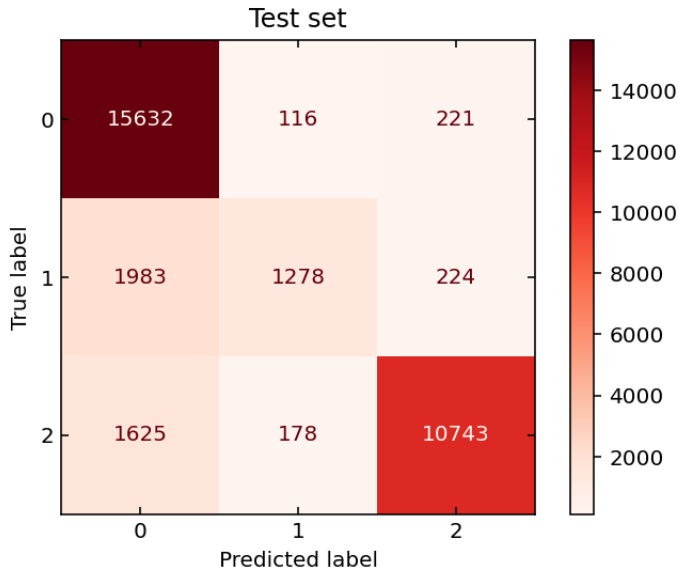


Fig. 6. Confusion matrix. 0: single AGN, 1: dual AGN with separation less than $0''.25$, 2: dual AGN with separation greater than $0''.25$.

performance of standard methods such as **SExtractor**¹ or the **IRAFStarFinder** module of **photutils** (Bradley et al. 2021). The deblending method adopted in **SExtractor** is based on multi-thresholding and works on any kind of object but it is unable to deblend components that are so close that no saddle is present in their profile. A better optimiser for point source detection is **IRAFStarFinder** that searches images for local density maxima that have a peak amplitude greater than a threshold above the local background and have a PSF FWHM similar to the input FWHM. Therefore, we measured the ROC curve of the standard method with **IRAFStarFinder** increasing 5, 10, 15, 20, 30 σ , and with a minimum detection separation from the limit imposed by the algorithm ($2 \times \text{FWHM}$). Obviously, these values are simply for the purpose of having a direct comparison between the two curves. We first calculated the true positives by applying the algorithm to all simulations with dual objects and then the false negatives by applying the algorithm to the single objects. For the calculation of TP, we considered the correct detection if the algorithm finds at least one source within $0''.8$ of the primary source. Detection of sources at larger distances this means that the network has detected other source present in the field (foreground objects). We note that this represents an upper limit to TPs since we consider a correct detection even if the algorithm detects possible contamination within $0''.8$ (clumps, structure in the galaxy). For the calculation of FNs, we applied the algorithm to single sources considering an incorrect prediction if it finds more than one source within $0''.8$. The ROC curves of the neural network and **IRAFStarFinder** predictions are shown in Fig. 5 as red line and with blue dots for increasing sigma, respectively. The ROC curve of the CNN lies above that of **IRAFStarFinder**, indicating that the CNN achieves higher completeness and precision across all detection thresholds. Classical methods do not reach the same level of completeness as the CNN without substantially increasing the number of false positives.

¹ <https://sextractor.readthedocs.io/en/latest/Introduction.html>

3.2. Confusion matrix

Figure 6 shows the confusion matrix, where the y-axis represents the true labels and the x-axis represents the predicted labels. The labels are defined as follows: 0 corresponds to a single AGN, 1 to a dual AGN with separation $< 0''.25$, and 2 to a dual AGN with separation larger than $0''.25$. If we consider dual candidates labelled as 1 or 2, we find 337 false positives, corresponding to approximately 2%, and a completeness of about 78%. This leads to a F1 score, defined as the harmonic mean of the precision and completeness (recall) scores, of 0.86. If, instead, we restrict the analysis to candidates labelled as 2 (dual AGN), the precision increases to 98.6%, with a completeness of 68% and a F1 score of 0.80. However, these global metrics are not necessarily meaningful in a real case scenario since the true distributions of source separations and magnitudes are unknown. In particular, in the limiting case where all dual AGN lie at separations $\sim 0''.25$, where some models predict the peak of their number density (Saeedzadeh et al. 2024), the completeness metric would lose significance, since many of them would not be detected. What is more important is to understand how precision and completeness vary across the physical parameter space.

3.3. Predictions of the model

As shown in Fig. 4, the network has four different end layers that predict the nature of the AGN, but also the exact location of the secondary source, $\Delta I_{E,21}$, $\Delta I_{E,H1}$. Figure 7 shows the network predictions, where the x- and y-axis are respectively the true value, and the difference between the network prediction and the true value. Panel a) presents the predictions of the separation between the two objects computed as $d = \sqrt{(x_0 - x_{\text{pred}})^2 + (y_0 - y_{\text{pred}})^2}$ where x_0 and y_0 are the central pixel of the primary source, while x_{pred} and y_{pred} are the coordinates of the secondary source predicted by the CNN. The plot only includes sources that are predicted to be dual sources. When the network is not able to detect a secondary component, it also cannot accurately assign a distance, increasing the scatter in the plot. We see that the prediction is correct within $\pm 0''.1$. Panels b) and c) show ΔI_E between the secondary and primary source ($\Delta I_{E,21}$) and ΔI_E between the host and primary ($\Delta I_{E,H1}$). The prediction of $\Delta I_{E,21}$ shows the scatter from the true value of $\sim \pm 0.5$ mag that increases when separations are small. This is mainly due to the fact that, at close separations, the network can no longer perform precise photometry.

3.4. Threshold to filter

In order to choose the best thresholds for filtering, it is also necessary to consider the number of dual AGN that we expect. According to simulations the number of dual AGN can vary between 0.1 and 10% of all AGN (Steinborn et al. 2016; Rosas-Guevara et al. 2019; Volonteri et al. 2022; Puerto-Sánchez et al. 2025), so being a class of rare objects, it is necessary to choose a metric such that the FPR are minimised compared to TPR. To do this, the output of the neural network can be used to filter out objects that most likely represent contaminants. We try to understand how completeness and precision vary as a function of network output by changing the thresholds of the output. In particular, we vary the threshold of each parameter for the prediction of all dual systems in bins of magnitude of the primary, secondary, and separation. This allows us to estimate the expected completeness in real observations for each region of parameter

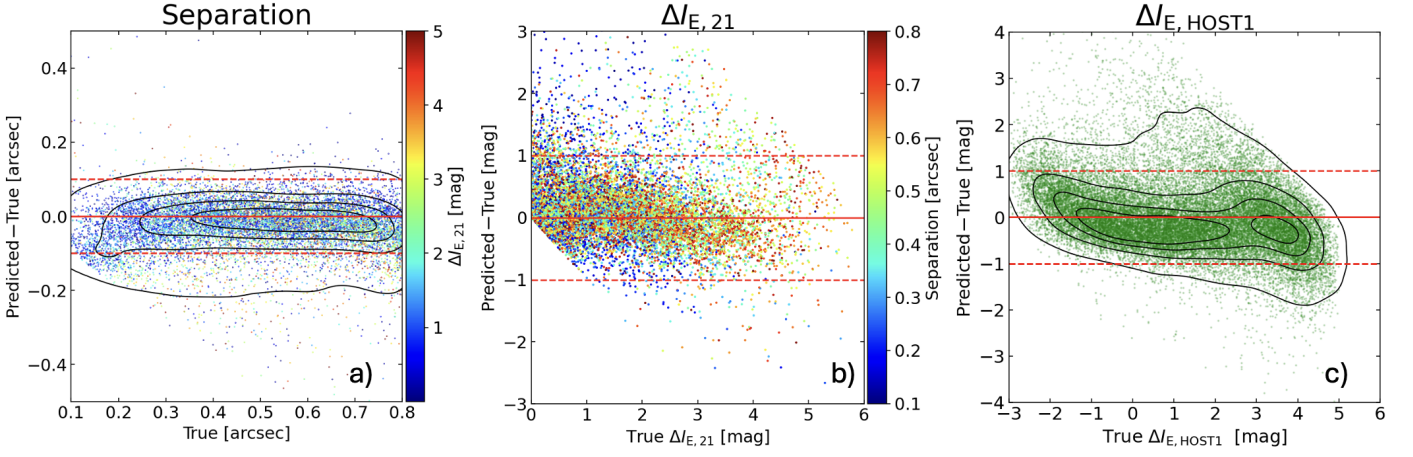


Fig. 7. Prediction of the CNN of objects classified as dual AGN. Scatter plots comparing true values (x-axis) with the difference of the predicted values and the true values (y-axis) for different parameters derived from the CNN. (a) Predicted separation (in arcsec) between the two sources. The colour bar represents the $\Delta I_{E,21}$ for the data points. (b) Prediction of the ΔI_E between the secondary and primary sources ($\Delta I_{E,21}$). The colour bar represents the separation in arcsec of the two sources. (c) Prediction of the magnitude difference (ΔI_E) between the host and primary source $\Delta I_{E,H1}$. Red dashed lines indicate perfect predictions and ± 0.1 in arcsec error margins (± 1 in px) and ± 0.5 in mag.

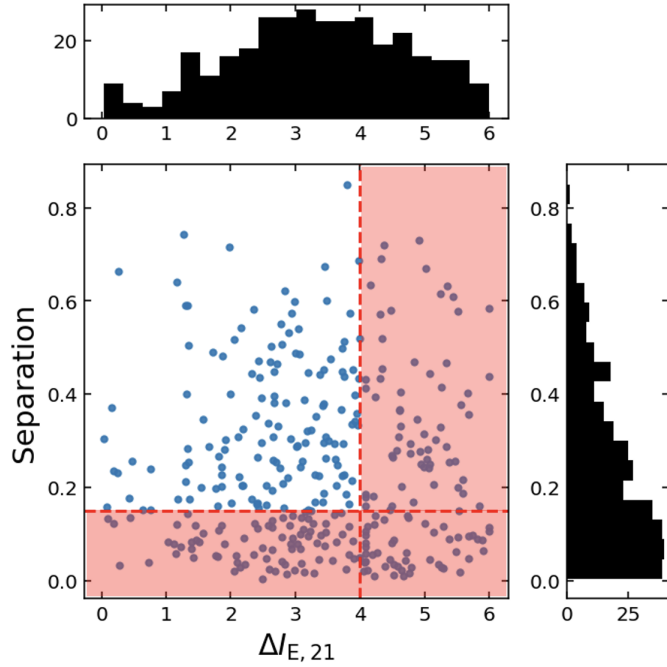


Fig. 8. CNN prediction of separation and the magnitude difference between the primary and secondary sources ($\Delta I_{E,21}$) for FP objects. The top and right histograms show the distribution of $\Delta I_{E,21}$ and of the separation of FPs, respectively. The red shaded regions indicate the excluded objects, defined by a separation smaller than $15''$ and $\Delta I_{E,21}$ higher than 4.

space. We perform the same analysis for systems labelled as single, in bins of magnitude of the primary and of the host, to estimate the number of false positives. For each trained network and for each value of the decision threshold applied to the network output, we calculate the TPR and FPR with Eqs. (1) and (2).

To determine the optimal classification threshold, we adopted a modified selection metric defined as

$$(\text{TPR } f_{\text{dual}} - \text{FPR } f_{\text{single}}) N_{\text{QSO}}, \quad (3)$$

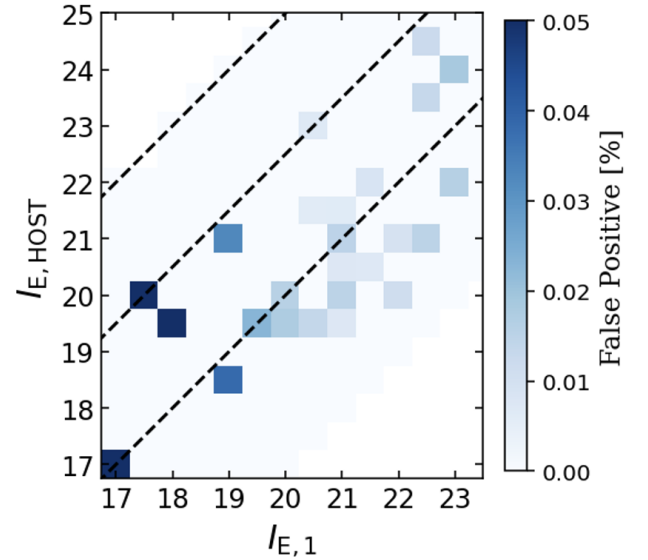


Fig. 9. FP rate in bins of magnitude of the host galaxy, $I_{E,\text{HOST}}$, and of magnitude of the primary AGN, $I_{E,1}$.

where f_{dual} and $f_{\text{single}} = 1 - f_{\text{dual}}$ represent the assumed fractions of dual and single AGN, respectively, and N_{QSO} is the total number of QSOs in the input sample. This metric approximates the expected number of true positives minus false positives, and was evaluated in the separation range of $0''.2$ to $0''.6$, where the number of dual AGN is expected to be higher. Assuming a dual AGN fraction of 1% and $N_{\text{QSO}} = 6000$, the resulting optimal thresholds is 0.9 and typically yield an expected net number of true detections (TP – FP) between 10 and 20. However, this estimate relies on several assumptions, in particular that the separation and magnitude distributions of dual AGN in the real data mirror those used in our simulations, an approximation that may not reflect reality.

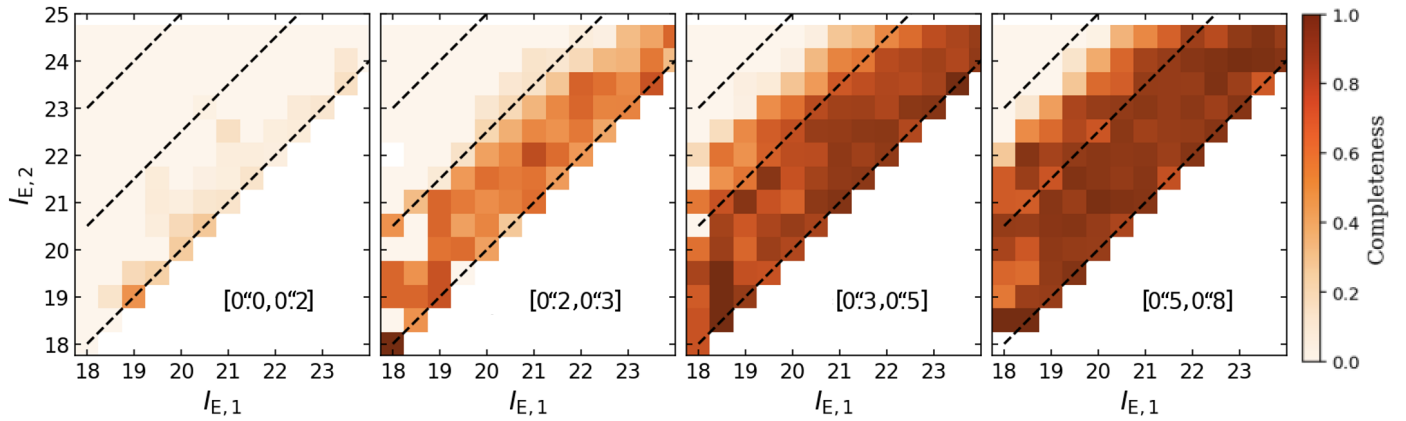


Fig. 10. Completeness in bins of $I_{E,1}$ and $I_{E,2}$ for different separation ranges. The black dashed lines represent flux ratios of 1, 10, and 100 between the primary and secondary sources.

3.4.1. Investigation of false positives

We investigated the behaviour of the CNN in relation to false positives, that are sources labelled as single but predicted to be class 1 or 2 by the network. Figure 8 shows the CNN predictions for $\Delta I_{E,21}$ and separation for the 337 false positives identified out of a total of 15 969 sources. Some of these FPs represent foreground contamination that may be present because we used stamps of real galaxies in which there is a non-zero probability that a foreground source is present. Another kind of FPs are galaxies with disturbed morphology, that may be misclassified. To better isolate probable FPs, we applied a posteriori physical threshold, selecting only those with $\Delta I_{E,21} < 4$ and separation $< 0''.15$. This additional cut allows us to eliminate more than half of the FPs though at the cost of also removing some TP. This choice is justified by physical considerations and helps to improve the overall robustness of the classification.

We studied the classification accuracy as a function of only the parameters available for single AGN: the magnitude of the primary source and that of the host galaxy. As shown in Fig. 9, the overall precision decreases when the host galaxy is particularly bright although it remains high (i.e., produces a low FP rate) with the chosen threshold.

3.5. Detection accuracy

We investigate how the completeness varies as a function of physical parameters, specifically the magnitudes of the primary and secondary components and their separation, using the selected thresholds. We apply the CNN with a classification threshold of 0.9, a minimum separation of $0''.15$, and $\Delta I_{E,21} < 4$ as described in Sect. 3.4.1.

Figure 10 shows the resulting completeness across this parameter space. At separations between $0''.2$ and $0''.3$, the CNN achieves an average completeness of approximately 60% in classifying dual sources with a contrast of up to a factor of ten. For larger separations, dual objects with $\Delta I_{E,21} < 4$ are detected with very high completeness. However, at separations greater than $0''.5$ – $0''.6$, standard deblending algorithms, such as IRAFStarFinder, perform better in terms of detection accuracy, but as described in Sect. 3.1, they do not reach the performances of the CNN.

4. Application to Q1 data

Starting from the initial dataset of approximately 15 000 AGN cutouts, we applied a series of selection criteria to refine the sample. We restricted the I_E range of the primary source to 18–23.5 and $\text{ISOAREA} < 450$, where the ISOAREA is defined as the area (in number of pixels) above the detection threshold, in order to remove very extended sources, and removed duplicate entries. Moreover, we excluded any QSO whose coordinates did not correspond to a detected source within $0''.3$. After applying these cuts, we ended up with 5993 cutouts, a sample size that is sufficiently to robustly test our method and to provide a realistic estimate of the number of dual systems to be followed-up with spectroscopic observations.

4.1. Application of the CNN

We first preprocessed the images by normalising the pixel values between 0 and 1. Each image was then centred on the pixel with the maximum flux, following the same procedure used during the training of the CNN on the simulated data (see Sect. 2.5). Finally, we applied the neural network on the preprocessed images. We found 265 candidates classified as dual systems. Using the threshold defined in Sects. 3.4 and Sect. 3.4.1, we identified a sample of 49 objects that meet the selection criteria, representing the $\sim 0.8\%$ of the initial sample. In this selected candidates, given the high precision of the neural network, we aimed to minimise contamination from objects with bright host galaxy structures. The main contaminants are represented by foreground sources, such as stars, and it is therefore necessary to estimate the expected number of such contaminants. The distribution of the separations predicted by the neural network is shown in Fig. 11.

4.2. Foreground contamination

Quantifying contamination from foreground stars within our candidate pair sample is inherently challenging. In this section, we provide a statistical estimate of the expected number of chance alignments with foreground sources along the line of sight, compared to systems that are more likely to be physically associated with the primary source, such as dual or lensed AGN or clumpy structure in the galaxy.

To estimate the probability of contamination from unrelated sources, we performed a simple statistical test by making ran-

dom square cutouts across the analysed fields of view. Specifically, we extracted 20 000 random $2'' \times 2''$ cutouts and ran **SExtractor** on each of them. The size of the cutouts was chosen to avoid truncating sources at separation of around $1''$. The estimated number of contaminants depends on both the limiting magnitude and the separation from the center within which sources are classified as contaminants. For example, by restricting to $I_E < 25$ and separations $< 0''.8$, **SExtractor** finds 605 contaminants. When extending the limits to $I_E < 26$ and separations $< 1''$, the number increases to 1650 out of 20 000, corresponding to approximately 8% contamination. If we consider only those with **CLASS STAR** > 0.5 , which would represent the true contaminants of dual or lensed AGN, **SExtractor** identifies 82 (0.4%: 24 out of ~ 6000) and 175 (0.87%: 52 out of ~ 6000) contaminants in the two respective thresholds.

We also provide estimates of the contaminants using the official catalogue provided by the *Euclid* pipeline (MER catalogue). For each MER tile catalogue, we randomly generated a set of RA and Dec positions uniformly distributed within the boundaries defined by the minimum and maximum RA and Dec of the tile. For each generated coordinate, we associated the closest object in the catalogue. We repeated this procedure 100 000 times. Under the previous thresholds, we find $\sim 7\%$ contamination for $I_E < 26$ and separation $< 1''$, and $\sim 3\%$ for $I_E < 25$ and separation $< 0''.8$. When using **PROB POINT LIKE** (the estimation uses a “star probability cube”) in the MER we found only 169 sources out of 100 000 images (0.17%).

It is important to note that all of the contaminant sources identified in our statistical test are detectable by the source extraction algorithm, as this method does not depend on the presence of a primary object. However, the most important question for our analysis is: how many of these contaminants would actually be detected by our CNN? Another noteworthy aspect is that, while it is possible to make sharp cuts in magnitude and separations for the contaminants when using **SExtractor**, the dual candidates (in this case, foreground contaminants) found by the CNN do not necessarily meet the criteria for magnitude and separation. This is because the CNN was trained with normalised images between 0 and 1 and with secondary sources at separations up to $0''.8$. Thus, it is not possible to include a cut in magnitude of the secondary, and most likely, it is not certain that the network only detects sources up to $0''.8$. Moreover, the network has been trained to detect point-like secondary objects (secondary AGN or lensed QSO), so it is not clear a priori how the network behaves with extended contaminants (e.g., galaxies). To estimate how many contaminants we expect from the neural network, we added to the random cutouts a point source with a I_E between 18 and the I_E of the contaminant (if present) following the same distribution of our input QSO sample. We compare the number of contaminants estimated by **SExtractor** with the number of contaminants estimated by the CNN. The network finds in total 118 contaminants out of 20 000 images, so around 30 contaminants out of the initial sample of ~ 6000 QSOs. We note from Fig. B.1 in Appendix B that, as expected, the network tends to consider as contaminants mostly systems with point source contaminants.

In the real case, the presence of a galaxy host can significantly affect the detectability of a nearby secondary source, especially when the host is brighter than the companion. In particular, faint secondary sources are more easily detected when no host is present. As a result, the number of contaminants estimated from random cutouts tends to exceed the number of pair sources identified around primary QSOs in the Q1. This difference is due to the presence of host galaxy structure in real sys-

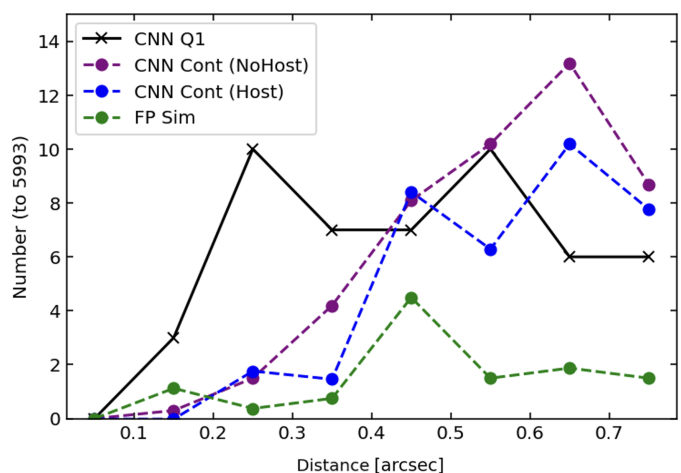


Fig. 11. Comparison between the number of dual candidates found by CNN and the expected contaminants. The number is normalized by the total number of AGN considered after the cuts (5993). Dual AGN candidates found by CNN in bins of separation (in black), number of expected contaminants classified dual by the CNN once we added a primary source without (purple) and with (blue) host galaxy, and false positives expected from simulations (in green).

tems, which lowers the efficiency of detecting nearby companions. We then included a host galaxy in the simulated cutouts, modelled with a Sérsic profile and a magnitude ranging between $I_{E,1} = -1$ to $I_{E,1} = +4$ and convolved with the PSF. When applying the CNN to these images, we find that the number of contaminants is reduced compared to the case without a host galaxy (blue line in Fig. 11).

Another important aspect involves the number of expected FPs from simulations. While the network has shown high precision overall, we observed that in some cases, single systems in simulations were misclassified as dual systems. To account for these different contaminants, we include the histogram of predicted distances for FPs in the simulations, shown in green in Fig. 11. In other words the green line represents the FPs that we expect from simulations. We note that the dual sources identified in the Q1 sample by the network with separations of $0''.2$ – $0''.4$ exceed the number expected from foreground objects and false positives caused by clumpy galaxy structure. This excess is likely indicative of a population of dual candidates (see Fig. 12). At larger separation ($> 0''.5$), the number of contaminants is consistent with the number of detected double sources. While this does not rule out the presence of dual AGN candidates at these higher separations, the probability of finding good candidates with secondary sources brighter than 25 mag is lower. As we address in Sect. 4.2.1, at larger separations it is possible to exploit the four *Euclid* bands to give a probability that the secondary source is an AGN.

4.2.1. Photometry and PSF subtraction

Euclid obtains images in four bands as described in Sect. 2.1, and the colours can in principle be used to study the nature of the sources. For systems with multiple entries in the MER catalogue, we can use MER fluxes to study the photometry with different apertures. To ensure reliable colour measurements, we implemented our own photometric pipeline, as accurate PSF subtraction is essential when analysing closely separated sources. While the MER catalogue provides photometric measurements using a range of aperture sizes, these are often affected by flux

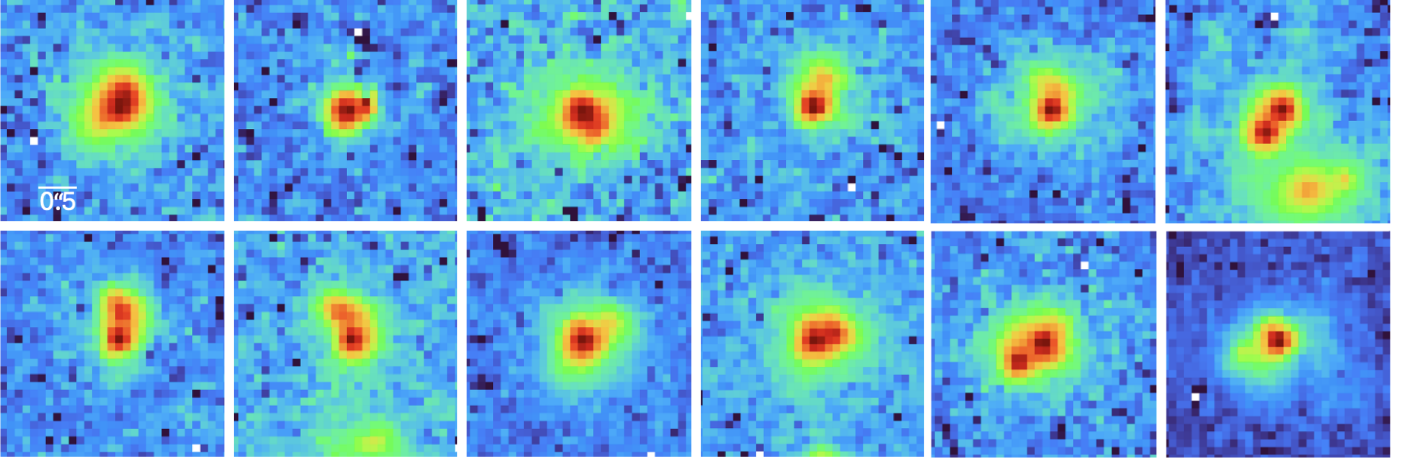


Fig. 12. Dual AGN candidates found by the network with separations between $0''.15$ and $0''.45$. The number of objects in this separation range exceeds the number expected from foreground objects and false positives caused by clumpy galaxy structure, making these candidates more reliable than those at larger separations. The size of each cutout is $3'' \times 3''$, and the flux is shown in logarithmic scale.

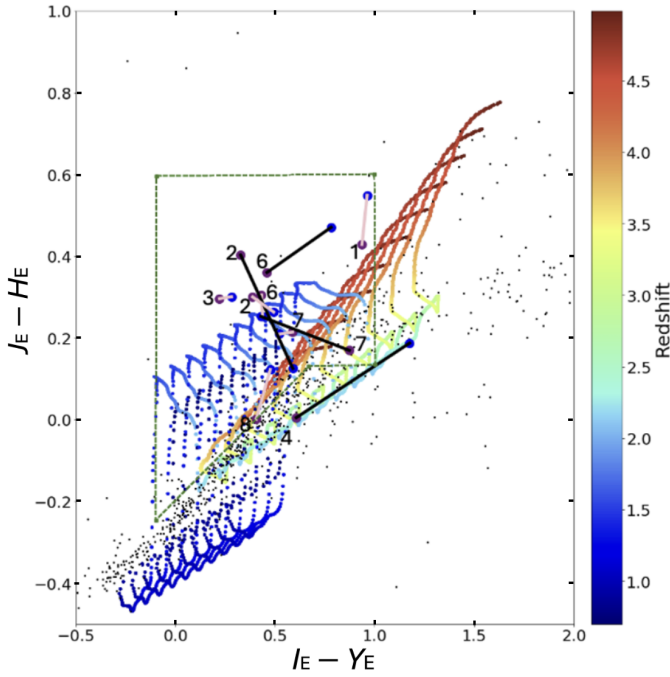


Fig. 13. Colour-colour diagram of dual AGN candidates in Fig. 14 overlaid with simulated AGN colour coded with redshift (Temple et al. 2021) and stars (black dots). The numbers correspond to the sources shown in Fig. 14. Each line connects the primary (purple dots) to the secondary (blue dot) source in a dual AGN candidate. Pink lines indicate photometry derived from MER catalogue using a 1 FWHM aperture, while black lines correspond to photometry obtained after PSF subtraction. The green dashed box indicates the AGN selection region defined in Euclid Collaboration: Matamoro Zatarain et al. (2025).

contamination of close sources, which is the case for our candidates. To perform the photometry on the candidate sample, we implemented a customized PSF-fitting algorithm for measuring the photometry of the blended sources. We first used `FittableImageModel` of `photutils.psf` on the VIS PSF of the MER catalogue mosaic at the source position. Then we used `IRAFStarFinder` on the VIS stamp, which has higher resolution, to search for point sources. Once we found the positions of the two sources, we measured the fluxes in the VIS and NISP

bands by using the appropriate PSF (from MER) and fixing the positions to within 1 pixel. We also included in the fit a Sérsic component convolved with the PSF of the corresponding band to model the host galaxy. The effective radius (r_{eff}) was allowed to vary between 2 and 3 pixels, the Sérsic index (n) between 2 and 4, the ellipticity in the range 0 and 1, and the position within a radius of 1 pixel from the center. We then summed the fluxes of each pixel in the model and using the ZERO MAG, we computed the magnitude of the two sources and of the host galaxy in each band.

4.2.2. Colours of the companion candidates

Figure 13 shows a colour-colour diagram ($J_E - H_E$ vs. $I_E - Y_E$) used to analyze the photometric properties of dual AGN candidates. The coloured points represent simulated AGN generated using the `qsogen` package from Temple et al. (2021), with redshift varying between 0.7 and 5 and extinction $E(B - V)$ ranging from 0 to 0.1. Black points correspond to stellar templates from the XSL_DR3_release catalogue (Verro et al. 2022), included to trace the stellar region in colour space. The green dashed box indicates the AGN selection region defined in Euclid Collaboration: Matamoro Zatarain et al. (2025), which is used to separate AGN from stars. For each dual AGN candidate pair, we compute and plot the colours of both the primary and secondary components. We measured the colours of all source pairs for which reliable photometry could be performed for all bands. In general, for dual sources with separations $> 0''.5$, we obtain good fits with errors on the flux of the two sources about 20%, which corresponds to an error on the magnitude of ~ 0.2 mag. To assess the consistency of our photometric method, we compared the colours derived from our PSF subtraction technique with those listed in the MER catalogue, where the two sources are recorded as distinct entries. Figure 14 presents seven source pairs in which both components fall within the AGN colour region in at least one of the two photometric methods (PSF subtraction or MER catalogue) and one source (J0420 – 4808) where the colours of both sources are consistent with being at the same redshift ($z \sim 2.5$). Such a comparison is only feasible for pairs with angular separations larger than $0''.5$. Only three dual systems exhibit both components within the AGN region (red box in Fig. 14), consistent with previous findings suggesting that most widely sepa-

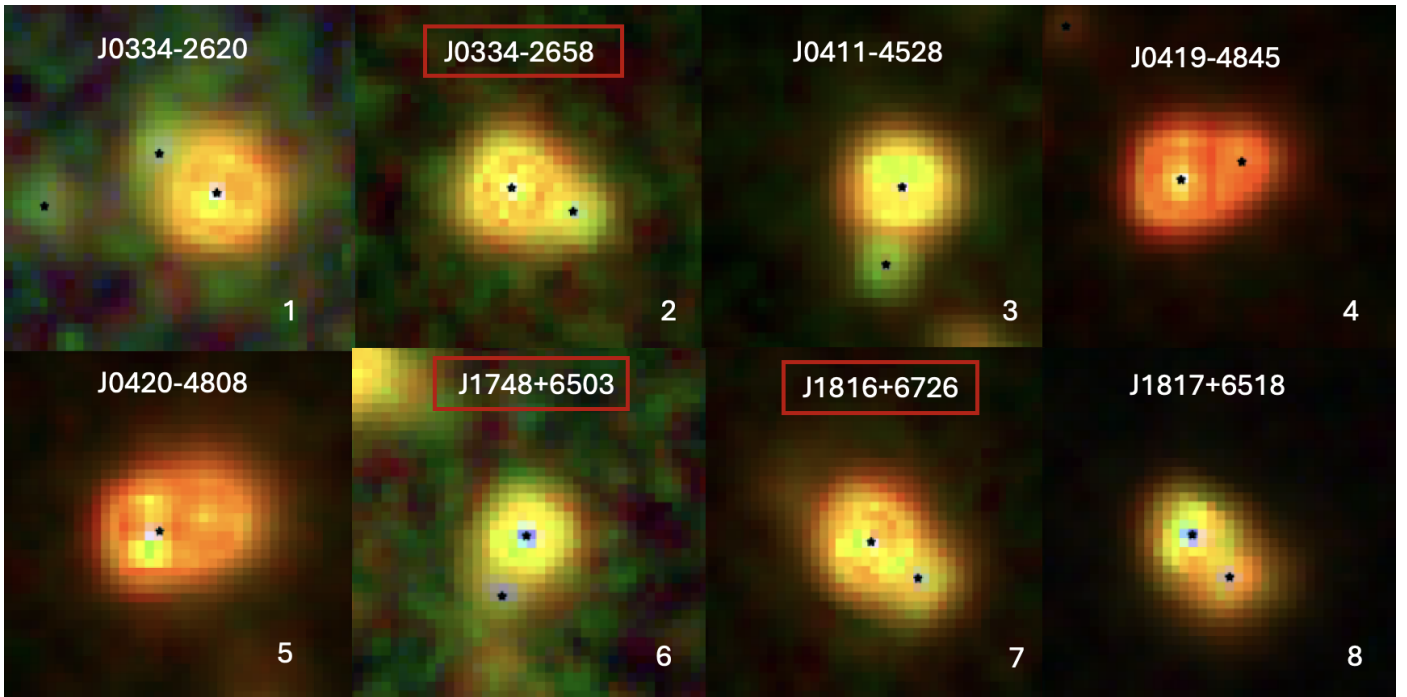


Fig. 14. Three-colour composite images (I_E , J_E , H_E) of dual AGN candidates derived from MER photometry or PSF subtraction. Systems enclosed in red boxes are those where both sources lie within the AGN region in both photometric methods. The black stars mark the positions of the input sources from the MER catalogue. The size of each cutout is $3'' \times 3''$.

rated pairs are likely due to line of sight (LOS) contaminations rather than physically associated AGN.

4.2.3. Visual inspection

We conducted a visual inspection of the QSO sample input to the network, categorising objects as follow:

0. Clear dual objects: These are well-separated pairs where both components are distinctly visible and compatible with being point sources.
1. Possible duals at low separations: Objects that could be dual sources but are at low separation ($< 0''.25$) or with asymmetric morphology, making identification less certain.
2. Contaminated objects: These include secondary sources that are not point-like sources or that are point-like but have a separation greater than $0''.8$.
3. Possible lens-like structures: Sources that exhibit morphological features like gravitational lensing effects such as arcs.
4. Single objects: Sources with no secondary components visible.
5. Objects with a bright host: QSOs where the presence of a luminous host galaxy that could affect the classification.

Out of the 5993 target in the input sample, we found 51 clear dual objects (class 0), 40 possible dual objects at low separation (class 1), about 340 classified as contaminations, 8 possible lenses, 5208 singles, and 336 associated with bright hosts. We compared the dual AGN candidates found by the CNN with the visual inspection. Thirty-two candidates belong to class 0 or 1, 8 belong to class 2, two are possible lenses (see Fig. C.1 in Appendix C), and 7 belong to class 4 or 5. We note that visual inspection was carried out in log scale on images normalized between 0 and 1, so we classified as dual objects with a very weak secondary when the network was trained to detect secondaries as faint as $I_E = 25$. Table 2 shows the number of TPs

and FPs based on visual inspection identified by the CNN across different threshold values. TPs and FPs refer to the class identified during the visual inspection (P: class 0, 1; N: other classes). The configuration adopted in this work is the first one which, as suggested by simulations, maximizes the difference between the TPs and FPs. For comparison, the other two configurations maximize TP/FP and the number of TPs but significantly increasing the number of FPs. We also applied visual inspection to simulations, finding that we were able to correctly classify only 4% of simulated dual objects between $0''.15$ and $0''.2$, while between $0''.25$ and $0''.3$ we classify correctly 60% of objects.

4.3. Additional space-based imaging

To test our results, we searched for higher resolution images of our dual AGN candidates obtained with facilities such as JWST and HST. In this case, we applied the network without using any threshold in ISOAREA and we defined as dual candidates all the objects for which the sum of the predicted probabilities for labels 1 and 2 is higher than 0.5 (lower than the threshold we used in Sect. 3.4). This approach allows us to identify cases that may represent true positives which were excluded by the CNN due to a too conservative threshold. Moreover, it provides a way to investigate potential FPs and understand the sources of misclassification and to assess whether the observed number of FPs is consistent with expectations from simulations. However, this comparison relies on the assumption that the simulated distributions of separations and the magnitudes of the primary, secondary, and host components accurately reflect the underlying real distributions, which may not always be the case. We confirm that when the thresholds are not applied the network does not always select sources with point-like secondary components in the data, as expected. Two sources with high resolution images were excluded from the initial selection due to ISOAREA values higher than 450 (both having known redshifts below 0.6) though they

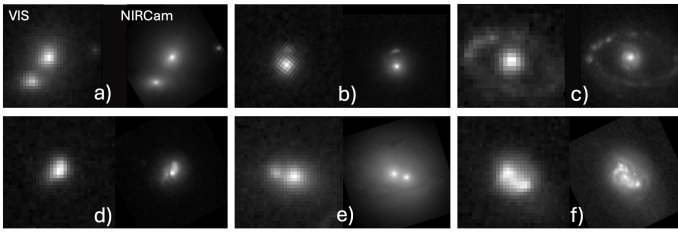


Fig. 15. VIS and NIRCam/JWST images of objects predicted dual without considering thresholds.

show clear indications of dual morphology (a and e in Fig. 15). In one case, we identify an arc-like feature that could be both a foreground lensed galaxy or a tidal feature (b). In two other cases, the network misclassifies the sources, since it is confused by the clumpiness and the irregular structure of the host galaxy (c, f). Another object presents a more ambiguous scenario (d, shown either in Fig. 12): it may be a dual AGN with a separation of $\sim 0''.2$, or alternatively a single AGN whose nuclear emission is covered by a dust lane, creating two distinct components. These examples illustrate some of the limitations of the network when dealing with complex morphologies and highlight the importance of high-resolution imaging for proper classification.

4.4. MER catalogue

Working directly on images and building a network made specifically for the dual AGN search is of paramount importance because in many cases the MER catalogue merges two objects with small separation into a single catalogue entry. It is interesting to understand which is the minimum distance between two point sources that allows the MER catalogue to distinguish the two objects and thus list them as two separate entries. To answer this, we calculated for each source in the catalogue the distance to the nearest source and studied in detail the pairs of sources with separations less than $0''.8$. In particular, we studied the minimum distance as a function of the I_E and Kron radius. We found that the MER catalogue is not effective for binary sources at distances less than $0''.8$ and with flux ratios > 2 . We checked which dual candidates predicted by the neural network are associated with two different entries in the MER catalogue. Of the 32 dual visual inspected sources of class 0 or 1, only one third have double entries in the MER catalogue.

4.5. Summary of the CNN application to real data

We found 265 candidates classified as dual systems by the CNN. Using the threshold defined in the previous Section, we identified a sample of 49 objects that fulfil the selection criteria, representing 0.8% of the initial sample. Of these 49 objects, 32 were classified according to the visual inspection as class 0 or 1, eight as class 2, two as class 3, and seven as class 4 or 5, as defined in Sect. 4.2.3. We showed that most systems with projected separations larger than $0''.5$ are likely to be LOS contaminants. However, *Euclid*-band photometry allows us to further investigate the nature of these systems. Among the objects with separations greater than $0''.5$ and available MER photometry, seven systems have both the primary and the secondary components falling within the AGN selection region defined by *Euclid* Collaboration: Matamoros Zatarain et al. (2025). After performing PSF subtraction and adding a host galaxy component when required, we found that only three of these systems lie within the AGN region. In addition to these three, the most reliable candi-

dates are those with smaller separations. These selected systems are reported in Fig. 12. All the dual candidates we have found, once we excluded foreground contaminants, represent around the 0.25% of the initial sample.

5. Discussion

Estimating the true fraction of dual AGN systems remains highly challenging, even when follow-up observations are available. Providing a robust estimate of the dual fraction is beyond the scope of this work. We note that the definition of dual fraction is not always the same in all works, and different definitions may lead to underestimation or overestimation. In this section, we explore how our results compare with cosmological simulations, particularly the HORIZON-AGN simulation (Volonteri et al. 2022), and with previous studies.

5.1. Comparison with simulations

The dual fraction, i.e., the number of multiple AGN over the total number of AGN, is one of the main predictions of the models of SMBH formation and BH-galaxy coevolution. Our results are still too preliminary to test these predictions, but it is interesting to compare the number of candidates we select with the results of these predictions.

The Horizon-AGN simulation (Volonteri et al. 2022) predicts a dual fraction of 5% when considering AGN with bolometric luminosities $L_{\text{bol}} > 10^{44} \text{ erg s}^{-1}$ and projected separations less than 30 kpc. From this we can obtain an order-of-magnitude estimate of the expected number of dual AGN in our sample by assuming that the dual fraction does not depend on redshift and bolometric luminosity. This assumption is very simplistic and known not to be true in most cases (Comerford et al. 2015; De Rosa et al. 2019), but can be used to obtain a rough estimate of the expected numbers. Given our parent sample of ~ 6000 AGN, this implies an expected number of ~ 300 dual AGN systems with the same selections. In this study, we focused on small projected separation ($< 0''.8$). This threshold excludes about 80% of the dual AGN population predicted by simulations, significantly reducing our expected number of detections. Moreover, our selection requires optical magnitude $I_E < 25$, which limits our sample to $L_{\text{bol}} > 10^{44} \text{ erg s}^{-1}$ at $z = 1$ and $L_{\text{bol}} > 10^{45} \text{ erg s}^{-1}$ at $z = 3$. This corresponds to only $\sim 1.8\%$ of the dual AGN population found in Volonteri et al. (2022). As a result, we expect to identify 5–6 dual AGN within our sample, at redshifts up to $z \sim 3$. It is important to note, however, that the Horizon-AGN simulations do not resolve dual AGN at physical separations below ~ 4 kpc, corresponding to $0''.5$. This limitation implies that our estimate is a lower bound, particularly at small separations where the majority of dual AGN are likely to reside. In fact, the most reliable candidates found in this work are at separations between $0''.2$ and $0''.45$.

This range of separations is sampled by ROMULUS25 simulations (Tremmel et al. 2017, 2019) presented in Saeedzadeh et al. (2024, Fig. 7), Chen et al. (2023, Astrid), and Steinborn et al. (2016, Magneticum), which show that dual AGN dominate at small separations. Rosas-Guevara et al. (2019) showed that there is not a clear trend in the distribution at small distances (< 15 kpc), while the dual AGN fraction increases around 20 and 25 kpc.

Table 2. Performance metrics for different threshold configurations which maximise TP-FP (first row), TP/FP (second row) and TP (third row). TP and FP refer to the class identified during the visual inspection (P: class 0, 1; N: other classes).

Threshold	Sep	$\Delta I_{E,21}$	$\Delta I_{E,H1}$	TP	FP	TP – FP	TP/FP	TPR (%)	FPR (%)
> 0.90	> 0''.15	< 4	> -2	32	7	25	4.5	35	0.12
> 0.80	> 0''.20	< 4	> 1	20	3	17	6.7	22	0.05
> 0.5	> 0''.0	< 5	> -2	48	141	-93	0.34	53	2.5

Table 3. Sample Selection

Category	Number of objects
QSO catalogue	15317
Selection and avoiding repetition	5993
D/L AGN (CNN w/o threshold)	265
D/L AGN (CNN w/ threshold)	49
" foreground contamination	~ 35
" D/L AGN cand. > 0''.45 (phot)	3
" D/L AGN cand. < 0''.45	~ 12
" lens cand.	2

5.2. Comparison with other works

Numerous studies have searched for dual AGNs using different selection techniques, including imaging and optical spectroscopy (Hennawi et al. 2006; More et al. 2016; Eftekharzadeh et al. 2017; Mannucci et al. 2022; Scialpi et al. 2024; Jiang et al. 2025), X-ray observations (Koss et al. 2012; Gross et al. 2019; Guainazzi et al. 2021; Sandoval et al. 2023) and radio surveys (Fu et al. 2015; Rubinur et al. 2018; Gross et al. 2019). Each of these approaches, however, is subject to selection biases. A major observational bias arises from the fact that dual AGNs in the late stages of galaxy mergers are often heavily obscured by dust (Ricci et al. 2017; Koss et al. 2018; Barrows et al. 2023). Another important limitation is due to the sensitivity and the angular resolution of the instruments employed, which probe different dual AGN populations and different physical scales. As shown in Fig. 24 of Pfeifle et al. (2024), dual AGN candidates reported in the literature are scarce at both high redshift ($z > 0.5$) and small separations (< 5 kpc). These unexplored regions of parameter space can be effectively probed thanks to the capabilities of *Euclid*. The dual AGN fractions reported in previous works are not directly comparable to our estimation, as they depend strongly on the adopted selection criteria as well as on the properties of the systems (e.g., separation, luminosity), but a comparison can still be instructive.

Barrows et al. (2023) reported a dual AGN fractions between 0.2% and 1%, depending on bolometric luminosity and separation, within the redshift range $0.08 < z < 0.58$ and projected separations ranging between 10 and 70 kpc. These values correspond to angular separations between $\sim 1''$ and $\sim 0''.6$, i.e. an order of magnitude larger than the physical scales probed in the present work. Li et al. (2024) estimated that the pair fraction increases with redshift, starting at approximately 4.5% at $z \sim 0.5$ and reaching around 22.9% at $z \sim 4.5$. This fraction drops to below 1% when considering only objects with projected separations smaller than $0''.8$ as in this work. They found the distribution of projected separations to be relatively flat, but did not identify any pairs with separations smaller than ~ 2 kpc.

Recently Silverman et al. (2020) found a dual QSO fraction of $(0.26 \pm 0.18)\%$ from $z = 3$ to $z = 1.5$ with separations of $0''.6$ and $4''$. Yu et al. (2011) found that 0.02%–0.06% of AGN are dual AGN with double-peaked narrow line features at redshifts

of $z \sim 0.5$ –1.2. Rosario et al. (2011) estimated a dual fraction between 0.32% and 0.5%, similar to Fu et al. (2011). Koss et al. (2012) estimated that, among AGN in the *Swift* BAT sample, the fraction of dual AGN at projected separations of less than 15 kpc is approximately 5% (8 out of 167 systems). In contrast, the corresponding dual AGN fraction among Seyfert galaxies in the SDSS is significantly lower, at only 0.25%. These estimates are all broadly comparable with our results.

5.3. Other architectures

We explored the use of Domain-Adversarial Neural Networks (DANN) for domain adaptation, a technique designed to help models generalize when the target domain data (i.e. the real data) is entirely unlabeled, as in our case (Ganin et al. 2015). The goal of DANN is to learn a mapping between domains such that the features extracted by the network are both discriminative for the task and invariant to domain shifts. This means that the features should have similar distributions in both the source (simulations) and target (data) domains, allowing the model to generalise well to the target domain (Ajakan et al. 2014; Huertas-Company et al. 2024; Belfiore et al. 2025; Ginolfi et al. 2025). However, our attempt did not provide any improvement on the real data, though the network was able to generalize the features of data and simulations. Minimising the discriminator, we found no improvement in terms of reducing FPs labelled with the visual inspection.

We also used one of the most popular architectures, Resnet18 (He et al. 2016), by adding the last layers for the prediction of system properties as described in Sect. 2.5. Although it achieved a training set accuracy of 99%, which was not achieved by the other networks, its performance on the test set and on the real data was worse than the network we used, producing an order of magnitude more FPs compared to our CNN, probably due to overfitting.

We also tried combining different simulations, increasing the size of the training set, but the performance remained similar.

6. Conclusions

In this work, we took advantage of the *Euclid* data release Q1, which, thanks to its wide area, high sensitivity and resolution, provides a unique opportunity to identify close AGN pairs. We developed a CNN trained on realistic simulations to search for dual or lensed AGN at sub-arcsec separations ($< 0''.8$). The CNN performs significantly better than classical detection methods, especially in the regime of small separations, and maintains a high completeness while finding a limited number of FPs. Applying our model to a sample of ~ 6000 QSOs, we identified a first set of promising dual/lensed AGN candidates. Once we estimated the number of foreground contaminants and expected FPs from the CNN, and identified two possible lens candidates, we found a dual fraction of $\sim 0.25\%$, with an overdensity of these systems at low separations (below $0''.45$) where contamination of foreground objects is lower compared to high separations. For sources at larger separations ($> 0''.45$) we performed a photo-

metric analysis to study colours relative to [Euclid Collaboration: Matamoro Zatarain et al. \(2025\)](#) diagrams and found that three objects could be compatible with being dual AGN, while most of the systems are most likely foreground contaminants. While our CNN demonstrates promising performance for future *Euclid* release, further progress requires refining the simulated training set, such as the exploitation of all four *Euclid* bands into the machine learning framework. This will not only improve the prediction of the network by providing more realistic examples, but also allow the CNN to estimate the probability that a secondary object is a foreground contaminant, such as a star, or another AGN. Larger samples and the full depth and coverage of future *Euclid* releases, such as DR1, should be used to test the scalability of the CNN. The main bottleneck remains in the visual inspection stage, which will not be feasible for future releases. However, because our CNN has been designed to maximise the number of TPs compared to FPs, and given that the expected fraction of dual AGN is relatively small ([Rosas-Guevara et al. 2019](#); [Volonteri et al. 2022](#)), the number of candidates should be manageable for visual inspection. Spectroscopic follow-up and higher-resolution imaging will be critical to confirm the nature of candidates, especially at sub-arcsec separations. Furthermore, future *Euclid* releases, complemented by data from *JWST*, *LSST*, and *Roman*, will be key to achieving a statistical census of dual AGN and constraining models of SMBH growth and galaxy evolution.

Acknowledgements. The Euclid Consortium acknowledges the European Space Agency and a number of agencies and institutes that have supported the development of *Euclid*, in particular the Agenzia Spaziale Italiana, the Austrian Forschungsförderungsgesellschaft funded through BMK, the Belgian Science Policy, the Canadian Euclid Consortium, the Deutsches Zentrum für Luft- und Raumfahrt, the DTU Space and the Niels Bohr Institute in Denmark, the French Centre National d'Etudes Spatiales, the Fundação para a Ciência e a Tecnologia, the Hungarian Academy of Sciences, the Ministerio de Ciencia, Innovación y Universidades, the National Aeronautics and Space Administration, the National Astronomical Observatory of Japan, the Nederlandse Onderzoekschool Voor Astronomie, the Norwegian Space Agency, the Research Council of Finland, the Romanian Space Agency, the State Secretariat for Education, Research, and Innovation (SERI) at the Swiss Space Office (SSO), and the United Kingdom Space Agency. A complete and detailed list is available on the *Euclid* web site (www.euclid-ec.org). This work has made use of the *Euclid* Quick Release Q1 data from the *Euclid* mission of the European Space Agency (ESA), 2025, <https://doi.org/10.57780/esa-2853f3b>. This publication was produced while attending the PhD program in PhD in Space Science and Technology at the University of Trento, Cycle XXXIX, with the support of a scholarship financed by the Ministerial Decree no. 118 of 2nd march 2023, based on the NRRP – funded by the European Union – NextGenerationEU – Mission 4 "Education and Research", Component 1 "Enhancement of the offer of educational services: from nurseries to universities" – Investment 4.1 "Extension of the number of research doctorates and innovative doctorates for public administration and cultural heritage". We acknowledge financial contribution from INAF Large Grant "Dual and binary supermassive black holes in the multi-messenger era: from galaxy mergers to gravitational waves" (Bando Ricerca Fondamentale INAF 2022), from the INAF project "VLT-MOONS" CRAM 1.05.03.07, from the French National Research Agency (grant ANR-21-CE31-0026, project MBH_waves), from the INAF Large Grant 2022 "The metal circle: a new sharp view of the baryon cycle up to Cosmic Dawn with the latest generation IFU facilities". This work has been partially financed by the European Union with the Next Generation EU plan, through PRIN-MUR project "PROMETEUS" (202223XPZM). A.F. acknowledges INAF Mini Grant 2024 "The pc-scale view of HII regions in M33". EB acknowledges funding through the INAF GO grant "A JWST/MIRI MIRACLE: Mid-IR Activity of Circumnuclear Line Emission" and the "Ricerca Fondamentale 2024" program (mini-grant 1.05.24.07.01). This research used the facilities of the Italian Center for Astronomical Archive (IA2) operated by INAF at the Astronomical Observatory of Trieste. FR acknowledges the support from the INAF Large Grant "AGN & *Euclid*: a close entanglement" Ob. Fu. 01.05.23.01.14.

References

- Acevedo Barroso, J. A., O'Riordan, C. M., Clément, B., et al. 2025, *A&A*, 697, A14
- Ajakan, H., Germain, P., Larochelle, H., Laviolette, F., & Marchand, M. 2014, arXiv e-prints, arXiv:1412.4446
- Assef, R. J., Stern, D., Noirot, G., et al. 2018, *ApJS*, 234, 23
- Barrows, R. S., Comerford, J. M., Stern, D., & Assef, R. J. 2023, *ApJ*, 951, 92
- Belfiore, F., Ginolfi, M., Blanc, G., et al. 2025, *A&A*, 694, A212
- Bennett, C. L., Larson, D., Weiland, J. L., & Hinshaw, G. 2014, *ApJ*, 794, 135
- Bertin, E. & Arnouts, S. 1996, *A&AS*, 117, 393
- Blecha, L. E. 2017, Modeling Dual Active Galactic Nuclei in Mergers: Unique Probes of Black Hole Inspiral and Growth, NSF Award Number 1715413, Directorate for Mathematical and Physical Sciences, Division Of Astronomical Sciences. 2017.
- Bradley, A. P. 1997, *Pattern Recognition*, 30, 1145
- Bradley, L., Sipőcz, B., Robitaille, T., et al. 2021, *astropy/photutils*: 1.3.0
- Capelo, P. R., Dotti, M., Volonteri, M., et al. 2017, *MNRAS*, 469, 4437
- Capelo, P. R., Volonteri, M., Dotti, M., et al. 2015, *MNRAS*, 447, 2123
- Chaussidon, E., Yèche, C., Palanque-Delabrouille, N., et al. 2023, *ApJ*, 944, 107
- Chen, N., Di Matteo, T., Ni, Y., et al. 2023, *MNRAS*, 522, 1895
- Chen, Y.-C., Hwang, H.-C., Shen, Y., et al. 2022, *ApJ*, 925, 162
- Ciurlo, A., Mannucci, F., Yeh, S., et al. 2023, *A&A*, 671, L4
- Comerford, J. M., Pooley, D., Barrows, R. S., et al. 2015, *ApJ*, 806, 219
- Crawshaw, M. 2020, arXiv e-prints, arXiv:2009.09796
- Croom, S. M., Smith, R. J., Boyle, B. J., et al. 2004, *MNRAS*, 349, 1397
- De Rosa, A., Vignali, C., Bogdanović, T., et al. 2019, *New A Rev.*, 86, 101525
- Eftekharzadeh, S., Myers, A. D., Hennawi, J. F., et al. 2017, *MNRAS*, 468, 77
- Euclid Collaboration: Aussel, H., Tereno, I., Schirmer, M., et al. 2025, *A&A*, submitted (Euclid Q1 SI), arXiv:2503.15302
- Euclid Collaboration: Cropper, M., Al-Bahlawan, A., Amiaux, J., et al. 2025, *A&A*, 697, A2
- Euclid Collaboration: Jahnke, K., Gillard, W., Schirmer, M., et al. 2025, *A&A*, 697, A3
- Euclid Collaboration: Lines, N. E. P., Collett, T. E., Walmsley, M., et al. 2025, *A&A*, in press (Euclid Q1 SI), <https://doi.org/10.1051/0004-6361/202554542>, arXiv:2503.15326
- Euclid Collaboration: Matamoro Zatarain, T., Fotopoulou, S., Ricci, F., et al. 2025, *A&A*, submitted (Euclid Q1 SI), arXiv:2503.15320
- Euclid Collaboration: Mellier, Y., Abdurro'uf, Acevedo Barroso, J., et al. 2025, *A&A*, 697, A1
- Euclid Collaboration: Rojas, K., Collett, T. E., Acevedo Barroso, J. A., et al. 2025, *A&A*, submitted (Euclid Q1 SI), arXiv:2503.15325
- Euclid Collaboration: Romelli, E., Kümmel, M., Dole, H., et al. 2025, *A&A*, in press (Euclid Q1 SI), <https://doi.org/10.1051/0004-6361/202554586>, arXiv:2503.15305
- Euclid Collaboration: Scaramella, R., Amiaux, J., Mellier, Y., et al. 2022, *A&A*, 662, A112
- Euclid Collaboration: Siudek, M., Huertas-Company, M., Smith, M., et al. 2025, *A&A*, accepted (Euclid Q1 SI), arXiv:2503.15312
- Euclid Collaboration: Walmsley, M., Holloway, P., Lines, N. E. P., et al. 2025, *A&A*, submitted (Euclid Q1 SI), arXiv:2503.15324
- Euclid Quick Release Q1. 2025, <https://doi.org/10.57780/esa-2853f3b>
- Flesch, E. W. 2023, *The Open Journal of Astrophysics*, 6, 49
- Fu, H., Myers, A. D., Djorgovski, S. G., & Yan, L. 2011, *ApJ*, 733, 103
- Fu, H., Myers, A. D., Djorgovski, S. G., et al. 2015, *ApJ*, 799, 72
- Gaia Collaboration: Prusti, T., de Bruijne, J. H. J., et al. 2016, *A&A*, 595, A1
- Ganin, Y., Ustinova, E., Ajakan, H., et al. 2015, arXiv e-prints, arXiv:1505.07818
- Ginolfi, M., Mannucci, F., Belfiore, F., et al. 2025, *A&A*, 693, A73
- Glikman, E., Langglin, R., Johnstone, M. A., et al. 2023, *ApJ*, 951, L18
- Gonzaga, S., Hack, W., Fruchter, A., & Mack, J. 2012, *The DrizzlePac Handbook*
- Gross, A., Fu, H., Djorgovski, S. G., et al. 2019, in *American Astronomical Society Meeting Abstracts*, Vol. 233, American Astronomical Society Meeting Abstracts #233, 243.26
- Guainazzi, M., De Rosa, A., Bianchi, S., et al. 2021, *MNRAS*, 504, 393
- He, K., Zhang, X., Ren, S., & Sun, J. 2016, in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1
- Hennawi, J. F., Strauss, M. A., Oguri, M., et al. 2006, *AJ*, 131, 1
- Hernquist, L. 1989, *Nature*, 340, 687
- Hopkins, P. F., Hernquist, L., Cox, T. J., & Kereš, D. 2008, *ApJS*, 175, 356
- Huertas-Company, M., Iyer, K. G., Angeloudi, E., et al. 2024, *A&A*, 685, A48
- Jiang, Y., Shen, Y., Liu, X., et al. 2025, arXiv e-prints, arXiv:2504.06415
- Jin, J.-J., Wu, X.-B., Fu, Y., et al. 2023, *ApJS*, 265, 25
- Khan, F. M., Preto, M., Berczik, P., et al. 2012, *ApJ*, 749, 147
- Kingma, D. P. & Ba, J. 2014, arXiv e-prints, arXiv:1412.6980
- Koss, M., Mushotzky, R., Treister, E., et al. 2012, *ApJ*, 746, L22
- Koss, M. J., Blecha, L., Bernhard, P., et al. 2018, *Nature*, 563, 214
- Lemon, C. A., Auger, M. W., McMahon, R. G., & Kposov, S. E. 2017, *MNRAS*, 472, 5023
- Li, J., Zhuang, M.-Y., Shen, Y., et al. 2024, arXiv e-prints, arXiv:2405.14980

- Lyke, B. W., Higley, A. N., McLane, J. N., et al. 2020, *ApJS*, 250, 8
- Mannucci, F., Pancino, E., Belfiore, F., et al. 2022, *Nature Astronomy*, 6, 1185
- Mannucci, F., Scialpi, M., Ciurlo, A., et al. 2023, *A&A*, 680, A53
- Merloni, A., Lamer, G., Liu, T., et al. 2024, *A&A*, 682, A34
- More, A., Oguri, M., Kayo, I., et al. 2016, *MNRAS*, 456, 1595
- Nagam, B. C., Acevedo Barroso, J. A., Wilde, J., et al. 2025, *A&A*, submitted, arXiv:2502.09802
- Paske, A., Gross, S., Massa, F., et al. 2019, arXiv e-prints, arXiv:1912.01703
- Pearce-Casey, R., Nagam, B. C., Wilde, J., et al. 2025, *A&A*, 696, A214
- Perna, M., Arribas, S., Lamperti, I., et al. 2025, *A&A*, 696, A59
- Pfeifle, R. W., Weaver, K. A., Secrest, N. J., Rothberg, B., & Patton, D. R. 2024, arXiv e-prints, arXiv:2411.12799
- Puerto-Sánchez, C., Habouzit, M., Volonteri, M., et al. 2025, *MNRAS*, 536, 3016
- Rakshit, S., Stalin, C. S., & Kotilainen, J. 2020, *ApJS*, 249, 17
- Ricci, C., Bauer, F. E., Treister, E., et al. 2017, *MNRAS*, 468, 1273
- Rosario, D. J., McGurk, R. C., Max, C. E., et al. 2011, *ApJ*, 739, 44
- Rosas-Guevara, Y. M., Bower, R. G., McAlpine, S., Bonoli, S., & Tissera, P. B. 2019, *MNRAS*, 483, 2712
- Rubín, K., Das, M., & Kharb, P. 2018, *Journal of Astrophysics and Astronomy*, 39, 8
- Ruder, S. 2017, arXiv e-prints, arXiv:1706.05098
- Saeedzadeh, V., Babul, A., Mukherjee, S., et al. 2024, *ApJ*, 975, 265
- Sandoval, B., Foord, A., Allen, S. W., et al. 2023, arXiv e-prints, arXiv:2312.02311
- Schramm, M. & Silverman, J. D. 2013, *ApJ*, 767, 13
- Scialpi, M., Mannucci, F., Marconini, C., et al. 2024, *A&A*, 690, A57
- Shu, Y., Kopylov, S. E., Evans, N. W., et al. 2019, *MNRAS*, 489, 4741
- Silverman, J. D., Tang, S., Lee, K.-G., et al. 2020, *ApJ*, 899, 154
- Steinborn, L. K., Dolag, K., Comerford, J. M., et al. 2016, *MNRAS*, 458, 1013
- Stern, D., Assef, R. J., Benford, D. J., et al. 2012, *ApJ*, 753, 30
- Storey-Fisher, K., Hogg, D. W., Rix, H.-W., et al. 2024, *ApJ*, 964, 69
- Temple, M. J., Hewett, P. C., & Banerji, M. 2021, *MNRAS*, 508, 737
- Tremmel, M., Karcher, M., Governato, F., et al. 2017, *MNRAS*, 470, 1121
- Tremmel, M., Quinn, T. R., Ricarte, A., et al. 2019, *MNRAS*, 483, 3336
- Verro, K., Trager, S. C., Peletier, R. F., et al. 2022, *A&A*, 660, A34
- Volonteri, M., Pfister, H., Beckmann, R., et al. 2022, *MNRAS*, 514, 640
- Yu, Q., Lu, Y., Mohayaee, R., & Colin, J. 2011, *ApJ*, 738, 92
- ¹ University of Trento, Via Sommarive 14, I-38123 Trento, Italy
- ² Dipartimento di Fisica e Astronomia, Università di Firenze, via G. Sansone 1, 50019 Sesto Fiorentino, Firenze, Italy
- ³ INAF-Osservatorio Astrofisico di Arcetri, Largo E. Fermi 5, 50125, Firenze, Italy
- ⁴ Department of Mathematics and Physics, Roma Tre University, Via della Vasca Navale 84, 00146 Rome, Italy
- ⁵ INAF-Osservatorio Astronomico di Roma, Via Frascati 33, 00078 Monteporzio Catone, Italy
- ⁶ STAR Institute, University of Liège, Quartier Agora, Allée du six Août 19c, 4000 Liège, Belgium
- ⁷ European Southern Observatory, Karl-Schwarzschild-Str. 2, 85748 Garching, Germany
- ⁸ Dipartimento di Fisica e Astronomia "Augusto Righi" - Alma Mater Studiorum Università di Bologna, via Piero Gobetti 93/2, 40129 Bologna, Italy
- ⁹ INAF-Osservatorio di Astrofisica e Scienza dello Spazio di Bologna, Via Piero Gobetti 93/3, 40129 Bologna, Italy
- ¹⁰ Centro de Astrobiología (CAB), CSIC-INTA, Cra. de Ajalvir Km. 4, 28850 – Torrejón de Ardoz, Madrid, Spain
- ¹¹ Max Planck Institute for Extraterrestrial Physics, Giessenbachstr. 1, 85748 Garching, Germany
- ¹² Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italy
- ¹³ Leiden Observatory, Leiden University, Einsteinweg 55, 2333 CC Leiden, The Netherlands
- ¹⁴ Kapteyn Astronomical Institute, University of Groningen, PO Box 800, 9700 AV Groningen, The Netherlands
- ¹⁵ Instituto de Astrofísica de Canarias, Vía Láctea, 38205 La Laguna, Tenerife, Spain
- ¹⁶ Instituto de Astrofísica de Canarias (IAC); Departamento de Astrofísica, Universidad de La Laguna (ULL), 38200, La Laguna, Tenerife, Spain
- ¹⁷ Université PSL, Observatoire de Paris, Sorbonne Université, CNRS, LERMA, 75014, Paris, France
- ¹⁸ Université Paris-Cité, 5 Rue Thomas Mann, 75013, Paris, France
- ¹⁹ Institute of Cosmology and Gravitation, University of Portsmouth, Portsmouth PO1 3FX, UK
- ²⁰ Institute of Space Sciences (ICE, CSIC), Campus UAB, Carrer de Can Magrans, s/n, 08193 Barcelona, Spain
- ²¹ Institut d'Estudis Espacials de Catalunya (IEEC), Edifici RDIT, Campus UPC, 08860 Castelldefels, Barcelona, Spain
- ²² Department of Physics, P.O. Box 64, University of Helsinki, 00014 Helsinki, Finland
- ²³ University of Applied Sciences and Arts of Northwestern Switzerland, School of Computer Science, 5210 Windisch, Switzerland
- ²⁴ Institut d'Astrophysique de Paris, 98bis Boulevard Arago, 75014, Paris, France
- ²⁵ ICL, Junia, Université Catholique de Lille, LITL, 59000 Lille, France
- ²⁶ Herzberg Astronomy and Astrophysics Research Centre, 5071 W. Saanich Rd. Victoria, BC, V9E 2E7, Canada
- ²⁷ Technical University of Munich, TUM School of Natural Sciences, Physics Department, James-Frank-Str. 1, 85748 Garching, Germany
- ²⁸ Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str. 1, 85748 Garching, Germany
- ²⁹ Institute of Physics, Laboratory of Astrophysics, Ecole Polytechnique Fédérale de Lausanne (EPFL), Observatoire de Sauverny, 1290 Versoix, Switzerland
- ³⁰ SCITAS, Ecole Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland
- ³¹ Institut de Ciències del Cosmos (ICCUB), Universitat de Barcelona (IEEC-UB), Martí i Franquès 1, 08028 Barcelona, Spain
- ³² Institució Catalana de Recerca i Estudis Avançats (ICREA), Pas-seig de Lluís Companys 23, 08010 Barcelona, Spain
- ³³ Institut de Ciències de l'Espai (IEEC-CSIC), Campus UAB, Carrer de Can Magrans, s/n Cerdanyola del Vallès, 08193 Barcelona, Spain
- ³⁴ Aix-Marseille Université, CNRS, CNES, LAM, Marseille, France
- ³⁵ Institut d'Astrophysique de Paris, UMR 7095, CNRS, and Sorbonne Université, 98 bis boulevard Arago, 75014 Paris, France
- ³⁶ Universitäts-Sternwarte München, Fakultät für Physik, Ludwig-Maximilians-Universität München, Scheinerstrasse 1, 81679 München, Germany
- ³⁷ Minnesota Institute for Astrophysics, University of Minnesota, 116 Church St SE, Minneapolis, MN 55455, USA
- ³⁸ School of Physical Sciences, The Open University, Milton Keynes, MK7 6AA, UK
- ³⁹ Dipartimento di Fisica "Aldo Pontremoli", Università degli Studi di Milano, Via Celoria 16, 20133 Milano, Italy
- ⁴⁰ INAF-IASF Milano, Via Alfonso Corti 12, 20133 Milano, Italy
- ⁴¹ University of Applied Sciences and Arts of Northwestern Switzerland, School of Engineering, 5210 Windisch, Switzerland
- ⁴² Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA, 91109, USA
- ⁴³ INAF-Osservatorio Astronomico di Brera, Via Brera 28, 20122 Milano, Italy
- ⁴⁴ IFPU, Institute for Fundamental Physics of the Universe, via Beirut 2, 34151 Trieste, Italy
- ⁴⁵ INAF-Osservatorio Astronomico di Trieste, Via G. B. Tiepolo 11, 34143 Trieste, Italy
- ⁴⁶ INFN, Sezione di Trieste, Via Valerio 2, 34127 Trieste TS, Italy
- ⁴⁷ SISSA, International School for Advanced Studies, Via Bonomea 265, 34136 Trieste TS, Italy
- ⁴⁸ Dipartimento di Fisica e Astronomia, Università di Bologna, Via Gobetti 93/2, 40129 Bologna, Italy
- ⁴⁹ INFN-Sezione di Bologna, Viale Berti Pichat 6/2, 40127 Bologna, Italy
- ⁵⁰ INAF-Osservatorio Astronomico di Padova, Via dell'Osservatorio 5, 35122 Padova, Italy
- ⁵¹ Dipartimento di Fisica, Università di Genova, Via Dodecaneso 33, 16146, Genova, Italy
- ⁵² INFN-Sezione di Genova, Via Dodecaneso 33, 16146, Genova, Italy

- ⁵³ Department of Physics "E. Pancini", University Federico II, Via Cinthia 6, 80126, Napoli, Italy
- ⁵⁴ INAF-Osservatorio Astronomico di Capodimonte, Via Moiariello 16, 80131 Napoli, Italy
- ⁵⁵ Dipartimento di Fisica, Università degli Studi di Torino, Via P. Giuria 1, 10125 Torino, Italy
- ⁵⁶ INFN-Sezione di Torino, Via P. Giuria 1, 10125 Torino, Italy
- ⁵⁷ INAF-Osservatorio Astrofisico di Torino, Via Osservatorio 20, 10025 Pino Torinese (TO), Italy
- ⁵⁸ European Space Agency/ESTEC, Keplerlaan 1, 2201 AZ Noordwijk, The Netherlands
- ⁵⁹ Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Avenida Complutense 40, 28040 Madrid, Spain
- ⁶⁰ Port d'Informació Científica, Campus UAB, C. Albareda s/n, 08193 Bellaterra (Barcelona), Spain
- ⁶¹ INFN section of Naples, Via Cinthia 6, 80126, Napoli, Italy
- ⁶² Dipartimento di Fisica e Astronomia "Augusto Righi" - Alma Mater Studiorum Università di Bologna, Viale Berti Pichat 6/2, 40127 Bologna, Italy
- ⁶³ Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK
- ⁶⁴ Jodrell Bank Centre for Astrophysics, Department of Physics and Astronomy, University of Manchester, Oxford Road, Manchester M13 9PL, UK
- ⁶⁵ European Space Agency/ESRIN, Largo Galileo Galilei 1, 00044 Frascati, Roma, Italy
- ⁶⁶ ESAC/ESA, Camino Bajo del Castillo, s/n., Urb. Villafranca del Castillo, 28692 Villanueva de la Cañada, Madrid, Spain
- ⁶⁷ Université Claude Bernard Lyon 1, CNRS/IN2P3, IP2I Lyon, UMR 5822, Villeurbanne, F-69100, France
- ⁶⁸ UCB Lyon 1, CNRS/IN2P3, IUF, IP2I Lyon, 4 rue Enrico Fermi, 69622 Villeurbanne, France
- ⁶⁹ Mullard Space Science Laboratory, University College London, Holmbury St Mary, Dorking, Surrey RH5 6NT, UK
- ⁷⁰ Departamento de Física, Faculdade de Ciências, Universidade de Lisboa, Edifício C8, Campo Grande, PT1749-016 Lisboa, Portugal
- ⁷¹ Instituto de Astrofísica e Ciências do Espaço, Faculdade de Ciências, Universidade de Lisboa, Campo Grande, 1749-016 Lisboa, Portugal
- ⁷² Department of Astronomy, University of Geneva, ch. d'Ecogia 16, 1290 Versoix, Switzerland
- ⁷³ INAF-Istituto di Astrofisica e Planetologia Spaziali, via del Fosso del Cavaliere, 100, 00100 Roma, Italy
- ⁷⁴ Université Paris-Saclay, CNRS, Institut d'astrophysique spatiale, 91405, Orsay, France
- ⁷⁵ INFN-Padova, Via Marzolo 8, 35131 Padova, Italy
- ⁷⁶ Aix-Marseille Université, CNRS/IN2P3, CPPM, Marseille, France
- ⁷⁷ Space Science Data Center, Italian Space Agency, via del Politecnico snc, 00133 Roma, Italy
- ⁷⁸ INFN-Bologna, Via Imerio 46, 40126 Bologna, Italy
- ⁷⁹ University Observatory, LMU Faculty of Physics, Scheinerstrasse 1, 81679 Munich, Germany
- ⁸⁰ Institute of Theoretical Astrophysics, University of Oslo, P.O. Box 1029 Blindern, 0315 Oslo, Norway
- ⁸¹ Department of Physics, Lancaster University, Lancaster, LA1 4YB, UK
- ⁸² Felix Hormuth Engineering, Goethestr. 17, 69181 Leimen, Germany
- ⁸³ Technical University of Denmark, Elektrovej 327, 2800 Kgs. Lyngby, Denmark
- ⁸⁴ Cosmic Dawn Center (DAWN), Denmark
- ⁸⁵ Max-Planck-Institut für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany
- ⁸⁶ NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA
- ⁸⁷ Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK
- ⁸⁸ Department of Physics and Helsinki Institute of Physics, Gustaf Hållströmin katu 2, University of Helsinki, 00014 Helsinki, Finland
- ⁸⁹ Université de Genève, Département de Physique Théorique and Centre for Astroparticle Physics, 24 quai Ernest-Ansermet, CH-1211 Genève 4, Switzerland
- ⁹⁰ Helsinki Institute of Physics, Gustaf Hållströmin katu 2, University of Helsinki, 00014 Helsinki, Finland
- ⁹¹ Laboratoire d'étude de l'Univers et des phénomènes eXtremes, Observatoire de Paris, Université PSL, Sorbonne Université, CNRS, 92190 Meudon, France
- ⁹² SKAO, Jodrell Bank, Lower Withington, Macclesfield SK11 9FT, United Kingdom
- ⁹³ Centre de Calcul de l'IN2P3/CNRS, 21 avenue Pierre de Coubertin 69627 Villeurbanne Cedex, France
- ⁹⁴ INFN-Sezione di Milano, Via Celoria 16, 20133 Milano, Italy
- ⁹⁵ Universität Bonn, Argelander-Institut für Astronomie, Auf dem Hügel 71, 53121 Bonn, Germany
- ⁹⁶ INFN-Sezione di Roma, Piazzale Aldo Moro, 2 - c/o Dipartimento di Fisica, Edificio G. Marconi, 00185 Roma, Italy
- ⁹⁷ Department of Physics, Institute for Computational Cosmology, Durham University, South Road, Durham, DH1 3LE, UK
- ⁹⁸ Université Paris Cité, CNRS, Astroparticule et Cosmologie, 75013 Paris, France
- ⁹⁹ CNRS-UCB International Research Laboratory, Centre Pierre Binétruy, IRL2007, CPB-IN2P3, Berkeley, USA
- ¹⁰⁰ Telespazio UK S.L. for European Space Agency (ESA), Camino bajo del Castillo, s/n, Urbanización Villafranca del Castillo, Villanueva de la Cañada, 28692 Madrid, Spain
- ¹⁰¹ Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Campus UAB, 08193 Bellaterra (Barcelona), Spain
- ¹⁰² School of Mathematics and Physics, University of Surrey, Guildford, Surrey, GU2 7XH, UK
- ¹⁰³ DARK, Niels Bohr Institute, University of Copenhagen, Jagtvej 155, 2200 Copenhagen, Denmark
- ¹⁰⁴ Waterloo Centre for Astrophysics, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada
- ¹⁰⁵ Department of Physics and Astronomy, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada
- ¹⁰⁶ Perimeter Institute for Theoretical Physics, Waterloo, Ontario N2L 2Y5, Canada
- ¹⁰⁷ Université Paris-Saclay, Université Paris Cité, CEA, CNRS, AIM, 91191, Gif-sur-Yvette, France
- ¹⁰⁸ Centre National d'Etudes Spatiales – Centre spatial de Toulouse, 18 avenue Edouard Belin, 31401 Toulouse Cedex 9, France
- ¹⁰⁹ Institute of Space Science, Str. Atomistilor, nr. 409 Măgurele, Ilfov, 077125, Romania
- ¹¹⁰ Dipartimento di Fisica e Astronomia "G. Galilei", Università di Padova, Via Marzolo 8, 35131 Padova, Italy
- ¹¹¹ Institut für Theoretische Physik, University of Heidelberg, Philosophenweg 16, 69120 Heidelberg, Germany
- ¹¹² Institut de Recherche en Astrophysique et Planétologie (IRAP), Université de Toulouse, CNRS, UPS, CNES, 14 Av. Edouard Belin, 31400 Toulouse, France
- ¹¹³ Université St Joseph; Faculty of Sciences, Beirut, Lebanon
- ¹¹⁴ Departamento de Física, FCFM, Universidad de Chile, Blanco Encalada 2008, Santiago, Chile
- ¹¹⁵ Universität Innsbruck, Institut für Astro- und Teilchenphysik, Technikerstr. 25/8, 6020 Innsbruck, Austria
- ¹¹⁶ Atlantis, University Science Park, Sede Bld 48940, Leioa-Bilbao, Spain
- ¹¹⁷ Instituto de Astrofísica e Ciências do Espaço, Faculdade de Ciências, Universidade de Lisboa, Tapada da Ajuda, 1349-018 Lisboa, Portugal
- ¹¹⁸ Cosmic Dawn Center (DAWN)
- ¹¹⁹ Niels Bohr Institute, University of Copenhagen, Jagtvej 128, 2200 Copenhagen, Denmark
- ¹²⁰ Universidad Politécnica de Cartagena, Departamento de Electrónica y Tecnología de Computadoras, Plaza del Hospital 1, 30202 Cartagena, Spain
- ¹²¹ Infrared Processing and Analysis Center, California Institute of Technology, Pasadena, CA 91125, USA

- ¹²² Department of Mathematics and Physics E. De Giorgi, University of Salento, Via per Arnesano, CP-I93, 73100, Lecce, Italy
¹²³ INFN, Sezione di Lecce, Via per Arnesano, CP-I93, 73100, Lecce, Italy
¹²⁴ INAF-Sezione di Lecce, c/o Dipartimento Matematica e Fisica, Via per Arnesano, 73100, Lecce, Italy

Appendix A: Comparison between data and simulations

We compare the statistical properties of our simulated AGN with those measured on observational data using the mean signal level, the standard deviation, and the Gini coefficient. The Gini coefficient is defined as

$$G = \frac{\sum_{i=1}^n \sum_{j=1}^n |x_i - x_j|}{2n^2 \bar{x}}, \quad (\text{A.1})$$

where x_i is the flux in pixel i , n is the total number of pixels, and \bar{x} is the mean flux. It defines the flux distribution in an image; higher values indicate that the flux is concentrated in the center (for instance, when the image is dominated by an unresolved AGN), while lower values correspond to AGN with more diffuse host galaxy. From Fig. A.1 we see that the simulations reproduce the observed distributions well, with both samples occupying the same region of parameter space.

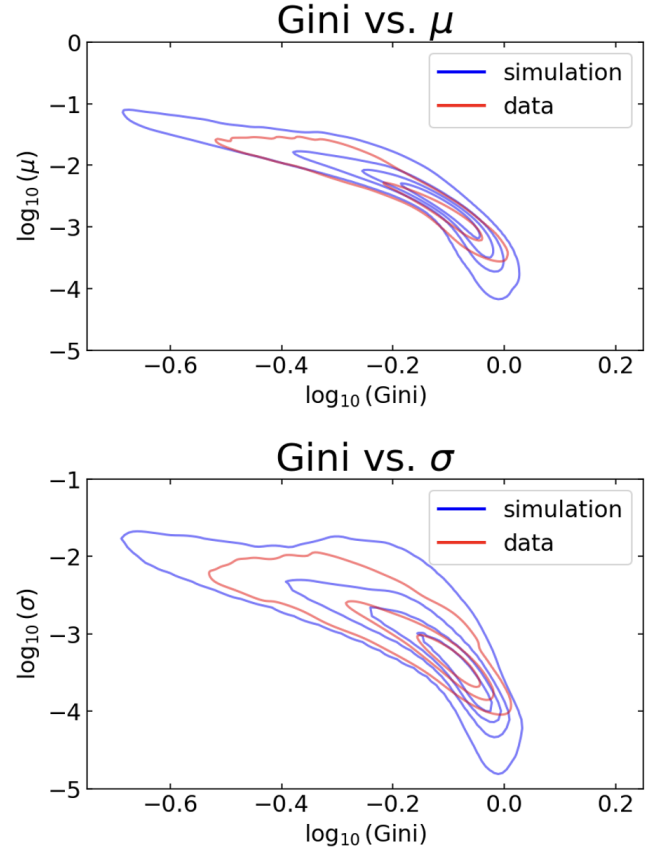


Fig. A.1. Comparison of the statistical properties of the simulations (blue) and the data (red). The upper panel shows the Gini coefficient vs. mean of the signal in logarithm. The lower panel shows the Gini coefficient vs. standard deviation of the signal in logarithm. In both panels, the contours represent kernel density estimates of the respective distributions with 5 levels.

Appendix B: CNN on foreground contamination

In Sect. 4.2 we gave an estimate of foreground contaminants. In order to compare the number of dual objects found in Q1 and the number of possible contaminants, it is necessary to use the same detection algorithm, in this case CNN. Therefore, once

we added a PSF to the random cutouts, it was possible to study how the network behaves in the case of extended contaminants sources. In Fig. B.1, we show some examples of the behavior of the network recognising when the contamination is point-like (left panel, dual) or extended and not considered as a contaminant (right panel, single).

Appendix C: Lens candidates

The network classified both the objects in Fig. C.1 as dual candidates. In the visual inspection, we identified features consistent with lensing and we classified both as class 3. The first system may represent a quadruply lensed configuration, while the second exhibits an arc-like structure that could also be interpreted as a ring galaxy, similar to the case shown in Fig. 15(c).

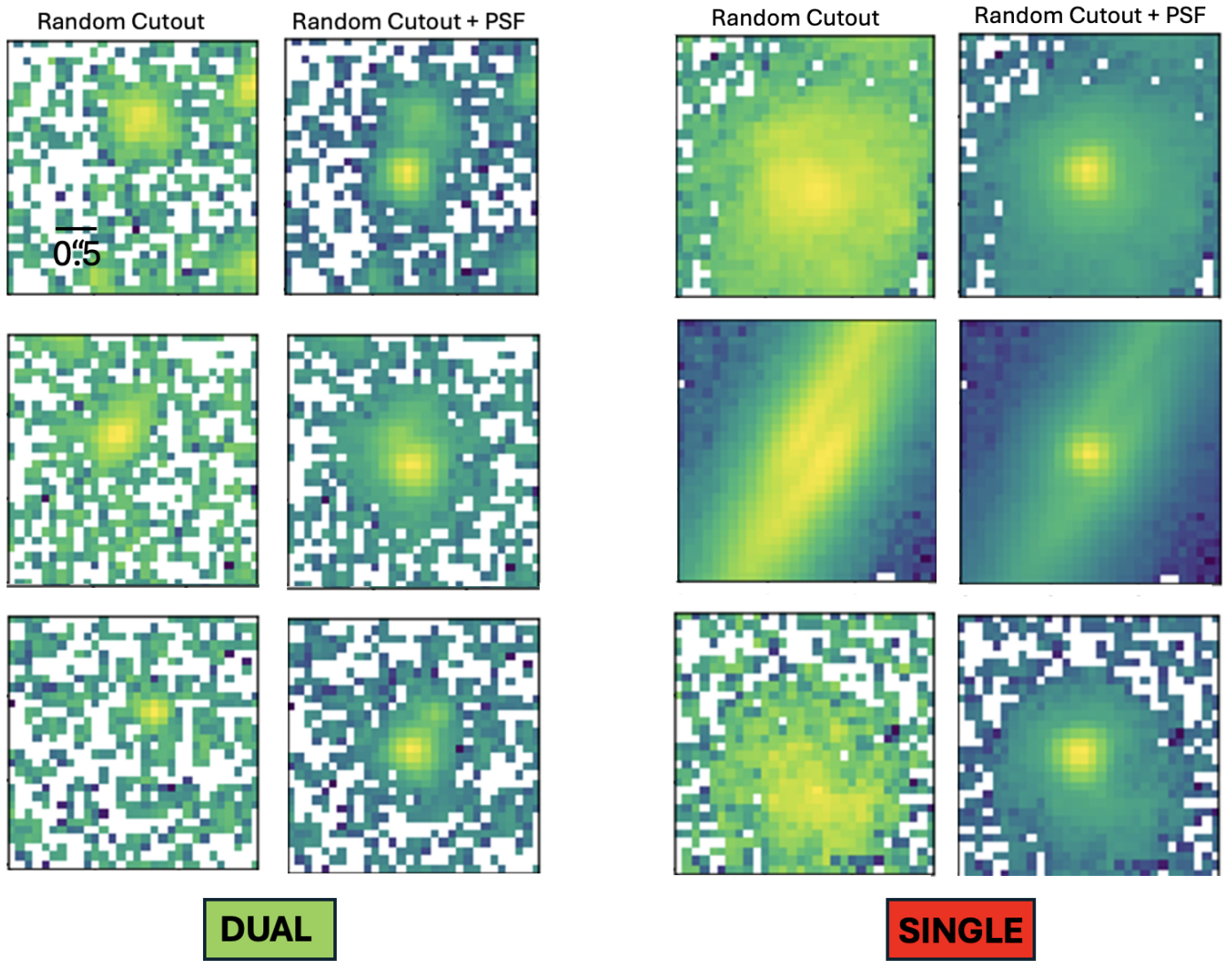


Fig. B.1. Application of the CNN to the foreground contaminations. The left panel shows examples predicted as dual systems; the right panel shows examples predicted as single systems. In both panels, the first column represents random cutouts and the second column the random cutouts + PSF.

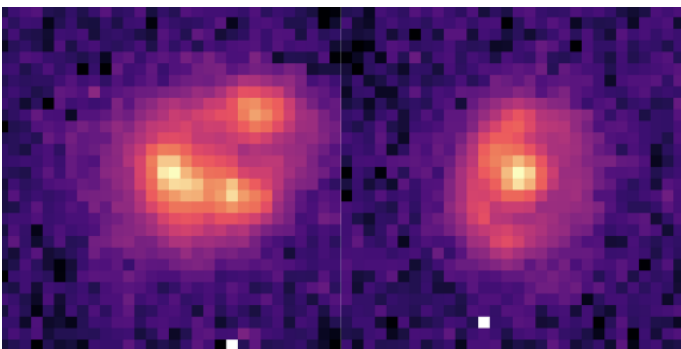


Fig. C.1. VIS images of objects predicted as dual by the CNN showing lens-like configurations.