

Euclid Quick Data Release (Q1)

XLI. AgileLens: A scalable CNN-based pipeline for strong gravitational lens identification

Euclid Collaboration: X. Xu^{1,2}, R. Chen¹, T. Li¹, A. R. Cooray¹, S. Schuldt^{3,4}, J. A. Acevedo Barroso⁵, D. Stern⁵, D. Scott⁶, M. Meneghetti^{7,8}, G. Despali^{9,7,8}, J. Chopra¹, Y. Cao¹, M. Cheng¹, J. Buda¹, J. Zhang¹, J. Furumizo¹, R. Valencia¹, Z. Jiang², C. Tortora¹⁰, N. E. P. Lines¹¹, T. E. Collett¹¹, S. Fotopoulou¹², A. Galan^{13,14}, A. Manjón-García¹⁵, R. Gavazzi^{16,17}, L. Iwamoto¹⁸, S. Kruk¹⁹, M. Millon²⁰, P. Nugent²¹, C. Saulder^{22,23}, D. Sluse²⁴, J. Wilde²⁵, M. Walmsley^{26,27}, F. Courbin^{25,28,29}, R. B. Metcalfe^{9,7}, B. Altieri¹⁹, A. Amara³⁰, S. Andreon³¹, N. Auricchio⁷, C. Baccigalupi^{32,33,34,35}, M. Baldi^{36,7,8}, A. Balestra³⁷, S. Bardelli⁷, P. Battaglia⁷, R. Bender^{22,23}, A. Biviano^{33,32}, E. Branchini^{38,39,31}, M. Brescia^{40,10}, S. Camera^{41,42,43}, V. Capobianco⁴³, C. Carbone⁴, V. F. Cardone^{44,45}, J. Carretero^{46,47}, S. Casas^{48,49}, M. Castellano⁴⁴, G. Castignani⁷, S. Cavuoti^{10,50}, A. Cimatti⁵¹, C. Colodro-Conde⁵², G. Congedo⁵³, C. J. Conselice²⁷, L. Conversi^{54,19}, Y. Copin⁵⁵, H. M. Courtois⁵⁶, M. Cropper⁵⁷, A. Da Silva^{58,59}, H. Degaudenzi⁶⁰, G. De Lucia³³, C. Dolding⁵⁷, H. Dole⁶¹, F. Dubath⁶⁰, X. Dupac¹⁹, S. Dusini⁶², S. Escoffier⁶³, M. Farina⁶⁴, R. Farinelli⁷, S. Farrens⁶⁵, S. Ferriol⁵⁵, F. Finelli^{7,66}, P. Fosalba^{67,68}, M. Frailis³³, E. Franceschi⁷, M. Fumana⁴, S. Galeotta³³, K. George⁶⁹, W. Gillard⁶³, B. Gillis⁵³, C. Giocoli^{7,8}, P. Gómez-Alvarez^{70,19}, J. Gracia-Carpio²², A. Grazian³⁷, F. Grupp^{22,23}, S. V. H. Haugan⁷¹, W. Holmes⁵, F. Hormuth⁷², A. Hornstrup^{73,74}, K. Jahnke⁷⁵, M. Jhabvala⁷⁶, B. Joachimi⁷⁷, S. Kermiche⁶³, A. Kiessling⁵, B. Kubik⁵⁵, M. Kümmel²³, M. Kunz²⁰, H. Kurki-Suonio^{78,79}, A. M. C. Le Brun⁸⁰, S. Ligi⁴³, P. B. Lilje⁷¹, V. Lindholm^{78,79}, I. Lloro⁸¹, G. Mainetti⁸², E. Maiorano⁷, O. Mansutti³³, S. Marcin⁸³, O. Marggraf⁸⁴, M. Martinelli^{44,45}, N. Martinet¹⁶, F. Marulli^{9,7,8}, R. J. Massey⁸⁵, E. Medinaceli⁷, S. Mei^{86,87}, M. Melchior⁸⁸, E. Merlin⁴⁴, G. Meylan⁸⁹, A. Mora⁹⁰, M. Moresco^{9,7}, L. Moscardini^{9,7,8}, R. Nakajima⁸⁴, C. Neissner^{91,47}, R. C. Nichol³⁰, S.-M. Niemi⁹², J. W. Nightingale⁹³, C. Padilla⁹¹, S. Paltani⁶⁰, F. Pasian³³, K. Pedersen⁹⁴, W. J. Percival^{95,96,97}, V. Pettorino⁹², G. Polenta⁹⁸, M. Poncet⁹⁹, L. A. Popa¹⁰⁰, F. Raison²², A. Renzi^{101,62}, J. Rhodes⁵, G. Riccio¹⁰, E. Romelli³³, M. Roncarelli⁷, R. Saglia^{23,22}, Z. Sakr^{102,103,104}, D. Sapone¹⁰⁵, M. Schirmer⁷⁵, P. Schneider⁸⁴, T. Schrabback¹⁰⁶, A. Secroun⁶³, G. Seidel⁷⁵, E. Sihvola¹⁰⁷, P. Simon⁸⁴, C. Sirignano^{101,62}, G. Sirri⁸, L. Stanco⁶², P. Tallada-Crespí^{46,47}, A. N. Taylor⁵³, I. Tereno^{58,108}, N. Tessore⁵⁷, S. Toft^{109,110}, R. Toledo-Moreo¹¹¹, F. Torradeflot^{47,46}, I. Tutusaus^{68,67,103}, L. Valenziano^{7,66}, J. Valiviita^{78,79}, T. Vassallo^{33,69}, G. Verdoes Kleijn¹¹², A. Veropalumbo^{31,39,38}, Y. Wang¹¹³, J. Weller^{23,22}, A. Zacchei^{33,32}, G. Zamorani⁷, F. M. Zerbi³¹, E. Zucca⁷, M. Ballardini^{114,115,7}, M. Bolzonella⁷, C. Burigana^{116,66}, R. Cabanac¹⁰³, M. Calabrese^{117,4}, A. Cappi^{118,7}, T. Castro^{33,34,32,119}, J. A. Escartin Vigo²², L. Gabarra¹²⁰, S. Hemmati¹¹³, J. Macias-Perez¹²¹, R. Maoli^{122,44}, J. Martín-Fleitas¹²³, N. Mauri^{51,8}, P. Monaco^{124,33,34,32}, A. A. Nucita^{125,126,127}, A. Pezzotta³¹, M. Pöntinen⁷⁸, I. Risso^{31,39}, V. Scottetz^{128,129}, M. Sereno^{7,8}, M. Tenti⁸, M. Tucci⁶⁰, M. Viel^{32,33,35,34,119}, M. Wiesmann⁷¹, Y. Akrami^{102,130}, I. T. Andika⁶⁹, G. Angora^{10,114}, S. Anselmi^{62,101,131}, M. Archidiacono^{3,132}, F. Atrio-Barandela¹³³, L. Bazzanini^{114,7}, P. Bergamini⁷, D. Bertacca^{101,37,62}, M. Bethermin¹³⁴, F. Beutler⁵³, L. Blot^{135,80}, S. Borgani^{124,32,33,34,119}, M. L. Brown²⁷, S. Bruton¹³⁶, A. Calabro⁴⁴, B. Camacho Quevedo^{32,35,33}, F. Caro⁴⁴, C. S. Carvalho¹⁰⁸, F. Cogato^{9,7}, S. Conseil⁵⁵, O. Cucciati⁷, S. Davini³⁹, G. Desprez¹¹², A. Díaz-Sánchez¹⁵, S. Di Domizio^{38,39}, J. M. Diego¹³⁷, P.-A. Duc¹³⁴, V. Duret⁶³, M. Y. Elkhachab^{124,33,34,32}, A. Enia⁷, Y. Fang²³, A. Finoguenov⁷⁸, A. Franco^{126,125,127}, K. Ganga⁸⁶, T. Gasparotto⁴⁴, E. Gaztanaga^{68,67,11}, F. Giacomini⁸, F. Gianotti⁷, G. Gozaliasl^{138,78}, M. Guidi^{36,7}, C. M. Gutierrez^{52,139}, A. Hall⁵³, C. Hernández-Monteagudo^{139,52}, H. Hildebrandt¹⁴⁰, J. Hjorth⁹⁴, J. J. E. Kajava^{141,142,143}, Y. Kang⁶⁰, V. Kansal^{144,145}, D. Karagiannis^{114,146}, K. Kiiveri¹⁰⁷, J. Kim¹²⁰, C. C. Kirkpatrick¹⁰⁷, F. Lepori¹⁴⁷, G. Leroy^{148,85}, G. F. Lesci^{9,7}, J. Lesgourgues⁴⁸, T. I. Liaudat¹⁴⁹, S. J. Liu⁶⁴, M. Magliocchetti⁶⁴, E. A. Magnier¹⁵⁰, F. Mannucci¹⁵¹, C. J. A. P. Martins^{152,153}, L. Maurin⁶¹, M. Miluzio^{19,154}, C. Moretti^{33,32,34}, G. Morgante⁷, K. Naidoo^{11,75}, A. Navarro-Alsina⁸⁴, S. Nesseris¹⁰², D. Paoletti^{7,66}, F. Passalacqua^{101,62}, K. Paterson⁷⁵, L. Patrizii⁸, A. Pisani⁶³, D. Potter¹⁵⁵, G. W. Pratt⁶⁵, S. Quai^{9,7}, M. Radovich³⁷, K. Rojas⁸³, W. Roster²², S. Sacquogna¹⁵⁶, M. Sahlén¹⁵⁷, D. B. Sanders¹⁵⁰, E. Sarpa^{35,119,33}, C. Scarlata¹⁵⁸, A. Schneider¹⁵⁵, M. Schultheis¹¹⁸, D. Sciotti^{44,45}, E. Sellentin^{159,160}, L. C. Smith¹⁶¹, K. Tanidis¹⁶², C. Tao⁶³,

F. Tarsitano^{163, 60}, G. Testera³⁹, R. Teyssier¹⁶⁴, S. Tosi^{38, 31, 39}, A. Troja^{101, 62}, A. Venhola¹⁶⁵, D. Vergani⁷,
G. Vernardos^{166, 167}, G. Verza^{168, 169}, S. Vinciguerra¹⁶, N. A. Walton¹⁶¹, A. H. Wright¹⁴⁰, and H. W. Yeung⁵³

(Affiliations can be found after the references)

June 30, 2026

ABSTRACT

We present an end-to-end, iterative pipeline for efficient identification of strong galaxy-galaxy lensing systems, applied to the *Euclid* Q1 imaging data. Starting from the VIS catalogues, we rejected point sources, applied a magnitude cut ($I_E \leq 24$) on the deflectors, and ran a pixel-level artefact-and-noise filter to build 96×96 pixel cutouts. The VIS+NISP colour composites were constructed with a VIS-anchored luminance scheme that preserves VIS morphology and NISP colour contrast. A VIS-only seed classifier supplies clear positives and typical impostors, which we used to curate a morphology-balanced negative set and augment scarce positives. Among the six compact convolutional neural networks (CNNs) that had been studied initially, the modified VGG16 (GlobalAveragePooling + 256/128 dense layers with the last nine layers trainable) exhibited the best performance. In our run, the training set grew from 27 seed lenses (augmented $67 \times$ to 1809) plus 2000 negatives to a colour dataset of 30 686 images. After three rounds of iterative fine-tuning, a human grading of the top 4000 candidates ranked by the final model yielded 441 Grade A/B candidate lensing systems, including 311 overlapping with the existing Q1 strong-lens catalogue, and 130 additional A/B candidates (9 As and 121 Bs) not previously reported. Independently, the model recovered 740 out of 905 (81.8%) candidate Q1 lenses within its top 20 000 predictions, considering off-centred samples. Candidates span $I_E \approx 17\text{--}24$ AB mag (median 21.3 AB mag) and are redder in $Y_E - H_E$ than the parent population, consistent with massive early-type deflectors. Each training iteration required about a week for a small team and the approach can easily be scaled to future wide-area *Euclid* releases. Subsequent works will be focussed on calibrating the selection function via lens injection, extending recall through uncertainty-aware active learning, and exploring multi-scale or attention-based neural networks with fast post hoc vetters that incorporate lens models into the classification.

Key words. Gravitational lensing: strong – Techniques: image processing – Methods: data analysis – Surveys

1. Introduction

The phenomenon of strong gravitational lensing, predicted by Einstein’s theory of general relativity, occurs when the gravitational potential of a massive foreground astronomical object, such as a galaxy or galaxy cluster, deflects and magnifies light from a background source. This deflection can lead to the formation of multiple, distorted images or (in cases of a precise alignment) characteristic structures such as Einstein rings. The observational properties of strong lensing systems enable wide astrophysical and cosmological applications. Given a relative alignment, the lensing features of the background depend entirely on the gravitational potential of the foreground, making it a useful probe of luminous and dark masses within the foreground (e.g. Koopmans et al. 2009; Treu 2010; Vegetti et al. 2024).

The lensed images are sensitive to small perturbations in the lensing potential, making it useful for probing dark matter substructures in the foreground (e.g. Dalal & Kochanek 2002; Koopmans 2005; Keeton & Moustakas 2009; Vegetti et al. 2010; Vegetti et al. 2014; Hezaveh et al. 2016; Gannon et al. 2025). The determination of the total foreground mass from lensing in combination with stellar mass analysis from photometry yields valuable constraints on the initial mass function (IMF) of galaxies outside the Local Group (e.g. Treu et al. 2010; Sonnenfeld et al. 2013; Sonnenfeld et al. 2015). Since multiple image copies of the same lensed background traverse different paths to Earth, there is often a time delay between different images. This difference is measurable if the background source is time variable and can be used to measure cosmological parameters such as the Hubble constant, H_0 (e.g. Wong et al. 2020; Rodney et al. 2021). Statistics characterising a large population of lensing systems can be used to constrain cosmological evolution and galaxy mass structures (e.g. Cen et al. 1994; Chae 2003; Sonnenfeld & Cautun 2021; Sonnenfeld 2021), along with well-defined selection criteria can further strengthen these constraints (Sonnenfeld 2022). Strong lensing can also be used to aid the observation of very high-redshift galaxies as it significantly magnifies and am-

plifies the light from the background source (e.g. Chirivì et al. 2020; Fudamoto et al. 2025). For all of its scientific value, the discovery of a greater number of strong lensing systems is highly desirable.

The *Euclid* Mission (*Euclid* Collaboration: Mellier et al. 2025), developed by the European Space Agency with contributions from NASA, is projected to survey most of the extragalactic sky and poised to observe billions of galaxies (*Euclid* Collaboration: Scaramella et al. 2022) with the space-based telescope. Initial estimates from simulations suggest that $\sim 170\,000$ strong lenses should be observed in the *Euclid* dataset (Collett 2015). Extrapolations from strong lenses identified in the preliminary results also show broadly consistent expectations (Acevedo Barroso et al. 2025; *Euclid* Collaboration: Walmsley et al. 2025). Although visual inspections have historically been utilised for finding lensing systems (e.g. Jackson 2008; Faure et al. 2008), it is infeasible for a survey of *Euclid*’s scale. Consequently, the development of efficient and reliable automated detection algorithms is imperative to systematically identify and catalogue strong lensing systems, thereby maximising the scientific return of the *Euclid* survey.

The detection of strong lenses in large-scale astronomical surveys presents several inherent challenges. Statistically, strong lensing events are exceptionally rare, occurring approximately once per $O(10^4)$ observed galaxies; nevertheless, *Euclid* is projected to observe a larger sample of lenses than any previous surveys (see figure 1 in *Euclid* Collaboration: Walmsley et al. 2025). Furthermore, lensing features exhibit considerable morphological diversity and are frequently confused with other astrophysical objects, such as interacting and ring galaxies. Consequently, it is essential to deploy classification algorithms with high precision and recall to minimise false positives while ensuring high completeness of the true lens sample.

Previous strong lensing surveys and associated detection efforts have employed a range of techniques. The initial discoveries were often serendipitous. Subsequent systematic searches, such as the Cosmic Lens All-Sky Survey (Myers et al. 2003) and the Sloan Lens ACS Survey (Bolton et al. 2006), utilised tar-

* e-mail: xinmeix@uci.edu

geted observational strategies combined with visual inspection or semi-automated algorithms based on predefined morphological or photometric criteria. Although these surveys yielded significant discoveries, their methodologies possess inherent limitations with respect to scalability and potential selection biases. More recent wide-field surveys, including the Kilo-Degree Survey (de Jong et al. 2013), the Dark Energy Survey (Dark Energy Survey Collaboration et al. 2016), and the Hyper Suprime-Cam SSP Survey (Aihara et al. 2018), have further underscored the need for advanced automated detection techniques commensurate with increasing data volumes.

The application of machine learning techniques in astronomy has seen considerable growth, offering powerful tools for data analysis and interpretation. Specifically, deep learning architectures, such as convolutional neural networks (CNNs; LeCun & Bengio 1995), have proven effective in tasks involving image classification from astronomical datasets (Carleo et al. 2019), such as morphological classification of SDSS galaxies (Zhu et al. 2019). Several studies have successfully deployed CNNs for the identification of strong lens candidates in observational survey data, frequently achieving performance metrics that surpass those of traditional algorithms in terms of accuracy and computational efficiency (Petrillo et al. 2017; Jacobs et al. 2017; Pourrahmani et al. 2018; Jacobs et al. 2019; Cheng et al. 2020; Li et al. 2020; Huang et al. 2021; Storfer et al. 2024).

Recently, a great deal of effort has been made to find strong lenses in the new *Euclid* data releases. Initial experiments were conducted on the *Euclid* ERO data to yield a preliminary validation of the lens finding pipelines (Pearce-Casey et al. 2025; Acevedo Barroso et al. 2025; Nagam et al. 2025). A much larger search was then conducted on the basis of the *Euclid* Q1 data, notably by the strong lens discovery engine (SLDE) series. SLDE A (Euclid Collaboration: Walmsley et al. 2025) encompasses an overview of SLDE research in *Euclid* Q1. SLDE B (Euclid Collaboration: Rojas et al. 2025) reported early discoveries based on spectroscopically confirmed high-velocity-dispersion galaxies, forming training samples for machine learning models in SLDE C (Euclid Collaboration: Lines et al. 2025b), which also evaluates their performance using real *Euclid* imaging data. SLDE D (Euclid Collaboration: Li et al. 2025) focuses on the double source plane lenses identified in Q1 and provides preliminary modelling for a forecast of cosmological parameters. SLDE E (Euclid Collaboration: Holloway et al. 2025) introduced a Bayesian ensemble approach that combines multiple lens classifiers to further optimise lens discovery in the DR1 discovery engine. SLDE F (Ecker et al. in prep.) presented additional bright and low redshift lens samples found in Q1 by looking into the *Gaia* cross-matched sources previously excluded in SLDE. In addition to CNN-based pipelines, vision transformers have been explored for strong lens detection, motivated by their capacity to model long-range spatial correlations (Vincken et al. in prep.), although their relative advantages over CNNs for *Euclid* data are not yet firmly established.

These prior investigations provide a robust foundation and compelling evidence for the utility of deep learning in the domain of strong lens detection. This paper aims to extend these efforts by developing and evaluating a CNN-based methodology specifically optimised for the characteristics of the *Euclid* survey data, addressing the critical need for high-fidelity lens identification in the era of large-scale astronomical surveys.

This paper is organised as follows. Section 2 gives an overview of the *Euclid* Q1 dataset, then describes our expectations and the way this work is meant to complement existing lens search efforts on the Q1 dataset. Section 3 presents the frame-

work of the automated strong-lens detection pipeline. We then detail the specific data generation procedure to procure the full dataset used for training and inference. These include: catalogue based pre-filtering, cutout generation and subsequent enhancements, pixel-level artefact and noise cleaning, balanced hard-negative curation, colour-set data composition, and the iterative fine-tuning of compact CNN approaches, with an emphasis on the modified VGG16 model that is our primary classifier. It ends with an explanation of the procedure applied for visual grading of the top-ranked candidates. Section 4 reports performance and validation: recovery of lenses from the Q1 imaging data and basic photometric trends of the discovered systems. Section 5 discusses the strengths and limitations of this approach, along with the selection bias of the detection pipeline, comparisons with simulation-heavy approaches, and scalability to future *Euclid* releases. We conclude with a brief summary and outlook in Sect. 6.

2. *Euclid* imaging data

2.1. Overview of *Euclid* VIS and NISP instruments

The *Euclid* Mission (Euclid Collaboration: Mellier et al. 2025) is equipped with two instruments for measuring electromagnetic radiation: the Visible Imager (VIS; Euclid Collaboration: Cropper et al. 2025) and the Near-Infrared Spectrometer and Photometer (NISP; Euclid Collaboration: Jahnke et al. 2025). When used together, these instruments deliver high-quality imaging data in the visible and near-infrared bands.

Of the two, the VIS instrument is optimally suited for resolving the morphology of astronomical sources due to its superior angular resolution. It observes in a single wide passband, approximately corresponding to the combined $r + i + z$ filters, and covers wavelength from 550 nm to 900 nm. The VIS focal plane comprises 36 charge-coupled devices (CCDs), each 4132×4096 pixels (≈ 16 megapixels per CCD), yielding an aggregate image size of approximately 609 megapixels. With a pixel scale of $0''.1$, the total field of view of VIS is approximately 0.54 deg^2 . Its combination of wide field, broadband filter, and long exposure time confers high sensitivity to faint sources across a broad spectral range, achieving a signal-to-noise ratio of $S/N \geq 10$ for sources with $I_E \leq 24.5$ (Euclid Collaboration: Cropper et al. 2025). In this work, all magnitudes are in AB.

Conversely, the NISP instrument is designed to provide detailed spectroscopic and photometric data. It images in three separate bands, Y_E , J_E , and H_E , collectively spanning the 950–2020 nm range (Euclid Collaboration: Schirmer et al. 2022). NISP employs 16 detectors, each with 2040×2040 pixels (approximately 4.16 megapixels). With a field of view comparable to VIS at 0.57 deg^2 , NISP offers about one-third the angular resolution, corresponding to a pixel scale of $0''.3$. It is also less sensitive than VIS, detecting $I_E \leq 24.5$ sources at a $S/N \geq 5$ in each band (Euclid Collaboration: Jahnke et al. 2025).

2.2. Characteristics of Q1 release imaging data

The full *Euclid* survey aims to observe approximately one-third of the celestial sphere using the VIS and NISP instruments over its eight-year mission duration. This work utilises a small portion of the full survey: the Quick Release 1 (Q1) dataset (Euclid Quick Release Q1 2025). Q1 comprises data from three deep survey fields (Euclid Collaboration: Aussel et al. 2025), covering a total area of 63.2 deg^2 , which represents approximately 0.45% of the mission's total planned sky coverage of $14\,000 \text{ deg}^2$. Based

on Q1 dataset, a broad range of scientific investigations were conducted in anticipation of the future data releases covering wider areas on the sky (e.g. [Euclid Collaboration: Enia et al. 2024](#); [Euclid Collaboration: Corcho-Caballero et al. 2025](#); [Euclid Collaboration: Holloway et al. 2025](#)).

2.3. Expected data volume and lens distribution

The survey fields for *Euclid* were selected to minimise contamination from Galactic stars and dust, thereby maximising the detection depth for extragalactic sources. As the most sensitive and accurate wide-field cosmological survey to date, *Euclid* is expected to detect roughly 6 billion galaxies ([Euclid Collaboration: Mellier et al. 2025](#)). According to the estimates of [Collett \(2015\)](#), this dataset is projected to include up to 170 000 strong gravitational lensing events. Scaling this estimate to the Q1 dataset suggests that ≈ 760 strong lenses may be present, of which ≈ 670 are expected after applying the $I_E < 24$ magnitude cut. Given that Q1 contains approximately 30 million sources ([Euclid Collaboration: Aussel et al. 2025](#)), the data volume renders manual identification impractical. Prior studies such as SLDE A employed a combination of citizen science, simulations, ML models, modelling, and expert inspection to identify lenses in Q1, yielding a catalogue of 500 expert-validated grade A/B lens candidates ([Euclid Collaboration: Walmsley et al. 2025](#)). SLDE C demonstrated the efficacy of automated methods for lens discovery in *Euclid* data, where the best performing model, Zoobot, detected 163 grade A/B lenses in its top 1000 ranked objects. The other four models used in SLDE C were trained on a combination of mostly simulated lenses and real lenses.

This work builds upon and complements previous works for three reasons. The GAIA_CROSSMATCH filter used to remove stars in SLDE A was later found by SLDE F to filter out galaxies, including numerous obvious gravitational lenses associated with mostly low-redshift, bright galaxies. Our study does not apply GAIA_CROSSMATCH for object filtering, ensuring that all possible lenses are kept. Secondly, SLDE A searched down to a magnitude of $I_E < 22.5$, which can potentially discard $\sim 27\%$ of foreground lensing galaxies with $I_E > 22.5$ per the forecast of [Collett \(2015\)](#). In this work, we extended the search to $I_E < 24.0$, theoretically reducing the percentage of lens discarded to $\sim 13\%$. Throughout this paper, magnitudes are expressed in the AB system unless explicitly stated otherwise. Finally, we implement a very different model-building approach compared with SLDE C where the strongest model Zoobot was trained from a general-purpose galaxy classification model using a transfer learning approach. In this work, we fine-tuned a CNN model from scratch with an iterative pipeline (Sect. 3.1).

One notable advantage of our approach lies in the construction of a fully independent pipeline, developed and executed from start to finish. This independence enhances confidence that the final combined catalogue – integrating results from multiple lens searches – is more comprehensive and less prone to systematic biases inherent to any single method. Such robustness is especially important for Q1, as it represents the first *Euclid* lens search and will form the foundation for subsequent catalogues and analyses.

3. Iterative fine-tuning pipeline

3.1. Pipeline architecture overview

An iterative, human-in-the-loop classification pipeline based on deep-learning methods is employed for efficient strong-lens de-

tection. The aim is to perform binary classification of galaxy cutouts into ‘lens’ and ‘non-lens’ using the model progressively refined by the pipeline. At its core is a CNN, widely recognised for extracting complex, hierarchical features from images and, thus, well suited to identifying the morphological signatures of gravitational lensing, such as arcs and Einstein rings.

For training, real *Euclid* imaging data are preferable to simulated images, ensuring the model learns features consistent with the observational characteristics of future data. While negative examples (i.e. non-lenses) are plentiful, the positive class (confirmed lens candidates) is intrinsically sparse and costly to assemble, owing to the rarity of lensing events and the need for expert annotation. A common approach in the literature is to rely on simulated lens images ([Delchambre et al. 2019](#); [Euclid Collaboration: Walmsley et al. 2025](#); [Metcalf et al. 2019](#); [Schuldt et al. 2022](#); [Thuruthipilly et al. 2022](#); [Sheng et al. 2022](#)). However, such simulations risk visual divergence from genuine *Euclid* data and can impair model generalisability – as demonstrated by [Pearce-Casey et al. \(2025\)](#), where models trained on simulations achieved high accuracy yet performed poorly on real *Euclid* observations. It is demonstrated in [Euclid Collaboration: Lines et al. \(2025a\)](#) that, by using real *Euclid* lens images and non-lens samples as a training set, the number of expected lenses improved 25–30%, in which 20% came from the use of real data compared to simulated methods.

To enhance model performance, we adopt an agile, iterative training procedure that uses simulated images only at the outset. First, a model pre-trained on simulated lenses assigns lensing probabilities to real objects in the Q1 dataset. From the highest-probability cases, experts select confident lens candidates (true positives) to assemble a positive training set of real strong lenses. Objects misidentified as lenses (false positives) and those classified as ‘non-lens’ inform the construction of the negative training set. This initial, all-real training set is then used to fine-tune a chosen CNN architecture end-to-end. Thereafter, the training set is updated iteratively from the model outputs and used to fine-tune the model in the next round. Here, fine-tuning denotes tuning selected layers of the network on progressively refined datasets so that it adapts to the nuances of real *Euclid* data. Since fine-tuning is one form of training, when the term training is used in this paper (e.g. as in the ‘training dataset’) it still should refer to the fine-tuning process where not all layers are allowed to vary, instead of the narrower definition of training where all layers can vary.

The pipeline’s dynamic design affords high flexibility in curating the training data. In each round, newly identified lenses are added to the positive set, while emerging false positives are incorporated into the negative set, which is balanced across different negative types (Sect. 3.1.3). Building a comprehensive negative set is as important as the quality of the positive images, since it increases the purity of positively classified objects. This workflow is repeated to convergence, yielding a robust classifier optimised for the identification of strong gravitational-lens systems (Fig. 1).

The early round seeds were chosen for their morphological clarity and high signal-to-noise ratio, to provide a clean foundation for iterative CNN training. Although a larger number of lenses were previously published, the independent seed set enables detection of strong lenses with morphologies distinct from prior catalogues, increasing the diversity and completeness of the final candidate list. Moreover, this work demonstrates the feasibility of expanding the number of strong lenses found through an iterative approach given limited training samples. The seed set was expanded through augmentation and regular-

isation was applied throughout to limit overfitting and shortcut learning.

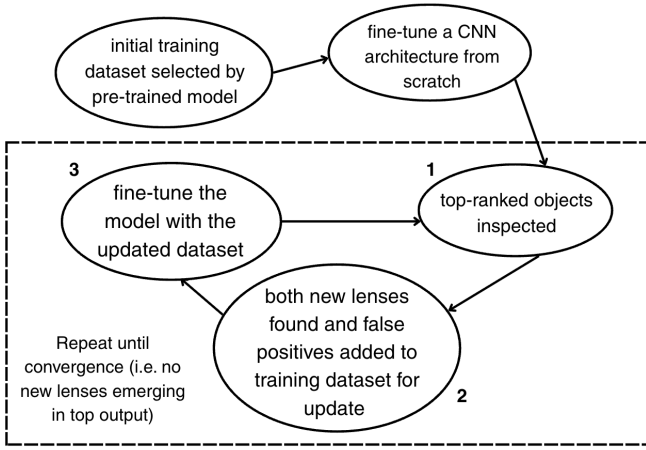


Fig. 1. Schematic representation of the iterative training pipeline. Each iteration comprises three steps: Step 1: Inspection of the top-ranked candidate objects in the model output; Step 2: Update of the training dataset through the inclusion of newly identified gravitational lens candidates and false positives; and Step 3: Fine-tuning of the model with the updated dataset. The process is repeated until convergence, i.e. when no additional lenses are identified among the top-ranked candidates.

Most current studies neither train their models over multiple rounds nor train on fully real *Euclid* data. Instead, they apply transfer learning from models trained for broader galaxy-classification tasks that are not purpose-built for lens identification. For example, Zoobot is a Bayesian CNN originally trained for general galaxy morphology and later adapted for lens detection by fine-tuning only the final layers, using a mixture of simulated and real objects in the training set (Euclid Collaboration: Lines et al. 2025b). The chief advantage of fine-tuning a related model is speed, motivated by tight time constraints: it can be faster than fine-tuning a new model across several rounds. However, we propose that training a dedicated model on *Euclid* data from the outset, for the sole purpose of lens-galaxy identification, allows all parameters to be fully utilised and optimised for this single task, improving both accuracy and runtime efficiency.

A well-performing single-channel model previously trained for lens identification (Li et al. 2023) was used to select *Euclid* VIS band images to form the initial training dataset. The details of the pre-trained model are elaborated in Sect. 3.1.2. Before running the model on *Euclid* I_E -band images, the *Euclid* Q1 data were subject to several necessary processing and cleaning stages, as outlined and reasoned in Sect. 3.1.1.

3.1.1. Image processing and cleaning

The Q1 data used in this work are composed of 344 MER tile mosaics and their corresponding catalogues (Euclid Collaboration: Romelli et al. 2025). We applied two filters based on the catalogue entries. The first, `POINT_LIKE_FLAG` $\neq 1$, excluded point-like sources, especially stars. The second criterion is $I_E \leq 24.0$ (converted from `FLUX_DETECTION_TOTAL` in μJy), since we have found that objects fainter than $I_E = 24$ would not be likely to yield detectable strong lenses.

In SLDE A, as a reference for comparison, the source selection included a cut at $I_E \leq 22.5$, with the 5th percentile corresponding to a `SEGMENTATION_AREA` of 259 pixels. Additional criteria include a selection limited to sources with `VIS_DET`

$= 1$ (i.e. detected in VIS), exclusion of sources with a *Gaia* cross-match identification, and a requirement that sources satisfy `MUMAX_MINUS_MAG` ≥ -2.6 and `MU_MAX` ≥ 15.0 – these latter two criteria remove stars and saturated objects, respectively. Furthermore, a threshold of `SPURIOUS_PROB` < 0.05 was imposed to filter out likely artefacts. This filtering resulted in a sample of 1 086 556 sources in the earlier study of selecting lenses in *Euclid* Q1 imaging data.

Each selected object was cropped into a 96×96 pixel cutout, corresponding to an angular size of $9''.6 \times 9''.6$. We chose a 96×96 cutout based on both lens morphology and network downsampling constraints. (Collett 2015) show that essentially all *Euclid*-discoverable strong lenses are expected to have Einstein radii $\theta_E \lesssim 3''$. With the *Euclid* VIS pixel scale of approximately $0''.1$ per pixel, this corresponds to a radius of ~ 30 pixels, implying that a cutout must be at least ~ 60 pixels across to fully contain the lensed features. We additionally include a margin to accommodate centring uncertainty, PSF wings, and surrounding context.

From a modelling perspective, 96 provides a convenient downsampling path: one factor-of-three reduction (e.g. a single 3×3 pooling with stride 3) maps $96 \rightarrow 32$, after which standard 2×2 pooling or strided operations cleanly reduce the dimensionality as $32 \rightarrow 16 \rightarrow 8 \rightarrow 4 \rightarrow 2 \rightarrow 1$. This keeps the cutout large enough to capture lens structure while avoiding the computational cost of 128×128 and the tight framing associated with 64×64 .

The relaxed filtering used in this study yielded a total of 5 461 976 sources with $I_E \leq 24.0$, subdivided into 2 594 180 with $I_E \leq 23.0$ and 2 867 796 with $23.0 < I_E \leq 24.0$ (Table 1). In our filtered cutouts, to enhance visual contrast for inspection, a pixel value clipping threshold at the 0.5% level was applied to each image.

Due to relaxed filtering steps, numerous instrumental artefacts persisted in our cutouts. These artefacts often mimic lens-like morphology, leading to a high false-positive rates in CNN models, as confirmed by preliminary tests. Although the application of a `SPURIOUS_PROB` < 0.05 filter, used in prior studies such as SLDE A, mitigates this issue partially, many artefacts remain and end up falsely flagged as candidates (see Fig. 2).

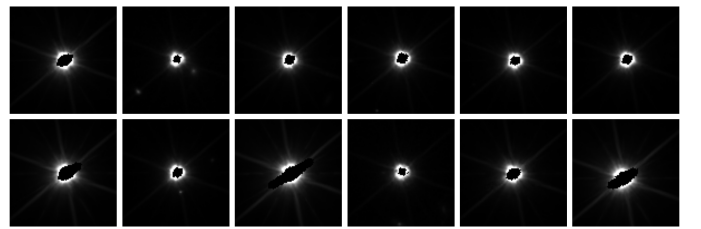


Fig. 2. Examples of randomly selected artefacts (displayed in I_E -band images) not removed by the `SPURIOUS_PROB` < 0.05 filter. In almost all cases, these are bright stars with centrally saturated pixels that are captured by L2 image flags.

To address this issue, we implemented an image-level artefact-removal algorithm to the VIS band images. Artefacts are broadly classified into two types: large and small defects. Large-defect removal is based on a 6×6 pixel sliding window with a stride of 1. If all pixels in the window exhibit identical intensity, the cutout is discarded. Small-defect removal employs a 2×2 pixel detector over the central 10×10 pixel region, a frequent location for compact artefacts.

Table 1. Number of detected sources per magnitude bin after successive quality filters.

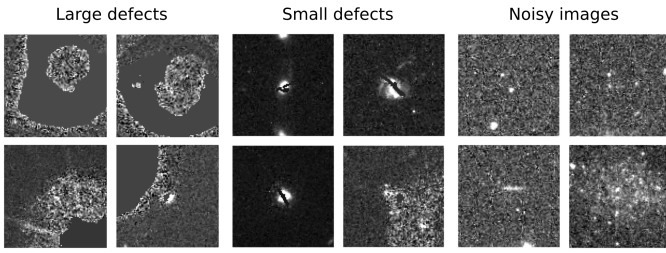
I_E	Point.	Point.+Spur.	Point.+Artefact
0–23	2 594 180	2 361 825	2 170 933
23–24	2 867 796	2 025 908	1 265 417

Notes. Column (2) contains sources with POINT_LIKE_FLAG $\neq$ 1.

Column (3) additionally applies the cut SPURIOUS_PROB $<$ 0.05.

Column (4) further includes the artefact and noise cleaning procedure.

In addition, noisy images were identified as a recurrent source of misclassification. To mitigate this, a final noise-based filter assesses the outer region of each image, excluding the central sector to preserve bright core features. Images with excessive background brightness in the outer region are removed.

**Fig. 3.** Examples of VIS band images images containing defects made of large and small pixel areas, and noisy images that are removed by our artefact removal algorithm.

The complete cleaning pipeline results in a well-vetted dataset suitable for strong lens identification. The statistics are outlined in Table 1, where we also compare the cleaning effect of SPURIOUS_PROB $<$ 0.05 and our artefact cleaner separately. The final *Euclid* Q1 dataset with minimal contamination includes 2 170 933 objects with $m \leq 23.0$ and 1 265 417 objects with $23.0 < m \leq 24.0$. A total of 3 436 350 detected sources were retained.

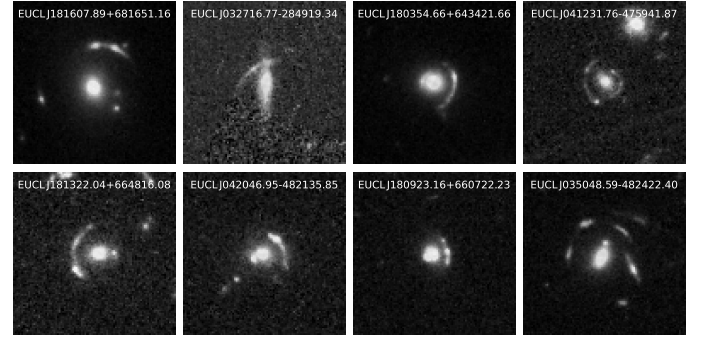
3.1.2. Pre-trained single-channel model

From the processed *Euclid* imaging cutouts, we selected an initial training sample using the pre-trained single-channel CNN described by Li et al. (2023). This model had been trained on 60 000 semi-simulated cutouts to distinguish lenses from non-lenses using I_E -band data. The positive class was constructed by pairing real galaxies from the *Hubble*/CANDELS survey (Grogin et al. 2011; Koekemoer et al. 2011) at various redshifts into foreground-background pairs. A singular isothermal ellipsoid (SIE) lensing potential was applied to distort the background galaxy, with its orientation and ellipticity aligned to the foreground light profile and a randomly sampled Einstein radius. The resulting lensed background was overlaid onto the foreground cutout to produce realistic strong-lensing configurations at *Hubble* resolution. These images were then convolved, resampled, and injected with noise to match *Euclid* quality. Two CNNs were trained independently on simulated VIS and NISP images, yielding AUC scores of 0.987 and 0.797, respectively. Owing to its superior performance, the I_E -band model was selected to perform binary classifications on the Q1 VIS cutouts.

3.1.3. Selection of positive and negative training samples

There are 73 825 I_E -band images positively classified (with an assigned prediction value greater than 0 with the sigmoid final layer) as lens candidates by the pre-trained single-channel model. We took quick observations of the top 10 000 candidates by looking at slices at different cut-offs, and we set 1000 as a reasonable cut-off for internal expert team inspection. Among the top 1000 images given the highest lens probability, we examined and selected true positives and false positives by giving human ratings. Confidently real lenses as true positives were selected to build the initial positive training set, and false positives for the negative training set.

In the initial training dataset selection stage of the pipeline, each object received ten expert votes between lens and non-lens. Twenty-seven clearly true positives were selected, with a random subset shown in Fig. 4. The false positives resemble the morphological patterns of lenses, with the resemblance decreasing at lower assigned probability values. The ~ 70 000 false-positive images mostly consist of four typical types, with unequal distribution: spiral galaxies, elliptical galaxies, edge-on galaxies, and diffraction spikes, separately shown in Fig. 5.

**Fig. 4.** Examples of confident strong lens images used for building the initial *Euclid* positive training set based out of VIS-only detection scheme used at Step 1. We augmented these initial detections from 27 candidate lenses to a total of ~ 2000 true positive cases using a combination of rotation, translation, and colour transformations.

In building the positive training set, the 27 true positives in I_E -band were converted to colour version (Sect. 3.2.1) by merging with Y_E -band and H_E -band imaging data. Then, they were augmented to around 2000 images using the methods outlined in Sect. 3.2.3.

For building the negative training set, although abundant data are naturally available, it is essential to limit its size to that of the positive set in the same round so that the model is not affected by potential bias in the training data. A strongly imbalanced ratio (e.g. 1:10 positive-to-negative) would cause the loss function to be dominated by negative examples, shifting the decision boundary towards predicting the majority class and suppressing recall of true lenses. Moreover, the predicted probabilities would reflect the artificial training prior rather than the astrophysical prior, leading to systematic underestimation of lens likelihoods. It is also crucial to maintain the diversity and balance of the negative training set to prevent systematic bias and improve model robustness. In fact, it has been demonstrated by C  n  meras et al. (2024) that a balanced negative training dataset for lens classification tasks leads to better model performance. Thus, although objects such as spiral arms dominate the false positives, we would need a balanced negative set for different types of false positives.

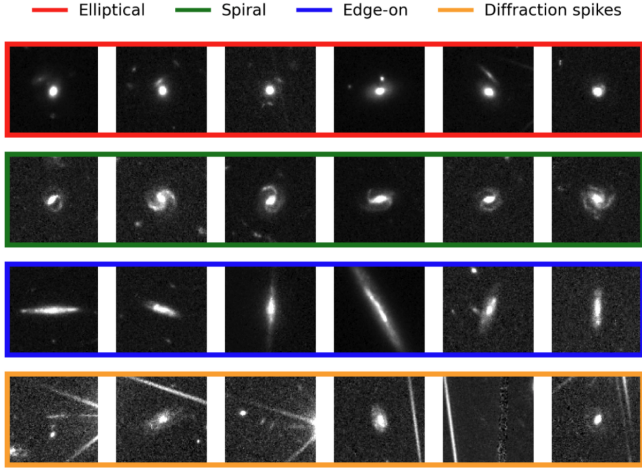


Fig. 5. Examples of false positives identified from the output of pre-trained model. These false positives consist of four typical types: elliptical galaxies (*top row*), spiral galaxies (*second row*), edge-on spiral galaxies (*third row*), and images containing diffraction spikes of a bright star close to a galaxy or another source (*bottom row*).

Table 2. Components and sizes of the initial *Euclid* imaging training set later used to fine-tune models and evaluate their performance.

Initial set	Component	Size
Positive	27 lens candidates, augmented 67-fold	1809
Negative	Balanced sample of:	2000
	• spiral-arm galaxies	
	• elliptical galaxies	
	• edge-on galaxies	
	• diffraction spikes	

We fine-tuned an EfficientNetB0-based CNN model (Tan & Le 2020) to help select false positive images. For each category (spiral arms, elliptical, edge-on galaxies, and diffraction spikes), we manually selected around 100 samples to build a balanced training set. The trained EfficientNetB0 model achieved a categorical accuracy of 90%, classifying the $\sim 70\,000$ false positives into four different groups. This classification does not intend to accurately identify specific galaxy types; rather, it divides the false positives into reasonably different groups regarding their distinct morphologies to build a balanced negative dataset. Next, 500 images were selected from each group, followed by quick visual confirmation, to build a negative training set of 2000 in size and these were also converted to colour images. The full statistics of the initial training set is presented in Table 2.

3.2. Data preparation

3.2.1. Construction of colour datasets

To adapt to applying a three-channel model in the iterative pipeline and for the purpose of improving model performance from data quality, we created colour versions for selected Q1 images and all training data. The colour images were generated by incorporating the H_E ($\sim 1.5\text{--}1.8\,\mu\text{m}$) and Y_E ($\sim 0.97\text{--}1.07\,\mu\text{m}$) bands from the *Euclid* NISP instrument. To construct colour composites of the original I_E dataset, we produced cutouts at identical sky positions in the H_E and Y_E mosaics using the source catalogues, and passed these cutouts to an algorithm to combine them into a single RGB image.

This algorithm comprised two main components: a colour-combination function and a luminance adjustment function. In the colour-combination function, each of the H_E , Y_E , and I_E cutouts were first normalised to the range $[0, 1]$, and then divided by their median pixel values prior to ensured a uniform black background, which improved visual contrast between the lens and the source galaxy. In accordance with the ordering of wavelengths for each of the bands, we composite the processed $H_E/Y_E/I_E$ cutouts into the R/G/B channels to form an intermediate colour cutout.

Due to the inherently lower spatial resolution of the NISP data compared to I_E , we applied an additional luminance correction step to project the morphological features of the I_E image onto the final colour image. A similar approach was also adopted in the work of Euclid Collaboration: Bergamini et al. (2025). Specifically, we extracted the luminance image of the intermediate colour cutout using the ITU-R Recommendation BT.709 standard¹, computed per pixel as $L = 0.2126R + 0.7152G + 0.0722B$. We then divided the I_E image by this extracted luminance image and multiplied the resulting pixel-wise normalisation factor to the intermediate colour cutout. This procedure adjusted the contrast of the colour image to match that of the higher resolution I_E data, yielding the final set of colour cutouts, stored as 96×96 pixel TIFF images with float32 precision.

3.2.2. Colour image cleaning

Following the creation of these colour version images, certain artefacts became apparent, particularly objects not covered by the corresponding H_E or Y_E bands. To maintain a consistent standard of data quality, we excluded all images lacking the complete set of three bands from the dataset. Up to this stage of filtering, we have retained 3 423 516 objects from the Q1 data, cleaned, in colour versions.

3.2.3. Colour image augmentation

In the data preparation stage of the first round of the iterative process, we applied image augmentation to the 27 true positives. We also applied augmentation to the increasing number of true positives in each subsequent round. This is a crucial step in the early stages of development, when true positives were exceedingly scarce. The augmentation procedures included horizontal and vertical flipping, affine transformations (e.g. zooming), and systematic variations in contrast, brightness, and saturation. Additional augmentations combined these operations, such as flipping with concurrent contrast and brightness adjustment, or simultaneous modifications of brightness, saturation, and contrast. These augmentations were implemented using *imgaug* sequences (Jung 2018), generating a diverse set of training examples with parameter variations typically in the range 10% to 20%. Using this protocol, we increased the number of positive images in each dataset by a factor of 67 (see Table 2). A full scheme of augmentations used is outlined in Appendix A for reproducibility.

3.3. CNN model training and fine-tuning

3.3.1. Selection of pre-trained architectures

To fine-tune a model from scratch, we selected six lightweight CNN architectures (Table 3) to best match the complexity and

¹ <https://www.itu.int/rec/R-REC-BT.709-6-201506-I>

dimensions of *Euclid* galaxy images. Each architecture was fine-tuned with the dataset in Table 2 using its optimal set of frozen layers, and then used to make predictions on our selected Q1 data. Their best performance was evaluated by counting the confident lenses flagged by experts from their top 128 predictions (as we found few lenses after this number), as summarised in Table 3. Among the six architectures, the VGG16 architecture (Simonyan & Zisserman 2015) achieved the best performance, and was thus selected as the model architecture for subsequent fine-tuning rounds in the iterative process.

Table 3. Performance of candidate CNN architectures fine-tuned on the initial training set.

Model name	Lenses found	Lens fraction
VGG16 (Simonyan & Zisserman 2015)	34	0.27
Xception (Chollet 2017)	27	0.21
ResNet50 (He et al. 2015)	24	0.19
InceptionV3 (Szegedy et al. 2015)	23	0.18
MobileNetV2 (Howard et al. 2017)	8	0.06
EfficientNetB0 (Tan & Le 2020)	6	0.05

Notes. ‘Lenses found’ gives the number of expert-graded, high-confidence strong-lens candidates among the top 128 highest-ranked objects predicted by each model. The ‘lens fraction’ is the corresponding fraction relative to the 128 inspected candidates.

3.3.2. Fine-tuning strategy (VGG16)

As can be seen from Table 3, there is still much room to increase the number of lenses found in top-ranked objects for VGG16. Although some progress has been achieved (34 lenses identified) with the initial training set, its diversity and volume can be further improved. The iterative fine-tuning process is designed to gradually enhance the performance of the model with a manually updated *Euclid* imaging dataset and careful adjustment of model parameters in each round.

To further increase the fraction of true positives in the top predictions, we specifically modified the original VGG16 architecture to better adapt to the complexity and dimensions of galaxy images. The original final layers of VGG16 were replaced with a customised layer set, including a GlobalAveragePooling2D layer (to reduce spatial dimensions), a 256-unit fully connected layer and a 128-unit fully connected layer (to learn high-level feature representations), and a sigmoid binary classification layer (which outputs classification results between 0 and 1). The modified model architecture is visualised in Fig. 6.

During each fine-tuning iteration, the number of trainable layers is determined based on both the fraction of lenses among the highest-ranked candidates and the model’s classification accuracy on a balanced test set. In the stage involving fine-tuning of the VGG16 architecture from scratch (see the top-right panel of Fig. 1), the optimal performance is achieved when the last nine layers are kept trainable. This is done by allowing the parameters of these layers to be updated while all preceding layers remain frozen. This iterative fine-tuning process is continued until a sufficiently large number of strong lens candidates appear among the top-ranked predictions.

3.3.3. Evaluation metrics and validation

In conventional supervised learning, model performance is typically assessed using held-out validation or test sets that contain labels unseen during training. Such an evaluation is effective only when the withheld data are sufficiently numerous, representative, and disjointed from the training pool. For strong gravitational lensing, however, this paradigm is not directly applicable. *Euclid* Q1 contains very few confirmed strong lenses, and many of the high-quality systems have been used (iteratively and without persistent identifiers) to refine the model itself. As a result, a large, purely independent test set of real *Euclid* lenses for a reliable recall measurement is difficult to assemble from Q1 alone. In Sect. 4.1.3, we perform an empirical cross-check against the random inspection sample described in Sect. 8.2 of *Euclid Collaboration: Walmsley et al. (2025)*, but its non-nested pre-selection, small lens count, and variance in human classification limit its use, so we adopt the recovery fraction relative to the forecast population of Collett (2015) as our forecast-normalised yield.

Because of this limitation, classical supervised-learning metrics such as accuracy, precision, and recall cannot be reported in their standard form. Instead, evaluation must reflect the operational context of the *Euclid* lens-search problem: identifying a small number of genuine lenses within millions of galaxy cutouts. Rather than relying on artificial balanced test sets, whose statistical properties do not resemble the *Euclid* population, we can therefore assess performance by examining how well the model prioritises likely lenses in its ranked candidate list.

Our primary evaluation metric is thus the recovery of known or visually confirmed lens candidates within the top- K ranked candidates, compared to the expected population of ~ 670 strong lenses in Q1 with $I_E < 24$ (Sect. 2.3). This approach provides a practical measure of the model’s capability: a method that successfully concentrates true lenses towards the top of the ranking directly reduces the human inspection effort required for scientific follow-up. Moreover, because this lens-search task is fundamentally dominated by extreme class imbalance, prioritisation-based evaluation more accurately captures real-world model behaviour than hypothetical accuracy-based metrics.

Another useful metric would be the measure of the continual improvement of the model throughout training: how successful iterations consistently increase the concentration of lens-like systems among the highest-ranked candidates. This progressive enhancement in ranking quality provides strong qualitative and quantitative evidence that the model learns more effective decision boundaries over time (elaborated in Sect. 3.3.4), even in the absence of a clean, fully independent *Euclid* test set.

3.3.4. Retraining and evaluation

Human vetting was performed in two modes. During the iterative training, we prioritised speed and use internal expert inspection of the top-ranked candidates, typically reviewing the top 4000 objects per iteration; this produced a labelled update to the training set on week timescales. For the final evaluation, we complemented the expert inspection with a larger scale grading campaign using established citizen-science style tooling, in which each object receives multiple independent votes. This two-stage design separates rapid iteration (to improve the ranking model) from broader community-level validation (to quantify grading consistency and provide higher confidence candidate labels).

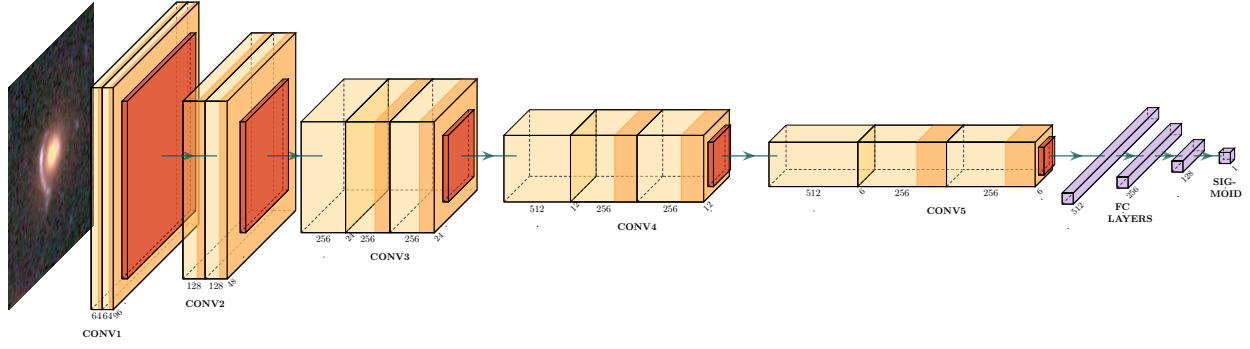


Fig. 6. Architecture of the modified VGG16 convolutional neural network used for lens classification. The original fully connected layers were replaced with a customised final structure, labelled in purple. This includes a GlobalAveragePooling2D layer to reduce spatial dimensions, followed by two dense layers with 256 and 128 units to capture high-level features, and a final sigmoid layer for binary classification.

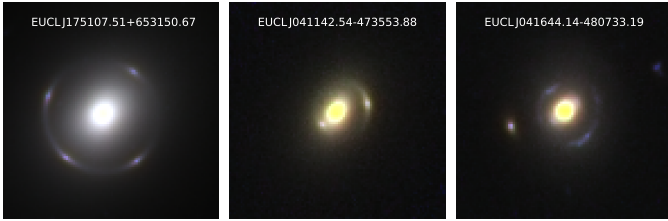


Fig. 7. Three newly identified (relative to the previous iterative round) Grade A strong lens candidates from the fine-tuned model in the second round that exhibit atypical ‘lemon’ and white colouring compared to most other lensing systems that show a colour difference between the foreground lens and the background source. The first system (left) is NGC 6505, a complete Einstein ring, discussed by O’Riordan et al. (2025) based on the visual identification of the ring.

In the first iteration, we started by inspecting the top-ranked objects (Step 1) in Fig. 1. In Step 2, there were 123 strong lens candidates selected and augmented to 8241 to optimise the positive training set. Around 900 new false positives were picked and added to the previous balanced negative training set – this formed an updated negative training set of approximately 8900 objects.

To further enhance the diversity of the positive training set, we created 100 hand-drawn lenses using the software Procreate² (Fig. C.1). The artistic lenses are purely experimental and are not derived from physical lens models or mass distributions. They were designed to test classifier sensitivity to extreme or atypical morphologies as a qualitative diversity regulariser. The artistic inspiration came from observing the dynamics and traits of the existing strong lens candidates found so far. Careful variations were designed to increase data diversity and robustness. These art lenses were added to the positive training set for fine-tuning the model. Among the approximately 1000 top-ranked objects of the fine-tuned model, by manual inspection, there were 229 strong lens candidates which were augmented and added to the positive training set. Among the 229 candidates, three new special lenses (relative to the previous round; Fig. 7) should be noted. These lenses display traits different from most other strong lenses, appearing more lemon-coloured and white based on our colour processing method. Specifically, the first one in Fig. 7 is NGC 6505, a complete Einstein ring, first identified by O’Riordan et al. (2025).

In the second iteration, the three lenses were augmented to 201, with a higher weight in the positive set to help the model grasp the traits of diverse lenses. There were dot-like objects, appearing in green with our colour grading method, among the false positives of the second iteration. We added them to the negative training set to allow more true positives – we pooled all the lenses we found up to that point and augmented them, creating a dataset of 30 686 images, with 15 343 each for positive and negative images. This dataset was used to produce the final iteration of the model.

In total, we fine-tuned the model three times, each resulting in noticeable improvements in performance (see the improving recall curve in Fig. 8; Sect. 4.1.3 elaborates on how we generated a list of lenses as ground truth in producing Fig. 8). For all training iterations, optimisation was performed using the Adam optimiser (Kingma & Ba 2014) with a learning rate of 5×10^{-5} , a batch size of 64, and a total of 15 epochs. During training, the model retained the final sigmoid activation so that outputs could be interpreted as probabilities, allowing the binary cross-entropy loss function to be applied effectively.

3.4. Inference on Q1 dataset

3.4.1. Linear output as raw scores

To evaluate the output of the neural network for predictive purposes, we found it to be more beneficial to remove the final sigmoid layer and use the raw, unbounded output of the neural network’s last dense layer; we will denote it as the so-called raw score in the rest of the paper. This is due to the fact that the sigmoid function compresses all values to a narrow range between 0 and 1, thereby losing granular information about the model’s confidence for predictions that fall within the region where the sigmoid function plateaus. This issue is particularly critical in the context of this paper due to covariate shift, where the statistical properties of the real-world images can differ significantly from the manually selected training images.

For instance, variations in instrumental signatures or unexpected artefacts can cause the distribution of the model’s output scores to drift. When this occurs, the sigmoid function maps these shifted scores into its saturation regions, making it impossible to distinguish between varying levels of model confidence. In contrast, the raw output retains the full magnitude of the model’s conviction and provides a wider distribution of scores. This makes the raw score a more sensitive input for ranking-based evaluation and allows for a more nuanced selection of a

² Procreate®, Savage Interactive Pty Ltd, <https://procreate.com>

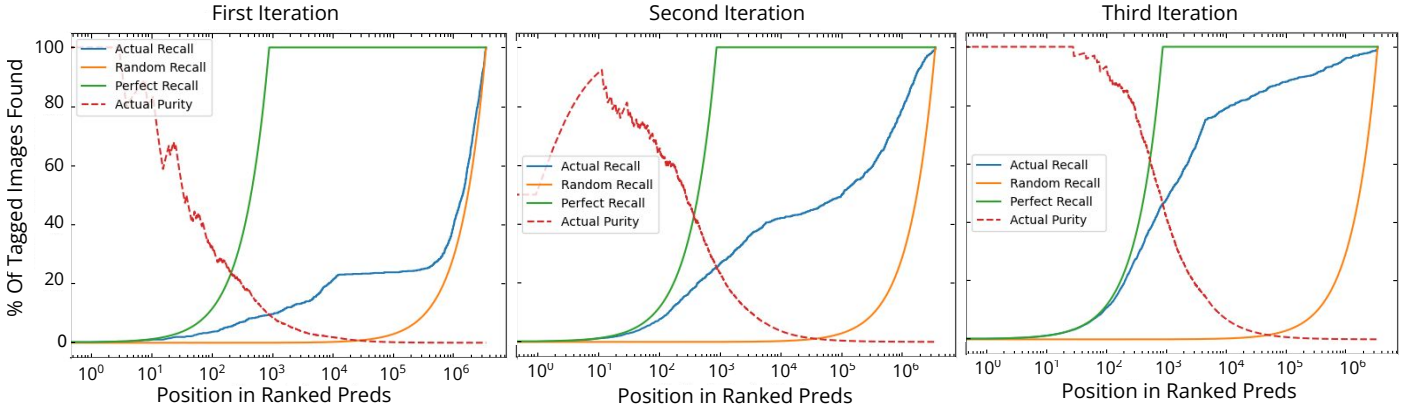


Fig. 8. Three rounds of the iterative model retraining process cross-validated with 905 sources (containing off-centred samples) from this work and SLDE A/B/F. The left-hand plot denotes the first iteration trained on the initial training dataset selected by the pre-trained model. The blue curve shows recall, that is, where in the ranked predictions we could find each of the confirmed lens candidates. The green curve denotes the recall of a perfect classifier, while the orange line denotes a completely random classifier. The red curve denotes the precision of the current model, that is, how well the model moves non-lenses to lower rankings. Throughout development, we found that the actual recall curve mostly overlaps with the perfect classifier in the top 100 predicted candidates.

decision threshold to better control the trade-off between precision and recall.

3.4.2. Human evaluation and original grading results

We selected the top 4000 candidates, without duplicates, ranked by a raw score, as evaluated from our final model, as lens presence declines drastically thereafter (Fig. 9). The choice of top 4000 was quantitatively validated using classification results, as explained in Sect. 4.1.1.

The top 4000 images were visually inspected and graded by lensing expert volunteers from the *Euclid* Collaboration. During the classification process, each participant was shown a randomly selected image from the top 4000 and prompted to assign it to one of the following categories: A+ (lens of individual scientific value), A (confidently a lens), B (probable lens, further information needed), C (lens-like, other explanations possible), or X (not a lens). The displayed images were in PNG format due to the platform’s compatibility. While the model input images were originally stored as TIFF files with float32 precision, we manually verified that the PNG versions used for classification exhibited no perceptible visual differences from the original TIFF files. Each image was evaluated by ten independent assessors, resulting in a total of 40 000 ratings.

For reference, in SLDE A, each candidate object likewise received ten expert votes and a calibration was performed by comparing each expert’s average grades with those of their peers on the same galaxies. If, for instance, an expert tended to assign systematically higher scores than the consensus (e.g. often giving candidate lens systems an A when the group average was B), their grades were adjusted downward to offset this optimism. However, this calibration method does not apply to this work, as there were 27 experts participating in the classification of our 4000 objects, each contributing to a different extent, from a ten classifications to all 4000. It is impractical to find groups with sufficiently large sample sizes that also received votes from the same sufficiently large number of experts. For instance, given an object that received 10 votes from 10 different experts, the grade of one expert is more optimistic than the grades of the other nine experts; the calibration method in SLDE A would then consider that expert as optimistic overall and discount their grade to all galaxies they classified. However, the conclusion of that expert

being generally optimistic can only be made when there is a sufficient set of objects which received votes from those exact same ten experts for comparison. Due to the vast difference in the number of classifications contributed by each expert, we applied a different calibration method in Sect. 3.4.3.

3.4.3. Grade aggregation

Based on the original classification results, on a quick inspection we found that many strong lenses that also received human ratings in SLDE A underwent a ‘downgrade’ with a mode-first rating processing method. For instance, if an object received two A, three B, one C, and four non-lens votes, the mode-first method would assign a non-lens classification, even though lens-like features are present, as indicated by the A/B/C votes. To avoid such a collective downgrade and to ensure lens candidates remain visible, we implemented the following post-processing threshold:

Algorithm 1 Grade aggregation threshold rule

- 1: **if** count(A+) is a majority **then**
 - 2: classify as grade A+
 - 3: **else if** count(A+) + count(A) is a majority **then**
 - 4: classify as grade A
 - 5: **else if** count(A+) + count(A) + count(B) is a majority **then**
 - 6: classify as grade B
 - 7: **else if** count(A+) + count(A) + count(B) + count(C) is a majority **then**
 - 8: classify as grade C
 - 9: **else**
 - 10: classify as ‘not a lens’
 - 11: **end if**
-

4. Results

4.1. Lens candidate identifications

4.1.1. Human evaluation results

The visual inspection was conducted within the Galaxy Judges project on the Zooniverse platform. The results of the processed ratings are shown in Table 4, which mitigates the downgrade effect of the mode-first processing method to some extent, as nu-

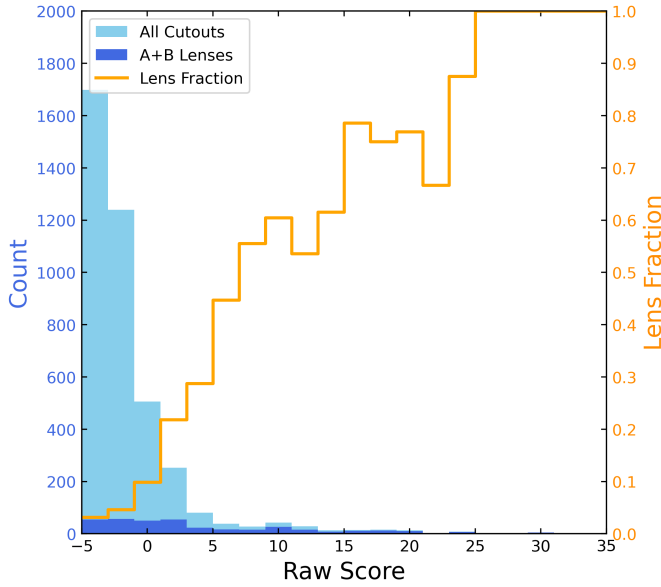


Fig. 9. Raw score of the top 4000 cutouts ranked by the model as light blue bars and the dark blue bars represent the identified A/B lens candidates within the 4000 cutouts. Orange line shows the fraction of lens candidates, computed as the ratio of dark blue over light blue in each bin. On the lower raw score end, the fraction of lenses declines to less than 0.05. This justifies the threshold for visual inspections as lenses beyond top 4000 are too rare for effective human inspections.

merous objects have moved to higher grades. Using the grade calibration procedure in Algorithm 1, we have further mitigated the downgrade effect to a large extent.

Table 4. Distribution of post-processed classification outcomes from *Euclid* expert ratings for the model’s top 4000 predicted objects.

Processed classification	Vote No.	Fraction (%)
Grade X: not a lens	3095	77.4
Grade C: lens-like feature	464	11.6
Grade B: probable lens	314	7.9
Grade A: confidently a lens	111	2.8
Grade A+: lens of science value	16	0.4
Total	4000	100.0

For clarity, lenses graded as A+ have been merged into Grade A. For further analysis, we will mostly consider grade A and grade B lens candidates since these candidates represent the more confident lens detections. The raw score distribution of these lens candidates are visualised in Fig. 9, which justifies the selection of only top 4000 ranked cutouts for visual inspections.

4.1.2. Variance in human evaluation results

For lenses recovered in this work and in SLDE A, we concatenated our ratings and their ratings to produce a summary of results in Fig. 10. Among the combined grades, there are 116 lenses that received a downgrade and 136 lenses that received an upgrade, resulting in a rough balance between grade changes. However, these discrepancies also indicate an inherent variance in human ratings, as demonstrated in Rojas et al. (2023). To investigate possible reasons for such variance, we show examples

of downgrade in Fig. B.1 and upgrade in Fig. B.2 as separate panels.

We investigated a number of reasons for the differences in human ratings between this work and the previous study. One cause of this discrepancy could be in image visualisation and processing. The use of VIS as the luminance component in the colour composites may have enhanced the perceived resolution of lens features. This could contribute to improved human grading accuracy and CNN performance, and highlights the importance of image construction methodology in lens identification pipelines. Additionally, the images in this work were visualised with percentile clipping, which can cause certain features to be fainter than that with a non-linear stretching function such as arcsinh. However, many of the images that dropped from Grade A to Grade B do not appear to be due to the presence of faint features (Fig. B.1). Since one image is displayed for each object, a panel consisting differently scaled images may influence the grades received. Moreover, two different groups of people contributed to the grading process of SLDE A and this work, sharing different expectations and experience. The large number of downgraded candidates demonstrates the essential role in accounting for rating variance when using human labels as a recognised benchmark for model performance. Recognising these differences should also help evaluate the true performance of a machine-learning model and potentially establish more robust metrics for the comparison of different machine-learning models.

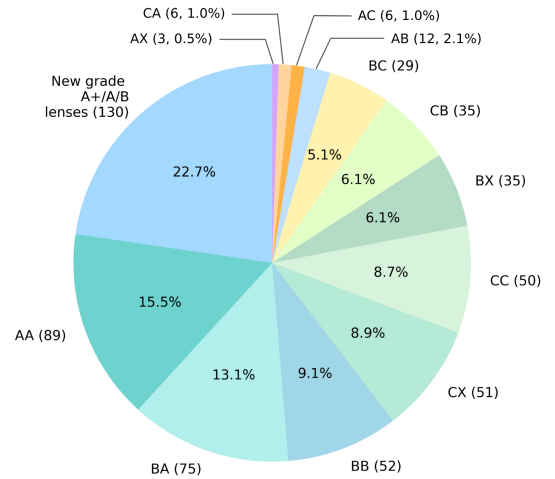


Fig. 10. Pie chart showing distribution of 130 new lenses identified and lenses previously discovered by SLDE A in the top 4000 predictions. Each grade is a concatenation of grade in this work and in SLDE A. For instance, the BA slice indicates that 75 (13.1%) lenses received grade B in this work but grade A in SLDE A.

4.1.3. Final model (VGG-16) performance

While the development of this model progressed independently of SLDE A, upon the release of their lens catalogue, we found it useful to validate our results. As model development progressed, we observed consistent improvements to both the recovery of known lenses and to the purity of the highest-ranked predictions (Fig. 8). The collection of lens samples used for validation is discussed next.

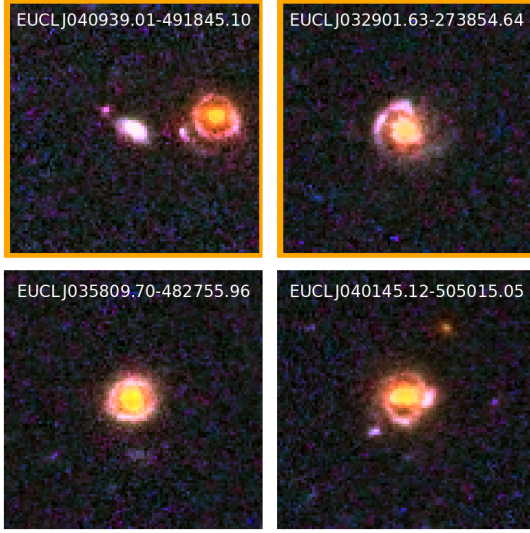


Fig. 11. Four new lenses within the flux magnitude thresholds used in SLDE A but not identified as lens candidates in that work. Objects outlined in orange were identified as lenses by previous models in SLDE C but rejected by citizen scientists in Space Warps, a classification project hosted on Zooniverse platform. The other two were poorly ranked by SLDE C machine learning models.

Following the classification procedure described earlier, we identified a total of 441 Grade A and B lenses among the top 4000 ranked objects. These 441 objects include 302 lenses previously classified as Grade A/B in SLDE A, and 9 lenses identified in SLDE F. This results in a total of 130 new lenses not previously identified in *Euclid* Q1 data.

To evaluate the final model in a manner suited to the extreme class imbalance of the *Euclid* lens-search problem, we focus on how effectively the classifier concentrates genuine lenses within its highest-ranked predictions. Because *Euclid* Q1 does not contain a sufficiently large or independent set of real, unused lenses suitable for traditional supervised-learning metrics, we instead assess performance using a curated reference sample of 905 high-confidence, expert-graded lens candidates (including 130 newly identified lens candidates in this work, the NGC 6505 lens, and 500/3/9 Grade A or B systems from SLDE A/B/F with associated duplicates and off-centred cutouts). Among the model’s top 1000 ranked candidates, 427 of these 905 known lens candidate images appear, meaning nearly half of all high-confidence lens candidate systems in Q1 lie within the top 0.03% of the 3.4 million cutouts. This demonstrates the model’s capability to remove non-lenses, offering substantial reduction in human inspection effort.

Broadening the threshold illustrates the model’s ability to recover an increasingly large fraction of true lenses while maintaining meaningful prioritisation. The top 4000 candidates cover 644 known lens images (including duplicates), capturing progressively more of the 905 lens reference sample. Removing the duplicates yields 441 distinct candidates in the top 4000 candidates, which is 65.8% of the 670 lenses expected for $I_E \leq 24$. This indicates that the model assigns elevated scores to a wide range of lens morphologies and continues to retrieve true systems even at lower confidence ranks. Together with the consistent improvement trends observed across training iterations (Fig. 8), this concentration-based evaluation shows that the final model provides a robust, practical, and effective tool for surfacing strong gravitational lenses in large-scale survey data.

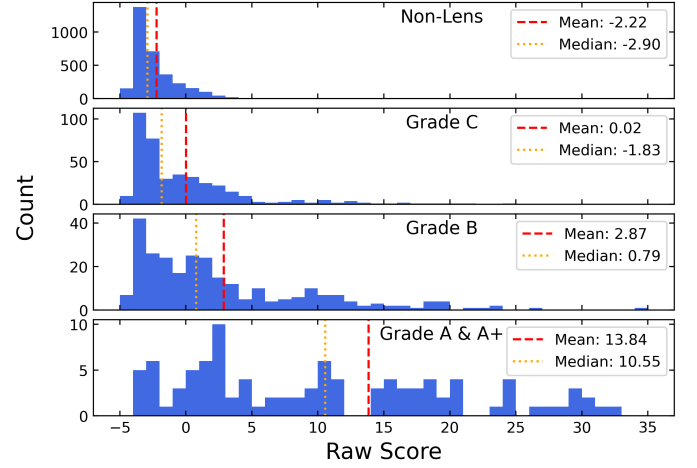


Fig. 12. Distribution of model prediction raw scores for candidates grouped by human-assigned grade: non-lens, C, B, and A. Each subplot shows a histogram of prediction values, with red dashed lines indicating the mean and yellow dotted lines indicating the median for that grade. As the grade increases, the mean and median of the prediction values shift towards the right, indicating a higher prediction value. Thus, we can infer that the prediction value from our model correlates with a higher grade.

As a cross-check against an externally graded sample, we matched our candidates to the 40 000 visually inspected random objects described in section 8.2 of *Euclid Collaboration: Walsley et al. (2025)*, of which 31 received grade B or above. That sample was drawn under the SLDE A selection and is non-nested with our parent sample, excluding the fainter and *Gaia*-matched regime that contributes many of our new candidates (Sect. 4.1.4). Of the 40 000, our model surfaces 12 of the 31 grade-B+ objects within the top 4000 inspected candidates; on re-grading, we assign 7 a grade of B or above and 5 a grade below B, and additionally grade 2 objects as B that were placed below B in the citizen-science pass. These disagreements are consistent with the variance in human evaluation discussed in Sect. 4.1.2.

As a population-normalised figure of merit, our 441 grade A/B candidates correspond to $441/670 = 65.8\%$ of the expected Q1 lens population after the $I_E < 24$ selection. This should not be interpreted as a calibrated completeness, since the denominator is a forecast rather than a known truth set. A full completeness measurement requires a selection-function analysis, for example through injection and recovery of simulated lenses in real *Euclid* backgrounds.

By analysing the model prediction scores and the human-assigned grades for the top 4000 candidates, we observe a clear correlation between the predicted probabilities and the human classifications (Fig. 12). Higher model scores are associated with a greater likelihood, as determined by experts, that a given object is a gravitational lens. This is consistent with the grading rubric used in the visual inspection process, where higher grades reflect more prominent and unambiguous lensing features. The observed correlation suggests that the model has successfully learnt to identify and rely on visual features similar to those used by human classifiers to assess lensing confidence – such as the presence of extended arcs and the geometric alignment between a foreground deflector and background sources.

4.1.4. Sample validation

To explain the source of the newly classified 139 grade A/B candidates (130 new candidates solely given in this paper and 9 candidates identified in both SLDE F and this paper) relative to SLDE A, an important factor to consider is the more inclusive filtering criterion applied in pre-selecting sources, which accounts for 52 new A/B lenses. If the same filtering from SLDE A is applied to these 52 candidates, 12 are removed by requiring no *Gaia* cross-match (which are generally sources with $I_E \lesssim 19$), then 38 are removed by the magnitude cut ($I_E \leq 22.5$), and finally SPURIOUS_PROB > 0.05 removed the last 2 candidates. Of these 12 new lenses collected by relaxing the *Gaia* cross-match criterion, 9 were also recovered in the work of SLDE F, which specifically examined sources previously removed by the *Gaia* cross-match filter applied in SLDE A, and found 68 grade A and B lens candidates relative to SLDE A. Taking SLDE F in consideration, the number of new A/B candidates in this work is adjusted to 130. The rest of the new candidates not removed by the filtering criterion ($139 - 52 = 87$ candidates) is mostly dominated by lower confidence grade B lens candidates, and can be a result of variance in human classification and the model's unique selection function.

An additional discrepancy between this work and SLDE A concerns the candidates examined in both studies that received significantly different scores. Specifically, this includes 4 candidates classified as Grade A in this work, where two were poorly ranked by the models in SLDE C and did not reach the visual inspection stage, while the other two were visually inspected but rejected by citizen scientists in Space Warps, a galaxy classification project on Zooniverse platform used in SLDE A. These four lenses are shown in Fig. 11. This divergence may be attributed to differences in visualisation schemes, where variations in intensity and colour scaling could lead to interpretation discrepancies or potentially overly optimistic classification.

4.1.5. Analysis of false positives and model robustness

Among the objects incorrectly identified as lenses by the model, several recurring morphological types emerge; namely, diffraction spikes, spiral galaxies (particularly barred spirals), isolated red elliptical galaxies, and galaxies with nearby companions. In many cases, these false positives can be attributed to structural similarities with genuine strong lenses. For example, spiral galaxies often exhibit a bright central bulge and a circular or ring-like morphology, features that may resemble lensing arcs. A representative panel of arbitrary false positives is shown in Fig. 13. Nevertheless, in the majority of cases, the model successfully prioritises clearly identifiable lenses over more ambiguous or misleading candidates.

These findings suggest that the model's susceptibility to false positives arises primarily from structural degeneracies, rather than random misclassification. While certain morphologies – such as barred spirals or ellipticals with nearby companions – can mimic lens-like features, the low prevalence of such cases among top-ranked candidates indicates that the model retains discriminative power even in morphologically ambiguous regimes. This behaviour reflects a degree of robustness and highlights the importance of incorporating domain-specific priors or multi-band information in future iterations to further suppress false detections, elaborated in Sect. 5.



Fig. 13. 20 false positives within the top 4000 classifications of the model. Within these we can see an example with a diffraction spike, several elliptical and spiral galaxies, with many containing evidence of recent mergers or an ongoing merger in the form of a tidal stream, as well as cut-outs containing multiple objects.

4.2. Characteristics of discovered lenses

One of the key distinguishing features of the lens candidates analysed is their distribution in I_E , inferred from FLUX_DETECTION_TOTAL. As shown in Fig. 14, the high-confidence lens candidates within $17 \leq I_E \leq 24$, have a median of 21.2, a mean of 21.1, and an interquartile range (IQR) of 20.3–22.0. The minimum and maximum observed values are 17.1 and 23.9, respectively. To further characterise the photometric properties of the sample, we examined the $Y_E - H_E$ colour as a function of I_E for the classified lenses, in comparison with the full parent population (Fig. 15). We find that lenses consistently exhibit higher $Y_E - H_E$ values, indicating redder colours relative to the general galaxy population. This trend is consistent with expectations that lensing galaxies are typically massive early-type systems. To statistically confirm these findings, we performed two Kolmogorov–Smirnov (KS) tests on the I_E and $Y_E - H_E$ distributions, yielding p -values of 3.5×10^{-64} and 8.1×10^{-118} , respectively. These extremely low p -values strongly reject the null hypothesis that the lensing galaxies are randomly drawn from the full galaxy population observed in the imaging data.

5. Discussion

5.1. Strengths and limitations of developmental pipeline

The greatest strength of the model development pipeline we present in this paper is its efficiency, requiring only limited involvement of human resources as such involvement is a bottleneck in large-scale astronomical surveys (Fortson et al. 2011). During the final two iterations of the model, after the key methods and functions were fully implemented, each iteration took approximately one week for two people. Each cycle showed noticeable improvements in model performance (Fig. 8).

This developmental pipeline is well suited to building a machine-learning model within a reasonable time frame when staffing for tasks such as image simulation, data processing, or human labelling is limited. Its efficiency rests on prioritising the removal of false positives from model outputs, thereby avoid-

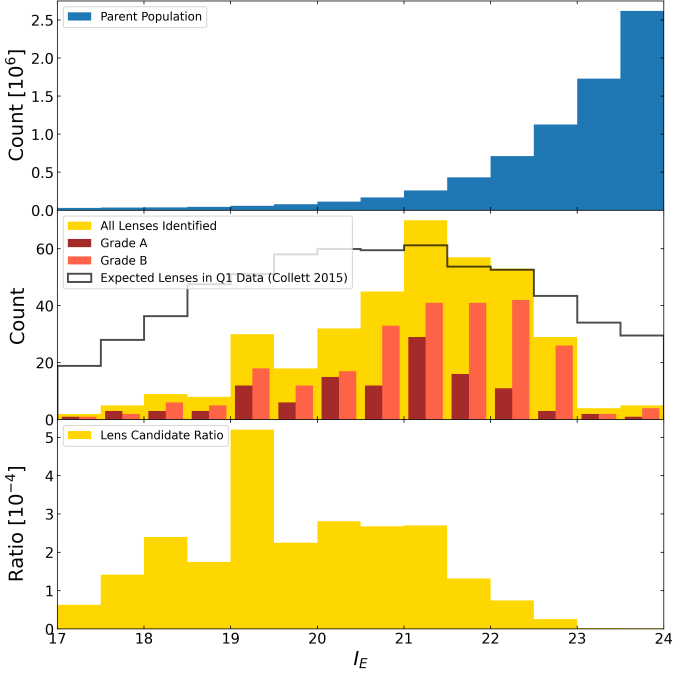


Fig. 14. Histogram of object counts as a function of detection magnitude. *Top:* Full set of inspected candidates with $I_E \leq 24.0$. shown as blue bars. *Middle:* Grade A and grade B lens candidates, respectively, shown as dark and light colour. Background bars are the sum of the two. For reference, we also show the forecast strong lens magnitude distribution from Collett (2015) normalised to *Euclid* Q1 area. *Bottom:* Ratio of lens candidates inside the parent population as a function of magnitude.

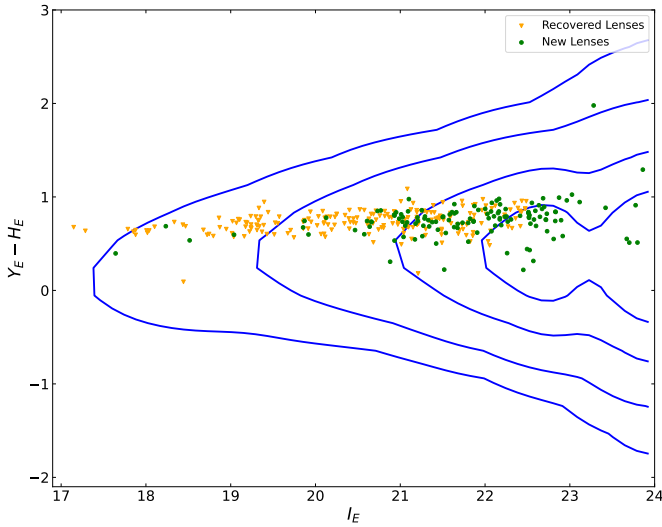


Fig. 15. $Y_E - H_E$ colour vs. I_E magnitude distribution for model-discovered lenses and the parent population. Orange triangles represent the lenses previously found in SLDE A, while green circles represent lenses not found in SLDE A. The blue contours represent the distribution of the parent population, with contour lines at 0.5σ , σ , 1.5σ , and 2σ confidence levels.

ing the need to inspect large numbers of objects by eye. With each fine-tuning step, the false-positive rate among the top candidates falls, so more true positives rise into the review set or the highest-ranked predictions. Previously identified true positives remain in this set throughout. In the early stages, false positives were common and true positives scarce, so the primary

bottleneck was assembling a robust sample of true positives. By design, the pipeline allows a small initial set of true positives to expand rapidly – often within one or two iterations – into a larger dataset that can be augmented and used to train the final model.

One limitation of our current approach is the potential for reduced recall in our final model. Due to our emphasis on precision and efficiency (Fig. 9), as well as the narrower scope of our human labelling (top 4000 here vs. top 10 000 in SLDE A), we likely miss considerable number of genuine lenses that appear lower in the rankings. While this deficiency can be mitigated by inspecting a larger range of candidates, a more fundamental issue arises from the iterative bootstrapping nature of the pipeline. Because the model learns from the lenses it finds, it risks overfitting to the specific characteristics of early detections (e.g. specific colours or brightness levels) at the expense of diversity. For instance, we observed that the model consistently assigned lower scores to lenses with atypical ‘whiter’ colouring. This sensitivity confirms that our pipeline does not sample the sky uniformly, but rather has a distinct selection bias that must be characterised before any physical conclusions can be drawn. Addressing this issue will require further investigation in future work, for example by incorporating a broader variety of lens types, such as edge-on systems and low-redshift lenses, into the training set to improve the robustness and diversity of the model.

5.2. Selection function

The population of the lenses detected in this work is a product of the true lens population and the network’s selection effect as a function of lensing configurations. Consequently, the distribution of candidates cannot be directly interpreted as the underlying astrophysical population without correction. For example, comparing our detections to the forecasts of Collett (2015) reveals a selection bias that peaked at $I_E \sim 21.5$. Understanding the selection function is thus crucial for enabling inference on population level (Sommenfeld 2022). We plan to determine this function by feeding simulated lenses into our network while keeping the network parameters frozen. These test images will be synthesised by compositing simulated lensed arcs with real galaxy cutouts drawn from Q1 data. The lensing configurations will be drawn from a uniform prior spanning the lens parameter space, rather than from an astrophysical prior, allowing us to comprehensively map the network’s decision boundaries. The resulting probability map will provide the correction weights necessary to reconstruct the parent population of strong lenses. Full details of the selection function and the associated lens modelling pipeline will be presented in our follow up work.

5.3. Comparison with traditional approaches

A common strategy to address the small number of confirmed lens candidates is to generate simulated images of lenses to compose the training sets for the model (Metcalf et al. 2019; Schuldt et al. 2023; Euclid Collaboration: Rojas et al. 2025; Acevedo Barroso et al. 2026). This is a particularly useful strategy for projects that rely on early-stage mission data, where there are few or virtually no images of lenses in the style of the target data. However, in simulating astronomical data there are many things to consider to create sufficiently realistic samples, such as morphological diversity, noise, PSF, and imaging artefacts; failing to adequately address these factors can significantly limit model performance (Bottrell et al. 2019; Huppenkothen et al. 2023; Pearce-Casey et al. 2025; Euclid Collaboration: Lines

et al. 2025b). As a result, the trained model will be limited by any differences found in the simulated data and the actual population of inference. In contrast to traditional methods, our model minimised reliance on simulated data and primarily utilised real data that the model found, which allowed us to focus on model development. The drawback of this element of our approach, however, is that model performance may be capped by the small volume of data, which would be insufficient to train a maximally accurate model. Quantitative measurement the selection function of the model will be necessary to fully characterise this insufficiency.

Additionally, our approach utilised a much faster human-rating process. Since efficiency was prioritised, we opted not to use citizen science platforms such as Space Warps and Galaxy Judges until evaluation of the final model was reached. Instead, for each iteration, the top 4000 images produced by the model underwent visual inspection by the authors with priority given to unambiguous lenses that could be readily identified by their distinct characteristics (beyond this range, lens-finding again becomes difficult for manual inspection as highlighted in Fig. 9). While this approach significantly reduced the time required to generate a labelled training set to be typically under one week per iteration, compared to 2–3 weeks using Galaxy Judges, it inherently risks the possibility of missing some less distinct lenses. As a result, the training sets in each iteration may under-represent less distinct lenses, potentially limiting our model’s scope of inference. Despite this, we believe that this was an effective solution for our goals. Our final model was able to make predictions with an exceptionally high purity in the top 1000 (427/1000 classified as Grade A/B lenses³) and recover most of the lenses found using more human resource-intensive approaches such as SLDE A. These results suggest that our efficiency-focused, iterative approach is a viable solution to future lens-finding initiatives, especially when human resources are relatively limited.

5.4. Implication for future work and *Euclid* data releases

The iterative model training approach we have presented here can be easily scaled to meet the demands of large datasets, including the upcoming *Euclid* DR1 data release (Euclid Collaboration: Mellier et al. 2025). Both the model and the developmental approach could be applied effectively to larger data volumes without modification. While the time and number of candidates to sort through would inevitably increase as data increase, this is also expected to result in a proportionally larger set of lenses found. For example, assuming one labeller could quickly sort through 4000 candidates within a week, such that inspection of tens of thousands by a larger groups such as the *Euclid* Strong Lensing Science Working Group is feasible, and over the six year long planned *Euclid* mission, the discovery of up to 170 000 lenses (Collett 2015) with this procedure realistic. Thus, the core advantage of this approach, mainly the efficiency in development under limited human resources, remains preserved even when scaled to the full scope of the *Euclid* Wide Survey.

Finally, it is important to note that although we inspected only the top 4000 lenses, Fig. 16 indicates that there are likely significantly more lenses beyond our range of inspection. The main challenge to address when looking beyond this range is the diminishing proportion of lenses, which returns to the original needle-in-a-haystack problem. Future work can be performed to inspect beyond the current range of the top 4000 rated lenses that

³ Lenses that were classified as Grade A/B by both this work and SLDE A/B/F, if available.

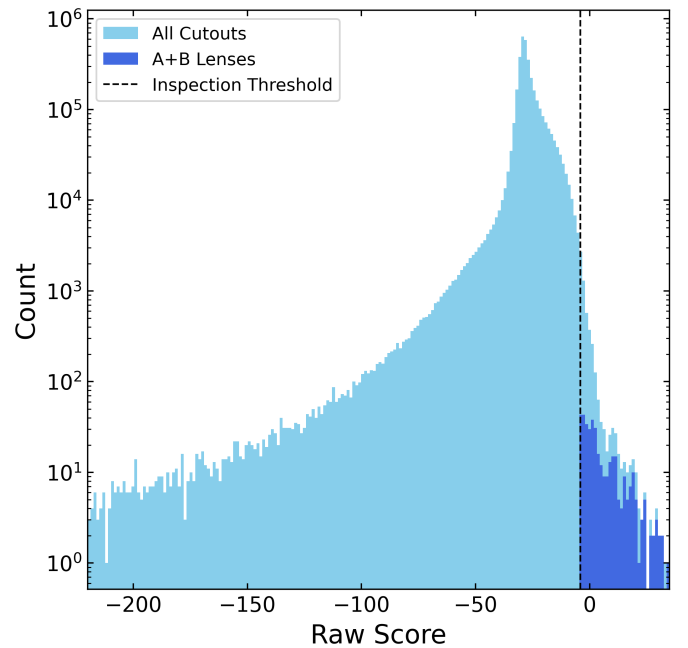


Fig. 16. Raw score distribution of all cut-outs and the identified lenses. The dashed line represents the inspection threshold of the top 4000 raw scores. All cutouts to the left of the dashed line were not inspected.

were the focus of this paper. Furthermore, as increasing samples of lenses are gathered from other *Euclid* lens search outcomes to train on, the model performance should continue to improve through the iterative and human-in-the-loop pipeline. In the meantime, it is crucial to adjust estimates of the proportion of lenses expected from a sky search such as *Euclid* taking into account the model capabilities and identification limitations.

6. Conclusions and future work

In this work, we present an end-to-end, automated pipeline for discovering strong gravitational lenses in the *Euclid* Q1 imaging, spanning a total imaging area of 63.1 deg² in three deep fields. Starting from the VIS instrument catalogues, we applied strict source and artefact vetting, mainly point-source identification to remove stars, and magnitude thresholding, as well as pixel-level, defect, and systematic effects filters. We built 96×96 pixel VIS+NISP colour cutouts using a VIS-anchored luminance scheme that preserves VIS morphology, while retaining NISP colour contrast. Rather than rely on fully simulated training datasets, we bootstrapped from a VIS-only seed classifier to harvest real *Euclid* positives and to curate a balanced, morphology-aware negative set that accounts for contamination from spiral arms, unlensed ellipticals, edge-on galaxies, and diffraction spikes. Among six lightweight CNN models, a modified VGG16 that included GlobalAveragePooling with 256/128 dense layers and a sigmoid with trainable last nine layers delivered the best precision at the top of the ranked list. Iterative fine-tuning grew the training set from 27 seed lenses to a balanced colour set of ~30.7k images (with 15 343 positive and 15 343 negative, respectively, with additional hard-negative mining and targeted diversification for underrepresented morphologies).

Applying the final model to Q1 data, the team arranged visual inspection of the top 4000 ranked candidates through the Galaxy Judges project on Zooniverse with exactly 10 independent votes per object. This procedure was aimed at repeating

the same analysis scheme as the one that was used previously to identify strong lensing systems in Q1 imaging data (Euclid Collaboration: Walmsley et al. 2025). These rankings resulted in 441 grade A/B systems, with 302 lenses overlapping with the previously reported candidates and an additional 130 Grade A/B candidates uncovered in this work. VGG16 model scores and human grades are monotonically correlated, with the lens yield strongly concentrated at the highest scores. Independently, we found that of the 905 high-confidence lens candidates in this work and SLDE A, B, F studies in combination and containing off-centred ones, 740 (81.8%) lie within this model’s top 20 000 predictions. Photometrically, recovered lenses span $I_E \approx 17\text{--}24$ (median 21.2) and exhibit redder $Y_E - H_E$ colour than the parent population, consistent with massive early-type deflectors. Operationally, each training iteration required roughly a week of effort by a small team, involving manual visual inspection and model retraining, because the pipeline steadily suppresses recurrent impostors so that true lenses rise into the top-ranked, human-reviewed subset.

The main contribution of this work is a practical, scalable workflow for rapidly adapting a dedicated CNN to *Euclid* survey images using iterative self-training on real Q1 cutouts, paired with a fast expert-vetting loop. This design produces a high-purity ranking on the target data domain, enables week-timescale iteration during early releases, and yields new lens candidates that complement existing Q1 discovery efforts. Confirming these systems as genuine lensing events remains challenging. The standard approach relies on spectroscopic observations to determine the redshifts of both the source and the lens, combined with detailed lens modelling. However, this method requires substantial observing time at ground-based facilities. Even multi-fibre spectrographs do not fully alleviate this limitation, because the lenses are widely distributed on the sky and the intrinsic lensing rate is low. In the future, *Euclid* grism data will provide a valuable complement, at least for the subset of systems whose lens or source redshifts lie within the grism sensitivity window.

The work presented here will be expanded in the future to include (i) a measurement of the survey-realistic selection function by injecting physically motivated simulated lenses into real *Euclid* backgrounds (native PSF, noise, and artefacts) and reporting calibrated precision–recall at fixed candidate budgets; (ii) a possible extension beyond the top 4000 via uncertainty-aware active learning (e.g. entropy/ensemble disagreement) to raise recall without inflating human load; (iii) a test with multi-scale/attention architectures and adding fast post-hoc lens-model vetters to prune residual impostors; and (vii) adaptation to *Rubin* and *Roman* (Gezari et al. 2022) with minimal additional labels, accompanied by a harmonised public candidates table (thumbnails, scores, and key metadata) for follow-up studies from the ground and space observatories.

The principal limitations of this pipeline are related to the completeness for faint or atypical morphologies and sensitivity to early training seeds and display stretch, as these effects can bias recall at fixed magnitude depth used for human classification. Moreover, our current evaluation is tied to a top- N candidate budget and a fully calibrated selection function capturing the completeness and purity versus lens and source properties remains to be established. Naturally, the selection function is a crucial ingredient in converting this or any other galaxy strong lens sample for inferences related to either cosmological parameters through statistics related to lensing or astrophysical properties of either the foreground lenses or background sources. Despite these caveats, the approach is immediately scalable to larger *Eu-*

clid releases and portable to *Rubin* and *Roman*. Furthermore, the code structure is suitable for public reuse.

Data availability

The full catalogue of classified lens candidates presented in this work (see Table E.1 for sample entries) is available in electronic form through the CDS/VizieR service.

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 BMIMI, 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 Swiss Space Office (SSO) at the State Secretariat for Education, Research, and Innovation (SERI), and the United Kingdom Space Agency. A complete and detailed list is available on the *Euclid* web site (www.euclid-ec.org/consortium/community/). 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>. XX, RC, TL and AC acknowledge the support from NASA ROSES Grant 12-EUCLID11-0004. SS has received funding from the European Union’s Horizon 2022 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 101105167–FASTIDIoUS.

References

- Acevedo Barroso, J. A., Clément, B., Courbin, F., et al. 2026, A&A, 706, A146
 Acevedo Barroso, J. A., O’Riordan, C. M., Clément, B., et al. 2025, A&A, 697, A14
 Aihara, H., Arimoto, N., Armstrong, R., et al. 2018, PASJ, 70, S4
 Bolton, A. S., Burles, S., Koopmans, L. V. E., Treu, T., & Moustakas, L. A. 2006, ApJ, 638, 703
 Bottrell, C., Hani, M. H., Teimoorinia, H., et al. 2019, MNRAS, 490, 5390
 Cănameras, R., Schuldt, S., Shu, Y., et al. 2024, A&A, 692, A72
 Carleo, G., Cirac, I., Cranmer, K., et al. 2019, Rev.Mod.Phys., 91
 Cen, R., Gott, III, J. R., Ostriker, J. P., & Turner, E. L. 1994, ApJ, 423, 1
 Chae, K.-H. 2003, MNRAS, 346, 746
 Cheng, T.-Y., Li, N., Conselice, C. J., et al. 2020, MNRAS, 494, 3750
 Chirivì, G., Yıldırım, A., Suyu, S. H., & Halkola, A. 2020, A&A, 643, A135
 Chollet, F. 2017, Xception: Deep Learning with Depthwise Separable Convolutions, arXiv:1610.02357
 Collett, T. E. 2015, ApJ, 811, 20
 Dalal, N. & Kochanek, C. S. 2002, ApJ, 572, 25
 Dark Energy Survey Collaboration, Abbott, T., Abdalla, F. B., et al. 2016, MNRAS, 460, 1270
 de Jong, J. T. A., Verdoes Kleijn, G. A., Kuijken, K. H., & Valentijn, E. A. 2013, Exp.Astro., 35, 25
 Delchambre, L., Krone-Martins, A., Wertz, O., et al. 2019, A&A, 622, A165
 Euclid Collaboration: Aussel, H., Tereno, I., Schirmer, M., et al. 2025, A&A, in press (Euclid Q1 SI), <https://doi.org/10.1051/0004-6361/202554610>, arXiv:2503.15302
 Euclid Collaboration: Bergamini, P., Meneghetti, M., Acebron, A., et al. 2025, A&A, in press (Euclid Q1 SI), <https://doi.org/10.1051/0004-6361/202554577>, arXiv:2503.15330
 Euclid Collaboration: Corcho-Caballero, P., Ascasibar, Y., Verdoes Kleijn, G., et al. 2025, A&A, in press (Euclid Q1 SI), <https://doi.org/10.1051/0004-6361/202554582>, arXiv:2503.15315
 Euclid Collaboration: Cropper, M., Al-Bahlawan, A., Amiaux, J., et al. 2025, A&A, 697, A2
 Euclid Collaboration: Enia, A., Bolzonella, M., Pozzetti, L., et al. 2024, A&A, 691, A175
 Euclid Collaboration: Holloway, P., Verma, A., Walmsley, M., et al. 2025, A&A, accepted (Euclid Q1 SI), arXiv:2503.15328
 Euclid Collaboration: Jahnke, K., Gillard, W., Schirmer, M., et al. 2025, A&A, 697, A3
 Euclid Collaboration: Li, T., Collett, T. E., Walmsley, M., et al. 2025, A&A, in press (Euclid Q1 SI), <https://doi.org/10.1051/0004-6361/202554543>, arXiv:2503.15327

- Euclid Collaboration: Lines, N. E. P., Collett, T. E., Holloway, P., et al. 2025a, A&A, in press (Euclid Q1 SI), <https://doi.org/10.1051/0004-6361/202558456>, arXiv:2512.05899
- Euclid Collaboration: Lines, N. E. P., Collett, T. E., Walmsley, M., et al. 2025b, A&A, in press (Euclid Q1 SI), <https://doi.org/10.1051/0004-6361/202554542>, arXiv:2503.15326
- 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, in press (Euclid Q1 SI), <https://doi.org/10.1051/0004-6361/202554605>, arXiv:2503.15325
- Euclid Collaboration: Romelli, E., Kümmel, M., Dole, H., et al. 2025, A&A, in press (Euclid Q1 SI), arXiv:2503.15305
- Euclid Collaboration: Scaramella, R., Amiaux, J., Mellier, Y., et al. 2022, A&A, 662, A112
- Euclid Collaboration: Schirmer, M., Jahnke, K., Seidel, G., et al. 2022, A&A, 662, A92
- Euclid Collaboration: Walmsley, M., Holloway, P., Lines, N. E. P., et al. 2025, A&A, in press (Euclid Q1 SI), <https://doi.org/10.1051/0004-6361/202555141>, arXiv:2503.15324
- Euclid Quick Release Q1. 2025, <https://doi.org/10.57780/esa-2853f3b>
- Faure, C., Kneib, J.-P., Covone, G., et al. 2008, ApJS, 176, 19
- Fortson, L., Masters, K., Nichol, R., et al. 2011, arXiv:1104.5513
- Fudamoto, Y., Sun, F., Diego, J. M., et al. 2025, Nat.Ast., 9, 428
- Gannon, C., Nierenberg, A., Benson, A., et al. 2025, Phys. Rev. D, 112, 023532
- Gezari, S., Bentz, M., De, K., et al. 2022, arXiv:2202.12311
- Grogin, N. A. et al. 2011, ApJS, 197, 35
- He, K., Zhang, X., Ren, S., & Sun, J. 2015, arXiv:1512.03385
- Hezaveh, Y., Dalal, N., Holder, G., et al. 2016, JCAP, 11, 048
- Howard, A. G., Zhu, M., Chen, B., et al. 2017, arXiv:1704.04861
- Huang, X., Storer, C., Gu, A., et al. 2021, ApJ, 909, 27
- Huppenkothen, D., Ntampaka, M., Ho, M., et al. 2023, arXiv:2310.12528
- Jackson, N. 2008, MNRAS, 389, 1311
- Jacobs, C., Collett, T., Glazebrook, K., et al. 2019, MNRAS, 484, 5330
- Jacobs, C., Glazebrook, K., Collett, T., More, A., & McCarthy, C. 2017, MNRAS, 471, 167
- Jung, A. B. 2018, imgaug: Image Augmentation Library, <https://github.com/aleju/imgaug>, [Online; accessed 2025-07-05]
- Keeton, C. R. & Moustakas, L. A. 2009, ApJ, 699, 1720
- Kingma, D. P. & Ba, J. 2014, arXiv:1412.6980
- Koekemoer, A. M. et al. 2011, ApJS, 197, 36
- Koopmans, L. V. E. 2005, MNRAS, 363, 1136
- Koopmans, L. V. E., Barnabe, M., Bolton, A., et al. 2009, in astro2010: The Astronomy and Astrophysics Decadal Survey, Vol. 2010, 159
- LeCun, Y. & Bengio, Y. 1995, in The Handbook of Brain Theory and Neural Networks, ed. M. A. Arbib (Cambridge, MA: MIT Press), 255–258
- Li, R., Napolitano, N. R., Tortora, C., et al. 2020, ApJ, 899, 30
- Li, T. et al. 2023, BAAS, 55, 12
- Metcalfe, R. B., Meneghetti, M., Avestruz, C., et al. 2019, A&A, 625, A119
- Myers, S. T., Jackson, N. J., Browne, I. W. A., et al. 2003, MNRAS, 341, 1
- Nagam, B. C., Acevedo Barroso, J. A., Wilde, J., et al. 2025, A&A, 702, A130
- O'Riordan, C. M., Oldham, L. J., Nersesian, A., et al. 2025, A&A, 694, A145
- Pearce-Casey, R., Nagam, B. C., Wilde, J., et al. 2025, A&A, 696, A214
- Petrillo, C. E., Tortora, C., Chatterjee, S., et al. 2017, A&A, 605, A52
- Pourrahmani, M., Nayyeri, H., & Cooray, A. 2018, ApJ, 856, 68
- Rodney, S. A., Brammer, G. B., Pierel, J. D. R., et al. 2021, Nat.Ast., 5, 1118
- Rojas, K., Collett, T., Ballard, D., Magee, M., et al. 2023, MNRAS, 526, 1065
- Schuldt, S., Cañameras, R., Shu, Y., et al. 2023, A&A, 671, A147
- Schuldt, S., Suyu, S. H., Cañameras, R., et al. 2022, A&A, 660, A14
- Sheng, X., Mao, S., Xu, D., et al. 2022, MNRAS, 517, 1171
- Simonyan, K. & Zisserman, A. 2015, arXiv:1409.1556
- Sonnenfeld, A. 2021, A&A, 656, A153
- Sonnenfeld, A. 2022, A&A, 659, A132
- Sonnenfeld, A. & Cautun, M. 2021, A&A, 651, A18
- Sonnenfeld, A., Gavazzi, R., Suyu, S. H., Treu, T., & Marshall, P. J. 2013, ApJ, 777, 97
- Sonnenfeld, A., Treu, T., Marshall, P. J., et al. 2015, ApJ, 800, 94
- Storfer, C., Huang, X., Gu, A., et al. 2024, ApJS, 274, 16
- Szegedy, C., Liu, W., Jia, Y., et al. 2015, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1–9
- Tan, M. & Le, Q. V. 2020, arXiv:1905.11946
- Thuruthipilly, H., Merten, J., Giocoli, C., & Meneghetti, M. 2022, MNRAS, 512, 1598
- Treu, T. 2010, ARA&A, 48, 87
- Treu, T., Auger, M. W., Koopmans, L. V. E., et al. 2010, ApJ, 709, 1195
- Vegetti, S., Birrer, S., Despali, G., et al. 2024, Space Sci. Rev., 220, 58
- Vegetti, S., Koopmans, L. V. E., Auger, M. W., Treu, T., & Bolton, A. S. 2014, MNRAS, 442, 2017
- Vegetti, S., Koopmans, L. V. E., Bolton, A., Treu, T., & Gavazzi, R. 2010, MNRAS, 408, 1969
- Wong, K. C., Suyu, S. H., Chen, G. C. F., et al. 2020, MNRAS, 498, 1420
- Zhu, X.-P., Dai, J.-M., Bian, C.-J., et al. 2019, ApSS, 364
- ¹ Department of Physics & Astronomy, University of California Irvine, Irvine CA 92697, USA
 - ² University of Southern California, 3551 Trousdale Parkway, Los Angeles, CA 90089, USA
 - ³ Dipartimento di Fisica "Aldo Pontremoli", Università degli Studi di Milano, Via Celoria 16, 20133 Milano, Italy
 - ⁴ INAF-IASF Milano, Via Alfonso Corti 12, 20133 Milano, Italy
 - ⁵ Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA, 91109, USA
 - ⁶ Department of Physics and Astronomy, University of British Columbia, Vancouver, BC V6T 1Z1, Canada
 - ⁷ INAF-Osservatorio di Astrofisica e Scienza dello Spazio di Bologna, Via Piero Gobetti 93/3, 40129 Bologna, Italy
 - ⁸ INFN-Sezione di Bologna, Viale Berti Pichat 6/2, 40127 Bologna, Italy
 - ⁹ Dipartimento di Fisica e Astronomia "Augusto Righi" - Alma Mater Studiorum Università di Bologna, via Piero Gobetti 93/2, 40129 Bologna, Italy
 - ¹⁰ INAF-Osservatorio Astronomico di Capodimonte, Via Moiariello 16, 80131 Napoli, Italy
 - ¹¹ Institute of Cosmology and Gravitation, University of Portsmouth, Portsmouth PO1 3FX, UK
 - ¹² School of Physics, HH Wills Physics Laboratory, University of Bristol, Tyndall Avenue, Bristol, BS8 1TL, UK
 - ¹³ Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str. 1, 85748 Garching, Germany
 - ¹⁴ Technical University of Munich, TUM School of Natural Sciences, Physics Department, James-Frank-Str. 1, 85748 Garching, Germany
 - ¹⁵ Departamento Física Aplicada, Universidad Politécnica de Cartagena, Campus Muralla del Mar, 30202 Cartagena, Murcia, Spain
 - ¹⁶ Aix-Marseille Université, CNRS, CNES, LAM, Marseille, France
 - ¹⁷ Institut d'Astrophysique de Paris, UMR 7095, CNRS, and Sorbonne Université, 98 bis boulevard Arago, 75014 Paris, France
 - ¹⁸ Center for Astrophysics | Harvard & Smithsonian, 60 Garden St., Cambridge, MA 02138, USA
 - ¹⁹ ESAC/ESA, Camino Bajo del Castillo, s/n., Urb. Villafranca del Castillo, 28692 Villanueva de la Cañada, Madrid, Spain
 - ²⁰ Université de Genève, Département de Physique Théorique and Centre for Astroparticle Physics, 24 quai Ernest-Ansermet, CH-1211 Genève 4, Switzerland
 - ²¹ Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, USA
 - ²² Max Planck Institute for Extraterrestrial Physics, Giessenbachstr. 1, 85748 Garching, Germany
 - ²³ Universitäts-Sternwarte München, Fakultät für Physik, Ludwig-Maximilians-Universität München, Scheinerstr. 1, 81679 München, Germany
 - ²⁴ STAR Institute, University of Liège, Quartier Agora, Allée du six Août 19c, 4000 Liège, Belgium
 - ²⁵ Institut de Ciències del Cosmos (ICCUB), Universitat de Barcelona (IEEC-UB), Martí i Franquès 1, 08028 Barcelona, Spain
 - ²⁶ David A. Dunlap Department of Astronomy & Astrophysics, University of Toronto, 50 St George Street, Toronto, Ontario M5S 3H4, Canada
 - ²⁷ Jodrell Bank Centre for Astrophysics, Department of Physics and Astronomy, University of Manchester, Oxford Road, Manchester M13 9PL, UK
 - ²⁸ Institutió Catalana de Recerca i Estudis Avançats (ICREA), Pas-seig de Lluís Companys 23, 08010 Barcelona, Spain
 - ²⁹ Institut de Ciències de l'Espai (IEEC-CSIC), Campus UAB, Carrer de Can Magrans, s/n Cerdanyola del Vallés, 08193 Barcelona, Spain
 - ³⁰ School of Mathematics and Physics, University of Surrey, Guildford, Surrey, GU2 7XH, UK

- 31 INAF-Osservatorio Astronomico di Brera, Via Brera 28, 20122 Milano, Italy
- 32 IFPU, Institute for Fundamental Physics of the Universe, via Beirut 2, 34151 Trieste, Italy
- 33 INAF-Osservatorio Astronomico di Trieste, Via G. B. Tiepolo 11, 34143 Trieste, Italy
- 34 INFN, Sezione di Trieste, Via Valerio 2, 34127 Trieste TS, Italy
- 35 SISSA, International School for Advanced Studies, Via Bonomea 265, 34136 Trieste TS, Italy
- 36 Dipartimento di Fisica e Astronomia, Università di Bologna, Via Gobetti 93/2, 40129 Bologna, Italy
- 37 INAF-Osservatorio Astronomico di Padova, Via dell'Osservatorio 5, 35122 Padova, Italy
- 38 Dipartimento di Fisica, Università di Genova, Via Dodecaneso 33, 16146, Genova, Italy
- 39 INFN-Sezione di Genova, Via Dodecaneso 33, 16146, Genova, Italy
- 40 Department of Physics "E. Pancini", University Federico II, Via Cinthia 6, 80126, Napoli, Italy
- 41 Dipartimento di Fisica, Università degli Studi di Torino, Via P. Giuria 1, 10125 Torino, Italy
- 42 INFN-Sezione di Torino, Via P. Giuria 1, 10125 Torino, Italy
- 43 INAF-Osservatorio Astrofisico di Torino, Via Osservatorio 20, 10025 Pino Torinese (TO), Italy
- 44 INAF-Osservatorio Astronomico di Roma, Via Frascati 33, 00078 Monteporzio Catone, Italy
- 45 INFN-Sezione di Roma, Piazzale Aldo Moro, 2 - c/o Dipartimento di Fisica, Edificio G. Marconi, 00185 Roma, Italy
- 46 Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Avenida Complutense 40, 28040 Madrid, Spain
- 47 Port d'Informació Científica, Campus UAB, C. Albareda s/n, 08193 Bellaterra (Barcelona), Spain
- 48 Institute for Theoretical Particle Physics and Cosmology (TTK), RWTH Aachen University, 52056 Aachen, Germany
- 49 Deutsches Zentrum für Luft- und Raumfahrt e. V. (DLR), Linder Höhe, 51147 Köln, Germany
- 50 INFN section of Naples, Via Cinthia 6, 80126, Napoli, Italy
- 51 Dipartimento di Fisica e Astronomia "Augusto Righi" - Alma Mater Studiorum Università di Bologna, Viale Berti Pichat 6/2, 40127 Bologna, Italy
- 52 Instituto de Astrofísica de Canarias, E-38205 La Laguna, Tenerife, Spain
- 53 Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK
- 54 European Space Agency/ESRIN, Largo Galileo Galilei 1, 00044 Frascati, Roma, Italy
- 55 Université Claude Bernard Lyon 1, CNRS/IN2P3, IP2I Lyon, UMR 5822, Villeurbanne, F-69100, France
- 56 UCB Lyon 1, CNRS/IN2P3, IUF, IP2I Lyon, 4 rue Enrico Fermi, 69622 Villeurbanne, France
- 57 Mullard Space Science Laboratory, University College London, Holmbury St Mary, Dorking, Surrey RH5 6NT, UK
- 58 Departamento de Física, Faculdade de Ciências, Universidade de Lisboa, Edifício C8, Campo Grande, PT1749-016 Lisboa, Portugal
- 59 Instituto de Astrofísica e Ciências do Espaço, Faculdade de Ciências, Universidade de Lisboa, Campo Grande, 1749-016 Lisboa, Portugal
- 60 Department of Astronomy, University of Geneva, ch. d'Ecogia 16, 1290 Versoix, Switzerland
- 61 Université Paris-Saclay, CNRS, Institut d'astrophysique spatiale, 91405, Orsay, France
- 62 INFN-Padova, Via Marzolo 8, 35131 Padova, Italy
- 63 Aix-Marseille Université, CNRS/IN2P3, CPPM, Marseille, France
- 64 INAF-Istituto di Astrofisica e Planetologia Spaziali, via del Fosso del Cavaliere, 100, 00100 Roma, Italy
- 65 Université Paris-Saclay, Université Paris Cité, CEA, CNRS, AIM, 91191, Gif-sur-Yvette, France
- 66 INFN-Bologna, Via Irnerio 46, 40126 Bologna, Italy
- 67 Institut d'Estudis Espacials de Catalunya (IEEC), Edifici RDIT, Campus UPC, 08860 Castelldefels, Barcelona, Spain
- 68 Institute of Space Sciences (ICE, CSIC), Campus UAB, Carrer de Can Magrans, s/n, 08193 Barcelona, Spain
- 69 University Observatory, LMU Faculty of Physics, Scheinerstr. 1, 81679 Munich, Germany
- 70 FRACTAL S.L.N.E., calle Tulipán 2, Portal 13 1A, 28231, Las Rozas de Madrid, Spain
- 71 Institute of Theoretical Astrophysics, University of Oslo, P.O. Box 1029 Blindern, 0315 Oslo, Norway
- 72 Felix Hormuth Engineering, Goethestr. 17, 69181 Leimen, Germany
- 73 Technical University of Denmark, Elektrovej 327, 2800 Kgs. Lyngby, Denmark
- 74 Cosmic Dawn Center (DAWN), Denmark
- 75 Max-Planck-Institut für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany
- 76 NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA
- 77 Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK
- 78 Department of Physics, P.O. Box 64, University of Helsinki, 00014 Helsinki, Finland
- 79 Helsinki Institute of Physics, Gustaf Hållströmin katu 2, University of Helsinki, 00014 Helsinki, Finland
- 80 Laboratoire d'étude de l'Univers et des phénomènes eXtremes, Observatoire de Paris, Université PSL, Sorbonne Université, CNRS, 92190 Meudon, France
- 81 SKAO, Jodrell Bank, Lower Withington, Macclesfield SK11 9FT, UK
- 82 Centre de Calcul de l'IN2P3/CNRS, 21 avenue Pierre de Coubertin 69627 Villeurbanne Cedex, France
- 83 University of Applied Sciences and Arts of Northwestern Switzerland, School of Computer Science, 5210 Windisch, Switzerland
- 84 Universität Bonn, Argelander-Institut für Astronomie, Auf dem Hügel 71, 53121 Bonn, Germany
- 85 Department of Physics, Institute for Computational Cosmology, Durham University, South Road, Durham, DH1 3LE, UK
- 86 Université Paris Cité, CNRS, Astroparticule et Cosmologie, 75013 Paris, France
- 87 CNRS-UCB International Research Laboratory, Centre Pierre Binétruy, IRL2007, CPB-IN2P3, Berkeley, USA
- 88 University of Applied Sciences and Arts of Northwestern Switzerland, School of Engineering, 5210 Windisch, Switzerland
- 89 Institute of Physics, Laboratory of Astrophysics, Ecole Polytechnique Fédérale de Lausanne (EPFL), Observatoire de Sauverny, 1290 Versoix, Switzerland
- 90 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
- 91 Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Campus UAB, 08193 Bellaterra (Barcelona), Spain
- 92 European Space Agency/ESTEC, Keplerlaan 1, 2201 AZ Noordwijk, The Netherlands
- 93 School of Mathematics, Statistics and Physics, Newcastle University, Herschel Building, Newcastle-upon-Tyne, NE1 7RU, UK
- 94 DARK, Niels Bohr Institute, University of Copenhagen, Jagtvej 155, 2200 Copenhagen, Denmark
- 95 Waterloo Centre for Astrophysics, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada
- 96 Department of Physics and Astronomy, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada
- 97 Perimeter Institute for Theoretical Physics, Waterloo, Ontario N2L 2Y5, Canada
- 98 Space Science Data Center, Italian Space Agency, via del Politecnico snc, 00133 Roma, Italy
- 99 Centre National d'Etudes Spatiales – Centre spatial de Toulouse, 18 avenue Edouard Belin, 31401 Toulouse Cedex 9, France
- 100 Institute of Space Science, Str. Atomistilor, nr. 409 Măgurele, Ilfov, 077125, Romania

- ¹⁰¹ Dipartimento di Fisica e Astronomia "G. Galilei", Università di Padova, Via Marzolo 8, 35131 Padova, Italy
- ¹⁰² Instituto de Física Teórica UAM-CSIC, Campus de Cantoblanco, 28049 Madrid, Spain
- ¹⁰³ Institut de Recherche en Astrophysique et Planétologie (IRAP), Université de Toulouse, CNRS, UPS, CNES, 14 Av. Edouard Belin, 31400 Toulouse, France
- ¹⁰⁴ Université St Joseph; Faculty of Sciences, Beirut, Lebanon
- ¹⁰⁵ Departamento de Física, FCFM, Universidad de Chile, Blanco Encalada 2008, Santiago, Chile
- ¹⁰⁶ Universität Innsbruck, Institut für Astro- und Teilchenphysik, Technikerstr. 25/8, 6020 Innsbruck, Austria
- ¹⁰⁷ Department of Physics and Helsinki Institute of Physics, Gustaf Hållströmin katu 2, University of Helsinki, 00014 Helsinki, Finland
- ¹⁰⁸ Instituto de Astrofísica e Ciências do Espaço, Faculdade de Ciências, Universidade de Lisboa, Tapada da Ajuda, 1349-018 Lisboa, Portugal
- ¹⁰⁹ Cosmic Dawn Center (DAWN)
- ¹¹⁰ Niels Bohr Institute, University of Copenhagen, Jagtvej 128, 2200 Copenhagen, Denmark
- ¹¹¹ Universidad Politécnica de Cartagena, Departamento de Electrónica y Tecnología de Computadoras, Plaza del Hospital 1, 30202 Cartagena, Spain
- ¹¹² Kapteyn Astronomical Institute, University of Groningen, PO Box 800, 9700 AV Groningen, The Netherlands
- ¹¹³ Caltech/IPAC, 1200 E. California Blvd., Pasadena, CA 91125, USA
- ¹¹⁴ Dipartimento di Fisica e Scienze della Terra, Università degli Studi di Ferrara, Via Giuseppe Saragat 1, 44122 Ferrara, Italy
- ¹¹⁵ Istituto Nazionale di Fisica Nucleare, Sezione di Ferrara, Via Giuseppe Saragat 1, 44122 Ferrara, Italy
- ¹¹⁶ INAF, Istituto di Radioastronomia, Via Piero Gobetti 101, 40129 Bologna, Italy
- ¹¹⁷ Astronomical Observatory of the Autonomous Region of the Aosta Valley (OAVdA), Loc. Lignan 39, I-11020, Nus (Aosta Valley), Italy
- ¹¹⁸ Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, Laboratoire Lagrange, Bd de l'Observatoire, CS 34229, 06304 Nice cedex 4, France
- ¹¹⁹ ICSC - Centro Nazionale di Ricerca in High Performance Computing, Big Data e Quantum Computing, Via Magnanelli 2, Bologna, Italy
- ¹²⁰ Department of Physics, Oxford University, Keble Road, Oxford OX1 3RH, UK
- ¹²¹ Univ. Grenoble Alpes, CNRS, Grenoble INP, LPSC-IN2P3, 53, Avenue des Martyrs, 38000, Grenoble, France
- ¹²² Dipartimento di Fisica, Sapienza Università di Roma, Piazzale Aldo Moro 2, 00185 Roma, Italy
- ¹²³ Aurora Technology for European Space Agency (ESA), Camino bajo del Castillo, s/n, Urbanización Villafranca del Castillo, Villanueva de la Cañada, 28692 Madrid, Spain
- ¹²⁴ Dipartimento di Fisica - Sezione di Astronomia, Università di Trieste, Via Tiepolo 11, 34131 Trieste, Italy
- ¹²⁵ Department of Mathematics and Physics E. De Giorgi, University of Salento, Via per Arnesano, CP-I93, 73100, Lecce, Italy
- ¹²⁶ INFN, Sezione di Lecce, Via per Arnesano, CP-I93, 73100, Lecce, Italy
- ¹²⁷ INAF-Sezione di Lecce, c/o Dipartimento Matematica e Fisica, Via per Arnesano, 73100, Lecce, Italy
- ¹²⁸ Institut d'Astrophysique de Paris, 98bis Boulevard Arago, 75014, Paris, France
- ¹²⁹ ICL, Junia, Université Catholique de Lille, LITL, 59000 Lille, France
- ¹³⁰ CERCA/ISO, Department of Physics, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, OH 44106, USA
- ¹³¹ Laboratoire Univers et Théorie, Observatoire de Paris, Université PSL, Université Paris Cité, CNRS, 92190 Meudon, France
- ¹³² INFN-Sezione di Milano, Via Celoria 16, 20133 Milano, Italy
- ¹³³ Departamento de Física Fundamental. Universidad de Salamanca. Plaza de la Merced s/n. 37008 Salamanca, Spain
- ¹³⁴ Université de Strasbourg, CNRS, Observatoire astronomique de Strasbourg, UMR 7550, 67000 Strasbourg, France
- ¹³⁵ Center for Data-Driven Discovery, Kavli IPMU (WPI), UTIAS, The University of Tokyo, Kashiwa, Chiba 277-8583, Japan
- ¹³⁶ California Institute of Technology, 1200 E California Blvd, Pasadena, CA 91125, USA
- ¹³⁷ Instituto de Física de Cantabria, Edificio Juan Jordá, Avenida de los Castros, 39005 Santander, Spain
- ¹³⁸ Department of Computer Science, Aalto University, PO Box 15400, Espoo, FI-00 076, Finland
- ¹³⁹ Universidad de La Laguna, Dpto. Astrofísica, E-38206 La Laguna, Tenerife, Spain
- ¹⁴⁰ Ruhr University Bochum, Faculty of Physics and Astronomy, Astronomical Institute (AIRUB), German Centre for Cosmological Lensing (GCCL), 44780 Bochum, Germany
- ¹⁴¹ Department of Physics and Astronomy, Vesilinnantie 5, University of Turku, 20014 Turku, Finland
- ¹⁴² Finnish Centre for Astronomy with ESO (FINCA), Quantum, Vesilinnantie 5, University of Turku, 20014 Turku, Finland
- ¹⁴³ Serco for European Space Agency (ESA), Camino bajo del Castillo, s/n, Urbanización Villafranca del Castillo, Villanueva de la Cañada, 28692 Madrid, Spain
- ¹⁴⁴ ARC Centre of Excellence for Dark Matter Particle Physics, Melbourne, Australia
- ¹⁴⁵ Centre for Astrophysics & Supercomputing, Swinburne University of Technology, Hawthorn, Victoria 3122, Australia
- ¹⁴⁶ Department of Physics and Astronomy, University of the Western Cape, Bellville, Cape Town, 7535, South Africa
- ¹⁴⁷ Departement of Theoretical Physics, University of Geneva, Switzerland
- ¹⁴⁸ Department of Physics, Centre for Extragalactic Astronomy, Durham University, South Road, Durham, DH1 3LE, UK
- ¹⁴⁹ IRFU, CEA, Université Paris-Saclay 91191 Gif-sur-Yvette Cedex, France
- ¹⁵⁰ Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, HI 96822, USA
- ¹⁵¹ INAF-Osservatorio Astrofisico di Arcetri, Largo E. Fermi 5, 50125, Firenze, Italy
- ¹⁵² Centro de Astrofísica da Universidade do Porto, Rua das Estrelas, 4150-762 Porto, Portugal
- ¹⁵³ Instituto de Astrofísica e Ciências do Espaço, Universidade do Porto, CAUP, Rua das Estrelas, PT4150-762 Porto, Portugal
- ¹⁵⁴ HE Space for European Space Agency (ESA), Camino bajo del Castillo, s/n, Urbanización Villafranca del Castillo, Villanueva de la Cañada, 28692 Madrid, Spain
- ¹⁵⁵ Department of Astrophysics, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland
- ¹⁵⁶ INAF - Osservatorio Astronomico d'Abruzzo, Via Maggini, 64100, Teramo, Italy
- ¹⁵⁷ Theoretical astrophysics, Department of Physics and Astronomy, Uppsala University, Box 516, 751 37 Uppsala, Sweden
- ¹⁵⁸ Minnesota Institute for Astrophysics, University of Minnesota, 116 Church St SE, Minneapolis, MN 55455, USA
- ¹⁵⁹ Mathematical Institute, University of Leiden, Einsteinweg 55, 2333 CA Leiden, The Netherlands
- ¹⁶⁰ Leiden Observatory, Leiden University, Einsteinweg 55, 2333 CC Leiden, The Netherlands
- ¹⁶¹ Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK
- ¹⁶² Center for Astrophysics and Cosmology, University of Nova Gorica, Nova Gorica, Slovenia
- ¹⁶³ Institute for Particle Physics and Astrophysics, Dept. of Physics, ETH Zurich, Wolfgang-Pauli-Strasse 27, 8093 Zurich, Switzerland
- ¹⁶⁴ Department of Astrophysical Sciences, Peyton Hall, Princeton University, Princeton, NJ 08544, USA
- ¹⁶⁵ Space physics and astronomy research unit, University of Oulu, Pentti Kaiteran katu 1, FI-90014 Oulu, Finland
- ¹⁶⁶ Department of Physics and Astronomy, Lehman College of the CUNY, Bronx, NY 10468, USA

- ¹⁶⁷ American Museum of Natural History, Department of Astrophysics, New York, NY 10024, USA
- ¹⁶⁸ International Centre for Theoretical Physics (ICTP), Strada Costiera 11, 34151 Trieste, Italy
- ¹⁶⁹ Center for Computational Astrophysics, Flatiron Institute, 162 5th Avenue, 10010, New York, NY, USA

Appendix A: Deterministic data augmentation scheme

We applied a fixed set of 67 deterministic image augmentations implemented using the `imgaug` library. Each transformation was applied independently to the training images. The augmentations consist of geometric operations (horizontal and vertical flips, scaling), photometric adjustments (brightness, contrast, saturation), and combinations thereof.

Appendix A.1: Geometric transformations

- G1: Horizontal flip (Fliplr)
- G2: Vertical flip (Flipud)
- G3: Horizontal + vertical flip
- G4: Affine scaling (scale = 1.1)
- G5: Affine scaling (scale = 1.2)

Appendix A.2: Contrast adjustments

- C1: LinearContrast 1.2
- C2: LinearContrast 1.1
- C3: LinearContrast 0.9
- C4: LinearContrast 0.8

Appendix A.3: Brightness adjustments

- B1: MultiplyBrightness 1.1
- B2: MultiplyBrightness 0.9

Appendix A.4: Saturation adjustments

- S1: MultiplySaturation 1.1
- S2: MultiplySaturation 0.9

Appendix A.5: Combined transformations

The remaining augmentations consist of deterministic combinations of flips with brightness, contrast, and saturation modifications.

- M1: Fliplr + LinearContrast 1.2
- M2: Flipud + LinearContrast 1.2
- M3: Fliplr + LinearContrast 1.1
- M4: Flipud + LinearContrast 1.1
- M5: Fliplr + Flipud + LinearContrast 1.2
- M6: Fliplr + Flipud + LinearContrast 1.1
- M7: Fliplr + MultiplyBrightness 1.1
- M8: Flipud + MultiplyBrightness 0.9
- M9: Fliplr + Flipud + MultiplyBrightness 0.9
- M10: Fliplr + Flipud + MultiplyBrightness 1.1
- M11: Fliplr + MultiplyBrightness 1.1 + LinearContrast 1.1
- M12: Flipud + MultiplyBrightness 0.9 + LinearContrast 1.1
- M13: Fliplr + Flipud + MultiplyBrightness 0.9 + LinearContrast 1.1
- M14: Fliplr + Flipud + MultiplyBrightness 1.1 + LinearContrast 1.1
- M15: Fliplr + MultiplyBrightness 1.1 + LinearContrast 1.2
- M16: Flipud + MultiplyBrightness 0.9 + LinearContrast 1.2
- M17: Fliplr + Flipud + MultiplyBrightness 0.9 + LinearContrast 1.2
- M18: Fliplr + Flipud + MultiplyBrightness 1.1 + LinearContrast 1.2
- M19: Fliplr + MultiplyBrightness 1.1 + LinearContrast 0.8
- M20: Flipud + MultiplyBrightness 0.9 + LinearContrast 0.8
- M21: Fliplr + Flipud + MultiplyBrightness 0.9 + LinearContrast 0.8
- M22: Fliplr + Flipud + MultiplyBrightness 1.1 + LinearContrast 0.8
- M23: Fliplr + MultiplyBrightness 1.1 + LinearContrast 0.9
- M24: Flipud + MultiplyBrightness 0.9 + LinearContrast 0.9
- M25: Fliplr + Flipud + MultiplyBrightness 0.9 + LinearContrast 0.9
- M26: Fliplr + Flipud + MultiplyBrightness 1.1 + LinearContrast 0.9
- M27: Fliplr + MultiplySaturation 1.1
- M28: Flipud + MultiplySaturation 0.9
- M29: Fliplr + Flipud + MultiplySaturation 0.9
- M30: Fliplr + Flipud + MultiplySaturation 1.1

- M31: Fliplr + MultiplySaturation 1.1 + MultiplyBrightness 1.1
- M32: Flipud + MultiplySaturation 0.9 + MultiplyBrightness 1.1
- M33: Fliplr + Flipud + MultiplySaturation 0.9 + MultiplyBrightness 1.1
- M34: Fliplr + Flipud + MultiplySaturation 1.1 + MultiplyBrightness 1.1
- M35: Fliplr + MultiplySaturation 1.1 + MultiplyBrightness 0.9
- M36: Flipud + MultiplySaturation 0.9 + MultiplyBrightness 0.9
- M37: Fliplr + Flipud + MultiplySaturation 0.9 + MultiplyBrightness 0.9
- M38: Fliplr + Flipud + MultiplyBrightness 1.1 + LinearContrast 0.9
- M39: Fliplr + MultiplySaturation 1.1 + LinearContrast 1.1
- M40: Flipud + MultiplySaturation 0.9 + LinearContrast 1.1
- M41: Fliplr + Flipud + MultiplySaturation 0.9 + LinearContrast 1.1
- M42: Fliplr + Flipud + MultiplySaturation 1.1 + LinearContrast 1.1
- M43: Fliplr + MultiplySaturation 1.1 + LinearContrast 1.2
- M44: Flipud + MultiplySaturation 0.9 + LinearContrast 1.2
- M45: Fliplr + Flipud + MultiplySaturation 0.9 + LinearContrast 1.2
- M46: Fliplr + Flipud + MultiplySaturation 1.1 + LinearContrast 1.2
- M47: Fliplr + MultiplySaturation 1.1 + LinearContrast 0.8
- M48: Flipud + MultiplySaturation 0.9 + LinearContrast 0.8
- M49: Fliplr + Flipud + MultiplySaturation 0.9 + LinearContrast 0.8
- M50: Fliplr + Flipud + MultiplySaturation 1.1 + LinearContrast 0.8
- M51: Fliplr + MultiplySaturation 1.1 + LinearContrast 0.9
- M52: Flipud + MultiplySaturation 0.9 + LinearContrast 0.9
- M53: Fliplr + Flipud + MultiplySaturation 0.9 + LinearContrast 0.9
- M54: Fliplr + Flipud + MultiplySaturation 1.1 + LinearContrast 0.9

Appendix B: Images with grading discrepancy

In Fig. B.1 we highlight the strong lens candidates that differed in rating from [Euclid Collaboration: Walmsley et al. \(2025\)](#).

Appendix C: 100 hand-drawn positive training set

In Fig. C.1 we highlight the 100 experimental hand-drawn lens simulations used to enhance the diversity of the positive training dataset in the second round of the iterative training process.

Appendix D: Plots for all classified A/B lenses within the top 4000 predictions

In Figs. D.1, D.2, D.3, D.4, and D.5 we show all 367 strong lens candidates that were classified as grade A/B lenses. Newly identified lenses are marked with green outlines. The other lenses, which were previously identified, are selected from having union grades between AA, AB, BA, and BB. For each union grade, it consists of grade received in this work and grade given in previous research. Note that in text we refer to a total A/B lens candidate sample of 441. The 441 includes 74 additional candidates that received AC, BC, AX, and BX union grades. These candidates are shown in Fig. B.1.

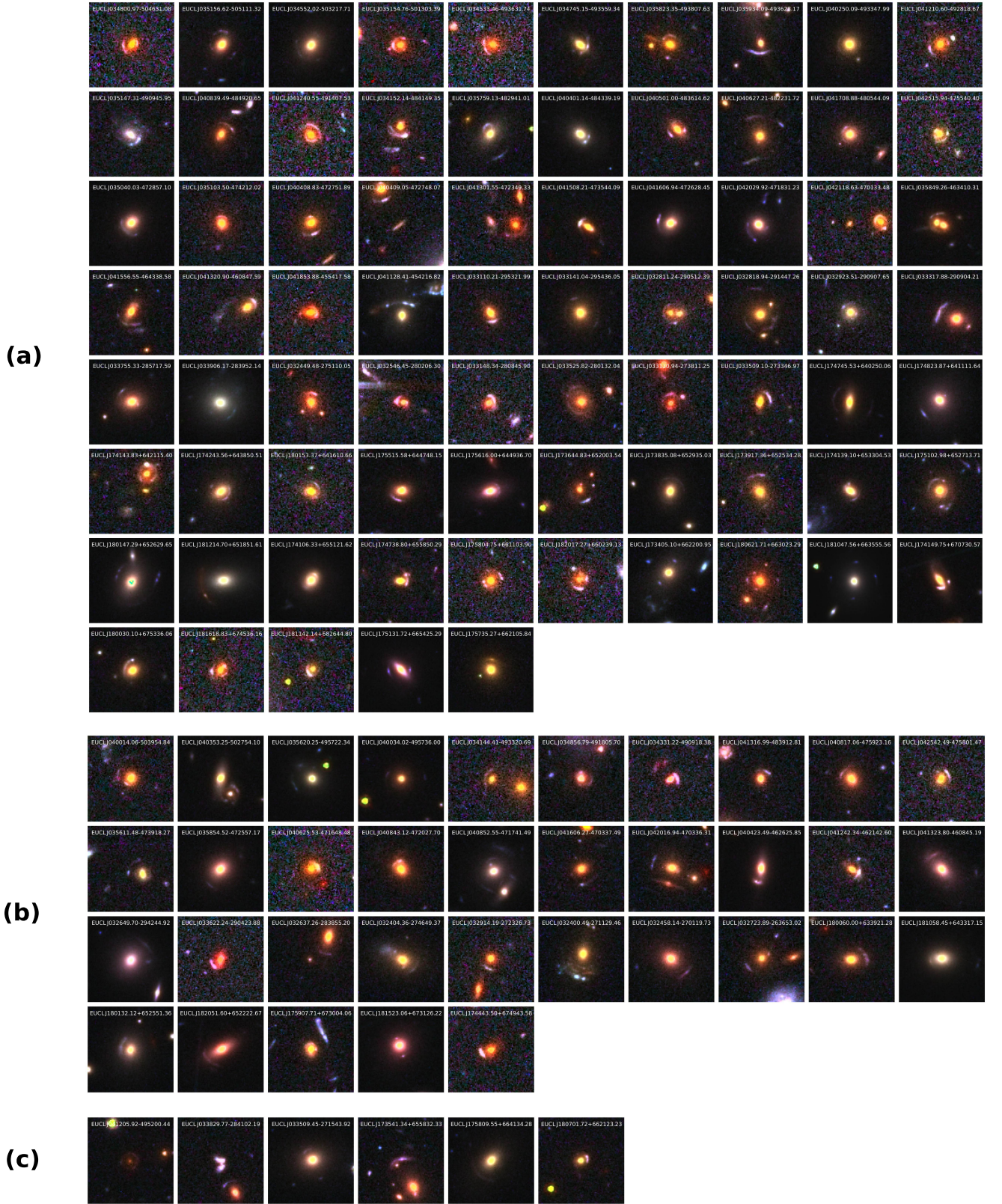


Fig. B.1. Centred lens candidates that received lower grades in this work compared with SLDE A. Panel (a): Objects that received Grade A in SLDE A but were graded as Grade B in this work. Panel (b): Objects that received Grade B in SLDE A but were graded as Grade C in this work. Panel (c): Objects that received Grade A in SLDE A but were graded as Grade C in this work.

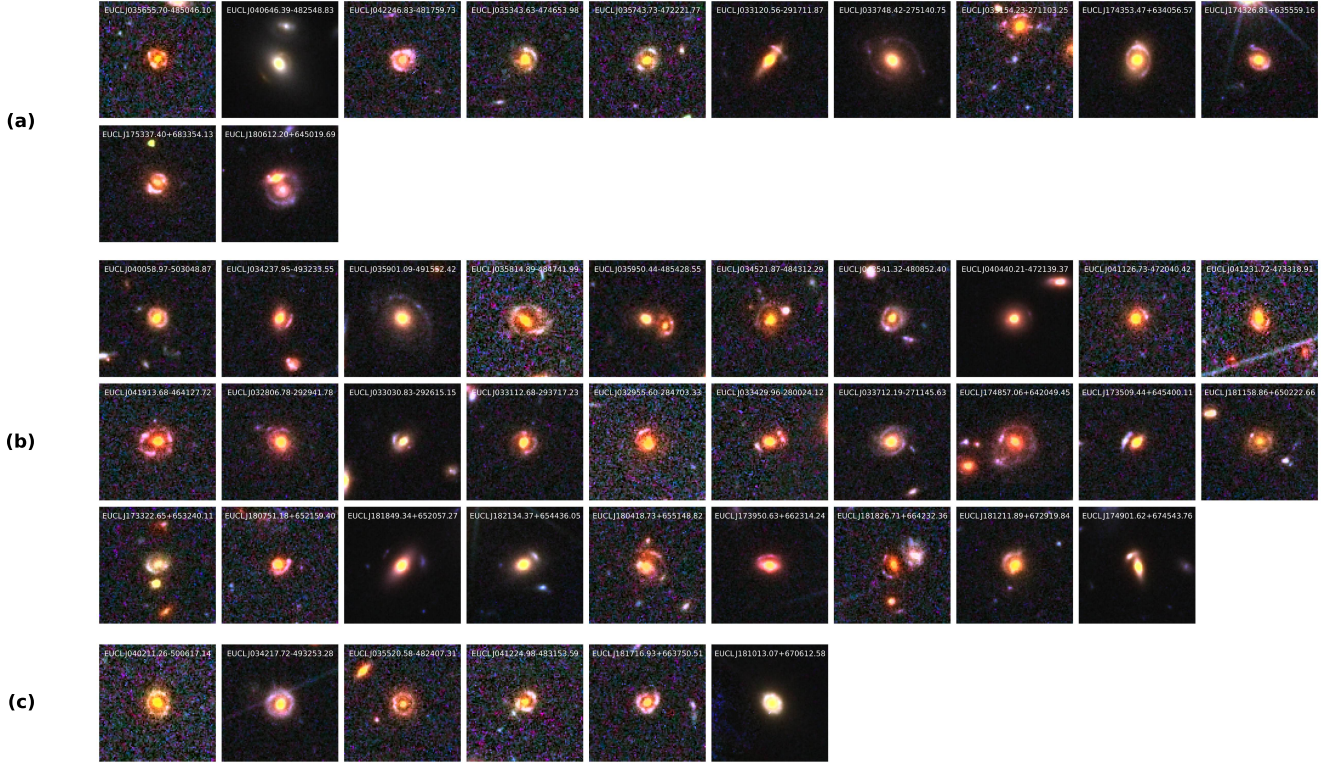


Fig. B.2. Centred lens candidates that received higher grades in this work compared with SLDE A. Panel (a): Objects that received Grade A in this work but were graded as Grade B in SLDE A. Panel (b): Objects that received Grade B in this work but were graded as C in SLDE A. Panel (c): Objects that received Grade A in this work but were graded as C in SLDE A.

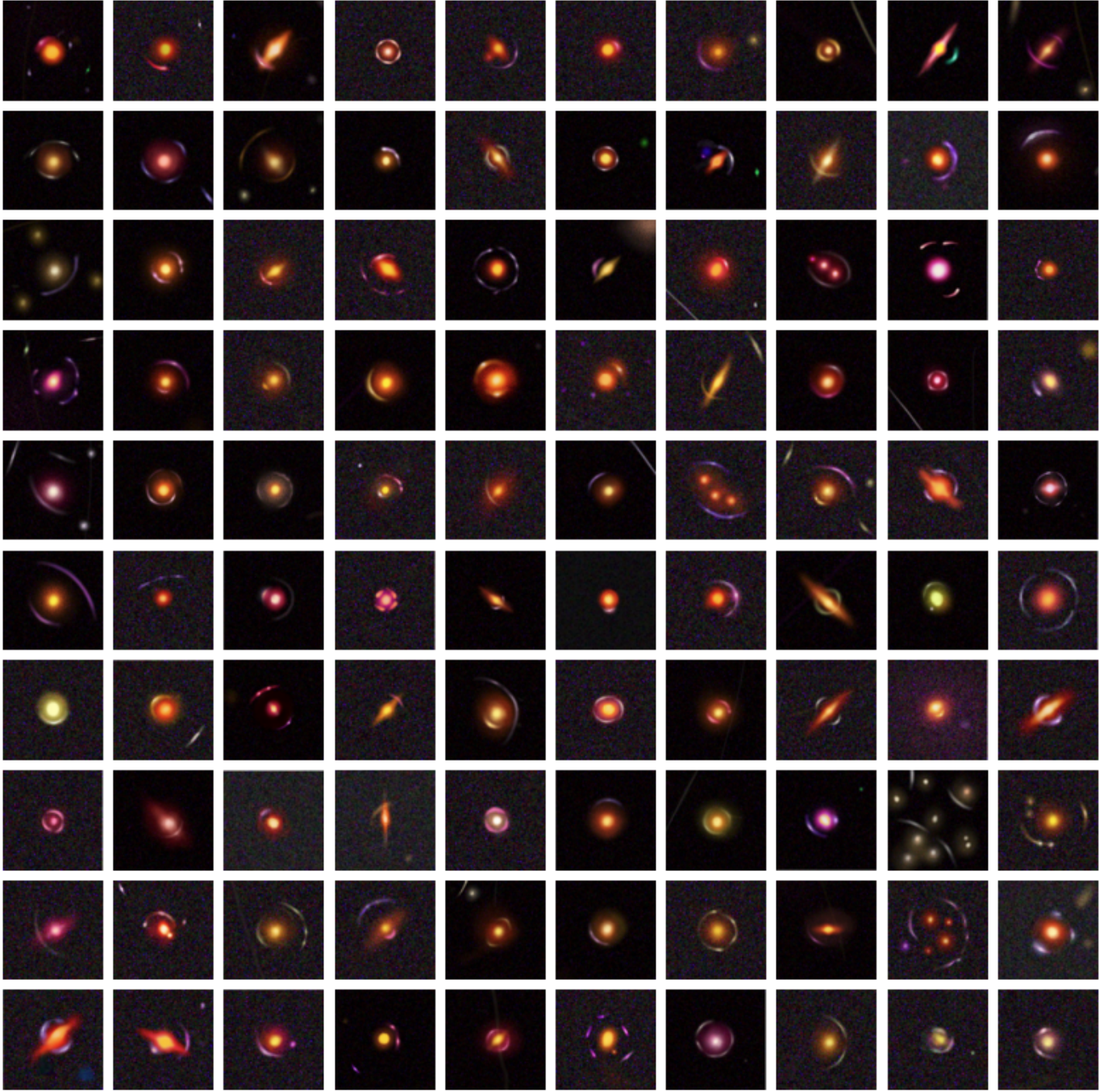


Fig. C.1. 100 experimental hand-drawn lens simulations used to enhance the diversity of the positive training dataset in the second round of the iterative training process. These images were created using Procreate and were based on visual features commonly exhibited among found lenses. These features include the red–orange colour of the lensing galaxy, purple colour of the background galaxy, various observed lensing configurations, etc.

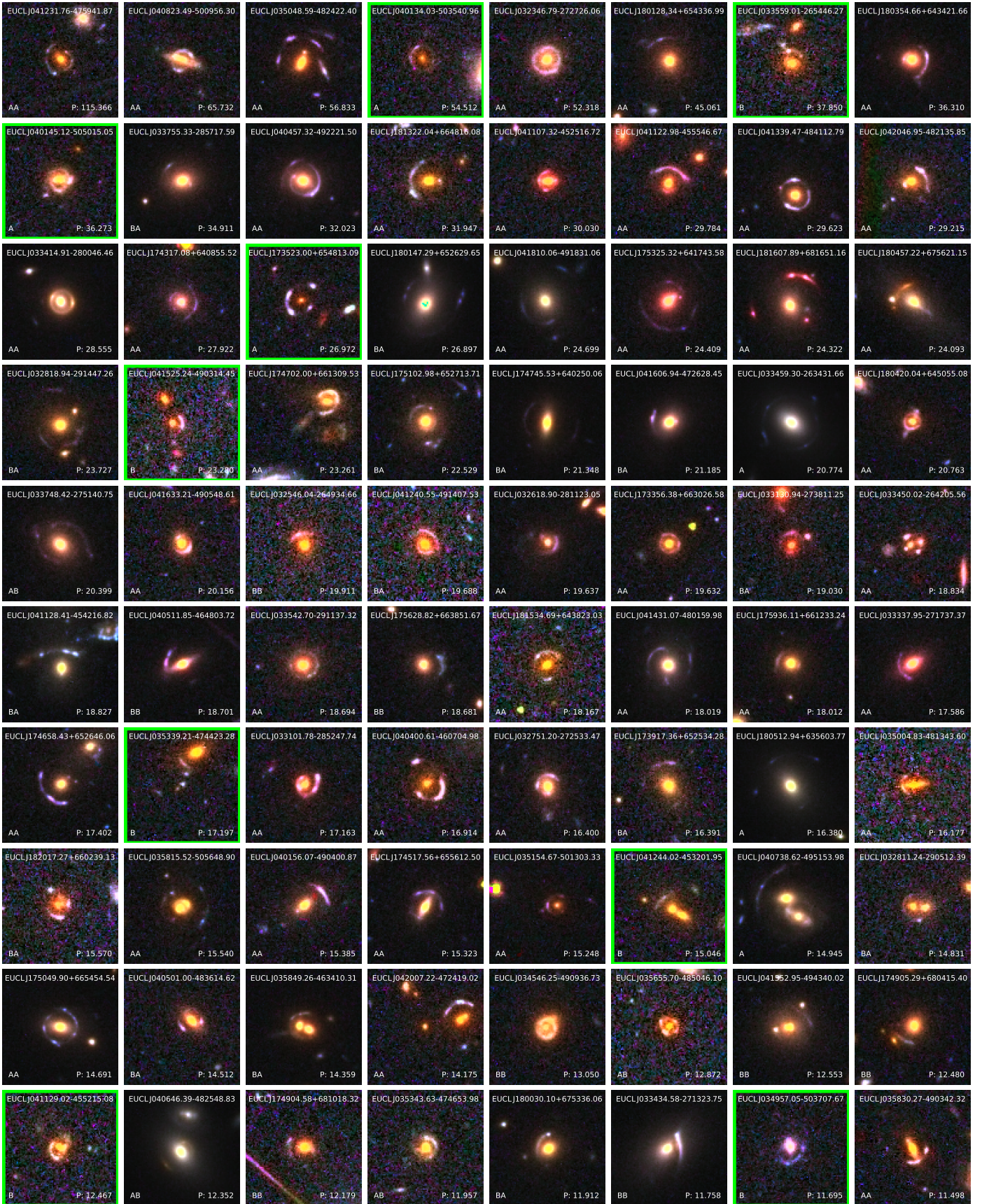


Fig. D.1. 367 classified grade A/B lens candidates (1–80). New lens candidates not included in SLDE A (Euclid Collaboration: Walmsley et al. 2025) are outlined in green. Images displayed are slightly enhanced with an arcsinh function; however, this function was only applied for display – the model itself was applied to linearly stretched images in training and inference.

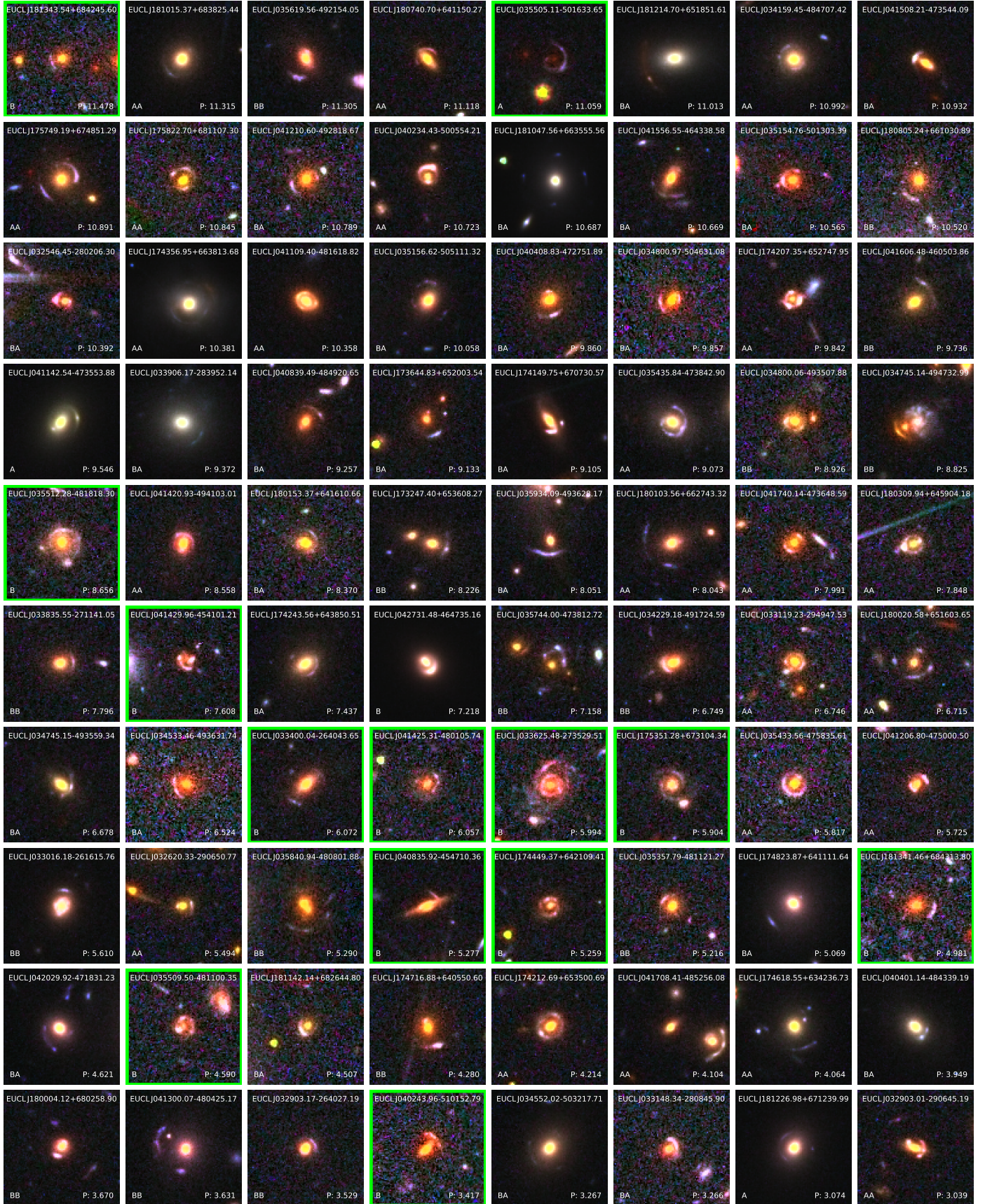


Fig. D.2. Continued: 367 classified grade A/B lens candidates (81–160).

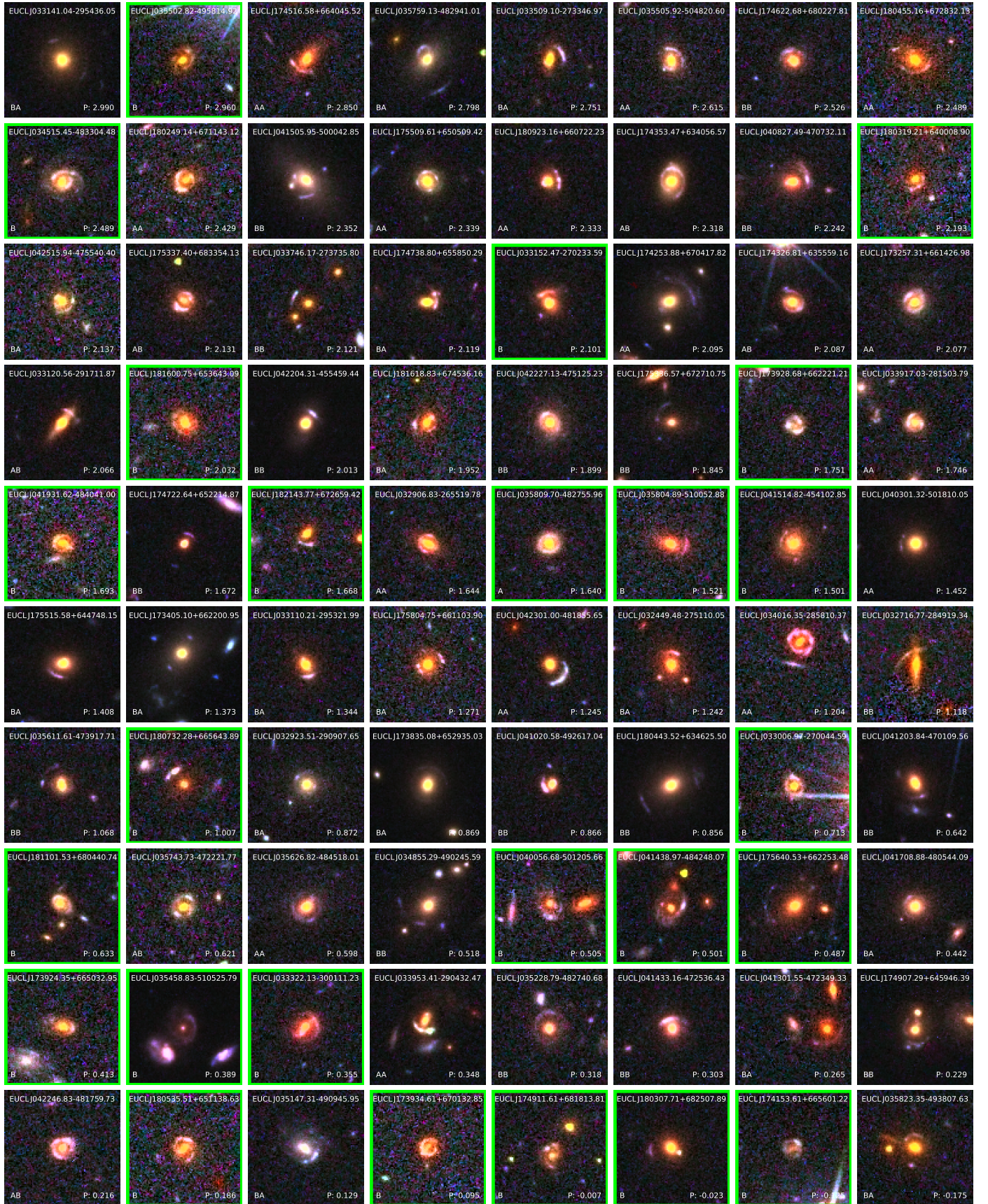


Fig. D.3. Continued: 367 classified grade A/B lens candidates (161–240).

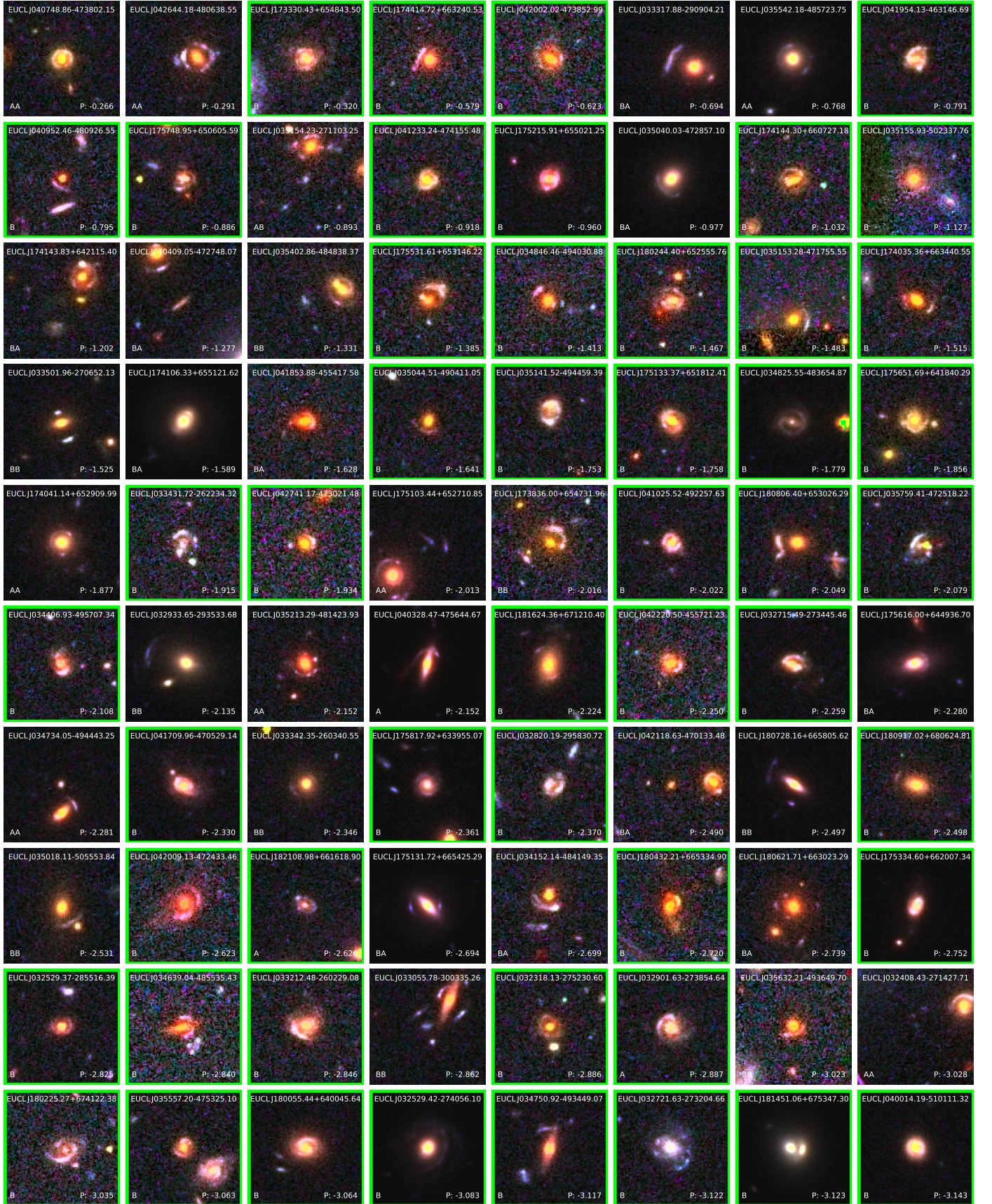


Fig. D.4. Continued: 367 classified grade A/B lens candidates (241–320).

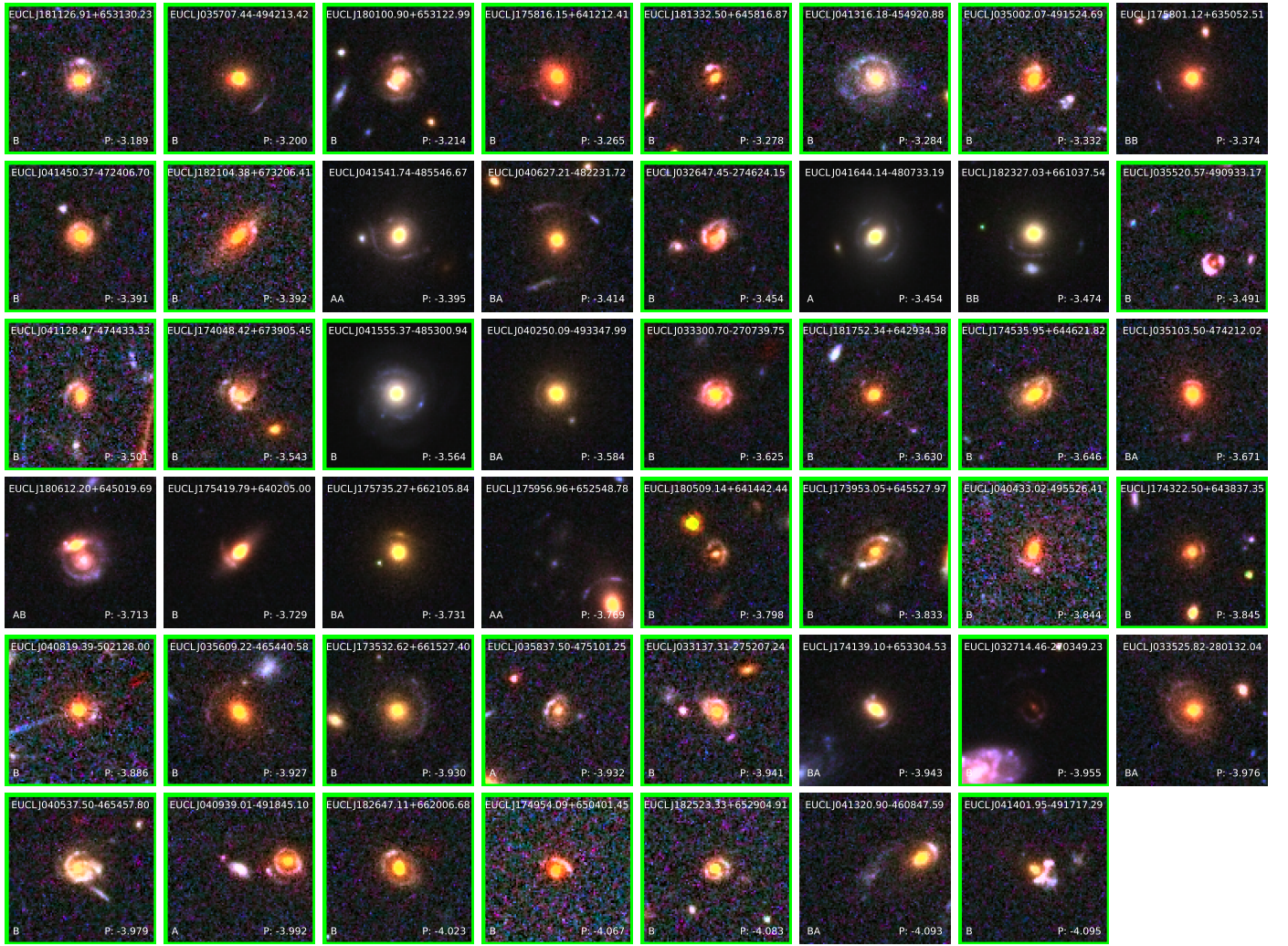


Fig. D.5. Continued: 367 classified grade A/B lens candidates (321–367).

Appendix E: Master table displaying RA, Dec, I_E , Y_E , J_E , and H_E magnitudes, raw prediction output from model, and citations

In Table E.1, ranked by raw scores, we tabulate the coordinates, *Euclid* magnitudes, classification grades, and modelled raw scores of grade A and B lens candidates.

Table E.1. Example entries of the master table of classified lenses depicted in D.

OBJ NAME	RA	Dec	I_E	Y_E	J_E	H_E	GRADE	RAW SCORE	NOTES
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
EUCLJ041231.76-475941.87	63.132342	-47.994963	21.82	20.94	20.55	20.11	AA	115.37	Previously found in L25,W25
EUCLJ040823.49-500956.30	62.097882	-50.165639	21.19	20.57	20.20	19.71	AA	65.73	Previously found in L25,W25
EUCLJ035048.59-482422.40	57.702479	-48.406222	20.40	19.96	19.55	19.10	AA	56.83	Previously found in L25,W25
EUCLJ040134.03-503540.96	60.391806	-50.594710	23.03	21.79	21.20	20.77	A	54.51	
EUCLJ032346.79-272726.06	50.944956	-27.457238	20.68	20.17	19.69	19.35	AA	52.32	Previously found in L25,W25
EUCLJ180128.34+654336.99	270.368101	65.726941	20.35	20.05	19.64	19.25	AA	45.06	Previously found in L25,W25
EUCLJ033559.01-265446.27	53.995875	-26.912852	21.64	20.78	20.31	19.86	B	37.85	
EUCLJ180354.66+643421.66	270.977739	64.572684	20.10	19.80	19.42	19.08	AA	36.31	Previously found in L25,W25
EUCLJ040145.12-505015.05	60.437996	-50.837513	21.42	20.68	20.28	19.95	A	36.27	Ranked poorly by L25 ML
EUCLJ033755.33-285717.59	54.480526	-28.954887	19.19	19.07	18.70	18.34	BA	34.91	Previously found in L25,W25
EUCLJ040457.32-492221.50	61.238847	-49.372639	19.39	19.13	18.85	18.43	AA	32.02	Previously found in L25,W25
EUCLJ181322.04+664816.08	273.341839	66.804466	21.36	20.57	20.17	19.74	AA	31.95	Previously found in L25,W25
EUCLJ041107.32-452516.72	62.780484	-45.421311	21.59	20.59	20.19	19.80	AA	30.03	Previously found in L25,W25
EUCLJ041122.98-455546.67	62.845764	-45.929631	21.25	20.46	20.14	19.61	AA	29.78	Previously found in L25,W25
EUCLJ041339.47-484112.79	63.414457	-48.686887	21.71	21.09	20.78	20.41	AA	29.62	Previously found in L25,W25
EUCLJ042046.95-482135.85	65.195645	-48.359958	21.09	20.45	20.07	19.67	AA	29.21	Previously found in L25,W25
EUCLJ033414.91-280046.46	53.562126	-28.012906	18.95	18.34		17.68	AA	28.55	Previously found in L25,W25
EUCLJ174317.08+640855.52	265.821153	64.148755	21.01	20.52	20.23	20.03	AA	27.92	Previously found in L25,W25
EUCLJ173523.00+654813.09	263.845831	65.803635	23.12	21.95	21.48	21.01	A	26.97	
EUCLJ180147.29+652629.65	270.447058	65.441569	18.44	18.49	18.06	18.40	BA	26.90	Previously found in L25,W25
EUCLJ041810.06-491831.06	64.541932	-49.308628	19.09	19.20	18.82	18.44	AA	24.70	Previously found in L25,W25
EUCLJ175325.32+641743.58	268.355520	64.295438	20.17	19.72	19.33	18.94	AA	24.41	Previously found in L25,W25
EUCLJ181607.89+681651.16	274.032869	68.280879	19.44	19.39	18.98	18.56	AA	24.32	Previously found in L25,W25
EUCLJ180457.22+675621.15	271.238421	67.939208	19.24	19.40	18.98	18.60	AA	24.09	Previously found in L25,W25
EUCLJ032818.94-291447.26	52.078907	-29.246461	19.97	19.61	19.20	18.78	BA	23.73	Previously found in L25,W25
EUCLJ041525.24-490314.45	63.855167	-49.054013	23.78	22.72	22.22	21.81	B	23.28	
EUCLJ174702.00+661309.53	266.758336	66.219313	22.96	22.31	21.84	21.66	AA	23.26	Previously found in L25,W25
EUCLJ175102.98+652713.71	267.762424	65.453808	20.25	19.88	19.48	19.07	BA	22.53	Previously found in L25,W25
EUCLJ174745.53+640250.06	266.939702	64.047238	19.30	18.76	18.37	17.97	BA	21.35	Previously found in L25,W25
EUCLJ041606.94-472628.45	64.028925	-47.441235	19.16	18.93	18.61	18.35	BA	21.19	Previously found in L25,W25

Notes. Currently depicting the top 18 lenses from Fig. D.1. The properties displayed: (1) Object id and name of each candidate; (2) Right ascension; (3) Declination; (4) I_E band magnitude, using FLUX_DETECTION_TOTAL; (5) Y_E -band magnitude, found using FLUX_Y_2FWHM_APER; (6) J_E -band magnitude, found using FLUX_J_2FWHM_APER; (7) H_E -band magnitude, found using FLUX_H_2FWHM_APER; (8) Processed grades of human experts classification. If there are two letters, then the object was previously identified and classified, with grade represented by the second letter; (9) Raw score output from the model in this work; (10) Citations and any additional notes. Two undisplayed properties are included in the full electronically available table in additional resources: an indication of whether the lens is new and an alternative name for each object composed of the mosaic tile and cutout id. Under notes, L25: [Euclid Collaboration: Lines et al. \(2025b\)](#), W25: [Euclid Collaboration: Walmsley et al. \(2025\)](#), R25: [Euclid Collaboration: Rojas et al. \(2025\)](#), and E25: Ecker et al. in prep. R25 and E25 not shown here.