Astronomy & Astrophysics

HOLISMOKES

VIII. High-redshift, strong-lens search in the Hyper Suprime-Cam Subaru Strategic Program[*]

Yiping Shu[1,2], Raoul Cañameras[1], Stefan Schuldt[1,3], Sherry H. Suyu[1,3,4], Stefan Taubenberger[1],
Kaiki Taro Inoue[5], and Anton T. Jaelani[6]

[1] Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str. 1, 85748 Garching, Germany
e-mail: ypshu@mpa-garching.mpg.de
[2] Ruhr University Bochum, Faculty of Physics and Astronomy, Astronomical Institute (AIRUB), German Centre for Cosmological Lensing, 44780 Bochum, Germany
[3] Technische Universität München, Physik Department, James-Franck Str. 1, 85748 Garching, Germany
[4] Institute of Astronomy and Astrophysics, Academia Sinica, 11F of ASMAB, No.1, Section 4, Roosevelt Road, Taipei 10617, Taiwan
[5] Faculty of Science and Engineering, Kindai University, Higashi-Osaka, 577-8502, Japan
[6] Astronomy Research Group and Bosscha Observatory, FMIPA, Institut Teknologi Bandung, Jl. Ganesha 10, Bandung 40132, Indonesia

ABSTRACT

We carry out a search for strong-lens systems containing high-redshift lens galaxies with the goal of extending strong-lensing-assisted galaxy evolutionary studies to earlier cosmic time. Two strong-lens classifiers are constructed from a deep residual network and trained with datasets of different lens-redshift and brightness distributions. We classify a sample of 5 356 628 pre-selected objects from the Wide-layer fields in the second public data release of the Hyper Suprime-Cam Subaru Strategic Program (HSC-SSP) by applying the two classifiers to their HSC *gri*-filter cutouts. Cutting off at thresholds that correspond to a false positive rate of 10^{-3} on our test set, the two classifiers identify 5468 and 6119 strong-lens candidates. Visually inspecting the cutouts of those candidates results in 735 grade-A or B strong-lens candidates in total, of which 277 candidates are discovered for the first time. This is the single largest set of galaxy-scale strong-lens candidates discovered with HSC data to date, and nearly half of it (331/735) contains lens galaxies with photometric redshifts above 0.6. Our discoveries will serve as a valuable target list for ongoing and scheduled spectroscopic surveys such as the Dark Energy Spectroscopic Instrument, the Subaru Prime Focus Spectrograph project, and the Maunakea Spectroscopic Explorer.

Key words. Galaxy: evolution – gravitational lensing: strong – methods: data analysis

1. Introduction

The strong gravitational lensing effect is a powerful and robust mass probe that can deliver precise and accurate measurements of the total mass (including dark matter) in the central regions of galaxies at extragalactic distances. Studies of strong-lens systems have successfully measured dark matter and stellar mass distributions and their evolution in distant galaxies, which have deepened our understanding of galaxy formation and evolution (e.g. Treu et al. 2006; Koopmans et al. 2006; Auger et al. 2010; Bolton et al. 2012a; Brewer et al. 2014; Shu et al. 2015, 2016c). Detections of dark-matter substructures beyond the local Universe and measurements of their masses from strong lensing have placed constraints on the sub-halo mass function and the nature of dark matter (e.g. Vegetti et al. 2010, 2012; Fadely & Keeton 2012; Nierenberg et al. 2014; Hezaveh et al. 2016; Inoue et al. 2016). In addition, the lensing magnification effect can be exploited to study high-redshift objects in detail by overcoming the sensitivity and/or resolution limitations of current

facilities (e.g. Christensen et al. 2012; Bussmann et al. 2013; Stark et al. 2015; Shu et al. 2016b; Marques-Chaves et al. 2017, 2018, 2020; Shu et al. 2022). Moreover, strongly lensed variable sources, such as quasars and supernovae (SNe), have evolved into an independent and compelling cosmological probe (e.g., Suyu et al. 2010, 2013, 2017; Grillo et al. 2018; Wong et al. 2020; Millon et al. 2020), which is one of the main motivations for our Highly Optimised Lensing Investigations of Supernovae, Microlensing Objects, and Kinematics of Ellipticals and Spirals (HOLISMOKES) programme (Suyu et al. 2020).

Various techniques have been developed to identify the intrinsically rare strong-lens systems. The most productive ones to date are imaging-based methods, which have discovered ≈400 confirmed strong-lens systems[1,2,3] (e.g. Browne et al. 2003; More et al. 2012; Stark et al. 2013; Sonnenfeld et al. 2018; Lemon et al. 2018; Shu et al. 2018b, 2019; Chan et al. 2020; Desira et al. 2022). In this work, we consider a strong-lens system as confirmed if multiple lensed images are detected and the

[1] http://admin.masterlens.org/index.php
[2] http://www-utap.phys.s.u-tokyo.ac.jp/~oguri/sugohi/
[3] https://research.ast.cam.ac.uk/lensedquasars/index.html

lens and source redshifts are spectroscopically measured. Over the past two decades, spectroscopy-based methods have heavily exploited large-scale spectroscopic surveys and discovered more than 200 confirmed strong-lens systems (e.g. Bolton et al. 2004, 2008; Treu et al. 2011; Brownstein et al. 2012; Courbin et al. 2012; Shu et al. 2016b,c, 2017; Oldham et al. 2017). Very recently, variability-based methods, which are particularly useful for discovering strongly lensed variable sources, have gained momentum and will undoubtedly play a crucial role in the ongoing and upcoming time-domain surveys (e.g. Kostrzewa-Rutkowska et al. 2018; Chao et al. 2020, 2021; Shu et al. 2021; Bag et al. 2022).

Although the total number of confirmed strong-lens systems have reached ≈ 600[1], many scientific applications call for more systems and a more thorough coverage of the phase space. For example, a lot of effort has been made to search for strongly lensed SNe, which is expected to provide tighter constraints on the Hubble constant compared with strongly lensed quasars (e.g. Oguri & Kawano 2003; Goldstein & Nugent 2017; Wojtak et al. 2019; Huber et al. 2022, 2021; Bayer et al. 2021; Ding et al. 2021). Two efficient approaches of catching such rare and short-lived lensing events are (1) cross-matching transient alerts from time-domain surveys with known strong-lens systems, and (2) carrying out dedicated monitorings of known strong-lens systems with high expected lensed SN rates (e.g. Shu et al. 2018a; Ryczanowski et al. 2020; Craig et al. 2021). Both of these approaches benefit greatly from discovering more strong-lens systems. Additionally, strong-lensing-assisted evolutionary analyses have so far been limited to low- and intermediate-redshift galaxies due to the lack of galaxy-galaxy strong-lens systems with high-redshift lens galaxies. Among all confirmed galaxy-galaxy strong-lens systems, only a handful contain lens galaxies at redshifts beyond 0.8 (e.g. Wong et al. 2014; Cañameras et al. 2017). On the other hand, high-redshift galaxies are crucial to understanding galaxy evolution as they are expected to undergo more frequent and vigorous transitions. Recently, the combination of wide-field imaging surveys and machine learning algorithms has led to a big leap in strong lens discoveries. A few thousand new strong-lens candidates have been uncovered by classifiers built upon supervised or unsupervised algorithms (e.g. Jacobs et al. 2019a; Petrillo et al. 2019; Cañameras et al. 2020, 2021; Huang et al. 2020, 2021; Li et al. 2020, 2021; Stein et al. 2021; Rojas et al. 2021; Savary et al. 2021). Future surveys, such as the *Rubin* Observatory Legacy Survey of Space and Time (LSST, Ivezić et al. 2019), *Euclid* (Laureijs et al. 2011), and the Chinese Space Station Optical Survey (CSS-OS, Zhan 2018), expect to deliver $\sim 10^5$ strong-lens systems (e.g. Collett 2015).

In this work, we focused on extending strong-lensing-assisted evolutionary analyses to earlier cosmic time by searching for high-redshift strong lenses in the Wide-layer data from the second public data release (PDR2) of the Hyper Suprime-Cam Subaru Strategic Program (HSC-SSP, Aihara et al. 2019). In Sect. 2, we describe the HSC-SSP PDR2 data and define our parent sample. Section 3 explains the construction and training of our two strong-lens classifiers based on a deep residual network, and the performance of the two classifiers is shown in Sect. 4. Discovered strong-lens candidates are presented in Sect. 5. Six candidates that show two sets of spectral features at different redshifts in auxiliary spectroscopic data are reported in Sect. 6. Discussions and conclusions are provided in Sects. 7 and 8. To compute the Einstein radii, we adopt a flat ΛCDM cosmology with $\Omega_{\rm m} = 1 - \Omega_\Lambda = 0.32$ (Planck Collaboration VI 2020) and $H_0 = 72\,{\rm km\,s^{-1}\,Mpc^{-1}}$ (Bonvin et al. 2017).

2. Data

In HSC-SSP PDR2, the Wide-layer data cover $\approx 300\,{\rm deg}^2$ to the nominal depths in all five filters (i.e. *grizy*) and additional $\approx 1100\,{\rm deg}^2$ in at least one filter and one exposure. For the PDR2 Wide layer, the median 5σ depths (for point sources) in *grizy* filters are 26.6, 26.2, 26.2, 25.3, and 24.5 mag and the median seeings in *grizy* filters are $0\farcs77$, $0\farcs76$, $0\farcs58$, $0\farcs68$, $0\farcs68$, respectively. A full overview of HSC-SSP PDR2 can be found in Aihara et al. (2019). For our high-redshift strong-lens search, we selected objects that are extended and likely located at high redshifts based on their $g - r$ and $g - i$ colours. To be more specific, we selected objects in the PDR2 Wide layer, that is the `pdr2_wide.forced` table, that satisfy the following criteria:

1. `isprimary` is `True`
2. `i_extendedness_value=1`
3. `[grizy]_sdsscentroid_flag` is `False`
4. `[grizy]_pixelflags_edge` is `False`
5. `[grizy]_pixelflags_interpolatedcenter` is `False`
6. `[grizy]_pixelflags_saturatedcenter` is `False`
7. `[grizy]_pixelflags_crcenter` is `False`
8. `[grizy]_pixelflags_bad` is `False`
9. `[grizy]_cmodel_flag` is `False`
10. `g_cmodel_mag` < 26.0
11. `r_cmodel_mag` < 26.0
12. `i_cmodel_mag` < 26.0
13. 0.6 < `g_cmodel_mag-r_cmodel_mag` < 3.0
14. 2.0 < `g_cmodel_mag-i_cmodel_mag` < 5.0

This query returns 5 356 628 unique HSC objects in total, which form the parent sample of this lens search project. Here, criteria 3–12 are used to remove objects with unreliable photometry (e.g. Tanaka et al. 2018; Schuldt et al. 2021a), and the colour-colour cuts in criteria 13–14 are directly taken from Jacobs et al. (2019b) to select red and potentially high-redshift galaxies. The HSC CModel photometry algorithm is presented in detail in Bosch et al. (2018). In summary, the single-filter imaging data of an object are fitted separately with an elliptical exponential model or with an elliptical de Vaucouleurs model, where each model is convolved with the point spread function (PSF). The CModel magnitude is subsequently computed from a composite model that is constructed as a linear combination of the previous exponential and de Vaucouleurs models, which best fit the imaging data. Since the CModel photometry is based on a reasonable analytical description of galaxy morphology, we expect it to provide more robust colour estimates than the fixed-aperture or Kron photometry that are also available in PDR2, especially for lens galaxies in strong-lens systems. We find that criteria 13–14 manage to substantially reduce the sample size and at the same time maintain a high completeness rate for high-redshift lens galaxies. Removing criteria 13–14 in the above query would have resulted in a sample of 79 577 619 unique extended objects, which in turn would have posed challenges to not only the final lens search but also the initial imaging data retrieval. On the other hand, Jacobs et al. (2019b) simulated 10 000 $z > 0.8$ elliptical galaxies with lensing features superimposed and found that $\gtrsim 90\%$ of the simulated lenses can be recovered with these two colour-colour cuts. In addition, we examined the colour distributions of strong-lens candidates discovered in the HSC footprint by the Survey of Gravitationally-lensed Objects in HSC Imaging (SuGOHI) project (Sonnenfeld et al. 2018, 2020; Wong et al. 2018; Chan et al. 2020; Jaelani et al. 2020). Every SuGOHI strong-lens candidate is assigned a grade of A (definite), B (probable), or C (possible) and a lens type from GG (galaxy-galaxy), GQ (galaxy-quasar), CG (cluster- or group-galaxy), or

CQ (cluster- or group-quasar). As we are particularly interested in galaxy-galaxy strong lenses, we focused on the 99 SuGOHI grade-A or B GG strong-lens candidates that have lens galaxies fulfilling criteria 1–12. The lens galaxies in those strong-lens candidates are primarily luminous red galaxies selected according to the criteria defined in Dawson et al. (2013). They span a wide redshift range from 0.2 to 1.0[4]. We note that candidates from Sonnenfeld et al. (2020) are not considered here because some GG strong-lens candidates therein are actually cluster- or group-scale lenses. Among the selected 99 SuGOHI strong-lens candidates, 92 (or ≈93%) further pass the colour-colour cuts in criteria 13–14. Limiting to the selected SuGOHI candidates with lens galaxy (spectroscopic or photometric) redshifts above 0.8, 4/5 (or 80%) pass the colour-colour cuts. Although the colour-colour cuts were originally defined in the photometric system of the Dark Energy Survey, we expect them to be similarly effective in the HSC photometric system given the minor difference between them (Abbott et al. 2021) and the encouraging results from the SuGOHI sample.

The HSC *gri*-filter cutouts (72 pixel × 72 pixel, 1 pixel = $0\rlap{.}{''}17$) centred on the 5 356 628 objects in our parent sample are retrieved from the PDR2 image cutout service. Photometry (CModel magnitudes from the `pdr2_wide.forced` table, Aihara et al. 2019) and photometric redshift (`photoz_best` from the `pdr2_wide.photoz_mizuki` table, Tanaka et al. 2018) for every object in the parent sample are also retrieved from the HSC CAS Search service. The parent sample covers roughly 960 deg^2.

3. Strong-lens classifier construction

We constructed our strong-lens classifiers based on the deep residual network, `deeplens_classifier`, pre-built in the `CMU DeepLens` package (Lanusse et al. 2018). Deep residual networks (resnets), a variation of convolutional neural networks, have become the current state-of-the-art imaging recognition algorithm, and `CMU DeepLens` adopts a specific resnet architecture proposed by He et al. (2016). Among the nine different lens-finding methods in the strong gravitational lens finding challenge (Metcalf et al. 2019), `CMU DeepLens` delivered the highest area under the receiver operating characteristic curve (AUROC) value, which is the most commonly used evaluation metric for classification problems. It is also top-ranked on TPR$_0$ and TPR$_{10}$, which correspond to the highest true positive rate reached before more than 0 and 10 false positives occur, respectively. We therefore chose `deeplens_classifier` from `CMU DeepLens` as our baseline model, and a full description of the network architecture can be found in Lanusse et al. (2018). The `deeplens_classifier` network is constructed such that it returns a number from 0 to 1 for every input system, which is referred to as the network score p_{resnet} in this work.

The `deeplens_classifier` network takes several parameters that determine how the actual training is done. In particular, `learning_rate` sets the initial learning rate, `learning_rate_steps` sets the number of learning rate updates during training, `learning_rate_drop` sets the amount by which the learning rate is updated, and `n_epochs` sets the total number of training epochs. For example, the network that delivered the highest AUROC value in the strong gravitational lens finding challenge had `learning_rate` = 0.001, `learning_rate_steps` = 3, `learning_rate_drop` = 0.1, and `n_epochs` = 120, which correspond to a starting learning rate of

0.001 that is multiplied by 0.1 every 40 epochs. We always use a `learning_rate_drop` of 0.1 for our classifiers.

In this work, we test two strong-lens classifiers. The main difference between the two is the properties of mock lenses in the training set. This allows us to investigate, among others, the impact of the training set on classifier performance. In addition, combining the results from the two classifiers yields a much more complete sample of strong-lens candidates, as we demonstrate later.

3.1. Classifier-1

3.1.1. Training and validation datasets

As the sample size of confirmed strong lenses is still small (of the order of 10^3), mock lens systems need to be created for training and validation. We tried to be as realistic as possible by using observed data of real galaxies to make the mock systems. Following Cañameras et al. (2021, C21 hereafter), we selected ≈80 000 galaxies from data release 14 of the Sloan Digital Sky Survey (SDSS, Abolfathi et al. 2018) that are also in the HSC footprint and have measured spectroscopic redshifts and velocity dispersions (Bolton et al. 2012b) as the lens sample. We directly took HSC *gri*-filter cutouts (72 pixel × 72 pixel) centred on those lens galaxies as the base layer. As a result, mock lens systems naturally include various observational effects, such as galaxy colour gradients, seeing variations, neighbouring and line-of-sight contaminants, and artefacts, that are also present in the parent sample. To further enlarge the lens sample, we rotated every galaxy in the lens sample by 90°, 180°, and 270°, and considered them as different lens galaxies. This implies that each galaxy in the lens sample is used four times at most. For the source sample, we used ≈1200 high signal-to-noise ratio (S/N) galaxies in the *Hubble* Ultra Deep Field with secure spectroscopic redshifts (Inami et al. 2017). We converted images of the selected source galaxies in HST bands (*F435W*, *F606W*, and *F775W*) to HSC *gri* filters using the method in Cañameras et al. (2021).

Similarly to procedures used in Cañameras et al. (2020, 2021) and Schuldt et al. (2021b), we modelled the effective lensing potential as two components: a projected lens mass component characterised by a singular isothermal ellipsoid (SIE) profile and an external shear. The axis ratio and position angle of the SIE profiles were set to values inferred from the lens surface-brightness distribution in the HSC *i* band. The external shear strength was randomly drawn from a Gaussian distribution with mean 0 and standard deviation 0.058 (e.g. Wong et al. 2011; Faure et al. 2011) and the position angle was randomly chosen from 0° to 180°. For every lens galaxy, we randomly paired it with a galaxy from the source sample that is at a redshift higher than the lens galaxy. The Einstein radius of the SIE profile can then be computed from the lens and source redshifts and the lens velocity dispersion. The selected source galaxy is randomly placed with a requirement that its centroid needs to be at a location with a total magnification of 5 or more. We used GLEE (Suyu & Halkola 2010; Suyu et al. 2012) to generate the lensed image of the source, which is further downsampled to the HSC pixel size and convolved with the PSF at the location of the lens provided by the HSC PSF picker. We required the brightest pixel in the lensed image to be brighter than the corresponding pixel in the base layer in either *g*- or *i*-band. Otherwise, we draw a new source position, generate the lensed image, and compare. This process can be iterated 40 times at most, after which point the brightness of the selected source galaxy is boosted by 0.5 mag in

---

[4] According to the spectroscopic or photometric redshifts available on http://www-utap.phys.s.u-tokyo.ac.jp/~oguri/sugohi/
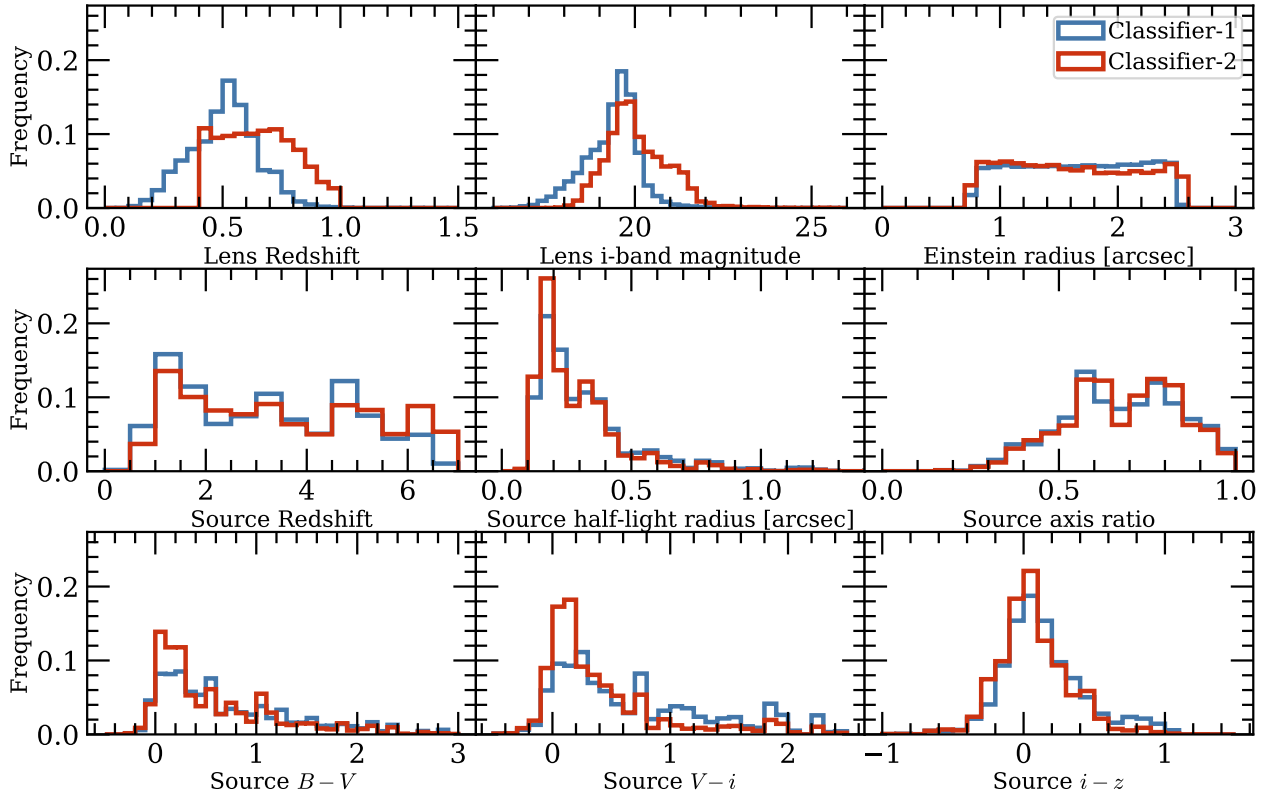
Fig. 1. Distributions of lens galaxy redshift, lens galaxy i−band magnitude, Einstein radius, source galaxy redshift, source galaxy half-light radius, source galaxy axis ratio, and source galaxy $B-V$, $V-i$, and $i-z$ colours for mock lenses in the training sets for Classifier-1 (blue) and Classifier-2 (red).

all three bands and the whole process is repeated. If the requirement is still not satisfied after boosting the selected source by 5 mag, a new source galaxy is selected from the source sample. Once the requirement is satisfied, the lensed image is added to the base layer to produce the composite image of a mock lens system.

For this classifier, we specifically selected 43 500 mock lens systems that produce a close to uniform Einstein radius distribution between 0″.75 and 2″.5 as positive examples. The Einstein radius is the single most important quantity of a strong-lens system, and is determined primarily by the lens galaxy mass with an additional dependence on the lens and source redshifts. We choose a uniform Einstein radius distribution so that the classifier is equally sensitive to galaxy-scale strong-lens systems with different image separations. We tried training with mock lenses that have more naturally distributed Einstein radii, that is starting from 0″.75 and decreasing towards larger radii. The corresponding classifier had a lower overall TPR and failed to recover some of the obvious strong-lens candidates with large Einstein radii in the test set. To ensure the translation invariance of the classifier, for each mock lens system we extracted a 60 pixel × 60 pixel gri cutout (roughly 10″ × 10″) randomly centred within ±5 pixels in both the RA and Dec directions of the centre of the original cutout (72 pixel × 72 pixel), and we refer to the 43 500 cutouts as the lens dataset. Considering that the largest Einstein radii of our mocks are 2″.5 and shifts up to 0″.85 in each direction are applied, 10″ × 10″ cutouts are needed and are sufficient to ensure all the lensing features are seen by the classifier. Using larger cutouts will presumably lead to classifier performance degradation as the chance of contamination due to irrelevant objects in the cutouts increases quadratically with the cutout size. As indicated by Fig. 1, the redshift distribution of lens galaxies in

this training set peaks at ≈0.55. The i-band magnitude distribution of lens galaxies peaks at ≈19.5 mag and drops rapidly towards the faint side. In fact, the magnitude distribution of the lens galaxies, which are all spectroscopically-observed galaxies in the SDSS surveys, is primarily due to SDSS selection effects. In SDSS-III, galaxies selected for spectroscopic observations are all brighter than $i = 19.9$ (Dawson et al. 2013), and the faint limit for galaxy target selection extends to $i \leq 21.8$ in SDSS-IV (Prakash et al. 2016). Distributions of several source galaxy properties are extracted from Beckwith (2005) and Inami et al. (2017) and shown in Fig. 1. We note that the source redshift distribution is biased because of the applied artificial source brightness boosting (by up to 5 mag) during the generation of mocks.

To construct the non-lens examples for training and validation, we first randomly select 48 213 objects from the parent sample. To further clean this subset, we cross-matched them with a sample of 10 241 known strong lenses and strong-lens candidates (referred to as the known strong lens compilation hereafter) compiled from the literature (e.g. Diehl et al. 2017; Sonnenfeld et al. 2018, 2020; Wong et al. 2018; Petrillo et al. 2019; Jacobs et al. 2019b,a; Chan et al. 2020; Jaelani et al. 2020; Huang et al. 2020, 2021; Cañameras et al. 2020, 2021; Li et al. 2020, 2021; Rojas et al. 2021; Savary et al. 2021) using a matching radius of 30 arcsec, and we removed the 114 matches. Considering the typical lensing rate of 10^{-4}–10^{-3} (e.g. Browne et al. 2003; Bolton et al. 2004; Oguri & Marshall 2010; Treu 2010), the remaining 48 099 objects are expected to be sufficiently pure. Among them, 43 500 objects are randomly selected as the final non-lens examples (to match the size of the lens dataset). Similarly, a random shift within ±5 pixels in both directions is applied simultaneously to the gri-filter cutouts of each non-lens example. The
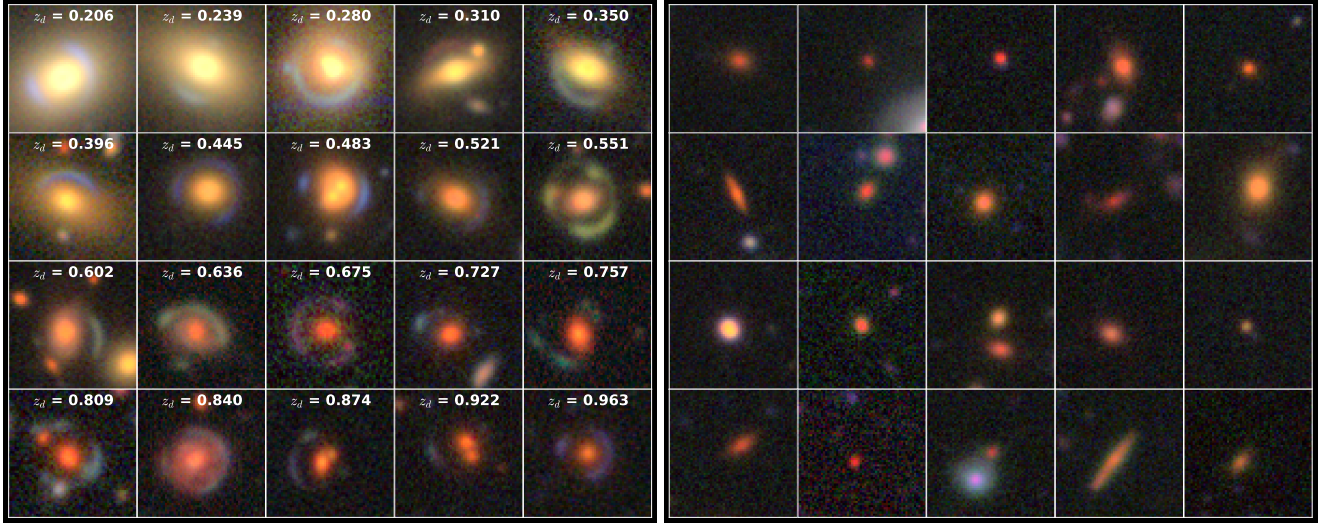
Fig. 2. Colour composite images of 20 mock lenses (*left*, ordered by lens galaxy redshift) and 20 non-lens examples (*right*) selected from the training set for Classifier-1.

shifted *gri*-filter cutouts of the 43 500 objects are trimmed to 60 pixel × 60 pixel and form the non-lens dataset.

The lens and non-lens datasets are merged into a single dataset, which is then randomly shuffled. 80% of the shuffled dataset is used for training and the remaining 20% is used for validation. Twenty mock lens systems and twenty non-lens systems randomly selected from the training set are shown in Fig. 2 as an illustration.

3.1.2. Test dataset

To construct the non-lens examples for the test set, we first randomly selected 53 570 objects from the parent sample. To further clean this subset, we cross-matched them with the known strong lens compilation from the previous step and the 43 500 non-lens examples used for training and validation using a matching radius of 30 arcsec, and we removed the 152 and 1649 matches. 50 000 objects were randomly selected from the remaining objects, and their *gri*-filter cutouts were trimmed to 60 pixel × 60 pixel and form the non-lens examples of the test set.

To construct the lens examples for the test set, we used strong lenses and strong-lens candidates from the SuGOHI project. The SuGOHI project has discovered 2002 strong lenses and strong-lens candidates based on HSC imaging data (Sonnenfeld et al. 2018, 2020; Wong et al. 2018; Chan et al. 2020; Jaelani et al. 2020), of which 1411 systems pass our selection criteria in Sect. 2 and are included in our parent sample. As we are particularly interested in our network's ability to discover galaxy-galaxy strong lenses, we only included 23 grade-A and 69 grade-B galaxy-galaxy strong-lens candidates from the 1411 SuGOHI systems in the test set. Again, their *gri*-filter cutouts are trimmed to 60 pixel × 60 pixel and form the lens examples of the test set. For the sake of simplicity, candidates from Sonnenfeld et al. (2020) are also not included in this step because some classified GG strong-lens candidates therein are actually cluster- or group-scale systems.

3.1.3. Network tuning

To quantify the network performance, we examined the true positive rate (TPR) and false positive rate (FPR). The TPR and FPR

are defined as follows:

$$
\mathrm{TPR} = \frac{\text{Number of lenses that are correctly classified as lenses}}{\text{Number of lenses in a dataset}},
$$
(1)

$$
\mathrm{FPR} = \frac{\text{Number of non-lenses that are mis-classified as lenses}}{\text{Number of non-lenses in a dataset}}.
$$
(2)

As mentioned previously, the network performance is usually measured by the AUROC metric for such a classification problem. The receiver operating characteristic (ROC) curve is the relation between TPR and FPR when the network score threshold varies from 0 to 1, and the AUROC is the integration of the ROC curve. For reference, a perfect classifier has an AUROC of 1.0, which is the best possible value, and a classifier that makes random predictions has an AUROC of 0.5.

For this classifier, we explore three different options for network parameters `learning_rate`, `learning_rate_steps`, and `n_epochs`. The first option corresponds to the default values that delivered the highest AUROC value in the Strong Gravitational Lens Finding Challenge, that is [0.001, 3, 120] (in the format of [`learning_rate`, `learning_rate_steps`, `n_epochs`]. The other two options are [0.01, 4, 160] and [0.1, 5, 200]. The network that is trained with [0.01, 4, 160] has the highest AUROC on the test dataset, and it was therefore chosen to be the final network for Classifier-1.

3.2. Classifier-2

3.2.1. Training and validation datasets

As the main focus of this work is finding high-redshift strong lenses, we experimented with a different training set that contains a higher fraction of high-redshift ($z \gtrsim 0.6$) lenses compared to the training set used for Classifier-1. We used the same procedures outlined in Sect. 3.1.1 to create mock lenses. The only difference is, at this point we manually adjusted the redshift distribution of the lens galaxies to a relatively uniform distribution from 0.4 to 1.0 (Fig. 1) when creating the mocks. Because the number of $z > 0.8$ galaxies in the lens sample is relatively small
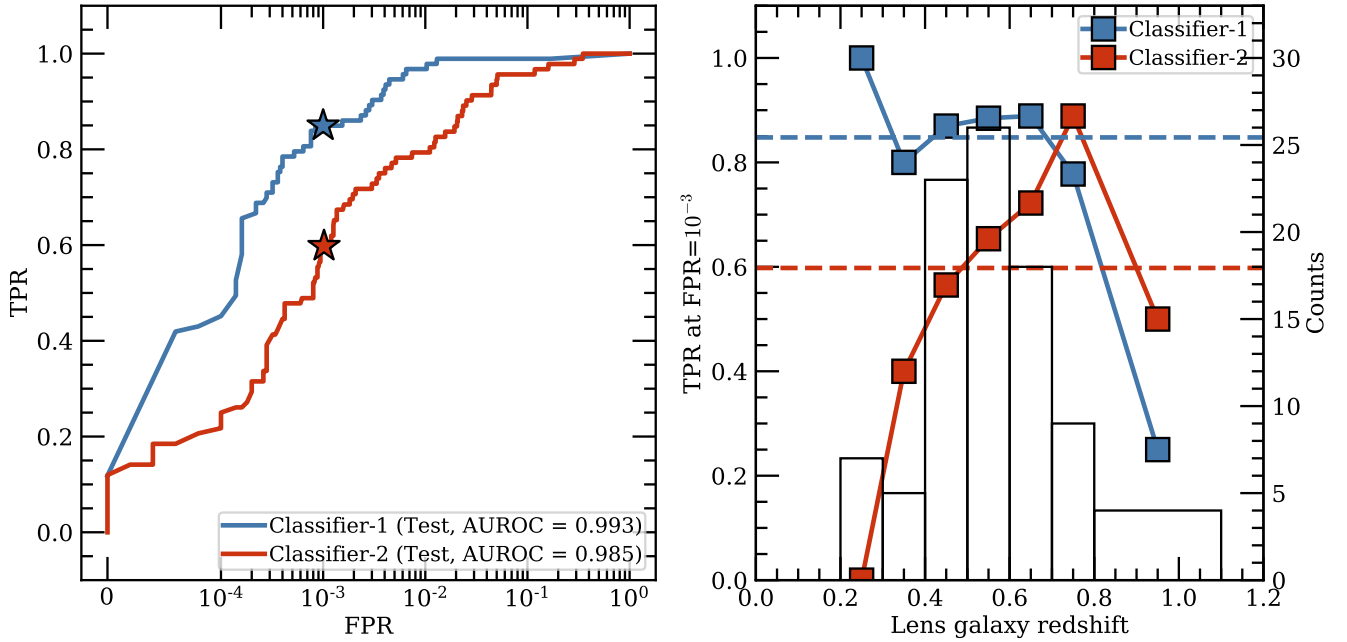
Fig. 3. Performances of the two classifiers. *Left:* ROC curves based on the test sets for Classifier-1 (blue) and Classifier-2 (red). The *x*-axis is scaled such that 0–10^{-4} is in a linear scale and 10^{-4}–1 is in a logarithmic scale. The two star symbols correspond to $\text{FPR} = 10^{-3}$. *Right:* TPR at an FPR of 10^{-3} as a function of lens galaxy redshift for Classifier-1 (blue) and Classifier-2 (red). The dashed lines indicate the overall TPRs of 0.85 and 0.60 for Classifier-1 and Classifier-2, respectively. Due to the small sample size, the last redshift bin is chosen to be 0.8–1.1. A histogram of the lens galaxy redshifts of the 92 SuGOHI strong-lens candidates in the test set is also shown (black).

and each galaxy was only used at most four times, the total number of mock lens systems was 28 500. We therefore augmented the mock lens sample by vertically flipping the cutouts of the 28 500 mock lens systems and considered them as new mock lens systems. 56 960 mock systems were then randomly selected from those 57 000 systems, which we used as the final sample of mock lenses. This new set of mocks has a similar close-to-uniform Einstein radius distribution but clearly contains a higher fraction of higher redshift and fainter lens galaxies, as indicated in Fig. 1. Source galaxy properties in these new mocks are not significantly different from those in the mocks for Classifier-1. There is a slightly higher fraction of source galaxies with smaller sizes or bluer $B - V$ and $V - i$ colours, most of which turn out to be at redshifts above 6. For the non-lens examples, we randomly selected another 56 960 objects from the parent sample that do not have counterparts in the known strong lens compilation and the test set for Classifier-1. The randomly shifted *gri*-filter cutouts (60 pixel × 60 pixel) of the 113 920 mock lenses and non-lens examples are merged into a single dataset, which is again randomly shuffled.

In addition, two pre-processing steps were introduced. We first took the square root of the absolute value of the dataset. Considering that the lensing features are generally fainter than the lens galaxies, especially in *r* and *i* filters, this square-root stretch step improves the contrast between the lens galaxy and lensing features, which has been found to improve the performance of the network (Cañameras et al., in prep.). Afterwards, we normalised the cutouts of every system in the dataset so that the brightest pixel in the individual filter always has a value of 1. Moreover, instead of one network, Classifier-2 is composed of ten networks that are trained with different training sets. This is achieved by implementing the k-fold cross-validation process. More specifically, the single dataset mentioned above was divided into ten chunks of equal size. Each of the ten chunks was used consecutively as the validation set, and the remaining

nine chunks were used to train a network. In total, ten networks are obtained, and the average of their output p_{resnet} is used as the final p_{resnet} for every input system.

3.2.2. Test dataset

The same 92 lens and 50 000 non-lens examples introduced in Sect. 3.1.2 were used to construct the test dataset for Classifier-2. The only difference is, their *gri*-filter cutouts also underwent the square-root stretch and normalisation steps.

3.2.3. Network tuning

Similarly, we considered the following three different options of network parameters `learning_rate`, `learning_rate_steps`, and `n_epochs`: $[0.001, 3, 120]$, $[0.01, 4, 160]$, and $[0.1, 5, 150]$. The set of ten networks that were trained with $[0.1, 5, 150]$ delivered the highest AUROC on the test dataset, and these were chosen as the final networks for Classifier-2.

4. Classifier performances

Figure 3 shows the ROC curves for Classifier-1 (blue) and Classifier-2 (red) based on the test dataset. Classifier-1 has an AUROC of 0.993 and Classifier-2 has an AUROC of 0.985. For reference, the highest AUROC reported in the strong gravitational lens finding challenge was 0.98 (Metcalf et al. 2019). Cañameras et al. (2020) obtained an AUROC of 0.985 and Huang et al. (2021) obtained an AUROC of 0.992. Although the AUROC values from different work cannot be directly compared because they are evaluated on different test sets, our AUROC values being in the ballpark of the highest values achieved by recent strong lens classifiers based on neural networks suggests that our two classifiers have been well trained.
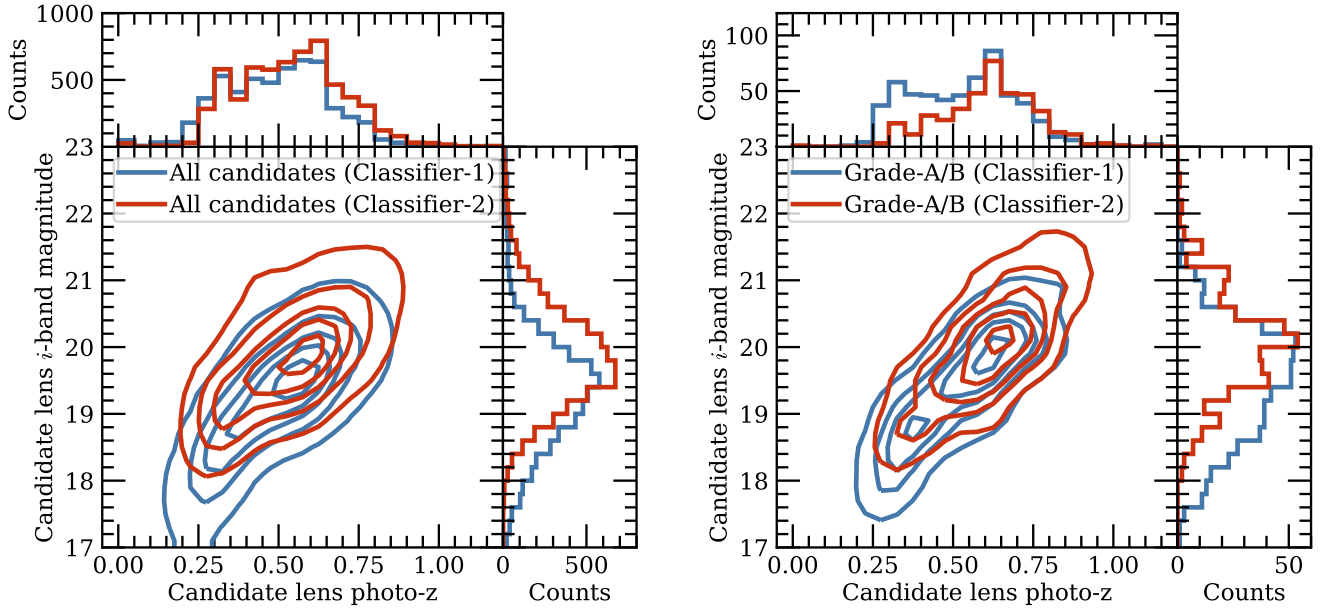
Fig. 4. Distributions of photometric redshift and i−band magnitude of the lens galaxies in strong-lens candidates found by our two classifiers (*left*) and the sub-samples that are classified as grade-A or grade-B after visual inspections (*right*). In both panels, the contours correspond to 10th, 30th, 50th, 70th, and 90th percentiles of the individual dataset.

For each classifier, we selected a $p_{\rm resnet}$ threshold that delivers an FPR of 10^{-3} as the fiducial threshold. Considering the typical strong-lensing rate of 10^{-4}–10^{-3} (e.g. Browne et al. 2003; Bolton et al. 2004; Oguri & Marshall 2010; Treu 2010), an FPR of 10^{-3} can ensure a reasonable balance between true positives and false positives. In addition, ≈6000 objects in our parent sample (with ≈5.36 million objects) are expected to pass the $p_{\rm resnet}$ threshold, which is still manageable in terms of visual inspections. For Classifier-1, the threshold is $p_{\rm resnet} = 0.9731$ and the corresponding TPR is 0.85. For Classifier-2, the threshold is $p_{\rm resnet} = 0.987$ and the corresponding TPR is 0.60. Breaking down into individual redshift bins, we find that the TPRs at an FPR of 10^{-3} for Classifier-1 are in agreement with its overall TPR of 0.85 for lens galaxy redshifts from 0.2 to 0.7, beyond which point it drops substantially to TPR = 0.25 in the redshift bin of 0.8–1.1 (Fig. 3). For Classifier-2, the TPRs for lens galaxy redshifts from 0.2 to 0.4 are lower than its overall TPR of 0.65, presumably because there is no lens galaxy in the training set that is below the redshift of 0.4 for Classifier-2. The TPR reaches the overall TPR level of 0.65 after the redshift of 0.4 and keeps increasing to almost 0.90 in the redshift bin of 0.7–0.8. In the redshift bin of 0.8–1.1, the TPR for Classifier-2 is 0.50. It becomes clear that even though the overall TPR for Classifier-2 is lower compared to Classifier-1, Classifier-2 is expected to outperform Classifier-1 in discovering strong-lens candidates with lens galaxy redshifts above 0.7. As is shown in the next section, this is further supported by the fact that Classifier-2 has discovered more high-redshift strong-lens candidates from the same parent sample.

5. Strong lens candidates in the HSC

5.1. Candidates from Classifier-1

Applying Classifier-1 to our parent sample returned 5468 unique objects with $p_{\rm resnet} \geq 0.9731$. This fraction, that is 5468/5 356 628 = 0.00102, is consistent with the FPR of 10^{-3} inferred from the test set, which suggests that Classifier-1 is not over-fitted. Those 5468 objects were considered as strong-lens candidates and passed to visual inspections. The photometric redshift and i-band magnitude distributions for the candidate lens galaxies are shown in Fig. 4 (red contours).

For the visual inspections, author Y. S. performed an initial check of all the 5468 objects and removed 1479 obvious non-lenses, which are mostly spiral galaxies, clearly isolated objects, and artefacts. Five authors (Y.S., R.C., S.S., S.H.S., and S.T.) then independently inspected the colour composite cutouts ($10'' \times 10''$, constructed from *gri* filters) with different scaling schemes and contrasts for the remaining 3989 objects and assigned an integer score between 0 and 3 to each system following the criteria adopted in Sonnenfeld et al. (2018), Cañameras et al. (2020, 2021). Specifically, score 3 corresponds to definite lenses with clear multiple images in configurations that a lens model can easily reproduce. Score 2 corresponds to probable lenses that have extended and distorted arcs but no clear signs of counter-images and/or would require a lens model to explain the configuration. Score 1 corresponds to possible lenses with single arcs far away from the central galaxy, and score 0 corresponds to non-lenses including spirals, ring galaxies, and everything else. The standard deviation of the scores from the five graders was computed for every system. We note that objects with high standard deviations usually show ambiguous arc-like features, which can be interpreted as either lensed background sources or spiral arms of the central galaxies. 531 objects with standard deviations above 0.75 were therefore re-graded by the five graders.

The visual-inspection scores were averaged over the five graders. 92 systems with average scores $\langle S \rangle \geq 2.5$ are considered as grade-A strong-lens candidates and 468 systems with $1.5 \leq \langle S \rangle < 2.5$ are considered as grade-B strong-lens candidates. Among the 5468 systems that were inspected, there are 78 grade-A or B SuGOHI galaxy-galaxy strong-lens candidates (again excluding candidates from Sonnenfeld et al. (2020) for the sake of simplicity), and 71 of them have average scores $\langle S \rangle \geq 1.5$. The recall of our visual-inspection procedure is therefore estimated to be 91%. The photometric redshift and i−band
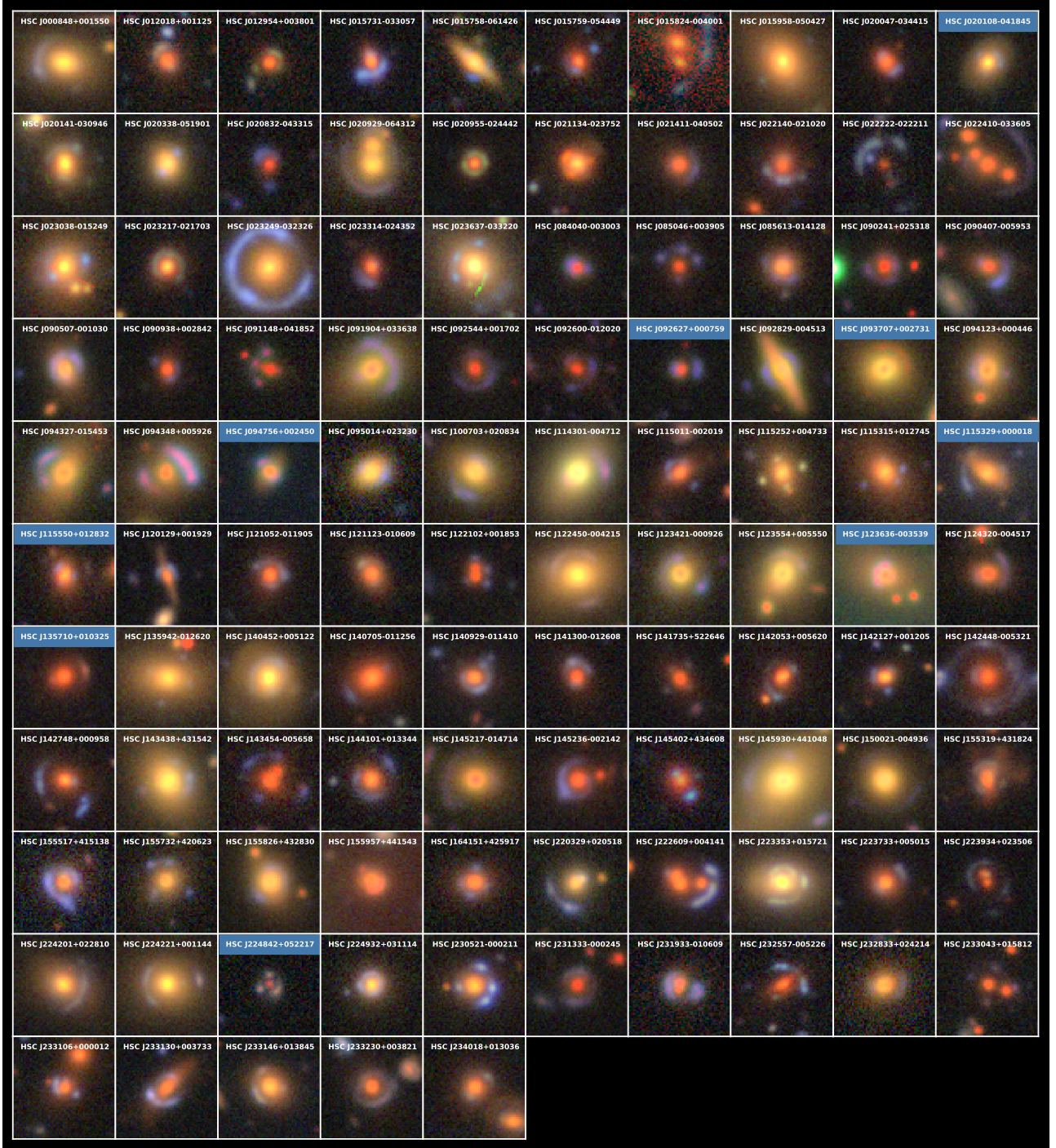
Fig. 5. Colour composite images ($10'' \times 10''$) of the 105 grade-A strong-lens candidates discovered by this work. Candidates with a blue background beneath the system name are new discoveries.

magnitude distributions for the lens galaxies in the 560 grade-A or B candidates are also shown in Fig. 4. Among them, 216 (39%) grade-A or B candidates contain lens galaxies at $z_\mathrm{d}^\mathrm{phot} \geq 0.6$ and 22 (4%) grade-A or B candidates contain lens galaxies at $z_\mathrm{d}^\mathrm{phot} \geq 0.8$.

5.2. Candidates from Classifier-2

Applying Classifier-2 to our parent sample returned 6119 unique objects with $p_\mathrm{resnet} \geq 0.987$, which is also consistent with the expectation of FPR $= 10^{-3}$. Among the 6119 candidates, 804

were also found by Classifier-1, so their visual-inspection scores were directly set to values from the previous round. Author Y.S. inspected the remaining 5315 candidates and removed 4175 candidates that appeared to be non-lenses. The remaining 1140 candidates were inspected by the same five graders independently. 233 candidates with standard deviations above 0.75 and average score above 1.0 were re-graded. Afterwards, the average visual-inspection scores were computed. In total, Classifier-2 discovers 69 grade-A ($\langle S \rangle \geq 2.5$) and 337 grade-B ($1.5 \leq \langle S \rangle < 2.5$) strong-lens candidates. Among the 6119 systems that were inspected, there are 55 grade-A or B SuGOHI

galaxy-galaxy strong-lens candidates, and 51 of them have average scores $\langle S \rangle \geq 1.5$. It confirms once again that the recall of our visual-inspection procedure is $\approx 92\%$.

Compared to Classifier-1, all 6119 candidates and the 406 grade-A or B candidates found by Classifier-2 tend to contain a higher fraction of higher-redshift or fainter lens galaxies (Fig. 4). There are 236 (58%) grade-A or B candidates with lens galaxies at $z_{\rm d}^{\rm phot} \geq 0.6$ and 32 (8%) grade-A or B candidates with lens galaxies at $z_{\rm d}^{\rm phot} \geq 0.8$. This confirms the finding in the previous section that Classifier-2 is more effective in discovering strong-lens systems with high-redshift or faint lens galaxies. The reported photometric redshift for one grade-B strong-lens candidate, HSC J100400+010320, is zero, which is believed to be a catastrophic outlier in the photometric-redshift estimation after checking its image.

5.3. The combined sample

Combining candidates from the two classifiers, we discover in total 105 grade-A and 630 grade-B strong-lens candidates, of which 56 grade-A and 175 grade-B candidates are found by both classifiers. Cross-matching with the known strong lens compilation suggests that 9 grade-A and 268 grade-B candidates are new discoveries. Figure 5 shows the colour composite images of the 105 grade-A candidates, with the new discoveries indicated by a blue background beneath the system name. Colour composite images of all grade-B candidates are shown in Fig. B.1. Lists of all grade-A and grade-B candidates are presented in Table B.1 and Table B.2.

There is considerable diversity in the lens and source populations in the discovered grade-A or B strong-lens candidates. The majority of them consist of a single elliptical lens galaxy surrounded by blue, extended lensing-like features, indicating star-forming source galaxies. Nonetheless, some candidates contain disc lens galaxies; for example, HSC J015758−061426, HSC J092829−004513, and HSC J144228+002105. Some candidates show orange or red lensing-like features from source galaxies with overall old stellar populations and/or noticeable dust attenuation; for example, HSC J021134−023752, HSC J093707+002731, and HSC J155957+441543. Some candidates show multiple lensed background sources as being compact; for example, HSC J115252+004733, HSC J122102+001853, and HSC J224842+052217. In addition, there are also some group-scale strong-lens candidates; for example, HSC J015824−004001, HSC J022410−033605, and HSC J222609+004141.

Nearly half of the discovered grade-A or B strong-lens candidates (331/735) contain lens galaxies with $z_{\rm d}^{\rm phot} \geq 0.6$, of which 4 grade-A and 129 grade-B candidates are new discoveries. 42 candidates contain lens galaxies with $z_{\rm d}^{\rm phot} \geq 0.8$, of which 1 grade-A and 12 grade-B candidates are new discoveries. According to Fig. 4, the candidate lens galaxies cover a broad magnitude range of 1–2 mag at a fixed redshift, indicating a span of 0.4–0.8 dex in lens galaxy mass.

5.4. Auxiliary spectroscopic data

We cross-matched our 735 grade-A or B strong-lens candidates with spectroscopic catalogues from SDSS-I (Abazajian et al. 2009), SDSS-III (Alam et al. 2015), SDSS-IV (Ahumada et al. 2020), the Master Lens Database[1], the SuGOHI project website[2], and a sample of spectroscopically-selected strong-lens candidates from Talbot et al. (2021) using a matching radius of 1″0, and we obtained spectroscopic redshifts for lens galaxies
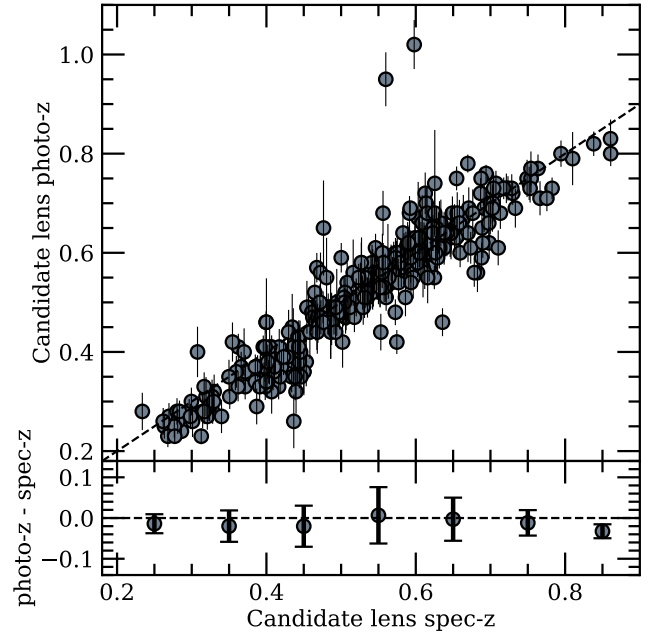

Fig. 6. Comparison between photometric redshifts and spectroscopic redshifts for 333 candidate lens galaxies that have measured spectroscopic redshifts (*Top*). The dashed black line is the one-to-one line. The mean and standard deviation of the differences between photometric redshifts and spectroscopic redshifts in seven redshift bins are shown in the bottom panel.

in 333 candidates and spectroscopic redshifts for source galaxies in 29 candidates. The HSC photometric redshifts for the 333 candidate lens galaxies are in excellent agreement with the corresponding spectroscopic redshifts in general. The differences between the photometric redshifts and spectroscopic redshifts have a mean of −0.008 and standard deviation of 0.06 in the redshift range of 0.23–0.86. Divided into seven redshift bins, the mean differences range from −0.032 to 0.007 (Fig. 6), smaller than the average photometric-redshift uncertainty of 0.036 for these 333 galaxies. Photometric redshifts for two candidate lens galaxies, HSC J000020−002051 (grade-B) and HSC J155957+441543 (grade-A), are significantly higher than the spectroscopic redshifts (by more than 0.3). For HSC J000020−002051, the potential lensing features are ≈3″ away from the candidate lens galaxy, so the HSC photometry should be reasonably accurate. We think its redshift discrepancy is likely due to a catastrophic failure in the photometric-redshift estimation, which is supported by the fact that the photometric redshift for the same galaxy in DESI Legacy Imaging Surveys Data Release 9 is 0.59 ± 0.03 (Dey et al. 2019), in agreement with the spectroscopic redshift of 0.560. For HSC J155957+441543, we think the photometric redshift is biased high due to the contamination from the candidate source galaxy, which is red in colour and is comparably as bright as the candidate lens galaxy in all five HSC filters. The photometric redshift for the same galaxy from the DESI Legacy Imaging Surveys is also over-estimated as 0.72 ± 0.10. Nevertheless, the overall agreement suggests that the photometric-redshift estimation for candidate lens galaxies in our sample is barely affected by the presence of surrounding potential lensing features. This is understandable as our visual inspection process preferentially picks out candidates that exhibit clear separations between the central galaxies and potential lensing features. Moreover, CModel photometry, instead of aperture photometry, is used for the photometric-redshift estimation (Tanaka et al. 2018),
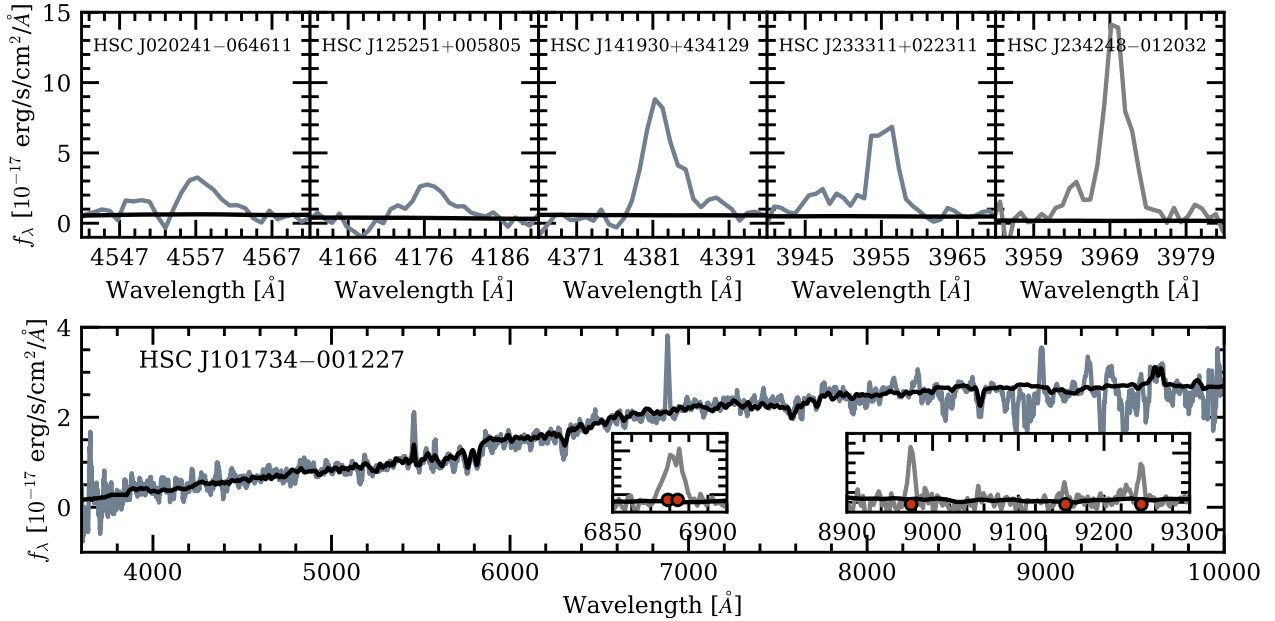
Fig. 7. SDSS spectra of the six strong-lens candidates with evidence of higher-redshift emission lines. In each panel, the grey line represents the observed spectrum and the black line represents the SDSS-provided best-fit model spectrum (only for the foreground lens galaxy). The *top row* shows 30 Å windows centred on the detected emission line for HSC J020241−064611, HSC J125251+005805, HSC J141930+434129, HSC J233311+022311, and HSC J234248−012032, and the *bottom row* shows the full optical spectrum for HSC J101734-001227. Several emission lines not associated with the redshift of the foreground galaxy (i.e. $z = 0.4647$) are shown in the zoomed-in images in the insets of the *bottom panel*. They are found to be coincident with the locations of [OII] doublet, Hβ, [OIII] 4960, and [OIII] 5008 at $z = 0.8457$.

in which substantial deblending from surrounding features is already involved. It also indicates that the photometric redshifts for the remaining candidate lens galaxies are likely reliable.

6. Notes on individual systems

We carried out visual inspections of the publicly available spectra of the 333 candidates identified in the previous section and found six cases where prominent emission lines not consistent with the redshift of the candidate lens galaxies are detected, suggesting superpositions of two objects along the same line of sight. We discuss those cases one by one in this section. We note, however, that this list is by no means complete, and interested readers are encouraged to conduct their own analyses.

HSC J020241−064611. This is a grade-B candidate according to our visual inspection. Two blue, arc-like features are found on the north and south sides of an orange, elliptical galaxy with a separation of ≈1″8 (Fig. B.1). A fibre-fed (2″ in diameter) spectrum from SDSS-III is available, which shows a high S/N emission line at 4557.2 Å on top of a $z = 0.5020$ early-type galaxy spectrum (Fig. 7). This line is obviously not coincident with any typical emission line at $z = 0.5020$. Shu et al. (2016a) interpreted this line as Lyα emission from a Lyα emitter (LAE) at $z = 2.7477$, and considered this system as a galaxy-LAE strong-lens candidate. This system was also classified, based on HSC data, as a grade-B candidate by Sonnenfeld et al. (2018), who resolved the two arc-like features after subtracting the foreground galaxy light. Combining imaging and spectroscopic evidence, we speculate that the two arc-like features are lensed images of a $z = 2.7477$ LAE. The SDSS-measured central velocity dispersion for the foreground galaxy is $156 \pm 25 \, \mathrm{km \, s^{-1}}$[15],

which corresponds to an Einstein radius of ≈0″48 ± 0″15 for a source at $z = 2.7477$ and a lens at $z = 0.5020$ with an isothermal total-mass profile. The estimated Einstein radius is ≈2.8σ lower than what is suggested from the image separation.

HSC J101734-001227. This is a grade-B candidate according to our visual inspection and was also classified as grade-B by C21. A red, elongated arc is located ≈1″6 west of an orange, elliptical galaxy, and there seems to be some hint of a counter image very close to the elliptical galaxy (Fig. B.1). A fibre-fed (2″ in diameter) spectrum from SDSS-III is available. The SDSS best-fit model suggests a redshift of 0.8457, which is primarily driven by several strong emission lines being coincident with [OII] doublet, Hβ, [OIII] 4960, and [OIII] 5008 at $z = 0.8457$. Nevertheless, it is noticed that some emission and absorption features in the spectrum cannot be explained by the best-fit model. Interestingly, we find that the second-best fit using galaxy templates at $z = 0.4647$ provided by SDSS can well reproduce those emission and absorption features (Fig. 7). It hence becomes clear that this particular line of sight contains two galaxies, one at $z = 0.4647$ and the other at $z = 0.8457$. Unfortunately we cannot estimate the Einstein radius because the SDSS-reported velocity dispersion is $850 \, \mathrm{km \, s^{-1}}$, indicating a failure in the measurement. Combining imaging and spectroscopic evidence, we speculate that the potential counter image and/or the elongated arc on the west are responsible for the detected [OII] doublet, Hβ, [OIII] 4960, and [OIII] 5008 at $z = 0.8457$.

---

[5] Starting from Data Release 9, SDSS provides two types of velocity dispersion. One is VDISP determined by fitting the observed spectrum with a linear combination of 24 eigenspectra. The other can be inferred

from VDISP_LNL, which is the velocity-dispersion likelihood function computed by fitting with a linear combination of five eigenspectra while marginalising over redshift uncertainties. As discussed in Shu et al. (2012) and Bolton et al. (2012b), velocity dispersions inferred from VDISP_LNL are more robust for SDSS-III galaxies, the spectra of which often have relatively low S/N. We therefore adopt the velocity dispersion inferred from VDISP_LNL in this work.

HSC J125251+005805. This is a grade-B candidate according to our visual inspection. A blue, elongated arc and a similarly blue blob are found on the northeast and southwest sides of an orange, elliptical galaxy with a separation of $\approx 1\overset{\prime\prime}{.}9$. A fibre-fed ($2''$ in diameter) spectrum from SDSS-III is available, which shows a high S/N emission line at 4176.4 Å on top of a $z = 0.5399$ early-type galaxy spectrum (Fig. 7). This line is obviously not coincident with any typical emission line at $z = 0.5399$. Shu et al. (2016a) interpreted this line as Lyα emission from an LAE at $z = 2.4345$, and considered this system as a galaxy-LAE strong-lens candidate. This system was also classified, based on HSC data, as a grade-B candidate by Wong et al. (2018). The SDSS-measured central velocity dispersion for the foreground galaxy is $203 \pm 40\,\mathrm{km\,s^{-1}}$, which corresponds to an Einstein radius of $\approx 0\overset{\prime\prime}{.}8 \pm 0\overset{\prime\prime}{.}3$ for a source at $z = 2.4345$ and a lens at $z = 0.5399$ with an isothermal total-mass profile. The estimated Einstein radius is in good agreement with the observed image separation. Combining imaging and spectroscopic evidence, we think that the blue arc and blob are indeed lensed images (in a cusp configuration) of a $z = 2.4345$ LAE.

HSC J141930+434129. This is a grade-B candidate according to our visual inspection. A blue, elongated arc is located $\approx 1\overset{\prime\prime}{.}5$ southwest of an orange, elliptical galaxy, but there is no decisive sign for any counter image in the HSC data (Fig. B.1). A fibre-fed ($2''$ in diameter) spectrum from SDSS-IV is available, which shows a high S/N emission line at 4381.3 Å on top of a $z = 0.5447$ early-type galaxy spectrum (Fig. 7). We verified that the detected line is present in the 1D spectra from three individual sub-exposures. This line is obviously not coincident with any typical emission line at $z = 0.5447$. It is also unlikely to be a low-redshift [OII] doublet, because no other strong emission is detected at wavelength positions that would correspond to Hβ, [OIII], and Hα. We hence interpret this line as Lyα emission at $z = 2.6030$. The SDSS-measured central velocity dispersion for the foreground galaxy is $200 \pm 40\,\mathrm{km\,s^{-1}}$, which corresponds to an Einstein radius of $\approx 0\overset{\prime\prime}{.}8 \pm 0\overset{\prime\prime}{.}3$ for a source at $z = 2.6030$ and a lens at $z = 0.5447$ with an isothermal total-mass profile. Combining imaging and spectroscopic evidence, we speculate that the detected Lyα emission is primarily from the blue arc on the southwest (due to scattering). If there is indeed a faint counter image close to the foreground galaxy, which is consistent with the Einstein radius estimation, it would also contribute to the detected Lyα emission.

HSC J233311+022311. This is a grade-B candidate according to our visual inspection. Two tangentially elongated blue blobs are located $\approx 1\overset{\prime\prime}{.}3$ southeast of an orange, elliptical galaxy, and there is no sign for any counter image in the HSC data (Fig. B.1). A fibre-fed ($2''$ in diameter) spectrum from SDSS-III is available, which shows a strong emission line at 3955.5 Å on top of a $z = 0.4716$ early-type galaxy spectrum (Fig. 7). This line is obviously not coincident with any typical emission line at $z = 0.4716$. Shu et al. (2016a) interpreted this line as Lyα emission from an LAE at $z = 2.2529$, and considered this system as a galaxy-LAE strong-lens candidate. This system was also classified, based on HSC data, as a grade-B candidate by Wong et al. (2018). The SDSS-measured central velocity dispersion for the foreground galaxy is $272 \pm 55\,\mathrm{km\,s^{-1}}$, which corresponds to an Einstein radius of $\approx 1\overset{\prime\prime}{.}4 \pm 0\overset{\prime\prime}{.}6$ for a source at $z = 2.2529$ and a lens at $z = 0.4716$ with an isothermal total-mass profile. Combining imaging and spectroscopic evidence, we speculate that the detected Lyα emission is primarily from the two blue blobs on the southeast. If there

is indeed a faint counter image close to the foreground galaxy, which is broadly consistent with the Einstein radius estimation, it would also contribute to the detected Lyα emission.

HSC J234248–012032. This is a grade-B candidate according to our visual inspection. A blue, elongated arc and a similarly blue blob are found on the northwest and southeast sides of an orange, elliptical galaxy with a separation of $\approx 2\overset{\prime\prime}{.}1$. A fibre-fed ($2''$ in diameter) spectrum from SDSS-III is available, which shows a high S/N emission line at 3970.1 Å on top of a $z = 0.5270$ early-type galaxy spectrum (Fig. 7). This line is obviously not coincident with any typical emission line at $z = 0.5270$. Shu et al. (2016a) interpreted this line as Lyα emission from an LAE at $z = 2.2649$, and considered this system as a galaxy-LAE strong-lens candidate. The SDSS-measured central velocity dispersion for the foreground galaxy is $271 \pm 44\,\mathrm{km\,s^{-1}}$, which corresponds to an Einstein radius of $\approx 1\overset{\prime\prime}{.}4 \pm 0\overset{\prime\prime}{.}4$ for a source at $z = 2.2649$ and a lens at $z = 0.5270$ with an isothermal total-mass profile. The estimated Einstein radius is in good agreement with the observed image separation. Combining imaging and spectroscopic evidence, we think that the blue arc and blob are indeed lensed images (in a cusp configuration) of a $z = 2.2649$ LAE.

7. Discussions

As already demonstrated in Sects. 4 and 5, Classifier-2 is more effective than Classifier-1 in the discovery of strong-lens systems with high-redshift or faint lens galaxies, which, essentially, is a result of differences in the training set and pre-processing steps. 60% and 28% of the mock lenses used for Classifier-2 are at redshifts above 0.6 and fainter than $i = 20.5$ mag, respectively, while these two fractions are only 24% and 4% for Classifier-1. In addition, the square-root stretch implemented only in Classifier-2 helps to better reveal lensing features in high-redshift lenses, which, by construction, require higher-redshift sources that appear fainter on average than sources in lower-redshift lenses. Interestingly, we find that including the two pre-processing steps (square-root stretch and normalisation) in Classifier-1 or removing them from Classifier-2 leads to worse performance in terms of AUROC. These findings highlight that the outcome of supervised machine learning techniques depends strongly on the training set and pre-processing procedures need to be chosen in accordance with the training set. We tested training classifiers with *griz*-filter (instead of *gri*) cutouts, but the performance was not as good as the two presented classifiers. More thorough discussions on the impact of the training set will be presented in Cañameras et al. (in prep.) and More et al. (in prep.).

According to the 0.85 TPR for Classifier-1 and $\approx 92\%$ visual-inspection recall, the 560 grade-A or B strong-lens candidates discovered by Classifier-1 suggest that, in our parent sample, there would be 716 strong lenses in total with properties similar to the 92 SuGOHI strong-lens candidates in our test set. Likewise, the 406 grade-A or B strong-lens candidates discovered by Classifier-2 with a TPR of 0.60 suggest a total number of 736 strong lenses. These two predictions agree well with each other, and they are also consistent with the 735 grade-A or B strong-lens candidates discovered by the two classifiers combined. From another perspective, 84 of the 92 SuGOHI candidates in our test set are recovered by the two classifiers combined, suggesting an overall recall of 91%. We therefore expect that $\gtrsim 90\%$ of all strong-lens candidates that are in our parent sample and have properties similar to the 92 SuGOHI strong-lens candidates have already been included in our lists of grade-A or B strong-lens candidates.

Collett ([2015](#)) made a prediction on the population of detectable galaxy-galaxy strong lenses in several imaging surveys. Although HSC was not considered there, we can use results for the LSST, relevant properties of which (including pixel scale, seeing distribution, and sky-brightness distribution) are similar to HSC, as an approximation. In particular, Collett ([2015](#)) forecasted that LSST can detect, over an area of $20\,000\ \mathrm{deg}^2$, 17 000 galaxy-galaxy strong lenses from the best single-epoch imaging and 39 000 galaxy-galaxy strong lenses from the final full stack of the survey. The nominal depths of LSST single-epoch and full-stack imaging are {25.0, 24.7, 24.0} and {27.4, 27.5, 26.8} in {g, r, i} filters (Ivezić et al. [2019](#)), which nicely bracket the depths of HSC PDR2. It hence suggests that the total number of detectable galaxy-galaxy strong lenses in HSC PDR2 is between 800 and 1900. In terms of high-redshift strong lenses, the forecast is that there will be between 180 and 190 $z_\mathrm{d} > 0.8$ strong lenses. We note that the actual number of detectable strong lenses is very sensitive to the adopted S/N threshold. Collett ([2015](#)) considered a lens system to be detectable if the total S/N, $\mathrm{SN_{TOT}}$, of the lensing features is 20 or higher in at least one band (along with three other conditions). If requiring $\mathrm{SN_{TOT}} > 30$, the forecasts for the total number of strong lenses and $z_\mathrm{d} > 0.8$ strong lenses in HSC PDR2 drop to 300–1200 and 80–110. On the other hand, Collett ([2015](#)) pointed out that their LSST forecasts are likely underestimated due to poorly constrained redshift and size distributions of source galaxies used in their simulation, especially on the faint end. The uncertainties were estimated to be at the level of ∼10%. It is unclear what fraction of the detectable strong lenses simulated in Collett ([2015](#)) can pass our selection criteria in Sect. [2](#). Nevertheless, we believe that the vast majority of our grade-A or B strong-lens candidates have $\mathrm{SN_{TOT}}$ substantially higher than 20 according to Figs. [5](#) and [B.1](#), and our single set of 735 grade-A or B strong-lens candidates (including 42 at $z_\mathrm{d}^\mathrm{phot} > 0.8$) represents ≳50% of all detectable strong lenses in HSC PDR2.

Prior to this work, there were two other projects that searched systematically for strong lenses in the HSC data. One of them is the SuGOHI project and the other is a project also done by us, that is C21. The SuGOHI project makes use of several different methods for lens search including automated algorithms (e.g. Sonnenfeld et al. [2018](#); Chan et al. [2020](#)) and crowdsourcing (e.g. Sonnenfeld et al. [2020](#)). C21 makes use of a resnet, similar to this work. Time-wise, the resnets used in C21 and this work are many orders of magnitude faster than the automated algorithms and crowdsourcing used in the SuGOHI project. Classifications of the 5.3 million objects in this work took ≈100 min, or ≈50 000 objects per minute. The classification speed of the methods used in the SuGOHI project is on the order of ∼10 s per object (K. Wong, priv. comm.).

A more fundamental distinction between the three projects is on the parent sample. The parent sample of this work contains galaxies (or more precisely speaking, extended objects) in the Wide layer of HSC PDR2 that satisfy certain magnitude and colour cuts defined in Sect. [2](#) (along with some quality flags). The parent sample in C21 is 62.5 million galaxies in the Wide layer of HSC PDR2 with an i−band Kron radius larger than $0\rlap{.}''8$. The parent samples in the SuGOHI project are more heterogeneous and selected not only from the Wide layer but also the HSC Deep and UltraDeep fields. In particular, the parent samples in Sonnenfeld et al. ([2018](#)), Wong et al. ([2018](#)), and Chan et al. ([2020](#)) are ≈500 000 luminous red galaxies selected for spectroscopic observations in SDSS-III. The parent sample in Sonnenfeld et al. ([2020](#)) is ≈300 000 galaxies with photometric

redshifts between 0.2 and 1.2 and stellar mass above $10^{11.2} M_\odot$. In our parent sample, 3 493 859 (65.2%) objects have i−band Kron radius smaller than $0\rlap{.}''8$ and 4 957 066 (92.5%) objects do not satisfy either of the two requirements in the SuGOHI project. As a result, approximately 3.4 million objects in our parent sample had not been classified by either the SuGOHI project or C21. In terms of high-redshift galaxies, our parent sample is much more complete than those in the other two projects. 80–90% of HSC PDR2 galaxies at redshifts above 0.8 are expected to be included by the colour-colour cuts in this work. In our parent sample, 1 402 958 objects have photometric redshifts above 0.8, of which only 524 078 (37.4%) have i−band Kron radius larger than $0\rlap{.}''8$. It suggests that the parent sample in C21 only included approximately one third of all HSC PDR2 galaxies at redshifts above 0.8. The total size of the parent samples in the SuGOHI project is only ≈800 000, and redshifts for the vast majority are below 0.8.

The SuGOHI project has discovered 497 grade-A or B strong-lens candidates, of which 248 are classified as galaxy-scale systems. For the following comparisons, galaxy-scale candidates from Sonnenfeld et al. ([2020](#)) are also included in this SuGOHI galaxy-scale sample, although some of them are actually cluster- or group-scale systems as pointed out in Sect. [2](#). C21 has discovered 467 grade-A or B strong-lens candidates, almost all of which are galaxy-scale systems. Similarly, almost all of the 735 grade-A or B strong-lens candidates discovered by this work are galaxy-scale systems. There are 132 candidates in common between this work and the SuGOHI project, and 302 candidates in common between this work and C21. Combining these three sample yields 1002 unique galaxy-scale strong-lens candidates, and 395 of the 735 (54%) grade-A or B strong-lens candidates in this work had not been discovered by the other two projects. Candidates in C21 and this work cover similar ranges in lens galaxy photometric redshift and i−band magnitude, while the SuGOHI galaxy-scale sample contains a higher fraction of candidates with lens photometric redshifts above ≈0.9 (Fig. [8](#)). In terms of numbers, 25 candidates in the SuGOHI galaxy-scale sample, 13 candidates in C21, and 11 candidates in our sample have lens photometric redshifts above ≈0.9. Nevertheless, we find that 13 of the 25 $z_\mathrm{d}^\mathrm{phot} > 0.9$ SuGOHI galaxy-scale candidates and 8 of the 13 $z_\mathrm{d}^\mathrm{phot} > 0.9$ C21 candidates do not fulfil our colour selection criteria defined in Sect. [2](#) (criteria 13–14) and are not included in our parent sample in the first place. Those candidates generally have bluer $g - i$ colours as a result of the contamination from the blue lensing features, especially in the g band. However, their photometric-redshift estimations appear not to be significantly affected by this type of contamination (see also Fig. [6](#)), likely because the photometric-redshift estimation is based on multiple colours and is therefore less sensitive to any bias in one particular band.

To further improve completeness with regard to discovering high-redshift strong lenses, a few options can be explored. The first is to improve the completeness in the parent sample. Although the colour-colour cuts used in this work are found to be already 80–90% complete in selecting high-redshift strong lenses, some known high-redshift strong-lens candidates are excluded due to contaminated photometry. On the other hand, the provided photometric redshifts do not seem to be strongly biased by lensing features in general. Combining the colour-colour criteria and a photometric-redshift selection should in principle result in a more complete parent sample. Moreover, the classifier may be further optimised. In this work, the classifiers were tuned to deliver high overall TPRs for strong-lens candidates covering a
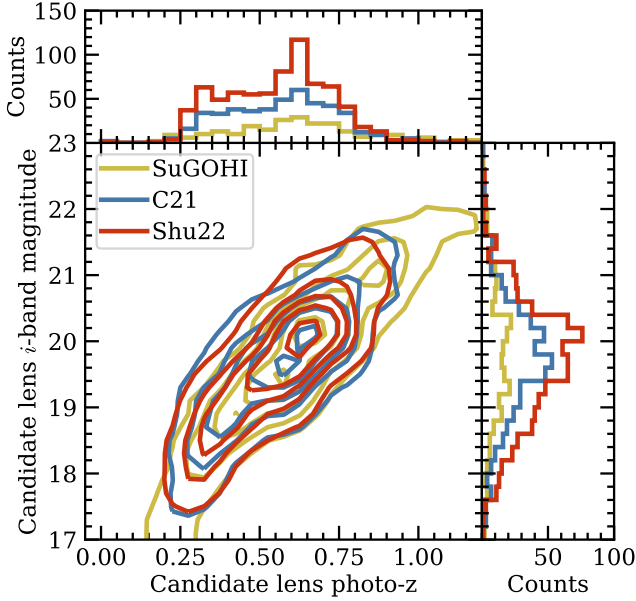
Fig. 8. Distributions of photometric redshift and i−band magnitude of the lens galaxies in galaxy-scale strong-lens candidates from the SuG-OHI project (yellow), C21 (blue), and this work (red). The contours correspond to 10th, 30th, 50th, 70th, and 90th percentiles of the individual samples. To make a fair comparison, we use `photoz_best` from the `pdr2_wide.photoz_mizuki` table for SuGOHI lens galaxies instead of the photometric redshifts provided by the SuGOHI project website[2].

wide redshift range from 0.2 to 1.1, and it has been shown that the TPRs can vary substantially in different redshift sub-ranges. One can consider optimising the classifier based on the performance on the redshift range of interest only.

8. Conclusions

In this work, we carried out a search for strong-lens systems consisting of high-redshift lens galaxies in the Wide layer data from HSC-SSP PDR2 with a sky coverage of ≈960 deg². We first applied several colour and magnitude cuts to reduce the sample size in HSC PDR2 from ≈80 million galaxies to ≈5.4 million galaxies. To further efficiently classify those galaxies, that is our parent sample, we constructed two strong-lens classifiers based on a deep residual network pre-built in the `CMU DeepLens` package. The two classifiers, Classifier-1 and Classifier-2, differ mainly in the training set and pre-processing procedures. After training, the two classifiers achieved AUROC values of 0.993 and 0.985 on a test dataset comprising real strong lenses and non-lenses. Applying each of the two classifiers to the gri-filter cutouts (60 pixel × 60 pixel, 1 pixel = $0.''17$) of the parent sample returned network scores p_{resnet} for individual galaxies in ≈100 min. Adopting p_{resnet} thresholds that correspond to an FPR of 10^{-3} based on the test set, Classifier-1 and Classifier-2 produced 5468 and 6119 unique strong-lens candidates, respectively. Five authors independently graded those strong-lens candidates based on visual inspections of the cutouts. According to the average visual-inspection scores, 560 candidates identified by Classifier-1 and 406 candidates identified by Classifier-2 are considered as grade-A or B (i.e. definite or probable) strong-lens candidates.

By combining the two samples, we discover in total 105 grade-A and 630 grade-B strong-lens candidates, which is the single largest set of galaxy-scale strong-lens candidates

discovered with HSC data to date. Among them, nine grade-A and 268 grade-B candidates are new discoveries. This list of 735 candidates is expected to include ≳90% of all strong-lens candidates that are in our parent sample and have properties similar to the test set. The candidate lens galaxies span a (photometric) redshift range from 0.2 to 1.0. Nearly half of the discovered candidates (331/735) contain lens galaxies with $z_{\mathrm{d}}^{\mathrm{phot}} \geq 0.6$, and 42 candidates contain lens galaxies with $z_{\mathrm{d}}^{\mathrm{phot}} \geq 0.8$. Despite having a lower overall TPR, Classifier-2 discovers a significantly higher fraction of high-redshift ($z_{\mathrm{d}}^{\mathrm{phot}} \geq 0.6$) lens galaxies compared to Classifier-1, which we attribute to differences in the training set and pre-processing procedures.

We obtained spectroscopic redshifts for lens galaxies in 333 candidates and spectroscopic redshifts for source galaxies in 29 candidates by cross-matching our candidates with spectroscopic catalogues in the literature. We found an excellent agreement between the HSC-reported photometric redshifts and the corresponding spectroscopic redshifts for the 333 candidate lens galaxies, indicating that the photometric redshifts for the remaining candidate lens galaxies are likely reliable. In addition, we noticed high S/N emission lines in publicly-available spectra of six candidates that are presumably from redshifts higher than those of the foreground galaxies. It is worth carrying out follow-up observations to determine the nature of the detected emission lines and lensing status of the six systems.

We will continue applying our classifiers to future HSC data releases to discover more strong-lens systems. Meanwhile, we will obtain follow-up spectroscopy to confirm the best-quality high-redshift strong-lens candidates from this search and turn them into a powerful probe for galaxy evolution at $z \gtrsim 0.8$. Our discoveries will also serve as a valuable target list for ongoing and scheduled spectroscopic surveys such as the Dark Energy Spectroscopic Instrument (DESI Collaboration 2016), the Subaru Prime Focus Spectrograph project (Takada et al. 2014), and the Maunakea Spectroscopic Explorer (The MSE Science Team et al. 2019). As demonstrated by this work, resnet-based algorithms are a promising approach for efficiently and effectively uncovering the ~10^5 strong-lens systems expected in forthcoming wide-field imaging surveys such as LSST, *Euclid*, and CSS-OS. All kinds of scientific applications enabled by strong lensing are expected to benefit from a larger and more complete population of strong-lens systems.

References

Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2009, ApJS, 182, 543

Abbott, T. M. C., Adamów, M., Aguena, M., et al. 2021, ApJS, 255, 20

Abolfathi, B., Aguado, D. S., Aguilar, G., et al. 2018, ApJS, 235, 42

Ahumada, R., Prieto, C. A., Almeida, A., et al. 2020, ApJS, 249, 3

Aihara, H., AlSayyad, Y., Ando, M., et al. 2019, PASJ, 71, 114

Alam, S., Albareti, F. D., Allende Prieto, C., et al. 2015, ApJS, 219, 12

Auger, M. W., Treu, T., Gavazzi, R., et al. 2010, ApJ, 721, L163

Bag, S., Shafieloo, A., Liao, K., & Treu, T. 2022, ApJ, 927, 191

Bayer, J., Huber, S., Vogl, C., et al. 2021, A&A, 653, A29

Beckwith, S. V. W. 2005, VizieR Online Data Catalog: II/258

Bolton, A. S., Burles, S., Schlegel, D. J., Eisenstein, D. J., & Brinkmann, J. 2004, AJ, 127, 1860

Bolton, A. S., Burles, S., Koopmans, L. V. E., et al. 2008, ApJ, 682, 964

Bolton, A. S., Brownstein, J. R., Kochanek, C. S., et al. 2012a, ApJ, 757, 82

Bolton, A. S., Schlegel, D. J., Aubourg, É., et al. 2012b, AJ, 144, 144

Bonvin, V., Courbin, F., Suyu, S. H., et al. 2017, MNRAS, 465, 4914

Bosch, J., Armstrong, R., Bickerton, S., et al. 2018, PASJ, 70, S5

Brewer, B. J., Marshall, P. J., Auger, M. W., et al. 2014, MNRAS, 437, 1950

Browne, I. W. A., Wilkinson, P. N., Jackson, N. J. F., et al. 2003, MNRAS, 341, 13

Brownstein, J. R., Bolton, A. S., Schlegel, D. J., et al. 2012, ApJ, 744, 41

Bussmann, R. S., Pérez-Fournon, I., Amber, S., et al. 2013, ApJ, 779, 25

Cañameras, R., Nesvadba, N. P. H., Kneissl, R., et al. 2017, A&A, 600, L3

Cañameras, R., Schuldt, S., Suyu, S. H., et al. 2020, A&A, 644, A163

Cañameras, R., Schuldt, S., Shu, Y., et al. 2021, A&A, 653, L6

Chan, J. H. H., Suyu, S. H., Sonnenfeld, A., et al. 2020, A&A, 636, A87

Chao, D. C. Y., Chan, J. H. H., Suyu, S. H., et al. 2020, A&A, 640, A88

Chao, D. C. Y., Chan, J. H. H., Suyu, S. H., et al. 2021, A&A, 655, A114

Christensen, L., Richard, J., Hjorth, J., et al. 2012, MNRAS, 427, 1953

Collett, T. E. 2015, ApJ, 811, 20

Courbin, F., Faure, C., Djorgovski, S. G., et al. 2012, A&A, 540, A36

Craig, P., O'Connor, K., Chakrabarti, S., et al. 2021, MNRAS, submitted [arXiv:2111.01680]

Dawson, K. S., Schlegel, D. J., Ahn, C. P., et al. 2013, AJ, 145, 10

DESI Collaboration (Aghamousa, A., et al.) 2016, ArXiv e-prints, [arXiv:1611.00036]

Desira, C., Shu, Y., Auger, M. W., et al. 2022, MNRAS, 509, 738

Dey, A., Schlegel, D. J., Lang, D., et al. 2019, AJ, 157, 168

Diehl, H. T., Buckley-Geer, E. J., Lindgren, K. A., et al. 2017, ApJS, 232, 15

Ding, X., Liao, K., Birrer, S., et al. 2021, MNRAS, 504, 5621

Fadely, R., & Keeton, C. R. 2012, MNRAS, 419, 936

Faure, C., Anguita, T., Alloin, D., et al. 2011, A&A, 529, A72

Goldstein, D. A., & Nugent, P. E. 2017, ApJ, 834, L5

Grillo, C., Rosati, P., Suyu, S. H., et al. 2018, ApJ, 860, 94

He, K., Zhang, X., Ren, S., & Sun, J. 2016, ArXiv e-prints [arXiv:1603.05027]

Hezaveh, Y. D., Dalal, N., Marrone, D. P., et al. 2016, ApJ, 823, 37

Huang, X., Storfer, C., Ravi, V., et al. 2020, ApJ, 894, 78

Huang, X., Storfer, C., Gu, A., et al. 2021, ApJ, 909, 27

Huber, S., Suyu, S. H., Noebauer, U. M., et al. 2021, A&A, 646, A110

Huber, S., Suyu, S. H., Ghoshdastidar, D., et al. 2022, A&A, 658, A157

Inami, H., Bacon, R., Brinchmann, J., et al. 2017, A&A, 608, A2

Inoue, K. T., Minezaki, T., Matsushita, S., & Chiba, M. 2016, MNRAS, 457, 2936

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

Jacobs, C., Glazebrook, K., Collett, T., More, A., & McCarthy, C. 2017, MNRAS, 471, 167

Jacobs, C., Collett, T., Glazebrook, K., et al. 2019a, ApJS, 243, 17

Jacobs, C., Collett, T., Glazebrook, K., et al. 2019b, MNRAS, 484, 5330

Jaelani, A. T., More, A., Oguri, M., et al. 2020, MNRAS, 495, 1291

Koopmans, L. V. E., Treu, T., Bolton, A. S., Burles, S., & Moustakas, L. A. 2006, ApJ, 649, 599

Kostrzewa-Rutkowska, Z., Kozłowski, S., Lemon, C., et al. 2018, MNRAS, 476, 663

Lanusse, F., Ma, Q., Li, N., et al. 2018, MNRAS, 473, 3895

Laureijs, R., Amiaux, J., Arduini, S., et al. 2011, ArXiv e-prints, [arXiv:1110.3193]

Lemon, C. A., Auger, M. W., McMahon, R. G., & Ostrovski, F. 2018, MNRAS, 479, 5060

Li, R., Napolitano, N. R., Tortora, C., et al. 2020, ApJ, 899, 30

Li, R., Napolitano, N. R., Spiniello, C., et al. 2021, ApJ, 923, 16

Marques-Chaves, R., Pérez-Fournon, I., Shu, Y., et al. 2017, ApJ, 834, L18

Marques-Chaves, R., Pérez-Fournon, I., Gavazzi, R., et al. 2018, ApJ, 854, 151

Marques-Chaves, R., Pérez-Fournon, I., Shu, Y., et al. 2020, MNRAS, 492, 1257

Metcalf, R. B., Meneghetti, M., Avestruz, C., et al. 2019, A&A, 625, A119

Millon, M., Galan, A., Courbin, F., et al. 2020, A&A, 639, A101

More, A., Cabanac, R., More, S., et al. 2012, ApJ, 749, 38

More, A., Verma, A., Marshall, P. J., et al. 2016, MNRAS, 455, 1191

More, A., Lee, C.-H., Oguri, M., et al. 2017, MNRAS, 465, 2411

Nierenberg, A. M., Treu, T., Wright, S. A., Fassnacht, C. D., & Auger, M. W. 2014, MNRAS, 442, 2434

Oguri, M., & Kawano, Y. 2003, MNRAS, 338, L25

Oguri, M., & Marshall, P. J. 2010, MNRAS, 405, 2579

Oldham, L., Auger, M. W., Fassnacht, C. D., et al. 2017, MNRAS, 465, 3185

Petrillo, C. E., Tortora, C., Vernardos, G., et al. 2019, MNRAS, 484, 3879

Planck Collaboration VI. 2020, A&A, 641, A6

Prakash, A., Licquia, T. C., Newman, J. A., et al. 2016, ApJS, 224, 34

Ratnatunga, K. U., Ostrander, E. J., Griffiths, R. E., & Im, M. 1995, ApJ, 453, L5

Rojas, K., Savary, E., Clément, B., et al. 2021, A&A, submited [arXiv:2109.00014]

Ryczanowski, D., Smith, G. P., Bianconi, M., et al. 2020, MNRAS, 495, 1666

Savary, E., Rojas, K., Maus, M., et al. 2021, A&A, submited [arXiv:2110.11972]

Schuldt, S., Suyu, S. H., Cañameras, R., et al. 2021a, A&A, 651, A55

Schuldt, S., Suyu, S. H., Meinhardt, T., et al. 2021b, A&A, 646, A126

Shu, Y., Bolton, A. S., Schlegel, D. J., et al. 2012, AJ, 143, 90

Shu, Y., Bolton, A. S., Brownstein, J. R., et al. 2015, ApJ, 803, 71

Shu, Y., Bolton, A. S., Kochanek, C. S., et al. 2016a, ApJ, 824, 86

Shu, Y., Bolton, A. S., Mao, S., et al. 2016b, ApJ, 833, 264

Shu, Y., Bolton, A. S., Moustakas, L. A., et al. 2016c, ApJ, 820, 43

Shu, Y., Brownstein, J. R., Bolton, A. S., et al. 2017, ApJ, 851, 48

Shu, Y., Bolton, A. S., Mao, S., et al. 2018a, ApJ, 864, 91

Shu, Y., Marques-Chaves, R., Evans, N. W., & Pérez-Fournon, I. 2018b, MNRAS, 481, L136

Shu, Y., Koposov, S. E, Evans, N. W., et al. 2019, MNRAS, 489, 4741

Shu, Y., Belokurov, V., & Evans, N. W. 2021, MNRAS, 502, 2912

Shu, X., Yang, L., Liu, D., et al. 2022, ApJ, 926, 155

Sonnenfeld, A., Gavazzi, R., Suyu, S. H., Treu, T., & Marshall, P. J. 2013, ApJ, 777, 97

Sonnenfeld, A., Chan, J. H. H., Shu, Y., et al. 2018, PASJ, 70, S29

Sonnenfeld, A., Verma, A., More, A., et al. 2020, A&A, 642, A148

Stark, D. P., Auger, M., Belokurov, V., et al. 2013, MNRAS, 436, 1040

Stark, D. P., Walth, G., Charlot, S., et al. 2015, MNRAS, 454, 1393

Stein, G., Blaum, J., Harrington, P., Medan, T., & Lukic, Z. 2021, ApJ, submitted [arXiv:2110.00023]

Suyu, S. H., & Halkola, A. 2010, A&A, 524, A94

Suyu, S. H., Marshall, P. J., Auger, M. W., et al. 2010, ApJ, 711, 201

Suyu, S. H., Hensel, S. W., McKean, J. P., et al. 2012, ApJ, 750, 10

Suyu, S. H., Auger, M. W., Hilbert, S., et al. 2013, ApJ, 766, 70

Suyu, S. H., Bonvin, V., Courbin, F., et al. 2017, MNRAS, 468, 2590

Suyu, S. H., Huber, S., Cañameras, R., et al. 2020, A&A, 644, A162

Takada, M., Ellis, R. S., Chiba, M., et al. 2014, PASJ, 66, R1

Talbot, M. S., Brownstein, J. R., Dawson, K. S., Kneib, J.-P., & Bautista, J. 2021, MNRAS, 502, 4617

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

The MSE Science Team, Babusiaux, C., Bergemann, M., et al. 2019, ArXiv e-prints, [arXiv:1904.04907]

Treu, T. 2010, ARA&A, 48, 87

Treu, T., Koopmans, L. V., Bolton, A. S., Burles, S., & Moustakas, L. A. 2006, ApJ, 640, 662

Treu, T., Dutton, A. A., Auger, M. W., et al. 2011, MNRAS, 417, 1601

Vegetti, S., Koopmans, L. V. E., Bolton, A., Treu, T., & Gavazzi, R. 2010, MNRAS, 408, 1969

Vegetti, S., Lagattuta, D. J., McKean, J. P., et al. 2012, Nature, 481, 341

Wojtak, R., Hjorth, J., & Gall, C. 2019, MNRAS, 487, 3342

Wong, K. C., Keeton, C. R., Williams, K. A., Momcheva, I. G., & Zabludoff, A. I. 2011, ApJ, 726, 84

Wong, K. C., Tran, K.-V. H., Suyu, S. H., et al. 2014, ApJ, 789, L31

Wong, K. C., Sonnenfeld, A., Chan, J. H. H., et al. 2018, ApJ, 867, 107

Wong, K. C., Suyu, S. H., Chen, G. C. F., et al. 2020, MNRAS, 498, 1420

Zhan, H. 2018, in 42nd COSPAR Scientific Assembly, 42, E1.16-4-18

Appendix A: Visual-inspection score comparisons with C21

Among all the network candidates from C21 and this work, 956 systems are in common and have been assigned visual-inspection scores twice by the same five graders. In this appendix, we discuss the variations in the visual-inspection scores for the same systems from round to round, which provides an idea on the robustness of our visual-inspection scores. We note that the visual-inspection processes between this work and C21 are slightly different. In C21, three images with different stretching and normalisation schemes for the same systems were provided to the graders, while four more images with different stretching and normalisation schemes for the same systems were provided in this work.

Inevitably, scores from each grader are not all identical. The biases for individual graders range from -0.18 to 0.12, and the typical dispersion is ~ 0.7 (Figure A.1). Encouragingly, the average score, which determines the final lens grade, has almost no bias (-0.01). Hence, for systems that have different average scores between this work and C21, our recommendation is to adopt the higher values so that a more complete list of candidates can be obtained.

Appendix B: Full lists of grade-A or B lenses
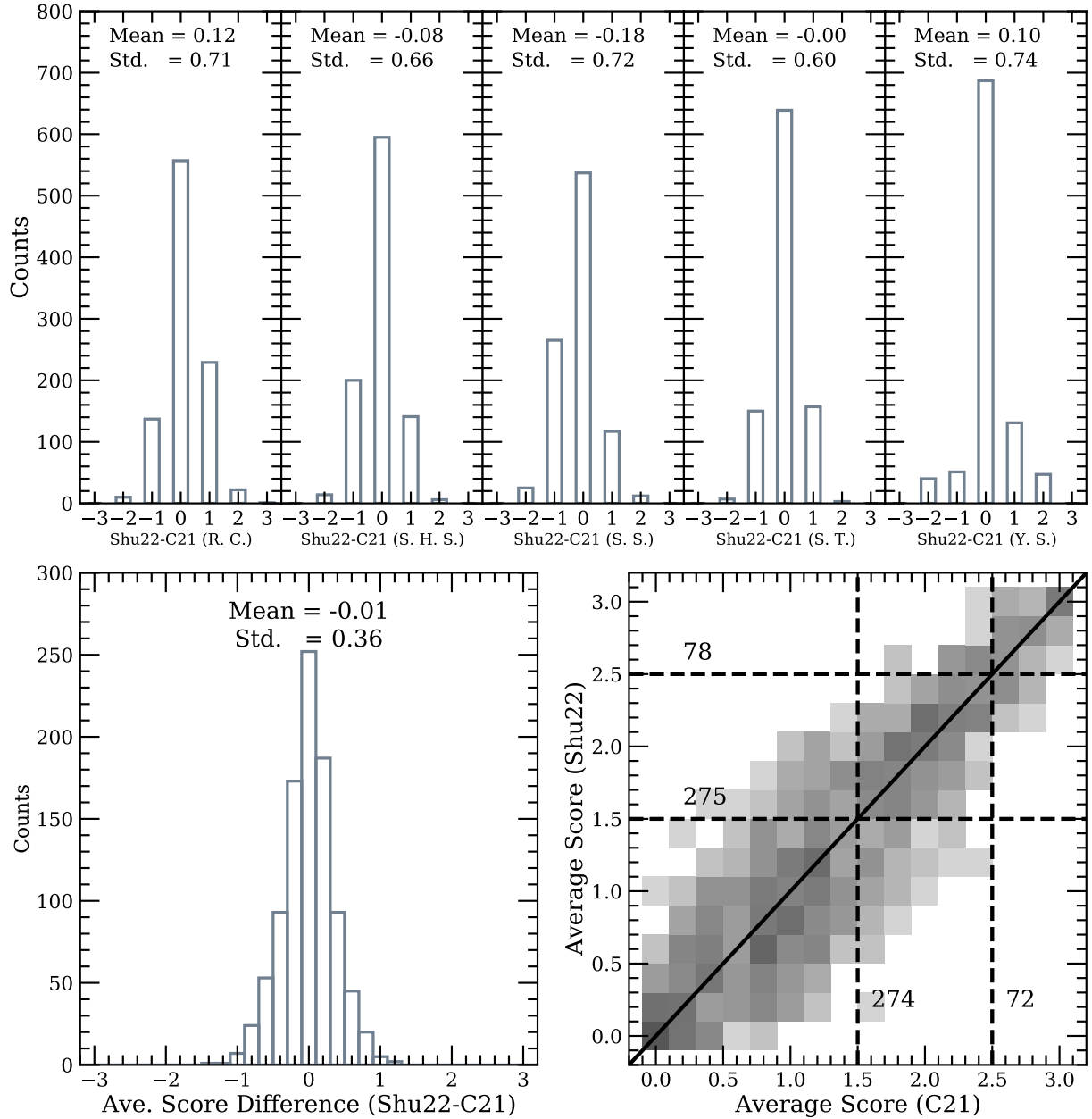
---

[6] http://sunguoyou.lamost.org/glc.html.

Fig. A.1: Comparisons on visual-inspection scores for the 956 systems that are in common between this work (i.e. Shu22) and C21. The top row shows the distributions of the difference in scores for the five graders (R. C., S. H. S., S. S., S. T., and Y. S.). The mean and standard deviations of the differences for individual graders are given in each sub-panel. The bottom left panel is the distribution of the difference in the average score, which has a mean of -0.01 and standard deviation of 0.36. The bottom right panel shows the 2D histogram of the average scores in C21 and Shu22. The solid black line is the one-to-one line, and the dashed black lines indicate thresholds that correspond to grade-A or B. According to average scores in C21, 72 are grade As and 274 are grade Bs. According to average scores in this work, 78 are grade As and 275 are grade Bs.

Table B.1: List of discovered grade-A strong-lens candidates.

Name	R.A.	Decl.	m_i	$z_\mathrm{d}^\mathrm{phot}$	$z_\mathrm{d}^\mathrm{spec}$	$z_\mathrm{source}^\mathrm{spec}$	Classifier	Score	Reference
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
HSC J000848+001550	2.20333	0.26412	18.46	0.35 ± 0.03	0.397	—	1	2.6	SuGOHI-5
HSC J012018+001125	20.07557	0.19048	20.14	0.63 ± 0.04	0.599	—	1,2	2.6	C21
HSC J012954+003801	22.47583	0.63363	20.27	0.67 ± 0.05	—	—	1	2.8	C21
HSC J015731−033057	29.38125	−3.51603	20.02	0.68 ± 0.04	0.621	—	1,2	2.8	SuGOHI-1
HSC J015758−061426	29.49429	−6.24057	18.96	0.35 ± 0.03	—	—	1,2	2.6	C21

NOTE. — Column 1 is the system name. Columns 2 and 3 are right ascension and declination (J2000) of the lens galaxy. Columns 4 and 5 are the i−band CModel magnitude and photometric redshift of the lens galaxy provided by the HSC catalogue. Columns 6 and 7 give the spectroscopic redshifts of the lens and source inferred from auxiliary data. Column 8 indicates the classifier(s) that finds the lens system. Column 9 is the average visual-inspection score of the lens system. Column 10 provides the paper that first discovered the system. Shu22 indicates a completely new discovery. Other relevant references are: Brownstein12: Brownstein et al. (2012); C21: Cañameras et al. (2021); Diehl17: Diehl et al. (2017); Huang20: Huang et al. (2020); Huang21: Huang et al. (2021); Jacobs17: Jacobs et al. (2017); Jacobs19b: Jacobs et al. (2019a); Li20: Li et al. (2020); More12: More et al. (2012); More16: More et al. (2016); More17: More et al. (2017); Petrillo19: Petrillo et al. (2019); Ratnatunga95: Ratnatunga et al. (1995); Shu16: Shu et al. (2016a); Sonnenfeld13: Sonnenfeld et al. (2013); Stark13: Stark et al. (2013); Stein21: Stein et al. (2021); SuGOHI-1: Sonnenfeld et al. (2018); SuGOHI-2: Wong et al. (2018); SuGOHI-4: Chan et al. (2020); SuGOHI-5: Jaelani et al. (2020); SuGOHI-6: Sonnenfeld et al. (2020). 'Guoyou Sun' corresponds to candidates identified by an amateur astronomer, Guoyou Sun, through visual inspections of HSC cutouts[6]. Systems with † and/or ‡ are independently discovered by Wong et al. (in prep.) and/or Jaelani et al. (in prep.). The full table is available at the CDS.

Table B.2: List of discovered grade-B strong-lens candidates.

Name	R.A.	Decl.	m_i	$z_\mathrm{d}^\mathrm{phot}$	$z_\mathrm{d}^\mathrm{spec}$	$z_\mathrm{source}^\mathrm{spec}$	Classifier	Score	Reference
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
HSC J000018+001617	0.07884	0.27158	19.91	0.63 ± 0.04	—	—	2	2.0	Shu22
HSC J000020−002051	0.08681	−0.34750	20.56	0.95 ± 0.05	0.560	—	1	2.2	Shu22
HSC J000106+010329†	0.27710	1.05827	20.07	0.73 ± 0.03	0.721	—	1,2	2.0	Shu22
HSC J000114+001619	0.31063	0.27214	19.41	0.67 ± 0.05	0.664	—	1	1.6	Shu22
HSC J000327+021020	0.86261	2.17248	20.18	0.72 ± 0.03	—	—	1	1.8	C21

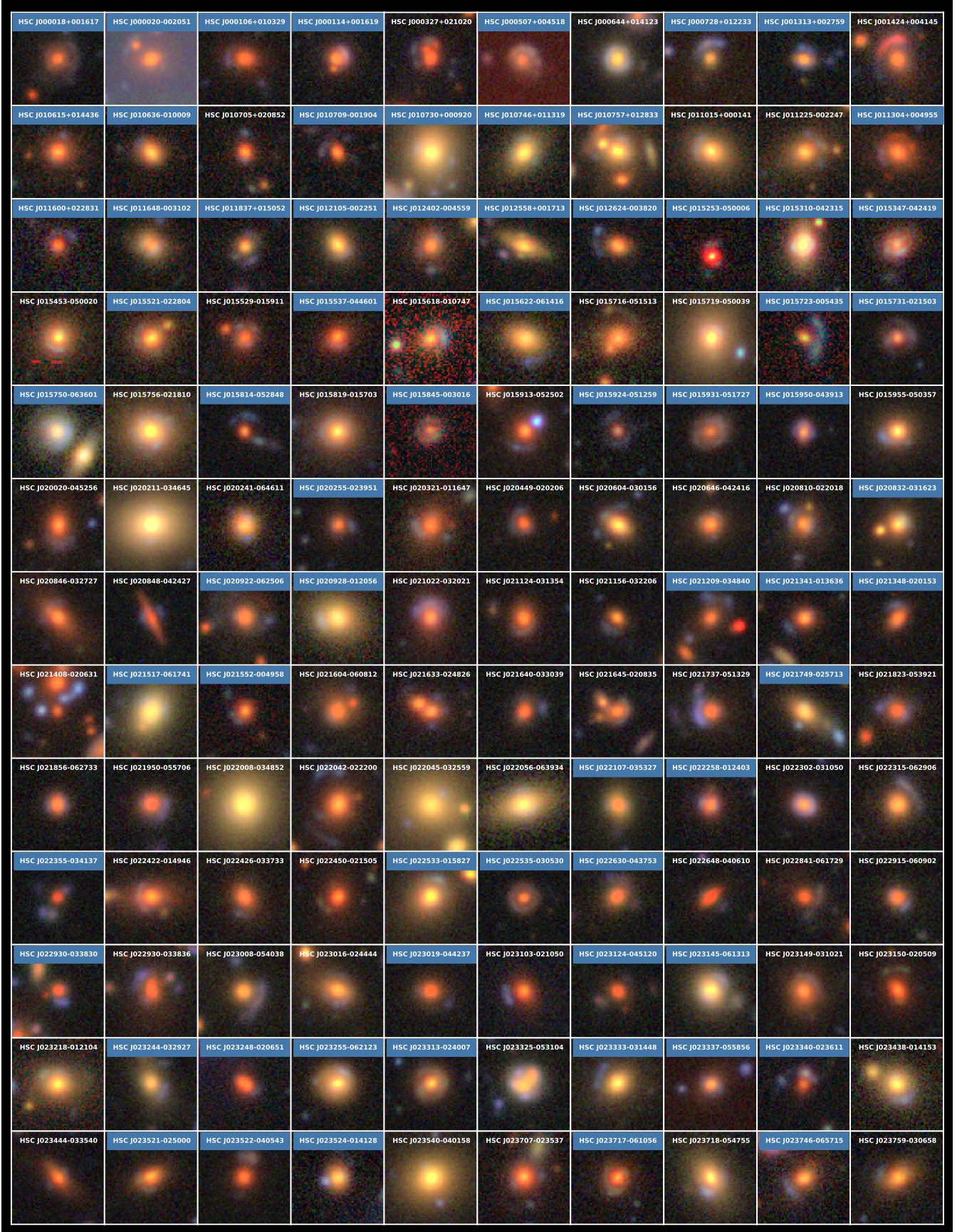NOTE. — The columns are the same as in Table B.1. The full table is available at the CDS.

Fig. B.1: Colour composite images ($10'' \times 10''$) of the 630 grade-B strong-lens candidates discovered by this work. Candidates with blue background beneath the system name are new discoveries.
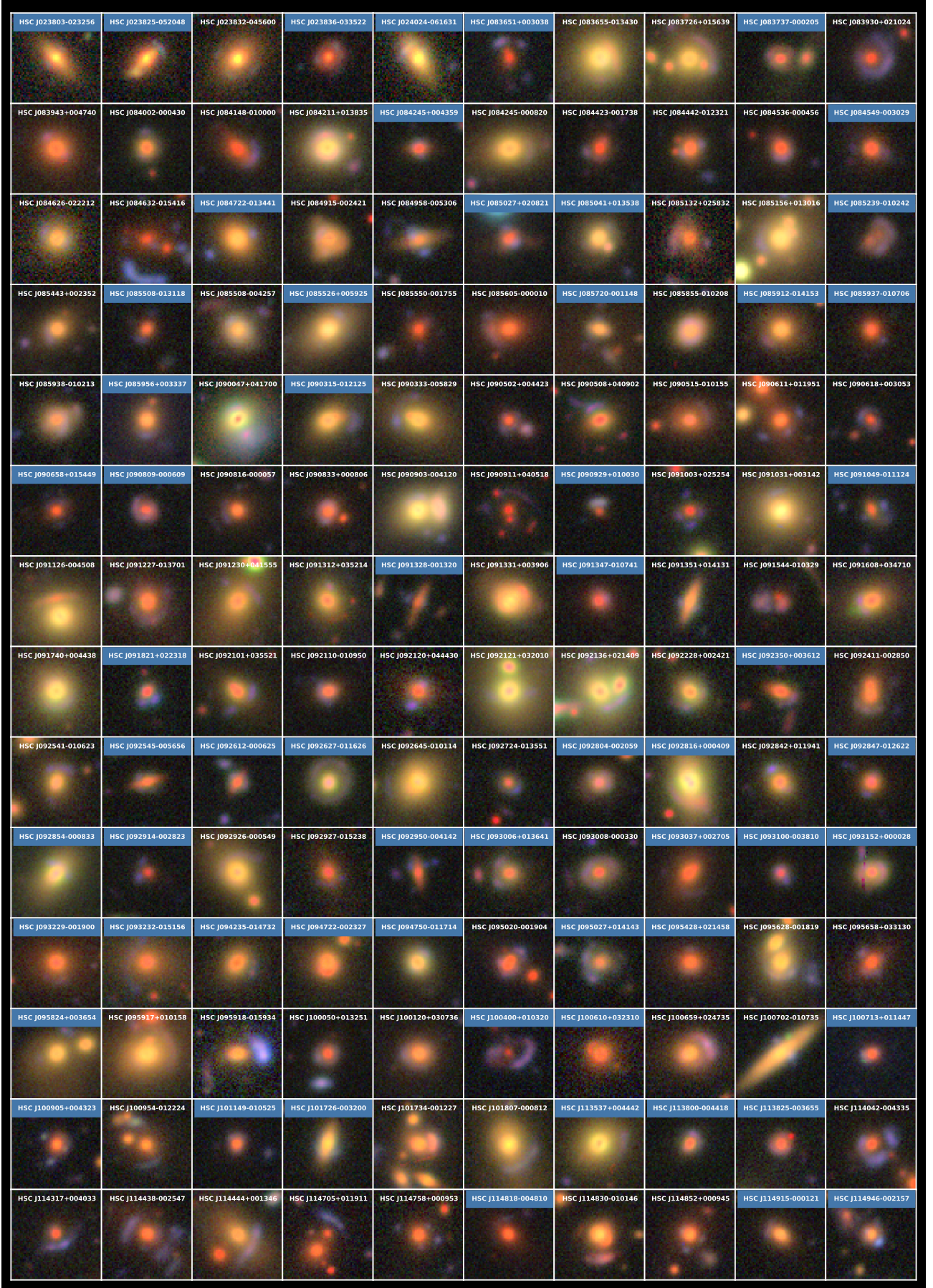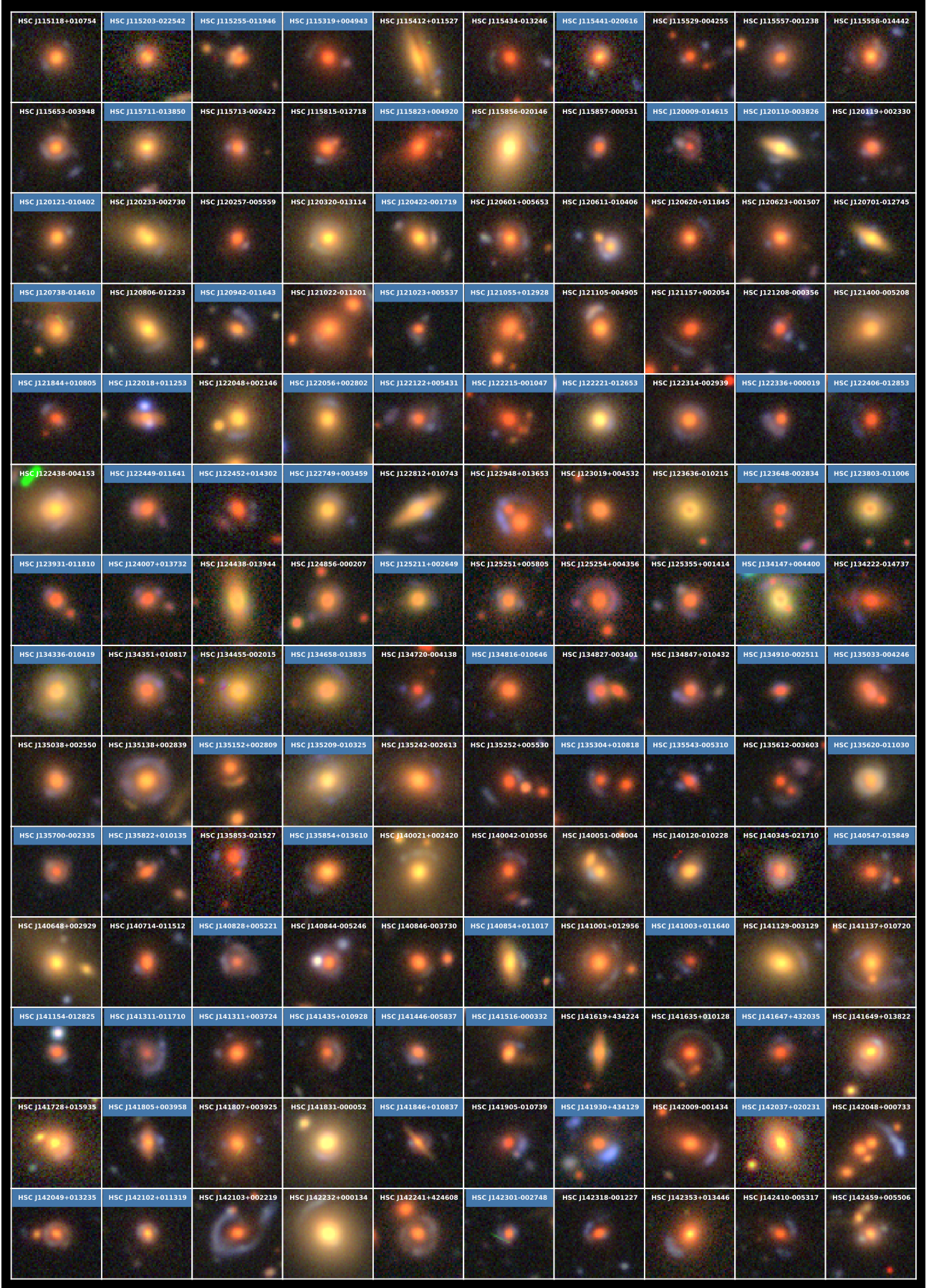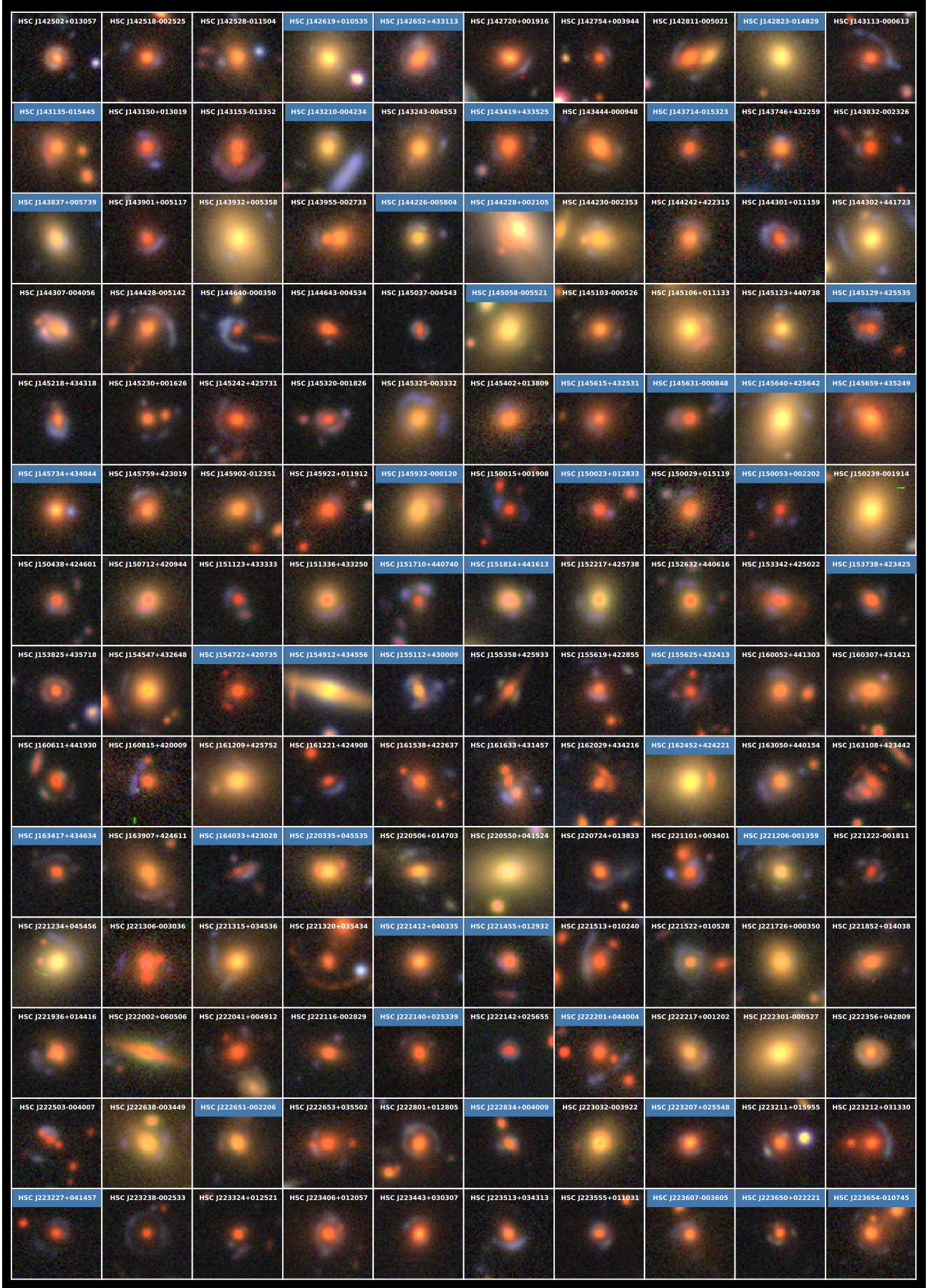
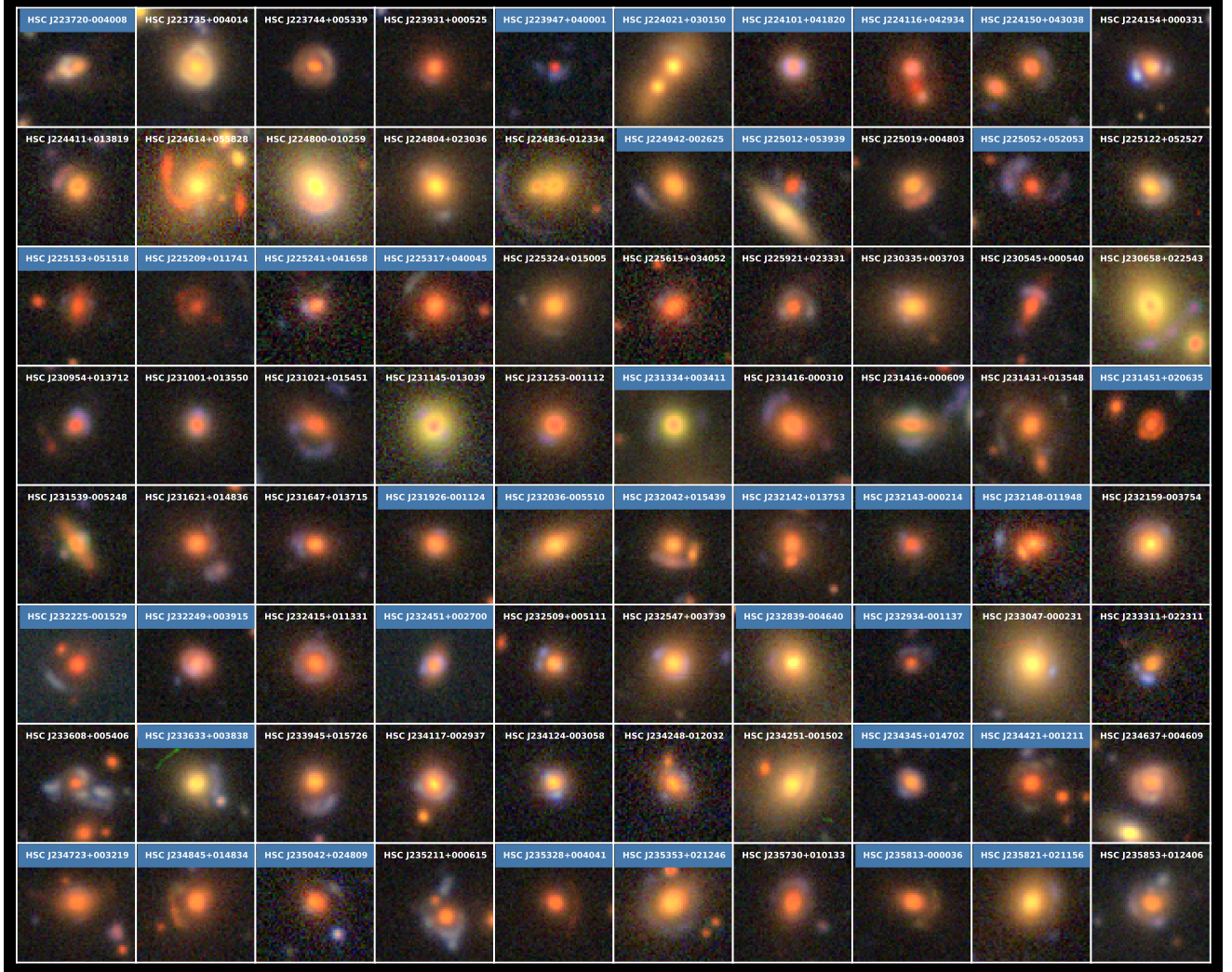Fig. B.1: continued.

Fig. B.1: continued.

Fig. B.1: continued.

Fig. B.1: continued.