Automated Detection of Merging Galaxies at $z = 0.25 - 1.0$ in the CLAUDS+HSC Survey Using Random Forests

by

Nathalie C. M. Thibert

A Thesis Submitted to Saint Mary's University, Halifax, Nova Scotia in Partial Fulfillment
of the Requirements for the Degree of MSc in Astronomy
(Department of Astronomy and Physics)

August 2018, Halifax, Nova Scotia

Approved: Dr. Marcin Sawicki

Supervisor

Approved: Dr. Ivana Damjanov

Examiner

Approved: Dr. Ian Short

Examiner

Date: August 23, 2018.

Acknowledgements

First of all, I would like to recognize my supervisor Dr. Marcin Sawicki. I have learned a great deal from him over the last two years about what it means to be a good scientist and colleague as part of a large international collaboration. I would like to thank my examining committee for their helpful comments and suggestions for improvement. I acknowledge NSERC and Saint Mary's University for their financial support during my degree.

To Anneya, Liz, Lejay, Gurpreet, and Thibaud for their scientific and emotional support throughout the last two years. I am forever indebted to both Anneya Golob and Dr. Andy Goulding for setting the time aside in their busy schedules to guide me through all aspects of this project. I would like to thank all of the students and faculty in the Astronomy Department for their unwavering support, especially Tiffany Fields who listened to me every step of the way, and Drs. Rob Thacker and Luigi Gallo for supporting me in their roles of graduate coordinator in times of need. I am immensely grateful of Florence Woolaver who has been a good friend during my time in the Department; I will miss our chats.

I thank my family, Stephanie, Ilona, and Rob Thibert for believing in me. Also to Ian Mulholland, Laura Lenkić, and Sarah Gallagher who guided me from afar. I express my love and gratitude to my partner Ethan Avila, who has helped me through both the highest and the lowest points of my degree. Without his support, I would not be where I am today. Finally, I would like to thank Mike Hirschbach and all of the instructors and students at Halifax Circus. They have been my family over the past two years and I will miss them as I start a new chapter in my life.

Contents

List of Figures

List of Tables

Abstract

Automated Detection of Merging Galaxies at $z = 0.25 - 1.0$ in the CLAUDS+HSC Survey Using Random Forests

by Nathalie C. M. Thibert

Using a sample of galaxies ($M_\star \geq 10^{10.5} M_\odot$) covering an effective area of ~ 20 deg^2 in the CLAUDS+HSC survey, we apply a Random Forest Classifier to automatically identify merger candidates in deep r-band images. We identify a largely pure, $\sim 90\%$ complete sample of mergers which we use to derive the evolution in the merger fraction from $0.25 \leq z_{\rm phot} \leq 1.0$. We parameterize the merger fraction evolution with a power law of the form $f_{\rm m} = f_0(1+z)^m$. Simulating the effects of increasing redshift on the detectability of mergers, we correct our merger fractions for incompleteness to obtain a local merger fraction of $f_0 = 1.0\% \pm 0.2\%$ and power-law index of $m = 2.3 \pm 0.4$, which is inconsistent with the mild or non-evolving merger scenario ($m < 1.5$) with 96.6% confidence. Finally, we estimate 0.3 merging events to occur per massive galaxy since $z = 1$.

August 23, 2018

Chapter 1

Introduction

There is overwhelming evidence that we live in an expanding Universe dominated by *non-baryonic* cold dark matter (ΛCDM, e.g., Spergel et al. 2003). A sizable fraction of the *baryonic* matter in our Universe is observed in massive ($\sim 10^{8-12} M_{\odot}$), luminous structures known as *galaxies*, which themselves are made up of millions to billions of stars, gas, and dust. Galaxies are thought to form and subsequently evolve over cosmic time through the hierarchical assembly of the even more massive ($10^{12-15} M_{\odot}$) dark matter halos in which they reside. Unfortunately, we cannot directly observe the merging of these dark matter halos and must instead turn to the galaxies themselves as probes of this dominant physical process behind the build up of massive structures in our Universe.

1.1 Mergers in the Framework of Galaxy Evolution

Studies of galaxy morphology date back to the first observations of galaxies (e.g., Hubble 1926). The appearance of a galaxy can help us to infer properties such as current star formation activity, galaxy mass, and local environment. Well-studied morphological types such as spiral and elliptical galaxies show clear trends in their properties, however, not all galaxies fall into these two categories. A number of galaxies in both the local and distant Universe show signs of an ongoing interaction with a companion galaxy. If these interaction events result in the two galaxies coalescing into a single system, they are called *mergers*. Mergers are thought to be a contributing effect to the mass build up and overall evolution of galaxies through cosmic time. If we observe two galaxies merging, then we can infer the same of their dark matter halos, providing us with an indirect probe of hierarchical structure formation.

Mergers typically involve two galaxies (called the *progenitors*) that, through mutual gravitational attraction, are drawn toward one another. Mergers can occur between two gas-rich galaxies (called a "wet" merger), two gas-poor galaxies (called a "dry" merger), or between one gas-rich and one gas-poor galaxy (called a "mixed" merger). At $z = 0.2 - 1.2$, the most common merging events are wet mergers (Lin et al. 2008; de Ravel et al. 2009). Dry and mixed mergers are less common over all epochs, but are seen to increase in importance in the local Universe.

The *modern merger hypothesis* is used to describe the merger scenario involving gas-rich progenitors (see Hopkins et al. 2006, 2008). The merger begins with an initial pass;

both galaxies are now in the same dark matter halo and they begin to lose angular momentum. Global star formation also increases during this phase. Long streams of gas and stellar material can be ejected from the galaxies resulting in a phenomenon called a *tidal tail*. Tidal tails are generally diffuse in structure and have low surface brightnesses when compared to the cores of the progenitors. The length of a tidal tail depends on several factors; for example, the stellar masses, gas fractions, and initial orbital parameters of the progenitors. In particular, the longest tidal tails are produced in wet mergers during a prograde encounter; i.e., when both the orbital axis of the system and the rotational axes of the individual progenitors are parallel (Wen & Zheng 2016). Mergers whose inital orbital parameters are not conducive to producing long tidal tails may instead show disturbed or asymmetric morphologies, clumps of ongoing star formation, or shorter streams of material in the envelope around the interacting pair.

After the initial pass, the galaxies will eventually lose enough angular momentum and coalesce into a single galaxy with a relaxed core. Timescales from initial pass to final coalescence are usually on the order of ~ 1 Gyr (see Figure 1 in Hopkins et al. 2008). After coalescence, black hole growth and feedback dominates the processing of gas in the system until star formation more or less ceases and a "red and dead" elliptical galaxy remains. Up until the final stage, signatures of merging activity may still be visible, provided that the images are deep enough (especially during the AGN phase; Hopkins et al. 2008). The above scenario, illustrated in Figure 1.1, is suggested by hydrodynamic simulations of gas-rich mergers (Hopkins et al. 2006).

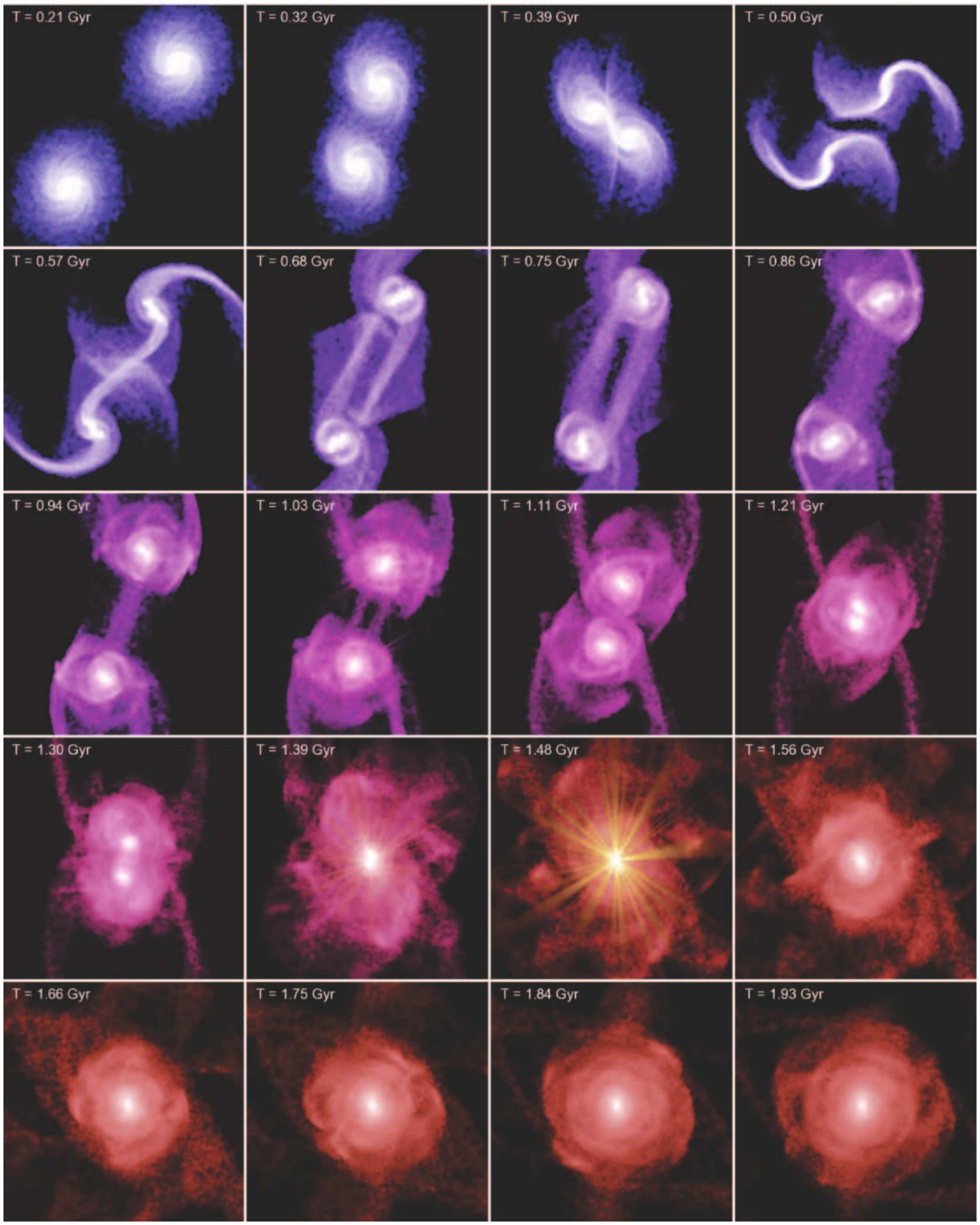


Figure 1.1: Figure 2 from Hopkins et al. (2006). The panels show a time sequence from a hydrodynamic simulation of a gas-rich merger. Quasar activity is denoted with bright point sources (see $T = 1.03, 1.39$, and 1.48 Gyr).

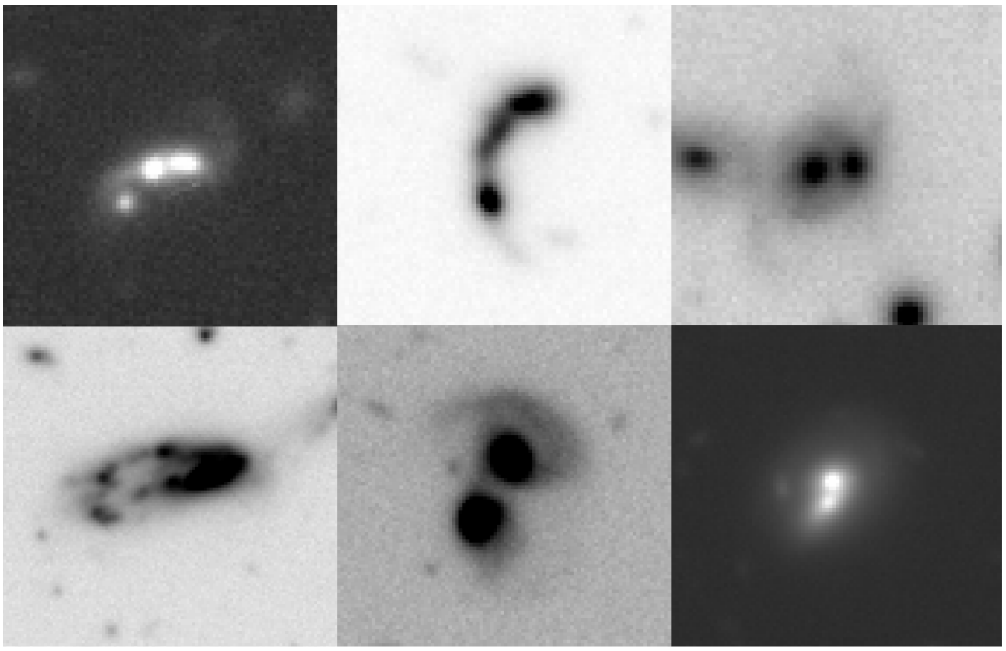


Figure 1.2: Examples of visually identified mergers in the r-band images from the Hyper Suprime-Cam on the Subaru Telescope. The galaxies shown here are meant to illustrate the diversity in the signatures of merging/interacting systems. For example, galaxies with tidal tails, clumps, double nuclei, and global asymmetries are all represented.

Figure 1.2 shows examples of merging galaxies in the well-studied COSMOS and SXDS fields. These mergers show a variety of morphologies and light profiles. These differences could be indicative of different progenitor types, mass ratios, gas fractions, or kinematics. Alternatively, these differences in morphology across mergers may also be caused by the galaxies being in different stages of the merger scenario outlined above. For example, short, wispy tidal features may be evidence of a late stage merger where most of the long tidal tails have already disappeared, or it could be evidence of specific morpho-kinematic properties of the progenitors. It is likely that when we observe a merger, both of the interpretations above are at play to some extent.

After undergoing a merger, global properties such as a galaxy's morphology and star formation activity are altered. The evolution of a galaxy over cosmic time, however, is not solely driven by *ex situ* effects such as interactions with its environment. Other factors such as *in situ* star formation, and black hole growth and feedback can also shape a galaxy over time (Lotz 2007; Conselice 2014). As of yet, the relative contributions of these processes to the growth and evolution of galaxies are not well constrained. We must therefore study how "important" these processes are as a function of redshift (i.e., study their history; Lotz 2007).

Since we cannot follow a single galaxy as it evolves over billions of years, we must take snapshots of the Universe at different epochs and study galaxy properties in a statistical sense. To study the importance of mergers in the evolution of galaxies, we identify which galaxies are merging and which are normal (non-interacting) at each epoch and use this to derive a *volume-averaged merger rate* ($\Gamma_{\rm merge}$ in units of mergers Gyr^{-1} Mpc^{-3}). The values for $\Gamma_{\rm merge}$ at each epoch are then used to study the *evolution* of the merger rate as a function of redshift. Seeing how this evolution behaves and comparing it with similarly derived quantities for star formation and AGN activity can help us to constrain the relative importance of these processes in the overall evolution of galaxies.

In order to derive $\Gamma_{\rm merge}$, we begin by observing a patch of the sky and counting the number of mergers we see at each epoch. We use these observations to derive a *merger fraction*. In the case where we assume the merger progenitors to be part of a single system,

the merger fraction is (Conselice 2014):

$$f_{\rm merge}(M_\star, z) = \frac{N_{\rm merge}}{N_{\rm tot}}, \tag{1.1}$$

where $N_{\rm merge}$ and $N_{\rm tot}$ are the number of mergers and total number of galaxies within a certain redshift bin z and stellar mass limit ($M > M_\star$). If, however, we assume the two progenitors are each their own system, the merger fraction becomes (Conselice 2014):

$$f'_{\rm merge}(M_\star, z) = \frac{2 \times N_{\rm merge}}{(N_{\rm tot} + N_{\rm merge})} = \frac{2 \times f_{\rm merge}}{(1 + f_{\rm merge})}. \tag{1.2}$$

The evolution of the merger fraction with redshift generally follows a power law of the form (Conselice 2014):

$$f_{\rm m} = f_0 \times (1+z)^m, \tag{1.3}$$

where m is the power-law index and f_0 is the fraction of mergers in the local Universe ($z \sim 0$). In recent works, the value for the power-law index m was found to range all the way from $m \sim 0$, corresponding to no evolution in the merger fraction; to $m \sim 4$, corresponding to a strong evolution in the merger fraction (see, for example: Le Fèvre et al. 2000; Conselice et al. 2003; Kartaltepe et al. 2007; Lotz et al. 2008a; Conselice et al. 2009; de Ravel et al. 2009; Lotz et al. 2011). Reasons for these discrepancies may include the use of different redshift ranges or stellar mass cuts, the inclusion of minor mergers (see Section 1.2.1), the precise methods used to identify mergers, and how incompleteness and other systematics are treated.

To fully understand the role of mergers in galaxy evolution, the above merger fraction should be converted into a merger rate (i.e., the number of mergers per unit time per unit volume for a given redshift bin and mass limit). The volume-averaged galaxy merger rate is given by (Lotz 2007; Conselice 2014):

$$\Gamma_{\rm merge}(M_\star, z) = n_{\rm merge} T_{\rm obs}^{-1} = \phi f^\star_{\rm merge} T_{\rm obs}^{-1}, \tag{1.4}$$

where $n_{\rm merge}$ is the number density of mergers, $T_{\rm obs}$ is the timescale during which a merger can be identified, ϕ is the total number density of galaxies (merging and non-merging), and $f^\star_{\rm merge}$ is either $f_{\rm merge}$ or $f'_{\rm merge}$ above.[1] Both galaxy number densities ($n_{\rm merge}$ and ϕ) are for redshift bin z and mass limit $M > M_\star$. Finally, once a merger rate is calculated for each redshift bin, we can derive the evolution in the galaxy merger rate $\Gamma_{\rm merge}$ over a range of epochs.

Transforming a merger fraction ($f_{\rm merge}$ or $f'_{\rm merge}$) into a volume-averaged galaxy merger rate ($\Gamma_{\rm merge}$) requires the knowledge of two quantities: (1) the number density of galaxies at different redshifts ϕ, and (2) the timescale of observability for a merging event $T_{\rm obs}$. Given a large enough survey, we can reasonably estimate ϕ, however, uncertainties introduced by large-scale density fluctuations, known as *cosmic variance* (Somerville et al. 2004), could cause the value of ϕ to be slightly different depending on which volume of the Universe is sampled. The timescale over which a specific type of merger is visible $T_{\rm obs}$ depends on the time it takes for a given merger signature to originate and subsequently disappear in

[1] It does not matter which merger fraction definition you use, so long as you make the appropriate adjustments when comparing your results to other studies.

an image.[2] For example, long tidal tails are only visible for a certain fraction of the total merger process.

To estimate these timescales, we can use simulations of mergers with different progenitor properties and initial conditions. The galaxy number density ϕ and observability timescale $T_{\rm obs}$ may also evolve with redshift (Lotz et al. 2011), and so understanding how they change can help us to obtain more accurate estimates of the merger rate. For example, Lotz et al. (2011) defines a cosmologically-averaged observability timescale:

$$\langle T_{\rm obs}(z) \rangle = \sum_{i,j} w_{i,j}(z) \times T_{i,j}, \tag{1.5}$$

where $w_{i,j}(z)$ and $T_{i,j}$ are the fraction of mergers at redshift z and the observability timescale, respectively, for mergers with baryonic mass ratio i and baryonic gas fraction j. These mass ratios and gas fractions can be calculated using the results of cosmological galaxy evolution models such as Croton et al. (2006); Somerville et al. (2008); Stewart et al. (2009). Ideally, we would like to probe all stages in the Hopkins et al. (2008) merger scenario equally from initial pass to final coalescence and beyond and use a combination of all merger stages to derive a total merger rate.

Some authors (e.g., Bundy et al. 2009; Conselice et al. 2009; López-Sanjuan et al. 2009; Jogee et al. 2009; Bridge et al. 2010) prefer to deduce the role of mergers by instead calculating a *fractional* merger rate $\Re_{\rm merge}$, which traces the number of merging events a galaxy undergoes over a range in lookback times (or equivalently, redshifts). In this case, the

[2] These "observability timescales" should not be confused with the total time it takes for a merger to occur since different merger signatures probe different parts of the total merger scenario.

fractional merger rate is not a volume-averaged quantity and is given by (see Lotz et al. 2011):

$$\mathfrak{R}_{\rm merge} = \frac{f_{\rm merge}}{\langle T_{\rm obs} \rangle}. \tag{1.6}$$

This realization of the merger rate is not dependent on the number density of galaxies.

In the present work, we will focus on deriving the *evolution in the galaxy-galaxy merger fraction* (Equation 1.3). The value for the merger fraction evolution is arguably more straightforward to compare with other studies as there are no assumptions on the galaxy number density or observability timescales imposed. We therefore defer the full derivation of the evolution in the volume-averaged galaxy merger rate to a future work and instead provide an rough estimate of the fractional merger rate $\mathfrak{R}_{\rm merge}$, while cautioning the reader that we cannot yet directly compare our merger fraction estimates with the evolution in star formation rate densities or AGN activity to obtain their relative importances.

1.2 Methods for Identifying Merging Galaxies

1.2.1 Close Pair Studies

Studies involving the identification of merging galaxies usually take one of two approaches. In the first, mergers are identified based on their morphologies and distributions of light in images (this is the approach that we will focus on in this work). The second approach is both widely used and conceptually simple and, although not the subject of this work, still deserves some attention– the identification of physically *close pairs* of galaxies.

A close pair consists of a more massive primary galaxy (M_1) and a less massive secondary

galaxy (M_2) that are not currently undergoing a merger, but are very likely to merge in the near future. Pairs are identified by their position on the sky (using a projected separation $R_{\rm proj}$), as well as through constraints on their relative velocities or redshifts determined using spectroscopy ($\Delta z \equiv |z_1 - z_2|/(1 + z_1)$). If spectroscopy is unavailable, photometric redshifts (sometimes referred to as photo-z's) are used. They must also satisfy a stellar mass ratio ($\mu = M_1/M_2$), which is used to distinguish between *major* and *minor* mergers. The definitions of major and minor mergers can vary from study to study, but they are roughly as follows (see, for example, Man et al. 2016): for a major merger μ is 1:1 − 4:1, and for a minor merger μ is 4:1 − 10:1. For examples of close pair studies, refer to: Patton et al. (1997); Le Fèvre et al. (2000); Patton et al. (2002); Kartaltepe et al. (2007); de Ravel et al. (2009); Bluck et al. (2009); Williams et al. (2011); Man et al. (2012); Newman et al. (2012); Man et al. (2016).

Using close pairs to identify potential mergers can be useful in cases where deep imaging data are unavailable and faint merger signatures lie below the detection limit of the telescope. Another advantage to using this method is that large samples of close pairs can be identified using only a few simple criteria. There are, however, a few caveats to using close pairs as a proxy for mergers. First, when using imaging data to derive projected separations, we are limited by the resolution of the telescope; in other words, we must have high spatial resolution to distinguish between multiple, nearby galaxies. This limitation is not specific to the close pair method. Any method of merger identification involving galaxy images is subject to limitations due to resolution. Second, obtaining spectra of merging galaxies can be difficult because of their small separations. Depending on the spectrometer, overlapping

slits or fiber collisions can sometimes occur, making it impossible to obtain spectra for both galaxies simultaneously. Finally, if using photometric redshifts, uncertainties in these values may lead to the contamination of samples by chance projections (Lotz et al. 2011). We must therefore be careful when using close pairs to infer merger activity since we are only assuming that the two galaxies will merge in the future.

1.2.2 Visual Identification

The simplest method by which mergers are identified is through *visual inspection*. In this method, human annotators look through galaxy images and flag potential merger candidates on the basis of features such as tidal tails, double nuclei, and global asymmetries. In general, the conductors of visual identification studies will choose a redshift range and limiting magnitude down to which they believe the imaging data accurately probe faint tidal features and other merger signatures. A set of criteria are then chosen such that each inspected galaxy falls into one of several predefined categories.

An example of a study that used visual identification to select mergers is Bridge et al. (2010). In their study, they used i-band images from the Canada−France−Hawaii Telescope Legacy Deep Survey (CFHTLS-Deep). They visually inspected 2 deg^2 of the sky in the redshift range $0.2 < z < 1.2$ down to a limiting magnitude of 22 in the i-band and identified about $1,600$ merging galaxies from a parent sample of about 25,000 galaxies. To select mergers, they predefined a set of criteria that they thought were indicative of merger activity; namely, short, medium, and long tidal tails, tidal bridges, and double nuclei. In general, by choosing several different merger signatures and by being able to inspect each galaxy

individually, a larger variety of mergers at different stages in the merger scenario are able to be selected. This, in theory, allows for a more robust derivation of the merger fraction (provided that sample completeness is well understood).

Another example of the human annotation of galaxy morphologies is the Galaxy Zoo Project, a citizen-based science project from which the morphologies of galaxies can be determined using the concept of a majority vote.[3] For example, the "Galaxy Zoo: Hubble" (GZH) Project of Willett et al. (2017) uses imaging data from the Hubble Space Telescope's Advanced Camera for Surveys (ACS). Pre-defined questions about the appearance of each galaxy are used to direct citizens down branches of a "decision tree," ultimately resulting in a morphological classification for a particular galaxy. In the case of the GZH Project, an average of ~ 200 human annotators provide responses to the questions for each galaxy, allowing for outlying responses to be statistically down-weighted. The Galaxy Zoo effort is not specifically tailored to identifying merging galaxies, but certainly includes them as a morphological criterion when posing their questions to the public.

1.2.3 A Motivation for Automatic Detection Methods

Visual identification of mergers, although simple, can be biased depending on the human annotator. In addition, the same annotator may not be able to reproduce the exact same results if given the same set of images a second time (see Section 3.3 of Bridge et al. 2010). Furthermore, as galaxy surveys become wider and deeper, the use of visual inspection as a means to identify mergers becomes very time consuming, even when considering citizen-

[3] https://data.galaxyzoo.org/

based efforts like Galaxy Zoo. We therefore need to devise automatic, computer-assisted methods for identifying interacting galaxies based on their morphologies and light distributions.

Fortunately, tools developed in the field of *supervised machine learning* have recently been popularized in Astronomy. In particular, a variety of these tools has been used to automatically classify the morphologies of galaxies based on samples for which the morphologies are already known. The success of such supervised machine learning approaches hinges on the definition of an informative set of *features* used to describe each object in the dataset. For example, a galaxy can be described by quantities such as its stellar mass, star formation rate, morphological classification, etc. Each of these quantities, or features, can be used together to describe a particular galaxy and help to differentiate it from other galaxies that might have different properties.

Several techniques have been developed in the last few decades which utilize *non-parametric* measures of galaxy morphology; in other words, none or very few assumptions about the distribution of light in a galaxy are made. This approach has proven to be very useful when quantifying the morphologies of interacting galaxies because, in general, their light distributions do not follow a distinct parametric form. Early studies involving non-parametric indicators of morphology used only one or two features at a time to separate mergers from the remainder of the galaxy population (see Abraham et al. 1996; Bershady et al. 2000; Conselice et al. 2000a; Abraham et al. 2003; Conselice et al. 2003; Lotz et al. 2004, 2008a; Cotini et al. 2013). Later studies introduced features specific to the detection of mergers (see Law et al. 2007; Hoyos et al. 2012; Freeman et al. 2013; Wen et al. 2014;

Pawlik et al. 2016; Peth et al. 2016; Wen & Zheng 2016). A few more recent studies used higher-dimensional feature spaces (i.e., three or more parameters at a time) to describe the morphologies of their galaxies (e.g., Hoyos et al. 2012; Freeman et al. 2013; Shamir et al. 2013; Elfattah et al. 2014; Cibinel et al. 2015; Peth et al. 2016; Goulding et al. 2018). In the higher-dimensional approach we are not limited to considering only one or two parameters at a time and can utilize a much more robust description of galaxy morphology.

In this work, we use a supervised machine learning approach, called a *Random Forest Classifier*, along with a collection of 14 features to automatically identify merging galaxies in deep r-band images from the Subaru Telescope. We use the results of our automatic classification to derive the evolution in the merger fraction $f_{\rm m}(M_\star, z)$ for galaxies with $M_\star \geq 10^{10.5} M_\odot$ and $0.25 \leq z_{\rm phot} \leq 1.0$.

Chapter 2

Catalog Details & Sample Selection

2.1 The CLAUDS+HSC Survey

The imaging and photometric data used in this work are drawn from preliminary data releases of two ongoing surveys, CLAUDS and HSC. In a collaboration between Japan, Taiwan, and Princeton University, the HSC (Hyper Suprime-Cam) instrument on the 8.2-m Subaru telescope on Mauna Kea was used to gather data in three separate fields: the *Wide*, *Deep*, and *UltraDeep* layers (see Aihara et al. 2017 for details on the HSC survey design).

The data we will use to find mergers belong to the *Deep* and *UltraDeep* fields, which cover an area of 26 deg^2 in five broadband filters (*grizy*) down to a limiting magnitude of $i \simeq 27$. The HSC survey *Deep* and *UltraDeep* layers overlap with several well-studied areas such as COSMOS (Scoville et al. 2007) and XMM-LSS (Pierre et al. 2004), allowing for

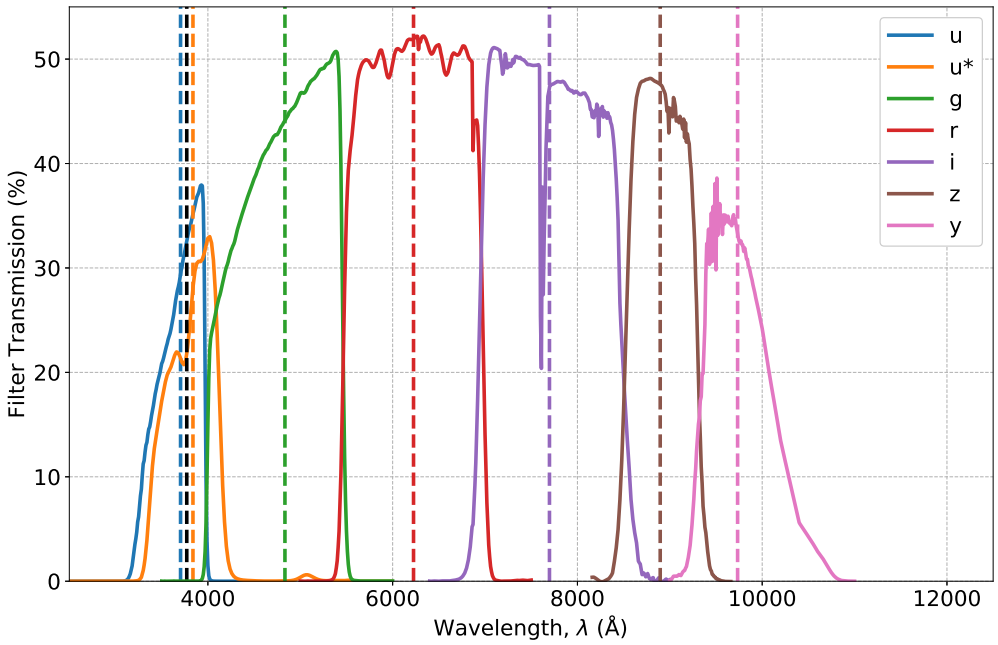


Figure 2.1: Filter transmission curves for the five broadband HSC filters, *grizy*, and the two broadband CFHT filters, u and u^*. Dashed lines mark the central wavelengths for each of the filters and are determined such that half of the area under each curve lie on either side of the line. The black dashed line marks the average between the central wavelengths of the u- and u^*-bands. Transmission curves include corrections for both instrument transmission and atmospheric opacity.

multiwavelength analyses.

The Canada-France-Hawaii Telescope (CFHT) Large Area U-band Deep Survey (CLAUDS; see overview paper from Sawicki et al., in prep) is a collaboration between researchers from Canada, France, and China. It provides u-band data from the MegaCam instrument on CFHT down to a limiting AB magnitude of $u = 27$ in a $\sim$20 deg^2 overlapping region within the HSC *Deep* and *UltraDeep* layers. By combining both the CLAUDS and HSC surveys over this 20 deg^2 area, we are able to obtain both deep images and more accurate photometric redshifts for use in further analyses.

Figure 2.1 shows the filter transmission curves for the five broadband HSC filters and two u-band filters from CLAUDS.[1] Table 2.1 summarizes the properties of the filters described above. In this study, we will use imaging data in the HSC r-band and photometrically derived quantities using all broadband filters (u or u^*, and $grizy$) for the morphological analysis of galaxies in the 20 deg^2 overlap region between the MegaCam and HSC data. We choose to use only the r-band in our analysis both for simplicity in calculations and, more importantly, because the imaging data in this band are deeper compared to the other CLAUDS+HSC filters, allowing us to probe the faint tidal signatures associated with some mergers.

Table 2.1: CLAUDS+HSC filter properties.

Filter	Central Wavelength (Å)
CFHT-u	3704.2
CFHT-u^*	3838.5
HSC-g	4834.0
HSC-r	6226.0
HSC-i	7697.4
HSC-z	8896.1
HSC-y	9734.0

[1]The newer u filter is preferred to the old u^* filter where available because of its larger area and throughput (Sawicki et al., in prep).

2.2 Sample Selection in CLAUDS+HSC

2.2.1 Parent Samples in the COSMOS, SXDS/XMM-LSS, ELAIS−N1, and DEEP 2−3 Fields

The primary galaxies we consider for our merger fraction estimate are drawn from a parent sample in the CLAUDS+HSC dataset, which covers $\sim$ 20 deg^2 over four separate fields: COSMOS, SXDS/XMM-LSS, ELAIS−N1, and DEEP 2−3. Each field covers $\sim 4-6$ deg^2 (see Sawicki et al., in prep). In the four fields, HSC r-band imaging is available at the HSC *Deep* depth ($r \approx 27$ mag). In addition to the *Deep* fields, there is also overlapping *UltraDeep* HSC r-band imaging (down to a limiting magnitude of $r \approx 28$ mag) available in the COSMOS and SXDS fields.[2] We preferentially use *UltraDeep* data where available (because it is deeper) and we consider there to be six fields in total where the COSMOS and XMM-LSS *Deep* layers do not include the overlapping regions from their corresponding *UltraDeep* layers, so as to eliminate any double-counting.

By using both the *Deep* and *UltraDeep* layers to calculate the merger fraction, we are assuming that the detectability of merger signatures is the same at both depths. This, in general, is not true because shallower data may cause merger signatures to disappear in the images. Merger fractions derived using the HSC *Deep* data may therefore be a lower limit since we could be missing mergers in the *Deep* data that would otherwise be detected in the *UltraDeep*. We do not treat this issue in the present work, however, one could quantify the difference in merger detectability across the *Deep* and *UltraDeep* layers by comparing

[2]The XMM-LSS field refers to the HSC *Deep* data, while the SXDS field refers to the smaller region overlapping with XMM-LSS for which *UltraDeep* data are also available.

results from the overlapping regions at both depths.

We require that there be photometric redshifts and stellar masses measured for our galaxies. Masses were calculated by Golob et al. (in prep) using BC03 stellar population models (Bruzual & Charlot 2003) and photometric data from six broadband CLAUDS+HSC filters (u or u^*, $grizy$) in the `LEPHARE` code of Ilbert et al. (2006), which uses a χ^2 minimization technique to determine the physical parameters of the best fitting model spectral energy distribution (SED). The photometric redshifts are determined using the photometric colors in the CLAUDS+HSC filters and a supervised machine learning algorithm (k-nearest-neighbours, see Section 2.2.2). The CLAUDS+HSC catalogs constructed by Golob et al. (in prep) include *both* stars and galaxies. They report the probability $P_{\rm star}$ that any particular object in the catalog is a star. For our galaxy population to remain reasonably complete, we only consider objects with $P_{\rm star} < 0.85$. This methodology and threshold in probability does not ensure 100% purity and so a small fraction of our objects are still likely to be stars.

In total, across the four *Deep* fields and two *UltraDeep* fields, the number of *galaxies* with measured masses and photometric redshifts is 6,735,580. Column *(3)* of Table 2.2 gives galaxy number counts and Figures 2.2−2.4 show the 2-dimensional projected spatial distributions of the galaxies across all six fields in the parent sample.

2.2.2 Targets for Merger Classification in CLAUDS+HSC

To study the evolution in the merger fraction, we consider a subset of the parent sample described above. We first require that fluxes be measured in each of six broadband filters (u or u^*, and $grizy$) and that their values be positive. These fluxes are converted to apparent

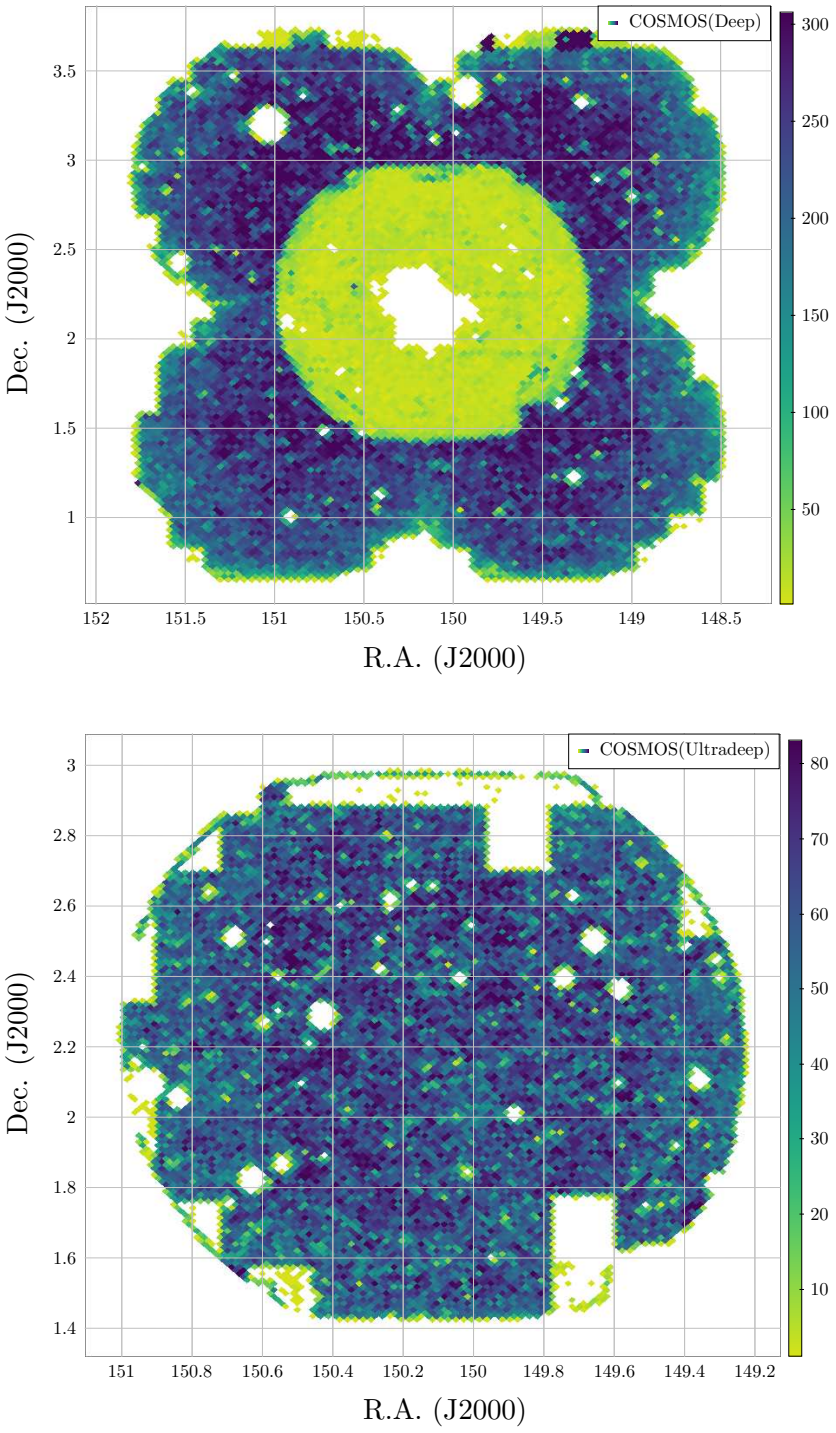


Figure 2.2: Projected 2-dimensional density plots of the distribution of galaxies in our parent sample for the COSMOS *Deep* (upper) and *UltraDeep* (lower) fields. The low density region of the *Deep* layer (centre of field) is where the *UltraDeep* overlaps.

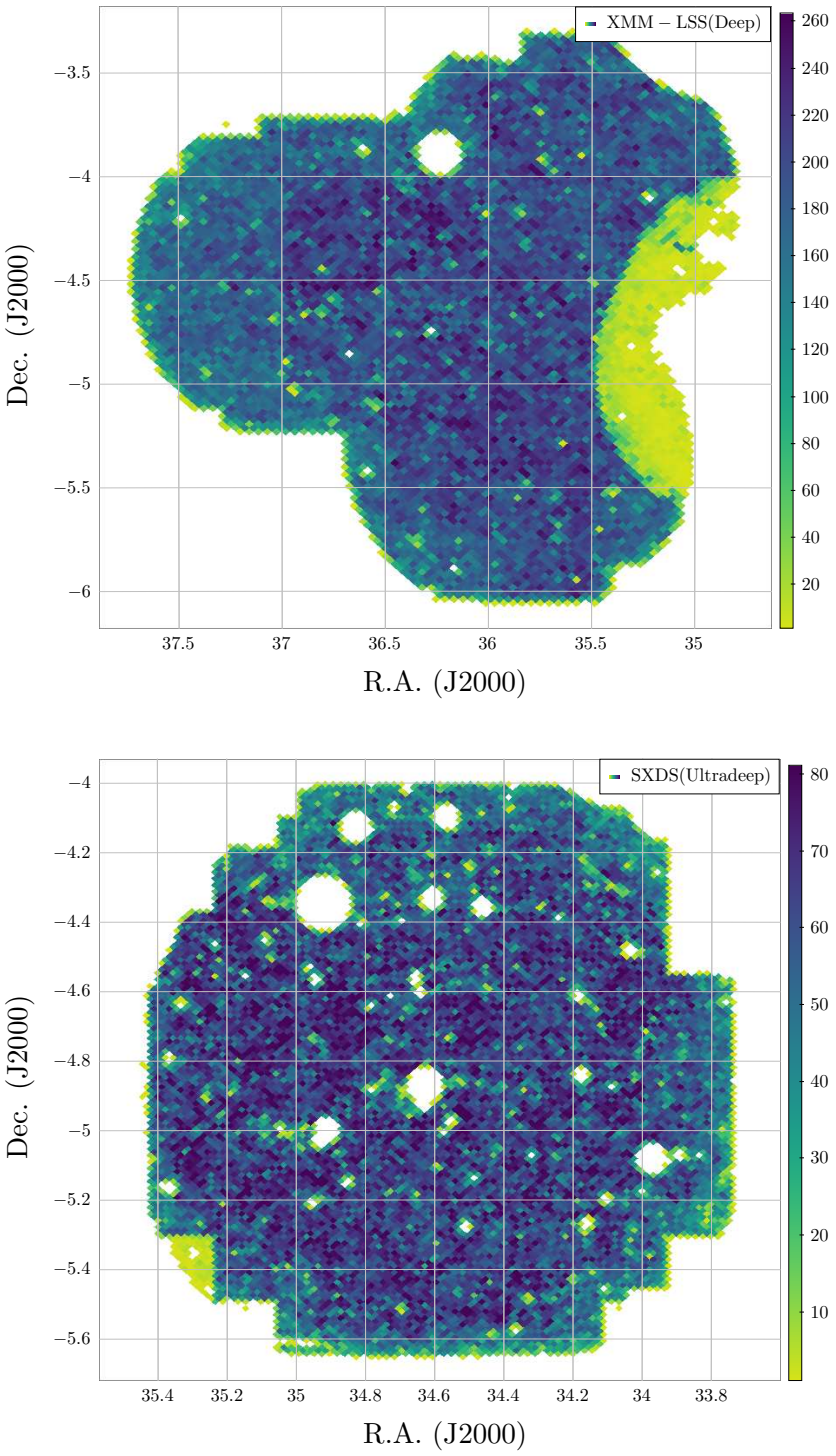


Figure 2.3: Same as Figure 2.2 but for the XMM-LSS *Deep* (upper) and SXDS *UltraDeep* (lower) fields. The low density region of the *Deep* layer (green) is where the *UltraDeep* overlaps.

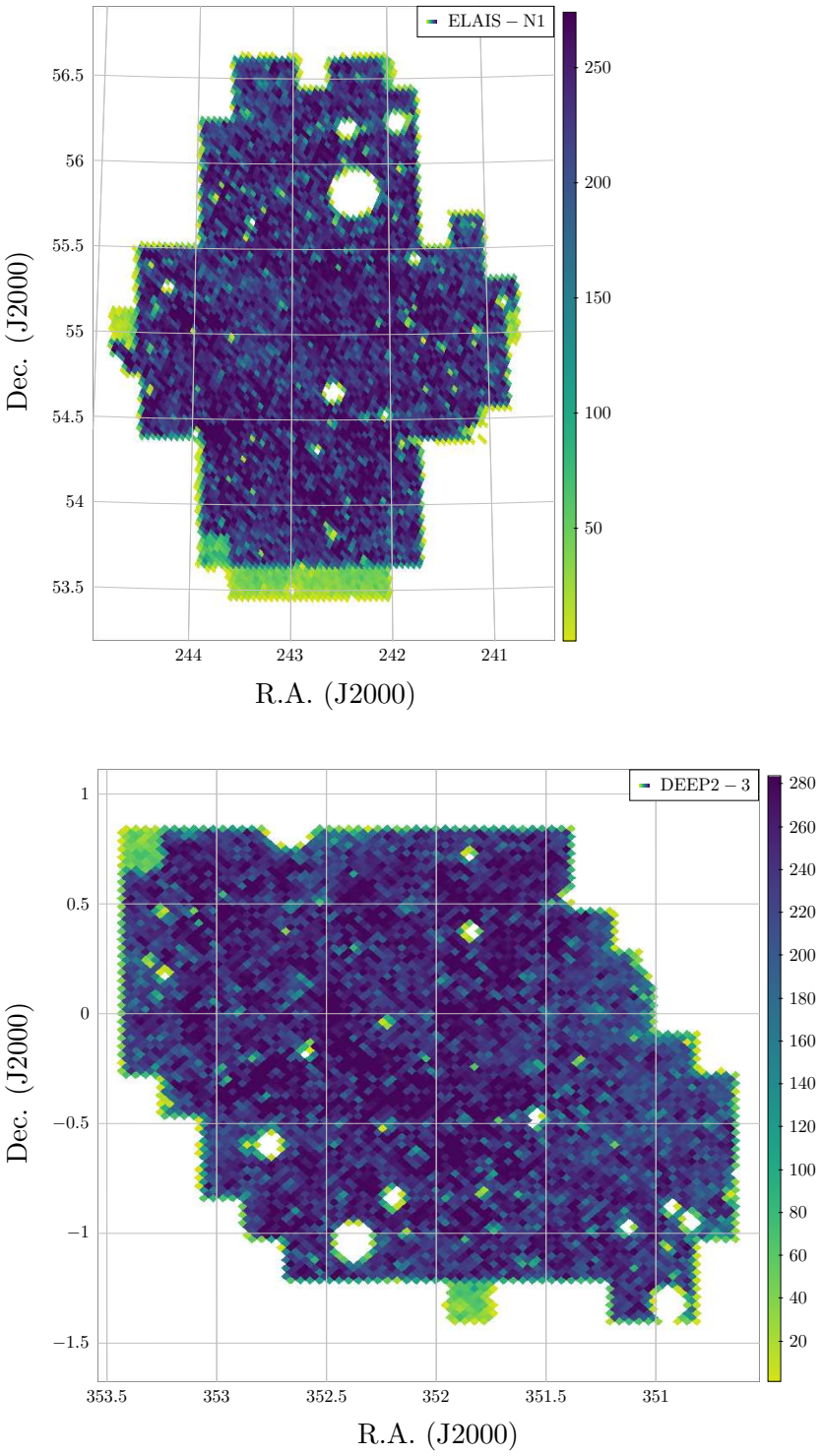


Figure 2.4: Same as Figure 2.2 but for the ELAIS−N1 (upper) and DEEP2−3 (lower) fields, which are both part of the HSC *Deep* layer.

Table 2.2: Overview of fields in the *Deep* and *UltraDeep* layers of the CLAUDS+HSC survey.

Field	HSC Survey Depth	Total # of galaxies ($M_\star$ & photo-z)[a]	# of galaxies (params)[b]	Final galaxy sample ($f_{\rm merge}$)[c]
COSMOS	*Deep*	1,795,578	12,940	11,788
COSMOS	*UltraDeep*	562,110	5,362	4,498
XMM-LSS	*Deep*	1,100,736	11,542	10,608
SXDS	*UltraDeep*	627,424	6,601	5,780
ELAIS−N1	*Deep*	1,292,927	16,597	14,885
DEEP 2−3	*Deep*	1,356,805	14,631	13,398
Total:		6,735,580	67,673	**60,957**

[a] This column represents the number of galaxies ($P_{\rm star} < 0.85$) in the *Deep* and *UltraDeep* layers of the survey for which stellar mass and photometric redshift information is available. This is the "parent sample" from which we choose our primary galaxies.

[b] This column gives the number of targets in the *Deep* and *UltraDeep* layers of the survey. They are the galaxies for which we attempt to calculate morphological parameters. See Section 2.2.2 for the criteria applied to the parent sample.

[c] This column gives the final number counts of galaxies used to estimate the merger fraction evolution. Not all galaxies in Column *(4)* had successful parameter estimates and so we leave them out of the analysis.

magnitudes in the AB system using the following formula:

$$m_{{\rm AB},x} = -2.5 \log_{10} \left(\frac{f_{\nu,x}}{\rm Jy} \right) + 8.90, \tag{2.1}$$

where $f_{\nu,x}$ is the flux of a galaxy in filter x in units of Janskys. After this conversion, the catalog is cleaned to include only galaxies brighter than $m_{{\rm AB},r} \leq 23$ mag in the HSC r-band. By imposing this brightness criterion, we have ensured that the galaxies in our sample are bright enough such that unambiguous visual classifications can be made. This is especially important while training the Random Forest (see Section 4.2.2). We also exclude objects that are within $20h^{-1}$ kpc of a bright ($r_{\rm AB} \leq 21$ mag) star. This ensures that we do not

include starlight which could contaminate the pixels immediately surrounding our galaxies.

Next, we apply a conservative mass cut by requiring that our galaxies be more massive than $M_\star \geq 10^{10.5}\ M_\odot$. We also consider only galaxies in the range of photo-z's between $0.25 \leq z_{\rm phot} \leq 1.0$. Our choice of mass cut will become apparent in Section 4.2.2 when we visually classify galaxies to train a computer to identify mergers. At the highest redshifts, we are unable to obtain unambiguous visual classifications and as a result our galaxy sample is incomplete at masses below this threshold. For consistency, we apply this mass cut across all redshifts. We choose our lower bound in redshift in part because the volume of the Universe probed at these redshifts in the CLAUDS+HSC survey is not sufficient to estimate a merger fraction that does not suffer from small number statistics. Arguably more important would be the accuracy of our photo-z's below $z_{\rm phot} \sim 0.25$. The method used by Golob et al. (in prep) to estimate the photometric redshifts uses a k-nearest-neighbour approach in a 5-dimensional color space ($u-g$, $g-r$, $r-i$, $i-z$, $z-y$) to obtain a probability density function (PDF) for each galaxy's redshift. Their results are based on a training sample of galaxies with known photo-z's derived by comparing the galaxies' photometry to template SEDs. In some cases, there are degeneracies where two or more prominent peaks are observed in the PDF and the most likely photo-z obtained by the algorithm is in fact that which corresponds to the incorrect peak.[3] We choose an upper bound of $z_{\rm phot} = 1.0$ to, again, decrease the number of catastrophic failures in the photo-z estimation, but also because of the accuracy of our stellar masses above this redshift. Golob et al. (in prep) showed that above $z_{\rm phot} \sim 1$, the stellar masses are much less certain because the part of the spectrum used to measure

[3]These *catastrophic failures* are a rare, but expected, result of supervised machine learning problems. In our case, we choose a redshift range for which the number of catastrophic outliers is decreased.

masses at higher redshift includes only the redder bands. This uncertainty in stellar mass at higher redshift is contributed to by a combination of fewer photometric data points with which to perform the fits, as well as shallower data available in the bandpasses with longer observed wavelengths (e.g., HSC z- and y-bands). We therefore decide to work with only those galaxies whose physical properties are accurately measured.

Some fields in the HSC r-band images suffer from low signal-to-noise. To accurately determine the morphological features of a galaxy, we need to be able to distinguish the majority of the galaxy light from that of the local background. If the signal-to-noise ratio is too low, then more pixels from the background could mistakenly be considered to belong to the galaxy, which can introduce artifacts into the morphological features. Several authors have attempted to quantify the effects of signal-to-noise on morphological parameter estimation; e.g., Conselice et al. (2000b) discuss galaxy asymmetry in this context. By visually inspecting a subsample of galaxies with various signal-to-noise levels, we choose to only include galaxies that satisfy:

$$\frac{f_\nu(r, 24 \text{ pix})}{\sigma_{f_\nu}(r, 24 \text{ pix})} > 14.0, \tag{2.2}$$

where $f_\nu(r, 24 \text{ pix})$ is the HSC r-band flux contained within a circular aperture 24 pixels in diameter, and $\sigma_{f_\nu}(r, 24 \text{ pix})$ is the error on that flux (from the catalogs of Golob et al., in prep). We use this as a rough estimate of the signal-to-noise ratio in our images.

We apply the above criteria to our parent sample to obtain a sample of galaxies in the CLAUDS+HSC catalog. This sample consists of 67,673 galaxies over 20 deg^2 and Column *(4)* in Table 2.2 outlines how many galaxies reside in each of the six fields. Finally,

only 60,957 galaxies could have their morphological parameters accurately measured (see Chapter 3)[4] and these galaxies are summarized in Column *(5)* of Table 2.2. Figures 2.5 and 2.6 show the apparent AB magnitude in the HSC r-band and the stellar mass, respectively, as a function of photometric redshift for the galaxies in our final catalog of galaxies (Column *(5)* of Table 2.2).

[4] As will become clear in Chapter 3, reasons for discarding $\sim$ 10% of our galaxy sample include "bad" pixels, crowded fields around the primary galaxy, and insufficient local background for parameter estimation.

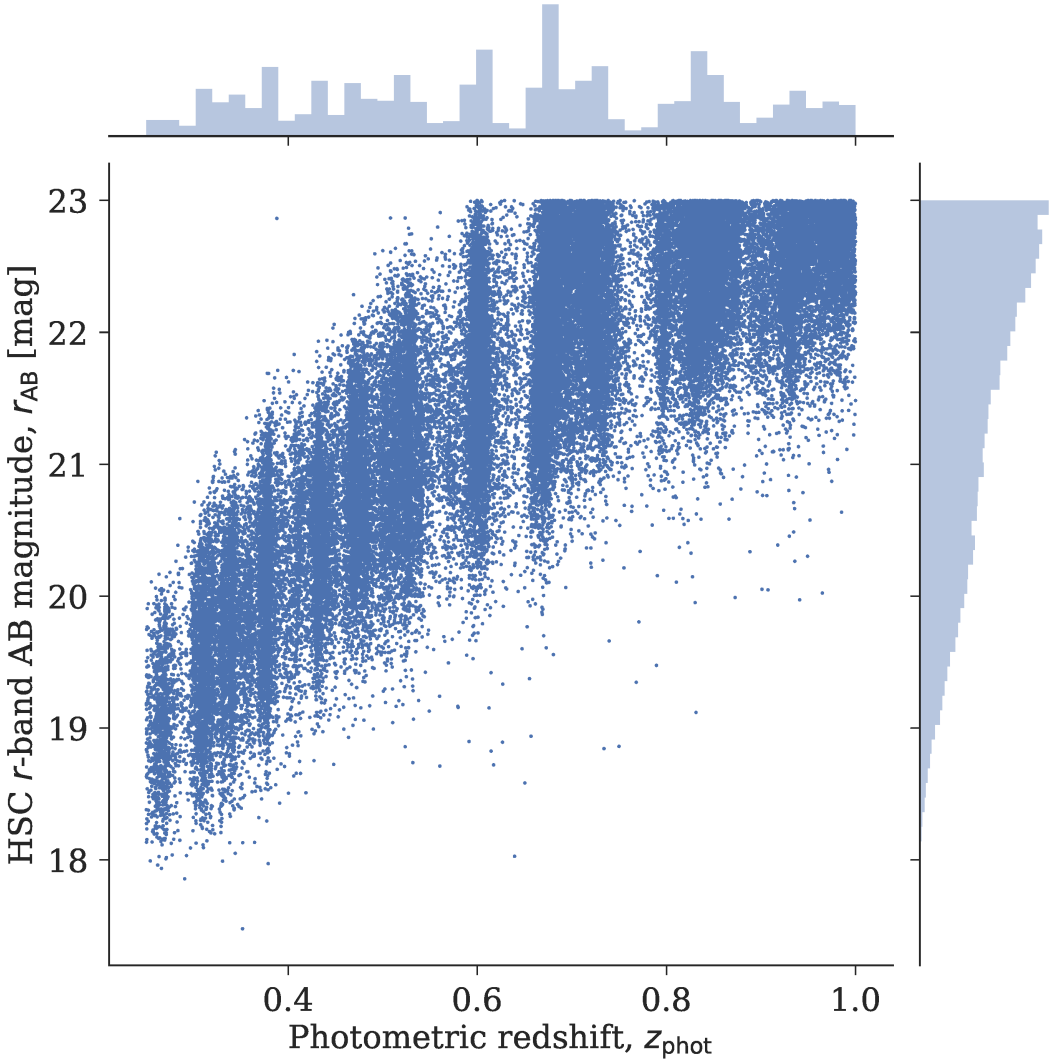


Figure 2.5: Apparent magnitude in the HSC r-band as a function of photo-z for the galaxies in the final sample (60,957 galaxies) used to estimate the merger fraction. The histograms show the distributions (normalized to unit area) of each parameter, respectively.

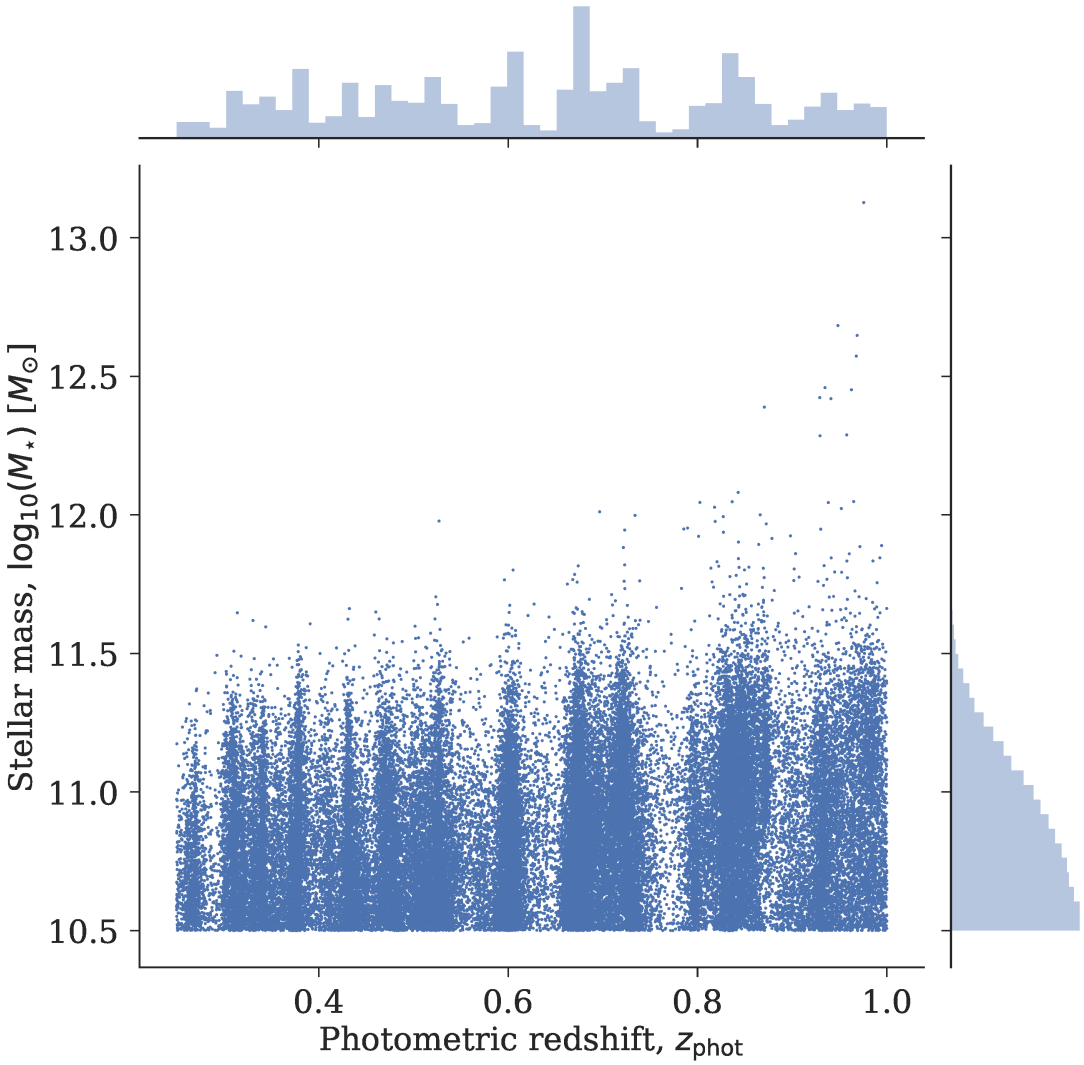


Figure 2.6: Same as Figure 2.5 but for the galaxy stellar mass as a function of photo-z.

Chapter 3

Non-parametric Measures of Galaxy Morphology

In this Chapter, we describe the methods used to obtain the input *features* to our Random Forest Classifier. We begin with a summary of how cutouts are created around our primary galaxies. We describe how we mask bad pixels and extract the pixels belonging to only the primary galaxy in the cutout. We then introduce the morphological parameters and describe how they are calculated for our galaxy cutouts. We outline the recipe used to model galaxy light profiles and produce residual images, on which several other morphological parameters can be measured. Finally, we describe a few popular approaches used to find mergers where only one or two morphological parameters are employed at a time. For many of our morphological parameter calculations, we modify code written as part of the STATMORPH package.[1]

[1] https://github.com/vrodgom/statmorph

3.1 Cutouts, Cleaning, and Segmentation Maps

3.1.1 Galaxy Cutouts in the Rest Frame

For each field in the CLAUDS+HSC survey (COSMOS/XMM-LSS/SXDS/ELAIS−N1/DEEP 2−3) the imaging data are split into several *tracts*, inside which there are a number of smaller *patches*. Each patch is roughly 0.04 deg^2 and can contain up to several thousand galaxies of different sizes, magnitudes, and redshifts. When considering a single galaxy, the imaging data we use is in the form of a `FITS` image for a particular *patch*.[2]

For each primary galaxy in our sample, we create a *cutout* around that particular galaxy within the full patch image so that it may be considered more or less on its own. A fixed height and width (in pixels) for all of the cutouts is a poor choice since galaxies at higher redshifts appear, in general, smaller. For example, if we were to choose a cutout size of 100×100 pixels, then galaxies at low redshifts may be too large and run off the cutout. We therefore decide to make our cutouts such that they are the same *physical* size; e.g., 100×100 (physical) kpc in the rest frame of the galaxy. To do this, we must use the centroid in units of pixel coordinates and the photometric redshift of the galaxy, both as calculated by Golob et al. (in prep). To convert from R.A. and Dec. to pixel coordinates, we use the World Coordinate System (WCS) information in the header of the `FITS` image for the corresponding patch and the `wcs_world2pix` function from the `astropy.wcs` package in `Python`. We assume a flat cosmology with $H_0 = 70$ km s^{-1} Mpc^{-1}, $\Omega_{\mathrm{m},0} = 0.3$, and

[2]The `FITS` images are 4200×4200 pixels in size, which corresponds to 4200 pixels $\times$ ($0.168''$/ pixel) = 0.196 deg on a side, or 0.038416 deg^2. Each patch, however, does not necessarily have data spanning this full area as the fields are not square.

$T_{\rm CMB} = 2.725$ K and apply the following conversion:

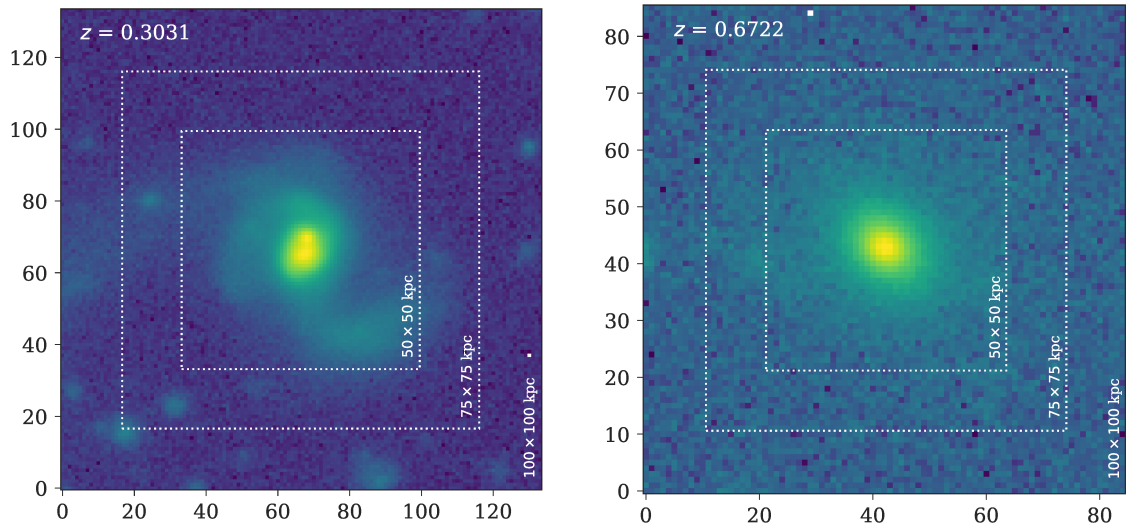
$$\delta\vartheta \text{ [pixels]} = \frac{D \cdot (1+z)^2}{d_L(z)} \times \frac{360^\circ}{2\pi \text{ rad}} \times \frac{3600''}{1^\circ} \times \frac{1 \text{ pixel}}{0.168''}, \tag{3.1}$$

where the first term is the *angular size* of the cutout in radians, and the last three terms are used to convert this size to pixels. In particular, the luminosity distance $d_L(z)$ is calculated using the `astropy.cosmology.FlatLambdaCDM.luminosity_distance` routine and converted to kpc, and D is the desired physical size of the cutout in kpc. The last term is referred to as the *pixel scale* and it is a property of the HSC images; in other words, each pixel in the images corresponds to $0.168''$ on the sky.

We begin by creating cutouts around each galaxy of size 100×100 kpc. This is the *maximum* size we use throughout the reduction of our imaging data. We also make use of the 75×75 kpc and 50×50 kpc cutouts later for steps such as local background estimation and morphological parameter calculations. Figure 3.1 shows examples of the three types of cutouts mentioned above for two galaxies of different redshifts in the HSC r-band data. The galaxy at redshift $z = 0.30$ is a visually classified merger, while the galaxy at redshift $z = 0.67$ is a visual non-merger.

3.1.2 Masking Bad Pixels and Estimating the Local Background

If we wish to determine which pixels belong to a particular galaxy, we must include a *threshold* in our calculations. Typically, this threshold is related to the level of the *local*



(a) (b)

Figure 3.1: Examples of cutouts for two galaxies *(a)* a merger with a double nucleus, and *(b)* a non-merger at different redshifts. Here, we show the 100×100 kpc cutout (full image) with white dotted lines denoting the bounding boxes for the 75×75, and 50×50 kpc cutouts, respectively. The labels along the axes correspond to the number of pixels in the image.

background in the region immediately around the galaxy in question.[3] In order to obtain accurate estimates of the local background around each primary galaxy, we must mask out two types of pixels: (1) those deemed "bad" by the CLAUDS+HSC imaging pipeline (Golob et al., in prep), and (2) those belonging to objects detected in the cutout. The "bad" pixels we choose to ignore in all calculations are flagged in a bitwise fashion by the pipeline as:

Description of Flag	Bit
Bad	0
Saturated	1
Suspect	7
No Data	8
Bright Object	9

[3] It is important to consider local values because background levels can vary quite drastically across images.

The bitwise value of 5 corresponds to pixels that belong to objects detected by the pipeline. We only exclude these pixels when dealing with estimates of the local background. Figure 3.2 shows an example of the bitwise masks in an area of the SXDS field. In this case, we can see detected objects (light blue), and several examples of "bad" pixels that would be masked in our calculations. At this stage, we check whether the 50×50 kpc cutout has more than 50% of its pixels flagged as "bad" (i.e., bitwise mask values belonging to the table above) when the original mask is rotated by 180° and added to itself.[4] If this is the case, we exclude the galaxy from further calculations. We do this to ensure that there is enough good data close to the primary galaxy so that meaningful morphological parameters can be obtained. Only 1259 (2%) of the primary galaxies had 50% or more of their cutout pixels masked due to bad, saturated, suspect, no data, or bright object pixels.

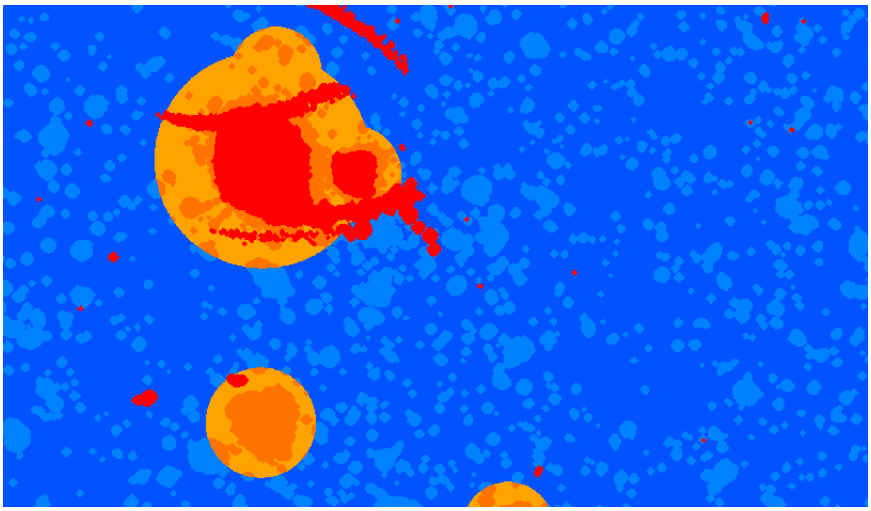
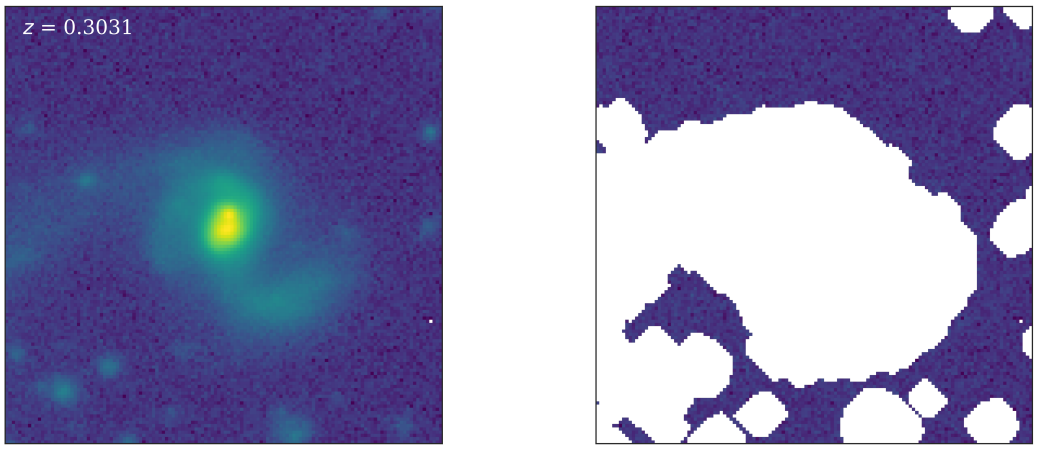


Figure 3.2: Examples of bitwise masks in the SXDS field. The final bitwise values in the mask image are a combination of all flags for a single pixel. Here, large orange circles are bright stars, smaller light blue objects are detections, darker blue is the background, and red are other "bad" pixels. The area shown corresponds to ~ 0.0025 deg^2.

[4]This rotation will become apparent when we consider calculations of the Asymmetry parameter (see Section 3.3.2).



(a) (b)

Figure 3.3: Example of *(a)* an unmasked, and *(b)* a masked cutout for determining the median and standard deviation in the local background. The cutouts shown here are 100 × 100 kpc in size.

In general, the local background is estimated by applying the bitwise masks (0, 1, 5, 7, 8, 9) to the desired cutout centred at the χ^2 centroid output by a SExtractor run on the combined *ugrizy* images (Golob et al., in prep). From this, we calculate values for the median and standard deviation ($\sigma_{\rm bkg}$) in the local background. These values become important when determining thresholds for image segmentation (see Section 3.1.3). Figure 3.3 shows an example of an unmasked and masked cutout.

3.1.3 Segmentation Using the *watershed* Method

To obtain accurate estimates of the morphological parameters for a galaxy, we must isolate *only* those pixels belonging to the primary galaxy in the cutout, a process called *segmentation*. In this work, we apply the *watershed* segmentation algorithm, which works in the

following way. Each object in the cutout is given a *marker* which tells the algorithm roughly where the local maxima are located; in other words, they provide the initial "seeds" to the watershed algorithm.[5] To define our markers, we search the Golob et al. (in prep) catalog within each 100×100 kpc cutout for objects from both the *UltraDeep* and *Deep* layers. This way, any objects not detected in the χ^2 images of the *UltraDeep* layer, but instead detected in the *Deep* layer, will also be included as markers for the watershed segmentation, making the list of initial seeds more complete.

Once the markers are defined, we create a *threshold image*, which is used to tell the watershed algorithm which pixels to consider when running the segmentation. The threshold image includes all pixels with fluxes above some value defined by the user, which is usually a multiple of the standard deviation in the local background of the image. In short, we multiply the image by -1 and use the *negative* values in the original cutout, along with the threshold image and the object markers, as inputs to the `skimage.morphology.watershed` routine in `Python`. The algorithm begins at the markers (which now more or less correspond to local minima[6]) and defines a *basin* in the area around each marker. The basins are then "flooded with water" until the water (or in this case, the cumulative pixel flux in a given area) from each marker-defined basin meets at what is called a "watershed line." It is this line that helps us to separate neighbouring objects which were detected as distinct objects by the SExtractor run on the χ^2 images in Golob et al. (in prep).

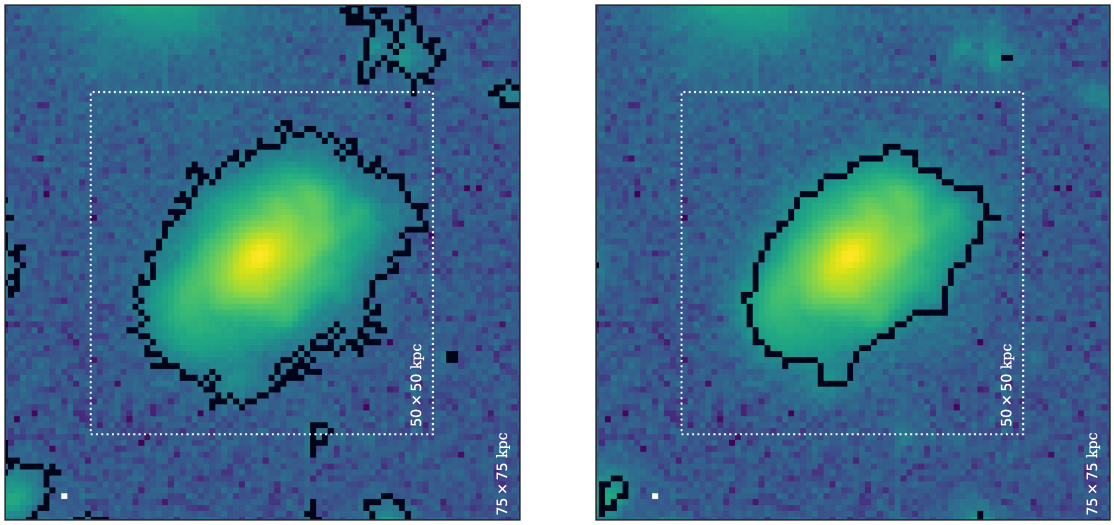
[5] It is extremely important that all objects within the cutout are given proper markers, or else the algorithm will not produce the correct results. For example, missing markers will cause objects to not be segmented and they will otherwise be treated as part of the "background."

[6] Since the markers used for the watershed are the centroids detected using the χ^2 images, the *true* centroids in the r-band may not necessarily correspond to the ones we use in the watershed algorithm. In most cases, the χ^2 centroids serve as good initial guesses to the object centroids, which we update later.

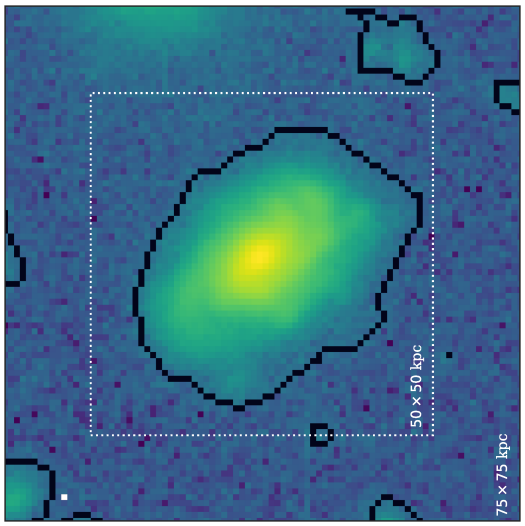
It is important to note that the resulting *segmentation maps* output by the watershed algorithm are highly dependent on the threshold image used. This step of the analysis involved several passes and visual inspection of a few thousand galaxies each time. The methods we ultimately choose are therefore thoroughly tuned to the properties of the HSC r-band images. In short, we use a combination of convolution, thresholding, and visual inspection to determine the optimal *order of operations* and *threshold values* to produce segmentation maps that closely resemble the shape of the primary. Convolution of imaging data involves a user-defined (2-dimensional) *kernel* which is applied to the image by means of an integral transform. This transformation acts to smooth out (or blur) features on low spatial scales. Without convolution, the resulting segmentation maps are filamentary and it is clear that the algorithm is "fitting the noise" (Figure 3.4*(a)*). It is very important that this not occur since we do not want background pixels to be confused with those of the primary galaxy; some morphological parameters can be artificially changed due to such unwanted contamination. Furthermore, convolution of the segmentation maps *after* watershed is performed on a thresholded image produces maps that are too smoothed (Figure 3.4*(b)*) and can either result in background contamination or a loss of flux of the primary segmentation map. We therefore decide to convolve our image first, which acts to smooth out the grainy nature in the outskirts of the galaxy. We then apply a threshold, which is dependent on the local background and use this convolved, thresholded image as the input to our watershed algorithm. We use a threshold where the pixel values in the convolved image are $> 1.5\sigma_{\rm bkg}$.[7]

The segmentation maps are also highly dependent on both the width and the extent of

[7] We visually inspected examples of segmentation maps using different thresholds between $\sim 1.0\sigma_{\rm bkg}$ and $2.5\sigma_{\rm bkg}$ to find the optimal value across most images.



(a) (b)



(c)

Figure 3.4: Resulting segmentation maps of a galaxy for three different cases described in the text. *(a)* Convolution is *not* included in the pipeline. In this case, the edges are not smooth. *(b)* Convolution with a Gaussian kernel is applied *after* the segmentation map of the galaxy is created. Here, the edges of the map are smoother, but they do not represent well where the galaxy light is. *(c)* When the image is first convolved, then thresholded, and finally segmented. This is the map we use for our remaining analyses. The 75×75 kpc cutouts are shown here.

the convolution kernel (i.e., the standard deviation, as well as the size of the array used to house the kernel). We apply a 2-dimensional Gaussian kernel to the images and find that smaller kernels allow for spatially small components to remain in the image, whereas larger kernels smooth out most features on small spatial scales as well as dilate the perimeters of the segmentation maps. Since all of our galaxies are not necessarily the same size, using the same size of kernel across all objects sometimes can produce wildly different results (e.g., a medium-sized kernel on a small galaxy will almost completely wash it out, whereas the same kernel on a large galaxy will produce a more filamentary segmentation map). Using kernel sizes roughly proportional to the size of the galaxy produces slightly better results, however, the best results across all images were produced by a single kernel whose standard deviation corresponded to the average seeing in the r-band images (FWHM $= 0.85''$). Using this kernel allowed us to only smooth down to the limitations of the imaging data themselves. This size gives us a standard deviation of $\sigma_{\rm kernel} \approx 2.15$ pixels. We choose the extent of the kernel array to be 5×5 pixels. Figure 3.4*(c)* shows what the final segmentation map looks like with this kernel and the above threshold.

The results of this segmentation are twofold. First, we recover the segmentation map of the primary galaxy, which allows us to isolate the pixels belonging to the object in question. In this work, we consider the "primary" segmentation map to be that of the central object in the cutout including all objects whose segmentation maps directly border that of the central. We choose this method because some galaxies contain several distinct object detections in the CLAUDS+HSC catalogs (each contributing to their own unique markers in the watershed segmentation step). For example, a merger with a double nucleus

could be seen as two separate objects in the catalog and the resulting segmentation maps would therefore be separate. In this case, we would be excluding any pixels from immediate companions, information that is crucial to the automated detection of mergers. Section 5.6 outlines potential pitfalls to using this method.

Secondly, we recover the segmentation maps of all other detected objects in the cutout. This is useful when considering object masks for some morphological parameters. The SExtractor masks shown in Figure 3.3 are very large and include much of the background. In addition, they overlap and are not segmented. With our segmentation, we are able to assign labels to each distinct object in the cutout and identify exactly which pixels belong to it. This also allows us to update the centroids of the objects and we do so by finding the flux-weighted centre of light in each segment using the `photutils.centroid.centroid_com` routine in `Python`. For the remainder of this work, the "centroids" of our primary galaxies refer to these updated, flux-weighted centroids, which we will denote as $\mathbf{r_c} \equiv (x_c, y_c)$. Figure 3.5 shows several examples of the results from applying the above-described methodology. We exclude galaxies from the primary sample (Column *(4)* in Table 2.2) that have more than 15% of the pixels in their primary segmentation map masked by the addition of the 0°- and 180°-rotated masks created using both the bad bits of Section 3.1.2 and the segmentation maps of other galaxies in the cutout (not including the primary). This is done so that we ensure there is still enough of the primary galaxy available in order to reasonably estimate the morphological parameters. By doing this, we only exclude another 1292 galaxies (2%) from our sample.

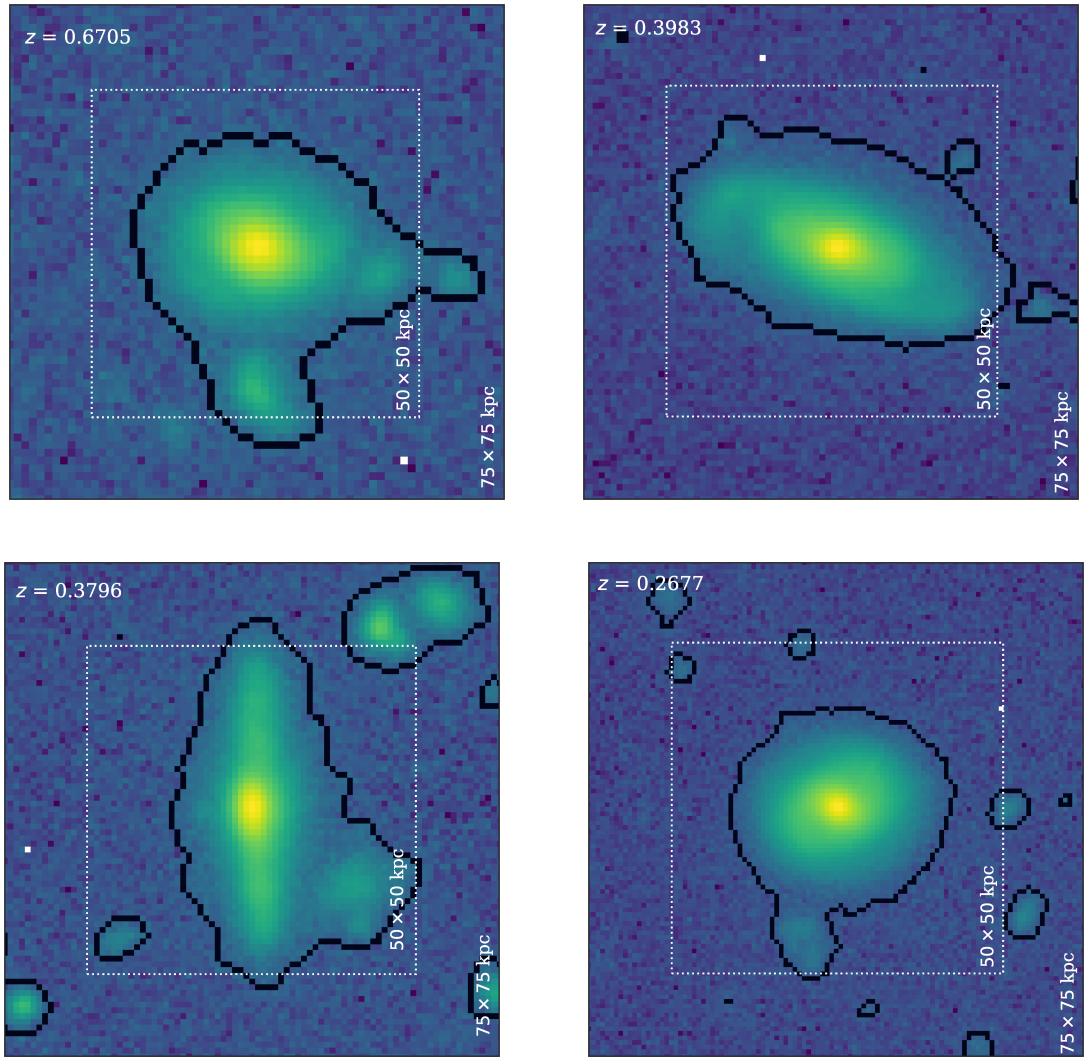


Figure 3.5: Examples of the resulting segmentation maps when we apply our watershed method to the HSC r-band images.

3.2 The Sérsic Index, n

One method of analysing the distribution of light in a galaxy is to fit the galaxy light profile with an analytic expression called a Sérsic profile (Sérsic 1963):

$$I(R) = I_0 \times \exp\left(-b(n) \times \left[(R/R_e)^{1/n} - 1\right]\right), \tag{3.2}$$

where I_0 is the value of the intensity at the centre of the galaxy, R is a variable describing the radial distance from the centre of the galaxy. In the above equation, R_e is the effective (or half-light) radius, which is determined to be the radius for which 50% of the total galaxy light is enclosed. The parameter $b(n)$ is then determined such that R_e is the effective radius. The parameter n is called the Sérsic index which describes the shape of the light profile (see Conselice 2014). The Sérsic fitting method is especially useful for distinguishing between normal galaxies that are bulge- or disk-dominated and is known as a *parametric* measure of galaxy morphology because it makes assumptions about the distribution of light in a galaxy before any measurements are performed (i.e., there is a well-defined centre and the light distribution is radially symmetric). In the case of mergers, however, the centroid of a galaxy's light distribution is not always well-defined and this distribution is not always radial in nature. Because of this, the Sérsic method alone is not a physically meaningful approach to classifying mergers and we must also use other methods to help describe the morphology of galaxies. Nevertheless, we include the Sérsic index as a morphological parameter and discuss how it is determined computationally in Section 3.5.1, where we make sure to convolve the model with an appropriate point spread function while fitting.

3.3 *CAS* Parameters

3.3.1 The Concentration Parameter, C

One question we may ask about a galaxy's light profile is how concentrated the light is in the central regions as opposed to the outskirts. The Concentration parameter C developed by Bershady et al. (2000) takes the ratio of the radii of two circular apertures that contain 80% and 20% of the total galaxy light, respectively:

$$C \equiv 5 \times \log_{10}\left(\frac{r_{80}}{r_{20}}\right). \tag{3.3}$$

The radii in the above equation are usually taken to be *Petrosian radii*. The Petrosian radius is the radius r_p of a circular aperture such that the following expression is satisfied (Petrosian 1976):

$$\eta(r_p) = \frac{I(r_p)}{\langle I(< r_p)\rangle}, \tag{3.4}$$

where $I(r_p)$ is the surface brightness at radius r_p, $\langle I(< r_p)\rangle$ is the average surface brightness within that same radius, and $\eta(r_p)$ is usually taken to be 0.2 (see Conselice 2014).

Operationally, we calculate the Petrosian radius by choosing $\eta(r_p) = 0.2$ and assuming a centroid defined by the flux-weighted centre of light within the primary segmentation map, $\mathbf{r_c} = (x_c, y_c)$, which includes all bordering segments (see Section 3.1.3). We use Brent's Method (Brent 1971; see `scipy.optimize.brentq`) to find the root r_p of the function $\frac{I(r_p)}{\langle I(<r_p)\rangle} - \eta(r_p) = 0$. We also use Brent's Method to calculate the values of r_{20} and r_{80} in Equation 3.3. The value we use for r_{100} (i.e., the radius inside which 100% of the galaxy

light is contained) is $1.5 \times r_p$. We then incrementally increase our aperture size until we reach the radii for which 20% and 80% of the light is accounted for to obtain r_{20} and r_{80}, respectively. The data we use to calculate C are the 50×50 kpc HSC r-band cutouts, masked by a combination of the bitwise masks (Section 3.1.2) and the watershed segments of all galaxies *not* belonging to the primary segmentation map (Section 3.1.3). Figure 3.6 shows an example of a galaxy cutout to illustrate the above method.

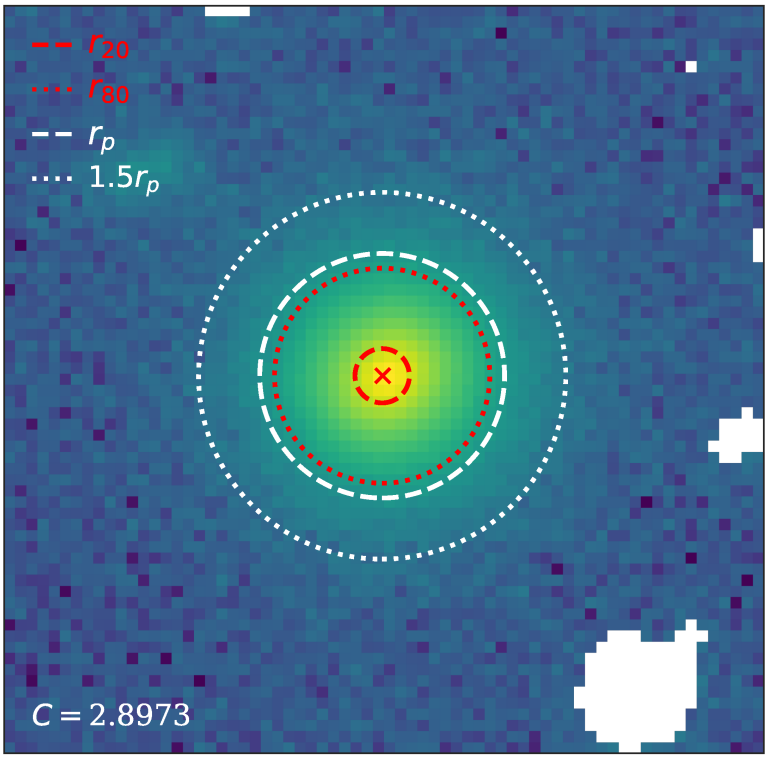


Figure 3.6: Example of a masked 50×50 kpc galaxy cutout illustrating the calculation of the Concentration parameter, C. The centroid of the galaxy $\mathbf{r_c}$ is marked with the red cross. The white dashed and dotted circles denote apertures of radii r_p and $1.5 \times r_p$, respectively. The red dashed and dotted circles denote apertures of radii r_{20} and r_{80}, respectively. The Concentration parameter for this galaxy is $C = 2.9$. A galaxy with a *lower* Concentration would exhibit a *smaller* gap between the r_{20} and r_{80} apertures.

The Concentration parameter has been shown to correlate strongly with the Sérsic index n (Conselice 2014). Furthermore, larger values of C indicate that more of the galaxy light is contained in the central regions and so we would expect galaxies such as ellipticals, lenticulars, and early-type spirals to have larger values of C. To calculate C, we need to assume that the galaxy has a well-defined centre, which is not always the case for merging galaxies. In addition, the Concentration parameter is known to behave poorly under degraded conditions; i.e., decreasing signal-to-noise and surface brightness, which occurs with increasing redshift (see, for example, Graham et al. 2001a,b, 2005).

3.3.2 The Asymmetry Parameter, A

Since not all galaxies are perfectly symmetric, especially those that are merging, it is useful to quantify the degree to which a galaxy is asymmetric. The Asymmetry parameter A was first defined by Abraham et al. (1996) to describe how a galaxy image behaves under a 180° rotation about its centre. Under such a rotation, we would expect highly symmetric galaxies to show minimal deviation from the original image. In contrast, we would expect highly asymmetric galaxies to show a large deviation from the original image if rotated.

The original definition for A was modified by Conselice et al. (2000a) to redefine the method by which the centre of rotation is obtained and to include corrections for high background noise levels. We use a slightly modified definition of the Conselice et al. (2000a)

formalism here (see also Conselice 2014):[8]

$$A = \min\left(\frac{\sum_{i,j}|I^0_{i,j} - I^{180}_{i,j}|}{\sum_{i,j}|I^0_{i,j}|}\right) - B_A. \tag{3.5}$$

In the above equation, the 75×75 kpc original and 180°-rotated images are denoted by $I^0_{i,j}$ and $I^{180}_{i,j}$, respectively.[9] We subtract the rotated image from the original and sum the resulting residual flux over all pixels in a circular aperture of radius $1.5r_p$ and centroid $\mathbf{r_c} = (x_c, y_c)$. We then normalize by the total flux in the original galaxy image. Following the Conselice et al. (2000a) definition, we choose the value for which the first term in Equation 3.5 is minimized. For this, we consider a 5×5 pixel grid of possible centroids around $\mathbf{r_c}$ and calculate the first term in Equation 3.5 for each of them, choosing the smallest value to include in the final calculation.

The second term B_A is defined as follows:

$$B_A = \min\left(\frac{\sum_{k,l}|B^0_{k,l} - B^{180}_{k,l}|}{S_{\text{skybox}}}\right). \tag{3.6}$$

The background image $B^0_{k,l}$ and 180°-rotated background image $B^{180}_{k,l}$ are also defined using the 75×75 kpc cutout[10]. We search the background image for all regions of sky background that are 32 pixels in size.[11] We use Equation 3.6 to calculate the average Asymmetry in all

[8]The difference in our definition arises in the normalization of the background correction factor. Instead of normalizing by the total galaxy flux $\sum_{i,j}|I_{i,j}|$ as in Conselice et al. (2000a), we normalize by the number of pixels (or area) we use to define our background region.

[9]These images are masked by the bad bits of Section 3.1.2 and the segmentation maps of all galaxies in the cutout except for the primary. The masks are rotated by 180° about $\mathbf{r_c}$ and are added to the original (un-rotated) masks.

[10]Here, we define our background images by applying the 0°- and 180°-degree rotated χ^2 object detection masks from Section 3.1.2 to the original image.

[11]If nothing suitable is found, we decrease the size of the box, but make sure the final sky region is not

backgrounds regions, normalizing by the size of the box $S_{\rm skybox}$, and then taking the value which minimizes B_A.

Figure 3.7 shows an example of a galaxy cutout, its 180°-rotated image, and the image resulting from the subtraction of the two. From the Figure, we can see that the Asymmetry parameter may be useful in identifying potentially merging galaxies since mergers often display disturbed, asymmetric morphologies.

3.3.3 The Smoothness Parameter, *S*

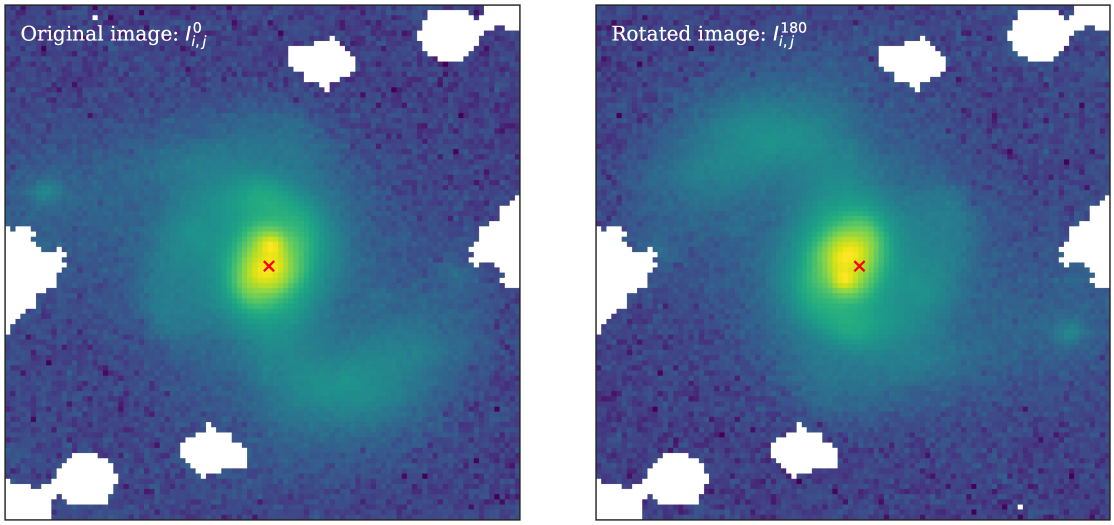
Another common descriptor of galaxy morphology is the Smoothness (sometimes called "clumpiness") parameter S. This parameter is used to describe the fraction of light in a galaxy that belongs to regions of high spatial frequency. In other words, S is a measure of how "smooth" (or "clumpy") a galaxy's light distribution is. In Lotz et al. (2004), the galaxy Smoothness is defined as follows:

$$S = \frac{\sum_{i,j \in A} |I_{i,j} - I^{\sigma}_{i,j}|}{\sum_{i,j \in A} |I_{i,j}|} - B_S. \tag{3.7}$$

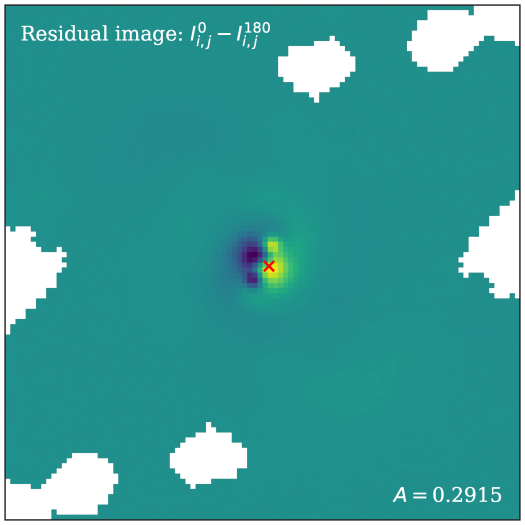
In the above definition, I is our 50×50 kpc galaxy cutout, where $I_{i,j}$ is the intensity of each pixel at position (i, j). The cutout is masked the same way as in our calculation of the Concentration parameter. The galaxy image is then smoothed by a 2D boxcar filter of width $\sigma = \frac{1}{6} r_p$ (as in Hambleton et al. 2011).[12] The smoothed image $I^{\sigma}_{i,j}$ is subtracted from the original image $I_{i,j}$ and is normalized by the total flux of the galaxy to emphasize clumpy

less than 3 pixels. Most sky regions have more than 20 pixels.

[12] Masking is taken into account in the convolution by using a *weight image* based on the convolution of the masked regions with the 2D boxcar filter.



(a) (b)



(c)

Figure 3.7: Example of a 75 × 75 kpc cutout for a merging galaxy with a double nucleus. The masking is the combined 0°- and 180°-rotated masks, and the red cross denotes the galaxy centroid $\mathbf{r_c}$ computed on the original image. *(a)* The original image $I^0_{i,j}$. *(b)* The 180°-rotated image $I^{180}_{i,j}$. *(c)* The residual image obtained by subtracting the previous two images ($I^0_{i,j} - I^{180}_{i,j}$). In this case, the value for the Asymmetry is $A = 0.3$. The double nucleus is very clear in panel *(c)*.

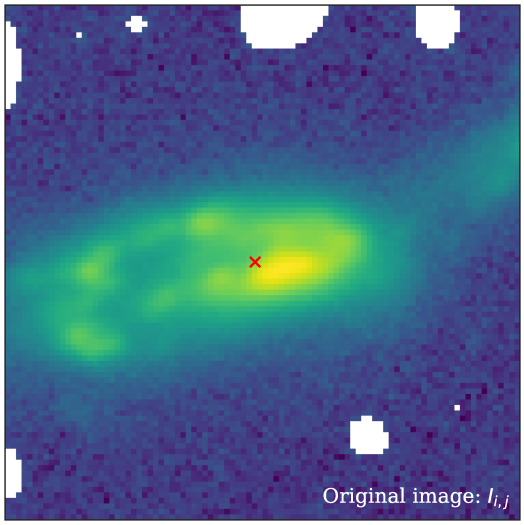
structures. The summations are performed over all pixels in a circular aperture A, which has a radius of $1.5r_p$ and centroid $\mathbf{r_c}$.

The second term of the above definition B_S is the average Smoothness in the background. We calculate B_S as follows:

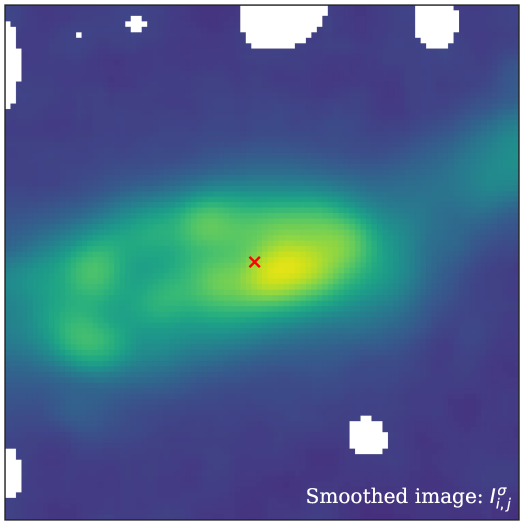
$$B_S = \frac{\sum_{k,l} |B_{k,l} - B^{\sigma}_{k,l}|}{A_B}. \tag{3.8}$$

Here, we isolate the background pixels in the 50×50 kpc cutout by applying the full χ^2 object detection masks (i.e., including bitwise value 5) to the original image. This gives us a local background image $B_{k,l}$ to which we apply the same 2D boxcar filter convolution as above to get $B^{\sigma}_{k,l}$. In this case, we normalize by the number of pixels A_B belonging to the background image.

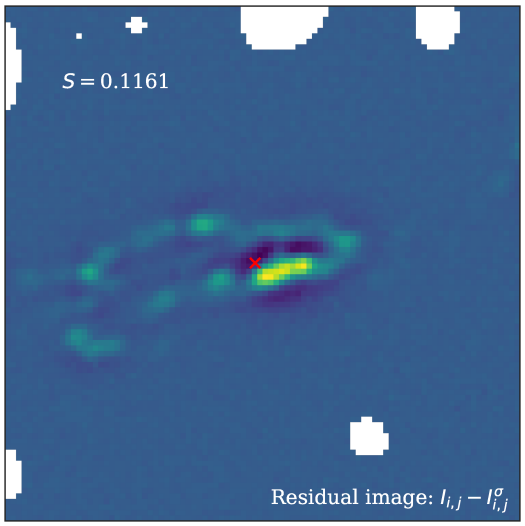
Galaxies with a high fraction of their light in high spatial frequencies will have larger values of S. For example, a galaxy with intense, clumpy star formation will have a larger S parameter than, say, an elliptical galaxy with a smoother light profile (provided that the rest-frame wavelength of the image is appropriately chosen so as to probe these features). In the case of merging galaxies, there can be starburst activity at certain stages in the merger scenario. The Smoothness parameter would therefore be useful in identifying clumpy star formation in these systems. To illustrate this, Figure 3.8 shows a galaxy image, the same image smoothed by a 2D boxcar filter, and the image resulting from the subtraction of the two.



(a)



(b)



(c)

Figure 3.8: Example of a 75×75 kpc cutout for a merging galaxy at $z = 0.33$ with clumpy star formation. The masking is the the bad bitwise values and the segmentation maps of all other objects except the primary. The red cross denotes the galaxy centroid $\mathbf{r_c}$ computed on the original image. *(a)* The original image $I^0_{i,j}$. *(b)* The smoothed image $I^{\sigma}_{i,j}$. *(c)* The residual image obtained by subtracting the previous two images $(I_{i,j} - I^{\sigma}_{i,j})$. In this case, the value for the Smoothness is $S = 0.12$. The star-forming clumps are very clear in panel *(c)*.

3.4 Gini and M_{20} Statistics

3.4.1 The Gini Parameter, G

First applied to galaxy images by Abraham et al. (2003), the Gini coefficient G (Lorenz 1905) is a commonly used statistic in econometrics for characterizing the distribution of wealth in a population. The coefficient itself is derived from a visual representation of the distribution, called the Lorenz curve (see Figure 3.9). In the context of galaxy images, the distribution of wealth within a population is replaced by the distribution of the galaxy's light across the individual pixels in the image. Therefore, each pixel represents a single member of the total population.

Mathematically, the Lorenz curve is defined as

$$L(p) = \frac{1}{\bar{X}} \int_0^p F^{-1}(u) du, \tag{3.9}$$

where $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$ is the mean value of the pixel flux over the n pixels in the image, p is the percentage of the total number of pixels in the image (where the pixels are in order of increasing pixel flux), and the integrand is the inverse of the cumulative distribution function of the pixel flux values (X_i's).

The Lorenz curve for a galaxy whose light is equally distributed across all pixels is shown by the dotted 1:1 line in Figure 3.9 (in other words, $L(p) = p$). Any unequal distribution of light would be seen as a deviation below this line of equality; the larger the deviation, the more unequal the distribution of light across the pixels in the image.

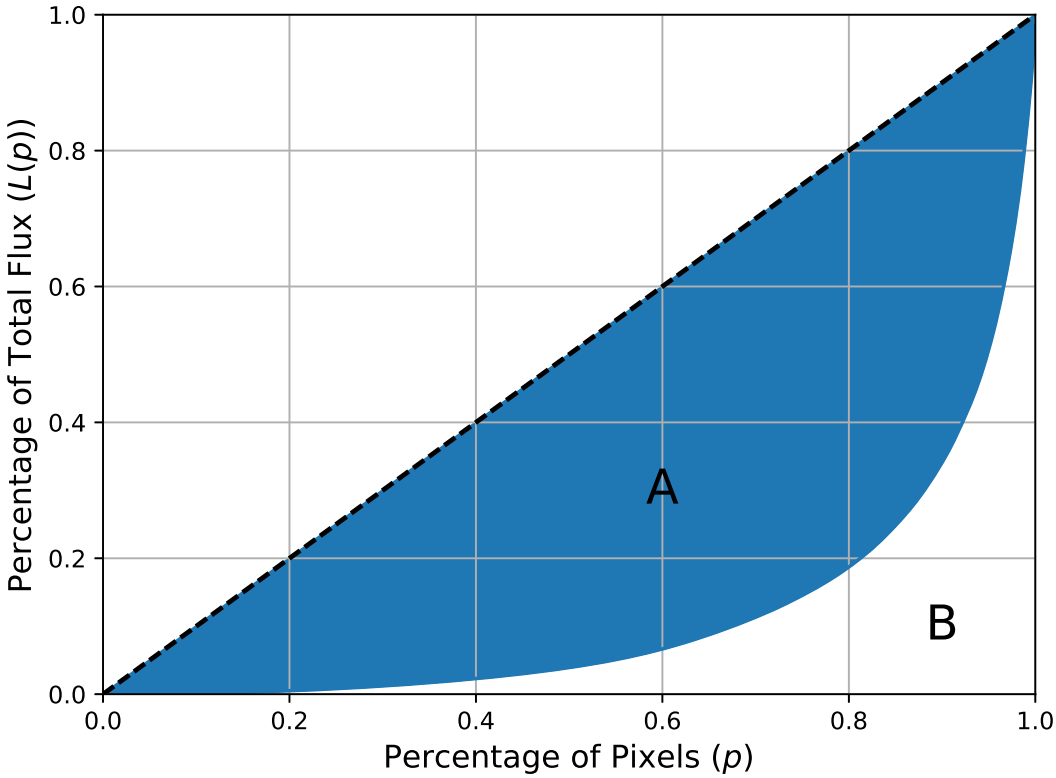


Figure 3.9: Lorenz curve for a galaxy in the CLAUDS+HSC catalog. The line of equality is shown as the dotted black line and the area between the line of equality and the Lorenz curve (lower boundary) for this particular galaxy is shaded in blue (also labelled A). The total area under the line of equality (1/2) is equivalent to adding the two areas labelled A and B.

Geometrically, the Gini coefficient G is the ratio between two areas: the area between the line of equality and the Lorenz curve (labelled A in Figure 3.9), and the total area under the line of equality (A + B = 1/2 in Figure 3.9). Therefore, $G = \mathrm{A}/(\mathrm{A}+\mathrm{B})$. Values for the Gini coefficient range between 0 and 1. A G of 0 indicates a perfectly equal distribution of galaxy light across the pixels in the image (i.e., Lorenz curve is the line of equality). A G of 1 indicates that all of the galaxy light is concentrated in a single pixel (i.e., Lorenz curve

is zero for all p except for $p = 1$ where $L(p) = 1$).[13]

Operationally, the Gini coefficient can be calculated by sorting the pixels in increasing order according to their pixel flux values (X_i's) and performing the following summation (Glasser 1962; Abraham et al. 2003):

$$G = \frac{1}{\bar{X} n(n-1)} \sum_{i=1}^{n} (2i - n - 1) X_i, \qquad \text{(for } n > 2\text{)}. \tag{3.10}$$

Lotz et al. (2004) modified the above definition for the Gini coefficient to account for the effects of cosmological surface brightness dimming (see Section 5.2) and decreasing signal-to-noise with increasing redshift. Therefore, in contrast to the Abraham et al. (2003) definition, the Lotz et al. (2004) definition allows for a direct comparison between low- and high-redshift galaxies. The Gini coefficient of Lotz et al. (2004) differs from that of Abraham et al. (2003) in that the absolute values of the pixel fluxes are used when performing summations:

$$G = \frac{1}{|\bar{X}| n(n-1)} \sum_{i=1}^{n} (2i - n - 1) |X_i|, \qquad \text{(for } n > 2\text{)}, \tag{3.11}$$

where, $|\bar{X}| = \frac{1}{n} \sum_{i=1}^{n} |X_i|$. Computationally, we apply Equation 3.11 to the primary segmentation map (including bordering segments) of each galaxy using the `photutils.gini` routine in `Python`.

In the local Universe, the Gini coefficient has been shown to correlate with the central Concentration parameter C with some scatter (Abraham et al. 2003). As a result, G can

[13] In the context of econometrics, the $G = 1$ case corresponds to a single person possessing all of the wealth of the society, and the rest of the population possessing no wealth.

be interpreted to first order as a generalized measure of the concentration of galaxy light. There are several advantages to using G as opposed to C in quantifying the distribution of light in merging galaxies. First, no aperture photometry is required for the calculation of G; measuring G requires only the pixel fluxes and no spatial information. Furthermore, in measuring G, it is not assumed that the galaxy has a well-defined centre, as is the case for C. Therefore, G can be measured for galaxies with highly asymmetric light profiles or even double nuclei, making it an attractive morphological indicator for merging systems.

3.4.2 The Moment of Light, M_{20}

Another non-parametric measure of galaxy morphology was developed by Lotz et al. (2004) to describe the spatial distribution of light for the brightest regions in a galaxy. This parameter, called M_{20}, employs the concept of the second-order moment of the light. In other words, M_{20} measures the spatial distribution of pixel fluxes within an image, whereby the flux in a given pixel is weighted by the square[14] of its distance to the galaxy centre. Given a segmentation map of a galaxy, the total second-order moment of the light is given by:

$$M_{\rm tot} = \sum_{i=1}^{n} M_i = \sum_{i=1}^{n} f_i \left[(x_i - x_c)^2 + (y_i - y_c)^2 \right], \tag{3.12}$$

where f_i is the flux in pixel i, (x_i, y_i) is the position of pixel i, and (x_c, y_c) is the position of the centre of the galaxy as determined iteratively to be the coordinates for which $M_{\rm tot}$ is minimized. This centroid is, in fact, exactly the same as the $\mathbf{r_c}$ determined during the segmentation step (Section 3.1.3).

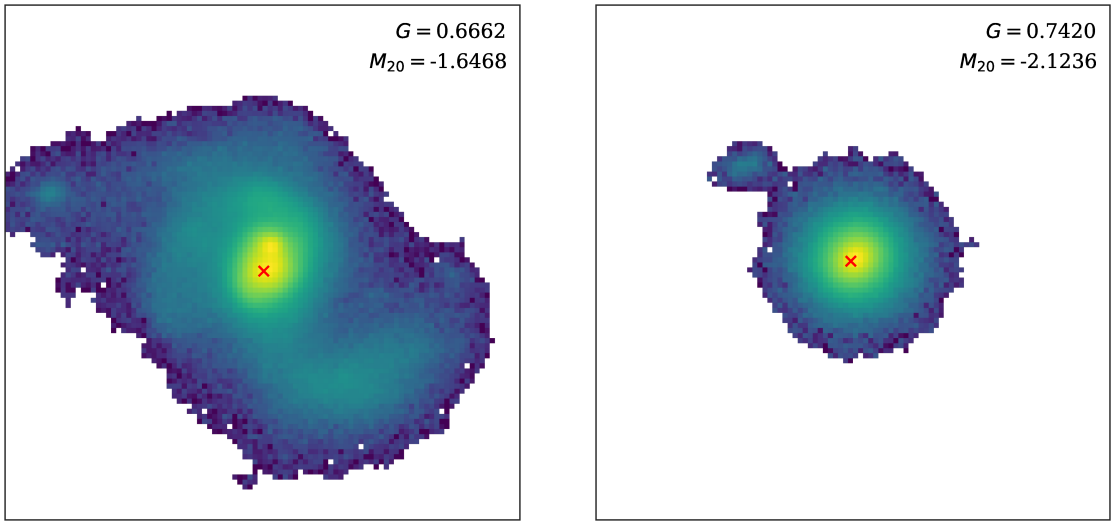
[14] Hence, the *second-order* moment.

The definition of the M_{20} parameter as described in Lotz et al. (2004) is as follows:

$$M_{20} = \log_{10}\left(\frac{\sum_i M_i}{M_{\rm tot}}\right), \qquad \text{while } \sum_i f_i < 0.2 f_{\rm tot}. \tag{3.13}$$

The above definition states that given a list of pixel fluxes in order of increasing brightness, only the brightest pixels are taken such that, together, they account for 20% of the total flux $f_{\rm tot}$ in the segmented galaxy image. The normalization by $M_{\rm tot}$ above removes any dependence of the M_{20} statistic on the total flux or galaxy size. Since the ratio inside the logarithm is always less than 1, M_{20} is always negative. The closer the brightest regions to the centre of the galaxy, the more negative M_{20}. Therefore, since we would expect galaxies with disturbed morphologies to have their brightest regions further away from the centre, their M_{20} parameters would be less negative compared to normal (non-interacting) galaxies.

Computationally, we determine M_{20} by measuring the image moments on both the original and thresholded (20% brightest) imaging data within the segmentation map of the primary galaxy using the `skimage.meaure.moments_central` routine in `Python`. These moments are used directly in Equation 3.13 to derive M_{20}. Figure 3.10 shows an example of a normal and merging galaxy with their corresponding segmentation maps, centroids, and values for G and M_{20}.



(a) (b)

Figure 3.10: Examples of the primary segmentation maps for *(a)* a merger, and *(b)* a non-merger. The cutouts are 75×75 kpc in extent. The red crosses denote the galaxy centroids $\mathbf{r_c} = (x_c, y_c)$ and the G and M_{20} values for each case are shown. As expected, the G value is lower for the merger since its flux is more equally distributed across the pixels in the primary segmentation map. The M_{20} value for the merger is also less negative because there are parts of its flux at higher distances from the centroid.

3.5 Profile Fitting and Residual Image Statistics

3.5.1 Sérsic Modelling and Residual Images

In a study performed by Hoyos et al. (2012), structural parameters of *residual images* were used to select mergers. The residual image of a particular galaxy is created by fitting the target galaxy with a smooth Sérsic profile (Equation 3.2) and subtracting this fit from the original image. The authors claimed that by using the morphological parameters of the residual images (as opposed to those of the original images), merger samples of comparable

or better statistical quality can be obtained and, in addition, more minor mergers can be selected. We choose to include three such residual image statistics, which we discuss in Section 3.5.2.

To obtain our residual images, we fit each galaxy with a 2-dimensional Sérsic profile using a Levenberg-Marquardt non-linear least squares minimization algorithm (`astropy.modelling.fitting.LevMarLSQFitter`). We initialize the model with a 2-dimensional Sérsic profile convolved with the point spread function (PSF) of the corresponding HSC r-band patch image.[15] The PSFs were computed as part of the HSC software pipeline `hscPipe` using the `PSFEx` software (Bertin 2011). More details on their computation can be found in Bosch et al. (2018). The galaxy models need to be convolved with a PSF so as to convert them to the same observational conditions as the data. To perform the fit, we use the data in the primary segmentation map with the standard initial parameters for all galaxies outlined in Table 3.1. After performing the fit, we subtract the model from the data to obtain the residual image. Figure 3.11 shows examples of elliptical, spiral, and merging galaxies along with their Sérsic models, and the resulting model-subtracted residual images.

3.5.2 Residual Image Statistics (RFF, A_{resid}, and S_{resid})

In an attempt to quantify what we can see visually in the model-subtracted residual images, Hoyos et al. (2011, 2012) defined the *Residual Flux Fraction*, RFF. This structural parameter measures the fraction of flux in the residual image that cannot be attributed to

[15] The HSC r-band PSFs were provided by A. Goulding of the HSC Collaboration.

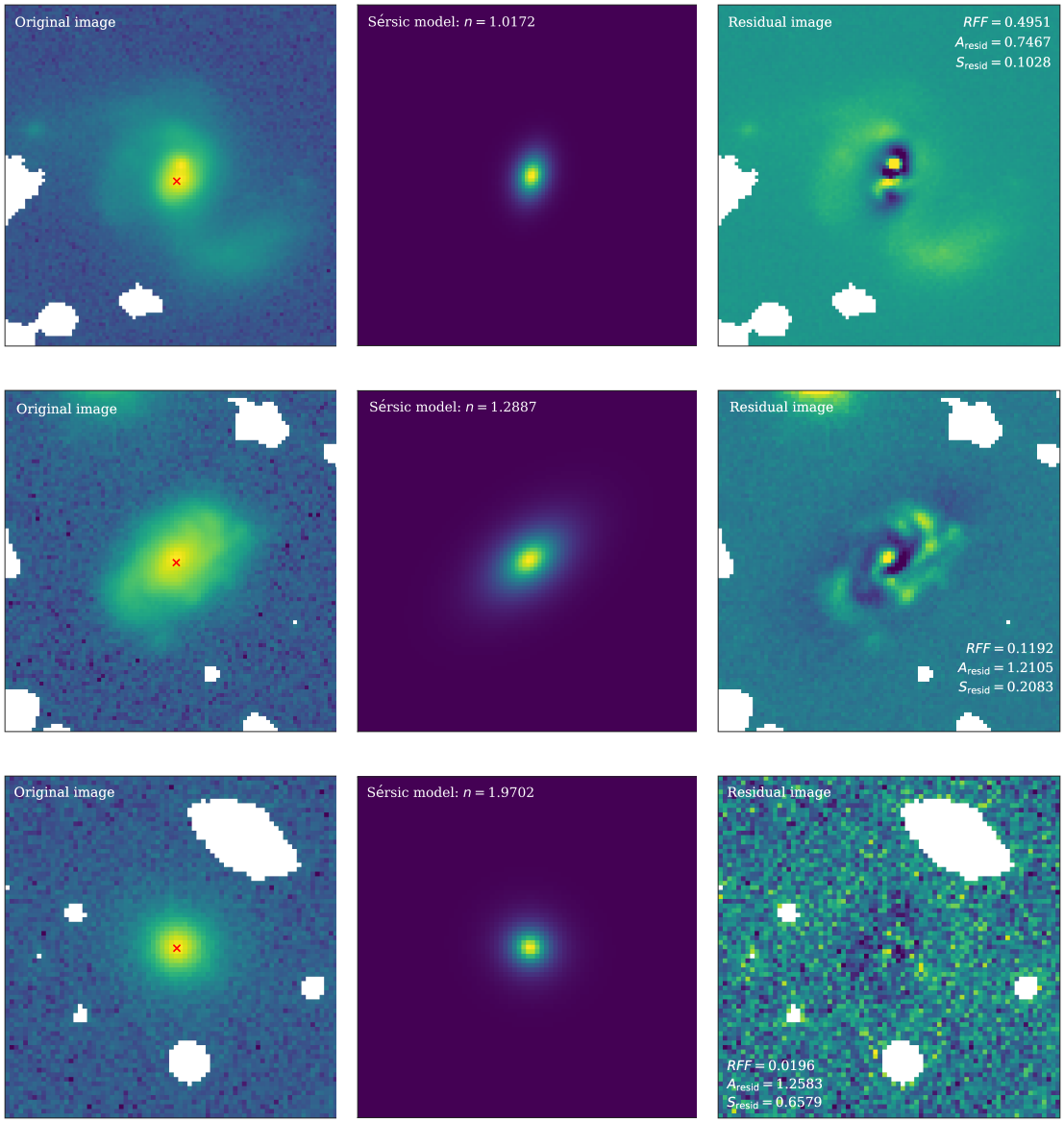


Figure 3.11: Examples of 75 × 75 kpc galaxy cutouts for a merger (first row), spiral non-merger (second row), and an elliptical non-merger (third row). The columns show the original image, the Sérsic model fit, and the model-subtracted residuals for each galaxy, respectively. Image stretches are chosen to emphasize the important features of each plot. The values for n, RFF, $A_{\rm resid}$ and $S_{\rm resid}$ are reported for each galaxy.

Table 3.1: Overview of the standard initial parameters used when calculating the 2-dimensional Sérsic profiles for our galaxies.

Parameter	Initial Value
Amplitude (I_0)	Max. pixel value in primary segmentation map
Effective radius (R_e)	$\sqrt{A_{\rm primary}/\pi}$, where $A_{\rm primary}$ is the # of pixels in the primary segmentation map[a]
Sérsic index (n)	2.5
Centroid (x_0, y_0)	Flux-weighted centre of mass $\mathbf{r_c}$ (i.e., M_{20} centroid from Sections 3.1.3 and 3.4.2)
Ellipticity (e)	0.5
Rotation angle (θ)	$\pi/2$

[a] Here, we are assuming that to first-order, the area of the primary segmentation map can be approximated with that of a circular region of area $A_{\rm primary} \approx \pi R_e^2$.

background noise fluctuations. It is calculated as follows:

$$RFF = \frac{\sum_{i,j\in A}|I_{i,j} - I_{i,j}^{\rm Sérsic}| - 0.8 \times \sum_{i,j\in A}\sigma_{i,j}^{\rm bkg}}{\sum_{i,j\in A} I_{i,j}^{\rm Sérsic}}, \tag{3.14}$$

where $I_{i,j}$ is the original image, $I_{i,j}^{\rm Sérsic}$ is the Sérsic model, A represents the area defined by the segmentation map of the galaxy in question, and $\sigma_{i,j}^{\rm bkg}$ is the standard deviation in the "local" background of the 50×50 kpc cutout.

Hoyos et al. (2012) also showed that if instead we use the residual images in Equations 3.5 and 3.7 for the Asymmetry and Smoothness parameters, we can emphasize merger signatures such as faint tidal tails and bright clumps since we would be removing the majority of light coming from the central bulge. Using the same methods as in Sections 3.3.2 and 3.3.3, we can calculate the Asymmetry and Smoothness parameters on the residual images (denoted as $A_{\rm resid}$ and $S_{\rm resid}$, respectively). In this case, however, we use the Petrosian radius and

centroid of the *original* galaxy image and only change the input "image" to be the model-subtracted residuals.[16]

Figure 3.11 also reports the values of n, RFF, $A_{\rm resid}$, and $S_{\rm resid}$ for each of the example galaxies. Notice that there is almost no residual flux in the case of the elliptical galaxy. It is dominated by a galactic bulge and so its light profile is well-described by a single-component, 2-dimensional Sérsic model. As expected, the elliptical galaxy has the lowest value for the RFF, whereas the spiral galaxy and merger both show higher values. In the case of the spiral galaxy, the Sérsic profile does not account for the complex spiral structure further out in the disk. As a result, these features are not removed by the model subtraction. If the spiral galaxy is relatively un-disturbed, then we would expect the spiral structure to be more or less symmetric upon rotation, a detail which would most likely manifest itself in the value of $A_{\rm resid}$. The residual image of the merger is very asymmetric and because of this, $A_{\rm resid}$ could, in theory, be a useful way of distinguishing a spiral galaxy from a merger in the residual images. In all, when comparing the residual image of the merger to the images of the non-mergers, especially the elliptical galaxy, it becomes apparent why residual image statistics could be very useful for identifying mergers.

3.6 Using Non-parametric Measures to Identify Mergers

A common practice to identify mergers using non-parametric methods is to choose two statistics and form a 2-dimensional parameter space (or plane) with each statistic represented by one axis. In this plane, each data point is an ordered pair containing the two morphological

[16] One could imagine that calculating r_p and the flux-weighted centroid on the residual image would lead to unphysical results in the case where the model-subtracted residuals yield mostly noise.

parameters for a single galaxy. Galaxies with different morphological types are thought to occupy different regions within this space and can therefore be separated from one another on the basis of their morphological parameters. Given a sample containing both normal and merging galaxies, a division within the plane can be made if the morphological types of the galaxies are already known (by using visual identification). After a criterion for division is created using a labelled sample of galaxies, the method can be applied to a sample for which the morphological types are unknown and be used to identify merging galaxies without the need to visually classify each one individually.

3.6.1 Projections of the *CAS* Space

One method for identifying merging galaxies is to project a 3-dimensional space defined by the C, A, and S parameters onto a 2-dimensional plane using only two of the parameters. All projections can be used, however, some are more meaningful than others when looking to identify mergers. For example, Conselice (2003) uses the Smoothness S and Asymmetry A parameters. In the $A - S$ space, normal galaxies are well fit by the following linear trend (Conselice 2003):

$$A_{\rm fit}(R) = (0.35 \pm 0.03)S(R) + (0.02 \pm 0.01), \tag{3.15}$$

where the structural parameters are computed from the R-band galaxy images. To identify mergers in this plane, we find galaxies that satisfy either of the following criteria (Conselice 2003; Cotini et al. 2013):

$$A > A_{\rm fit} + 3\sigma \quad \text{or} \quad A > 0.35, \tag{3.16}$$

where the mean dispersion for the linear fit is $\sigma = 0.035$. Therefore, in this formalism, mergers are considered to be galaxies with either very high Asymmetry values, or Asymmetry values that deviate substantially from the correlation in Equation 3.15.

3.6.2 The $G - M_{20}$ Plane

Another commonly used 2-dimensional parameter space of morphological indicators is the $G - M_{20}$ plane of Lotz et al. (2004). In their pilot study, Lotz et al. (2004) plotted the Gini and M_{20} values for samples of both normal local galaxies (from Frei et al. 1996) and ultraluminous infrared galaxies (ULIRGs, Borne et al. 2000).[17] They found that the normal galaxies in their sample followed a linear trend in the $G - M_{20}$ plane. In this space, elliptical and lenticular galaxies tend to have high G and low M_{20} values, while late-type spirals and irregulars show lower values for G and higher values for M_{20} (see Figure 9 in Lotz et al. 2004). They also found that the ULIRG population was separated from the normal galaxies in this plane.

In a later work, Lotz et al. (2008a) proposed a division between mergers and normal galaxies using a sample of galaxies at higher redshifts ($0.2 < z < 1.2$). They claimed that most mergers were captured using the following criterion:

$$G > -0.14M_{20} + 0.33. \tag{3.17}$$

[17] ULIRGs can be used as a rough proxy for mergers because they often show signatures of recent or ongoing merger activity (see, for example, Wu et al. 1998; Borne et al. 2000; Conselice et al. 2000b).

Chapter 4

Random Forest Classifiers

In this Chapter, we introduce the theory behind supervised machine learning, specifically Random Forest Classifiers, and how they can be applied in practice to our sample of galaxies in the CLAUDS+HSC survey. We discuss the simple example of a single Decision Tree Classifier, which can be thought of as one building block for the more complex Random Forest. We explain how we choose a sample with which to train our Random Forest and describe the data pre-processing and training methodology. We then explore how the features calculated in Chapter 3 can be supplied to our classifier, producing probabilities that the given galaxies are mergers. We assess the performance of our classifier by defining several commonly-used statistics. Finally, we compare our classifier to the 1- and 2-dimensional approaches of Section 3.6 as a motivation for a higher-dimensional approach.

4.1 Introduction to Supervised Machine Learning

One of the most common problems that can be solved in the framework of supervised machine learning is that of *classification*. Classification problems take advantage of the idea that each object in a dataset can be associated with a specific *class* of objects and their association with any particular class is dependent on a list of descriptive *features*. In this work, we wish to separate galaxies into one of two groups, *merging* and *non-merging*, based on their visual morphology. Using the features introduced in Chapter 3 which describe the morphologies our galaxies, we can train a computer to automatically classify each galaxy in our sample.

All supervised classification algorithms follow the same general "recipe" (see, for example, Murphy 2012; Ivezić et al. 2014). In short, a sample of objects for which the classes (sometimes called *labels*) are already known is identified and a feature space motivated by the context of the problem is defined. This sample is called the *training set* and it is used to build a model for an automated classifier. Once a model has been built (a process sometimes referred to as "training a classifier"), a *test set* of objects whose features have been calculated, but whose labels are unknown to the computer, are fed through the classifier and assigned to one of the predefined groups. In some cases, the classification algorithm will output the *probability* that a particular object belongs to one of the predefined groups. The results of the automated classification are dependent on the list of features used to describe each object, as well as the completeness of the training set; in other words, how well the objects in the training set represent each of the classes.

Mergers have been identified using supervised classification techniques such as random forests (Freeman et al. 2013; Goulding et al. 2018), support vector machines (Cibinel et al. 2015), artificial neural networks (ANN, Naim et al. 1997), and convolutional neural networks (CNN, Ackermann et al. 2018). In this work, we apply a *Random Forest Classifier (RFC)* to our sample of galaxies in the CLAUDS+HSC dataset.

4.2 Random Forests

4.2.1 Random Forest Classifier Theory

Random Forests are an extension to the simpler, base algorithm called a *Decision Tree Classifier*. Decision Tree Classifiers work in a top-down hierarchical fashion, where all objects in the sample begin in the same, mixed pool. The model is first built using a labelled training set. In our case, we use a training set of visually-inspected galaxies whose morphological classifications (merger or non-merger) are already known. At each step in the algorithm, the sample experiences a binary split along the axis of a single feature. Both the feature and feature value is chosen by the algorithm at each step so as to minimize a statistic called "Gini".[1] The Gini statistic therefore measures the "quality" of a split and is defined as:

$$G = \sum_{i}^{k} p_i(1 - p_i), \tag{4.1}$$

where p_i denotes the probability that a point with class i will be found in the dataset.

[1]The definition for this statistic is, in fact, that of the Gini used in econometrics from Section 3.4.1, however, it should *not* be confused with the morphological parameter Gini G used to describe the distribution of flux across the pixels in a galaxy image.

The sample continues to be split until a certain threshold is met and the classification assigned to any terminal (or *leaf*) node is that with the higher relative fraction of objects from the training set. In other words, if for a particular leaf node more than 50% of the objects are visually classified as mergers, then the classification of that node would be "merger." Splitting the tree until each object occupies its own leaf node is not only computationally expensive, but will also lead to overfitting of the training set and result in a less accurate classifier when it is applied to a test set. To avoid this, we can reduce the complexity (i.e., the depth of the tree) by requiring, for example, that each leaf node contain only a certain number or class of objects.

Another method used to reduce model complexity is that of a Random Forest. In this case, we train our classifier on several labelled training sets instead of just one, thereby producing a set of multiple decision trees (hence "forest"). To obtain the final classification, we would take an average of the results from all Decision Trees in the forest. The training sets for each individual tree in the forest are taken to be bootstrapped (i.e., randomly sampled with replacement) subsets of a parent training set. Furthermore, at each step in the building of the trees, a randomly selected subsample of features is used to make the split (hence "random"). The number of features used at each split is defined by the user and is usually kept small compared to the total number of features in order to reduce overfitting (Ivezić et al. 2014).

In Figure 4.1, we show an example of a Decision Tree trained using galaxies in the CLAUDS+HSC dataset. For example, if using 1000 Decision Trees to construct a Random Forest, we would need to average the results from 1000 trees similar to the one shown. Due

to this, the *classification* for a particular object determined by a Random Forest is the *probability* that it belongs to a particular class. In our case, the RFC would output the probability that a galaxy is a merger, $P_{\rm merge} \in [0, 1]$. As an example, if using 1000 Decision Trees and the resulting $P_{\rm merge}$ for a galaxy is 0.75, this means that 750 of the 1000 Decision Trees classified that galaxy as a merger.

4.2.2 Engineering a Training Set

The first step in applying a Random Forest Classifier to the CLAUDS+HSC dataset is to identify a *clean* sample of galaxies with which to train the algorithm. It is very important to "feed" our classifier galaxies for which the visual classifications are known with close to 100% accuracy. An algorithm supplied with uncertain classifications will only "confuse" it and not give reliable results in return.

We decide to calculate our merger fraction evolution from $0.25 \leq z \leq 1.0$ by splitting our data into *five* equally spaced redshift bins:

$$\texttt{zbin1} : 0.25 \leq z < 0.4,$$

$$\texttt{zbin2} : \ 0.4 \leq z < 0.55,$$

$$\texttt{zbin3} : 0.55 \leq z < 0.7,$$

$$\texttt{zbin4} : \ 0.7 \leq z < 0.85, \text{ and}$$

$$\texttt{zbin5} : 0.85 \leq z \leq 1.0.$$

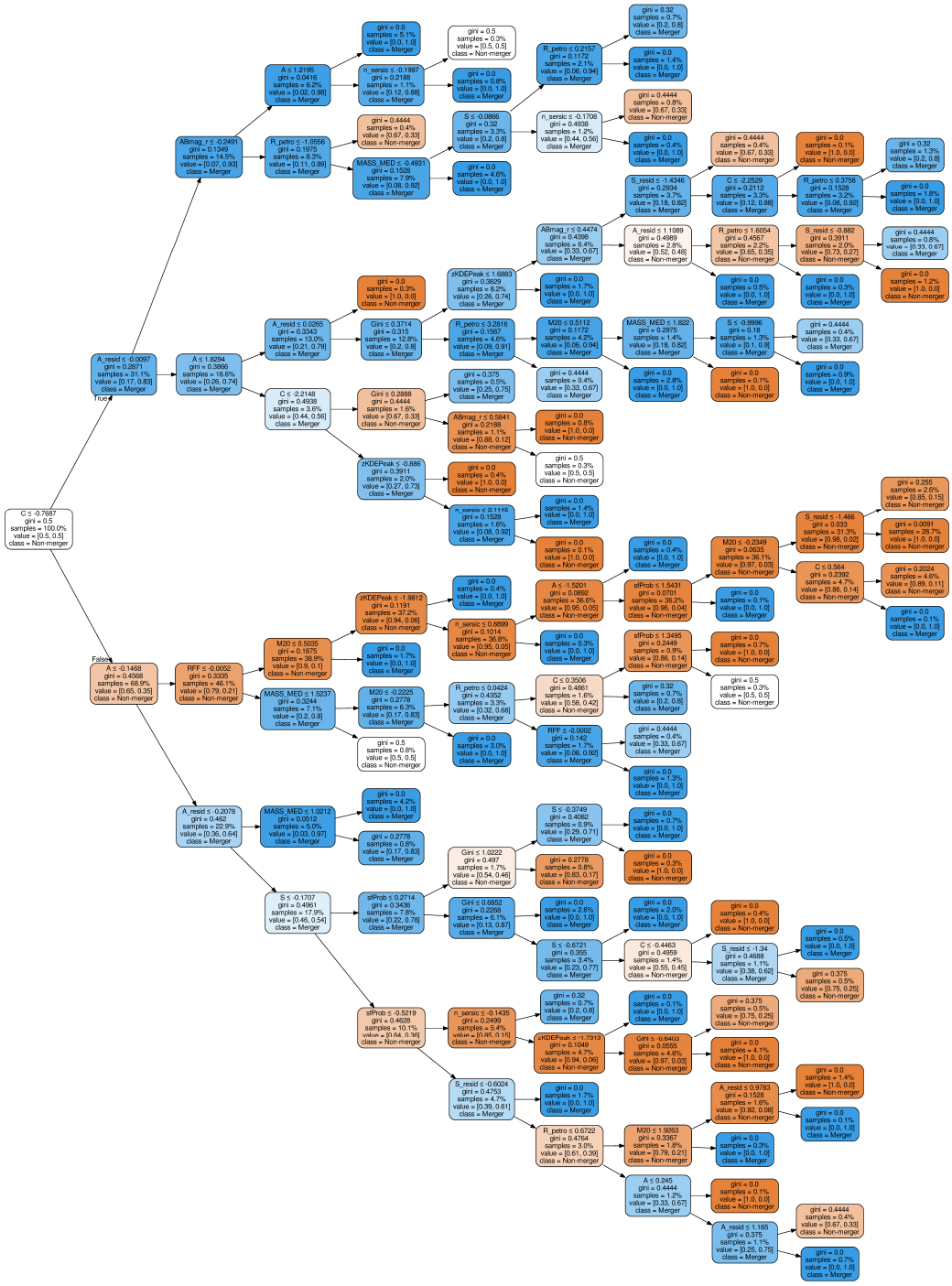


Figure 4.1: Example of a tree resulting from training a Decision Tree Clasifier on a subsample of the CLAUDS+HSC dataset. The colour of each node represents the majority class (i.e., blue for $> 50\%$ mergers and orange for $> 50\%$ non-mergers). The darker the colour, the more pure the class of each node. This tree respresents the process described in Section 4.2.1 of the text.

We visually inspected several thousand galaxies across the five redshift bins in the HSC *UltraDeep* r-band images to populate our training set with visually unambiguous mergers and non-mergers.[2] The visual classifications used in this work were performed by the author (Nathalie C. M. Thibert). To minimize the bias associated with a single human annotator, visual classifications performed by two citizens on a subset of the data on two separate occasions (Mahassen-Hawraa El-Sayegh during the summer of 2017, and Robert Thibert during the summer of 2018) were used to calibrate the responses of the main annotator. We use the `ds9`[3] software product for visual classification, which allows the user to vary the image stretch and emphasize morphological abnormalities.

To maintain the robustness of our classifier to identifying mergers out to the highest redshifts in our sample, we make sure to include roughly equal numbers of galaxies ($\sim 15\%$ mergers and $\sim 85\%$ non-mergers) across each redshift bin. We also make sure to include *both* spiral-like (face- and edge-on) and elliptical non-mergers so as to probe the full parameter space of non-merging galaxies. In particular, we found that including only elliptical galaxies in the training set biased the RFC to believe that non-mergers were only those galaxies which are bright, symmetric, elliptical in shape, and possess low star-formation activity. Everything else regardless of visual classification, especially the spiral population, was considered to be a merger. In total, we identify 918 galaxies (136 mergers, and 782 non-mergers) to use in training our Random Forest Classifier. Table 4.1 gives an overview of the number counts of galaxies found in each bin, separated by their visual classification.

[2] Patch images used for visual identification were downloaded from the CANFAR pages under the following directory: `/clauds/coupon/s16a_udeep_deep_depth.ext_v1.0/deepCoadd/HSC-R/`

[3] http://ds9.si.edu/site/Home.html

Table 4.1: Overview of the visually identified galaxies used to train our Random Forest.

Visual Classification	`zbin1` $[0.25, 0.4)$	`zbin2` $[0.4, 0.55)$	`zbin3` $[0.55, 0.7)$	`zbin4` $[0.7, 0.85)$	`zbin5` $[0.85, 1.0]$
Mergers	28	21	29	30	28
Non-mergers (total)	192	141	134	173	142
Star-forming non-mergers[a]	98	53	57	69	91
Quiescent non-mergers[b]	94	88	77	104	51
Total galaxies:	220	162	163	203	170

[a] Visual non-mergers for which the Golob et al. (in prep) star-forming probability $P_{\rm sf} \geq 0.5$. It is important to note that not all of the galaxies in this group will show spiral or disk-like structure.

[b] Visual non-mergers for which the Golob et al. (in prep) star-forming probability $P_{\rm sf} < 0.5$.

Figure 4.2 shows the HSC r-band magnitudes of the training set galaxies as a function of their photometric redshifts. The (normalized) redshift distributions across both samples are similar, whereas an underabundance of visual mergers at faint magnitudes ($r_{\rm AB} \gtrsim 22.5$ mag) is evident. This results from the fact that we were unable to obtain unambiguous visual classifications closer to the magnitude limit of our sample. See Section 5.6 for a discussion on how this may affect our results. We present the 50×50 kpc HSC r-band cutouts for each galaxy in our training set in Appendix A.

4.2.3 Pre-Processing and Training Methodology

In this Section, we outline the methodology used to train our Random Forest Classifier using the training set introduced in Section 4.2.2. In machine learning problems, it is common to re-scale the input features to an algorithm such that they follow a particular distribution. Some algorithms assume the data is roughly Gaussian in form and can behave poorly if this is not the case. For this reason, we use our final CLAUDS+HSC

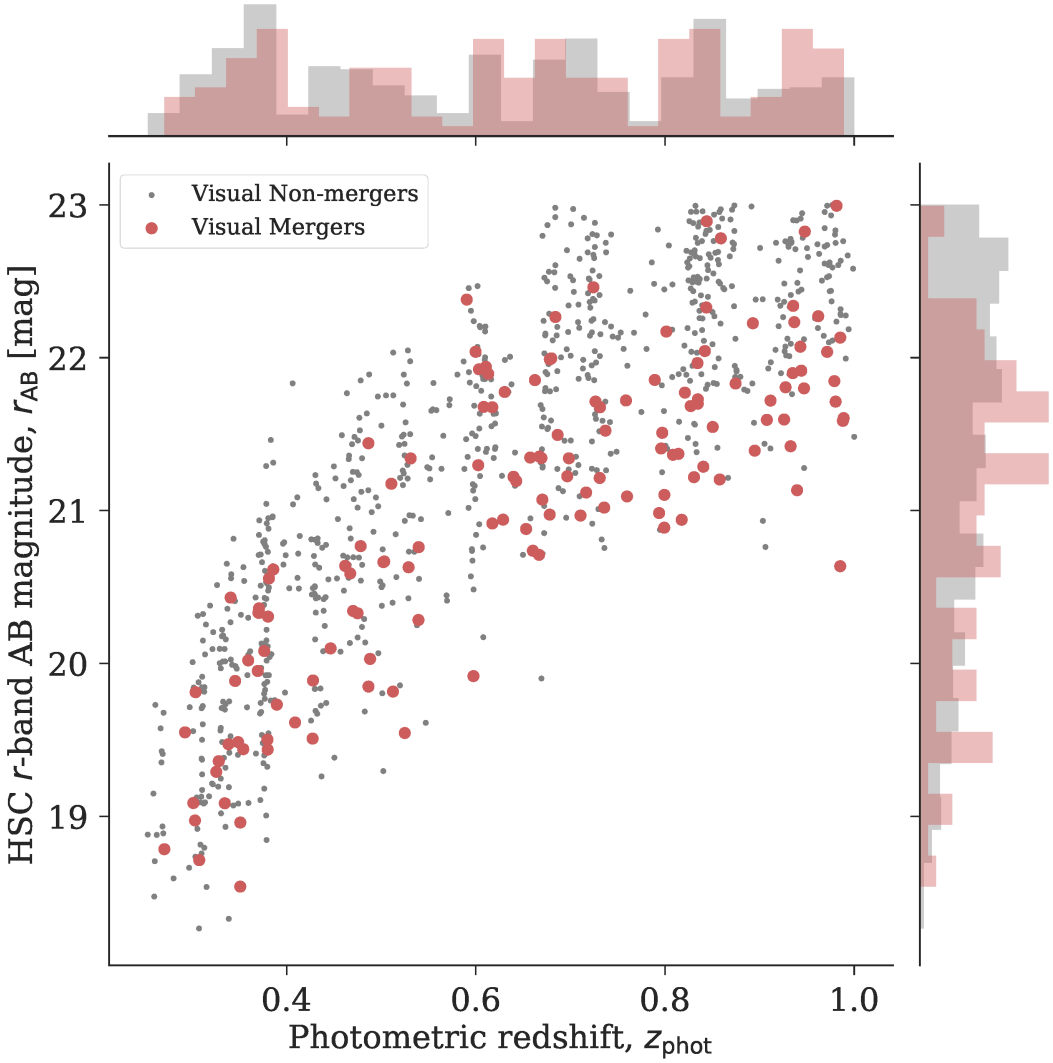


Figure 4.2: Apparent magnitude in the HSC r-band as a function of photo-z for the visually classified galaxies in our training sample (918 galaxies). The marginal histograms show the (normalized) distributions of each parameter, respectively. The mergers (red points) follow roughly the same distribution in parameter space to the non-mergers (grey points), however, less mergers are identified at fainter magnitudes in each redshift bin.

sample (60,957 galaxies) to define a mapping $X \rightarrow X'$ such that each feature follows a standard normal distribution with mean of zero and unit variance. We use the routine `sklearn.preprocessing.StandardScaler` to achieve this. We then apply this scaler to our training set galaxies to transform their features to this new space.

Next, we split our training sample of 918 galaxies into two groups: 70% for training the algorithm (547 non-mergers and 95 mergers), and 30% withheld for testing the classifier's performance (235 non-mergers and 41 mergers). We do not touch this 30% of galaxies until the testing stage in Section 4.2.4. We make sure to "stratify" our training sample such that the relative fractions of visual mergers and non-mergers stays constant in each of the two groups ($\sim 15\%$ mergers and $\sim 85\%$ non-mergers).

Machine learning problems can suffer from biases due to *class imbalance* in the training set. In other words, if there are too many of one class of objects in the training set (in this case, non-mergers), the algorithm will assume that most of the time, guessing the dominant class will yield the correct answer.[4] There are two ways to think about treating this imbalance: (1) we can remove non-mergers until there are equal counts of both classes; or (2) we can increase the number of mergers to match that of the non-mergers. Taking the first approach causes us to lose information on our non-merging population; in general, the larger the training set, the better. On the other hand, taking the second approach proves difficult in practice because by nature, mergers are rare and it takes time to find them in

[4]To visualize this, think about the weather patterns in Los Angeles, California. Most of the time, it is sunny and therefore if you guess that it will be sunny, then you will generally be correct. In machine learning, using this approach results in a less robust classification and does not treat rare occurences. In our case, if we were to guess that each galaxy is a non-merger, we would be correct most of the time, but we wouldn't be finding any mergers! This is why it is important to treat class imbalance.

our images. We therefore choose to take an approach halfway between (1) and (2), which is to *down-weight* our non-mergers and *up-boost* our mergers so that we end up with 50% of each class.

Down-weighting our non-merger population involves randomly sampling galaxies (without replacement) from the 547 non-mergers used for training. Doing this, and retaining $\sim 70\%$ of the original non-mergers, we are left with 380 non-mergers. In order to up-boost our merger population without needing to inspect more images, we create new "fake" mergers using the features of the ones we already have. We create new mergers in a statistical manner by taking the following approach: (1) Take all 95 mergers used in training and find the standard deviation in each of their re-scaled features. This gives us a rough estimate of the error on each feature for *just* the mergers. (2) Consider a *single* merger and its features. For each *feature*, we create a normal distribution from which we randomly draw three new values for that feature. The value of the feature from the original merger is used as the mean, and the error derived from the entire merger sample (from (1)) is used as the standard deviation for the normal distribution. (3) We then repeat this process of randomly sampling three new sets of features for each original merger to obtain a merger sample up-boosted by a factor of 3 (i.e., 95 original mergers becomes: $95 + 95 \times 3 = 380$ mergers). We have now treated the class imbalance and have equal numbers of non-mergers and mergers to use for training our classifier. By taking the above approach, we aim to fill out the parameter space of the merger population, something that would not result from simply making direct copies of the mergers we already have. This, however, assumes that our features are both independent and Gaussian in nature, which may not necessarily be true.

Finally, we train our Random Forest Classifier with the following input: the 380 up-boosted mergers and 380 down-weighted non-mergers with their re-scaled features and labels (i.e., visual classification). We use the `sklearn.ensemble.RandomForestClassifier` routine to perform the fit. The input features and user-supplied *hyperparameters* for our forest are summarized in Table 4.2. For our features, we include all parametric and non-parametric morphological indicators discussed in Chapter 3 as well as the Petrosian radius r_p, stellar mass $M_\star$, photometric redshift $z_{\rm phot}$, r-band apparent magnitude $r_{\rm AB}$, and star-forming probability $P_{\rm sf}$ of the galaxy. The star-forming probabilities for each galaxy are calculated by Golob et al. (in prep) using a Support Vector Machine to choose the star-forming/quiescent decision boundary in a 3-dimensional parameter space defined by the rest-frame $u-r$ and $r-y$ galaxy colours and the photometric redshifts from their k-nearest-neighbours technique. We choose to add these additional features so as to give the RFC as much information about the galaxy as possible. In total, we input 14 features to our Random Forest Classifier and are therefore searching for mergers in a *14-dimensional parameter space* defined by both the morphological parameters we calculate using the HSC r-band images and the photometrically derived physical properties of our galaxies from Golob et al. (in prep).

The hyperparameter `n_estimators` is the number of Decision Trees we use to make up our forest. We therefore `bootstrap` resample our training set 1000 times and take the average of the 1000 resulting Decision Trees to obtain our merger probabilities. We employ the *gini* `criterion` of Equation 4.1 as a measure of the quality of each split. At each split, the algorithm randomly chooses 3 features (`max_features`) to consider. As mentioned before,

Table 4.2: Input features and hyperparameters used to initiate our `RandomForestClassifier`.[a]

Hyperparameter	Value
`input features`	$M_\star$; $z_{\rm phot}$; $r_{\rm AB}$; r_p; n; G; M_{20}; RFF; C; A; S; $A_{\rm resid}$; $S_{\rm resid}$; $P_{\rm sf}$
`n_estimators`	1000
`criterion`	*gini*
`max_features`	3
`max_depth`	10
`min_samples_split`	7
`bootstrap`	True
`warm_start`	False
`class_weight`	*balanced*

[a] The hyperparameters chosen in this work follow closely those chosen in Goulding et al. (2018).

the decision of which feature as well as its value is that which minimizes *gini*. The threshold criterion which stops the tree from splitting any further is chosen to be whichever of the following is satisfied first: (1) the minimum number of galaxies in a node (`min_samples_split`) becomes less than 7, or (2) the depth of the tree (`max_depth`) reaches a maximum value of 10 splits. As mentioned before, both `max_depth` and `min_samples_split` act to reduce model complexity and overfitting. We specify that we have treated class imbalance by setting the `class_weight` hyperparameter to *balanced*. Finally, the `warm_start` hyperparameter requires that a new forest be fit each time the `RandomForestClassifier` function is called. Generally, these hyperparameters can be optimized to produce the best possible fit to the data, however, we choose to keep them fixed for simplicity.

4.2.4 Testing the Forest

Now that we have trained our Random Forest Classifier, there are several ways in which we can assess how well it is performing. This step of the machine learning process is important as it lets us know roughly how accurate our algorithm will be when applied to the full dataset.

4.2.4.1 Feature Importances

First, we can look at the relative importances of each input feature to determining the probability that each object is a merger. Figure 4.3 lists the features in decreasing level of importance. In particular, the Residual Flux Fraction RFF is the most important feature when determining merger probability, closely followed by the M_{20} parameter. Together, they account for almost 40% of the total (cumulative) importance. Features such as the photometric redshift, r-band magnitude, and stellar mass are of lesser importance. In other words, if we wished to only consider features that contribute to $\sim 90\%$ of the total importance, then we could safely discard the four least important features (G, $z_{\rm phot}$, $r_{\rm AB}$, and $M_{\star}$).

4.2.4.2 Merger Probabilities

We now use our withheld 30% (276 galaxies) of the training set to assess the performance of our classifier. We run each object through the forest *without* their labels to obtain the probability $P_{\rm merge}$ that they are merging, being sure to re-scale the data using the same scaler defined in Section 4.2.3. We compare these probabilities to our visual identifications

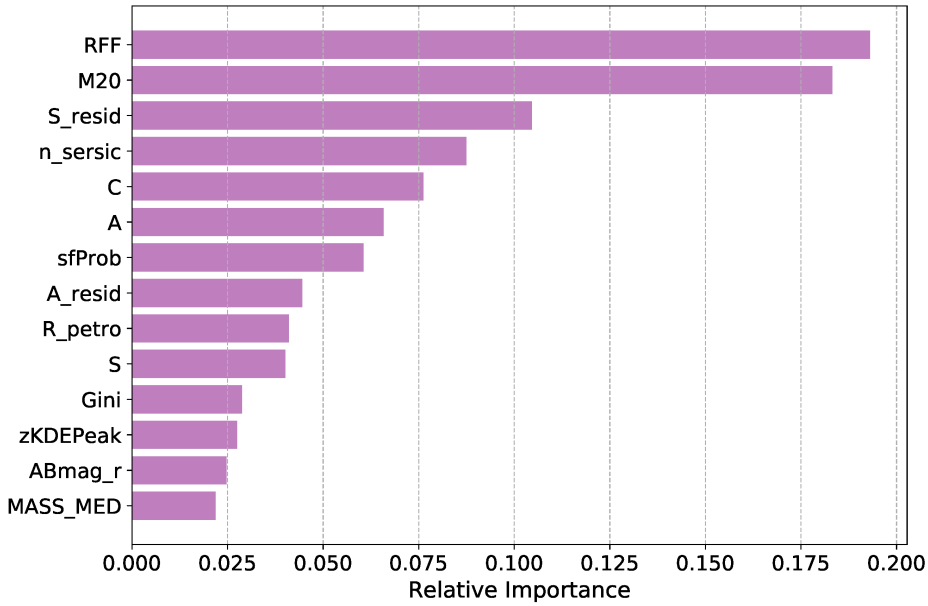


Figure 4.3: Rank ordered feature importances for our `RandomForestClassifier`. The RFF and M_{20} parameters are the most important features.

in Figure 4.4, where the distribution of merger probabilities is plotted and colour-coded by visual identification. From the Figure, we can see that most visual non-mergers (grey histogram) lie at lower merger probabilities $P_{\rm merge}$, whereas most visual mergers (red histogram) lie at higher values of $P_{\rm merge}$. A clear bimodality between the two classes, however, is not apparent and there is a sizeable overlap between the two samples, especially at the highest merger probabilities. The distributions shown in Figure 4.4 can be used to infer the distributions we will obtain when running our RFC on the full sample of 60,957 galaxies. We can assume that the distributions will be similar provided that we assume our training sample probes the same parameter space as the full galaxy population.

Since the Random Forest returns the *probability* that an object belongs to a particular class, we must make a cut in probability in order to obtain a binary classification: merger

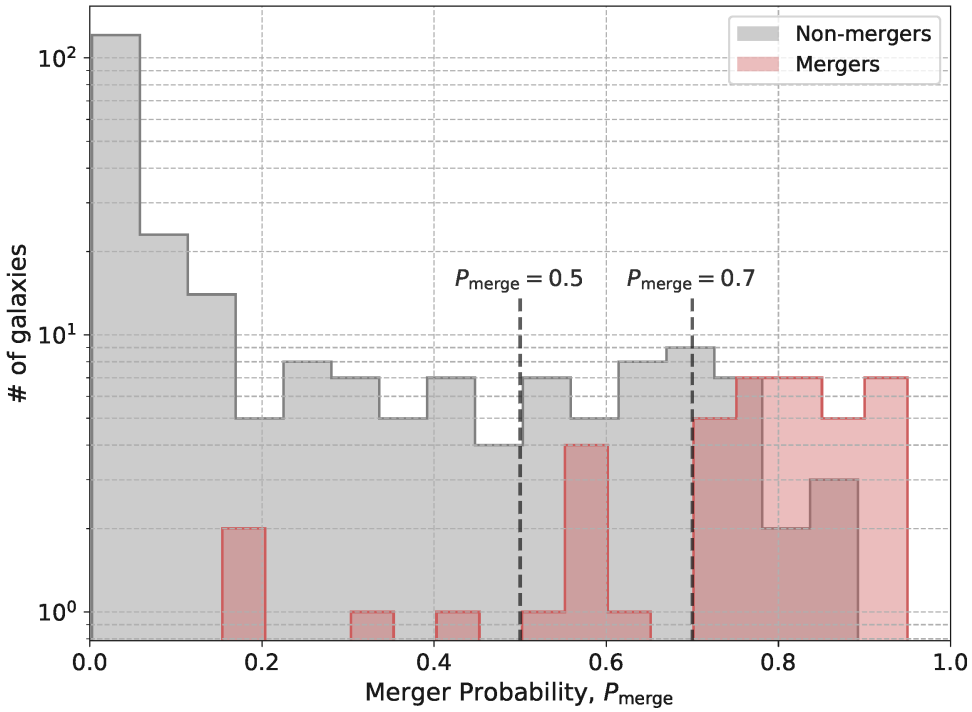


Figure 4.4: Distribution of merger probabilities $P_{\rm merge}$ for the test set (30% withheld). The 235 visual non-mergers (grey) tend to lower values of $P_{\rm merge}$, and the 41 visual mergers (red) are seen at higher values of $P_{\rm merge}$. The vertical axis is logarithmic so as to emphasize the merger population and we mark two probabilities $P_{\rm merge} = 0.5$ and $P_{\rm merge} = 0.7$ which we use in later Sections.

or non-merger. The question now is *where* to make this cut. The answer to this depends on what our goals are. Specifically, we must ask ourselves whether we prefer a sample of mergers that is *pure* and contains little to no non-mergers, or whether we wish to have a sample that is *complete*; i.e., most of the mergers are included. In reality, we want a little of *both*. We don't want to discard too many mergers since this is what we are interested in finding, however, we also don't want to include too many non-mergers because our "merger" sample would be highly contaminated. Ultimately, we opt for a *complete* merger sample as opposed to a pure one. In Section 4.2.4.4, we explain how to treat this contamination by

choosing cuts in probability as a guide for further visual inpection.

4.2.4.3 Measures of Classifier Performance

We can use the results of Figure 4.4 to assess how well our Random Forest Classifier will perform on the full CLAUDS+HSC dataset. As an example, say we choose a cut in merger probability of $P_{\rm merge} = 0.5$. Galaxies with probabilities above (below) this value are therefore classified to be "mergers" ("non-mergers") by our Random Forest. We can compare the Random Forest's labels to our *true* labels from visual classification in a *confusion matrix*.

Figure 4.5 shows the confusion matrix for a cut of $P_{\rm merge} = 0.5$. In this case, 194 ($\sim 82.5\%$) of the 194+41 = 235 visual non-mergers were correctly classified by the algorithm. Similarly, 37 ($\sim 90\%$) of the $37 + 4 = 41$ visual mergers were correctly classified. These two populations are represented by the diagonal elements of the confusion matrix. The off-diagonal elements correspond to the objects that were misclassified by the algorithm using our cut of $P_{\rm merge} = 0.5$. Ideally, we would like these elements to both be as close to zero as possible.

We can define several performance statistics related to the numbers of diagonal and off-diagonal elements of this matrix, where the notation for each element is as follows:

True Positives (TP): # of visual mergers identified as mergers by the RFC,

True Negatives (TN): # of visual non-mergers identified as non-mergers by the RFC,

False Negatives (FN): # of visual mergers identified as non-mergers by the RFC, and

False Positives (FP): # of visual non-mergers identified as mergers by the RFC.

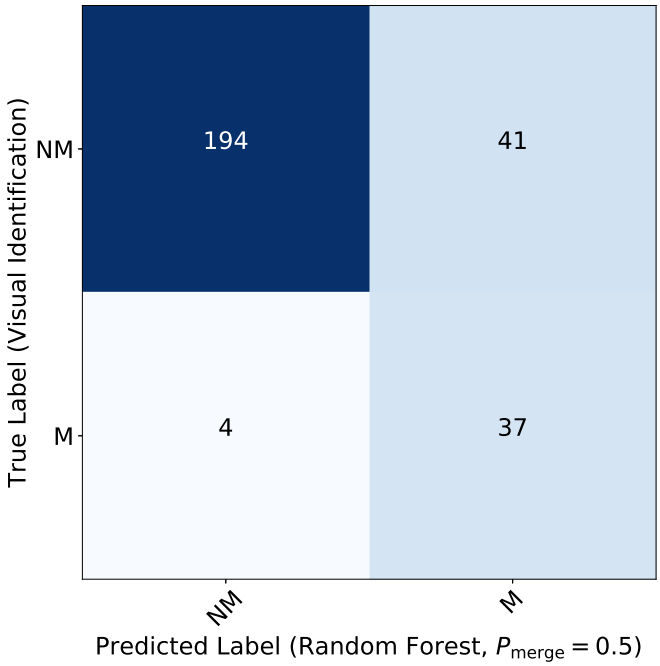


Figure 4.5: Confusion matrix for the test set (30% withheld). Mergers and non-mergers are denoted as "M" and "NM", respectively. Notice the high number of misclassified visual non-mergers (i.e., false positives− 41 galaxies). Although this only accounts for a small fraction of the non-merging population, it exceeds the number of correctly-classified mergers (i.e., true positives− 37 galaxies) leading to a high level of contamination.

Accuracy Score

In *binary* classification problems such as our own, an "accuracy score" can be defined using the Jaccard similarity coefficient (Jaccard 1901):

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|}, \qquad \text{where } 0 \leq J \leq 1. \tag{4.2}$$

For our problem, A in the above equation denotes the set of visual (ground truth) classifications, whereas B denotes the set of classifications predicted by the Random Forest. Therefore, the numerator in Equation 4.2 is the number of galaxies for which both the

visual classification and the RFC prediction match. For $P_{\rm merge} = 0.5$, this is given by: $TN + TP = 194 + 37 = 231$ galaxies. The denominator in Equation 4.2 is simply the total number of galaxies in the test set (276 galaxies). For a cut of $P_{\rm merge} = 0.5$, the accuracy score is 0.8370, or 83.7%. In general, we would like this number to be as close to 1 as possible.

Completeness

The *completeness* measures the fraction of true (visual) mergers that are correctly classified by the algorithm. It is sometimes also referred to as the recall ratio, sensitivity, or true positive rate (TPR). It is calculated as follows:

$$\text{completeness} = \frac{TP}{TP + FN}. \tag{4.3}$$

For the example of $P_{\rm merge} = 0.5$, the completeness is $\sim 90\%$. We want the value for the completeness measure to be as close to 100% as possible.

Reliability

The *reliability* measures the fraction of misclassified non-mergers to the total number of non-mergers:

$$\text{reliability} = \frac{FP}{TN + FP}. \tag{4.4}$$

For $P_{\rm merge} = 0.5$, the reliability is 0.17; in other words, we are incorrectly classifying 17% of our non-merging population. In general, we want its value to be as close to 0% as possible.

Purity

The *purity*, sometimes also called precision, is defined as:

$$\text{purity} = \frac{TP}{TP + FP}. \tag{4.5}$$

It tells us what fraction of galaxies above our cut in P_{merge} are actually true (visual) mergers. In the case of $P_{\text{merge}} = 0.5$, the "merger" sample defined by the algorithm (i.e., everything above P_{merge} of 0.5) is only 47% pure. We want this value to be as close to 100% as possible.

Contamination

In a similar way to the purity, we can define the *contamination* of the "merger" sample defined by the algorithm:

$$\text{contamination} = \frac{FP}{FP + TP}. \tag{4.6}$$

In fact, the contamination is simply 1− *purity*. Therefore, it tells us what fraction of galaxies above our cut in P_{merge} are true (visual) non-mergers. For $P_{\text{merge}} = 0.5$, the contamination is 53% and in general, we want the contamination to be close to 0%. The purity and contamination sum to 100%.

F1 Score

The $F1$ score is another diagnostic of the performance of a binary classifier. It is defined as the harmonic mean of the *completeness* and *purity* measures:

$$F1 = \frac{2}{\frac{1}{\text{completeness}} + \frac{1}{\text{purity}}}. \tag{4.7}$$

In general, we want the value of $F1$ to be as close to 1 as possible; in other words, we want our merger sample to be both 100% complete and 100% pure. Therefore, if we wish to optimize our merger sample for *both* purity and completeness, choosing the cut in P_{merge} for which the $F1$ score is maximized is a sensible approach. For our example of $P_{\text{merge}} = 0.5$, the $F1$ score is 0.62.

Logistic Loss

In binary classification problems, the *logistic loss* (Log Loss), also called the *cross entropy loss*, is a statistic used to represent the amount by which the distribution of predicted values deviates from the distribution of true labels. To this end, we needn't make a cut in probability and the Log Loss value for a particular Random Forest Classifier will be the same no matter the cut in probability.

Mathematically, the Log Loss is defined as (see, for example, Murphy 2012):

$$\text{Log Loss} = -(y \ln(p) + (1 - y) \ln(1 - p)), \tag{4.8}$$

where y is the visual classification of a particular galaxy (either 0 for a non-merger, or 1 for a merger), and p is the probability that the same galaxy is a merger. We want the value for the Log Loss to be as close to 0.0 as possible. For our classifier, Log Loss = 0.3038. Much work went into engineering a well-sampled training set in order to minimize this value along

with optimizing the other measures. For example, we needed to treat an underabundance of star-forming non-mergers at higher redshifts, which caused the Log Loss value to be larger.

Area Under Curve (AUC)

The final performance measure we consider is the *area under the curve* (AUC). This statistic is given by the area under the receiver operating characteristic (ROC) curve, which compares the true positive rate (or *completeness*) and the false positive rate (FPR) for different cuts in probability for a given classifier. The false positive rate is given by: FPR $= FP/(FP+TN)$. Graphically, the ROC curve for our Random Forest Classifier is shown in Figure 4.6.

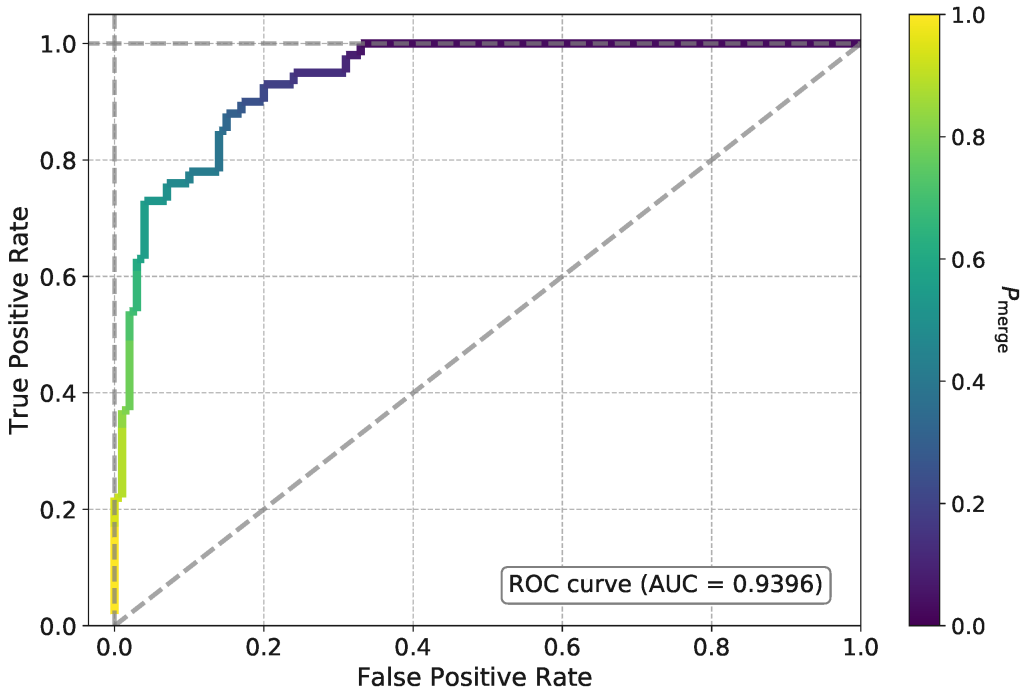


Figure 4.6: ROC curve for our Random Forest Classifier. The curve is colour-coded by the cut in P_{merge} which returns the corresponding values for the FPR and TPR. The dotted lines are for reference.

By definition, when the cut is $P_{\rm merge} = 1.0$, there are no galaxies lying above this value and thus, FPR = TPR = 0. Similarly, when the cut is $P_{\rm merge} = 0.0$, all galaxies lie above this value and thus, FPR = TPR = 1. In the case of a perfect classifier, the area under the ROC curve would be exactly equal to 1. In other words, the TPR would be equal to 1 for all positive values of the FPR. For our classifier, we find AUC = 0.9396.

4.2.4.4 Thresholding $P_{\rm merge}$ to Treat Contamination

The results obtained using supervised methods such as Random Forests are always subject to some level of cross-contamination between the classes. Ultimately, we need to choose a cut in $P_{\rm merge}$ to obtain samples of merging and non-merging galaxies to use in our estimate of the merger fraction evolution. The above performance measures can help us understand the levels of misclassification in both samples so that we may apply the proper corrections. In Figure 4.7 we show how each statistic behaves when we vary the position of the cut in $P_{\rm merge}$. We also list the numerical values for each of the statistics at intervals between $0.1 \leq P_{\rm merge} \leq 0.9$ in Table 4.3 for reference.

As stated before, the best classifier is that which *maximizes* the accuracy score, completeness, purity, and $F1$ score; while *minimizing* the reliability and contamination. We consider two cuts: $P_{\rm merge} = 0.5$ and $P_{\rm merge} = 0.7$. Just above $P_{\rm merge} = 0.7$, the $F1$ score (red dotted line in Figure 4.7) reaches a maximum. This tells us that both the purity and completeness are optimized at this value. The levels of purity and contamination are roughly equal around $P_{\rm merge} = 0.6$. Choosing a cut at $P_{\rm merge} = 0.7$ might seem a good choice at first, however, the value for the completeness ($\sim 76\%$) motivates us to also consider a cut

Table 4.3: Overview of performance statistics for the RFC of this work at various cuts in $P_{\rm merge}$. The boldface numbers highlight the statistics for two values of $P_{\rm merge}$ (0.5 and 0.7) discussed in the text.

$P_{\rm merge}$	Accuracy Score $J = \frac{\lvert A \cap B \rvert}{\lvert A \cup B \rvert}$	Completeness $\frac{TP}{TP+FN}$	Reliability $\frac{FP}{TN+FP}$	Purity $\frac{TP}{TP+FP}$	Contamination $\frac{FP}{FP+TP}$	$F1$ Score[a]
0.1	0.6558	1.0	0.4043	0.3015	0.6985	0.4633
0.2	0.7283	0.9512	0.3106	0.3482	0.6518	0.5098
0.3	0.7681	0.9512	0.2638	0.3861	0.6139	0.5493
0.4	0.8043	0.9268	0.2170	0.4270	0.5730	0.5846
0.5	**0.8370**	**0.9024**	**0.1745**	**0.4744**	**0.5256**	**0.6218**
0.6	0.8514	0.7805	0.1362	0.5	0.5	0.6095
0.7	**0.9058**	**0.7561**	**0.0681**	**0.6596**	**0.3404**	**0.7045**
0.8	0.9058	0.4634	0.0170	0.8261	0.1739	0.5938
0.9	0.8768	0.1707	0.0	1.0	0.0	0.2917

[a] $F1$ Score: $2/(1/\text{completeness} + 1/\text{purity})$.

of $P_{\rm merge} = 0.5$ for which this value greatly improves ($\sim 90\%$). This can also be seen in Figure 4.4, where by choosing this cut, the number of true positives is increased by almost 20%. We will use these cuts in Section 5.1 to obtain a relatively complete ($\sim 90\%$) and close to 100% pure sample of mergers using our full sample of galaxies in CLAUDS+HSC.

4.3 Further Motivation for a Multi-dimensional Approach

In Section 3.6, we introduced two popular 2-dimensional approaches to classifying mergers: the $A-S$ plane of Conselice (2003), and the $G-M_{20}$ plane of Lotz et al. (2004, 2008b). We also discussed the 1-dimensional cut in Asymmetry $A > 0.35$ (Conselice 2003). In this Section, we will assess the performance of these methods in comparison to our Random Forest Classifier to further motivate the need for multiple dimensions when classifying mergers.

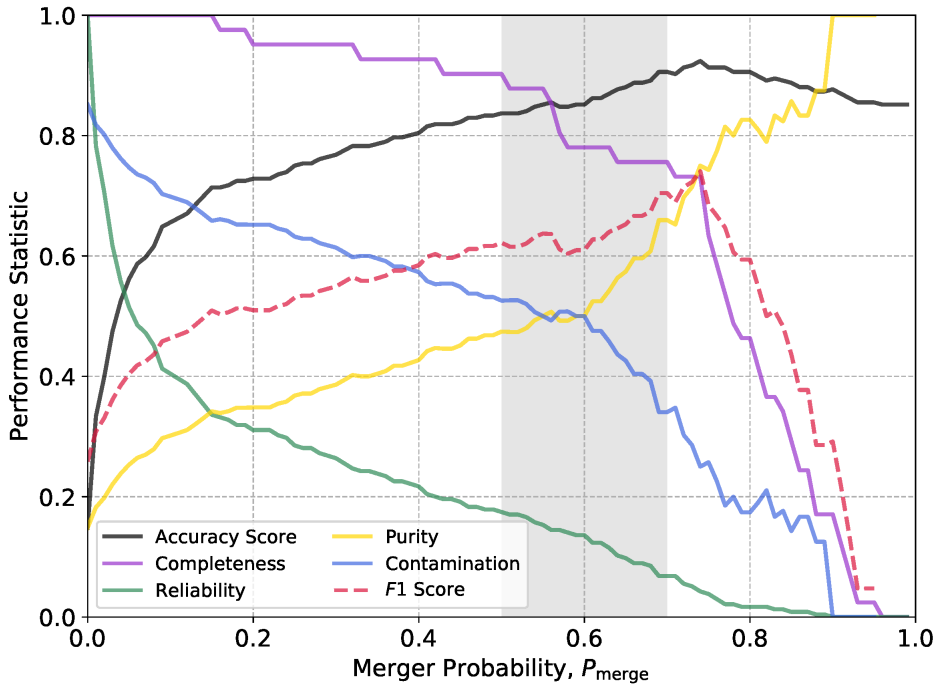


Figure 4.7: Performance statistic values as a function of the cut in $P_{\rm merge}$ for our Random Forest Classifier. The region between $0.5 \leq P_{\rm merge} \leq 0.7$ is shaded for reference. See text for the interpretation of each performance measure.

We begin with the $A - S$ plane which, for our training sample of galaxies, is shown in Figure 4.8. Visually classified mergers (shown in red) tend to have higher values of the Asymmetry parameter, which is to be expected. The horizontal grey dot-dashed line in Figure 4.8 shows $A = 0.35$. By the Conselice (2003) definition, only mergers are supposed to lie above this line, however, we can clearly see this is not the case. The dotted grey line denotes the position of the Conselice (2003) relation which also includes the vertical shift by 3σ (Equations 3.15 and 3.16). Clearly, our data do not abide by this relation, which is to be expected since we used slightly different definitions for both A and S than did Conselice (2003).

To directly compare the $A-S$ method to our Random Forest, we attempt to recreate the Conselice (2003) relation using our data. Following their methods, we first choose a relatively clean sample of non-mergers. Here, we choose a subsample of our non-mergers since there is quite a bit of contamination to the merger sample from non-mergers at $A > 0.35$. Specifically, we choose the 84% of non-mergers with the lowest A values (i.e., $A < 0.62$). Next, we fit a linear model to the non-mergers in the $A-S$ plane by performing 1000 random bootstrap resamplings (with replacement) of the data and minimizing the χ^2 error. We obtain the best fit parameters (slope and intercept) using the $16^{\rm th}$, $50^{\rm th}$, and $84^{\rm th}$ percentiles of the 1000 slopes and intercepts and derive a 68% confidence interval on the fit. For the confidence interval, we sample our bootstrapped fits along the abscissa and choose the $16^{\rm th}$ and $84^{\rm th}$ percentiles in the Asymmetry parameter at each value for the galaxy Smoothness. Our relationship between A and S (measured by our techniques of Section 3.3.2 and 3.3.3) from the HSC r-band images is therefore given by:

$$A_{\rm fit}(r) = (1.42 \pm 0.10)S(r) + (0.19 \pm 0.01). \tag{4.9}$$

In Figure 4.8, the best fit to the non-merger sample and its confidence interval is shown by the grey shaded region. Following Conselice (2003), we assume a dispersion in the *linear* model (σ in Equation 3.16) to be the average of $0.5 \times (A_{\rm fit}^{\rm 84th} - A_{\rm fit}^{\rm 16th})$ for all points along the fit. Here, $A_{\rm fit}^{i\rm th}$ denotes the value of the $i^{\rm th}$ percentile of the bootstrapped fits, the values of which are sampled along S. They are, in fact, the edges of the confidence intervals. The black dashed line in Figure 4.8 shows the best fit line shifted by 3 times the average

dispersion, which we derive to be $\sigma = 0.0172$.

From Figure 4.8, there is less contamination from non-mergers in the sample of galaxies that lie above our $A - S$ relation, when compared to the simple $A > 0.35$ cut. In contrast, our $A - S$ relation produces a less complete sample of mergers than the $A > 0.35$ cut. Table 4.4 outlines the performance statistics of our $A - S$ relation and the $A > 0.35$ cut compared to two cuts in probability on our Random Forest Classifier: $P_{\rm merge} \geq 0.5$ and $P_{\rm merge} \geq 0.7$. In general, our RFC performs much better than the $A - S$ plane and $A > 0.35$ methods. Although the $A > 0.35$ method returns a higher merger sample completeness than the RFC with $P_{\rm merge} \geq 0.7$, it suffers with almost 75% contamination from non-mergers.

We now look at the $G - M_{20}$ plane, which is shown in Figure 4.9 for our training sample of galaxies. Visually classified non-mergers show lower values of the M_{20} parameter, which is to be expected since they are usually more concentrated than mergers (especially in the case of bulge-dominated non-mergers). In Figure 4.9, the Lotz et al. (2008b) definition for mergers is plotted with the grey dotted line. Our galaxy sample does not follow this trend since we use images in a different wavelength, and with a different spatial resolution (Lotz et al. 2008b used data from *HST*, while we use ground-based imaging). We also use a different method for generating the segmentation maps of our galaxies.

Just as we did in the $A - S$ plane, we can recalibrate the $G - M_{20}$ relation to our data. Lotz et al. (2008b) followed a similar approach to the one described above for the $A - S$ plane to obtain their $G - M_{20}$ relation and so we apply it here as well. To fit the sequence of non-mergers, we choose the 84% of visually-classified non-mergers with the lowest M_{20} values and bootstrap this sample 1000 times to obtain the best fit and 68% confidence intervals.

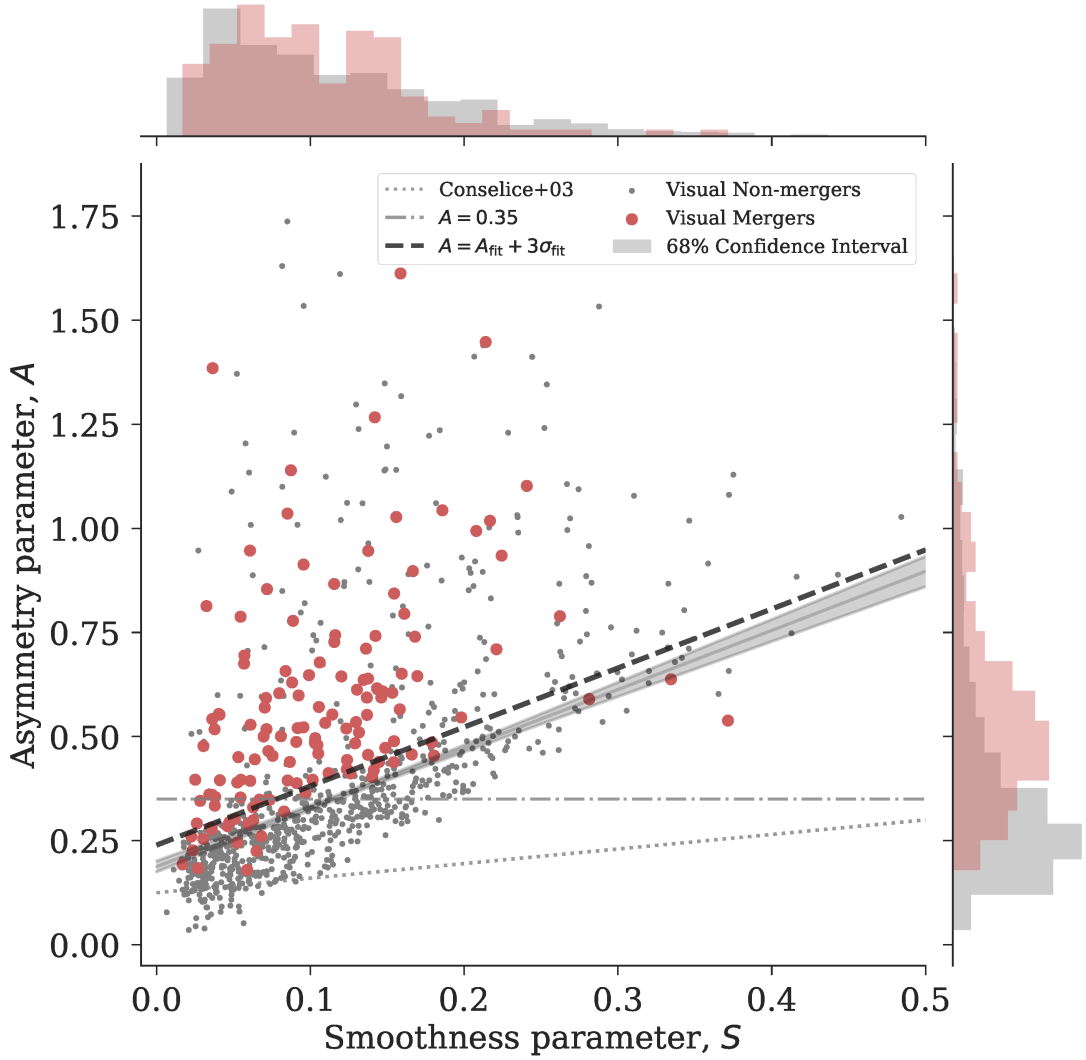


Figure 4.8: Asymmetry A vs. Smoothness S for our training set galaxies. Visually identified mergers (red points) and non-mergers (grey points) more or less occupy different regions in this parameter space, but with high levels of contamination. Distributions of each parameter, separated by visual classification, are shown in the (normalized to unit area) marginal histograms. The Conselice (2003) $A-S$ relation (grey dotted line, Equations 3.15 and first of Equation 3.16) and $A = 0.35$ cut (grey dot-dashed line) do not identify reasonably clean samples of mergers and non-mergers. Recalibrating the $A-S$ relation for our galaxies, we recover the best fit with 68% confidence intervals to the non-mergers (grey shaded) and the final $A-S$ relation for our galaxies $A_{\rm fit} = 1.42S + 0.19 + 3\sigma$, where $\sigma = 0.017$ (black dashed line). The non-merger sample used to derive this relation uses galaxies with $A < 0.62$. See text for details.

Our recalibrated fit to the non-mergers in the $G - M_{20}$ plane using the HSC r-band images is given by:

$$G_{\rm fit}(r) = (-0.26 \pm 0.02)M_{20}(r) + (0.11 \pm 0.04). \tag{4.10}$$

Deriving a dispersion in the linear fit using the same methods as above, we find $\sigma = 0.0063$. The best fit line in Equation 4.10 vertically shifted by 3σ is shown as the black dashed line in Figure 4.9. This recalibrated expression works slightly better for our galaxies than the original Lotz et al. (2008b) definition, however, there is still a significant amount of overlap between the merging and non-merging samples. Table 4.4 gives the performance statistics by using our recalibrated $G - M_{20}$ method. Out of all the classification schemes we have discussed, it performs the worst with high levels of contamination, low levels of completeness, and low levels of purity in the merger sample.

In Table 4.4, we can see just how well our RFC performs when compared to lower dimensional approaches. We obtain higher values for the accuracy score and lower values of the reliability statistic. Our method also returns a less contaminated (more pure) merger sample for both cuts in $P_{\rm merge}$. Furthermore, our values for the $F1$ score are higher, our Log Loss is lower, and our AUC is higher than in the lower dimensional methods. The only instance when a lower dimensional approach appears to outperform our method is in the completeness of the merger sample. As mentioned above, for a simple cut in the Asymmetry of $A > 0.35$, the completeness is 83% while our RFC for $P_{\rm merge} = 0.5$ returns a completeness of only 76%. Recall, however, that we want to optimize both the completeness and purity simultaneously (i.e., maximize the $F1$ score). In the case of the $A > 0.35$ classifier, the

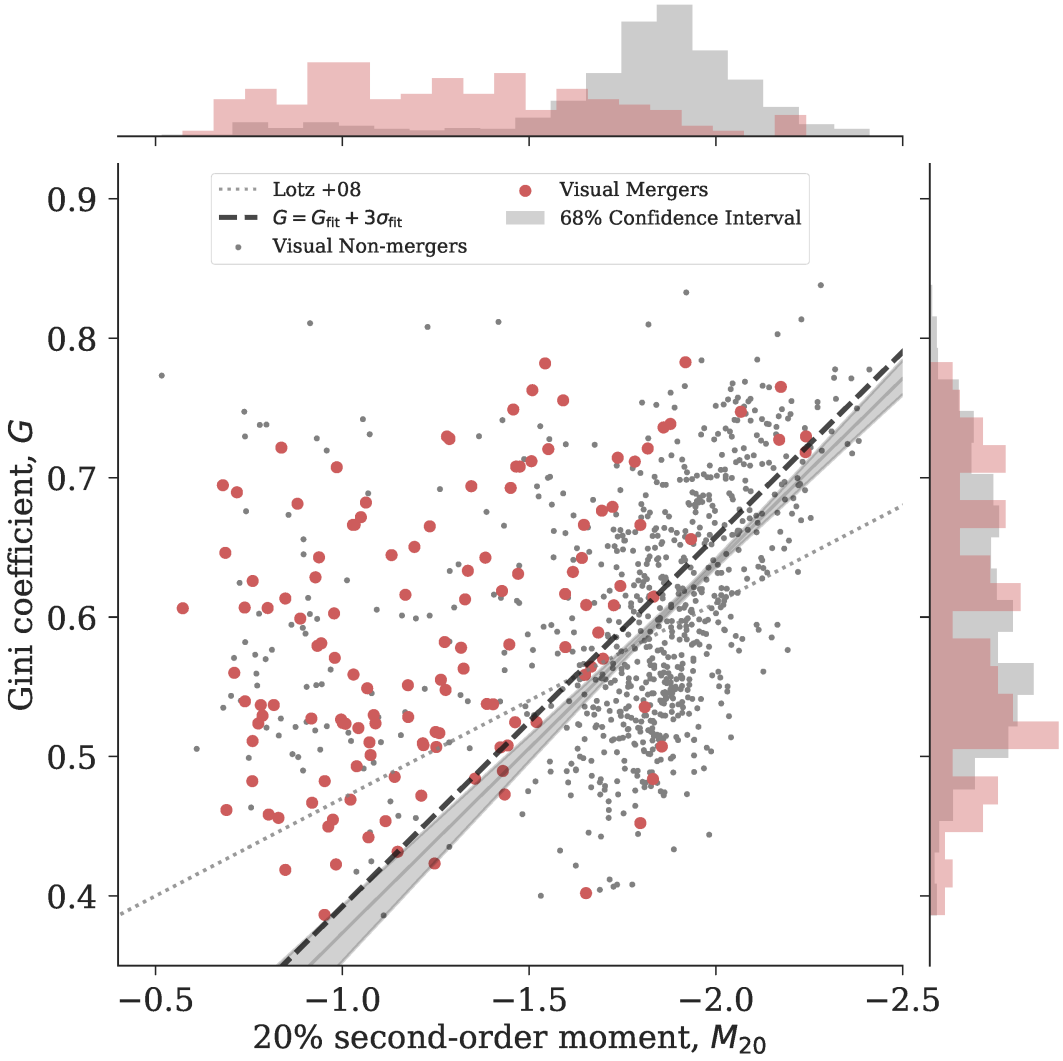


Figure 4.9: Same as Figure 4.8, but for our training set galaxies in the $G - M_{20}$ plane. The Lotz et al. (2008b) relation (Equation 3.17) is shown by the grey dotted line and clearly does not apply to our galaxies. Our recalibrated fit in the $G - M_{20}$ plane and 68% confidence intervals for the non-mergers are shown as the grey shaded region. Our final $G - M_{20}$ relation, $G_{\rm fit} = -0.26M_{20} + 0.11 + 3\sigma$, where $\sigma = 0.006$ is shown by the black dashed line. There is still quite a bit of contamination from non-mergers above this line.

merger sample is $< 30\%$ pure, while our RFC with $P_{\rm merge} = 0.5$ returns a much higher purity of 66% and our $F1$ score is indeed much higher.

By examining these results, it becomes clear why 1- and 2-dimensional approaches to merger classification are *not* very effective. If we were limited to a 2-dimensional approach, however, the results of our RFC inform us that using only the two most important features to define a parameter space (i.e., RFF vs. M_{20}) might produce more reliable results than those presented above. Alternatively, even combining the $A - S$ and $G - M_{20}$ methods in an $A - M_{20}$ relation may result in a more accurate separation of mergers from the rest of the galaxy population.

Mergers are very diverse in their morphologies and using only one or two statistics to describe them neglects a wealth of information that could be used to contribute to a more robust classification scheme. This is why the multi-dimensional approach, in our case we use 14 features and thus 14 dimensions, is much more effective at finding mergers.

Table 4.4: Overview of performance statistics for the Random Forest Classifier of this work compared to lower dimensional classification schemes discussed in the text. For reference, we boldface the statistics of the two best classifiers in each column.

Classifier Name	Accuracy Score	Completeness	Reliability	Purity	Contamination	$F1$ Score	Log Loss	AUC
RFC ($P_{\rm merge} \geq 0.5$)[a]	**0.8370**	**0.9024**	**0.1745**	**0.4744**	**0.5256**	**0.6218**	**0.3038**	**0.9396**
RFC ($P_{\rm merge} \geq 0.7$)[a]	**0.9058**	0.7561	**0.0681**	**0.6596**	**0.3404**	**0.7045**	**0.3038**	**0.9396**
$A > A_{\rm fit} + 3\sigma_{\rm fit}$	0.7756	0.5	0.5	0.3768	0.6232	0.4297	7.7507	0.7802
$A > 0.35$	0.6438	**0.8309**	0.3887	0.2710	0.7290	0.4087	12.3033	0.7211
$G > G_{\rm fit} + 3\sigma_{\rm fit}$	0.5980	0.5	0.5	0.2547	0.7453	0.3375	13.8835	0.7185

[a] We display the performance statistics for the two cuts in $P_{\rm merge}$ we identify from Figure 4.7 in Section 4.2.4.3.

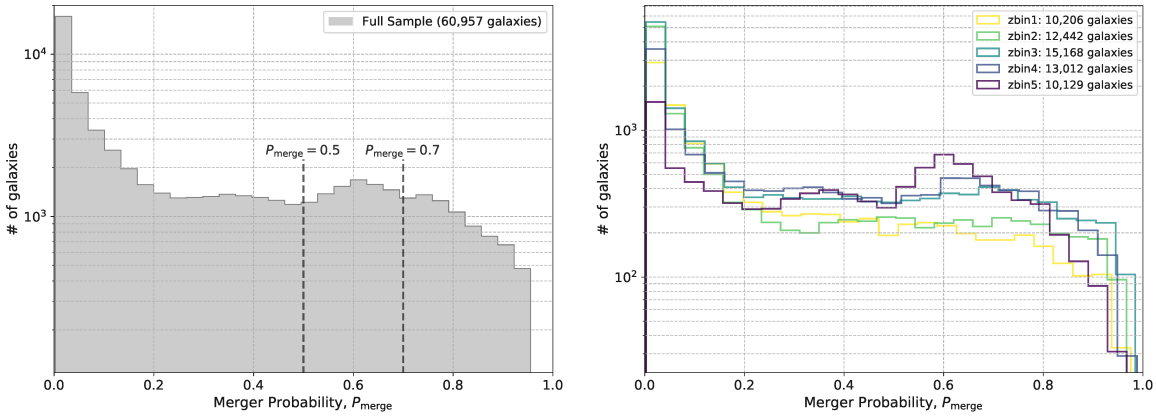
Chapter 5

Results & Discussion

In this Chapter, we examine the results of applying our Random Forest Classifier to the full $\sim$ 20 deg^2 *Deep* and *UltraDeep* layers of the CLAUDS+HSC survey. We present the distribution of merger probabilities for the entire dataset and explain how these probabilities are used to derive a merger fraction evolution. We outline the technique used to correct our fractions to account for incompleteness in the merger sample at higher redshifts. We present our incompleteness corrected merger fraction evolution, examine how it fits into current findings in the literature, and estimate the fractional merger rate. Finally, we discuss the major caveats of our work along with suggestions for improvement.

5.1 Applying the Forest to the Full $\sim$ 20 deg^2

Using the Random Forest Classifier we trained in Chapter 4, we can obtain the merger probabilities $P_{\rm merge}$ for all 60,957 galaxies in our sample. The results of running each galaxy



(a) (b)

Figure 5.1: *(a)* Distribution of merger probabilities $P_{\rm merge}$ for our full sample of 60,957 CLAUDS+HSC galaxies. The lines at $P_{\rm merge} = 0.5$ and $P_{\rm merge} = 0.7$ are for reference. *(b)* Distributions of $P_{\rm merge}$ in each of the 5 redshift bins: $z = 0.25 - 0.4$, $z = 0.4 - 0.55$, $z = 0.55 - 0.7$, $z = 0.7 - 0.85$, and $z = 0.85 - 1.0$.

through our classifier are shown in Figure 5.1. Below $P_{\rm merge} = 0.5$, there are 44,393 galaxies. This accounts for $\sim 73\%$ of the total galaxy sample. By the arguments in Section 4.2.4.3, we consider this sample to contain only non-mergers. In reality, Figure 4.4 tells us that about 2% of these galaxies are likely mergers but for our purposes we ignore this.

Above $P_{\rm merge} = 0.5$, recall that we have roughly 50% contamination from non-mergers in the validation sample. We treat this contamination by visually inspecting *all galaxies* with $P_{\rm merge} \geq 0.7$ (7,550 galaxies) and a randomly chosen *subset* of 500 galaxies (100 in each redshift bin) with $0.5 \leq P_{\rm merge} < 0.7$. We choose to inspect only a subset of the galaxies in this range, where the total number of galaxies is 9,014, since visual inspection is time consuming and we expect $\lesssim 20\%$ of the galaxies to be visual mergers. Using the subset of galaxies with $0.5 \leq P_{\rm merge} < 0.7$, we can easily infer the number of true mergers in this range.

Table 5.1: Overview of galaxy number counts used to derive $f_{\rm merge}$ for our sample of CLAUDS+HSC galaxies. The errors on $f_{\rm merge}$ are given by $\sqrt{N_{\rm merge}}/N_{\rm tot}$.

Redshift	$N_{\rm tot}$	$P_{\rm merge} \geq 0.5^a$		$P_{\rm merge} \geq 0.7^b$		With Visual Inspection^c	
		$N_{\rm merge}$	$f_{\rm merge}$ (%)	$N_{\rm merge}$	$f_{\rm merge}$ (%)	$N_{\rm merge}$	$f_{\rm merge}$ (%)
$0.25 - 0.4$	10,206	2,008	19.67±0.44	918	8.99±0.30	260	2.55±0.16
$0.4 - 0.55$	12,442	2,589	20.81±0.41	1,376	11.06±0.30	254	2.04±0.13
$0.55 - 0.7$	15,168	3,805	25.09±0.41	1,975	13.02±0.29	350	2.31±0.12
$0.7 - 0.85$	13,012	3,932	30.22±0.48	1,840	14.14±0.33	310	2.38±0.14
$0.85 - 1.0$	10,129	4,230	41.76±0.64	1,441	14.23±0.37	412	4.07±0.20

[a] Results if we consider all galaxies with $P_{\rm merge} \geq 0.5$ to be mergers.
[b] Results if we consider all galaxies with $P_{\rm merge} \geq 0.7$ to be mergers.
[c] Results if we apply the visual inspection methods of Section 5.1 to treat contamination from non-mergers above $P_{\rm merge} \geq 0.5$. We refer to this as the *un-corrected* merger fraction.

In Table 5.1, we list the number counts of galaxies and mergers in each of the 5 redshift bins, along with the resulting merger fraction in three cases: (1) if we were to consider all galaxies with $P_{\rm merge} \geq 0.5$ to be mergers, (2) if we were to consider all galaxies with $P_{\rm merge} \geq 0.7$ to be mergers, and (3) if we apply the visual inspection described above. Comparing case (3) to the previous two, it is clear that there are high levels of contamination at larger values for the merger probability and the resulting merger fractions are quite different if we do not treat this impurity. It is important to consider that even though we needed to visually inspect just over 8,000 galaxies to obtain a pure merger sample, this corresponds to only $\sim 13\%$ of our total galaxy sample. Therefore, using a RFC as a *triage* to yield a much smaller sample of galaxies for further visual confirmation is one way to think of the methodology presented in this work.

5.2 Incompleteness Corrections

At higher redshifts, galaxies become both fainter in magnitude and smaller on the sky, which makes it increasingly difficult to resolve intricate nuclear structures and detect faint tidal features. Consequently, an interacting galaxy at high reshift may in fact look quite similar to one that is not interacting at all and we will be underestimating the number of mergers as we look further back in time. In order to account for this issue of *incompleteness*, we need to estimate the number of mergers we are missing at each redshift and adjust our values for the merger fraction accordingly. We could consider the opposite effect as well, where with increasing redshift non-merging galaxies may begin to occupy the same parameter space as the mergers, thereby artificially *increasing* our merger fraction. In Section 5.1 we added a second layer of visual identification to eliminate the majority of this contamination.

To correct for incompleteness, we take a sample of relatively *local* mergers and artificially redshift them to simulate how their morphologies and merger probabilities change with increasing redshift. We then ask, for each new "mock" galaxy, whether or not they would be detected as mergers by our Random Forest Classifier. This gives us a correction factor c which we can apply to our un-corrected merger fraction $f_{\rm merge}$ at each redshift bin to obtain the true, incompleteness corrected, merger fraction:

$$f_{\rm merge,corr} = f_{\rm merge} \times \frac{1}{c}. \tag{5.1}$$

We visually inspected ~ 450 galaxies between $0.1 \leq z_{\rm phot} < 0.25$ that, when run through our RFC, returned merger probabilities of $P_{\rm merge} \geq 0.45$. We understand that applying

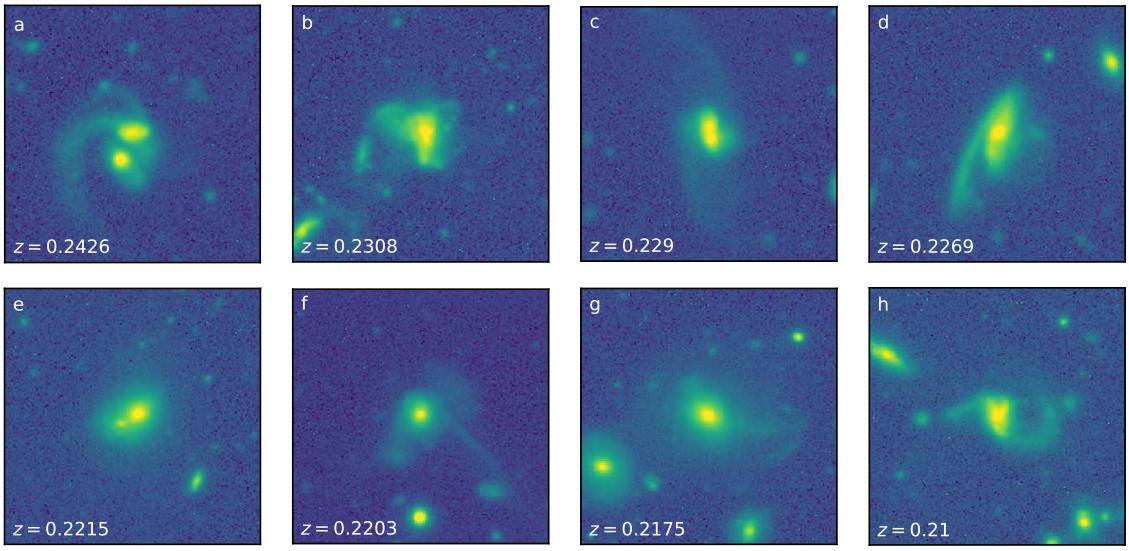


Figure 5.2: HSC r-band images of the 8 low-redshift mergers used to derive our corrections for incompleteness.

our classifier to objects in a different redshift bin than the ones we used to train it is "bad practice", however, to first-order, this quickly gives us a small sample of potentially interacting galaxies from which we choose our visually unambiguous mergers. From this sample, we identify 8 galaxies which show clear signs of merging. Figure 5.2 shows each galaxy along with their photometric redshifts.

We make sure to include mergers with various signatures such as tidal tails, double nuclei, and global asymmetries so as to probe as much of the parameter space of mergers as possible for this correction. We note that the galaxy population at higher redshift is different from that at lower redshift (e.g., there are no massive elliptical galaxies undergoing minor mergers with less massive companions, and the fractions of wet, dry, and mixed mergers are different). By choosing a sample of low-redshift galaxies with which to derive a correction

factor, we are assuming that to first-order we would observe these types of galaxies at higher redshifts, which is most likely not the case. To treat this properly, we could use images of merging galaxies from simulations such as EAGLE (McAlpine et al. 2016) or Illustris (Nelson et al. 2015). We caution the reader that the corrections derived here are only to be viewed as a rough estimate.

We take the following approach to derive the correction factor c at each redshift bin. First, for simplicity we choose to define our redshifts to lie at the centre of each bin. In other words, we force all of our low-redshift galaxies in Figure 5.2 to be at $z = 0.175$ to define a common starting point. The remaining bin centres are therefore at: `zbin1`: $z = 0.325$; `zbin2`: $z = 0.475$; `zbin3`: $z = 0.625$; `zbin4`: $z = 0.775$; and `zbin5`: $z = 0.925$. We take 100×100 pixel cutouts of each low-redshift merger in the HSC r-band (this corresponds to roughly 50×50 kpc at $z = 0.175$). We run each merger through our watershed segmentation pipeline so all other galaxies in the cutout may be masked. In Figure 5.3, we show the resulting masked (with zeros) cutout of galaxy *(h)* from Figure 5.2. It is this image (and several others like it) from which we start when simulating galaxies at each of the higher redshifts.

For each of the 8 galaxies, we manually choose a 100×100 pixel background sky region in the corresponding HSC r-band patch image. This way, when we simulate our high-redshift galaxies, we can re-insert them into the same background we observe for the original galaxy. Next, we consider two effects that redshift has on a galaxy: the apparent magnitude increases, and the angular size on the sky decreases (i.e., the galaxy appears both fainter and smaller).

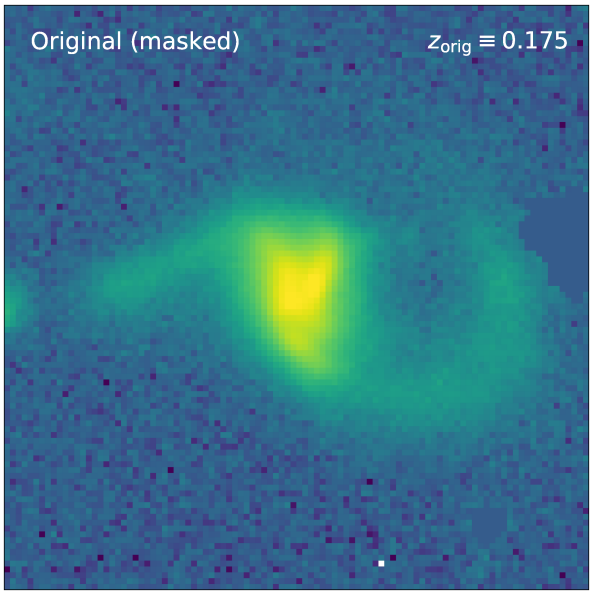


Figure 5.3: Example of a 100×100 pixel ($\sim 50 \times 50$ kpc) HSC r-band image for one galaxy used in our correction for incompleteness. We mask our image with zeros for simplicity. This is the image we artificially redshift.

To make our galaxies fainter, we use the following cosmological relation for *surface brightness* dimming in the monochromatic case (see, for example, Peacock 1999):

$$\mu_{\rm dim} = \mu_{\rm orig} \frac{(1+z_{\rm old})^3}{(1+z_{\rm new})^3}, \tag{5.2}$$

where $z_{\rm old}$ is the original redshift of the galaxy before any artificial redshifting (for all galaxies we use $z_{\rm old} = 0.175$), $z_{\rm new}$ is the redshift at which we wish to simulate our galaxy, and $\mu_{\rm orig}$ and $\mu_{\rm dim}$ are the surface brightnesses (i.e., flux per unit solid angle on the sky) of the original and artificially dimmed galaxies, respectively. In our case, the galaxy image

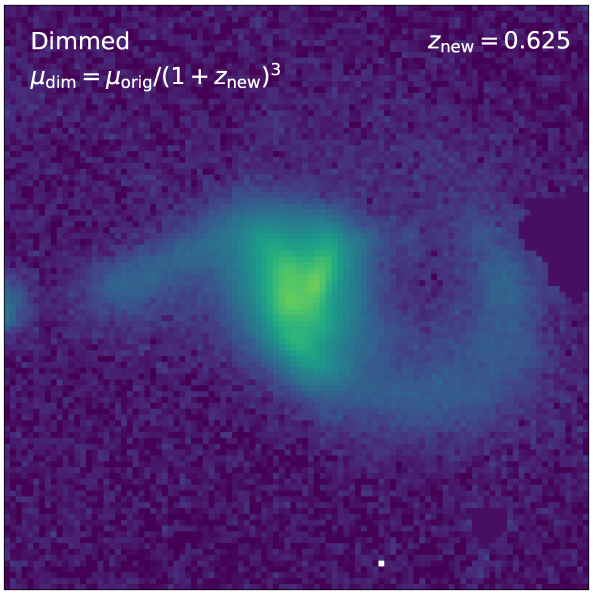


Figure 5.4: Same galaxy as in Figure 5.3, but dimmed to a redshift of $z_{\rm new} = 0.625$. We use Equation 5.3 to dim the fluxes in each pixel. Here, we have *only* applied the dimming to the original image and so it is still the same size (100×100 pixels).

pixels are in flux units (J s^{-1} cm^2 Hz^{-1}), so this relation simply becomes:[1]

$$f_{{\rm dim},i} = f_{{\rm orig},i} \frac{(1+z_{\rm old})^3}{(1+z_{\rm new})^3}, \tag{5.3}$$

where i denotes a single pixel in the image. Figure 5.4 shows the same galaxy as in Figure 5.3, but dimmed to the redshift of our third bin, $z_{\rm new} = 0.625$.

Next, we must resize our galaxies. We do this by rebinning the pixels in our images so that the resulting image is smaller by a factor of the ratio between the *scale length* at the new redshift compared to that of the original. At $z = 0.175$, assuming a flat cosmology with

[1]We note that on our original pass of this methodology, our surface brightness dimming calculations assumed that $z_{\rm old} = 0$, which is not entirely true since we use $z = 0.175$ for our scale lengths in the rebinning process. Due to time constraints, we did not correct this. The flux would, in reality, be ~ 1.6 times greater at each redshift. If this were to cause more galaxies to be classified as mergers, then our correction factor would be lower, thereby resulting in a slightly shallower merger fraction evolution.

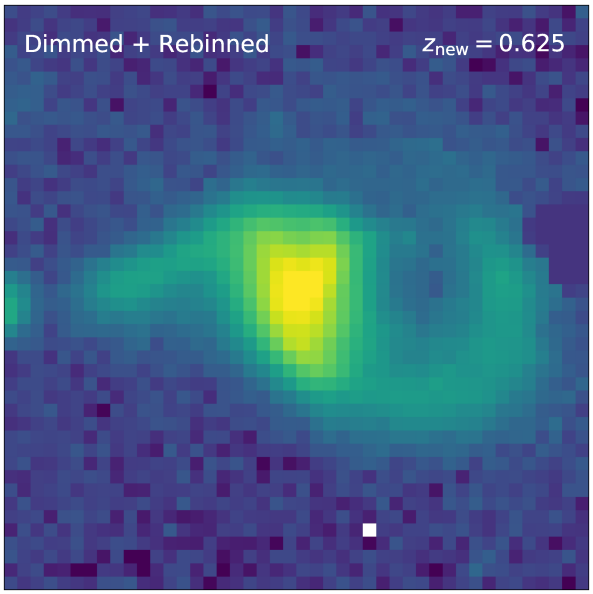


Figure 5.5: Same galaxy as in Figure 5.3, but rebinned to reflect the angular size of the same galaxy at a redshift of $z_{\rm new} = 0.625$. Here, we have applied *both* the rebinning and dimming to the original image. The image shown is 44×44 pixels.

$H_0 = 70$ km s^{-1} Mpc^{-1} and $\Omega_{\rm m,0} = 0.3$, this scale length is 2.967 kpc/$''$. If, for example, we choose $z_{\rm new} = 0.625$, then the scale length is 6.811 kpc/$''$ and our rebinned image needs to be smaller by a factor of ~ 2.3; in other words, our 100×100 pixel image is rebinned to 44×44 pixels. We rebin our images conserving total galaxy *flux*.[2] Figure 5.5 shows the dimmed, rebinned image for our example galaxy *(h)*.

It is important that we "observe" our model galaxy under the same conditions as the original. This is so that we can directly apply our software pipeline for feature calculation to the new model galaxies and be certain that the features we measure are more or less calibrated. To acheive this, we pad our rebinned images with zeros to resize them to 100×100 pixels and then convolve them with the HSC r-band PSF of the patch from which

[2]Code used for rebinning was taken from: http://martynbristow.co.uk/wordpress/blog/rebinning-data/

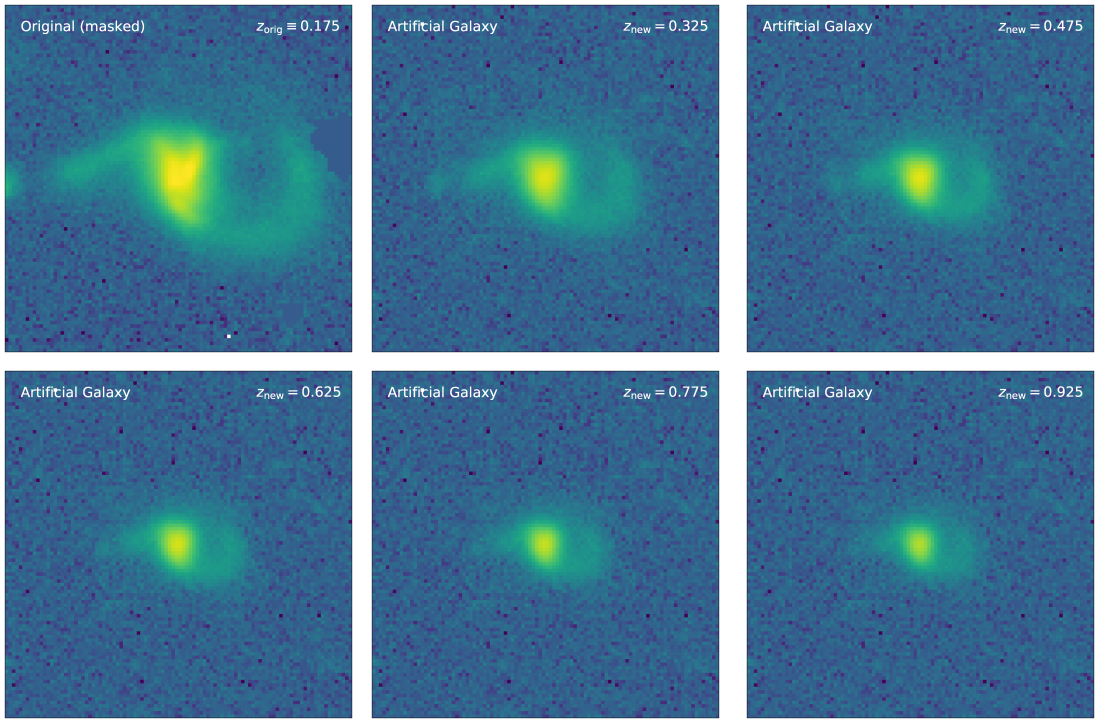


Figure 5.6: Examples of our artificially redshifted galaxies in each of the 5 redshift bins for galaxy *(h)* in Figure 5.3. We plot the original image again for reference. The simulated images here are the result of the dimming, rebinning, convolution, and insertion into a local background of the original galaxy image.

the original galaxy was taken. This simulates the effects due to the telescope optics and observing conditions. We then re-insert this convolved galaxy image into the background defined above to obtain the final artificially redshifted galaxy images. We simulate each of our 8 galaxies across the 5 redshift bins to obtain 40 new images in total. Figure 5.6 shows the final artificially redshifted galaxy images for our example galaxy *(h)* in each of the 5 redshift bins. As the redshift increases, the nuclear structure in this galaxy is completely lost, and there is less definition in the tidal tail as it becomes smaller and fainter. In the highest redshift bin, one might incorrectly visually identify this galaxy to be a spiral non-merger.

In this work, we estimate our correction factor using only the r-band images. By doing this, we are neglecting the *bandpass shifting* that occurs with increasing redshift. At our lowest redshifts, the observed-frame r-band also corresponds roughly to the rest-frame r-band. This follows from the definition of redshift (see Ryden 2016):

$$z = \frac{\lambda_{\rm obs} - \lambda_{\rm em}}{\lambda_{\rm em}} \qquad \text{or} \qquad \lambda_{\rm obs} = (z+1)\lambda_{\rm em}, \tag{5.4}$$

where $\lambda_{\rm obs}$ is the wavelength at which we observe a galaxy, and $\lambda_{\rm em}$ is the rest-frame wavelength of that same galaxy. Therefore at higher redshifts ($z > 0$), $\lambda_{\rm obs} > \lambda_{\rm em}$ and we are no longer probing the rest-frame r-band when we use the (observed-frame) r-band images in our artificial redshifting pipeline. We should instead be using images of *shorter* wavelength, artificially redshifting them, and re-inserting them into the r-band sky background to simulate high-redshift mergers.

The final step in obtaining our correction factor c in each redshift bin is to take our simulated galaxies, calculate their new morphological features, and run them through our RFC to obtain a merger probability $P_{\rm merge}$. We must change the r-band apparent magnitude $r_{\rm AB}$ for each galaxy according to (see, Ryden 2016):

$$m_{\rm AB} = M_{\rm AB} + 5\log_{10}\left(\frac{d_L}{1{\rm Mpc}}\right) + 25, \tag{5.5}$$

where d_L is the luminosity distance corresponding to each redshift bin. We also assume for simplicity that the stellar mass $M_\star$ and star-forming probability $P_{\rm sf}$ do not change for

each of our simulated galaxies. To first order, the majority of a galaxy's stellar mass will be captured by the brightest components of the galaxy and very little will reside in faint tidal structures. We would therefore likely detect a similar stellar mass at each of the higher redshifts. Furthermore, since we are not using a physical quantity, but rather a probability, to represent the star formation activity in our galaxies we assume that it will not change drastically with redshift for a particular galaxy. This, in reality, is not quite true since the specific star formation rate of galaxies evolves with redshift (e.g., Noeske et al. 2007) and this could, in turn, affect the value of $P_{\rm sf}$. To treat this correctly, we would need to apply an evolutionary model to the specific star formation rates of Golob et al. (in prep) and convert these redshifted quantities into the appropriate artificially redshifted star-forming probabilities.

The methodology described above is a gross oversimplification of other well-known artificial redshifting pipelines used in the literature (e.g., the FERENGI code of Barden et al. 2008). Here, we did not account for the effects of *bandpass shifting and stretching* with increasing redshift. We therefore caution the reader that the results from our treatment are only meant as a rough estimate to the true correction that should be applied.

We summarize the results of our incompleteness corrections in Table 5.2 where we report the merger probabilities for each of our 8 galaxies across the 5 redshift bins. We also report the resulting correction factors c (and $1/c$) which we will use in Equation 5.1 to correct our values of $f_{\rm merge}$ for incompleteness. Here, we assume that any galaxy with a merger probability of $P_{\rm merge} \geq 0.5$ would be successfully recovered by our algorithm and thus classified as a merger. Recall that in order to obtain our merger sample in Section 5.1, we

Table 5.2: Merger probabilities for each of our 40 simulated mergers. The instances where a galaxy would be missed by our criterion ($P_{\rm merge} \geq 0.5$) are boldfaced. In the last row, we list the correction factors c for each of the redshift bins.

Galaxy	$P_{\rm merge}$				
	zbin1	zbin2	zbin3	zbin4	zbin5
(a)	0.8852	0.8758	0.8844	0.8843	0.8706
(b)	0.8387	0.8351	0.8223	0.6790	0.6713
(c)	0.8283	0.8180	0.7907	0.6596	0.6120
(d)	0.6987	0.7530	0.7180	0.6317	0.5851
(e)	0.7379	0.6279	0.5158	**0.4429**	**0.4278**
(f)	0.6355	0.5682	**0.4417**	**0.4270**	**0.4202**
(g)	0.5976	0.5329	**0.3864**	**0.3721**	**0.3637**
(h)	0.8235	0.8072	0.7631	0.6114	0.5237
Correction c:[a]	1.00 ± 0.35	1.00 ± 0.35	0.75 ± 0.31	0.63 ± 0.28	0.63 ± 0.28
$1/c$:[b]	1.00 ± 0.35	1.00 ± 0.35	1.33 ± 0.54	1.60 ± 0.72	1.60 ± 0.72

[a] To obtain the error in each c, we use $\sqrt{N}/N_{\rm tot}$, where N is the number of mergers with $P_{\rm merge} \geq 0.5$ and $N_{\rm tot}$ is the total number of mergers considered, which is 8.

[b] We use standard Gaussian error propagation to derive the uncertainties on $1/c$. See, for example, Taylor (1982).

added a layer of visual identification after applying our RFC. Therefore, by assuming a cut in $P_{\rm merge}$ of 0.5, we are subject to high levels of contamination by visual non-mergers. Since the author is biased, we would need to conduct a blind study in which annotators unaware of the context of the problem provide visual classifications.

To illustrate the importance of applying incompleteness corrections, notice that we derive a correction factor of $c = 5/8$ in our two highest redshift bins. This means that within the uncertainties we could be missing between $\sim 10-65\%$ of the mergers at these redshifts, thus causing the evolution in our merger fraction to appear much shallower than it is in reality. The consequences of excluding a correction could skew our understanding of the true role

mergers play in galaxy evolution. A shallower evolution would imply that there were not as many mergers occuring in the past when in reality, we are just not able to detect them.

In this work, we use both HSC *Deep* and *UltraDeep* data. We must therefore consider that the incompleteness correction factor c will be different in each of the two layers simply because the depth of the images affects our ability to detect faint tidal structures. For example, a tidal tail detected in the *UltraDeep* image may not be detected for the same galaxy in the corresponding *Deep* image. Our correction factor should, in reality, be weighted by the relative contributions (i.e., fractional effective survey area) of individual correction factors calculated for each depth:

$$c_{\rm tot} = \frac{A_{\rm UD}}{A_{\rm tot}} c_{\rm UD} + \frac{A_{\rm D}}{A_{\rm tot}} c_{\rm D}, \tag{5.6}$$

where $c_{\rm UD}$ and $c_{\rm D}$ are the individual correction factors for the *UltraDeep* and *Deep* fields, respectively; $A_{\rm tot}$ is the total effective area of the combined survey (~ 20 deg^2); and $A_{\rm UD}$ and $A_{\rm D}$ are the effective areas of the *UltraDeep* and *Deep* layers, respectively.

In this work, we have used 8 galaxies to correct for incompleteness; 2 of which belong to the *UltraDeep* layer, and 6 of which belong to the *Deep* layer. We do not apply the above correction here, however, in the future our correction factor would not only benefit from a larger sample of low-redshift galaxies, but also from corrections calculated for each of the two HSC layers separately to account for this difference in detectability with survey depth.

5.3 The Corrected Merger Fraction Evolution

5.3.1 Power Law Modelling and Confidence Intervals

Using the results of our incompleteness corrections, we can correct our values for the merger fraction in each redshift bin. In Table 5.3, we report the values of the merger fraction in the un-corrected and incompleteness corrected cases. The uncertainties on the un-corrected fractions correspond to only the $\sqrt{N_{\rm merge}}/N_{\rm tot}$ errors on the number counts. The uncertainties on the incompleteness corrected fractions use both $\sqrt{N}/N_{\rm tot}$ errors (for both $f_{\rm m}$ and c) and standard Gaussian error propagation (Taylor 1982).

To model the evolution in the merger fraction, we fit a power law to the data in Table 5.3. We use a least squares algorithm to minimize the χ^2 statistic:

$$\chi^2 = \left(\frac{f_i - f_i^{\rm model}}{\sigma_i} \right)^2 , \tag{5.7}$$

where f_i is the merger fraction at redshift bin centre i, $f_i^{\rm model}$ is the value of the power-law model (Equation 1.3), and σ_i is the uncertainty in the merger fraction. We use the `Python` routine `lmfit.minimize` to perform our fits. The modelling step returns the following best fit parameters of the power law: the local merger fraction f_0, and the power-law index m which describes the *strength* (or steepness) of the evolutionary model. We report the values of these parameters for both the un-corrected and corrected merger fractions (using *only* the data points and their associated errors) at the bottom of Table 5.3. The merger fraction evolution in the corrected case is, as expected, much steeper than the un-corrected evolution.

Table 5.3: Overview of merger fractions in the un-corrected and incompleteness corrected (Equation 5.1) cases. We also report the best fit parameters f_0 and m for each case.

Redshift	$f_{\rm merge}$ (%) (Un-corrected)	$f_{\rm merge,corr}$ (%) (Corrected)
`zbin1`: $0.25 \leq z < 0.4$	2.55±0.16	2.55 ± 0.91
`zbin2`: $0.4 \leq z < 0.55$	2.04±0.13	2.04 ± 0.73
`zbin3`: $0.55 \leq z < 0.7$	2.31±0.12	3.08 ± 1.27
`zbin4`: $0.7 \leq z < 0.85$	2.38±0.14	3.81 ± 1.72
`zbin5`: $0.85 \leq z \leq 1.0$	4.07±0.20	6.51 ± 2.93
Local Merger Fraction (%), f_0:	1.403±0.394	0.949±0.263
Power Law Index, m:	1.271±0.511	2.509±0.684

We can calculate a 68% confidence interval on the power-law model to the incompleteness corrected merger fraction $f_{\rm merge,corr}$. To do this, we perform 1000 bootstrap resamplings (randomly chosen with replacement) of the 8 galaxies we use to derive our incompleteness corrections. At each iteration, we recalculate the correction factors c, the corrected merger fractions $f_{\rm merge,corr}$, along with their uncertainties in each of the redshift bins. We model each bootstrapped sample with a power law as we did above and calculate the $16^{\rm th}, 50^{\rm th}$, and $84^{\rm th}$ percentiles for all 1000 fits to obtain a new best fit line and confidence inverval.

In Figure 5.7 we show our un-corrected (open red circles) and incompleteness corrected (closed red circles) merger fractions. The errors on the un-corrected fractions are much smaller than the corrected points and so we do not plot them here (see instead Table 5.3). The power-law model using only the corrected data points and errors is shown by the red dashed line and its best fit parameters are listed on the plot. Each grey line corresponds to one of the 1000 bootstrapped models, which we used to derive the confidence interval on

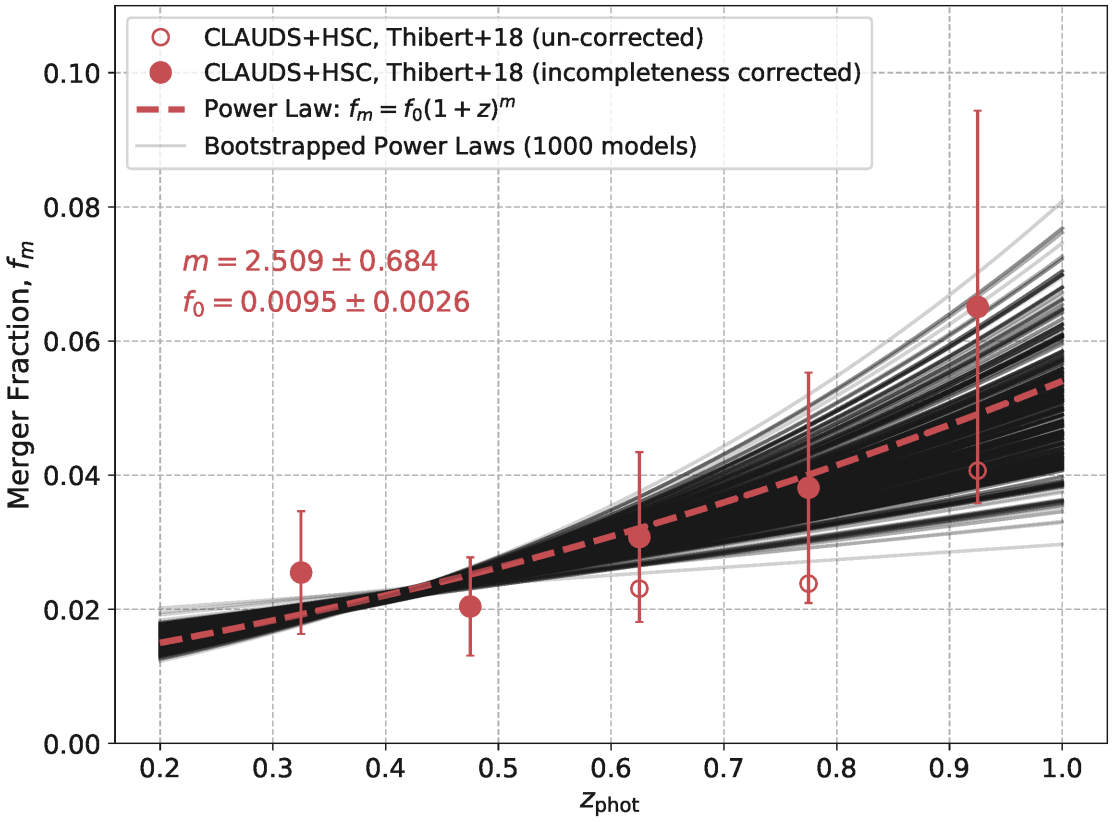


Figure 5.7: The evolution in the merger fraction for our sample of galaxies in the CLAUDS+HSC survey. The un-corrected (open circles) and corrected (closed circles) mergers fractions are shown, along with the power-law model to the corrected merger fraction (red dashed line). The 1000 bootstrapped models used to derive the confidence interval in the fit are shown as grey lines. See text for details on the uncertainties in the corrected merger fraction.

the fit. The large spread in models at higher redshifts is due to this resampling process.

5.3.2 The Evolution in the Merger Fraction from $0.25 \leq z \leq 1.0$

In Figure 5.8, we present the major results of this study: the incompleteness corrected evolution in the merger fraction for galaxies in the CLAUDS+HSC survey with $M_\star \geq 10^{10.5} M_\odot$ and $0.25 \leq z_{\rm phot} \leq 1.0$. The red points, errorbars, and power-law model are the same as

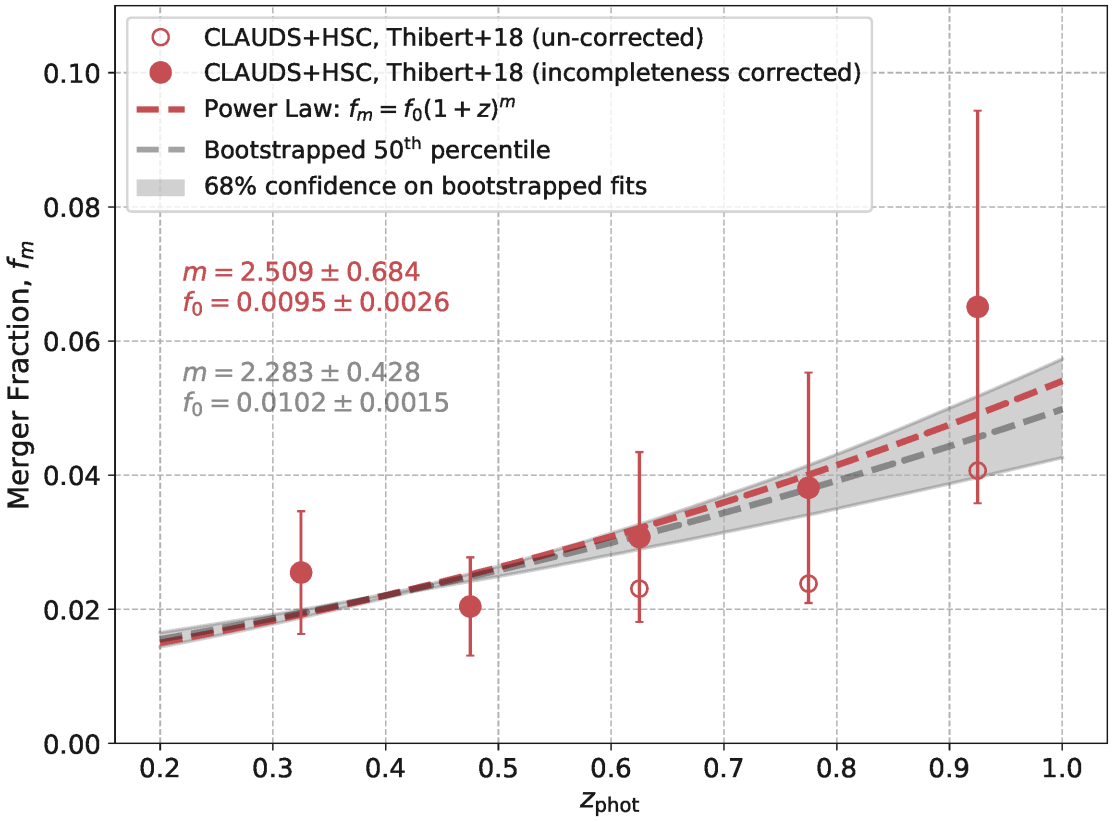


Figure 5.8: The merger fraction evolution for galaxies in the CLAUDS+HSC survey with $M_\star \geq 10^{10.5} M_\odot$ and $0.25 \leq z_{\rm phot} \leq 1.0$. The red open (closed) circles denote the un-corrected (incompleteness corrected) merger fractions we derive in this work. The red dashed line is the power-law model to the corrected merger fraction. The grey dashed line and shaded region denote the 50$^{\rm th}$ percentile and 68% confidence interval on the 1000 bootstrapped fits shown in Figure 5.7.

in Figure 5.7. The results of our bootstrap resampling in Section 5.3.1 are shown as the grey dotted line and the grey shaded regions, which denote the 50$^{\rm th}$ percentile and the 68% confidence interval on the bootstrapped models, respectively. In Figure 5.8, we list the best fit parameters for the corrected merger fraction evolution (using only the points and errors) in red text, and the parameters from the bootstrap resampling in grey text. If we do not bootstrap our incompleteness correction sample, we obtain a slightly higher value for the

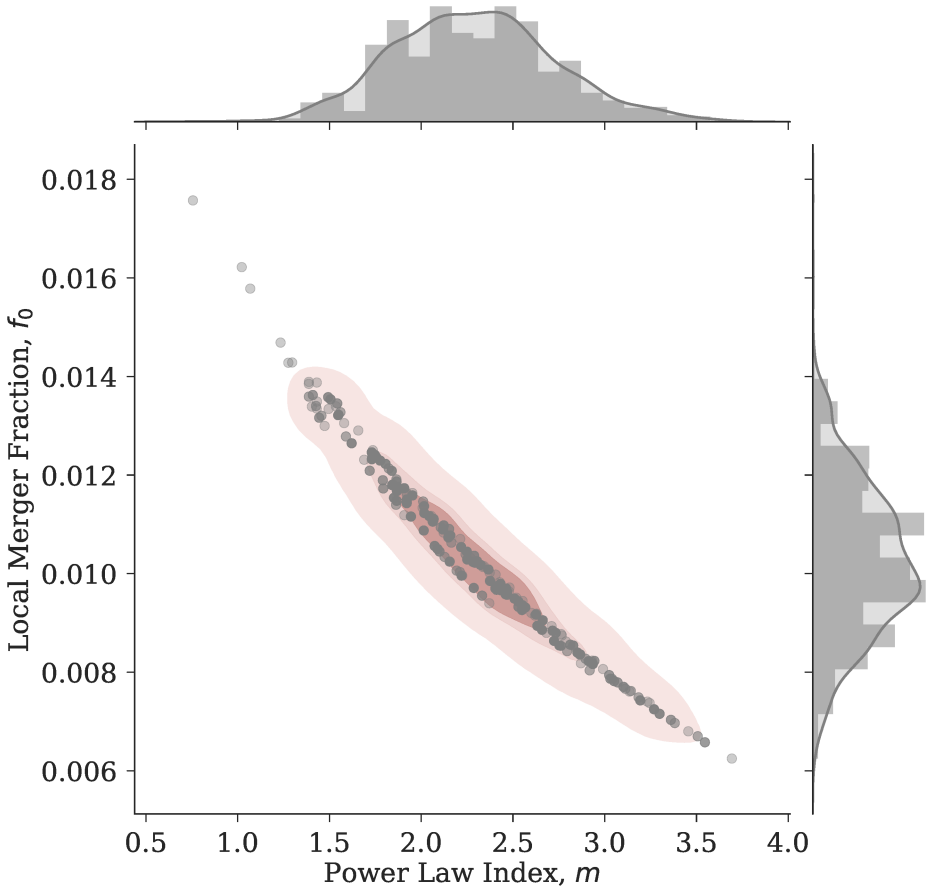


Figure 5.9: Local merger fraction f_0 vs. power-law index m for the 1000 bootstrapped fits shown in Figure 5.7. The grey points show the values themselves and the pink contours show the 68^{th}, 84^{th}, and 99^{th} percentiles of the data. The normalized 1-dimensional distributions (grey histograms and Gaussian kernel density estimators) for each best fit parameter are shown in the margins. The Pearson correlation coefficient for these two parameters is $r = -0.989$.

power-law index m, however, these two methods agree within uncertainty.

The best fit parameters of the 50^{th} percentile fit are determined using the 1000 parameters for each of the individual bootstrapped models. The errors on the parameters denote 1 standard deviation in the distribution of bootstrapped values. To further visualize the results of the bootstrap, we plot our 1000 resampled local merger fractions f_0 as a function of

the power-law indices m in Figure 5.9. The best fit parameters show a strong negative correlation to one another, with a Pearson product-moment correlation coefficient of $r = -0.989$ (see Kaufman & Rousseeuw 1990). We can therefore constrain our values for f_0 and m even more than just using the 1-dimensional uncertainties. This constraint is shown by the pink contours in Figure 5.9, which denote the $68^{\rm th}$, $84^{\rm th}$, and $99^{\rm th}$ percentiles, respectively.

5.4 Comparison with Other Studies

We now investigate how our results for the merger fraction evolution fit into findings from the current literature. We compile the f_0 and m values from several other studies and compare them to our own in Table 5.4. We include examples of studies which select mergers using visual inspection, morphological parameters, and kinematic close pairs and discuss the first six entries in the Table further below.

The first study we consider is that of Bridge et al. (2010) who visually inspected 27,000 galaxies over $\sim$2 deg^2 in the CFHTLS-Deep survey. They used CFHT MegaCam i-band images down to a limiting magnitude of $i_{\rm Vega} = 21.9$ mag to derive the merger fraction evolution for galaxies with redshifts $0.1 \leq z \leq 1.2$ and masses $M_\star \geq 10^{9.5} M_\odot$. They corrected their merger fraction for incompleteness by artificially redshifting 54 low-redshift galaxies using a similar approach to what we take here. Bridge et al. (2010) derives a stronger evolution in the merger fraction ($m = 3.31 \pm 0.22$) than what we find and they estimate a local merger fraction of $f_0 = 0.015 \pm 0.002$, which agrees with our results within the uncertainties. Their power-law index is only consistent with our incompleteness corrected,

non-bootstrapped index $m = 2.51 \pm 0.68$, likely due to the size of our error bars. One question we may ask, however, is why the results of Bridge et al. (2010) show such low fractional errors on their power-law index when compared to our own, despite the fact that our sample of galaxies is $\gtrsim 2\times$ larger and our survey covers $\sim 10\times$ the area. The answer lies in the incompleteness corrections. We noticed that the errors presented in the incompleteness corrected merger fraction of Bridge et al. (2010) (their Figure 6) were very similar to those of their un-corrected fraction and so it is likely that they did not include the uncertainties from their incompleteness corrections. Since the uncertainties in the merger fraction are dominated by those of the incompleteness corrections, their *true* best fit parameters would likely show larger fractional errors, which could affect the values they derive. We therefore caution the reader when interpreting the values of Bridge et al. (2010) since their treatment of errors is likely not as statistically robust as it could be.

In a study by Lotz et al. (2008b), they applied their novel $G - M_{20}$ selection technique to deep rest-frame B-band *HST* ACS images of galaxies in the All-Wavelength Extended Groth Strip International Survey (AEGIS, Davis et al. 2007). They derived incompleteness corrected merger fractions using galaxies brighter than $0.4L_B^*$ with redshifts of $0.2 < z < 1.2$, where L_B^* is the characteristic luminosity in the B band. They do not fit their values to the power-law form for the merger fraction evolution, but claim that their results suggest little to no evolution in the merger fraction out to $z \sim 1.2$, the fraction staying roughly constant at $10 \pm 2\%$ in this range. As suggested in Bridge et al. (2010), their data is consistent with a power-law index of $m < 1$. In the case of both our corrected merger fractions, the results from Lotz et al. (2008b) are inconsistent with our own.

Conselice et al. (2003) measured the merger fraction evolution on a sample of galaxies in the WFPC2 and NICMOS Hubble Deep Field North. They used rest-frame optical images and applied the $A > 0.35$ cut to select mergers. They investigated various stellar mass limits ($\log_{10}(M_\star) \geq 8,\ 9,\ 10\ M_\odot$) and redshift ranges out to $z \sim 3$. Here, we present the results from the limits which best match our own ($z \lesssim 1$ and $M_\star \geq 10^{10} M_\odot$). They do not apply a completeness correction directly to their merger fraction, but instead apply corrections to their Asymmetry values to account for decreasing signal-to-noise, resolution, angular size, and brightness with increasing redshift. The best fit value for their merger fraction evolution index $m = 1.7 \pm 0.3$ is shallower than our own, but still agrees within uncertainty.

Conselice et al. (2009) uses a combination of the $A - S$ plane and $A > 0.35$ selection criterion to identify mergers in the Extended Groth Strip and COSMOS surveys. They use deep HST ACS $F814W$ (I band) images of 21,902 galaxies with stellar masses $M_\star > 10^{10} M_\odot$ and redshifts $0.2 < z < 1.2$ to derive the evolution in the merger fraction. They address the effects of redshift on their morphological parameters: i.e., cosmological surface brightness dimming, the evolution of the CAS parameters with redshift, and the result of probing different rest-frame wavelengths when measuring the parameters in only the $F814W$ band. They also correct for contamination of non-merging galaxies in the parameter space of mergers in the CAS space. Their best fit parameters are given by $f_0 = 0.025 \pm 0.005$ and $m = 2.3 \pm 0.4$. Although their value for the local merger fraction does not agree with our own, the power-law index is in very good agreement with the value we derive from the bootstrapped incompleteness samples. The absolute value for the merger fraction at all redshifts (and locally) will depend on the exact *definition* we use to select mergers. This

can certainly change the normalization (i.e., f_0) of the power law across different studies, while having no effect on the slope m itself. It is therefore not disconcerting in this case that our values of f_0 differ, when the power-law index still agrees because it is the value of m which reveals to us the relative role of mergers in the evolution of galaxies at different redshifts.

Using the close pair method we discussed in Section 1.2.1, Patton et al. (2002) measures the evolution in the merger fraction using a sample of 4,184 galaxies from the Canadian Network for Observational Cosmology (CNOC2) Field Galaxy Redshift Survey (Yee et al. 2000) with redshifts $0.12 \leq z \leq 0.55$ and absolute magnitudes $M_B \leq -18$. They impose the following criteria on the maximum projected separation and line-of-sight velocity difference: $R_{\rm proj} \leq 50h^{-1}$kpc, and $\Delta v \leq 500$ km s^{-1}. They derive a merger fraction evolution power-law index of $m = 2.3 \pm 0.7$, which is in close agreement to our own. In a separate study, de Ravel et al. (2009) use a sample of 6,464 galaxies to select physically close, spectroscopically confirmed galaxy pairs in the VIMOS VLT Deep Survey (VVDS, Le Fevre et al. 2004). They measure the evolution in the pair fraction for varying levels of both the projected separation ($R_{\rm proj} = 20 - 100h^{-1}$ kpc) and line-of-sight velocity difference ($\Delta v \leq 500,\ 1000,\ 2000$ km s^{-1}). They also consider the effects on the pair fraction when considering faint ($M_B \leq -18$) and bright ($M_B \leq -18.77$) galaxy samples. We report the results they obtain by following the same selection criteria as Patton et al. (2002). In the case of their faint sample, the best fit parameters they obtain are $f_0 = 0.0219 \pm 0.0118$ and $m = 4.07 \pm 0.95$. Their local pair fraction agrees with our local merger fraction, however, their fractional uncertainties are over 50%. Their power-law index agrees with only our non-bootstrapped

value within uncertainty.

We must be careful when comparing the results of close pair studies to those of morphological studies for various reasons. For example, the observability timescales for close pairs are different from morphologically disturbed galaxies (see Section 3.4 in Conselice et al. 2009). Furthermore, close pairs are only a proxy for systems that *might* merge in the future, whereas morphoogical disturbances are a nearly 100% sure sign of a merger in progress. The number counts derived from close pairs and morphologically disturbed systems would not only be different, but would be probing slightly different physical scenarios. The evolution in the close pair fraction is therefore not directly comparable to the results we derive from morphological studies. Nevertheless, if we assume to first order that the fractions derived from morphological and pair studies are both more or less probing the rate with which galaxies merge, then any differences in number counts should only affect the value derived for the local merger fraction f_0. We should not expect the value of m to change drastically in either case, provided that this assumption is true.

There are several other studies which derive the evolution in the merger and close pair fraction– e.g., Le Fèvre et al. 2000; Cassata et al. 2005; Bridge et al. 2007; De Propris et al. 2007; Kampczyk et al. 2007; Kartaltepe et al. 2007; Lin et al. 2008; Bundy et al. 2009; Chou et al. 2011; Man et al. 2012; Xu et al. 2012; López-Sanjuan et al. 2013; Lackner et al. 2014; López-Sanjuan et al. 2015; Wen & Zheng 2016; Mundy et al. 2017.

Considering all of the studies presented here, our value for the power-law index m agrees well with most within uncertainty, however some close pair studies derive a much shallower evolution than our own. From studies involving CDM models, we would expect a merger

fraction dependence $\propto (1+z)^3$ (see Gottlöber et al. 2001 and discussion in Berrier et al. 2006). Most studies in Table 5.4 including our corrected, non-bootstrapped merger fraction agree with this theoretical prediction, however the average evolution across these studies is $m \approx 2.4$. This could imply that either the simulations need to be modified, or the high-redshift end of the merger fraction evolution (including the necessary incompleteness corrections) need to be constrained further.

5.5 Interpretation & the Fractional Merger Rate $\mathfrak{R}_{\mathbf{merge}}$

We derive the evolution in the merger fraction with the purpose of understanding how galaxies evolve through cosmic time. Firstly, we can determine the significance of our results to supporting/refuting any one of the merger scenarios. Our bootstrapped, incompleteness corrected merger fraction evolves as $\propto (1+z)^{2.28\pm0.43}$, suggesting that it might fall into the category of "moderate" evolution. If we wish to rule out the "mild or non-evolving" merger scenario, which is sometimes characterized by $m < 1.5$ (see Bridge et al. 2010), we can estimate the significance of our result using standard Gaussian statistics. We find that our value for the power-law index is *inconsistent* with a mild or non-evolving merger scenario with 96.6% confidence; in other words, at the 2.15σ confidence level. Similarly, our corrected merger fraction evolution without the bootstrap ($m = 2.509\pm0.684$) is statistically inconsistent with a power-law index of $m < 1.5$ with 93.1% confidence, or at the 1.82σ confidence level.

Table 5.4: Summary of current results in the literature for the merger fraction evolution. Some studies do not report f_0.

Study	Method	Merger Fraction Evolution		Redshift Range	Mass/Luminosity Limit[a]
		Local f_0 (%)	Index m		
Bridge et al. (2010)	Visual Inspection	1.5±0.2	3.31±0.22	$0.1 < z < 1.2$	$10^{9.5} M_\odot$
Lotz et al. (2008b)	$G - M_{20}$ plane	–	< 1	$0.2 < z < 1.2$	$0.4 L_B^*$
Conselice et al. (2003)	$A > 0.35$	~ 0	1.7±0.3	$z \lesssim 1$	$10^{10} M_\odot$
Conselice et al. (2009)	CAS space	2.5±0.5	2.3±0.4	$0.2 < z < 1.2$	$10^{10} M_\odot$
Patton et al. (2002)	Kinematic Close Pairs	–	2.3 ± 0.7	$0.12 \leq z \leq 0.55$	$-21 \leq M_B \leq -18$
de Ravel et al. (2009)	Kinematic Close Pairs	2.19 ± 1.18	4.07 ± 0.95	$z \lesssim 1$	$M_B \leq -18$
Le Fèvre et al. (2000)	Visual Inspection	2.1 ± 0.4	3.4 ± 0.6	$0.05 \leq z \leq 1.2$	$M_B \leq -17.5$
Le Fèvre et al. (2000)	Pair Counts	1.9 ± 0.4	3.25 ± 0.63	$0.05 \leq z \leq 1.2$	$m_{B,AB} \geq 20.5$
Cassata et al. (2005)	Visual, n, CAS	2.2 ± 0.2	2.2 ± 0.3	$0 < z < 2.5$	$i < 24.5$
Bridge et al. (2007)	Kinematic Close Pairs	7.7 ± 4.5	2.12 ± 0.93	$0.2 < z < 1.3$	$-21 \leq M_B \leq -19$
Kampczyk et al. (2007)	Simulations, Visual	–	3.8 ± 1.2	$z \lesssim 0.7 - 1.2$	$I_{AB} < 24$
Kartaltepe et al. (2007)	Kinematic Close Pairs	–	3.1 ± 0.1	$z < 1.2$	$M_V \lesssim -19.8$
Lin et al. (2008)	Kinematic Close Pairs	–	0.41 ± 0.20	$0.2 < z < 1.2$	$-21 < M_B < -19$
Bundy et al. (2009)	Kinematic Close Pairs	–	1.6 ± 1.6	$z \lesssim 1.2$	$10^{11} M_\odot$
Chou et al. (2011)	Visual Identification	–	2.0 ± 0.3	$z \lesssim 0.7$	$10^{9.5} M_\odot$
Man et al. (2012)	Kinematic Close Pairs	7.0 ± 4.0	0.6 ± 0.5	$0 \leq z \leq 3$	$10^{11} M_\odot$
Xu et al. (2012)	Kinematic Close Pairs	1.3 ± 0.1	2.2 ± 0.2	$z \lesssim 1$	$10^{10} - 10^{11.5} M_\odot$
López-Sanjuan et al. (2013)	Kinematic Close Pairs	–	3.95 ± 0.12	$0.9 < z < 1.8$	$10^{10} - 10^{10.5} M_\odot$
Lackner et al. (2014)	Automated Detection	–	3.8 ± 0.9	$0.25 < z \leq 1.0$	$10^{10.6} M_\odot$
López-Sanjuan et al. (2015)	Kinematic Close Pairs	0.43 ± 0.05	2.7 ± 0.5	$z \leq 1$	$M_B \leq -20 - 1.1z$
Wen & Zheng (2016)	$A_o - D_o$ plane	0.64 ± 0.13	2.0 ± 0.4	$0.2 \leq z \leq 1.0$	$10^{9.5} M_\odot$
Mundy et al. (2017)	Kinematic Close Pairs	2.4 ± 0.4	$0.85^{+0.19}_{-0.20}$	$0.005 < z < 3.5$	$10^{10} M_\odot$
This work (corrected)	RFC+Visual Inspection	0.95±0.26	2.51±0.68	$0.25 \leq z \leq 1.0$	$10^{10.5} M_\odot$
This work (bootstrap)	RFC+Visual Inspection	1.02±0.15	2.28±0.43	$0.25 \leq z \leq 1.0$	$10^{10.5} M_\odot$

[a] The selection of mergers based on a mass limit vs. a luminosity/absolute magnitude limit may affect the results and so the reader should be wary when directly comparing studies which use different selection techniques.

Our value for m is an especially important addition to the current literature since we determine the morphologies of galaxies out to higher redshift, which helps to further constrain the merger fraction evolution at these epochs. In addition we use an automated approach to detect potential mergers, a method which is becoming increasingly popular as data volumes are sure to grow with future ground-based efforts such as the Large Synoptic Survey Telescope (LSST) and the Thirty Meter Telescope (TMT).

We can use our results to estimate the fractional merger rate $\Re_{\rm merge}$. Recall that the fractional merger rate in Equation 1.6 was dependent on an estimate of the timescale of observability of merger signatures. Ideally, we would derive an average, redshift-dependent observability timescale $\langle T_{\rm obs}(z) \rangle$ as is done in Lotz et al. (2011) using simulations. For simplicity we use the results from their Figure 9 to crudely estimate a (constant) timescale of $T_{\rm obs} \approx 0.5$ Gyr, despite the fact that it is highly dependent on both redshift and the methodology used to select mergers. This estimate is also used in Jogee et al. 2009 for their derivation of the merger rate. Applying this timescale to our merger fraction, the fractional merger rate becomes:

$$\Re_{\rm merge} = \frac{f_{\rm merge}}{T_{\rm obs}} = 2 \times f_0 (1+z)^m \quad [{\rm Gyr}^{-1}]. \tag{5.8}$$

Using the results from our bootstrap resampling in the above Equation, we derive the evolution in the fractional merger rate to be $\propto (1+z)^{2.28 \pm 0.43}$ implying that the significance of merging events increases with lookback time.

We can also estimate the *number* of merging events a galaxy would undergo in the range of lookback times covered by our study. We convert our redshifts to lookback times t_L using (see Hogg 1999):

$$t_L = t_H \int_0^z \frac{dz'}{(1+z')E(z')}, \tag{5.9}$$

where $t_H = 1/H_0 = 1/(70 \text{ km s}^{-1} \text{ Mpc}^{-1}) \approx 14$ Gyr is the Hubble time, and $E(z)$ is the time derivative of the logarithm of the cosmological scale factor $a(t)$ (Hogg 1999). We use the `Python` routine `astropy.cosmology.Planck13.lookback_time` which uses cosmological results from the Planck telescope (Planck Collaboration et al. 2014) to perform this conversion for each sampled redshift between $0.25 \leq z \leq 1.0$. Our study covers a range of lookback times between $t_L = 3.0$ Gyr at $z = 0.25$ and $t_L = 7.9$ Gyr at $z = 1.0$.

To estimate the number of mergers occuring between $0.25 \leq z \leq 1.0$, we must integrate the fractional merger rate $\mathfrak{R}_{\text{merge}}$ between $t_L = 3.0$ Gyr and $t_L = 7.9$ Gyr. Applying the Simpson's rule for numerical integration (`scipy.integrate.simps` in `Python`) to our fractional merger rate we find that galaxies with masses greater than $10^{10.5} M_\odot$ undergo roughly 0.287 merging events from redshift $z = 1.0$ to $z = 0.25$. This result is consistent with the value derived by López-Sanjuan et al. (2009) who uses *HST* imaging to select mergers based on morphological asymmetries. They derive a value of ~ 0.2 merging events for galaxies with masses $M_\star \geq 10^{10} M_\odot$ and redshifts between $0 < z < 1$. If we apply these same limits to our integration, we find 0.367 merging events per massive galaxy since $z = 1$. They also note that more massive systems tend to experience more merging events as was found in Bluck et al. (2009) for $M_\star \geq 10^{11} M_\odot$ galaxies. Both López-Sanjuan et al. (2009)

and Bluck et al. (2009) report the expected number of mergers between $0 < z < 3$. For $M_\star \geq 10^{10} M_\odot$ galaxies, López-Sanjuan et al. (2009) derive $1.2^{+0.4}_{-0.2}$ mergers to occur since $z = 3$. Similarly, for $M_\star \geq 10^{11} M_\odot$ galaxies, Bluck et al. (2009) expect $1.8^{+0.6}_{-0.4}$ mergers to occur between $0 < z < 3$. If we extrapolate our fractional merger rate evolution to this range, we derive ~ 1.2 merging events since $z = 3$. The results we derive for our mass range are therefore in accordance with findings in the current literature. Generalizing our results in terms of the local merger fraction f_0 and the observability timescale $T_{\rm obs}$ to obtain the number of merging events over the limits of our survey ($0.25 \leq z \leq 1.0$):

$$N_{\rm m}(0.25, 1) = \frac{f_0}{T_{\rm obs}} \times 14.07 \text{ Gyr}. \tag{5.10}$$

5.6 Caveats of the Methodology & Future Work

There are several caveats to the results of this work, which we address individually below:

(1) Biases in the calculation of segmentation maps. The most obvious bias introduced by our segmentation pipeline is the inclusion of all bordering segments when assigning pixels to the primary. As mentioned before, we do this so as to not exclude any close companions with distinct detections in the CLAUDS+HSC catalog. Keep in mind that these companions could either be physically associated with the primary (true mergers/close pairs), or simply be galaxies in chance projections or crowded fields. We did attempt to minimize crowded fields by imposing a maximum level of masking allowed in our primary segmentation maps (see Section 3.1.3). Given that the RFF and M_{20} parameters are the

two most important features in our forest, the inclusion of false companions to the primary segmentation map would artificially increase both RFF and M_{20} and likely cause these systems to occupy a similar parameter space to the merging population. Through our secondary level of visual inspection, we eliminate most of these misclassifications, however even then, biases in the human annotator could increase the inclusion of these systems, thus artificially increasing the merger fraction we derive.

Another, more subtle bias in our segmentation step involves the quality of the imaging data as well as the morphological nature of the galaxies themselves. Lotz et al. (2004) assigned pixels to their galaxies on the basis of a threshold defined by the mean surface brightness at the Petrosian radius ($\mu(r_p)$). Law et al. (2007) found that for more nebulous galaxies, the segmentation maps created using the same method as Lotz et al. (2004) tended to include more sky (background) pixels at the Petrosian radius. As a result, the Gini coefficients for more nebulous galaxies were artificially increased due to the added noise in the segmentation maps. Even though we apply more robust methods for galaxy segmentation in the current study, we too employ a threshold when assigning pixels to our galaxies. Our methods are therefore likely subject to some level of background contamination. We also note that bright galaxies with more concentrated light profiles (which also manifests itself as a higher value for the probability $P_{\rm star}$ from Golob et al., in prep) are subject to PSF effects. In this case, galaxy light is artificially spread to larger distances from the galaxy centroid, resulting in a "halo" effect. This extra light causes the segmentation maps of these galaxies to appear more extended. Since several of our most important morphological parameters are calculated using segmentation maps, we must understand the potential biases and exercise

caution when interpreting parameters derived using them.

(2) Biases while performing visual identifications. In this work, we rely heavily on visual classification in order to generate a training sample (Section 4.2.2) and further clean our sample of potential mergers to estimate the merger fraction (Section 5.1). For the bulk of this work, only one human annotator (the author) provided classifications; the citizen efforts were used as a rough calibration of the author's responses and were otherwise not used to produce the final results. The issues we face are similar to those discussed in Bridge et al. (2010). In some cases, the responses of the author and the citizens were wildly different. In addition, when reclassifying galaxies several months later, the author's responses were sometimes contradictory to their own previous classifications. Therefore, including more annotators per galaxy and using a weighted vote to obtain the final morphological classification, much like in the Galaxy Zoo project, would reduce outlying classifications from any single person. Goulding et al. (2018) minimizes the issue of opposing annotator responses in the context of the images on which the classifications are performed. In their study, they perform visual classifications using rest-frame 3-colour images. By using more than a single band at a time, they are able to emphasize merger signatures probed by different rest-frame wavelengths; e.g., clumpy star formation appears at bluer wavelengths, while the diffuse envelope around a late stage merger would most likely be seen at redder wavelengths. In cases where multiple merger signatures are at play, it would be useful to include more than one rest-frame wavelength so that important structures are not lost in the visual identification process.

(3) Using only the r-band images to calculate morphological parameters and ignoring the evolution of our morphological parameters with redshift. In this work, we use a single band (the HSC r-band) to probe galaxy morphology across different epochs. When considering galaxies at different redshifts, we must remember that observing all galaxies in a single bandpass does not mean that we are observing them all at the same rest-frame wavelength. By only using the r-band to estimate our morphological parameters, we introduce a bias where *different* populations of stars are being probed in each of our galaxies. In other words by using the r-band, lower redshift galaxies will be seen at redder rest-frame wavelengths, which roughly corresponds to older stellar populations. Conversely, higher redshift galaxies will be seen at bluer rest-frame wavlengths and we will be probing younger stellar populations in these galaxies.

It has been shown that the values of morphological parameters can change slightly depending on the rest-frame wavelength probed– see, for example, Hibbard & Vacca (1997); Windhorst et al. (2002); Papovich et al. (2003); Taylor-Mager et al. (2007); Conselice et al. (2003, 2008, 2009). There are two ways we can treat this bias. The first would be to simply use different bands at different redshifts. For example, if we wish to observe our galaxies at a rest-frame wavelength of $\sim$ 4000 Å, then we need to use the g-band for galaxies at redshift $z \sim 0.2$, the r-band for galaxies at redshift $z \sim 0.55$, the i-band for galaxies at redshift $z \sim 0.9$, and so on. This treatment assumes that a large suite of broadband filters is available, which in the case of the HSC survey is true, however, the imaging data in the redder filters suffer from decreasing signal-to-noise. In addition, the limiting magnitudes of each band in the HSC imaging data are different and so we would need to carefully define

our parent galaxy sample such that we are probing the same depths in each of the bands. The differing quality of imaging data across bandpasses would also result in the need to carefully redefine the segmentation pipeline thresholds for each filter.

In works such as Conselice et al. (2003, 2009), a different approach is taken where the same filter is used to calculate morphological parameters of all the galaxies, but instead a *morphological k-correction* is applied to the parameters themselves. This is done by estimating the amount by which a parameter changes based on the rest-frame wavelength used to perform the measurement. For example, Conselice et al. (2008) find that at $z < 0.75$, if we wish to probe rest-frame optical wavelengths, we must change the Asymmetry parameter according to $\delta A_{k-\mathrm{corr}}/\lambda = -0.30\mu\mathrm{m}^{-1}$. In the future, similar approaches to this could be taken to correct the values for our parameters.

(4) Considering only a finite feature space with which to train our RFC. In this study, we use 14 features to describe the morphologies and general properties of our galaxies. The benefit to feature engineering in machine learning is that the dimensionality of the problem, and therefore computational time, is significantly reduced. In theory, we would like to include as much information about our galaxies as possible. One approach to achieving this could be to add more features to our pipeline. In a few recent studies, physically motivated features were defined with the specific goal of finding mergers. For example, in an effort to select merging galaxies with long tidal tails, Wen et al. (2014) introduced two new statistics which are sensitive to asymmetric features in the outskirts of galaxies: the *outer asymmetry* A_{o} and the *centroid deviation* D_{o}. Pawlik et al. (2016) modified the Asymmetry parameter A by removing its dependence on the pixel flux. They claim

that their new parameter, the *shape asymmetry* A_S, is more sensitive to faint, asymmetric tidal features. Freeman et al. (2013) devised a new set of non-parametric morphological indicators called the *multimode* M, *intensity* I, and *deviation* D statistics. The creation of these new parameters was motivated by claims that the C, A, S, G, and M_{20} parameters do not perform well under decreasing resolution, galaxy size, and signal-to-noise (i.e., with increasing redshift, see Conselice et al. 2000b; Lotz et al. 2004). Including these statistics in our Random Forest would certainly provide more information on the morphologies and light distributions of our galaxies. This would allow us to obtain a cleaner sample of mergers at the training stage, thereby reducing the amount of secondary visual inspection needed.

By choosing a finite list of features with which to describe our galaxies, we are limiting ourselves to only identifying galaxies which fall into certain regions of the parameter space we have already defined. To first order, this is a reasonable approach, however by doing this, we are assuming that *all* mergers fit into predefined categories within our carefully engineered feature space, which may in fact not be the case. In the last year, a study conducted by Ackermann et al. (2018) addressed this issue by using another supervised machine learning approach: the combination of transfer learning and a deep convolutional neural network (CNN, see their paper for details on the algorithm). In short, they trained their algorithm using a sample of galaxies from the 7$^{\text{th}}$ Data Release of the Sloan Digital Sky Survey (SDSS DR7, Abazajian et al. 2009), labelled using visual identifications from the Galaxy Zoo project. For their input, they simply used the RGB JPEG images of their galaxies from the SDSS online image cutout service.[3] In the case of a CNN, the features used

[3] http://skyserver.sdss.org/dr12/en/help/docs/api.aspx

to identify mergers are *automatically* identified by the algorithm to be those which provide the best results. By doing this, the algorithm uses *all* of the information contained in the galaxy image. They claim that their method outperforms traditional automated detection methods returning an $F1$ score of 0.97 (i.e., a completeness of 0.96 and a purity of 0.97). One could imagine applying similar methods to pseudo rest-frame three-colour images of the galaxies in the CLAUDS+HSC catalog to obtain a more accurate estimate of the evolution in the merger fraction.

(5) Training our RFC on galaxies which probe a limited parameter space. In engineering a training sample, the basic assumption is that its properties well represent those of the sample to which the trained classifier will be applied. In Figure 4.2 of Section 4.2.2, we acknowledged that the galaxies in our training set did *not* probe the exact same parameter space as the galaxies in our parent sample; in particular, fainter galaxies (both mergers and non-mergers) were underrepresented at all redshifts. Faint, high-redshift mergers do not show the complex structures indicative of merger activity and to a human annotator, they would be visually ambiguous. To treat this issue and improve the results of our classifier, we could simulate galaxies with fainter magnitudes at high redshift using similar methods to Section 5.2 and include them as part of our training sample. This was suggested in a different context by Golob et al. (in prep).

(6) Using a small number of objects for the incompleteness corrections. As we saw in Section 5.4, our results agreed with most others within uncertainty, however, this is not surprising since the errors on our best fit parameters are so high. The uncertainty in our merger fraction evolution is dominated by the uncertainties in the incompleteness

corrections. Therefore, in order to place better constraints on the evolution of the merger fraction, we would need to include more galaxies in our incompleteness corrections. For example, in our highest redshift bin $c = 5/8$ (i.e., 5 of 8 galaxies were detected as mergers). We currently derive an error of $\sigma_c = \sqrt{5}/8 = 0.28$ and from Gaussian error propagation $\sigma_{1/c} = 0.72$ (see Table 5.2). Say we were to find 10× as many galaxies with which to perform our incompleteness corrections. Then if we assume that 50 out of the 80 mergers would be recovered by the algorithm, the corresponding errors would significantly improve ($\sigma_c = 0.09$, $\sigma_{1/c} = 0.23$). The fractional error on the merger fraction in the highest redshift bin would decrease from 45% to only 15%. If we simulate the effects of this increase in sample size on our data by simply changing the numbers of galaxies in our incompleteness sample, the fractional error for the bootstrapped best fit parameter f_0 would decrease from $\sim 15\%$ to $\sim 6\%$, while that of the power-law index m would decrease from $\sim 19\%$ to $\sim 7\%$. Even though this is a very crude approximation, it is clear that increasing the number of galaxies in our imcompleteness sample would drastically improve our results. In the future, this improvement could be accomplished by extending our search for unambiguous mergers at low-redshift to the HSC *Wide* layer, which covers $\sim$ 1400 deg^2 in its entirety. Even though the data in this layer of the survey are much shallower ($r \approx 26$ mag) than the *Deep* and *UltraDeep* layers ($r \approx 27$ and $r \approx 28$ mag, respectively), this would likely have little effect on galaxies at low redshift above our mass limit.

(7) Ignoring values derived for the local merger fraction f_0 when fitting a power law. Conselice et al. (2009) investigated the use of a prior at $z \sim 0$ in the derivation of the merger fraction evolution. Specifically, while performing their power-law fit, they

did not include the local merger fraction f_0 as a free parameter, but rather forced it to be an empirically derived quantity found using low-redshift galaxy samples. In their case, they used a value of $f_0 = 0.009$ from De Propris et al. (2007). By including f_0 as a prior, the value for their power-law index changed from $m = 2.3 \pm 0.2$ to $m = 3.8 \pm 0.2$, which illustrates the importance of including local estimates for the merger fraction in derivations of the total merger fraction evolution. As a next step, we could include such a prior to investigate its effect on our merger fraction evolution at higher redshifts. Alternatively, if not used as a prior it would still be beneficial to include an estimate of the local merger fraction as an independent data point since small volumes in the *Deep* and *UltraDeep* layers at low redshift cause our merger sample to be incomplete below $z < 0.25$.

Chapter 6

Conclusions

Using a sample of massive ($M_\star \geq 10^{10.5} M_\odot$) galaxies over $\sim$ 20 deg^2 in the *Deep* and *UltraDeep* layers of the CLAUDS+HSC survey, we combined an automated Random Forest classification algorithm with visual inspection to derive the evolution in the galaxy-galaxy merger fraction from $0.25 \leq z_{\rm phot} \leq 1.0$. The morphological features of our galaxies were calculated using deep HSC r-band images and a sample of galaxies with unambiguous visual classifications was used to train a Random Forest Classifier. Using the probabilities that each of our galaxies are undergiong a merger, we compared the performance of our Random Forest Classifier to standard 1- and 2-dimensional approaches to merger classification. Finally, we simulated the effects of redshift on the detectability of merger signatures in our galaxy images and applied corrections for incompleteness to obtain our final estimates of the evolution in both the merger fraction and the fractional merger rate.

The major results of this work are summarized as follows:

1. We compared the results of our Random Forest Classifier for two carefully chosen thresholds in the merger probability ($P_{\rm merge} = 0.5$ and 0.7) to those obtained using the $A-S$ techniques of Conselice et al. (2003) and the $G-M_{20}$ technique of Lotz et al. (2008b). From this, we have shown that automated, higher dimensional classification schemes perform much better than a few more traditional 1- and 2- dimensional approaches to merger selection. This is especially outstanding considering the fact that we compared these tradiional methods to our "raw" classifier, without the second layer of visual identification we later imposed.

2. We applied our Random Forest Classifier to the full ~ 20 deg^2 in the CLAUDS+HSC dataset, and visually inspected merger candidates with $P_{\rm merge} \geq 0.5$ to obtain a more or less pure, $\sim 90\%$ complete sample of interacting galaxies. Assuming that the merger fraction evolution is parameterized by a power law of the form $f_{\rm m} = f_0(1+z)^m$ we derived two sets of best fit parameters from our incompleteness corrected merger fraction evolution. We found in the two cases $(f_0, m) = (0.0095 \pm 0.0026, 2.509 \pm 0.684)$ for just the data points and their associated errors, and $(f_0, m) = (0.0102 \pm 0.0015, 2.283 \pm 0.428)$ for 1000 bootstrap resampled realizations of the galaxy sample we used for incompleteness corrections.

3. The results we obtained for massive ($M_\star \geq 10^{10.5} M_\odot$) galaxies were statistically inconsistent with a mild or non-evolving ($0 < m \lesssim 1.5$) merger fraction evolution with $\sim 93-97\%$ confidence ($\approx 2\sigma$ confidence level). Although we cannot rule out the

possibility of a mildly evolving merger fraction within our uncertainties, our values may suggest that there were more mergers in the past. To first order, we assumed a constant observability timescale of $T_{\rm obs}$ =0.5 Gyr and found that the fractional rate with which galaxies merge is directly proportional to our estimates of the merger fraction, suggesting that the role of mergers in the distant Universe was, to some degree, more significant than it is at the present epoch. Integrating the fractional merger rate with lookback time, we find that massive ($M_\star \geq 10^{10.5} M_\odot$) galaxies undergo ~ 0.29 merging events between $0.25 \leq z \leq 1.0$.

The methods presented in this work can be viewed as a *pilot study* to test the feasilibity of accurate merger fraction estimates using the CLAUDS+HSC imaging data and photometrically derived quantities. As discussed, there is much that can be done to improve the results of this work. With future HSC data releases, the images in the *Deep* and *UltraDeep* layers will become much deeper. We will be able to probe to fainter apparent magnitudes, lower masses, and higher redshifts, allowing us to further constrain the evolution in the galaxy-galaxy merger fraction and merger rate.

Bibliography

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

Abraham, R. G., van den Bergh, S., Glazebrook, K., et al. 1996, ApJS, 107, 1

Abraham, R. G., van den Bergh, S., & Nair, P. 2003, ApJ, 588, 218

Ackermann, S., Schawinski, K., Zhang, C., Weigel, A. K., & Turp, M. D. 2018, MNRAS , 479, 415

Aihara, H., Arimoto, N., Armstrong, R., et al. 2017, ArXiv e-prints

Barden, M., Jahnke, K., & Häußler, B. 2008, ApJS , 175, 105

Berrier, J. C., Bullock, J. S., Barton, E. J., et al. 2006, ApJ , 652, 56

Bershady, M. A., Jangren, A., & Conselice, C. J. 2000, AJ, 119, 2645

Bertin, E. 2011, in Astronomical Society of the Pacific Conference Series, Vol. 442, Astronomical Data Analysis Software and Systems XX, ed. I. N. Evans, A. Accomazzi, D. J. Mink, & A. H. Rots, 435

Bluck, A. F. L., Conselice, C. J., Bouwens, R. J., et al. 2009, MNRAS, 394, L51

Borne, K. D., Bushouse, H., Lucas, R. A., & Colina, L. 2000, ApJL, 529, L77

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

Brent, R. P. 1971, The Computer Journal, 14, 422

Bridge, C. R., Appleton, P. N., Conselice, C. J., et al. 2007, ApJ , 659, 931

Bridge, C. R., Carlberg, R. G., & Sullivan, M. 2010, ApJ, 709, 1067

Bruzual, G. & Charlot, S. 2003, MNRAS, 344, 1000

Bundy, K., Fukugita, M., Ellis, R. S., et al. 2009, ApJ , 697, 1369

Cassata, P., Cimatti, A., Franceschini, A., et al. 2005, MNRAS , 357, 903

Chou, R. C. Y., Bridge, C. R., & Abraham, R. G. 2011, AJ , 141, 87

Cibinel, A., Le Floc'h, E., Perret, V., et al. 2015, ApJ, 805, 181

Conselice, C. J. 2003, ApJS, 147, 1

Conselice, C. J. 2014, ARA&A, 52, 291

Conselice, C. J., Bershady, M. A., Dickinson, M., & Papovich, C. 2003, AJ, 126, 1183

Conselice, C. J., Bershady, M. A., & Gallagher, III, J. S. 2000a, A&A, 354, L21

Conselice, C. J., Bershady, M. A., & Jangren, A. 2000b, ApJ, 529, 886

Conselice, C. J., Rajgor, S., & Myers, R. 2008, MNRAS , 386, 909

Conselice, C. J., Yang, C., & Bluck, A. F. L. 2009, MNRAS, 394, 1956

Cotini, S., Ripamonti, E., Caccianiga, A., et al. 2013, MNRAS, 431, 2661

Croton, D. J., Springel, V., White, S. D. M., et al. 2006, MNRAS , 365, 11

Davis, M., Guhathakurta, P., Konidaris, N. P., et al. 2007, ApJ , 660, L1

De Propris, R., Conselice, C. J., Liske, J., et al. 2007, ApJ , 666, 212

de Ravel, L., Le Fèvre, O., Tresse, L., et al. 2009, A&A, 498, 379

Elfattah, M. A., El-Bendary, N., Elsoud, M. A. A., Platoš, J., & Hassanien, A. E. 2014, Principal Component Analysis Neural Network Hybrid Classification Approach for Galaxies Images, ed. A. Abraham, P. Krömer, & V. Snášel (Cham: Springer International Publishing), 225–237

Freeman, P. E., Izbicki, R., Lee, A. B., et al. 2013, MNRAS, 434, 282

Frei, Z., Guhathakurta, P., Gunn, J. E., & Tyson, J. A. 1996, AJ, 111, 174

Glasser, G. J. 1962, Journal of the American Statistical Association, 57, 648

Gottlöber, S., Klypin, A., & Kravtsov, A. V. 2001, ApJ , 546, 223

Goulding, A. D., Greene, J. E., Bezanson, R., et al. 2018, PASJ , 70, S37

Graham, A. W., Driver, S. P., Petrosian, V., et al. 2005, AJ, 130, 1535

Graham, A. W., Erwin, P., Caon, N., & Trujillo, I. 2001a, ApJL, 563, L11

Graham, A. W., Trujillo, I., & Caon, N. 2001b, AJ, 122, 1707

Hambleton, K. M., Gibson, B. K., Brook, C. B., et al. 2011, MNRAS , 418, 801

Hibbard, J. E. & Vacca, W. D. 1997, AJ , 114, 1741

Hogg, D. W. 1999, ArXiv Astrophysics e-prints

Hopkins, P. F., Hernquist, L., Cox, T. J., et al. 2006, ApJS, 163, 1

Hopkins, P. F., Hernquist, L., Cox, T. J., & Kereš, D. 2008, ApJS, 175, 356

Hoyos, C., Aragón-Salamanca, A., Gray, M. E., et al. 2012, MNRAS, 419, 2703

Hoyos, C., den Brok, M., Verdoes Kleijn, G., et al. 2011, MNRAS , 411, 2439

Hubble, E. P. 1926, ApJ , 64

Ilbert, O., Arnouts, S., McCracken, H. J., et al. 2006, A&A , 457, 841

Ivezić, Ž., Connolly, A., VanderPlas, J., & Gray, A. 2014, Statistics, Data Mining, and Machine Learning in Astronomy: A Practical Python Guide for the Analysis of Survey Data, Princeton Series in Modern Observational Astronomy (Princeton University Press)

Jaccard, P. 1901, Bull Soc Vaudoise Sci Nat, 37, 241

Jogee, S., Miller, S. H., Penner, K., et al. 2009, ApJ , 697, 1971

Kampczyk, P., Lilly, S. J., Carollo, C. M., et al. 2007, ApJS , 172, 329

Kartaltepe, J. S., Sanders, D. B., Scoville, N. Z., et al. 2007, ApJS, 172, 320

Kaufman, L. & Rousseeuw, P. J. 1990, Finding groups in data : an introduction to cluster analysis, Wiley series in probability and mathematical statistics (New York: Wiley), a Wiley-Interscience publication.

Lackner, C. N., Silverman, J. D., Salvato, M., et al. 2014, AJ , 148, 137

Law, D. R., Steidel, C. C., Erb, D. K., et al. 2007, ApJ, 656, 1

Le Fèvre, O., Abraham, R., Lilly, S. J., et al. 2000, MNRAS, 311, 565

Le Fevre, O., Vettolani, G., Maccagni, D., et al. 2004, in Astrophysics and Space Science Library, Vol. 301, Astrophysics and Space Science Library, ed. M. Plionis, 7

Lin, L., Patton, D. R., Koo, D. C., et al. 2008, ApJ, 681, 232

López-Sanjuan, C., Balcells, M., Pérez-González, P. G., et al. 2009, A&A , 501, 505

López-Sanjuan, C., Cenarro, A. J., Varela, J., et al. 2015, A&A , 576, A53

López-Sanjuan, C., Le Fèvre, O., Tasca, L. A. M., et al. 2013, A&A , 553, A78

Lorenz, M. O. 1905, Publications of the American Statistical Association, 9, 209

Lotz, J. M. 2007, in Astronomical Society of the Pacific Conference Series, Vol. 380, Deepest Astronomical Surveys, ed. J. Afonso, H. C. Ferguson, B. Mobasher, & R. Norris, 467

Lotz, J. M., Davis, M., Faber, S. M., et al. 2008a, ApJ, 672, 177

Lotz, J. M., Davis, M., Faber, S. M., et al. 2008b, ApJ, 672, 177

Lotz, J. M., Jonsson, P., Cox, T. J., et al. 2011, ApJ, 742, 103

Lotz, J. M., Primack, J., & Madau, P. 2004, AJ, 128, 163

Man, A. W. S., Toft, S., Zirm, A. W., Wuyts, S., & van der Wel, A. 2012, ApJ, 744, 85

Man, A. W. S., Zirm, A. W., & Toft, S. 2016, ApJ, 830, 89

McAlpine, S., Helly, J. C., Schaller, M., et al. 2016, Astronomy and Computing, 15, 72

Mundy, C. J., Conselice, C. J., Duncan, K. J., et al. 2017, MNRAS , 470, 3507

Murphy, K. P. 2012, Machine Learning: A Probabilistic Perspective. Adaptive Computation and Machine Learning

Naim, A., Ratnatunga, K. U., & Griffiths, R. E. 1997, ApJ, 476, 510

Nelson, D., Pillepich, A., Genel, S., et al. 2015, Astronomy and Computing, 13, 12

Newman, A. B., Ellis, R. S., Bundy, K., & Treu, T. 2012, ApJ, 746, 162

Noeske, K. G., Weiner, B. J., Faber, S. M., et al. 2007, ApJ , 660, L43

Papovich, C., Giavalisco, M., Dickinson, M., Conselice, C. J., & Ferguson, H. C. 2003, ApJ, 598, 827

Patton, D. R., Pritchet, C. J., Carlberg, R. G., et al. 2002, ApJ, 565, 208

Patton, D. R., Pritchet, C. J., Yee, H. K. C., Ellingson, E., & Carlberg, R. G. 1997, ApJ, 475, 29

Pawlik, M. M., Wild, V., Walcher, C. J., et al. 2016, MNRAS, 456, 3032

Peacock, J. A. 1999, Cosmological physics (Cambridge university press)

Peth, M. A., Lotz, J. M., Freeman, P. E., et al. 2016, MNRAS, 458, 963

Petrosian, V. 1976, ApJL, 209, L1

Pierre, M., Valtchanov, I., Altieri, B., et al. 2004, JCAP, 9, 011

Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2014, A&A , 571, A16

Ryden, B. 2016, Introduction to cosmology (Cambridge University Press)

Scoville, N., Aussel, H., Brusa, M., et al. 2007, ApJS, 172, 1

Sérsic, J. L. 1963, Boletin de la Asociacion Argentina de Astronomia La Plata Argentina, 6, 41

Shamir, L., Holincheck, A., & Wallin, J. 2013, Astronomy and Computing, 2, 67

Somerville, R. S., Hopkins, P. F., Cox, T. J., Robertson, B. E., & Hernquist, L. 2008, MNRAS , 391, 481

Somerville, R. S., Lee, K., Ferguson, H. C., et al. 2004, The Astrophysical Journal Letters, 600, L171

Spergel, D. N., Verde, L., Peiris, H. V., et al. 2003, ApJS , 148, 175

Stewart, K. R., Bullock, J. S., Wechsler, R. H., & Maller, A. H. 2009, ApJ , 702, 307

Taylor, J. R. 1982, An introduction to error analysis: Mill Valley

Taylor-Mager, V. A., Conselice, C. J., Windhorst, R. A., & Jansen, R. A. 2007, ApJ , 659, 162

Wen, Z. Z. & Zheng, X. Z. 2016, ApJ, 832, 90

Wen, Z. Z., Zheng, X. Z., & An, F. X. 2014, ApJ, 787, 130

Willett, K. W., Galloway, M. A., Bamford, S. P., et al. 2017, MNRAS , 464, 4176

Williams, R. J., Quadri, R. F., & Franx, M. 2011, ApJL, 738, L25

Windhorst, R. A., Taylor, V. A., Jansen, R. A., et al. 2002, ApJS , 143, 113

Wu, H., Zou, Z. L., Xia, X. Y., & Deng, Z. G. 1998, A&AS, 132, 181

Xu, C. K., Zhao, Y., Scoville, N., et al. 2012, ApJ , 747, 85

Yee, H. K. C., Morris, S. L., Lin, H., et al. 2000, ApJS , 129, 475

Appendix A

Training Set Galaxies

The Figures in this Appendix correspond to the 50×50 kpc HSC r-band cutouts for each visually identified galaxy in the training set of Section 4.2.2. We divide them by redshift bin and order them in terms of their Golob et al. (in prep) star-forming probabilities $P_{\rm sf}$ (decreasing order). We report the photometric redshifts from Golob et al. (in prep) and merger probabilities $P_{\rm merge}$ for each galaxy, which are obtained by running the RFC on the full 20 deg^2 CLAUDS+HSC catalog. Each galaxy is also annotated with "Merger" or "Non-merger" to denote the author's visual classification. A square-root stretch has been applied to the images so that both nuclear structure and tidal features may be visible.[1]

[1] We chose a common stretch factor for all galaxies for simplicity, however, in reality an `arcsinh` stretch best emphasizes nuclear structures without oversaturating the galaxy centre, while a logarithmic stretch best emphasizes diffuse tidal features in the galaxy outskirts. The square-root stretch is roughly midway between these two. Therefore, some galaxies here may not show all of their defining features, even though they were visible in the `ds9` software under different stretches.

A.1 $0.25 \leq z_{\rm phot} < 0.4$

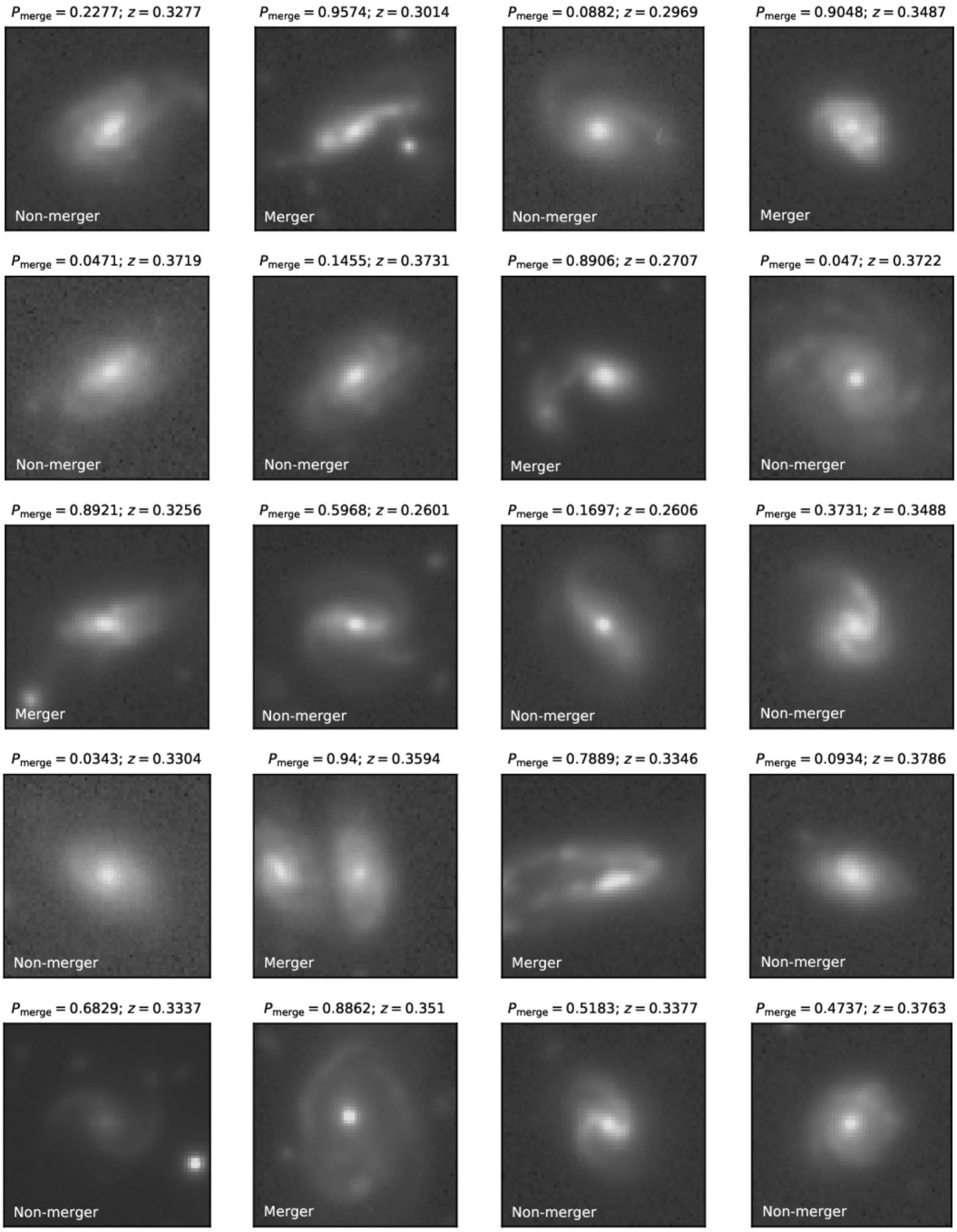


Figure A.1: Training set galaxies in the 1$^{\rm st}$ redshift bin ($0.25 \leq z_{\rm phot} < 0.4$).

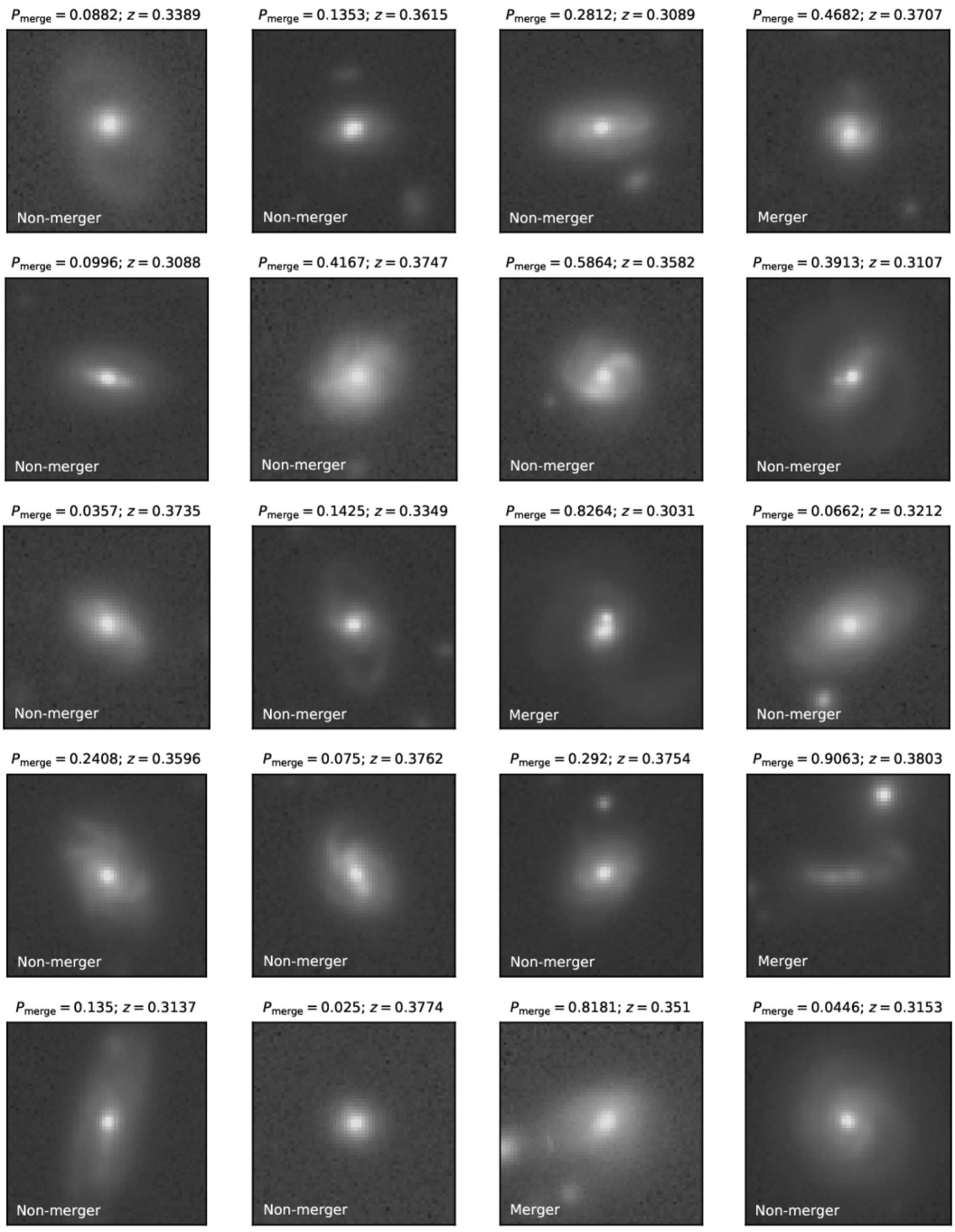


Figure A.1−continued.

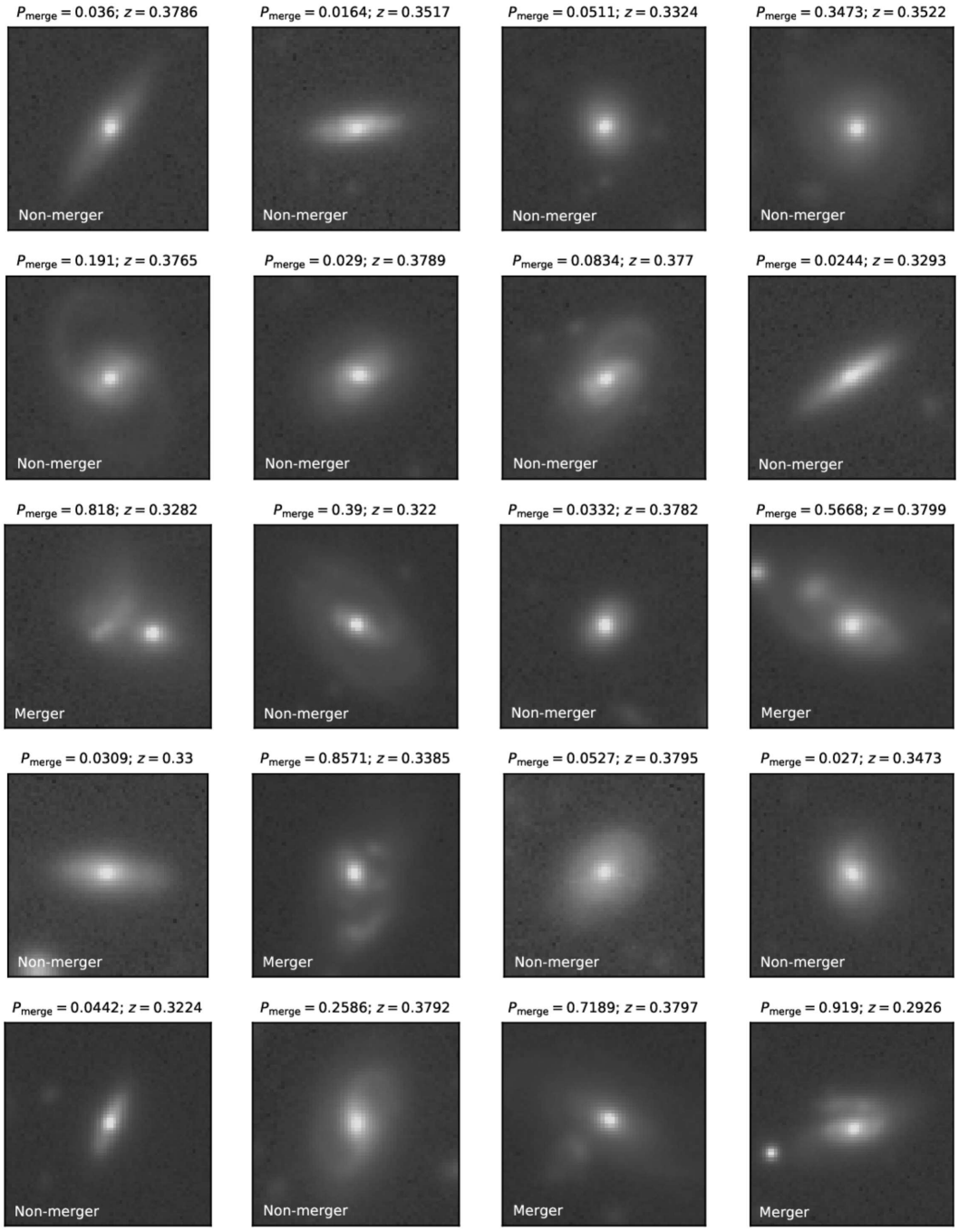


Figure A.1−continued.

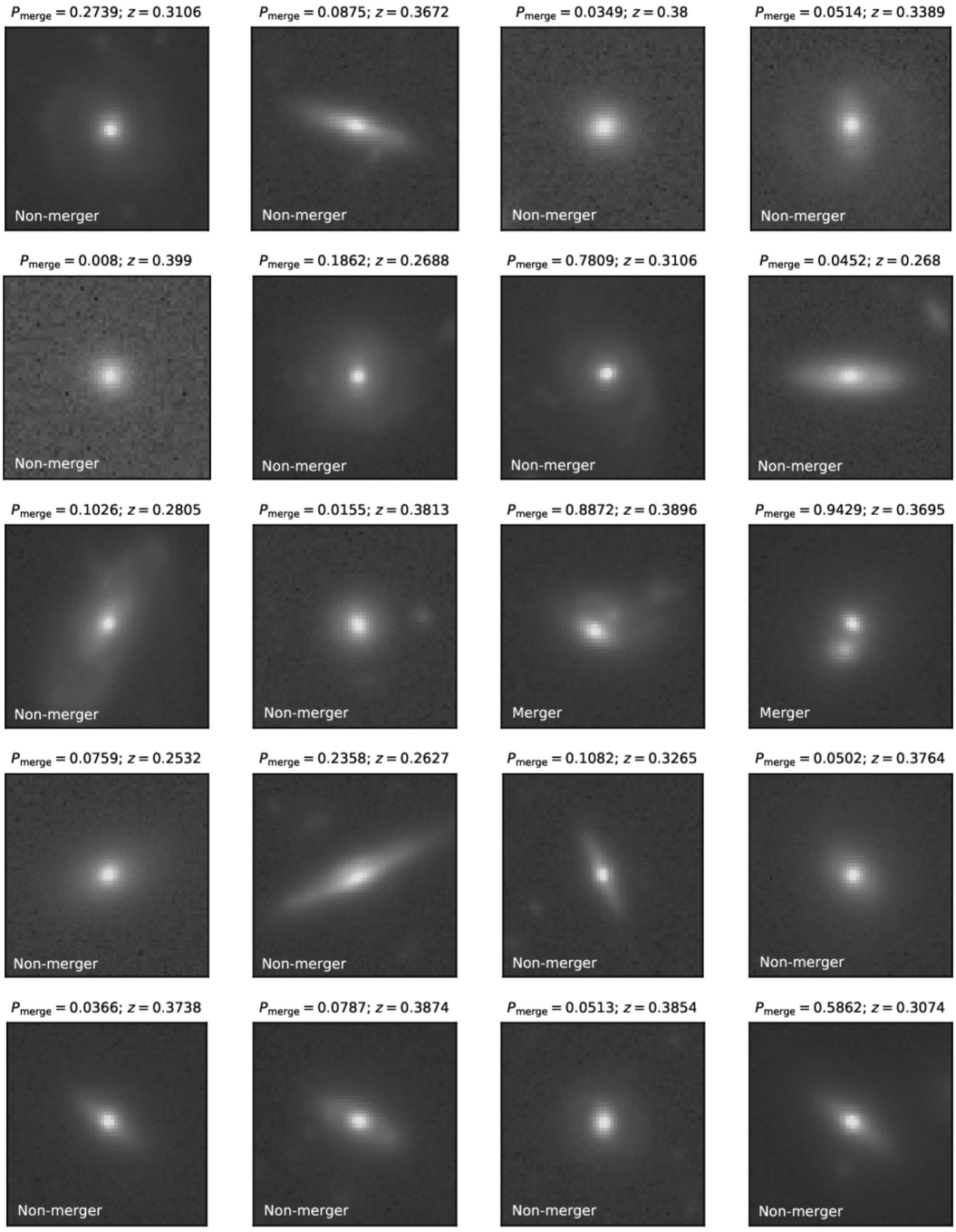


Figure A.1−continued.

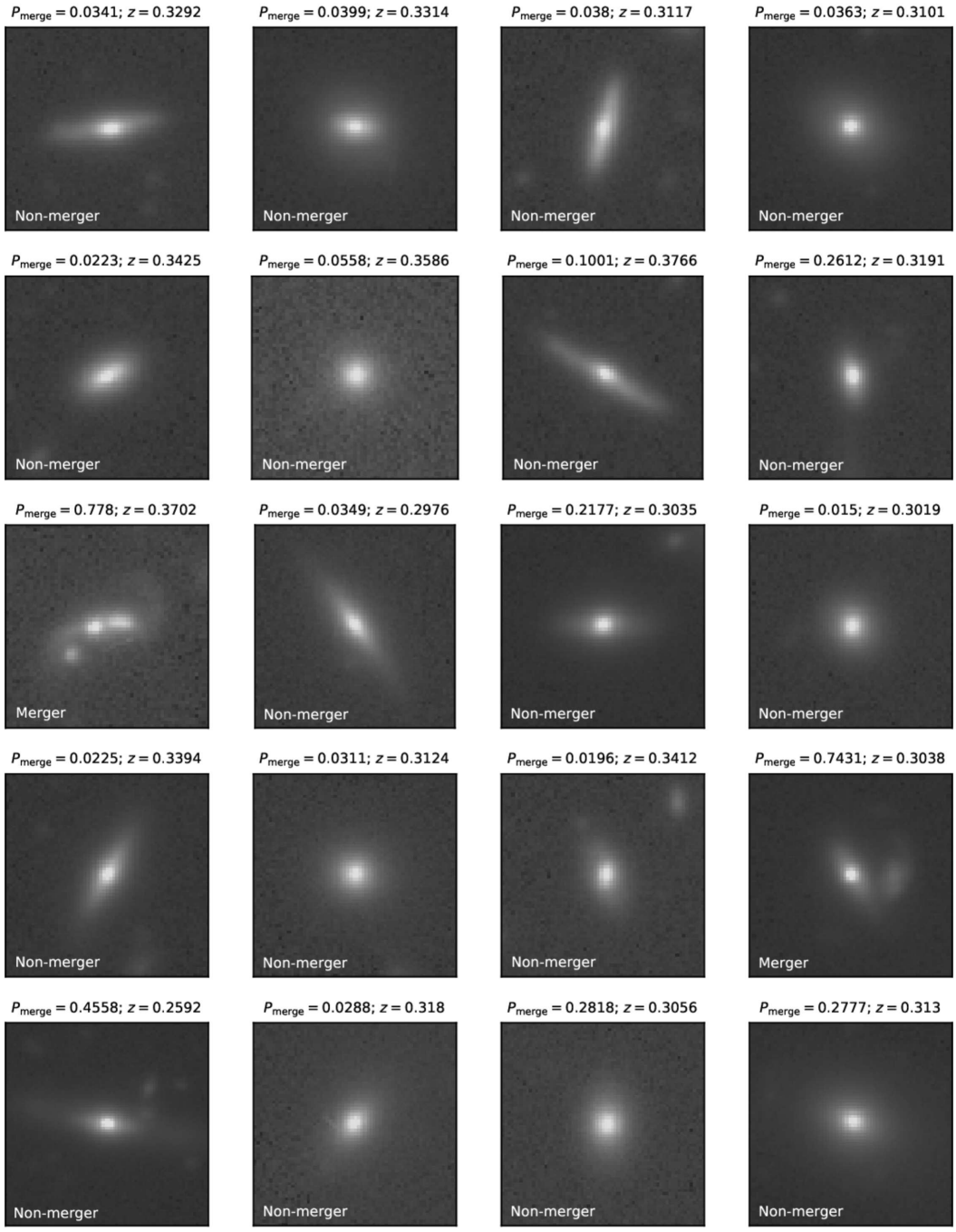


Figure A.1—continued.

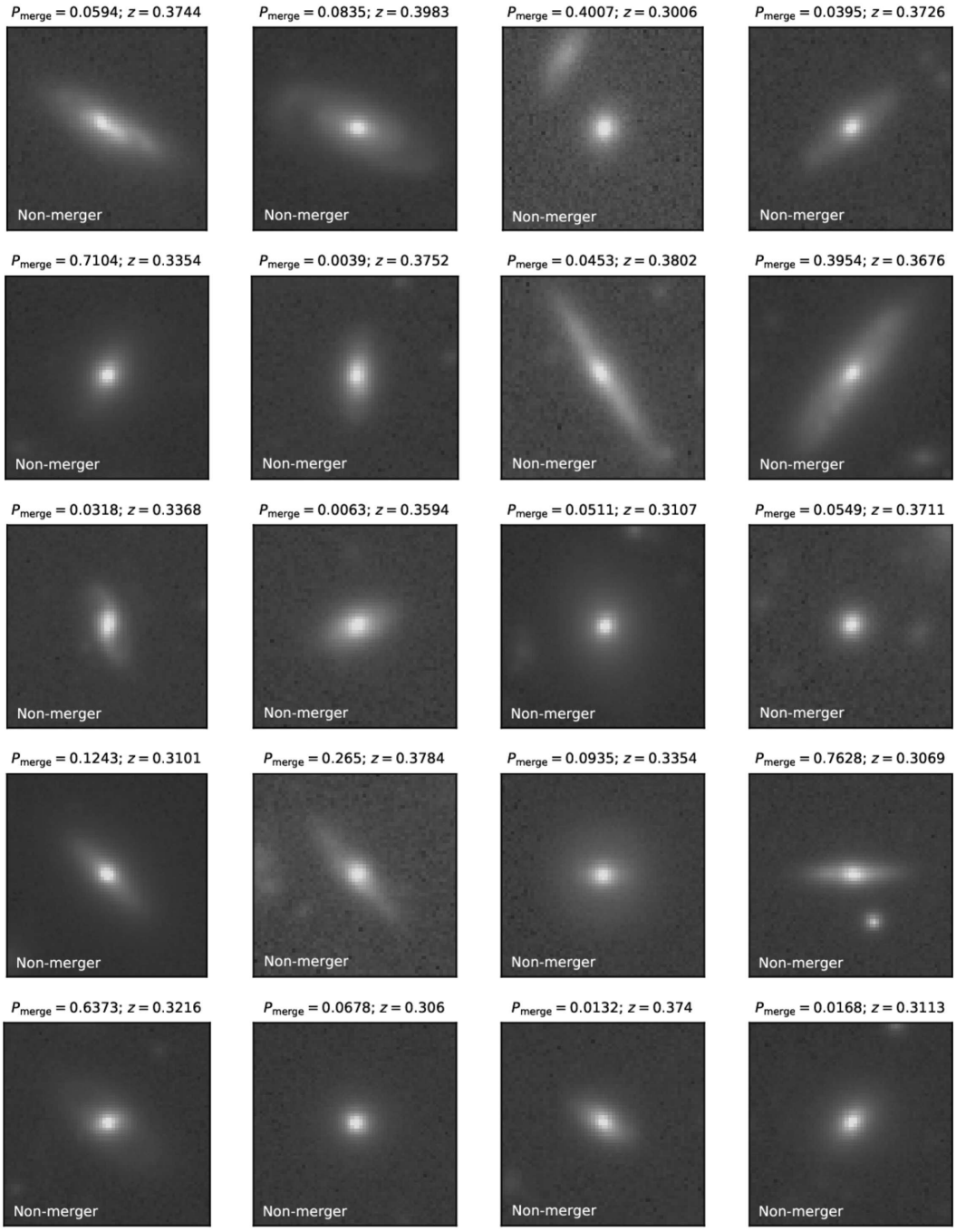


Figure A.1−continued.

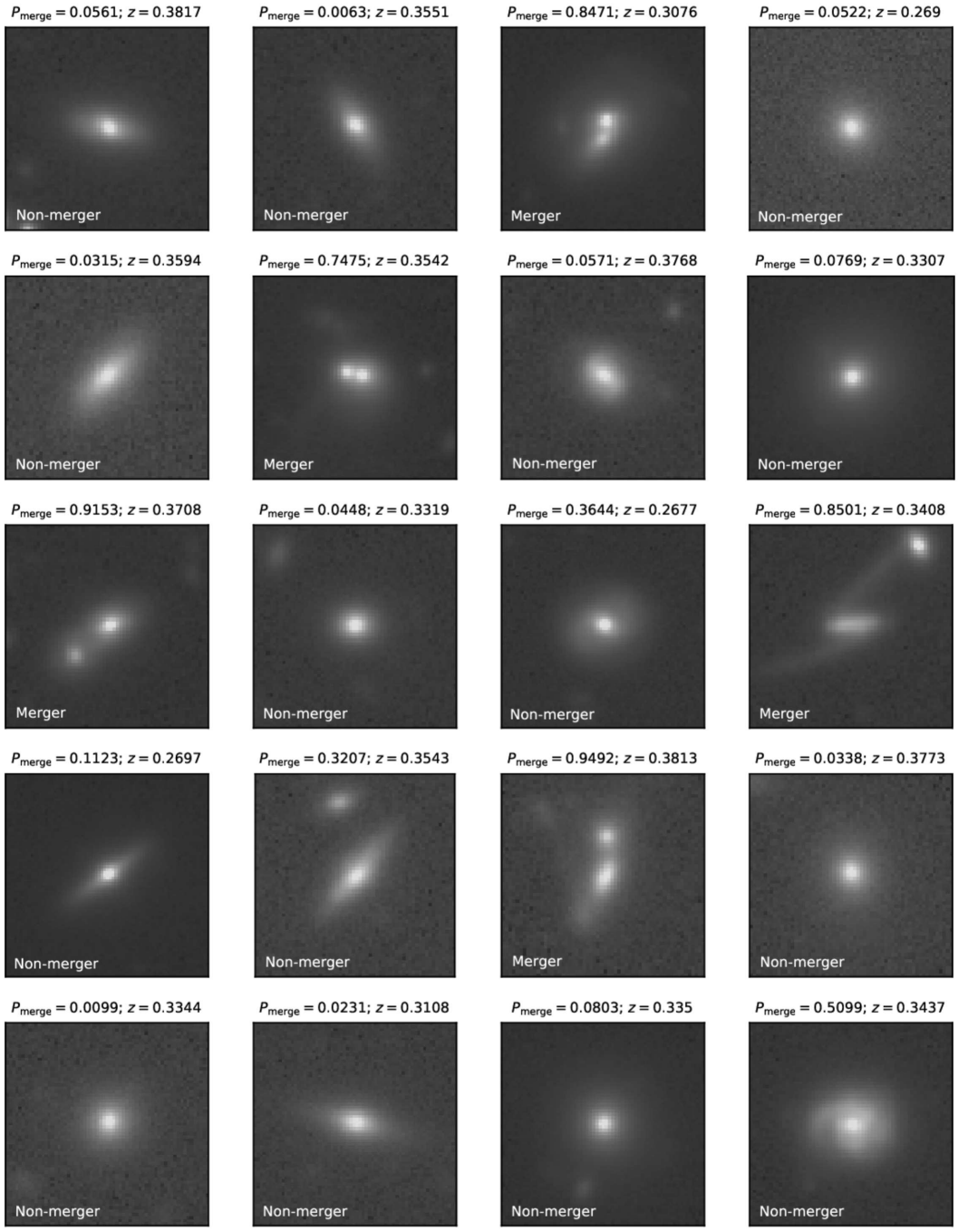


Figure A.1−continued.

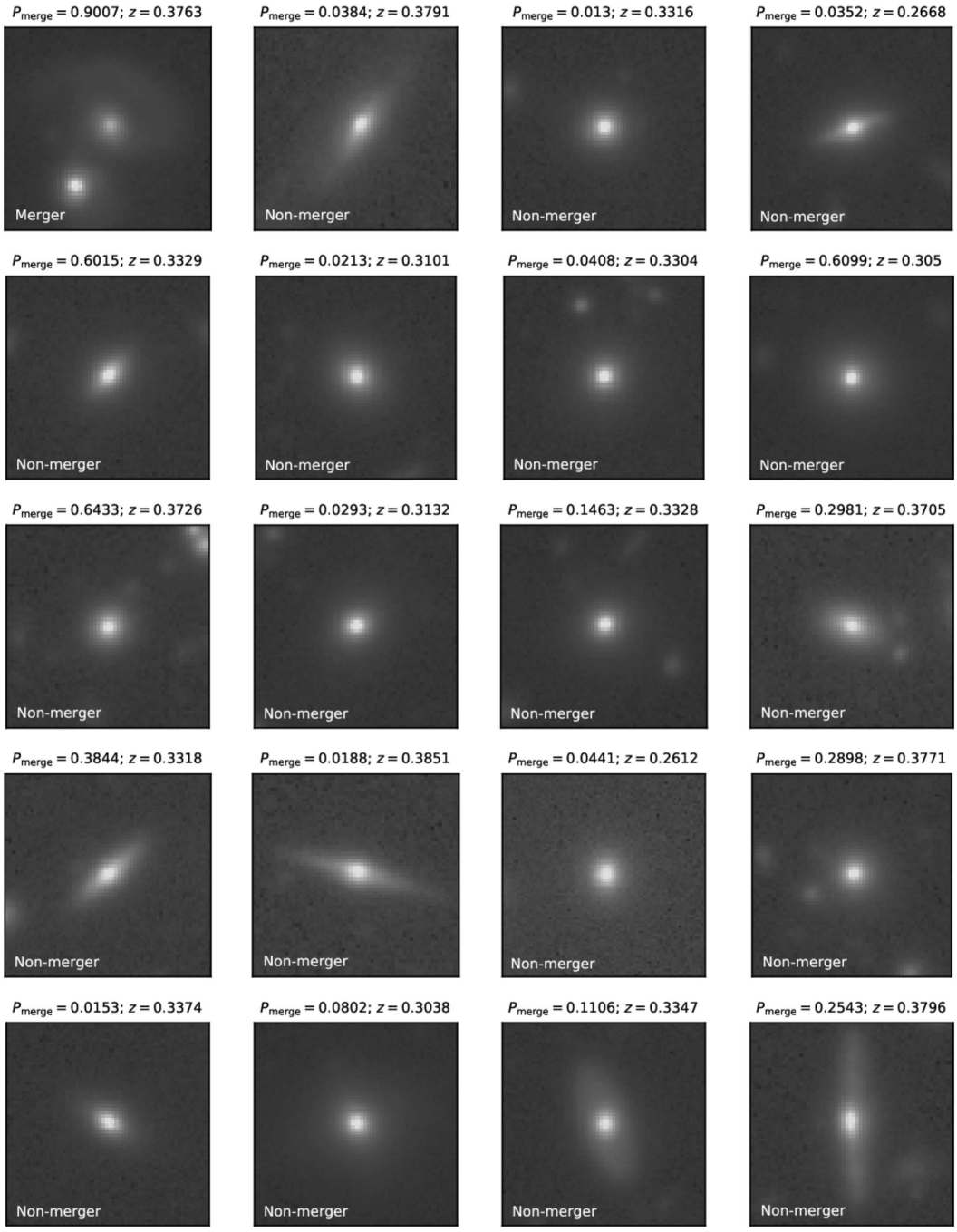


Figure A.1—continued.

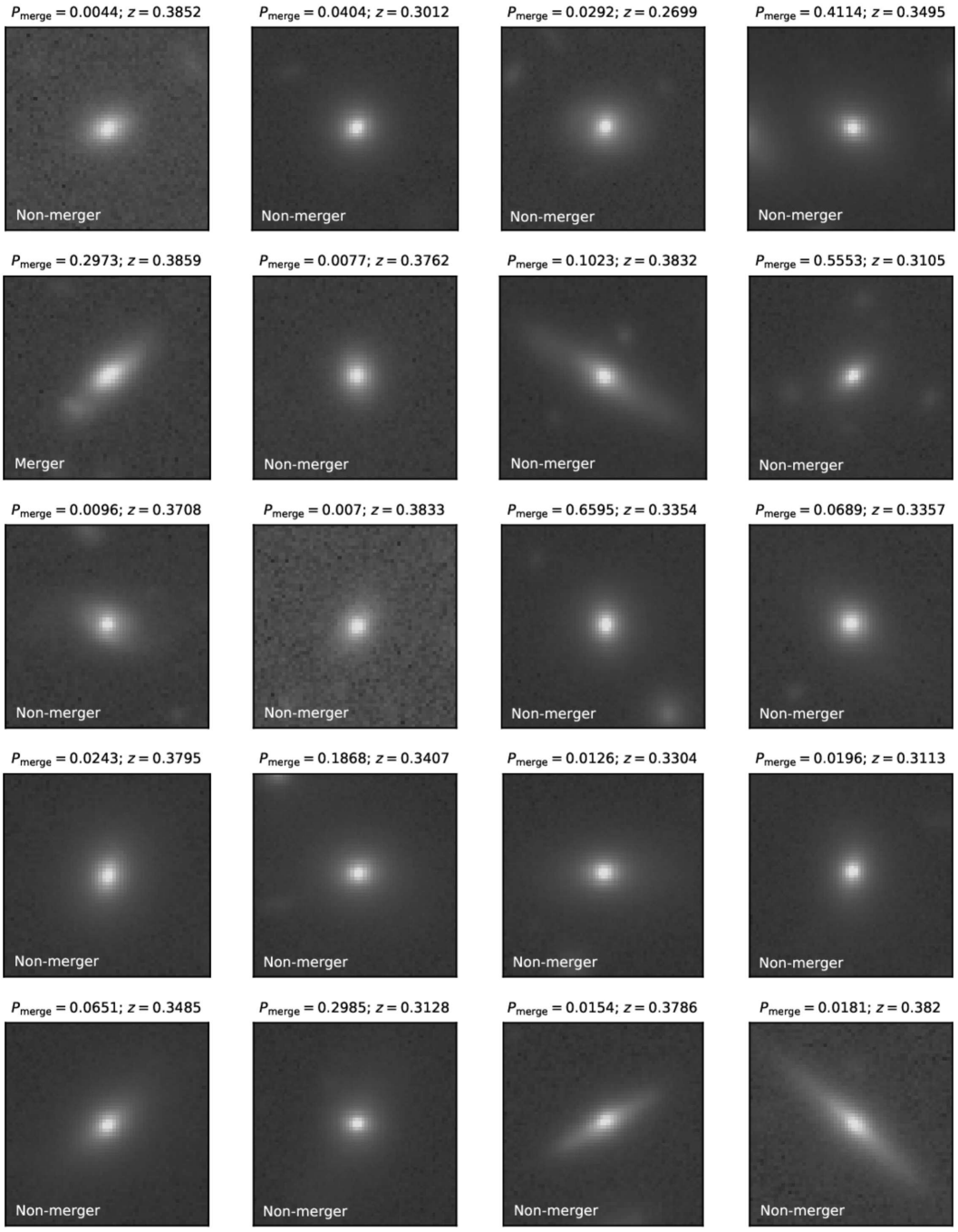


Figure A.1—continued.

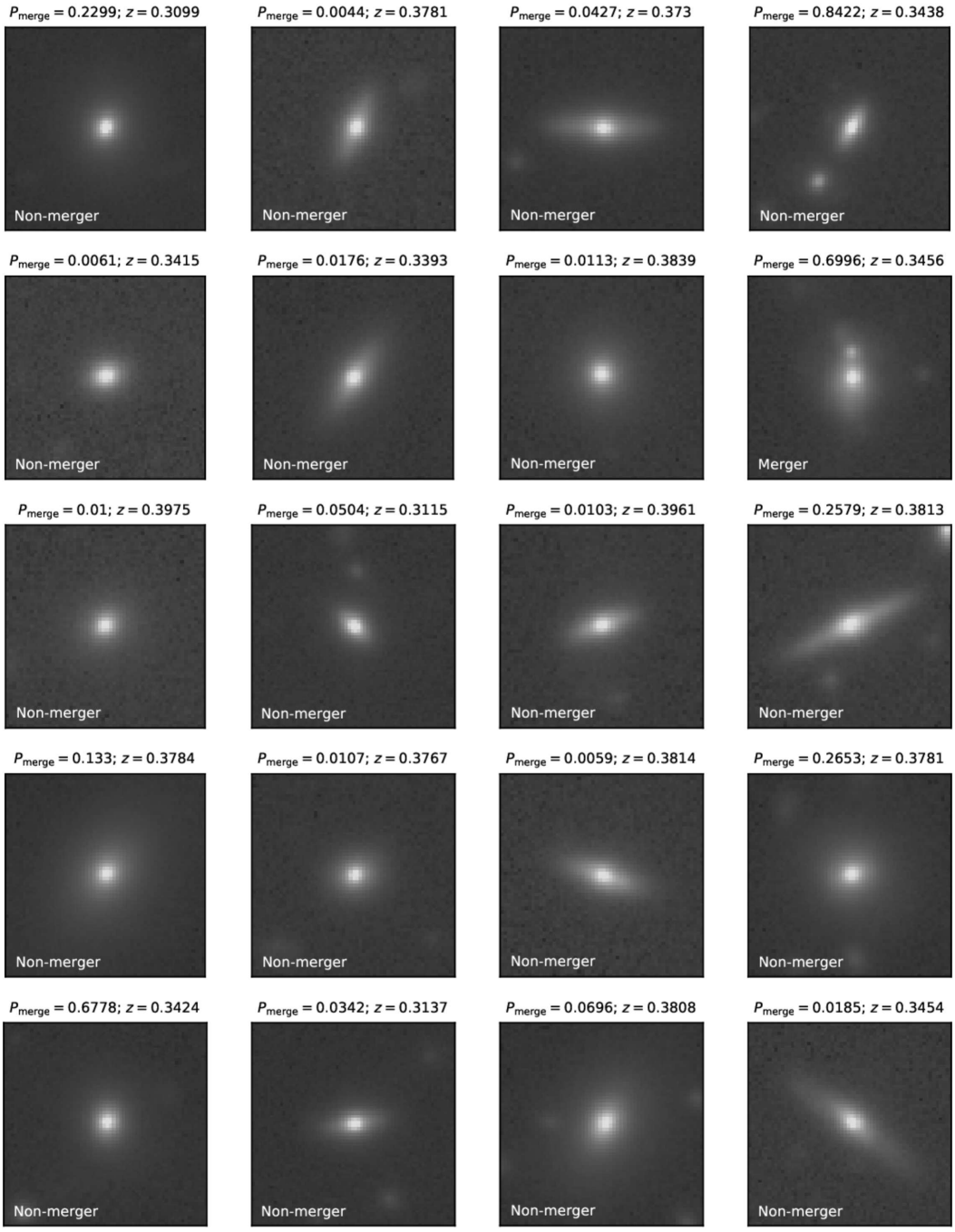


Figure A.1—continued.

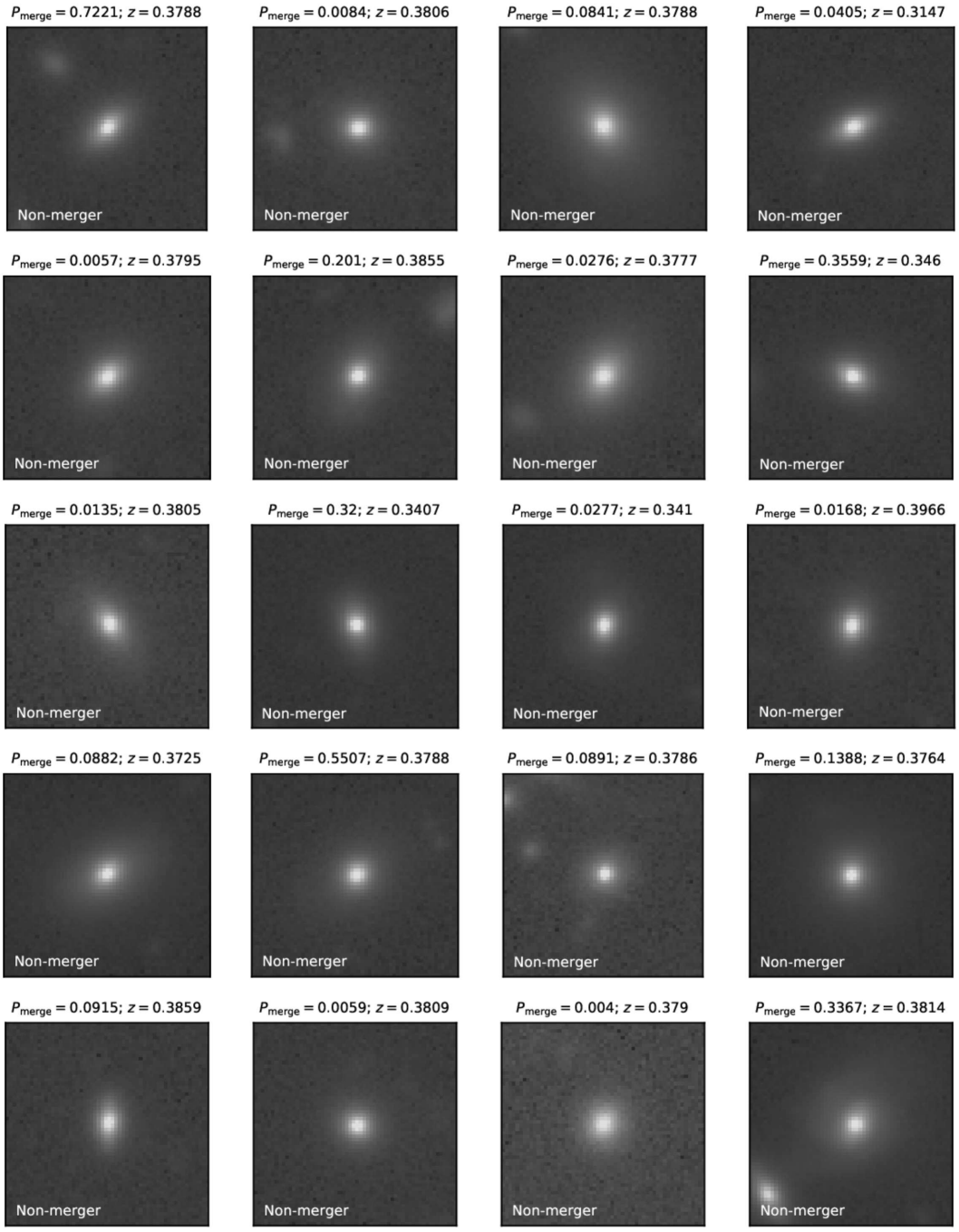


Figure A.1−continued.

A.2 $0.4 \leq z_{\mathrm{phot}} < 0.55$

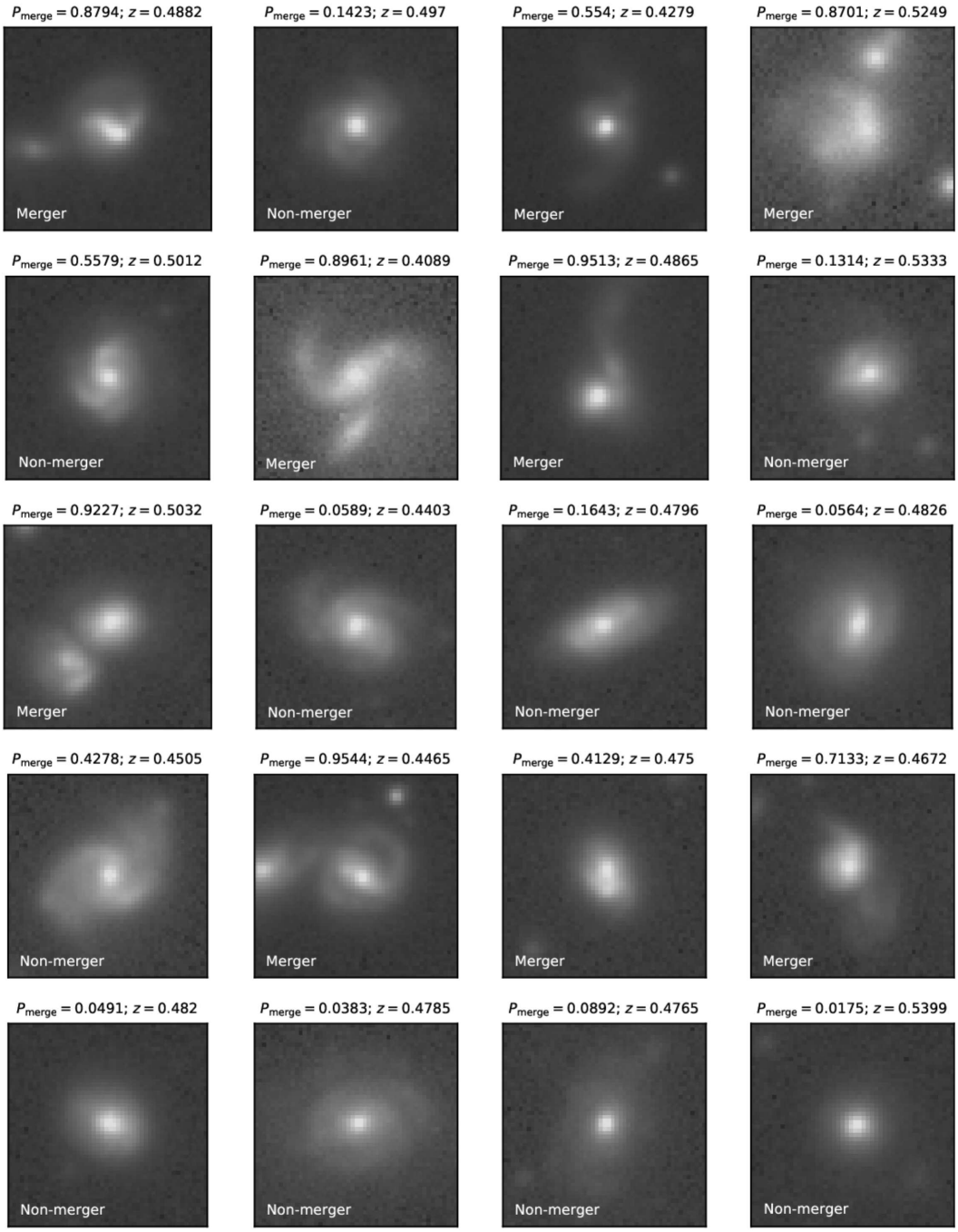


Figure A.2: Training set galaxies in the 2$^{\mathrm{nd}}$ redshift bin ($0.4 \leq z_{\mathrm{phot}} < 0.55$).

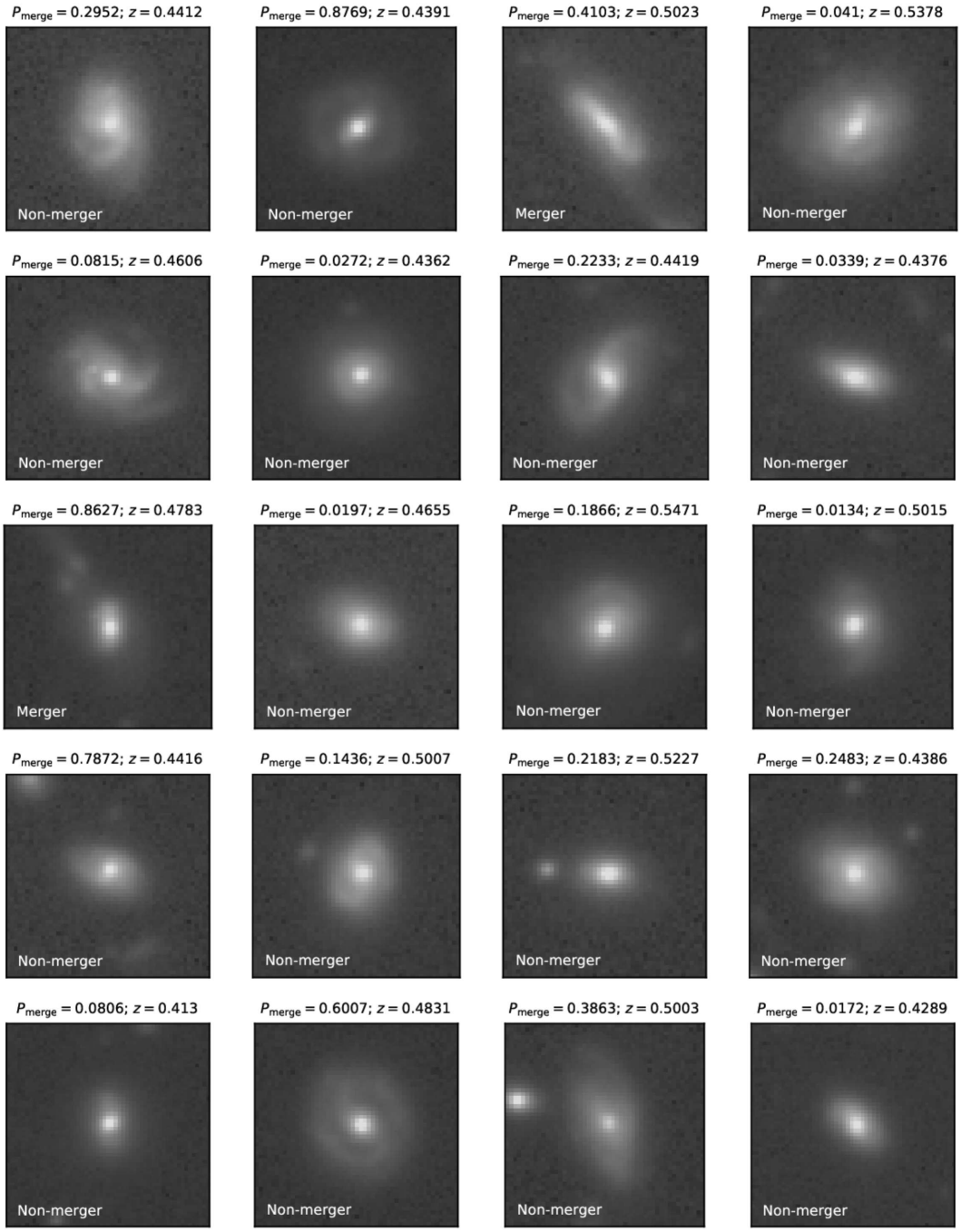


Figure A.2−continued.

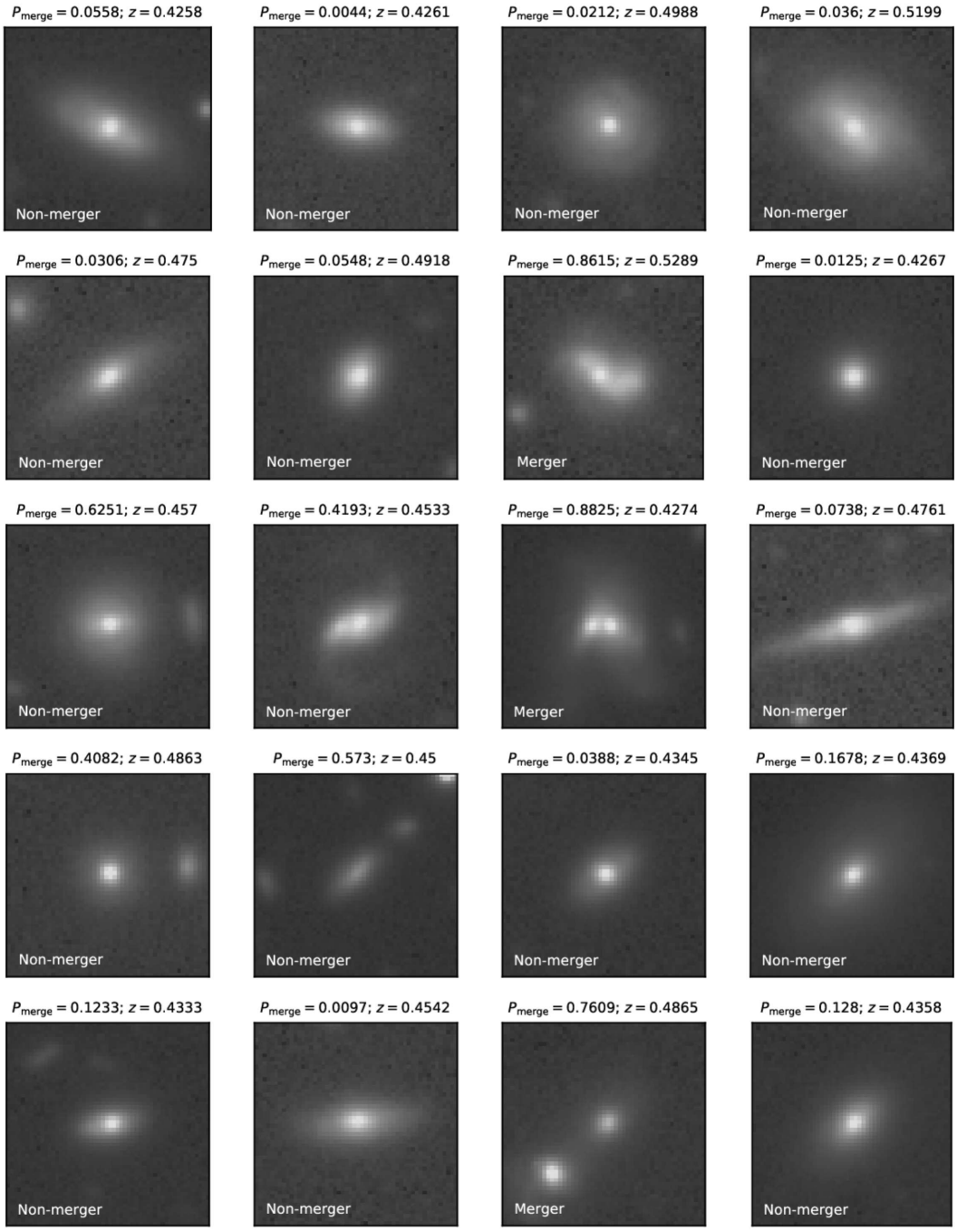


Figure A.2—continued.

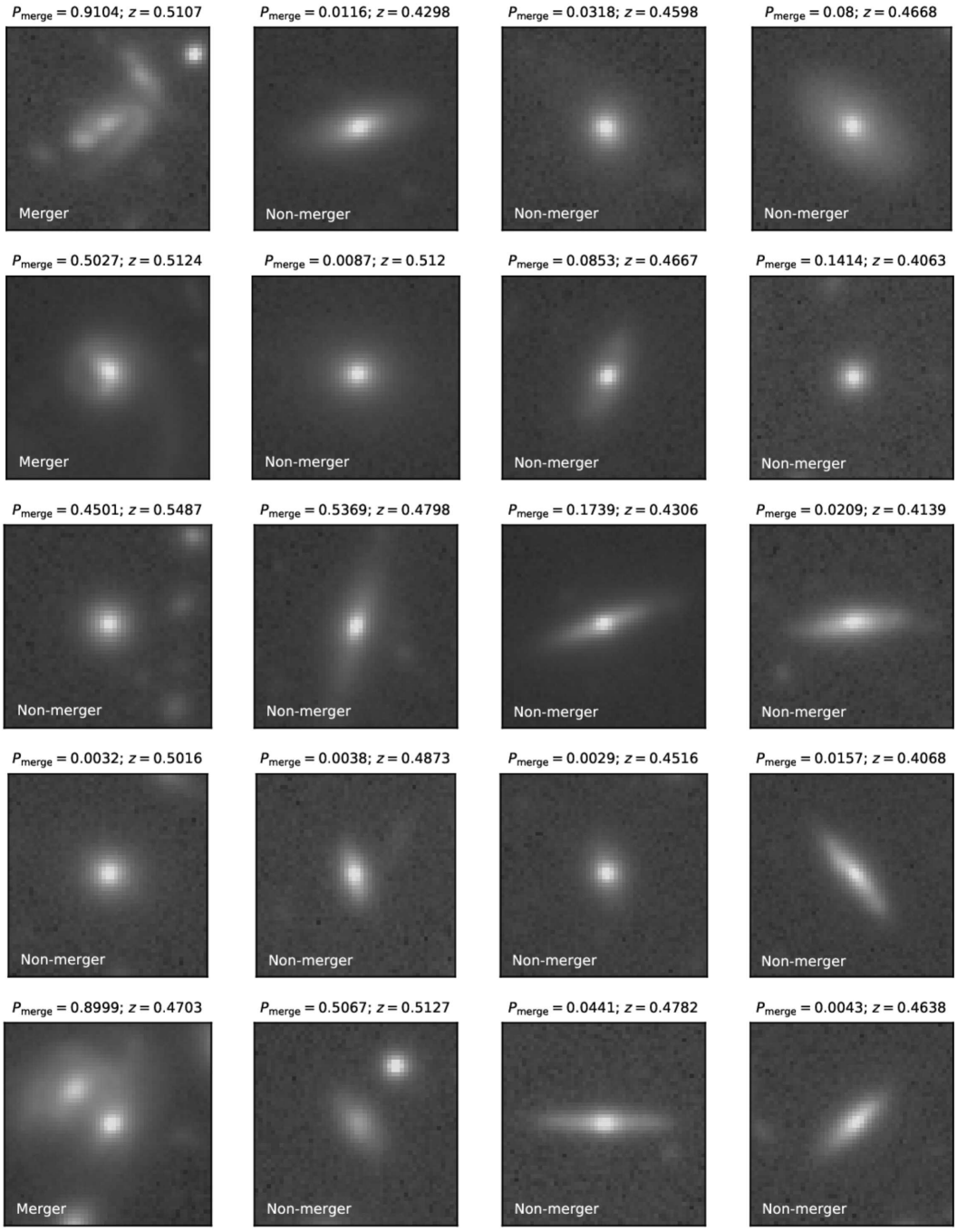


Figure A.2—continued.

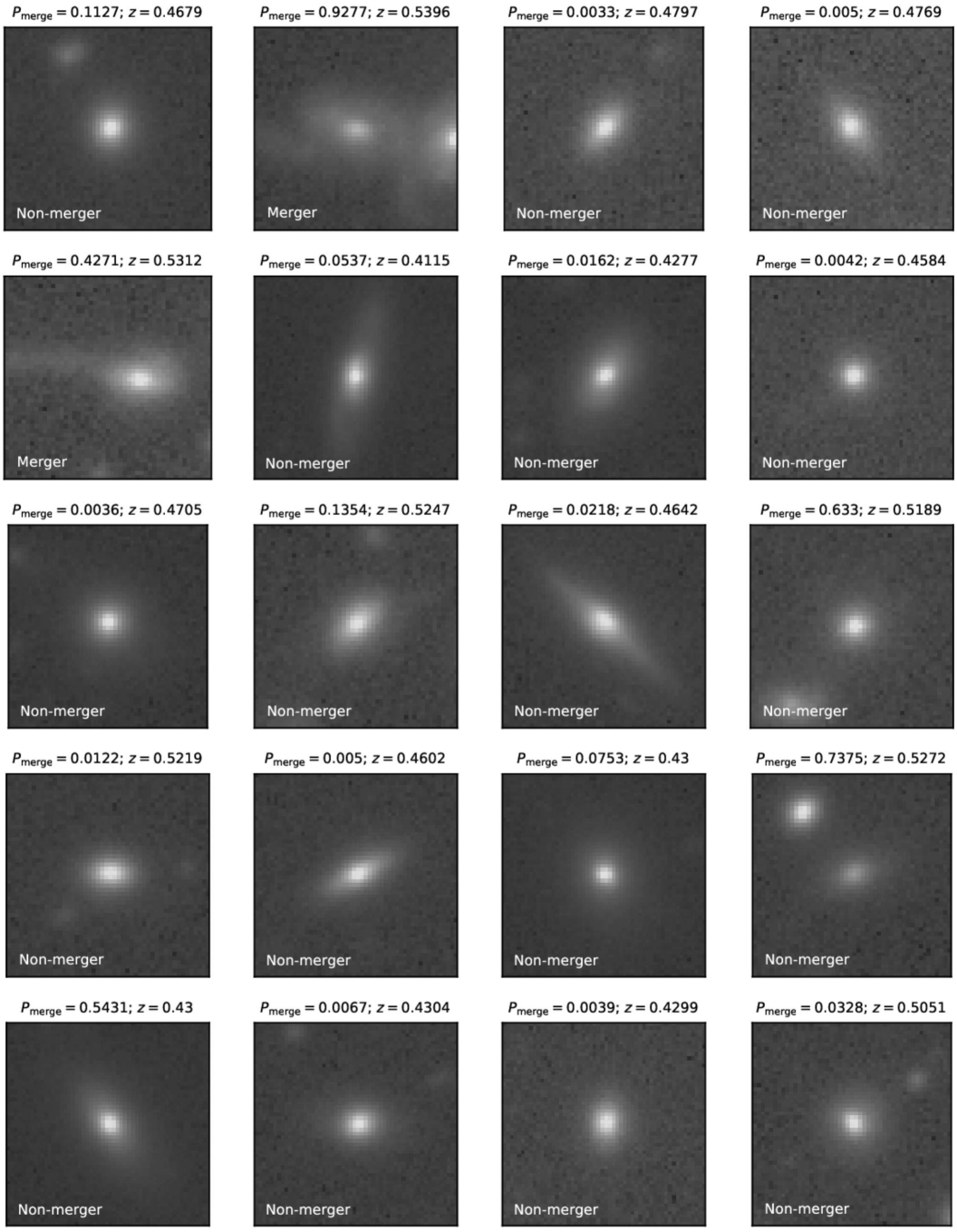


Figure A.2—continued.

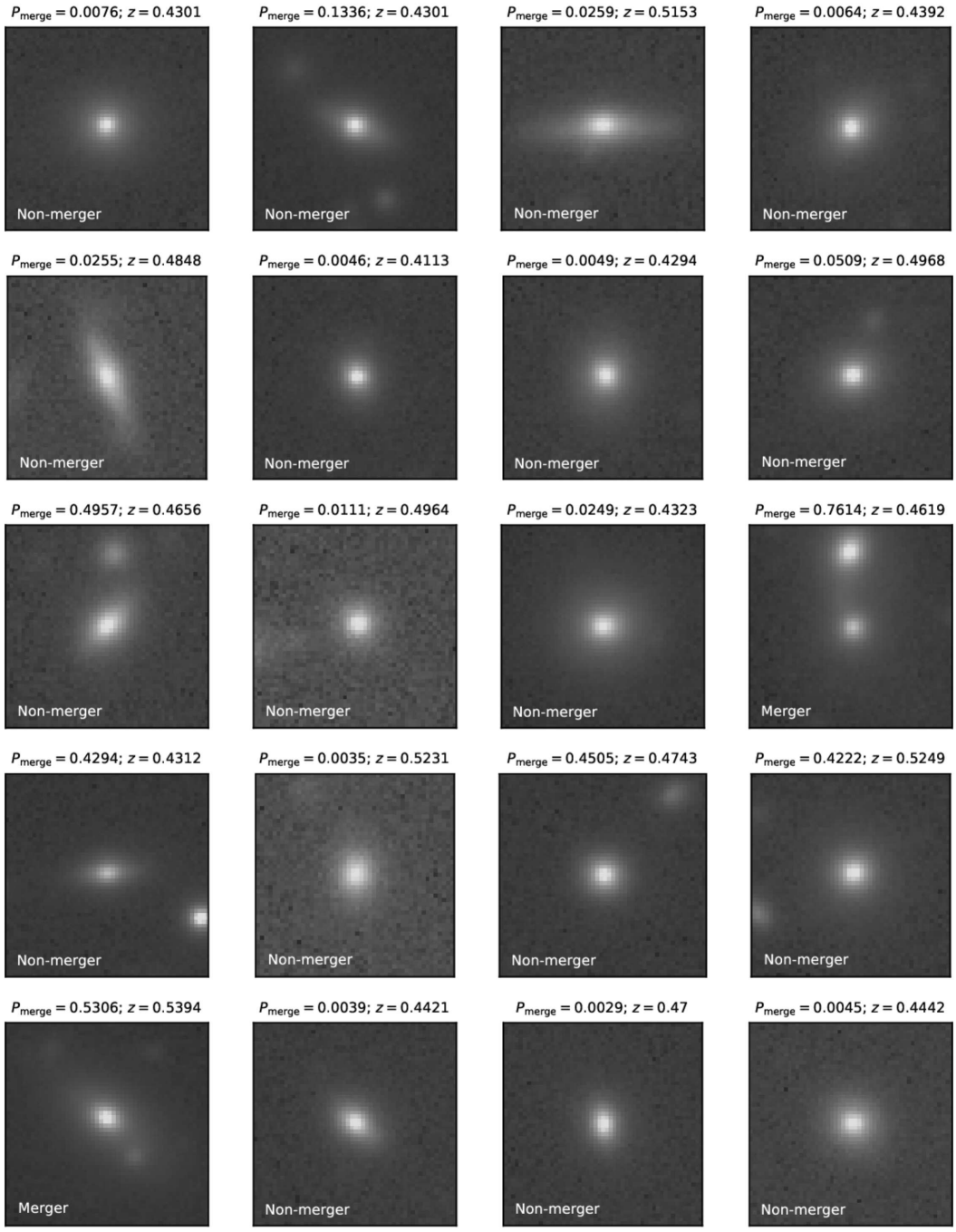


Figure A.2—continued.

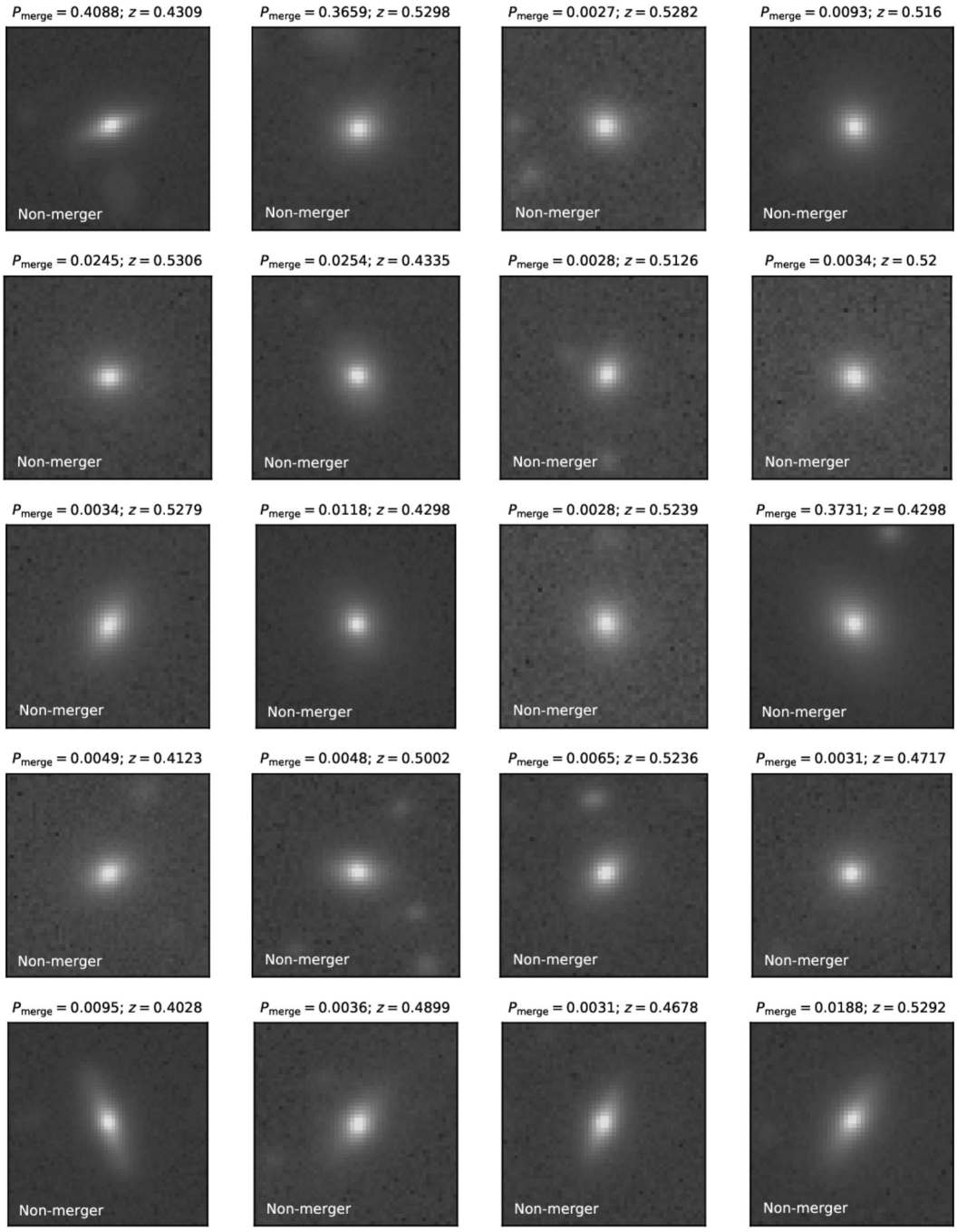


Figure A.2—continued.

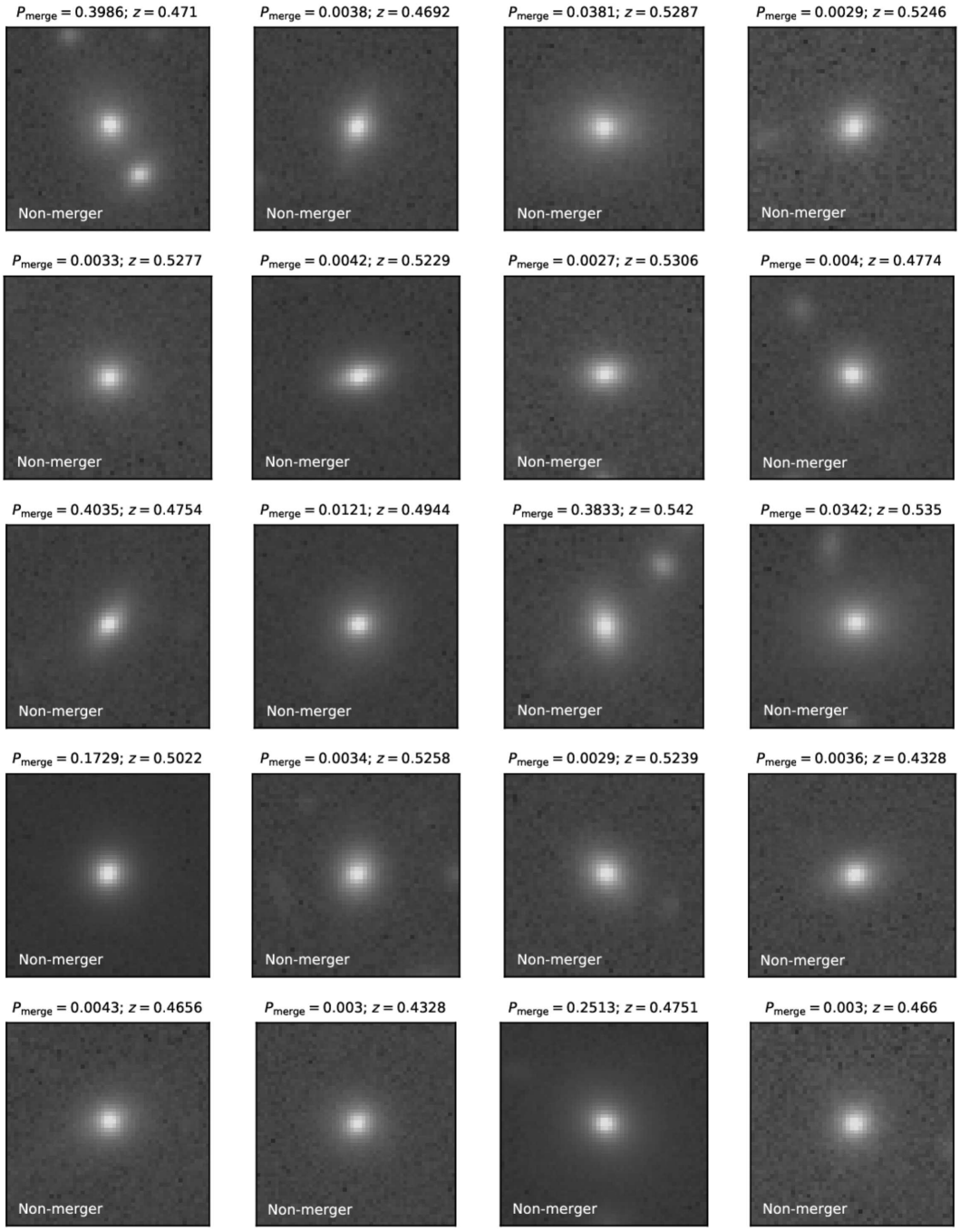


Figure A.2—continued.

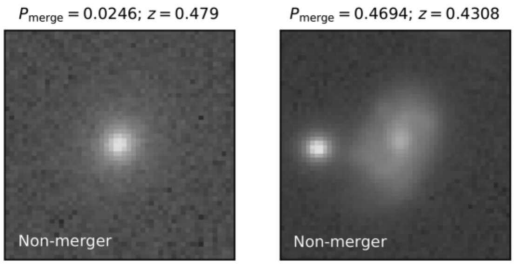


Figure A.2−continued.

A.3 $0.55 \leq z_{\mathbf{phot}} < 0.7$

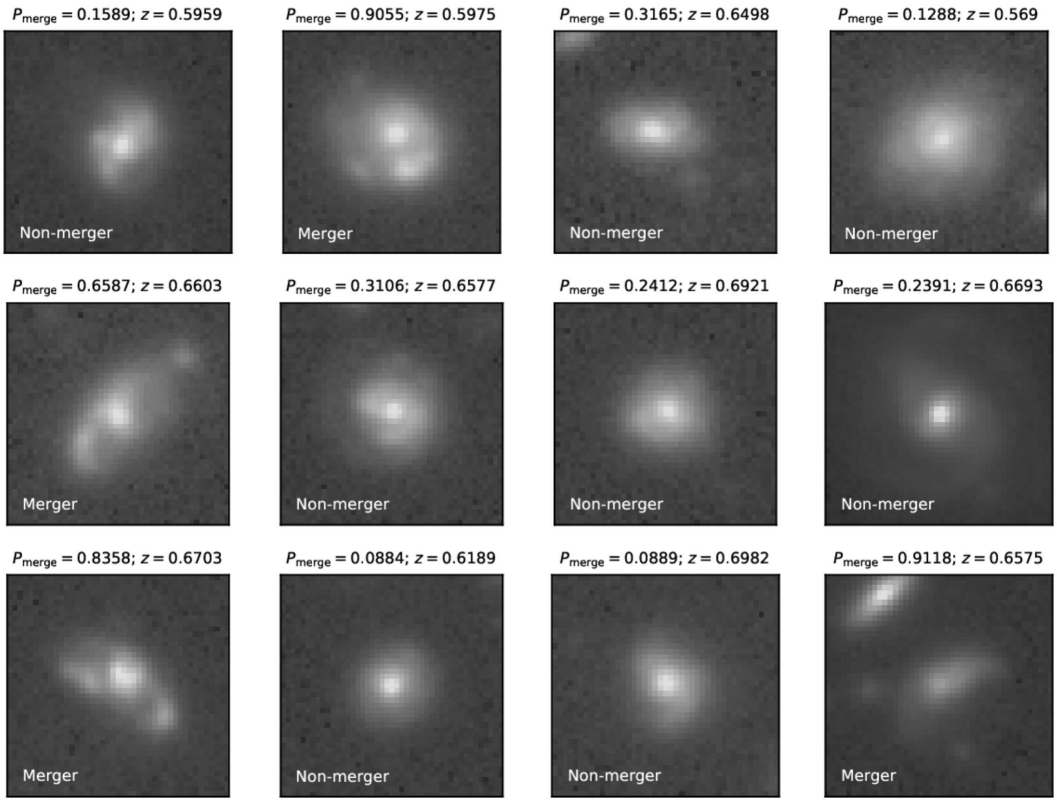


Figure A.3: Training set galaxies in the 3$^{\mathrm{rd}}$ redshift bin ($0.55 \leq z_{\mathrm{phot}} < 0.7$).

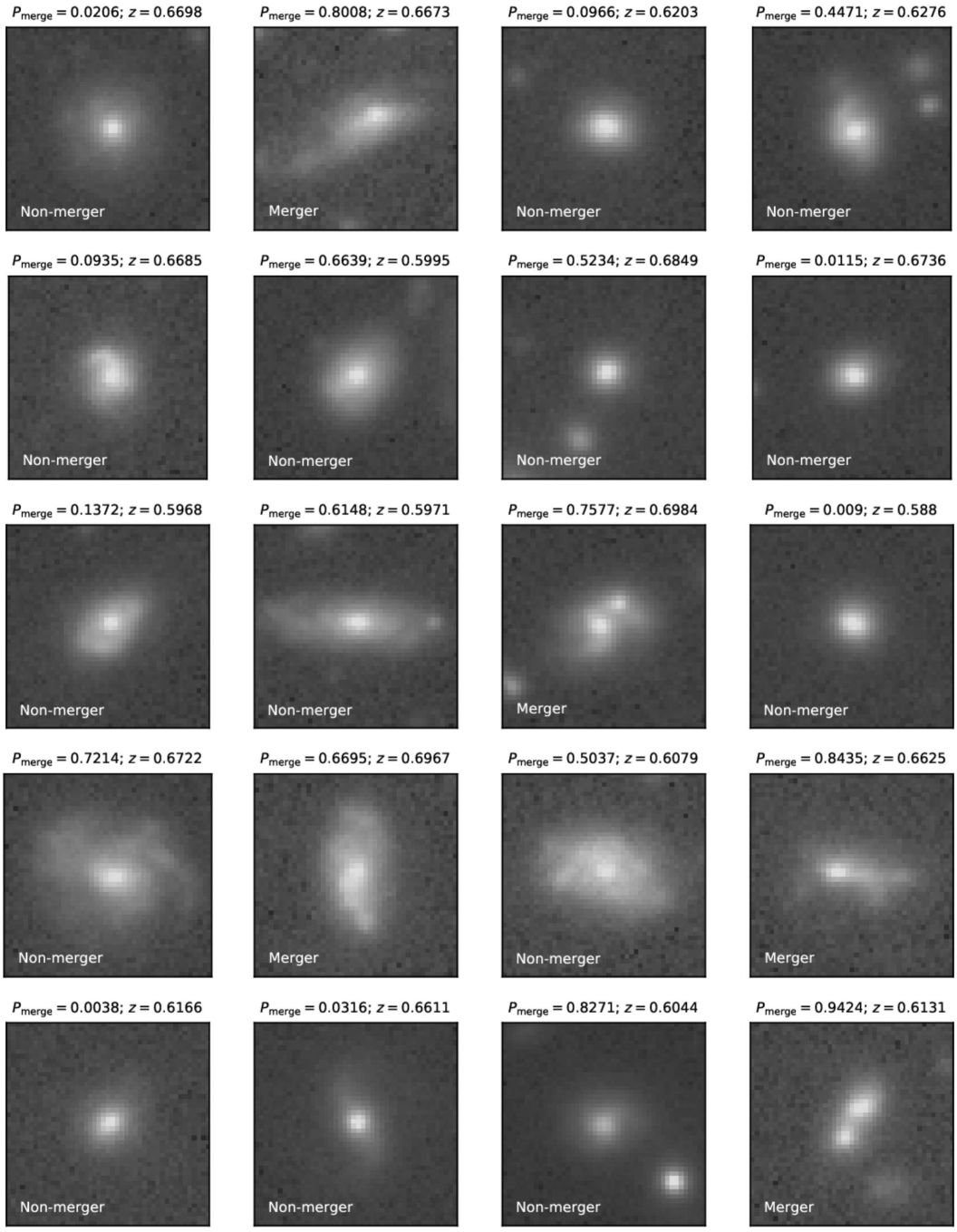


Figure A.3—continued.

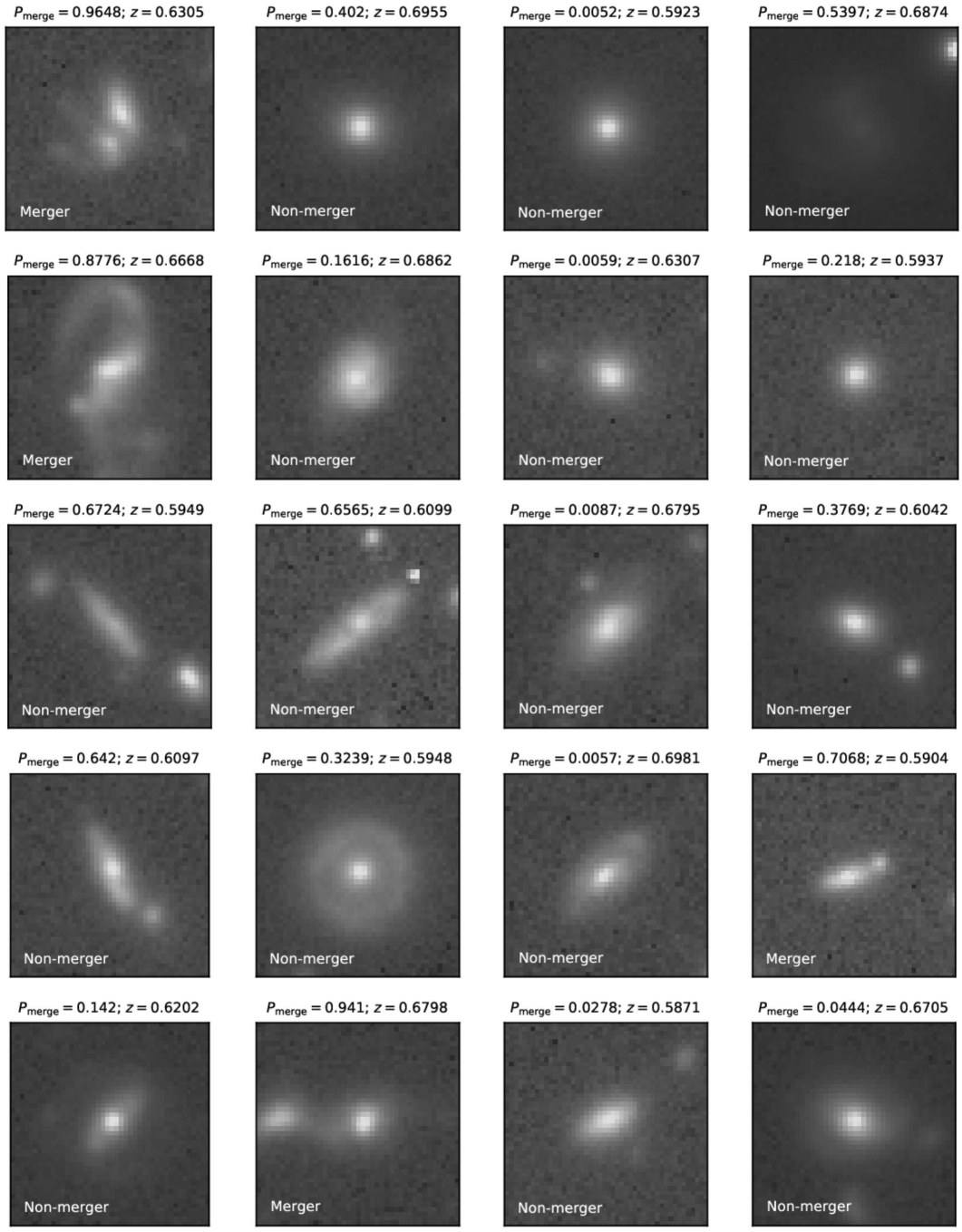


Figure A.3—continued.

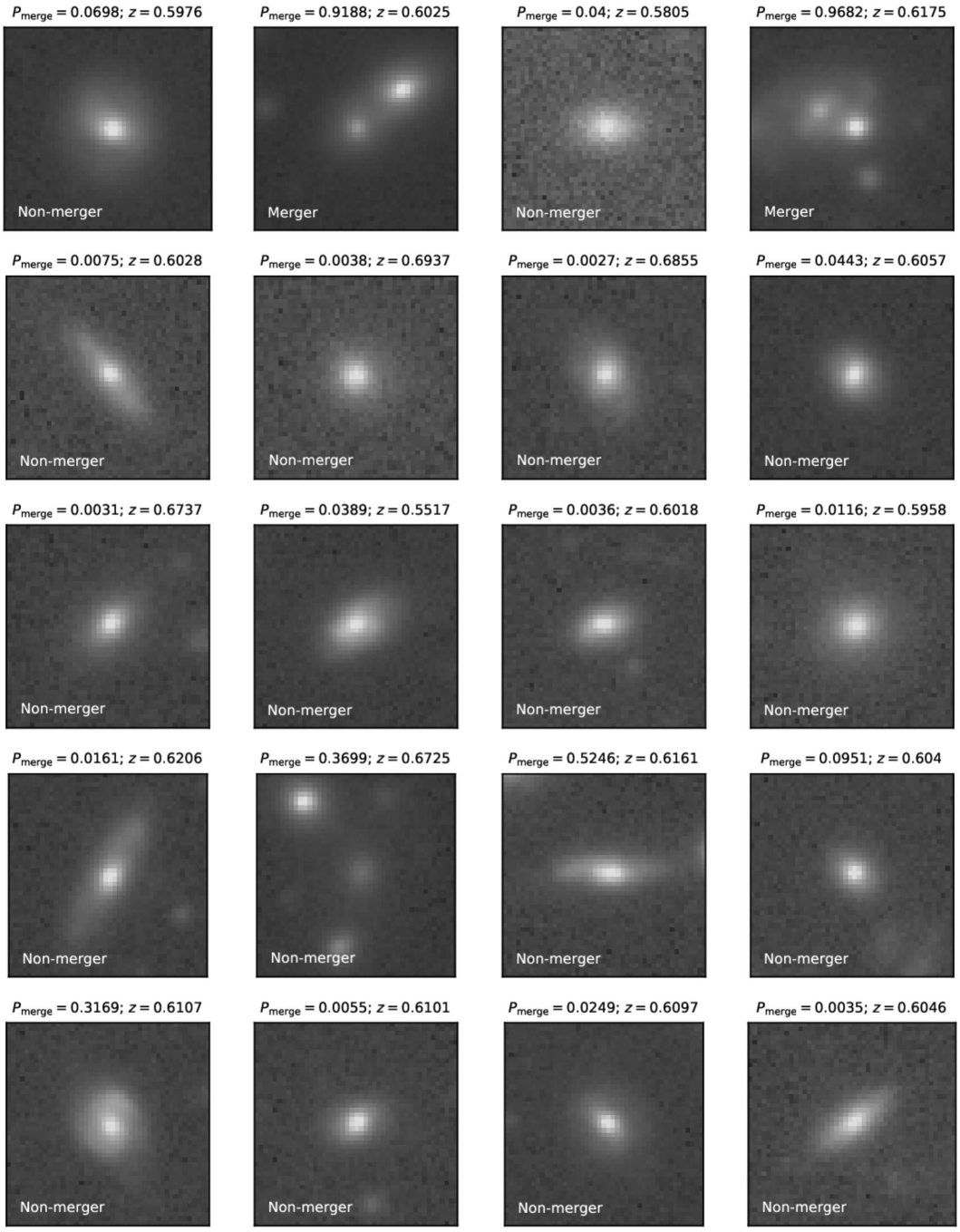


Figure A.3—continued.

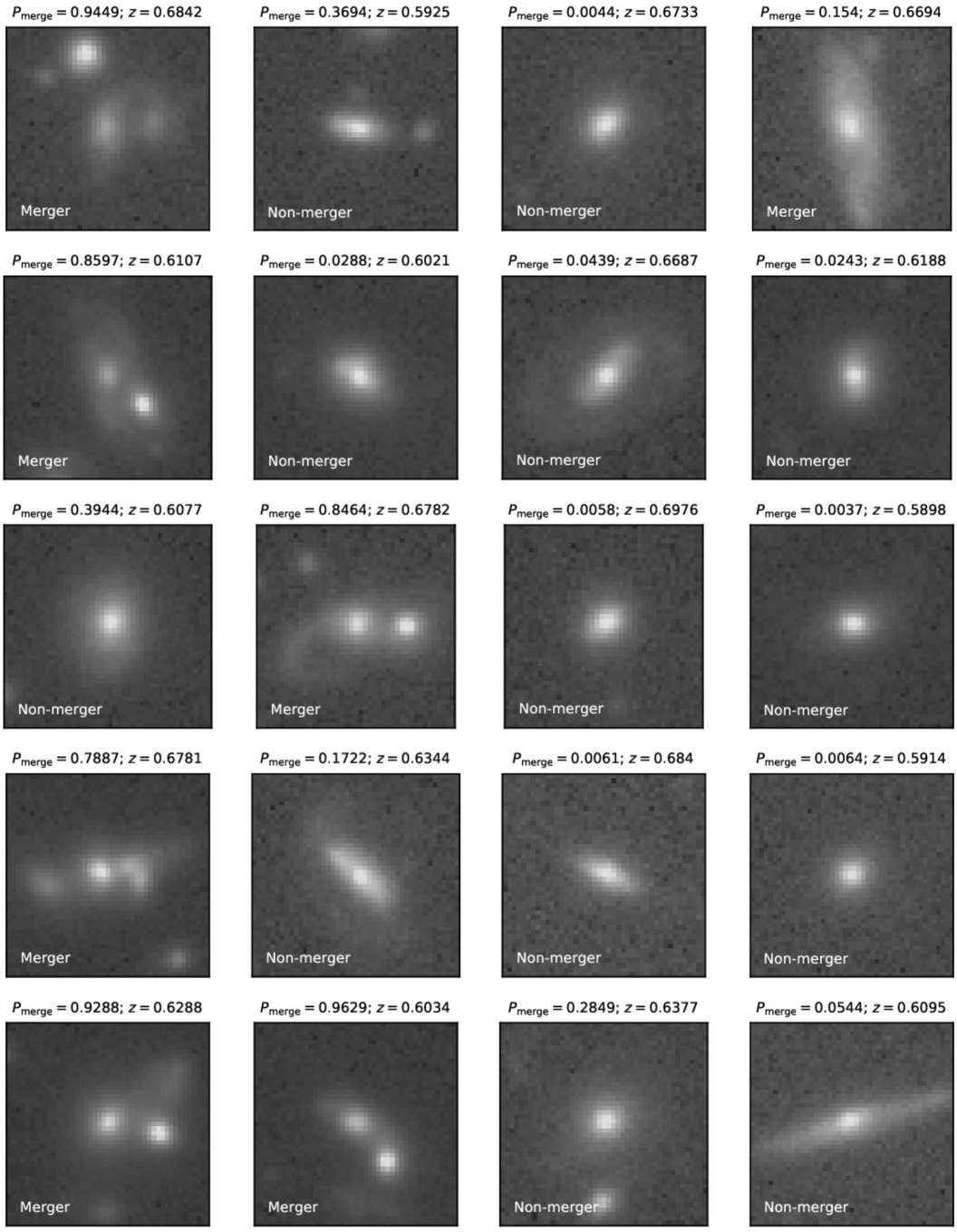


Figure A.3—continued.

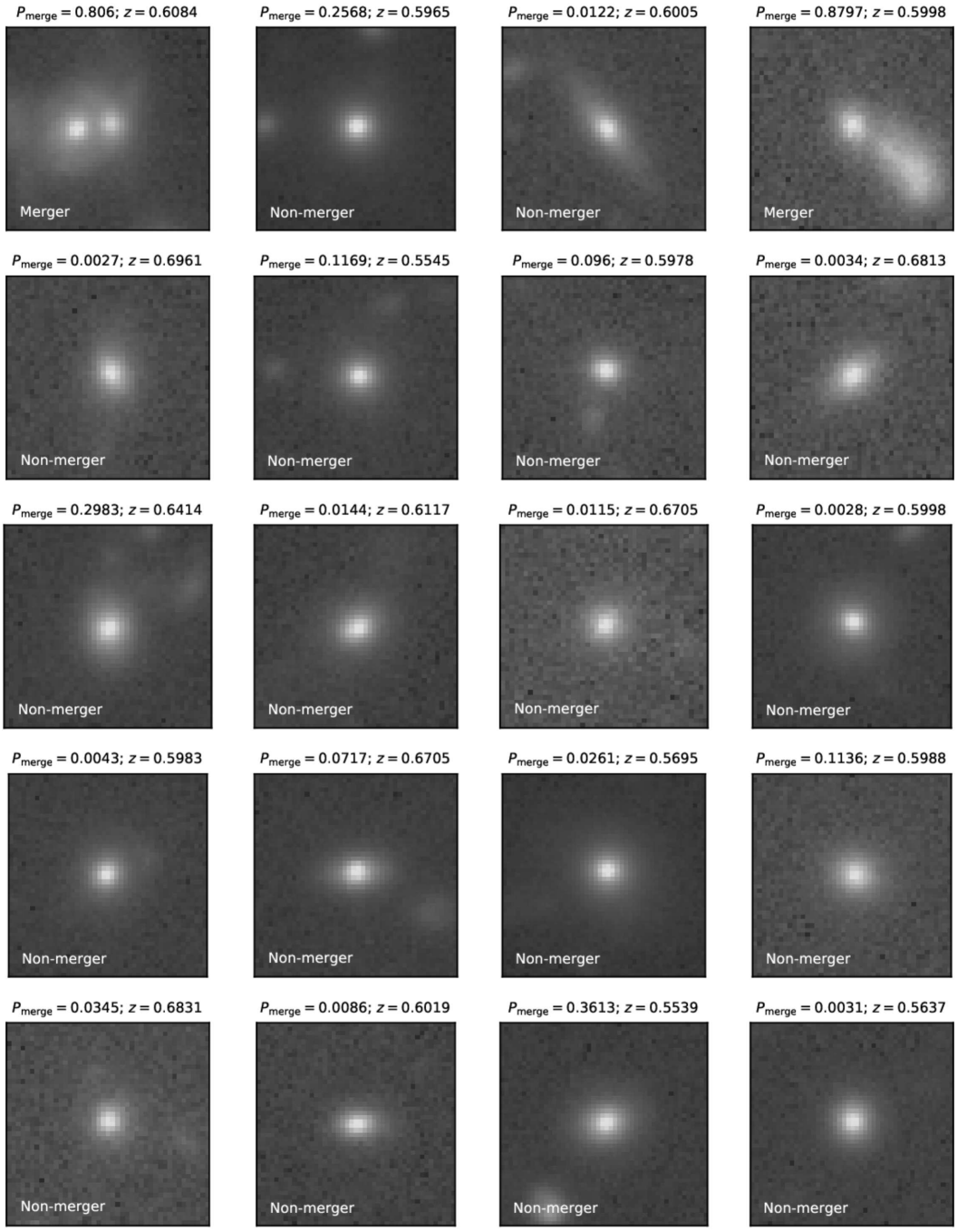


Figure A.3—continued.

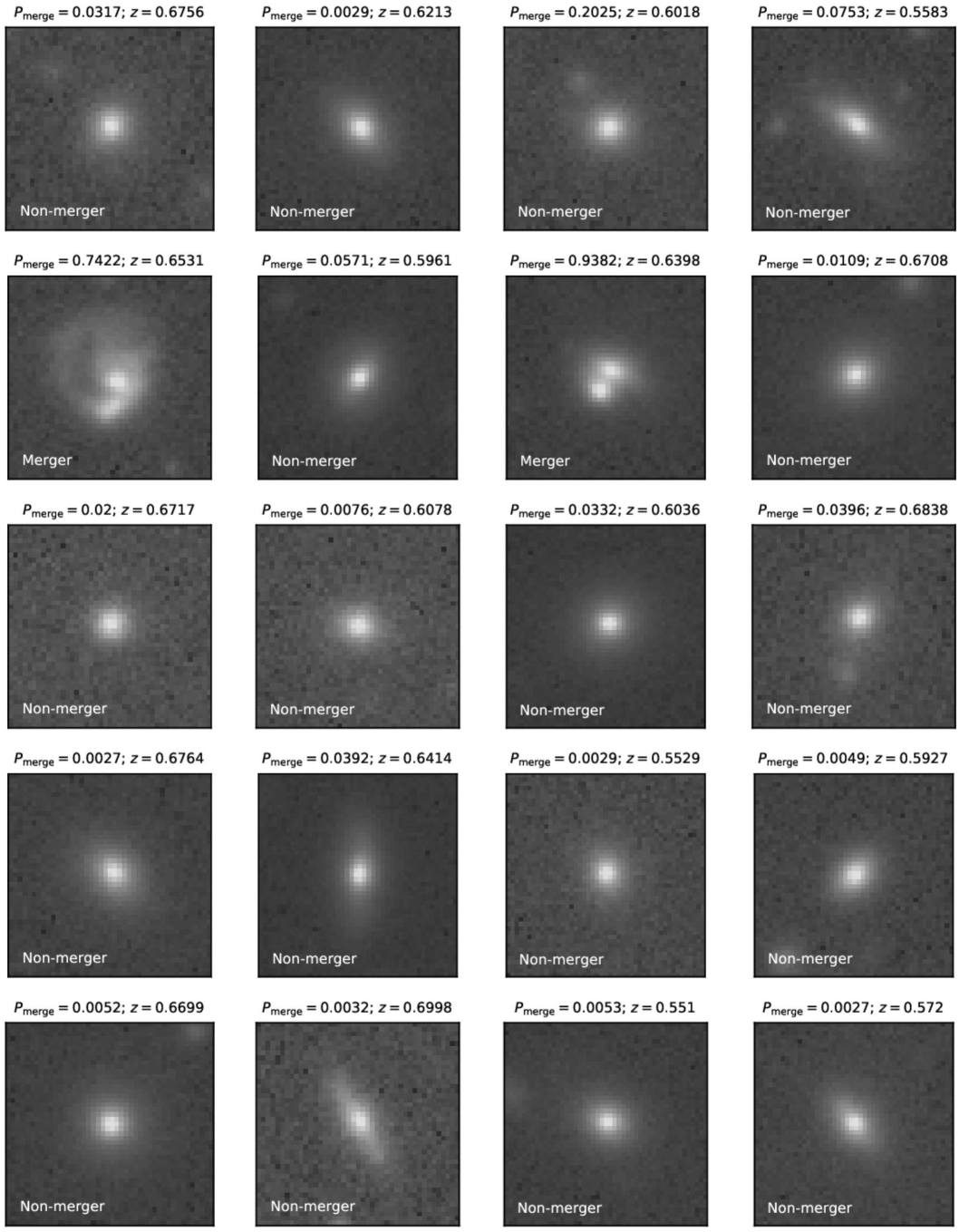


Figure A.3—continued.

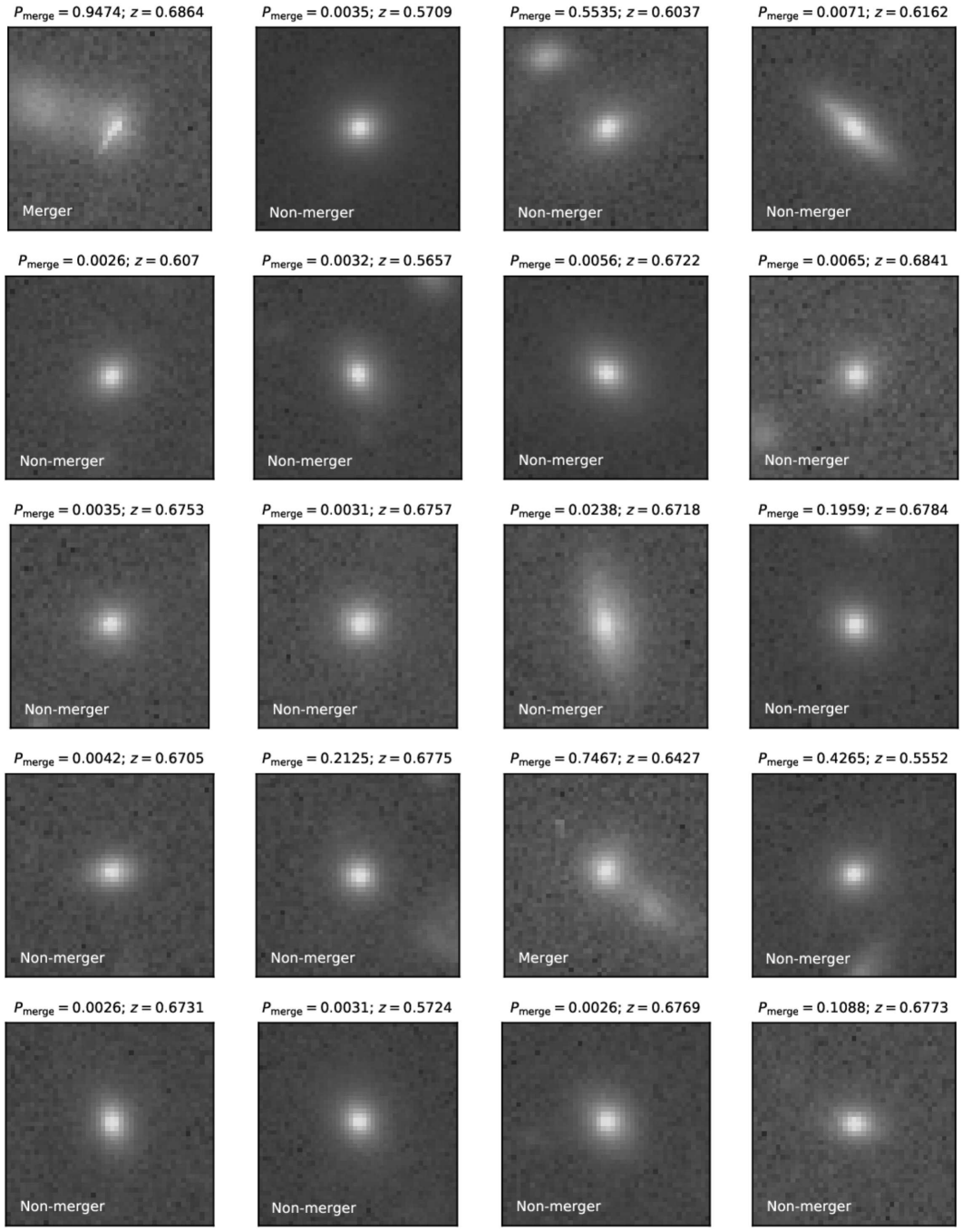


Figure A.3—continued.

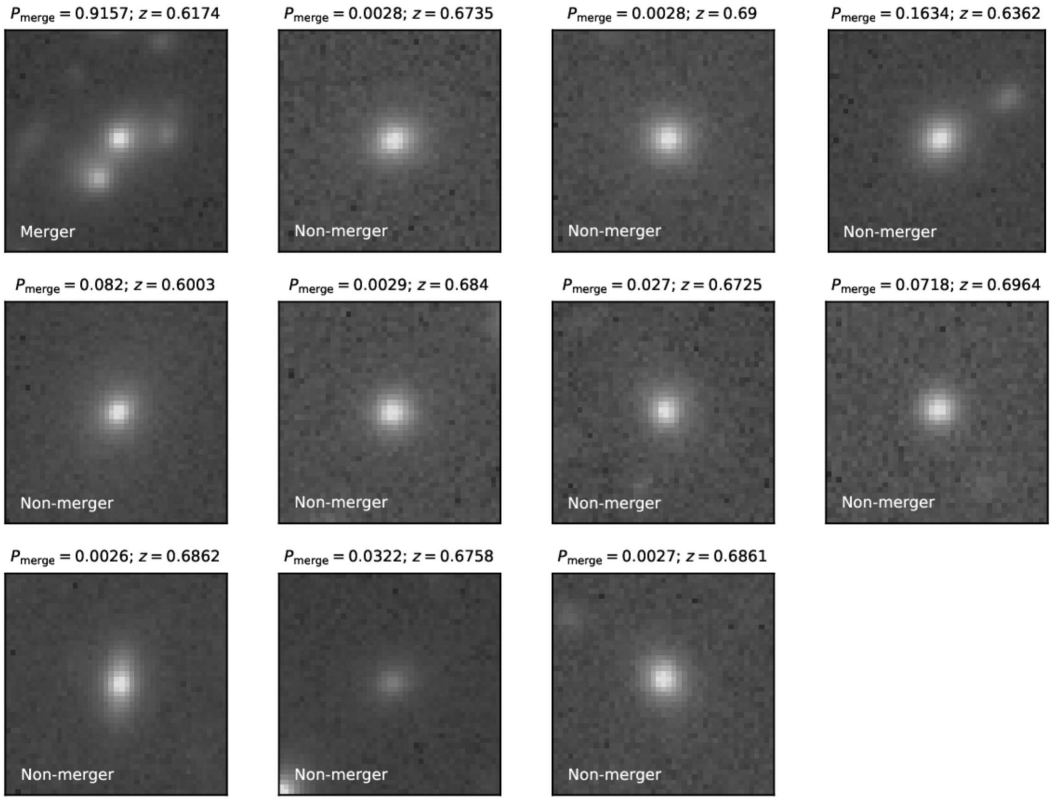


Figure A.3—continued.

A.4 $0.7 \leq z_{\mathbf{phot}} < 0.85$

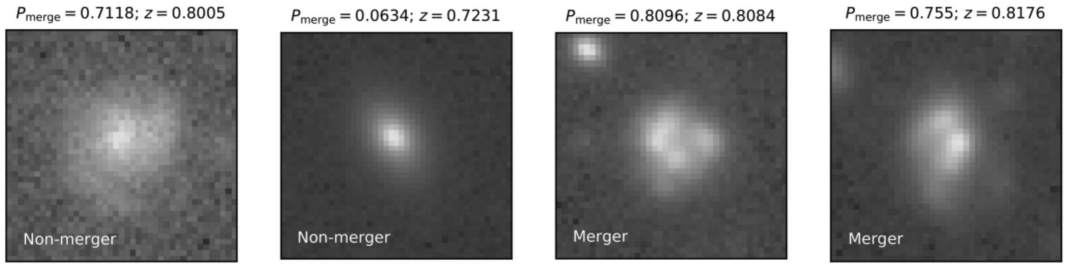


Figure A.4: Training set galaxies in the 4^{th} redshift bin ($0.7 \leq z_{\text{phot}} < 0.85$).

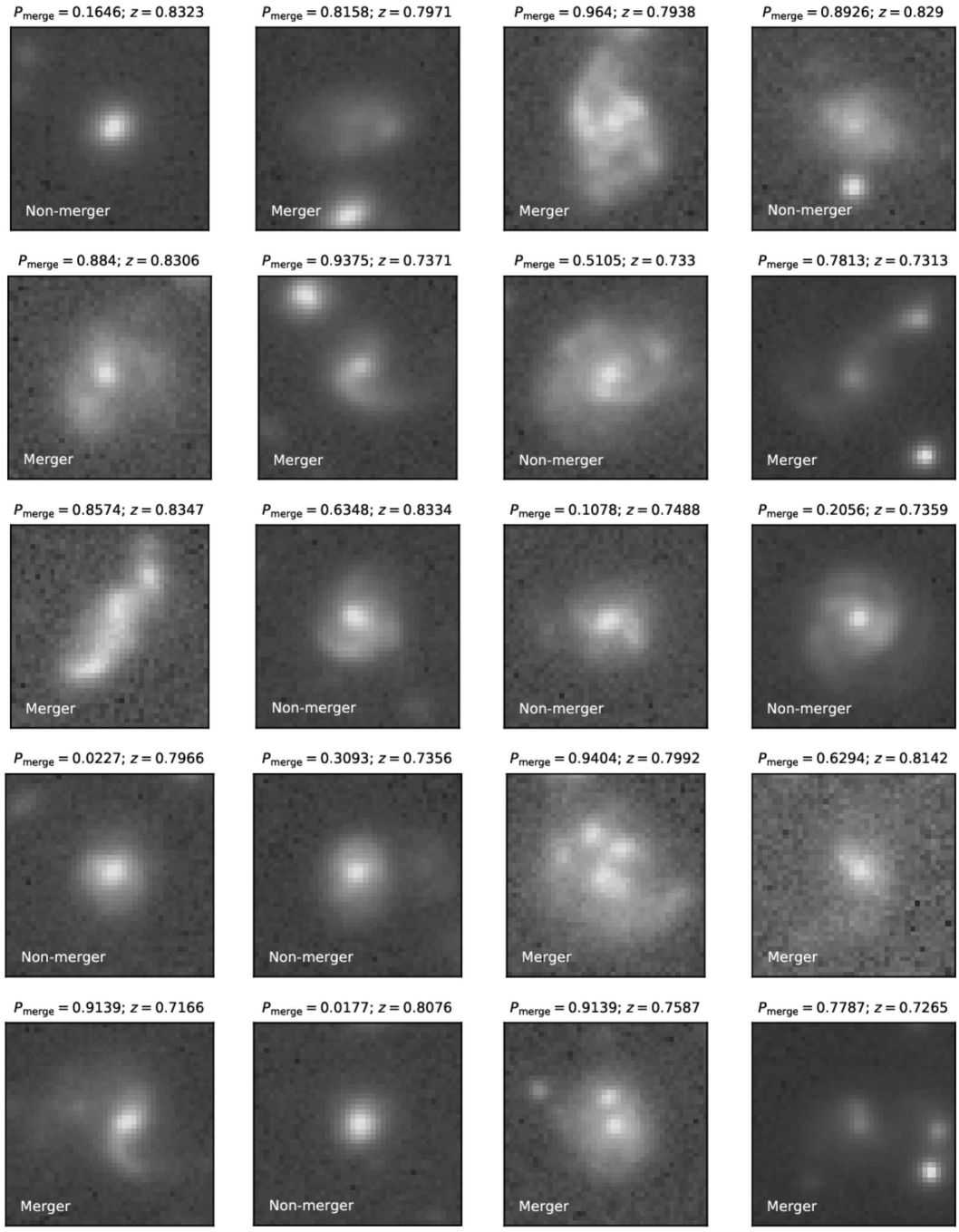


Figure A.4—continued.

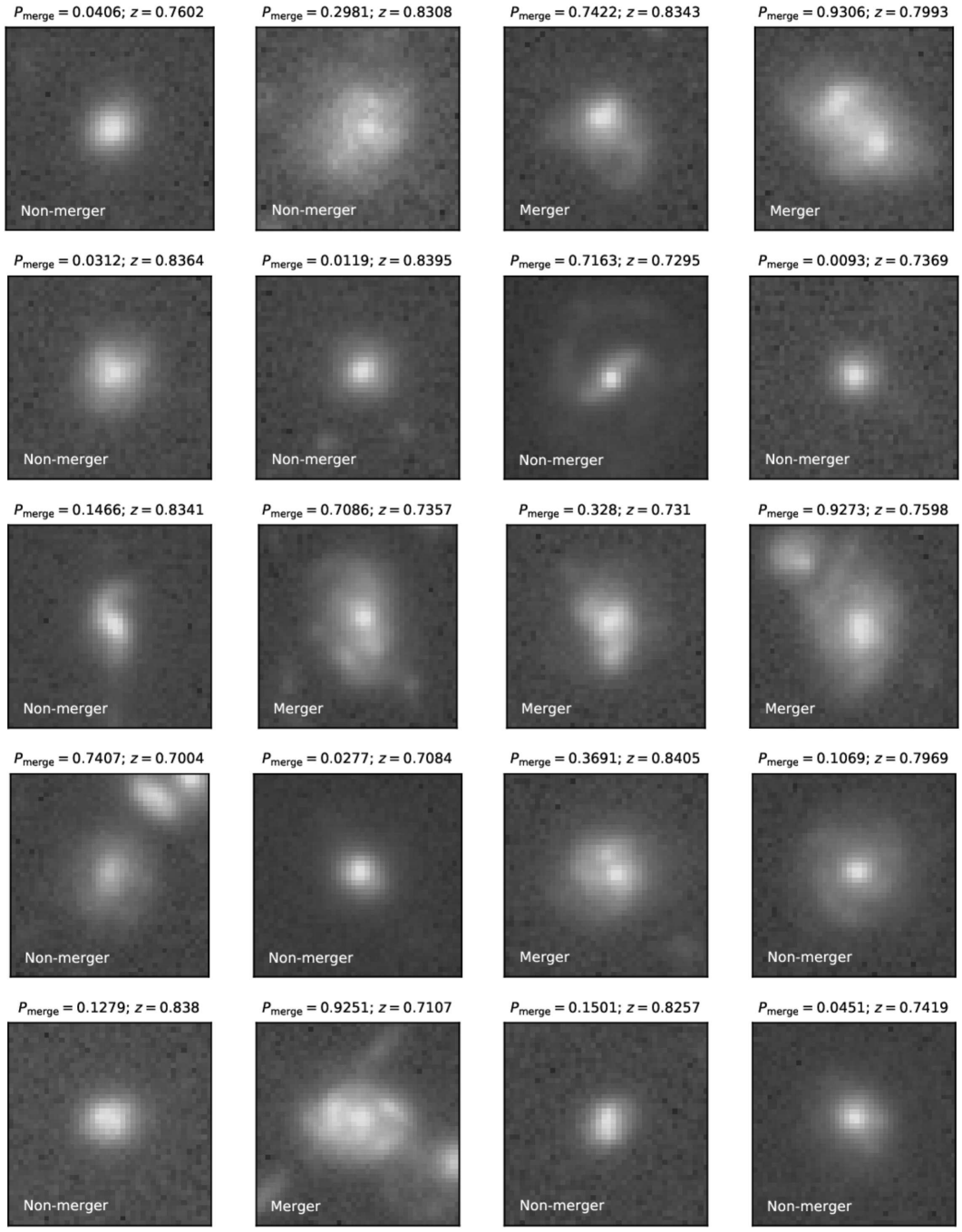


Figure A.4—continued.

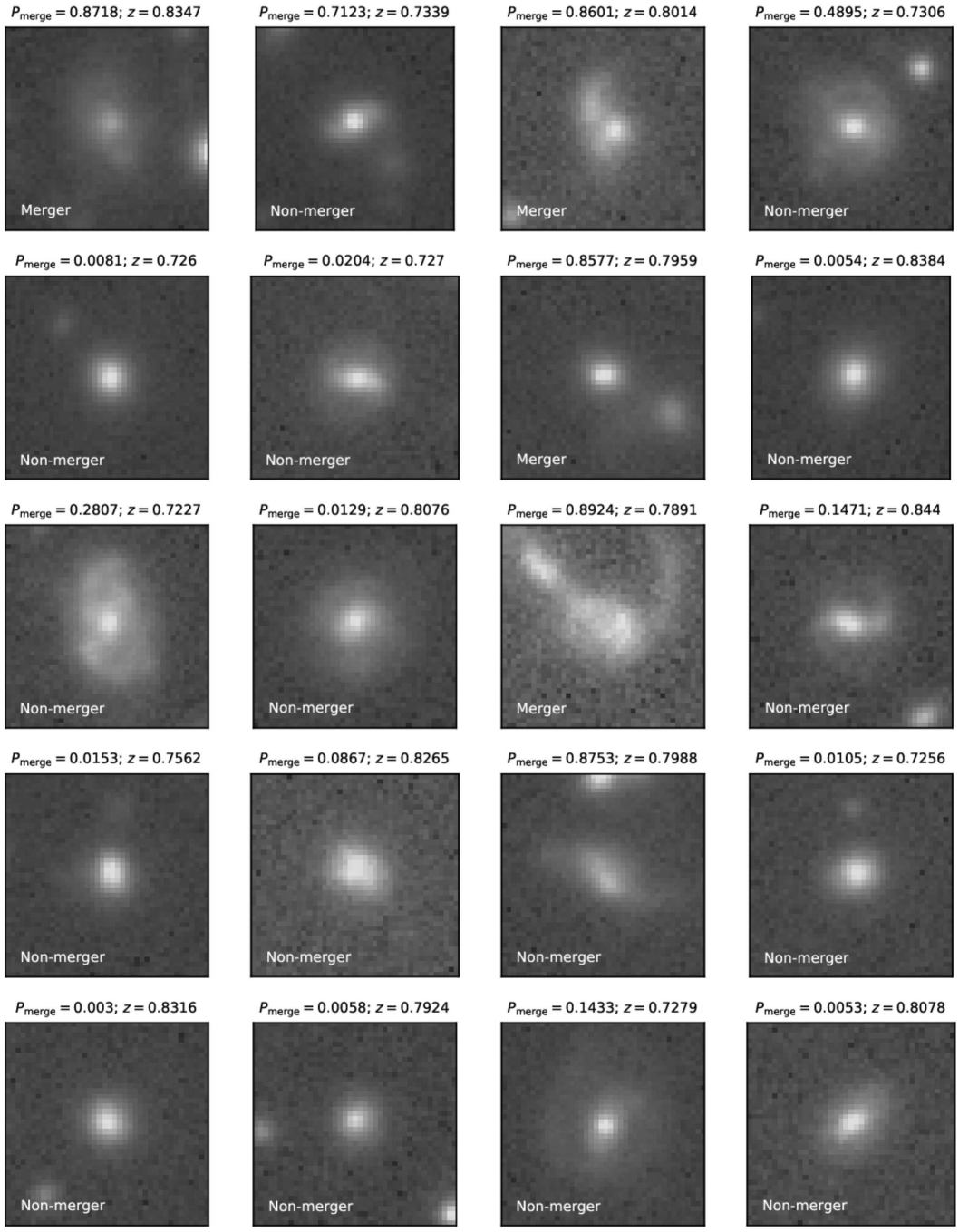


Figure A.4—continued.

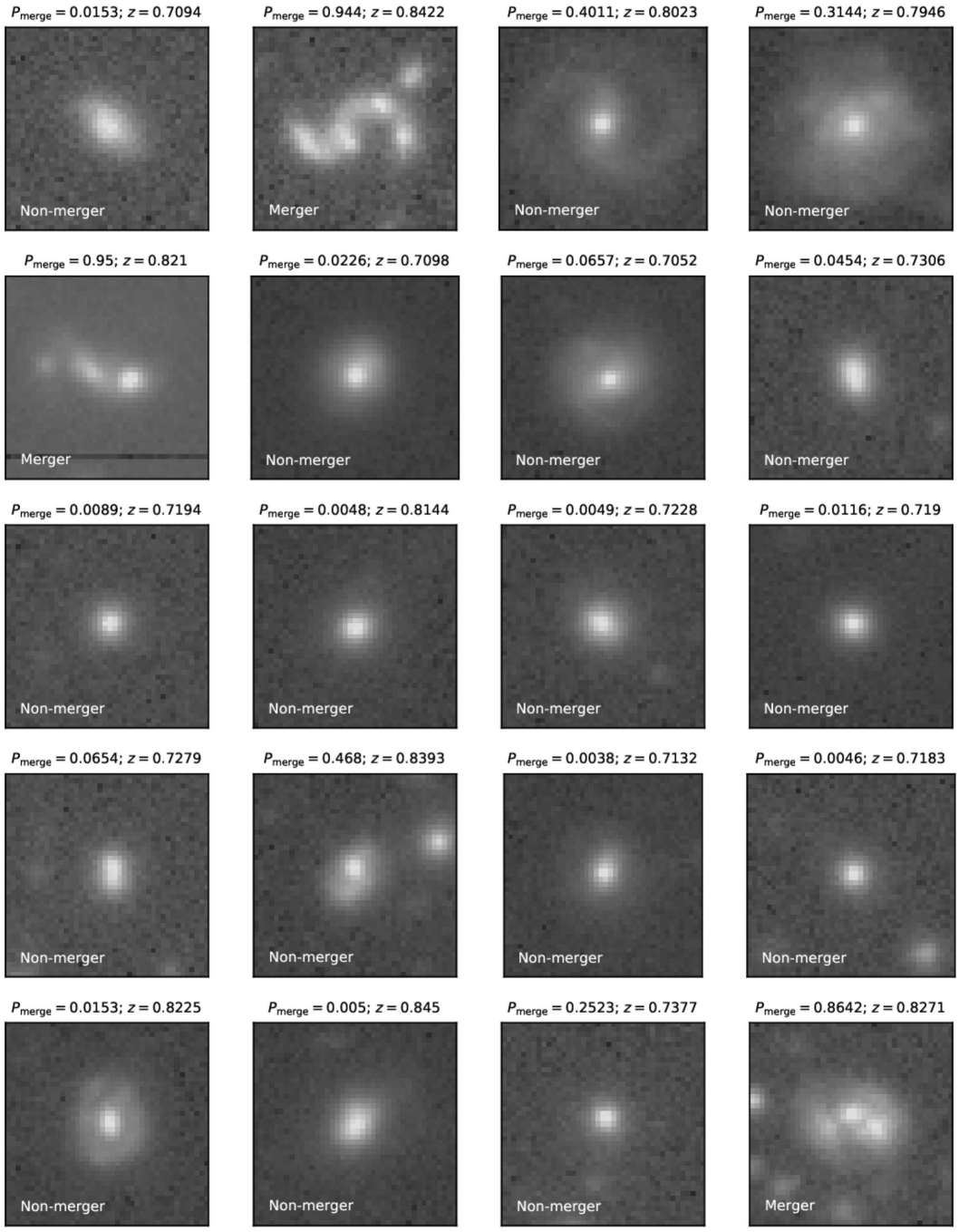


Figure A.4—continued.

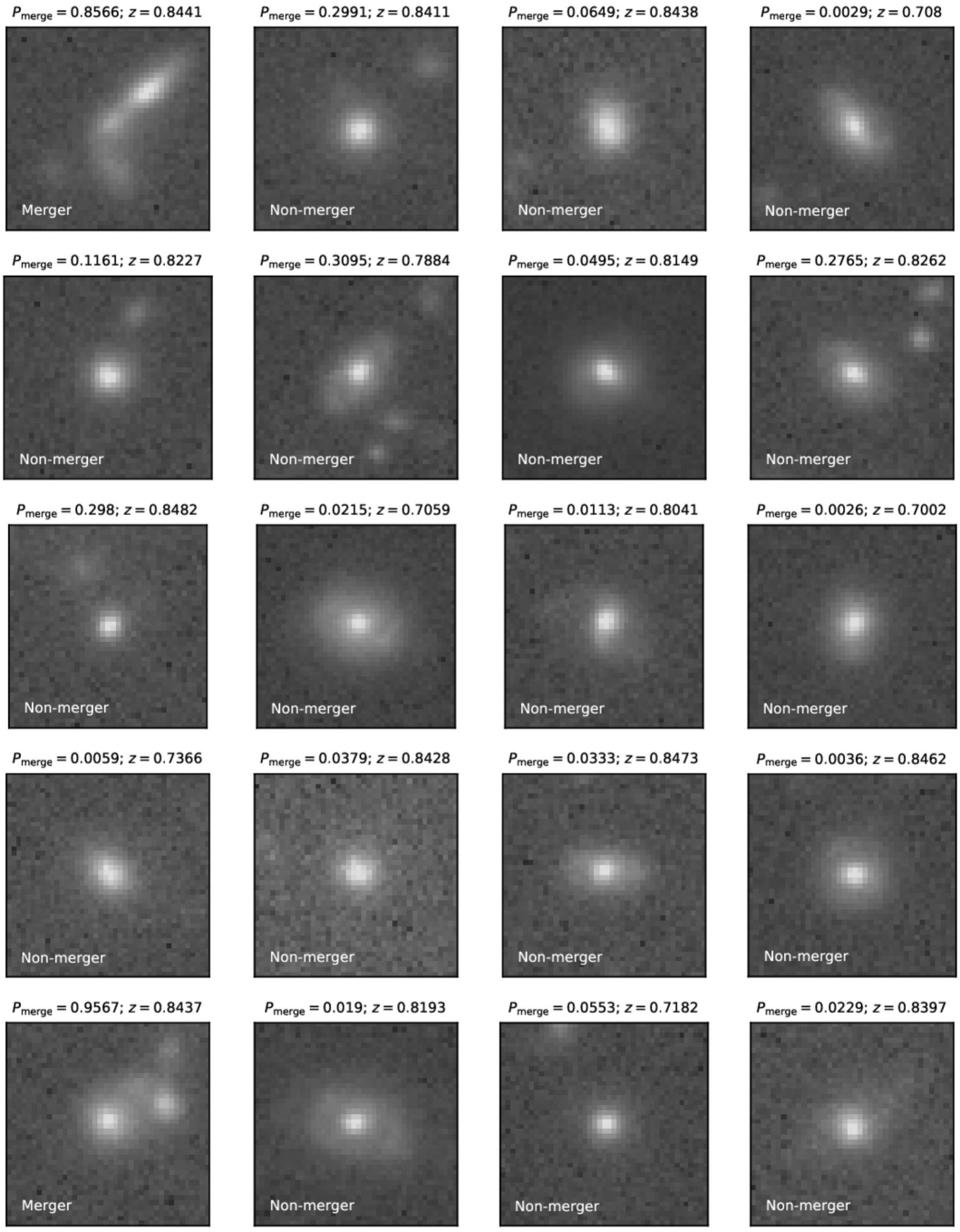


Figure A.4—continued.

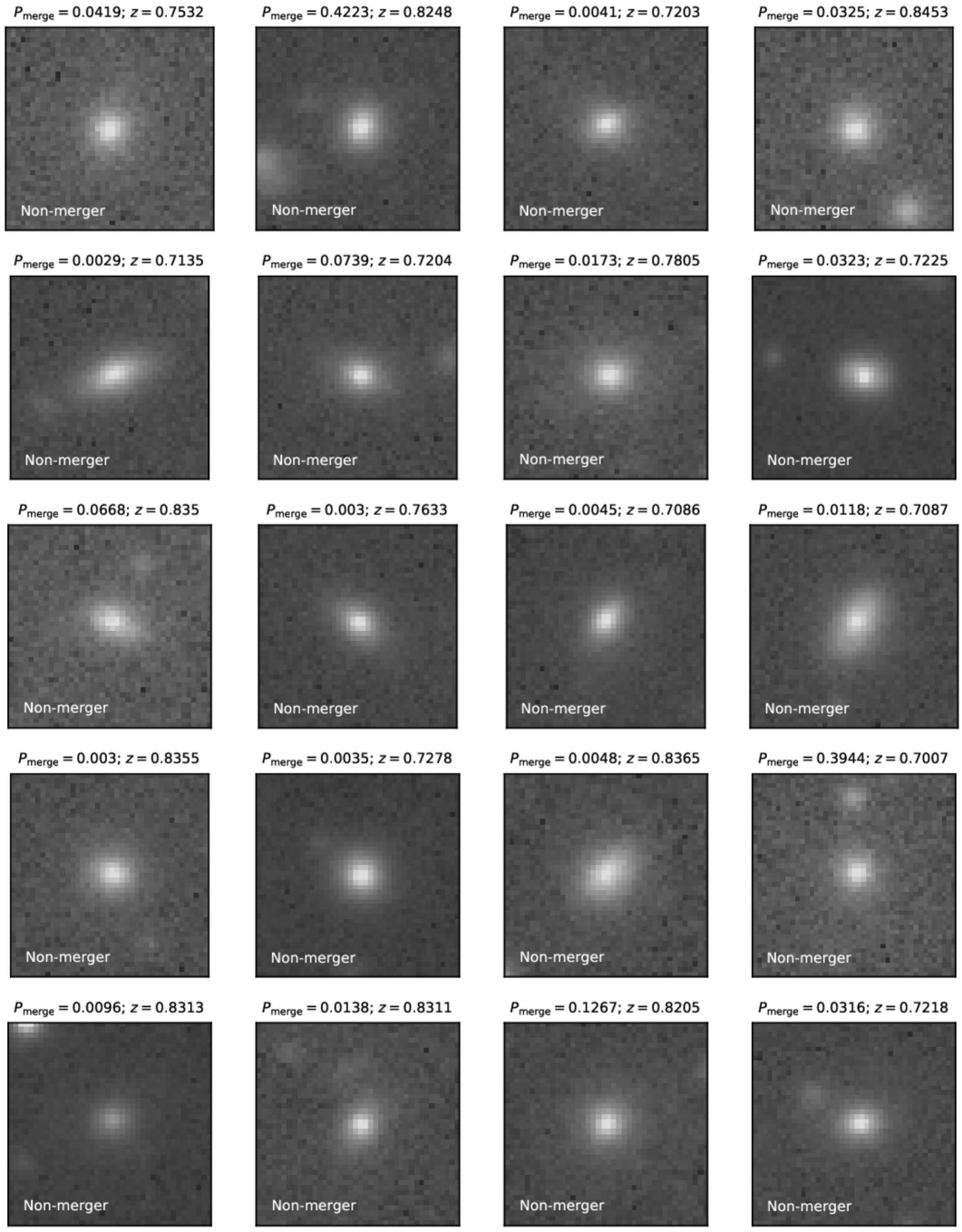


Figure A.4—continued.

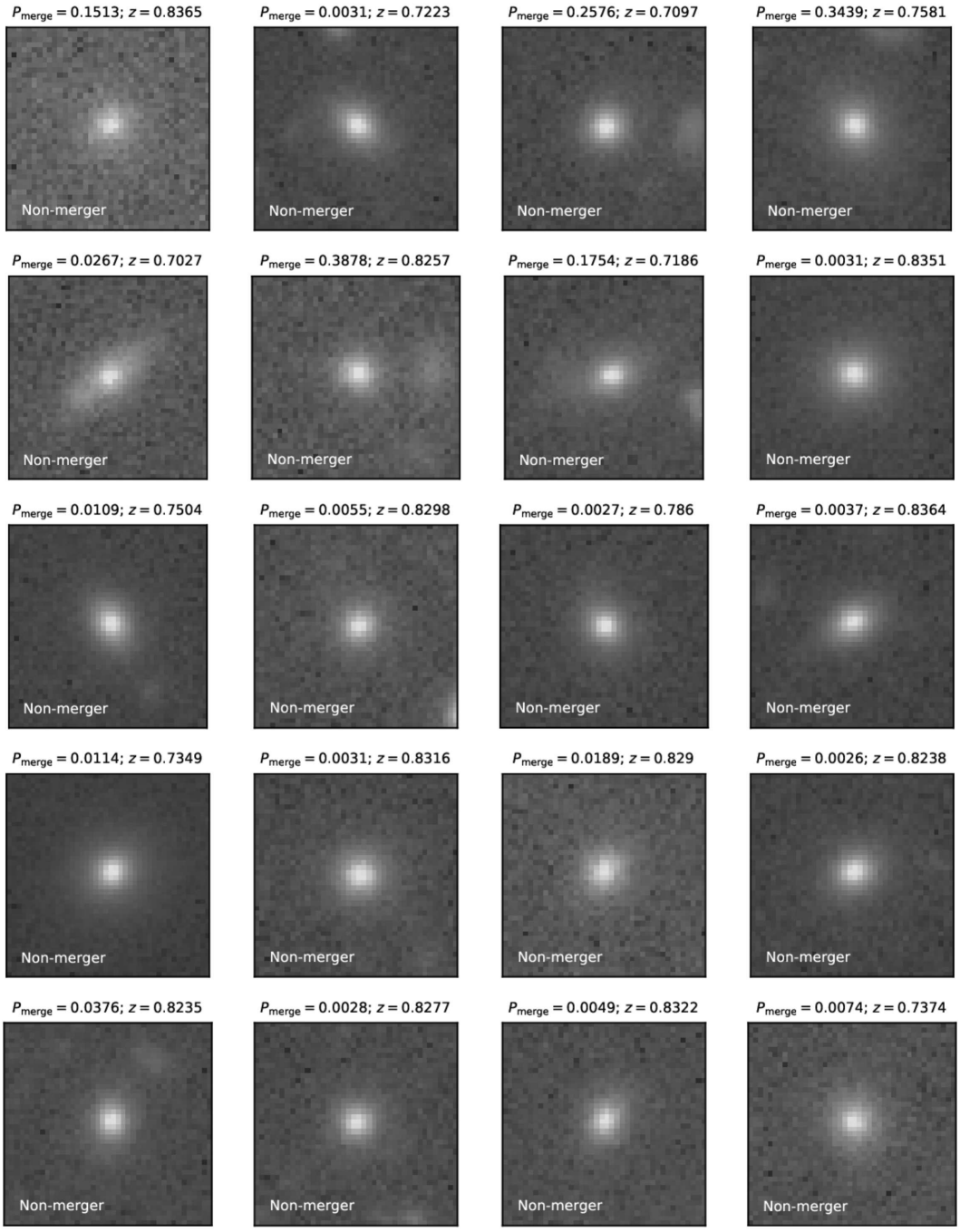


Figure A.4—continued.

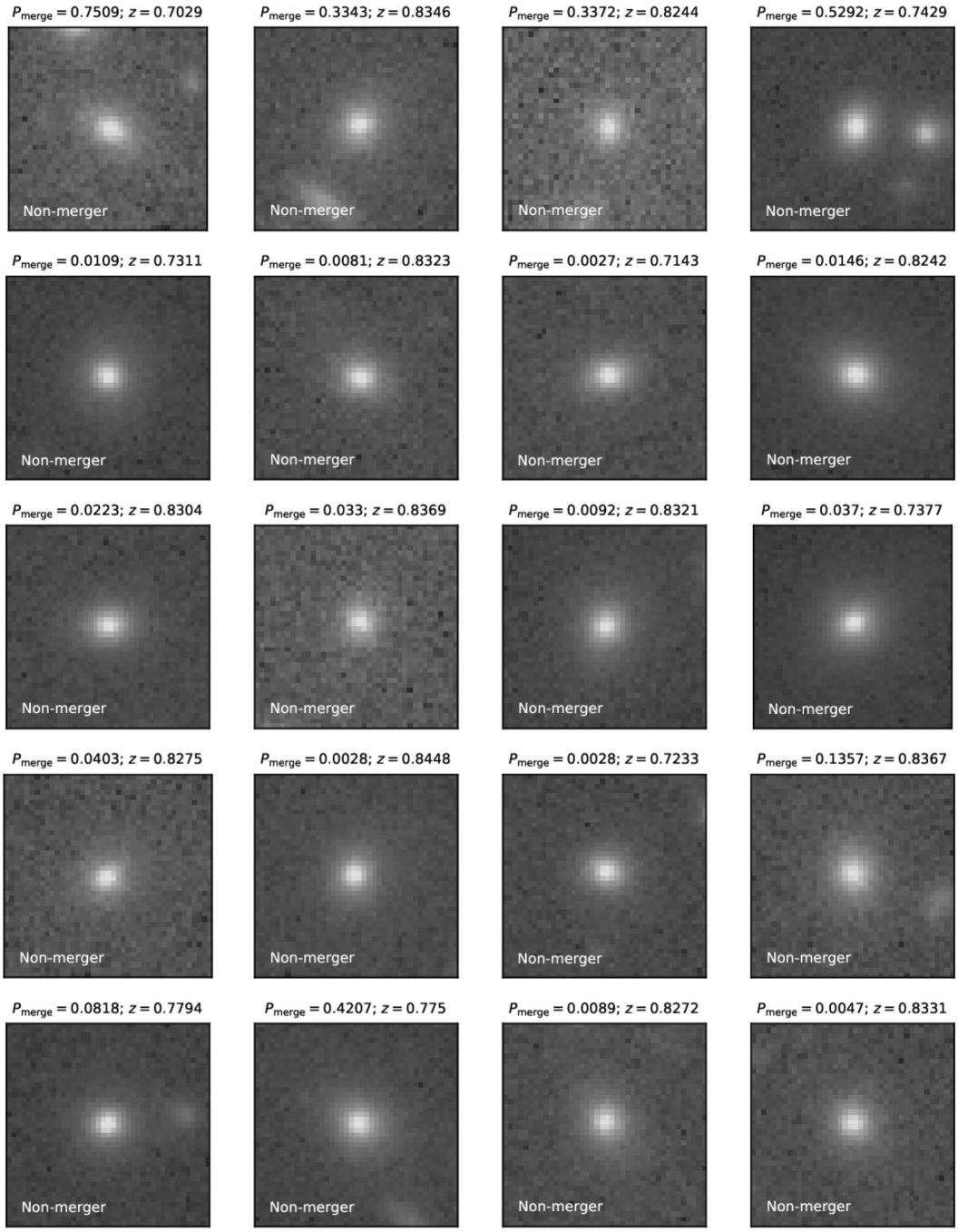


Figure A.4—continued.

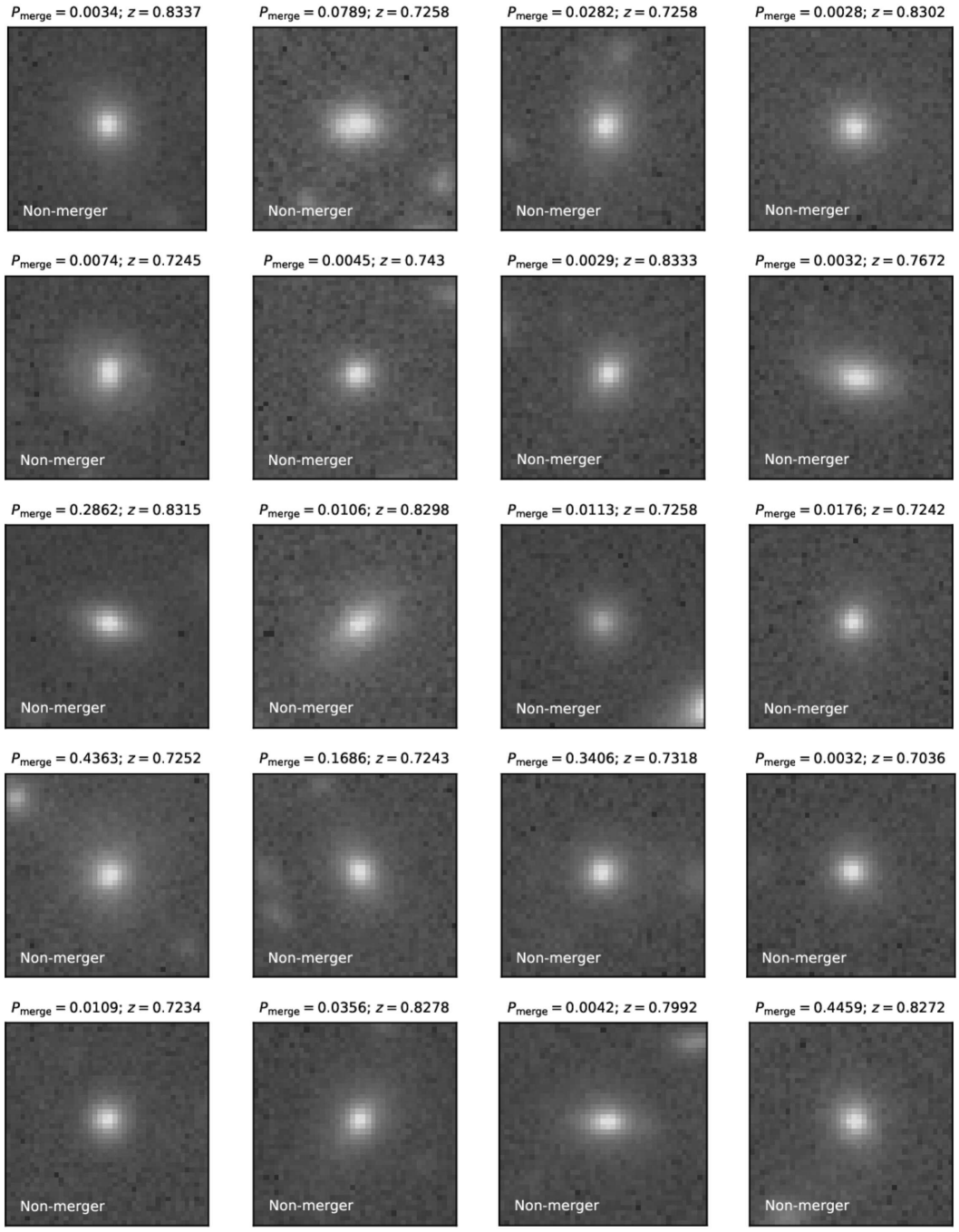


Figure A.4—continued.

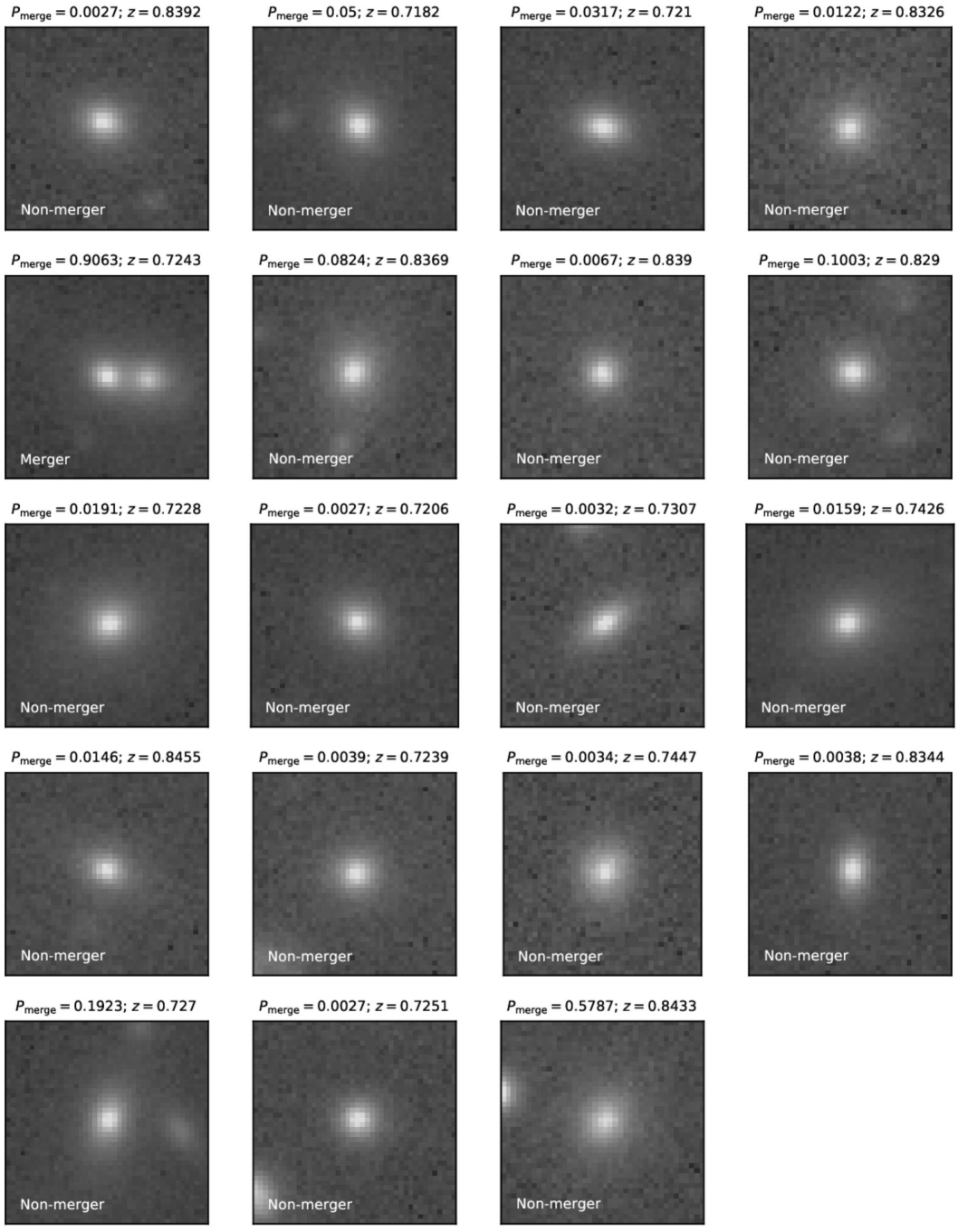


Figure A.4—continued.

A.5 $0.85 \leq z_{\rm phot} \leq 1.0$

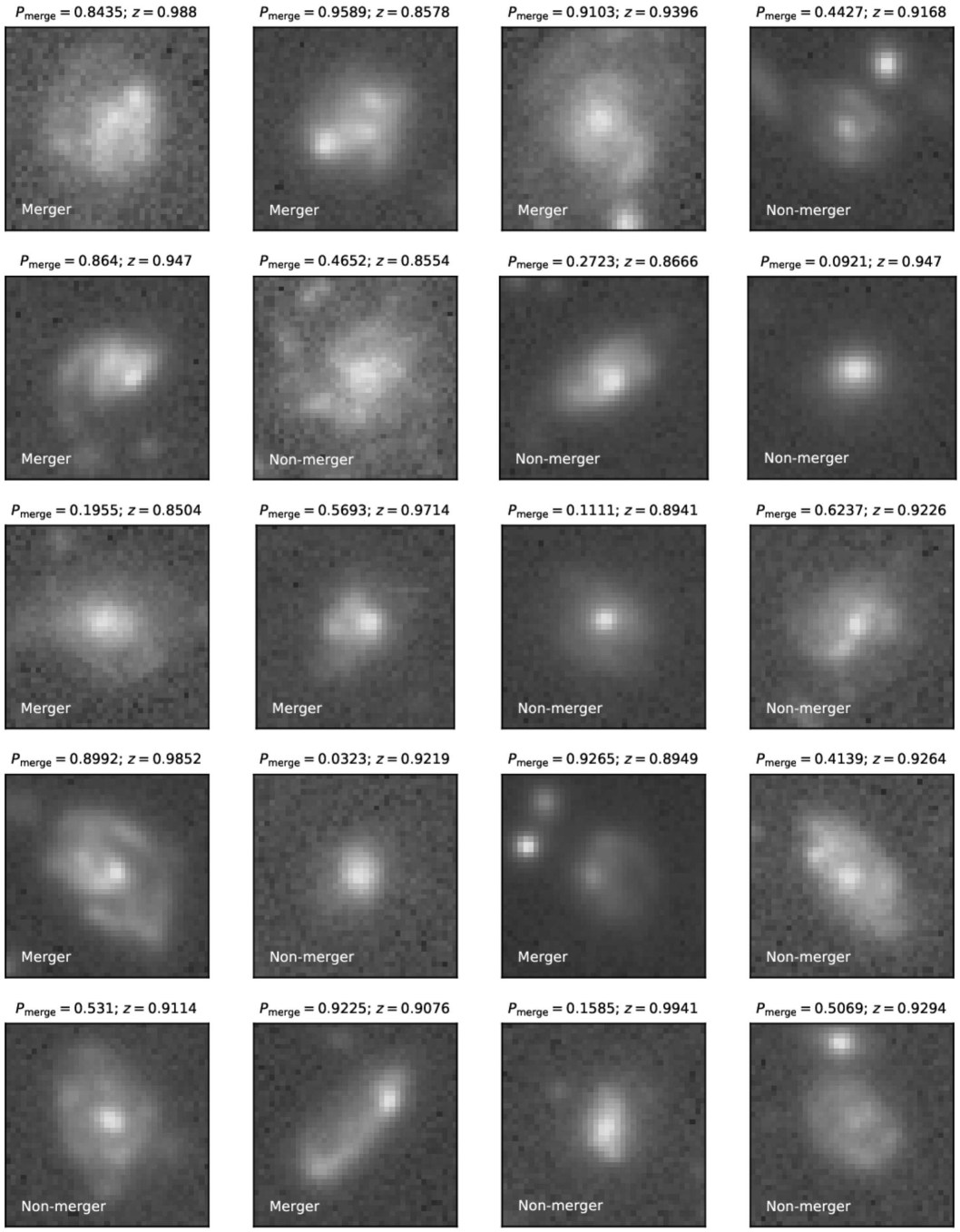


Figure A.5: Training set galaxies in the 5$^{\rm th}$ redshift bin ($0.85 \leq z_{\rm phot} \leq 1.0$).

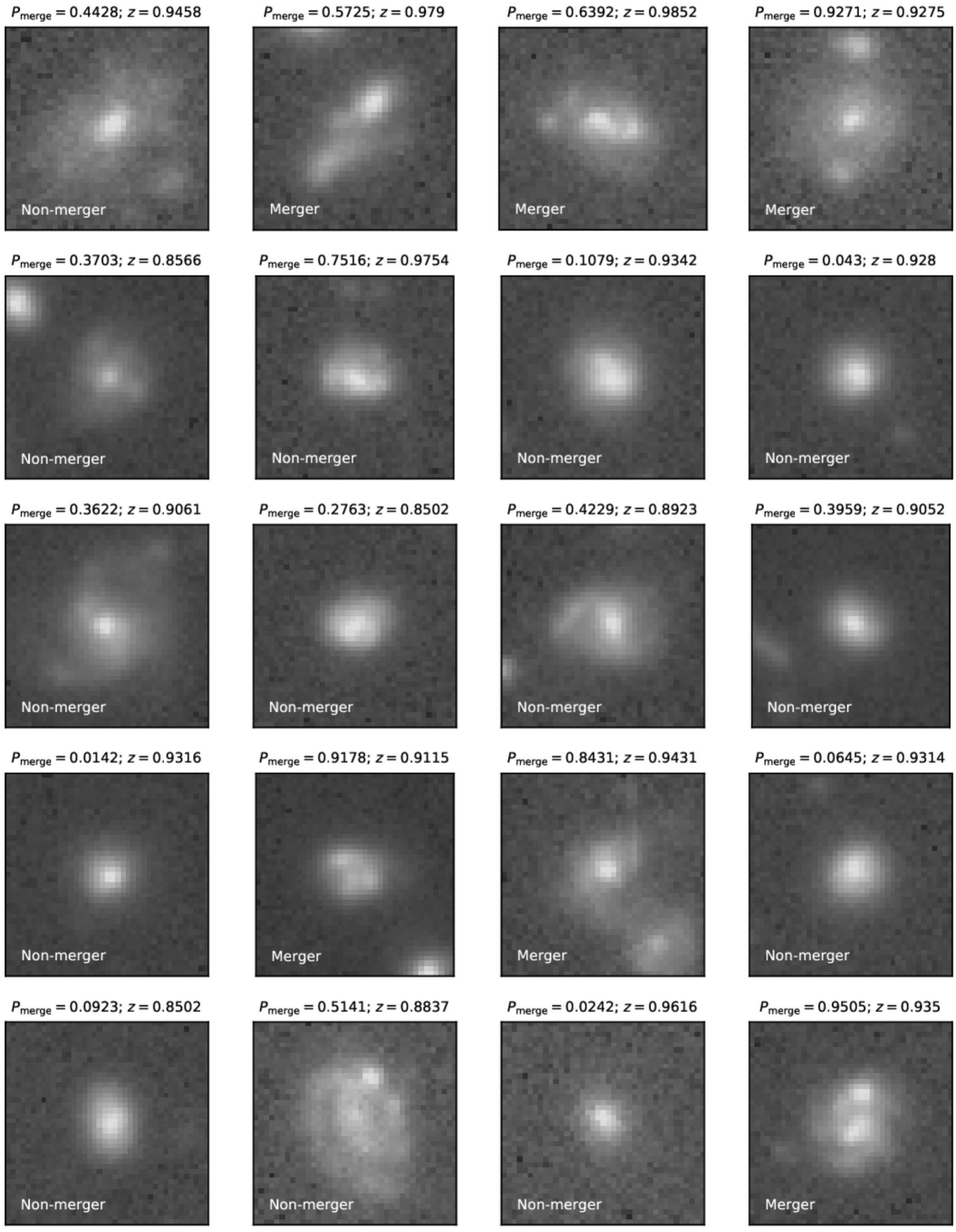


Figure A.5−continued.

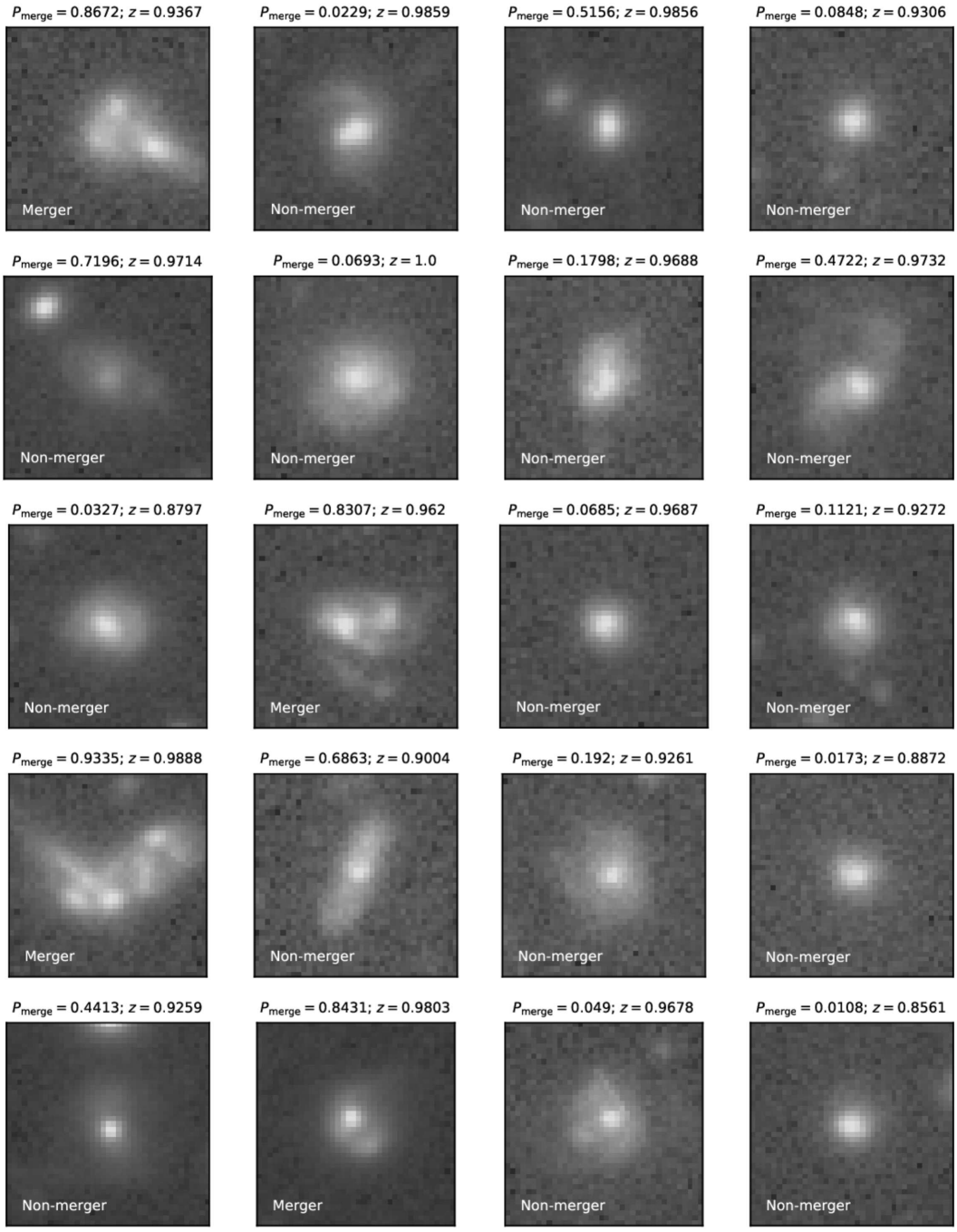


Figure A.5—continued.

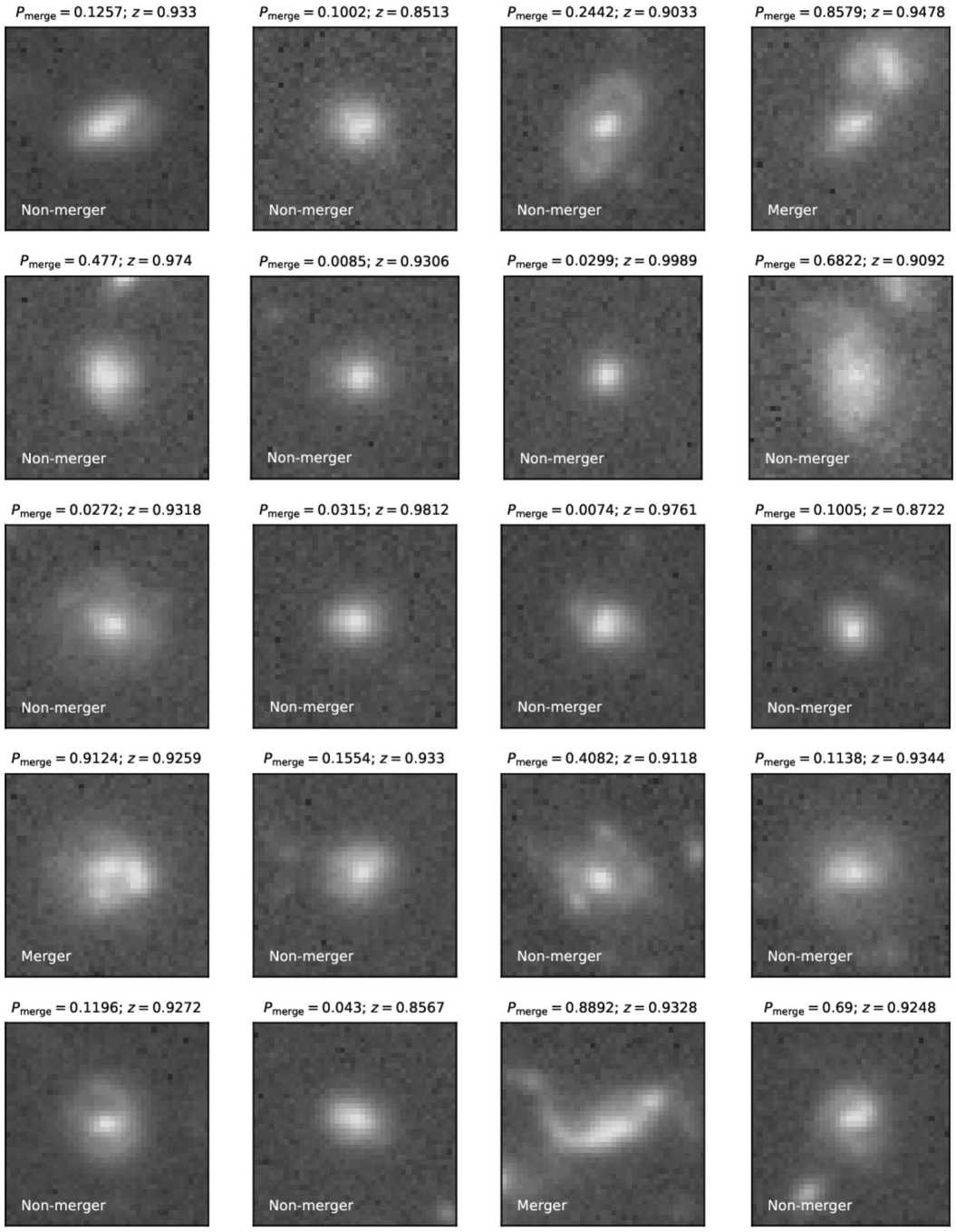


Figure A.5—continued.

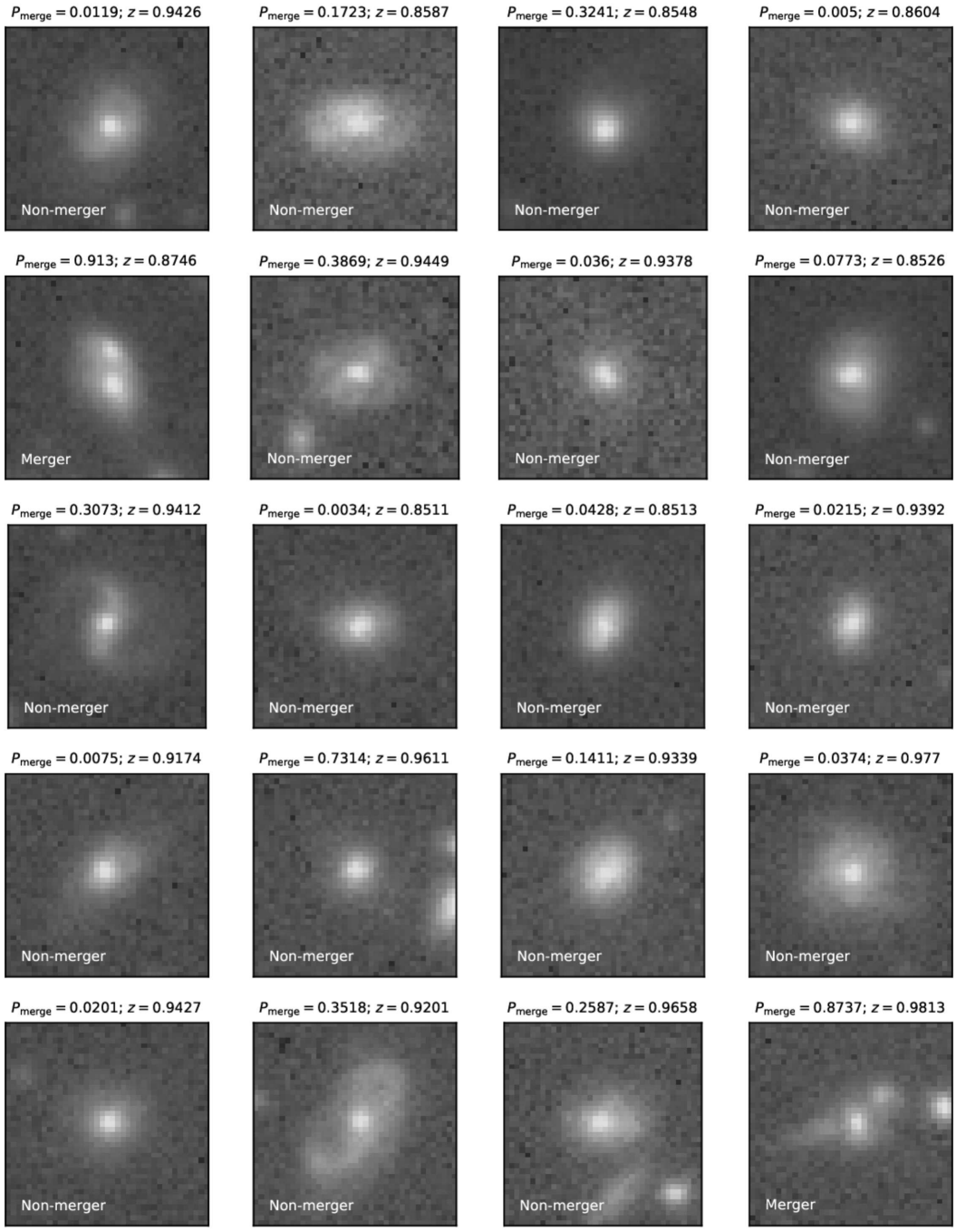


Figure A.5—continued.

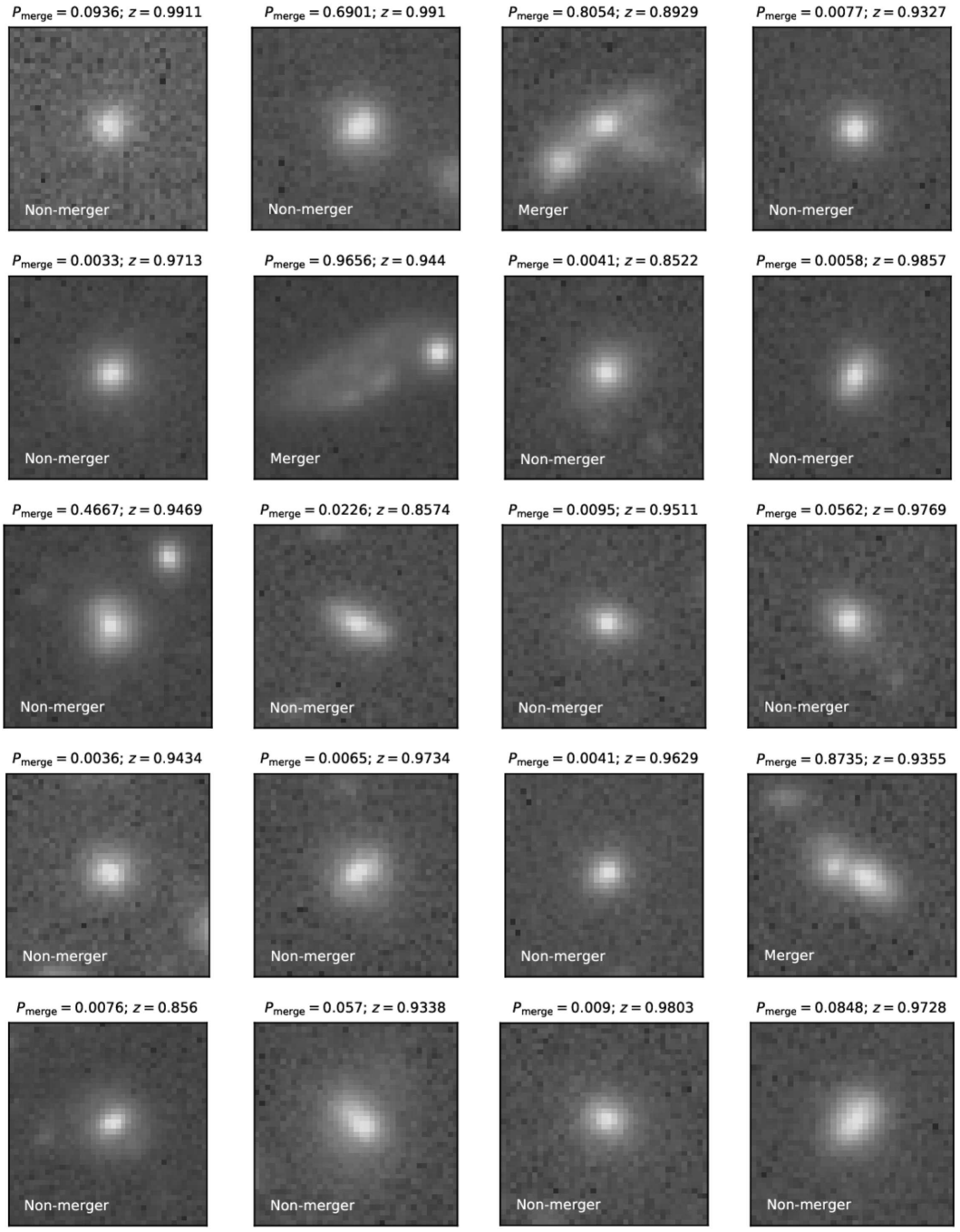


Figure A.5—continued.

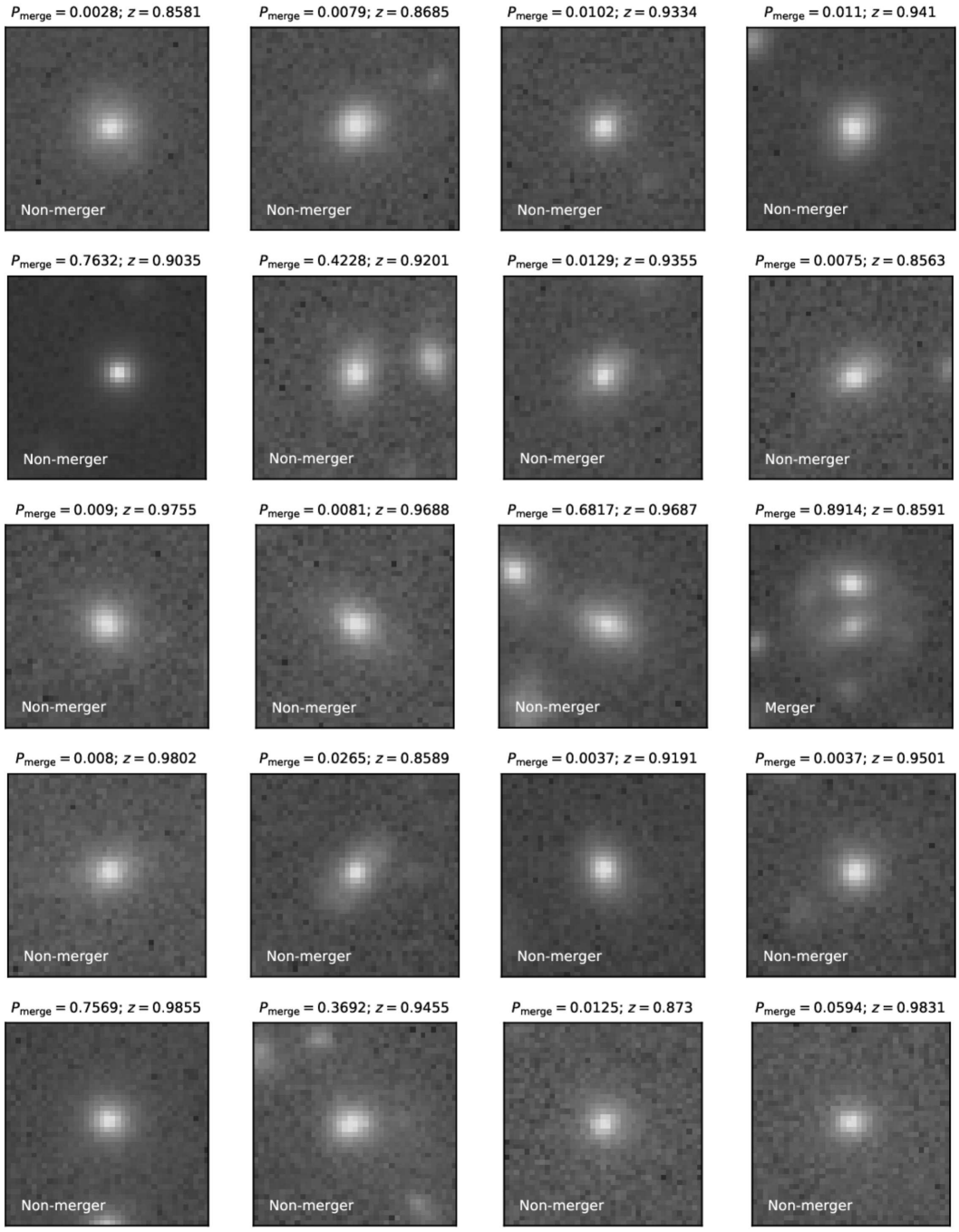


Figure A.5—continued.

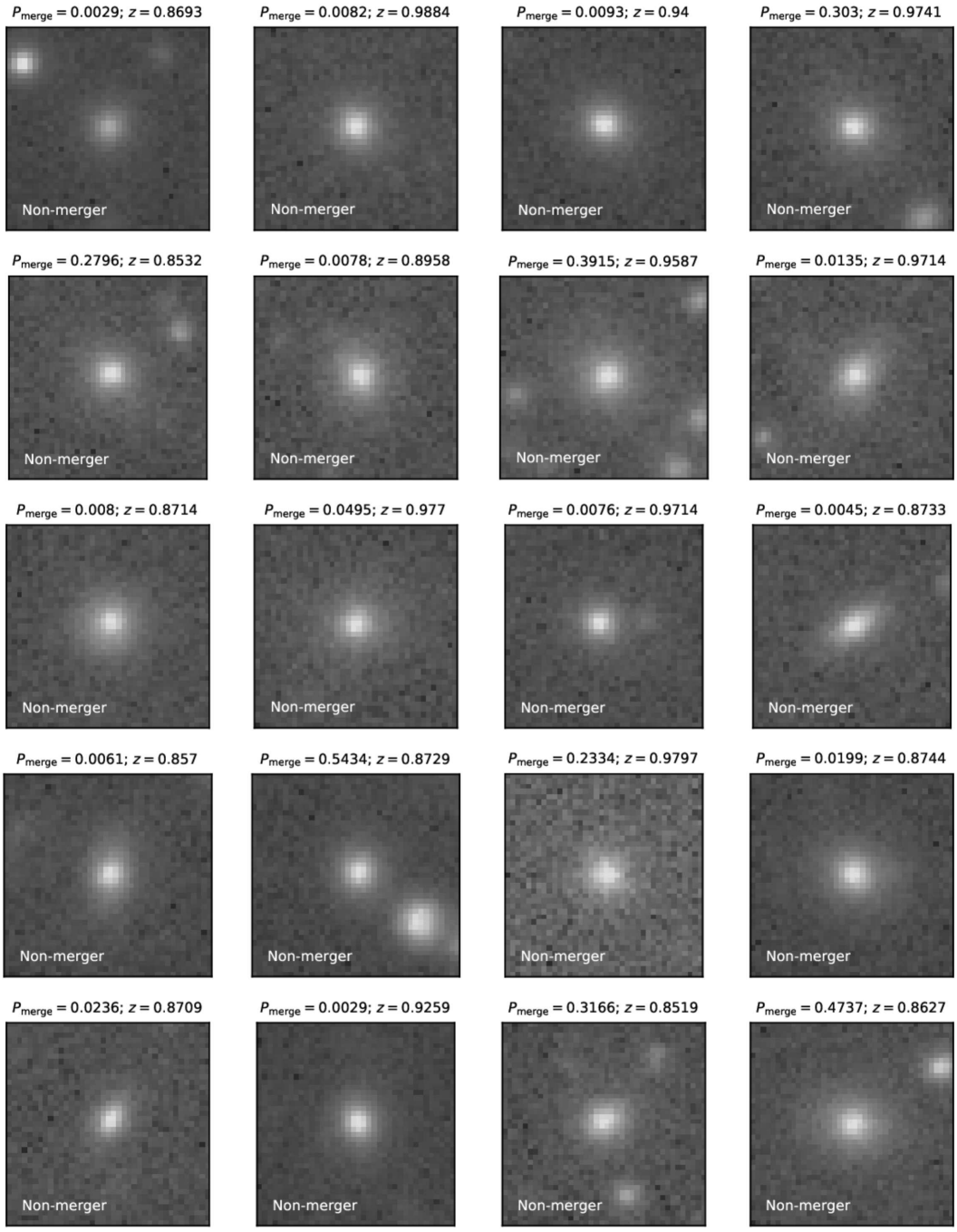


Figure A.5—continued.

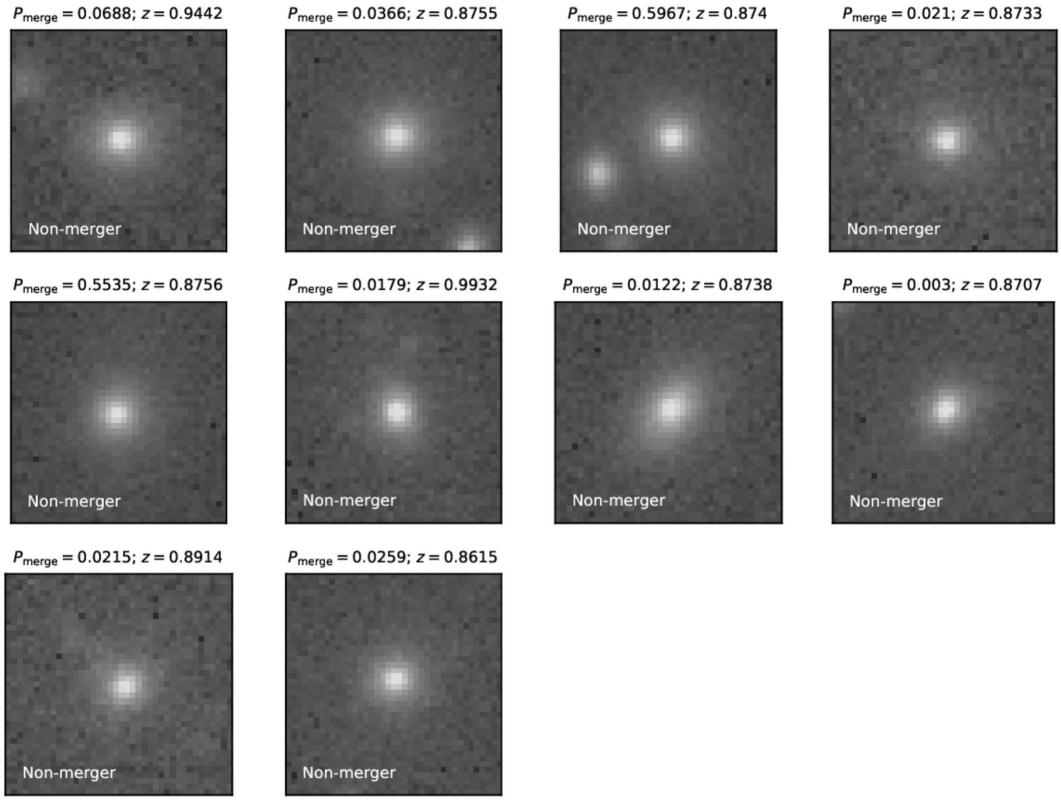


Figure A.5−continued.