Toward Complete Merger Identification at Cosmic Noon with Deep Learning

Aimee Schechter
Department of Astrophysical and Planetary Sciences
University of Colorado Boulder
Boulder, CO 80309
aimee.schechter@colorado.edu

Aleksandra Ćiprijanović
Computational Science and AI Directorate
Fermi National Accelerator Laboratory
Batavia, IL 60510
Department of Astronomy and Astrophysics
University of Chicago
Chicago, IL 60637
aleksand@fnal.gov

Rebecca Nevin
Computational Science and AI Directorate
Fermi National Accelerator Laboratory
Batavia, IL 60510
rnevin@fnal.gov

Julie Comerford
Department of Astrophysical
and Planetary Sciences
University of Colorado Boulder
Boulder, CO 80309
julie.comerford@colorado.edu

Xuejian Shen
TAPIR, California Institute of Technology
Pasadena, CA 91106
Kavli Institute for Astrophysics and Space Research
Massachusetts Institute of Technology
Cambridge, MA 02139
xuejian@mit.edu

Aaron Stemo
Department of Physics and Astronomy
Vanderbilt University
Nashville, TN 37235
aaron.m.stemo@vanderbilt.edu

Laura Blecha
Department of Physics
University of Florida
Gainesville, FL 32611
lblecha@ufl.edu

Abstract

As we enter the era of large imaging surveys such as *Roman*, *Rubin*, and *Euclid*, a deeper understanding of potential biases and selection effects in optical astronomical catalogs created with the use of ML-based methods is paramount. This work focuses on a deeper understanding of the performance and limitations of deep learning-based classifiers as tools for galaxy merger identification. We train a ResNet18 model on mock Hubble Space Telescope CANDELS images from the IllustrisTNG50 simulation. Our focus is on a more challenging classification of galaxy mergers and nonmergers at higher redshifts $1 < z < 1.5$, including minor mergers and lower mass galaxies down to the stellar mass of $10^8 M_\odot$. We demonstrate, for the first time, that a deep learning model, such as the one developed in this work, can successfully identify even minor and low mass mergers even at these redshifts. Our model achieves overall accuracy, purity, and completeness of 73%. We show that some galaxy mergers can only be identified from certain

39th Conference on Neural Information Processing Systems (NeurIPS 2025) Workshop: Machine Learning and the Physical Sciences (ML4PS).

observation angles, leading to a potential upper limit in overall accuracy. Using Grad-CAMs and UMAPs, we more deeply examine the performance and observe a visible gradient in the latent space with stellar mass and specific star formation rate, but no visible gradient with merger mass ratio or merger stage.

1 Introduction

Galaxies grow through cosmic time hierarchically. In addition to growing in total mass, this process also forms new stars, changes galaxy morphologies, and can trigger active galactic nucleus (AGN) activity (e.g., [3, 20, 11, 32, 34, 33, 12, 23, 4]). To fully understand the role of galaxy mergers in star formation and AGN activity, we need large catalogs of merging galaxies at different merger stages and across a range of stellar masses and merger mass ratios. However, this can be challenging as many non-parametric methods are calibrated at low-z and are targeted at identifying major mergers (merger mass ratio $\mu \geq$ 1/4) among high mass galaxies ($M_\star \gtrsim 10^{10} M_\odot$). These methods quantify the distributions of light in an image to measure how concentrated, clumpy, or asymmetric the distribution is [10, 17]. The merger stage adds another complication. Finding early-stage mergers with two identifiable nuclei and faint features like tidal tails is easier for non-parametric methods (e.g., tidal tails causing asymmetry) than finding late-stage mergers near coalescence. Close pair analyses, which find galaxies that are within a given separation in physical and velocity space, can identify early-stage mergers, but only looking at early stages leaves out half of the merging population. Combining non-parametric methods through Linear Discriminant Analysis (LDA) or a Random Forest (RF) (e.g., [22, 31, 27, 37]) is one way to get a more accurate and diverse catalog of mergers.

Convolutional Neural Networks (CNNs) offer an even more flexible method for finding mergers at different merger stages and redshifts since they have the ability to utilize all features present in galaxy images. They have already been applied to multiple mock and real imaging survey datasets (e.g., [7, 39, 5, 18, 26]). Still, the majority of these studies are focused on lower redshifts (all are at $z < 1$ except [38] at $z = 2$ and [26] at $3 < z < 5$), and higher mass galaxies (all above $M_\star = 10^9 M_\odot$ except [26]). We apply our CNN to a sample that includes both higher-z galaxies ($1 < z < 1.5$) and lower mass galaxies ($M_\star > 10^8 M_\odot$). Our aim is that by using machine learning (ML) rather than visual identification, we can avoid biases of identifying only more obvious mergers. Additionally, CNNs are not restricted to any redshift or mass range, and could potentially be able to identify even less visible merger features (such as those in minor mergers or high-z galaxies). We aim to use interpretive tools, such as examining important regions of the image with Gradient-weighted Class Activation Mapping [Grad-CAM; 28] and the latent space with UMAPs [19], to better understand how to identify high-z mergers and why our network made its decisions.

2 Data

Cosmological simulations have proven to be useful tools in training ML algorithms for merger identification since there is a ground truth that is separate from any other merger identification tools, including by-eye classification. This work uses the IllustrisTNG cosmological magnetohydrodynamical simulation suite [24, 21]. We use the smallest TNG50 box with 50 comoving Mpc per side. Its $\sim 0.1\,\mathrm{kpc}$ spatial resolution and $\sim 8 \times 10^4 M_\odot$ baryonic mass resolution provide a diverse set of galaxy morphologies, stellar masses, and details that may be important for distinguishing mergers from nonmergers such as star-forming clumps. We utilize the definition of a merger from [25] and apply a minimum mass cutoff: any subhalo (galaxy) at least 1000 times the baryonic mass resolution with two direct progenitors in the previous time snapshot is classified as a merger. We use two full snapshots (which include full physics outputs necessary to run radiative transfer): $z = 1$ and $z = 1.5$. We apply a 500Myr time window centered around those snapshots, and anything that merges at any point in that window is considered a merger. All images are taken at the two central snapshots, which means that our sample includes early-stage mergers that merge later in the window but have not yet merged at the central snapshot, and late-stage mergers that merge early in the window and are near coalescence at the central snapshot. For each merging galaxy found, we find a corresponding, mass-matched non-merging galaxy in the same snapshot.

To create our mock images, we first use the radiative transfer code SKIRT that includes dust and AGN [version 9; 2, 1, 8, 9] (for details see [35] and [29, 30]). Each extracted galaxy is observed from

Accuracy	Purity	Completeness	Brier Score	ECE	AUC
$73.0 \pm 0.4\%$	$74.0 \pm 0.01\%$	$72.0 \pm 0.01\%$	0.19 ± 0.01	0.08 ± 0.03	0.8 ± 0.01

Table 1: Mean and standard deviation of Accuracy, Purity, Completeness, Brier Score, ECE, and AUC for our model's three random seeds.

Accuracy					
All Mergers	Major	Minor	Early Stage	Late Stage	Nonmergers
$71.9 \pm 1.0\%$	$75.8 \pm 0.8\%$	$68.3 \pm 0.7\%$	$79.6 \pm 0.7\%$	$66.0 \pm 0.01\%$	$74.0 \pm 0.01\%$

Table 2: Mean and standard deviation of our model's accuracy in three random seeds broken down by different subsets of galaxies.

six viewpoints. We then filter the wavelengths down to those of the *HST* CANDELS F814W, F160W, and F606W filters [15] for our three-channel input. Additionally, we rebin to the camera's pixel scale and convolve with the PSF from the Tiny Tim software [16] in that filter (following[22]). The CNN must be able to distinguish between merging galaxies and background sources, so we place our mock galaxies in realistic environments by creating cutouts from real CANDELS mosaics that are not centered on any sources. These cutouts also introduce background noise that gives a 5σ limiting magnitude of 26.5. Any galaxy that had issues producing a reliable radiatively transferred image was thrown out, but we kept the matched counterpart because it did not drastically change the balance of the dataset.

Our dataset is split into training (70%), validation (15%), and test (15%) sets. All viewpoints of any given galaxy are in the same set. The images are normalized between 0 and 1 with a log stretch. The training set includes data augmentation through rotation up to 30° and a vertical or horizontal flip, which makes the total number of images in the training set 5940 mergers and 5916 nonmergers, the validation set 630 mergers and 624 nonmergers, and the test set 630 mergers and 630 nonmergers. Example images are shown in the bottom right of Figure 1.

3 Methods

We use the ResNet18 architecture [14], with pre-trained weights from Zoobot2.0.2 [36] in PyTorch. We set the initial learning rate 10^{-5}, and employ an exponential learning rate decay of 0.5 and cross-entropy loss with the Adam optimizer. This is a binary classification (merger or nonmerger), so after convolutional layers, we change the head to have 2 output nodes. Early stopping is triggered at epoch 48 when the validation set loss does not improve by at least 0.0005 for 5 epochs. The epoch with the lowest validation loss (epoch 43) is used to save the best model[1].

We examine the performance of our model using standard metrics such as accuracy, completeness, and purity. To further examine the behavior of our trained model, we utilize GradCAMs [28], which use the gradients heading into the final convolutional layer of a network to identify the key pixels for a given class. UMAP [19] is a dimensionality reduction technique, which we use to examine the high-dimensional latent space of our model (penultimate layer). By seeing how close different galaxies in our test set appear on a UMAP, we can determine what physical processes the CNN may have recognized.

4 Results

We train three models with different random seed initializations to ensure model stability. The mean accuracy, completeness, and purity of all three random seeds are shown in Table 1. All plots are from Seed 626, as it has both the highest accuracy and completeness. All seeds in our model classifying galaxies with $M_\star > 10^8 M_\odot$ at $1 < z < 1.5$ have accuracy of $\sim 73\%$, similar to models reported for galaxies at $M_\star > 10^9$ at $0.1 < z < 1$ in[18]. We examine Grad-CAMs, UMAPs, and the effect of observation angle to dig into what may have been causing difficulties in classifying galaxies.

The Grad-CAMs (Figure 1, bottom left) show that the network focuses on the galaxy when it makes its decision. The central galaxy is highlighted when activating the predicted class, and the edges are highlighted for the non-predicted class. This is true for mergers and nonmergers. The Grad-CAMs are

[1]Data and code are available at `https://github.com/alschechter/NeurIPSCosmicNoonMergerID`.

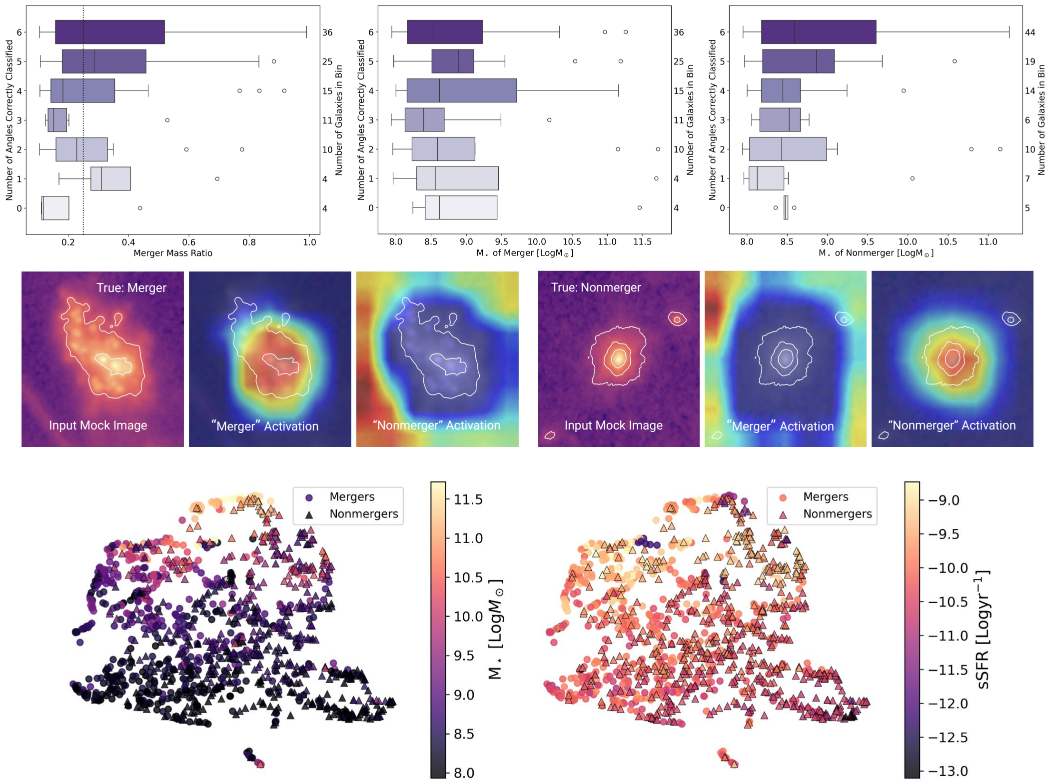

Figure 1: *Top row:* Box plots showing the 25th-75th percentile (inter-quartile range, IQR) of the data inside the box, with the whiskers stretching to the furthest point within 1.5 times the IQR. The remaining points are outside of that range. These show the overall distributions of galaxies by merger mass ratio (*left*; dotted line separates major and minor mergers) and stellar masses (*center and right*), which are identified correctly from a given number of observing angles. *Middle row*: Grad-CAM showing that the network focuses on the central galaxy when making its prediction. *Bottom row*: UMAPs showing the true class by shape (nonmergers triangle, mergers circle) and colored by stellar mass (*left*) and sSFR (*right*), respectively, which show a clear gradient.

not different enough between classes to draw conclusions about specific image features, but do offer confidence that the network has learned not to focus on background noise or sources, an important quality of successful merger identification using CNNs [7].

We examine UMAPs of the test set images in terms of multiple physical quantities of the galaxies. The mergers (circles) primarily live on the bottom side of the UMAP, and nonmergers (triangles) on the top. However, there is a lot of overlap in the middle. We see a clear gradient in the UMAP when colored by the stellar mass of each galaxy (Figure 1, bottom center). The low-mass galaxies are on the bottom of the UMAP, with stellar mass increasing towards the top. No stellar mass information was provided during training, but the network was able to recognize this physically meaningful quantity. There is again a clear gradient in the UMAP when colored by specific star formation rate (sSFR; Figure 1, bottom right). The less star-forming galaxies are on the bottom right with sSFR increasing towards the upper left, even though no sSFR information was input to the model. Because of this gradient, we speculate some of the nonmergers misclassified as mergers may be due to high, clumpy star formation, likely in the nonmergers seen in the top left. There were no obvious trends for UMAPs relative to the merger stage or merger mass ratio.

There may be a ceiling of around 85% accuracy for merger identification due to some morphological disturbances not being visible from every observation angle, and some clumpy nonmergers being indistinguishable from minor mergers [6]. Each galaxy in our test set is observed from six different angles. We examine the number of viewing angles from which a given galaxy is correctly identified, as a function of merger mass ratio and stellar mass, using box plots. On the top left panel of Figure 1, we can see that almost all mergers were identified correctly from at least one angle. Major mergers ($\mu \geq 1/4$) can cause large morphological disruptions, and thus we expect them to be easier to identify than minor mergers ($1/10 < \mu < 1/4$). Overall, as the mass ratio increases, the merger is identified correctly more often. However, we note that not all major mergers are correctly identified from more than three viewpoints. The misclassification of major mergers could be due to a major merger between lower-mass galaxies, and thus it is harder to identify than a merger between high-mass galaxies. Alternatively, if one of the galaxies is large and spheroidal, it could be blocking the companion galaxy, making it invisible from some angles. Finally, a late-stage merger can be tricky to identify even among major mergers, but CNNs have been proven to be capable of it [5, 13]. We also note that though minor mergers can be difficult to reliably identify, the majority of minor mergers were correctly identified at four or more angles (out of six possible), proving that CNNs can be a path forward in fully understanding the role of all mergers in galaxy evolution.

The lack of a trend in the stellar mass plots for both mergers and nonmergers (Figure 1 top center and right) is promising: with the right training data, CNNs enable the identification of less obvious, minor mergers among galaxies with stellar mass $M_\star < 10^9 M_\odot$. Important to note for our analysis is that the train, validation, and test sets all include more low-mass than high-mass galaxies and more minor mergers than major mergers. This reflects the hierarchical structure that is expected in any 50 Mpc^3 box of either simulated or real data: far more low-mass galaxies than high-mass galaxies. We speculate that when the network incorrectly classifies a high-mass, major merger, it is because it does not see as many examples of these mergers during training.

5 Conclusions and Outlook

We used a CNN trained on mock *HST* CANDELS images from the IllustrisTNG simulation at $z \sim 1$ to identify a wide range of galaxy mergers, including masses down to $M_\star = 10^8 M_\odot$ and merger mass ratios down to $\mu = 1/10$. The network has a final accuracy of $\sim 73\%$. A few of our highest mass merging galaxies were incorrectly classified, and in the future, we could potentially correct this by combining low-mass galaxies from TNG50 with a sample of high-mass galaxies from the larger simulation box size, TNG100, to create a more balanced and larger training set. UMAPs show us that the network is sensitive to the stellar mass and specific star formation rates of the galaxies. Our nonmerger sample is currently only mass-matched, and with a bigger box size, we could also find SFR-matched nonmergers to break the reliance on SFR. By building a training set with similar numbers of major and high-mass mergers as minor and low-mass mergers, we could potentially improve the distinction between mergers and nonmergers for all subcategories of galaxies.

Acknowledgments and Disclosure of Funding

A.L.S. and J.M.C. acknowledge support from NASA's Astrophysics Data Analysis program, grant number 80NSSC21K0646, and NSF AST-1847938. XS acknowledges the support from the NASA theory grant JWST-AR-04814. The work of A.S. was supported by the National Science Foundation MPS-Ascend Postdoctoral Research Fellowship under grant No. 2213288. L.B. acknowledges support from the NASA Astrophysics Theory program, grant 80NSSC22K0808, and NSF AAG 2307171. A.Ć: This work was produced by FermiForward Discovery Group, LLC under Contract No. 89243024CSC000002 with the U.S. Department of Energy, Office of Science, Office of High Energy Physics. Publisher acknowledges the U.S. Government license to provide public access under the (DOE Public Access Plan).

We acknowledge the Deep Skies Lab as a community of multi-domain experts and collaborators who've facilitated an environment of open discussion, idea generation, and collaboration. This community was important for the development of this project.

Author Contributions: The following authors contributed in different ways to the manuscript. Schechter: Writing manuscript, all ML code and analysis. Ćiprijanović: editing manuscript, mentoring, and overseeing ML code and analysis. Shen: Radiative transfer, writing section 2. Nevin: Editing manuscript, mentoring, and overseeing ML code and analysis. Comerford: Editing manuscript, mentoring, and overall direction of paper. Stemo: Assistance with mock image creation. Blecha: Initial ideas and direction of paper.

Glossary of Symbols: z redshift; $M_\odot$ solar mass = 2×10^{30}kg; $M_\star$ stellar mass; μ stellar mass ratio $M_{\star,1}/M_{\star,2}$; pc parsec = 3×10^{16}m

References

[1] M. Baes and P. Camps. SKIRT: The design of a suite of input models for Monte Carlo radiative transfer simulations. *Astronomy and Computing*, 12:33–44, September 2015. ADS Bibcode: 2015A&C....12...33B.

[2] Maarten Baes, Joris Verstappen, Ilse De Looze, Jacopo Fritz, Waad Saftly, Edgardo Vidal Pérez, Marko Stalevski, and Sander Valcke. EFFICIENT THREE-DIMENSIONAL NLTE DUST RADIATIVE TRANSFER WITH SKIRT. *The Astrophysical Journal Supplement Series*, 196(2):22, October 2011.

[3] Joshua E. Barnes and Lars E. Hernquist. Fueling Starburst Galaxies with Gas-rich Mergers. *The Astrophysical Journal*, 370:L65, April 1991. Publisher: IOP ADS Bibcode: 1991ApJ...370L..65B.

[4] R. Scott Barrows, Julia M. Comerford, Daniel Stern, and Roberto J. Assef. A Census of WISE-selected Dual and Offset AGNs Across the Sky: New Constraints on Merger-driven Triggering of Obscured AGNs. *The Astrophysical Journal*, 951:92, July 2023. ADS Bibcode: 2023ApJ...951...92B.

[5] Robert W. Bickley, Connor Bottrell, Maan H. Hani, Sara L. Ellison, Hossen Teimoorinia, Kwang Moo Yi, Scott Wilkinson, Stephen Gwyn, and Michael J. Hudson. Convolutional neural network identification of galaxy post-mergers in UNIONS using IllustrisTNG. *Monthly Notices of the Royal Astronomical Society*, 504:372–392, June 2021. ADS Bibcode: 2021MNRAS.504..372B.

[6] Robert W. Bickley, Scott Wilkinson, Leonardo Ferreira, Sara L. Ellison, Connor Bottrell, and Debarpita Jyoti. The effect of image quality on galaxy merger identification with deep learning. *Monthly Notices of the Royal Astronomical Society*, 534:2533–2550, November 2024. Publisher: OUP ADS Bibcode: 2024MNRAS.534.2533B.

[7] Connor Bottrell, Maan H. Hani, Hossen Teimoorinia, Sara L. Ellison, Jorge Moreno, Paul Torrey, Christopher C. Hayward, Mallory Thorp, Luc Simard, and Lars Hernquist. Deep learning predictions of galaxy merger stage and the importance of observational realism. *Monthly Notices of the Royal Astronomical Society*, 490:5390–5413, December 2019. Publisher: OUP ADS Bibcode: 2019MNRAS.490.5390B.

[8] Peter Camps and Maarten Baes. SKIRT: an Advanced Dust Radiative Transfer Code with a User-Friendly Architecture. *Astronomy and Computing*, 9:20–33, March 2015. arXiv:1410.1629 [astro-ph].

[9] Peter Camps and Maarten Baes. SKIRT 9: redesigning an advanced dust radiative transfer code to allow kinematics, line transfer and polarization by aligned dust grains, March 2020. arXiv:2003.00721 [astro-ph].

[10] Christopher J. Conselice. The Relationship between Stellar Light Distributions of Galaxies and Their Formation Histories. *The Astrophysical Journal Supplement Series*, 147:1–28, July 2003. Publisher: IOP ADS Bibcode: 2003ApJS..147....1C.

[11] Tiziana Di Matteo, Volker Springel, and Lars Hernquist. Energy input from quasars regulates the growth and activity of black holes and their host galaxies. *Nature*, 433:604–607, February 2005. ADS Bibcode: 2005Natur.433..604D.

[12] Sara L. Ellison, David R. Patton, Luc Simard, and Alan W. McConnachie. GALAXY PAIRS IN THE SLOAN DIGITAL SKY SURVEY. I. STAR FORMATION, ACTIVE GALACTIC NUCLEUS FRACTION, AND THE LUMINOSITY/MASS-METALLICITY RELATION. *The Astronomical Journal*, 135(5):1877–1899, May 2008.

[13] Leonardo Ferreira, Sara L. Ellison, David R. Patton, Shoshannah Byrne-Mamahit, Scott Wilkinson, Robert Bickley, Christopher J. Conselice, and Connor Bottrell. Galaxy evolution in the post-merger regime I – Most merger-induced in-situ stellar mass growth happens post-coalescence, October 2024. arXiv:2410.06356.

[14] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep Residual Learning for Image Recognition, December 2015. arXiv:1512.03385 [cs].

[15] Anton M. Koekemoer, S. M. Faber, Henry C. Ferguson, Norman A. Grogin, Dale D. Kocevski, David C. Koo, Kamson Lai, Jennifer M. Lotz, Ray A. Lucas, Elizabeth J. McGrath, Sara Ogaz, Abhijith Rajan, Adam G. Riess, Steve A. Rodney, Louis Strolger, Stefano Casertano, Marco Castellano, Tomas Dahlen, Mark Dickinson, Timothy Dolch, Adriano Fontana, Mauro Giavalisco, Andrea Grazian, Yicheng Guo, Nimish P. Hathi, Kuang-Han Huang, Arjen van der Wel, Hao-Jing Yan, Viviana Acquaviva, David M. Alexander, Omar Almaini, Matthew L. N. Ashby, Marco Barden, Eric F. Bell, Frédéric Bournaud, Thomas M. Brown, Karina I. Caputi, Paolo Cassata, Peter J. Challis, Ranga-Ram Chary, Edmond Cheung, Michele Cirasuolo, Christopher J. Conselice, Asantha Roshan Cooray, Darren J. Croton, Emanuele Daddi, Romeel Davé, Duilia F. de Mello, Loic de Ravel, Avishai Dekel, Jennifer L. Donley, James S. Dunlop, Aaron A. Dutton, David Elbaz, Giovanni G. Fazio, Alexei V. Filippenko, Steven L. Finkelstein, Chris Frazer, Jonathan P. Gardner, Peter M. Garnavich, Eric Gawiser, Ruth Gruetzbauch, Will G. Hartley, Boris Häussler, Jessica Herrington, Philip F. Hopkins, Jia-Sheng Huang, Saurabh W. Jha, Andrew Johnson, Jeyhan S. Kartaltepe, Ali A. Khostovan, Robert P. Kirshner, Caterina Lani, Kyoung-Soo Lee, Weidong Li, Piero Madau, Patrick J. McCarthy, Daniel H. McIntosh, Ross J. McLure, Conor McPartland, Bahram Mobasher, Heidi Moreira, Alice Mortlock, Leonidas A. Moustakas, Mark Mozena, Kirpal Nandra, Jeffrey A. Newman, Jennifer L. Nielsen, Sami Niemi, Kai G. Noeske, Casey J. Papovich, Laura Pentericci, Alexandra Pope, Joel R. Primack, Swara Ravindranath, Naveen A. Reddy, Alvio Renzini, Hans-Walter Rix, Aday R. Robaina, David J. Rosario, Piero Rosati, Sara Salimbeni, Claudia Scarlata, Brian Siana, Luc Simard, Joseph Smidt, Diana Snyder, Rachel S. Somerville, Hyron Spinrad, Amber N. Straughn, Olivia Telford, Harry I. Teplitz, Jonathan R. Trump, Carlos Vargas, Carolin Villforth, Cory R. Wagner, Pat Wandro, Risa H. Wechsler, Benjamin J. Weiner, Tommy Wiklind, Vivienne Wild, Grant Wilson, Stijn Wuyts, and Min S. Yun. CANDELS: The Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey—The Hubble Space Telescope Observations, Imaging Data Products, and Mosaics. *The Astrophysical Journal Supplement Series*, 197:36, December 2011. Publisher: IOP ADS Bibcode: 2011ApJS..197...36K.

[16] John E. Krist, Richard N. Hook, and Felix Stoehr. 20 years of Hubble Space Telescope optical modeling using Tiny Tim. In *Optical Modeling and Performance Predictions V*, volume 8127, pages 166–181. SPIE, September 2011.

[17] Jennifer M. Lotz, Joel Primack, and Piero Madau. A New Nonparametric Approach to Galaxy Morphological Classification. *The Astronomical Journal*, 128:163–182, July 2004. Publisher: IOP ADS Bibcode: 2004AJ....128..163L.

[18] B. Margalef-Bentabol, L. Wang, A. La Marca, C. Blanco-Prieto, D. Chudy, H. Domínguez-Sánchez, A. D. Goulding, A. Guzmán-Ortega, M. Huertas-Company, G. Martin, W. J. Pearson, V. Rodriguez-Gomez, M. Walmsley, R. W. Bickley, C. Bottrell, C. Conselice, and D. O'Ryan. Galaxy merger challenge: A comparison study between machine learning-based detection methods. *Astronomy & Astrophysics*, 687:A24, July 2024. arXiv:2403.15118 [astro-ph].

[19] Leland McInnes, John Healy, and James Melville. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction, September 2020. arXiv:1802.03426 [stat].

[20] J. Christopher Mihos and Lars Hernquist. Gasdynamics and Starbursts in Major Mergers. *The Astrophysical Journal*, 464:641, June 1996. Publisher: IOP ADS Bibcode: 1996ApJ...464..641M.

[21] Dylan Nelson, Annalisa Pillepich, Volker Springel, Rüdiger Pakmor, Rainer Weinberger, Shy Genel, Paul Torrey, Mark Vogelsberger, Federico Marinacci, and Lars Hernquist. First results from the TNG50 simulation: galactic outflows driven by supernovae and black hole feedback. *Monthly Notices of the Royal Astronomical Society*, 490:3234–3261, December 2019. ADS Bibcode: 2019MNRAS.490.3234N.

[22] R. Nevin, L. Blecha, J. Comerford, and J. Greene. Accurate Identification of Galaxy Mergers with Imaging. *The Astrophysical Journal*, 872:76, February 2019. ADS Bibcode: 2019ApJ...872...76N.

[23] David R. Patton, Sara L. Ellison, Luc Simard, Alan W. McConnachie, and J. Trevor Mendel. Galaxy pairs in the Sloan Digital Sky Survey - III. Evidence of induced star formation from optical colours. *Monthly Notices of the Royal Astronomical Society*, 412:591–606, March 2011. ADS Bibcode: 2011MNRAS.412..591P.

[24] Annalisa Pillepich, Dylan Nelson, Volker Springel, Rüdiger Pakmor, Paul Torrey, Rainer Weinberger, Mark Vogelsberger, Federico Marinacci, Shy Genel, Arjen van der Wel, and Lars Hernquist. First results from the TNG50 simulation: the evolution of stellar and gaseous discs across cosmic time. *Monthly Notices of the Royal Astronomical Society*, 490:3196–3233, December 2019. ADS Bibcode: 2019MNRAS.490.3196P.

[25] Vicente Rodriguez-Gomez, Shy Genel, Mark Vogelsberger, Debora Sijacki, Annalisa Pillepich, Laura V. Sales, Paul Torrey, Greg Snyder, Dylan Nelson, Volker Springel, Chung-Pei Ma, and Lars Hernquist. The merger rate of galaxies in the Illustris simulation: a comparison with observations and semi-empirical models. *Monthly Notices of the Royal Astronomical Society*, 449:49–64, May 2015. ADS Bibcode: 2015MNRAS.449...49R.

[26] Caitlin Rose, Jeyhan S. Kartaltepe, Gregory F. Snyder, Marc Huertas-Company, L. Y. Aaron Yung, Pablo Arrabal Haro, Micaela B. Bagley, Laura Bisigello, Antonello Calabrò, Nikko J. Cleri, Mark Dickinson, Henry C. Ferguson, Steven L. Finkelstein, Adriano Fontana, Andrea Grazian, Norman A. Grogin, Benne W. Holwerda, Kartheik G. Iyer, Lisa J. Kewley, Allison Kirkpatrick, Dale D. Kocevski, Anton M. Koekemoer, Jennifer M. Lotz, Ray A. Lucas, Lorenzo Napolitano, Casey Papovich, Laura Pentericci, Pablo G. Pérez-González, Nor Pirzkal, Swara Ravindranath, Rachel S. Somerville, Amber N. Straughn, Jonathan R. Trump, Stephen M. Wilkins, and Guang Yang. CEERS Key Paper. IX. Identifying Galaxy Mergers in CEERS NIRCam Images Using Random Forests and Convolutional Neural Networks. *The Astrophysical Journal*, 976:L8, November 2024. Publisher: IOP ADS Bibcode: 2024ApJ...976L...8R.

[27] Caitlin Rose, Jeyhan S. Kartaltepe, Gregory F. Snyder, Vicente Rodriguez-Gomez, L. Y. Aaron Yung, Pablo Arrabal Haro, Micaela B. Bagley, Antonello Calabró, Nikko J. Cleri, M. C. Cooper, Luca Costantin, Darren Croton, Mark Dickinson, Steven L. Finkelstein, Boris Häußler, Benne W. Holwerda, Anton M. Koekemoer, Peter Kurczynski, Ray A. Lucas, Kameswara Bharadwaj Mantha, Casey Papovich, Pablo G. Pérez-González, Nor Pirzkal, Rachel S. Somerville, Amber N. Straughn, and Sandro Tacchella. Identifying Galaxy Mergers in Simulated CEERS NIRCam Images Using Random Forests. *The Astrophysical Journal*, 942:54, January 2023. Publisher: IOP ADS Bibcode: 2023ApJ...942...54R.

[28] Ramprasaath R. Selvaraju, Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, and Dhruv Batra. Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization. *International Journal of Computer Vision*, 128(2):336–359, February 2020. arXiv:1610.02391 [cs].

[29] Xuejian Shen, Mark Vogelsberger, Dylan Nelson, Annalisa Pillepich, Sandro Tacchella, Federico Marinacci, Paul Torrey, Lars Hernquist, and Volker Springel. High-redshift JWST predictions from IllustrisTNG: II. Galaxy line and continuum spectral indices and dust attenuation curves. *Monthly Notices of the Royal Astronomical Society*, 495:4747–4768, July 2020. ADS Bibcode: 2020MNRAS.495.4747S.

[30] Xuejian Shen, Mark Vogelsberger, Dylan Nelson, Sandro Tacchella, Lars Hernquist, Volker Springel, Federico Marinacci, and Paul Torrey. High-redshift predictions from IllustrisTNG - III. Infrared luminosity functions, obscured star formation, and dust temperature of high-redshift galaxies. *Monthly Notices of the Royal Astronomical Society*, 510:5560–5578, March 2022. Publisher: OUP ADS Bibcode: 2022MNRAS.510.5560S.

[31] Gregory F. Snyder, Vicente Rodriguez-Gomez, Jennifer M. Lotz, Paul Torrey, Amanda C. N. Quirk, Lars Hernquist, Mark Vogelsberger, and Peter E. Freeman. Automated Distant Galaxy Merger Classifications from Space Telescope Images using the Illustris Simulation. *Monthly Notices of the Royal Astronomical Society*, 486(3):3702–3720, July 2019. arXiv:1809.02136 [astro-ph].

[32] Volker Springel, Tiziana Di Matteo, and Lars Hernquist. Black Holes in Galaxy Mergers: The Formation of Red Elliptical Galaxies. *The Astrophysical Journal*, 620:L79–L82, February 2005. ADS Bibcode: 2005ApJ...620L..79S.

[33] Volker Springel, Tiziana Di Matteo, and Lars Hernquist. Modelling feedback from stars and black holes in galaxy mergers. *Monthly Notices of the Royal Astronomical Society*, 361:776–794, August 2005. Publisher: OUP ADS Bibcode: 2005MNRAS.361..776S.

[34] Volker Springel and Lars Hernquist. Formation of a Spiral Galaxy in a Major Merger. *The Astrophysical Journal*, 622:L9–L12, March 2005. ADS Bibcode: 2005ApJ...622L...9S.

[35] Mark Vogelsberger, Dylan Nelson, Annalisa Pillepich, Xuejian Shen, Federico Marinacci, Volker Springel, Rüdiger Pakmor, Sandro Tacchella, Rainer Weinberger, Paul Torrey, and Lars Hernquist. High-redshift JWST predictions from IllustrisTNG: dust modelling and galaxy luminosity functions. *Monthly Notices of the Royal Astronomical Society*, 492:5167–5201, March 2020. ADS Bibcode: 2020MNRAS.492.5167V.

[36] Mike Walmsley, Campbell Allen, Ben Aussel, Micah Bowles, Kasia Gregorowicz, Inigo Val Slijepcevic, Chris J. Lintott, Anna M. m Scaife, Maja Jabłońska, Kosio Karchev, Denise Lanzieri, Devina Mohan, David O'Ryan, Bharath Saiguhan, Crisel Suárez, Nicolás Guerra-Varas, and Renuka Velu. Zoobot: Adaptable Deep Learning Models for Galaxy Morphology. *Journal of Open Source Software*, 8(85):5312, May 2023.

[37] Scott Wilkinson, Sara L. Ellison, Connor Bottrell, Robert W. Bickley, Shoshannah Byrne-Mamahit, Leonardo Ferreira, and David R. Patton. The limitations (and potential) of non-parametric morphology statistics for post-merger identification. *Monthly Notices of the Royal Astronomical Society*, 528:5558–5585, March 2024. Publisher: OUP ADS Bibcode: 2024MNRAS.528.5558W.

[38] A. Ćiprijanović, A. Lewis, K. Pedro, S. Madireddy, B. Nord, G. N. Perdue, and S. M. Wild. DeepAstroUDA: semi-supervised universal domain adaptation for cross-survey galaxy morphology classification and anomaly detection. *Machine Learning: Science and Technology*, 4:025013, June 2023. Publisher: IOP ADS Bibcode: 2023MLS&T...4b5013C.

[39] A. Ćiprijanović, G. F. Snyder, B. Nord, and J. E. G. Peek. DeepMerge: Classifying high-redshift merging galaxies with deep neural networks. *Astronomy and Computing*, 32:100390, July 2020. ADS Bibcode: 2020A&C....3200390C.