
GaMPEN: An ML Framework for Estimating Galaxy Morphological Parameters and Quantifying Uncertainty

Aritra Ghosh^{1,2} C. Megan Urry^{3,2} Amrit Rau⁴ Laurence Perreault-Levasseur^{5,6,7} Miles Cranmer⁸
Kevin Schawinski⁹ Dominic Stark⁹ Chuan Tian^{3,2} Ryan Ofman¹

Abstract

We introduce a novel machine learning framework for estimating the Bayesian posteriors of morphological parameters for arbitrarily large numbers of galaxies. The Galaxy Morphology Posterior Estimation Network (GaMPEN) estimates values and uncertainties for a galaxy’s bulge-to-total light ratio (L_B/L_T), effective radius (R_e), and flux (F). GaMPEN also uses a Spatial Transformer Network (STN) to automatically crop input galaxy frames to an optimal size before determining their morphology. Training and testing GaMPEN on galaxies simulated to match $z < 0.75$ galaxies in Hyper Suprime-Cam Wide images, we demonstrate that GaMPEN can accurately quantify uncertainties and estimate parameters. GaMPEN is the first machine learning framework for determining posterior distributions of multiple morphological parameters and is also the first application of an STN to optical imaging in astronomy.

1. Introduction

Galaxy morphology has been shown to be related to many fundamental properties of the galaxy and its environment, including galaxy mass, star formation rate, stellar kinematics, merger history, cosmic environment, the influence of supermassive black holes (e.g., Bender et al., 1992; Tremaine

et al., 2002; Pozzetti et al., 2010; Schawinski et al., 2014; Huertas-Company et al., 2016; Powell et al., 2017; Dimauro et al., 2022). Studying the morphology of large samples of galaxies at different redshifts is crucial in order to understand the physics of galaxy formation and evolution.

Over the last decade, Convolutional Neural Networks (CNNs) have become increasingly popular for determining galaxy morphology. From early attempts at using a CNN to classify galaxies morphologically (e.g., Dieleman et al., 2015; Huertas-Company et al., 2015) to the largest CNN produced morphology catalogs currently available (Cheng et al., 2021; Vega-Ferrero et al., 2021), most CNNs have provided broad, qualitative classifications rather than numerical estimates of morphological parameters. By contrast, Tuccillo et al. (2018) used a CNN to estimate the parameters of a single-component Sérsic fit, though without uncertainties. Meanwhile, the computation of full Bayesian posteriors for different morphological parameters is crucial for drawing scientific inferences that account for uncertainty and thus are indispensable in the derivation of robust scaling relations or tests of theoretical models using morphology.

In this contribution, we introduce GaMPEN (the Galaxy Morphology Posterior Estimation Network), a novel machine learning framework that estimates the Bayesian posteriors for three morphological parameters: the bulge-to-total light ratio (L_B/L_T), the effective radius (R_e), and the total flux (F). GaMPEN can also automatically crop the input image frames to an optimal size before morphology determination. This helps GaMPEN focus on the galaxy of interest at the center of each cutout – while cropping out most secondary galaxies in the input frame. This allows us to apply GaMPEN to a wide range of redshifts without having to worry about optimal cutout sizes – not a trivial task when applying a machine learning (ML) algorithm to a new survey without predetermined R_e measurements.

To have a robust understanding of the performance, bias, and limitations of GaMPEN, we train and test GaMPEN on simulations of galaxy images—the only situation where we have access to the “ground truth” morphological parameters of the galaxies. We match our simulations to the observations of the Hyper Suprime-Cam (HSC) Wide survey

¹Department of Astronomy, Yale University, New Haven, CT, USA ²Yale Center for Astronomy and Astrophysics, New Haven, CT, USA ³Department of Physics, Yale University, New Haven, CT, USA ⁴Department of Computer Science, Yale University, New Haven, CT, USA ⁵Department of Physics, Université de Montréal, Montréal, Canada ⁶Mila - Quebec Artificial Intelligence Institute, Montréal, Canada ⁷Center for Computational Astrophysics, Flatiron Institute, New York, NY, USA ⁸Department of Astrophysical Sciences, Princeton University, Princeton, NJ, USA ⁹Modulos AG, Technoparkstr. 1, CH-8005, Zurich, Switzerland. Correspondence to: Aritra Ghosh <aritra.ghosh@yale.edu; aritraghsh09@gmail.com>.

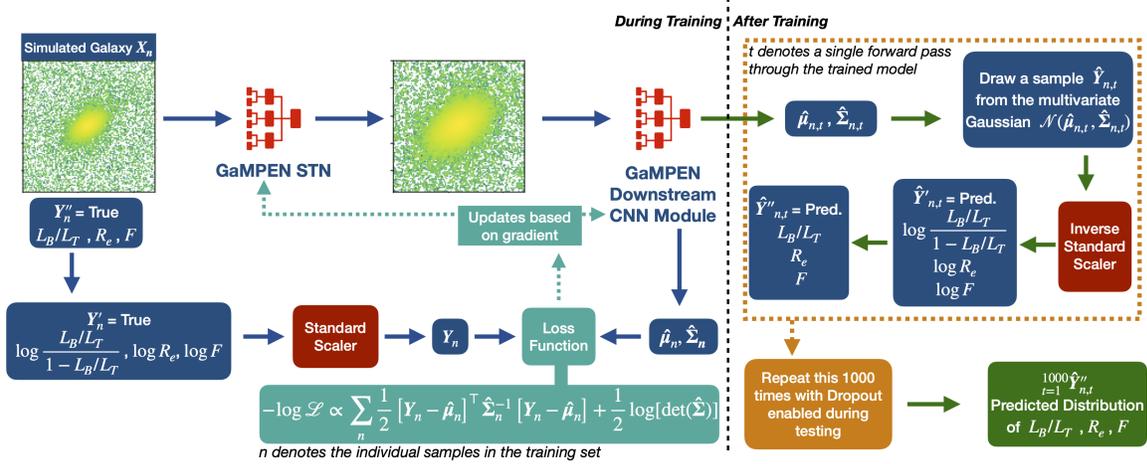


Figure 1. Diagram outlining the training (left) and inference (right) phases of the GaMPEN workflow.

(Aihara et al., 2018), as this is an obvious application.

2. Description of the Framework

The architecture of GaMPEN consists of an upstream Spatial Transformer Network (STN) followed by a downstream CNN, as shown in Figure 5 (Appendix A). STNs, introduced by Jaderberg et al. (2015), are learnable modules that can be inserted in CNNs and explicitly allow for the spatial manipulation of data within the CNN. In GaMPEN, the STN applies a two-dimensional affine cropping transformation to the input image, and the transformed image is then passed to the CNN. Since the STN can be trained with standard back-propagation, the entire GaMPEN framework can be trained end-to-end without any separate supervision required for the STN. The input image, once transformed by the STN, is passed to the downstream CNN module, as depicted in Figure 1. This downstream module predicts the posterior distribution of the bulge-to-total light ratio, effective radius, and total flux for each input galaxy.

2.1. Prediction of Posteriors

Two primary sources of error contribute to the uncertainties in the parameters predicted by GaMPEN. The first arises from errors inherent to the input imaging data (e.g., noise and PSF blurring), and this is commonly referred to as aleatoric uncertainty. The second error comes from the limitations of the model being used for prediction (e.g., the number of free parameters in GaMPEN, the amount of training data, etc.); this is referred to as epistemic uncertainty. GaMPEN accounts for both these kinds of uncertainties in its predictions.

In order to obtain epistemic uncertainties, we treat the trained model itself as a random variable – because, intuitively, there are many possible models that could be trained from the same training data, \mathcal{D} . To predict the posterior, we

need to marginalize over these possible models; and for that marginalization, we use the Monte-Carlo Dropout (MCD) technique as introduced by Gal & Ghahramani (2016). During inference, we feed every test image to the trained GaMPEN framework 1000 times and collect the outputs. Each forward pass through GaMPEN samples the approximate parameter posterior.

To account for aleatoric uncertainties, GaMPEN predicts the parameters of a multivariate Gaussian distribution $\mathcal{N}(\mu, \Sigma)$ for every input image. μ and Σ are the mean and covariance matrix of the multivariate Gaussian distribution, respectively. Although we would like to use GaMPEN to predict aleatoric uncertainties, the covariance matrix, Σ , is not known *a priori*. Instead, we train GaMPEN to learn these values by minimizing the negative log-likelihood of the output parameters, which is shown at the bottom of Figure 1. The predicted posterior distributions for a randomly chosen galaxy in the test set is shown in Figure 2.

2.2. Training & Inference

We train GaMPEN on realistic simulations of galaxies created using GalSim (Rowe et al., 2015) to match the properties of $z < 0.75$ galaxies in the HSC Wide Survey. We convolved these simulated galaxies with a representative point spread function (PSF) and added representative noise based on real HSC-W images. In order to ensure that the values predicted by GaMPEN are always physical ($0 \leq L_B/L_T \leq 1$; $R_e > 0$; $F > 0$) and have similar ranges, we pass the training labels through the logit/log transformations followed by a standard scaler transformation as depicted in Figure 1. During the inference, the inverse of these transformations are applied to the predicted values.

One of the most critical adjustable parameters is the dropout rate – higher dropout rates generally lead networks to estimate higher epistemic uncertainties. Thus, we train GaM-

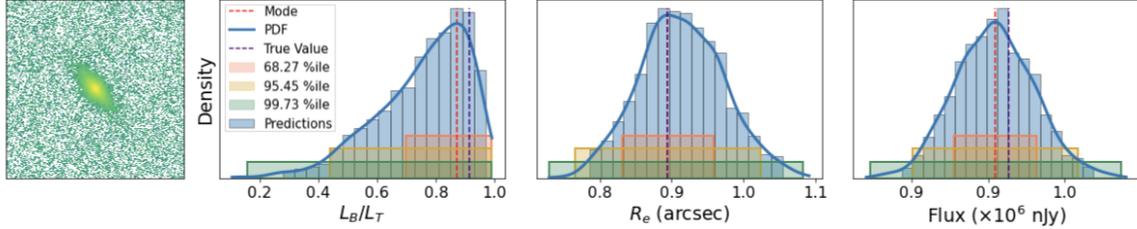


Figure 2. Examples of predicted posterior distributions for a randomly chosen simulated galaxy. The blue shaded histogram shows the predictions from GaMPEN and the blue solid lines show the associated probability distribution functions estimated by kernel density estimation. These are used to calculate the confidence intervals shown in the figure.

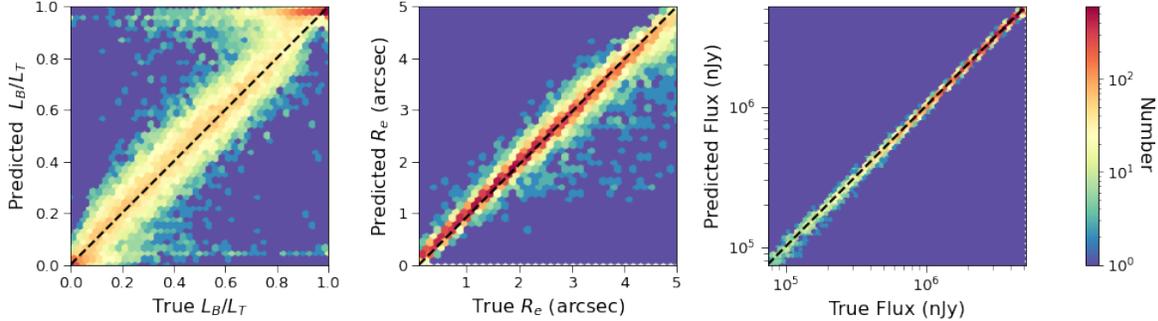


Figure 3. The true values of the galaxy parameters plotted against the most probable values predicted by GaMPEN. The black dashed line marks the $y = x$ diagonal on which perfectly recovered parameters should lie. The color of each hexagon corresponds to the number of galaxies it contains, as indicated by the colorbar on the right.

PEN with different dropout rates and then calculate their coverage probabilities, defined as the fraction of the validation set galaxies where the true value lies within a particular confidence interval. From our experiments, as shown in Figure 7 (Appendix A), a dropout rate of 7×10^{-4} yields coverage probabilities very close to their corresponding confidence levels, resulting in accurately calibrated posteriors. It is important to note that the inclusion of the full covariance matrix in the loss function (instead of just the diagonal values) allowed us to incorporate the relationships between the different output variables and achieve simultaneous calibration of the coverage probabilities for all three variables.

Using an 80-10-10 train-validation-test split, we trained GaMPEN on the simulated galaxies. The network was trained using stochastic gradient descent and its hyperparameters were tuned using the loss obtained on the validation set. To obtain the posterior distribution of the output variables, as outlined in Figure 1, we feed each input image, in the test set 1000 times into the trained GaMPEN framework with dropout enabled. During each iteration, we collect the predicted set of $\hat{\mu}_{n,t}, \hat{\Sigma}_{n,t}$ for the t^{th} forward pass. For each forward pass, we draw a sample from the multivariate normal distribution $\mathcal{N}(\hat{\mu}_{n,t}, \hat{\Sigma}_{n,t})$. The distribution generated by the collection of all 1000 forward passes represents the predicted posterior distribution.

3. Results

We first verify that the STN in GaMPEN correctly learns to transform the input images. Figures 8 and 9 (Appendix A) show examples of the transformations applied by the STN of a trained GaMPEN framework to simulated and real HSC data. As can be seen, the STN learns to apply an optimal amount of cropping for each input galaxy and focus on the galaxy of interest at the center of the cutout.

Parameter Name	68.27% Conf. Level	95.45% Conf. Level	99.73% Conf. Level
L_B/L_T	71.8%	96.9%	98.9%
R_e	68.1%	95.9%	98.3%
F	78.7%	98.2%	99.9%
Mean	72.9%	97.0%	99.0%

Table 1. Coverage Probabilities Obtained on the Test Set. In Table 1, we report the coverage probabilities that GaMPEN achieves on the test set. Clearly, GaMPEN produces well-calibrated and accurate posteriors, consistently close to the claimed confidence levels.

Figure 3 shows the most probable values (i.e., modes of the predicted distributions) predicted by GaMPEN versus the true values for the test set. Most galaxies are clustered around the line of equality, showing that the most probable values of the distributions predicted by GaMPEN closely track the true values of the parameters. Note that among the larger deviations evident in Figure 3, are predictions

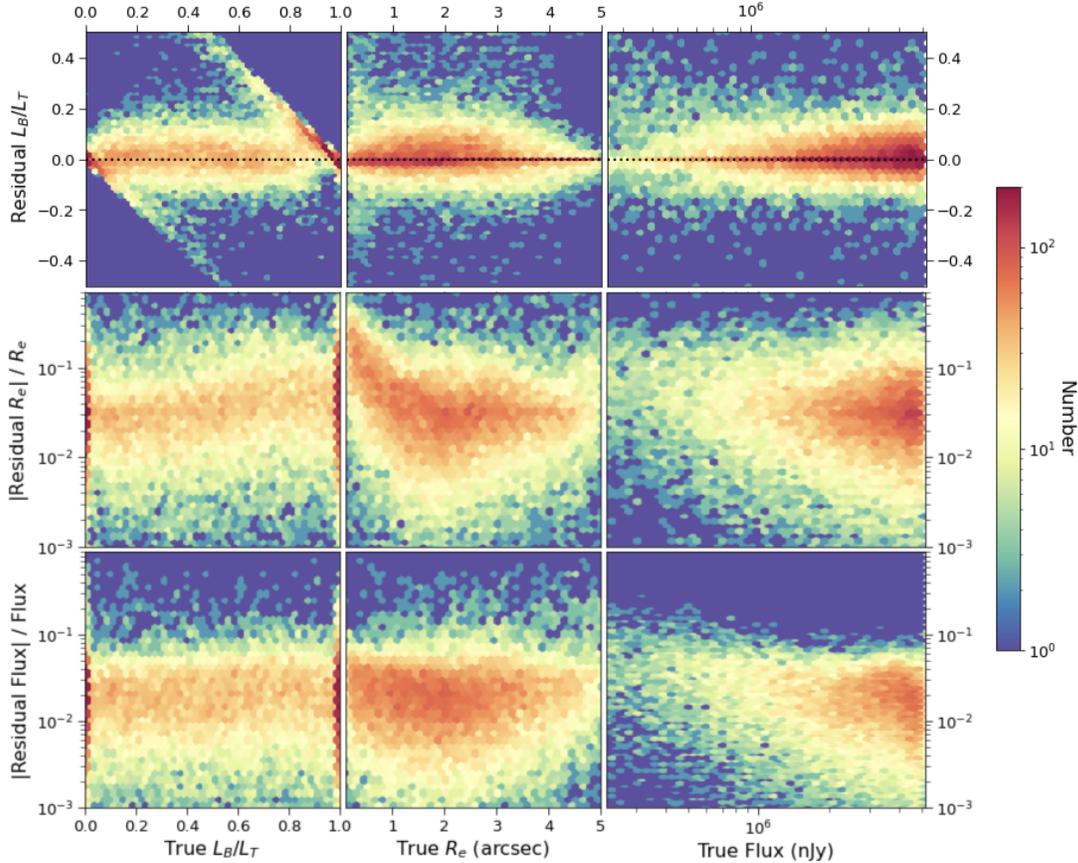


Figure 4. Residuals of GaMPEN predicted parameter values plotted against the true values. The residual for each parameter is defined as the difference between the most probable predicted value and the true value. The color of each hexagonal bin corresponds to the number of galaxies it contains, as shown by the colorbar on the right. To keep the y-axis dimensionless, for R_e and F , we plot fractional residuals.

near the limits of L_B/L_T ; we explore this further below. In Figure 6 (Appendix B), we show the residual distribution for the three parameters predicted by GaMPEN. The GaMPEN prediction of the bulge-to-total ratio is, in $\sim 68.27\%$ of cases, within 0.1 of the true value – the “typical” error. The typical error in effective radius is 0.17 arcsec. Typical uncertainties in the flux are at the 0.1 – 1% level.

Figure 4 shows the residuals for each variable when plotted against the true values of the output variables. The majority of galaxies are clustered uniformly in areas of low residuals. There are a few other notable features in this figure. In the top left panel, the L_B/L_T residuals are highest near the limits of L_B/L_T . This is not very surprising, given that it is inherently difficult to accurately determine L_B/L_T when one component strongly dominates over the other. Larger residuals in the predictions near the limits of L_B/L_T lead to the feature seen in the top left panel of the above figure. We can use a parameter transformation to mitigate this edge effect, as described in Appendix B.

The center panel shows that for small values of effective radius, $R_e < 1.0$ arcsec, there is an increase in the magnitude of the residuals. Similarly, the bottom-right two panels show

that the residuals of R_e and F are systematically higher for faint galaxies, $F < 10^6$ nJy. In other words, GaMPEN systematically becomes less accurate at predicting the radii of galaxies when their sizes become comparable to the seeing of the HSC-Wide Survey (g-band median FWHM ~ 0.85 arcsec). Similarly, GaMPEN finds it more challenging to predict the sizes and fluxes of fainter galaxies, just as one would expect.

The primary advantage of a Bayesian ML framework like GaMPEN is its ability to predict the posterior distributions instead of just point estimates. Thus, we would expect such a network to inherently produce higher uncertainties in regions of the parameter space where residuals are higher. Figure 10 (Appendix A) show that GaMPEN indeed correctly predicts broader distributions in regions with the largest residuals.

4. Conclusions and Future Direction

We have demonstrated that we can use GaMPEN to predict well-calibrated and accurate posterior distributions for morphological parameters of galaxies. The galaxies with the

largest residuals are smaller and/or fainter and/or have one morphological component completely dominating over the other — situations where morphological analysis is inherently difficult. GaMPEN correctly accounts for this by predicting correspondingly higher uncertainties (i.e., broader distributions) in all these above situations. We have also outlined how the use of an STN allows GaMPEN to crop out secondary galaxies present in the cutout and focus on galaxies at the center. We are currently applying GaMPEN to real HSC galaxies (preliminary results in Appendix C). We aim to make GaMPEN public by the Fall of 2022.

References

- Aihara, H., Armstrong, R., Bickerton, S., Bosch, J., Coupon, J., Furusawa, H., Hayashi, Y., Ikeda, H., Kamata, Y., Karoji, H., Kawanomoto, S., Koike, M., Komiyama, Y., Lang, D., Lupton, R. H., Mineo, S., Miyatake, H., Miyazaki, S., Morokuma, T., Obuchi, Y., Oishi, Y., Okura, Y., Price, P. A., Takata, T., Tanaka, M. M., Tanaka, M., Tanaka, Y., Uchida, T., Uruguchi, F., Utsumi, Y., Wang, S.-Y., Yamada, Y., Yamanoi, H., Yasuda, N., Arimoto, N., Chiba, M., Finet, F., Fujimori, H., Fujimoto, S., Furusawa, J., Goto, T., Goulding, A., Gunn, J. E., Harikane, Y., Hattori, T., Hayashi, M., Hełminiak, K. G., Higuchi, R., Hikage, C., Ho, P. T. P., Hsieh, B.-C., Huang, K., Huang, S., Imanishi, M., Iwata, I., Jaelani, A. T., Jian, H.-Y., Kashikawa, N., Katayama, N., Kojima, T., Konno, A., Koshida, S., Kusakabe, H., Leauthaud, A., Lee, C.-H., Lin, L., Lin, Y.-T., Mandelbaum, R., Matsuoka, Y., Medezinski, E., Miyama, S., Momose, R., More, A., More, S., Mukae, S., Murata, R., Murayama, H., Nagao, T., Nakata, F., Niida, M., Niikura, H., Nishizawa, A. J., Oguri, M., Okabe, N., Ono, Y., Onodera, M., Onoue, M., Ouchi, M., Pyo, T.-S., Shibuya, T., Shimasaku, K., Simet, M., Speagle, J., Spergel, D. N., Strauss, M. A., Sugahara, Y., Sugiyama, N., Suto, Y., Suzuki, N., Tait, P. J., Takada, M., Terai, T., Toba, Y., Turner, E. L., Uchiyama, H., Umetsu, K., Urata, Y., Usuda, T., Yeh, S., and Yuma, S. First data release of the Hyper Suprime-Cam Subaru Strategic Program. *Publications of the Astronomical Society of Japan*, 70(SP1), 1 2018. ISSN 0004-6264. doi: 10.1093/pasj/psx081. URL <https://academic.oup.com/pasj/article/doi/10.1093/pasj/psx081/4494171>.
- Bender, R., Burstein, D., and Faber, S. M. Dynamically hot galaxies. I - Structural properties. *The Astrophysical Journal*, 399:462, 11 1992. ISSN 0004-637X. doi: 10.1086/171940. URL <http://adsabs.harvard.edu/doi/10.1086/171940>.
- Cheng, T.-Y., Conselice, C. J., Aragón-Salamanca, A., Aguena, M., Allam, S., Andrade-Oliveira, F., Annis, J., Bluck, A. F. L., Brooks, D., Burke, D. L., Carrasco Kind, M., Carretero, J., Choi, A., Costanzi, M., da Costa, L. N., Pereira, M. E. S., De Vicente, J., Diehl, H. T., Drlica-Wagner, A., Eckert, K., Everett, S., Evrard, A. E., Ferrero, I., Fosalba, P., Frieman, J., García-Bellido, J., Gerdes, D. W., Giannantonio, T., Gruen, D., Gruendl, R. A., Gschwend, J., Gutierrez, G., Hinton, S. R., Hollowood, D. L., Honscheid, K., James, D. J., Krause, E., Kuehn, K., Kuropatkin, N., Lahav, O., Maia, M. A. G., March, M., Menanteau, F., Miquel, R., Morgan, R., Paz-Chinchón, F., Pieres, A., Plazas Malagón, A. A., Roodman, A., Sanchez, E., Scarpine, V., Serrano, S., Sevilla-Noarbe, I., Smith, M., Soares-Santos, M., Suchyta, E., Swanson, M. E. C., Tarle, G., Thomas, D., and To, C. Galaxy morphological classification catalogue of the Dark Energy Survey Year 3 data with convolutional neural networks. *Monthly Notices of the Royal Astronomical Society*, 507(3):4425–4444, 9 2021. ISSN 0035-8711. doi: 10.1093/mnras/stab2142. URL <https://academic.oup.com/mnras/article/507/3/4425/6327560>.
- Dieleman, S., Willett, K. W., and Dambre, J. Rotation-invariant convolutional neural networks for galaxy morphology prediction. *Monthly Notices of the Royal Astronomical Society*, 450(2):1441–1459, 6 2015. ISSN 1365-2966. doi: 10.1093/mnras/stv632. URL <http://academic.oup.com/mnras/article/450/2/1441/979677/Rotationinvariant-convolutional-neural-networks>.
- Dimauro, P., Daddi, E., Shankar, F., Cattaneo, A., Huertas-Company, M., Bernardi, M., Caro, F., Dupke, R., Häubler, B., Johnston, J., Evelyn, E., Cortesi, A., Mei, S., and Peletier, R. Coincidence between morphology and star-formation activity through cosmic time: The impact of the bulge growth. *Monthly Notices of the Royal Astronomical Society*, 2022. doi: 10.1093/MNRAS/STAC884. URL <https://academic.oup.com/mnras/advance-article/doi/10.1093/mnras/stac884/6563709>.
- Gal, Y. and Ghahramani, Z. Dropout as a Bayesian approximation: Representing model uncertainty in deep learning. In *33rd International Conference on Machine Learning, ICML 2016*, volume 3, pp. 1651–1660. PMLR, 6 2016. ISBN 9781510829008. URL <https://proceedings.mlr.press/v48/gall16.html>.
- Huertas-Company, M., Gravet, R., Cabrera-Vives, G., Pérez-González, P. G., Kartaltepe, J. S., Barro, G., Bernardi, M., Mei, S., Shankar, F., Dimauro, P., Bell, E. F., Kocevski, D., Koo, D. C., Faber, S. M., and Mcintosh, D. H. A CATALOG OF VISUAL-LIKE MORPHOLOGIES IN THE 5 CANDELS FIELDS USING DEEP LEARNING. *The Astrophysical Journal Supplement Series*, 221(1):8, 10 2015. ISSN 1538-4365. doi: 10.1088/0067-0049/

- 221/1/8. URL <https://iopscience.iop.org/article/10.1088/0067-0049/221/1/8>.
- Huertas-Company, M., Bernardi, M., Pérez-González, P. G., Ashby, M. L. N., Barro, G., Conselice, C., Daddi, E., Dekel, A., Dimauro, P., Faber, S. M., Grogin, N. A., Kartaltepe, J. S., Kocevski, D. D., Koekemoer, A. M., Koo, D. C., Mei, S., and Shankar, F. Mass assembly and morphological transformations since $z \approx 3$ from CANDELS. *Monthly Notices of the Royal Astronomical Society*, 462(4):4495–4516, 11 2016. ISSN 0035-8711. doi: 10.1093/mnras/stw1866. URL <https://academic.oup.com/mnras/article-lookup/doi/10.1093/mnras/stw1866>.
- Jaderberg, M., Simonyan, K., Zisserman, A., and Kavukcuoglu, K. Spatial transformer networks. *Advances in Neural Information Processing Systems*, 28: 2017 – 2025, 2015. URL <http://proceedings.neurips.cc/paper/2015/file/33ceb07bf4eeb3da587e268d663aba1a-Paper.pdf>.
- Powell, M. C., Urry, C. M., Cardamone, C. N., Simmons, B. D., Schawinski, K., Young, S., and Kawakatsu, M. MORPHOLOGY AND THE COLOR–MASS DIAGRAM AS CLUES TO GALAXY EVOLUTION AT $z \approx 1$. *The Astrophysical Journal*, 835(1):22, 1 2017. ISSN 1538-4357. doi: 10.3847/1538-4357/835/1/22. URL <https://iopscience.iop.org/article/10.3847/1538-4357/835/1/22>.
- Pozzetti, L., Bolzonella, M., Zucca, E., Zamorani, G., Lilly, S., Renzini, A., Moresco, M., Mignoli, M., Casata, P., Tasca, L., Lamareille, F., Maier, C., Meneux, B., Halliday, C., Oesch, P., Vergani, D., Caputi, K., Kovač, K., Cimatti, A., Cucchiati, O., Iovino, A., Peng, Y., Carollo, M., Contini, T., Kneib, J.-P., Le Fèvre, O., Mainieri, V., Scodeggio, M., Bardelli, S., Bongiorno, A., Coppa, G., de la Torre, S., de Ravel, L., Franzetti, P., Garilli, B., Kampczyk, P., Knobel, C., Le Borgne, J.-F., Le Brun, V., Pellò, R., Perez Montero, E., Ricciardelli, E., Silverman, J. D., Tanaka, M., Tresse, L., Abbas, U., Bottini, D., Cappi, A., Guzzo, L., Koekemoer, A. M., Leauthaud, A., Maccagni, D., Marinoni, C., McCracken, H. J., Memeo, P., Porciani, C., Scaramella, R., Scarlata, C., and Scoville, N. zCOSMOS – 10k-bright spectroscopic sample. *Astronomy & Astrophysics*, 523:A13, 11 2010. ISSN 0004-6361. doi: 10.1051/0004-6361/200913020. URL <http://www.aanda.org/10.1051/0004-6361/200913020>.
- Rowe, B., Jarvis, M., Mandelbaum, R., Bernstein, G., Bosch, J., Simet, M., Meyers, J., Kacprzak, T., Nakajima, R., Zuntz, J., Miyatake, H., Dietrich, J., Armstrong, R., Melchior, P., and Gill, M. GalSim: The modular galaxy image simulation toolkit. *Astronomy and Computing*, 10:121–150, 4 2015. ISSN 22131337. doi: 10.1016/j.ascom.2015.02.002. URL <https://linkinghub.elsevier.com/retrieve/pii/S221313371500013X>.
- Schawinski, K., Urry, C. M., Simmons, B. D., Fortson, L., Kaviraj, S., Keel, W. C., Lintott, C. J., Masters, K. L., Nichol, R. C., Sarzi, M., Skibba, R., Treister, E., Willett, K. W., Wong, O. I., and Yi, S. K. The green valley is a red herring: Galaxy Zoo reveals two evolutionary pathways towards quenching of star formation in early- and late-type galaxies. *Monthly Notices of the Royal Astronomical Society*, 440(1): 889–907, 5 2014. ISSN 1365-2966. doi: 10.1093/mnras/stu327. URL <http://academic.oup.com/mnras/article/440/1/889/1749989/The-green-valley-is-a-red-herring-Galaxy-Zoo>.
- Simard, L., Trevor Mendel, J., Patton, D. R., Ellison, S. L., and McConnachie, A. W. A catalog of bulge+disk decompositions and updated photometry for 1.12 million galaxies in the Sloan digital sky survey. *Astrophysical Journal, Supplement Series*, 196(1):11, 9 2011. ISSN 00670049. doi: 10.1088/0067-0049/196/1/11. URL <http://stacks.iop.org/0067-0049/196/i=1/a=11>.
- Tremaine, S., Gebhardt, K., Bender, R., Bower, G., Dressler, A., Faber, S. M., Filippenko, A. V., Green, R., Grillmair, C., Ho, L. C., Kormendy, J., Lauer, T. R., Magorrian, J., Pinkney, J., and Richstone, D. The Slope of the Black Hole Mass versus Velocity Dispersion Correlation. *The Astrophysical Journal*, 574(2): 740–753, 8 2002. ISSN 0004-637X. doi: 10.1086/341002. URL <https://iopscience.iop.org/article/10.1086/341002>.
- Tuccillo, D., Huertas-Company, M., Decencière, E., Velasco-Forero, S., Domínguez Sánchez, H., and Dimauro, P. Deep learning for galaxy surface brightness profile fitting. *Monthly Notices of the Royal Astronomical Society*, 475(1):894–909, 3 2018. ISSN 0035-8711. doi: 10.1093/mnras/stx3186. URL <http://academic.oup.com/mnras/article/475/1/894/4725057>.
- Vega-Ferrero, J., Domínguez Sánchez, H., Bernardi, M., Huertas-Company, M., Morgan, R., Margalef, B., Aguena, M., Allam, S., Annis, J., Avila, S., Bacon, D., Bertin, E., Brooks, D., Carnero Rosell, A., Carrasco Kind, M., Carretero, J., Choi, A., Conselice, C., Costanzi, M., da Costa, L. N., Pereira, M. E. S., De Vicente, J., Desai, S., Ferrero, I., Fosalba, P., Frieman, J., García-Bellido, J., Gruen, D., Gruendl, R. A., Gschwend, J., Gutierrez, G., Hartley, W. G., Hinton, S. R., Hollowood, D. L., Honscheid, K., Hoyle, B., Jarvis, M., Kim, A. G., Kuehn,

K., Kuropatkin, N., Lima, M., Maia, M. A. G., Menanteau, F., Miquel, R., Ogando, R. L. C., Palmese, A., Paz-Chinchón, F., Plazas, A. A., Romer, A. K., Sanchez, E., Scarpine, V., Schubnell, M., Serrano, S., Sevilla-Noarbe, I., Smith, M., Suchyta, E., Swanson, M. E. C., Tarle, G., Tarsitano, F., To, C., Tucker, D. L., Varga, T. N., and Wilkinson, R. D. Pushing automated morphological classifications to their limits with the Dark Energy Survey. *Monthly Notices of the Royal Astronomical Society*, 506 (2):1927–1943, 7 2021. ISSN 0035-8711. doi: 10.1093/mnras/stab594. URL <https://academic.oup.com/mnras/article/506/2/1927/6156625>.

A. Supplemental Figures

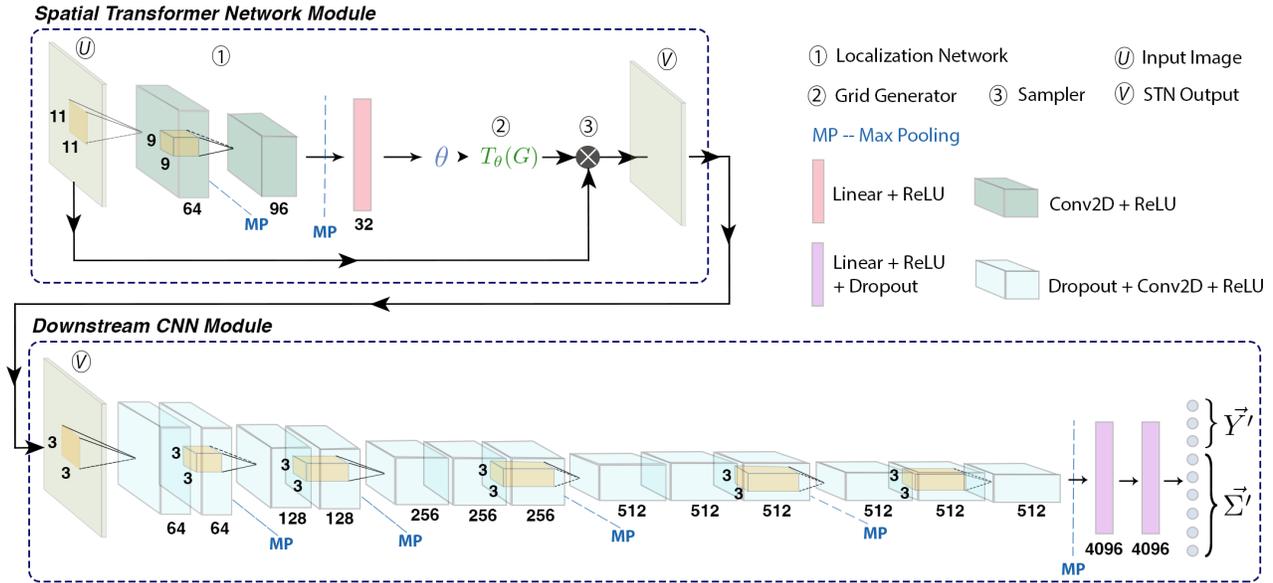


Figure 5. A schematic diagram of the Galaxy Morphology Posterior Estimation Network. GaMPEN’s architecture consists of a downstream CNN module preceded by an upstream STN module. The CNN module empowers GaMPEN to estimate posterior distributions of galaxy morphology parameters. The upstream STN module trains without any extra supervision and learns to apply appropriate cropping transformations to the input image before passing it on to the CNN. The numbers below each layer refer to the number of filters/neurons in each layer. The yellow boxes inside the convolutional layers show the kernel and the number beside it refers to the corresponding kernel size. Only one kernel is shown per set of convolutional layers; all other layers in the set have kernels of the same size. Conv2D and ReLU refer to Convolutional Layers and Rectified Linear Units, respectively.

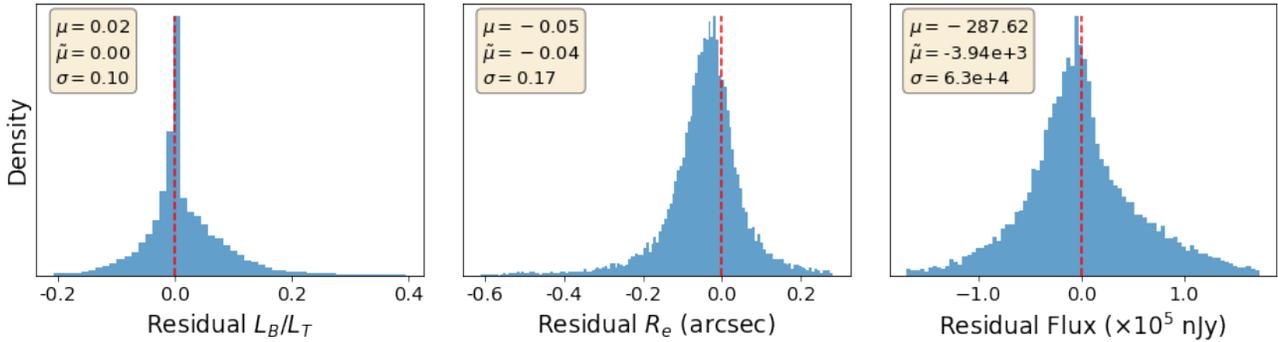


Figure 6. Histograms of residuals for all galaxies in the testing set. We define the residuals as the difference between the true value and the most probable value predicted by GaMPEN. The dashed vertical line represents $x = 0$, denoting cases with perfectly recovered parameter values. The mean (μ), median ($\tilde{\mu}$), and standard deviation (σ) of each residual distribution are listed in each panel.

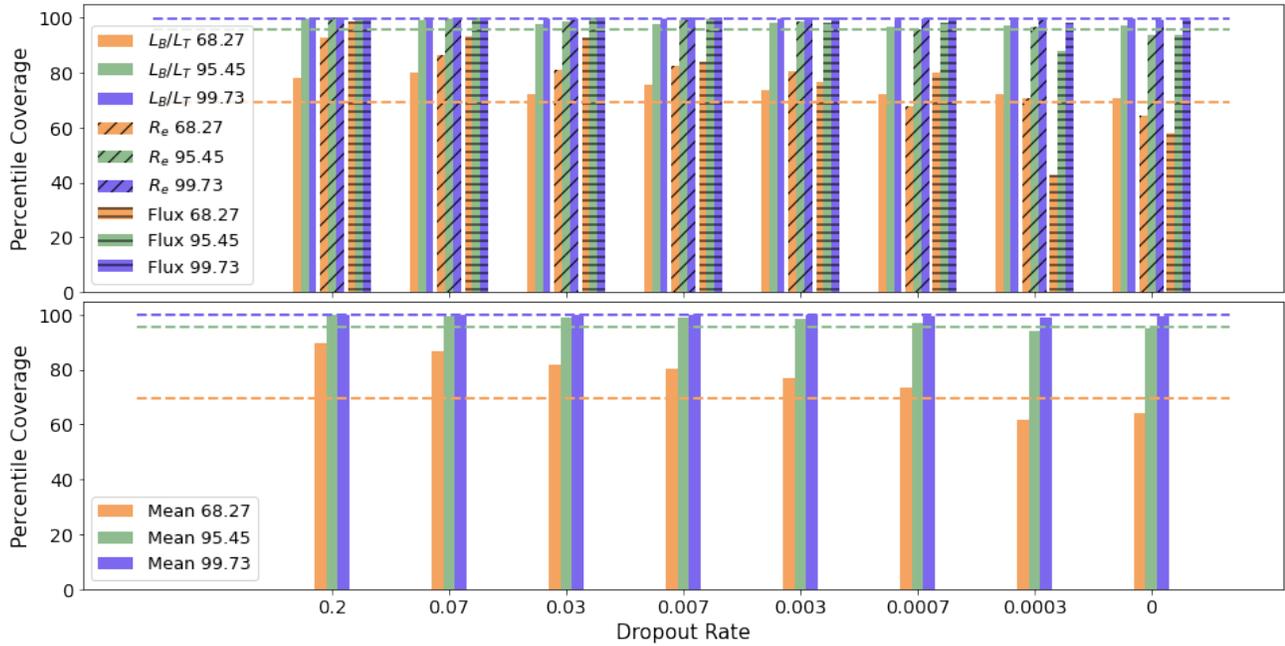


Figure 7. The calculated percentile coverage probabilities for different dropout rates. The top row shows the probabilities for each output variable individually, while the bottom row shows the probabilities averaged over the three variables. The coverage probabilities are defined as the percentage of the total test examples where the true value of the parameter lies within a particular confidence interval of the predicted distribution. A dropout rate of 7×10^{-4} leads to coverage probabilities very close to their corresponding confidence levels.

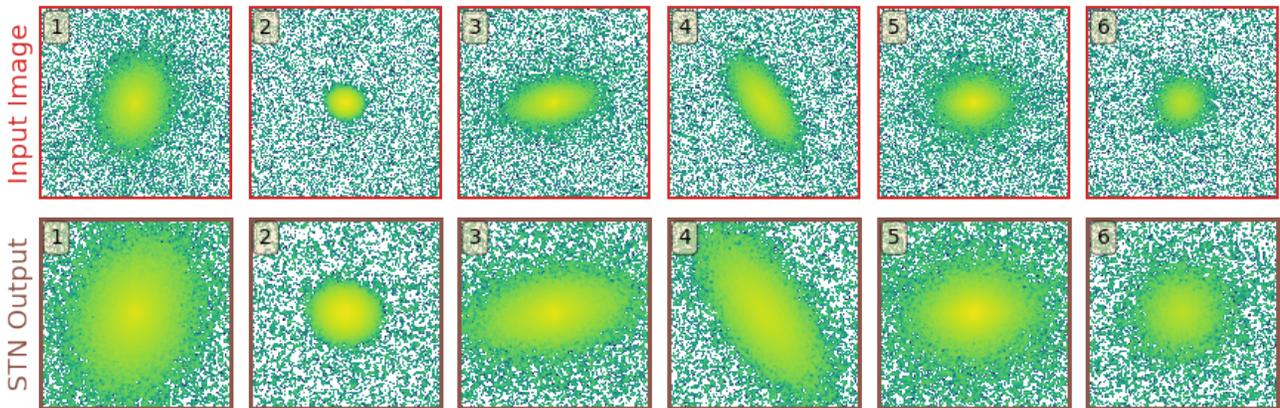


Figure 8. Examples of the transformation applied by the STN to six randomly selected simulated galaxy images. The top row shows the input galaxy images, and the bottom row shows the corresponding output from the STN. The numbers in the top-left yellow boxes help correspond the output images to the input images. As can be seen, the STN learns to apply an optimal amount of cropping for each input galaxy.

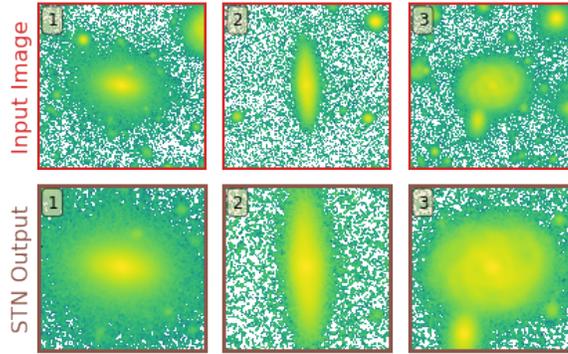


Figure 9. Examples of the transformation applied by a trained STN to real HSC-Wide g-band galaxies. The STN helps the downstream CNN to focus on the galaxy of interest at the center of the cutout by cropping out most secondary galaxies present in the input frame.

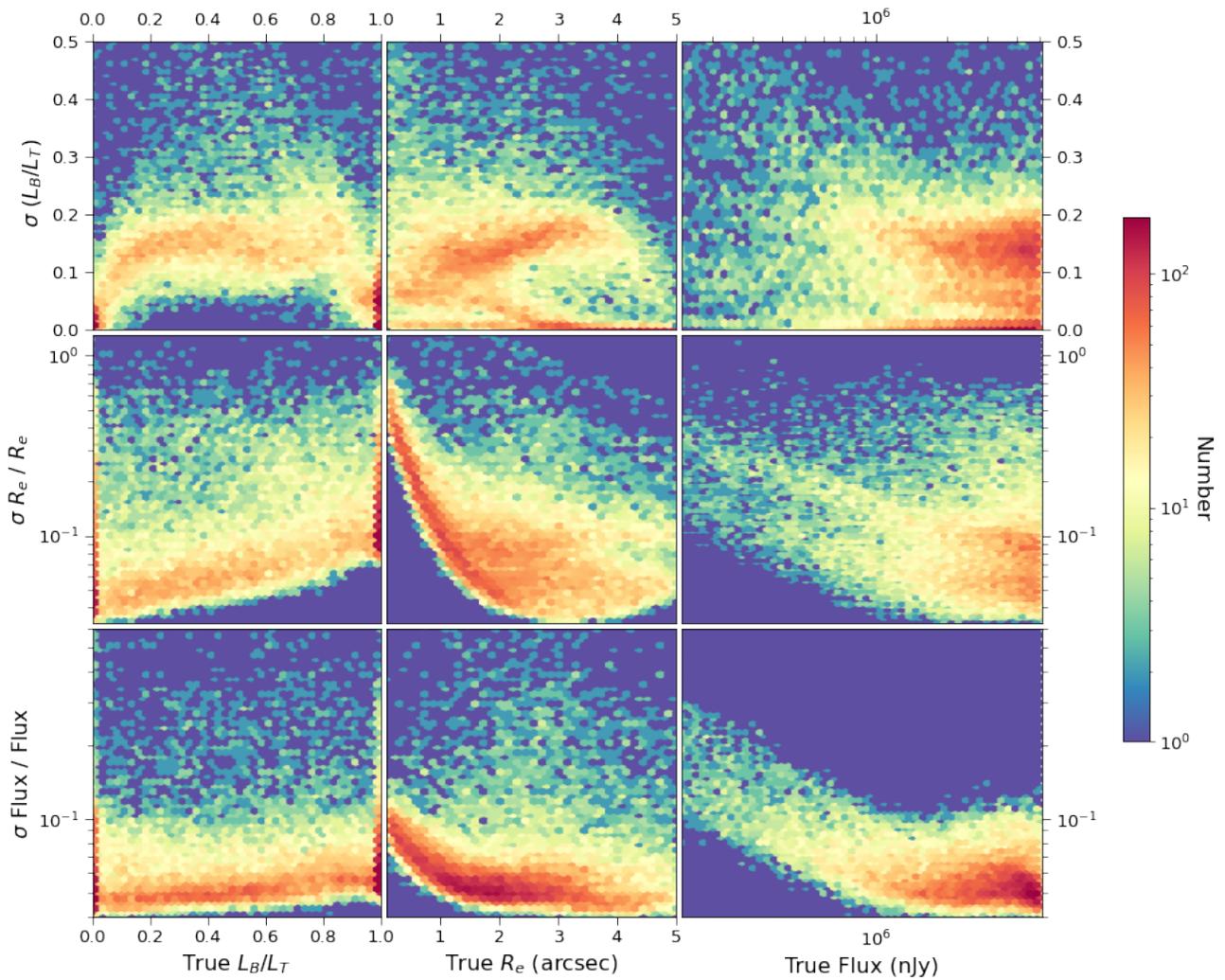


Figure 10. Uncertainties predicted by GaMPEN for each parameter plotted against the true values. The σ for each parameter is defined as the width of the 68.27% confidence interval. Note that we plot fractional uncertainties for radius and flux in order to make the y-axis dimensionless for all three rows.

B. Qualitative Transformation of GaMPEN Predictions

Given that we know GaMPEN residuals are higher for certain regions of the parameter space, we explore how using only qualitative labels in those regions (instead of quantitative predictions) affects the overall residual values. The labeling is informed by the results of §3. The labels are applied based on the predicted values of GaMPEN because we will not have access to true values of the parameters when applying GaMPEN to previously unanalyzed real galaxies. This is crucial given that when we apply GaMPEN to real data, techniques like this will provide us practical tools to deal with predictions in regions of the parameter-space where we know GaMPEN to be less accurate.

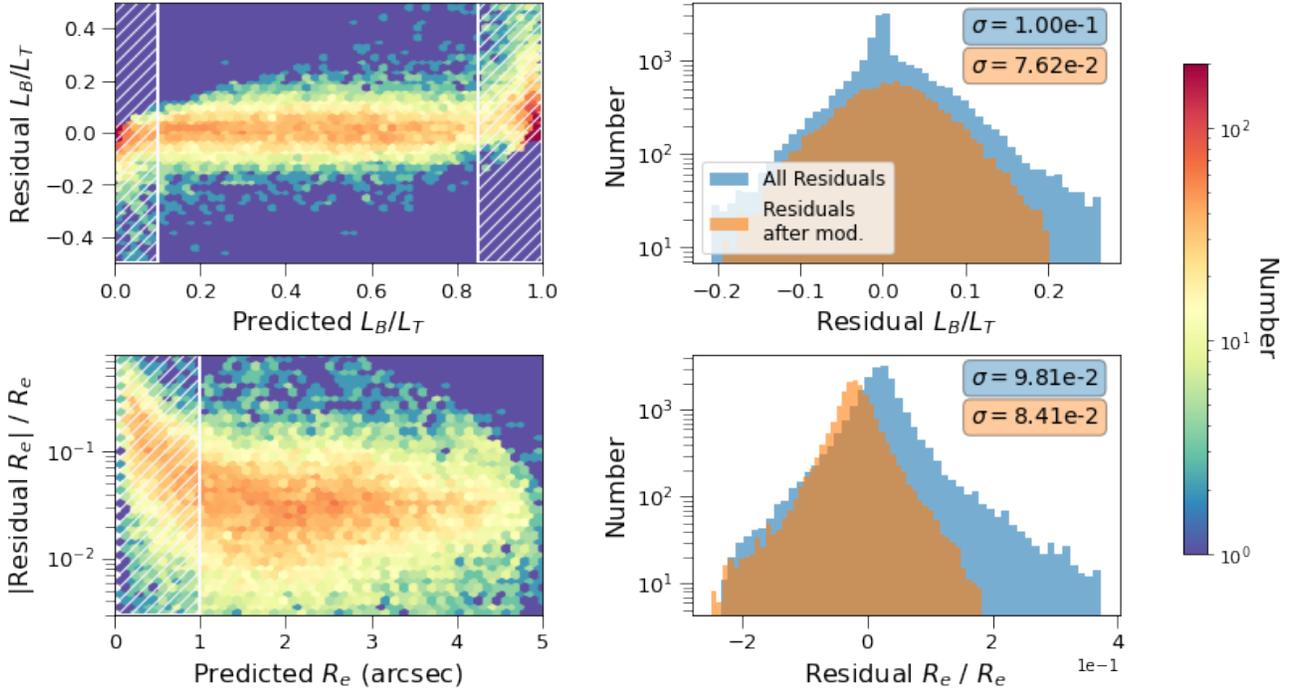


Figure 11. The left panels show the residuals for bulge-to-total light ratio and radius plotted against their predicted values. The white dashed regions show the parameter-space where we replace the quantitative predictions with qualitative flags. Each corresponding histogram on the right shows the distribution of residuals before and after the transformation of output values.

For the bulge-to-total ratio, we retain GaMPEN’s numerical predictions for $0.1 < L_B/L_T < 0.85$, but label the more extreme galaxies as “highly bulge-dominated” ($L_B/L_T \geq 0.85$) or “highly disk-dominated” ($L_B/L_T \leq 0.1$). The top left panel of Figure 11 shows the two labeled regions (white-shaded grid), which is where the residuals are highest. The right panel of the top row shows the residual distributions including and excluding the extreme cases. As indicated by the standard deviation (top right corner), removing these extreme cases eliminates the largest errors in the predicted values of L_B/L_T . We also checked the accuracy of our assigned labels, and show the confusion matrix in Figure 12. From this, we calculate the net accuracy of our extreme L_B/L_T labels to be $\approx 99\%$.

We apply similar labels to small predicted values of the effective radius. As shown in the bottom row of Figure 11, we flag galaxies with $R_e < 1.0$ arcsec with the label “galaxy with $R_e < 1$ arcsec” in place of the exact numerical value. This reduces the typical error for R_e , as shown in the histogram on the right. We calculate the accuracy of this label to be $\sim 97\%$.

Thus, replacing GaMPEN’s quantitative predictions in certain small regions of the parameter space with qualitative flags results in a reduction of the typical residuals as well as highly accurate qualitative predictions.

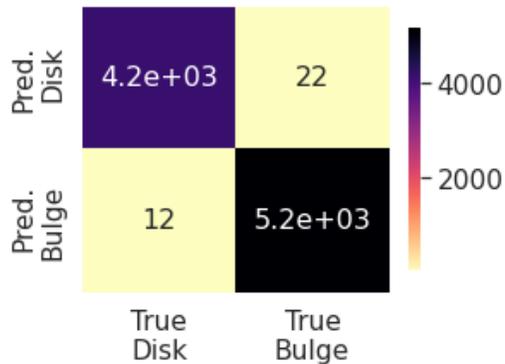


Figure 12. Confusion matrix between the labels we assign when GaMPEN predicts extreme bulge-to-total ratios, $L_B/L_T < 0.1$ or > 0.85 , and their true L_B/L_T values. The number in each block shows how many galaxies correspond to that panel, resulting in an overall accuracy $> 99\%$.

C. Preliminary Results of Application on Real Data

We take the GaMPEN model trained on simulations and further fine-tune it using a small amount of real data. We select $z < 0.25$ HSC-W g-band galaxies with secure redshifts and no imaging issues (such as cosmic ray hits), and then cross-match these with the Simard et al. (2011) catalog, which had performed bulge + disk decomposition using Sloan Digital Sky Survey (SDSS) imaging. Since the depth reached by SDSS is much shallower than HSC-Wide, we also Galfit ~ 2000 additional galaxies to determine their structural properties. We use the Simard et al. (2011) values and our own fits to represent the “true” parameter values. The residuals obtained by comparing these values to the most probable values of the GaMPEN predictions is shown in Figures 13 and 14. As can be seen from both the figures, these preliminary results are extremely promising and closely resemble the results obtained on the simulated galaxies (Figures 4 and 6).

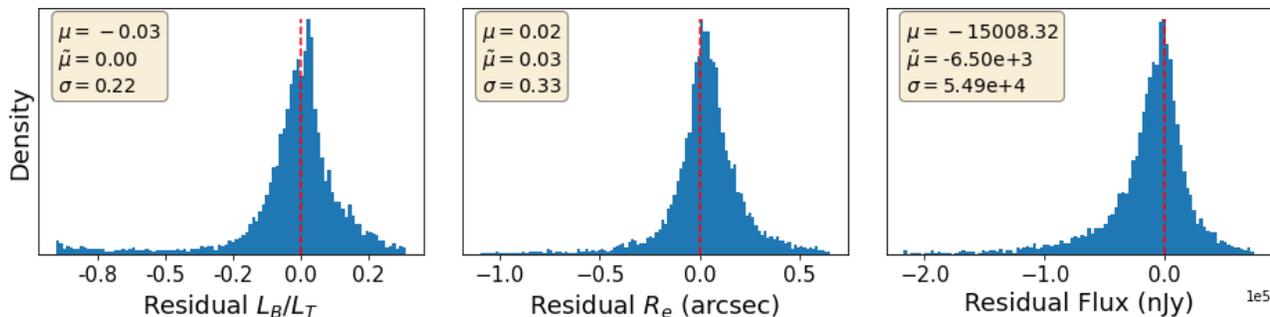


Figure 13. Histograms of residuals for real HSC-W $z < 0.75$ galaxies. We define the residuals as the difference between the true value and the most probable value predicted by GaMPEN. The dashed vertical line represents $x = 0$, denoting cases with perfectly recovered parameter values. The mean (μ), median ($\tilde{\mu}$), and standard deviation (σ) of each residual distribution are listed in each panel.

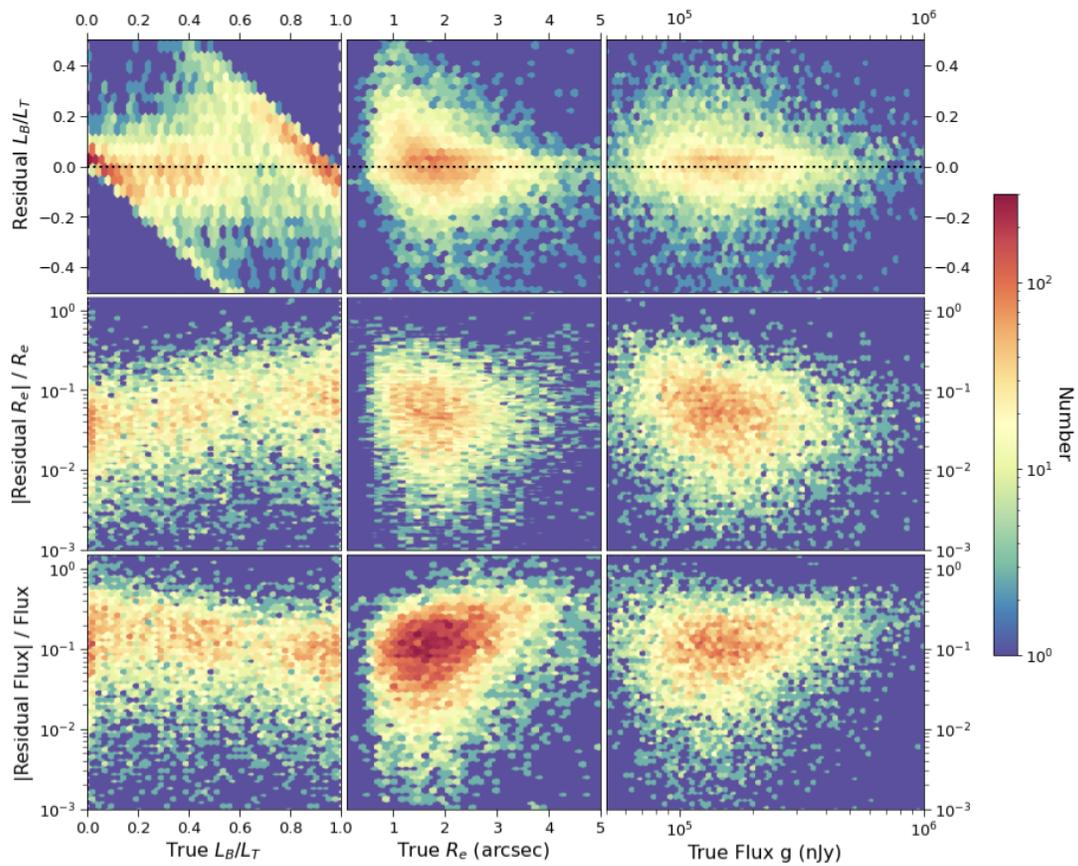


Figure 14. Residuals of GaMPEN predicted parameter values plotted against the true values. The residual for each parameter is defined as the difference between the most probable predicted value and the true value. The color of each hexagonal bin corresponds to the number of galaxies it contains, as shown by the colorbar on the right.