From Simulations to Surveys: Domain Adaptation for Galaxy Observations

Kaley Brauer
Center for Astrophysics
Harvard University
kaley_brauer@cfa.harvard.edu

Aditya Prasad Dash
Physics and Astronomy
University of California, Los Angeles
aditya55@physics.ucla.edu

Meet J. Vyas
International Centre for Space and Cosmology
Ahmedabad University
meet.v2@ahduni.edu.in

Ahmed Salim
Department of Computing
Universiti Teknologi Malaysia
taha-20@live.utm.my

Stiven Briand Massala
Université Paris-Saclay, CentraleSupélec, ENS Paris-Saclay,
CNRS, LMPS - Laboratoire de Mécanique Paris-Saclay
massala934@gmail.com

Abstract

Large photometric surveys will image billions of galaxies, but we currently lack quick, reliable automated ways to infer their physical properties like morphology, stellar mass, and star formation rates. Simulations provide galaxy images with ground-truth physical labels, but domain shifts in PSF, noise, backgrounds, selection, and label priors degrade transfer to real surveys. We present a preliminary domain adaptation pipeline that trains on simulated TNG50 galaxies and evaluates on real SDSS galaxies with morphology labels (elliptical/spiral/irregular). We train three backbones (CNN, $E(2)$-steerable CNN, ResNet-18) with focal loss and effective-number class weighting, and a feature-level domain loss $\mathcal{L}_D$ built from `GeomLoss` (entropic Sinkhorn OT, energy distance, Gaussian MMD, and related metrics). We show that a combination of these losses with an OT-based "top-k soft matching" loss that focuses $\mathcal{L}_D$ on the worst-matched source–target pairs can further enhance domain alignment. With Euclidean distance, scheduled alignment weights, and top-k matching, target accuracy (macro F1) rises from $\sim$46% ($\sim$30%) at no adaptation to $\sim$87% ($\sim 62.6\%$), with a domain AUC near 0.5, indicating strong latent-space mixing.

https://github.com/ahmedsalim3/galaxy-da

1 Introduction

Machine learning is now routine in the physical sciences, yet transferring models across domains remains challenging and scientifically consequential. For galaxies, simulated images carry true physical labels but differ from real observations in PSF, noise, backgrounds, selection functions, and demographics. Naively transferring such models can bias physical inferences, shifting the early/late-type mix, distorting mass–SFR demographics, or contaminating scaling relations, so robust domain adaptation must preserve calibrated, physically meaningful predictions (and uncertainties) for population studies and hypothesis testing.

Domain adaptation has emerged as a widely used and powerful paradigm for addressing classification tasks in which source and target domains follow different data distributions [39; 36; 44; 8; 47; 40; 22; 7; 13; 27], with successful applications across various domains [18; 42; 21; 24; 11]. Astronomers have similarly adopted these techniques to infer astrophysical and cosmological parameters [43; 32; 17; 1; 35; 41]. With expanding simulation capabilities and survey infrastructure, the need for robust domain-transfer methods is increasingly timely, motivating updates to techniques based on time-domain projects such as ZTF [4] and Pan-STARRS1 [15] in preparation for projects including the Vera C. Rubin Observatory [29], the Nancy Grace Roman Space Telescope [2], and the Euclid mission [37], which will provide millions of nightly alerts.

In this work, we prototype a pipeline to study how different domain-adaptation techniques can narrow this gap in a realistic setting by training on mock observations of simulated galaxies from TNG50 SKIRT [34] and testing on real SDSS observations with Galaxy Zoo–derived morphology labels [26]. We frame this as a covariate- (or coverage-) shift problem in which the conditional label distribution is approximately stable, $p_S(y \mid x) \approx p_T(y \mid x)$, while the input and selection distributions differ, $p_S(x) \neq p_T(x)$, making our setup representative of many "train on simulations, test on surveys" applications. Previous studies have reported promising results [30; 31; 6; 49; 48], yet room remains for methodological development. We also explore a suite of distance metrics for quantifying domain shift by updating the GeomLoss Library [14] with new metrics.
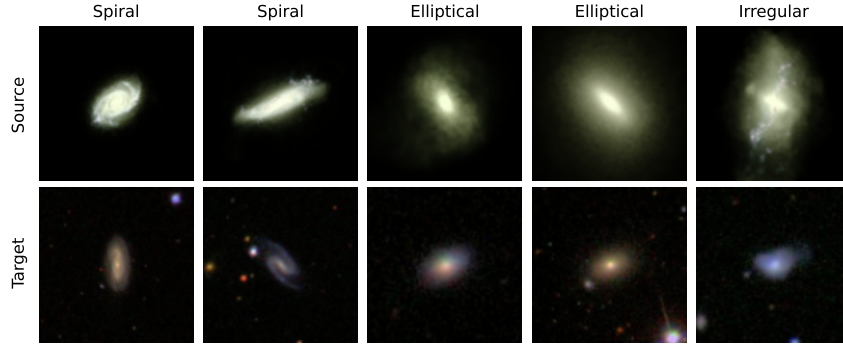
2   Data

Figure 1: **Example source and target data.** Our source data are mock observations of galaxies from the TNG50 simulation with known galaxy properties. Our target data are real observations of galaxies. In this proof of concept, we use SDSS observations at $z \sim 0$ with known morphologies. Real data differ from simulated data in point spread function, noise, selection function, demographics, and more.

Source (TNG50 Simulation).   Our source set comprises mock observations of galaxies from the Illustris TNG50 simulation suite [28; 33]. We use SKIRT synthetic images and morphology classifications at z=0 and $z \approx 0.05$ [34], restricted to $M_\star > 10^{9.5} \, M_\odot$, with quality cuts (`flag=0`, `sn_per_pixel>2.5`; `flag_sersic=0` when Sérsic parameters are used). Images have four bands (g,r,i,z) and are "idealized"; they are convolved with a Gaussian PSF and have sky noise added per [34]. The dataset includes 3232 galaxies. We split 90/10 train/test and augment via one horizontal flip and four rotations, yielding 25,856 total images. For future work, we also produce labels for stellar mass, star formation rate (defining "star-forming" as specific SFR $> 10^{-11}$ yr^{-1}), metallicity, and whether or not the galaxy includes an active galactic nuclei (AGN; galaxies with actively accreting black holes, defined as $\lambda_{\mathrm{Edd}} \simeq \dot{M}/\dot{M}_{\mathrm{Edd}} > 10^{-3}$).

Target (SDSS Survey).   Our target set comprises 6416 Sloan Digital Sky Survey (SDSS) galaxy images [46] with morphologies labeled by Galaxy Zoo volunteers, using debiased vote fractions from [19]. We select three morphology classes – elliptical, spiral, and irregular – by applying strict thresholds on these fractions: ellipticals require high smoothness, strong "no spiral" votes, and low edge-on probability; spirals require strong spiral-arm and feature/disk votes while excluding highly edge-on systems; irregulars require strong votes for "odd/irregular," "merger," or "disturbed" features. Objects flagged as artifacts are removed. Additional labels for stellar mass, star formation, and

AGN come from [38; 26; 25]; following the BPT scheme [3] as in [38], we label BPTCLASS = 1 galaxies as star-forming, BPTCLASS = 3 or 4 as AGN hosts, and exclude BPTCLASS = 0 or 2.

The three morphology classes are highly imbalanced (spirals dominate, with ellipticals and especially irregulars rarer), motivating a focal loss with effective-number class weights (Section 3). We also assume simulation-based and Galaxy Zoo morphologies provide compatible labels for the same classes, justified by strict Galaxy Zoo cuts and visual similarity of source and target images.

3   Methods

Architectures.   We compare three backbone architectures: (i) a small convolutional network (CNN) with two conv blocks, each with two 3×3 conv-BN-ReLU layers and 2×2 max pooling, followed by a three-layer MLP classifier; (ii) an $E(2)$-steerable CNN (ESCNN) implemented with `escnn` [45] using a discrete rotation group C_8 and four group-equivariant convolutional blocks; and (iii) a `torchvision resnet18` [20] backbone pretrained on ImageNet [10], where we replace the final fully-connected layer by a task-specific MLP head and fine-tune only the top residual blocks, keeping the rest frozen. Unless otherwise stated we train for 200 epochs (with early stopping) with batch size 128 using AdamW.

Supervised loss and class imbalance.   All models use MLP classifier head and are trained with a supervised loss on the labeled source domain. Because the three morphological classes are highly imbalanced, we use a focal loss $\mathcal{L}_{\text{sup}}$ in place of cross-entropy, with focusing parameter $\gamma=2$ and α determined from per-class weights computed from the "effective number of samples"[9]. We also use class scales to boost minority class predictions. Unlike class weights (used in the loss), these scales are applied to logits to boost predictions. The scales are computed from the source training data using the formula $\text{scale}_c = \sqrt{\max(\text{count})/\text{count}_c}$, normalized so the majority class has scale 1.0, then moderated to reduce aggressiveness. The scales are implemented as a learnable parameter in the model, initialized from these data-driven values and applied multiplicatively to the logits before the final classification output.

Domain adaptation objectives.   Let z_s and z_t denote the feature embeddings (just before the classifier head) for source and target samples in a mini-batch. We introduce a novel combination of domain-alignment terms $\mathcal{L}_D$ computed on L2-normalized features using differentiable distance-based losses implemented with `GeomLoss` [14].

In particular, we experiment with (i) entropic optimal transport (Sinkhorn divergence), (ii) the energy distance, and (iii) Gaussian MMD, corresponding to `SamplesLoss` methods "sinkhorn", "energy", and "gaussian" (with blur parameter σ) respectively. We extend the GeomLoss library [14] by implementing a comprehensive suite of 46 distinct distance and similarity measures derived from the taxonomy of [5]. This expansion augments the standard Euclidean formulations with metrics organized into eight distinct families: *L_p Minkowski, L_1, Intersection, Inner Product, Squared-chord, Squared L_2 (χ^2), Shannon's Entropy*, and *Combination*.

Recognizing that the topology of the embedding space fundamentally constrains cross-domain alignment, we empirically validate the choice of metric by benchmarking model performance across 12 representative measures: *inner product, cosine, Jaccard, Dice coefficient, Kumar–Hassebrook, harmonic mean, Euclidean, Manhattan, Chebyshev, Minkowski, Gower*, and *average L_1–L_∞*. These choices let us probe how different metrics on the latent space affect alignment and downstream performance.

On top of this, we add an explicit optimal-transport alignment loss that combines global OT, soft matching, and a top-k penalty.

Using P^λ, we define the full alignment loss:

$$\mathcal{L}_{\text{OT}}(z_s, z_t) \;=\; \lambda_{\text{OT}}\, d_\lambda(p_s, p_t) \;+\; \lambda_{\text{match}}\, \text{MSE}\big(z_s,\; P^\lambda z_t\big) \;+\; \lambda_{\text{topk}}\, \frac{1}{k}\sum_{\ell=1}^{k} d_{(\ell)},$$

where $C_{ij} = \|z_{s,i} - z_{t,j}\|_2^2$ is the pairwise cost matrix, $d_\lambda(p_s, p_t)$ is an entropically regularized OT distance computed from a Sinkhorn transport plan P^λ, $d_i = \min_j C_{ij}$ and $d_{(1)} \geq \cdots \geq d_{(k)}$ are the k largest per-source closest-target distances.

3

This composite formulation couples global OT alignment, soft barycentric matching, and a focused top-k penalty on poorly aligned instances. Top-k losses have proved effective in supervised alignment [12]; here we use the same idea to emphasize the hardest-to-align source features.

Our training objective combines the supervised and domain-alignment terms as

$$\mathcal{L} = \lambda_{\text{sup}} \, \mathcal{L}_{\text{sup}} + \lambda_D \, \mathcal{L}_D + \lambda_{\text{OT}} \, \mathcal{L}_{\text{OT}},$$

where each λ is a non-negative scalar that controls the relative weights of the losses (with $\lambda_{\text{OT}}{=}0$ for experiments without OT/top-k). We evaluate several strategies for choosing $(\lambda_{\text{sup}}, \lambda_D)$:

- **Baseline.** $\lambda_D{=}0$, training only with $\mathcal{L}_{\text{sup}}$.
- **Fixed-λ.** We set $\lambda_D{=}\lambda$ and $\lambda_{\text{sup}}{=}1$, tuning λ over a small grid; $\lambda{=}0.1$ is used for the main experiments.
- **Trainable weights.** We use trainable parameters η_1, η_2 and set $\lambda_{\text{sup}} = \sigma(\eta_1)$, $\lambda_D = \sigma(\eta_2)$, learning the alignment strength jointly with the network. [23]
- **Blur schedule + trainable weights.** For Sinkhorn-based alignment, we anneal the blur parameter from σ_{initial} to σ_{final} over training epochs, enabling coarse-to-fine OT alignment.
- **Domain-adversarial training (DANN).** As a complementary baseline, we implement a GRL-based domain discriminator [16] acting on (z_s, z_t) (`DAAdversarialTrainer`). The GRL scale λ_{adv} is tied to `lambda_da` (set to 0.1 in the main experiments).

Top-k soft domain alignment. Using the composite OT loss defined above, we test a restricted top-k / soft-matching variant that focuses the alignment on the hardest-to-match examples. Standard batch-level alignment treats all source–target pairs equally, which can encourage mismatched couplings (e.g. spirals aligned to ellipticals) when the domains differ. Our OT term instead combines a transport-induced soft matching with a sparse top-k penalty on the worst-aligned source features, encouraging ellipticals to align with ellipticals and spirals with spirals rather than with arbitrary batch neighbors. This is the key modification that yields our best domain-alignment performance in our experiments, including all three backbones.

Training procedure. All models are trained for up to 200 epochs with early stopping based on the training loss (patience of 10 epochs). We use a batch size of 128 and resize all images to $224{\times}224$ pixels. For domain-adaptation methods, we include a warmup phase of 20 epochs during which only the supervised focal loss is applied, allowing the model to establish a reasonable feature representation before introducing domain-alignment terms. We optimize using AdamW with a weight decay of 10^{-2} and apply gradient clipping with a maximum norm of 10.0 to stabilize training. The learning rate follows a cosine-annealing schedule, starting at 10^{-4} and decaying to 10^{-6} over the full training horizon, reducing the step size as training progresses while preserving relatively large updates early on. Source and target images are normalized separately using domain-specific statistics computed from their respective training sets, preserving the natural distributional differences that domain adaptation aims to bridge.

4 Results

Accuracy improvements on target domain. Figure 2 illustrates 200-epoch training dynamics for baseline, adversarial, and Euclidean models, selecting Euclidean as a representative metric for brevity. The left panel displays target-domain accuracy (solid lines) and macro–F1 (dashed lines). The performance trajectories reveal a clear distinction: the Baseline (grey) exhibits significant instability and terminates early due to loss degradation, whereas the domain adaptation methods show a sharp performance jump after the 20-epoch warmup phase, maintaining high stability thereafter. The loss plot (right panel) further distinguishes the trainable Euclidean variants (purple and yellow), which converge to a distinct negative loss range compared to the fixed parameters.

At peak performance (epoch 197), the Euclidean method with trainable weights (η_1, η_2) achieved the best results (**87.3% accuracy, 0.626 macro–F1**), a dramatic improvement over the baseline (46.8%, 0.298). This quantitative boost is visualized in the confusion matrices (Fig. 3), which demonstrate significantly cleaner class separation after adaptation. While DANN (86.5%) and other Euclidean variants yielded competitive gains, the fully trainable weighting scheme consistently provided the
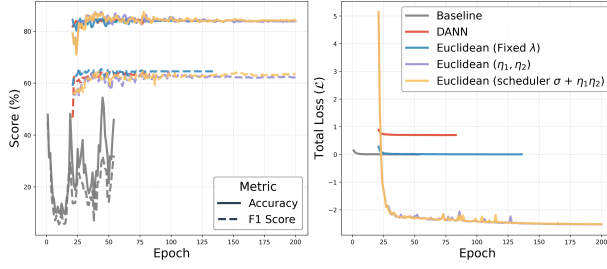
Figure 2: **Training Dynamics.** Comparison of stability and loss convergence.
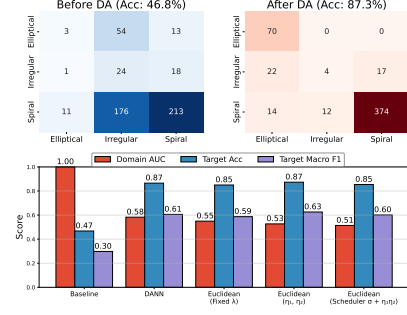

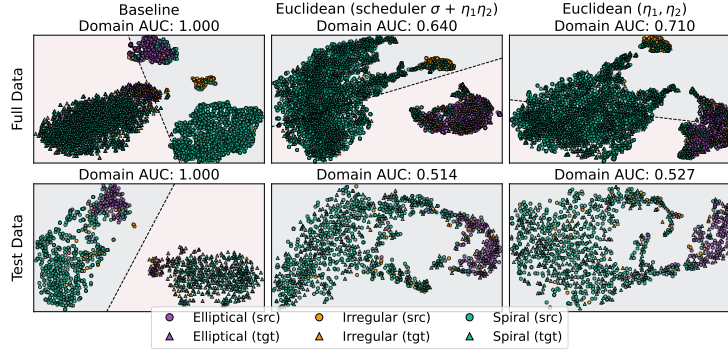Figure 3: **Target Accuracy.** Confusion matrix and metrics.

Figure 4: **Latent space embedding.** The Baseline (left, AUC=1.00) shows distinct domain separation, while Euclidean variants (center/right, AUC≈0.51) achieve effective domain alignment where source and target distributions are indistinguishable. In an ideal domain-invariant representation, AUC = 0.5.

most stable alignment. We find that top-k soft-matching with ResNet effectively aligns classes across domains, driving these improvements in both accuracy and F1.

5   Discussion & Next Steps

This work is the precursor to a domain-adapted predictive model to leverage hundreds of thousands of mock observations from the Illustris simulations at different redshifts with the ultimate goal of interpreting upcoming Rubin galaxy observations.

Next steps include (1) **Richer labels:** extending from three-way morphology to a multi-task setting (e.g., star-forming vs. quenched, AGNs, stellar mass.) (2) **Irregular class:** improving the rare irregular class with targeted augmentations and stronger imbalance handling. (3) **Latent space and distance metrics:** probing the learned feature space under different distance metrics in $\mathcal{L}_D$ to understand when particular metrics yield better physical alignment between simulations and observations, and to guide the choice of future architectures. (4) **Learning rate scheduler:** incorporating distance-aware schedulers that modulate the learning rate using alignment signals such as Sinkhorn cost, top-k distances, or barycentric mismatch, and quantifying their impact on convergence stability and domain-alignment. (5) **Alternative architectures:** testing equivariant transformers or graph-based models to see how different inductive biases affect the geometry of the learned latent space.

Acknowledgments and Disclosure of Funding

References

[1] S. AGARWAL, A. ĆIPRIJANOVIĆ, AND B. D. NORD, *Neural Network Prediction of Strong Lensing Systems with Domain Adaptation and Uncertainty Quantification*, arXiv e-prints, (2024), p. arXiv:2411.03334.

[2] R. AKESON, L. ARMUS, E. BACHELET, V. BAILEY, AND L. B. ET AL., *The wide field infrared survey telescope: 100 hubbles for the 2020s*, 2019.

[3] J. A. BALDWIN, M. M. PHILLIPS, AND R. TERLEVICH, *Classification parameters for the emission-line spectra of extragalactic objects*, Publications of the Astronomical Society of the Pacific, 93 (1981), p. 5.

[4] E. C. BELLM, S. R. KULKARNI, M. J. GRAHAM, R. DEKANY, AND R. M. S. ET AL., *The zwicky transient facility: System overview, performance, and first results*, Publications of the Astronomical Society of the Pacific, 131 (2019), p. 018002.

[5] S.-H. CHA, *Comprehensive survey on distance/similarity measures between probability density functions*, City, 1 (2007), p. 1.

[6] 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 (2023), p. 025013.

[7] N. COURTY, R. FLAMARY, D. TUIA, AND A. RAKOTOMAMONJY, *Optimal transport for domain adaptation*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 39 (2014), pp. 1853–1865.

[8] G. CSURKA, *Domain adaptation for visual applications: A comprehensive survey*, arXiv preprint arXiv:1702.05374, (2017).

[9] Y. CUI, M. JIA, T.-Y. LIN, Y. SONG, AND S. BELONGIE, *Class-balanced loss based on effective number of samples*, in 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019, pp. 9260–9269.

[10] J. DENG, W. DONG, R. SOCHER, L.-J. LI, K. LI, AND L. FEI-FEI, *Imagenet: A large-scale hierarchical image database*, in 2009 IEEE Conference on Computer Vision and Pattern Recognition, 2009, pp. 248–255.

[11] R. DING, J. LIU, K. HUA, X. WANG, X. ZHANG, M. SHAO, Y. CHEN, AND J. CHEN, *Leveraging data mining, active learning, and domain adaptation for efficient discovery of advanced oxygen evolution electrocatalysts*, Science Advances, 11 (2025), p. eadr9038.

[12] Y. FAN, S. LYU, Y. YING, AND B.-G. HU, *Learning with average top-k loss*, in Neural Information Processing Systems, 2017.

[13] A. FARAHANI, S. VOGHOEI, K. RASHEED, AND H. R. ARABNIA, *A brief review of domain adaptation*, Advances in data science and information engineering: proceedings from ICDATA 2020 and IKE 2020, (2021), pp. 877–894.

[14] J. FEYDY, T. SÉJOURNÉ, F.-X. VIALARD, S.-I. AMARI, A. TROUVE, AND G. PEYRÉ, *Interpolating between optimal transport and mmd using sinkhorn divergences*, in The 22nd International Conference on Artificial Intelligence and Statistics, 2019, pp. 2681–2690.

[15] H. A. FLEWELLING, E. A. MAGNIER, K. C. CHAMBERS, J. N. HEASLEY, AND C. H. ET AL., *The pan-starrs1 database and data products*, The Astrophysical Journal Supplement Series, 251 (2020), p. 7.

[16] Y. GANIN AND V. LEMPITSKY, *Unsupervised domain adaptation by backpropagation*, in Proceedings of the 32nd International Conference on International Conference on Machine Learning - Volume 37, ICML'15, JMLR.org, 2015, p. 1180–1189.

[17] S. GILDA, A. DE MATHELIN, S. BELLSTEDT, AND G. RICHARD, *Unsupervised Domain Adaptation for Constraining Star Formation Histories*, Astronomy, 3 (2024), pp. 189–207.

[18] H. GUAN AND M. LIU, *Domain adaptation for medical image analysis: A survey*, IEEE Transactions on Biomedical Engineering, 69 (2021), pp. 1173–1185.

[19] R. E. HART, S. P. BAMFORD, K. W. WILLETT, K. L. MASTERS, C. CARDAMONE, C. J. LINTOTT, R. J. MACKAY, R. C. NICHOL, C. K. ROSSLOWE, B. D. SIMMONS, AND R. J. SMETHURST, *Galaxy zoo: comparing the demographics of spiral arm number and a new method for correcting redshift bias*, Monthly Notices of the Royal Astronomical Society, 461 (2016), pp. 3663–3682.

[20] K. HE, X. ZHANG, S. REN, AND J. SUN, *Deep residual learning for image recognition*, in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770–778.

[21] J. JAKUBIK, M. MUSZYNSKI, M. VÖSSING, N. KÜHL, AND T. BRUNSCHWILER, *Unsupervised Domain Adaptation for Geospatial Semantic Segmentation*, in AGU Fall Meeting Abstracts, vol. 2023, Dec. 2023, pp. IN53A–07.

[22] G. KANG, L. JIANG, Y. YANG, AND A. HAUPTMANN, *Contrastive adaptation network for unsupervised domain adaptation*, 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), (2019), pp. 4888–4897.

[23] A. KENDALL, Y. GAL, AND R. CIPOLLA, *Multi-task learning using uncertainty to weigh losses for scene geometry and semantics*, CoRR, abs/1705.07115 (2017).

[24] J. LI, R. XU, X. LIU, J. MA, B. LI, Q. ZOU, J. MA, AND H. YU, *Domain adaptation based object detection for autonomous driving in foggy and rainy weather*, IEEE Transactions on Intelligent Vehicles, 10 (2025), pp. 900–911.

[25] C. LINTOTT, K. SCHAWINSKI, S. BAMFORD, A. SLOSAR, K. LAND, D. THOMAS, E. EDMONDSON, K. MASTERS, R. C. NICHOL, M. J. RADDICK, A. SZALAY, D. ANDREESCU, P. MURRAY, AND J. VANDENBERG, *Galaxy zoo 1: data release of morphological classifications for nearly 90,0000 galaxies*, Monthly Notices of the Royal Astronomical Society, 410 (2011), pp. 166–178.

[26] C. J. LINTOTT, K. SCHAWINSKI, A. SLOSAR, K. LAND, S. BAMFORD, D. THOMAS, M. J. RADDICK, R. C. NICHOL, A. SZALAY, D. ANDREESCU, P. MURRAY, AND J. VANDENBERG, *Galaxy zoo: morphologies derived from visual inspection of galaxies from the sloan digital sky survey*, Monthly Notices of the Royal Astronomical Society, 389 (2008), pp. 1179–1189.

[27] X. LIU, C. YOO, F. XING, H. OH, G. FAKHRI, J.-W. KANG, AND J. WOO, *Deep unsupervised domain adaptation: A review of recent advances and perspectives*, APSIPA Transactions on Signal and Information Processing, (2022).

[28] D. NELSON, A. PILLEPICH, V. SPRINGEL, R. PAKMOR, R. WEINBERGER, S. GENEL, P. TORREY, M. VOGELSBERGER, F. MARINACCI, AND L. HERNQUIST, *First results from the TNG50 simulation: galactic outflows driven by supernovae and black hole feedback*, MNRAS, 490 (2019), pp. 3234–3261.

[29] NSF-DOE VERA C. RUBIN OBSERVATORY, *Legacy Survey of Space and Time Data Preview 1 [Data set]*, (2025).

[30] S. PANDYA, P. PATEL, B. D. NORD, M. WALMSLEY, AND A. ĆIPRIJANOVIĆ, *Sidda : Sinkhorn dynamic domain adaptation for image classification with equivariant neural networks*, Machine Learning: Science and Technology, 6 (2025), p. 035032.

[31] S. PANDYA, P. PATEL, F. O, AND J. BLAZEK, *E(2) equivariant neural networks for robust galaxy morphology classification*, 2023.

[32] H. PARUL, S. GLEYZER, P. REDDY, AND M. W. TOOMEY, *Domain Adaptation in Application to Gravitational Lens Finding*, , 990 (2025), p. 47.

[33] A. PILLEPICH, D. NELSON, V. SPRINGEL, R. PAKMOR, P. TORREY, R. WEINBERGER, M. VOGELSBERGER, F. MARINACCI, S. GENEL, A. VAN DER WEL, AND L. HERNQUIST, *First results from the TNG50 simulation: the evolution of stellar and gaseous discs across cosmic time*, MNRAS, 490 (2019), pp. 3196–3233.

[34] V. RODRIGUEZ-GOMEZ, G. F. SNYDER, J. M. LOTZ, D. NELSON, A. PILLEPICH, V. SPRINGEL, S. GENEL, R. WEINBERGER, S. TACCHELLA, R. PAKMOR, P. TORREY, F. MARINACCI, M. VOGELSBERGER, L. HERNQUIST, AND D. A. THILKER, *The optical morphologies of galaxies in the IllustrisTNG simulation: a comparison to Pan-STARRS observations*, MNRAS, 483 (2019), pp. 4140–4159.

[35] A. RONCOLI, A. ĆIPRIJANOVIĆ, M. VOETBERG, F. VILLAESCUSA-NAVARRO, AND B. NORD, *Domain Adaptive Graph Neural Networks for Constraining Cosmological Parameters Across Multiple Data Sets*, arXiv e-prints, (2023), p. arXiv:2311.01588.

[36] K. SAITO, D. KIM, P. TETERWAK, S. SCLAROFF, T. DARRELL, AND K. SAENKO, *Tune it the right way: Unsupervised validation of domain adaptation via soft neighborhood density*, in 2021 IEEE/CVF International Conference on Computer Vision (ICCV), 2021, pp. 9164–9173.

[37] R. SCARAMELLA, J. AMIAUX, Y. MELLIER, C. BURIGANA, AND C. S. C. ET AL., *Euclid preparation*, Astronomy & Astrophysics, 662 (2022), p. A112.

[38] K. SCHAWINSKI, C. M. URRY, S. VIRANI, P. COPPI, S. P. BAMFORD, E. TREISTER, C. J. LINTOTT, M. SARZI, W. C. KEEL, S. KAVIRAJ, C. N. CARDAMONE, K. L. MASTERS, N. P. ROSS, D. ANDREESCU, P. MURRAY, R. C. NICHOL, M. J. RADDICK, A. SLOSAR, A. S. SZALAY, D. THOMAS, AND J. VANDENBERG, *Galaxy zoo: The fundamentally different co-evolution of supermassive black holes and their early- and late-type host galaxies*, The Astrophysical Journal, 711 (2010), pp. 284–302.

[39] C. SHUI, Q. CHEN, J. WEN, F. ZHOU, C. GAGNÉ, AND B. WANG, *A novel domain adaptation theory with jensen–shannon divergence*, Knowledge-Based Systems, 257 (2022), p. 109808.

[40] B. SUN, J. FENG, AND K. SAENKO, *Correlation alignment for unsupervised domain adaptation*, in Domain adaptation in computer vision applications, Springer, 2017, pp. 153–171.

[41] P. SWIERC, M. ZHAO, A. ĆIPRIJANOVIĆ, AND B. NORD, *Domain Adaptation for Measurements of Strong Gravitational Lenses*, arXiv e-prints, (2023), p. arXiv:2311.17238.

[42] D. TUIA, C. PERSELLO, AND L. BRUZZONE, *Domain adaptation for the classification of remote sensing data: An overview of recent advances*, IEEE Geoscience and Remote Sensing Magazine, 4 (2016), pp. 41–57.

[43] R. VILALTA, K. D. GUPTA, D. BOUMBER, AND M. M. MESKHI, *A general approach to domain adaptation with applications in astronomy*, Publications of the Astronomical Society of the Pacific, 131 (2019), p. 108008.

[44] M. WANG AND W. DENG, *Deep visual domain adaptation: A survey*, Neurocomputing, 312 (2018), p. 135–153.

[45] M. WEILER AND G. CESA, *General E(2)-Equivariant Steerable CNNs*, in Conference on Neural Information Processing Systems (NeurIPS), 2019.

[46] K. W. WILLETT, C. J. LINTOTT, S. P. BAMFORD, K. L. MASTERS, B. D. SIMMONS, K. R. V. CASTEELS, E. M. EDMONDSON, L. F. FORTSON, S. KAVIRAJ, W. C. KEEL, T. MELVIN, R. C. NICHOL, M. J. RADDICK, K. SCHAWINSKI, R. J. SIMPSON, R. A. SKIBBA, A. M. SMITH, AND D. THOMAS, *Galaxy zoo 2: detailed morphological classifications for 304122 galaxies from the sloan digital sky survey*, Monthly Notices of the Royal Astronomical Society, 435 (2013), pp. 2835–2860.

[47] G. WILSON AND D. J. COOK, *A survey of unsupervised deep domain adaptation*, ACM Trans. Intell. Syst. Technol., 11 (2020).

[48] A. ĆIPRIJANOVIĆ, D. KAFKES, K. D. S. JENKINS, G. N. PERDUE, S. MADIREDDY, T. JOHNSTON, AND B. NORD, *Domain adaptation techniques for improved cross-domain study of galaxy mergers*, FERMILAB-CONF-20-582-SCD, (2020).

[49] A. ĆIPRIJANOVIĆ, D. KAFKES, G. SNYDER, F. J. SÁNCHEZ, G. N. PERDUE, K. PEDRO, B. NORD, S. MADIREDDY, AND S. M. WILD, *Deepadversaries: examining the robustness of deep learning models for galaxy morphology classification*, Machine Learning: Science and Technology, 3 (2022), p. 035007.