Towards Galaxy Foundation Models with Hybrid Contrastive Learning

Mike Walmsley [1] **Inigo Val Slijepcevic** [1] **Micah Bowles** [1] **Anna M. M. Scaife** [1]

Abstract

New astronomical tasks are often related to earlier tasks for which labels have already been collected. We adapt the contrastive framework BYOL to leverage those labels as a pretraining task while also enforcing augmentation invariance. For large-scale pretraining, we introduce GZ-Evo v0.1, a set of 96.5M volunteer responses for 552k galaxy images plus a further 1.34M comparable unlabelled galaxies. Most of the 206 GZ-Evo answers are unknown for any given galaxy, and so our pretraining task uses a Dirichlet loss that naturally handles unknown answers. GZ-Evo pretraining, with or without hybrid learning, improves on direct training even with plentiful downstream labels (+4% accuracy with 44k labels). Our hybrid pretraining/contrastive method further improves downstream accuracy vs. pretraining or contrastive learning, especially in the low-label transfer regime (+6% accuracy with 750 labels).

1. Introduction

Contrastive learning frameworks are typically evaluated by their ability to learn representations useful for downstream tasks given minimal labelled data (Le-Khac et al., 2020). All existing applications of contrastive learning to astronomical images follow this pattern (Hayat et al., 2021; Sarmiento et al., 2021; Stein et al., 2022). However, in practice, astronomers need not start from a blank slate.

New astronomical tasks are often related to earlier tasks for which labels have already been collected. In the context of galaxy morphology (i.e. visual appearance), new versions of the Galaxy Zoo citizen science project (Masters et al., 2019) gradually shift the questions asked of volunteers ('does this galaxy have a bar' to 'how large is the bar', 'is the galaxy merging' to 'what kind of merger', etc.). The responses to previous questions are therefore likely to be useful in predicting the answers to new questions.

We propose using the responses to all questions in all previous Galaxy Zoo versions to create generalizable representations useful for answering new questions, in a hybrid approach that jointly combines both contrastive learning and supervised pretraining. Specifically, in this work:

1. We assemble the largest and broadest galaxy morphology dataset to date, with 552k labelled and 1.34m unlabelled galaxies from five telescopes and four Galaxy Zoo campaigns

2. We adapt the contrastive framework BYOL (Grill et al., 2020) to additionally calculate a supervised regression loss that forces the representation to solve 206 galaxy morphology tasks while remaining invariant to contrastive augmentations.

We are motivated by CS literature emphasizing the practical value of broad (typically self-supervised) pretraining at scale (Hendrycks et al., 2019; Bommasani et al., 2021; Goyal et al., 2021; Azizi et al., 2021). We hope that our hybrid approach will ultimately create foundation models useful for as-yet-unknown galaxy morphology tasks.

2. Data

The key properties of telescope images are angular resolution (how sharp sources appear) and depth (how bright sources must be to distinguish themselves from the background). Images are grouped into surveys according to the telescope used and the operating conditions. Galaxies identified in those surveys may then be labelled by Galaxy Zoo volunteers - members of the public answering a series of questions about the features of each galaxy via a browser. Below, we briefly summarise the properties of the surveys from which draw our images and the Galaxy Zoo campaigns (project versions) that labelled them.

Galaxy Zoo 2 (GZ2, Willett et al. 2013) labelled images from the Sloan Digital Sky Survey DR7 (Abazajian et al., 2009). This survey was the first to which deep learning was applied (Dieleman et al., 2015) and remains a common reference dataset. GZ2 volunteers were asked whether galaxies had features including bars, bulges, and spiral arms. We include the 209,239 galaxies in the GZ2 main sample.

[*]Equal contribution [1]Department of Physics and Astronomy, University of Manchester, Manchester, UK. Correspondence to: Mike Walmsley <michael.walmsley@manchester.ac.uk>.

Galaxy Zoo Legacy Survey (LegS, Walmsley et al. in prep.) labels images from the DESI Legacy Imaging Surveys (Dey et al., 2019). The individual surveys are DECaLS, MzLS, and BASS, each producing comparable images. LegS images are deeper than those in GZ2, revealing fainter galaxy details. Most (305k) of our volunteer labels for this campaign are drawn from Galaxy Zoo DECaLS (Walmsley et al., 2022a) which labelled a non-random subset of DECaLS images. The remainder (60k) are DECaLS images labelled subsequently. Volunteers were asked similar questions to GZ2, with an additional focus on fainter details (e.g. is the bar strong or weak, is the galaxy disturbed). We also include the 1.34 million nearby (redshift $z < 0.1$, Zou et al. 2019) galaxies not labelled by volunteers as a large unlabelled subset of well-resolved galaxies for contrastive learning.

Galaxy Zoo Hubble (Willett et al., 2017) and Galaxy Zoo CANDELS (Simmons et al., 2017) labelled galaxy images from Hubble Space Telescope surveys (Grogin et al., 2011; Griffith et al., 2012). Despite the improved resolution of a space-based telescope, the target galaxies are far more distant (redshifts $1 \leq z \leq 3$ vs. $z \leq 0.4$ for LegS and $z \leq 0.15$ for GZ2) and hence appear less well-resolved (obscuring fine details) and relatively faint. However, these distant galaxies appear earlier in their evolution and show unique features such as starforming clumps. Volunteers were asked new questions about these clumps and fewer questions about detailed morphology. We include the 100k and 50k galaxies from the GZ Hubble and GZ CANDELS primary samples, respectively.

For our downstream task, we use volunteer labels from Galaxy Zoo Rings (Walmsley et al. in prep.). Volunteers labelled a selected subset of LegS galaxies as being ringed or not using the Galaxy Zoo Mobile swipe interface. 10 volunteers label each image and the majority response is used as a binary label. This is a new task - Galaxy Zoo volunteers do not label galaxies as ringed in any of our pretraining campaigns. Rings includes 93k galaxies and is roughly balanced between rings and non-rings.

We group the 552k galaxies labelled during the campaigns above with the 1.34m unlabelled nearby LegS galaxies and name the aggregate dataset **GZ-Evo v0.1**. GZ-Evo is the largest collection of volunteer galaxy labels to date, with 96.5 million volunteer answers. GZ-Evo is publicly available at www.github.com/mwalmsley/pytorch-galaxy-datasets.

3. Method

GZ-Evo has two properties that motivate our approach. First, the questions volunteers were asked vary between campaigns according to the science goals of the organisers and the nature of the images (Sec. 2). Even where questions

do not change, subtle details like the tutorial and the answer icons cause significant distribution shifts (Walmsley et al., 2022a). We therefore choose to treat answers as non-exchangeable between campaigns. Across GZ-Evo, there are 206 possible answers to 65 questions. We hypothesize that learning to jointly predict all answers may help build a representation that generalizes to new tasks, and do so using a custom Dirichlet loss function (Sec. 3.1).

Second, as often in astronomy, we have plentiful (1.34m) unlabelled images. We would like to exploit these images to help solve downstream tasks while also using the substantial (552k) set of labelled images. We therefore introduce a hybrid pretraining/contrastive approach (Sec. 3.2).

3.1. Multi-Question Multi-Campaign Dirichlet Loss

Our aim to learn from volunteers answering different questions on different images, where any particular image has only a small subset of possible questions answered.

Walmsley et al. (2022a) introduced the Dirichlet loss function in the context of managing heteroskedastic votes for questions in the same Galaxy Zoo campaign.

$$\mathcal{L}_q = \int \text{Multi}(k|\rho, N)\text{Dirichlet}(\rho|\alpha)d\alpha \qquad (1)$$

where, for some target question q, k is the (vector) counts of responses (successes) of each answer, N is the total number of responses (trials) to all answers, and ρ is the (vector) probabilities of a volunteer giving each answer. ρ is drawn from $\text{Dirichlet}(\rho|\alpha)$, where the model predicts the Dirichlet concentrations α. Multinomial and Dirichlet distributions are conjugates and hence the integral is analytic. Intuitively, this loss corresponds to the odds of observing k heads (votes for an answer) after N coin flips (volunteers asked) assuming a (model-predicted) distribution for the bias of that coin. The Dirichlet-Multinomial distribution allows models to flexibly express uncertainty through wider posteriors and confidence through narrower posteriors.

Assuming answers to different questions are independent, the loss may be applied to multiple questions via the sum

$$\log \mathcal{L} = \sum_q \mathcal{L}_q(k_q, N_q, f_q^w) \qquad (2)$$

where, for question q, N_q is the total answers, k_q is the observed votes for each answer, and f_q^w is the values of the output units corresponding to those answers (which we interpret as the Dirichlet α parameters in Eqn. 1).

To extend to multiple campaigns, we note that this loss naturally handles questions with no answers as $p(a = 0|\alpha, N = 0) = 1$ for all α and hence $\frac{\partial \mathcal{L}}{\partial \alpha} = 0$, meaning unanswered questions do not affect the training gradients. We can therefore construct a multi-campaign vote count vector K where

K_i is the votes for answer i and i indexes all answers across all questions *across all campaigns*. For a galaxy labelled in any single campaign, K_i is 0 for any answer a_i to any question not asked in that campaign. Every answer is always predicted but the prediction only affects training if votes for that answer are non-zero. Intuitively, this corresponds to having zero recorded votes to questions not asked. Questions are effectively treated as orthogonal prediction tasks using the same representation.

3.2. Adapting BYOL with Supervised Projection Head

We would like to guide the contrastive learning process using Galaxy Zoo labels where available (via our loss function above). Unfortunately, all supervised contrastive frameworks we identified (Khosla et al., 2020; Tian et al., 2020; Wang et al., 2021; Zhao et al., 2021) assume a classification context, often mirroring the positive/negative paired views concept of e.g. Sohn (2016) with a contrastive supervised loss that rewards/penalizes similar representations for pairs of the same/different class. Our volunteer responses are of the form k of N volunteers gave answer a to question q and are not easily reduced to e.g. majority class labels due to heteroskedasticity (typically $5 < N < 40$) and the overlapping nature of the questions. We instead use BYOL (Grill et al., 2020), an unsupervised contrastive framework which relies only on views of the same object, and extend it to a supervised context.

Briefly, BYOL consists of two encoders, where the target encoder f_ξ is an exponential weighted average of the online encoder f_θ. Each encoder projects strongly-augmented image views to representations y_ξ and y_θ and then (via fully-connected heads) to projections z_ξ and z_θ. The online encoder has a further fully-connected head $q_{\theta,\,\mathrm{con}}$ which predicts the projection of the target encoder $q_{\theta,\,\mathrm{con}}(z_\theta) = \hat{z}_\xi$. To train, the parameters θ of the online encoder are updated to best predict the representation of the target encoder when both encoders are fed different strongly-augmented views of the same image. The parameters ξ of the target encoder are then updated to the new rolling average of θ.

To guide the representation learned by BYOL, we add a new fully-connected prediction head to the online encoder. This head $q_{\theta,\,\mathrm{sup}}(y_\theta) = \hat{K}$ predicts the volunteer votes K from the online-encoded representation, analogously to the final layers of a CNN in the standard supervised context. The head outputs are constrained via sigmoid activation to the range (1, 100) such that they may be interpreted as the parameters of a Dirichlet distribution (Sec. 3.1).

We combine the contrastive prediction head loss with the new supervised prediction head loss via a normalised weighted sum, with hyperparameter λ weighting contrastive

invariance and supervised predictive performance.

$$\mathcal{L} = \mathcal{L}_{\mathrm{con}}(q_{\theta,\,\mathrm{con}}(z_\theta, z_\xi)) + \lambda \mathcal{L}_{\mathrm{Dirichlet}}(q_{\theta,\,\mathrm{sup}}(y_\theta), K)/\mathcal{L}_{\mathrm{con}} \tag{3}$$

3.3. Training and Evaluation Details

Our aim is to learn a representation that maximises performance on the downstream task (classifying galaxies as ringed or not-ringed, with labels from Galaxy Zoo Mobile, Sec. 2), without using downstream labels until the representation is trained.

We experiment with three representation-learning approaches on GZ-Evo: purely-contrastive BYOL ('Contrastive'); purely-supervised pretraining on all GZ-Evo answers with our Dirichlet loss function ('Pretrained'); and our hybrid contrastive/supervised approach ('Hybrid', BYOL with the added supervised prediction head and Dirichlet loss function). For each experiment, we evaluate the quality of our learned representation by training a linear classifier to solve the downstream rings task on that (now frozen) representation. Our baseline is directly training a supervised model on the downstream labels ('Direct').

The Pretrained and Direct classifiers use Zoobot (Walmsley et al., 2022b), a version of EfficientNetB0 (Tan & Le, 2019) with the Dirichlet loss above and standard (as opposed to contrastive-style) augmentations of flips, rotations, and non-central crops. Pretrained learns from all 206 GZ-Evo answers and the representation is used as input features for the linear classifier, while Direct is directly trained on only the ring labels. For BYOL's encoder (with or without the Dirichlet supervised head), we experiment with EfficientNetB0 and ResNet18 (He et al., 2016).

All models use Adam (Kingma & Ba, 2015) with standard hyperparameters (learning rate 0.001, $\beta_1 = 0.9$, $\beta_2 = 0.999$). We divide GZ-Evo into fixed 70/10/20 train/val/test sets (for rings, 50/50 train/test) and reserve the test sets for future work. We evaluate downstream linear accuracy with 5-fold validation on the ring train set. All models are trained to convergence, as measured by supervised validation loss (on rings, for direct training, or on GZ-Evo, for pretraining) or contrastive loss. All models use color images downscaled to 128x128 pixels.

4. Results & Discussion

Our hybrid pretrained/contrastive approach matches or outperforms purely-contrastive and purely-pretrained approaches on all downstream ring dataset sizes (Fig 1). In the low-label transfer regime (761 downstream labels), our hybrid approach achieves a 6% accuracy increase (relative error reduction of 28%) over pretraining or contrastive learn-
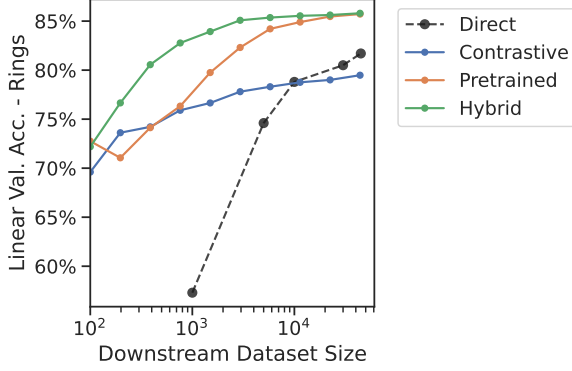
Figure 1. For our downstream task, classifying ring galaxies, our hybrid pretrained/contrastive approach matches or outperforms purely-pretrained and purely-contrastive approaches at all downstream dataset sizes.

ing alone. As additional downstream labels are collected, hybrid training provides consistent but decreasing improvements, with pretraining eventually matching hybrid learning at the limit of our dataset size (44k). Both methods outperform direct training. Purely-contrastive learning provides a strong baseline, outperforming direct training up to approx. 7k downstream labels, but is not competitive with pretraining.

Fig. 2 compares the downstream accuracy and contrastive loss of the purely-contrastive vs. hybrid approach. The purely-contrastive approach better solves the contrastive problem (achieving a lower contrastive loss) but the learned representation is worse for solving the downstream task. Considering the encoder, downstream accuracy is particularly disconnected from contrastive loss with EfficientNetB0 - unless our hybrid second head is added, in which case downstream accuracy is similar to ResNet18. This emphasizes that good contrastive solutions do not necessarily imply good downstream accuracy, and that hybrid supervised training can help guide those solutions.

5. Conclusion

New galaxy morphology tasks are often related to earlier tasks for which labels have already been collected. We experiment with pretraining on these labels, in combination with contrastive learning, to solve new downstream tasks. We first introduce GZ-Evo, a collection of 96.5M volunteer responses on the morphology of 552K galaxy images with a further 1.34M comparable unlabelled images. The responses span a wide range of related prediction tasks - there are 206 possible telescope/instruction/answer combinations, but any given galaxy only has responses for a minority of these. We jointly learn to predict all tasks using a Dirichlet
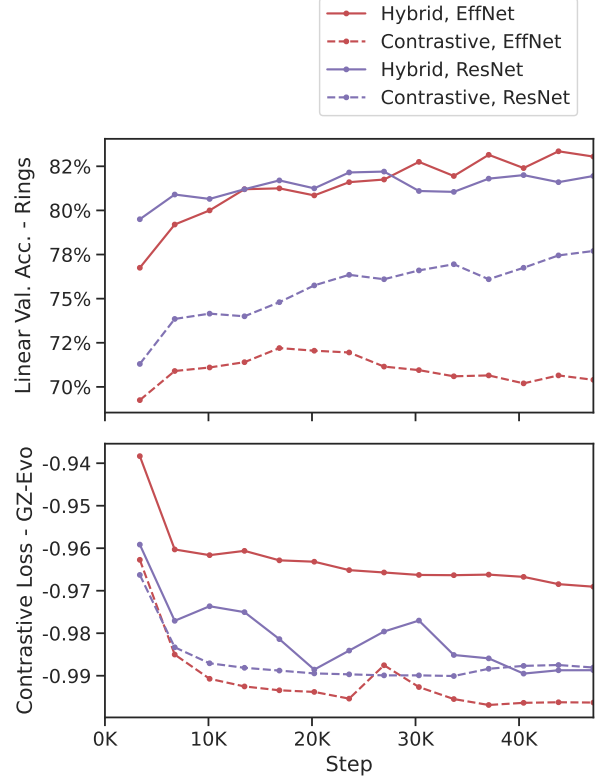

Figure 2. The representation learned by our hybrid pretrained/contrastive approach achieves higher downstream accuracy when classifying ring galaxies than a purely-constrastive approach (BYOL). The purely-contrastive approach achieves a lower (better) contrastive loss but *worse* downstream performance - particularly when using EffNetB0 as an encoder. With the hybrid approach, ResNet18 performs similarly to EffNet.

loss function that naturally handles missing responses. Pretraining EfficientNetB0 on all GZ-Evo tasks then applying a simple linear classifier to the resulting representation outperforms direct training on our downstream task (identifying ringed galaxies), even with 44k downstream labels.

We then add a new supervised prediction head to the BYOL contrastive learning framework, allowing us to learn from both the supervised pretraining task (where GZ-Evo labels exist) and the standard contrastive invariance loss (all galaxies). Our hybrid approach matches or outperforms purely-pretraining or purely-contrastive approaches, particularly in the low-labels regime (approx. 750 downstream labels).

We hope that this simple hybrid method will help astronomers leverage the last decade of volunteer effort to answer new science questions.

6. Acknowledgements

References

Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., Allam, S. S., Prieto, C. A., An, D., Anderson, K. S. J., Anderson, S. F., Annis, J., Bahcall, N. A., Bailer-Jones, C. A. L., Barentine, J. C., Bassett, B. A., Becker, A. C., Beers, T. C., Bell, E. F., Belokurov, V., Berlind, A. A., Berman, E. F., Bernardi, M., Bickerton, S. J., Bizyaev, D., Blakeslee, J. P., Blanton, M. R., Bochanski, J. J., Boroski, W. N., Brewington, H. J., Brinchmann, J., Brinkmann, J., Brunner, R. J., Budavári, T., Carey, L. N., Carliles, S., Carr, M. A., Castander, F. J., Cinabro, D., Connolly, A. J., Csabai, I., Cunha, C. E., Czarapata, P. C., Davenport, J. R. A., de Haas, E., Dilday, B., Doi, M., Eisenstein, D. J., Evans, M. L., Evans, N. W., Fan, X., Friedman, S. D., Frieman, J. A., Fukugita, M., Gänsicke, B. T., Gates, E., Gillespie, B., Gilmore, G., Gonzalez, B., Gonzalez, C. F., Grebel, E. K., Gunn, J. E., Györy, Z., Hall, P. B., Harding, P., Harris, F. H., Harvanek, M., Hawley, S. L., Hayes, J. J. E., Heckman, T. M., Hendry, J. S., Hennessy, G. S., Hindsley, R. B., Hoblitt, J., Hogan, C. J., Hogg, D. W., Holtzman, J. A., Hyde, J. B., Ichikawa, S.-i., Ichikawa, T., Im, M., Ivezic, Z., Jester, S., Jiang, L., Johnson, J. A., Jorgensen, A. M., Jurić, M., Kent, S. M., Kessler, R., Kleinman, S. J., Knapp, G. R., Konishi, K., Kron, R. G., Krzesinski, J., Kuropatkin, N., Lampeitl, H., Lebedeva, S., Lee, M. G., Lee, Y. S., Leger, R. F., Lépine, S., Li, N., Lima, M., Lin, H., Long, D. C., Loomis, C. P., Loveday, J., Lupton, R. H., Magnier, E., Malanushenko, O., Malanushenko, V., Mandelbaum, R., Margon, B., Marriner, J. P., Martínez-Delgado, D., Matsubara, T., McGehee, P. M., McKay, T. A., Meiksin, A., Morrison, H. L., Mullally, F., Munn, J. A., Murphy, T., Nash, T., Nebot, A., Neilsen, E. H., Newberg, H. J., Newman, P. R., Nichol, R. C., Nicinski, T., Nieto-Santisteban, M., Nitta, A., Okamura, S., Oravetz, D. J., Ostriker, J. P., Owen, R., Padmanabhan, N., Pan, K., Park, C., Pauls, G., Peoples, J., Percival, W. J., Pier, J. R., Pope, A. C., Pourbaix, D., Price, P. A., Purger, N., Quinn, T., Raddick, M. J., Fiorentin, P. R., Richards, G. T., Richmond, M. W., Riess, A. G., Rix, H.-W., Rockosi, C. M., Sako, M., Schlegel, D. J., Schneider, D. P., Scholz, R.-D., Schreiber, M. R., Schwope, A. D., Seljak, U., Sesar, B., Sheldon, E., Shimasaku, K., Sibley, V. C., Simmons, A. E., Sivarani, T., Smith, J. A., Smith, M. C., Smolčić, V., Snedden, S. A., Stebbins, A., Steinmetz, M., Stoughton, C., Strauss, M. A., SubbaRao, M., Suto, Y., Szalay, A. S., Szapudi, I., Szkody, P., Tanaka, M., Tegmark, M., Teodoro, L. F. A., Thakar, A. R., Tremonti, C. A., Tucker, D. L., Uomoto, A., Vanden Berk, D. E., Vandenberg, J., Vidrih, S., Vogeley, M. S., Voges, W., Vogt, N. P., Wadadekar, Y., Watters, S., Weinberg, D. H., West, A. A., White, S. D. M., Wilhite, B. C., Wonders, A. C., Yanny, B., Yocum, D. R., York, D. G., Zehavi, I., Zibetti, S., and Zucker, D. B. the Seventh Data Release of the Sloan Digital Sky Survey. *The Astrophysical Journal Supplement Series*, 182(2):543–558, 2009. ISSN 0067-0049. doi: 10.1088/0067-0049/182/2/543. URL http://stacks.iop.org/0067-0049/182/i=2/a=543?key=crossref.bc07496fb06b943bcf82755687fa84b4.

Azizi, S., Mustafa, B., Ryan, F., Beaver, Z., Freyberg, J., Deaton, J., Loh, A., Karthikesalingam, A., Kornblith, S., Chen, T., Natarajan, V., and Norouzi, M. Big Self-Supervised Models Advance Medical Image Classification. In *Proceedings of the IEEE International Conference on Computer Vision*, pp. 3458–3468, 1 2021. ISBN 9781665428125. doi: 10.1109/ICCV48922.2021.00346. URL http://arxiv.org/abs/2101.05224.

Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M. S., Bohg, J., Bosselut, A., Brunskill, E., Brynjolfsson, E., Buch, S., Card, D., Castellon, R., Chatterji, N., Chen, A., Creel, K., Davis, J. Q., Demszky, D., Donahue, C., Doumbouya, M., Durmus, E., Ermon, S., Etchemendy, J., Ethayarajh, K., Fei-Fei, L., Finn, C., Gale, T., Gillespie, L., Goel, K., Goodman, N., Grossman, S., Guha, N., Hashimoto, T., Henderson, P., Hewitt, J., Ho, D. E., Hong, J., Hsu, K., Huang, J., Icard, T., Jain, S., Jurafsky, D., Kalluri, P., Karamcheti, S., Keeling, G., Khani, F., Khattab, O., Koh, P. W., Krass, M., Krishna, R., Kuditipudi, R., Kumar, A., Ladhak, F., Lee, M., Lee, T., Leskovec, J., Levent, I., Li, X. L., Li, X., Ma, T., Malik, A., Manning, C. D., Mirchandani, S., Mitchell, E., Munyikwa, Z., Nair, S., Narayan, A., Narayanan, D., Newman, B., Nie, A., Niebles, J. C., Nilforoshan, H., Nyarko, J., Ogut, G., Orr, L., Papadimitriou, I., Park, J. S., Piech, C., Portelance, E., Potts, C., Raghunathan, A., Reich, R., Ren, H., Rong, F., Roohani, Y., Ruiz, C., Ryan, J., Ré, C., Sadigh, D., Sagawa, S., Santhanam, K., Shih, A., Srinivasan, K., Tamkin, A., Taori, R., Thomas, A. W., Tramèr, F., Wang, R. E., Wang, W., Wu, B., Wu, J., Wu, Y., Xie, S. M., Yasunaga, M.,

You, J., Zaharia, M., Zhang, M., Zhang, T., Zhang, X., Zhang, Y., Zheng, L., Zhou, K., and Liang, P. On the Opportunities and Risks of Foundation Models, 8 2021. URL http://arxiv.org/abs/2108.07258.

Dey, A., Schlegel, D. J., Lang, D., Blum, R., Burleigh, K., Fan, X., Findlay, J. R., Finkbeiner, D., Herrera, D., Juneau, S., Landriau, M., Levi, M., McGreer, I., Meisner, A., Myers, A. D., Moustakas, J., Nugent, P., Patej, A., Schlafly, E. F., Walker, A. R., Valdes, F., Weaver, B. A., Yèche, C., Zou, H., Zhou, X., Abareshi, B., Abbott, T. M., Abolfathi, B., Aguilera, C., Alam, S., Allen, L., Alvarez, A., Annis, J., Ansarinejad, B., Aubert, M., Beechert, J., Bell, E. F., BenZvi, S. Y., Beutler, F., Bielby, R. M., Bolton, A. S., Briceño, C., Buckley-Geer, E. J., Butler, K., Calamida, A., Carlberg, R. G., Carter, P., Casas, R., Castander, F. J., Choi, Y., Comparat, J., Cukanovaite, E., Delubac, T., DeVries, K., Dey, S., Dhungana, G., Dickinson, M., Ding, Z., Donaldson, J. B., Duan, Y., Duckworth, C. J., Eftekharzadeh, S., Eisenstein, D. J., Etourneau, T., Fagrelius, P. A., Farihi, J., Fitzpatrick, M., Font-Ribera, A., Fulmer, L., Gänsicke, B. T., Gaztanaga, E., George, K., Gerdes, D. W., Gontcho, S. G. A., Gorgoni, C., Green, G., Guy, J., Harmer, D., Hernandez, M., Honscheid, K., Huang, L., James, D., Jannuzi, B. T., Jiang, L., Joyce, R., Karcher, A., Karkar, S., Kehoe, R., Jean-Paul, K., Kueter-Young, A., Lan, T. W., Lauer, T., Le Guillou, L., van Suu, A. L., Lee, J. H., Lesser, M., Levasseur, L. P., Li, T. S., Mann, J. L., Marshall, B., Martínez-Vázquez, C. E., Martini, P., du Mas des Bourboux, H., McManus, S., Meier, T. G., Ménard, B., Metcalfe, N., Muñoz-Gutiérrez, A., Najita, J., Napier, K., Narayan, G., Newman, J. A., Nie, J., Nord, B., Norman, D. J., Olsen, K. A., Paat, A., Palanque-Delabrouille, N., Peng, X., Poppett, C. L., Poremba, M. R., Prakash, A., Rabinowitz, D., Raichoor, A., Rezaie, M., Robertson, A. N., Roe, N. A., Ross, A. J., Ross, N. P., Rudnick, G., Safonova, S., Saha, A., Javier Sánchez, F., Savary, E., Schweiker, H., Scott, A., Seo, H. J., Shan, H., Silva, D. R., Slepian, Z., Soto, C., Sprayberry, D., Staten, R., Stillman, C. M., Stupak, R. J., Summers, D. L., Tie, S. S., Tirado, H., Vargas-Magaña, M., Katherina Vivas, A., Wechsler, R. H., Williams, D., Yang, J., Yang, Q., Yapici, T., Zaritsky, D., Zenteno, A., Zhang, K., Zhang, T., Zhou, R., and Zhou, Z. Overview of the DESI legacy imaging surveys. *The Astronomical Journal*, 157(5):168, 4 2019. ISSN 23318422. doi: 10.3847/1538-3881/ab089d. URL http://arxiv.org/abs/1804.08657.

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, 2015. ISSN 0035-8711. doi: 10.1093/mnras/stv632. URL http://arxiv.org/abs/1503.07077.

Goyal, P., Caron, M., Lefaudeux, B., Xu, M., Wang, P., Pai, V., Singh, M., Liptchinsky, V., Misra, I., Joulin, A., and Bojanowski, P. Self-supervised Pre-training of Visual Features in the Wild, 2021. URL http://arxiv.org/abs/2103.01988https://t.co/HNcnBgves7?amp=1&s=09.

Griffith, R. L., Cooper, M. C., Newman, J. A., Moustakas, L. A., Stern, D., Comerford, J. M., Davis, M., Lotz, J. M., Barden, M., Conselice, C. J., Capak, P. L., Faber, S. M., Kirkpatrick, J. D., Koekemoer, A. M., Koo, D. C., Noeske, K. G., Scoville, N., Sheth, K., Shopbell, P., Willmer, C. N., and Weiner, B. The advanced camera for surveys general catalog: Structural parameters for approximately half a million galaxies. *Astrophysical Journal, Supplement Series*, 200(1), 2012. ISSN 00670049. doi: 10.1088/0067-0049/200/1/9.

Grill, J. B., Strub, F., Altché, F., Tallec, C., Richemond, P. H., Buchatskaya, E., Doersch, C., Pires, B. A., Guo, Z. D., Azar, M. G., Piot, B., Kavukcuoglu, K., Munos, R., and Valko, M. Bootstrap your own latent a new approach to self-supervised learning. In *Advances in Neural Information Processing Systems*, San Diego, 6 2020. Neural information processing systems foundation. ISBN 2006.07733v3. URL https://arxiv.org/abs/2006.07733v3.

Grogin, N. A., Kocevski, D. D., Faber, S. M., Ferguson, H. C., Koekemoer, A. M., Riess, A. G., Acquaviva, V., Alexander, D. M., Almaini, O., Ashby, M. L., Barden, M., Bell, E. F., Bournaud, F., Brown, T. M., Caputi, K. I., Casertano, S., Cassata, P., Castellano, M., Challis, P., Chary, R. R., Cheung, E., Cirasuolo, M., Conselice, C. J., Cooray, A. R., Croton, D. J., Daddi, E., Dahlen, T., Davé, R., De Mello, D. F., Dekel, A., Dickinson, M., Dolch, T., Donley, J. L., Dunlop, J. S., Dutton, A. A., Elbaz, D., Fazio, G. G., Filippenko, A. V., Finkelstein, S. L., Fontana, A., Gardner, J. P., Garnavich, P. M., Gawiser, E., Giavalisco, M., Grazian, A., Guo, Y., Hathi, N. P., Häussler, B., Hopkins, P. F., Huang, J. S., Huang, K. H., Jha, S. W., Kartaltepe, J. S., Kirshner, R. P., Koo, D. C., Lai, K., Lee, K. S., Li, W., Lotz, J. M., Lucas, R. A., Madau, P., McCarthy, P. J., McGrath, E. J., McIntosh, D. H., McLure, R. J., Mobasher, B., Moustakas, L. A., Mozena, M., Nandra, K., Newman, J. A., Niemi, S. M., Noeske, K. G., Papovich, C. J., Pentericci, L., Pope, A., Primack, J. R., Rajan, A., Ravindranath, S., Reddy, N. A., Renzini, A., Rix, H. W., Robaina, A. R., Rodney, S. A., Rosario, D. J., Rosati, P., Salimbeni, S., Scarlata, C., Siana, B., Simard, L., Smidt, J., Somerville, R. S., Spinrad, H., Straughn, A. N., Strolger, L. G., Telford, O., Teplitz, H. I., Trump, J. R., Van Der Wel, A., Villforth, C., Wechsler, R. H., Weiner, B. J., Wiklind, T., Wild, V., Wilson, G., Wuyts, S., Yan, H. J., and Yun,

M. S. Candels: The Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey. *Astrophysical Journal, Supplement Series*, 197(2):35, 12 2011. ISSN 00670049. doi: 10.1088/0067-0049/197/2/35. URL https://iopscience.iop.org/article/10.1088/0067-0049/197/2/35https://iopscience.iop.org/article/10.1088/0067-0049/197/2/35/meta.

Hayat, M. A., Stein, G., Harrington, P., Lukić, Z., and Mustafa, M. Self-supervised Representation Learning for Astronomical Images. *The Astrophysical Journal Letters*, 911(2):L33, 4 2021. ISSN 2041-8205. doi: 10.3847/2041-8213/abf2c7. URL https://doi.org/10.3847/2041-8213/abf2c7.

He, K., Zhang, X., Ren, S., and Sun, J. Deep Residual Learning for Image Recognition. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, number 3, pp. 770–778. IEEE Computer Society, 12 2016. ISBN 9781467388504. doi: 10.1109/CVPR.2016.90. URL http://arxiv.org/abs/1512.03385.

Hendrycks, D., Mazeika, M., Kadavath, S., and Song, D. Using self-supervised learning can improve model robustness and uncertainty, 6 2019. ISSN 23318422. URL http://arxiv.org/abs/1906.12340.

Khosla, P., Teterwak, P., Wang, C., Sarna, A., Tian, Y., Isola, P., Maschinot, A., Liu, C., and Krishnan, D. Supervised contrastive learning. In *Advances in Neural Information Processing Systems*. Neural information processing systems foundation, 4 2020. URL https://arxiv.org/abs/2004.11362v5.

Kingma, D. P. and Ba, J. L. Adam: A method for stochastic optimization. In *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*. International Conference on Learning Representations, ICLR, 12 2015.

Le-Khac, P. H., Healy, G., and Smeaton, A. F. Contrastive Representation Learning: A Framework and Review. *IEEE Access*, 8:193907–193934, 2020. ISSN 21693536. doi: 10.1109/access.2020.3031549. URL http://arxiv.org/abs/2010.05113http://dx.doi.org/10.1109/ACCESS.2020.3031549https://arxiv.org/abs/2010.05113?s=09.

Masters, K. L., Lintott, C. J., Hart, R. E., Kruk, S. J., Smethurst, R. J., Casteels, K. V., Keel, W. C., Simmons, B. D., Stanescu, D. O., Tate, J., and Tomi, S. Galaxy Zoo: Unwinding the winding problem - Observations of spiral bulge prominence and arm pitch angles suggest local spiral galaxies are winding. *Monthly Notices of the Royal Astronomical Society*, 487(2):1808–1820, 4 2019.

ISSN 13652966. doi: 10.1093/mnras/stz1153. URL http://arxiv.org/abs/1904.11436.

Sarmiento, R., Huertas-Company, M., Knapen, J. H., Sánchez, S. F., Domínguez Sánchez, H., Drory, N., and Falcón-Barroso, J. Capturing the Physics of MaNGA Galaxies with Self-supervised Machine Learning. *The Astrophysical Journal*, 921(2):177, 4 2021. ISSN 0004-637X. doi: 10.3847/1538-4357/ac1dac. URL http://arxiv.org/abs/2104.08292.

Simmons, B. D., Lintott, C., Willett, K. W., Masters, K. L., Kartaltepe, J. S., Häußler, B., Kaviraj, S., Krawczyk, C., Kruk, S. J., McIntosh, D. H., Smethurst, R. J., Nichol, R. C., Scarlata, C., Schawinski, K., Conselice, C. J., Almaini, O., Ferguson, H. C., Fortson, L., Hartley, W., Kocevski, D., Koekemoer, A. M., Mortlock, A., Newman, J. A., Bamford, S. P., Grogin, N. A., Lucas, R. A., Hathi, N. P., McGrath, E., Peth, M., Pforr, J., Rizer, Z., Wuyts, S., Barro, G., Bell, E. F., Castellano, M., Dahlen, T., Dekel, A., Ownsworth, J., Faber, S. M., Finkelstein, S. L., Fontana, A., Galametz, A., Grützbauch, R., Koo, D., Lotz, J., Mobasher, B., Mozena, M., Salvato, M., and Wiklind, T. Galaxy Zoo: Quantitative visual morphological classifications for 48 000 galaxies from CANDELS. *Monthly Notices of the Royal Astronomical Society*, 464(4):4420–4447, 2017. ISSN 13652966. doi: 10.1093/mnras/stw2587.

Sohn, K. Improved deep metric learning with multi-class N-pair loss objective. In *Advances in Neural Information Processing Systems*, volume 29, pp. 1857–1865, 2016.

Stein, G., Blaum, J., Harrington, P., Medan, T., and Lukic, Z. Mining for strong gravitational lenses with self-supervised learning. *The Astrophysical Journal*, 932 (107), 2022. URL http://arxiv.org/abs/2110.00023.

Tan, M. and Le, Q. V. EfficientNet: Rethinking model scaling for convolutional neural networks. In *36th International Conference on Machine Learning, ICML 2019*, pp. 10691–10700, Long Beach, CA, 5 2019. Proceedings of Machine Learning Research. ISBN 9781510886988. URL http://arxiv.org/abs/1905.11946.

Tian, Y., Sun, C., Poole, B., Krishnan, D., Schmid, C., and Isola, P. What makes for good views for contrastive learning? In *Advances in Neural Information Processing Systems*. Neural information processing systems foundation, 5 2020. URL https://arxiv.org/abs/2005.10243v3.

Walmsley, M., Lintott, C., Tobias, G., Kruk, S. J., Krawczyk, C., Willett, K., Bamford, S., Keel, W., Kelvin, L. S., Fortson, L., Masters, K., Mehta, V., Simmons, B., Smethurst, R. J., Baeten, E. M. L., and Macmillan, C. Galaxy

Zoo DECaLS: Detailed Visual Morphology Measurements from Volunteers and Deep Learning for 314,000 Galaxies. *Monthly Notices of the Royal Astronomical Society*, 509(3):3966–3988, 12 2022a. URL https://arxiv.org/abs/2102.08414.

Walmsley, M., Scaife, A. M. M., Lintott, C., Lochner, M., Etsebeth, V., Géron, T., Dickinson, H., Fortson, L., Kruk, S., Masters, K. L., Mantha, K. B., and Simmons, B. D. Practical galaxy morphology tools from deep supervised representation learning. *Monthly Notices of the Royal Astronomical Society*, 513(2):1581–1599, 5 2022b. ISSN 0035-8711. doi: 10.1093/mnras/stac525. URL https://academic.oup.com/mnras/article/513/2/1581/6539338.

Wang, P., Han, K., Wei, X. S., Zhang, L., and Wang, L. Contrastive Learning based Hybrid Networks for Long-Tailed Image Classification. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 943–952. IEEE Computer Society, 3 2021. ISBN 9781665445092. doi: 10.1109/CVPR46437.2021.00100. URL https://arxiv.org/abs/2103.14267v1.

Willett, K. W., Lintott, C. J., Bamford, S. P., Masters, K. L., Simmons, B. D., Casteels, K. R. V., Edmondson, E. M., Fortson, L. F., Kaviraj, S., Keel, W. C., Melvin, T., Nichol, R. C., Jordan Raddick, M., Schawinski, K., Simpson, R. J., Skibba, R. A., Smith, A. M., and Thomas, D. Galaxy Zoo 2: Detailed morphological classifications for 304 122 galaxies from the sloan digital sky survey. *Monthly Notices of the Royal Astronomical Society*, 435(4):2835–2860, 2013. ISSN 00358711. doi: 10.1093/mnras/stt1458.

Willett, K. W., Galloway, M. A., Bamford, S. P., Lintott, C. J., Masters, K. L., Scarlata, C., Simmons, B. D., Beck, M., Cardamone, C. N., Cheung, E., Edmondson, E. M., Fortson, L. F., Griffith, R. L., Häußler, B., Han, A., Hart, R., Melvin, T., Parrish, M., Schawinski, K., Smethurst, R. J., and Smith, A. M. Galaxy Zoo: Morphological classifications for 120 000 galaxies in HST legacy imaging. *Monthly Notices of the Royal Astronomical Society*, 464(4):4176–4203, 2 2017. ISSN 13652966. doi: 10.1093/mnras/stw2568. URL https://academic.oup.com/mnras/article-lookup/doi/10.1093/mnras/stw2568.

Zhao, X., Vemulapalli, R., Mansfield, P. A., Gong, B., Green, B., Shapira, L., and Wu, Y. Contrastive Learning for Label Efficient Semantic Segmentation. In *Proceedings of the IEEE International Conference on Computer Vision*, pp. 10603–10613. Institute of Electrical and Electronics Engineers (IEEE), 12 2021. ISBN 9781665428125. doi: 10.1109/ICCV48922.2021.01045. URL https://arxiv.org/abs/2012.06985v4.

Zou, H., Gao, J., Zhou, X., and Kong, X. Photometric Redshifts and Stellar Masses for Galaxies from the DESI Legacy Imaging Surveys. *The Astrophysical Journal Supplement Series*, 242(1):8, 2019. ISSN 0067-0049. doi: 10.3847/1538-4365/ab1847. URL http://dx.doi.org/10.3847/1538-4365/ab1847.