Detecting Tidal Features using Self-Supervised Representation Learning

Alice Desmons [1] **Sarah Brough** [1] **Francois Lanusse** [2]

Abstract

Low surface brightness substructures around galaxies, known as tidal features, are a valuable tool in the detection of past or ongoing galaxy mergers. Their properties can answer questions about the progenitor galaxies involved in the interactions. This paper presents promising results from a self-supervised machine learning model, trained on data from the Ultradeep layer of the Hyper Suprime-Cam Subaru Strategic Program optical imaging survey, designed to automate the detection of tidal features. We find that self-supervised models are capable of detecting tidal features and that our model outperforms previous automated tidal feature detection methods, including a fully supervised model. The previous state of the art method achieved 76% completeness for 22% contamination, while our model achieves considerably higher (96%) completeness for the same level of contamination.

1. Introduction

The currently accepted model of the Universe, known as the Lambda Cold Dark Matter (ΛCDM) Cosmological Model, postulates that galaxies evolve through a process which is referred to as the 'hierarchical merger model', wherein the growth of the universe's highest-mass galaxies is dominated by merging with lower-mass galaxies (e.g. Lacey and Cole 1994; Cole et al. 2000; Robotham et al. 2014; Martin et al. 2018). During the merging process, the extreme gravitational forces involved cause stellar material to be pulled out from the galaxies, forming diffuse non-uniform regions of stars in the outskirts of the galaxies, known as tidal features. These tidal features contain information about the merging history of the galaxy, and can thus be used to study the galaxy evolution process.

In order to draw accurate and statistically robust conclusions about this evolution process, we require a large sample of galaxies exhibiting tidal features. One thing that makes this difficult is the extremely low surface brightness of tidal features, which can easily reach $\mu_r \geq 27$ mag arcsec^{-2}. With the next generation of wide-field optical imaging surveys reaching new limiting depths, such as the Vera C Rubin Observatory's Legacy Survey of Space and Time (LSST; Ivezić et al. 2019) which is predicted to reach $\mu_r \sim 30.1$ mag arcsec^{-2} (Martin et al., 2022), assembling a statistically significant sample of galaxies with tidal features is becoming more feasible. One challenge associated with surveys like LSST, due to commence in 2024 and run for 10 years, is the amount of data predicted to be released, with LSST predicted to output over 500 petabytes of imaging data including billions of galaxies (Ivezić et al., 2019). Current tidal feature detection and classification is primarily achieved through visual identification (e.g. Tal et al. 2009; Sheen et al. 2012; Atkinson et al. 2013; Hood et al. 2018; Bílek et al. 2020; Martin et al. 2022), but this amount of data is virtually impossible to classify visually by humans, even using large community based projects such as Galaxy Zoo (Lintott et al., 2008; Darg et al., 2010), and hence we are in urgent need of a tool that can automate this classification task and isolate galaxies with tidal features.

With the promising recent results of machine learning in galaxy classification tasks (e.g. Hocking et al. 2018; Diaz et al. 2019; Pearson et al. 2019; Snyder et al. 2019; Walmsley et al. 2019; Cavanagh and Bekki 2020; Martin et al. 2020), we turn to machine learning to construct a model which can take galaxy images as input, convert them into representations - low-dimensional maps which preserve the important information in the image - and output a classification based on whether the galaxy possesses tidal features. We use a recently developed machine learning method that is essentially a middle-point between supervised and unsupervised learning, known as Self-Supervised machine Learning (SSL; He et al. 2019; Chen et al. 2020a;b;c; Chen and He 2020). Such models do not require labelled data for the training of the encoder, which learns to transform images into meaningful low-dimensional representations, but can perform classification when paired with a linear classifier and a small labelled dataset. Instead of labels, SSL models

[1]School of Physics, University of New South Wales, NSW 2052, Australia [2]AIM, CEA, CNRS, Université Paris-Saclay, Université Paris Diderot, Sorbonne Paris Cité, F-91191 Gif-sur-Yvette, France. Correspondence to: Alice Desmons <a.desmons@unsw.edu.edu.au>.

rely on augmentations to learn under which conditions the output low-dimensional representations should be invariant. These types of models have been successfully used for a variety of astronomical applications (e.g. Hayat et al. 2021; Stein et al. 2022; Slijepcevic et al. 2022; Walmsley et al. 2022; Wei et al. 2022; Ćiprijanović et al. 2023; Huertas-Company et al. 2023; Slijepcevic et al. 2023) Compared to supervised models, self-supervised models are also much easier to adapt to perform new tasks, and apply to datasets from different astronomical surveys (Ćiprijanović et al., 2023) making this kind of model perfect for our goal of applying the tool developed using HSC-SSP data to future LSST data.

2. Methods

2.1. Sample Selection

The dataset used for this work is sourced from the Ultradeep (UD) layer of the HSC-SSP Public Data Release 2 (PDR2; Aihara et al. 2019) for deep galaxy images. We use the Ultradeep field, which spans an area of $3.5 \deg^2$ and reaches a surface brightness depth of $\mu_r \sim 28.0$ mag arcsec^{-2} as it reaches depths faint enough to detect tidal features.

We assemble an unlabelled dataset of $\sim$44,000 galaxies by parsing objects in the HSC-SSP PDR2 database using an SQL search and only selecting objects which have at least 3 exposures in each band and have i-band magnitudes $15 < i < 20$ mag. We set a faint magnitude limit of 20 mag to ensure that objects are bright enough for tidal features to be visible. We access the HSC-SSP galaxy images using the 'Unagi' Python tool (Huang et al., 2019) which, given a galaxy's right ascension and declination, allows us to create multi-band 'HSC cutout' images of size 128×128 pixels, or 21×21 arcsecs, centred around each galaxy. Each cutout is downloaded in five (g, r, i, z, y) bands.

For the training of the linear classifier we require a small labelled dataset of galaxies with and without tidal features. We use the HSC-SSP UD PDR2 dataset assembled by Desmons et al. (2023) composed of 211 galaxies with tidal features and 641 galaxies without tidal features. These galaxies were selected from a volume-limited sample from the cross-over between then Galaxy and Mass Assembly survey (Driver et al., 2011) and HSC-SSP with spectroscopic redshift limits $0.04 \leq z \leq 0.2$ and stellar mass limits $9.50 \leq \log_{10}(M_\star/M_\odot) \leq 11.00$ and have i-band magnitudes in the range $12.8 < i < 21.6$ mag. To increase the size of our tidal feature training sample we classified additional galaxies from our HSC-SSP PDR2 unlabelled dataset of $\sim$ 44,000 objects, according to the classification scheme outlined in Desmons et al. (2023). Our final labelled sample contains 760 galaxies, 380 with tidal features, labelled 1, and 380 without, labelled 0. We split our labelled dataset

set into training, validation, and testing datasets composed of 600, 60, and 100 galaxies respectively.

2.2. Image Pre-processing and Augmentations

Before the images are augmented and fed through the model we apply a pre-processing function to normalise the images. The augmentations we use for this project are:

- **Orientation:** We randomly flip the image across each axis (x and y) with 50% probability.

- **Gaussian Noise**: We sample a scalar from $\mathcal{U}(1,3)$ and multiply it with the median absolute deviation of each channel (calculated over 1000 training examples) to get a per-channel noise σ_c. We then introduce Gaussian noise sampled from $\sigma_c \times \mathcal{N}(0,1)$ for each channel.

- **Jitter and Crop:** For HSC-SSP images we crop the 128×128 pixel image to the central 109×109 pixels before randomly cropping the image to 96×96 pixel. Random cropping means the image center is translated, or 'jittered', along each respective axis by i, j pixels where $i, j \sim \mathcal{U}(-13,13)$ before cropping to the central 96×96 pixels.

2.3. Model Architecture

The model we utilise to perform classification of tidal feature candidates consists of two components; a self-supervised model used for pre-training, and a 'fine-tuned' model used for classification. All models described below are built using the TensorFlow framework (Abadi et al., 2016).

2.3.1. THE SELF-SUPERVISED ARCHITECTURE

For our task of classifying tidal feature candidates we use a type of self-supervised learning known as Nearest Neighbour Contrastive Learning of visual Representations (NNCLR; Dwibedi et al. 2021). We closely follow Dwibedi et al. (2021) in designing the architecture and training process for our model. The model was compiled using the Adam optimiser (Kingma and Ba, 2015) and trained for 25 epochs on our unlabelled dataset of $\sim$ 44,000 HSC-SSP PDR2 galaxies.

2.3.2. THE FINE-TUNED ARCHITECTURE

The fine-tuned model is a simple linear classifier which takes galaxy images as input and converts them to representations using the pre-trained self-supervised encoder. These representations are passed through a 'Dense' layer with a sigmoid activation, which outputs a single number between 0 and 1. This fine-tuned model was compiled using the Adam optimiser (Kingma and Ba, 2015) and a binary cross entropy loss. It was trained for 50 epochs using the labelled

training set of 600 HSC-SSP galaxies. Training was completed within $\sim$ 1 minute using a single GPU.

2.3.3. THE SUPERVISED ARCHITECTURE

To draw conclusions about the suitability of self-supervised models for the detection and classification of tidal features, we compare our results with those of a fully supervised model. We do not construct this model from scratch, but instead use the published model designed by Pearson et al. (2019) to classify merging galaxies. The output layer was changed from two neurons with softmax activation, to a single neuron with sigmoid activation. The network was compiled using the Adam optimiser (Kingma and Ba, 2015) with the default learning rate and loss of the network was determined using binary cross entropy. We additionally changed the input image dimension from 64×64 pixels with three colour channels to 96×96 pixels with five colour channels to ensure extended tidal features remain visible. We train this fully supervised network from scratch using the labelled training set of 600 HSC-SSP galaxies.

2.4. Model Evaluation

To evaluate our model performance we use the true positive rate (also known as recall or completeness) and false positive rate (also known as fall-out or contamination). The true positive rate (TPR) ranges from 0 to 1 and is defined as the fraction of galaxies correctly classified by the model as having tidal features with respect to the total number of galaxies with tidal features. The false positive rate (FPR) also ranges from 0 to 1 and is defined as the fraction of galaxies incorrectly classified by the model as having tidal features with respect to the total number of galaxies without tidal features.

In addition to using the TPR for a given FPR to evaluate our model, we also use the area under the receiver operating characteristic (ROC) curve, or ROC AUC, to evaluate performance.

3. Results

3.1. Self-Supervised vs. Supervised Performance

Figure 1 illustrates the testing set ROC AUC for a supervised and self-supervised network as a function of the number of labels used in training for our HSC-SSP dataset. Each point represents the ROC AUC averaged over ten runs using the same training, validation, and testing sets for each run. We average the ROC AUC over the 10 runs and remove outliers further than 3σ from the mean. Our SSL model maintains high performance across all amounts of labels used for training, having ROC AUC $= 0.911 \pm 0.002$ when training on the maximum number of labels and only dropping to ROC AUC $= 0.89 \pm 0.01$ when using only

50 labels for training. The supervised model also maintains its performance regardless of label number, but only reaches ROC AUC $= 0.867 \pm 0.004$ when training on the maximum number and ROC AUC $= 0.83 \pm 0.01$ when using 50 labels for training. This figure not only shows
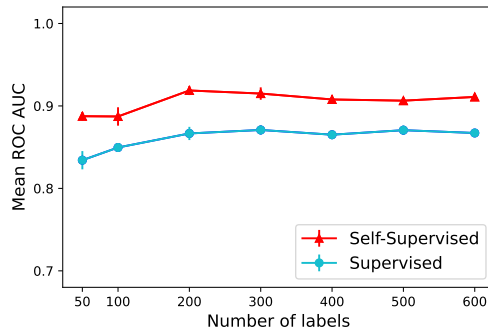

Figure 1. Average ROC AUC as a function of the number of HSC-SSP labels used for training for a supervised (blue) and self-supervised (red) model. Each point is an average of ten runs.

that an SSL model can be used for the detection of tidal features with good performance, but also that it performs consistently better than the supervised network regardless of the number of training labels. We also calculated the average TPR reached by the self-supervised model on the testing set for a given FPR $= 0.2$, averaging over 10 runs and removing outliers. When training using 600 labels, the model reaches TPR $= 0.94 \pm 0.01$, and this only drops to TPR $= 0.90 \pm 0.01$ when using a mere 50 labels for training.

3.2. Detection of Tidal Features

One advantage of self-supervised models over supervised models is the ability to use just one labelled example to find examples of similar galaxies from the full dataset. By using just one image from our labelled tidal feature dataset as a query image, and the encoded 128-dimensional representations from the self-supervised encoder, we can perform a similarity search that assigns high similarity scores to images which have similar representations to the query image. This is demonstrated in Figure 2 where we select a random galaxy with tidal features from our training sample and perform a similarity search with the 44,000 unlabelled HSC-SSP galaxies. In Figure 2 the query image is shown on the right alongside the 24 galaxies which received the highest similarity scores. This figure shows the power of self-supervised learning, where using only a single labelled example, we can find a multitude of other tidal feature candidates.

We can also visualise how the model organises the galaxy images in representation space, by using Uniform Manifold
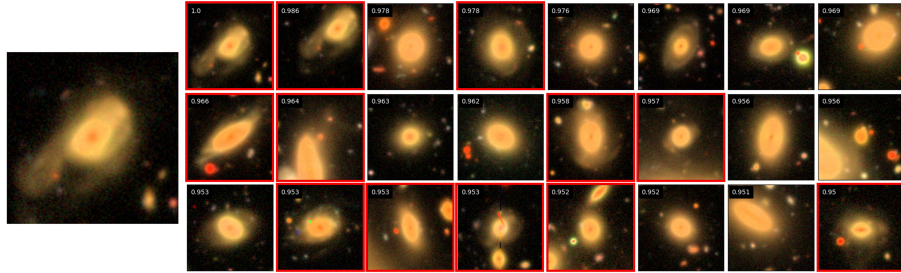
Figure 2. Results from a similarity search using a random galaxy with tidal features as a query image, displayed on the left, alongside the top 24 galaxies with the highest similarity scores for each similarity search on the right. The similarity score is displayed in the top left corner for each image. The red outlines indicate images containing galaxies which would be visually classified as hosting tidal features, regardless of whether this galaxy is the central object in the image.
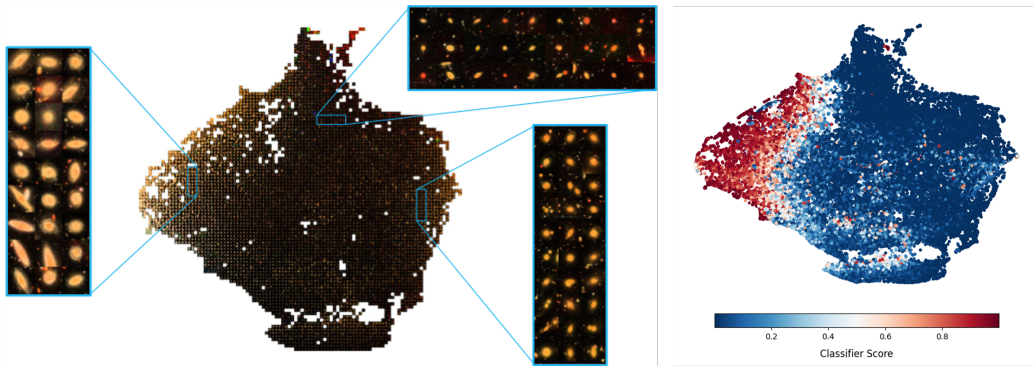

Figure 3. Left: 2D UMAP projection of the self-supervised representations. Made by binning the space into 100×100 cells and randomly selecting a sample from that cell to plot in the corresponding cell location. Right: The same 2D UMAP projection without binning, coloured according the scores assigned to each galaxy by the linear classifier.

Approximation and Projection (UMAP; McInnes et al. 2018) which reduces the encoded representations to an easier to visualise 2 dimensional projection. Figure 3 illustrates this 2D projection, created by binning the space into 100×100 cells and randomly selecting a sample from that cell to plot in the corresponding cell location. We also enquire whether the scores given to galaxies by the linear classifier are related to the galaxies' positions in the UMAP projection, by colouring the UMAP plot according the scores given to each galaxy by the linear classifier, shown in the right panel of Figure 3. We find that the majority of galaxies which were assigned a high classifier score, indicating a high likelihood of tidal features, are located on the left side of the UMAP projection plot. This reinforces the idea that the encoded representations contain meaningful information about tidal features.

4. Discussion and Conclusions

In this work, we have shown that SSL models composed of a self-supervised encoder and linear classifier can not only be used to detect galaxies with tidal features, but can do so

reaching both high completeness (TPR = 0. 94 ± 0.1) for low contamination (FPR = 0.20) and high area under the ROC curve (ROC AUC = 0.91 ± 0.002). This means that such models can be used to isolate the majority of galaxies with tidal features from a large sample of galaxies, thus drastically reducing the amount of visual classification needed to assemble a large sample of tidal features. One major advantage of this model over other automated classification methods, is that this level of performance can be reached using only 600 labelled training examples, and only drops mildly when using a mere 50 labels for training maintaining ROC AUC = 0.89 ± 0.01 and TPR = 0.90 ± 0.1 for FPR = 0.2. This makes SSL models easy to re-train on data from different surveys with minimal visual classification needed. Following Stein et al. (2021), we emphasise the usefulness of being able to perform a similarity search using just the self-supervised encoder and one example of a galaxy with tidal features to find other galaxies with tidal features from a dataset of tens of thousands of galaxies.

The level of comparison that can be carried out with respect to the results obtained here and other works is limited due

to the scarcity of similar works. There is only one study focusing on the detection of tidal features using machine learning, namely the work of Walmsley et al. (2019) who used a supervised network to identify galaxies with tidal features from the Wide layer of the Canada-France-Hawaii Telescope Legacy Survey (Gwyn, 2012). Walmsley et al. (2019) found that their method outperformed other automated methods of tidal feature detection, reaching 76% completeness (or TPR) and 22% contamination (or FPR). Our SSL model, trained on 600 galaxies performs considerably better, reaching a completeness of 96% for the same contamination percentage. Most importantly, our model consistently outperforms a fully supervised model trained on the same data, reaching ROC AUC = 0.911 ±0.002 while the fully supervised model only reaches a maximum ROC AUC of 0.864 ± 0.004.

The code use to create, train, validate, and test the SSML model, along with instructions on loading and using the pre-trained model as well as training the model using different data can be downloaded from GitHub[1].

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mane, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viegas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., and Zheng, X. (2016). TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. *arXiv e-prints*, page arXiv:1603.04467.

Aihara, H., AlSayyad, Y., Ando, M., Armstrong, R., Bosch, J., Egami, E., Furusawa, H., Furusawa, J., Goulding, A., Harikane, Y., Hikage, C., Ho, P. T. P., Hsieh, B.-C., Huang, S., Ikeda, H., Imanishi, M., Ito, K., Iwata, I., Jaelani, A. T., Kakuma, R., Kawana, K., Kikuta, S., Kobayashi, U., Koike, M., Komiyama, Y., Li, X., Liang, Y., Lin, Y.-T., Luo, W., Lupton, R., Lust, N. B., MacArthur, L. A., Matsuoka, Y., Mineo, S., Miyatake, H., Miyazaki, S., More, S., Murata, R., Namiki, S. V., Nishizawa, A. J., Oguri, M., Okabe, N., Okamoto, S., Okura, Y., Ono, Y., Onodera, M., Onoue, M., Osato, K., Ouchi, M., Shibuya, T., Strauss, M. A., Sugiyama, N., Suto, Y., Takada, M., Takagi, Y., Takata, T., Takita, S., Tanaka, M., Terai, T., Toba, Y., Uchiyama, H., Utsumi, Y., Wang, S.-Y., Wang, W., and Yamada, Y. (2019). Second data release of the Hyper Suprime-Cam Subaru Strategic Program. *PASJ*, 71(6):114.

Atkinson, A. M., Abraham, R. G., and Ferguson, A. M. N. (2013). Faint Tidal Features in Galaxies within the Canada-France-Hawaii Telescope Legacy Survey Wide Fields. *ApJ*, 765(1):28.

Bílek, M., Duc, P.-A., Cuillandre, J.-C., Gwyn, S., Cappellari, M., Bekaert, D. V., Bonfini, P., Bitsakis, T., Paudel, S., Krajnović, D., Durrell, P. R., and Marleau, F. (2020). Census and classification of low-surface-brightness structures in nearby early-type galaxies from the MATLAS survey. *MNRAS*, 498(2):2138–2166.

Cavanagh, M. K. and Bekki, K. (2020). Bars formed in galaxy merging and their classification with deep learning. *A&A*, 641:A77.

Chen, T., Kornblith, S., Norouzi, M., and Hinton, G. (2020a). A Simple Framework for Contrastive Learning of Visual Representations. *arXiv e-prints*, page arXiv:2002.05709.

Chen, T., Kornblith, S., Swersky, K., Norouzi, M., and Hinton, G. (2020b). Big Self-Supervised Models are Strong Semi-Supervised Learners. *arXiv e-prints*, page arXiv:2006.10029.

---

[1]https://github.com/LSSTISSC/Tidalsaurus

Chen, X., Fan, H., Girshick, R., and He, K. (2020c). Improved Baselines with Momentum Contrastive Learning. *arXiv e-prints*, page arXiv:2003.04297.

Chen, X. and He, K. (2020). Exploring Simple Siamese Representation Learning. *arXiv e-prints*, page arXiv:2011.10566.

Ćiprijanović, A., Lewis, A., Pedro, K., Madireddy, S., Nord, B., Perdue, G. N., and Wild, S. M. (2023). DeepAstroUDA: Semi-Supervised Universal Domain Adaptation for Cross-Survey Galaxy Morphology Classification and Anomaly Detection. *arXiv e-prints*, page arXiv:2302.02005.

Cole, S., Lacey, C. G., Baugh, C. M., and Frenk, C. S. (2000). Hierarchical galaxy formation. *MNRAS*, 319(1):168–204.

Darg, D. W., Kaviraj, S., Lintott, C. J., Schawinski, K., Sarzi, M., Bamford, S., Silk, J., Proctor, R., Andreescu, D., Murray, P., Nichol, R. C., Raddick, M. J., Slosar, A., Szalay, A. S., Thomas, D., and Vandenberg, J. (2010). Galaxy Zoo: the fraction of merging galaxies in the SDSS and their morphologies. *MNRAS*, 401(2):1043–1056.

Desmons, A., Brough, S., Martínez-Lombilla, C., De Propris, R., Holwerda, B., and López Sánchez, Á. R. (2023). Galaxy and mass assembly (GAMA): Comparing visually and spectroscopically identified galaxy merger samples. *MNRAS*.

Diaz, J. D., Bekki, K., Forbes, D. A., Couch, W. J., Drinkwater, M. J., and Deeley, S. (2019). Classifying the formation processes of S0 galaxies using Convolutional Neural Networks. *MNRAS*, 486(4):4845–4862.

Driver, S. P., Hill, D. T., Kelvin, L. S., Robotham, A. S. G., Liske, J., Norberg, P., Baldry, I. K., Bamford, S. P., Hopkins, A. M., Loveday, J., Peacock, J. A., Andrae, E., Bland-Hawthorn, J., Brough, S., Brown, M. J. I., Cameron, E., Ching, J. H. Y., Colless, M., Conselice, C. J., Croom, S. M., Cross, N. J. G., de Propris, R., Dye, S., Drinkwater, M. J., Ellis, S., Graham, A. W., Grootes, M. W., Gunawardhana, M., Jones, D. H., van Kampen, E., Maraston, C., Nichol, R. C., Parkinson, H. R., Phillipps, S., Pimbblet, K., Popescu, C. C., Prescott, M., Roseboom, I. G., Sadler, E. M., Sansom, A. E., Sharp, R. G., Smith, D. J. B., Taylor, E., Thomas, D., Tuffs, R. J., Wijesinghe, D., Dunne, L., Frenk, C. S., Jarvis, M. J., Madore, B. F., Meyer, M. J., Seibert, M., Staveley-Smith, L., Sutherland, W. J., and Warren, S. J. (2011). Galaxy and Mass Assembly (GAMA): survey diagnostics and core data release. *MNRAS*, 413(2):971–995.

Dwibedi, D., Aytar, Y., Tompson, J., Sermanet, P., and Zisserman, A. (2021). With a Little Help from My Friends: Nearest-Neighbor Contrastive Learning of Visual Representations. *arXiv e-prints*, page arXiv:2104.14548.

Gwyn, S. D. J. (2012). The canada-france-hawaii telescope legacy survey: Stacked images and catalogs. *AJ*, 143(2).

Hayat, M. A., Stein, G., Harrington, P., Lukić, Z., and Mustafa, M. (2021). Self-supervised Representation Learning for Astronomical Images. *ApJl*, 911(2):L33.

He, K., Fan, H., Wu, Y., Xie, S., and Girshick, R. (2019). Momentum Contrast for Unsupervised Visual Representation Learning. *arXiv e-prints*, page arXiv:1911.05722.

Hocking, A., Geach, J. E., Sun, Y., and Davey, N. (2018). An automatic taxonomy of galaxy morphology using unsupervised machine learning. *MNRAS*, 473(1):1108–1129.

Hood, C. E., Kannappan, S. J., Stark, D. V., Dell'Antonio, I. P., Moffett, A. J., Eckert, K. D., Norris, M. A., and Hendel, D. (2018). The Origin of Faint Tidal Features around Galaxies in the RESOLVE Survey. *ApJ*, 857(2):144.

Huang, S., Li, J.-X., Lanusse, F., and Bradshaw, C. (2019). Unagi. https://github.com/dr-guangtou/unagi.git.

Huertas-Company, M., Sarmiento, R., and Knapen, J. (2023). A brief review of contrastive learning applied to astrophysics. *arXiv e-prints*, page arXiv:2306.05528.

Ivezić, Ž., Kahn, S. M., Tyson, J. A., Abel, B., Acosta, E., Allsman, R., Alonso, D., AlSayyad, Y., Anderson, S. F., Andrew, J., Angel, J. R. P., Angeli, G. Z., Ansari, R., Antilogus, P., Araujo, C., Armstrong, R., Arndt, K. T., Astier, P., Aubourg, É., Auza, N., Axelrod, T. S., Bard, D. J., Barr, J. D., Barrau, A., Bartlett, J. G., Bauer, A. E., Bauman, B. J., Baumont, S., Bechtol, E., Bechtol, K., Becker, A. C., Becla, J., Beldica, C., Bellavia, S., Bianco, F. B., Biswas, R., Blanc, G., Blazek, J., Blandford, R. D., Bloom, J. S., Bogart, J., Bond, T. W., Booth, M. T., Borgland, A. W., Borne, K., Bosch, J. F., Boutigny, D., Brackett, C. A., Bradshaw, A., Brandt, W. N., Brown, M. E., Bullock, J. S., Burchat, P., Burke, D. L., Cagnoli, G., Calabrese, D., Callahan, S., Callen, A. L., Carlin, J. L., Carlson, E. L., Chandrasekharan, S., Charles-Emerson, G., Chesley, S., Cheu, E. C., Chiang, H.-F., Chiang, J., Chirino, C., Chow, D., Ciardi, D. R., Claver, C. F., Cohen-Tanugi, J., Cockrum, J. J., Coles, R., Connolly, A. J., Cook, K. H., Cooray, A., Covey, K. R., Cribbs, C., Cui, W., Cutri, R., Daly, P. N., Daniel, S. F., Daruich, F., Daubard, G., Daues, G., Dawson, W., Delgado, F., Dellapenna, A., de Peyster, R., de Val-Borro, M., Digel, S. W., Doherty, P., Dubois, R., Dubois-Felsmann, G. P., Durech, J., Economou, F., Eifler, T., Eracleous, M., Emmons, B. L., Fausti Neto, A., Ferguson, H., Figueroa,

E., Fisher-Levine, M., Focke, W., Foss, M. D., Frank, J., Freemon, M. D., Gangler, E., Gawiser, E., Geary, J. C., Gee, P., Geha, M., Gessner, C. J. B., Gibson, R. R., Gilmore, D. K., Glanzman, T., Glick, W., Goldina, T., Goldstein, D. A., Goodenow, I., Graham, M. L., Gressler, W. J., Gris, P., Guy, L. P., Guyonnet, A., Haller, G., Harris, R., Hascall, P. A., Haupt, J., Hernandez, F., Herrmann, S., Hileman, E., Hoblitt, J., Hodgson, J. A., Hogan, C., Howard, J. D., Huang, D., Huffer, M. E., Ingraham, P., Innes, W. R., Jacoby, S. H., Jain, B., Jammes, F., Jee, M. J., Jenness, T., Jernigan, G., Jevremović, D., Johns, K., Johnson, A. S., Johnson, M. W. G., Jones, R. L., Juramy-Gilles, C., Jurić, M., Kalirai, J. S., Kallivayalil, N. J., Kalmbach, B., Kantor, J. P., Karst, P., Kasliwal, M. M., Kelly, H., Kessler, R., Kinnison, V., Kirkby, D., Knox, L., Kotov, I. V., Krabbendam, V. L., Krughoff, K. S., Kubánek, P., Kuczewski, J., Kulkarni, S., Ku, J., Kurita, N. R., Lage, C. S., Lambert, R., Lange, T., Langton, J. B., Le Guillou, L., Levine, D., Liang, M., Lim, K.-T., Lintott, C. J., Long, K. E., Lopez, M., Lotz, P. J., Lupton, R. H., Lust, N. B., MacArthur, L. A., Mahabal, A., Mandelbaum, R., Markiewicz, T. W., Marsh, D. S., Marshall, P. J., Marshall, S., May, M., McKercher, R., McQueen, M., Meyers, J., Migliore, M., Miller, M., Mills, D. J., Miraval, C., Moeyens, J., Moolekamp, F. E., Monet, D. G., Moniez, M., Monkewitz, S., Montgomery, C., Morrison, C. B., Mueller, F., Muller, G. P., Muñoz Arancibia, F., Neill, D. R., Newbry, S. P., Nief, J.-Y., Nomerotski, A., Nordby, M., O'Connor, P., Oliver, J., Olivier, S. S., Olsen, K., O'Mullane, W., Ortiz, S., Osier, S., Owen, R. E., Pain, R., Palecek, P. E., Parejko, J. K., Parsons, J. B., Pease, N. M., Peterson, J. M., Peterson, J. R., Petravick, D. L., Libby Petrick, M. E., Petry, C. E., Pierfederici, F., Pietrowicz, S., Pike, R., Pinto, P. A., Plante, R., Plate, S., Plutchak, J. P., Price, P. A., Prouza, M., Radeka, V., Rajagopal, J., Rasmussen, A. P., Regnault, N., Reil, K. A., Reiss, D. J., Reuter, M. A., Ridgway, S. T., Riot, V. J., Ritz, S., Robinson, S., Roby, W., Roodman, A., Rosing, W., Roucelle, C., Rumore, M. R., Russo, S., Saha, A., Sassolas, B., Schalk, T. L., Schellart, P., Schindler, R. H., Schmidt, S., Schneider, D. P., Schneider, M. D., Schoening, W., Schumacher, G., Schwamb, M. E., Sebag, J., Selvy, B., Sembroski, G. H., Seppala, L. G., Serio, A., Serrano, E., Shaw, R. A., Shipsey, I., Sick, J., Silvestri, N., Slater, C. T., Smith, J. A., Smith, R. C., Sobhani, S., Soldahl, C., Storrie-Lombardi, L., Stover, E., Strauss, M. A., Street, R. A., Stubbs, C. W., Sullivan, I. S., Sweeney, D., Swinbank, J. D., Szalay, A., Takacs, P., Tether, S. A., Thaler, J. J., Thayer, J. G., Thomas, S., Thornton, A. J., Thukral, V., Tice, J., Trilling, D. E., Turri, M., Van Berg, R., Vanden Berk, D., Vetter, K., Virieux, F., Vucina, T., Wahl, W., Walkowicz, L., Walsh, B., Walter, C. W., Wang, D. L., Wang, S.-Y., Warner, M., Wiecha, O., Willman, B., Winters, S. E., Wittman, D., Wolff, S. C., Wood-Vasey,

W. M., Wu, X., Xin, B., Yoachim, P., and Zhan, H. (2019). LSST: From Science Drivers to Reference Design and Anticipated Data Products. *ApJ*, 873(2):111.

Kingma, D. P. and Ba, J. (2015). Adam: A Method for Stochastic Optimization. *3rd International Conference for Learning Representations*, page arXiv:1412.6980.

Lacey, C. and Cole, S. (1994). Merger Rates in Hierarchical Models of Galaxy Formation - Part Two - Comparison with N-Body Simulations. *MNRAS*, 271:676.

Lintott, C. J., Schawinski, K., Slosar, A., Land, K., Bamford, S., Thomas, D., Raddick, M. J., Nichol, R. C., Szalay, A., Andreescu, D., Murray, P., and Vandenberg, J. (2008). Galaxy Zoo: morphologies derived from visual inspection of galaxies from the Sloan Digital Sky Survey. *MNRAS*, 389(3):1179–1189.

Martin, G., Bazkiaei, A. E., Spavone, M., Iodice, E., Mihos, J. C., Montes, M., Benavides, J. A., Brough, S., Carlin, J. L., Collins, C. A., Duc, P. A., Gómez, F. A., Galaz, G., Hernández-Toledo, H. M., Jackson, R. A., Kaviraj, S., Knapen, J. H., Martínez-Lombilla, C., McGee, S., O'Ryan, D., Prole, D. J., Rich, R. M., Román, J., Shah, E. A., Starkenburg, T. K., Watkins, A. E., Zaritsky, D., Pichon, C., Armus, L., Bianconi, M., Buitrago, F., Busá, I., Davis, F., Demarco, R., Desmons, A., García, P., Graham, A. W., Holwerda, B., Hon, D. S. H., Khalid, A., Klehammer, J., Klutse, D. Y., Lazar, I., Nair, P., Noakes-Kettel, E. A., Rutkowski, M., Saha, K., Sahu, N., Sola, E., Vázquez-Mata, J. A., Vera-Casanova, A., and Yoon, I. (2022). Preparing for low surface brightness science with the Vera C. Rubin Observatory: Characterization of tidal features from mock images. *MNRAS*, 513(1):1459–1487.

Martin, G., Kaviraj, S., Devriendt, J. E. G., Dubois, Y., and Pichon, C. (2018). The role of mergers in driving morphological transformation over cosmic time. *MNRAS*, 480(2):2266–2283.

Martin, G., Kaviraj, S., Hocking, A., Read, S. C., and Geach, J. E. (2020). Galaxy morphological classification in deep-wide surveys via unsupervised machine learning. *MNRAS*, 491(1):1408–1426.

McInnes, L., Healy, J., and Melville, J. (2018). UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *arXiv e-prints*, page arXiv:1802.03426.

Pearson, W. J., Wang, L., Trayford, J. W., Petrillo, C. E., and van der Tak, F. F. S. (2019). Identifying galaxy mergers in observations and simulations with deep learning. *A&A*, 626:A49.

Robotham, A. S. G., Driver, S. P., Davies, L. J. M., Hopkins, A. M., Baldry, I. K., Agius, N. K., Bauer, A. E., Bland-Hawthorn, J., Brough, S., Brown, M. J. I., Cluver, M., De

7

Propris, R., Drinkwater, M. J., Holwerda, B. W., Kelvin, L. S., Lara-Lopez, M. A., Liske, J., López-Sánchez, Á. R., Loveday, J., Mahajan, S., McNaught-Roberts, T., Moffett, A., Norberg, P., Obreschkow, D., Owers, M. S., Penny, S. J., Pimbblet, K., Prescott, M., Taylor, E. N., van Kampen, E., and Wilkins, S. M. (2014). Galaxy And Mass Assembly (GAMA): galaxy close pairs, mergers and the future fate of stellar mass. *MNRAS*, 444(4):3986–4008.

Sheen, Y.-K., Yi, S. K., Ree, C. H., and Lee, J. (2012). Post-merger Signatures of Red-sequence Galaxies in Rich Abell Clusters at z ¡~0.1. *ApJs*, 202(1):8.

Slijepcevic, I. V., Scaife, A., Walmsley, M., and Bowles, M. R. (2022). Learning useful representations for radio astronomy "in the wild" with contrastive learning. In *Machine Learning for Astrophysics*, page 53.

Slijepcevic, I. V., Scaife, A. M. M., Walmsley, M., Bowles, M., Wong, O. I., Shabala, S. S., and White, S. V. (2023). Radio Galaxy Zoo: Building a multi-purpose foundation model for radio astronomy with self-supervised learning. *arXiv e-prints*, page arXiv:2305.16127.

Snyder, G. F., Rodriguez-Gomez, V., Lotz, J. M., Torrey, P., Quirk, A. C. N., Hernquist, L., Vogelsberger, M., and Freeman, P. E. (2019). Automated distant galaxy merger classifications from Space Telescope images using the Illustris simulation. *MNRAS*, 486(3):3702–3720.

Stein, G., Blaum, J., Harrington, P., Medan, T., and Lukić, Z. (2022). Mining for Strong Gravitational Lenses with Self-supervised Learning. *ApJ*, 932(2):107.

Stein, G., Harrington, P., Blaum, J., Medan, T., and Lukic, Z. (2021). Self-supervised similarity search for large scientific datasets. *arXiv e-prints*, page arXiv:2110.13151.

Tal, T., van Dokkum, P. G., Nelan, J., and Bezanson, R. (2009). The Frequency of Tidal Features Associated with Nearby Luminous Elliptical Galaxies From a Statistically Complete Sample. *AJ*, 138(5):1417–1427.

Walmsley, M., Ferguson, A. M. N., Mann, R. G., and Lintott, C. J. (2019). Identification of low surface brightness tidal features in galaxies using convolutional neural networks. *MNRAS*, 483(3):2968–2982.

Walmsley, M., Slijepcevic, I., Bowles, M. R., and Scaife, A. (2022). Toward Galaxy Foundation Models with Hybrid Contrastive Learning. In *Machine Learning for Astrophysics*, page 29.

Wei, S., Li, Y., Lu, W., Li, N., Liang, B., Dai, W., and Zhang, Z. (2022). Unsupervised Galaxy Morphological Visual Representation with Deep Contrastive Learning. *PASP*, 134(1041):114508.