Detecting Low Surface Brightness Galaxies with Mask R-CNN

Caleb Levy
cmlevy@colgate.edu
Department of Physics and Astronomy
Colgate University
Hamilton, NY 13346

Aleksandra Ćiprijanović
aleksand@fnal.gov
Fermi National Accelerator Laboratory

Alex Drlica-Wagner
kadrlica@fnal.gov
Fermi National Accelerator Laboratory
Kavli Institute for Cosmological Physics
Department of Astronomy and Astrophysics
University of Chicago

Burçin Mutlu-Pakdil
burcinmp@uchicago.edu
Kavli Institute for Cosmological Physics
Department of Astronomy and Astrophysics
University of Chicago

Brian Nord
nord@fnal.gov
Fermi National Accelerator Laboratory
Kavli Institute for Cosmological Physics
Department of Astronomy and Astrophysics
University of Chicago

Dimitrios Tanoglidis
dtanoglidis@uchicago.edu
Department of Astronomy and Astrophysics
University of Chicago

Abstract

Low surface brightness galaxies (LSBGs), galaxies that are fainter than the dark night sky, are famously difficult to detect. Nonetheless, studies of these galaxies are essential to improve our understanding of the formation and evolution of low-mass galaxies. In this work, we train a deep learning model using the Mask R-CNN framework on a set of simulated LSBGs inserted into images from the Dark Energy Survey (DES) Data Release 2 (DR2). This deep learning model is combined with several conventional image pre-processing steps to develop a pipeline for the detection of LSBGs. We apply this pipeline to the full DES DR2 coadd image dataset, and preliminary results show the detection of 22 large, high-quality LSBG candidates that went undetected by conventional algorithms. Furthermore, we find that the performance of our algorithm is greatly improved by including examples of false positives as an additional class during training.

1 Introduction

The limited sensitivity of astronomical observations leads to lower efficiency in detecting faint astronomical systems (Disney, 1976; Bothun et al., 1997). In particular, low surface brightness galaxies (LSBGs), galaxies with central brightness fainter than the night sky, have proven to be very difficult to detect with conventional algorithms. These systems are extremes of the galactic population, and thus provide a good test for the standard models of cosmology and galaxy formation. To address these observational issues, it is essential to develop efficient and effective techniques for identifying and cataloging very faint galaxies.

Fourth Workshop on Machine Learning and the Physical Sciences (NeurIPS 2021).

Recently, a catalog of 23,790 LSBGs was identified in the first three years of data from the Dark Energy Survey (DES; DES Collaboration, 2005; DES Collaboration et al., 2018) using results from `SourceExtractor` (Bertin and Arnouts, 1996) and a conventional support vector machine classifier (Tanoglidis et al., 2021b). This labelled dataset has subsequently been used to train a Convolutional Neural Network (CNN) for separating LSBGs from artifacts (Tanoglidis et al., 2021a), showing the promise of machine vision for studying the low-surface-brightness universe. The focus of this work is to leverage the power of machine vision, specifically through the Mask R-CNN framework (He et al., 2018), to train a model to detect large, faint LSBGs that are missed by conventional algorithms.

2 Mask R-CNN Model

In this work, we utilize the instance segmentation and object detection algorithm Mask R-CNN. We describe this model briefly here, and refer the interested reader to the original paper for a more technical discussion (He et al., 2018). A high-level depiction of the algorithm is shown in Fig. 1. The input image is first fed into the “backbone" of the network, a pre-trained CNN, in this case ResNet, whose output is a feature map. These feature maps are then fed into a Region Proposal Network (RPN) whose output is a set of Regions of Interest (RoIs) where potential objects are located. These regions are identified from a CNN trained to identify *anchor boxes* of various sizes at each position. The regions with the highest probabilities of containing an object are then passed into the RoIAlign method along with the feature maps from the backbone to produce RoI feature maps of the same size. The final part of the algorithm is a simultaneous three-fold process: classification of the object in the

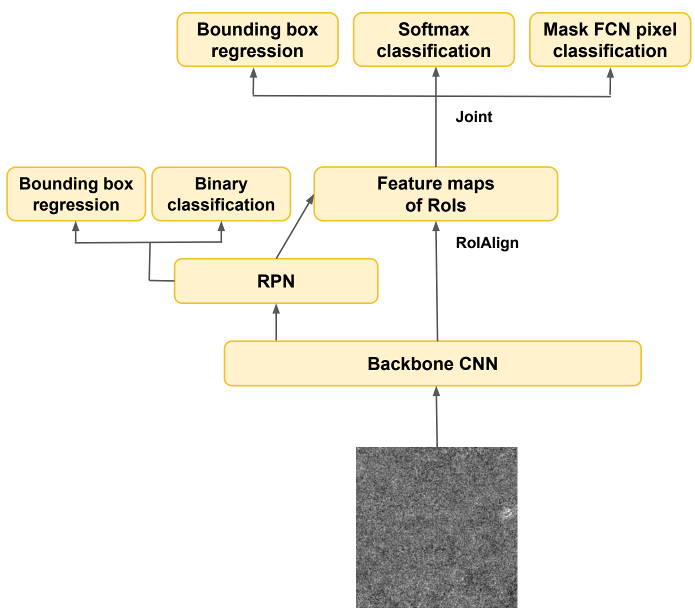

Figure 1: Schematic description of the Mask R-CNN algorithm.

RoI with a softmax classifier, regression to learn the best bounding box coordinates, and per-pixel classification with a Fully Convolutional Network (FCN) to create an object mask. The final output of the algorithm is a bounding box region, a per-pixel mask, an object class, and a confidence score for each detected object.

2.1 Image Processing and Training

Training the Mask R-CNN algorithm requires a set of images along with per-pixel labels of training objects. To train a model for detecting LSBGs, we injected simulated galaxies into 94 random 10,000 pixel $\times$ 10,000 pixel DES images. A simulated dataset provides the distinct advantage of having pre-labeled ground truth masks, bypassing the requirement of manually labelling a dataset. Artificial

LSBGs are modeled with GalSim (Rowe et al., 2015), an open-source simulation toolkit. GalSim can simulate galaxies from a variety of simple parametric models, and we used a Sersić profile for our artificial LSBGs.

After injecting 16 LSBGs per tile, each tile was divided into sixteen 2,500 pixel $\times$ 2,500 pixel quadrants with each containing one LSBG. In addition to the LSBG class, we also included a training class for very large objects identified from the DES DR2 catalog (DES Collaboration and et al., 2021). Objects in this class were identified as those whose masked area exceeded 10,000 pixels (~ 0.2 deg^2) in angular size. The purpose of this was to remove contamination coming from the periphery of bright sources by training the algorithm to specifically detect this class of object as distinct from LSBGs. The next step is to pre-process images before training in order to accentuate low surface brightness features. This occurs in three steps:

1. Replace pixels associated with all sources detected in the DES DR2 catalog with noise.
2. Convolve the image with a Gaussian kernel.
3. Bin the image by a factor of 10×10.

The model was trained on 94 tiles with 1034 simulated LSBGs over 60 epochs. In total 1504 LSBGs were injected with $\sim 70\%$ of these assigned to a training set, $\sim 15\%$ left for validation, and $\sim 15\%$ used for testing. Each epoch consisted of 150 steps, with 15 steps at the end of each for validation.

The final step was to test the model on the reserved test set of 282 simulated LSBGs. While examining the quantitative results of this process is useful, we focused our efforts more on a qualitative analysis of the model. After looking at the model's output on the entire test set, we were confident in the model's performance and decided to apply the model to the full DES dataset. The detection pipeline is shown schematically in Fig. 2.

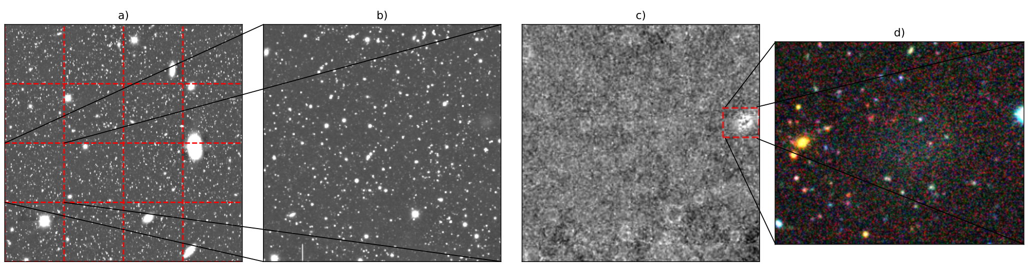

Figure 2: Example of a successful LSBG detection using the pipeline described in Section 2.1. Panel a) represents the full, unprocessed image along with a red grid outline indicating the cutouts. Panel b) shows a zoomed version of the indicated cutout before processing. Panel c) represents the fully processed version of panel b), showing the accentuation of the LSBG in the right of the image. The region identified by the Mask R-CNN is outlined in red. Panel d) shows the bounded detection region zoomed in from the DESI Legacy Imaging Surveys Sky Viewer.[1]

3 Results

In this section we summarize the results of running the detection pipeline on all 10,169 coadd image tiles in DES DR2. The model returned 13,336 objects classified as LSBGs of which $\sim 11{,}000$ have been visually inspected. This visual inspection identified 22 high confidence LSBGs and 19 low confidence LSBGs by comparing to images of known LSBGs in DES. Fig. 4 shows the zoomed in bounding box of 4 out of the 22 high confidence LSBG candidates in the DESI Legacy Imaging Surveys Sky Viewer (the same as panel (d) in Fig. 2). These galaxies are generally larger (with most having angular sizes > 20 arcsec) and lower surface brightness than those detected in previous DES LSBG catalogs (e.g., Tanoglidis et al., 2021b).

The vast majority of the model's output labelled as LSBGs were artifacts. The most prominent artifact class consisted of Galactic cirrus, starlight scattered off of interstellar dust in our Galaxy. Contamination from Galactic cirrus is a well-known systematic when studying the low surface

[1] http://legacysurvey.org/

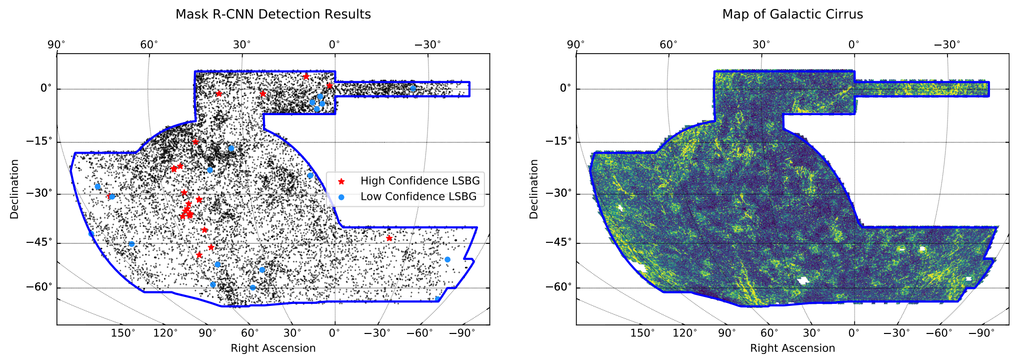

Figure 3: Left panel: Sky map of all Mask R-CNN model detections (black dots) along with our new high confidence (red stars) and low confidence (blue points) LSBG detections. Right panel: Sky map of Galactic cirrus, where brighter regions represent higher detection densities. In both panels, the blue outline indicates the region of observed by DES.

brightness universe (e.g., when studying faint tidal features around distant galaxies; Cortese et al., 2010). The model also detected a number of tidal features associated with galaxies and galaxy mergers. After visually inspecting the images, we conclude that this is likely because these features are not completely characterized by `SourceExtractor` in the DES pipeline, and thus when the compact regions of the sources are removed, the tidal features in the periphery remain and are amplified by our pre-processing. We hoped to prevent this from occurring with the binary dilation step in the image pre-processing, which was implemented to cover as much of the compact sources in the mask as possible. However, we were not able to completely remove all features from these compact sources, hence their prominence in the model's output. The same is true for bright stars and large galaxies. Our initial models classified a high volume of these artifacts as LSBGs; however, adding instances of these false positives to the training set greatly improved the purity of the results.

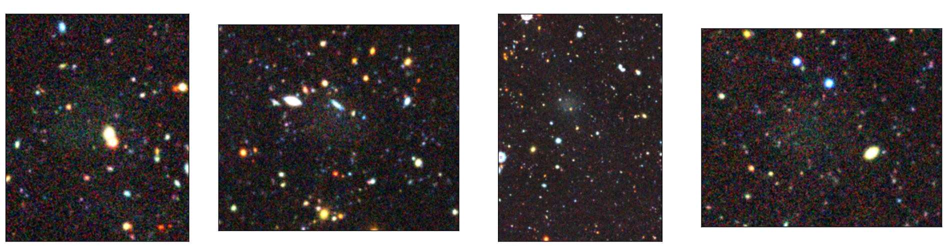

Figure 4: Four high confidence LSBG candidates from detection ran on 10,169 tiles of DES DR2 using the Mask R-CNN Model.

In Fig. 3, we plot the spatial distribution of all our detections (with the high and low confidence LSBGs highlighted) as compared to a map of the Galactic cirrus in the DES footprint.[2] As mentioned previously, our visual inspection showed a high incidence of Galactic cirrus in the pipeline output. We therefore decided to compare our results to a map of the Galactic cirrus, which showed a very high visual correlation in most areas, confirming the model's apparent propensity for detecting cirrus. It is interesting to note that some regions appear to visually lack a correlation to the map. Specifically, the left and lower right of the DES footprint appear underdense in our detection results while the map of Galactic cirrus suggests high densities in these regions. This is likely a result of the fact that these regions are closer to the Galactic plane and thus the model's detections are most likely dominated by artifacts of very bright, nearby stars. Another intriguing feature of Fig. 3 is the cluster of high confidence LSBG detections in the region around (RA, DEC) $= (55^\circ, -35^\circ)$. This is

[2]Cirrus map courtesy of the DES Collaboration (private communication).

approximately the location of the nearby Fornax galaxy cluster. Tanoglidis et al. (2021b) showed prominent clustering of red LSBGs around the Fornax cluster, a fact that our model verifies.

4 Conclusions

In this work we applied the deep learning framework Mask R-CNN to the problem of detecting large low surface brightness galaxies in the Dark Energy Survey Data Release 2 using a simulated, labelled dataset of LSBGs generated with GalSim with a range of physical parameters. We trained on 94 tiles with 16 LSBGs simulated per tile. We then developed a detection pipeline which we tested on $\sim 1\%$ of the DES dataset (100 tiles). We ran the pipeline on the entire DR2 dataset (10,169 tiles), identifying 22 new high-quality LSBGs. These galaxies were generally larger and lower surface brightness than those detected by conventional algorithms. In addition, we found the model to be very efficient at detecting Galactic cirrus and tidal features. The spatial distribution of all the detections appears to have a strong visual correlation with maps of Galactic cirrus. We have thus demonstrated that the Mask R-CNN framework could be very beneficial for detecting low surface brightness features.

Broader Impact

In this work, we used a well-established object detection/image segmentation algorithm for detecting astrophysical objects. On societal scales, object detection algorithms are known to have a significant impact, especially when trained for human detection, like facial recognition. The application of these algorithms to the physical sciences is limited so far, and so studying them in this context provides a unique opportunity to better understand their shortcomings. This can then help alleviate some of the biases inherent in these algorithms.

Acknowledgments and Disclosure of Funding

This manuscript has been supported by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy, Office of Science, Office of High Energy Physics.

References

Bertin, E., Arnouts, S., 1996. SExtractor: Software for source extraction. A&AS 117, 393–404.

Bothun, G., Impey, C., McGaugh, S., 1997. Low-surface-brightness galaxies: Hidden galaxies revealed. Publications of the Astronomical Society of the Pacific 109, 745.

Cortese, L., Bendo, G., Isaak, K., Davies, J., Kent, B., 2010. Diffuse far-infrared and ultraviolet emission in the ngc 4435/4438 system: Tidal stream or galactic cirrus. Monthly Notices of the Royal Astronomical Society: Letters 403.

DES Collaboration, 2005. The Dark Energy Survey. arXiv e-prints , astro–ph/0510346astro-ph/0510346.

DES Collaboration, et al., 2021. The dark energy survey data release 2. 2101.05765.

DES Collaboration, NOAO Data Lab, et al., 2018. The Dark Energy Survey: Data Release 1. ApJS 239, 18. 1801.03181.

Disney, M.J., 1976. Visibility of galaxies. Nature 263, 573–575.

He, K., Gkioxari, G., Dollár, P., Girshick, R., 2018. Mask r-cnn. 1703.06870.

Rowe, B.T.P., Jarvis, M., Mandelbaum, R., Bernstein, G.M., Bosch, J., et al., 2015. GALSIM: The modular galaxy image simulation toolkit. Astronomy and Computing 10, 121–150. 1407.7676.

Tanoglidis, D., Ćiprijanović, A., Drlica-Wagner, A., 2021a. DeepShadows: Separating low surface brightness galaxies from artifacts using deep learning. Astronomy and Computing 35, 100469. 2011.12437.

Tanoglidis, D., Drlica-Wagner, A., Wei, K., Li, T.S., Sánchez, J., DES Collaboration, et al., 2021b. Shadows in the Dark: Low-surface-brightness Galaxies Discovered in the Dark Energy Survey. ApJS 252, 18. 2006.04294.

Checklist

1. For all authors...
   (a) Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope? [Yes]
   (b) Did you describe the limitations of your work? [Yes]
   (c) Did you discuss any potential negative societal impacts of your work? [N/A]
   (d) Have you read the ethics review guidelines and ensured that your paper conforms to them? [Yes]
2. If you are including theoretical results...
   (a) Did you state the full set of assumptions of all theoretical results? [N/A]
   (b) Did you include complete proofs of all theoretical results? [N/A]
3. If you ran experiments...
   (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? [No]
   (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? [Yes]
   (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? [N/A]
   (d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? [Yes]
4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets...
   (a) If your work uses existing assets, did you cite the creators? [Yes]
   (b) Did you mention the license of the assets? [No]
   (c) Did you include any new assets either in the supplemental material or as a URL? [N/A]
   (d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? [No]
   (e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? [N/A]
5. If you used crowdsourcing or conducted research with human subjects...
   (a) Did you include the full text of instructions given to participants and screenshots, if applicable? [N/A]
   (b) Did you describe any potential participant risks, with links to Institutional Review Board (IRB) approvals, if applicable? [N/A]
   (c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? [N/A]