

Generative Adversarial Networks for Galaxy Classification Using Convolutional Neural Networks

Aryan Singh
University of Massachusetts, Amherst
arysingh@umass.edu

Polina Petrova
University of Massachusetts, Amherst
ppetrova@umass.edu

1. Proposal

1.1. Group Members

Our group consists of Aryan Singh and Polina Petrova. Aryan will be training a Super Resolution Generative Adversarial Network (SRGAN) to enhance image quality in the dataset and implementing a conditional Generative Adversarial Network (cGAN) to create additional images for each galaxy class. This involves selecting appropriate values for layer size, learning rate, regularization parameters, and loss function. Polina will be responsible for implementing the Convolutional Neural Network (CNN) for galaxy classification, including designing the architecture, choosing layers and configuring activation functions tailored to galaxy image data. Polina will also manage the training process, including correcting for overfitting. Each group member will complete hyperparameter tuning on their respective components of the model to achieve optimal accuracy, ensuring balanced trade-offs between model complexity and performance. Together, we will complete a high-level analysis of our GAN-CNN pipeline based on selected evaluation criteria and observe the effectiveness of classifying test instances on a classical cGAN compared to a SR-cGAN.

1.2. Motivation

We are addressing the problem of accurately classifying galaxies using CNNs where there is an imbalance of data within the classes. In such a situation, a cGAN can help generate additional data. This problem is interesting because data imbalance is a common real-world problem and leads to overcomplication of training algorithms. We will explore how cGANS can be used to mitigate this issue.

1.3. Literature Review

We will examine several key papers to provide context and background for our research. Works such as *Deep generative models for galaxy image simulations* [7] and *Forging new worlds: high-resolution synthetic galaxies with chained generative adversarial networks* [3] will provide

a foundational understanding of the use of GANs as applied to galaxy morphology classification problems. Additionally, we will explore papers such as *Galaxy classification: a deep learning approach for classifying Sloan Digital Sky Survey images* [4], *Star-galaxy classification using deep convolutional neural networks* [5], and *Neural computation as a tool for galaxy classification: methods and examples* [6], which will inform our approach to constructing a functional CNN and integrate it with our implementation of GAN.

1.4. Data

We plan to use the Galaxy Zoo 2 dataset, which is publicly available on Kaggle. The dataset includes approximately 250,000 images labeled by volunteers based on characteristics like shape, size, and smoothness. If required, we plan to use Google Colab to accelerate the training process, particularly for GANs as they require significant computational power.

1.5. Approach

We propose a combined GAN-CNN model aimed at improving galaxy classification by augmenting limited datasets with synthetic images. The goal is to address the challenge of scarce labeled data without relying on large, deep networks. First, we use a SRGAN to improve the image resolution, then we train the cGAN on these enhanced images to produce additional images for each galaxy class. Afterwards, we will build the CNN using a feed-forward architecture, which has been proven effective in image classification tasks. We will modify the architecture as needed, depending on its performance on the galaxy data, ensuring that the CNN is robust and capable of learning from the images effectively. The CNN will first be trained on real data for baseline performance, then with both real and synthetic images to improve classification of underrepresented classes. Finally, we will evaluate and compare the model's performance with real data, cGAN-generated data, and SRGAN-enhanced images.

1.6. Evaluation Metrics

To evaluate the effectiveness of our model, we will use standard quantitative metrics such as accuracy, precision, recall, and F-1 score, which will help us fine-tune our model. These metrics will be plotted as a visual representation of the GAN-CNN performance for each class label. It will also represent the CNN's performance when trained on the original dataset versus the augmented dataset. We may also use statistical significance tests, such as t-tests, to ensure that performance improvements are not due to random chance. These evaluations will be replicated using cGAN and SR-GAN, providing a statistical comparison between the two models.

References

- [1] M. K. Cavanagh, K. Bekki, and B. A. Groves. Morphological classification of galaxies with deep learning: comparing 3-way and 4-way cnns. *Monthly Notices of the Royal Astronomical Society*, 506(1):659–676, 2021.
- [2] S. Dieleman, K. W. Willett, and J. Dambre. Rotation-invariant convolutional neural networks for galaxy morphology prediction. *Monthly Notices of the Royal Astronomical Society*, 450(2):1441–1459, 2015.
- [3] L. Fussell and B. Moews. Forging new worlds: high-resolution synthetic galaxies with chained generative adversarial networks. *Monthly Notices of the Royal Astronomical Society*, 485(3):3203–3214, 2019. [1](#)
- [4] S. Gharat and Y. Dandawate. Galaxy classification: a deep learning approach for classifying sloan digital sky survey images. *Monthly Notices of the Royal Astronomical Society*, 511(4):5120–5124, 2022. [1](#)
- [5] E. J. Kim and R. J. Brunner. Star–galaxy classification using deep convolutional neural networks. *Monthly Notices of the Royal Astronomical Society*, 464(4):4463–4475, 2016. [1](#)
- [6] Ofer Lahav, A. Nairn, Laerte Sodré, and M. C. Storrie-Lombardi. Neural computation as a tool for galaxy classification: methods and examples. *Monthly Notices of the Royal Astronomical Society*, 283(1):207–221, 1996. [1](#)
- [7] F. Lanusse, R. Mandelbaum, S. Ravanbakhsh, C.-L. Li, P. Freeman, and B. Póczos. Deep generative models for galaxy image simulations. *Monthly Notices of the Royal Astronomical Society*, 504(4):5543–5555, 2021. [1](#)
- [8] K. Sharma, A. Kembhavi, A. Kembhavi, T. Sivarani, S. Abraham, and K. Vaghmare. Application of convolutional neural networks for stellar spectral classification. *Monthly Notices of the Royal Astronomical Society*, 491(2):2280–2300, 2019.