

Semi-Supervised Active Learning Guided Detection System (SSALD)

Mid-way deliverables for CS-460

Aniket Nath¹² Diptarko Choudhury¹²

¹School of Physical Sciences
National Institute of Science Education and Research, HBNI

²School of Computer Sciences
National Institute of Science Education and Research, HBNI

March 10, 2023

Introduction I

Problem Statement

When two galaxies merge, their nuclear bulge hosting the SMBH, come closer and interact with each other, forming dual or even in certain cases multiple nuclei galaxies. Often, merger process triggers accretion disk around the SMBHs, leading to the formation of an Active Galactic Nuclei (AGN) [5]. We are trying to find dual AGNs in SDSS database using ML algorithms.

Related Works

- GOTHIC [2]: A Graph Boosted iterated Hill Climbing model to find two or more closely separated nucleus to detect dual or multiple active galaxies.
- Self-Supervised Representation learning for Astronomical Images[4].

Methodology

Steps

We aim to use an Active Learning guided self-supervised algorithm to solve our problem. In the following sections, we discuss our strategy in a step-by-step fashion.

- Supervised Learning
- Unsupervised and Semi-Supervised Learning
- Self-Supervised Learning
- Active Learning

VICReg

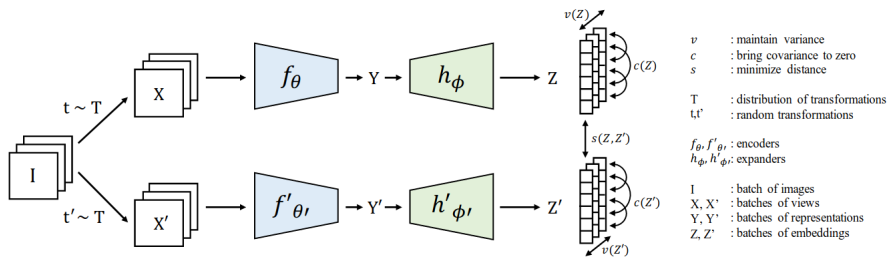


Figure: Schematics of VICReg Algorithm [6]

Results

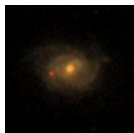


Figure: Query

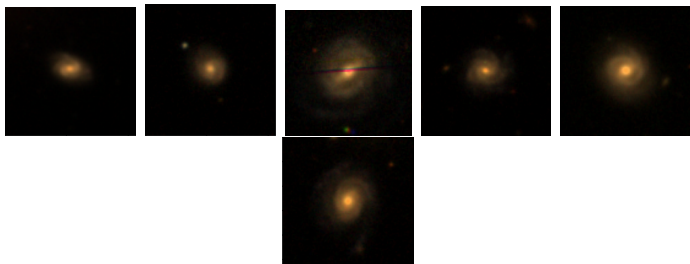


Figure: Results

Results

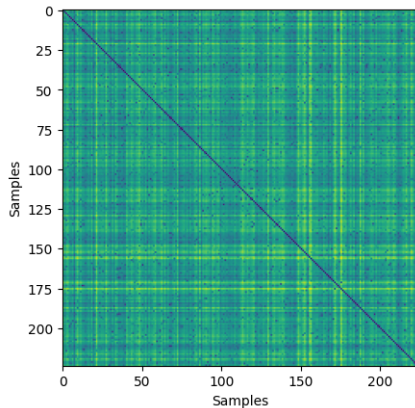


Figure: Correlations between the samples

Conclusion

Our findings till now are promising. The self-supervised backbone seems to be trained. The upcoming works are -

- Running the data miner to mine over 10 Million samples from the SDSS database.
- Finding a way to include spectroscopic data in our pipeline.
- Defining a metric to understand the VICReg training process better.
- Defining a custom loss to make the transition from SSL to Active Learning smoother.

References I

- [1] A. Stemo et al., "A Catalog of 204 Offset and Dual Active Galactic Nuclei (AGNs): Increased AGN Activation in Major Mergers and Separations under 4 kpc," *ApJ*, vol. 923, no. 1, p. 36, Dec. 2021, doi: 10.3847/1538-4357/ac0bbf.
- [2] A. Bhattacharya, N. C. P., M. Das, A. Paswan, S. Saha, and F. Combes, "Automated Detection of Double Nuclei Galaxies using GOTHIC and the Discovery of a Large Sample of Dual AGN." *arXiv*, Nov. 14, 2022. doi: 10.48550/arXiv.2011.12177.
- [3] K. Rubinur, M. Das, and P. Kharb, "Searching for dual active galactic nuclei," *Journal of Astrophysics and Astronomy*, vol. 39, p. 8, Feb. 2018, doi: 10.1007/s12036-018-9512-y.
- [4] M. A. Hayat, G. Stein, P. Harrington, Z. Lukić, and M. Mustafa, "Self-Supervised Representation Learning for Astronomical Images," *ApJL*, vol. 911, no. 2, p. L33, Apr. 2021, doi: 10.3847/2041-8213/abf2c7.
- [5] I. Shlosman, M. C. Begelman, and J. Frank, "The fuelling of active galactic nuclei," *Nature*, vol. 345, no. 6277, Art. no. 6277, Jun. 1990, doi: 10.1038/345679a0.
- [6] A. Bardes, J. Ponce, and Y. LeCun, "VICReg: Variance-Invariance-Covariance Regularization for Self-Supervised Learning." *arXiv*, Jan. 28, 2022. doi: 10.48550/arXiv.2105.04906.
- [7] David R Patton, Kieran D Wilson, Colin J Metrow, Sara L Ellison, Paul Torrey, Westley Brown, Maan H Hani, Stuart McAlpine, Jorge Moreno, Joanna Woo, Interacting galaxies in the IllustrisTNG simulations - I: Triggered star formation in a cosmological context, *Monthly Notices of the Royal Astronomical Society*, Volume 494, Issue 4, June 2020, Pages 4969–4985, <https://doi.org/10.1093/mnras/staa913>
- [8] Gimeno, G. N., Díaz, R. J., and Carranza, G. J., "Catalog of Double Nucleus Disk Galaxies", *ijz The Astronomical Journal*/ijz, vol. 128, no. 1, pp. 62–67, 2004. doi:10.1086/421371.
- [9] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017, doi: 10.1145/3065386.
- [10] A. Dosovitskiy et al., "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale." *arXiv*, Jun. 03, 2021. doi: 10.48550/arXiv.2010.11929.
- [11] Lupton, R., "Preparing Red-Green-Blue Images from CCD Data", *ijz Publications of the Astronomical Society of the Pacific*/ijz, vol. 116, no. 816, pp. 133–137, 2004. doi:10.1086/382245.

References II

[12] "Self-supervised learning: The dark matter of intelligence."

<https://ai.facebook.com/blog/self-supervised-learning-the-dark-matter-of-intelligence/> (accessed Mar. 10, 2023).

[13] "Illustration of active learning selection strategy." https://www.researchgate.net/figure/Illustration-of-active-learning-selection-strategy_fig2_360562440