Learning Dark Matter Representation From Strong Lensing Images Through Self-Supervision

Yashwardhan A. Deshmukh
Computer Engineering Department,
NMIMS University
Mumbai, India
yaashwardhan@gmail.com

Sergei Gleyzer
Department of Physics & Astronomy,
University of Alabama,
Tuscaloosa, AL 35401, USA
sgleyzer@ua.edu

Kartik Sachdev
RWTH Aachen University
Munich, Germany
kartik.sachdev@rwth-aachen.de

Michael W. Toomey
Center for Theoretical Physics,
Massachusetts Institute of Technology,
Cambridge, MA 02139, USA
mtoomey@mit.edu

Abstract

Gravitational lensing is one of the most important probes of dark matter and has recently seen a surge in applications of machine learning techniques. This is typically studied in the context of supervised learning, but given the upcoming influx of gravitational lensing data from Euclid and LSST, manual labeling for deep learning tasks has become an unsustainable approach. To address this challenge, self-supervised learning (SSL) emerges as a scalable solution. By leveraging unlabeled strong lensing data to learn feature representations, self-supervised models have the potential to enhance our understanding of dark matter via the effect of its substructure in strong lensing images. This work implements contrastive learning, Bootstrap Your Own Latent (BYOL), Simple Siamese (SimSiam), and self-distillation with no labels (DINO) using ResNet50 and Vision Transformer (ViT) networks, to acquire unsupervised embeddings for strong lensing images simulated for different dark matter models: ultra-light axions, cold dark matter, and halos without substructure. The learned representations of the encoder are fine-tuned using supervision and applied to classification and regression tasks which are also benchmarked against a fully supervised, ResNet50 baseline. Our results show that the self-supervised methods can consistently outperform their supervised counterparts.

1 Introduction

One of the most pressing problems in physics today is the identity of dark matter. Originally detected through its gravitational effects, there has yet to be any detection of dark matter due to interactions with the Standard Model. Indeed, numerous experiments have constrained a large swath of parameter space for the most compelling models [1, 2, 3, 4, 5, 6]. A promising avenue to study dark matter, and possibly take steps towards establishing its identity, has been the use of strong galaxy-galaxy lensing as a probe of dark matter substructure. This is interesting because the distribution and morphology of dark matter are intimately related to the underlying microphysical model. While this has been known for some time, recently machine learning algorithms have seen success in application to this problem, see for example [7, 8, 9, 10, 11]. In another direction, [12] explored a range of unsupervised techniques, which also found that integrating an unsupervised model with a supervised one resulted in

Machine Learning and the Physical Sciences Workshop, NeurIPS 2023.

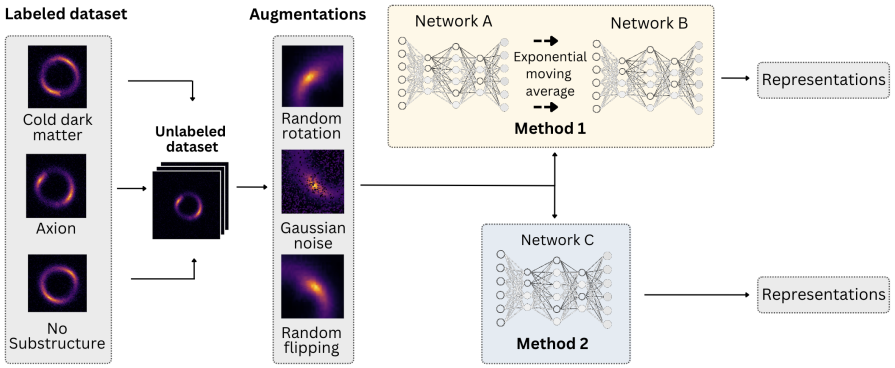

Figure 1: A schematic overview of implemented self-supervised methods (see text for discussion).

higher performance for the task of classification and helped in elucidating the nature of substructure. These approaches have all been successful in effectively learning meaningful embeddings that capture the intricate details of dark matter substructure.

Recently, in a slightly different context of lens finding, [13] applied semi-supervised learning for the identification of strong lenses. They used a data set that consisted of only a small amount of labeled training data. Indeed, as lensing data becomes more prevalent, the high cost and human bias associated with labeling make supervised learning increasingly unscalable. In a similar application, [14] used a self-supervised contrastive learning model trained to identify strong lensing images from other unlensed galaxies. The results showed that the self-supervised approach outperformed traditional supervised methods. To date, only simple Convolutional Neural Networks (CNNs) have been used with the contrastive learning SSL technique in application to strong lensing datasets. In this work we aim to test contrastive learning [15] in the context of the imprints of dark matter substructure on strong lensing images while using more complex architectures as well as to expand our scope to more modern SSL methods such as BYOL [16], SimSiam [17], and DINO [18]. We pre-train and finetune these methods for downstream tasks for 1) predicting the class of simulated gravitational lensing images and 2) determining the particle mass for axion dark matter directly from lensing images.

2 Methods

This section gives a brief introduction to the parts of self-supervised methods relevant to this work and the overall pretraining and finetuning procedure used in our benchmark.

2.1 Self-Supervised Pretraining

In the domain of unsupervised learning, contrastive learning [15] has become very popular due to its effectiveness in learning representations from large amounts of unlabelled data [19]. Specifically, it leverages the idea of learning similar and dissimilar representations of data by contrasting positive and negative examples. Figure 1 mentions a network C, which is an encoder along with a projected head and a linear neural network layer, implemented to learn representations through solving rotation and Gaussian noise augmentation pretext tasks. We experimented with this SSL method by using two encoders, ResNet50 and a custom Vision Transformer (ViT) [20] with 12 transformer blocks and a kernel size of 32. Normalized Temperature-scaled Cross Entropy (NT-Xent) loss[21], designed to minimize the distance between similar items and maximize the distance between dissimilar items in the embedded space, was used for contrastive learning.

Bootstrap Your Own Latent (BYOL) [16] is a recent advancement in the domain of unsupervised learning. Unlike traditional contrastive learning methods, BYOL does not require the positive and negative pairs of images, it rather trains two networks parallelly, the target network and the online networks; represented as network A (online) and network B (target) as seen in Figure 1. The target network's parameters are updated as a moving average of the online network's parameters. For BYOL, we implemented random flipping and random rotation as an augmentation.

Self-distillation with no labels (DINO) [18] is a modern self-supervised method in which we provide two different augmentations (local and global views) to the teacher and student networks, as seen in Figure 1 as network A and network B respectively. The network is trained to reduce the cross-entropy loss for matching the similarity between the two views. Similar to BYOL, network B is an exponential moving average of network A.

Another self-supervised method we implemented was Simple Siamese (SimSiam) [17]. In contrast to other learning techniques, it does not require a momentum encoder, memory bank of negative samples, or large batch sizes that are generally used to avoid the risk of collapsing solutions [22]. It uses a single encoder network, and the same image with two randomly augmented views which are matched using negative cosine similarity. We used `keras` [23] to train contrastive learning and BYOL models on two NVIDIA T4 GPUs and `lightly` [24] to train DINO and SimSiam models on NVIDIA A100 GPUs.

2.2 Supervised Finetuning

Our goal is to compare fine-tuned self-supervised methods with their supervised baseline on simulated strong lensing images for two tasks (a) image classification to identify different types of dark matter candidates, and (b) regression to detect the mass of the axion from a given image. Initially, an encoder network is trained without labels to learn representations of the data, using one of the above-mentioned SSL methods at a time. A suitable neural network featuring fully connected layers is employed on top of the encoder for supervised fine-tuning. For classification tasks, the network uses three neurons in the final layer, whereas for regression tasks, it utilizes a single neuron in the last layer. We test our methods for both tasks on respective holdout test datasets consisting of 15,000 images, the results of which are mentioned in Section 3.

2.3 Datasets

In this analysis, we have constructed two data sets: Model A and Model B which we have made to resemble mock simulations of galaxy-galaxy strong lensing images from observations with HST- and Euclid-like surveys, respectively. We generate our simulations with the publicly available package `lenstronomy`[25] where we create single channel images sized 64×64. Our background galaxies, which are lensed, are modeled with a Sersic light profile. On the dark matter side, we generate three classes. Our first-class represents lensing from standard cold dark matter (CDM) where our main halo, which is modeled with a spherical isothermal profile, includes dark matter subhaloes which are drawn from the standard subhalo mass distribution function (see [7] for further details). Beyond CDM, we also simulate the lensing signature of ultra-light axion dark matter. Specifically, our axion simulations correspond to a particle mass $\sim 10^{-23}$ eV where the substructure is highly suppressed and where the main substructure observable is topological defects in the main halo; namely vortex substructure [7]. The final dark matter model we study consists of one with the absence of any substructure. Despite not being a realistic model, i.e. observations clearly preclude this possibility, since we are working with simulations, it serves as a useful scenario to contrast with other dark matter realizations in testing machine learning models – particularly because its signature is not expected to be degenerate with CDM or the axion. When generating our simulations we construct 30,000 images per class and an additional test set of 5,000 images per class.

3 Results

We now present the results for the classification of different models of dark matter. The results for Model A and Model B are presented in Tables 1 and 2, respectively. We trained all the supervised baselines with the same augmentation and training scheme as used in the corresponding supervised finetuning with random weight intialization instead of pretrained self-supervised weights. From these results, it is clear that, with the exception of one example, our self-supervised algorithms improve the performance over the supervised alone. In the one exception, ResNet50 with BYOL for Model B, the performance is effectively equivalent between the supervised and unsupervised iterations.

Going beyond classification, we also study the application of self-supervised algorithms for the task of regressing on the mass of the axion. That is, we take the images that we have simulated and train our algorithms to output the known mass of the axion. We present our results in Fig.2 for both our

Table 1: Model A Test Set Classification Results (Macro AUC)

Method	Encoder	Supervised	Self-Supervised with finetuning
Contrastive (Rotational Pretext)	ResNet50	0.960	**0.993**
Contrastive (Rotational Pretext)	ViT	0.977	**0.990**
Contrastive (Gaussian Noise Pretext)	ResNet50	0.960	**0.993**
Contrastive (Gaussian Noise Pretext)	ViT	0.977	**0.993**
BYOL	ResNet50	0.990	**0.993**
SimSiam	ResNet50	0.983	**0.995**
DINO	ResNet50	0.996	**0.997**

Table 2: Model B Test Set Classification Results (Macro AUC)

Method	Encoder	Supervised	Self-Supervised with finetuning
Contrastive (Rotational Pretext)	ResNet50	0.967	**0.997**
Contrastive (Rotational Pretext)	ViT	0.983	**0.987**
Contrastive (Gaussian Noise Pretext)	ResNet50	0.967	**0.997**
Contrastive (Gaussian Noise Pretext)	ViT	0.983	**0.990**
BYOL	ResNet50	**0.963**	0.960
SimSiam	ResNet50	0.984	**0.996**
DINO	ResNet50	0.998	**0.998**

baseline supervised algorithm, ResNet50, and for various self-supervised algorithms. In the case of Model A, we find that all self-supervised algorithms perform about the same but our baseline algorithm is noticeably worse, particularly for axion masses $\sim 10^{-23}$ eV. For Model B, the baseline model performance is significantly worse. In this case, we find that the self-supervised Contrastive (Gaussian Noise pretext) model does noticeably worse than the other self-supervised algorithms, though its still performs better than our baseline model. Despite this, the three other self-supervised algorithms demonstrate great performance at reconstructing the mass of the axion over supervised baselines from our simulations.

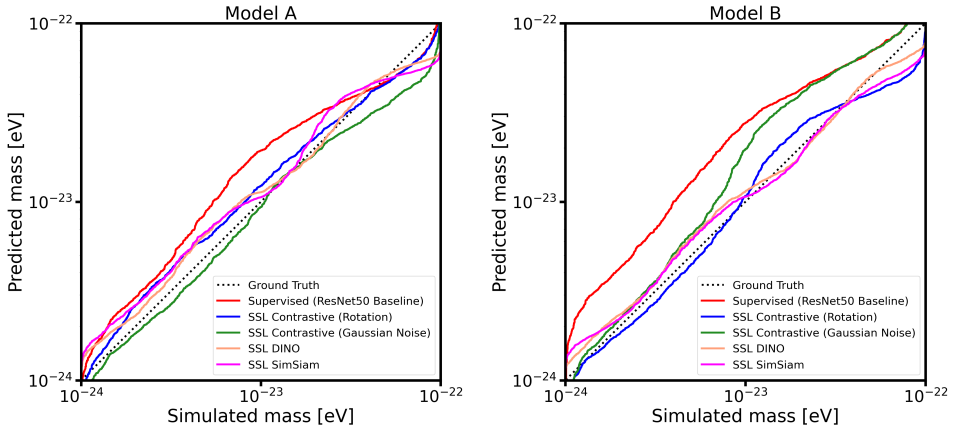

Figure 2: Predicted mass versus simulated mass for Model A (left panel) and Model B (right panel).

4 Conclusion and Limitations

In this work, we found that the learned representation showed robust performance on both classification and regression with the results outperforming (in one case, only on par with) the supervised learning techniques. Our work is an attempt to further the research efforts in applying self-supervised learning to astrophysical data sets which typically involve large quantities of unlabeled data. Of course, the specific example that we consider is somewhat unrealistic, in that dark matter in reality would only be one thing, but our examples serve as a concrete example. A learned representation is generic and could be reused with a task-specific prediction head for different applications, reducing the efforts to train a whole network from scratch for newer tasks. As of now, our work only focuses on a handful of self-supervised methods and neural network architectures. To date, there are more

recent networks architectures like Convolutional vision Transformer (CvT) [26], Swin Transformer [27], CrossFormer [28] to name a few and several options for self-supervised learning techniques such as VICReg [29], EsViT [30] etc., but we leave the implementation of these for future work.

5 Acknowledgments

Y. D. and K.S. were participants in the Google Summer of Code 2023 program. S.G. was supported in part by U.S. National Science Foundation award No. 2108645. M.W.T. was partially supported in part by U.S. National Science Foundation award No. 2108866. Portions of this work were conducted in MIT's Center for Theoretical Physics and partially supported by the U.S. Department of Energy under grant Contract Number DE-SC0012567.

References

[1] A. K. Drukier, Katherine Freese, and D. N. Spergel. Detecting Cold Dark Matter Candidates. *Phys. Rev.*, D33:3495–3508, 1986.

[2] Mark W. Goodman and Edward Witten. Detectability of Certain Dark Matter Candidates. *Phys. Rev.*, D31:3059, 1985. [,325(1984)].

[3] D. S. Akerib et al. Results from a search for dark matter in the complete LUX exposure. *Phys. Rev. Lett.*, 118(2):021303, 2017.

[4] Xiangyi Cui et al. Dark Matter Results From 54-Ton-Day Exposure of PandaX-II Experiment. *Phys. Rev. Lett.*, 119(18):181302, 2017.

[5] E. Aprile et al. Dark Matter Search Results from a One Ton-Year Exposure of XENON1T. *Phys. Rev. Lett.*, 121(11):111302, 2018.

[6] Morad Aaboud et al. Constraints on mediator-based dark matter and scalar dark energy models using $\sqrt{s} = 13$ TeV pp collision data collected by the ATLAS detector. *JHEP*, 05:142, 2019.

[7] Stephon Alexander, Sergei Gleyzer, Evan McDonough, Michael W. Toomey, and Emanuele Usai. Deep learning the morphology of dark matter substructure. *The Astrophysical Journal*, 893(1):15, apr 2020.

[8] Siddharth Mishra-Sharma and Ge Yang. Strong Lensing Source Reconstruction Using Continuous Neural Fields. In *39th International Conference on Machine Learning Conference*, 6 2022.

[9] Gemma Zhang, Siddharth Mishra-Sharma, and Cora Dvorkin. Inferring subhalo effective density slopes from strong lensing observations with neural likelihood-ratio estimation. *Monthly Notices of the Royal Astronomical Society*, 517(3):4317–4326, 10 2022.

[10] Thuruthipilly, Hareesh, Zadrozny, Adam, Pollo, Agnieszka, and Biesiada, Marek. Finding strong gravitational lenses through self-attention - study based on the bologna lens challenge. *A&A*, 664:A4, 2022.

[11] Ana Diaz Rivero and Cora Dvorkin. Direct Detection of Dark Matter Substructure in Strong Lens Images with Convolutional Neural Networks. *Phys. Rev. D*, 101(2):023515, 2020.

[12] Stephon H. S. Alexander, Sergei Gleyzer, Hanna Parul, Pranath Reddy, Michael W. Toomey, Emanuele Usai, and Ryker von Klar. Decoding dark matter substructure without supervision. *arXiv: Cosmology and Nongalactic Astrophysics*, 2020.

[13] Stephen Sheng, Keerthi Vasan G C, Chi Po P Choi, James Sharpnack, and Tucker Jones. An unsupervised hunt for gravitational lenses. In Gustau Camps-Valls, Francisco J. R. Ruiz, and Isabel Valera, editors, *Proceedings of The 25th International Conference on Artificial Intelligence and Statistics*, volume 151 of *Proceedings of Machine Learning Research*, pages 9827–9843. PMLR, 28–30 Mar 2022.

[14] George Stein, Jacqueline Blaum, Peter Harrington, Tomislav Medan, and Zarija Lukić. Mining for strong gravitational lenses with self-supervised learning. *The Astrophysical Journal*, 932(2):107, jun 2022.

[15] R. Hadsell, S. Chopra, and Y. LeCun. Dimensionality reduction by learning an invariant mapping. In *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*, volume 2, pages 1735–1742, 2006.

[16] Jean-Bastien Grill, Florian Strub, Florent Altch'e, Corentin Tallec, Pierre H. Richemond, Elena Buchatskaya, Carl Doersch, Bernardo Ávila Pires, Zhaohan Daniel Guo, Mohammad Gheshlaghi Azar, Bilal Piot, Koray Kavukcuoglu, Rémi Munos, and Michal Valko. Bootstrap your own latent: A new approach to self-supervised learning. *ArXiv*, abs/2006.07733, 2020.

[17] Xinlei Chen and Kaiming He. Exploring simple siamese representation learning. *arXiv preprint arXiv:2011.10566*, 2020.

[18] Mathilde Caron, Hugo Touvron, Ishan Misra, Hervé Jégou, Julien Mairal, Piotr Bojanowski, and Armand Joulin. Emerging properties in self-supervised vision transformers. *Proceedings of the International Conference on Computer Vision (ICCV)*, 2021.

[19] Mathilde Caron, Ishan Misra, Julien Mairal, Priya Goyal, Piotr Bojanowski, and Armand Joulin. Unsupervised learning of visual features by contrasting cluster assignments. *34th Conference on Neural Information Processing Systems (NeurIPS 2020)*, abs/2006.09882, 2020.

[20] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, Jakob Uszkoreit, and Neil Houlsby. An image is worth 16x16 words: Transformers for image recognition at scale. *ArXiv*, abs/2010.11929, 2020.

[21] Kihyuk Sohn. Improved deep metric learning with multi-class n-pair loss objective. In D. Lee, M. Sugiyama, U. Luxburg, I. Guyon, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 29. Curran Associates, Inc., 2016.

[22] Tobias Höppe, Agnieszka Miszkurka, and Dennis Bogatov Wilkman. [re] understanding self-supervised learning dynamics without contrastive pairs. In *ML Reproducibility Challenge 2021 (Fall Edition)*, 2022.

[23] Francois Chollet et al. Keras, 2015.

[24] Igor Susmelj, Matthias Heller, Philipp Wirth, Jeremy Prescott, and Malte Ebner et al. Lightly. *GitHub. Note: https://github.com/lightly-ai/lightly*, 2020.

[25] Simon Birrer, Adam Amara, and Alexandre Refregier. Gravitational lens modeling with basis sets. *The Astrophysical Journal*, 813(2):102, nov 2015.

[26] Haiping Wu, Bin Xiao, Noel Codella, Mengchen Liu, Xiyang Dai, Lu Yuan, and Lei Zhang. Cvt: Introducing convolutions to vision transformers. *arXiv preprint arXiv:2103.15808*, 2021.

[27] Ze Liu, Yutong Lin, Yue Cao, Han Hu, Yixuan Wei, Zheng Zhang, Stephen Lin, and Baining Guo. Swin transformer: Hierarchical vision transformer using shifted windows. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2021.

[28] Wenxiao Wang, Lu Yao, Long Chen, Binbin Lin, Deng Cai, Xiaofei He, and Wei Liu. Crossformer: A versatile vision transformer hinging on cross-scale attention, 2021.

[29] Adrien Bardes, Jean Ponce, and Yann LeCun. VICReg: Variance-invariance-covariance regularization for self-supervised learning. In *International Conference on Learning Representations*, 2022.

[30] Jin Li, Yaoming Wang, XIAOPENG ZHANG, Yabo Chen, Dongsheng Jiang, Wenrui Dai, Chenglin Li, Hongkai Xiong, and Qi Tian. Progressively compressed auto-encoder for self-supervised representation learning. In *The Eleventh International Conference on Learning Representations*, 2023.