Decoding Dark Matter Substructure without Supervision

Stephon Alexander and Michael W. Toomey
Department of Physics and Brown Theoretical Physics Center
Brown University
Providence, RI 02912

Sergei Gleyzer, Hanna Parul, and Ryker Von Klar
Department of Physics & Astronomy
University of Alabama
Tuscaloosa, AL 35401

Pranath Reddy
Birla Institute of Technology
Science, Pilani - Hyderabad Campus
Telangana, India

Emanuele Usai
Department of Physics
Brown University
Providence, RI 02912

Abstract

The identity of dark matter remains one of the most pressing questions in physics today. While many promising dark matter candidates have been put forth over the last half-century, to date the true identity of dark matter remains elusive. While it is possible that one of the many proposed candidates may turn out to be dark matter, it is at least equally likely that the correct physical description has yet to be proposed. To address this challenge, novel applications of machine learning can help physicists gain insight into the dark sector from a *theory agnostic* perspective. In this work we demonstrate the use of unsupervised machine learning techniques to infer the presence of substructure in dark matter halos using galaxy-galaxy strong lensing simulations in a proof-of-principle application.

1 Introduction

Since the discovery of dark matter, physicists have been searching the entirety of cosmic history, from experiments at colliders to observations of the cosmic microwave background, for fingerprints that might reveal the identity of dark matter. Most models work under the assumption that the dark sector interacts, typically only very weakly, with the standard model, e.g. weakly interacting massive particles (WIMPs) [1] and axions [2–4]. However, these models have avoided efforts at direct and indirect detection [5–22], including searches at colliders [23, 24]. To date the only evidence for dark matter comes from its gravitational interactions [25–28]. In light of this, it makes sense to explore avenues to identify dark matter via its gravitational imprints.

Third Workshop on Machine Learning and the Physical Sciences (NeurIPS 2020), Vancouver, Canada.

A promising means to identify the nature of dark matter is to study substructure in dark matter halos, i.e. overdensities within the larger dark matter halo. Different models come into their own on *subgalactic* scales which allow current and future observational programs to start to constrain potential dark matter candidates [29–31]. While it is possible to study larger substructures such as ultra-faint dwarf galaxies (see for example [32]), subhalos on smaller scales suffer from suppressed star formation, making manifest the need for a gravitational probe. Promising directions to identify substructure gravitationally include tidal streams [33–36] and astrometric observations [37–40]. Another avenue to consider is strong gravitational lensing which has seen encouraging results in detecting the existence of substructure from strongly lensed quasars [41–43], high resolutions observations with the Atacama Large Millimeter/submillimeter Array [44] and extended lensing images [45–51].

Recently, promising results have been achieved with supervised machine learning algorithms for identifying dark matter substructure properties with simulated galaxy-galaxy strong lensing images. These include applications of convolutional neural networks (CNNs) for inference of population level properties of substructure [52], classification of halos with and without substructure [53, 54] and between dark matter models with disparate substructure morphology [53], as well as classifying between lenses with different subhalo mass cut-offs [55]. In a similar spirit to these strong lensing studies, Ref. [56] used a CNN to classify simulated astrometic signatures of a population of quasars as being consistent with the presence of dark matter substructure in the Milky Way. Another interesting direction, as pointed out in Ref. [53], is the application of unsupervised machine learning techniques to this challenge.

In this work we present a proof-of-principle application for using unsupervised machine learning in the context of identifying dark matter substructure with strong gravitational lensing. Specifically, we utilize autoencoders, a type of unsupervised machine learning algorithm for anomaly detection (AD). We show that after training adversarial [57], variational [58], and deep convolutional autoencoders [59] on simulated strong lensing images *without* substructure, it is possible to identify data *with* substructure as anomalous for further analysis. The application of an unsupervised machine learning algorithm can serve as a first step in an analysis pipeline of strong lensing systems. Data which are flagged as anomalous can be followed-up with a Bayesian likelihood analysis or a dedicated supervised machine learning algorithm to identify the type of substructure observed.

2 Simulation of Strong Lensing Images

There is currently only a small sample of (real) strong galaxy-galaxy lensing images. Nonetheless, simulated images can serve as a proof of principle data set to benchmark different algorithms before the influx of data from Euclid and the Vera C. Rubin Observatory where we can expect thousands of high quality lensing images [60, 61]. In this work the strong lensing images were simulated with the `PyAutoLens` package [62, 63]. We simulate data for strong lensing with no substructure and substructure from two disparate types of dark matter - subhalos of CDM and vortices of superfluid dark matter. Images are composed 150×150 pixels with a pixel scale of $0.5''$/pixel. Informed by real strong galaxy-galaxy lensing images, we further include background and noise in our simulations such that the lensing arcs have a maximum signal-to-noise ratio (SNR) of ~ 20 [64]. We further include modifications induced by a point-spread function approximated by an Airy disk with a first zero-crossing at $\sigma_{\rm psf} \lesssim 1''$.

We consider two models, Model A and Model B, for this analysis. The difference between the two models is that all simulated images for Model A are held at fixed redshift while Model B allows the lensed and lensing galaxy redshifts to float over a range of values. An additional difference is the SNR in both models. Images for Model A have SNR ≈ 20 where Model B is constructed such that simulations produce images where $10 \lesssim \mathrm{SNR} \lesssim 30$. It is clear then that Model B is a noticeably more difficult training set. In practice though, since the redshifts of galaxies can be measured, one can approximate Model A by binning lensing images by redshift prior to analysis.

3 Network Training and Performance Metrics

For training we use 25,000 samples with no substructure and 2,500 validation samples per class for training the unsupervised models. The models are implemented using the `PyTorch` [65] package and

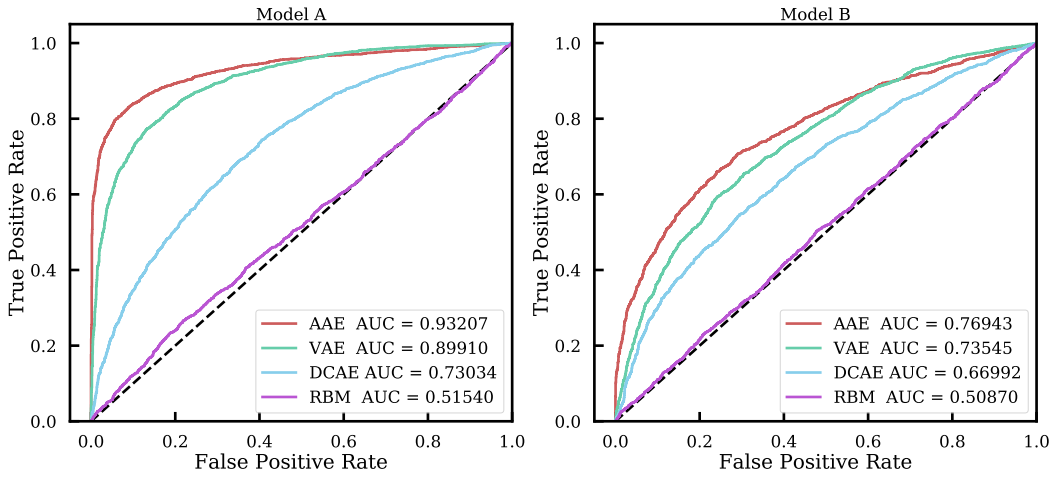

Figure 1: ROC-AUC curve for unsupervised algorithms. Plots on left correspond to Model A, plots on the right correspond to Model B.

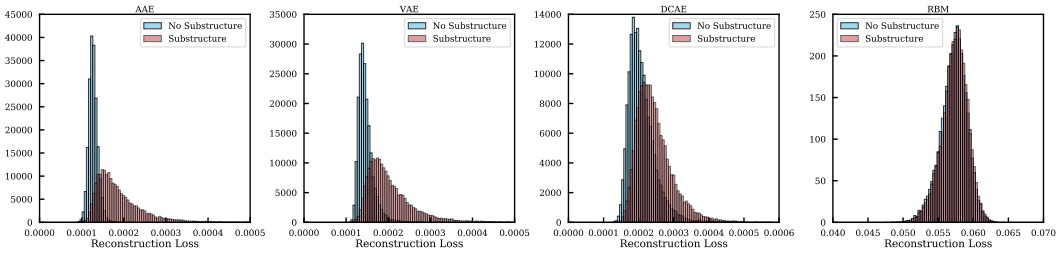

Figure 2: Reconstruction loss for architectures with Model A. Left to right: AAE, VAE, DCAE, and RBM.

are run on a single NVIDIA Tesla K80 GPU for 500 epochs. We utilize the area under the receiver operating characteristic (ROC) curve (AUC) as a metric for classifier performance for all our models. For unsupervised models the ROC values are calculated for a set threshold of the reconstruction loss. Additionally, for unsupervised models we use the Wasserstein distance, a measure of similarity between probability distributions, as an additional performance metric.

4 Results

For this work four different unsupervised architectures were studied in the context of anomaly detection - a deep convolutional autoencoder (DCAE), convolutional variational autoencoder (VAE), an adversarial autoencoder (AAE) and a restricted Boltzmann machine (RBM). The ROC curves for all four architectures applied to Models A & B are shown in Figure 1.

As expected for an algorithm whose design is not optimized for image inputs, the RBM model has the poorest performance and proves to be unsuccessful in learning anything significant from our training set. This is evident from the AUC values slightly above 0.5 for both Models A and B. Although the RBM model does not distinguish well between images with differing substructure, its low reconstruction loss implies that it does succeed in reconstructing the general morphology of the lens.

Next, we turn to the autoencoder models. First, we consider the DCAE model. With an AUC of ≈ 0.730 for Model A, our DCAE shows good discriminating power at distinguishing images with and without substructure. The improvement over the RBM is also clearly visible in the distribution of reconstruction losses in Figure 2. However, the added complexity of the Model B data set significantly degrades DCAE performance where the AUC decreases to ≈ 0.670.

	Architecture	AUC	$\mathbf{W}_1$
Model A			
	AAE	0.93207	0.22112
	VAE	0.89910	0.22533
	DCAE	0.73034	0.26566
	RBM	0.51054	1.27070
Model B			
	AAE	0.76943	0.15563
	VAE	0.73545	0.16617
	DCAE	0.66992	0.23653
	RBM	0.50870	1.25804

Table 1: Performance of architectures used in this analysis. W_1 is the average 1^{st} Wasserstein distance for images without substructure.

Next we consider the variational autoencoder (VAE) which is trained with MSE and KL divergence loss functions. Furthermore, we implement KL cost annealing where only the reconstruction loss is used during the first 100 epochs and then the weight of KL divergence loss is gradually increased from 0 to 1. With an AUC ≈ 0.899 our VAE achieves performance much closer to that of *ResNet* for Model A, but it also suffers from a lower AUC of ≈ 0.769 for Model B in a more difficult training scenario. The increase in performance relative to the DCAE, and RBM is easy to see from the distribution of reconstruction losses in Figure 2 where images with substructure have noticeably higher average reconstruction loss.

Finally, we consider an adversarial autoencoder (AAE) model trained on samples with no substructure with MSE as reconstruction loss and cross-entropy loss for the discriminator network. The AAE reaches top performance of our architectures with an AUC of ≈ 0.903 for Model A and ≈ 0.769 for Model B. As with the VAE model, the reconstruction loss, see Figure 2, makes it evident how much better the architecture performance is compared to the others.

As an additional performance metric for unsupervised architectures we calculate the average Wasserstein distance from our validation data set - results are compiled in Table 1. We first do this for the no substructure class. As the Wasserstein distance is a geodesic between distributions, smaller values correspond to distributions that are more similar. We find that the AAE and VAE achieve the best performance and that of the DCAE is $\sim 20\%$ larger by comparison. The ability of the RBM to reconstruct the inputs is significantly degraded compared to the autoencoder models. Together, these results are consistent with calculated AUC values for the architectures.

It is interesting to also consider the Wasserstein distance for data with substructure, i.e. a class that the models were not trained on. All the autoencoders consistently have smaller distances for no substructure compared to that calculated from substructure. Furthermore, the AAE and VAE seem to be slightly better at reconstructing images for subhalo substructure versus vortices. This may be a result of higher symmetry from the effects of subhalo substructure.

An interesting question to consider is whether unsupervised algorithms trained on the no substructure class have the ability to disentangle images with vortices from those with subhalos. From the distribution of reconstruction losses for vortices and subhalos separately, for both models, there isn't significant constraining power between the two types of substructure. As an additional check, we trained the AAE model on the subhalo data to establish if there is any ability to distinguish between the two different dark matter substructure models. We find that the autoencoder models are not able to distinguish well between the two substructure classes. Interestingly, we find the average reconstruction loss for the no substructure class to be lower than that for subhalos. We affirm this result by calculating the Wassertein distances for our three classes of substructure, finding the no substructure class to yield the lowest values for both models. This implies that the unsupervised models used as anomaly detectors are able to accurately distinguish between the no substructure and substructure scenarios, but are not able to accurately distinguish between the different types of substructures (ie. differing types of anomalies). This is of course, a more naturally suited task for dedicated supervised machine learning algorithms. One can further imagine an analysis pipeline where unsupervised models are first used to identify potential anomalies for further analysis that can

be performed with dedicated supervised machine learning classification algorithms or other more traditional approaches.

5 Discussion & Conclusion

In this work we show that dark matter substructure can serve as a useful probe to constrain models of dark matter. Concretely, different models of dark matter can lead to disparate morphology which can manifest themselves observationally via their imprint on extended lensing arcs. It has been previously shown that machine learning may even have the power to identify dark matter models based on signatures unique to substructure [53]. In this work we have expanded on this idea, now utilizing the power and versatility of unsupervised machine learning algorithms to identify the presence of dark matter substructure in lensing images.

We have extended beyond the results from [53], where *ResNet* was trained as a supervised multi-class classifier, to unsupervised models. Specifically, we have considered a restricted Boltzmann machine, deep convolutional autoencoder, convolutional variational autoencoder, and an adversarial autoencoder on two data sets with increasing complexity.

We found that the RBM model does not perform well for this task. Though a construction of a deep belief network [66] may help to disentangle substructures, likely the downfall of the RBM can be traced back to the loss of spatial information when the data is compressed to one dimension for training. Inspired by our earlier results, we trained three types of convolutional autoencoders on data containing no substructure. Both our AAE and VAE performed very well in distinguishing between images with substructure on Model A, achieving high AUC scores. The added complexity of Model B proved more challenging for both architectures - though they still have strong discriminating power. Lastly, our implementation of a DCAE shows moderate success on both data sets.

By calculating the average Wasserstein distances for our data we further quantified the performance of our architectures. We found that the AAE and VAE produced the most faithful reconstruction of the input images, the DCAE is moderately degraded, while the RBM is significantly worse at reconstruction of the data. Furthermore, by calculating the distances for the data with substructures, we determined that Wasserstein distance is on average larger for these data - consistent with our expectations. We also find that the AAE and VAE give a smaller geodesic distance between the input and reconstruction for images with subhalos compared to those with vortices. An explanation for this can be attributed to higher symmetry for a population of subhalos, i.e. its lensing signature is more consistent with no substructure.

The results we obtained support the conclusion that unsupervised models can be a useful first step in an analysis pipeline to establish the most anomalous (potentially promising) sources for further follow up with dedicated supervised machine algorithms for further classification or with standard Bayesian likelihood analysis techniques. The advantage of such a first step is that no model-specific knowledge was assumed while looking for anomalous sources.

Broader Impact

This work serves to augment the understanding and application of machine learning in cosmology - which is still very much in its initial stages. This work serves to increase the accessibility to those interested in applications of machine learning for strong lensing applications around the globe as our simulation dataset and analysis pipeline is open sourced. Given the computational requirements of our implementation, those who have limited access to computing power may be at a disadvantage.

References

[1] Gary Steigman and Michael S. Turner. Cosmological Constraints on the Properties of Weakly Interacting Massive Particles. *Nucl. Phys. B*, 253:375–386, 1985.

[2] John Preskill, Mark B. Wise, and Frank Wilczek. Cosmology of the Invisible Axion. *Phys. Lett.*, B120:127–132, 1983. [,URL(1982)].

[3] L. F. Abbott and P. Sikivie. A Cosmological Bound on the Invisible Axion. *Phys. Lett.*, B120:133–136, 1983. [,URL(1982)].

[4] Michael Dine and Willy Fischler. The Not So Harmless Axion. *Phys. Lett.*, B120:137–141, 1983. [,URL(1982)].

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

[6] Mark W. Goodman and Edward Witten. Detectability of Certain Dark Matter Candidates. *Phys. Rev.*, D31:3059, 1985.

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

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

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

[10] Francis Froborg and Alan R Duffy. Annual Modulation in Direct Dark Matter Searches. *arXiv e-prints*, page arXiv:2003.04545, March 2020.

[11] Fermi LAT Collaboration. Limits on dark matter annihilation signals from the Fermi LAT 4-year measurement of the isotropic gamma-ray background. 2015(9):008, September 2015.

[12] Alex Geringer-Sameth, Savvas M. Koushiappas, and Matthew G. Walker. Comprehensive search for dark matter annihilation in dwarf galaxies. 91(8):083535, April 2015.

[13] A. Albert, R. Alfaro, C. Alvarez, J. D. Álvarez, R. Arceo, and et al. Dark Matter Limits from Dwarf Spheroidal Galaxies with the HAWC Gamma-Ray Observatory. *ApJ*, 853(2):154, February 2018.

[14] VERITAS Collaboration. Dark matter constraints from a joint analysis of dwarf Spheroidal galaxy observations with VERITAS. *Phys. Rev. D*, 95(8):082001, April 2017.

[15] Javier Rico. Gamma-Ray Dark Matter Searches in Milky Way Satellites—A Comparative Review of Data Analysis Methods and Current Results. *Galaxies*, 8(1):25, March 2020.

[16] MAGIC Collaboration. Limits to dark matter annihilation cross-section from a combined analysis of MAGIC and Fermi-LAT observations of dwarf satellite galaxies. 2016(2):039, February 2016.

[17] IceCube Collaboration. Search for Neutrinos from Dark Matter Self-Annihilations in the center of the Milky Way with 3 years of IceCube/DeepCore. *arXiv e-prints*, page arXiv:1705.08103, May 2017.

[18] The Super-Kamiokande Collaboration. Search for neutrinos from annihilation of captured low-mass dark matter particles in the Sun by Super-Kamiokande. *arXiv e-prints*, page arXiv:1503.04858, March 2015.

[19] N. Du, N. Force, R. Khatiwada, E. Lentz, R. Ottens, L. J. Rosenberg, G. Rybka, G. Carosi, N. Woollett, D. Bowring, A. S. Chou, A. Sonnenschein, W. Wester, C. Boutan, N. S. Oblath, R. Bradley, E. J. Daw, A. V. Dixit, J. Clarke, S. R. O'Kelley, N. Crisosto, J. R. Gleason, S. Jois, P. Sikivie, I. Stern, N. S. Sullivan, D. B. Tanner, G. C. Hilton, and ADMX Collaboration. Search for Invisible Axion Dark Matter with the Axion Dark Matter Experiment. *Phys. Rev. Lett*, 120(15):151301, April 2018.

[20] Peter W. Graham, Igor G. Irastorza, Steven K. Lamoreaux, Axel Lindner, and Karl A. van Bibber. Experimental Searches for the Axion and Axion-Like Particles. *Annual Review of Nuclear and Particle Science*, 65:485–514, October 2015.

[21] Kristjan Kannike, Martti Raidal, Hardi Veermäe, Alessandro Strumia, and Daniele Teresi. Dark Matter and the XENON1T electron recoil excess. *arXiv e-prints*, page arXiv:2006.10735, June 2020.

[22] Jatan Buch, Manuel A. Buen-Abad, JiJi Fan, and John Shing Chau Leung. Galactic Origin of Relativistic Bosons and XENON1T Excess. *arXiv e-prints*, page arXiv:2006.12488, June 2020.

[23] 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.

[24] Sirunyan, A. M. and Tumasyan, A. and Adam, W. and Asilar, E. and Bergauer, T., and et al. Search for new physics in the monophoton final state in proton-proton collisions at $\sqrt{\{s\}}$=13 TeV. *Journal of High Energy Physics*, 2017(10):73, October 2017.

[25] Planck Collaboration. Planck 2015 results. XIII. Cosmological parameters. *A&A*, 594(A13):63, 2016. arXiv.

[26] L. Anderson, E. Aubourg et al. The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: baryon acoustic oscillations in the Data Releases 10 and 11 Galaxy samples . *MNRAS*, 441(1):24–62, 2014. MNRAS.

[27] C. Heymans, L. van Waerbeke et al. CFHTLenS: the Canada–France–Hawaii Telescope Lensing Survey . *MNRAS*, 427(1):146–166, 2012. MNRAS.

[28] Douglas Clowe, Anthony Gonzalez, and Maxim Markevitch. Weak lensing mass reconstruction of the interacting cluster 1E0657-558: Direct evidence for the existence of dark matter. *Astrophys. J.*, 604:596–603, 2004.

[29] Matthew R. Buckley and Annika H. G. Peter. Gravitational probes of dark matter physics. *Phys. Rept.*, 761:1–60, 2018.

[30] Alex Drlica-Wagner et al. Probing the Fundamental Nature of Dark Matter with the Large Synoptic Survey Telescope. *arXiv e-prints*, page arXiv:1902.01055, February 2019.

[31] Joshua Simon et al. Testing the Nature of Dark Matter with Extremely Large Telescopes. *Bull. Am. Astron. Soc*, 51(3):153, May 2019.

[32] A. Drlica-Wagner et al. Eight Ultra-faint Galaxy Candidates Discovered in Year Two of the Dark Energy Survey. *Astrophys. J.*, 813(2):109, 2015.

[33] Wai-Hin Wayne Ngan and Raymond G. Carlberg. Using gaps in N-body tidal streams to probe missing satellites. *Astrophys. J.*, 788:181, 2014.

[34] R. G. Carlberg. Modeling GD-1 Gaps in a Milky Way Potential. 820(1):45, March 2016.

[35] Jo Bovy. Detecting the Disruption of Dark-Matter Halos with Stellar Streams. *Phys. Rev. Lett.*, 116(12):121301, March 2016.

[36] Denis Erkal, Vasily Belokurov, Jo Bovy, and Jason L. Sand ers. The number and size of subhalo-induced gaps in stellar streams. *MNRAS*, 463(1):102–119, November 2016.

[37] Siddharth Mishra-Sharma, Ken Van Tilburg, and Neal Weiner. Power of halometry. *Phys. Rev. D*, 102(2):023026, 2020.

[38] Ken Van Tilburg, Anna-Maria Taki, and Neal Weiner. Halometry from Astrometry. *JCAP*, 07:041, 2018.

[39] Robert Feldmann and Douglas Spolyar. Detecting Dark Matter Substructures around the Milky Way with Gaia. *MNRAS*, 446:1000–1012, 2015.

[40] Robyn E. Sanderson, Carlos Vera-Ciro, Amina Helmi, and Joren Heit. Stream-subhalo interactions in the Aquarius simulations. *arXiv e-prints*, page arXiv:1608.05624, August 2016.

[41] S. Mao and P. Schneider. Evidence for Substructure in lens galaxies? *MNRAS*, 295:587–594, 1998. arXiv.

[42] J.W. Hsueh et al. SHARP - IV. An apparent flux ratio anomaly resolved by the edge-on disc in B0712+472. *MNRAS*, 469(3):3713–3721, 2017. arXiv.

[43] N. Dalal and C.S. Kochanek. Direct Detection of CDM Substructure. *ApJ*, 572:25–33, 2002. arXiv.

[44] Y.D. Hezaveh et al. Detection of Lensing Substructure Using ALMA Observations of the Dusty Galaxy SDP.81. *ApJ*, 823(1):37–56, 2016. arXiv.

[45] S. Vegetti, L. V. E. Koopmans, A. Bolton, T. Treu, and R. Gavazzi. Detection of a dark substructure through gravitational imaging. *MNRAS*, 408(4):1969–1981, November 2010.

[46] S. Vegetti, D. J. Lagattuta, J. P. McKean, M. W. Auger, C. D. Fassnacht, and L. V. E. Koopmans. Gravitational detection of a low-mass dark satellite galaxy at cosmological distance. *Nature*, 481(7381):341–343, January 2012.

[47] S. Vegetti, L. V. E. Koopmans, M. W. Auger, T. Treu, and A. S. Bolton. Inference of the cold dark matter substructure mass function at z = 0.2 using strong gravitational lenses. *MNRAS*, 442(3):2017–2035, August 2014.

[48] E. Ritondale, S. Vegetti, G. Despali, M.W. Auger, L.V.E. Koopmans, and J.P. McKean. Low-mass halo perturbations in strong gravitational lenses at redshift z $\sim$ 0.5 are consistent with CDM. *MNRAS*, 485(2):2179–2193, 2019.

[49] S. Vegetti and L.V.E. Koopmans. Bayesian strong gravitational-lens modelling on adaptive grids: objective detection of mass substructure in Galaxies. *MNRAS*, 392(3):945–963, 2009. arXiv.

[50] L.V.E. Koopmans. Gravitational imaging of cold dark matter substructures. *MNRAS*, 363(4):1136–1144, 2005. Oxford Journals.

[51] S. Vegetti and L.V.E. Koopmans. Statistics of mass substructure from strong gravitational lensing: quantifying the mass fraction and mass function. *MNRAS*, 400:1583–1592, 2009. arXiv.

[52] Johann Brehmer, Siddharth Mishra-Sharma, Joeri Hermans, Gilles Louppe, and Kyle Cranmer. Mining for Dark Matter Substructure: Inferring subhalo population properties from strong lenses with machine learning. *Astrophys. J.*, 886(1):49, Nov 2019.

[53] Stephon Alexander, Sergei Gleyzer, Evan McDonough, Michael W. Toomey, and Emanuele Usai. Deep Learning the Morphology of Dark Matter Substructure. *Astrophys. J.*, 893:15, 2020.

[54] 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.

[55] Sreedevi Varma, Malcolm Fairbairn, and Julio Figueroa. Dark Matter Subhalos, Strong Lensing and Machine Learning. *arXiv e-prints*, page arXiv:2005.05353, May 2020.

[56] Kyriakos Vattis, Michael W. Toomey, and Savvas M. Koushiappas. Deep learning the astrometric signature of dark matter substructure. *arXiv e-prints*, page arXiv:2008.11577, August 2020.

[57] Alireza Makhzani, Jonathon Shlens, Navdeep Jaitly, Ian Goodfellow, and Brendan Frey. Adversarial Autoencoders. *arXiv e-prints*, page arXiv:1511.05644, November 2015.

[58] Diederik P Kingma and Max Welling. Auto-Encoding Variational Bayes. *arXiv e-prints*, page arXiv:1312.6114, December 2013.

[59] Jonathan Masci, Ueli Meier, Dan Cireşan, and Jürgen Schmidhuber. Stacked convolutional auto-encoders for hierarchical feature extraction. In Timo Honkela, Włodzisław Duch, Mark Girolami, and Samuel Kaski, editors, *Artificial Neural Networks and Machine Learning – ICANN 2011*, pages 52–59, Berlin, Heidelberg, 2011. Springer Berlin Heidelberg.

[60] Aprajita Verma, Thomas Collett, Graham P. Smith, Strong Lensing Science Collaboration, and the DESC Strong Lensing Science Working Group. Strong Lensing considerations for the LSST observing strategy. *arXiv e-prints*, page arXiv:1902.05141, February 2019.

[61] Masamune Oguri and Philip J. Marshall. Gravitationally lensed quasars and supernovae in future wide-field optical imaging surveys. *MNRAS*, 405(4):2579–2593, July 2010.

[62] J. W. Nightingale and S. Dye. Adaptive semi-linear inversion of strong gravitational lens imaging. *MNRAS*, 452(3):2940–2959, September 2015.

[63] J. W. Nightingale, S. Dye, and Richard J. Massey. AutoLens: automated modeling of a strong lens's light, mass, and source. *MNRAS*, 478(4):4738–4784, August 2018.

[64] Adam S. Bolton, Scott Burles, Léon V. E. Koopmans, Tommaso Treu, Raphaël Gavazzi, Leonidas A. Moustakas, Randall Wayth, and David J. Schlegel. The Sloan Lens ACS Survey. V. The Full ACS Strong-Lens Sample. *ApJ*, 682(2):964–984, August 2008.

[65] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Kopf, Edward Yang, Zachary DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. Pytorch: An imperative style, high-performance deep learning library. In H. Wallach, H. Larochelle, A. Beygelzimer, F. dAlché-Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems 32*, pages 8024–8035. Curran Associates, Inc., 2019.

[66] Geoffrey E. Hinton, Simon Osindero, and Yee-Whye Teh. A fast learning algorithm for deep belief nets. *Neural Computation*, 18(7):1527–1554, 2006. PMID: 16764513.