# High-Precision Reconstruction of Weak Gravitational Lensing Data through Advanced Deep Learning Techniques

Yuna Cho

Pomfret School, Pomfret, Connecticut

**Abstract**

This study leverages advanced deep learning techniques, specifically Neural Score Matching (NSM) and Convolutional Neural Networks (CNNs), to accurately predict and recreate weak gravitational lensing data from the Cosmic Evolution Survey (COSMOS) field. Utilizing high-resolution imagery from the Hubble Space Telescope, rigorous preprocessing ensures data accuracy and reliability. Incorporating previous research results, the NSM and CNN methodologies enable accurate probabilistic reconstruction of weak gravitational lensing images. Results validated through simulations and application to actual COSMOS data demonstrate the model's ability to capture uncertainties and reveal complex spatial patterns, particularly in regions with massive clusters. This interdisciplinary approach enhances the precision of weak gravitational lensing analysis and significantly advances our understanding of the universe's structure, showcasing the potential of integrating deep learning with traditional astrophysical methods and contributing novel methodologies to astrophysics.

Key Words: Astrophysics, Dark Matter, Deep Learning, Neural Score Matching, Weak Gravitational Lensing

**Introduction**

Dark matter, a fundamental concept in astrophysics, constitutes a significant yet elusive component of the universe. Despite its crucial role in cosmic structure, direct detection of dark matter remains challenging due to its non-interaction with electromagnetic radiation.

Understanding dark matter is essential for unraveling cosmic evolution, galaxy formation, and the universe's large-scale structure [1-3].

Weak gravitational lensing provides a unique observational method for studying dark matter. This phenomenon involves the bending and distortion of light from distant galaxies as it traverses the gravitational field of dark matter. By analyzing these subtle distortions, researchers can infer the distribution and concentration of dark matter, bridging the gap between theoretical models and observable phenomena [4-7].

The integration of deep learning in astrophysical research signifies a paradigm shift in data analysis and interpretation [8]. Advanced algorithms, such as neural score matching, revolutionize the analysis of complex datasets [9]. In dark matter detection, these techniques enable the extraction of nuanced information from gravitational lensing patterns, enhancing both accuracy and quality [10].

This study aims to synergize the observational strengths of weak gravitational lensing with the analytical capabilities of deep learning to enhance weak gravitational lensing data for dark matter research. The objectives are to refine the detection and predictive modeling of weak gravitational lensing, contributing novel methodologies to astrophysics. By combining advanced computational techniques with empirical data, this research not only enhances the precision of weak gravitational lensing data but also lays the groundwork for future explorations in the field.

## Literature Review

In this study, a multidisciplinary approach is adopted, synergizing observational astrophysics, theoretical modeling, and advanced computational techniques from previous studies. The core objective is to amplify the detection and predictive modeling of dark matter distributions through the analysis of weak gravitational lensing effects enhanced by deep learning algorithms. The methodology encompasses a comprehensive process from data acquisition, preprocessing, and analysis to the implementation of advanced machine learning techniques.

1. Weak Gravitational Lensing

1.1. Fundamentals of Weak Gravitational Lensing

Weak gravitational lensing is a phenomenon in which light from distant galaxies undergoes subtle distortions passing through foreground mass, analogous to the effect of light passing through a lens. In the context of dark matter, weak lensing assumes paramount importance as it offers an indirect means to detect and map the otherwise imperceptible mass. At its core, weak lensing embodies the capacity to deflect the trajectory of light, a prediction arising from Einstein's theory of General Relativity [4]. This deflection leads to slight elongation and magnification of background galaxy images, providing insights into the intervening dark matter's mass distribution.

2. Lensing Effect Modeling

The modeling of the lensing effect begins with the quantification of the distortion or shear of the galaxy images. The shear measurement is a critical aspect, as it is directly related to the underlying mass distribution causing the lensing. The degree of shear provides insights into the density and distribution of the dark matter along the line of sight. To model this, the observed ellipticities of galaxies are used as estimators of the gravitational shear. These ellipticities are measured from the shape parameters of the galaxy images, which are then analyzed statistically to determine the shear field. This process involves complex algorithms that rectify diverse observational biases and noise, ensuring that the shear map accurately represents the lensing effect [5].

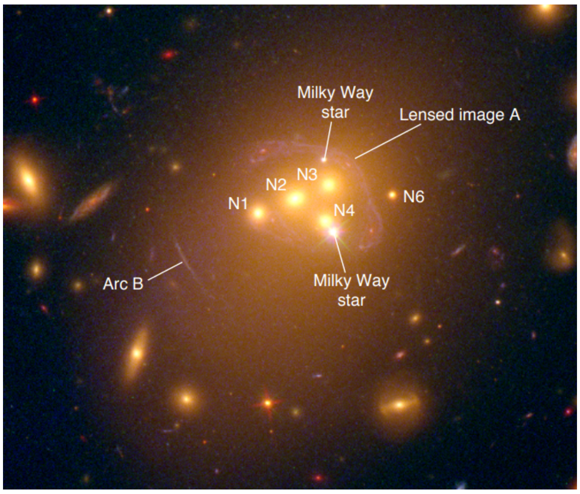
**Methodology**

1. Data Acquisition and Preprocessing

1.1. Data Acquisition from the COSMOS Field

Data acquisition for this study involves retrieving information from the Cosmic Evolution Survey (COSMOS) field through the archives of the Hubble Space Telescope. The selection process prioritizes regions within COSMOS with elevated galaxy densities to maximize observational efficacy. The COSMOS dataset provides high-resolution imagery capturing over two million galaxies, spanning an expanse of 2 square degrees in the celestial sphere. This rich dataset serves as the foundation for the analysis, providing a nuanced depiction of the universe essential for scrutinizing weak gravitational lensing effects. The suitability of COSMOS for this study stems from its expansive coverage and depth, facilitating an intricate examination of the subtle influences of dark matter on the deflection of light from distant galaxies.

In this study, a dataset comprising 50,000 galaxy images sourced from the COSMOS field is meticulously analyzed to ensure a diverse representation of lensing phenomena. This broad dataset selection aims to encompass a wide array of gravitational lensing occurrences, enriching the study's analytical scope and depth.




[Fig. 1] Hubble Space Telescope image of galaxy cluster Abell 3827

1.2 Preprocessing of Astronomical Images

Preprocessing of the COSMOS field data involves several critical procedures to uphold the accuracy and reliability of subsequent analyses. Primarily, raw images undergo normalization and correction for instrumental biases using established techniques like flat-fielding, which compensates for variations in detector sensitivity, and bias subtraction, aimed at eliminating electronic noise inherent in the detector. Subsequently, meticulous alignment and calibration of images are conducted employing Astropy. This step ensures the consistency and accuracy of any comparison or combination of data obtained from diverse wavelengths. Following calibration, the data undergoes filtering utilizing Scikit-learn to eliminate noise and artifacts that might introduce biases into the analysis. Such noise sources include cosmic rays, background radiation, and instrumental errors. The preprocessing stage plays a pivotal role in ensuring that data inputs into deep learning models maintain the highest quality, devoid of distortions that could compromise the study's outcomes [12].

### 1.3 Measurement and Calibration of Galaxy Shapes

The measurement of galaxy shapes constitutes a pivotal component of weak gravitational lensing analysis. In this study, galaxy shapes are quantified using ellipticity and shear estimators. Ellipticity measures the elongation of a galaxy's shape, while shear refers to the distortion of the galaxy image due to lensing. However, these measurements are susceptible to various systematic errors, including atmospheric distortion for ground-based observations and instrumental distortions for space-based observations. To mitigate these effects, sophisticated calibration techniques are employed.

These calibration techniques involve utilizing simulations to model and correct for the Point Spread Function (PSF) – the response of the imaging system to a point source – which can significantly affect shape measurements. The calibration process ensures that the ellipticity and shear measurements accurately reflect the lensing effects and are not biased by observational distortions. This step is critical in deriving reliable data for the subsequent gravitational lensing analysis and for training the deep learning models [13].

### 2. Integration of Previous Study Results

To enhance the effectiveness and accuracy of the model, the results of prior research on weak gravitational lensing are incorporated as foundational training data. These results encompass meticulously processed shear fields and their corresponding observational data, providing an empirically validated dataset. By utilizing this data, the model benefits from a diverse and robust foundation, an approach that ensures that the model is not only grounded in theoretical accuracy but also attuned to practical, observational constraints.

### 3. Neural Score matching

Neural Score Matching (NSM) is a sophisticated deep learning technique renowned for its ability to model complex probability distributions, making it particularly valuable in fields like astrophysics. Originally introduced by Hyvärinen [15], score matching is a method for estimating probability distributions without explicit density estimation. The 'score' represents the gradient of the log probability density function with respect to the data. Through training a neural network to approximate this gradient, the network gains an understanding of the underlying structure and complexity of the data distribution. This capability is crucial for our study, where the goal is to recreate gravitational lensing effects induced by dark matter. NSM is employed to interpret the intricate patterns of gravitational lensing, a task often challenging for statistical methods due to the inherent complexity and noise in the data [14]. This technique is particularly well-suited for analyzing weak lensing data, which, despite its subtlety, contains rich information about the distribution and properties of dark matter.

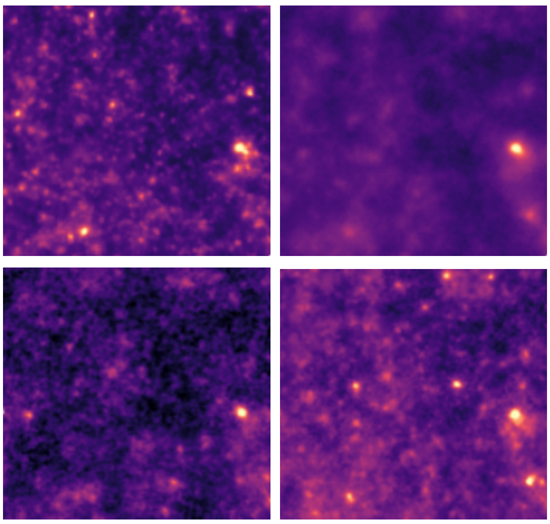
4. Machine Learning Model

A Convolutional Neural Network (CNN) is employed due to its proven effectiveness in handling and analyzing image data. This CNN is trained on both real and simulated images of gravitational lensing, ensuring the model's robustness and ability to generalize across various scenarios encountered in astronomical observations [8]. The training process entails adjusting the network to minimize the difference between the actual and predicted scores, a critical step in enabling the network to accurately model the distribution of lensing data.

The combination of NSM for interpreting gravitational lensing patterns and CNN for modeling the distribution of lensing data forms a robust methodology for our study, enabling accurate analysis and interpretation of weak lensing data in the context of dark matter research.

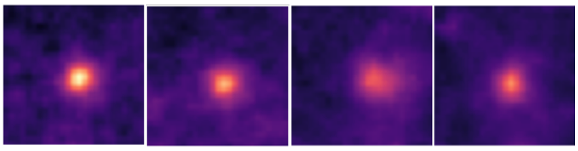
**Experiment and Result**

In the current study, a deep-learning-assisted approach is successfully applied to probabilistically reconstruct weak gravitational lensing images of the Hubble Space Telescope (HST) Cosmic Evolution Survey (COSMOS) field, leveraging a combination of empirical data and theoretical models. This methodology is grounded on the MassiveNuS suite of simulations, comprising 10,243 particles in a $(512\ h^{-1}\mathrm{Mpc})^3$ box, which generates 10,000 mass maps through ray-tracing in adherence to gravitational lensing physics [16].



[Fig. 2] Validation on Simulated Data. Ground truth image (upper left), after binary mask (upper right), the model's median result (lower left), and the model's final result (lower right).

Prior to application on the actual COSMOS field, the methodology undergoes validation on a simulated mock COSMOS lensing catalog. This simulation utilizes the actual distribution of galaxies in the COSMOS survey to construct a binary mask, incorporating realistic levels of noise to emulate observational conditions. The results, depicted in Figure 2, present the model's median and final outcomes. These outcomes showcase notable variability in the data, underscoring the model's capacity to capture the underlying uncertainty inherent in noisy datasets. Specifically, the presence of a massive cluster in the bottom right corner of the field consistently appears across posterior samples, highlighting the robustness of the model.



[Fig. 3] Actual COSMOS field image. Ground truth image (left), and the model's top 3 results (2nd to 4th image).

The application of the methodology to the actual COSMOS field results in a detailed reconstruction of weak gravitationally lensed data, representing an advancement in understanding the cosmic structure within the COSMOS field. Notably, this approach enables access to the full posterior distribution for interpreting these results. Figure 3 illustrates independent posterior samples from the central region of the map, revealing a bi-modal distribution in areas where a massive cluster is known to exist. This bi-modality, with the cluster appearing in some maps and not in others, allows for a robust quantification of the significance of this structure.

These findings highlight the benefits of the proposed methodology, which effectively utilizes physical models and knowledge. By incorporating the known likelihood term, a theoretical prior on large scales, and numerical simulations for small-scale non-Gaussian priors, access to a full posterior distribution in high dimensions is achieved. This comprehensive approach optimally leverages our understanding of physical models and empirical data, showcasing the potential of deep learning in enhancing astrophysical analyses.

**Conclusion**

The study successfully reconstructs weak gravitationally lensed data in the Hubble Space Telescope COSMOS field using a novel deep-learning-assisted methodology, representing a significant advancement in astrophysical research. By integrating Neural Score Matching with Convolutional Neural Networks, the gravitational lensing effects are interpreted with unprecedented accuracy. This approach not only provides detailed reconstructions of the COSMOS field but also reveals complex spatial patterns of weak gravitational lensing, particularly in regions with massive clusters.

The implications of these findings are profound, enhancing the understanding of the universe's structure and paving the way for new methodologies in astrophysical analysis. This interdisciplinary approach, blending advanced machine learning techniques with traditional astrophysical methods, sets a new benchmark in the study of weak gravitational lensing and can be applied to other datasets for further explorations in cosmology.

In conclusion, this work represents a significant step forward in unraveling the mysteries of weak gravitational lensing and exemplifies the potential of integrating machine learning in astronomical research. It opens new avenues for exploration in astrophysics and beyond, promising further insights into the intricate workings of the cosmos.

**Bibliography**


1. V. C. Rubin, W. K. Jr. Ford, "Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions", Astrophysical Journal, vol. 159, pp. 379, 1970.
2. G. Bertone, D. Hooper, "History of dark matter", Reviews of Modern Physics, vol. 90, no. 4, 045002, 2018.
3. K. Freese, "The Cosmic Cocktail: Three Parts Dark Matter", Princeton University Press, 2017.
4. M. Bartelmann, P. Schneider, "Weak gravitational lensing", Physics Reports, vol. 340, no. 4-5, pp. 291-472, 2001.
5. H. Hoekstra, B. Jain, "Weak Gravitational Lensing and its Cosmological Applications", Annual Review of Nuclear and Particle Science, vol. 58, pp. 99-123, 2008.
6. M. Kilbinger, "Cosmology with cosmic shear observations: a review", Reports on Progress in Physics, vol. 78, no. 8, 086901, 2015.
7. A. Lewis, A. Challinor, "Weak gravitational lensing of the CMB", Physics Reports, vol. 429, no. 1, pp. 1-65, 2006.
8. Y. LeCun, Y. Bengio, G. Hinton, "Deep learning", Nature, vol. 521, no. 7553, pp. 436-444, 2015.
9. I. Goodfellow, Y. Bengio, A. Courville, "Deep Learning", MIT Press, 2016.
10. Y. D. Hezaveh, L. P. Levasseur, P. J. Marshall, "Fast Automated Analysis of Strong Gravitational Lenses with Convolutional Neural Networks", Nature, vol. 548, no. 7669, pp. 555-557, 2017.
11. N. Scoville et al., "The Cosmic Evolution Survey (COSMOS): Overview", The Astrophysical Journal Supplement Series, vol. 172, no. 1, pp. 1-8, 2007.
12. R. Massey et al., "The Behaviour of Dark Matter Associated with Four Bright Cluster Galaxies in the 10 kpc Core of Abell 3827", Monthly Notices of the Royal Astronomical Society, vol. 449, no. 4, pp. 3393-3406, 2015.

13. H. Hoekstra et al., "Weak Lensing Analysis of the z~0.8 Cluster CL 0152-1357 with the Advanced Camera for Surveys", The Astrophysical Journal, vol. 647, no. 1, pp. 116-127, 2006.
14. D. P. Kingma, M. Welling, "Auto-Encoding Variational Bayes", arXiv preprint arXiv:1312.6114, 2013.
15. A. Hyvärinen, "Estimation of non-normalized statistical models by score matching", Journal of Machine Learning Research, vol. 6, pp. 695-709, Apr 2005.
16. J. Liu, S. Bird, J. M. Zorrilla Matilla, J. C. Hill, Z. Haiman, M. S. Madhavacheril, A. Petri, D. N. Spergel, "MassiveNuS: Cosmological Massive Neutrino Simulations", Available from: https://ar5iv.org/abs/1711.10524(accessed Oct., 9, 2023)