Iranian Journal of Astronomy and Astrophysics
Vol. 12, No. 4, Autumn 2025, 347–357
©Available online at http://ijaa.du.ac.ir
DOI: 10.22128/ijaa.2026.3309.1246
Online ISSN: 2383–403X

Iranian Journal of Astronomy and Astrophysics

Research Paper

Gravitational Lens Detection in Simulated LSST Data: A Comprehensive Feature Extraction Approach

Seyedeh Nahid Hasani[1] · Nasibe Alipour[2] · Reza Saffari*[3]
[1] Department of Physics, University of Guilan, 41335-1914, Rasht, Iran;
E-mail: nahid_hasani@webmail.guilan.ac.ir
[2] Department of Physics, University of Guilan, 41335-1914, Rasht, Iran;
E-mail: alipourrad@guilan.ac.ir
[3] Department of Physics, University of Guilan, 41335-1914, Rasht, Iran;
*E-mail: rsk@guilan.ac.ir

Received: 30 April 2026; **Accepted:** 18 July 2026; **Published:** 28 July 2026

Abstract. Identifying strong gravitational lensing systems is a major challenge in modern astronomical surveys. This research presents a machine learning-based method to distinguish lensed objects from non-lensed objects using multiband data. To this end, a comprehensive set of physically interpretable features is extracted from simulated data across the g, r, i, and z bands. These features capture statistical, spatial, amplitude, and frequency information. Then, using the XGBoost algorithm, we identify the features that have the greatest impact on the classification process. The model used in this study primarily relies on indicators of structural complexity (such as spread, entropy, and kurtosis) and asymmetry (such as skewness), especially in the g band. Our analysis shows that the proposed model achieves strong performance, with an accuracy of 0.88 and a TSS score of 0.75. Additionally, the Area Under the Curve (AUC) of 0.94 demonstrates the model's high ability to distinguish between gravitational lenses and non-lensed systems. Overall, the proposed framework can provide a reliable approach for identifying gravitational lens candidates.

Keywords: Gravitational Lens Detection, Machine Learning, Feature Engineering.

1 Introduction

According to the theory of general relativity, the gravity caused by mass causes light rays to deviate from their straight path. Therefore, if a mass or distribution of matter (such as a black hole, etc.) is placed between the observer and the light source, according to the relations of general relativity, this mass acts as an optical converging lens and causes the observer to receive more light rays than if there were no mass distribution. For this reason, the mass distribution that causes light to deviate is called a "gravitational lens"[1]. The first example of a true gravitational lens was a double-image quasar named (Q0975+561A; B), which was observed by Walsh and his colleagues in 1979 [2].

The study of gravitational lenses provides unique information about the distribution of matter that bends light, information that is independent of the light emitted by that matter.

Corresponding author

This feature enables us to understand the mass model of galaxies [3–5] and to investigate the structure of dark matter[6–8]. In addition, monitoring time-varying gravitational lenses is a powerful tool for estimating the Hubble constant [9–13] and investigating models of the expansion of the universe and dark energy [14,15]. On the other hand, the phenomenon of lensing magnification allows us to use powerful lenses as cosmic telescopes and observe objects that are not normally visible, such as quasi-stellar host galaxies and silent radio quasars [16–18].

Despite their scientific importance, gravitational lenses have posed a challenge in astrophysics, and many attempts have been made to identify them. Initially, traditional detection methods [19] focused on observing arcs in image data [20] or spectroscopic information [21], a process that involved meticulous manual inspection of each image, requiring large teams of people and a long time. With the generation of astronomical data in recent decades through large-scale sky surveys such as the Sloan Digital Sky Survey (SDSS) [22], the Hubble Space Telescope (HST) [23–25], the Large Synoptic Survey Telescope (LSST) [26], and others, traditional methods of detecting gravitational lenses have become less efficient. The expansion of computer science has introduced new methods, such as machine learning, that play an important role in advancing physics, especially in the identification of gravitational lenses. This technology encompasses a wide range of applications, including face and object recognition, speech recognition, search engines, and medical diagnosis [27].

Gravitational lens detection methods using artificial intelligence can be divided into two general categories. (a) Hand-crafted Features-based Methods: In this method, a set of statistical, morphological, and color features is first extracted from images, and these features are given as input to classical classifiers such as Support Vector Machine (SVM), Random Forest, Extreme Gradient Boosting (XGBoost), and Artificial Neural Networks (ANN). Research using tools such as Arcfinder [28] and Ringfinder [29] for multiband data to identify lensing systems, as well as other programs designed to detect arc-like geometric patterns [30,31], can generally be considered as part of hand-crafted features-based methods. (b) Deep Learning-based Methods: In this method, features are automatically extracted from raw images in Convolutional Neural Networks (CNN). There is significant research on the detection and classification of gravitational lenses that have been able to detect gravitational lenses using CNN [32–37] on image data. Since some supernovae may also have lenses [38–41], several studies have followed a multifaceted approach [38,39,41–43] to consider other objects that may also contain lenses. For this purpose, in addition to the image, they also examine its time series [38,39,41], and by examining the time series, they have been able to identify a set of gravitational lenses related to supernovae. Among machine learning algorithms, XGBoost [44] is one of the most popular and widely used algorithms, which has attracted significant attention due to its high accuracy and fast speed. This algorithm builds multiple decision trees in sequence, where each new tree attempts to correct the errors of the previous trees. Due to preventing overfitting and the ability to find complex nonlinear relationships between features, XGBoost has shown good performance in astrophysical classification tasks such as gravitational lens detection.

Statistical and morphological features are excellent tools that help us describe the shape and distribution of light in astronomical images. Parameters such as skewness, elongation, signal energy, and frequency features can detect subtle changes in light caused by gravitational bending. In this paper, we present a method that automatically detects gravitational lenses by extracting a comprehensive set of statistical, temporal, and frequency features. We extract a set of key features from images and show that these features can be used to identify gravitational lenses. Our proposed method, focusing on these features, is a fast and efficient solution for detecting gravitational lenses.

This study is organized as follows: Section 2 provides an overview of the simulated data. Section 3 details the feature extraction methodology and the machine learning-based classification approach. Section 4 evaluates the performance of the classification model on the simulation data and presents the resulting metrics and findings. Finally, Section 5 summarizes the concluding remarks of this article.

2 Data

The Vera C. Rubin Observatory is an advanced astronomical observatory whose primary mission is to conduct a comprehensive and systematic survey of the southern sky. One of the goals of this science project, now known as the Legacy Survey of Space and Time (LSST), is to study dark matter and dark energy through gravitational lensing [26]. In this study, we used the LSST-wide dataset, a simulated lensed and non-lensed gravity dataset [41]. This dataset was generated using the open-source deeplenstronomy package [45] based on the lenstronomy framework [46]. The LSST dataset consists of simulated reconstructions of LSST survey images generated using the LSSTCam instrument model [47]. The simulated data consists of 4992 lensed and 4968 non-lensed images, each dataset containing four images in the (griz) bands: g-band (green), r-band (red), i-band (near infrared), and z-band (infrared). The size of all images in these four datasets is 45 $\times$ 45 pixels (Table 1).

Table 1: A summary of the simulated dataset used in this study.

Dataset	**Camera**	**Description**	**Lenses**	**Non-lenses**
LSST-wide	LSSTCam	4-band 45 $\times$ 45 px	4,992	4,968

3 Method

We develop a machine learning model for gravitational lens detection and use the XGBoost (Extreme Gradient Boosting) algorithm [44], which operates based on decision trees and gradient boosting. We use simulated lens and non-lens data extracted from the LSST dataset reported in [41], and Figure 1 shows an example of these data. In this study, we apply a feature-engineering approach to extract physical features from the four photometric bands g, r, i, and z. For each band, we derive 20 features that fall into four categories: statistical features describing pixel-intensity distribution (mean, variance, skewness, and kurtosis); frequency-domain features modeling the complexity of spatial structures (frequency centroid and entropy); time/space-domain features such as energy and zero-crossing rate; and amplitude-related features. Table 2 provides a summary of these features. In total, we extract 80 features from the four bands. To prepare the data, we standardize all features, and the resulting standardized feature vectors are then fed into the model.

All 80 extracted features are then fed into the XGBoost classifier, which analyzes them and learns how each one contributes to distinguishing lens images from non-lens images. During training, the algorithm evaluates and ranks the features by their importance, and we find that 20 of them have the strongest influence and play the primary role in the final prediction. We divide the dataset into two parts: 80% is used for training the model, and 20% is used for testing. The model then performs a binary classification to decide whether each input image is a gravitational lens or not.

To rigorously evaluate the performance of the proposed classifier, we employ a comprehensive set of metrics derived from the confusion matrix, following standard definitions in the

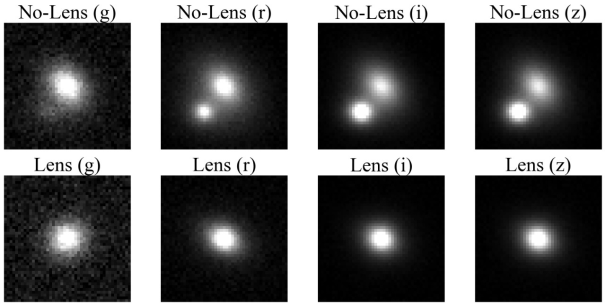

Figure 1: Display of gravitational lensed and no-lensed images in four bands (g,r,i,z). Top row is no-lensed images and bottom row is gravitational lensed images.

Table 2: Summary of the features extracted from images of gravitational lenses and non-lenses.

Type	**Description**
Statistical features	Mean, Standard Deviation, Variance RMS, Skewness, Kurtosis Maximum, Minimum, Median
Frequency features	Spectral Centroid, Spectral Spread, Spectral Energy Peak Frequency, Spectral Entropy
Time/Space domain features	Zero Crossing Rate, Signal Energy Peak-to-Peak Amplitude
Amplitude domain features	Maximum Amplitude Mean Amplitude Standard Deviation of Amplitude

literature [48]. Specifically, we compute Accuracy, Precision (Precision_+ and Precision_-), Recall (Recall_+ and Recall_-), and the F1-score (F_1^+ and F_1^-) for both the lens (positive) and non-lens (negative) classes. These metrics provide a detailed insight into the model's ability to correctly identify true positives while minimizing false alarms. Furthermore, to assess the overall separability of the model independent of class prevalence, we utilize the True Skill Statistic (TSS) [49–53]. This metric is particularly useful in imbalanced classification problems, as it simultaneously accounts for the true positive rate and the false positive rate, thereby providing a more reliable indication of the classifier's intrinsic skill than accuracy alone. Additionally, we evaluate the model's discriminative power across all classification thresholds using the Receiver Operating Characteristic (ROC) curve and its corresponding Area Under the Curve (AUC). Unlike metrics derived from the confusion matrix, which are computed at a fixed decision threshold, the ROC curve illustrates the trade-off between the true positive rate and the false positive rate over the entire range of possible thresholds. The AUC metric summarizes the model's overall ability to distinguish between positive and negative classes, providing a robust and threshold-independent measure of the classifier's discriminative capacity. Therefore, ROC/AUC complements the confusion-matrix-based metrics by assessing the ranking quality and separability of the model beyond a single fixed classification threshold.

4 Results

In this study, we use a binary classification approach to distinguish simulated gravitationally lensed images from non-lensed images. Lensing galaxies are typically old and have red spectral energy distributions that exhibit stronger emission in longer-wavelength photometric bands, specifically the near-infrared i-band and the infrared z-band. In contrast, background sources typically have blue spectral energy distributions. Consequently, the relative intensity and visibility of these objects differ across the g, r, i, and z bands. To preserve these wavelength-dependent variations, we perform feature extraction independently for each band.

In total, we extract 20 physical features for each of the four photometric bands (g, r, i, z), resulting in 80 features extracted from the photometric bands that serve as inputs to the model. Using the Gain criterion, the model assigns an importance score to each feature and identifies the 20 features that contribute the most to distinguishing gravitational lenses from non-lenses. Features with higher scores demonstrate stronger discriminative power.

Figure 2 illustrates the relative importance of each photometric band in the proposed model. As shown, the g band exhibits the highest cumulative importance score, followed by the r, z, and i bands. This distribution indicates that features extracted from shorter wavelengths (specifically the g band) contain more discriminative information for identifying gravitational lenses. This finding aligns with the physical interpretation that lensing arcs and distorted structures are more prominently visible in shorter wavelength observations, where the contrast between the lensing galaxy and the background source is enhanced. The contribution of longer-wavelength bands (r, i, z) provides complementary information that helps refine the classification, particularly in distinguishing morphological complexities that may not be apparent in a single band.

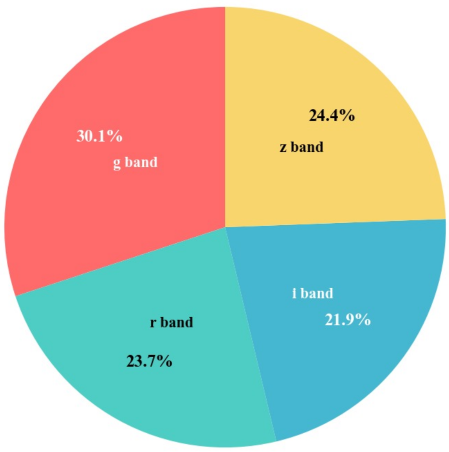

Figure 2: Relative importance of bands in the proposed model.

Table 3 lists these top-ranked features, which contribute more effectively to the model's ability to distinguish gravitational lenses compared with the remaining features. As shown, for classifying simulated gravitational-lens images, the model relies primarily on indicators of structural complexity (such as spread, entropy, and kurtosis) and asymmetry (such as skewness). The most influential feature is the spread in the g band, with an importance value of 0.041. This indicates that gravitational-lensing systems at shorter wavelengths exhibit a broader pattern of frequency dispersion. Since these wavelengths are characteristic of younger stellar populations, shorter-wavelength bands are more likely to host lensing systems. The high importance values of kurtosis (0.038) and entropy (0.030) in the same band further confirm that these systems show intensity distributions with shorter tails and greater disorder, driven by their arc-like structures.

In addition, the presence of dominant features in the r, i, and z bands demonstrates that multi-wavelength information plays a complementary role in enhancing the model's discriminability. This effect is especially important for describing the optical contrast between the lensing galaxy and the background source. The model essentially operates by comparing the photometric bands and examining how the morphology of the system changes across wavelengths. For example, the z band shows a compact, bright central object at longer wavelengths, while the g band reveals more diffuse structures at shorter wavelengths. From this discrepancy, the model concludes that two distinct objects are present and a lensing event has likely occurred.

The high feature importance scores reported in Table 3 reveal that frequency-domain descriptors, such as 'Spread' and 'Entropy,' are particularly informative as they capture the light complexity and dispersion directly associated with gravitational arcs and elongated structures. Conversely, statistical indices like 'Kurtosis' and 'Skewness' distinguish the heterogeneous, distorted nature of gravitational lenses from simpler objects—such as stars or elliptical galaxies—by quantifying deviations from normality and the degree of asymmetry. This strategic combination of structural and statistical features enables the model to accurately identify complex spatial patterns resulting from light bending, thereby enhancing the classification accuracy and generalizability of the lens systems by leveraging complementary information across different bands.

We evaluate the model's performance using the confusion matrix and the receiver operating characteristic (ROC) curve. Figure 3 presents the confusion matrix for the simulated dataset. As illustrated, the classifier correctly identifies 899 lensed images as positive samples (TP) and 845 non-lensed images as negative samples (TN), while misclassifying 149 non-lensed images as lensed (FP) and 99 lensed images as non-lensed (FN). The model demonstrates strong overall performance, achieving an accuracy of 0.88. For the positive class, $\mathrm{Precision}_+ = 0.86$ and $\mathrm{Recall}_+ = 0.90$ indicate that the model successfully identifies most true lensing cases while maintaining a relatively low false-positive rate, resulting in an F_1^+-score of 0.88. The negative class also shows reliable performance, with $\mathrm{Precision}_- = 0.89$, $\mathrm{Recall}_- = 0.85$, and an F_1^--score of 0.87. Based on these results, we also compute the TSS index to assess the overall separability of the model. All of these metric values, together with their corresponding formulas, are summarized in Table 4. Furthermore, the discriminative capability of the classifier is illustrated by the ROC curve in Figure 4. The curve (solid orange line) achieves an AUC of 0.94, significantly outperforming the random classifier (dashed diagonal line) and demonstrating the model's strong ability to distinguish between lensed and non-lensed images.

Overall, this pattern indicates that the model identifies gravitational lenses mainly based on morphological irregularities caused by arcs and multi-peaked brightness structures. This means the model does not seek simple round stars, but rather areas with high entropy, high dispersion, and unusual kurtosis; signs that are all more prominent in the g band because lensing arcs are clearer in blue light. Therefore, it can be said that the model is successfully able to identify "structural complexity" and "asymmetry" in images and consider them as indicators of gravitational lensing.

Table 3: Top 20 important features for gravitational lens detection based on feature importance scores.

Rank	**Type**	**Feature**	**Importance**	**Band**
1	Frequency	Spread	0.0406	g
2	Statistical	Kurtosis	0.0379	g
3	Frequency	Entropy	0.0304	g
4	Amplitude	Maximum	0.0250	r
5	Statistical	Skewness	0.0233	g
6	Statistical	Kurtosis	0.0204	r
7	Statistical	Skewness	0.0181	i
8	Statistical	Maximum	0.0180	z
9	Frequency	Peak	0.0180	r
10	Statistical	Skewness	0.0177	r
11	Statistical	Skewness	0.0175	z
12	Statistical	Kurtosis	0.0173	i
13	Frequency	Peak	0.0166	g
14	Frequency	Entropy	0.0164	z
15	Frequency	Centroid	0.0155	z
16	Frequency	Peak	0.0152	z
17	Frequency	Entropy	0.0149	i
18	Time/Space	Energy	0.0148	i
19	Amplitude	STD	0.0139	z
20	Frequency	Energy	0.0135	z

5 Conclusion

In this study, we aim to propose a machine learning–based method for classifying gravitational lenses. We use the XGBoost algorithm and extract its input features from the images. These features include physically grounded and interpretable characteristics, such as statistical, spatial, amplitude-based, and frequency-domain properties. We apply the proposed method to simulated LSST data. A key advantage of this approach is its transparency in the decision-making process. Since the extracted features are based on physical properties, the model's decisions are highly interpretable. Our analyses show that "structural complexity" and "asymmetry," especially in shorter wavelength bands, play a critical role in lens detection. Furthermore, the results demonstrate that the method successfully classifies images with an accuracy of 0.88, a TSS value of 0.75, and an AUC of 0.94. These findings provide a clear and understandable classification mechanism and establish the method as a robust tool for astrophysical applications.

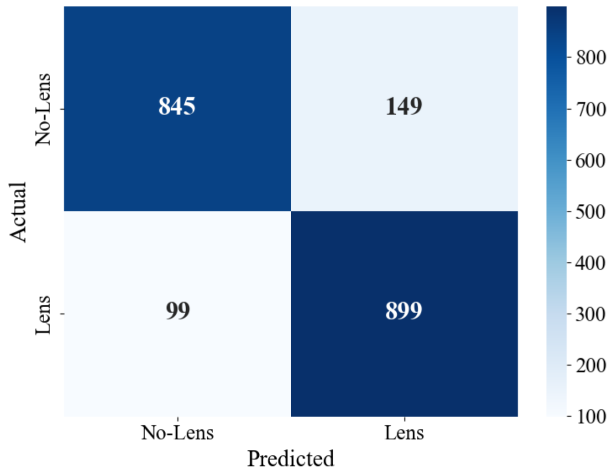

Figure 3: The confusion matrix of the presented model based on statistical features for the simulated dataset at the LSST level.

Table 4: Classification performance metrics and their mathematical definitions.

Metric	Definition	LSST-wide
Recall (positive and negative)	$\text{Recall}_+ = \frac{\text{TP}}{\text{TP}+\text{FN}}$	0.90
	$\text{Recall}_- = \frac{\text{TN}}{\text{TN}+\text{FP}}$	0.85
Precision (positive and negative)	$\text{Precision}_+ = \frac{\text{TP}}{\text{TP}+\text{FP}}$	0.86
	$\text{Precision}_- = \frac{\text{TN}}{\text{TN}+\text{FN}}$	0.89
F_1-score (positive and negative)	$\text{F}_1^+ = \frac{2\times\text{Precision}_+\times\text{Recall}_+}{\text{Precision}_+ + \text{Recall}_+}$	0.88
	$\text{F}_1^- = \frac{2\times\text{Precision}_-\times\text{Recall}_-}{\text{Precision}_- + \text{Recall}_-}$	0.87
Accuracy	$\frac{\text{TP}+\text{TN}}{\text{TP}+\text{TN}+\text{FP}+\text{FN}}$	0.88
True Skill Statistic (TSS)	$\text{TSS} = \frac{\text{TP}}{\text{TP}+\text{FN}} - \frac{\text{FP}}{\text{FP}+\text{TN}}$	0.75

Acknowledgment

We gratefully thank the anonymous referee for very helpful comments and suggestions to improve the manuscript.

Authors' Contributions

All authors have the same contribution.

Data Availability

The dataset used in this study is publicly available in the DeepGraviLens repository at https://github.com/nicolopinci/deepgravilens. The data are based on simulated datasets

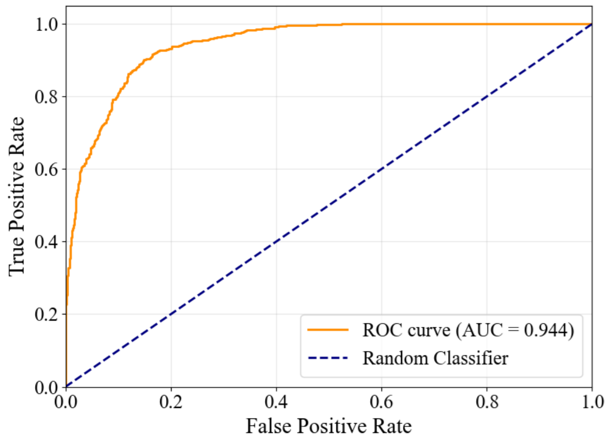

Figure 4: Receiver Operating Characteristic (ROC) curve of the classifier model for the simulated data, with an AUC of 0.94 indicating a high ability to discriminate between lens and non-lens images.

at the LSST level, generated using the tools and parameters described in the repository documentation.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Ethical Considerations

The authors have diligently addressed ethical concerns, such as informed consent, plagiarism, data fabrication, misconduct, falsification, double publication, redundancy, submission, and other related matters.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not for profit sectors.

References

[1] Meylan, G., & et al. 2006, Gravitational Lensing: Strong, Weak and Micro. astro-ph.

[2] Walsh, D., Carswell, R. F., & et al. 1979, Nature, 279, 381.

[3] Treu, T., Koopmans, L. V. E. 2002, ApJ, 575, 87.

[4] Treu, T., Koopmans, L. V. E. 2003, MNRAS, 343, L29.

[5] Treu, T., Koopmans, L. V. E. 2004, ApJ, 611, 739.

[6] Mao, S., Schneider, P. 1998, MNRAS, 295, 587.

[7] Metcalf, R. B., Madau, P. 2001, ApJ,563, 9.

[8] Vegetti, S., Lagattuta, D. J., & et al. 2012, Nature, 481, 341.

[9] Refsdal, S., 1964b, MNRAS, 128, 307.

[10] Vuissoz, C., & et al., 2007, A&A , 464, 845.

[11] Kochanek, C. S., Schechter, P. L., 2004, Measuring and Modeling the Universe. Cambridge Univ. Press, Cambridge, p. 117.

[12] Suyu, S. H. & et al., 2014, ApJ, 788, L35.

[13] Goobar, A. & et al., 2016, Science, 356, 291.

[14] Bonvin, V., & et al., 2017, MNRAS, 465, 4914.

[15] Wong, K. C., & et al., 2020, MNRAS, 498, 1420.

[16] Claeskens, J.F., Sluse, D., & et al., 2006, A&A, 451, 865.

[17] Jackson, N., 2011, ApJ, 739, L28.

[18] Jackson, N., & et al., 2015, MNRAS, 454, 287.

[19] Zwicky, F., 1937, Phys. Rev, 51, 679.

[20] Alard, C., 2006, arXiv e-prints, astro–ph/0606757.

[21] Faure, C., & et al., 2008, ApJS, 176, 19.

[22] James, E., Gunn., & et al., 2006, AJ, 131, 2332.

[23] Williams, R., E., & et al., 1996, AJ, 112, 1335.

[24] Williams, R., E., & et al., 2000, AJ, 120, 2735.

[25] Koekemoer, A., & et al., 2007, ApJS, 172, 196.

[26] Tyson, J., Anthony., 2002, arXiv, astro-ph/0302102.

[27] Mohammed, M., & et al.,2016, Machine learning: algorithms and applications, CRC Press.

[28] Seidel, G., Bartelmann, M., 2007, A&A, 472, 341.

[29] Gavazzi, R., Marshall, P. J., & et al., 2014, ApJ, 785, 144.

[30] Lenzen, F., Schindler, S., Scherzer, O., 2004, A&A, 416, 391.

[31] More, A., Cabanac, R., More, S., & et al., 2012, ApJ, 749, 38.

[32] Petrillo, C., Tortora, C., Chatterjee, S., & et al., 2017, MNRAS, 472, 1129.

[33] Pearson, J., Pennock, C., Robinson, T., 2018, Emerg Sci, 2:1.

[34] Petrillo, C., Tortora, C., Vernardos, G., & et al., 2019, MNRAS, 484, 3879.

[35] Petrillo, C., Tortora, C., Chatterjee, S., & et al., 2019, MNRAS, 482, 807.

[36] Can˜ameras, R., Schuldt, S., Suyu, S., & et al., 2020, A&A, 644, 163.

[37] Savary, E., Rojas, K., Maus, M., & et al., 2022, A&A, 666, 1.

[38] Morgan, R., Nord, B., Bechtol, K., & et al., 2022, ApJ, 927, 109.

[39] Morgan, R., Nord, B., Bechtol, K., & et al., 2022, arXiv:2204.05924.

[40] Kodi Ramanah, D., Arendse, N., Wojtak, R., 2022, MNRAS, 512, 5404.

[41] PinciroliVagoN, O., Fraternali, P., 2023, Neural Comput. Appl., 35, 19253.

[42] Marshall, P. J., Hogg, D. W., Moustakas, L. A., & et al., 2009, ApJ, 694, 924.

[43] Khramtsov, V., Sergeyev, A., Spiniello, C., & et al., 2019, A&A, 632, 56.

[44] Chen, T., He, T., & et al., 2015, R package version 0.4-2, 1, 1.

[45] Morgan, R., Nord, B., & et al., 2021, J. Open Source Softw., 6, 2854.

[46] Birrer, S., Amara, A., 2018, Phys. Dark Univ., 22, 189.

[47] Stalder, B., & et al., 2020, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 11447.

[48] Fawcett, T., 2006, P. R. L., 27, 861.

[49] Raboonik, A., & et al., 2017, ApJ., 834, 11.

[50] Alipour, N., & et al., 2019, ApJS., 243, 20.

[51] Ghader, H., & et al., 2025, ApJS., 277, 10.

[52] Hassani, Z., & et al., 2025, ApJS., 279, 27.

[53] Hassani, Z., & et al., 2026, The J. Supercomputing, 82, 326.