Proceedings of CONF-CIAP 2026 Symposium: International Conference on Atomic Magnetometer and Applications
DOI: 10.54254/2753-8818/2026.HZ29784

Exoplanet Detection Techniques: From Classical Methods to Modern Machine Learning Approaches

Xiaotong Wang

The High School Affiliated to Renmin University of China, Beijing, China
xulina313@126.com

Abstract. Since the discovery of 51 Pegasi b in 1995, exoplanet research has evolved from serendipitous radial-velocity detections to large-scale surveys employing transit photometry, microlensing, astrometry, and high-contrast direct imaging. Each technique probes different physical regimes, enabling the measurement of planetary masses, radii, orbital architectures, and atmospheric compositions. However, as instrument precision and data volume continue to grow, traditional detection algorithms struggle with noise, degeneracy, and the massive data throughput of modern facilities. Recent progress in machine learning,especially deep convolutional and generative models,has begun to transform this field, improving sensitivity and automation across all detection modalities. This review provides a chronological and conceptual overview of exoplanet detection methods, highlighting how data-driven frameworks are reshaping exoplanet discovery and characterization. The paper concludes with a discussion of key challenges, interpretability issues, and prospects for future space and ground-based missions.

Keywords: exoplanet detection methods, exoplanet overview, direct imaging, Machine learning, AI

1. Introduction

The search for planets beyond the Solar System stands among the most significant scientific endeavors of the past half-century. Exoplanets provide laboratories for testing theories of planetary formation, migration, and atmospheric physics under conditions far beyond those available in our own system. The first confirmed detection,51 Pegasi b, discovered via radial-velocity (RV) measurements by Mayor and Queloz [1],opened a new era in astrophysics. Within a decade, hundreds of giant planets had been detected, mostly hot Jupiters on short-period orbits. The launch of NASA's Kepler mission in 2009 extended this revolution, revealing that planets are ubiquitous and that small, terrestrial-sized worlds are common [2].

Each detection method,transit photometry, RV spectroscopy, microlensing, astrometry, and direct imaging,targets a distinct domain of planetary parameter space. Transit and RV techniques dominate in detecting close-in planets, while microlensing and astrometry probe distant or low-mass regimes. Direct imaging, in contrast, is uniquely capable of isolating a planet's light from its host star, permitting spectroscopic characterization of atmospheres.

DOI: 10.54254/2753-8818/2026.HZ29784

As the field matured, it faced two key challenges. First, increasing precision magnified the impact of stellar noise and instrumental systematics. Second, survey missions began producing terabytes of time-series and imaging data that exceeded the capacity of manual analysis. These pressures fostered the adoption of statistical inference, Bayesian modeling, and most recently machine learning (ML). ML algorithms,especially deep convolutional neural networks (CNNs) and generative adversarial networks (GANs),can learn complex nonlinear relationships between features and signals, enhancing both sensitivity and reliability.

This review traces the chronological evolution of exoplanet detection techniques, summarizing the principles of each method and the modern role of data-driven frameworks. The organization is as follows: Section 2 reviews indirect methods (transit, RV, microlensing, and astrometry); Section 3 examines high-contrast direct imaging; Section 4 describes the integration of ML across detection domains; Section 5 discusses current challenges; Section 6 outlines prospects for future missions; and Section 7 offers concluding remarks.

2. Indirect detection techniques

2.1. Transit photometry

The transit method identifies periodic dips in stellar brightness when a planet crosses the stellar disk. The fractional flux decrease

$$\Delta F/F \approx (R_p/R_\star)^2 \tag{1}$$

relates the planetary and stellar radii. From the light-curve shape one can infer the orbital period, inclination, and,combined with RV data,planetary density.

Early ground-based surveys such as HATNet and SuperWASP demonstrated feasibility, but space telescopes brought the precision required for Earth-size detections. Kepler observed over 150,000 stars with 30-ppm photometric precision, revealing thousands of confirmed planets [2]. Its successor TESS [3] continues all-sky coverage of bright nearby stars.

Systematic errors,instrumental drifts, pointing jitter, and stellar variability,necessitate complex detrending. Algorithms like the Trend Filtering Algorithm [4] and Gaussian-process regression model correlated noise, while Box-Least Squares searches identify periodic transits. However, distinguishing true transits from false positives (eclipsing binaries, pulsations) increasingly relies on ML classifiers. Google's AstroNet CNN [5] achieved over 95 % accuracy in classifying Kepler light curves. Similar models now vet TESS candidates, dramatically reducing manual effort.

The transit method's bias toward edge-on geometries means only a small fraction of planetary systems are detectable, but its combination with follow-up spectroscopy provides planetary radii, masses, and atmospheric spectra via transmission spectroscopy [6].

2.2. Radial-velocity spectroscopy

The RV technique measures periodic Doppler shifts in stellar absorption lines caused by an orbiting planet's gravitational influence. The RV semi-amplitude is

DOI: 10.54254/2753-8818/2026.HZ29784

$$K=\left(\frac{2\pi G}{P}\right)^{1/3}\frac{M_p \sin i}{(M_\star+M_p)^{2/3}}\frac{1}{\sqrt{1-e^2}}, \tag{2}$$

where P is orbital period and e eccentricity. For Sun-like stars, Earth induces only 9 cm s^{-1},below current instrumental limits,whereas Jupiter yields 12 m s^{-1}.

Spectrographs such as HARPS, HIRES, and ESPRESSO have pushed precision to 30 cm s^{-1} [7], allowing detection of Neptune-mass planets. The principal limitation arises from stellar activity,spots and convective motions that mimic RV shifts. ML tools have become essential for disentangling these signals. For example, Haywood et al. [8] applied Gaussian-process regression to model correlated stellar noise; Rajpaul et al. [9] developed simultaneous photometric-RV modeling; and Dumusque et al. [10] used neural networks to predict activity indicators from spectral features, improving planet yield.

RV surveys complement transit data by providing true masses and eccentricities, and their synergy enables density-based classification into rocky, gaseous, and water-rich worlds.

2.3. Gravitational microlensing

Microlensing exploits the magnification of a background source when a foreground lens star,and possibly its planet,passes near the line of sight [11]. The planetary perturbation lasts from hours to days, producing distinctive deviations in the light curve. The technique is most sensitive to planets at a few astronomical units and can detect Earth-mass bodies around distant or faint stars.

Surveys such as OGLE, MOA, and KMTNet monitor millions of stars toward the Galactic bulge, identifying several dozen planetary events [12]. Because events are singular and non-repeatable, rapid identification is critical. Automated ML pipelines now classify microlensing light curves in real time. Kim et al. [13] used convolutional and recurrent neural networks (RNNs) to flag planetary anomalies, reducing response time for follow-up telescopes. The upcoming Nancy Grace Roman Space Telescope will perform a dedicated microlensing survey expected to detect thousands of cold planets [14].

2.4. Astrometry

Astrometric detection measures a star's tiny positional wobble on the sky plane due to orbiting planets. The angular amplitude

$$\alpha=\frac{M_p}{M_\star}\frac{a_p}{d} \tag{3}$$

depends on the planetary mass M_p, semi-major axis a_p, and stellar distance d. For a Jupiter analogue at 10 pc, $\alpha \approx 0.5$mas. Space missions provide the required stability: Hipparcos pioneered the approach; ESA's Gaia now achieves micro-arcsecond precision [15]. Gaia DR4 is expected to yield several thousand new giant planets.

Analyzing astrometric data involves high-dimensional covariance modeling. Bayesian inference and ML regression (e.g., random forests for outlier detection) improve robustness against attitude-

DOI: 10.54254/2753-8818/2026.HZ29784

model errors and correlated noise. Combining astrometry with RVs will provide full three-dimensional orbits, breaking the sini degeneracy and refining planetary mass estimates.

3. Direct imaging

3.1. Principles

Direct imaging isolates a planet's photons from its host star's overwhelming brightness, a contrast typically between 10^5 and 10^9 in the near-infrared. Success demands both high angular resolution and precise starlight suppression. Adaptive optics (AO) correct atmospheric turbulence in real time, while coronagraphs block the central stellar core. Post-processing then removes residual speckles,quasi-static patterns caused by optical imperfections,that otherwise mask faint companions [16].

Direct imaging is uniquely suited to young, self-luminous giant planets at wide separations (≳10 AU). Landmark discoveries include HR 8799 bcde [17] and β Pictoris b [18]. Modern facilities,VLT/SPHERE [19], Gemini/GPI [20], Subaru/SCExAO, and JWST's NIRCam and MIRI coronagraphs,extend sensitivity toward cooler and lower-mass planets.

3.2. Differential-imaging algorithms

Because residual speckles dominate images, differential-imaging techniques exploit temporal, spectral, or spatial diversity:

• Angular Differential Imaging (ADI): the field rotates with parallactic angle while the optical system remains fixed [21]. Median subtraction or principal-component analysis (PCA-ADI) isolates non-rotating speckles.

• Spectral Differential Imaging (SDI): uses wavelength-dependent speckle scaling, distinguishing them from achromatic planetary signals.

• Reference Differential Imaging (RDI): subtracts a reference star's point-spread function (PSF).

• Locally Optimized Combination of Images (LOCI): constructs a linear combination of reference frames minimizing residuals in localized zones.

• ANDROMEDA and KLIP: maximum-likelihood and eigenimage decomposition approaches.

These methods dramatically enhance contrast but can self-subtract part of the planetary signal. Selecting the number of PCA modes or optimization zones remains an empirical trade-off between noise suppression and throughput.

3.3. High-contrast instrumentation

Coronagraph designs,Lyot, vortex, and phase-mask,achieve inner working angles down to 0.1 arcsec. Extreme AO systems employ deformable mirrors with thousands of actuators. Wavefront-sensing algorithms such as electric-field conjugation further reduce residuals. The MIRI coronagraphs aboard JWST extend imaging to mid-infrared wavelengths (10–15 μm), ideal for mature cold Jupiters like GJ 832 b [22]. Data volumes from such instruments require automated, statistically rigorous pipelines, motivating the transition to machine learning.

DOI: 10.54254/2753-8818/2026.HZ29784

4. Machine learning in exoplanet science

4.1. Motivation and overview

Machine learning has permeated nearly every stage of exoplanet detection and characterization. Unlike parametric models requiring explicit equations, ML algorithms infer mappings directly from data. Their ability to learn nonlinear boundaries makes them ideal for distinguishing subtle planetary signals from correlated noise. They also enable rapid real-time analysis critical for large surveys.

The first applications involved transit light-curve classification. Shallue and Vanderburg [5] trained a deep CNN on Kepler data to distinguish planets from eclipsing binaries with >98 % precision. Later, Ansdell et al. [23] applied transfer learning to TESS data, achieving similar results with smaller training sets. For RV data, Hara et al. [24] developed Agatha, a Bayesian neural network that identifies multi-planet signals in noisy time series.

In direct imaging, Cantalloube et al. [25] launched the Exoplanet Imaging Data Challenge, benchmarking ML and classical algorithms on SPHERE and GPI data. Approaches combining PCA with CNNs or GANs achieved superior sensitivity compared with traditional ADI.

4.2. Deep learning for direct imaging

In imaging, supervised CNNs classify small pixel patches as “planet” or “background.” The network learns translation-invariant spatial features,ideal for distinguishing circular PSF-like signals from random speckles. Training requires labeled datasets: either real detections or simulated injections of fake planets into residual frames [26]. Data augmentation (rotation, flipping, noise variation) improves generalization.

GANs further expand training datasets by generating realistic synthetic examples. A discriminator trained against such fakes develops enhanced sensitivity to subtle anomalies. Variational autoencoders (VAEs) and transformer architectures have recently been explored for unsupervised anomaly detection, identifying outlier patches without explicit labels.

Performance metrics include precision, recall, F_1 score, and receiver-operating-characteristic (ROC) curves. On benchmark datasets, CNNs achieve accuracies ≈ 90 %, while GAN-based discriminators reach similar levels with lower false-positive rates. Visual interpretability tools like Grad-CAM highlight the image regions influencing classification, verifying that the network focuses on planetary PSF cores rather than noise.

4.3. Machine learning in other domains

• Transit Surveys: ML classifiers expedite vetting of millions of light curves. The Autovetter and ExoMiner frameworks [27] use neural and tree-based models for candidate ranking.

• Radial Velocities: Gaussian-process kernels augmented by neural networks capture stellar jitter correlations [9].

• Microlensing: RNNs flag short planetary deviations in high-cadence photometry.

• Astrometry: ML regressors detect subtle periodic motion in noisy Gaia data.

• Atmospheric Retrieval: Neural emulators accelerate forward models, mapping spectra to atmospheric parameters orders of magnitude faster than traditional radiative-transfer inversion [28].

5. Challenges and limitations

Despite remarkable progress, several obstacles constrain ML-based exoplanet detection.

DOI: 10.54254/2753-8818/2026.HZ29784

• Limited Labeled Data: Genuine exoplanet detections are rare compared with negative samples. Simulated injections may not perfectly reproduce real noise statistics, leading to domain mismatch. Generative models and unsupervised learning mitigate but do not fully solve this.

• Interpretability: Deep networks often act as black boxes. Explainable-AI (XAI) techniques,saliency maps, SHAP values,are increasingly applied to ensure astrophysical consistency.

• Overfitting and Generalization: Models trained on one instrument (e.g., SPHERE) may perform poorly on another (e.g., GPI). Domain-adaptation and transfer-learning strategies attempt to overcome this by retraining on small labeled subsets.

• Physical Constraints: Data-driven algorithms may ignore conservation laws or PSF symmetry. Physics-informed neural networks (PINNs) embed optical or orbital equations directly into loss functions, combining ML flexibility with physical validity.

• Computational Demands: Training 3D CNNs on multi-wavelength cubes requires GPUs and extensive memory. Efficient architectures (MobileNet, UNet) and on-the-fly data compression are under development to support real-time pipelines.

• Bias and False Positives: ML classifiers may reproduce training-set biases. Rigorous cross-validation, injection-recovery tests, and human verification remain essential before publication.

6. Future prospects

6.1. Next-generation facilities

The next decade will witness dramatic expansion in observational capabilities:

• Roman Space Telescope (launch ≈ 2027) will combine wide-field microlensing and coronagraphy, testing contrasts of 10^{-9} [14]. ML-driven wavefront control and real-time speckle suppression are integral to mission design.

• Extremely Large Telescopes (ELT, TMT, GMT): 30–40 m apertures with advanced AO will achieve milliarcsecond resolution. Automated ML pipelines will be mandatory for terabyte-scale nightly data.

• JWST and Ariel: will extend atmospheric spectroscopy; ML retrieval frameworks will handle the high dimensionality of molecular parameter spaces.

• Habitable Worlds Observatory (HWO): a proposed flagship mission targeting reflected-light Earth analogs; exascale data analysis will depend on AI-based adaptive scheduling and classification.

6.2. Synergy between physics and data science

The frontier lies in hybrid models combining physical understanding with data-driven learning:

• Bayesian Deep Learning: provides probabilistic outputs and credible intervals, allowing robust propagation of uncertainties.

• End-to-End Differentiable Pipelines: link raw pixel data to physical parameters through neural surrogates of PSF modeling and instrument calibration.

• Active Learning: allows ML models to request new training data from human experts, efficiently refining boundaries between real and spurious signals.

• Federated Learning: enables cross-observatory training without sharing proprietary data, ensuring consistency across missions.

As telescopes become increasingly autonomous, ML algorithms will not only identify candidates but also optimize exposure times, prioritize targets, and even recalibrate instruments in situ.

DOI: 10.54254/2753-8818/2026.HZ29784

6.3. Scientific implications

Improved detection efficiency will refine estimates of planetary occurrence rates, informing models of planet formation and migration. Large, uniform ML analyses of Kepler, TESS, and Gaia data already reveal that planets are more common around low-mass stars and that compact multi-planet systems dominate short periods [29]. Direct imaging enhanced by deep learning will characterize colder populations bridging the gap between hot Jupiters and free-floating planets, providing the missing link between core accretion and gravitational instability formation pathways.

7. Conclusion

Exoplanet detection has progressed from precision radial-velocity spectroscopy to machine-learning-assisted, multi-technique discovery. Each method,transit, RV, microlensing, astrometry, and direct imaging,illuminates a different dimension of planetary diversity. The convergence of high-contrast instrumentation, statistical inference, and deep learning has ushered in an era of intelligent astronomy. ML accelerates discovery, enhances reliability, and opens the door to automated observatories capable of self-calibrating and self-classifying their data streams.

Yet challenges remain: interpretability, generalization, and the limited physical grounding of current models. The future of exoplanet science lies in synergizing theory and data, physics and computation. As upcoming missions probe deeper into the habitable zone, data-driven algorithms will be the key to recognizing the faint signatures of new worlds orbiting distant stars.

References

[1] Mayor, M., & Queloz, D. 1995, A Jupiter-mass companion to a solar-type star. Nature, 378, 355.
[2] Borucki, W. J., Koch, D. G., Basri, G., et al. 2010, Kepler Planet-Detection Mission: Introduction and First Results. Science, 327, 977.
[3] Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2015, Transiting Exoplanet Survey Satellite (TESS). JATIS, 1, 014003.
[4] Kovács, G., Bakos, G., & Noyes, R. 2005, A trend filtering algorithm for wide-field variability surveys. MNRAS, 356, 557.
[5] Shallue, C. J., & Vanderburg, A. 2018, Identifying Exoplanets with Deep Learning: A Five-planet Resonant Chain around Kepler-80 and an Eighth Planet around Kepler-90. AJ, 155, 94.
[6] Charbonneau, D., Brown, T. M., Latham, D. W., & Mayor, M. 2000, Detection of Planetary Transits Across a Sun-like Star. ApJ, 529, L45.
[7] Pepe, F., Cristiani, S., Rebolo, R., et al. 2021, ESPRESSO at VLT. On-sky performance and first results.. A&A, 645, A96.
[8] Haywood, R. D., Collier Cameron, A., et al. 2020, MNRAS, 496, 4769.
[9] Rajpaul, V., Aigrain, S., Osborne, M. A., et al. 2015, A Gaussian process framework for modelling stellar activity signals in radial velocity data. MNRAS, 452, 2269.
[10] Dumusque, X., et al. 2021, Three years of HARPS-N high-resolution spectroscopy and precise radial velocity data for the Sun. A&A, 648, A103.
[11] Mao, S., & Paczynski, B. 1991, Gravitational Microlensing by Double Stars and Planetary Systems. ApJ, 374, L37.
[12] Gaudi, B. S. 2012, Microlensing Surveys for Exoplanets. ARA&A, 50, 411.
[13] Kim, D., Chung, S.-J., & Kim, H. 2018, The Korea Microlensing Telescope Network Microlensing Event Finder, ApJ, 869, 55.
[14] Penny, M. T., Gaudi, B. S., Kerins, E., et al. 2019, Predictions of the WFIRST Microlensing Survey. I. Bound Planet Detection Rates, ApJS, 241, 3.
[15] Perryman, M. A. C., Hartman, J., Bakos, G. A., & Lindegren, L. 2014, Astrometric Exoplanet Detection with Gaia, ApJ, 797, 14.
[16] Hinkley, S., Vigan, A., Kasper, M., et al. 2021, ExoFrontiers: Charting the Frontiers of Exoplanet Discovery and Characterization, ExoFrontiers, 5–1.

[17] Marois, C., Macintosh, B., Barman, T., et al. 2008, Direct Imaging of Multiple Planets Orbiting the Star HR 8799, Science, 322, 1348.
[18] Lagrange, A.-M., Bonnefoy, M., Chauvin, G., et al. 2010, A Giant Planet Imaged in the Disk of the Young Star β Pictoris, Science, 329, 57.
[19] Beuzit, J.-L., Vigan, A., Mouillet, D., et al. 2019, SPHERE: The Exoplanet Imager for the Very Large Telescope, A&A, 631, A155.
[20] Chilcote, J., Konopacky, Q., Macintosh, B., et al. 2022, The Gemini Planet Imager: Upgrades and On-sky Performance after the Gemini North Move, Proc. SPIE, 12184, 121841T.
[21] Marois, C., Lafrenière, D., Doyon, R., et al. 2006, Angular Differential Imaging: A Powerful High-contrast Imaging Technique, ApJ, 641, 556.
[22] Currie, T., Biller, B., Lagrange, A.-M., et al. 2023, Directly Imaged Exoplanets and Circumstellar Disks, in Protostars and Planets VII, 799.
[23] Ansdell, M., Ioannou, Y., Osborn, H. P., et al. 2018, The Exoplanet Explorers Project. II. Discovery of a Hot Sub-Saturn Planet Orbiting HD 73344, AJ, 158, 138.
[24] Hara, N., Delisle, J.-B., Unger, N., et al. 2022, Kima: An Automatic Bayesian Framework for Exoplanet Detection in Radial Velocity Data, A&A, 658, A171.
[25] Cantalloube, F., Christiaens, V., Cantero, C., et al. 2024, Advances in Exoplanet High-contrast Imaging with the SPHERE+ Instrument, Proc. SPIE, 13097, 1309713.
[26] Flasseur, O., Bodrito, T., Mairal, J., et al. 2024, PACOME: A Probabilistic Framework for Exoplanet Detection in Direct Imaging, A&A, 527, 1534.
[27] Valizadegan, H., et al. 2022, Exo-MAST: An Integrated Portal for Exoplanet Data Access and Analysis, ApJS, 258, 39.
[28] Zingales, T., & Waldmann, I. P. 2018, ExoGAN: Retrieving Exoplanetary Atmospheres Using Deep Convolutional Generative Adversarial Networks.. AJ, 156, 268.
[29] Mulders, G. D., Pascucci, I., & Apai, D. 2018, The Exoplanet Population Observation Simulator. I. The Inner Edges of Planetary Systems. AJ, 156, 24.