International Journal of Scientific Research and Engineering Development-– Volume 8 Issue 5, Sep-Oct 2025
Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

How Artificial Intelligence Will Help Humans to Study The Habitable Planet in The Universe

Manav Akbari[1], Mr Yagnesh Makwana [2]
[1](B Tech in Computer Science Engineering, Atmiya University, Rajkot, India
Email: akbarimanav04@gmail.com)
[2](Faculty of Engineering and Technology (CE), Atmiya University, Rajkot, India
Email: Yagnesh.makwana@atmiyauni.ac.in)

Abstract: The search for habitable worlds beyond Earth is one of humanity's grand quests. As telescopes produce increasing amounts of data, artificial intelligence (AI) and machine learning methods are essential for identifying, filtering, and characterizing potentially habitable exoplanets. This paper explores how artificial intelligence contributes to each step of the detection process-from raw data processing and candidate selection to analysing atmospheres and scoring habitability. We discuss existing successes, challenges, and future possibilities. The aim is to show that AI is not just helpful but necessary for speeding up the discovery of habitable planets.

Keyword - AI in space exploration, Exoplanet detection using AI, AI for astrobiological exploration.

1. Introduction: Human curiosity to find Earth like Habitable planet

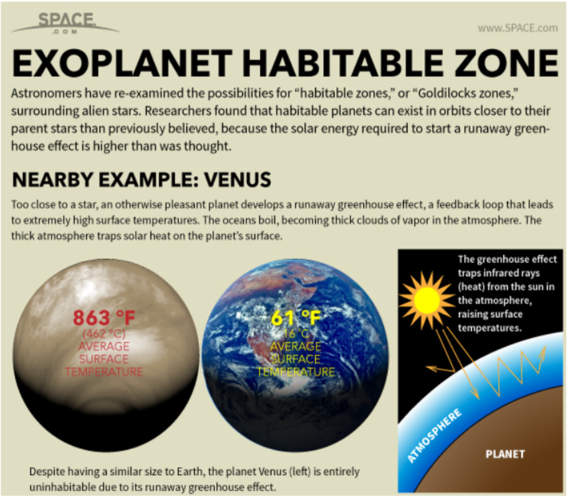
Human curiosity about life beyond Earth has driven missions such as Kepler, TESS, JWST, and upcoming telescopes. These missions generate vast amounts of data, often containing weak or noisy signals of exoplanets hidden in light curves, spectra, images, and time series. Traditional methods—manual inspection, thresholding, and hand-designed criteria—are becoming less effective and prone to mistakes.

Artificial Intelligence (AI), particularly machine learning (ML) and deep learning (DL), can automatically detect patterns, identify potential candidates, and interpret subtle signals. In this paper, we trace how AI is changing each stage of exoplanet and habitability research. We also discuss limitations and future directions.

2. Background: Habitable zones, planets, and Astrobiology

2.1 The Concept of Habitable Zone

The habitable zone (sometimes called the Goldilocks zone) is the area around a star where conditions might allow liquid water to exist on a planet's surface. Since liquid water is believed to be essential for life, the habitable zone is a primary consideration when evaluating exoplanet habitability. However, this criterion is oversimplified. Factors like atmospheric composition, internal heating, magnetic fields, and other planetary characteristics significantly influence actual habitability.



2.2 Exoplanet Detection Methods

Transit method: measuring periodic dips in a star's brightness caused by a planet passing in front of it.

Radial velocity (wobble) method: detecting slight shifts in a star's spectrum due to the gravitational pull from an orbiting planet.

Direct imaging: using powerful instruments to sometimes capture an image of a planet by blocking starlight.

Gravitational micro lensing: utilizing lensing effects when a foreground star-planet system passes in front of a more distant star.

Each method generates data (light curves, spectra, and images) that must be analyzed, filtered, and verified. Traditional methods often discard borderline cases or need human review, risking the omission of subtle or rare signals.

3. AI / Machine Learning in Exoplanet Discovery & Classification

AI has become an integral part of exoplanet research. Here's how:

3.1 Transit Detection & Signal Extraction

AI models, particularly convolutional neural networks (CNNs) and other deep learning architectures, can detect weak periodic transit signals more effectively than traditional algorithms, especially in noisy data. For example:

Shallue & Vanderburg's work used CNNs to find exoplanet transits in Kepler data, surpassing traditional least-squares methods. Recent efforts merge real and synthetic light curves to train models more effectively, improving sensitivity to longer or fainter transits. A comparative study using Kepler data evaluated multiple supervised ML algorithms (Random Forest, K-Nearest Neighbours, Decision Tree, and Logistic Regression) and found that Random Forest achieved about 99.8% accuracy for exoplanet classification after balancing classes. These techniques facilitate automatic candidate selection and reduce false negatives.

3.2 Reducing False Positives & Validation

One challenge is that many "signals" are false positives (e.g., stellar variability, instrument noise, binary star eclipses). AI models assist by:

Using anomaly detection or auto encoder-based methods to filter out false positives. For example, the ArtAe model employs neural networks and auto encoders to validate Kepler and TESS candidates, achieving around 93-94% accuracy in distinguishing real from spurious signals. Training on known false positives and actual candidates helps models learn features that differentiate genuine exoplanets. Ensemble learning and cross-validation help reduce over fitting and enhance generalization. Thus, AI helps refine the candidate list and spotlight the most promising ones for further study.

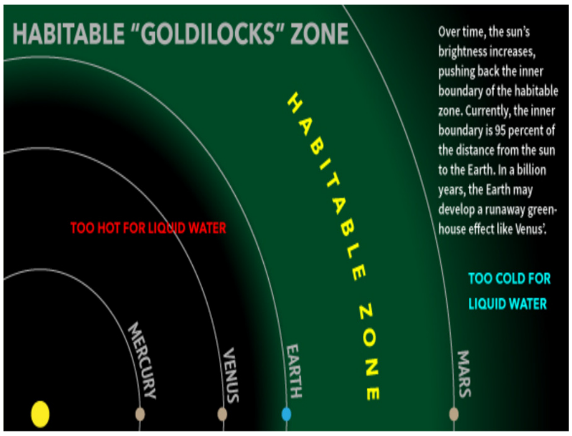
3.3 Atmospheric Retrieval & Composition Interface

Once a candidate planet is detected, scientists strive to characterize its atmosphere (gas composition, temperature, pressure, clouds). AI plays a key role here:

Machine learning models can invert spectral observations to estimate atmospheric parameters (e.g., detecting water, methane, CO_2). Some studies create large synthetic spectral datasets to train models for retrieval tasks. Techniques like neural networks, Gaussian processes, and Bayesian inference speed up what would otherwise be lengthy, iterative modelling. In high-contrast imaging (direct imaging), ML is applied to suppress stellar noise, separate planetary signals, and reduce false positives. For instance, a recent study used ML techniques on integral field spectroscopy data to enhance detection limits for directly imaged exoplanets.

Consequently, AI helps convert raw spectral data into physical and chemical profiles of exoplanet atmospheres.

3.4 Habitability Scoring and Classification



With many candidate planets and their atmospheric data, the next step is to score or classify planets based on their potential for habitability. AI assists with:

Habitability metrics: Traditional metrics include the Earth Similarity Index (ESI) and advanced ones like Constant Elasticity of Exponential Stellar Attractiveness (CEESA). Some recent studies combine computational intelligence with optimization to create new habitability indices.

Machine learning classification: Models trained on known and simulated planets can categorize new candidates into habitability classes (e.g., habitable, marginal, non-habitable).

For instance, one study applied logistic regression and random forest classifiers to features from Kepler archive data to predict the habitability of orbiting exoplanets. The review "Assessing Exoplanet Habitability through Data-driven Approaches" highlights trends in applying machine learning, feature engineering, and domain-specific constraints to evaluate habitability. These methods assist in ranking or prioritizing exoplanet candidates that deserve more thorough investigation.

4. Integrative Workflow: From Data to Discovery

Here's a potential pipeline showing how AI aids the journey from raw data to identifying a promising habitable planet:

I. Data acquisition (space telescopes, ground telescopes)

II. Pre-processing & cleaning (remove instrument noise, adjust light curves)

III. Transit / signal detection using AI models (CNNs, RNNs, and ensemble methods)

IV. False positive filtering & validation (auto encoders, anomaly detection, ensemble classifiers)

V. Atmospheric retrieval (inverse modelling using ML to map spectra to gas composition and temperature)

VI. Habitability scoring / classification (applying metrics and ML classification)

VII. Selection of prioritized candidates for follow-up observations

VIII. Iterative improvement: feedback from actual observations refines the AI models Since each stage is complex, AI enables automation, better scalability, and the discovery of signals that human methods might overlook.

5. Case Studies & Recent Achievements

A recent Nature paper showed that even basic machine learning models, when engineered thoughtfully, can significantly aid exoplanet classification tasks. The systematic review A Systematic Review of Machine Learning and Deep Learning Techniques for Exoplanet Detection (2025) examines state-of-the-art approaches and highlights future challenges. AI tools developed at institutions like UNIGE and UniBE have been applied to detect exoplanets and lessen the workload in large surveys. In a 2024 astronomy and astrophysics paper, machine learning improved detection limits in direct imaging by processing spectral datasets and minimizing false positives. These successes show that AI is advancing from proof-of-concept to a vital part of research on exoplanets.

6. Challenges, Limitations & Risks

While AI is powerful, it comes with challenges and risks:

Data biases and imbalance: There are many more negative cases than confirmed exoplanets, leading to potential model over fitting or misclassification. Synthetic data augmentation can help address this.

Explain-ability /interpretability: Deep models can act as "black boxes." In astronomy, scientists prefer models that are interpretable or hybrid methods that combine domain knowledge.

False positives / overconfidence: Relying too heavily on AI predictions without human validation can lead to incorrect conclusions.

Generalization to new instruments: Models trained on Kepler data may not work as well on TESS, JWST, or future instruments unless carefully modified.

Computational cost: Training on large datasets and performing spectral inversions can require significant computing resources.

Physical realism: AI models must adhere to physical constraints (e.g., chemistry, thermodynamics). Purely data-driven models may suggest unrealistic interpretations if not checked.

Addressing these issues requires hybrid approaches, active learning (with human involvement), uncertainty quantification, and thorough model validation.

7. Outlook & Future Directions

Looking ahead, AI will continue to enhance its influence in several ways:

Transfer learning & domain adaptation: Models that can adjust from one instrument or star system to another will help improve generalization quickly.

Active learning / human involvement: AI can suggest ambiguous cases to experts for annotation, progressively enhancing model performance.

Integration with robotics / probes: When human or robotic probes visit moons or planets (e.g., Europa, Titan), on-board AI can manage exploration, sample selection, and on-site characterization.

Multimodal AI: Combining light curves, spectra, images, and Astrometric data into unified models will provide richer insights.

Explainable AI in astronomy: Emphasizing models whose decisions can be traced and justified will build more scientific trust.

Crowdsourcing + citizen science + AI: Hybrid systems where AI flags candidates and citizen scientists or astronomers refine them.

Better habitability simulators: AI could simulate planetary climate, interior, and atmospheric changes, bridging the gap between data and theory. These directions promise to speed up the search for not just exoplanets, but also habitable exoplanets.

8. Conclusion

Combination of high-sensitivity instruments and vast amounts of astronomical data pushes us to adopt artificial intelligence and machine learning methods. From detection to atmospheric analysis to habitability scoring, AI streamlines and enhances the workflow, making the search for habitable worlds more efficient and less reliant on manual methods. Although challenges remain interpretability, data biases, and instrument generalization ongoing research is gradually overcoming them. In the coming decade, AI is likely to be crucial in guiding us to the most promising habitable planets in the universe and potentially, to the discovery of life beyond Earth.

9. References & Further Reading

Below are reviews, and sources used in crafting this paper. You can click to verify or dive deeper:

1. Machine Learning in Exoplanet Detection (SSRN) — review and hybrid proposals SSRN
2. Machine Learning for Exoplanet Detection: A Comparative Analysis Using Kepler Data (2025) arXiv
3. Searching for Exoplanets Using Artificial Intelligence (Oxford / MNRAS) OUP Academic
4. Deep Learning Exoplanets Detection by Combining Real and Synthetic PMC
5. A Systematic Review of Machine Learning and Deep Learning Techniques for Exoplanet Detection (2025) ResearchGate
6. Assessing Exoplanet Habitability through Data-driven Approaches (2023) Astrobiology
7. Feature Engineering for False Positive Exoplanet Prediction: A Deep Learning Approach (ArtAe) IJISAE
8. Machine Learning techniques in studies of the interior of exoplanets (mixture density networks) A&A Publishing

9. Machine learning for exoplanet detection in high-contrast imaging (2024) A&A Publishing+1
10. Predicting the Habitability of Exoplanets Using Statistical Modelling & ML IAEME
11. CEESA meets machine learning: A Constant Elasticity model for habitability ScienceDirect+1

