Exoplanet Detection from Light Curves using Deep Learning

Rushank Tripathi, Hitesh S
Department of Computer Science and Engineering,
Amrita School of Computing,Amrita Vishwa Vidyapeetham
Bengaluru, India
bl.en.u4cse22051, bl.en.u4cse20058

***Abstract*—Utilizing cutting-edge machine learning techniques, the project "Exoplanet Detection from Light Curves using Deep Learning" seeks to detect exoplanets from the periodic dimming patterns seen in star light curves. Light curves are created by tracking a star's brightness over time, with exoplanets producing minute changes when they pass in front of the star. These light curves are analyzed and possible exoplanetary transits are classified using deep learning models, namely convolutional neural networks (CNNs) and recurrent neural networks (RNNs). Preprocessing the light curve data, creating and training deep learning models, and assessing the models' accuracy, precision, and recall are all part of the project. This method improves the effectiveness of finding new planets and advances space by automating exoplanet detection and exploration.**

***Keywords*- Exoplanet Detection, Light Curves, Deep Learning, Neural Networks, CNNs, RNNs, Transit Method, Time Series Analysis, Machine Learning, Space Exploration.**

I. Introduction

One of the main goals of contemporary astronomy has been the finding of exoplanets, or planets circling stars outside of our solar system. Finding these far-off worlds offers important information about the genesis of planets, their habitability, and the hunt for extraterrestrial life. Radial velocity measurements, direct imaging, and the transit method—which uses a periodic decrease of a star's brightness to suggest a planet crossing its face—are examples of traditional exoplanet discovery techniques. With the help of space missions like Kepler, TESS (Transiting Exoplanet Survey Satellite), and CoRoT, thousands of exoplanets have been detected thanks to the transit method, which has proven to be very successful among these. However, it takes a lot of time and is prone to human mistake to manually analyze the enormous amount of light curve data generated by these missions. In order to tackle this difficulty, deep learning methods have become a powerful tool to automate and enhance exoplanet detection. In pattern recognition tasks including image classification, audio processing, and medical diagnostics, deep learning—a subset of machine learning—has demonstrated impressive results. Deep learning may effectively detect exoplanetary transits by identifying minute variations in light curves when applied to astronomical data. Given their ability to identify sequential patterns in the data and extract valuable features, Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) are especially well-suited for studying light curves. RNNs, which focus on sequential data processing, capture temporal dependencies in light curves, whereas CNNs, which are frequently used in image recognition, assist in identifying important transit properties. Exoplanet candidates can be robustly classified thanks to the combination of these models.

Creating an automated deep learning-based method to identify exoplanets from stellar light curves is the main goal of this project. This entails a few crucial steps: Data Collection Preprocessing: The project uses datasets from the Kepler and TESS missions, among other publically accessible sources. To improve model performance, raw light curve data is preprocessed using techniques like noise removal, normalization, and augmentation. Model Development Training: To identify transit signatures, CNNs and RNNs are trained using labeled light curve data. To increase the accuracy of classification, different architectures and hyperparameter tuning strategies are used.

Evaluation Optimization: Metrics including accuracy, precision, recall, and F1-score are used to evaluate the model's performance. To improve resilience, fine-tuning techniques like data augmentation and dropout regularization are used.

Deployment Automation: To analyze real-time astronomical data and enable continuous and effective exoplanet detection, the trained model can be included into an automated pipeline.

This research intends to improve the precision and effectiveness of transit-based detection methods by incorporating deep learning into exoplanet identification, hence decreasing the necessity for laborious manual analysis. In addition to speeding up the discovery process, automation of exoplanet identification frees up researchers to concentrate on describing recently discovered planets rather than laboriously sorting through enormous databases. Additionally, this study shows the potential of artificial intelligence in space travel and advances the subject of computational astrophysics. To sum up, the combination of deep learning with astronomy offers a novel method for finding exoplanets. We may improve our knowledge of planetary systems, find new worlds more quickly, and get closer to the age-old question of whether we are alone in the universe by utilizing machine learning techniques.

II. Literature Survey

[1] Exoplanets can be detected using various observational techniques, with **radial velocity (RV)** playing a crucial role in revealing planetary system architecture and measuring planet mass and orbits. RV is essential for detecting Earth-like planets but faces challenges due to complex, correlated instrumental and astrophysical noise. Additionally, efficiently exploring RV models' large parameter space is difficult. RV data analysis involves detecting planetary signals in unevenly sampled, multivariate time series. This review aims to introduce RV analysis challenges to nonspecialists and unify existing advanced approaches. Identifying areas for improvement can enhance the accuracy of exoplanet detection using RV methods.

[2] This study introduces a machine learning-based approach for exoplanet detection using the transit method. By leveraging the tsfresh library, 789 features were extracted from light curves, which were then used to train a gradient boosting classifier (LightGBM). The method was tested on K2 campaign 7 data with artificial transit signals, showing competitive performance compared to conventional box least-squares fitting. It achieved an AUC of 0.948 for Kepler data, with a recall of 0.96. For TESS data, the model attained 98 percent accuracy and identified planets with 82 percent recall at 63 percent precision. This approach is computationally efficient and comparable to state-of-the-art deep learning models.

[3] The study of exoplanetary systems can help us understand the formation and evolution of the solar system itself and search for terrestrial planets in the habitable and extrasolar lives in exoplanetary systems. Exoplanets have become an important area of astrophysics in the last two decades. This paper reviews five different methods to detect exoplanets, including direct imaging, astrometry, radial velocity, transit event observation, and microlensing. These approaches could expand the sample size of exoplanets and further our understanding of the types, formation and evolution of exoplanets.

[4] A space-based mid-infrared (MIR) nulling interferometer is proposed to detect and characterize exoplanets by measuring their thermal emissions. A Monte Carlo instrument simulator was used to estimate detection yields for exoplanets within 20 pc of the Sun. A four 2m-aperture design could detect 550 exoplanets, including 25–45 rocky planets in habitable zones (eHZ), while larger apertures improve yields. Most small, temperate exoplanets are found around M dwarfs, and varying wavelength ranges has minimal impact. This competes with single-aperture reflected light missions, enabling the study of planetary atmospheres. A multi-visit strategy could further increase habitable planet detections.

[5] Exoplanet detection is a key research area, traditionally relying on methods like Radial Velocity, Transit Method, and Direct Imaging. This study explores Artificial Intelligence for detection, using machine learning models such as Decision Trees, SVMs, Logistic Regression, Random Forest, MLP, and CNNs. An Ensemble-CNN model is introduced, outperforming baseline models with 99.62 percent accuracy. Performance is evaluated using Accuracy, Precision, Sensitivity, and Specificity. The research bridges Astronomy and AI, aiding scientists and machine learning experts in exoplanet detection. It has significant implications for physicists, cosmologists, and industry professionals.

[6] Direct imaging of exoplanets is challenging due to high contrast and small angular separation from their host stars. Traditional post-processing methods rely on self-observed nuisance models, limiting detection sensitivity. This study proposes a supervised deep learning approach, leveraging an archive of multiple observations to improve nuisance modeling. Unlike reference differential imaging, the method is highly non-linear and avoids direct image-to-image subtraction. Tested on VLT/SPHERE data, it outperforms the PACO algorithm, especially when angular diversity is low. This approach enhances detection sensitivity and robustness, improving exoplanet imaging capabilities.

[7] Exoplanets are planets beyond our solar system, with detection crucial for finding Earth-like habitable worlds. This study explores machine learning classification using time-series light intensity data from the NASA Kepler mission (Campaign 3). Various classification algorithms were tested, with K-Nearest Neighbors (KNN) on SMOTE-balanced data achieving 98.20 percent accuracy and 98 percent F1-score. To address real-time data limitations, models were trained on reduced data (1 percent–75 percent), with KNN still achieving 90.1 percent accuracy at 1 percent data. The results highlight the potential of machine learning in efficient exoplanet detection using minimal data.

[8] Astrometry is less affected by stellar activity than radial velocity, making it suitable for detecting Earth-mass exoplanets in habitable zones of solar-type stars. This study re-evaluates detection limits for 55 F-G-K stars targeted by the THEIA astrometric mission using realistic stellar activity simulations. Results show worse detection limits than previous estimates, but with low uncertainty in fitted parameters. Stellar activity has minimal impact except for nearby stars like Cen A and B. The best targets are stars with close habitable zones, while subgiants may require a longer mission for detection.

[9] Direct imaging of exoplanets requires high contrast and angular resolution, aided by adaptive optics, coronagraphs, and advanced post-processing algorithms. This study introduces

PACOME, a new algorithm that combines multi-epoch observations to enhance detection sensitivity by accounting for Keplerian orbital motion. PACOME extends the maximum likelihood framework of PACO, improving detection and estimating orbital elements with error bars. Tested on VLT/SPHERE data, it detects previously undetectable exoplanets, with sensitivity scaling as (number of epochs). Applied to HR 8799, it successfully identifies all four known exoplanets. PACOME is efficient, automated, and enhances exoplanet discovery.

[10] present a fuzzy logic-based database engine designed to improve MongoDB's capabilities in handling imprecise or uncertain data. The authors discuss the problem of managing fuzzy data, which is common in real-world applications such as decision-making systems and natural language processing. They suggest integrating a fuzzy engine within MongoDB to support operations such as fuzzy querying, fuzzy searching, and fuzzy reasoning. The paper shows how this kind of approach enhances the flexibility of dealing with approximate values and information uncertainty in MongoDB. Another potential benefit of using fuzzy logic is that it will contribute to more intuitive and human-like processes of data retrieval and decision.

[11] Brown dwarfs and exoplanets share atmospheric characteristics, but key molecules like chromium hydride (CrH) remain unconfirmed in exoplanets. CrH is crucial for temperature probing and spectral classification in brown dwarfs. Recent low-resolution studies suggested CrH in WASP-31b, a hot Jupiter. Using high-resolution data from GRACES/Gemini North and UVES/VLT, this study confirms CrH detection at 5.6 significance. This marks the first high-resolution detection of a metal hydride in an exoplanet. The finding advances understanding of metal hydrides in exoplanet atmospheres.

The research on optimizing the performance of search and query focuses on scalability, efficiency, and security in modern databases. Key areas of improvements include hybrid systems that include the strengths of NoSQL and traditional databases, followed by the integration of complex techniques such as machine learning, Particle Swarm Optimization, and blockchain. These innovations target the problem of handling big, unstructured data at low latency and high throughput. Additionally, the adoption of advanced indexing strategies and encryption methods ensures fast data retrieval and secure transactions, especially in sensitive environments such as healthcare. Overall, the aim is to streamline data processing, reduce redundancy, and provide more personalized, accurate search results in complex, real-time applications.

REFERENCES

[1] Hara, N.C. and Ford, E.B., 2023. Statistical methods for exoplanet detection with radial velocities. Annual Review of Statistics and Its Application, 10(1), pp.623-649.

[2] Malik, A., Moster, B.P. and Obermeier, C., 2022. Exoplanet detection using machine learning. Monthly Notices of the Royal Astronomical Society, 513(4), pp.5505-5516.

[3] Dai, Z., Ni, D., Pan, L. and Zhu, Y., 2021, September. Five methods of exoplanet detection. In Journal of Physics: Conference Series (Vol. 2012, No. 1, p. 012135). IOP Publishing.

[4] Quanz, S.P., Ottiger, M., Fontanet, E., Kammerer, J., Menti, F., Dannert, F., Gheorghe, A., Absil, O., Airapetian, V.S., Alei, E. and Allart, R., 2022. Large Interferometer For Exoplanets (LIFE)-I. Improved exoplanet detection yield estimates for a large mid-infrared space-interferometer mission. Astronomy Astrophysics, 664, p.A21.

[5] Priyadarshini, I. and Puri, V., 2021. A convolutional neural network (CNN) based ensemble model for exoplanet detection. Earth Science Informatics, 14(2), pp.735-747.

[6] Bodrito, T., Flasseur, O., Mairal, J., Ponce, J., Langlois, M. and Lagrange, A.M., 2024. MODELCO: Exoplanet detection in angular differential imaging by learning across multiple observations. Monthly Notices of the Royal Astronomical Society, 534(2), pp.1569-1596.

[7] Bahel, V. and Gaikwad, M., 2022, July. A study of light intensity of stars for exoplanet detection using machine learning. In 2022 IEEE Region 10 Symposium (TENSYMP) (pp. 1-5). IEEE.

[8] Meunier, N. and Lagrange, A.M., 2022. A new estimation of astrometric exoplanet detection limits in the habitable zone around nearby stars. Astronomy Astrophysics, 659, p.A104.

[9] Dallant, J., Langlois, M., Flasseur, O. and Thiébaut, É., 2023. PACOME: Optimal multi-epoch combination of direct imaging observations for joint exoplanet detection and orbit estimation. Astronomy Astrophysics, 679, p.A38.

[10] Visser, K., Bosma, B. and Postma, E., 2022. Exoplanet detection with Genesis. Journal of Astronomical Instrumentation, 11(03), p.2250011.

[11] Flagg, L., Turner, J.D., Deibert, E., Ridden-Harper, A., de Mooij, E., MacDonald, R.J., Jayawardhana, R., Gibson, N., Langeveld, A. and Sing, D., 2023. ExoGemS Detection of a metal hydride in an exoplanet atmosphere at high spectral resolution. The Astrophysical Journal Letters, 953(2), p.L19.