International Journal of Scientific Research in Science, Engineering and Technology
Print ISSN: 2395-1990 | Online ISSN : 2394-4099 (www.ijsrset.com)
doi : https://doi.org/10.32628/IJSRSET

A Deep Learning Approach for Identification and Classification of Exoplanets

Mr. Arbaaz Mohammad Shaikh, Miss. Zarinabegam Mundargi
Department of Data Science, Zeal College of Engineering and Research Pune, Maharashtra, India

Article Info
Volume 9, Issue 2
Page Number : 459-465

Publication Issue :
March-April-2022

Article History
Accepted :15 April 2022
Published: 30 April 2022

ABSTRACT

More than a million stars have been observed in the last decade for transiting planets. The manual interpretation of candidates for exoplanets is labor-intensive and susceptible to human mistake, the consequences of which are difficult to measure. A neural network, rather than the more traditional ways of discovering exoplanet candidates, is being used in large scale planetary search operations. Neural networks, often known as ”deep learning” or ”deep nets,” are designed to offer a computer understanding of a given problem by teaching it to identify patterns. As a result, Earth-sized planets orbiting Sunlike stars have eluded NASA's Kepler Space Telescope, which was built to find out how common they are. Individual candidates for planets will need to be automatically and precisely assessed for their likelihood of becoming planets, even at low signal-to-noise ratios. Deep learning, a family of machine learning methods that has recently become state-of-the-art in a wide range of problems, is used in this paper to classify probable planet signal. Deep convolutional neural networks are used to identify whether an exoplanet transiting the star is real or a false positive caused by an astrophysical or instrumental event. There is a huge research gap between the identification of such exoplanets at such highest level of complex artificial intelligence model. The major gap analysis is the time required to identify the planet is an exoplanet or not an exoplanet. The proposed model also adds the state of art extension algorithm which is capable of classifying the exoplanets into standard types of the planets. The goal of this project is to analyze data from the Kepler telescope and create a state of art model that can identify exoplanets and classify them into different types of planets time efficiently and with high accuracy.

Keywords: Exoplanets Exoplanets Detection, Machine Learning, Deep learning, Convolutional Neural Network Transit Theory, Kepler

I. INTRODUCTION

Our solar system's all planets orbit around the sun. The Exoplanet Exploration Program of NASA studies planets that orbit other stars and calls them exoplanets. With telescopes, exoplanets are extremely difficult to view, because of the stars they orbit, they are obscured by their own light. [1] For many years, astronomers have looked into the possibility of discovering extra-solar planets. An exoplanet orbiting Van Maanen 2 has been theorised since 1917, but its existence could not be proven due to restricted technologies at the time.. Humanity has been searching for a new home on exoplanets for many decades in the hope of finding life elsewhere in the universe. In 2009, NASA launched its Kepler Space Telescope to search for Earth-sized planets orbiting in the habitable zone of their star. The "transit method" is the most common way to look for planets. [2] The

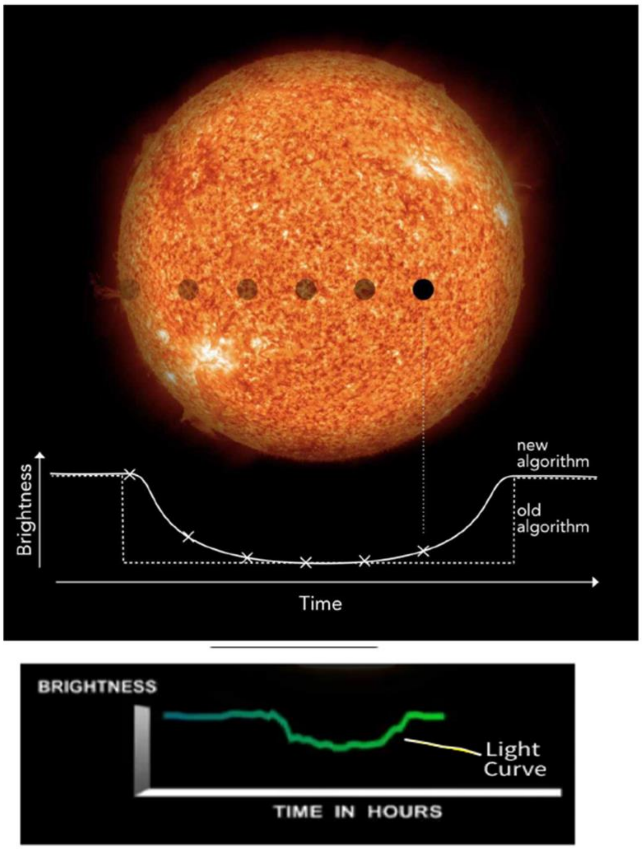

Fig 1. Kepler Telescope Transit Methodology

Telescope's ability to see the star dims when the planet passes in front of it, blocking the star's light. The starlight re-emerges as the planet crosses the star. [3] Using the flux intensity of 3198 stars, we are trying to determine if there is a planet orbiting a star in deep space. More exoplanets will be found thanks to the use of Deep Neural Networks and Support Vector Machines than have ever been before. [4] K2 and CoRoT, as well as ground-based observations, have found 3513 exoplanets to date. Planet hunting surveys like TESS, PLATO, and LSST are planning to expand the thresholds that restrict current photo-metric surveys by sampling brighter stars at quicker cadences and over broader fields of view. About 15percent of stars in the solar system contain a planet with an orbital period of between 5 and 50 days, according to Kepler's first four-year survey. The discovery of such small Earth-sized planets is difficult because the transit depth, 100 ppm for a solar type star, hits the limit of existing photo-metric surveys and is below the typical stellar variability. [5] Stellar variability is evident in about 25percent of the 133 030 main-sequence Kepler stars and ranges between 950 ppm (5th percentile) and 22 700 ppm (95th percentile) with periodicity between 0.2 and 70d. The processing of data in the future needs to be both sensitive to Earth-like planets and robust to stellar variability.[6]

II. OBJECTIVE AND MOTIVATION

When it comes to astronomy and exploration of the cosmos, there is no such thing as a limit to what can be accomplished. Exoplanets are on the agenda for a number of leading space agencies, including NASA, ISRO, JAXA, and CSA. To determine the exoplanets based on different properties of exoplanets using deep learning technique. Most of the work has been carried out using old - outdated methods only. There is so much to explore in the space and with the new missions to be launched soon can help the humanity to explore the new worlds in the space with the skill of Data Science.

III. REVIEW OF LITERATURE

The most common techniques for discovering planets use least-squares optimization, grid-search, or matched filter approaches to maximise the correlation between data and a basic transit model (Kov'acs et al., 2002[10]; Jenkins et al., 2002[11]; Carpano et al., 2003[12]; Petigura et al., 2013). [13]. Convoluting data with the appropriate filter maximizes the SNR of a transit detection. As a result, kernels are handmade to match the human user's perception of the ideal filter in the situation of varied transit forms. CNNs have already been utilised to solve comparable kernel optimization problems using deep learning techniques like convolutional neural networks (CNNs) (Krizhevsky et al. 2012) [14].

A deep neural network for automated Kepler TCE vetting system is presented in this research. Our model uses light curves as inputs and is trained on a set of Kepler TCEs classified by humans. Neural networks have already been used to solve a range of astronomical issues. Pearson et al. (2017)[15] employed a neural network to detect planet transits in simulated light curves. Deep learning, a form of representation learning, employs computer layers to develop progressively complex features - , which are particularly - beneficial for classification issues (Lecun et al. 2015)[16]. Deep image classification models may begin by detecting simple edge features, which can subsequently be used to detect curves and corners and so on until the model's final feature layer can differentiate between complicated objects. Deep neural networks, which are a sort of deep learning model, have lately become the state-of-the art in a number of tasks (e.g. picture classification, Krizhevsky et al. 2012)[17], and typically outperform models built with hand-designed features.

This Paper [18] proposes ASTRONET, a deep learning architecture to look for exoplanets that are habitable based on different characteristics. The paper refers to the Habitability of Exoplanets using Deep Learning techniques such as ASTRONET a combination of Astrology and Deep neural networks. Models are mainly focused on the habitability, atmosphere and the climatic nature of the exoplanets. The deep learning methodologies used in this paper can be improved and be helpful for more accuracy of the exoplanets identification.

This paper[19] refers to find similar planets using Deep Neural Networks and Support Vector Machines by using the dataset of flux intensity of distant stars. Discovering Exoplanets in Deep Space using Deep Learning Algorithms can be more diversified, since variety of techniques can deliver different results. The paper proposes the hybrid architecture where combination of both Machine learning and Deep learning comes in to achieve better results on the basis of exoplanets habitability dataset.

This Paper[20] proposes CNN technique of deep learning and Time Series to search for most capable candidates for exoplanets. Searching for Exoplanets using Artificial Intelligence is now became a trend. The AI gives the human cognitive power to the model which trains the model in the similar way as the neurons in the human body trains the brain. This helps the model to boost the efficiency and credibility while exploring the different exoplanets. The Deep convolutional layers are helpful in identifying and differentiation in the different classes such as exoplanets. This paper contributes more on the exploring and searching different solar bodies present in the space and identifies the exoplanets.

The literature review is based on the study around several high level journals, articles, books and genuine sources. This study indicates the previous work done in this area of space exploration and research in exoplanets. The study focuses mainly on the existence of the exoplanets based on its basic features and the results generated have less accuracy and cannot be sure if the planet is real exoplanet or a false exoplanet. The study also suggest that all the previous work requires a time efficient and highly accurate model which have the capability of identifying and classifying the exoplanets at the same time in just a few seconds. This gap analysis is carried out as the major update in this research and can be a state of art model which can have the capability of identifying and classifying the exoplanets based on the data provided by the telescope.

IV. METHODOLOGY

Exoplanets can be found using a variety of techniques. For the most part, exoplanets discovered using transit photometry and Doppler spectroscopy have been located within the tidal locking zone. However, these methods have a clear observational bias favouring planets close to the star. There have been a number of instances where multiple planets have been found orbiting the star. [7] About 1 A "Earth-sized" planet in the habitable zone can be found orbiting one of the roughly one in five Sun-like stars. 11 billion Earth-sized planets in the Milky Way could theoretically exist if there are 200 billion stars in the Milky Way, rising to 40 billion if planets orbiting the many red dwarf stars are included in this estimate. An exoplanet is one that is not in our own solar system, but rather one that orbits another star. Several techniques, such as radial velocity, transits, direct imaging, and microlensing, can be used to locate these exoplanets.[8] A. Kepler Data Analysis The third quarter of Kepler data is used to calculate the precision of Kepler light curves for all planet-hosting stars. Averaging over tiny time windows reduces the large-scale fluctuations produced by stellar variability by calculating the scatter in the data as the average standard deviation from 10 h bins. Some of the transit depths of tiny planets currently exceed the limits of the Kepler telescope and even fall below its noise scatter.[9] A least-squares optimization tries to reduce the mean squared error (MSE) between data and a model. Since the transit parameters are unknown a priori, a simplified transit model is developed utilising a box function. Least-squares optimizers are sensitive to finding local minima (see Fig. 2) while trying to minimise the MSE and, hence, can result in inaccurate transit detections unless the global solution can be identified. [10] The data is correlated using a basic box model using a least squares optimization in the past transit detection procedures. Using the parameters that were used to generate the data, the global answer is displayed in red. Randomly initialized parameters are used in the blue line to indicate a local solution to the problem. Local minima in the parameter space of least squares algorithms can lead to false transit detections. To simulate variability, the data and model were created using a box transit function combined with a sinusoidal systematic trend. [11-12]. B.

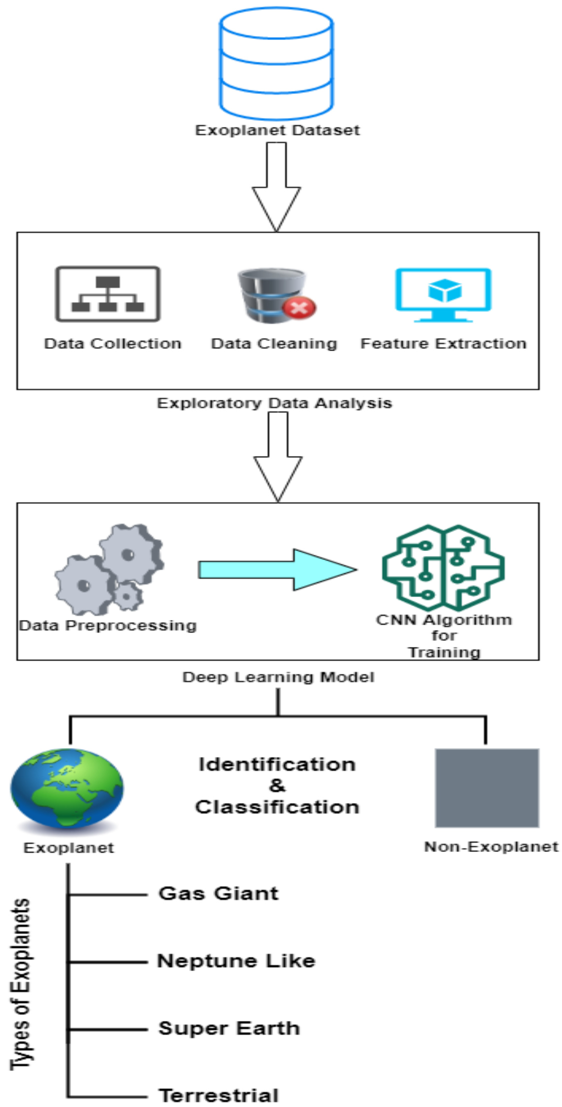

Fig 2. Proposed Model Architecture

Kepler Data Preprocessing the utilisation of time series data allows for the creation of small value offset overlapping patches from the time series. Transit must be analysed in terms of its geometric form. According to the preceding statement that a light curve's shape is a result of its physical qualities during transit and its systematic effects (parameters in our case), we are interested in seeing how these parameters can alter a light curve's shape (e.g., star spots). [19] The "Transit" method was the primary means by which the Kepler Space Telescope searched for planets. In the diagram below, a planet is shown orbiting a star. The graph shows that the starlight intensity decreases as a result of the planet partially obscuring it. Once the planets pass in front of the star, the brightness of the star returns to its original value. [20] A unique identifier called 'KOI' was assigned to each piece of Kepler data (Kepler Object of Interest). The maximum number of hash fingerprints and chunking methods that can be used to reduce the system's overhead. The pre-processed data only consist of required and necessary features, the unnecessary features have been drooped as well as during data cleaning all the garbage data is also been dropped. A total of 37

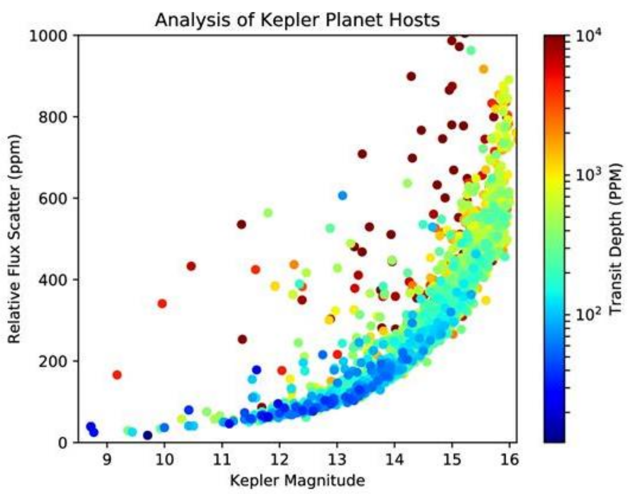

Fig 3. Analysis of Kepler planet hosts

planetary and stellar features were discovered by analysing each of the Kepler data attributes. [21] After the data pre-processing, the data consists of 37 feature columns and 8000 rows of exoplanet dataset. Since this size of data is sufficient for modelling but before below are some visualization of how the data looks like, which gives a better idea about the structure of the data.

V. DEEP LEARNING MODEL

A. Convolutional Neural Network Model:

The structure of CNN algorithm includes three layers. First is the extraction layer of features in which each neuron's input is directly connected to its previous layer's local ready fields and local features are extracted. The spatial relationship between it and other features will be shown once those local features are extracted. The other layers are feature map layer; every feature map in this layer is a plane, the weight of the neurons in one plane are same. The feature plans structure makes use of the function called sigmoid. This function known as activation function of the CNN, which makes the feature map have shift in difference. In the CNN each convolution layer come after a computing layer and it's usage is to find the local average as well as the second extract; this extraction of all the features is unique structure which increases the accuracy. It is computationally possible to mimic how the human brain solves issues in a similar fashion to an convolutional neural network. Many neuronal units are linked together, and the linkages and synapse can be strengthened or inhibited based on the current activation state. Restricted Boltzmann Machines (RBMs) are laid-out "neurons" (see Figure 5), each of which includes a collection of input parameters, each of which is connected with a weight to show the relevance of one input parameter compared to another. [18] The classification of the exoplanets based on their features is another high complexity work which can be carried out through neural network architecture with proper balancing of the hidden layers. The classification is precisely divided into 4 standard types of exoplanets: Gas giant, Neptune-Like, Super Earth, Terrestrial. The deep learning have the capability to identify and classify the exoplanets at a same time efficiently and with better accuracy.

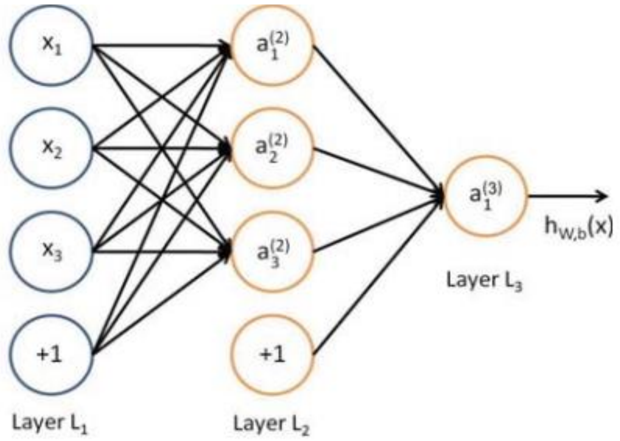

Fig 4. CNN Model Architecture

B. Time Series Model:

Detecting exoplanets using the Transit technique is possible with Artificial Intelligence. This program utilises time series data collected by telescopes during a transit. When data is collected over a period of time, we can immediately see that it is simply a time series. There is less brightness and more time passing during transportation. The visualization below of the image demonstrates the impact.

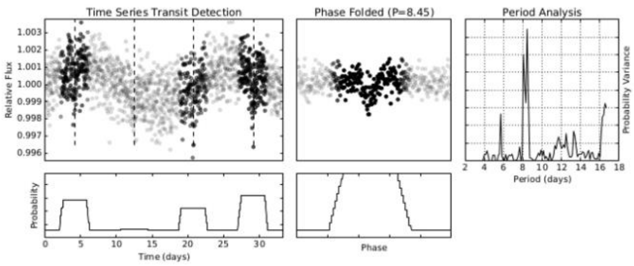

Fig 5. Time Series Data Exploration

VI. CONCLUSION

In this research a deep learning system for detecting planetary transits in Kepler Telescope light-curves were presented. In order to validate the approach, genuine light-curves were used to compare it to previous research. Astronomical activity and planet size and orbital period can influence the geometry of exoplanet transits. Thus, a basic template cannot capture the delicate features, especially if the signal is below the noise or there are significant systematic. The photo-metric properties of a transiting exoplanet can be learned using an convolutional neural network. In a handful of seconds, deep learning can analyse millions of light curves. It is possible to only qualitatively evaluate a candidate signal in the sense that neural networks are only capable of discovering transits in certain subsets of time series. We use light curves from the Kepler mission to test our deep nets and find periodic transits that are very close to the genuine period without the need for model fit. The proposed model eliminates the gap between the identification and classification of the exoplanets which consumes lot of time and man power. The model is capable of identifying the exoplanets and classifying them into different types of planets within a second of time. The proposed model is time efficient and out performs all the previous work in this area with respect to the identification and classification of the exoplanets. Machine learning techniques give an artificially intelligent platform that can learn subtle aspects from big data sets in a more efficient manner than a human can. More advanced learning techniques (such as long short-term memory and PReLU) will be investigated in the future in order to improve the detection's robustness against noise. It's currently possible to optimise the network's architecture using advance deep learning techniques, which is actively being researched.

VII. REFERENCES

[1]. Parviainen, H. PYTRANSIT: fast and easy exoplanet transit modelling in PYTHON. Monthly Notices of the Royal Astronomical Society 2015, 450, 3233–3238.

[2]. Hebrard, G., et al. "Observation of the full 12-hour-long transit of ´ the exoplanet HD 80606b-Warm-Spitzer photometry and SOPHIE spectroscopy." Astronomy Astrophysics 516 (2010): A95.

[3]. Akeson, R. L., et al. ”The NASA exoplanet archive: data and tools for exoplanet research.” Publications of the Astronomical Society of the Pacific 125.930 (2013): 989.

[4]. Morton, Timothy D., et al. ”False positive probabilities for all Kepler objects of interest: 1284 newly validated planets and 428 likely false positives.” The Astrophysical Journal 822.2 (2016): 86.

[5]. Nair, Vinod / Hinton, Geoffrey E.(2010): Rectified Linear Units Improve Restricted Boltzmann MachinesIn: Proceedings of the 27th International Conference on Machine Learning (ICML10)807-814.

[6]. Petigura, E. A. / Marcy, G. W. / Howard, A. W.(2013): A Plateau in the Planet Population below Twice the Size of Earth69.

[7]. He, K. / Zhang, X. / Ren, S. / Sun, J.(2015): Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification.

[8]. Aigrain, S. / Parviainen, H. / Roberts, S. / Reece, S. / Evans, T.(2017): Robust, open-source removal of systematic in Kepler data.

[9]. Abadi, Mart´ın u.a. (2015): Tensorflow: Large-Scale Machine Learning on Heterogeneous Systems, Software available from tensorflow.org.

[10]. Kov´acs G., Zucker S., Mazeh T., 2002, AA, 391, 369.

[11]. Jenkins J. M., Caldwell D. A., Borucki W. J., 2002, ApJ, 564, 495

[12]. Carpano S., Aigrain S., Favata F., 2003, AA, 401, 743.

[13]. Petigura E. A., Marcy G. W., Howard A. W., 2013, ApJ, 770, 69.

[14]. Krizhevsky A., Sutskever I., Hinton G. E., 2012, in Pereira F., Burges C. J. C., Bottou L., Weinberger K. Q., eds, , Advances in Neural Information Processing Systems 25. Curran Associates, Inc., pp 1097–1105.

[15]. Pearson, K. A., Palafox, L., Griffith, C. A. 2017, ArXiv e-prints, arXiv: 1706.04319.

[16]. Lecun, Y., Bengio, Y., Hinton, G. 2015, Nature, 521, 436.

[17]. Aigrain, S. / Parviainen, H. / Roberts, S. / Reece, S. / Evans, T.(2017): Robust, open-source removal of systematic in Kepler data.

[18]. Rutuja Jagtap, Uzella Inamdar, Shivani Dere, Maziya Fatima. Habitability of Exoplanets using Deep Learning MIT ADT University, Pune, India 26 May 2021.

[19]. Mohd Aijaj Khan Department of Electronics Engineering Madhav Institute of Technology and Sci Gwalior, India Discovering Exoplanets in Deep Space using Deep Learning Algorithms 19 June 2021.

[20]. Kyle A. Pearson, Leon Palafox, and Caitlin A. Griffith. Lunar and Planetary Laboratory, University of Arizona, 1629 East, USA 23 October 2017.