Research Article
Open Access
CrossMark

Exoplanets detection using lite convolutional neural networks (LCNN)

Abstract

This research comprehensively reviews the historical trajectory of exoplanet discovery. Delving into significant milestones, we highlight the pivotal role in shaping our understanding of exoplanetary systems. Building upon this historical foundation, we propose an innovative approach employing a Lite Convolutional Neural Networks (LCNN) model for exoplanet detection utilizing the Kepler Dataset. By fusing advancements in machine learning, particularly CNNs, with the rich domain of exoplanetary studies, this research represents a pivotal stride towards automated, efficient, and precise exoplanet detection. Our LCNN model demonstrates exceptional performance, achieving a training accuracy of 76.92% and an outstanding testing accuracy of 99.12%. This notable accuracy differential indicates successful model generalization and promises reliable exoplanet identification. The study not only enriches our grasp of exoplanetary history but also underscores the transformative potential of machine learning in furthering our cosmic exploration.

Keywords: exoplanet, convolutional neural networks, lite convolutional neural networks

Volume 7 Issue 4 - 2023

Al-Mahmud Al-Mamun,[1] Md Rasel Hossain,[2] Mst Mahfuza Sharmin[3]

[1]Department of Computer Science and Engineering, Islamic University, Bangladesh
[2]Department of Arts, National University, Bangladesh
[3]Department of Mathematics, National University, Bangladesh

Correspondence: Al-Mahmud Al-Mamun, Department of computer science and engineering, Islamic University, Bangladesh, Email almamun.o@gmail.com

Received: November 10, 2023 | **Published:** November 23, 2023

Introduction

Exploring exoplanets has emerged as a cornerstone of modern astrophysics in recent decades. This research embarks on a dual journey through time and technology: a meticulous review of the historical narrative of exoplanetary discovery, followed by the design and implementation of cutting-edge machine-learning techniques to detect exoplanets within the Kepler Dataset.

In Section II, we delve into the captivating history of exoplanets. We traverse the annals of astronomy, commencing with early theories and speculations about the existence of exoplanets. From the revolutionary discovery of 51 Pegasi b in 1995, a pivotal moment in this scientific saga, to the present-day robust detections and characterizations, we trace the milestones that have defined this transformative field. Understanding this historical context is essential for appreciating the incredible advancements in exoplanet detection methodologies.

In Section III, we transition to a detailed discussion of our methodology. Leveraging the knowledge from the historical review, we design and apply a Lite Convolutional Neural Network (LCNN) model to detect exoplanets using the Kepler Dataset. CNNs are powerful tools in machine learning, particularly effective in image recognition tasks due to their ability to capture intricate patterns and features.

The subsequent section, Section IV, is dedicated to the meticulous analysis of results obtained through the LCNN model. We present our findings regarding training and testing accuracy, shedding light on the efficacy and potential of this machine-learning approach in exoplanet detection.

This research aims to contribute by incorporating the insights gleaned from modern technological advancements to the ever-expanding knowledge of exoplanetary systems and further integrating machine learning in unraveling the mysteries of the cosmos.

Review of exoplanets

At night, when we look at the sky, the stars make us think about an interesting question - "Are we alone in the universe?" The discovery of the other planets in the Solar System, the Milky Way galaxy, and other galaxies in the universe make this question deepened and more profound. Astronomers have suspected for a long time that other stars in the universe have orbiting planets like the Earth. The planetary object (like the Earth) that usually orbits the star (like the Sun) outside our solar system is the Exoplanet or Extrasolar planet. The term 'Exo' from Greek ἔξω means external or outside. Astronomers have declared more than 4000; at least 1,000 or more are awaiting confirmation. The disclosure of exoplanets has intensified the possibility of extraterrestrial life.

Short history review

For thousands of years, scientists and philosophers suspected that exoplanets existed, but there was no way to prove it. In the 1800s, several detection claims were rejected by astronomers. The first evidence was noted in 1917 that a possible exoplanet orbited the Van Maanen 2 but needed to be recognized. Van Maanen 2 is the third white dwarf (after Sirius B and Procyon B) discovered by Adriaan van Maanen,[1] and it is the first example of solitary 14.1 light-years away.

In 1988, the first extrasolar planet called 'γ Cephei Ab' (in short Tadmor) was detected by a Canadian astronomer's team led by Bruce Campbell, Gordon Walker, and Stephenson Yang. Anthony Lawton and P. Wright announced that γ Cephei Ab existed in 1989,[2] but it has not been confirmed. In 2002, Artie P. Hatzes et al.,[3] confirmed it exists and is approximately 45 light-years away.[3] In 1990, 'PSR B1257+12' was detected by Aleksander Wolszczan.[4] It is orbiting the host star Pulsar PSR B1257+12 and is 2,300 light-years away. It was the first confirmed detection of exoplanets in 1992.

In 1995, 51 Pegasi b[5] became the first Extrasolar planet orbiting 51 Pegasi, approximately 50 light-years away in the constellation of Pegasus. 51 Pegasi is a main-sequence star like the Sun, and 51 Pegasi b is a hot Jupiter exoplanet with an orbiting time of 4.2 days. Hot Jupiters are gas-giant class exoplanets that are physically similar to Jupiter but have short orbital periods ($P < 10$ days). Those have the close presence of stars and high temperatures of the surface atmosphere. The hot Jupiters are easy to detect using the radial-velocity method.

47 Ursae Majoris b was the first Jupiter-like planet with a long orbiting period discovered in 1996.[6] It is orbiting at 2.11 AU from the star 47 Ursae Majoris with 0.049 eccentricities. 47 Ursae Majoris b has a second companion orbiting at 3.39 AU with 0.220 ± 0.028 eccentricity. Upsilon Andromedae was the first multiple-planetary system that contained three Jupiter-like planets. The planets' masses are 0.687, 1.97, and 3.93 MJ, and were announced in 1996, 1999, and 1999 respectively.

In 2006, HD 69830 was the first triple planetary system around a Sun-like star without any Jupiter-like planets.[7] The outermost planet masses, 18 times that of Earth, appear to be in the habitable zone. GJ 357 d was discovered in 2019; it is a planet considered "Super-Earth" in the circumstellar habitable zone.[8] Kepler-1649 c was found in 2020 and is an Earth-size exoplanet in the habitable zone.[9] In 2020, several planets within the habitable zone, including TOI-700 d (Earth-sized), KOI-456.04 of Kepler-160, and GJ 3470 c (Figures 1,2).

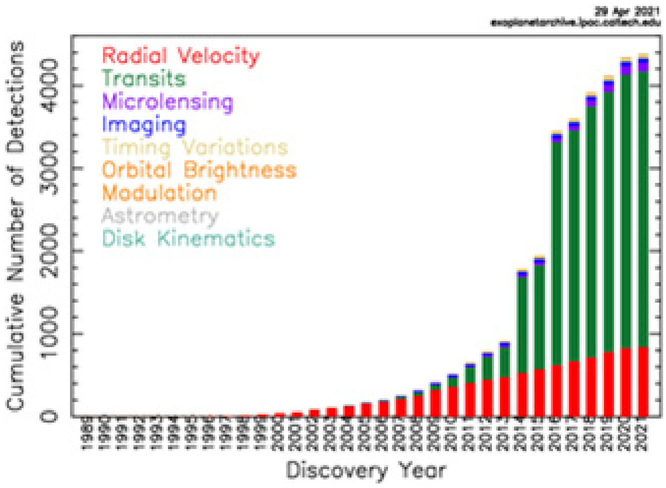

Figure 1 Cumulative exoplanet detections per years.

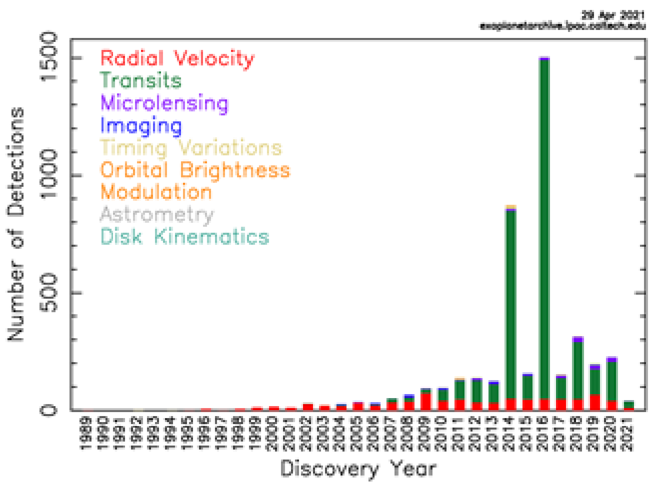

Figure 2 Exoplanet detections per years.

Methodology

The objective of this study is to utilize LCNNs to build an automated system for detecting exoplanets, with a specific emphasis on achieving a high level of accuracy. The methodology adopted for this research involves the following sequential steps (Figure 3).

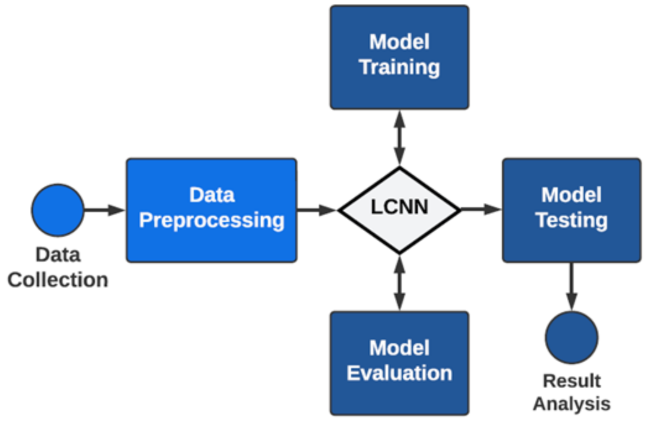

Figure 3 Workflow diagram.

Dataset details

In this research, we use **Exoplanet Hunting in Deep Space** dataset from Kaggle.[10] The dataset about the flux variations (light intensity) observed in numerous stars is essential for exoplanet detection and analysis. Each star is labeled with a binary classification: 2 signifies a confirmed presence of at least one exoplanet, including systems with multiple planets, while 1 indicates stars without confirmed exoplanets. The observations were meticulously recorded, encompassing 5087 rows and 3198 columns, with the first column denoting the label vector and subsequent columns representing flux values at distinct time intervals. The dataset comprises 37 confirmed exoplanet stars and 5050 non-exoplanet stars in the training set, five confirmed exoplanet stars, and 565 non-exoplanet stars in the test set.

Trainset

Rows/Observations:	5087
Columns/Features:	3198
Confirmed Exoplanet-Stars:	37
Non-Exoplanet-Stars:	5050

Testset

Rows/Observations:	570
Columns/Features:	3198
Confirmed Exoplanet-Stars:	5
Non-Exoplanet-Stars:	565

The dataset is derived from meticulously cleaned observations collected by the NASA Kepler space telescope, a vital tool in the ongoing mission to discover and study exoplanetary systems. The data primarily originates from Campaign 3, comprising over 99% of the dataset, with additional confirmed exoplanet-stars included from other campaigns to enrich its diversity. The meticulous preparation of the dataset took place in late-summer 2016, ensuring its accuracy and relevance for scientific analysis. Campaign 3 was selected for its perceived likelihood of containing accurately labeled exoplanets, ensuring the dataset's integrity. NASA generously open-sourced the original Kepler Mission data, which is hosted at the Mikulski Archive. Advanced de-noising algorithms were employed to eliminate artifacts generated by the telescope, resulting in a refined dataset. The data, stored in the .fits format, is readily available online, allowing enthusiasts and researchers, with the guidance of seasoned astrophysicists, to access and explore the dataset for their research endeavors (Figures 4,5).

Citation: Al-Mamun AM, Hossain MR, Sharmin MM. Exoplanets detection using lite convolutional neural networks (LCNN). *Material Sci & Eng.* 2023;7(4):191–194. DOI: 10.15406/mseij.2023.07.00223

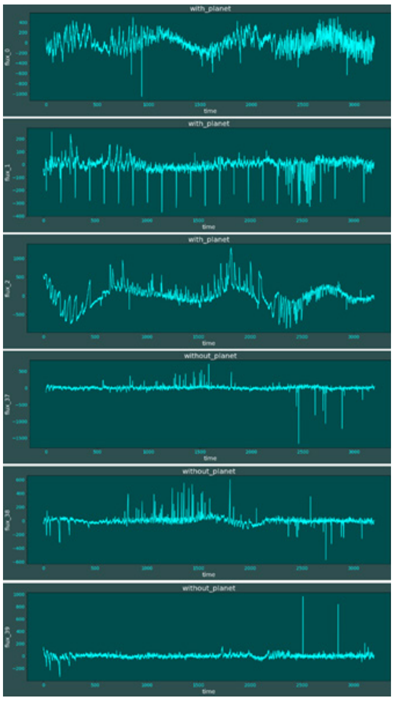

Figure 4 General FLUX values.

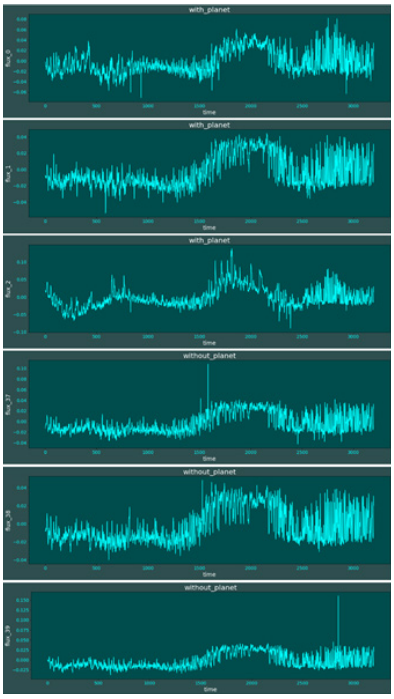

Figure 5 Scaling FLUX values.

CNN architecture

Our Convolutional Neural Network (CNN) architecture is implemented to extract meaningful features from images of Kepler Labelled Time Series items. CNNs excel in image classification tasks due to their ability to capture spatial patterns and textures. Our model's architecture comprises two Conv1D layers, two MaxPooling1D layers, a Flatten layer, three dense layers, and three Dropout layers. We employ rectified linear units (ReLU) as activation functions for the input layer and use ReLU and sigmoid for the output layer.

The convolutional layers play a vital role in feature extraction by applying a set of learnable filters to the input images. These filters detect edges, textures, and other visual features. The pooling layers reduce the dimensionality of the feature maps, summarizing the most relevant information while preserving important features. The fully connected layers leverage the extracted features for the final detection, mapping these learned features to the corresponding exoplanet hunting (see Table 1 Model - Sequential).

Table 1 Model – Sequential

Layer (type)	Output shape	Parameters
reshape (Reshape)	(None, 3197, 1)	0
conv1d (Conv1D)	(None, 3197, 128)	384
max_pooling1d (MaxPooling1D)	(None, 1598, 128)	0
dropout (Dropout)	(None, 1598, 128)	0
conv1d_1 (Conv1D)	(None, 1597, 128)	32896
max_pooling1d_1 (MaxPooling1)	(None, 798, 128)	0
dropout_1 (Dropout)	(None, 798, 128)	0
flatten (Flatten)	(None, 102144)	0
dense (Dense)	(None, 128)	13074560
dropout_2 (Dropout)	(None, 128)	0
dense_1 (Dense)	(None, 128)	16512
dense_2 (Dense)	(None, 1)	129

Total Parameters: 13,124,481

Trainable Parameters: 13,124,481

Non-trainable Parameters: 0

Training and testing

The training of the CNN model involves minimizing a suitable loss function to gauge the disparity between the predicted class probabilities and the actual labels. The model's parameters are iteratively updated by backpropagating the gradients through the network and adjusting the weights accordingly. The learning rate, determining the step size in parameter updates, is a crucial hyperparameter that requires careful tuning for optimal training performance. The training process typically spans a fixed number of epochs; in our case, we utilized 30 epochs, each representing a complete pass through the training dataset. As the epochs progress, the CNN model's performance on the training data gradually improves, converging towards a state where the accuracy stabilizes (Figure 6).

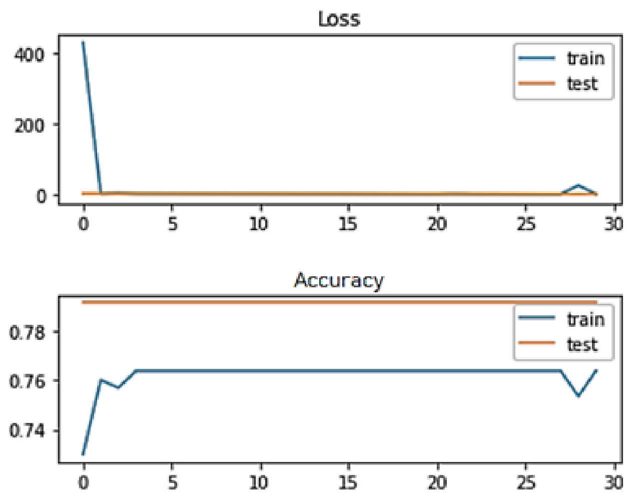

Figure 6 Train and test results.

Citation: Al-Mamun AM, Hossain MR, Sharmin MM. Exoplanets detection using lite convolutional neural networks (LCNN). *Material Sci & Eng.* 2023;7(4):191–194. DOI: 10.15406/mseij.2023.07.00223

Train: loss: 2.1925 and accuracy: 0.7692

Test: loss: 1.9257 and accuracy: 0.9912

Result analysis

Our LCNN model achieved a training accuracy of 0.7692 (76.92%) and an impressive testing accuracy of 0.9912 (99.12%). The results signify a strong ability of our model to generalize and accurately identify exoplanets (Figure 7).

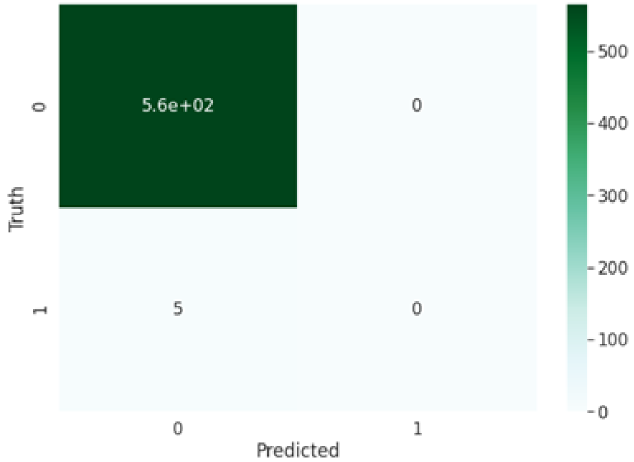

Figure 7 Prediction metrics.

Accuracy: 0.991

Precision: 0.000

Recall: 0.000

F1 Score: 0.000

ROC AUC: 0.500

Balanced

Accuracy: 0.500

Sensitivity: 0.000

Specificity: 1.000

Geometric Mean: 0.000

The 99.12% testing accuracy indicates that our model can successfully classify exoplanets with high precision. Furthermore, the 76.92% training accuracy suggests that our LCNN has effectively learned features and patterns from the training data, demonstrating its ability to capture the underlying structure of the dataset.

The accuracy gap between the training and testing suggests minimal overfitting, implying that the model generalizes well to previously unseen data. The LCNN architecture, tailored to exploit spatial relationships in the dataset, seems effective in extracting meaningful features related to exoplanet detection. Augmenting the dataset or exploring similar datasets might improve the model's robustness and broaden its applicability.

Conclusions

In this article, we present a comprehensive review of the historical trajectory of exoplanet discovery, and we propose a Lite Convolutional Neural Networks (LCNN) model for exoplanet detection utilizing the Kepler Dataset. LCNN's exceptional testing accuracy of 99.12% showcases its potential as a powerful tool for exoplanet detection. Ongoing efforts to optimize the model and validate its results will be crucial for its successful integration into real-world applications within astrophysics. After performing our experiment, our work reached its goals. And according to the result, our work is successful.

Acknowledgment

First of all, we thank the Almighty; without His mercy, it is not possible to succeed. We would prefer to specify our sincere feelings for our family, friends, and teachers for the incorporation, steerage, inspiration, and constructive suggestions that helped us during the preparation and attainment of this work.

Conflicts of interest

The author declares there are no conflicts of interest.

References

1. Van Maanen A, Two faint stars with large proper motion, publications of the astronomical society of the pacific. 1917;29(172):258–259.
2. Boss A. The crowded universe: the search for living planets, Basic Books; 2009.
3. Artie P. Hatzes, William D. Cochran, et al. A planetary companion to γ Cephei A. *The Astrophysical Journal*. 2003;599(2).
4. Wolszczan A, Frail DA, A planetary system around the millisecond pulsar PSR1257 + 12, *Nature.* 1992;355(6356):145–147.
5. Mayor M. The HARPS search for southern extra-solar planets: XVIII. An Earth-mass planet in the GJ 581 planetary system. *Astronomy and Astrophysics*. 2009;507(1):487–494.
6. Gregory, Philip C, Fischer, et al. A Bayesian periodogram finds evidence for three planets in 47 Ursae Majoris. *Monthly Notices of the Royal Astronomical Society.* 2010;403(2):731–747.
7. Beichman CA, et al. An excess due to small grains around the nearby k0 v star HD 69830: Asteroid or Cometary Debris?. *The Astrophysical Journal*. 2005;626(2):1061–1069.
8. Luque R, Pallé E. Planetary system around the nearby M dwarf GJ 357 including a transiting, hot, Earth-sized planet optimal for atmospheric characterization. *Astronomy & Astrophysics*. 2019;628:A39.
9. Vanderburg Andrew. A habitable-zone earth-sized planet rescued from false positive status. *The Astrophysical Journal*. 2020;893(1):L27.
10. Exoplanet hunting in deep space.