Identification of Exoplanets using Machine Learning

Deepanshu Tyagi[1*], Renu Chaudhary[2], Gurpreet Kaur[3], Rekha Kumari[4]

[1*]Scholar, Department of Computer Science , HMR Institute Of Technology And Management, 110036, Delhi, India

[2] Faculty, Department of Information Technology, HMR Institute Of Technology And Management, 110036, Delhi, India

[3,4] Faculty, Department of Computer Science & Engineering, HMR Institute Of Technology And Management, 110036, Delhi, India

* Corresponding author

doi: https://doi.org/10.21467/proceedings.178.20

ABSTRACT

Over the past decade more than a million stars have been observed as part of efforts to identify transiting planets. Identifying potential exoplanet candidates requires manual interpretation which is both labor intensive and susceptible to human error the consequences of which are often difficult to quantify. This paper describes a novel approach that utilizes a neural network to automate the identification of potential exoplanet candidates in large planetary search programs a method that contrasts with traditional manual techniques. A combination of classification models and diverse datasets are employed to assess the probability that an object is an exoplanet. The first step of the process involves cleaning raw astronomical images obtained from reliable sources removing extraneous objects and disturbances introduced by noise and electromagnetic radiation. Once the data is cleaned categorization models are applied to analyze and classify the objects. This approach significantly reduces the need for manual intervention and enhances the accuracy and efficiency of exoplanet detection. By automating the detection process it enables more extensive planetary search programs to operate more effectively thus contributing to the advancement of exoplanet research. This paper offers a more reliable and scalable solution to the identification of exoplanet candidates.

Keywords- Exoplanet, Detection, HCI, Post Processing, Algorithms.

1 INTRODUCTION

Astronomy is one of the oldest natural sciences of human civilization Throughout history astronomy has influenced religion guided explorers determined food production schedules and fuelled philosophical questions about our existence and role in the universe Our curiosity about the stars is a natural extension of wondering if there is another planet in another solar system capable of supporting life The mission to detect planets outside our solar system known as exoplanets is leading to truly new discoveries The discovery and examination of exoplanets represent a significant domain within astronomy Exoplanets have become a key focus of research especially since the first exoplanet orbiting a Sun-like star was discovered in 1995 Since then the study of exoplanets has expanded rapidly and become one of the fastest growing disciplines in astronomy Detecting exoplanets has traditionally been a time consuming task requiring highly experienced experts with access to specialized equipment These experts relied on their training intelligence diligence and team expertise to search for exoplanets using images collected by ground based observatories and satellite based telescopes such as Hubble A new era has begun in the hunt for exoplanets In recent years a new generation of modern satellites such as Kepler has been launched to optimize and improve scientific observations and data production related to exoplanet detection These satellites capture images and process them using proven astronomical techniques and methods to produce a wide range of high

Proceedings DOI: 10.21467/proceedings.178; Series: AIJR Proceedings; ISSN: 2582-3922; ISBN: 978-81-984081-8-1

quality data Astronomers and astrophysicists can query this data to determine if the object of interest is indeed an exoplanet While detecting exoplanets sounds simple it is not Exoplanets are difficult to see directly because they are often obscured by the bright light of their host stars Therefore astronomers must rely on various labor intensive methods to detect and monitor exoplanets Most of these methods are time consuming and require substantial human effort The methodologies explored in this study aim to enhance the precision of exoplanet detection by integrating multiple algorithms to refine forecasting models This approach will benefit researchers looking to improve their investigations and further the understanding of exoplanets The study of exoplanets is significant not only for the advancement of astronomy but also for addressing fundamental questions about the existence of life beyond our solar system The exploration of exoplanets ranging from Earth-like entities to more exotic planets provides crucial insights into planetary formation and evolution as well as the potential for habitability beyond Earth These insights are crucial as the limited sample size of our solar system's eight planets underscores the need for broader exploration of exoplanetary systems The discovery of exoplanets and their characteristics such as Jupiter-sized planets in close orbits to their host stars supports theories of dynamic planetary movements during formation Ultimately the study of exoplanets compels us to confront an existential question Are we the sole inhabitants of the cosmos Although current technological limitations prevent direct searches for extraterrestrial life the identification of potentially habitable exoplanets narrows the quest and offers tantalizing possibilities for answering this age-old question

2 CONCEPTS IN USE

The study uses a number of cutting-edge ideas to provide accurate and effective imaging and analysis. In order to identify weak objects close to brilliant sources, such exoplanets or stellar partners, High-Contrast Imaging (HCI) suppresses the overpowering brightness of the primary source. This procedure is further aided with a coronagraph, which blocks starlight to allow for the observation of surrounding dim structures. Real-time correction of atmospheric distortions using adaptive optics (AO) guarantees crisp, detailed pictures from ground-based telescopes. Image enhancement is aided by the Point Spread Function (PSF), which models the distribution of light from point sources in an optical system. The common astronomical file format, FITS, is used to store and interpret data, while Angular Differential Imaging (ADI) improves detection accuracy by employing rotational subtraction techniques to isolate dim objects from noise and distortions.

2.1 HIGH CONTRAST IMAGING

Detecting an Earthlike planet orbiting a Sun-like star using imaging is currently not feasible due to technological limitations. The light reflected by Earth-like planets is approximately ten billion times fainter than the light emitted by Sun-like stars, and existing instruments cannot detect such weak signals near such bright sources. This limitation also applies to giant planets like Jupiter, which reflect about 100 times more light than Earth-like planets. Similarly, exoplanets orbiting large and bright stars outside our galaxy remain hidden due to the intense luminosity of their host stars. The closer an exoplanet is to its host star, the more challenging it becomes to detect.Overcoming this challenge requires the development of imaging systems capable of detecting faint objects near bright point sources such as stars.

Characterizing exoplanets through spectroscopy must overcome the high contrast between the planet and its star. This challenge is particularly difficult for Earth-like planets where the visible contrast is approximately ten to the power of minus ten and the angular difference is zero point one arcseconds in a ten parsec system. Several high contrast imaging techniques have been developed to block excess light from stars around which an exoplanet orbits. In addition to suppressing starlight these techniques help reduce unnecessary noise in the image. These imaging methods are applied both at the instrument level and during the processing of raw images using various post processing algorithms. The use of telescopes in space has opened new frontiers in high contrast imaging due to their stability and location above the Earths atmosphere which eliminates significant image noise

2.2 CORONAGRAPH

Telescopes have the ability to magnify stars easily. However when astronomers aim to observe faint objects like planets near stars they face the challenge of the stars overwhelming brightness. To overcome this obstacle they use a specialized instrument called a coronagraph. A coronagraph is a device mounted on a telescope designed to block direct starlight allowing observation of nearby objects that would otherwise be hidden by the stars intense glare. While traditional coronagraphs focus on studying the Suns corona a new type of instrument known as a stellar coronagraph has been developed. These instruments detect extrasolar planets interstellar disks around nearby stars and phenomena like quasars with active galactic nuclei. The concept behind a coronagraph is inspired by natural events like solar eclipses when the moons shadow covers the brighter layers of the Sun revealing its fainter corona. Mimicking this phenomenon a coronagraph uses a circular mask within the telescope to block most sunlight. Advancements in coronagraph technology now allow scientists to directly observe light emitted by distant exoplanets.

2.3 HOW IT WORKS?

The coronagraph operates by capturing uniformly illuminated light through the telescopes aperture where a central secondary mirror is located. This light passes through a lens instead of a traditional camera or detector and encounters a focal plane mask. The mask absorbs much of the central light redirecting the remaining light to the edges of the telescopes pupil often forming visible rings. Next the Lyot stop another important component blocks these light rings from the central star while allowing light from surrounding sources to pass through for final imaging by the lens. This technique improves image contrast making it particularly useful for studying exoplanets and debris disks around stars. However achieving optimal coronagraphic imaging presents challenges. Atmospheric turbulence affects light entering the telescopes pupil even in space based observatories like the Hubble Space Telescope. Additionally optical imperfections or fluctuations caused by temperature changes over time disrupt the light entering the coronagraph reducing its ability to remove starlight effectively. Therefore successful imaging of exoplanets and debris disks with a coronagraph requires adaptive optics whether the telescope is in space or on the ground

2.4 ADAPTIVE OPTICS

In optical systems challenges such as component displacements imperfections in elements or misalignments can cause internal performance degradation while external factors like atmospheric conditions can also impact performance. Careful optical system design can reduce these issues to some extent but adaptive optics provides a method to improve the efficiency of optical systems by correcting incoming wavefront distortions. This is achieved by adjusting mirrors to counteract the distortions. Adaptive optics is used in various fields including astronomical telescope to counter atmospheric distortions. It is also applied in microscopyoptical manufacturing and retinal imaging systems to minimize optical aberrations. The core principleof adaptive optics involves measuring wavefront distortions and using correction devices such as liquid crystal displays to compensate for these errors. Figure one below describes thestructure of an adaptive optics.

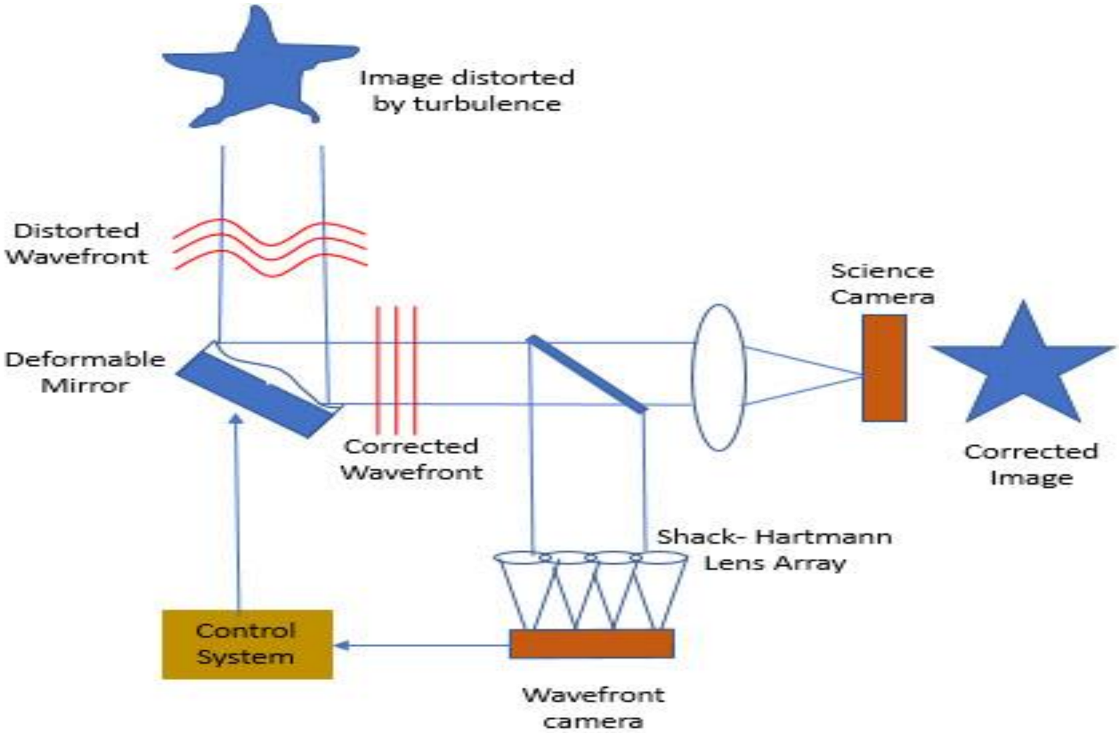

Figure 1: Adaptive Optics System

2.5 POINT SPREAD FUNCTION

The ideal point spread function characterizes the three dimensional diffraction pattern of light originating from an extremely small point source within the sample. This light is directed to the image plane through a high numerical aperture objective. The point spread function serves as a fundamental component When light emanates from such a point the lens captures a fraction of it and directs it to a corresponding point in the image plane. Instead of converging precisely to a single point the light waves intersect at the focal point and interfere with each other generating a diffraction pattern consisting of concentric rings of light surrounding a central bright disk in the xy plane. The diameter of this disk is determined by the numerical aperture providing a way to estimate the resolution of the lens by examining the Airy disk named after George Biddell Airy. The central bright region of the Airy disk and the surrounding concentric rings correspond to the peaks of intensity in the distribution.

Proceedings DOI: 10.21467/proceedings.178
ISBN: 978-81-984081-8-1

Series: AIJR Proceedings
ISSN: 2582-3922

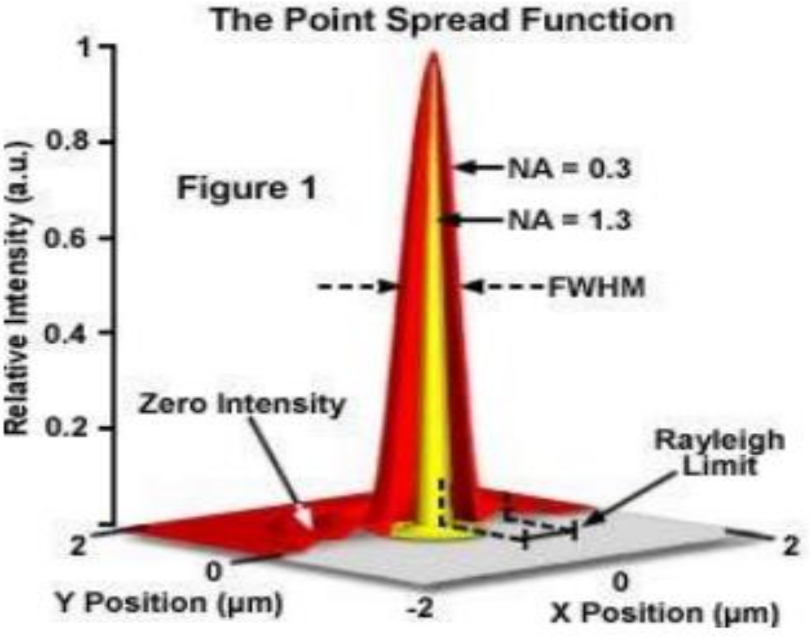

Figure 2: Point Spread Function

2.6 ANGULAR DIFFERENTIAL IMAGING

Angular differential imaging is a technique used in high contrast imaging to reduce quasi static speckle noise and improve the detection of nearby companions. In angular differential imaging a sequence of images is captured using an elevation azimuth telescope with the field derotator disabled. This setup allows the instrument and telescope optics to stay aligned while enabling the field of view to rotate relative to the instrument. Each image in the sequence is matched with a reference point spread function generated from other appropriately selected images within the same sequence. Subtracting this reference point spread function from each image helps remove quasi stationary structures. The remaining images are then rotated and combined to flatten them. Angular differential imaging acts as a calibration method for point spread functions notably reducing quasi static structures. The process involves capturing image sequences with an elevation azimuth telescope typically with the instrument rotator either off or adjusted to align with the telescope optics. This alignment ensures stability in the quasi static structure throughout the sequence by causing a gradual rotation of the field of view relative to the instrument. While the field of view rotates over time the point spread function remains fixed. As the field of view rotates during exposure the attached structures spread azimuthally affecting the intensity of companion peaks. This effect can be minimized by using short exposures and optimizing aperture photometry. Following data reduction and image registration the reference point spread function is subtracted from each image to remove quasi static structures. This subtraction given sufficient field of view rotation preserves the signal from potential companions. Finally all image differences are aligned and combined using the median.

The following points describe the importance of angular differential imaging
one It can help determine whether all solar systems behave similarly by observing exoplanets. Based on current observations the majority of star systems do not resemble our solar system
two Discovering exoplanets benefits humanity by expanding the search area to include a large region for inhabited worlds. This also increases the possibility that we are not alone
three It helps address unanswered questions about the cosmos that lead us to question our own existence
four In addition to discovering exoplanets the search for new planets and asteroids has led to a broader taxonomy of celestial entities expanding and organizing the domain of study

2.7 FITS FILE

The flexible image transport system is an open standard defining a digital file format valuable for the storage transmission and manipulation of data. It accommodates multidimensional arrays such as two dimensional images or tables. Primarily used in astronomy the flexible image transport system stands as the most common digital file format in the field. Designed specifically for astronomical data the flexible image transport system includes features such as photometric and spatial calibration data along with metadata about image provenance.

3 PYTHON LIBRARIES USED

The code for this study was developed and implemented using the Python programming language, which was chosen for its huge library and adaptability. To improve functionality and streamline intricate procedures, a number of Python modules were used. Pandas and NumPy, two data manipulation tools included in these packages, made managing and processing data more efficient. Data and outcomes were visualized in an understandable and significant way using Matplotlib and Seaborn. Models were designed, trained, and evaluated using frameworks like TensorFlow and PyTorch for deep learning and object identification tasks. Additionally, Scikit-learn enabled a variety of machine learning applications, while OpenCV was used for image processing. Python was the perfect choice for this study because of the way these libraries combined to expedite development.

3.1 MATPLOTLIB

Matplotlib serves as a graphics library for Python, along with its numerical mathematics extension, NumPy. It offers an object-oriented application programming interface (API) for integrating graphics into various applications, utilizing common GUI tools like Tkinter, wxPython, Qt, or GTK+. Additionally, it provides a procedural interface called "pylab", akin to MATLAB, although its usage is discouraged. Matplotlib is commonly utilized within SciPy.

3.2 NUMPY

NumPy is a Python library designed to enhance support for handling large, multidimensional arrays andmatrices, alongside a wide array of advanced mathematical functions tailored for manipulating such arrays. Specifically targeting the CPython reference implementation of Python, which is a bytecode interpreter lacking optimization, NumPy aims to bridge the gap between mathematical algorithms written in Python and their compiled counterparts. However, algorithms developed for this Python version typically exhibit slower execution speeds compared to their compiled equivalents.

3.3 ASTROPY

Astropy represents a compilation of software packages developed in Python, tailored specifically for astronomical applications. This collection includes a core package of astronomical utilities, provided free of charge. Astropy serves as a pivotal tool as Python gains traction among astronomers, fostering seamless compatibility among different Python-based astronomy packages. Notably, Astropy is incorporated into various prominent Python distributions, including Linux, macOS, Anaconda Python Distribution, Enthought Canopy, and the Ureka package managers.

3.4 SCIPY

SciPy stands as an open source Python library essential for scientific and engineering computing endeavors. Offering a wide array of modules, SciPy encompasses functionalities for optimization, linear algebra, integration, interpolation, special functions, FFT, signal and image processing, ODE solvers, and various other tasks prevalent in science and engineering domains. Built upon NumPy's array object, SciPy is an integral component of the NumPy stack, alongside tools like Matplotlib, pandas, and SymPy. This cohesive stack caters to users familiar with applications such as MATLAB, GNU Octave, and Scilab, thus earning it the moniker "SciPy stack."

3.5 PANDAS

Pandas is a Python library dedicated to data processing and analysis tasks. It offers a range of data structures and functionalities specifically tailored for handling numerical tables and time series data. Released under the three-clause BSD license, Pandas enables users to manipulate datasets freely. Its name originates from "panel data," an econometric concept referring to datasets containing observations of the same entities across various time intervals.

3.6 SCIKIT – LEARN

Scikit-learn, previously known as scikits.learn and also referred to as sklearn, is a freely available machine learning library tailored for the Python programming language. It encompasses a diverse range of algorithms for classification, regression, and clustering tasks, incorporating methods such as support vector machines, random forests, gradient boosting, k-means, and DBSCAN. Designed to seamlessly integrate with Python's numerical and scientific libraries NumPy and SciPy, Scikit-learn primarily utilizes Python for implementation, leveraging NumPy for efficient linear algebra and array operations. While some fundamental algorithms are coded in Python to enhance performance, others, like support vector machines and logistic regression, utilize Cython wrappers around LIBSVM and LIBLINEAR for computational efficiency. Despite this, Scikit-learn maintains compatibility with various other Python libraries including matplotlib, plotly, numpy, pandas, and scipy. On the other hand, scikit-image, previously known as scikits.image, emerges as an open-source image processing library crafted for Python. It offers an extensive array of algorithms for tasks like segmentation, geometric transformations, color space manipulation, filtering, feature detection, and more. Just like Scikit-learn, scikit-image is designed to synergize seamlessly with NumPy and SciPy, utilizing Python predominantly for its implementation, with select core algorithms optimized for efficiency.

3.7 PHOTUTILS

Photutils serves as a supplementary package to Astropy, primarily offering functionalities for detecting and conducting photometry on astronomical sources. As an open-source Python package, it operates under the 3-clause BSD license.

3.8 SYMPY

SymPy stands as an open-source Python library dedicated to symbolic computing, offering computer algebra capabilities. It can be utilized either as a standalone application, integrated within other applications, or accessed in real-time online through platforms like SymPy Live or SymPy Gamma. One of its notable advantages is its straightforward installation and maintenance process, as it is entirely written in Python with minimal dependencies. This simplicity, combined with a clear and expandable codebase within a familiar programming language, renders SymPy accessible to users with varying levels of expertise. Covering a wide range of mathematical domains, from basic arithmetic to calculus, algebra, discrete mathematics, and even quantum physics, SymPy also provides the capability to format calculation results into LaTeX code. Licensed under the new BSD license, SymPy is freely available, with its library divided into a core and numerous optional modules. As of now, the SymPy core comprises approximately 260,000 lines of code.

3.9 TENSORFLOW

TensorFlow represents a freely available open-source software library designed for machine learning applications. While versatile in its utility, TensorFlow places particular emphasis on tasks related to training and executing deep neural networks. Functioning as a symbolic math library grounded in data flow and differentiable programming, it finds extensive use within Google, serving both research and production purposes. Initially developed by the Google Brain team for internal use, TensorFlow was subsequently released under the Apache License 2.0 in 2015.

4 POST PROCESSING ALGORITHMS

DirectThe direct imaging of extrasolar planets has emerged as a significant focus within contemporary astronomy and biology. This endeavor necessitates the concurrent utilization of an ultra-high-performance Adaptive Optics (AO) system alongside a differential imaging camera to effectively eliminate stellar flux. Moreover, sophisticated post-processing techniques are indispensable for achieving the requisite level of detection sensitivity.A pivotal aspect of exoplanet imaging involves the application of post-processing to discern planets amidst noisy data, resulting in substantial sensitivity enhancements compared to previous methodologies. Advances in data analysis techniques throughout the decade have facilitated this process, particularly for ground-based observations, and have become a prerequisite for space-based observatories like WFIRST, HabEx, and LUVOIR. Modern post-processing methodologies predominantly rely on feature reduction techniques such as principal component analysis or sparse decomposition to effectively remove scattered starlight from the focal plane, thereby revealing underlying planetary signals. However, challenges persist, including the risk of over-subtraction of planetary signals, the necessity to fine-tune numerous parameters, and the difficulty in establishing robust confidence limits for detection validity.

Furthermore, disagreements among independent experts regarding the validity of dataset identification may arise, even when employing the same post-processing algorithm.

Following is the basic formula working behind the post processing algorithms: -

Ai = image at instance i

where, i = 1, 2, 3,---n

B = point Spread Function (PSF)

Ci = image without noise

= Ai - B

Di = de-rotated Ci images

= de-rotation (Ci)

E = result of superimposition

= median (Di)

Here we discuss the post processing techniques that we have used in our project to help us detect an exoplanet:

4.1 PCA (PRINCIPAL COMPONENT ANALYSIS)

As datasets grow in size, interpreting them becomes increasingly challenging. Principal component analysis (PCA) offers a solution by reducing the dimensionality of these datasets, enhancing interpretability while minimizing loss of information. PCA achieves this by generating new, uncorrelated variables—termed principal components—that sequentially maximize variance. The process involves solving the eigenvalue/eigenvector problem, with the resulting principal components being defined by the dataset itself rather than predetermined, rendering PCA an adaptive data analysis technique. This reduction in computational complexity accelerates the performance of machine learning algorithms..There are three types of PCA:

1. Simple PCA.
2. Full Frame PCA.
3. Annular PCA.

Principal Component Analysis or PCA is a widely used technique for dimensionality reduction of the large data set as displayed in Figure 3.

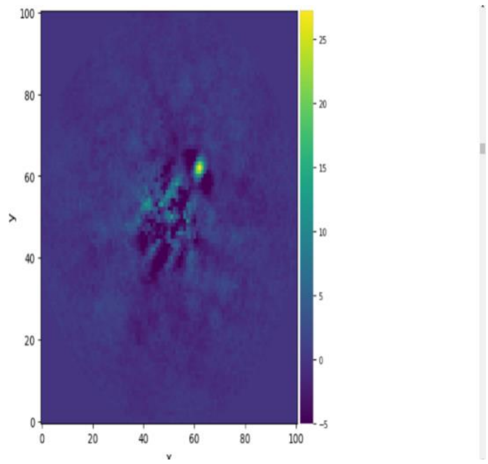

Figure 3: Output of Full Frame PCA

We can exploit the range of rotation by using annular PCA approximations and applying a PA threshold for different annuli as displayed in Figure 4.

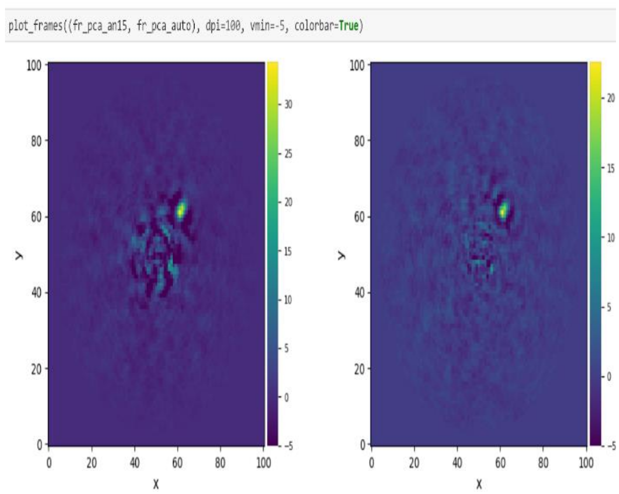

Figure 4: Output of Annular PCA

4.2 LEAST SQUARE APPROXIMATION

The least-square approximation method offers a way to find an approximate solution for a system of linear equations when an exact solution is unavailable. In typical situations where there are many more constraints (n) than variables (d), existing exact methods usually take a time complexity of O(nd^2) to find the solution vector. Here, we present two randomized algorithms that can provide accurate relative error estimates for both the optimal value and the solution vector of the least-squares approximation problem, and they do so more efficiently than existing exact algorithms. Both algorithms involve preprocessing the data using a random Hadamard transform[7]. Then, one algorithm randomly selects constraints and solves the reduced problem using those constraints uniformly, while the other uses a sparse random projection and solves the reduced problem using those projected coordinates. In both cases, solving the reduced problem yields a relative error estimate. If n is significantly larger than d, the approximate solution can be computed in O(ndlnd) time. Figure 5 displays the output of LSA algorithm. .

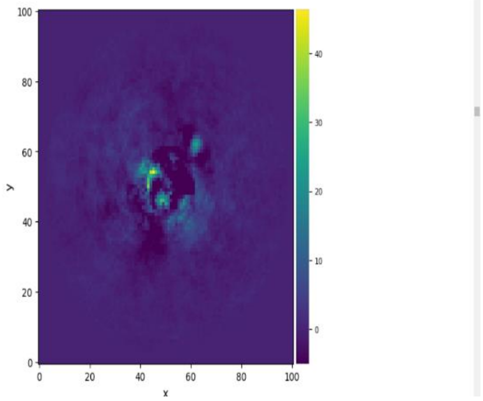

Figure 5: Output of Least Square Approximation Algorithm

4.3 PAIRWISE FRAME DIFFERENCING

Frame discrimination is a technique where a computer checks the difference between two video frames. If the pixels have changed, something in the image has probably changed (like moved). Most techniques work with some degree of blurring and thresholding, two distinct motions of noise. Because the shot can also vary when the lighting conditions in the room (and the camera's autofocus, brightness correction, etc), change as displayed in figure 6.

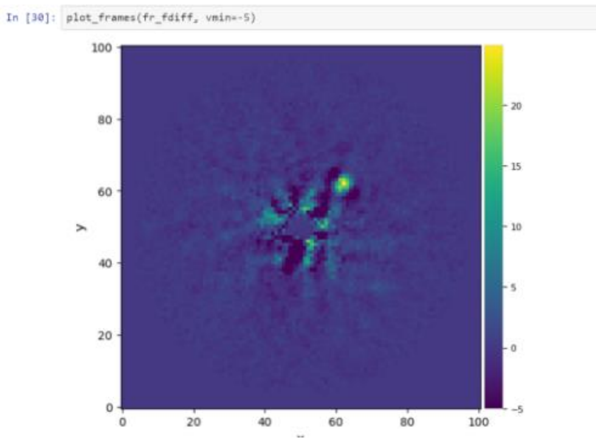

Figure 6: Output of Pairwise Frame Differencing Algorithm

4.4 LLSG

Data Data processing is integral to high-contrast exoplanet imaging, akin in importance to selecting the appropriate coronagraph or wavefront guidance system, and intricately linked with the chosen observation strategy. A cutting-edge technique in angular[8] differential imaging (ADI) data processing revolves around principal component analysis (PCA) algorithms. PCA serves as a method for subspace projection to create a reference point spread function (PSF), which, when subtracted from scientific data, enhances the detectability of potential companions. However, PCA encounters challenges in constructing this reference PSF from scientific data, particularly due to its susceptibility to non-Gaussian noise in lower-dimensional orthogonal subspaces. Drawing inspiration from recent advancements in machine learning, such as robust

Proceedings DOI: 10.21467/proceedings.178
ISBN: 978-81-984081-8-1

Series: AIJR Proceedings
ISSN: 2582-3922

PCA, our objective is to introduce a localized subspace projection technique that outperforms current PCA-based post-processing algorithms in near real-time companion detection. This advancement holds significant promise for future exoplanetary studies. Leveraging recent progress in machine learning, we utilize random low-rank approximation methods combined with trigger thresholding to locally decompose the ADI image sequence into low-rank, sparse, and Gaussian (LLSG)[9] noise components. This localized triple decomposition effectively separates starlight and associated speckle noise from the planetary signal, which predominantly remains sparse. Figure 7 displays the output of LLSG algorithm.

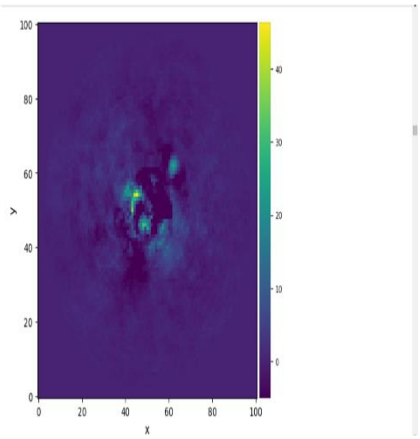

Figure 7: Output of LLSG Algorithm

4.5 ADI MEDIAN SUBTRACTION

There are two ways to implement this algorithm. The first method calculates the median of all frames and subtracts it from each individual frame. Median averages the statistical variations of each frame, while subtraction removes the static speckle pattern[10]. Another, slightly more complex method builds a PSF model for each frame i of the data cube with the following procedure: Frame i is divided into rings. For each ring, two next and two previous frames are selected that match the rotation threshold. – The 4 selected frames are combined with the median to obtain the PSF model, which is then subtracted from frame i. By limiting the number of frames in the PSF structure of the model, the temporal variation of the data cube is better accounted for. Figure 8 displays the output of LLSG algorithm.

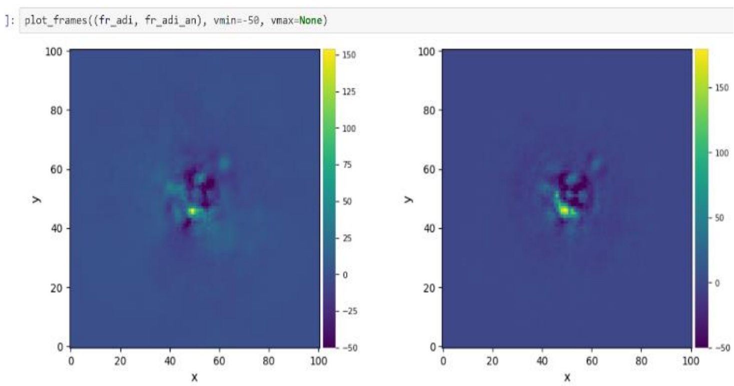

Figure 8: Output of ADI Mean Subtraction Algorithm

5 SNR MAP

Signal To Noise Ratio Map is generated from the data cube. Figure 9 dispays the formula of the SNR: -

$$S/N \equiv \frac{\overline{x}_1 - \overline{x}_2}{s_2\sqrt{1+\frac{1}{n_2}}},$$

Figure 9: Formula of SNR

The main idea is to test a given speckle against the background resolution elements (at the same angular separation or radial distance from the center). Figure 10 displays the SNR Map.

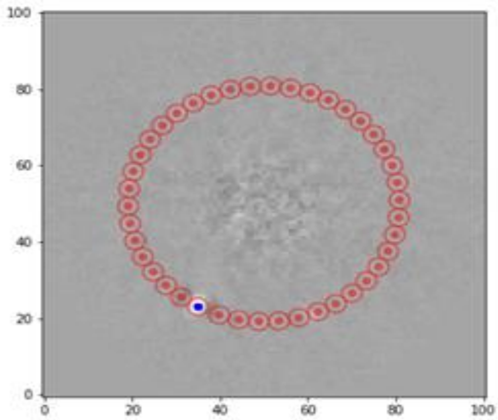

Figure 10: SNR Map

6 DETECTION

Detection is done by plotting a combination of PP algorithm and SNR map. Finding the exact coordinates of the companion is an important part of the detection.The output of the post-processing algorithm is used together with the SNR map to find the coordinates of the exoplanet's compatible companion (if one exists) as in Figure 11. Those coordinates are then used to create a surface map as displayed in Figure 12, the result of which helps determine the existence of an exoplanet.

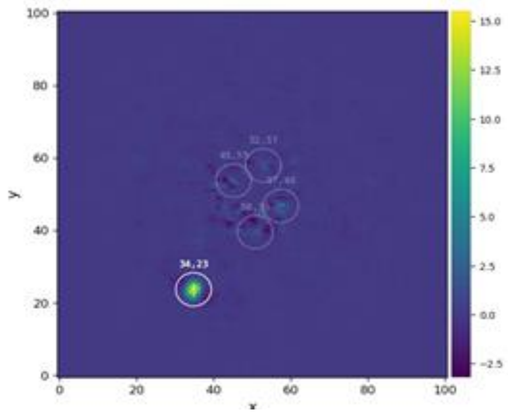

Figure 11: Detected companion and its coordinates

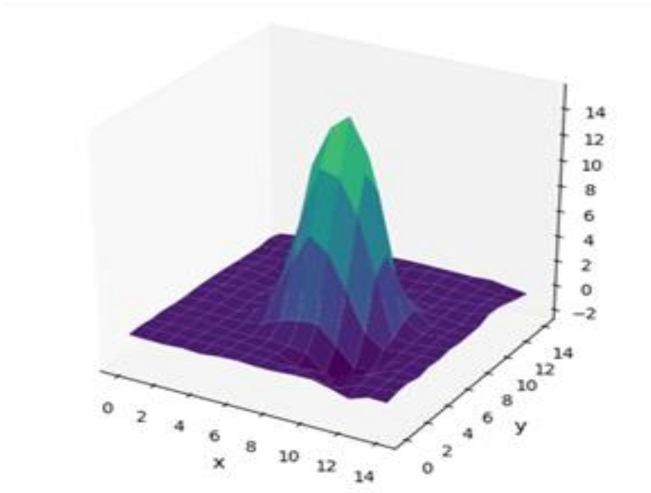

Figure 12: Surface map of detected companion

7 RESULT AND DISCUSSION

In conclusion the work conducted in this study holds significant importance for the field of astronomy Interest in the exploration of exoplanets continues to grow with time The development of advanced technologies and techniques enables more accurate and efficient predictions regarding exoplanets The future of humanity extends beyond Earth and for that it is essential to focus on exploring the universe and identifying new planets where humans could potentially settle in the future Exoplanets play a crucial role in this search A home is one of humanity's fundamental needs and this study contributes to addressing this need by helping to identify potential new homes in the vastness of space The integration of multiple machine learning models within this framework represents an effective approach to advancing the discovery of exoplanets By combining various methodologies this approach not only optimizes the efficiency of identifying and localizing target objects but also improves the accuracy of predictions The framework's ability to quickly adapt to evolving data ensures that it remains a valuable tool in exoplanet research consistently generating promising candidates for further exploration This iterative process accelerates the pace of discovery and establishes a strong foundation for understanding planetary systems beyond our own The study's findings provide significant insights that can be built upon in future research to uncover more about the universe and the potential for habitable planets

8 CONCLUSION

In conclusion the utilization of multiple machine learning models within the framework presents a powerful strategy for advancing exoplanet discovery By combining various methodologies this approach not only optimizes the efficiency of target identification and localization but also increases the reliability of predictions The ability to quickly adapt to evolving data landscapes ensures that the framework remains at the forefront of exoplanetary research continually generating promising candidates for further investigation This iterative process accelerates the pace of discovery and lays a solid foundation for unraveling the mysteries of planetary systems beyond our own Additionally the automation of the exoplanet detection process significantly reduces the need for manual intervention and human error This enhances the accuracy and efficiency of planetary search programs thereby contributing to the advancement of exoplanet research As the study of exoplanets expands the approach developed in this paper offers a more reliable and scalable solution for identifying exoplanet candidates and strengthens the ongoing quest to find planets capable of supporting life beyond our solar system Through the integration of machine learning techniques the future of exoplanet detection is set to become more efficient effective and comprehensive enabling the scientific community to push the boundaries of our understanding of the universe

9 Declarations

9.1 Competing Interests

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

9.2 Publisher's Note

AIJR remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

How to Cite

Deepanshu Tyagi, Renu Chaudhary, Gurpreet Kaur, Rekha Kumari (2025). Identification of Exoplanets using Machine Learning. *AIJR Proceedings*, 169-182. https://doi.org/10.21467/proceedings.178.20

REFERENCES

[1] C. J. Shallue and A. Vanderburg, "Exoplanet detection via machine learning: A fast, neural network driven approach," *arXiv preprint*, vol. 1, no. 1, pp. 1-10, Dec. 2017. https://arxiv.org/abs/1712.05044

[2] A. Vanderburg, et al., "Astronet: An artificial intelligence system to identify exoplanets in light curves," *The Astronomical Journal*, vol. 155, no. 1, pp. 22, Jan. 2018. https://iopscience.iop.org/article/10.3847/1538-3881/aa9e09

[3] K. A. Pearson, et al., "TransitNet: Detecting periodic features using deep learning," *Monthly Notices of the Royal Astronomical Society*, vol. 481, no. 4, pp. 5098–5109, Dec. 2018. https://academic.oup.com/mnras/article/481/4/5098/5145839

[4] T. Zingales, et al., "Exoplanet: Detection of exoplanets using deep learning," *Monthly Notices of the Royal Astronomical Society*, vol. 479, no. 1, pp. 1453–1462, Sep. 2018. https://academic.oup.com/mnras/article/479/1/1453/4985702

[5] J. L. Coughlin, et al., "Robovetter: Automated light curve classification for Kepler planet candidates," *The Astrophysical Journal*, vol. 820, no. 2, pp. 136, Mar. 2016. https://iopscience.iop.org/article/10.3847/0004-637X/820/2/136

[6] S. E. Thompson, et al., "Machine learning for the discovery of Kepler transiting exoplanets," *The Astronomical Journal*, vol. 155, no. 1, pp. 38, Jan. 2018. https://iopscience.iop.org/article/10.3847/1538-3881/aa9e10

[7]] D. J. Armstrong, et al., "An automated vetting of 2.4 million time-series light curves from the Kepler mission," *Monthly Notices of the Royal Astronomical Society*, vol. 478, no. 3, pp. 3640–3653, Aug. 2018. https://academic.oup.com/mnras/article/478/3/3640/4995237

[8] C. Marois, D. Lafreniere, R. Doyon, B. Macintosh, and D. Nadeau, "Angular differential imaging: A powerful high-contrast imaging technique," *arXiv preprint*, vol. 1, no. 1, pp. 1-10, Dec. 2005. https://arxiv.org/abs/astro-ph/0512335

[9] M. Ansdell, et al., "K2 variable catalog: Variable stars and eclipsing binaries in K2 campaigns 1-8," *The Astronomical Journal*, vol. 156, no. 2, pp. 84, Aug. 2018. https://iopscience.iop.org/article/10.3847/1538-3881/aacb21

[10] D. J. Armstrong, et al., "Machine learning for TESS: A case study with eclipsing binaries," *Monthly Notices of the Royal Astronomical Society*, vol. 488, no. 3, pp. 3406–3423, Sep. 2019. https://academic.oup.com/mnras/article/488/3/3406/5549173

[11] A. Giuliani, et al., "Deep neural networks for direct exoplanet detection," *Astronomy & Astrophysics*, vol. 631, pp. A90, Nov. 2019. https://www.aanda.org/articles/aa/full_html/2019/11/aa36784-19/aa36784-19.html

[12] C. J. Shallue and A. Vanderburg, "Identifying exoplanets with deep learning: A five-planet resonant chain around Kepler-80 and an eighth planet around Kepler-90," *The Astronomical Journal*, vol. 155, no. 2, pp. 94, Feb. 2019. https://iopscience.iop.org/article/10.3847/1538-3881/aa9e09