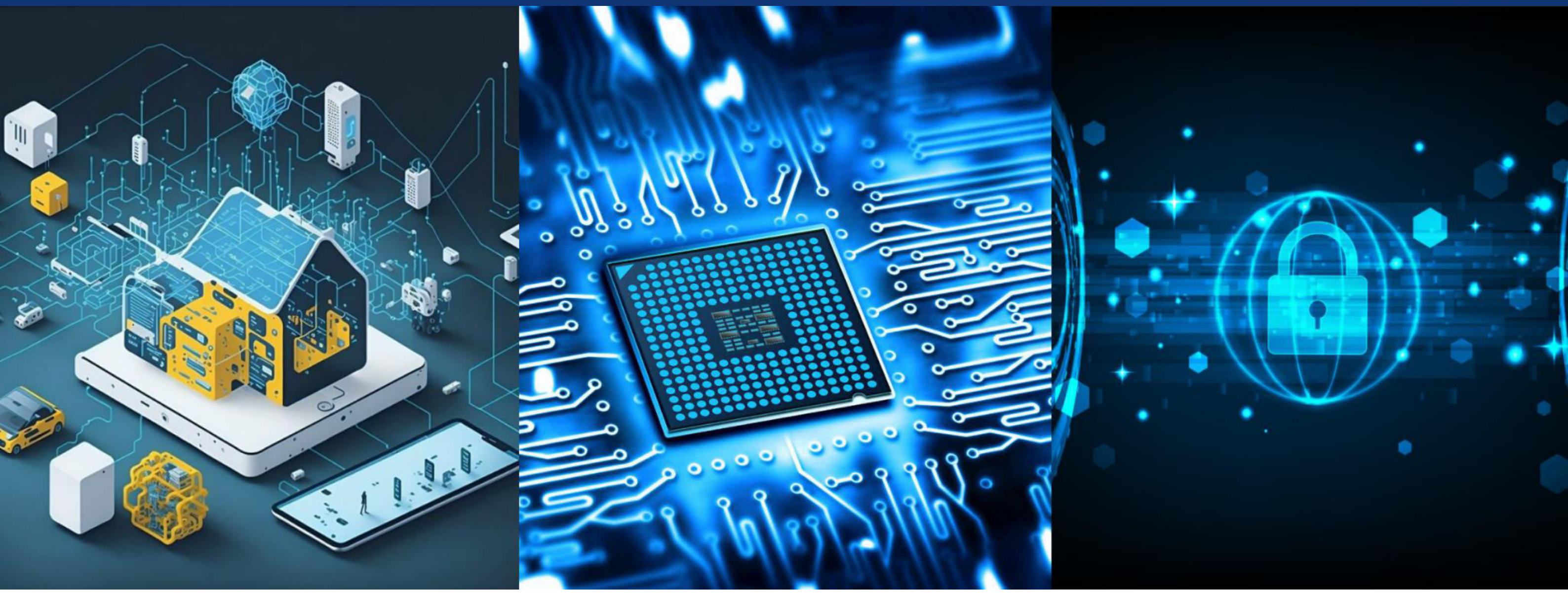
IJIRCCE

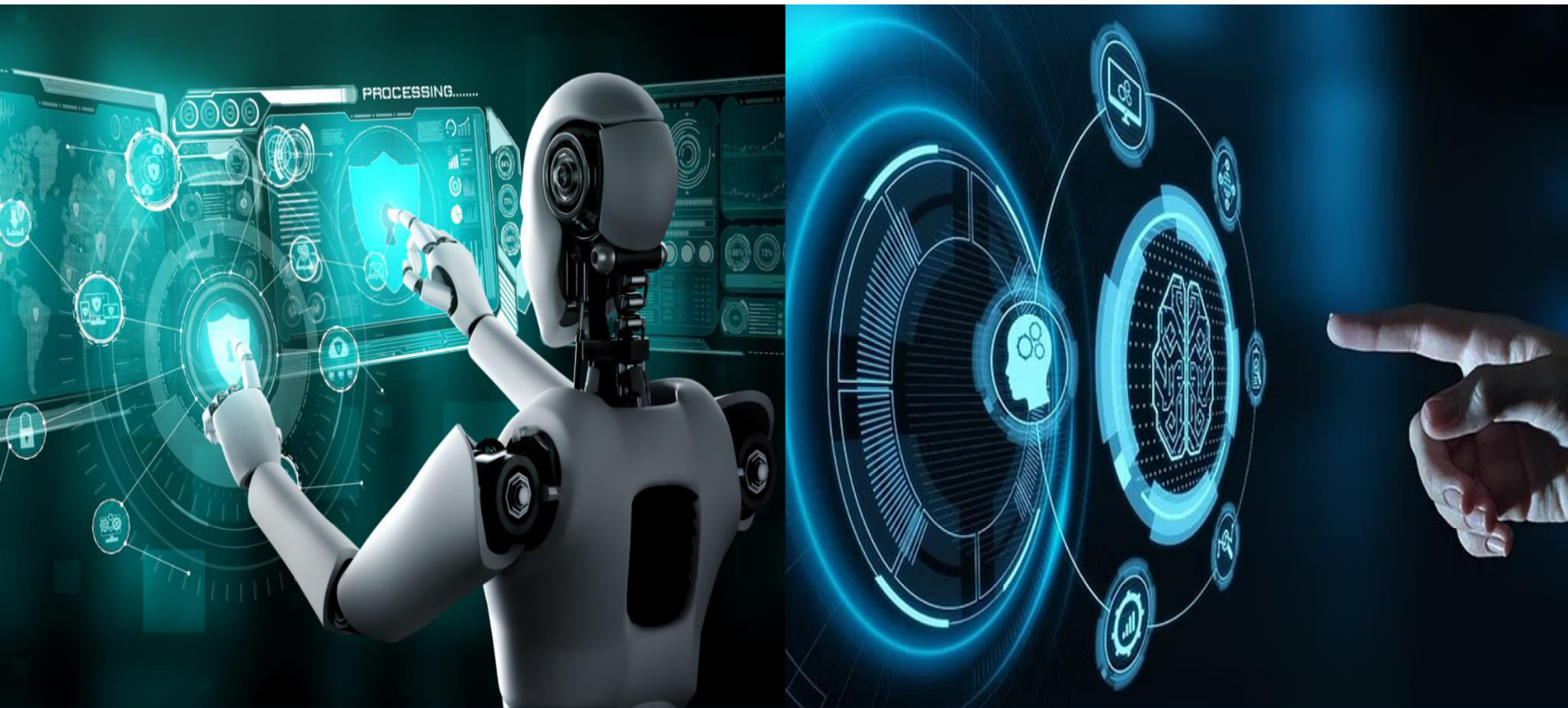
ISSN(O): 2320-9801

ISSN(P): 2320-9798



International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.771

Volume 14, Issue 4, April 2026

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

EXOENSEMBLE: An Approach to Exoplanet Spectrum Extraction

Kollu Poojitha[1], Mohammed Abdul Kalam Khan[2], Ms.N.Musrat Sultana[3], Dr.K.Rajitha[4], Dr.V. Subbaramaiah[5]

Student, Department of Computer Science and Engineering, Mahatma Gandhi Institute of Technology, Gandipet, India[1,2]

Assistant Professor, Department of Computer Science and Engineering, Mahatma Gandhi Institute of Technology, Gandipet, India[3,4,5]

ABSTRACT: Exoplanetary atmospheric characterization is a critical challenge in modern astronomy due to the extremely weak signals and complex noise present in observational data. Traditional analytical methods often fail to accurately extract spectral features under such conditions. This paper introduces **EXOENSEMBLE**, a hybrid ensemble-based framework that integrates advanced signal processing, Shared-Weight 1D Convolutional Neural Networks (CNN), and Rational Quadratic Neural Networks (RQ-NN) for robust spectrum reconstruction. The system performs multi-instrument data fusion and applies rigorous calibration techniques to recover faint atmospheric signatures. By leveraging deep ensemble learning with Gaussian Log-Likelihood constraints, EXOENSEMBLE produces both mean spectra and uncertainty estimates, enabling reliable detection of atmospheric compounds such as water vapor, methane, and carbon dioxide. Experimental insights demonstrate improved denoising, generalization, and uncertainty quantification compared to traditional methods, making the framework highly suitable for next-generation exoplanet missions.

KEYWORDS: Exoplanets, Atmospheric Retrieval, Deep Learning, CNN, Ensemble Learning, Spectral Reconstruction, Noise Modeling, Uncertainty Quantification, Astronomical Data Processing, Machine Learning in Astronomy.

I. INTRODUCTION

Exoplanetary research has emerged as a key domain in astronomy, focusing on understanding planetary formation, atmospheric composition, and the potential for extraterrestrial life. During planetary transits, a small portion of starlight passes through the exoplanet's atmosphere, encoding valuable spectral information. However, these signals are extremely weak and often obscured by instrument noise and stellar variability. Traditional methods such as curve fitting and PCA struggle to handle nonlinear and time-dependent noise patterns. Recent advancements in machine learning have shown promise in improving detection and analysis accuracy. In this context, hybrid systems that combine signal processing and deep learning are essential for improving spectral reconstruction and reliability. This work proposes **EXOENSEMBLE**, a unified framework that integrates deep learning, ensemble modeling, and mathematical noise handling for accurate exoplanet spectrum extraction and atmospheric analysis.

II. LITERATURE SURVEY

[1] Enhanced Classification of Exoplanets Using Machine Learning: Integrating Feature Selection, Ensemble Algorithms, and Explainable AI Insights by Ishvarya G, Kavya Sree Kammari, Sharanya Vanraj Thambi, and Manju Venugoplan (2024, IEEE ICCCNT) employs machine learning algorithms for exoplanet classification with enhanced accuracy through feature selection techniques. The study utilizes KOI planet candidates data and applies chi-square method for feature selection, while employing PCA for data visualization.

[2] Active Learning Framework for Anomaly Detection in Astronomical Data by Ghadiyaram Hanumath Srinivasa Dixit, Jitendra Parmar, Atul Kumar Verma, Raghavendra Mishra, Gopal Behera (2024, IEEE ICCCNT) introduces a

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

novel machine learning framework designed to efficiently detect anomalies in large-scale astronomical datasets. The study addresses the challenge of limited labeled data in astronomy by employing an active learning approach that iteratively selects the most informative samples for labeling, thereby improving detection accuracy while minimizing manual effort.

[3] Reconstructing Atmospheric Parameters of Exoplanets Using Deep Learning by Flavio Giobergia, Alkis Koudounas, and Elena Baralis (2023, IEEE AICT) presents a multi-target probabilistic regression approach combining deep learning and inverse modeling techniques within a multimodal architecture to extract atmospheric parameters from exoplanets. The methodology addresses the complex problem of estimating atmospheric parameters that best fit observed spectra within specified atmospheric models.

[4] Chandrayaan-3: Challenges and Opportunities of using AI to Detect and Identify Exoplanets by M Benisha, Gautham Gopan, and Mithileysh Sathiyanarayanan (2023, IEEE ICCAMS) presents a survey on opportunities and challenges of using AI for exoplanet detection and identification. Drawing inspiration from the successful Chandrayaan-3 mission by ISRO, the paper discusses how AI played crucial roles in spacecraft launching, navigation, prediction, anomaly detection, mission optimization, data processing, and communication.

[5] Astronomical Image Processing: Exoplanet Detection by M.S. Prasad, Shivani Verma, and Yulia A. Shichkina (2023, IEEE SCM) focuses on time-series data analysis for space data through interactive light-curves examination. The system utilizes data from NASA Exoplanet Archive and missions like Kepler and TESS, employing Box Least Squares Methods for periodogram analysis to detect transiting signals. The pre-processed normalized flux data generates Transit models and Cadence Masks from BLS periodograms.

[6] Classification of Images with Linear Objects in All-Sky Astronomical Survey Data using Convolutional Neural Networks" by Lenka Kali kov, Peter Butka (2023, IEEE) proposes a deep-learning method to automatically identify and classify linear features—such as asteroid trails, satellite streaks, meteor tracks or cosmic ray events—in all-sky astronomical survey images. The work leverages convolutional neural networks (CNNs) to scan large volumes of wide-field imaging data, extracting patterns associated with linear objects against complex celestial backgrounds.

[7] Low-Rank Plus Sparse Trajectory Decomposition for Direct Exoplanet Imaging by Simon Vary, Hazan Daglayan, Laurent Jacques, and P.-A. Absil (2023, IEEE ICASSP) proposes a direct imaging method for exoplanet detection using a combined low-rank plus structured sparse model. The approach develops a dictionary of possible effective circular trajectories for planets during observation time, computed efficiently using rotation and convolution operations.

[8] Exoplanet Detection Using Feature Engineering with Ensemble Learning by G. Venkata Sai Rakesh, M. Jahnavi Bhuvana Chandrika, and Ch. Venkata Rami Reddy (2023, IEEE ICPCSN) presents a novel approach to detecting exoplanets via the transit method by first crafting meaningful features from raw astronomical data and then applying ensemble machine-learning techniques. The study uses six base classifiers — Decision Tree, Support Vector Classifier, K-Nearest Neighbour, Random Forest, Multi-Layer Perceptron, and others — and then combines selected top performers (SVC, KNN, RF, MLP) through a majority-voting ensemble to boost detection accuracy.

III. PROBLEM DEFINITION

Extracting exoplanetary spectra is highly challenging due to extremely weak transit signals (50–200 ppm) buried in strong, time-dependent noise from instrumental effects and stellar variability. Nonlinear, correlated noise patterns and inconsistencies across multi-instrument datasets further complicate analysis. Conventional methods like PCA and statistical filtering struggle to preserve subtle spectral features and lack robust uncertainty quantification. There is therefore a critical need for an intelligent, data-driven system capable of accurate denoising, reliable spectral reconstruction, multi-instrument data integration, and systematic uncertainty estimation.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

IV. PROPOSED SYSTEM

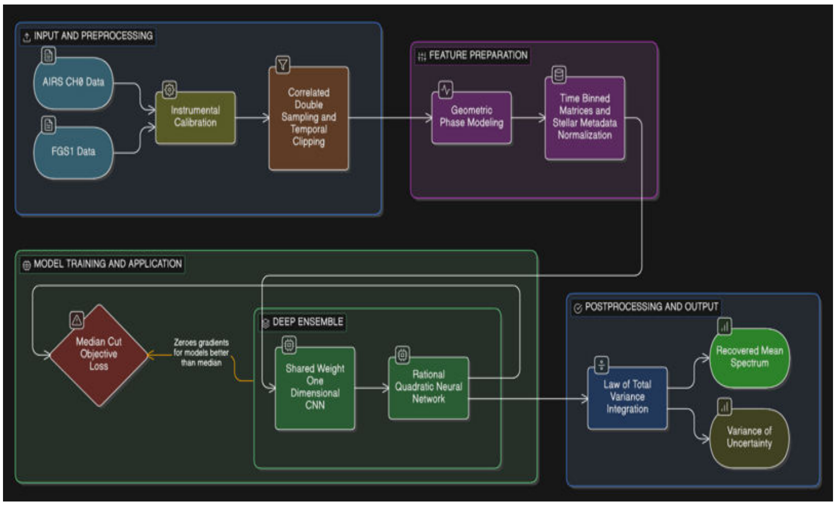
EXOENSEMBLE is a hybrid ensemble framework for exoplanetary spectral extraction and atmospheric characterization. It combines an advanced calibration pipeline — including dark correction, flat fielding, and noise reduction — with Shared-Weight 1D CNNs for temporal feature extraction and Rational Quadratic Neural Networks for prototype-based spectral reconstruction. Processing multi-instrument data from FGS1 (visible photometric) and AIRS-CH0 (infrared spectroscopic) channels, the system employs a deep ensemble trained with Gaussian Log-Likelihood constraints to jointly predict the mean spectrum (μ) and variance (σ^2), ensuring robust uncertainty quantification.

V. EXISTING SYSTEM

Existing approaches combine machine learning, deep learning, and statistical techniques with notable limitations. Traditional models such as KNN, SVM, and Random Forests handle classification tasks but struggle with noisy astronomical data. Deep learning methods improve transit detection but often lack robust uncertainty modeling. Statistical techniques like PCA and curve fitting fail with nonlinear, correlated noise and risk losing subtle spectral features. Sparse and decomposition-based methods are computationally expensive and require extensive manual tuning. Collectively, these systems suffer from poor noise modeling, limited generalization, weak uncertainty estimation, and no integrated multi-instrument processing.

VI. DESIGN AND METHODOLOGY

EXOENSEMBLE follows a four-stage pipeline: preprocessing, feature extraction, model training, and postprocessing. Raw spectroscopic and photometric data are first calibrated using dark correction, flat fielding, and noise filtering to remove sensor-level distortions. The cleaned data is then transformed into structured time-series features through transit signal extraction and binning. A hybrid model combining Shared-Weight 1D CNNs for temporal feature extraction with Rational Quadratic Neural Networks for spectral reconstruction forms the core, trained using a Gaussian Log-Likelihood objective across a deep ensemble for robust uncertainty quantification. Finally, statistical aggregation via the Law of Total Variance produces the reconstructed spectrum and uncertainty estimates, ensuring accuracy and scalability for real-world astronomical applications.



The above Figure shows the **architecture diagram** illustrates the EXOENSEMBLE data processing and reconstruction pipeline for exoplanetary spectra. It begins with Input and Preprocessing of spectroscopic and photometric data, followed by Feature Preparation utilizing algorithmic extraction. The core is the Model Training and Application block, where an ensemble integrating Shared-Weight Convolutional Neural Networks and Rational Quadratic Neural Network components processes the data. Finally, the Postprocessing and Output block

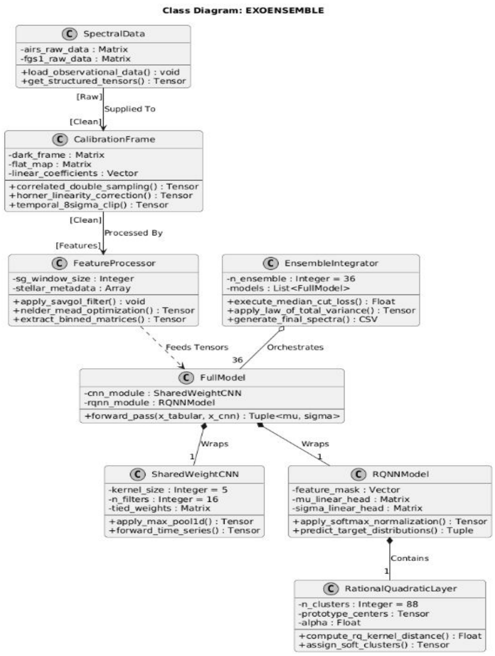
© 2026 IJIRCCE | Volume 14, Issue 4, April 2026 DOI: 10.15680/IJIRCCE.2026.1404103

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

mathematically binds the ensemble predictions to generate the Recovered Mean Spectrum and the Variance of Uncertainty.

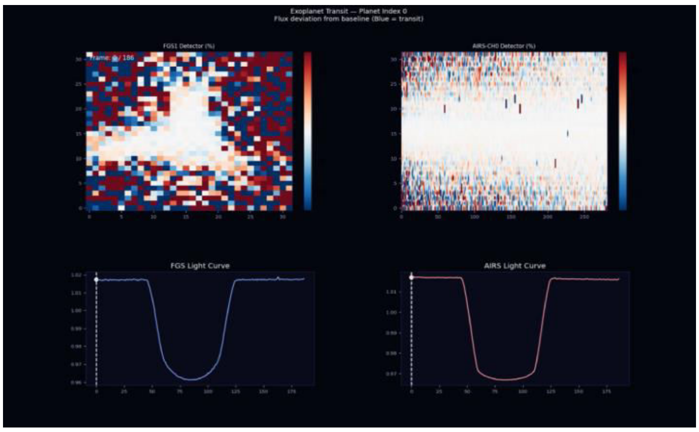


The Figure outlines the EXOENSEMBLE system, which is divided into three functional stages: Data Processing (calibration and geometric feature extraction), Modeling & Ensemble (combining Shared-Weight CNN logic and Rational Quadratic Neural Network soft-clustering for spectral reconstruction and targeted uncertainty estimation), and Reconstruction & Evaluation (fusing the total ensemble spectrum, calculating epistemic uncertainty, and assessing performance metrics).

VII. IMPLEMENTATION

7.1 Data Preprocessing Implementation

The preprocessing stage cleans and prepares raw FGS1 and AIRS-CH0 data using CDS for drift removal, ADC correction, and Horner-based linearity adjustment. Noise and anomalies are removed through sigma clipping, dark subtraction, and 8σ outlier filtering. The signal is then smoothed, transit boundaries are extracted, and features are normalized for stable model input.



© 2026 IJIRCCE | Volume 14, Issue 4, April 2026 DOI: 10.15680/IJIRCCE.2026.1404103

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

7.2 Model Training Implementation

The training stage uses a hybrid ensemble of Shared-Weight 1D CNN and RQ-NN for feature extraction and spectral reconstruction. Optimized with AdaBelief and a cosine learning schedule, a 36-model ensemble is trained using a median-cut loss to ensure robustness and accurate uncertainty estimation.

7.3 System Validation and Inference Implementation

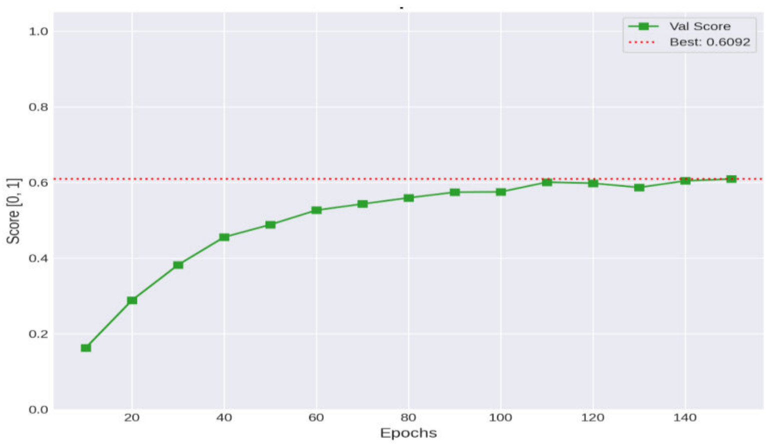
During inference, predictions from all 36 ensemble models are combined, where the mean spectrum is obtained through averaging and uncertainty is computed using the Law of Total Variance to capture both aleatoric and epistemic components. Model performance is evaluated using Gaussian Log-Likelihood, with weighted scoring across instruments, producing a normalized and reliable measure of prediction accuracy.

VIII.TESTING AND RESULTS

8.1 Stable Ensemble Convergence

Metric	Value
Best Training Loss (GLL)	1.35
Best Validation Loss (GLL)	1.64
Epoch at Best Performance	150
Final Validation Score	0.6092

The training and validation loss curves demonstrated in Figure confirm controlled and stable convergence. The Median-Cut Objective Loss successfully prevented ensemble collapse—individual model training losses converge at distinctly varied rates, confirming that ensemble member diversity was actively maintained throughout the 300+ epoch training process



The above figure shows the validation score of 0.6092 indicates accurate spectral reconstruction with reliable uncertainty estimation, while the small gap between training and validation loss confirms good generalization. Overall, the results demonstrate effective noise handling and robust recovery of faint exoplanet signals.

IX. CONCLUSION AND FUTURE SCOPE

This work presents EXOENSEMBLE, a robust hybrid framework for exoplanetary spectrum extraction and atmospheric characterization from highly noisy observational data. By integrating advanced signal preprocessing, Shared-Weight 1D

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

CNN, Rational Quadratic Neural Networks, and a deep ensemble strategy, the system effectively captures complex noise patterns and reconstructs accurate spectral signals. The incorporation of the Law of Total Variance enables reliable uncertainty quantification, improving scientific interpretability. Experimental results demonstrate stable convergence, strong generalization, and enhanced capability in detecting faint atmospheric signatures such as H_2O, CH_4, and CO_2. Overall, EXOENSEMBLE provides a scalable and mathematically grounded solution for next-generation exoplanet analysis.

REFERENCES

[1] I. G, K. S. Kammari, S. V. Thambi, and M. Venugopalan, "Enhanced Classification of Exoplanets Using Machine Learning: Integrating Feature Selection, Ensemble Algorithms, and Explainable AI Insights," 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT), Kamand, India, 2024, pp. 1-7. DOI: https://ieeexplore.ieee.org/document/10725900

[2] G. H. S. Dixit, J. Parmar, A. K. Verma, R. Mishra, and G. Behera, "Active Learning Framework for Anomaly Detection in Astronomical Data," 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT), Kamand, India,2024,pp.17. DOI: https://ieeexplore.ieee.org/document/10726168

[3] M. S. Prasad, S. Verma, and Y. A. Shichkina, "Astronomical Image Processing: Exoplanet Detection," 2023 XXVI International Conference on Soft Computing and Measurements (SCM), Saint Petersburg, RussianFederation,2023,pp.1-5. DOI: https://ieeexplore.ieee.org/document/10159069

[4] S. Vary, H. Daglayan, L. Jacques, and P.-A. Absil, "Low-Rank Plus Sparse Trajectory Decomposition for Direct Exoplanet Imaging," ICASSP 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Rhodes Island, Greece, 2023, pp. 1-5. DOI: https://ieeexplore.ieee.org/document/10096197

[5] G. V. S. Rakesh, M. J. B. Chandrika, C. V. R. Reddy, and M. Suneetha, "Exoplanet Detection Using Feature Engineering with Ensemble Learning," 2023 3rd International Conference on Pervasive Computing and Social Networking (ICPCSN), Salem, India, 2023, pp. 118-123.. DOI: https://ieeexplore.ieee.org/document/10266186

[6] M. Benisha, G. Gopan, and M. Sathiyanarayanan, "Chandrayaan-3: Challenges and Opportunities of using AI to Detect and Identify Exoplanets," 2023 International Conference on New Frontiers in Communication, Automation, Management and Security (ICCAMS), Bangalore,India,2023,pp.1-6. DOI: https://ieeexplore.ieee.org/document/10525805

[7] A. Malik, B. P. Moster, and C. Obermeier, "Exoplanet detection using machine learning," Monthly Notices of the Royal Astronomical Society, vol. 513, no. 4, pp. 5505-5516, Dec. 2021. DOI: https://ieeexplore.ieee.org/document/10084656

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING