EXPLAINABLE REAL-TIME CNN-VIT MODEL FOR MULTISPECTRAL CROP STRESS DETECTION IN PRECISION AGRICULTURE

Project Reference No.: *49S_BE_0787*

College	:	*Christ Deemed to be University, Kengeri Campus, Bengaluru*
Branch	:	*Department of Artificial Intelligence and Machine Learning*
Guide(s)	:	*Dr. Manjunatha Swamy C*
Student(s)	:	*Mr. Aldrin P Thomas*
		Mr. Muhamed Nabhan
		Mr. Sanyu Reji Kuriakose

Keywords:

Precision Agriculture, Crop Stress Detection, Vision Transformer, Convolutional Neural Network, Explainable AI, Sentinel-2

Introduction:

Agriculture plays a crucial role in ensuring food security and economic stability, especially in developing countries like India. However, crop productivity is signifi cantly affected by environmental stress factors such as drought, nutrient deficiency, pest infestation, and water imbalance. Traditional crop monitoring methods rely heavily on manual field inspections, which are time-consuming, labor-intensive, and not scalable for large agricultural areas. Recent advancements in remote sensing and artificial intelligence have enabled au tomated crop health monitoring using satellite imagery and deep learning models. Sentinel-2 multispectral imagery provides valuable spectral information across vis ible, near-infrared (NIR), and short-wave infrared (SWIR) bands, allowing computation of vegetation indices such as NDVI, NDWI, NDRE, and MSI. These indices help in detecting vegetation vigor, chlorophyll content, and water stress. Deep learning architectures such as Convolutional Neural Networks (CNNs) and Vision Transformers (ViTs) have demonstrated strong performance in image-based classification tasks. Integrating these models with multispectral satellite data creates an opportunity for scalable and intelligent precision agriculture systems.

Existing agricultural monitoring approaches either focus solely on satellite-based vegetation analysis or on plant-level disease detection, but rarely integrate both into a unified explainable framework. Furthermore, many AI models function as black boxes, limiting interpretability and adoption in real-world agricultural scenarios.

To address the identified gap, this project proposes an explainable dual-model framework that integrates a Vision Transformer (ViT) for satellite-based crop 2 stress detection and a Convolutional Neural Network (CNN) for leaf-level disease identification. The system computes vegetation indices from Sentinel-2 imagery, performs classifi cation using deep learning architectures, and provides visual explanations through Grad-CAM and attention maps. By combining large-scale remote sensing with plant-level diagnostics, the framework ensures comprehensive precision agriculture monitoring. This structured integration forms the foundation of the proposed GeoVision system

Objectives:

1. To analyze Sentinel-2 multispectral imagery and compute vegetation indices for crop health assessment.
2. To develop a Vision Transformer-based model for satellite-level crop stress detection.
3. To implement a CNN-based model for plant-level disease classification.
4. To integrate Explainable AI techniques such as Grad-CAM and attention visualization.
5. To deploy the system through a user-friendly web-based interface

Methodology:

This project follows a quantitative and experimental research methodology inte grating remote sensing analytics, deep learning modeling, and explainable AI tech niques for precision agriculture. The study combines multispectral satellite-based crop stress detection with plant-level disease classification and chatbot-assisted advisory support. The overall framework consists of data acquisition, preprocessing, feature engi neering, model development, optimization, evaluation, explainability integration, and deployment through a unified web-based system. The proposed methodology ensures both scientific rigor and real-world applicability.

The overall integrated system architecture combining satellite analysis and plant disease detection is shown in Figure 2. The system consists of two primary pipelines: • Vision Transformer-based satellite crop stress detection • CNN-based plant disease classification with chatbot integration.

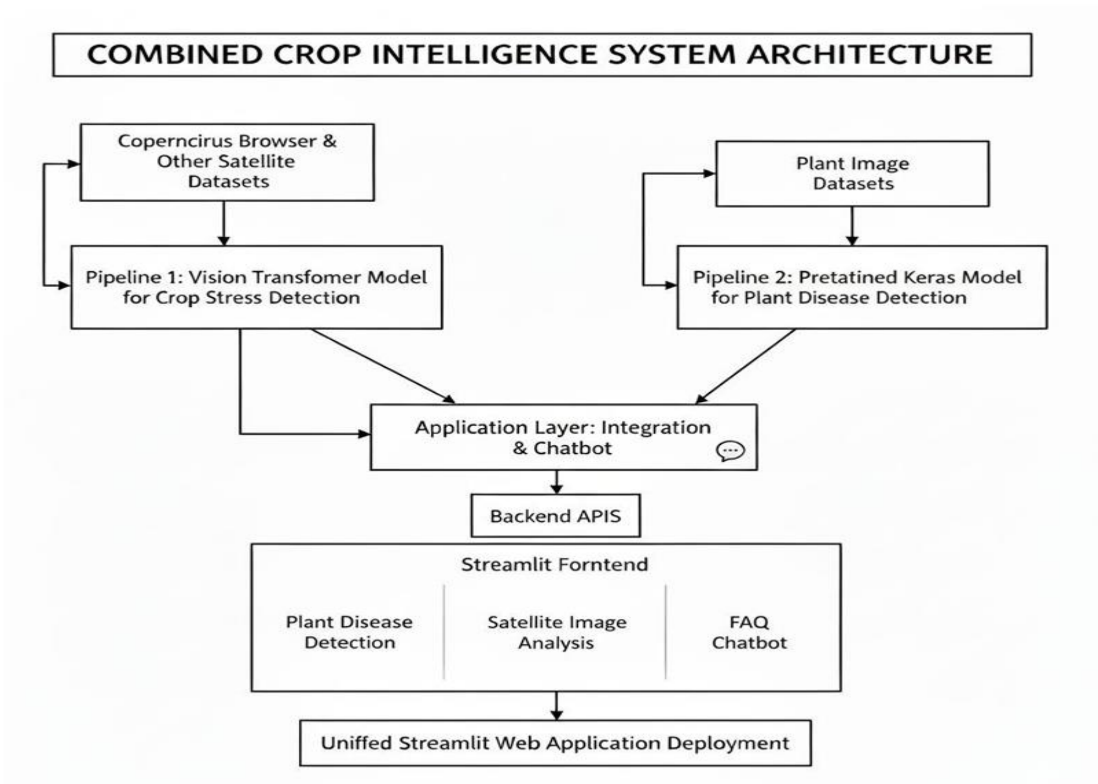

Figure 1*: Combined Crop Intelligence System Architecture*

Result and Conclusion:

The Vision Transformer model trained on multispectral satellite imagery achieved an overall classification accuracy of approximately 87% on the EuroSAT dataset. The CNN-based plant disease detection model achieved an accuracy of over 92% on the leaf image dataset. This project successfully developed an explainable dual-model precision agricul ture framework that integrates Vision Transformer–based multispectral satellite analysis with CNN-based plant disease detection. The proposed system combines large-scale crop stress monitoring with plant-level disease classification, creating a unified approach to agricultural health assessment. The Vision Transformer model effectively captured global spatial patterns from satellite imagery, while the CNN 25 model accurately identified localized disease features in leaf images. The frame work achieved strong classification performance across both tasks, supported by reliable accuracy, precision, recall, and F1-score metrics. Furthermore, the integration of Explainable AI techniques enhanced model transparency by highlighting critical regions influencing predictions, thereby improving interpretability and trustworthiness. In addition to strong model performance, the system demonstrated practical deployment feasibility through a user-friendly web-based interface enabling satellite area selection, vegetation index computation, and AI-assisted guidance. This deployment-oriented design illustrates the scalability and real-world applicability of the framework. Overall, the proposed GeoVision framework contributes toward intelligent, scalable, and sustainable agricultural monitoring solutions by integrating deep learning, multispectral analysis, and explainability into a cohesive precision agriculture system

References:

[1] Murugavalli, S., & Gopi, R. (2025). Plant leaf disease detection using vision transformers for precision agriculture. PMC Journal, 12, PMC12216567

[2] Alshammari, S. A. (2024). TLDViT: A vision transformer model for tomato leaf disease classification. International Journal of Advanced Computer Science and Applications, 15(12), 841 849.

[3] Haque, M. A., Rahman, S., & Islam, M. (2025). An enhanced vision transformer network for efficient and accurate crop disease detection. Expert Systems with Applications, 238, 121365.

[4] Naseer, R., Shashidhar, V. S., & Shreya, S. B. (2025). Explainable AI for enhancing decision making in precision agriculture. International Journal of Engineering Research and Technology, 14(5), 156–172.

[5] Kawakura, S., Hirafuji, M., Ninomiya, S., & Shibasaki, R. (2023). Visual analysis of agricultural workers using explainable artificial intelligence based on class activation mapping. European Journal of Artificial Intelligence and Machine Learning, 2(1), 12–19.

[6] Ahmed, M. T., & Rahman, S. (2025). A systematic review of explainable artificial intelligence for spectroscopic agricultural quality assessment. Computers and Electronics in Agriculture, 210, 108456.

[7] Zhang, R., Wu, X., Li, J., Zhao, P., Zhang, Q., Wuri, L., Zhang, D., Zhang, Z., & Yang, L. (2025). A bibliometric review of deep learning in crop monitoring: Trends, challenges, and future perspectives. Frontiers in Artificial Intelligence, 8, 1636898.

[8] Jeong, S., Ko, J., & Yeom, J. M. (2024). Deep learning-enhanced remote sensing-integrated crop modeling approach for rice yield estimation and mapping. Agricultural Systems, 215, 103842.
[9] Wolfe, J. (2023). Deep learning in agricultural remote sensing applications. NV5 Geospatial Software Technical Report, 6317, 1–12.
[10] Wang, A. X., Tran, C., Desai, N., Lobell, D., & Ermon, S. (2018). Deep transfer learning for crop yield prediction with remote sensing data. Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies, 1–5.
[11] Kondaveeti, H. K., Kumaravelu, N. K., Vandrangi, S. K., Ruddarraju, R. R., Samarasinha, A., & Mamidi, R. (2025). Evaluation of deep learning models using explainable AI for crop disease classification. Scientific Reports, 15, 1430.

Future Scope:

The future scope of this project includes:

1. Edge Deployment in Rural Areas: The current system is based on a cloud-based processing system that requires a stable internet connection. One of the most practical future directions is to fine-tune the deep learning models to execute on low-cost edge computing devices (like a Raspberry Pi). This will enable the farmers in remote locations where there is poor or no internet connectivity to apply real-time offline crop health tests.
2. Mobile Application Integration: Although the existing prototype will utilize a web-based Streamlit interface, the construction of a specific smartphone application will greatly enhance the user experience. Disease diagnosis would be instantly done by a farmer who would easily use his phone cameras to take photos of the leaves and then use a mobile application which would further help in diagnosing the disease. Also, the app may leverage the inbuilt GPS on the device to automatically retrieve the satellite position of the user without having to input the coordinates manually.
3. Combination with IoT-Based Soil Sensors: Satellite imagery can be an outstanding wide-area measurement of crop moisture and health, yet it can also be confirmed with ground-level data. The framework could be integrated with Internet of Things (IoT) soil sensors in the fields to enable the framework to compare satellite predictions (such as NDWI) with the real physical measurements of soil moisture and temperature, which would be very precise in providing irrigation recommendation