INTELLIGENT AGRICULTURAL ADVISORY SYSTEM USING TRANSFORMER MODEL & KNOWLEDGE BASE

Project Reference No.: *49S_MTECH_0005*

College : *Reva University, Bengaluru*
Branch : *Department of Computer Science Engineering*
Guide : *Dr. Laxmi B Rananavare*
Student : *Mr. Viraj D Talathi*

Keywords:
Retrieval Augmented Generation systems Data Efficient Image Transformers, Plant leaf disease classification, Vrikshayurveda, Intelligent Systems.

Introduction:
Agriculture plays a vital role in the food security and livelihood support of agrarian societies. Vegetables such as tomatoes and crops such as cassava play a crucial role as food and raw materials for industries. However, the productivity of these crops is greatly influenced by diseases such as leaf diseases. Leaf diseases have a great impact on the reduction of crop yields and have economic consequences for farmers. Disease diagnosis is one of the major challenges in the field of agriculture, as there is limited availability of experts and the diseases have high visual similarities. Various diseases have the same texture, vein patterns, and lesion structures. With the development of deep learning techniques such as Convolution Neural Networks, the detection of diseases in crops is possible. However, the local perception of CNN is limited to the local perception of the image. Vision Transformers have the potential to handle global spatial dependencies in images. However, current systems lack important environmental factors such as soil and climatic conditions and lack the incorporation of traditional knowledge systems such as Vrikshayurveda. This work proposes a crop-agnostic system, providing disease-specific recommendations using both conventional agricultural practices and traditional knowledge systems.

Objectives:
1. To develop a transformer based, crop-agnostic system for leaf disease classification.
2. To capture global spatial dependencies using self-attention mechanisms.
3. To incorporate soil and climatic factors for context-aware analysis.
4. To integrate modern treatments with Vrikshayurveda-based knowledge using RAG.
5. To enable accessibility through multilingual support and text-to-speech.

Methodology:
The methodology presents a pipeline for crop disease diagnosis combined with advisory development. The model can be used for the diagnosis and development of advisories for tomato and cassava leaf diseases, among others, using a unified approach. The data is obtained from mendeley & Kaggle, where the classes for the diseases are labelled. The pre-processing steps include resizing, normalization, and augmentation. The class balancing

method is also used for the model's generalization. The data is split into training, validation, and test sets. The Data Efficient Image Transformer model uses pre-trained weights and is fine-tuned using the prepared data with a supervised learning approach. Training process involves optimizing the loss function with backpropagation. The model's performance is evaluated using metrics like accuracy, precision, recall, and F1 score. The model's performance is also evaluated against other internal and external datasets with similar preprocessing. Feature representations are analysed using dimensionality reduction techniques such as t-SNE or UMAP. The system is designed to be scalable and extendable to additional crops with minimal modifications. A hybrid module has been designed to provide recommendations. Modern treatment details are stored and retrieved in JSON format. A RAG module has been designed to fetch Ayurvedic knowledge from Vrikshayurveda.

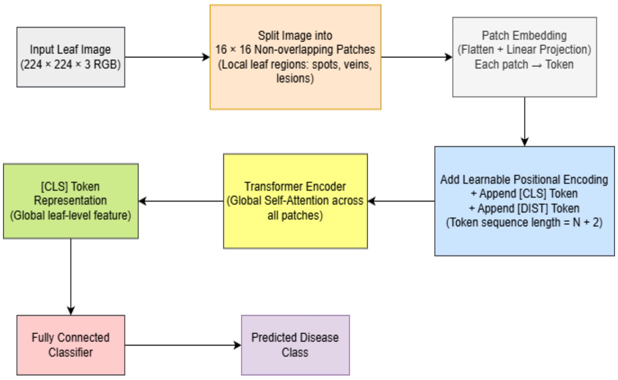

Figure 1: DEIT Architecture

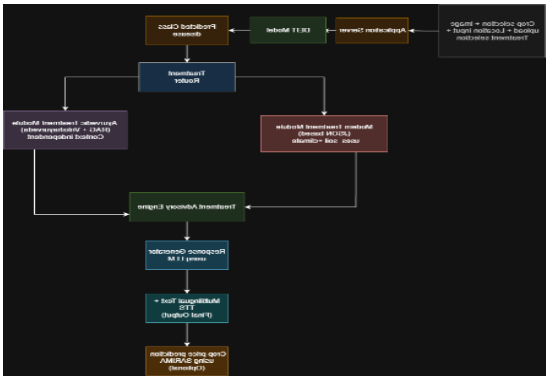

Figure 2: System Architecture

Environmental factors such as soil and climate are also considered. Recommendations are also refined based on the context. Multilingual support with text-to-speech functionality is also provided.

Result and Conclusion:

In conclusion, the proposed system shows strong performance. The accuracy for the tomato dataset is 97.86%. The precision, recall, and F1-score for the model on the tomato dataset are 97.83%, 97.87%, and 97.84%, respectively. For the cassava dataset, the accuracy for the model is 93.10%. The precision, recall, and F1-score for the model on the cassava dataset are 93.10%. These results indicate effective classification even for visually similar disease classes. The system achieves near real-time performance with an average latency of approximately 650ms. The integration of environmental factors improves the relevance and accuracy of treatment recommendations. This work is implemented as a simulation of a real-world agricultural advisory system. The outcome validates the feasibility of integrating AI-based classification with hybrid knowledge systems. From an industry perspective, the system is relevant for precision agriculture and decision support applications. It can help farmers, agronomists, and agritech platforms improve crop management. The project has provided valuable insights for designing scalable AI architectures. It has also improved knowledge about how machine learning can be combined with domain-specific knowledge. The project has shown how essential it is to balance technical innovation with usability.

References:

1. R. Alexander, A. Mary Valentina Janet, R. Vanidha Sri, S. Sasirekha, A. Kathyani and V. Shakthi Priya, "A Multilingual LLM-Driven Digital Twin Framework for Climate Resilient Precision Agriculture," 2025 6th International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India, 2025, pp. 1794-1805, doi: 10.1109/ICESC65114.2025.11212629.
2. R. Sheikh, D. Kumar, S. K. Singh, A. J. Hati and J. U. Kidav, "AgroBuddy: Empowering Indian Farmers Through Precision Farming Chatbot," 2025 3rd International Conference on Communication, Security, and Artificial Intelligence (ICCSAI), Greater Noida, India, 2025, pp. 1627-1632, doi: 10.1109/ICCSAI64074.2025.11064149.
3. S. Senthil Pandi, P. Kumar, R. M. Suchindhar and A. Naresh Kumar, "Improving Crop Yield Through Leaf Disease Detection and Pests Recommendations," 2025 International Conference on Knowledge Engineering and Communication Systems (ICKECS), Chickballapur, India, 2025, pp. 1-6, doi: 10.1109/ICKECS65700.2025.11034936.
4. V. D. Talathi and L. B. Rananavare, "Transformer Based Pipeline for Tomato Leaf Disease Detection and Intelligent Knowledge Augmentation," *2025 IEEE International Conference on Emerging Trends in Engineering and Computing (ETECOM)*, Riffa, Bahrain, 2025, pp. 1-6, doi: 10.1109/ETECOM66111.2025.11319083.

Future Scope:

The future scope of this project includes:

1. Further work is needed on improving robustness against domain shifts and variations.
2. Optimization of RAG pipeline which can lead to faster response time for the retrieval process.
3. Model compression methods may be used to allow deployment on edge devices.
4. The system may be tested for its usability in actual world situations, including low-resource settings
5. User feedback may be integrated into the system for its continuous improvement.
6. The framework may be used for other plant disease classification problems.
7. Expansion of the knowledge base can enhance coverage of treatment recommendations.
8. Implementing Explainability methods can be to improve model interpretability.
9. Integration with cloud-based platforms can enable scalable deployment.
10. Exploring fine-tuning strategies to improve performance on imbalanced datasets.
11. Incorporating uncertainty estimation to improve reliability of predictions.