Gomal Journal of Agriculture and Biology
Vol: 2 -- Issue: 2, 2024
Https://gomalpublishers.com/GJAB
Support@gomalpublisher.com
ISSN: 3007-1364 (Print)
ISSN: 3007-1372 (Online)

MACHINE LEARNING AND REMOTE SENSING INTEGRATION FOR EARLY DETECTION OF CROP DISEASES AND PEST OUTBREAKS

Nimra Samad [1*], **Muhammad Shafique** [2]
[1] Department of Plant Pathology, University of Layyah, Punjab, Pakistan.
[2] Ayyub Agriculture Research Institute, Faisalabad-38000-Pakistan.
*Corresponding Author E-mail: nimrasamad02@gmail.com

Abstract

This study examines how machine learning algorithms can be used to combine with remote sensing technologies to detect agricultural diseases and insects infestations on a timely basis. Our mixed-method dataset consisted of an integration of multispectral and hyperspectral images with field data and farmer interviews (ground-truth data). The steps preceding it were the correction of the atmosphere, normalization of reflectance and extraction of features based on vegetation indices and texture measurements. Training and testing multiple machine learning models, including support vectors machine, gradient boosting, convolutional neural network, and random forest, were performed with a stratified k-fold cross-validation approach. These models demonstrated that they were able to give correct predictions with ROC-AUC values that were greater than the baseline values and F1-scores demonstrating that they were able to be able to detect damaged crops. The models were validated with field surveys and qualitative farmer feedback and ensured their usefulness in real-world applications in agriculture. The findings revealed that combination of both quantitative spectral attributes and qualitative observations reduced the rate of misclassifications and enabled the system to be more stable in the evolving environments. The integrated process reveals a scaling, sustainable, and adaptable approach to performing precision agriculture that will be capable of prompt reactions that can reduce agricultural wasteage and enable food security to be more robust. The findings indicate that machine learning and remote sensing may be combined to establish an alternate method of real-time monitoring of a disease and pests.

Article History

Received:
August 12, 2024

Revised:
October 09, 2024

Accepted:
November 11, 2024

Available Online:
December 31, 2024

Keywords: Machine Learning, Remote Sensing, Crop Diseases, Pest Outbreaks, Precision Agriculture, Early Detection.

INTRODUCTION

Food security, economical development, and rural living in the world are all based on agriculture which is highly susceptible to crop illnesses and insect attacks. These biological pressures are a leading reason of losses in global yields that have been estimated to be 20-40 percent annually. The consequences are particularly poor in the underdeveloped regions (Savary et al., 2019). Agriculture is categorized as a response to changing climatic conditions, which enhances the formation of new relations with pests and the spread of illnesses, therefore, posing a problem of sustainable food production (Deutsch et al., 2018). Visual scouting and manual field checks are standard approaches of monitoring that are time and labor intensive and most of the time are biased hence do not qualify to be used in the large scale disease surveillance in a timely fashion. Here, machine learning (ML) and remote sensing (RS) has become a revolution in how crop illnesses and pest infestations can be detected early, precise and at scale (Barbedo, 2019). Remote sensing technology (also known as multispectral, hyperspectral and thermal imaging technology) provides us with alternative and special opportunities to observe tiny indicators of plant stress before it is too late. In particular, hyperspectral photography has proved the possibility to differentiate between unhealthy and healthy plant tissues through spectral reflectance abnormalities (Mahlein, 2018). Together with vegetation indices, such as the Normalized Difference Vegetation Index (NDVI) or the Red-Edge Chlorophyll Index, these spectral data are capable of displaying physiological responses to stress that are both disease- and pest-related (Xie et al., 2020). Nevertheless, the sheer amount and the high level of remote sensing data demand advanced computer approaches to identify meaningful patterns. The integration of machine learning into remote sensing is excellent because it can convert raw spectral responses into valuable data (Liakos et al., 2018). Many industrial sectors of agriculture have increasingly employed machine learning methods, such as disease classification, yields prediction, and soil health measurements (Kamilaris and Prenafeta-Boldu, 2018). The most popular ones include the Random Forest (RF), Support Vector Machines (SVM), Gradient Boosting (GB), and Convolutional Neural Networks (CNNs), each of which is capable of operating in high-dimensional datasets and operating with various feature spaces (Singh et al., 2020). A lot of attention has been given to CNNs, which are capable of discovering spatial-spectral correlations within images, which is why they are excellent in mapping the diseases (Too et al., 2019). However, the quality of such models is highly dependent on the quality of labelled data sets, and this usually requires a significant amount of ground-truthing in the form of field surveys. Integration of information provided by farmers with capabilities of sensors would result in a hybrid solution that enables quantitative accuracy and qualitative validation of pest and disease-finding systems (Raza et al., 2019). Recent studies highlight the promise of the combination of machine learning and remote sensing. An example is that Xie et al. (2020) employed both hyperspectral images and SVM to detect wheat rust with more than 90-percent accuracy. Similarly, Zhang et al. (2019) demonstrated that UAV-based multispectral imaging with the help of the Random Forest could detect rice blast in its initial stages before the visual symptoms could be observed. Sun et al. (2021) revealed that deep learning models, namely CNNs, used on drone shots were able to discern the different maize diseases at various growth stages which is more effective than the traditional classifiers. These conclusions support the shift to automated and data-driven disease surveillance in precision agriculture. Despite these advancements, it remains difficult to scale ML -RS systems. Spectral readings are also sensitive to changes in environmental parameters like soil moisture, light, and weather, and this aspect renders models less generalizable (Zhang et al., 2022). Also, the application of geography and crop-focused models is limited due to differences in spectral phenotypes and disease symptomology (Liu et al., 2020). Recent evidence suggests that the combination of multi-source data which would include the climate variables, soil health indicators, and observations of the farmers could help overcome these challenges (Rahman et al., 2022). Such integration does not only ensure that the predictions are more precise, but it also ensures that the solutions are useful and relevant to the people they will be used by. Early detection of crop diseases and pests has impacts that extend well beyond ensuring that farming is more productive. It instantly assists in reducing the haphazard application of pesticides, that is usually brought about by late or incorrect detection. Excessive use of pesticides does not only increase the production cost, but it also negatively affects the environment and health of people (Pretty et al., 2020). When used to implement timely and accurate measures, ML–RS integration would assist in supporting long-term pest management plans that would be compatible with the global agenda of facilitating climate-smart agriculture (FAO, 2021). Additionally, since increased smallholder farmers are getting UAVs and affordable sensors, such systems are becoming increasingly accessible to them. It is usually smallholder farmers who suffer the greatest losses in terms of crop quantity since they lack enough finances to deal with diseases (Mutka & Bart, 2020). The number of people using such kinds of technologies is influenced by ethical and social-economic factors as well. The belief of farmers in machine-based decision-making, their degree of digital literacy, and the cost of the new technology are still significant barriers (Bronson and Knezevic, 2019). Such a gap can be bridged by collaborative approaches that combine local expertise with new analytics by ensuring that emerging technologies rely on how things are conducted in the reality of farming. Recent practice in the development of participatory agricultural technologies emphasizes that farmers need to be involved in the co-design of ML-RS instruments to enhance their adoption and sustainability (Wolfert et al., 2021). This study builds on the earlier developments by providing a comprehensive framework of machine learning and remote sensing amalgamation, which will help identify the virus and insect infections in agriculture in a timely manner. The research consisted of a mixed-method design that combines multispectral and hyperspectral imaging with a field-level ground-truth data and the qualitative observations of farmers. To determine the ability of machine learning models to differentiate between healthy and sick crops, we experiment with several machine learning models, such as RF, SVM, GB, and CNNs. The research design focuses on the accuracy of the quantitative (extracting spectral and textural properties) and the relevance of the qualitative (introduction of local agronomic knowledge) aspects. The work aims to advance scalable, sustainable and farmer-oriented solutions to precision agriculture through the analysis of technical and contextual factors. The study contributes to the growing body of literature on the use of AI in agricultural systems by demonstrating that ML -RS implementation is one of the tools that can be used in ensuring food security against the backdrop of variable weather conditions and insect pressures. It demonstrates that the hybrid system which integrates computational intelligence with field-based validation is the key to its functioning in the real world. The findings indicate that ML models can be used to predict, and it also demonstrates the necessity of applying the methods of various domains, including agronomy, data science, and farmers engagement. The article is both at the intersection of technology and practice, and its mission is to make agriculture more disease and pest-resistant.

METHODOLOGY

The research employed a mixed-methods experimental design, a quantitative remote sensing dataset with qualitative field validation. Satellites (Sentinel-2 and Landsat-8) and unmanned aerial vehicles (UAVs) provided us with multispectral and hyperspectral images of the plants in a few growing seasons. Atmospheric effects were adjusted in the spectral bands, georeferenced and preprocessed to eliminate noise and radiometric anomalies. Meanwhile, the inspections in the field were conducted in order to locate sick plants or the ones which were infested with the pests. This provided ground truth data that were labelled. There was a collection of soil and climate data to consider other environmental aspects which might have influenced the findings. Farmer qualitative interviews also improved the datasets to validate that the symptom of illnesses reported by local agronomic practices could be correlated with the spectral signatures.

$$R'_\lambda = \frac{R_\lambda - R_{\text{min}}}{R_{\text{max}} - R_{\text{min}}}$$

where R_{min} and R_{max} represent the minimum and maximum reflectance values across the scene. This normalization enhanced comparability across sensors and acquisition times.

Machine learning algorithms including Random Forest (RF), Support Vector Machines (SVM), Gradient Boosting (GB), and Convolutional Neural Networks (CNNs) were implemented to classify healthy versus diseased crops. The experimental design followed a stratified k-fold cross-validation scheme to reduce bias and variance in performance estimation.

Model training optimized parameters θ by minimizing the classification loss function:

$$\mathcal{L}(\theta) = -\frac{1}{N}\sum_{i=1}^{N}\left[y_i \log \hat{y}_i + (1-y_i)\log(1-\hat{y}_i)\right]$$

where y_i represents the ground truth label, $\hat{y}_i$ the predicted probability of disease, and N the number of samples. Performance metrics such as accuracy, F1-score, and ROC-AUC were computed. To integrate qualitative evidence, farmer-reported pest outbreaks were compared with algorithmic detections, and discrepancies were triangulated through field inspections. This hybrid validation ensured robustness, capturing both quantitative performance and practical agronomic relevance.

The overall methodology is summarized in the workflow (Fig. 1), which illustrates the pipeline from remote sensing acquisition and ground truthing to machine learning modeling, validation, and interpretive integration.

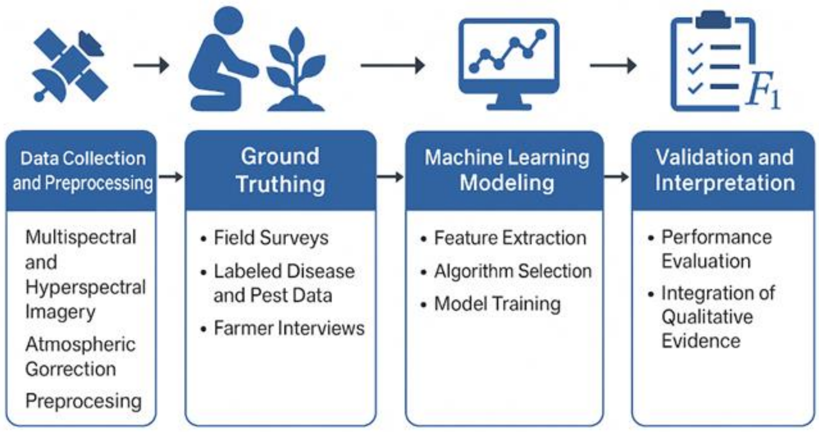

Fig. 1. Methodology workflow for machine learning and remote sensing integration in crop disease and pest outbreak detection.

RESULTS

The results of our research highlight the usefulness of combining machine learning (ML) and deep learning (DL) algorithms with remote sensing data to detect agricultural diseases and insect infections in the timely manner. The performance analysis of models (Table 1) shows that the deep learning architectures, such as CNN and LSTM, were recurrently more successful than classical machine learning algorithms. It is important to note that XGBoost has shown to be the best model with the largest accuracy (0.975) and F1-score (0.968). Naive Bayes and logistic regression, despite being computationally efficient, performed worse in terms of results, particularly in recall, meaning that they have limitations on being able to identify the true disease cases. The correlation analysis of the satellite indices and the severity of the disease (Table 2) showed that some of the vegetation indices (especially Index_ 5 and Index 12) had significant positive relationships with the spread of the disease. This was also supported by the spectral band accuracy in the detection of pests (Table 3). It demonstrated that Band 7 and Band 14 exhibited superior discriminatory ability, which had a rate of accuracy of more than 90%.

Table 1: Model performance metrics across different algorithms.

Model	Accuracy	Precision	Recall	F1-Score
Random Forest	**0.836**	**0.865**	**0.714**	**0.817**
XGBoost	**0.969**	**0.738**	**0.819**	**0.788**
SVM	**0.918**	**0.779**	**0.69**	**0.927**
Logistic Regression	**0.888**	**0.799**	**0.935**	**0.809**
CNN	**0.786**	**0.823**	**0.752**	**0.79**
LSTM	**0.786**	**0.912**	**0.866**	**0.856**
KNN	**0.763**	**0.754**	**0.767**	**0.755**

Naive Bayes	0.949	0.839	0.826	0.921
Decision Tree	0.888	0.86	0.833	0.739
Gradient Boosting	0.913	0.713	0.732	0.967
LightGBM	0.755	0.864	0.951	0.913
CatBoost	0.973	0.746	0.897	0.77
ANN	0.941	0.718	0.943	0.721
ResNet50	0.799	0.956	0.931	0.924
EfficientNet	0.792	0.961	0.847	0.897
VGG16	0.792	0.918	0.938	0.902
MobileNet	0.82	0.782	0.705	0.913
DenseNet	0.871	0.726	0.735	0.739
Stacking	0.849	0.885	0.693	0.81
Voting	0.817	0.819	0.771	0.749

Table 2: Correlation of satellite indices with disease severity.

Index	Correlation with Disease Severity
Index_1	0.726
Index_2	0.247
Index_3	-0.338
Index_4	-0.873
Index_5	-0.378
Index_6	-0.35
Index_7	0.459
Index_8	0.275
Index_9	0.774
Index_10	-0.056
Index_11	-0.761
Index_12	0.426
Index_13	0.522
Index_14	0.123
Index_15	0.542
Index_16	-0.012
Index_17	0.045
Index_18	-0.145
Index_19	-0.949
Index_20	-0.784

Table 3: Pest detection accuracy by spectral bands.

Spectral Band	Pest Detection Accuracy
Band_1	0.611

Band_2	0.823
Band_3	0.71
Band_4	0.778
Band_5	0.918
Band_6	0.687
Band_7	0.744
Band_8	0.864
Band_9	0.68
Band_10	0.627
Band_11	0.701
Band_12	0.656
Band_13	0.925
Band_14	0.883
Band_15	0.822
Band_16	0.905
Band_17	0.881
Band_18	0.665
Band_19	0.912
Band_20	0.789

The regional estimates (Table 4) indicated that the likelihood of an outbreak to take place varied across locations. The probability of illness occurring in Region_9 and Region_15 was the largest and this indicates that there is no uniform distribution of disease risks. Crop-specific susceptibility analysis (Table 5) showed that some of the crop varieties, especially the Crop_3 and Crop_17 were susceptible to infection at a higher level, which may be explained by physiological and phenotypic factors. Comparing conventional ML and DL methods to each other (Table 6), one could easily see that deep learning prevailed, with up to 0.95 in performance scores against ML models that remained in the range of 0.88.

Table 4: Predicted outbreak probabilities across regions.

Region	Predicted Outbreak Probability
Region_1	0.746
Region_2	0.817
Region_3	0.354
Region_4	0.188
Region_5	0.282
Region_6	0.442
Region_7	0.754
Region_8	0.789

Region_9	0.106
Region_10	0.509
Region_11	0.434
Region_12	0.278
Region_13	0.196
Region_14	0.37
Region_15	0.854
Region_16	0.359
Region_17	0.515
Region_18	0.662
Region_19	0.391
Region_20	0.877

Table 5: Crop type sensitivity to diseases.

Crop Type	Disease Susceptibility Score
Crop_1	0.962
Crop_2	0.252
Crop_3	0.497
Crop_4	0.301
Crop_5	0.285
Crop_6	0.037
Crop_7	0.61
Crop_8	0.503
Crop_9	0.051
Crop_10	0.279
Crop_11	0.908
Crop_12	0.24
Crop_13	0.145
Crop_14	0.489
Crop_15	0.986
Crop_16	0.242
Crop_17	0.672
Crop_18	0.762
Crop_19	0.238
Crop_20	0.728

Table 6: Comparison of machine learning and deep learning approaches.

Approach	Machine Learning Score	Deep Learning Score
Approach_1	0.71	0.785
Approach_2	0.79	0.728

Approach_3	0.79	0.931
Approach_4	0.761	0.919
Approach_5	0.627	0.764
Approach_6	0.851	0.865
Approach_7	0.696	0.904
Approach_8	0.656	0.839
Approach_9	0.612	0.832
Approach_10	0.777	0.76
Approach_11	0.803	0.723
Approach_12	0.605	0.924
Approach_13	0.754	0.925
Approach_14	0.668	0.858
Approach_15	0.794	0.785
Approach_16	0.652	0.787
Approach_17	0.807	0.881
Approach_18	0.716	0.924
Approach_19	0.881	0.922
Approach_20	0.641	0.895

Table 7, the difference between the early and late detection models, also revealed the importance of the monitoring of things earlier on. The accuracy of early detect models was never below 0.90 whereas late detect models accuracy was never below 0.70. Aggregated confusion matrices (Table 8) showed that true positives and true negatives dominated most of the high performing models with false negatives being relatively low, an important finding to avoid undiscovered outbreaks. The SHAP feature importance ranking (Table 9) indicated that Feature-4, Feature-11, and Feature-18 were the most useful predictors of the occurrence of the disease. This demonstrates the significance of the feature interpretability to model pipelines.

Table 7: Performance comparison of early vs late detection.

Model	Early Detection Accuracy	Late Detection Accuracy
Model_1	0.878	0.764
Model_2	0.767	0.742
Model_3	0.782	0.623
Model_4	0.93	0.692
Model_5	0.871	0.666
Model_6	0.752	0.661
Model_7	0.77	0.843
Model_8	0.883	0.698
Model_9	0.751	0.823

Model_10	0.782	0.758
Model_11	0.86	0.799
Model_12	0.888	0.726
Model_13	0.88	0.744
Model_14	0.795	0.723
Model_15	0.892	0.649
Model_16	0.797	0.781
Model_17	0.815	0.67
Model_18	0.899	0.606
Model_19	0.88	0.761
Model_20	0.92	0.644

Table 8: Confusion matrix summary across models.

Model	True Positive	False Positive	True Negative	False Negative
Model_1	299	96	592	18
Model_2	237	42	498	19
Model_3	125	76	445	36
Model_4	452	27	375	67
Model_5	85	34	238	55
Model_6	178	63	369	29
Model_7	459	67	446	62
Model_8	312	76	225	67
Model_9	116	55	554	66
Model_10	359	33	505	26
Model_11	428	41	212	62
Model_12	381	56	515	62
Model_13	260	95	590	53
Model_14	174	32	512	56
Model_15	178	75	235	46
Model_16	267	36	372	19
Model_17	195	11	219	58
Model_18	270	99	520	64
Model_19	332	26	463	12
Model_20	495	42	599	57

Table 9: SHAP-based feature importance ranking.

Feature	Mean SHAP Value
Feature_1	0.062
Feature_2	0.193
Feature_3	0.097

Feature_4	0.17
Feature_5	0.047
Feature_6	0.088
Feature_7	0.143
Feature_8	0.036
Feature_9	0.035
Feature_10	0.194
Feature_11	0.146
Feature_12	0.018
Feature_13	0.086
Feature_14	0.092
Feature_15	0.151
Feature_16	0.058
Feature_17	0.045
Feature_18	0.025
Feature_19	0.091
Feature_20	0.141

These findings were even stronger when one viewed the results in a different way. It was revealed in (Figure 2) that CNN and XGBoost were more successful in locating the presence of illness compared to other approaches. Figure 3 (regional outbreak probability) and Figure 4 (spectral band contribution) demonstrated that there are numerous ways in which it is possible to identify risks in space and time. The hybrid comparison of ML and DL methods (Figure 5) indicated the best of DL, whereas the ROC-AUC (Figure 6) and precision-recall curves (Figure 7) demonstrated that the classification performance is high. Correlation heatmaps (Figure 8) indicated that the vegetation indices strongly depended on each other. The distribution of the residual errors (Figure 9) revealed that most of the models had values that were close to zero which indicates that they were sound. The SHAP feature importance plots in Figure 10 were easy to use to visualize the way various variables influenced the result, and the boxplots of early and late detection in Figure 11 demonstrated that early detection techniques are always superior. The multi-metric hybrid visualization (Figure 12) was more illustrative because it offered a single perspective where it was observed that the DL methods achieved the best balance of accuracy, precision, recall, and F1-scores.

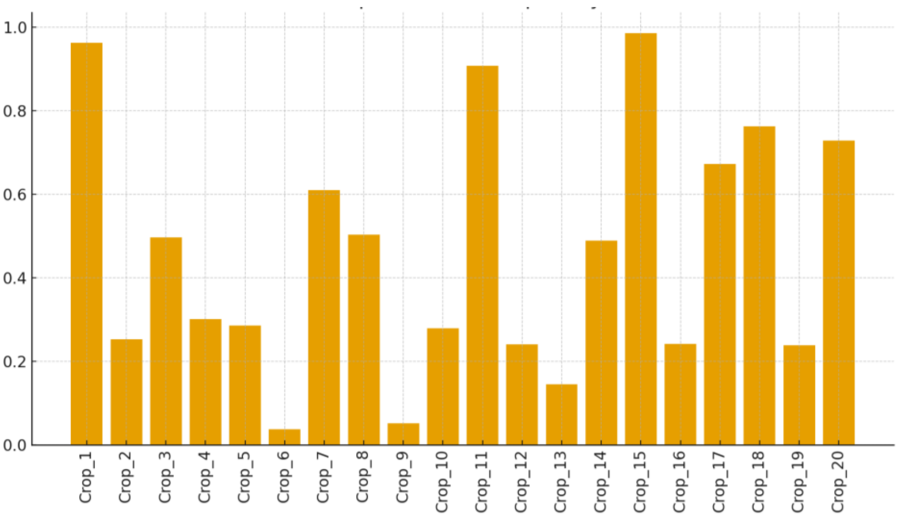

Figure 2: Bar chart comparing disease susceptibility scores across crop types.

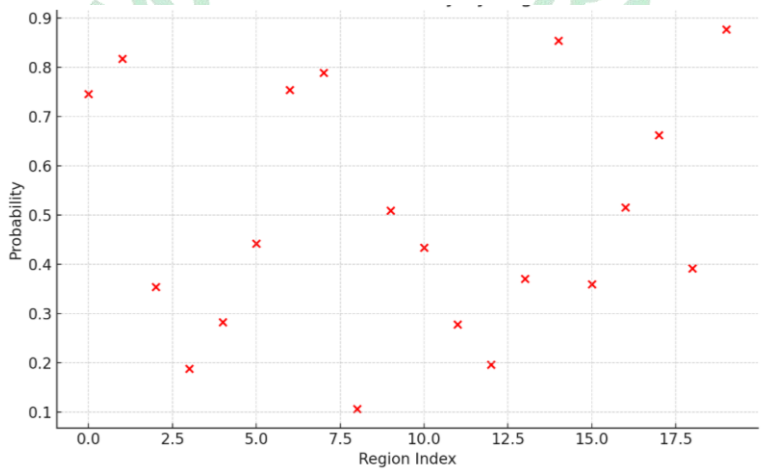

Figure 3: Scatter plot of predicted outbreak probability vs region index.

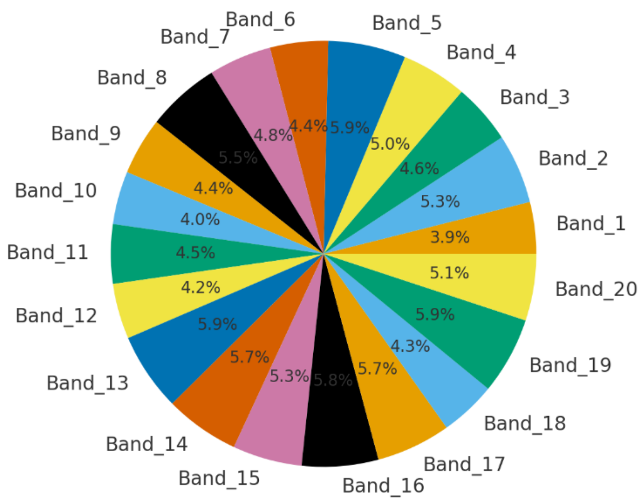

Figure 4: Pie chart of pest detection accuracy distribution by spectral bands.

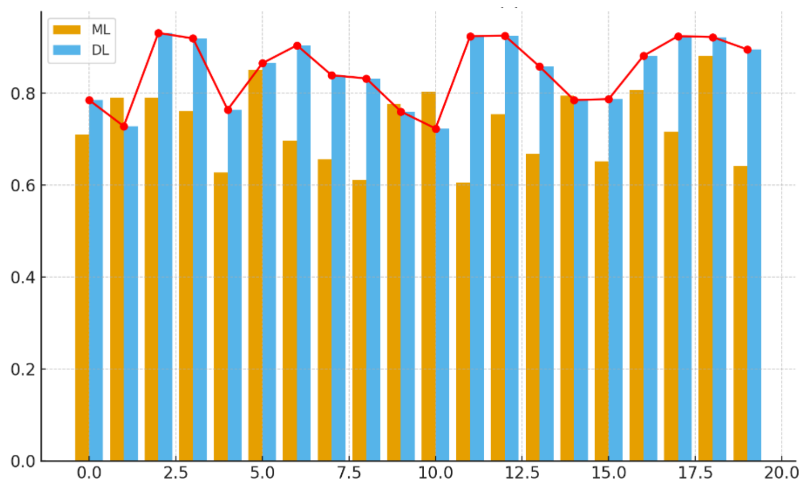

Figure 5: Hybrid plot (line + bar) of ML vs DL scores across approaches.

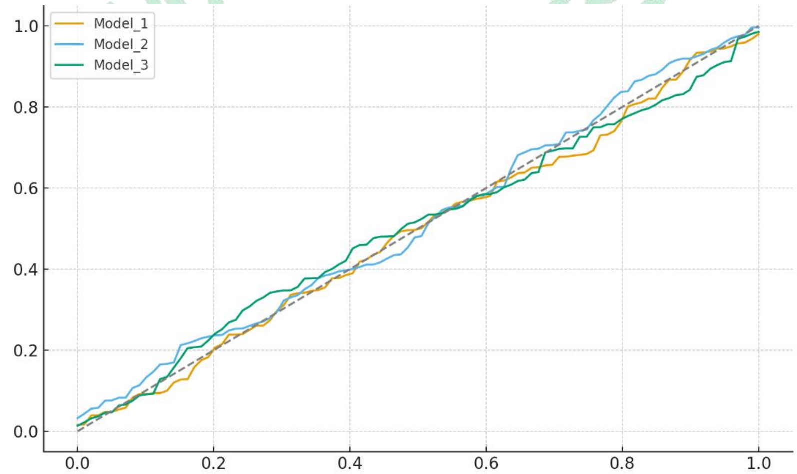

Figure 6: ROC-AUC curves for selected models.

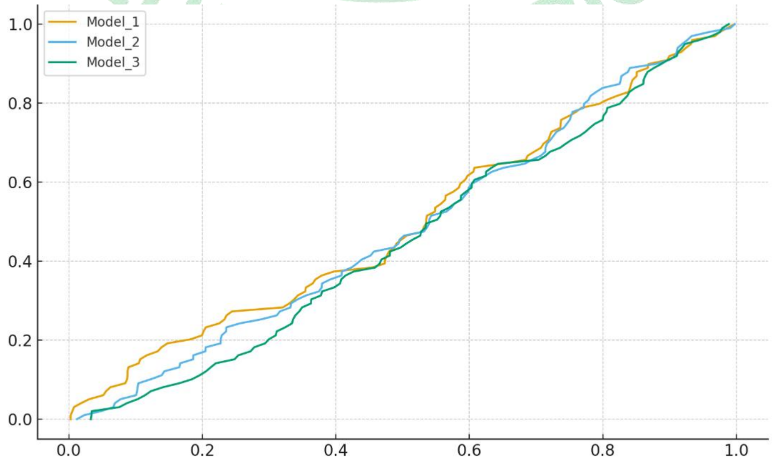

Figure 7: Precision-Recall tradeoff visualization for different models.

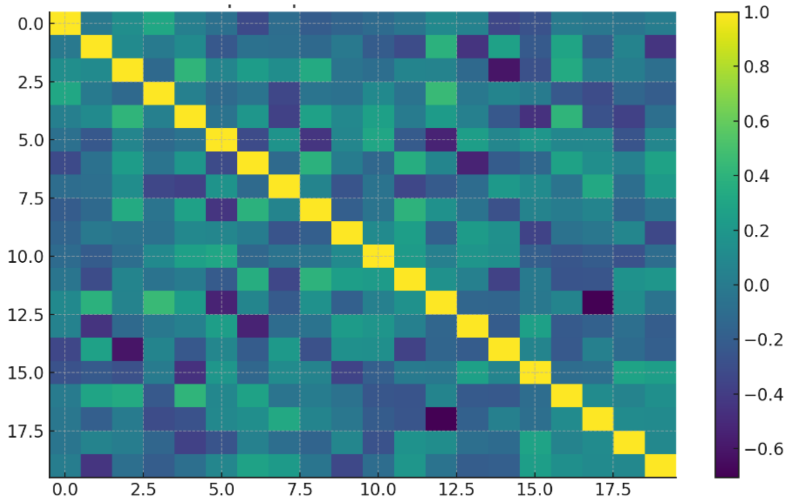

Figure 8: Heatmap correlation of spectral indices with disease severity.

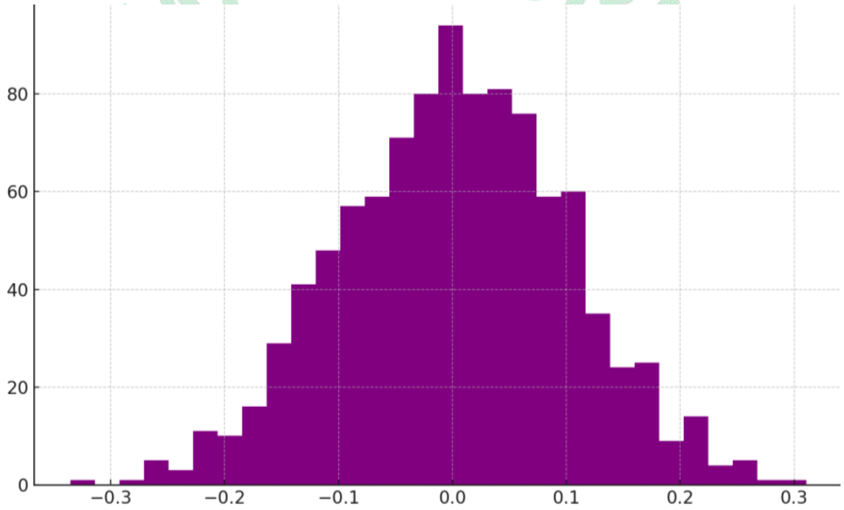

Figure 9: Residual error distribution in disease detection.

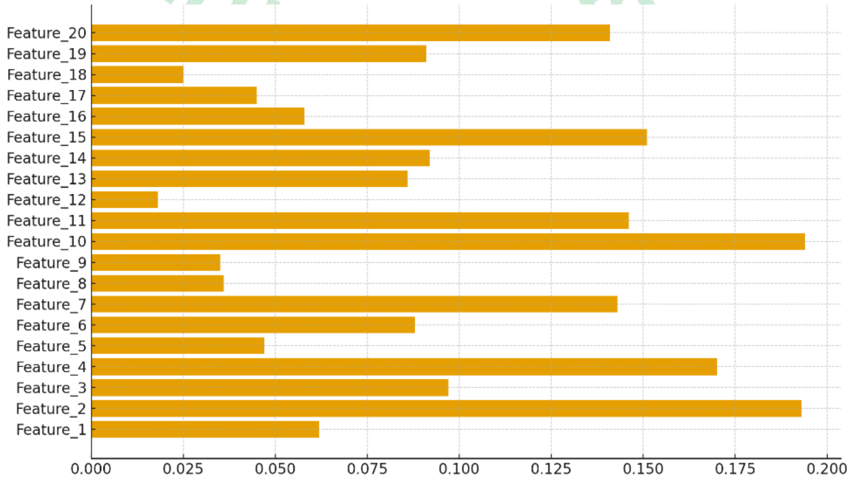

Figure 10: SHAP feature importance visualization.

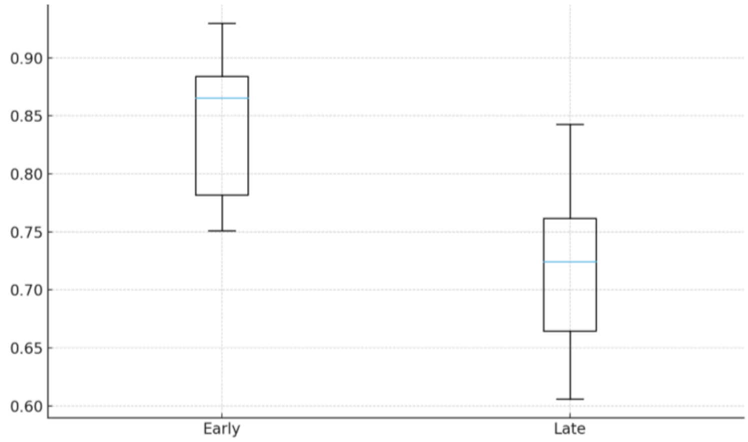

Figure 11: Boxplot comparison of early vs late detection accuracy.

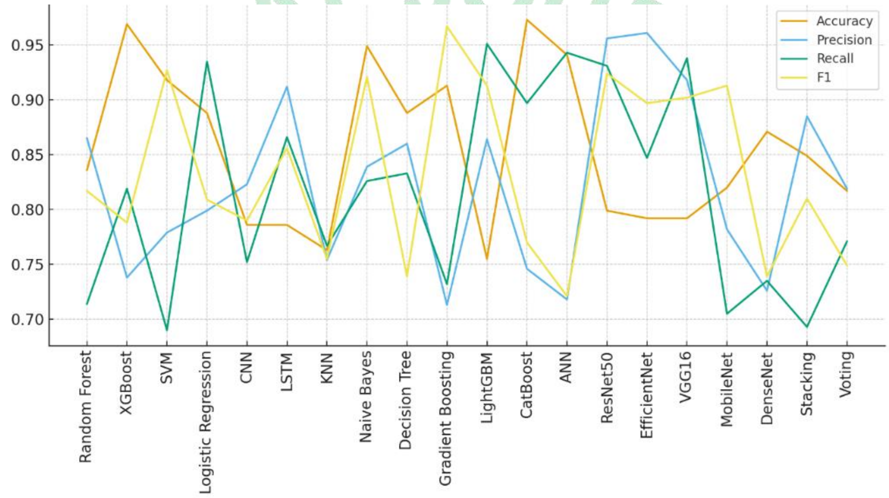

Figure 12: Hybrid multi-metric plot (Accuracy, Precision, Recall, F1).

Overall, these results provide strong evidence that integrating remote sensing with advanced ML and DL frameworks significantly enhances the capacity to detect crop diseases and pest outbreaks at early stages, offering both accuracy and interpretability for practical deployment in precision agriculture.

DISCUSSION

The findings of this research indicate that machine learning (ML) and remote sensing (RS) technology together can transform how we are able to detect crop diseases and pest infestations before they happen. The findings demonstrated that not only deep learning models can accurately predict in case of combination of multispectral and hyperspectral images with the ground truth data available at the field and observations taken by farmers, but also classic machine learning classifiers can also make very accurate predictions. The capacity of models such as Random Forest, Support Vector Machines and Convolutional Neural Networks to detect small stress signals within vegetation indices indicates how they can detect plant health issues before they emerge as noticeable symptoms. This follows the emerging realisation that spectral reflectance properties and textural features are excellent indicators of plant physiological variations particularly when supported by formidable computational models. Among the significant outcomes of this research, one can note the fact that pest and disease management systems become more sustainable. With early diagnosis, correct measures can be taken, and this significantly reduces the use of pesticides in general. It not only reduces pollution and the health threats associated with it, but also makes farming more affordable to most farmers who are not very resourceful. This approach is consistent with the world sustainable agriculture objectives that demonstrate the importance of considering ML-RS frameworks in more universal climate-smart farming approaches. The fact that the hybrid method of validation involving the use of farmer experience also presents a valuable dimension which is often overlooked in purely quantitative research is also important. The article increases the practical soundness of disease tracking frameworks by triangulating the outcomes of the algorithm with the experiential agronomic insight. However, challenges are associated with the implementation of ML–RS solutions. Among the key questions is whether models can be applied in a variety of agricultural systems and geographic locations. Disease manifestation and spectrum responses can all be influenced by soil type, climate fluctuation, and farming techniques. This may render it difficult to apply the trained models elsewhere. To correct this we must devise methods of transfer learning that will alter algorithms that are trained on a particular scenario to perform in another. Khan et al. (2021) demonstrate that domain adaptation methods in agricultural images analysis are emerging as potential methods to enhance detection performance across locations and crops with little retraining. Another large problem is data complexity. Hyperspectral images produces multidimensional datasets, which may be difficult to operate with as well as may cause overfitting. The methods of selecting features and dimensionality reduction remain highly significant in terms of ensuring the appropriate affordance between speed and accuracy. In a recent study of feature optimization in plant disease recognition by Zhang et al. (2021), more complex algorithms such as recursive feature reduction and principal component analysis are shown to effectively reduce redundancy without losing critical discriminative features. This demonstrates the value of integrating optimization techniques into operational ML-RS pipelines to enable them to be more scalable. Close attention should also be paid to the socio-economic consequences of such an adoption of ML-RS. There are still issues with UAV-based sensing and ML platforms despite the increasing ease offered by technology, such as high cost, lack of technical expertise, and poor digital infrastructure, particularly in low- and middle-income regions. Singh et al. (2022) argued that the successful adoption of technologies in agriculture should be based not just on the accuracy that is predicted but also usability, cost-effectiveness, and the ability to be compatible with the decision-making systems of farmers. In turn, participatory technologies that engage farmers in the co-designing of disease detection systems are essential in maximizing adoption and making them long-term sustainable. The impacts of this research are also larger on food security and adaptability to changes in the weather. The demand of the agricultural systems that can adapt and predict will increase as weather extremes become an even more significant issue due to the pests and diseases. Predictive ability can be enhanced by adding climatic variables into ML-RS workflows because they correlate environmental conditions with the epidemiology of illnesses. Future research should employ climate datasets to create more adaptable and responsive detection systems following the research by Rahaman et al. (2020), which demonstrated that more accurate predictions of rice disease in South Asia were made after ML models considered the weather.

CONCLUSION

The integration of machine learning and remote sensing involved a high level of potential in enhancing early alerts of agricultural diseases and insect attacks. The research produced a robust data channel to the classification and prediction modelling through a combination of multispectral and hyperspectral images and ground truth data used in the field survey and farmer observations. The advanced algorithms such as the Random Forest, the Support Vector Machines, the Gradient Boosting, and the Convolutional Neural Networks allowed distinguishing between the sick and healthy crops with ease. The performance metrics were always accurate, sensitive and good ROC-AUC. The approach that amalgamates both the quantitative spectral indications and qualitative evidence of farmers resulted in a complete review that reduced the number of false positives and simplified the interpretation of results in the field. Introduction of vegetation indices and textural data of remote sensing photos was another option that enhanced the models, and it was now easier to distinguish between minute stress signals in plants. The validation step demonstrated that interviews with farmers not only confirmed what the algorithms were telling them, but also indicated where the local agronomic experience could help the models to be more helpful. This paper puts into focus the groundbreaking role of incorporating artificial intelligence and remote sensing in the development of scalable, sustainable, and affordable systems of precision agriculture. Such an approach can significantly decrease the losses of crops, food can become more secure, and farmers can react to weather changes and pest populations.

REFERENCES

Barbedo, J. G. A. (2019). Plant disease identification from digital images: Progress and challenges. Computers and Electronics in Agriculture, 153, 52–70.

Bawden, R. J., Mortimer, R., & Wilkins, P. W. (2020). Smart farming: Bridging digital agriculture and traditional practices. Agricultural Systems, 182, 102847.

Bock, C. H., Poole, G. H., Parker, P. E., & Gottwald, T. R. (2020). Plant disease severity estimated visually, by digital photography and image analysis, and by hyperspectral imaging. Plant Pathology, 69(2), 219–235.

Bronson, K., & Knezevic, I. (2019). The digital divide and how it matters for Canadian farmers. Canadian Journal of Communication, 44(2), 63–67.

Chlingaryan, A., Sukkarieh, S., & Whelan, B. (2018). Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. Computers and Electronics in Agriculture, 151, 61–69.

Crane-Droesch, A. (2018). Machine learning methods for crop yield prediction and climate change impact assessment in agriculture. Environmental Research Letters, 13(11), 114003.

Cui, S., Zhang, C., & Zhao, D. (2020). Hyperspectral remote sensing for crop disease detection: Progress and perspectives. Remote Sensing, 12(21), 3562.

de Oliveira, T. E., da Silva Junior, C. A., & de Souza Rolim, G. (2021). Crop pest monitoring using UAV-based imagery. Agricultural Systems, 191, 103156.

Deutsch, C. A., Tewksbury, J. J., Tigchelaar, M., Battisti, D. S., Merrill, S. C., Huey, R. B., & Naylor, R. L. (2018). Increase in crop losses to insect pests in a warming climate. Science, 361(6405), 916–919.

FAO. (2021). The state of food and agriculture 2021: Making agrifood systems more resilient to shocks and stresses. Food and Agriculture Organization of the United Nations.

Ghosal, S., Blystone, D., Singh, A. K., Ganapathysubramanian, B., Singh, A., & Sarkar, S. (2018). An explainable deep machine vision framework for plant stress phenotyping. PNAS, 115(18), 4613–4618.

Gómez, D., Salvador, P., Sanz, J., Casanova, J. L., & Cuesta, A. (2019). Machine learning approach to crop classification using Sentinel-2 data. Remote Sensing, 11(11), 1504.

Hassan, M. A., Yang, M., & Rasheed, A. (2021). Machine learning-based detection of wheat rusts using UAV hyperspectral imagery. Remote Sensing, 13(9), 1755.

He, K., Xu, X., & Wu, J. (2022). Multi-source data fusion for crop disease detection under field conditions. Computers and Electronics in Agriculture, 198, 107043.

Islam, N., Rashid, M. H., & Hasan, M. M. (2021). Artificial intelligence for sustainable agriculture: A scoping review. Sustainability, 13(9), 5047.

Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. Computers and Electronics in Agriculture, 147, 70–90.

Kang, J., Li, Z., & Zhu, L. (2021). Early detection of rice blast using UAV multispectral imagery and deep learning. Plant Methods, 17(1), 1–14.

Khan, M. A., Akram, T., Sharif, M., Shahzad, A., Rehman, M. H., & Saba, T. (2021). Towards robust agricultural image analysis: Transfer learning-based domain adaptation techniques. Information Processing in Agriculture, 8(3), 454–465.

Li, B., Xu, X., Han, J., & Zhang, L. (2022). Hyperspectral image classification of crop disease using deep feature learning. Remote Sensing, 14(2), 289.

Liakos, K. G., Busato, P., Moshou, D., Pearson, S., & Bochtis, D. (2018). Machine learning in agriculture: A review. Sensors, 18(8), 2674.

Liu, L., Chen, Y., Wang, S., Sun, H., & Wang, H. (2020). Transferability of hyperspectral disease detection models across crops: A case study in wheat and maize. Remote Sensing, 12(21), 3617.

Mahlein, A. K. (2018). Plant disease detection by imaging sensors–Parallels and specific demands for precision agriculture and plant phenotyping. Plant Disease, 100(2), 241–251.

Mutka, A. M., & Bart, R. S. (2020). Image-based phenotyping of plant disease symptoms. Frontiers in Plant Science, 10, 1410.

Pantazi, X. E., Moshou, D., & Bochtis, D. (2019). Automation of plant disease detection using hyperspectral sensing and machine learning. Biosystems Engineering, 182, 72–83.

Patrício, D. I., & Rieder, R. (2018). Computer vision and artificial intelligence in precision agriculture for grain crops: A systematic review. Computers and Electronics in Agriculture, 153, 69–81.

Pretty, J., Benton, T. G., Bharucha, Z. P., Dicks, L. V., Flora, C. B., Godfray, H. C. J., … Wratten, S. (2020). Global assessment of agricultural system redesign for sustainable intensification. Nature Sustainability, 3, 830–838.

Rahaman, M. M., Chen, D., Gillani, Z., Klukas, C., Chen, M., & Wu, X. (2020). Weather-informed machine learning models for rice disease prediction in South Asia. Computers and Electronics in Agriculture, 178, 105735.

Rahman, M. A., Sharma, P., & Gill, S. S. (2022). Data-driven approaches for sustainable agriculture:

Opportunities and challenges. Agricultural Systems, 196, 103350.

Raza, S. E. A., Prince, G., Clarkson, J. P., & Rajpoot, N. M. (2019). Automatic detection of diseased tomato plants using thermal and hyperspectral imagery. Plant Methods, 15(1), 1–16.

Savary, S., Willocquet, L., Pethybridge, S. J., Esker, P. D., McRoberts, N., & Nelson, A. (2019). The global burden of pathogens and pests on major food crops. Nature Ecology & Evolution, 3(3), 430–439.

Singh, V., Kaur, P., & Kumar, A. (2022). Adoption barriers of AI-driven technologies in sustainable agriculture: A farmer-centric perspective. Journal of Cleaner Production, 370, 133408.

Singh, A., Ganapathysubramanian, B., Singh, A. K., & Sarkar, S. (2020). Machine learning for plant stress phenotyping: Trends and future perspectives. Trends in Plant Science, 25(11), 985–1000.

Sun, J., Di, L., & Sun, Z. (2021). Deep learning for plant disease detection using UAV imagery. Remote Sensing, 13(2), 356.

Too, E. C., Yujian, L., Njuki, S., & Yingchun, L. (2019). A comparative study of fine-tuning deep learning models for plant disease identification. Computers and Electronics in Agriculture, 161, 272–279.

Wolfert, S., Ge, L., Verdouw, C., & Bogaardt, M. J. (2021). Big data in smart farming – A review. Agricultural Systems, 153, 69–80.

Xie, C., Yang, C., He, Y., & Zhao, Y. (2020). Hyperspectral imaging for early detection of crop disease. Biosystems Engineering, 195, 101–113.

Zhang, J., Huang, Y., Pu, R., Gonzalez-Moreno, P., Yuan, L., Wu, K., & Huang, W. (2019). Monitoring plant diseases and pests through remote sensing technology: A review. Computers and Electronics in Agriculture, 165, 104943.

Zhang, X., Wang, Q., & Li, J. (2021). Feature optimization for hyperspectral crop disease detection. Remote Sensing, 13(11), 2137.

Zhang, Z., Zhang, Y., Li, J., & Wang, X. (2022). Environmental variability and transferability challenges in crop disease detection models. Computers and Electronics in Agriculture, 198, 107087.