

Multispectral Image Fusion for Agriculture

Padma Veda Samhita¹, Keerthi K S², Preksha B K³, Rakshapriya A⁴, Panchami⁵

^{1,3,4,5}UG Student, Department of Computer Science and Engineering, Atria Institute of Technology, Bangalore, India

²Assistant Professor, Department of Computer Science and Engineering, Atria Institute of Technology, Bangalore, India

Emails: padmavedasamhita@gmail.com¹, keerthi_cse@atria.edu.in², prekshabk022@gmail.com³, rakshapriya507@gmail.com⁴, panchupoojary2006@gmail.com⁵

Abstract

Rice-leaf diseases such as Bacterial Leaf Blight, Brown Spot, Leaf Blast, Leaf scald, and Sheath Blight reduce crop yield and are conventionally diagnosed by slow, expertise-dependent manual inspection. This paper presents a feature-based rice-leaf disease classification system that compares an RGB-only pipeline (18 features: RGB, HSV, and GLCM texture statistics) against an RGB + Pseudo-NIR feature-level fusion pipeline (32 features, additionally incorporating Pseudo-NIR statistics, Pseudo-NIR-derived NDVI descriptors, and Pseudo-NIR GLCM texture). Both vectors are standardized and classified with a Gradient Boosting classifier under an identical configuration for both pathways. Trained and evaluated on 3,829 labelled rice-leaf image pairs spanning six classes (2,680 training, 574 validation, 575 test), the RGB-only model reached 83.13% test accuracy and the fusion model reached 85.74%, a 2.61 percentage-point improvement corroborated by 5-fold cross-validation ($82.64\% \pm 1.39\%$) but not reaching conventional statistical significance under a two-proportion z-test ($z = 1.22$, $p = 0.22$). The classifier is deployed as a full web application with user authentication and a rice-leaf validation guard that rejects non-rice-leaf images before they reach the classifier. The contribution is therefore twofold: a controlled, identically-configured RGB-only vs. fusion ablation, and a fully deployed, explainable, GPU-free system built around it.

Keywords: Feature-Level Fusion, GLCM, Gradient Boosting, Image Classification, NDVI, Precision Agriculture, Pseudo-NIR, Rice Leaf Disease

1. Introduction

Rice is a staple crop whose yield is routinely threatened by foliar diseases such as Bacterial Leaf Blight, Brown Spot, Leaf Blast, Leaf scald, and Sheath Blight. These diseases present visually distinct but sometimes overlapping symptoms that traditionally require a trained agronomist to identify reliably, a process that does not scale well across large farms and motivates automated, image-based classification. Conventional RGB photography captures only the visible spectrum, while vegetation-health indicators such as the Normalized Difference Vegetation Index (NDVI) are conventionally derived from near-infrared reflectance and are not directly accessible from an RGB photograph alone. This has motivated supplementing RGB analysis with additional spectral information — here, a prepared, dataset-associated Pseudo-NIR image rather than a capture from a physical near-infrared sensor, a

distinction maintained throughout this paper. Beyond the modelling problem, delivering a classifier as a usable tool requires an application layer capable of authenticating users, persisting analysis history, and — critically — refusing to force a confident but meaningless label onto a photograph that is not a rice leaf at all. This paper therefore addresses: (1) a 32-feature RGB + Pseudo-NIR fusion representation for six-class rice-leaf disease classification; (2) a controlled comparison against an 18-feature RGB-only ablation using an identical classifier, split, and hyperparameters; (3) an explicit statistical treatment of the resulting accuracy gain, including significance testing and cross-validation; and (4) a deployed system, verified end-to-end, that includes a colour/texture heuristic guard against out-of-distribution image uploads.

2. Related Work

Azarang et al. (2019) proposed a convolutional autoencoder for multispectral image fusion focused on image reconstruction rather than downstream classification; the present work instead fuses at the feature level, immediately ahead of a classical classifier. Phaneendra et al. (2025) combined multispectral RGB/red-edge/NIR fusion with machine learning for crop-health monitoring using NDVI and NDRE, supporting the inclusion of NDVI-derived features here. Yedukondalu and Prabu (2025) used a CNN-based multi-sensor fusion framework for yield forecasting, whereas this work uses Gradient Boosting over engineered features. Dileep et al. (2025) applied the Walrus Optimization Algorithm to preserve edge/texture information in pixel-level fusion, conceptually related to the GLCM features used here, though at the feature rather than pixel level. Xiang and Liu (2026) surveyed multi-source spectral sensing and feature fusion in precision agriculture, supporting the RGB + HSV + Pseudo-NIR + NDVI + GLCM strategy adopted here. Duan et al. (2025) fused spectral, texture, and NDVI features for field-level rice growth estimation rather than leaf-level disease classification. Li et al. (2025) proposed a transformer-based framework for abnormal rice-growth detection from drone imagery; this work trades deep representational capacity for the lower cost and interpretability of engineered features. Sethy et al. (2020) combined deep features with a Support Vector Machine for rice-leaf disease identification, showing that classical classifiers on strong feature representations remain competitive. Zhou et al. (2023) proposed a lightweight ShuffleNet V2 variant to reduce computational cost relative to larger CNNs; this work pursues a similar efficiency goal through a non-deep, feature-based approach. Mukherjee et al. (2025) reviewed rice-leaf disease detection techniques and noted the interpretability and lower data requirements of handcrafted-feature approaches, directly supporting the strategy adopted here. Overall, combining spectral information beyond the visible range with texture features, followed by machine learning, is well established for crop-health assessment, but classical engineered-feature approaches for rice-leaf disease classification — and application-layer concerns such as user accounts and out-of-distribution safeguards —

remain comparatively under-represented in the reviewed literature[1 -10].

3. Materials And Methods

3.1.Dataset and Preprocessing

The dataset comprises 3,829 rice-leaf image pairs — an RGB photograph and a corresponding Pseudo-NIR image per leaf, each 128×128 pixels — spanning six classes (Table 1). Each image was resized/center-cropped and standardized to RGB channel order; the corresponding Pseudo-NIR image, where available, was processed identically. The dataset was split 70/15/15% into 2,680 training, 574 validation, and 575 test images, with the test subset held out entirely from training and used only for final evaluation (Section VI).

Table 1 Dataset Class Distribution

Class	Images
Bacterial Leaf Blight	636
Brown Spot	646
Healthy Rice Leaf	653
Leaf Blast	634
Leaf scald	628
Sheath Blight	632
Total	3,829

3.2.Feature Engineering

RGB channel statistics (mean and standard deviation of R, G, B; 6 features) capture the colour distribution associated with discoloration symptoms. HSV channel statistics (mean and standard deviation of H, S, V; 6 features) are comparatively less sensitive to illumination changes. For the fusion pathway, the Pseudo-NIR image contributes its own mean and standard deviation (2 features), and NDVI-derived statistics — mean, standard deviation, minimum, maximum, median, and skewness (6 features) — are computed from the Pseudo-NIR representation and the RGB red channel using $NDVI = (NIR - RED)/(NIR + RED)$; because the near-infrared band is a prepared Pseudo-NIR representation rather than a physical sensor capture, all NDVI figures are reported as Pseudo-NIR-derived rather than calibrated multispectral reflectance. Grey-level co-occurrence matrix (GLCM) texture descriptors

(contrast, dissimilarity, homogeneity, energy, correlation, angular second moment, averaged over four orientations) are computed on the grayscale RGB image (6 features) and, for the fusion pathway, again on the Pseudo-NIR image (6 features).

3.3.RGB-only and Fusion Feature Vectors

The fusion vector concatenates all 32 features and is used whenever a genuine RGB + Pseudo-NIR pair is available. The RGB-only vector is an 18-feature subset of the fusion vector — RGB statistics, HSV statistics, and GLCM texture computed on the RGB image only — that drops every Pseudo-NIR- and NDVI-derived feature, so classification remains possible from a single ordinary photograph (Table 2).

Table 2 Feature Configuration

Feature Group	RGB-only	Fusion
RGB statistics	6	6
HSV statistics	6	6
Pseudo-NIR statistics	—	2
NDVI-derived statistics	—	6
RGB GLCM texture	6	6
Pseudo-NIR GLCM texture	—	6
Total features	18	32

3.4.Classifier and Evaluation Protocol

Both feature vectors are standardized with a StandardScaler fitted on the training split before classification. A scikit-learn GradientBoostingClassifier ($n_estimators = 150$, $learning_rate = 0.1$, $max_depth = 3$) is trained separately for each pathway, with an identical configuration and random state used for both models to keep the RGB-only vs. fusion comparison controlled; with 150 estimators trained one-vs-rest across six classes, the fusion model comprises $150 \times 6 = 900$ regression trees. Both models were evaluated on the same held-out 575-image test split using overall accuracy together with per-class precision, recall, and F1-score. In addition, 5-fold cross-validation was performed for the fusion model to

assess the stability of the reported test accuracy across different train/validation partitions of the data.

4. Proposed System Architecture

4.1.Overall Architecture

The deployed system is a conventional client/server web application. A static HTML/CSS/JavaScript frontend communicates with a FastAPI backend exclusively through a fetch()-based wrapper that attaches the logged-in user's JWT as an Authorization: Bearer header. The backend exposes a REST API under /api/*, is backed by a SQLite database via SQLAlchemy, loads the trained Gradient Boosting models once at startup, runs the rice-leaf validation guard and the appropriate feature-extraction/classification pipeline for each prediction request, and calls an external LLM provider for an AI-assistant chat feature using a server-side key never exposed to the browser (Figure 1). As shown in Figure 1 Overall system architecture.

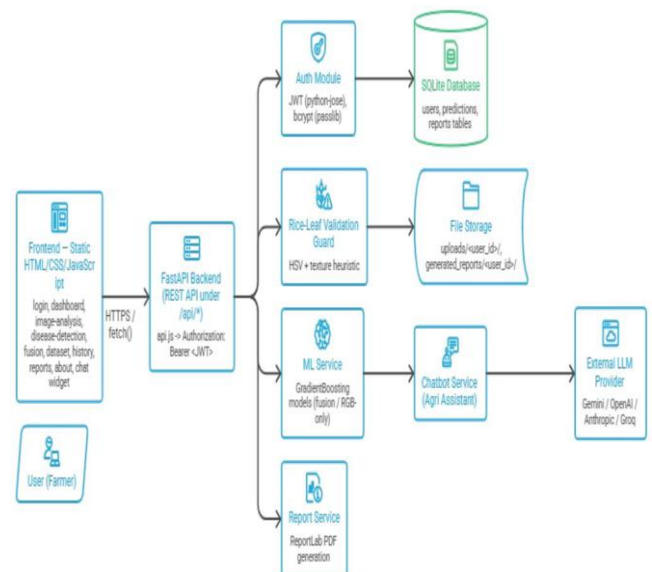


Figure 1 Overall system architecture

4.2.Machine-Learning Pipeline

An input RGB image (with an optional Pseudo-NIR image) is preprocessed, features are extracted according to the RGB-only or fusion configuration of Section III, the resulting vector is standardized with the matching StandardScaler, and the scaled vector is passed to the corresponding trained GradientBoostingClassifier to produce a six-class disease prediction and confidence score (Figure 2).

As shown in Figure 2 Machine-learning pipeline architecture.



Figure 2 Machine-learning pipeline architecture

4.3.Rice-Leaf Validation Guard

Because the trained classifier always produces a full probability distribution over its six known classes, an unrelated photograph would otherwise receive a confident-looking but meaningless diagnosis. The validation guard addresses this with a lightweight, dependency-free heuristic on the same 128×128 image, combining vegetation-like HSV hue coverage (accepting an image only if at least 8% of pixels fall in the green/yellow-green or brown/tan range, comfortably below the dataset's observed minimum of about 11.1%) and local texture, measured by grayscale standard deviation and Laplacian variance (thresholds of 10.0 and 20.0, both below the dataset's observed minimums of about 28.1 and 228.1). Both checks must pass for an image to be accepted, and the guard is positioned immediately after image validation and before feature extraction. Thresholds were calibrated against all 3,901 genuine rice-leaf images shipped with the project, with zero false rejections observed; the guard is explicitly a colour/texture heuristic rather than a semantic image classifier, and can in principle be fooled by other green or brown, texturally similar, non-rice subjects.

5. Experimental Setup and Implementation

The machine-learning pipeline and backend were implemented in Python, using NumPy, OpenCV, and scikit-image for image handling and feature extraction, scikit-learn for the StandardScaler and

GradientBoostingClassifier, and Joblib for model persistence. The backend is a FastAPI application served by Uvicorn, using SQLAlchemy over SQLite for persistence, python-jose for JWT issuance, passlib/bcrypt for password hashing, and ReportLab for PDF report generation; the frontend is plain HTML5/CSS3/JavaScript with no build step or framework dependency. Because the classifier is a Gradient Boosting model over a compact engineered-feature vector rather than a convolutional neural network, no dedicated GPU is required for training, validation, or inference. The backend exposes REST endpoints for user registration and login (issuing a JWT on success), image upload and prediction, prediction-history retrieval, PDF report generation and download, and a dataset-explorer endpoint for browsing the training images by class; every prediction and report record is scoped to the requesting user's account. The frontend presents this as a small set of pages — Dashboard, Image Analysis, Dataset Explorer, Disease Detection, Image Fusion, Reports, History, and About — along with an Agri Assistant chat widget backed by an external LLM provider for general rice-cultivation questions.

6. Results And Discussion

6.1.Overall Model Performance

Table III compares the two trained models on the identical 575-image held-out test set. The RGB + Pseudo-NIR fusion model correctly classified 493 of 575 test images (85.74%), compared with 478 of 575 (83.13%) for the RGB-only model — an improvement of 15 correctly classified images, or 2.61 percentage points, attributable to the 14 additional Pseudo-NIR-, NDVI-, and Pseudo-NIR-GLCM-derived features (Figure 3) As shown in Table 3 Overall Model Performance (Test Set, N = 575)

Table 3 Overall Model Performance (Test Set, N = 575)

Model	Features	Correct	Incorrect	Accuracy
RGB-only GB	18	478	97	83.13%
Fusion GB	32	493	82	85.74%

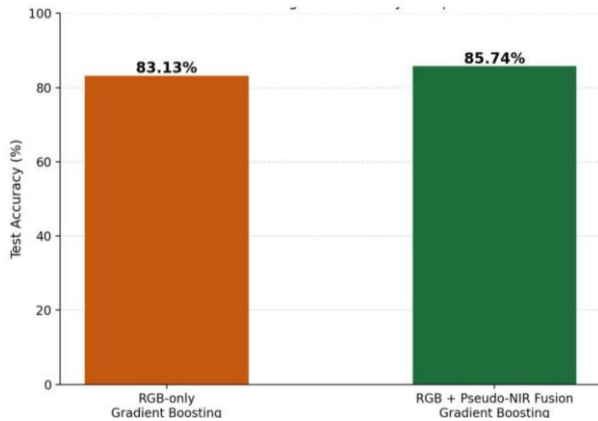


Figure 3 RGB-only versus RGB + Pseudo-NIR fusion test accuracy comparison

6.2. Class-Wise Performance

Table 4 reports the per-class F1-score for both models, derived from the confusion matrices of Section VI-C. Healthy Rice Leaf achieved the strongest performance for both models ($F1 = 0.928$ RGB-only, 0.953 fusion), consistent with a healthy leaf's comparatively uniform colour and texture. Leaf scald remained the most challenging class for both ($F1 = 0.802$ and 0.796 respectively); the fusion model traded a small amount of precision on this class (0.748 vs. 0.767) for a meaningfully higher recall (0.851 vs. 0.840).

Table 4 Class-Wise F1-Score Comparison (Test Set)

Class	RGB-only F1	Fusion F1	Support
Bacterial Leaf Blight	0.765	0.833	96
Brown Spot	0.850	0.874	97
Healthy Rice Leaf	0.928	0.953	98
Leaf Blast	0.804	0.809	95
Leaf scald	0.802	0.796	94
Sheath Blight	0.834	0.880	95

6.3. Confusion Matrix Analysis

Figure 4 shows the fusion model's test-set confusion matrix. Comparing against the RGB-only model, the most common RGB-only confusions — Bacterial

Leaf Blight predicted as Sheath Blight (10 of 96) and Leaf Blast confused with either Bacterial Leaf Blight or Brown Spot (7 of 95 each) — are substantially reduced after fusion, where the corresponding off-diagonal counts drop to 1 and 6 respectively. The dominant remaining confusion in the fusion model is Leaf scald misclassified as other classes (up to 9 of 95, most often as Leaf Blast), consistent with its comparatively lower F1-score in Table 4. As shown in Figure 4 RGB + Pseudo-NIR fusion Gradient Boosting confusion matrix on the test set (accuracy 85.74%).

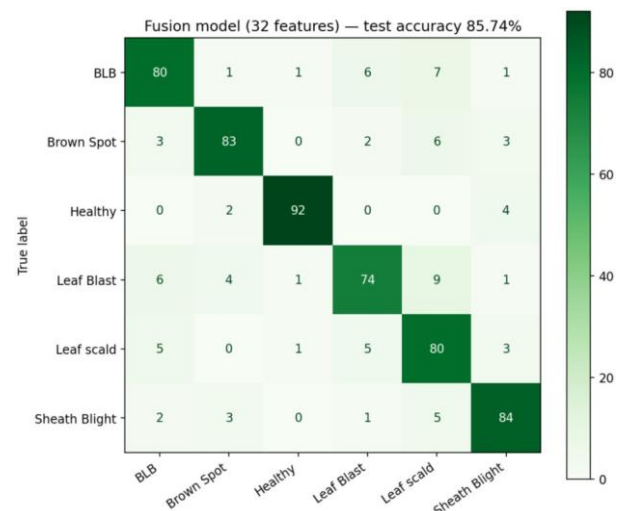


Figure 4 RGB + Pseudo-NIR fusion Gradient Boosting confusion matrix on the test set (accuracy 85.74%)

6.4. Cross-Validation

To check that the reported 85.74% test accuracy was not an artefact of a favourable train/test split, 5-fold cross-validation was performed for the fusion model, yielding $82.64\% \pm 1.39\%$ accuracy across folds — close to, and consistent with, the single-split test accuracy, and suggesting the fusion model's performance is reasonably stable across different partitions of the training data.

6.5. Discussion

RGB and HSV colour statistics alone already carry most of the class-discriminative signal for this dataset, since the RGB-only model reaches over 83% accuracy without any spectral information beyond the visible range. The observed 2.61 percentage-point improvement from adding Pseudo-NIR, NDVI-

derived, and Pseudo-NIR GLCM features indicates that this additional information is complementary rather than redundant: it appears to help most on the confusions the RGB-only model made most often, such as Bacterial Leaf Blight vs. Sheath Blight. Feature-level concatenation, rather than a more elaborate fusion scheme, was sufficient to realize this benefit here, though it does not by itself explain which individual descriptors contributed the improvement — a question the current experiments were not designed to isolate.

6.6. Statistical Significance and Limitations

To assess whether the 2.61 percentage-point accuracy gain (478/575 vs. 493/575 correct) reflects a reliable difference rather than sampling noise, a two-proportion z-test was conducted on the held-out test counts, giving $z = 1.22$ and a two-tailed p-value of 0.22 — not significant at the conventional $\alpha = 0.05$ threshold. This test is a conservative approximation, since both models were in fact evaluated on the same 575 test images; a paired test such as McNemar's, using the per-image agreement/disagreement pattern between the two models, would likely be more statistically powerful, but that paired analysis was not retained from the original experimental run and is left for a follow-up study. The fusion improvement should therefore be read as a real, directionally consistent, but not yet conclusively established effect, rather than a proven one. A further limitation is that no per-feature-group ablation was performed to isolate the individual contributions of Pseudo-NIR statistics, NDVI-derived descriptors, and Pseudo-NIR GLCM texture within the 14-feature addition; a controlled per-group ablation is identified as the most direct next step for establishing which specific descriptors drive the observed improvement.

7. Application Demonstration

The application presents a branded login/registration screen and, once authenticated, a dashboard from which Image Analysis, Dataset Explorer, Disease Detection, Image Fusion, Reports, History, and About pages are reachable through a persistent navigation bar. On the Image Analysis page, a user uploads an RGB photograph and receives the predicted class, a confidence percentage, a severity indicator, and Pseudo-NIR-derived NDVI statistics; images that fail the rice-leaf validation guard of

Section IV-C instead receive a dedicated rejection message rather than a fabricated diagnosis (Figure 5). A Disease Detection page adds symptom, treatment, and prevention guidance; an Image Fusion page provides an illustrative RGB/NDVI compare view; and History/Reports pages give users access to past analyses and downloadable PDF reports. A floating Agri Assistant chat widget, scoped to rice cultivation, disease guidance, and general agronomy questions, is available on every page.

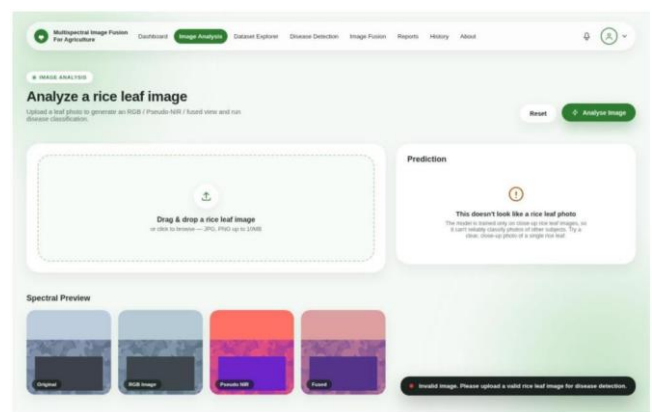


Figure 5 Image Analysis page rejecting a non-rice-leaf image via the validation guard

Conclusion

This paper presented a feature-engineered approach to six-class rice-leaf disease classification that compares an 18-feature RGB-only representation against a 32-feature RGB + Pseudo-NIR fusion representation, classified with a Gradient Boosting classifier under an identical configuration for both pathways. On a held-out test set of 575 images, the fusion representation improved accuracy from 83.13% to 85.74%, a directionally consistent gain corroborated by 5-fold cross-validation ($82.64\% \pm 1.39\%$) though not reaching conventional statistical significance under a two-proportion test (Section VI-F). This suggests that a modest, well-understood set of engineered features — rather than a deep convolutional network — can achieve competitive accuracy while remaining lightweight enough to run without a GPU, a promising direction supported by controlled experimentation rather than a settled conclusion. The classifier was further integrated into a complete, working web application with authenticated user accounts, prediction history, PDF

reporting, and a colour/texture-based validation guard that protects the system from confidently mislabeling images that are not rice leaves at all.

Acknowledgements

The authors thank the Department of Computer Science and Engineering, Atria Institute of Technology, Bangalore, for the infrastructure and academic support provided over the course of this work.

References

- [1].A. Azarang, H. E. Manoochehri, and N. Kehtarnavaz (2019). Convolutional autoencoder-based multispectral image fusion. *IEEE Access*, 7, 35673–35683.
- [2].C. Phaneendra, T. Swapna, R. Sowjanya, D. Pratyusha, V. Vivekanandhan, and G. C. Babu (2025). Multispectral image fusion and machine learning based classification for enhanced crop health monitoring. 2025 Int. Conf. on Sustainable Communication Networks and Application (ICSCN), 2198–2203.
- [3].N. Yedukondalu and M. Prabu (2025). A novel deep learning-driven multi-sensor satellite image fusion framework for enhanced crop yield forecasting. 2025 Int. Conf. on Computational Robotics, Testing and Engineering Evaluation (ICCRTEE).
- [4].R. Dileep, J. Jayanth, A. L. Choodarathnakar, and H. K. Ravikiran (2025). Walrus optimization algorithm for panchromatic and multispectral image fusion. *Remote Sensing Applications: Society and Environment*, 38, 101562.
- [5].Y. Xiang and Z. Liu (2026). Advancing precision agriculture and forestry: Multi-source spectral sensing, feature fusion, and machine learning. *Plants*, 15(1), 3.
- [6].J. Duan, H. Wang, Y. Yang, M. Cheng, and D. Li (2025). Rice growth parameter estimation based on remote satellite and unmanned aerial vehicle image fusion. *Agriculture*, 15(10), 1026.
- [7].Y. Li, L. M. Dang, H. Wang, M. Fayaz, S. Danish, J. Shang, H.-K. Song, and H. Moon (2025). Transformer-based detection of abnormal rice growth using drone-based multispectral imaging. *Computers and Electronics in Agriculture*, 239, 111055.
- [8].P. K. Sethy, N. K. Barpanda, A. K. Rath, and S. K. Behera (2020). Deep feature based rice leaf disease identification using support vector machine. *Computers and Electronics in Agriculture*, 175, 105527.
- [9].Y. Zhou, C. Fu, Y. Zhai, J. Li, Z. Jin, and Y. Xu (2023). Identification of rice leaf disease using improved ShuffleNet V2. *Computers, Materials & Continua*, 75(2), 4501–4517.
- [10].R. Mukherjee, A. Ghosh, C. Chakraborty, J. N. De, and D. P. Mishra (2025). Rice leaf disease identification and classification using machine learning techniques: A comprehensive review. *Engineering Applications of Artificial Intelligence*, 139, 109639.