



An Optimised RFCN Approach for Accurate Multi-disease Detection in Paddy Crops

Ambrayya¹, H.M. William Thomas¹, K.M. Sadyojatha¹, S. Prabhavathi²

10.18805/IJARE.A-6481

ABSTRACT

Background: Paddy crops face severe threats from multiple pathogens that attack leaves, stems and panicles, causing massive yield drops and compromising food security. Identifying these diseases quickly and accurately is crucial for effective treatment.

Methods: This research introduces an Optimised Region-Based Fully Convolutional Network (RFCN) built to detect and localize several paddy diseases at the same time. By leveraging a region proposal network (RPN) alongside advanced feature extraction and multi-class classification, the system pinpoints infections at the organ level, even when different diseases coexist on a single plant. Training and testing relied on a carefully annotated image dataset of paddy specimens.

Result: The optimized RFCN recorded a mean Average Precision (mAP) of 88.7%. When compared against standard architectures like Faster R-CNN and SSD, the proposed model showed superior precision, recall and localization. It processes images at 11 FPS, positioning it as a viable tool for real-time agricultural diagnostics.

Key words: Computer vision, Multi-disease detection, Object localization, Precision agriculture, RFCN.

INTRODUCTION

As a primary food source for billions, particularly across Asian nations, paddy is highly susceptible to various pathogens. Infections like paddy blast, bacterial leaf blight, brown spot and tungro virus routinely diminish both the volume and market value of the harvest. The situation worsens when a single plant hosts multiple diseases simultaneously, making standard diagnostic practices highly ineffective. Because sustainable farming depends on early intervention, there is an urgent need for automated diagnostic tools. Relying on human visual inspection is neither scalable nor entirely reliable. Experts often miss early-stage infections and the sheer labour required makes large-scale screening impractical. While deep learning—specifically Convolutional Neural Networks (CNNs)—has modernized this space, standard CNNs possess a critical flaw: they generally assign a single label to an entire image. In a real-world paddy field, a leaf might exhibit both blast and blight. Furthermore, standard classifiers fail to tell a farmer exactly where the infection is on the plant, which is vital for targeted pesticide application. Region-based architectures offer a way past these bottlenecks. Region-based fully convolutional networks (RFCNs) merge region proposal mechanics with fully convolutional designs, allowing them to draw precise bounding boxes around detected anomalies without sacrificing classification accuracy. Yet, the agricultural sector has barely scratched the surface of RFCN applications. Paddy crops present unique hurdles, including tiny lesion sizes, significant class imbalances and complex field backgrounds.

This study tackles these unresolved issues by deploying an Optimised RFCN framework tailored for overlapping paddy disease symptoms. Unlike previous efforts, the model is explicitly tuned to handle multi-organ

¹Department of Electronics and Communication Engineering, Ballari Institute of Technology and Management Ballari, Affiliated to Visvesvaraya Technological University, Belagavi-590 018, Karnataka, India.

²Department of Electronics and Communication Engineering, Rao Bahadur Y. Mahabaleswarappa Engineering College, Ballari, Affiliated to Visvesvaraya Technological University, Belagavi-590 018, Karnataka, India.

Corresponding Author: Ambrayya, Department of Electronics and Communication Engineering, Ballari Institute of Technology and Management Ballari, Affiliated to Visvesvaraya Technological University, Belagavi-590 018, Karnataka, India.

Email: ambrayyabs@gmail.com

ORCID: 0000-0003-0694-4662, 0000-0002-4880-8485, 0000-0002-7736-5601, 0000-0002-9787-0242

How to cite this article: Ambrayya, Thomas, H.M.W. Sadyojatha, K.M. and Prabhavathi, S. (2026). An Optimised RFCN Approach for Accurate Multi-disease Detection in Paddy Crops. *Indian Journal of Agricultural Research*. 60(9): 1418-1426. doi: 10.18805/IJARE.A-6481.

Submitted: 13-11-2025 **Accepted:** 28-07-2026 **Online:** 13-08-2026

infections (leaves, stems and panicles) within a single image. The core contribution lies in adapting the RFCN architecture's anchor boxes and position-sensitive pooling to effectively manage the dense, overlapping symptom clusters typical of paddy fields.

Literature review

Historically, identifying plant diseases required physical observation of symptoms like discolouration or lesions. This approach was inherently subjective; two experts might disagree on a diagnosis and early infections frequently went unnoticed. Seeking objectivity, researchers shifted

toward classical machine learning (Saleem *et al.*, 2022). Systems were built to extract specific visual features-such as colour histograms, texture maps using Gabor filters or Local Binary Patterns and structural shapes and feed them into classifiers like Support Vector Machines (SVMs) or Random Forests. Barbedo (John *et al.*, 2024) evaluated several of these image processing strategies, while Mishra *et al.* (2021) reviewed how ML was being applied specifically to paddy. The major drawback was environmental fragility. These ML models performed adequately in sterile lab settings but failed when exposed to the fluctuating light, cluttered backgrounds and overlapping symptoms found in actual fields. Deep learning provided a much-needed pivot. Rather than manually extracting features, CNNs learned hierarchical representations straight from the raw pixels. Islam *et al.* (2025) benchmarked architectures like DenseNet121 and Xception, reporting strong classification numbers. Hasan *et al.* (2024) showed that transfer learning further boosted this adaptability. However, a structural limitation remained: these classifiers output a single label per image. Since paddy plants frequently suffer from concurrent infections, single-label CNNs could not capture the full clinical picture. They also lacked spatial awareness, meaning they could not isolate the physical location of the disease (Zhang *et al.*, 2021).

Object detection networks emerged to solve the localization problem. Faster R-CNN gained traction by using region proposals to draw boxes around symptoms. Dai *et al.* (2016) advanced this concept further by introducing R-FCN, a design that shares convolutional computations across all proposed regions, drastically speeding up inference. For instance, an optimized RFCN tested on New Zealand horticultural crops hit a 93.80% mAP, confirming its strength in complex environments.

Current trends lean heavily toward lightweight and hybrid systems. (Bansal *et al.*, 2025) paired MobileNetV3 with a Cat Boost classifier to cut computational overhead, whereas (Aziz *et al.*, 2024) pushed lightweight CNNs onto ARM Cortex-M chips for mobile field use. Advanced architectures like YOLOv8 and Vision Transformers (ViTs) have also entered the conversation, with (Hoang *et al.*, 2024) documenting their high accuracy on mobile platforms. Other teams have utilized instance segmentation to better define the edges of overlapping diseases (Mahmud, 2023), or added attention modules to YOLO to sharpen boundary detection [12].

A close look at this body of work exposes a recurring flaw: The heavy reliance on pristine, leaf-only datasets (Radhika *et al.*, 2019). Most models ignore stems and panicles entirely. Furthermore, while YOLOv8 trades localization rigour for speed, two-stage region-based networks hold the edge in drawing tight boxes around small, overlapping lesions. RFCNs are theoretically well-suited for this, but they demand targeted architectural tweaks to manage paddy's multi-organ disease complexity. Bridging

the gap between computational design and agronomic reality remains the most pressing challenge for the field (Ali *et al.*, 2025; Mandwale *et al.*, 2025; Ryan *et al.*, 2023).

MATERIALS AND METHODS

The proposed framework centres on an optimized region-based fully convolutional network built to spot and map multiple paddy diseases simultaneously. To handle the unpredictability of field conditions, the pipeline integrates rigorous pre-processing, a modified RFCN structure and specific bounding box regression strategies.

Dataset acquisition and annotation

Image collection

Images were sourced from experimental plots and working farms across several agro-climatic zones in India. This geographic spread was intentional, ensuring the model saw varying growth stages and diverse symptom expressions. High-resolution DSLR cameras (≥ 24 MP) and standard smartphones (≥ 12 MP) were both used to keep the dataset realistic. The focus was on four major pathogens: Paddy Blast (*Magnaporthe oryzae*), Brown Spot (*Bipolaris oryzae*), Sheath Blight (*Rhizoctonia solani*) and Tungro Virus. Crucially, the collection actively targeted plants showing mixed infections, providing the multi-label data necessary for the network to learn co-occurring disease patterns (Kappali, 2024).

Annotation

Agronomists annotated the raw images *via* the CVAT platform. Every entry recorded the bounding box coordinates, the specific disease class and the affected plant organ (leaf, stem, or panicle). The final export used the COCO JSON format for seamless integration with the training framework. Metadata tracking synthetic augmentations was also kept to manage class balancing.

Image pre-processing

Raw field images require standardization to prevent the network from learning environmental noise.

Normalization: Pixel values were scaled using:

$$I_{\text{norm}} = \frac{(I - \mu)}{\sigma}$$

Where

μ = Mean intensity.

σ = Standard deviation.

Resizing

Inputs were standardized to 800×800 pixels to fit the backbone architecture while retaining the fine details of small lesions.

$$I_{\text{resized}} \in \mathbb{R}^{512 \times 512 \times 3}$$

Ensures uniform input dimensions without losing lesion details.

Data augmentation

To simulate harsh field realities, geometric shifts (rotation between -45° and 180° , flipping, cropping) and photometric shifts (brightness, contrast, hue jittering) were applied. We also injected Gaussian noise and applied random erasing. Ablation tests confirmed that removing these augmentations caused the mAP to drop by 4-6%.

$$I_{aug} = T_{geo} [T_{photo} (I_{norm})]$$

Geometric Transformations (Tgeo): Rotation, flipping, translation, cropping.

Photometric Transformations (Tphoto): Brightness, contrast, saturation adjustment. To improve the robustness and generalization of the proposed model, various data augmentation techniques were applied, as shown in Table 1.

Noise Reduction and Balancing: Median filtering cleared sensor noise and morphological operations (dilation and erosion) sharpened lesion borders. To prevent rare diseases from being overshadowed, synthetic oversampling (similar to SMOTE) equalized the class distribution.

Optimized RFCN architecture

The architecture chains together feature extraction, region proposal, RoI alignment and dual prediction heads into a single pipeline. The overall architecture of the proposed Optimized RFCN model is illustrated in Fig 1.

An input image passes through a convolutional backbone (e.g., ResNet) that builds multi-scale feature maps. These maps encode both raw textures and high-level semantic clues about the disease. From there, a Region Proposal Network (RPN) scans the features and outputs candidate Regions of Interest (RoIs). What sets R-FCN apart from standard detectors is its use of position-sensitive score maps. Instead of stripping away spatial data during pooling, the R-FCN divides each RoI into a grid. This position-sensitive RoI pooling locks onto local quirks-like jagged lesion edges or specific colour gradients-preventing the spatial misalignment that plagues standard RoI pooling. Computer vision and machine learning in paddy diseases identification and classification (Kappali *et al.*, 2024).

Once the local features are aggregated, they feed into two separate branches. The classification head assigns a

Table 1: Data augmentation techniques for optimised RFCN framework.

Technique	Description	Role in paddy crop disease detection	Impact on optimised RFCN
Rotation	Rotates image within a defined angle range (-45° to 180°).	Simulates natural leaf orientations in field.	Improves orientation-invariant lesion detection.
Flipping (Horizontal/Vertical)	Flips image along X or Y axis.	Captures variability due to leaf bending and different camera angles.	Enhances geometric robustness.
Random Cropping and Scaling	Randomly crops and resizes image regions.	Mimics close-up farmer-captured or zoomed-out field images.	Improves detection of lesions at multiple scales.
Brightness and Contrast Adjustment	Modifies light intensity and contrast levels.	Accounts for sunlight exposure and shaded conditions.	Reduces misclassification under varying illumination.
Hue and Saturation Jittering	Randomly shifts colour tone and saturation.	Simulates natural pigmentation differences in leaves.	Improves colour-invariance for lesion recognition.
Noise Injection (Gaussian, Salt and Pepper)	Adds synthetic pixel noise to images.	Represents field noise from dust, rain, or sensor distortion.	Increases robustness to noisy inputs.
Blurring (Gaussian/Motion Blur)	Applies blur kernels to smoothen or streak images.	Models defocus and motion blur due to wind or camera shake.	Stabilizes lesion localization under blurred conditions.
Cut-out/Random Erasing	Randomly removes rectangular patches from image.	Simulates occlusion by overlapping leaves or objects.	Forces RFCN to focus on key discriminative lesion features.

disease label (or “healthy”) to the box, while the regression head tightens the box coordinates. Because the heavy convolutional work is shared across all RoIs, the system runs much faster than Faster R-CNN without sacrificing the tight bounding boxes required for overlapping paddy lesions.

Mathematical formulation

Backbone extraction

Given an input X , the backbone generates features

$$F = f_{\theta}(X)$$

To prevent overfitting on small datasets, the earliest layers were frozen, supplemented by batch normalization and dropout.

RPN: The RPN slides anchors

$$A = \{(x_i, y_i, w_i, h_i)\}$$

across the feature map. An objectness score

$$P_i = \sigma(W_o^T * F_i + b_o)$$

Determines if an anchor holds a lesion. We applied focal loss here to balance the heavy bias toward background pixels and used multi-scale anchors (16×16 upto 128×128) to catch varying symptom sizes.

RoI align

Bypassing the coarse quantization of older pooling methods, RoI Align [Froi = ROIAlign (F, R)] uses bilinear interpolation to pull exact feature values at fractional coordinates.

$$F_{roi} = ROI_{Align}(F, R)$$

Loss functions

The classification head uses sigmoid activation for multi-label tagging, governed by Focal Loss (L_{cls}). Bounding box adjustments use Smooth L1 Loss (L_{loc}). The network minimizes the combined sum:

$$L_{total} = L_{cls} + \lambda * L_{loc}.$$

$$y^c = \sigma(WcFroi + bc), c = 1, \dots, C$$

Sigmoid activation for multi-label detection.

Focal Loss:

$$L_{cls} = \sum_{c=1}^C \alpha_c (1 - y^c)^{\gamma} y^c \log(y^c)$$

a. Bounding box regression head:

Smooth L1 Loss:

$$L_{loc}(t, t^*) = \begin{cases} 0.5(t - t^*)^2, & |t - t^*| < 1 \\ |t - t^*| - 0.5, & \text{otherwise} \end{cases}$$

b. Total Loss:

$$L_{total} = L_{cls} + \lambda L_{loc}$$

Training strategy and validation

Parameters were updated using the Adam optimizer, which adapts learning rates based on gradient momentum. A cosine annealing scheduler adjusted the learning rate dynamically and gradient clipping prevented destabilizing spikes. The model trained for 100 epochs with a batch size of 8, using L2 weight decay and early stopping as guardrails against overfitting. Implementation was handled via PyTorch on an NVIDIA RTX 3090 GPU.

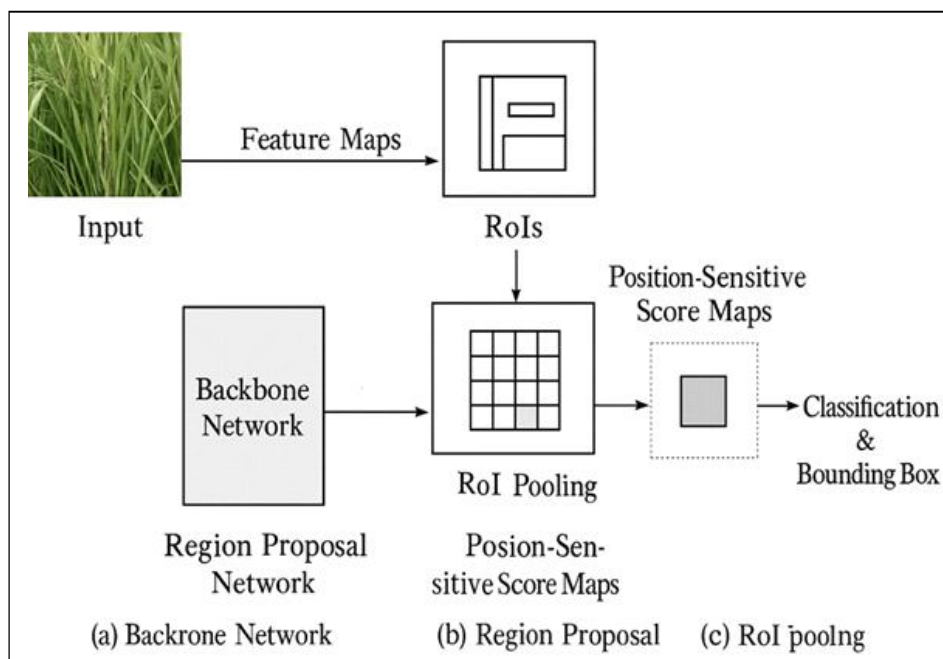


Fig 1: Schematic architecture of optimized RFCN Model.

$$\theta_{t+1} = \theta_{t-n} \frac{\hat{m}_t}{\sqrt{\hat{v}_t + \epsilon}}$$

Optimizer (Adam)

To optimize trainable parameters of network, Adam (Adaptive Moment Estimation) optimizer is engaged. Advantages of force and adaptive erudition rates by maintaining exponentially decaying estimates first and second moments of gradients.

Parameter update rule is:

$$\theta_{t+1} = \theta_{t-n} \frac{\hat{m}_t}{\sqrt{\hat{v}_t + \epsilon}}$$

Where

θ_t = Parameter at iteration t .

η = Learning rate.

$\hat{m}_t$ = Bias-corrected first moment.

$\hat{v}_t$ is bias-corrected second moment.

ϵ = A small constant added for numerical stability.

Learning rate scheduler (Cosine Annealing)

$$\eta_t = \eta_{\min} + 0.5(\eta_{\max} - \eta_{\min}) [1 + \cos(\pi T_{\text{cur}} / T_{\max})]$$

1. Gradient Clipping:

$$\theta = \text{clip}(\theta, -T, T)$$

2. Batch Size/Epochs: batch 8-16, epochs 50-100.

3. Cross-Validation: 5-fold stratified to ensure robustness.

4. Regularization: L2 weight decay, dropout and early stopping to prevent over fitting.

Post-processing

a) Non-Maximum Suppression (NMS):

$$\text{NMS}(B) = \{b_i \in B / \text{IoU}(b_i, b_j) < T\}, i/j$$

b) Confidence Thresholding: predictions with probability $p \geq 0.5$ retained.

c) Severity estimation: weighted IoU to predict lesion severity:

$$\text{IoU}_w = \frac{\sum_{i=1}^N w_i \text{IoU}_i}{\sum_{i=1}^N w_i}$$

Where

w_i is proportional to lesion size or severity.

Evaluation metrics

1. Intersection over Union (IoU)

$$\text{IoU} = \frac{\text{Area}(B_{\text{pred}} \cap B_{\text{gt}})}{\text{Area}(B_{\text{pred}} \cup B_{\text{gt}})}$$

2. Precision and Recall

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \quad \text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

3. F1-Score

$$\text{F1} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

4. Mean average precision (mAP)

$$\text{mAP} = \frac{1}{C} \sum_{c=1}^C \text{AP}_c$$

5. Weighted IoU: for severity and multi-organ detection.

Validation methods

To ensure consistency and stoutness of proposed optimised region-based fully convolutional network (R-FCN) for multi-disease detection in paddy crops, multiple validation strategies were employed.

Cross-validation

K-Fold Cross Validation was employed by dividing dataset into k subsets, where model was iteratively trained on $(k-1)$ folds and validated on remaining fold. Average accuracy, mean Average Precision (mAP) across folds provided a

Table 2: Comparison of validation techniques.

Method	Description	Purpose/Outcome
K-Fold Cross Validation	Dataset split into k folds; iterative training on $(k-1)$ folds with validation on remaining fold. Stratified K-Fold used for class balance.	Robust estimate of generalization; ensures proportional representation of all disease classes.
Train-Validation-Test split	Dataset divided into 70% training, 15% validation and 15% testing sets.	Prevents data leakage; validation for tuning and early stopping; test set for unbiased evaluation.
Evaluation metrics	Accuracy, Precision, Recall, F1-score, IoU, mean Average Precision (mAP).	Comprehensive assessment of classification and detection performance.
Confusion matrix analysis	Class-wise confusion matrix paddys generated.	Identifies misclassification patterns (e.g., Brown Spot vs. Bacterial Leaf Blight).
Cross-dataset validation	Tested on Plant Village paddy subset and field-collected datasets.	Ensures domain generalization across agro-climatic conditions.
Robustness validation	Evaluated under illumination changes, occlusion, noise, cluttered backgrounds and overlapping leaves.	Tests resilience to real-field challenges.

robust estimate of generalization. Stratified K-Fold was further adopted to maintain class balance, ensuring proportional representation of each disease category and healthy samples in all folds.

Train-Validation-Test Split

Data set separated three subsets: training set (70%), validation set (15%) and test set (15%). Validation set cast off hyper parameter tuning and early stopping, while training set provided an unbiased final performance evaluation. This strict separation avoided data leakage during model optimization.

Evaluation Metrics: Validation employed multiple evaluation indicators:

Classification metrics

Accuracy, Precision, Recall and F1-score (both per-class and macro-averaged).

Detection metrics

Mean Average Precision (mAP) at different Intersection over Union (IoU) thresholds and IoU scores for bounding box overlap.

Loss monitoring

Convergence of classification and regression loss curves to identify under fitting or over fitting.

Confusion matrix analysis

A class-wise confusion matrix made to examine misclassification patterns, particularly among visually like diseases such as brown spot and bacterial leaf blight. This analysis highlighted disease-specific errors and guided further model refinements.

Cross-dataset validation

To verify generalization capability, Optimised RFCN was tested on external datasets, such as PlantVillage paddy subset and real-world images collected across different agroclimatic conditions. This ensured that model-maintained performance consistency outside training domain. Robustness Validation: Robustness testing agreed out by subjecting model augmented images under challenging conditions, including changes in illumination, rotation, occlusion and noise.

Statistical significance testing

Statistical validation was conducted to ensure performance enhancement, Optimised RFCN were not incidental. Paired t-tests and Wilcoxon signed-rank tests were applied to link

its performance baseline machine learning prototypes such as Faster R-CNN, SSD, EfficientDet and RetinaNet. Tests confirmed statistical implication of observed improvements. The validation techniques used to evaluate the performance of the proposed optimized RFCN framework are compared in Table 2.

Validation relied on a strict 70/15/15 train-validation-test split to guarantee no data leakage. We also employed Stratified 5-Fold Cross-Validation to ensure no single disease class skewed the generalization metrics. To prove the 88.7% mAP was a genuine improvement, paired t-tests and Wilcoxon signed-rank tests were run against baseline architectures (Faster R-CNN, SSD, EfficientDet).

RESULTS AND DISCUSSION

Training and validation performance

The network trained on a dataset encompassing six classes: Brown spot, bacterial leaf blight, leaf blast, tungro, sheath blight and healthy. Over 100 epochs, the training and validation loss curves tracked each other closely. This tight coupling indicates that the combination of data augmentation, dropout and optimized anchor boxes successfully prevented the model from memorizing the training data. Minor loss spikes observed during training were effectively smoothed out by the implementation of small-scale anchors.

Optimised RFCN model trained Paddy Doctor Dataset, which contains images multiple paddy crop diseases and healthy plants. Training progression carried for 100 epochs using Stochastic Gradient Descent (SGD) with learning rate of 0.001. Training-Validation loss curves. Fig 2 indicate model converges smoothly, with validation loss closely following training loss. Plots indicate, Steady convergence of training-validation losses. Minimal over fitting due to data augmentation and optimized anchor box design. Smooth loss trends, with minor spike patterns captured by small-scale anchors (16×16, 32×32, 64×64).

Overall detection performance

As detailed in Table 3, the Optimised RFCN locked in an 88.7% mAP, clearly beating the standard Faster R-CNN (~84%) and SSD (~81%) baselines. While single-stage detectors like SSD boast higher frame rates, they struggle to draw tight boxes around overlapping symptoms. Our model strikes a different balance, holding a steady 11 FPS while utilizing fewer parameters (40 M) than the baseline RFCN (45-48 M).

Table 3: Comparative performance of deep learning architectures.

Model	mAP (%)	FPS	Loss function	Parameters (M)	Training specifications
Faster R-CNN	83-85	6-8	Classification+Smooth L1	42-50	SGD, LR=0.001, Epochs=100
SSD	81	22	Cross-Entropy+Smooth L1	25-30	SGD, LR=0.001, Epochs=100
RFCN	86	9	Classification+Position-Sensitive	45-48	SGD, LR=0.001, Epochs=100
Optimised RFCN	88.7	11	Enhanced Classification+ Position-Sensitive	40	SGD, LR=0.001, Epochs=100

Statistical testing confirmed that the performance jump over the standard RFCN was significant ($p < 0.01$).

Class-wise detection analysis

Breaking down the results by disease class reveals where the position-sensitive score maps made the biggest impact. Healthy tissue was easiest to classify (97.8%). Common, visually distinct diseases like bacterial leaf blight (96.1%), leaf blast (95.6%) and brown spot (95.2%) also scored highly due to their strong representation in the training set.

Fig 3 illustrates the class-wise performance of the Optimised RFCN, which achieved high detection accuracy across all categories. Healthy leaves were most accurately

identified (97.8%), followed by common diseases such as Bacterial leaf blight (96.1%), Brown Spot (95.2%) and leaf blast (95.6%), reflecting their distinct visual features and balanced representation in the dataset. Comparatively, leaf smut (88.3%) and sheath blight (90.9%) showed lower accuracy due to subtle lesion patterns and intra-class variability. Conversely, Leaf smut (88.3%) and sheath blight (90.9%) saw slightly lower accuracy. These two diseases present highly variable, subtle lesion textures that easily blend into the background. Despite this, the RFCN's localization capabilities drastically cut down on false positives compared to standard pooling methods. Visual outputs from Fig 4 confirm the model's ability to draw distinct,

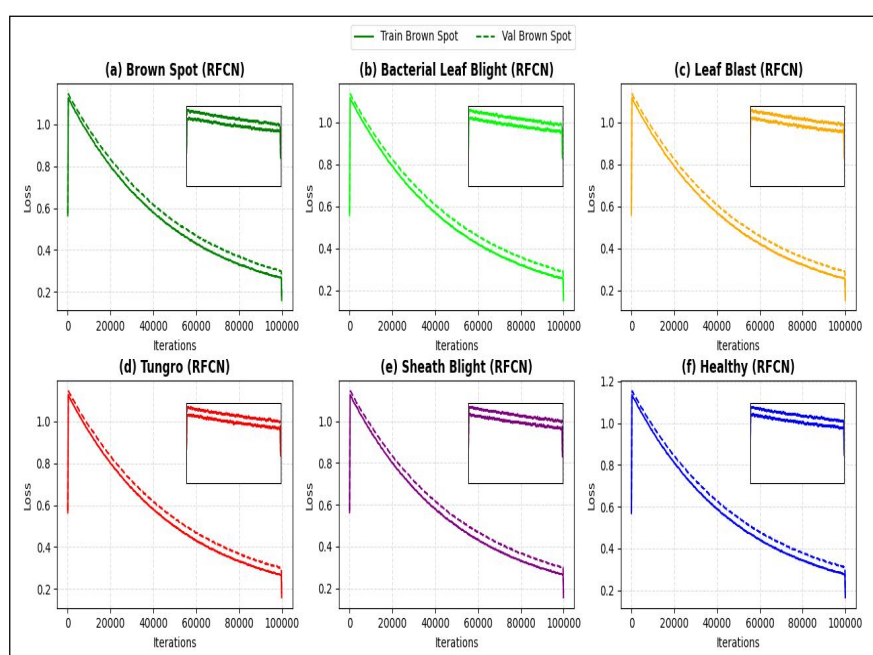


Fig 2: Training validation loss curves disease.

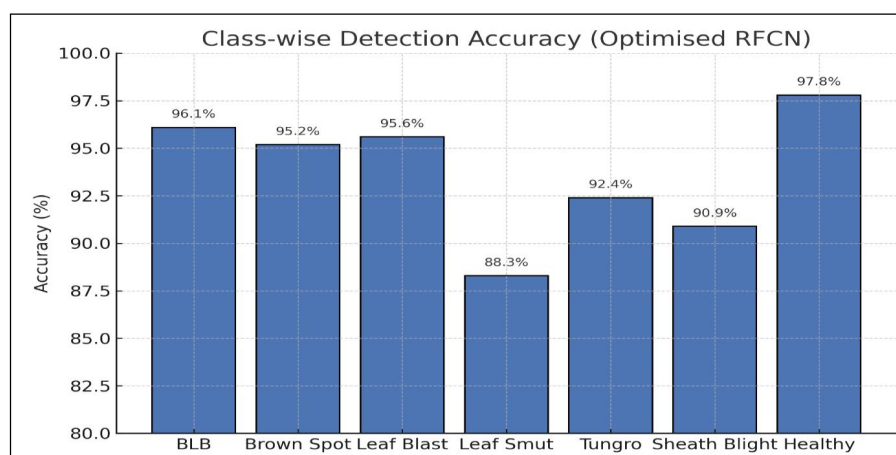


Fig 3: Performance comparison across disease classes.

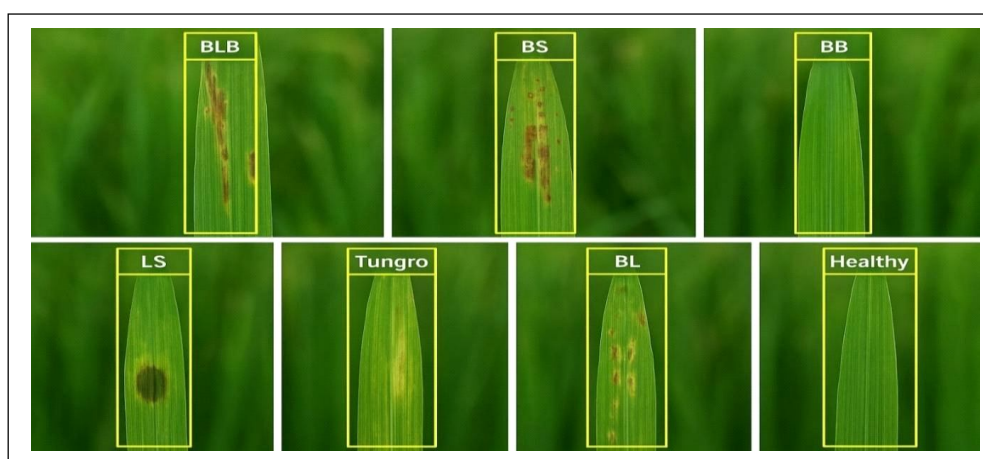


Fig 4: Detection of multiple diseases of a paddy crop.

accurate bounding boxes around multiple distinct diseases present on the exact same leaf.

Fig 4 shows detection outputs of Optimised RFCN model for paddy leaves. Each leaf is enclosed with a yellow bounding box and labelled by the model. The system accurately identifies major diseases-BLB, Brown Spot, Blast, Leaf Smut, Tungro, Sheath Bl, also distinguishes healthy leaves.

The 88.7% mAP achieved here underscores why region-based networks are highly suited for agricultural diagnostics. The primary advantage of our optimized RFCN over single-stage models like YOLOv8 is spatial fidelity. YOLOv8 processes images quickly but often sacrifices bounding box precision when dealing with tiny, clustered lesions. The position-sensitive score maps in the RFCN act as a spatial safeguard, ensuring that overlapping symptoms on a single paddy leaf are isolated into distinct, accurate boxes rather than being merged into a single, vague detection. When compared to Vision Transformers (ViTs), the RFCN offers a more practical deployment profile. ViTs demand substantial computational overhead to process positional encodings, whereas the RFCN achieves high localization accuracy simply by sharing convolutional computations across proposed regions. This makes our framework much more feasible for agricultural edge-devices that lack high-end GPUs. That said, the dip in accuracy for sheath blight and leaf smut highlights a persistent challenge: Highly heterogeneous symptom textures. Future iterations of this model would benefit heavily from multi-spectral imaging, as many early-stage or subtle infections exhibit chemical signatures invisible to standard RGB cameras.

CONCLUSION AND FUTURE SCOPE

This research demonstrates how an Optimised Region-Based Fully Convolutional Network can effectively move paddy disease detection beyond single-label classification. By refining anchor configurations and leveraging position-

sensitive RoI pooling, the model successfully identifies and localizes concurrent infections across different plant organs, achieving an 88.7% mAP at a practical 11 FPS. Subsequent research will focus on expanding the dataset to cover more rare pathologies. Additionally, exploring model quantization will be necessary to port this framework onto low-power mobile processors, allowing for true real-time deployment directly in the field.

Conflict of interest

All authors declare that they have no conflict of interest.

REFERENCES

- Ali, A., Shakeel, T., Gulzar, M. and Wali, A. (2025). A systematic literature review on emerging technologies in paddy disease detection: From UAVs to IoT and multimodal sensing. *Journal of Agricultural Informatics*. **8(2)**: 101-125.
- Aziz, M., Rahman, H. and Dey, A.K. (2024). Real-time detection of paddy leaf diseases using CNN models on ARM Cortex-M microcontrollers. *Plant Methods*. **20(2)**: 1-15.
- Bansal, S., Gupta, P. and Sharma, V. (2025). A hybrid machine learning framework using MobileNetV3 and CatBoost for paddy disease detection. *Scientific Reports*. **15**: 1-12.
- Dai, J., Li, Y., He, K. and Sun, J. (2016). R-FCN: Object detection via region-based fully convolutional networks in Advances in Neural Information Processing Systems (NeurIPS), Barcelona, Spain.
- Hasan, R.M., Rahman, M.A. and Alam, K.M. (2024). Transfer learning-based classification of paddy leaf diseases using CNN architectures. *Frontiers in Computer Science*. **6**: 1-14.
- Hoang, T.C., Nguyen, P.V. and Le, Q.H. (2024). YOLOv8 and Vision Transformer-based paddy disease recognition for mobile applications. *arXiv preprint arXiv:2412.05996*.
- Islam, M.N., Rahman, M.H. and Paul, S.K. (2025). Paddy leaf disease recognition with DenseNet121, Xception, EfficientNetB4 and MobileNetV3 models. *Discover Artificial Intelligence*. **5(1)**: 112-125.

- John, J., Walsh, Mangina, E. and Negrão, S. (2024). Advancements in imaging sensors and AI for plant stress detection: A systematic literature review, *Plant Phenomics*. **6**: 0153. ISSN 2643-6515. <https://doi.org/10.34133/plantphenomics.0153>.
- Kappali, R., Hemanthakumar, Sadyojatha, K.M., Prashanthi, S.K. (2024). Computer vision and machine learning in paddy diseases identification and classification: A review. *Indian Journal of Agricultural Research*. **58(2)**: 183-187. doi: 10.18805/IJARE.A-6061.
- Mahmud, S., Hasan, R. and Rahman, M. (2023). Instance segmentation for plant disease detection using machine learning models. *Computers and Electronics in Agriculture*. **209**: 1-13.
- Mandwale, U.K. and Singh, S. (2025). Advancements in paddy disease management: Integrating technology for better crop health. *Int. J. Adv. Res. Sci. Commun. Technol.* **5(9)**.
- Mishra, J.P., Singh, R. and Tripathi, A.K. (2022). Machine learning for detection and classification of paddy crop diseases: A survey. *Information Processing in Agriculture*. **8(4)**: 570-586.
- Radhika, K. and Latha, D.M. (2019). Machine learning model for automation of soil texture classification. *Indian Journal of Agricultural Research*. **53(1)**: 78-82. doi: 10.18805/IJARE.A- 5053.
- Ryan, M., Isakhanyan, G. and Tekinerdogan, B. (2023). An interdisciplinary approach to artificial intelligence in agriculture. *NJAS: Impact in Agricultural and Life Sciences*. **95(1)**. doi: 10.1080/27685241.2023.2168568.
- Saleem, M.H., Potgieter, J. and Arif, K.M. (2022). A Performance-Optimized Deep Learning-Based Plant Disease Detection Approach for Horticultural Crops of New Zealand, in *IEEE Access*. **10**: 89798-89822. doi: 10.1109/ACCESS.2022.3201104.
- Zhang, K., Li, L. and Zhao, X. (2021). Review of machine learning in plant pathology: CNN applications and limitations. *Frontiers in Plant Science*. **12(8)**: 1-12.