

Cotton Leaf Disease Detection and Classification Using Deep Learning: A Comprehensive Critical Review

Chirag Kaneria¹, Dr. Hemant Patel²

¹ PG Scholar, School of Engineering and Technology, Dr. Subhash University, Junagadh, Gujarat, India

² Professor and Head, CES/IT Department, School of Engineering and Technology, Dr. Subhash University, Junagadh, Gujarat, India

¹chirag.kaneria@dsuni.ac.in ²hemant.patel@dsuni.ac.in

Abstract - Cotton is an important agricultural crop, but its production is significantly affected by various leaf diseases that can reduce yield and quality. Deep learning (DL) and computer vision (CV) techniques have emerged as effective approaches for automated disease detection and classification. This paper presents a comprehensive critical review of deep learning-based methods for cotton leaf disease detection and classification. A systematic analysis of 36 studies published between 2016 and 2026 is conducted, focusing on datasets, preprocessing techniques, model architectures, evaluation protocols, and reported performance. The review examines the application of convolutional neural networks, transfer learning, object detection, and advanced deep learning models for identifying cotton leaf diseases. The findings indicate that although high classification performance has been reported, results vary considerably due to differences in datasets, disease classes, imaging conditions, and evaluation methodologies. Key research challenges include limited field-based datasets, class imbalance, variations in environmental conditions, model interpretability, and deployment constraints on resource-limited devices. Based on these findings, the paper identifies research gaps and proposes directions toward robust, explainable, and computationally efficient systems for real-world cotton disease diagnosis.

Key words -- Cotton leaf disease, Deep learning, Convolutional neural networks, Disease detection, Transfer learning, Precision agriculture.

1. INTRODUCTION

1.1 Agricultural Context

Cotton is a major commercial crop supporting fibre production, textile manufacturing and agricultural livelihoods. Disease and pest pressure can reduce plant health, yield potential and fibre quality. The reviewed literature considers problems including bacterial blight, target spot, powdery mildew, cotton leaf curl, aphids, armyworm and other fungal or viral symptoms. Automated visual diagnosis is attractive because it can support rapid screening over large areas. General research in agricultural deep learning has demonstrated the growing role of image-based methods for automated crop monitoring and disease

recognition [33], while early plant-disease studies established the feasibility of deep learning for image-based diagnosis [35].

1.2 Limitations of Conventional Diagnosis and Emergence of Computational Approaches

Manual diagnosis depends on human expertise and can be affected by observer experience, symptom similarity, early-stage symptoms, illumination, leaf orientation, occlusion and complex field backgrounds. Laboratory confirmation can improve objectivity but is slower and more expensive for routine large-scale screening. These limitations motivate image-based computational approaches. Previous plant-disease research has demonstrated the use of image segmentation, colour information and soft-computing techniques for automated disease recognition [32].

Early plant-disease systems relied on colour transformation, thresholding, segmentation, texture descriptors and conventional classifiers [32]. CNNs subsequently enabled automatic hierarchical feature learning, with early large-scale demonstrations showing the potential of deep learning for image-based plant disease detection [35]. Deep learning subsequently became increasingly important in agricultural applications [33]. Transfer learning became particularly attractive because cotton disease datasets are usually much smaller than large generic computer-vision datasets. Research on plant disease and plant classification has shown that transfer learning and fine-tuning can effectively adapt pretrained representations to domain-specific agricultural image tasks [31][36]. Recent cotton-specific work has moved toward hybrid CNNs, ensembles, attention, YOLO-family detection, lightweight networks, XAI and field-oriented systems [1–10].

1.3 Central Research Question and Contributions of This Review

The important research question is no longer simply whether deep learning can classify cotton leaf images. The literature already demonstrates that capability under many experimental settings. The stronger question is whether a model can remain accurate, calibrated, explainable, computationally efficient and reliable when tested on

heterogeneous images collected across farms, cultivars, seasons, devices and environmental conditions. Factors associated with image characteristics, dataset composition and acquisition conditions can substantially influence plant-disease recognition performance [34].

This review contributes: (i) a cotton-specific critical synthesis; (ii) a dataset-centric analysis of provenance, augmentation and leakage; (iii) comparison of classification, detection, segmentation and severity tasks; (iv) analysis of recent attention, XAI and lightweight approaches; (v) identification of research gaps; and (vi) a future multi-stage research framework suitable for field-oriented cotton disease diagnosis.

2. METHODOLOGY

2.1 Review Design and Sources

The review follows a structured critical literature review methodology, supplemented by verification of bibliographic records and recent publisher/indexed records for key studies. Studies are examined along the dimensions shown in Table I. Reported performance is preserved as reported by the individual study and is not treated as directly comparable across different datasets or protocols. This caution is consistent with broader reviews of deep learning in agriculture and plant disease recognition, which emphasize the importance of datasets, preprocessing, experimental design and evaluation conditions [33][34]. Figure 1 Literature Review Workflow and Study Selection Process.

Literature was searched using Scopus, Web of Science, IEEE Xplore, ScienceDirect, SpringerLink and Google Scholar using combinations of the keywords “cotton leaf disease,” “cotton disease detection,” “cotton disease classification,” “deep learning,” “CNN,” “transfer learning,” “YOLO,” “segmentation,” “severity estimation,” and “explainable AI.” Studies were included when they addressed automated cotton disease or pest recognition using image processing, machine learning or deep learning and were published in English between 2016 and 2026. Duplicate records, studies unrelated to cotton disease detection, non-research articles and papers lacking sufficient methodological information were excluded.

2.2 Review Dimensions

Each study was examined along the dimensions summarized in Table 1, covering disease/pest categories, dataset properties, preprocessing, model architecture, localization capability, severity estimation, validation protocol, computational efficiency, explainability, and deployment readiness.

Dimension	Aspects Examined
Disease/pest categories	Disease classes, pests, class structure
Dataset	Source, size, original versus augmented images, field/lab setting
Preprocessing	Resize, normalization, enhancement, segmentation, augmentation
Model	CNN, transfer learning, ensemble, attention, transformer, YOLO
Localization	Bounding boxes, lesion localization and segmentation
Severity	Severity grades, infected-area estimation
Validation	Train/test, cross-validation, cross-dataset and external field tests
Efficiency	Parameters, FLOPs, inference time, memory
XAI	Grad-CAM, attention maps and explanation validation
Deployment	Web, mobile, edge and real-time use

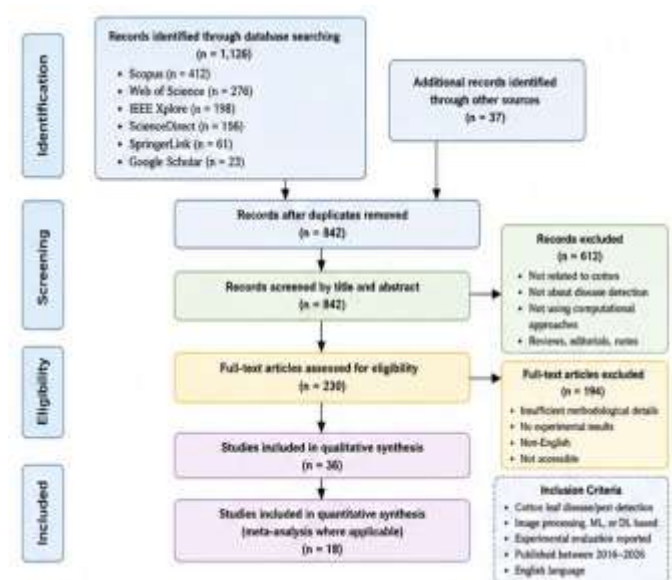


Fig -1 Literature Review Workflow and Study Selection Process.

Table-1: Review Dimensions

2.3 Limitations

This review is based on a defined set of supplied cotton disease literature, supplemented by verification of bibliographic records, rather than an exhaustive, independently executed systematic search across all indexing databases; some relevant studies may therefore have been missed. The synthesis is narrative-critical rather than a formally registered systematic review with predefined inclusion/exclusion protocols, meta-analytic pooling, or a documented PRISMA-style flow diagram.

Numerical performance values reported by the original studies are preserved as reported for comparison purposes, even though this review argues throughout that such values are not directly comparable across differing datasets, splits, and metrics. This limitation is particularly important because factors influencing plant disease recognition can include dataset size, image quality, disease presentation, acquisition conditions and experimental design [34].

Finally, because much of the underlying literature originates from 2024–2026, some very recent publications may not yet be indexed or may have been updated after this review was prepared.

3. RESULTS

3.1 Task Taxonomy

Classification is appropriate for clean leaf-dominant images and rapid screening. Detection and segmentation become more important when field images contain multiple leaves, clutter, occlusion, shadows or small lesions. Previous work on cotton leaf lesion identification and natural-environment segmentation demonstrates the importance of moving beyond simple image-level classification when disease regions must be identified under realistic conditions [9], [17]. Severity estimation requires more than a class label because management decisions depend on the amount and stage of infection [18]. Table 2 summarizes the major task categories, their typical outputs, suitable approaches, and evaluation metrics for cotton disease detection.

Table 2: Task Taxonomy for Cotton Disease Detection

Task	Typical Output	Suitable Approaches	Key Metrics
Image classification	One disease/pest label	CNN, transfer learning, transformer	Accuracy, balanced accuracy, Macro-F1, recall

Multi-label classification	Multiple co-occurring stresses	CNN/transformer	Per-class precision/recall/F1, mAP
Object detection	Bounding boxes	YOLO, RT-DETR, Faster R-CNN	mAP, precision, recall, F1
Semantic segmentation	Leaf/lesion masks	U-Net, U-Net++, DeepLab	IoU, Dice
Severity estimation	Percentage/grade	Segmentation + regression/classification	MAE, RMSE, correlation, weighted kappa
Explainability	Attribution/saliency maps	Grad-CAM, Grad-CAM++, SHAP	Expert agreement, localization overlap, faithfulness

3.2 Dataset Characteristics and Critical Issues

The supplied literature uses public datasets, researcher-collected images and combined datasets. Public datasets improve accessibility and reproducibility, while field-collected datasets can capture realistic variation. However, many studies provide limited information about cultivar, geography, season, acquisition device and the number of truly independent plants. Similar concerns have been highlighted in broader plant-disease deep-learning research, where dataset composition and acquisition conditions can strongly influence model performance [33], [34].

A key concern is that a large reported image count may result from augmenting a much smaller set of original photographs. Augmentation can improve training diversity but cannot create new biological or geographic variability. Therefore, papers should report original image count, augmented image count, augmentation operations and the point at which splitting was performed.

Leakage may occur when augmented copies of one original image appear in both training and test partitions, when multiple photographs of the same plant are split across partitions, or when images from the same acquisition session are present in both sets. The preferred rule is to split original images before augmentation and, when possible, split at

plant, plot, farm, field-visit or acquisition-session level. These dataset-design considerations are particularly important in agricultural image recognition because apparent model performance can be affected by non-independent samples and dataset-specific visual patterns [34].

3.3 Evolution of Deep Learning Methods

Traditional methods established important principles for lesion segmentation and disease-region analysis. Colour transforms, thresholding and handcrafted descriptors can be useful in controlled settings, but their performance can be sensitive to illumination, background and disease stage. They also depend heavily on segmentation quality. Traditional image segmentation and soft-computing methods have previously been applied to plant leaf disease detection [32], while cotton-specific natural-environment segmentation research further demonstrates the importance of handling realistic backgrounds and illumination conditions[17]. The evolution of cotton leaf disease detection has progressed from handcrafted image-processing techniques to machine learning, deep learning, object detection, explainable AI, and lightweight edge-oriented models, as summarized in Figure 2.

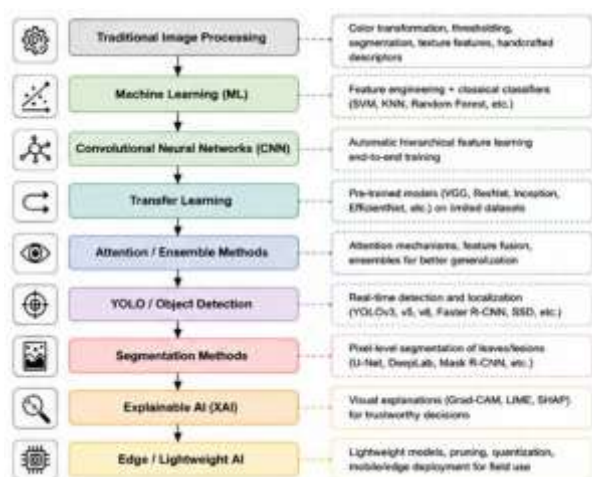


Fig -2. Evolution of Cotton Leaf Disease Detection Techniques

CNNs became dominant because they learn colour, texture, shape and spatial representations directly from images. Early deep-learning research demonstrated the effectiveness of CNN-based plant disease recognition [35], and subsequent agricultural surveys documented the rapid adoption of deep learning for crop and disease-related applications [33]. Cotton-specific studies have subsequently developed customized CNN architectures, including CottonLeafNet [16]. The supplied literature reports many customized CNN systems in the high-90% range under their

respective experimental conditions. Their major weakness is dependence on dataset quality and sufficient variation.

Transfer learning reuses representations learned from large datasets and adapts them to cotton disease images. Transfer learning has been widely investigated for plant disease identification [31], and studies have demonstrated that fine-tuning pretrained networks can provide strong performance in agricultural image classification [36]. Cotton-specific research has investigated deep transfer learning [12], DenseNet-121 transfer learning [14], ensemble transfer learning [11], and fine-tuned transfer-learning systems with practical web deployment [19]. The most frequently investigated families include VGG, ResNet, Inception, Xception, DenseNet, EfficientNet and MobileNet.

Johri et al. compared several such architectures and reported 99.96% validation accuracy for EfficientNetB3 [3]. This is strong evidence of model capability, but not proof that EfficientNetB3 is universally superior because the evaluation is dataset-specific.

Kanakala et al. combined fine-tuned VGG16 and InceptionV3 using stacking with logistic regression and reported 98.56% accuracy for six classes [1]. Ensemble approaches can reduce individual-model mistakes and combine complementary feature representations, but they frequently raise computational and memory requirements [11], [22].

Haque et al. proposed CottonLeafNet, an attention-enhanced CNN ensemble with lightweight backbones, multi-head self-attention and Grad-CAM. It achieved 98.33% on D1, 99.43% on D2 and 99.08% on a unified eight-class dataset, with cross-dataset evaluation and a real-time web application [2]. Other recent research has also investigated explainable deep-learning approaches for cotton leaf disease classification [23]. The study represents an important movement from isolated accuracy reporting toward interpretability and generalization.

Attention mechanisms and lightweight architectures are also becoming increasingly important. Mo and Wei integrated synthetic data augmentation with decoupled focused self-attention for lightweight cotton disease detection [24], exhibiting how attention can be added while retaining an emphasis on computational economy.

Object detection is better suited to field images where leaves or lesions do not occupy the whole image. Recent cotton studies use YOLO variants, including YOLOv11, and combined detection approaches. Chinnadurai and Selvakumar reported a framework combining YOLOv5 and Faster R-CNN for real-time cotton disease diagnosis [4]. Hu et al. investigated an efficient YOLOv11-based multi-class

detector [5]. More recent comparative work has also investigated several YOLO generations for cotton disease detection and localization [29]. These approaches indicate a shift from “what disease is present?” toward “where is the disease?”

MobileNet-family architectures and lightweight CNNs reduce computation and are promising for mobile and edge devices. Lightweight CNN research has demonstrated that reduced-complexity architectures can achieve strong cotton disease classification performance [15], [20]. RF-CottNet and other lightweight approaches emphasize the need to report parameters, FLOPs, inference time, model size and hardware rather than accuracy alone [25]. Compression methods such as pruning, quantization and knowledge distillation can further reduce deployment cost. Pruning has already been investigated specifically for cotton disease identification [10], while resource-efficient cotton disease models have incorporated mobile-oriented architectures and quantization-aware strategies [25].

Naeem et al. created CLDP-CNN with two datasets, including field photos, and reported 99.62% accuracy on a secondary dataset and 99.78% accuracy on the original field dataset [6]. This is significant since it specifically takes field data into account. Deep learning techniques have been used in other studies to detect and predict cotton diseases [21].

Ram et al. investigated federated learning for cotton disease detection in Indian agricultural environments and reported 97.01% test accuracy with four clients and 95.92% with eight clients, illustrating a different route toward distributed and privacy-aware agricultural learning [7].

3.4 Comparative Analysis of Representative Studies

Reported values should be interpreted as evidence of methodological capability rather than as a universal leaderboard. The studies use different datasets, disease classes, splits and metrics; therefore, numerical ranking across them is scientifically inappropriate. Table 3 presents a comparative analysis of representative studies, including their approaches, reported results, key contributions, and principal cautions.

Table 3: Comparative Analysis of Representative Studies

Study	Year	Approach	Reported Result	Key Contribution	Main Caution
Kanakala et al. [1]	2026	VGG16 + Inception	98.56% accuracy	Hybrid ensemble	Dataset-specific result

		V3 stacking			
Haque et al. [2]	2026	Attention CNN ensemble + Grad-CAM	98.33–99.43% accuracy	XAI, cross-dataset, web deployment	Domain shift remains
Johri et al. [3]	2024	Transfer-learning comparison	99.96% validation accuracy	Broad architecture comparison	Dataset-specific validation
Chinnadurai & Selvakumar [4]	2025	YOLOv5 + Faster R-CNN	Real-time detection framework	Localization + speed	Annotation quality
Hu et al. [5]	2025	YOLOv11 detector	94.6% mAP@0.5; 148 FPS	Efficient multi-class detection	Not comparable with accuracy
Naeem et al. [6]	2026	CLDP-CNN transfer learning	99.78% field; 99.62% secondary	Field dataset + web app	Broader external testing needed
Ram et al. [7]	2026	Efficient NetB0 + FedSCAFFOLD	97.01% / 95.92%	Federated learning	Client/data configuration
Rai & Pahuja [11]	2024	Ensemble transfer learning	Not reported in the reviewed source	Ensemble transfer learning	Dataset/protocol dependence
Rajasekar et al. [12]	2021	Deep transfer learning	Not reported in the reviewed source	Cotton disease transfer learning	Dataset-specific validation

			d source		
Odukoya et al. [13]	2023	Segmentation + ML	Not reported in the reviewed source	Fungal disease segmentation	Dataset/environment dependence
Arathi & Dulhare [14]	2023	DenseNet-121 transfer learning	Not reported in the reviewed source	Transfer-learning classification	Dataset-specific validation
Peyal et al. [15]	2022	Lightweight CNN	Not reported in the reviewed source	Reduced computational complexity	Hardware-level evaluation needed
Singh et al. [16]	2023	CottonLeafNet	98.33–99.43%	Cotton-specific deep neural network	Generalization remains important
Zhang et al. [17]	2018	Natural-environment segmentation	Not reported in the reviewed source	Disease-leaf segmentation	Segmentation/generalization
Parikh et al. [18]	2016	Detection + severity estimation	Not reported in the reviewed source	Severity analysis	Early literature/dataset limitations
Islam et al. [19]	2023	Fine-tuned deep learning	Not reported in the reviewed source	Web application	External validation

Islam et al. [20]	2022	Lightweight deep learning	Not reported in the reviewed source	Efficient cotton disease model	Deployment validation
Aslam et al. [22]	2025	Multi-CNN synergistic learning	Not reported in the reviewed source	Multi-CNN framework	Complexity
Kaur et al. [23]	2025	Explainable deep learning	Not reported in the reviewed source	XAI for cotton disease	Explanation validation
Mo & Wei [24]	2025	StyleGAN2-ADA + attention	Not reported in the reviewed source	Lightweight attention	Generalization
Wang et al. [25]	2025	Resource-efficient network	Not reported in the reviewed source	Mobile/edge-oriented design	Device-level validation
Salot et al. [26]	2025	ML early detection	Not reported in the reviewed source	Early-stage disease detection	Deep-learning extension needed
Pavates et al. [29]	2026	YOLO variants	Not reported in the reviewed source	YOLO-based detection	Cross-dataset testing

3.5 Research Gaps

Ten major research gaps were identified across the reviewed literature: limited dataset diversity; laboratory-to-

field generalization; leakage and augmentation dependency; non-standardized evaluation; limited XAI validation; computational complexity; disease severity estimation; early-stage disease detection; limited multimodal sensing; and continuous learning. The recommended direction for future work is to construct multi-location datasets across cultivars, seasons and devices; use independent field validation; split original samples before augmentation; define benchmark datasets and common reporting protocols; validate explanations against expert lesion masks; use lightweight architectures and compression; estimate infected-area percentage or severity grades; build early-stage subsets; investigate RGB plus thermal, multispectral and environmental information; and monitor model drift under controlled incremental learning. Dataset diversity and environmental variability are consistent concerns in broader plant-disease deep-learning research [33], [34]. Figure 3 summarizes the relationship between the major research gaps identified in the literature and corresponding future research directions for robust cotton disease detection systems. Table 4 summarizes the ten major research gaps identified in the reviewed literature and the corresponding recommended research directions.

Table 4: Research Gaps in Cotton Disease Detection

Gap	Problem	Recommended Direction
1 — Limited dataset diversity	One dataset or limited geographic coverage	Multi-location datasets across cultivars, seasons and devices
2 — Laboratory-to-field generalization	Shadows, occlusion, clutter, blur and natural symptom variation	Independent field validation
3 — Leakage and augmentation dependency	Augmented data may overstate effective diversity	Split original samples before augmentation; plant/plot/farm-level separation
4 — Non-standardized evaluation	Different datasets and metrics prevent fair comparison	Benchmark datasets and common reporting protocol

5 — Limited XAI validation	Heat maps often lack lesion-level validation	Compare explanations with expert masks and faithfulness tests
6 — Computational complexity	Large ensembles/transformers may not suit field devices	Lightweight models, pruning, quantization, distillation [10], [15], [20], [25]
7 — Disease severity	Class label alone omits infection amount	Estimate infected-area percentage, disease severity grade, lesion ratio and progression over time [18]
8 — Early-stage disease	Subtle symptoms are difficult to distinguish	Early-stage/mild-infection subsets and recall evaluation [26]
9 — Limited multimodal sensing	Most systems use RGB alone	RGB + thermal/multispectral + environmental information
10 — Continuous learning	Appearance varies across seasons and locations	Incremental learning with feedback and drift monitoring


Fig -3. Research Gap-to-Future Direction Framework

4. SUMMARY AND FUTURE RESEARCH

4.1 Proposed Multi-Stage Research Architecture

The literature supports a multi-stage system rather than a single classifier. The following architecture is proposed as a research direction; it is not claimed as an experimentally validated system. The stages are: (1) image acquisition using smartphone, fixed camera or drone; (2) image-quality assessment for blur, brightness and resolution with a recapture decision; (3) leaf/plant detection using a YOLO-family or lightweight detector, supported by recent cotton object-detection studies [4], [5], [29]; (4) leaf/lesion segmentation using U-Net/U-Net++ or lightweight segmentation, motivated by cotton lesion and natural-environment segmentation studies [9], [17]; (5) disease-region localization using attention or lesion localization; (6) disease classification using EfficientNet, ConvNeXt, MobileViT or a lightweight ensemble; (7) severity estimation from infected-area percentage and severity grade, building on earlier cotton severity-estimation work [18]; (8) XAI using Grad-CAM, Grad-CAM++ or SHAP, consistent with recent explainable cotton disease research [2], [23]; (9) decision support combining disease, confidence, severity and management guidance; and (10) edge/mobile deployment using a quantized or pruned model on an Android or edge device, consistent with resource-efficient cotton disease research [10], [15], [20], [25]. The proposed framework is presented in Figure 4. The functions of the ten stages are summarized in Table 5.

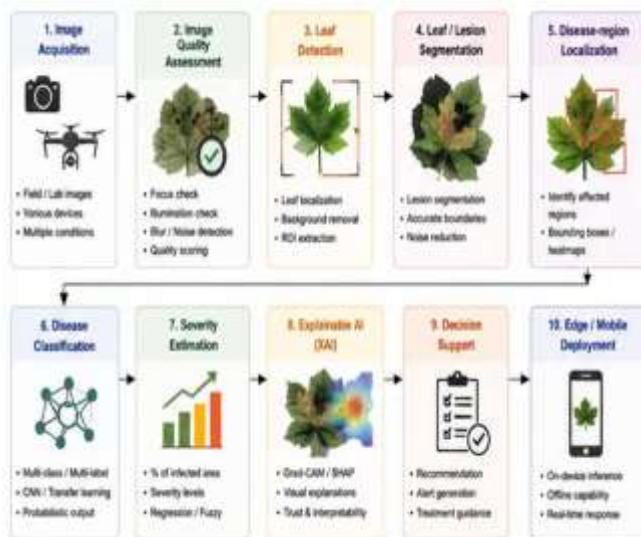


Fig -4. Proposed Multi-Stage Cotton Leaf Disease Detection Framework

Table 5: Proposed Multi-Stage Architecture

Stage	Function
1. Image acquisition	Smartphone, fixed camera or drone

2. Image-quality assessment	Blur, brightness, resolution and recapture decision
3. Leaf/plant detection	YOLO-family or lightweight detector
4. Leaf/lesion segmentation	U-Net/U-Net++ or lightweight segmentation
5. Disease-region localization	Attention/lesion localization
6. Disease classification	EfficientNet/ConvNeXt/Mobile ViT or lightweight ensemble
7. Severity estimation	Infected-area percentage and severity grade
8. Explainable AI	Grad-CAM/Grad-CAM++/SHAP
9. Decision support	Disease + confidence + severity + management guidance
10. Edge/mobile deployment	Quantized/pruned model on Android/edge device

4.2. Recommended Experimental Methodology for Future Work

1) Dataset Design: Collect images across multiple farms, cultivars, growth stages, seasons, devices, illumination conditions and disease severities. Record metadata that allows plant-level and farm-level splitting. Expert annotation should include disease class, leaf boundary, lesion region and severity where feasible. These requirements directly address the dataset and environmental factors known to influence plant disease recognition [34] and the broader agricultural deep-learning literature [33].

2) Data Partitioning: Use separate training, validation, internal-test, cross-dataset and independent-field-test partitions. The critical rule is to perform partitioning before augmentation and keep all augmented versions of an original image within the same partition.

3) Evaluation: For classification report accuracy, balanced accuracy, precision, recall, Macro-F1, weighted-F1, confusion matrix and Cohen's kappa. For detection report precision, recall, F1, mAP@0.5, mAP@0.5:0.95 and inference time. For segmentation report IoU, Dice and per-class performance. For severity report MAE/RMSE, correlation, weighted kappa or grade-wise F1. For deployment report parameters, FLOPs, model size, preprocessing time, inference time, RAM and device. For XAI report localization overlap, expert agreement and

faithfulness. The recommended data-partitioning strategy is summarized in Table 6.

Table 6: Recommended Data Partitioning

Partition	Purpose
Training	Learn model parameters
Validation	Architecture and hyperparameter selection
Internal test	Same-distribution performance
Cross-dataset test	Transfer to different dataset
Independent field test	Real-world generalization across site/season/device

Table 7 summarizes the minimum evaluation metrics and reporting requirements for each task.

Table 7: Minimum Evaluation Reporting By Task

Task	Minimum Reporting
Classification	Accuracy, balanced accuracy, precision, recall, Macro-F1, weighted-F1, confusion matrix, Cohen's kappa
Detection	Precision, recall, F1, mAP@0.5, mAP@0.5:0.95, inference time
Segmentation	IoU, Dice, per-class performance
Severity	MAE/RMSE, correlation, weighted kappa or grade-wise F1
Deployment	Parameters, FLOPs, model size, preprocessing time, inference time, RAM and device
XAI	Localization overlap, expert agreement and faithfulness

4.3 Priority Research Questions

RQ1: How can heterogeneous cotton datasets be constructed to improve geographical and environmental generalization?

RQ2: Can leaf and lesion segmentation before classification improve robustness under complex field backgrounds?

RQ3: Can lightweight attention models approach the performance of large ensembles with substantially lower computational cost?

RQ4: Does XAI reliably highlight biologically relevant disease regions rather than background artifacts?

RQ5: Can RGB, thermal and environmental information improve early-stage disease detection?

RQ6: Can knowledge distillation transfer large-model performance to mobile/edge models?

RQ7: Can controlled continuous learning adapt to seasonal and geographical changes?

RQ8: Does calibrated uncertainty improve the safety of agricultural decision support?

4.4. Critical Synthesis

The literature shows a clear progression: conventional image processing → machine learning → CNN classification → transfer learning → hybrid/ensemble learning → attention mechanisms → object detection and segmentation → XAI → lightweight/edge deployment → integrated agricultural decision support. This progression is consistent with the broader development of deep learning in agriculture [32], [33], [35] and the increasingly specialized cotton disease literature [1–10], [11–25], [29].

The scientific challenge has therefore shifted from proving that CNNs can classify cotton disease to proving that systems remain reliable under realistic field variability. High benchmark accuracy is valuable, but it is only one component of a deployable agricultural system.

The strongest recent studies improve the evidence base by adding field images, cross-dataset evaluation, cross-validation, attention, XAI, real-time inference or distributed learning [2], [6], [7], [23], [25]. Nevertheless, broad external validation, severity estimation, explanation validation and standardized reporting remain insufficiently developed. A future contribution should therefore prioritize robustness and evidence quality over another marginal improvement in single-dataset accuracy.

4.5. Recommended Novelty Direction

A strong future research direction is: “An Explainable, Lightweight, Multi-Stage Deep Learning Framework for Real-Time Cotton Leaf Disease Detection, Localization, and Severity Estimation.” Its novelty should come from integration and rigorous validation rather than simply proposing another CNN. The study should combine field-oriented data, leakage-safe partitioning, detection/segmentation, an efficient attention classifier, severity estimation, XAI validation, cross-dataset testing and edge deployment. Recent lightweight, attention and resource-efficient studies provide a foundation for this direction [23], [24], [25]. If multimodal data are available, thermal or environmental information can be evaluated as an additional modality.

5. CONCLUSION

Deep learning has substantially advanced automated cotton leaf disease recognition. CNNs and transfer-learning architectures such as VGG, ResNet, Inception, Xception, DenseNet, EfficientNet and MobileNet have produced strong results in the reviewed literature [3], [11], [12], [14],

[16], [19], [31], [36]. Hybrid and ensemble models can improve complementary feature representation [1], [11], [22], while attention and XAI increase interpretability [2], [23], [24]. Recent work is also moving toward YOLO-based localization [4], [5], [29], lightweight inference [10], [15], [20], [25], field datasets [6], [9], cross-dataset evaluation and distributed learning [7].

However, reported accuracy values must be interpreted cautiously. Differences in dataset composition, augmentation, class definitions, validation protocols and environmental conditions make direct numerical ranking scientifically inappropriate [33], [34]. The major unresolved challenges are dataset diversity, field robustness, leakage-safe evaluation, early-stage detection, severity estimation, explanation validity, computational efficiency and adaptation over time.

The future of cotton disease detection should therefore move from isolated accuracy-oriented classifiers toward complete intelligent agricultural decision-support systems that can detect, localize, classify, quantify and explain disease under real field conditions. The proposed multi-stage architecture provides a practical research pathway toward that goal.

REFERENCES

- [1] S. Kanakala, R. Ch. Manoranjini, P. Ravi, M. Riyazuddin, and K. Naresh, "Hybrid deep learning approach for cotton leaf disease detection and management using fine-tuned VGG16 and Inception v3 models," *International Journal of Systematic Innovation*, vol. 10, no. 2, Article 2008, 2026.
- [2] M. E. Haque, M. T. H. Saykat, A. Al-Imran, A. H. Siam, J. Uddin, and D. Ghose, "An attention enhanced CNN ensemble for interpretable and accurate cotton leaf disease classification," *Scientific Reports*, vol. 16, Article 4476, 2026.
- [3] P. Johri, S. Kim, K. Dixit, P. Sharma, B. Kakkar, Y. Kumar, J. Shafi, and M. F. Ijaz, "Advanced deep transfer learning techniques for efficient detection of cotton plant diseases," *Frontiers in Plant Science*, vol. 15, Article 1441117, 2024.
- [4] S. Chinnadurai and S. Selvakumar, "Image-based cotton leaf disease diagnosis using YOLO and Faster R-CNN techniques," *Scientific Reports*, vol. 15, Article 45785, 2025.
- [5] F. Hu et al., "Investigation of an efficient multi-class cotton leaf disease detection algorithm that leverages YOLOv11," *Sensors*, vol. 25, Article 4432, 2025.
- [6] M. Naeem et al., "Automatic prediction of cotton leaf's diseases using deep learning techniques," *Scientific Reports*, 2026.
- [7] D. Ram, A. Avadhanam, M. Gyanchandani, A. Rasool, and N. Khare, "Federated deep learning for cotton leaf disease detection in Indian agricultural environments using EfficientNet B0 and FedSCAFFOLD," *Discover Artificial Intelligence*, vol. 6, Article 614, 2026.
- [8] M. S. Memon, P. Kumar, and R. Iqbal, "Meta deep learn leaf disease identification model for cotton crop," *Computers*, vol. 11, Article 102, 2022.
- [9] R. F. Caldeira, W. E. Santiago, and B. M. Teruel, "Identification of cotton leaf lesions using deep learning techniques," *Sensors*, vol. 21, no. 9, Article 3169, 2021.
- [10] D. Zhu, Q. Feng, J. Zhang, and W. Yang, "Cotton disease identification method based on pruning," *Frontiers in Plant Science*, vol. 13, Article 1038791, 2022.
- [11] C. K. Rai and R. Pahuja, "An ensemble transfer learning-based deep convolution neural network for the detection and classification of diseased cotton leaves and plants," *Multimedia Tools and Applications*, vol. 83, pp. 1–34, 2024.
- [12] V. Rajasekar, K. Venu, S. R. Jena, R. J. Varthini, and S. Ishwarya, "Detection of cotton plant diseases using deep transfer learning," *Journal of Mobile Multimedia*, vol. 18, pp. 307–324, 2021.
- [13] O. H. Odukoya, S. Aina, and F. W. Dégbéssé, "A model for detecting fungal diseases in cotton cultivation using segmentation and machine learning approaches," *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 1, pp. 1–10, 2023.
- [14] B. Arathi and U. N. Dulhare, "Classification of cotton leaf diseases using transfer learning-DenseNet-121," in *Proc. Third International Conference on Advances in Computer Engineering and Communication Systems (ICACECS 2022)*, pp. 393–405, 2023.
- [15] H. I. Peyal et al., "Cotton leaf disease detection and classification using lightweight CNN architecture," in *Proc. 2022 12th International Conference on Electrical and Computer Engineering (ICECE)*, pp. 413–416, 2022.
- [16] P. Singh et al., "CottonLeafNet: Cotton plant leaf disease detection using deep neural networks," *Multimedia Tools and Applications*, vol. 82, pp. 37151–37176, 2023.
- [17] J. H. Zhang et al., "Automatic image segmentation method for cotton leaves with disease under natural environment," *Journal of Integrative Agriculture*, vol. 17, no. 8, pp. 1800–1814, 2018.
- [18] A. Parikh, M. S. Raval, C. Parmar, and S. Chaudhary, "Disease detection and severity estimation in cotton plant from unconstrained images," in *Proc. IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, pp. 594–601, 2016.
- [19] M. M. Islam et al., "A deep learning model for cotton disease prediction using fine-tuning with smart web application in agriculture," *Intelligent Systems with Applications*, vol. 20, Article 200278, 2023.
- [20] M. T. Islam, M. M. Rahman, and M. M. Hasan, "Lightweight deep learning model for cotton leaf disease detection," *Smart Agricultural Technology*, vol. 2, Article 100040, 2022.
- [21] A. Rehman, N. Akhtar, and O. H. Alhazmi, "Monitoring and predicting cotton leaf diseases using deep learning approaches

- and mathematical models,” *Scientific Reports*, vol. 15, Article 22570, 2025.
- [22] A. Aslam et al., “Multi-convolutional neural networks for cotton disease detection using synergistic deep learning paradigm,” *PLOS ONE*, vol. 20, pp. 1–25, 2025.
- [23] G. Kaur et al., “Explainable AI for cotton leaf disease classification: A metaheuristic-optimized deep learning approach,” *Food Science & Nutrition*, 2025.
- [24] H. Mo and L. Wei, “Lightweight detection of cotton leaf diseases using StyleGAN2-ADA and decoupled focused self-attention,” *Journal of King Saud University Computer and Information Sciences*, vol. 37, Article 41, 2025.
- [25] Z. Wang et al., “Resource-efficient cotton network: A lightweight deep learning framework for cotton disease and pest classification,” *Plants*, vol. 14, Article 2082, 2025.
- [26] P. Salot et al., “Cotton leaf analysis based early plant disease detection using machine learning,” *Journal of Integrated Science and Technology*, vol. 13, 2025.
- [27] M. Sharma and R. Kumar, “Precision agriculture through deep learning for cotton leaf disease detection using advanced CNN architecture,” in *Proc. 2026 International Conference on Emerging Trends and Innovations in ICT (ICEI)*, 2026.
- [28] S. S. Kumar et al., “Hybrid deep learning-based framework for multi-level classification and detection of cotton leaf diseases,” *International Journal of Drug Delivery Technology*, vol. 16, no. 20s, Article 84, 2026.
- [29] A. Pavates et al., “Cotton care: YOLO-powered cotton leaf disease detection,” *Journal of Artificial Intelligence Machine Learning and Neural Network*, pp. 153–167, 2026.
- [30] “Deep transfer learning for comprehensive diagnosis of cotton leaf pathologies,” *Medical Mycology Pathogenesis*, Article 108257, 2025.
- [31] J. Chen et al., “Using deep transfer learning for image-based plant disease identification,” *Computers and Electronics in Agriculture*, vol. 173, Article 105393, 2020.
- [32] V. Singh and A. K. Misra, “Detection of plant leaf diseases using image segmentation and soft computing techniques,” *Information Processing in Agriculture*, vol. 4, no. 1, pp. 41–49, 2017.
- [33] A. Kamilaris and F. X. Prenafeta-Boldú, “Deep learning in agriculture: A survey,” *Computers and Electronics in Agriculture*, vol. 147, pp. 70–90, 2018.
- [34] J. G. A. Barbedo, “Factors influencing the use of deep learning for plant disease recognition,” *Biosystems Engineering*, vol. 172, pp. 84–91, 2018.
- [35] S. P. Mohanty, D. P. Hughes, and M. Salathé, “Using deep learning for image-based plant disease detection,” *Frontiers in Plant Science*, vol. 7, Article 1419, 2016.
- [36] A. Kaya et al., “Analysis of transfer learning for deep neural network based plant classification models,” *Computers and Electronics in Agriculture*, vol. 158, pp. 20–29, 2019.