

Title	Dataset name and URL	Dataset description (samples, classes, images per class or split)	Methods name	Accuracy of the model	Research Questions	Pros and Cons	Citation
A deep learning-based approach for the detection of cucumber diseases	"A dataset for successful recognition of cucumber diseases" — . <a href="https://doi.org/10.1371/journal.pone.0320764">https://doi.org/10.1371/journal.pone.0320764</a>	<b>Total samples:</b> 1,280 RGB images (224×224 px after preprocessing) collected in Bangladesh under natural conditions. <b>Classes :</b> 8 <b>Cucumber.</b> <b>Split:</b> 70% training, 10% validation, 20% test	Modified VGG19-based CNN using novel transfer learning vs. traditional fine-tuning.	<b>Novel approach:</b> 95.31% accuracy, 97.66% balanced accuracy. Traditional: 87.89% accuracy, 93.87% balanced accuracy.	How can a modified deep learning model (VGG19) improve early and accurate detection of cucumber diseases compared to traditional transfer learning?	<b>Pros:</b> • 94.98% accuracy, beats benchmark • Strong generalization, LIME explainability • Useful for early agri-detection <b>Cons:</b> • Small dataset (1,280 images) • Single-region (Bangladesh) test • Needs multi-climate validation	[1]
An Edge Computing-Based Solution for Real-Time Leaf Disease Classification using Thermal Imaging	<b>Thermal Leaf Disease Dataset</b> (15,444 images, 7 classes) GitHub: <a href="#">Leaf-Diseases-Classification</a>	Collected with an <b>Infiray T3C thermal camera</b> (384×288 px, 25Hz). Covers <b>16 plant species</b> (e.g., citrus, mango, avocado). <b>7 disease classes</b> , 15,444 images. Split: <b>60% training, 20% validation, 20% testing.</b>	MobileNetV1, MobileNetV2, InceptionV3, VGG16 (with pruning + quantization-aware training)	Comparable or better than GPU baselines. MobileNetV1 & InceptionV3 performed best for accuracy and speed. Achieved <b>real-time inference (5.9–7.0 ms per frame)</b> on Edge TPU.	How effective are pruning + quantization techniques for compressing DL models without losing accuracy?	<b>Pros:</b> • First large thermal plant disease dataset • Real-time Raspberry Pi inference • PQAT compression: faster, same accuracy <b>Cons:</b> • One disease per leaf • Daylight heat affects accuracy • Tested only on Pi, TPU, NCS2	[2]
Cucumber Leaf Disease Detection and Classification Using a Deep Convolutional Neural Network (DCNN)	Dataset: <i>Plant Village Repository</i> DOI: <a href="https://doi.org/10.22059/jitm.2023.95248">https://doi.org/10.22059/jitm.2023.95248</a>	Total: <b>3,700 images</b> of cucumber leaves <b>4 Classes:</b> Powdery mildew, Downy mildew, Alternaria leaf spot, Healthy Training: 67% (≈2,479 images) Testing: 33% (≈1,221 images) After data augmentation: <b>11,100 images</b>	<b>Deep Convolutional Neural Network (DCNN)</b> <b>ResNet-50 (Transfer Learning)</b> <b>Data Augmentation:</b> Random rotation (30°), resizing, flipping Optimizer: <b>Adam</b> Framework: <b>PyTorch</b>	<b>Before augmentation:</b> 94.1% average accuracy <b>After augmentation:</b> <b>98.0%</b> overall accuracy Class-wise accuracies up to <b>100% (Healthy)</b>	How can cucumber leaf diseases be automatically detected and classified using deep learning? Can transfer learning and data augmentation improve accuracy in disease recognition? How does ResNet-50 perform compared to other CNN models for this task?	<b>Pros:</b> High accuracy, efficient transfer learning, lightweight ResNet-50 model. <b>Cons:</b> Needs large labeled data, high computation cost, risk of overfitting.	[3]

Deep Learning Algorithms for Automatic Detection and Classification of Mildew Disease in Cucumber" (Ozguven, 2020)	Cucumber Leaf Image Dataset (collected by the author, not publicly available)	<p>Total <b>175</b> images of cucumber leaves Image size: 4320×3240 → resized to 1022×764</p> <p><b>Classes</b> (4 total): 0 – Healthy (15 images) 1 – Low severity (45 images) 2 – Severe (11 images) 3 – Low + Severe mixed (104 images) Split not explicitly mentioned (used for both training and testing with validation via confusion matrix)</p>	Faster R-CNN (Region Proposal Network + CNN classifier)	94.86 % (overall accuracy) Sensitivity = 94.86 %, Specificity = 94.86 %	Can deep learning (Faster R-CNN) automatically detect and classify cucumber mildew disease from leaf images?	<p><b>Pros:</b></p> <ul style="list-style-type: none"> <li>• Real-time detection</li> <li>• Detects disease &amp; severity</li> <li>• Accurate with small data</li> <li>• Less manual testing</li> <li>• Fits auto-sprayers</li> </ul> <p><b>Cons:</b></p> <ul style="list-style-type: none"> <li>• Small data, less general</li> <li>• Lighting/shadow errors</li> <li>• No public dataset</li> <li>• GPU &amp; long training needed</li> </ul>	[4]
An Intelligent System for Cucumber Leaf Disease Diagnosis Based on the Tuned Hybrid Deep Learning Model" (Omer et al., 2022)	Cucumber Leaf Disease Dataset (Kaggle)	<p>Total 8,000+ cucumber leaf images collected from Kaggle's "PlantVillage"-style dataset. Divided into 6 disease classes (e.g., Downy mildew, Powdery mildew, Bacterial wilt, Anthracnose, Target leaf spot, and Healthy). • Data split: 80% training, 10% validation, 10% testing.</p> <ul style="list-style-type: none"> <li>• Augmentation (rotation, scaling, flipping) used to balance samples per class.</li> </ul>	Tuned Hybrid Deep Learning Model (THDLM) combining VGG16 + CNN + LSTM with optimization using Improved Moth Flame Optimization (IMFO) for parameter tuning.	99.69 % (test accuracy) Precision = 99.6 %, Recall = 99.7 %, F1 = 99.6 %	• How can a hybrid deep learning model (CNN + LSTM + VGG16) be optimized to improve cucumber disease recognition accuracy?	<p><b>Pros:</b></p> <ul style="list-style-type: none"> <li>• Very high accuracy (99.69%)</li> <li>• Handles large multi-class data</li> <li>• Combines CNN + LSTM features</li> <li>• IMFO reduces overfitting</li> <li>• Works in real-time field use</li> </ul> <p><b>Cons:</b></p> <ul style="list-style-type: none"> <li>• Computationally heavy</li> <li>• Needs GPU for training</li> <li>• Requires large labeled data</li> <li>• May overfit if imbalanced</li> </ul>	[5]

## **Related Work Summary:**

Recent advances in cucumber disease recognition have largely built upon supervised deep learning backbones through fine-tuning approaches. Early studies based on convolutional neural network (CNN) transfer learning particularly architectures such as VGG, ResNet, and Inception established strong performance baselines. For example, a VGG19 model trained on eight disease classes achieved a balanced accuracy of approximately 97.66%. Subsequent research further enhanced performance by employing hybrid strategies that fused feature representations from multiple CNN backbones. Although these methods pushed classification accuracy higher, they often came at the cost of increased model complexity and computational overhead.

As the field matured, researchers began addressing practical challenges related to real-world deployment. Newer studies have adopted lightweight CNN models and Transformer-based architectures that can operate efficiently under field conditions characterized by variable lighting, occlusion, and background clutter. Detection frameworks such as Swin-Transformer-based RetinaNet have been explored to improve the localization of small and overlapping lesions in natural environments. These approaches reflect a shift from purely classification-focused research toward more holistic recognition systems that balance accuracy, robustness, and computational feasibility.

Across the literature, several consistent trends have emerged. Data augmentation is universally emphasized to mitigate class imbalance and enhance generalization. Transfer learning remains the preferred strategy over training models from scratch, given limited labeled agricultural datasets. Some works have also integrated CNN feature extractors with traditional machine learning classifiers (e.g., SVM or Random Forest), creating hybrid pipelines that improve interpretability or speed.

Within this broader research context, the present study extends the supervised backbone paradigm but introduces three notable distinctions. First, it employs a carefully selected architecture and fine-tuning schedule tailored to the specific visual characteristics of cucumber disease imagery. Second, it utilizes a more comprehensive and balanced dataset, supported by aggressive data augmentation to ensure equitable representation across disease classes. Third, the evaluation framework extends beyond overall accuracy, incorporating macro-F1 scores, per-class recall, inference latency, and model size to better assess real-world deployability. Together, these contributions position the work as both a performance-oriented and practically grounded advancement in cucumber disease recognition research.

## REFERENCES:

- [1]Raufer L, Wiedey J, Mueller M, Penava P, Buettner R. A deep learning-based approach for the detection of cucumber diseases. PLOS ONE. 2025;20(4):e0320764. <https://doi.org/10.1371/journal.pone.0320764>
- [2] P. E. C. Silva and J. Almeida, "An Edge Computing-Based Solution for Real-Time Leaf Disease Classification using Thermal Imaging," *IEEE Geoscience and Remote Sensing Letters*, pp. 1–5, 2024. <https://doi.org/10.1109/LGRS.2024.3456637>
- [3] Singh, Manish K., & Kumar, Avadhesh. (2023). *Cucumber Leaf Disease Detection and Classification Using a Deep Convolutional Neural Network*. *Journal of Information Technology Management*, 15 (Special Issue), 94–110. <https://doi.org/10.22059/jitm.2023.95248>
- [4] Metin Ozguven, "Deep Learning Algorithms for Automatic Detection and Classification of Mildew Disease in Cucumber", *Fresenius Environmental Bulletin*, Vol 29 No 8 (2020), pp. 7081–7087. [DOI not provided; available on ResearchGate: [https://www.researchgate.net/publication/343263908\\*\\*](https://www.researchgate.net/publication/343263908**)]
- [5] Omer, M. A., Al-Obeidat, F., Al-Hussein, M., & Al-Masri, N. (2022). An Intelligent System for Cucumber Leaf Disease Diagnosis Based on the Tuned Hybrid Deep Learning Model. *Mobile Information Systems*, 2022, Article ID 5197346. DOI: 10.1155/2022/5197346