TomDetLeaf: A Realistic Multi-Source Dataset for Real-Time Tomato Leaf Detection

Yassmine Ben Dhiab[1], Mohamed Ould-Elhassen Aoueileyine[2], Abdallah Namoun[3]*, Ridha Bouallegue[4]
Innov'COM Laboratory-Higher School of Communication of Tunis (SUP'COM), University of Carthage, Tunis, Tunisia
AI Center-Faculty of Computer and Information Systems, Islamic University of Madinah, Madinah 42351, Saudi Arabia
Ibn Khaldoun University (UIK), Soukra, Ariana 2036, Tunisia

***Abstract*—Plant diseases remain a major threat to crop productivity, especially where timely diagnosis is difficult. This paper introduces TomDetLeaf, a new annotated dataset designed for tomato leaf detection in diverse agricultural environments, supporting the development of generalizable deep learning models for edge AI deployment. Unlike existing datasets such as PlantVillage, which consist mainly of single-leaf images captured under controlled conditions, TomDetLeaf integrates heterogeneous sources including the Taiwan dataset, climate-controlled greenhouses, hydroponic systems and farm environments. The dataset combines single-leaf and multi-leaf images, realistic backgrounds and varying illumination, addressing a key gap that limits the real-world robustness of current models. To demonstrate its utility, we trained and evaluated YOLOv8 on both the original Taiwan dataset and our proposed TomDetLeaf. Results show that YOLOv8 trained on TomDetLeaf achieved 88.3% mAP@0.5, 81.8% precision, and 82.7% recall, exceeding the Taiwan-subset baseline of 77.4% mAP@0.5, 81.6% precision, and 67.6% recall. This validates the contribution of TomDetLeaf in improving detection accuracy and generalization under realistic conditions. By providing a diverse, deployment-ready dataset, this work bridges the gap between theoretical benchmarks and practical real-time applications.**

***Keywords*—Tomato leaf detection; smart agriculture; dataset; tomato leaf dataset; real-time inference; Edge AI; object detection**

I. Introduction

Ensuring tomato crop health is vital for maintaining sustainable agricultural production and global food security [1]. Tomato plants are highly susceptible to multiple foliar diseases that cause severe yield losses and economic constraints for farmers worldwide [2], [3]. Early and accurate detection of diseased leaves is therefore critical for timely intervention and disease management. Conventional detection methods, such as expert visual inspections or laboratory analysis, are limited by subjectivity, high costs and the lack of accessibility in rural or resource-constrained regions [4].

In recent years, the combination of deep learning (DL) and computer vision has shown promising results for plant disease recognition. Convolutional neural networks (CNNs), trained on large-scale datasets such as PlantVillage, have achieved remarkable classification accuracy, often exceeding 99% [5]. However, these models typically rely on synthetic or laboratory-controlled images with plain backgrounds and uniform lighting, conditions that do not reflect the complexity of real-world agricultural environments. As a result, models trained on such data may perform well in academic benchmarks but fail to generalize when deployed in greenhouses, hydroponic systems or open-field farms [6].

To address these limitations, there is a pressing need for annotated datasets that capture the diversity of real-world conditions, including single and multiple leaves per image, heterogeneous backgrounds, variable lighting and naturally occurring noise [7]. Without such datasets, AI systems risk providing artificially inflated results during testing while underperforming in practical deployment scenarios.

Despite the success of existing datasets such as PlantVillage [8], their controlled acquisition conditions, uniform backgrounds, isolated leaves and synthetic lighting often lead to inflated accuracy that does not generalize to real-world deployment. Models trained on such data frequently fail when tested in farms, greenhouses, or hydroponic environments where multiple leaves overlap, backgrounds vary and environmental noise is present [9]. This critical gap between laboratory benchmarks and field performance motivates the development of TomDetLeaf, a dataset explicitly designed to capture real-world complexity and support the creation of robust detection models for edge deployment [10].

This work introduces TomDetLeaf, a curated dataset designed to bridge this gap by enabling robust tomato leaf detection in diverse contexts. Unlike existing resources, TomDetLeaf merges images from multiple sources, including the Taiwan dataset, a climate-controlled greenhouse and hydroponic environments and it is complemented with distractor images to enhance model generalization [11]. Each tomato leaf in the dataset is annotated for detection tasks, providing a realistic foundation for developing edge-ready models. To demonstrate its effectiveness, we trained YOLOv8 on TomDetLeaf and compared its performance with the same model trained on the original Taiwan dataset [12]. Results show that TomDetLeaf enables a substantial improvement in both mean Average Precision (mAP) and recall, while achieving stronger robustness in real-world scenarios.

Beyond dataset development, this work sets the foundation for an end-to-end pipeline combining detection and classification. In future work, the detected tomato leaves will be cropped and passed to a lightweight classification model capable of distinguishing between healthy and multiple disease categories in real time. The pipeline will be designed for integration into autonomous platforms, such as agricultural robots or UAVs, enabling continuous monitoring of crops directly in the field [13], [14]. This vision positions TomDetLeaf not only as a

*Corresponding authors.

dataset, but as a key enabler of embedded AI solutions for smart agriculture.

The main contributions of this paper are summarized as follows:

- Development of TomDetLeaf, a new annotated dataset for tomato leaf detection, integrating images from controlled and real-world environments with diverse backgrounds, lighting and leaf arrangements.
- Cleaning and curation of existing data sources, including the Taiwan dataset, by removing low-quality and duplicate samples to enhance reliability.
- Incorporation of greenhouse, hydroponic and distractor images, ensuring robustness against non-ideal and heterogeneous agricultural conditions.
- Comprehensive evaluation with YOLOv8, showing significant performance improvements over baseline datasets, confirming the dataset's value for real-time deployment in edge-based agricultural applications.

By filling the dataset gap and validating its utility through rigorous experiments, TomDetLeaf lays the groundwork for reliable end-to-end systems where detection of tomato leaves can serve as the first step toward subsequent disease classification and full integration into autonomous agricultural platforms.

To situate the dataset within a complete decision workflow, Fig. 4 presents the conceptual architecture of the intended system. It comprises image acquisition, tomato-leaf detection with YOLOv8, cropping of detected leaves, a lightweight disease classification module and a final decision layer that reports health status and disease type. The present study empirically evaluates the detection component; the classification component is outlined as future work.

II. PROPOSED DATASET: TOMDETLEAF

The TomDetLeaf dataset was designed to provide a realistic and diverse benchmark for tomato leaf detection, with the specific objective of improving the generalization of deep learning models in real-world agricultural contexts. Unlike existing datasets, which are often captured under controlled laboratory conditions with plain backgrounds, TomDetLeaf integrates heterogeneous sources and environmental conditions [15], [16]. This deliberate design ensures that the dataset reflects the variability of actual agricultural fields, where lighting, leaf density and background noise are far less predictable. The construction process follows a structured pipeline that includes data selection, cleaning, integration of new sources, annotation and the addition of distractor images [17].

A. Taiwan Subset and Cleaning

The first source of TomDetLeaf is the publicly available Taiwan tomato dataset [18], which contains both single-leaf and multi-leaf images. From this dataset, only a relevant subset was retained. The raw images were manually cleaned to remove duplicates, low-quality samples and blurred images that could negatively affect model learning. This cleaning process ensured that the retained data maintained reliability and quality [19]. The Taiwan subset serves as the foundation of TomDetLeaf, contributing a balanced mixture of isolated and overlapping tomato leaves under varying capture conditions.

B. Greenhouse Captures

To complement the Taiwan subset, we collected new high-resolution images from a climate-controlled greenhouse. These images introduce diversity in terms of background textures, controlled yet variable lighting and different plant growth stages. Including greenhouse data ensures that the dataset better reflects the conditions of commercial tomato cultivation systems, where disease detection systems are increasingly deployed. The controlled acquisition process in this environment also provided consistent labeling opportunities, while maintaining realistic challenges such as shadows, leaf overlap and reflections [20].

C. Hydroponic Captures

Another significant contribution to TomDetLeaf comes from hydroponic environments, where tomato plants are grown in nutrient solutions without soil. The images captured in this context differ substantially from those in both controlled greenhouse and field scenarios. Hydroponic systems often include artificial lighting, reflective surfaces and pipes or structural elements in the background, which add to the complexity of the detection task [20]. Incorporating these samples enhances the dataset's heterogeneity and improves the likelihood that trained models will generalize to a broad spectrum of agricultural settings [21].

D. Distractor Images

In addition to tomato-specific samples, TomDetLeaf includes a curated set of distractor images consisting of non-tomato leaves, stems and other irrelevant agricultural objects. These negative examples are essential to reduce model bias and prevent false positives during deployment. In real-world conditions, object detection systems are frequently exposed to cluttered backgrounds containing various types of vegetation and objects unrelated to tomato crops. By incorporating distractor images into the training process, TomDetLeaf strengthens the ability of models to distinguish tomato leaves from irrelevant elements, thereby improving robustness in practical applications.

E. Annotation Process

All images were consistently annotated using the Roboflow platform. Bounding boxes were applied to each tomato leaf with a single unified label, Tomato-leaf. The decision to annotate only one class simplifies the dataset for the detection stage, which we consider the foundation of an end-to-end pipeline. This modular approach allows subsequent classification models to focus exclusively on cropped leaf regions, where disease type or health status can later be determined. The consistency of annotations across heterogeneous sources ensures clarity and usability for researchers and developers aiming to deploy detection systems on embedded devices.

F. Dataset Overview

Fig. 1 presents the complete construction pipeline of TomDetLeaf, while Fig. 2 shows representative examples of images drawn from both the Taiwan dataset and the proposed TomDetLeaf dataset. The latter clearly demonstrates its

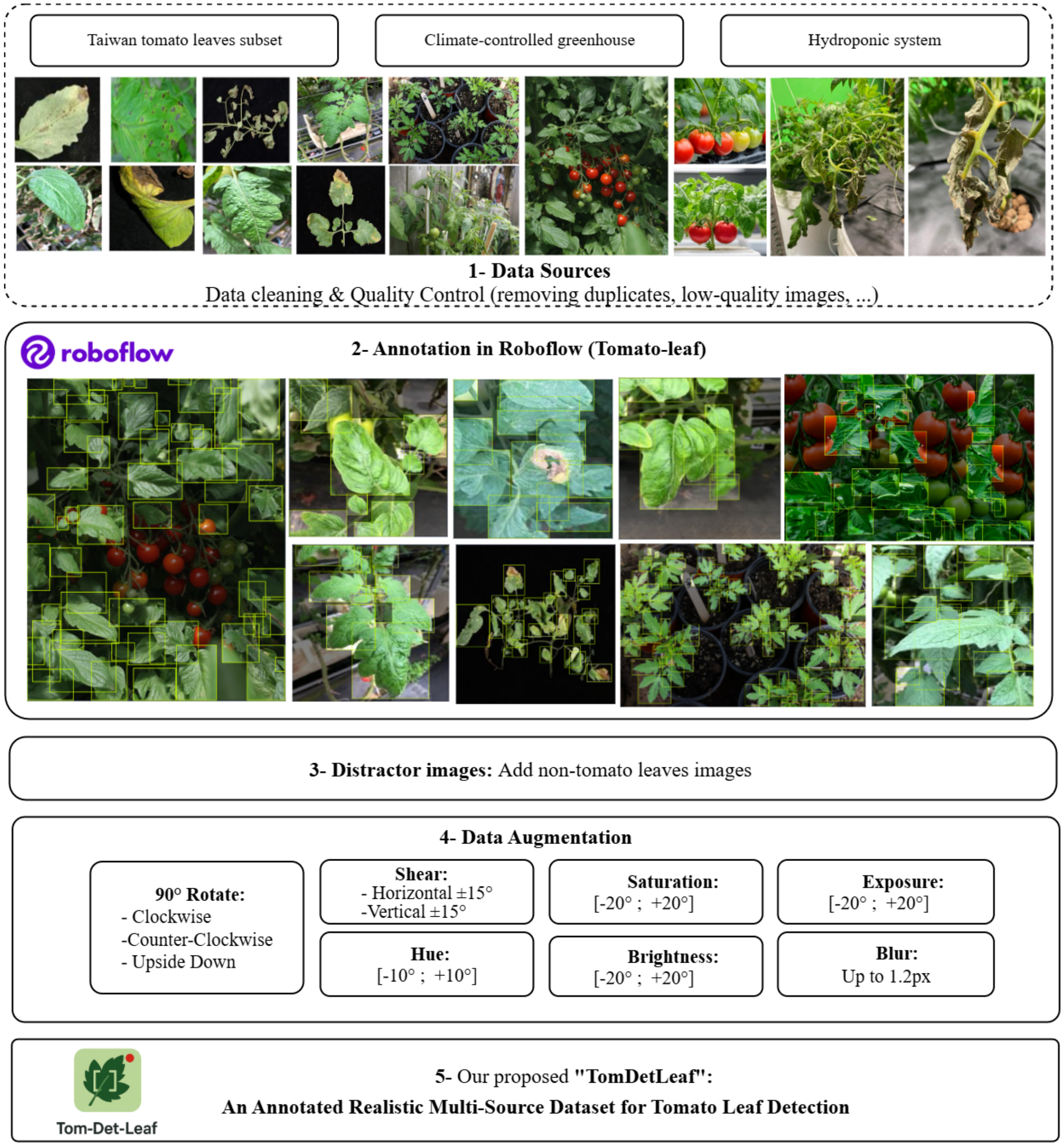


Fig. 1. Construction pipeline of the proposed TomDetLeaf dataset. The dataset integrates cleaned subsets of the Taiwan dataset with new captures from climate-controlled greenhouses and hydroponic systems, complemented by distractor images and consistently annotated for tomato leaf detection.

enhanced variability, combining single and multiple leaves, natural and artificial backgrounds and diverse environmental conditions. This heterogeneity makes TomDetLeaf a unique and valuable resource for training and evaluating models intended for TinyML and Edge AI deployments in agriculture [22], [23].

G. Dataset Statistics

Table I summarizes the composition of TomDetLeaf after cleaning. Counts reflect the final pool used prior to splitting.

The present study addresses leaf detection only, i.e. all leaves share a single Tomato-leaf class. The healthy/diseased counts in Table I were not used during training or evaluation; they are reported to show coverage across environments and

TABLE I. TomDetLeaf Composition by Source after Cleaning. H: Healthy, D: Diseased

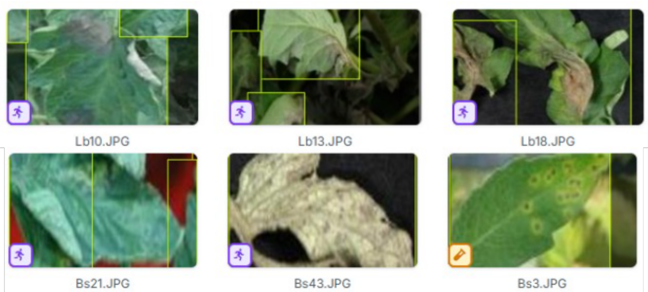
Source	Images	H	D
Taiwan subset (cleaned)	417	63	354
Greenhouse captures	182	78	104
Hydroponic captures	127	74	53
Distractor scenes (negatives)	34	–	–
Total	**760**	**215**	**511**

to motivate the forthcoming classification release.

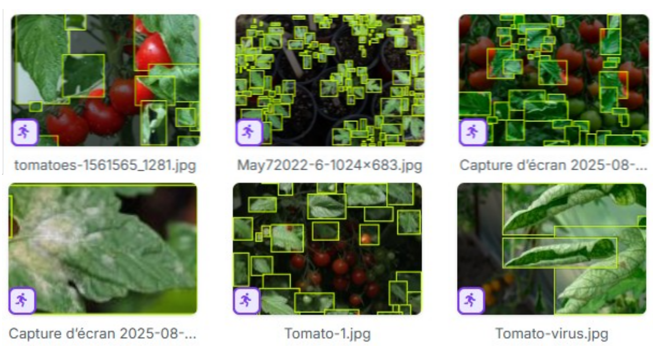
To indicate scene complexity at a glance, TomDetLeaf (non-augmented) comprises approximately one third single-leaf images and two thirds multi-leaf images (see Table II).

TABLE II. SINGLE- VS. MULTI-LEAF COMPOSITION OF TOMDETLEAF (NON-AUGMENTED)

Scene type	**Share**	**Images (of 760)**
Single-leaf	≈33%	≈250
Multi-leaf	≈67%	≈510



(a) Samples from the Taiwan dataset



(b) Samples from the proposed TomDetLeaf dataset

Fig. 2. Examples of tomato leaf images: (a) Samples from the Taiwan dataset, (b) Samples from the proposed TomDetLeaf dataset. The TomDetLeaf dataset integrates single and multiple leaves, diverse backgrounds and images from real-world environments such as farms, climate-controlled greenhouses and hydroponics.

PlantVillage (tomato subset) [8] is a large classification corpus of single-leaf images captured on largely uniform backgrounds with image-level labels only (no bounding boxes). By design it is not detection-ready and would require re-annotation to train a detector. In contrast, TomDetLeaf provides bounding boxes for a single Tomato-leaf class, mixes single- and multi-leaf scenes (approximately one third vs. two thirds), and includes complex, real-world backgrounds across greenhouse and hydroponic settings. Consequently, direct mAP comparison to PlantVillage is not meaningful; it serves instead as a scale reference for classification. Key differences are summarized in Table III.

III. EXPERIMENTAL SETUP

To evaluate the effectiveness of the proposed TomDetLeaf dataset, we conducted experiments using YOLOv8, a state-of-the-art object detection model widely adopted for lightweight and real-time applications. YOLOv8 was selected due to its balance between accuracy and computational efficiency, making it suitable for assessing detection performance in agricultural edge scenarios. The implementation was based on the Ultralytics YOLOv8 framework with PyTorch backend.

TABLE III. TOMDETLEAF VS. PLANTVILLAGE (TOMATO SUBSET)

Aspect	TomDetLeaf	PlantVillage–Tomato
Primary task	Detection	Classification
Annotation	Boxes (1 class)	Image labels
Leaf multiplicity	Single+Multi (≈1/3:≈2/3)	Single only
Backgrounds	Complex, real-world	Plain, uniform
Environments	Greenhouse, hydroponic, farm	Controlled
Images (raw)	760	18,164
Detection-ready	Yes	No (re-annotation)

A. Datasets for Comparison

Two datasets were considered in this study:

1) Taiwan dataset: A publicly available dataset of tomato plants [18], annotated to include a single bounding-box class labeled Tomato-leaf. Only the relevant subset was retained, with noisy, poor-quality and duplicate images removed.

2) TomDetLeaf dataset (Proposed): A newly curated dataset combining selected and cleaned images from the Taiwan dataset with additional images collected from climate-controlled greenhouses and hydroponic environments. The dataset captures a wider variety of conditions, including single-leaf and multi-leaf settings, diverse backgrounds and real-world environmental variability. All images were annotated with a single class Tomato-leaf.

For both datasets the detector is health-agnostic (single Tomato-leaf class); the healthy/diseased composition in Table I does not alter the training labels used here.

B. Training Setup

Both datasets were trained under identical conditions to ensure a fair comparison. Images were resized to **225×225** and the data were split **70%/20%/10%** into training, validation, and test sets (**532/152/76** images, respectively). Training ran for **100** epochs with a batch size of **64**, using the **Adam** optimizer (initial learning rate **0.001**). Data augmentation was applied to the *training* split only (random horizontal flip, mild geometric transforms and photometric jitter) and performed online at load time, yielding an effective **3**× increase in training samples per epoch; this corresponds to **1,596** effective training images and an overall effective dataset size of **1,824** when counting augmented samples, while validation and test sets remained unchanged.

C. Evaluation Metrics

The performance of the object detection models was evaluated using three standard metrics: Precision, Recall and mean Average Precision (mAP) [24], [25]. These metrics are widely adopted in computer vision benchmarks and provide complementary insights into model reliability and robustness in real-world scenarios.

Precision measures the ability of the model to correctly identify tomato leaves among all predicted bounding boxes. It is defined as:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{1}$$

where, TP denotes true positives (correct detections) and FP false positives (incorrect detections). A higher precision indicates that fewer non-leaf objects are misclassified as leaves.

Recall quantifies the proportion of correctly detected tomato leaves among all ground-truth annotations:

$$\text{Recall} = \frac{TP}{TP + FN} \tag{2}$$

where, FN represents false negatives (missed detections). High recall is particularly critical in agricultural monitoring since missing diseased leaves may prevent timely intervention.

Mean Average Precision (mAP) provides a holistic measure that balances precision and recall. It is based on the area under the Precision–Recall (PR) curve. For a single class, the Average Precision (AP) is computed as:

$$AP = \int_0^1 P(R), \mathrm{d}R \tag{3}$$

where, $P(R)$ is the precision as a function of recall. For multi-class tasks, the mean of AP across all classes yields the mAP. In practice, different intersection-over-union (IoU) thresholds are applied to determine whether a predicted bounding box matches a ground-truth object.

$mAP@50$ corresponds to AP calculated at a fixed IoU threshold of 0.5, following the Pascal VOC standard. It is relatively lenient, considering a detection correct if the overlap with ground truth is at least 50%.

$mAP@90$ applies a stricter threshold of 0.9, requiring near-perfect overlap between predicted and ground-truth bounding boxes.

$mAP@[.5:.95]$, the COCO standard, averages AP over IoU thresholds ranging from 0.5 to 0.95 in increments of 0.05. This provides a more comprehensive assessment of detection quality across varying levels of localization.

By reporting precision, recall and mAP under multiple IoU thresholds, we provide a complete picture of detection performance, highlighting both the ability to identify leaves accurately (precision) and to avoid missing them (recall), as well as the localization accuracy of bounding boxes (mAP). Unless otherwise stated, reported mAP corresponds to mAP@0.5 (Pascal VOC).

TABLE IV. Detection Performance of YOLOv8 Trained on Taiwan and TomDetLeaf Datasets

Dataset	Precision (%)	Recall (%)	mAP@0.5 (%)
Taiwan dataset	81.6	67.6	77.4
TomDetLeaf (Proposed)	81.8	82.7	88.3

IV. Results

Table IV presents the comparative performance of YOLOv8 [26] trained on the Taiwan dataset and on the proposed TomDetLeaf dataset.

The results demonstrate that TomDetLeaf enables significantly stronger detection performance, with an improvement of +10.9% mAP@0.5 and +15.1% recall compared to the Taiwan dataset. While precision remained similar, recall improved markedly, indicating that models trained on TomDetLeaf capture a larger proportion of true positives under diverse conditions.

An important observation from our experiments is the significant improvement in recall (+15.1%) when training YOLOv8 on TomDetLeaf compared to the Taiwan dataset. In agricultural disease monitoring, recall is particularly critical: missing a diseased leaf can allow pathogens to spread unchecked, whereas a false alarm (lower precision) may only result in an unnecessary inspection. Thus, the higher recall achieved by TomDetLeaf-trained models underscores its value for real-world disease detection systems, where early and reliable identification of all potentially diseased leaves is more important than minimizing false positives.

Qualitative results are illustrated in Fig. 3, showing YOLOv8 predictions on complex real-world tomato plant images. Models trained on TomDetLeaf generalize better to challenging backgrounds, overlapping leaves and mixed plant structures, highlighting the importance of dataset diversity for real-world deployment.

A. *Error Analysis and Cross-Dataset Checks*

We examined generalization and typical failure modes under the same evaluation protocol as Table IV. A model trained on TomDetLeaf maintains its performance on complex scenes and attains even higher scores on the Taiwan test set, which largely contains simpler, single-leaf images and plain backgrounds. Conversely, a model trained on the Taiwan subset degrades on TomDetLeaf due to background clutter, occlusion and mixed structures.

TABLE V. Cross-Dataset Evaluation (mAP@0.5, Precision, Recall)

Training set	Evaluation set	mAP@0.5 (%)	Precision (%)	Recall (%)
TomDetLeaf	TomDetLeaf	88.3	81.8	82.7
TomDetLeaf	Taiwan	92.1	86.4	88.9
Taiwan	Taiwan	77.4	81.6	67.6
Taiwan	TomDetLeaf	65.2	79.8	55.1

Qualitative inspection indicates that false negatives are primarily associated with heavy overlap/occlusion in multi-leaf clusters (approximately 40% of missed instances), small

Fig. 3. YOLOv8 detection results on sample images from the proposed TomDetLeaf dataset. The model demonstrates accurate detection in complex backgrounds, validating the robustness of the dataset.

leaves (height $< 32\,\text{px}$; approximately 30%), backlit or low-contrast foliage against bright backgrounds (approximately 20%), and truncation at image borders (approximately 10%). False positives arise mainly from confusion with petioles/stems and tomato fruit (approximately 50%), non-tomato foliage present in aisles or background (approximately 30%) and specular reflections from hydroponic hardware (approximately 20%). We also observed instances where nearby leaves are suppressed during non-maximum suppression in very crowded scenes, which selectively reduces recall in multi-leaf images, the scenario TomDetLeaf was designed to represent.

PlantVillage provides image-level labels only; it is therefore unsuitable for a formal detection benchmark without additional bounding-box annotation. A qualitative probe on single-leaf PlantVillage images suggested that the TomDetLeaf-trained detector localizes the visible leaf reliably, but we do not report mAP for this corpus.

V. Discussion

Fig. 4 summarizes the envisaged end-to-end pipeline for tomato-leaf monitoring. The detector isolates candidate leaves that are subsequently provided to a lightweight classifier. This modular design supports embedded deployment because detection and classification can be optimized independently. The experiments reported here target the detection stage.

In addition, the cross-dataset results in Table V show that models trained on TomDetLeaf transfer reliably to simpler domains (Taiwan), whereas the inverse transfer degrades markedly, underscoring the value of training on realistic, cluttered scenes.

Across all capture settings, both healthy and diseased foliage are present (Table I), which reduces environment–label confounding and supports the observed generalization.

Together with the single- versus multi-leaf composition (Table II), this coverage limits scene-type bias and helps explain the improved recall on complex images.

The experimental results highlight the importance of dataset diversity and realism for developing object detection models that generalize to real-world agricultural environments. Although the Taiwan dataset is commonly used in plant leaf detection tasks, its limited variability in terms of background conditions and image diversity constrained the performance of YOLOv8, particularly in recall (67.6%). This weakness became evident when testing in complex scenarios with multiple overlapping leaves, heterogeneous lighting and natural backgrounds. Models trained on such restricted datasets tend to achieve high accuracy in controlled settings but fail to maintain robustness in deployment contexts.

By contrast, the proposed TomDetLeaf dataset demonstrated a marked improvement in detection performance, raising recall to 82.7% and overall mAP to 88.3%. This represents a relative gain of 10.9 percentage points in mAP and 15.1 percentage points in recall compared to the Taiwan dataset. The improvement in recall is particularly significant, as it indicates that fewer leaves are missed during inference, which is critical in applications such as early disease diagnosis where every missed detection can affect the effectiveness of downstream classification and decision-making.

The strength of TomDetLeaf lies in its deliberate design to reflect realistic agricultural conditions. By merging cleaned

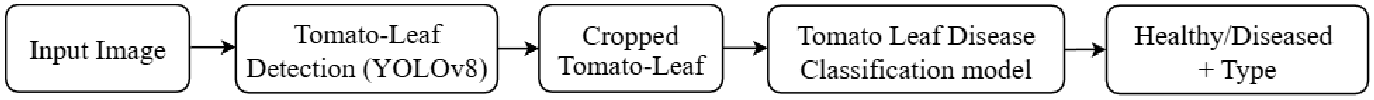


Fig. 4. Conceptual end-to-end pipeline for tomato-leaf monitoring. The system comprises image acquisition, detection with YOLOv8, cropping of detected leaves, a lightweight disease classification module and a final decision reporting health status and disease type. The present work evaluates the detection stage; the classification stage is reserved for future work.

subsets of the Taiwan dataset with additional images collected from climate-controlled greenhouses and hydroponic systems, the dataset introduces multiple layers of complexity: single-leaf and multi-leaf compositions, variations in background and differences in lighting and plant growth environments. Training on such data enables models to encounter challenging scenarios that mimic real-world deployment, thereby enhancing robustness and generalization.

Another important aspect is that TomDetLeaf focuses on leaf-level detection as a first stage of an end-to-end pipeline for plant health monitoring. In practice, robust leaf detection forms the foundation for subsequent disease classification, which can be performed on cropped leaf regions. This approach is well aligned with embedded and TinyML systems, where modular detection–classification pipelines allow real-time inference on resource-constrained edge devices such as robots or UAVs [27], [28]. Therefore, the dataset not only improves detection metrics but also provides a realistic training ground for the next step of integrating lightweight disease classification models [29].

Two practical limitations are worth noting. First, the current release uses a single detection class without disease labels; these labels are planned for the next version. Second, images are resized to 225×225, which can under-represent very small leaves; increasing the input resolution, adopting multi-scale training, or tuning the NMS IoU threshold are straightforward mitigations for future experiments.

Finally, qualitative results confirm that TomDetLeaf-trained models are able to detect tomato leaves under highly cluttered, natural conditions, with consistent bounding-box precision even in the presence of fruits, stems and overlapping vegetation. This demonstrates the practical advantage of training with diverse and realistic data, bridging the gap between controlled research benchmarks and real-world deployment in smart agriculture [30].

VI. Conclusion and Future Work

This work introduced TomDetLeaf, a customized and diverse tomato leaf detection dataset designed to improve the robustness and generalization of object detection models for real-world agricultural applications. By integrating cleaned subsets of existing datasets with newly collected images from climate-controlled greenhouses and hydroponic systems, TomDetLeaf provides a richer representation of realistic field conditions, including single- and multi-leaf scenarios, varied lighting and heterogeneous backgrounds. Experimental evaluation with YOLOv8 demonstrated that models trained on TomDetLeaf outperformed those trained on the Taiwan dataset, achieving an mAP of 88.3% and recall of 82.7%, compared to 77.4% and 67.6%, respectively. These results underline the importance of dataset realism in enabling reliable real-time deployment. Cross-dataset checks further confirmed that TomDetLeaf-trained detectors generalize to simpler domains, while reverse transfer degrades on cluttered scenes.

The dataset is intended as a foundation for modular end-to-end pipelines in smart agriculture, where reliable leaf detection constitutes the first stage of disease monitoring on embedded and edge devices. To facilitate downstream benchmarking, we also report per-source healthy/diseased counts and the single- versus multi-leaf composition to support balanced, reproducible splits. By improving detection accuracy and recall, TomDetLeaf contributes to enabling practical, low-cost solutions for autonomous platforms such as robots and UAVs.

Future work will extend this contribution in several directions. First, we will evaluate the dataset using additional state-of-the-art detection models to benchmark its generalizability across architectures. Second, we are preparing an expanded annotated version of TomDetLeaf that includes disease-specific labels for tomato leaves, allowing simultaneous detection and classification of healthy and diseased samples. Finally, we aim to implement and validate a complete end-to-end pipeline, combining detection and classification on embedded hardware platforms to demonstrate real-time disease diagnosis in realistic agricultural scenarios.

References

[1] Rehman, Abdul, et al. "Sustainable agricultural practices for food security and ecosystem services." Environmental Science and Pollution Research 29.56 (2022): 84076-84095.

[2] Shukla, Smriti, et al. "Challenges faced by farmers in crops production due to fungal pathogens and their effect on Indian economy." Fungal diversity, ecology and control management. Singapore: Springer Nature Singapore, 2022. 495-505.

[3] Singhal, Pankhuri, et al. "Mixed infection of plant viruses: diagnostics, interactions and impact on host." Journal of Plant Diseases and Protection 128.2 (2021): 353-368.

[4] Thangaraj, Rajasekaran, et al. "Artificial intelligence in tomato leaf disease detection: a comprehensive review and discussion." Journal of Plant Diseases and Protection 129.3 (2022): 469-488.

[5] Dhaka, Vijaypal Singh, et al. "A survey of deep convolutional neural networks applied for prediction of plant leaf diseases." Sensors 21.14 (2021): 4749.

[6] Xu, Mingle, et al. "Embracing limited and imperfect training datasets: opportunities and challenges in plant disease recognition using deep learning." Frontiers in Plant Science 14 (2023): 1225409.

[7] Cheng, Hsueh-Hung, et al. "Identifying tomato leaf diseases under real field conditions using convolutional neural networks and a chatbot." Computers and Electronics in Agriculture 202 (2022): 107365.

[8] Hughes, David, and Marcel Salathé. "An open access repository of images on plant health to enable the development of mobile disease diagnostics." arXiv preprint arXiv:1511.08060 (2015).

[9] Bhargava, Anuja, et al. "Plant leaf disease detection, classification, and diagnosis using computer vision and artificial intelligence: A review." IEEE access 12 (2024): 37443-37469.

[10] Akbar, Jalal Uddin Md, et al. ”A comprehensive review on deep learning assisted computer vision techniques for smart greenhouse agriculture.” IEEE Access 12 (2024): 4485-4522.

[11] Fuentes-Peñailillo, Fernando, et al. ”New generation sustainable technologies for soilless vegetable production.” Horticulturae 10.1 (2024): 49.

[12] Hussain, Muhammad. ”Yolov5, yolov8 and yolov10: The go-to detectors for real-time vision.” arXiv preprint arXiv:2407.02988 (2024).

[13] Lytridis, Chris, et al. ”An overview of cooperative robotics in agriculture.” Agronomy 11.9 (2021): 1818.

[14] Kim, Jeongeun, et al. ”Unmanned aerial vehicles in agriculture: A review of perspective of platform, control, and applications.” Ieee Access 7 (2019): 105100-105115.

[15] Nobori, Tatsuya, and Kenichi Tsuda. ”The plant immune system in heterogeneous environments.” Current Opinion in Plant Biology 50 (2019): 58-66.

[16] Thangaraj, Rajasekaran, et al. ”Artificial intelligence in tomato leaf disease detection: a comprehensive review and discussion.” Journal of Plant Diseases and Protection 129.3 (2022): 469-488.

[17] Zhu, Zheng, et al. ”Distractor-aware siamese networks for visual object tracking.” Proceedings of the European conference on computer vision (ECCV). 2018.

[18] Huang, Mei-Ling, and Ya-Han Chang. ”Dataset of tomato leaves.” Mendeley Data 1.10.17632 (2020).

[19] García, Salvador, et al. ”Big data preprocessing: methods and prospects.” Big data analytics 1.1 (2016): 9.

[20] Malabadi, Ravindra B., et al. ”Greenhouse farming: Hydroponic vertical farming-Internet of Things (IOT) Technologies: An updated review.” World Journal of Advanced Research and Reviews 23.02 (2024): 2634-2686.

[21] Ragaveena, S., A. Shirly Edward, and U. Surendran. ”Smart controlled environment agriculture methods: A holistic review.” Reviews in Environmental Science and Bio/Technology 20.4 (2021): 887-913.

[22] Hizem, Moez, et al. ”Reliable ECG Anomaly Detection on Edge Devices for Internet of Medical Things Applications.” Sensors 25.8 (2025): 2496.

[23] Dhiab, Yassmine Ben, et al. ”Edge-based human activity recognition: A novel approach using spectral analysis and deep learning.” 2024 International Wireless Communications and Mobile Computing (IWCMC). IEEE, 2024.

[24] Goutte, Cyril, and Eric Gaussier. ”A probabilistic interpretation of precision, recall and F-score, with implication for evaluation.” European conference on information retrieval. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005.

[25] Yue, Yisong, et al. ”A support vector method for optimizing average precision.” Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval. 2007.

[26] Hussain, Muhammad. ”YOLO-v1 to YOLO-v8, the rise of YOLO and its complementary nature toward digital manufacturing and industrial defect detection.” Machines 11.7 (2023): 677.

[27] Elhanashi, Abdussalam, et al. ”Advancements in TinyML: Applications, limitations, and impact on IoT devices.” Electronics 13.17 (2024): 3562.

[28] Jouini, Oumayma, et al. ”A survey of machine learning in edge computing: Techniques, frameworks, applications, issues, and research directions.” Technologies 12.6 (2024): 81.

[29] Khan, Zohaib, Yue Shen, and Hui Liu. ”ObjectDetection in Agriculture: A Comprehensive Review of Methods, Applications, Challenges, and Future Directions.” Agriculture 15.13 (2025): 1351.

[30] Soussi, Abdellatif, et al. ”Smart sensors and smart data for precision agriculture: a review.” Sensors 24.8 (2024): 2647.