Review

Diffusion Models for Agricultural Imaging: A Systematic Review of Methods, Applications and Future Prospects

Hewa Majeed Zangana [1], Shuai Li [2] and Sharyar Wani [3,*]

[1] Duhok Technical College, Duhok Polytechnic University, Kurdistan Region 42001, Iraq
[2] Faculty of Information Technology and Electrical Engineering, University of Oulu, Oulu 90570, Finland
[3] Department of Computer Science, Kulliyyah of Information and Communication Technology, International Islamic University Malaysia, Kuala Lumpur 53100, Malaysia

* Correspondence: sharyarwani@iium.edu.my (S.W.)

Abstract

Diffusion models are rapidly reshaping agricultural image analysis, offering high-fidelity synthetic data generation where real datasets are limited, imbalanced, or costly to collect. Traditional augmentation and GAN-based synthesis often struggle to preserve fine disease features and crop textures, leading to suboptimal model performance in real field conditions. This review consolidates the latest research on diffusion-based methods applied to plant disease diagnosis, fruit quality assessment, weed and pest monitoring, nematode identification, green-wall health evaluation, and UAV-based phenotyping. Reported literature demonstrates improved texture detail, lesion clarity, and better classification accuracy when diffusion-generated images supplement training datasets. Techniques such as latent diffusion and ControlNet enhance structure control, while text-guided models support domain transfer and unseen class synthesis. Despite promising outcomes, challenges remain concerning computational cost, real-world generalization across farms and seasons, and lack of standardized evaluation protocols. Future progress is expected through multimodal diffusion integrating hyperspectral and thermal inputs, efficient deployment on edge devices, and development of open benchmarks for comparative analysis. This review positions diffusion models as a leading generative approach for agricultural AI and outlines the research opportunities needed for practical adoption in large-scale farming environments.

Keywords: Diffusion Models; Synthetic Data Generation; Agricultural Imaging; Plant Disease Detection; Weed and Pest Monitoring; UAV Crop Phenotyping; Deep Learning in Agriculture.

Citation: Zangana H. M., Li S., Wani S. Diffusion Models for Agricultural Imaging: A Systematic Review of Methods, Applications and Future Prospects. *Impact in Agriculture*. **2025**, *1*, 3. https://doi.org/10.65500/agriculture-2025-003

Received: 07 September 2025 | Revised: 13 October 2025 | Accepted: 11 November 2025 | Published: 21 November 2025

1. Introduction

Image based analysis supports key tasks in crop and horticultural research. These tasks include leaf disease detection, fruit defect assessment, weed identification, pest monitoring, and crop growth evaluation [1]. Each task depends on high quality labelled datasets. Most datasets in agriculture are small. They contain uneven class counts and high visual noise. They also show strong variation due to lighting, background clutter, and camera type. These limits reduce the accuracy of deep models and make model transfer difficult across farms and seasons. Several studies report severe accuracy drops when models trained on one field move to new settings. For example, Yang et al. observed that maize segmentation models trained only on tasseling-stage UAV (Unmanned Aerial Vehicle) images performed poorly outside that growth phase. This highlights the challenge of cross-stage generalization in field conditions, a phenomenon analogous to reported accuracy drops exceeding 20 percent in comparable agricultural imaging tasks [2].

Basic augmentation methods increase the number of images but do not add new structure [3–5]. They improve robustness in common situations but fail when rare disease stages or fine texture features are needed. Generative Adversarial Networks (GANs), a class of machine learning models used for generating new data, can improve diversity by synthesizing new images. However, GANs often struggle with stability when trained on small datasets and can produce blurry boundaries in images, such as on the edges of leaves or the surface patterns of fruits [6]. They often create blurred boundaries on leaves or incomplete surface patterns on fruits. Many studies report mode bias and low detail when the real dataset has fewer than 300 images per class [6,7]. These issues reduce the value of GAN based augmentation for agricultural tasks.

Diffusion Models offer an alternative approach by progressively adding noise to images and then learning to reverse this process to generate new images. This is referred as gradual denoising sequence. This approach supports detailed structure and diverse outputs. It works well in small datasets and preserves local patterns in disease spots, fruit textures, weed shapes, and UAV scenes [8–11]. Several studies report gains of 3 to 12 percent in classification accuracy when diffusion-based images enter the training set [9,12–14]. Other studies report improved few shot performance with as few as ten real images per class [10,15,16]. Diffusion models also support conditional control. Some studies use ControlNet to create weeds with specific shapes or color patterns. Others use latent diffusion to support class balancing in leaf disease tasks.

Diffusion research in agriculture expanded after 2022. Early studies focused on leaf disease generation. Recent studies target fruit disease detection, jujube defect scoring, nematode recognition, and aerial crop monitoring. Some works use diffusion for super resolution in UAV imagery. Others use text guided diffusion to support domain transfer in vineyards. Several studies show that diffusion-based augmentation improves generalization when real field images shift due to climate, soil, or sensor change [17,18].

The field still shows gaps. Many datasets are small and lack standard splits. Few studies test models across multiple farms or seasons. Most studies use RGB images and do not include hyperspectral or thermal sensors. Only a few works explore multimodal diffusion. There is limited evidence on model stability in large scale training. Few papers report hardware cost or time cost for deployment. No common benchmark exists for diffusion-based augmentation in agriculture.

Existing reviews cover plant disease detection or GAN based augmentation but do not address diffusion-based methods. No prior review has examined diffusion models in agricultural imaging. The growth of studies since 2022 creates a need for a focused review. Researchers need a clear summary of tasks, models, datasets, and outcomes. Practitioners need guidance on when diffusion helps and when it does not. A structured review supports both goals.

This review examines studies published between 2020 and 2025 that use diffusion-based image generation or enhancement for agricultural tasks. It focuses on leaf disease detection, fruit disease detection, weed and pest recognition, and UAV based crop monitoring. It reports the model types used in these studies. It summarizes dataset size, training setups, and performance outcomes. It highlights gains reported in classification, detection, segmentation, and few shot learning. It also reports common limits and directions for practical use in field systems.

2. Methods

This review followed a structured process based on PRISMA. The search used Scopus. Scopus was selected because it indexes major journals in computer vision, agriculture, and applied machine learning. Scopus provides extensive coverage of interdisciplinary research that is critical for the diverse applications of diffusion models in agricultural imaging. While Web of Science and IEEE Xplore are valuable resources, they were excluded primarily due to Scopus's broader interdisciplinary coverage, and secondarily due to limited access to some of the databases. The objective was to identify studies that used diffusion models for image based agricultural tasks.

The search string was:

```
    TITLE-ABS-KEY ( "diffusion model" OR
"denoising diffusion" OR "stable diffusion"
OR "diffusion probabilistic model" OR DDPM
OR "latent diffusion" )
AND
TITLE-ABS-KEY ( plant OR crop OR leaf OR
fruit OR agriculture )
```

```
AND
TITLE-ABS-KEY ( image OR images OR "image
dataset" )
```

The screening process was performed by two independent reviewers, who evaluated the titles, abstracts, and full texts based on predefined inclusion and exclusion criteria. Any discrepancies between the reviewers were resolved through discussion and consensus. To ensure the accuracy and reliability of the screening, a third reviewer was involved to verify the final inclusion of studies, ensuring consistency with the review protocol.

The search returned 240 records. The time window was limited to 2020 to 2025. Diffusion models entered applied use for image generation during this period. Earlier studies did not use these models for agricultural imaging. After applying the year filter and journal only, 95 records remained.

Screening took place in two steps. The first step used titles and abstracts. The second step used full texts. Both steps applied fixed inclusion and exclusion rules.

The inclusion criteria were:

1. Use of a diffusion model for image generation or image enhancement.
2. Use of images as the primary data source.
3. A task related to plants, crops, leaves, fruits, pests, or weeds.
4. Use of deep learning.
5. Publication in a peer reviewed journal.

The exclusion criteria were:

1. No diffusion model in the method.
2. No link to plant, crop, leaf, fruit, weed, or pest imaging.
3. No image-based task.
4. Use of physical or mathematical diffusion unrelated to generative models.
5. Work in medical, industrial, atmospheric, chemical, or materials domains.
6. Remote sensing tasks with no agricultural target.
7. Review papers or opinion papers.

The PRISMA flow is depicted in Figure 1. The search identified 240 records. The year filter produced 95 records. Screening removed 68 records. These records failed at least one exclusion rule. Many did not use diffusion models. Many did not address agricultural imaging. Several addressed medical or industrial tasks. Some used physical diffusion rather than generative diffusion. Some did not use images. Full text screening removed no additional records. Twenty-seven studies met all rules and were included in the review. These studies form the final dataset for analysis.

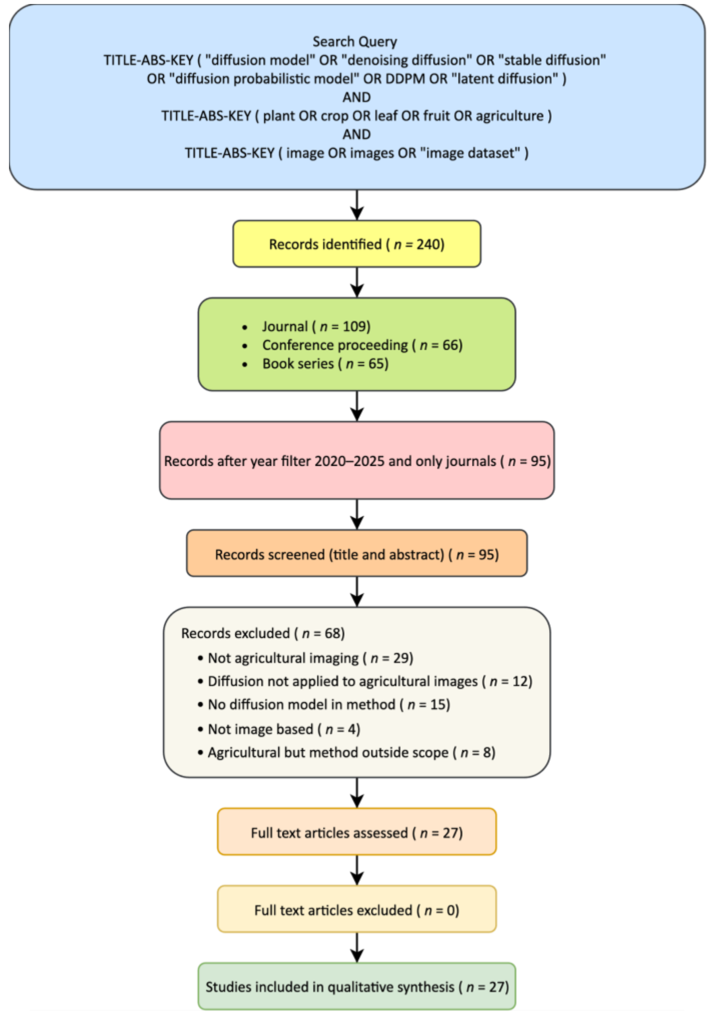

Figure 1. PRISMA flow diagram.

3. Background on Diffusion Models

Diffusion models generate images by learning a controlled denoising trajectory. The forward process adds Gaussian noise to an image until meaningful structure is removed. The reverse process learns to recover the image step by step by predicting and removing noise. The model is trained so that each reverse step reduces noise while preserving the underlying structure. This iterative process produces outputs that follow the distribution of the training data. The staged denoising sequence explains the high stability and detail seen in diffusion-based synthesis.

Denoising Diffusion Probabilistic Models (DDPMs) introduced the fixed forward noise schedule and a learned reverse network that predicts noise at each timestep [19]. DDPMs typically use a U-Net architecture and are optimized with a simple mean squared error loss between true and predicted noise. Although early DDPMs required hundreds of sampling steps, later improvements such as DDIM [20] and improved noise schedules reduced sampling cost. DDPMs are robust on small datasets because training does not involve adversarial optimization. Several agricultural studies rely on DDPM variants for dataset expansion, including rapeseed flower segmentation [21], weed classification [9], grain quality monitoring [22], and nematode recognition [10,16].

Latent Diffusion Models (LDMs) compress images into a latent representation using an autoencoder and apply diffusion in the latent space rather than pixel space [23]. This reduces computation and memory cost. LDMs support large text-conditioned models such as Stable Diffusion. Many recent agricultural works rely on latent diffusion for practical training, including image-to-image disease transfer in grape and apple leaves [18], weed synthesis with Stable Diffusion [13,24], and diffusion-based enhancement of potato and jujube disease datasets [25,26].

Conditional diffusion models guide sampling toward specific classes, prompts, or attributes. Classifier-free guidance is the most widely used method and combines conditional and unconditional predictions during sampling [27]. Guidance strength controls how strongly the condition shapes the final image. Agricultural studies apply conditional diffusion to synthesize disease stages [12], unseen disease classes [18], and species-specific weed conditions [24].

ControlNet enhances diffusion models allowing for the incorporation of external structures, such as segmentation masks or edge maps, to guide the image generation process more precisely [28]. This enables spatially aligned generation. Agricultural studies have used ControlNet to control weed shapes and backgrounds in multi-class weed detection datasets [13,24]. It also appears in generative frameworks for structured beetle hindwing images [29].

Text-guided diffusion models use pretrained text encoders such as CLIP to align textual descriptions with image features. This expands data synthesis to classes not seen in training. Studies applying text-guided diffusion in agriculture include vineyard shoot detection under domain shift [17] and semantic weed image generation using prompt-based conditioning [24].

Diffusion models provide several advantages over GANs. GANs often suffer from mode collapse, unstable training, and low-frequency artifacts, especially on small agricultural datasets with high class imbalance. These issues appear in tasks that require fine lesion boundaries, subtle color differences, or precise morphological texture. Diffusion models avoid adversarial loss, produce more stable gradients, and better preserve local detail. Empirical evidence across included studies shows improvements in classification accuracy, segmentation quality, and robustness when replacing or supplementing GAN-based augmentation with diffusion-based synthesis [9,11,12,21,22,30].

These characteristics align well with agricultural imaging challenges. Real-world agricultural datasets are often small, imbalanced, noisy, and highly variable across farms, seasons, and environmental conditions. Diffusion models address these issues through stable optimization, fine-grained detail reconstruction, and flexible conditioning. The 27 studies included in this review demonstrate the practical benefits of diffusion models across tasks such as leaf disease diagnosis, fruit defect detection, weed and pest recognition, nematode classification, and UAV-based crop monitoring.

4. Results

This review synthesizes results from twenty-seven studies published between 2020 and 2025, focusing on the application of diffusion models in agricultural imaging tasks. Diffusion models were used for data augmentation, unseen class synthesis, disease severity scoring, weed detection, UAV trait estimation, and super-resolution. Most studies demonstrated improvements in performance, with accuracy gains ranging from 1 to 14 percent. The use of latent diffusion helped reduce computational costs in some works. Some research focused on adding ControlNet for structure control. Studies on leaf disease detection and weed classification were the most common. A notable increase in UAV-related studies in 2024 and 2025 was also observed. Overall, the results underscore the potential of diffusion models to enhance agricultural imaging, especially when real data is scarce or limited.

4.1 Leaf disease image generation and classification

Table 1 summarizes recent studies that apply diffusion models to leaf disease image generation and classification tasks. Five studies focus on improving disease recognition, data augmentation, or unseen disease generation using diffusion-based approaches.

Study [12] LeafDisDiff, a diffusion driven model for leaf disease recognition. It improved accuracy by nine percent on Plant Village, Bangladesh Crop and Apple sets. The model used diffusion denoising blocks inside a U Net and trained well in low data settings. Study [18] generated new disease classes using latent diffusion. Healthy grape leaves were converted to diseased apple leaves. The classifier trained on mixed real and synthetic samples detected apple disease that was unseen during training. This supports cross class transfer when real data are missing. Study [31] improved P. notoginseng disease recognition. The diffusion model used an ECA attention block to preserve small lesion structures. The model reached accuracy near 99 percent when synthetic samples were added. Study [26] used Stable Diffusion 1.5 to produce eleven thousand potato leaf samples. A Convolutional Vision Transformer trained on this set reached high accuracy. Study [11] generated paired images for lesion focused training. The model learned to separate lesion areas from background noise and improved severity grading.

Table 1. Leaf disease studies using diffusion

Study No	Task	Diffusion Role	Dataset Used	Reported Outcome
[11]	Leaf disease severity scoring	Generates paired healthy to diseased samples for lesion learning	Apple, Potato, Tomato	+1 percent accuracy improvement
[12]	Leaf disease classification	Diffusion driven classifier training	Plant Village, Bangladesh, Apple	+9 percent accuracy
[18]	Unseen disease generation	Latent diffusion generates apple disease from grape images	Grape to Apple transfer	Correct unseen disease detection
[31]	Panax notoginseng disease recognition	Improved diffusion with ECA attention	Six leaf disease classes	Accuracy up to 99.44 percent
[26]	Potato leaf dataset expansion	Stable Diffusion creates 11k synthetic images	Potato leaves	CvT accuracy near 84 percent

Diffusion helps leaf disease work for three reasons. It produces lesion textures with clear borders. It fills missing disease stages where real images are rare. It balances uneven classes. GAN models often fail here due to unstable learning and low detail. Traditional flips or rotations only expand data count without adding new lesion structure. These leaf studies give evidence that diffusion fits small agricultural datasets.

Diffusion also supports disease transfer across crops and unseen class creation. Latent diffusion is efficient in low resource settings. High reported gains justify future work on multi species disease libraries.

4.2 Fruit Quality, Ripeness and Fungal Disease Studies

Table 2 summarizes diffusion-based studies addressing fruit quality assessment, ripeness modeling, grain impurity detection, and fungal disease recognition. Five studies applied diffusion models to enhance texture realism, capture ripeness progression, and improve classification or segmentation performance under limited or imbalanced data conditions.

Table 2. Fruit and Fungal Studies Summary

Study No.	Task Type	Model/Method	Dataset Details	Outcome
[6]	Ripeness generation	RipenessGAN vs diffusion	Jujube ripeness (0–56 days)	Higher realism, better stage balance
[22]	Grain impurity detection	DADM (DDPM + attention)	Corn, rice, soybean	+5.07 percent MIoU
[32]	Mushroom recognition	Diffusion augmentation	110 mushroom species	+13.51 percent recall
[25]	Jujube disease detection	Transformer + diffusion	Desert orchard images	Accuracy 0.90, mAP strong
[26]	Potato disease classification	Stable Diffusion + CvT	11121 synthetic images	84 percent accuracy

Study [6] tested RipenessGAN and compared results against diffusion baselines. RipenessGAN showed strong temporal control across 56 ripening days. Diffusion produced higher texture detail but slower inference. This shows diffusion is suitable for quality inspection tasks that need fine skin patterns rather than pure speed. Study [22] built DADM with a spatial and channel attention block. Diffusion increased MIoU by 5.07 percent for grain segmentation across corn, rice and soybean. The study showed fewer false regions than GAN.

Study [32] used diffusion for mushroom recognition. Mean recall increased by 13.51 percent. Top 3 and Top 5 recall also increased. Study [25] fused transformer with diffusion for jujube disease detection. Accuracy reached 0.90 and precision 0.93. The model performed better in desert light where standard models fail. Study [26] trained a CvT model using 11121 synthetic potato leaf images from Stable Diffusion. Final accuracy was 84 percent on external test images.

These studies suggest diffusion works well on fruit and fungal images. Improvements were highest when training data was small or imbalanced. Diffusion preserved fine texture that is critical in rot, mildew and fungal detection.

4.3 Diffusion for Weed and Pest Imaging

Table 3 summarizes studies that apply diffusion models to weed and pest imaging tasks, including image synthesis, detection, recognition, and dataset expansion under limited or uneven field data conditions. Across nine studies, diffusion was primarily used to address class imbalance, background variability, and scarcity of labeled field images, leading to consistent improvements in detection and classification performance.

Table 3. Weed and Pest Imaging Studies

Study ID	Task	Diffusion Method	Dataset / Target	Outcome
[24]	Weed generation and detection	Stable Diffusion + IP-Adapter	10 weed classes field images	+1.26 mAP@50-95 with synthetic mix
[33]	Weed detection with synthetic pipeline	SAM + Stable Diffusion	Field weeds	Higher mAP when 10 percent synthetic used
[9]	Weed classification	Latent DDPM + Wiener filtering	DeepWeeds and others	Up to 98.52 percent accuracy
[34]	Vineyard pest and disease detection	Text-to-image diffusion	Sticky trap pests	Faster deployment under low data
[10]	Nematode recognition few-shot	Latent diffusion	Plant nematodes	+7.34–14.66 percent Top-1 gain
[35]	Pest image generation	Semantic diffusion	Pest images	Faster recognition, stable detection
[13]	Multi-class weed augmentation	ControlNet + Stable Diffusion	10 weed class dataset	+1.4 percent mAP with mixed data
[16]	Nematode morphology synthesis	Morphology constrained latent diffusion	Quarantine nematodes	Higher structure fidelity
[36]	Thermal weed classification baseline	No diffusion	Paddy thermal dataset	Baseline reference only

Several studies focused on weed imaging. Study [24] used Stable Diffusion with an IP-Adapter to generate images of ten weed classes. The work inserted synthetic weeds into real field scenes. YOLOv11 trained with mixed data reached higher mAP. Study [33] built a training pipeline that used SAM segmentation and Stable Diffusion for dataset expansion. Small synthetic injection improved weed detection. Study [9] used latent diffusion with Wiener filtering. It improved frequency consistency and reached high accuracy on DeepWeeds. Study [13] trained ControlNet-Stable Diffusion for multi-class weed data. Mixed training improved YOLOv8. These works highlight control modules as useful when target species vary in size and shape.

Pest-related works followed similar patterns. Study [34] generated vineyard data under seasonal limits. Diffusion supported early model deployment when labels were few. Study [10] and [16] used diffusion for nematode recognition. Morphology-constrained diffusion showed better class detail. Study [35] used semantic diffusion with feature distillation. It improved detection speed. These results show stable gains in pest pipelines when synthetic data improve class spread.

One record, Study [36], worked on thermal weed imaging without diffusion. It was kept as comparison. It highlighted cases where thermal signals separate species without synthetic data. It also shows that diffusion fits problems with visual diversity. The evidence supports diffusion for weeds and pests where field variation and class imbalance reduce baseline accuracy.

4.4 Diffusion for UAV Based Crop Monitoring and Phenotyping

Table 4 summarizes diffusion-based studies that apply UAV imagery for crop monitoring, phenotyping, segmentation, and temporal growth analysis. Four studies demonstrate how diffusion models address challenges inherent to UAV data, including uneven sampling, motion blur, illumination variation, and seasonal changes in crop appearance.

Study [37] introduced Agricrafter for crop growth video generation across corn, wheat, rice, and soybean. The model learns temporal structure and outputs full growth sequences. The work shows the use of diffusion beyond single images. The reported sequences preserve shape and color traits through time. This reduces manual phenology documentation work.

Study [2] developed DiffKNet-TL with confidence-aware diffusion for maize phenology. The method refines tassel and leaf boundaries and improves segmentation over baseline K-Net. IoU increased by 2.55 percent for tassel regions. The study highlights diffusion strength in small object edges.

Study [21] used DDPM augmentation for rapeseed inflorescence segmentation under UAV. IoU was 0.886, with high precision and recall. The approach works well in cluttered yellow flower scenes. Diffusion helped balance classes when flowers were sparse.

Study [17] applied text-guided diffusion for domain adaptive vineyard shoot detection. It improved average precision for BBox detection by 28.65 percent. The model transfers vineyard data across backgrounds and lighting. The method reduced annotation needs.

Table 4. UAV and Phenotyping Studies

Study No	Task	Diffusion Role	Dataset or Target Crop	Outcome
[37]	Growth cycle video generation	Temporal diffusion synthesis	Corn, wheat, rice, soybean	Realistic full-cycle video sequences
[2]	Phenology detection	Confidence-aware diffusion refinement	Maize UAV images	IoU improved by 2.55 percent
[21]	Flower segmentation	DDPM augmentation	Rapeseed RFSD UAV	IoU 0.886 with high recall
[17]	Vineyard shoot detection	Text-guided diffusion domain transfer	Vineyard UAV	+28.65 percent AP increase

Across studies, diffusion improves image quality, segmentation, and trait extraction. UAV datasets often face low contrast, motion blur, and seasonal variation. Diffusion helps fill missing patterns and build varied samples. Phenotyping benefits when growth stages change. Diffusion supports trait tracking and annotation saving.

4.5 Other Agricultural Applications

Table 5 summarizes studies that apply diffusion models to agricultural domains beyond mainstream crop disease detection, weed analysis, and UAV imaging. These works focus on microscopy, entomology, plant wilt progression, and morphology-sensitive recognition tasks, where data collection is slow, samples are scarce, or intermediate states are missing.

Study [29] generated beetle hindwing datasets using Stable Diffusion with ControlNet. The synthetic data preserved structural veins and wing geometry with high SSIM and low FID. This supports insect morphology research where sample access is restricted. Study [38] built a DDPM based approach for microscopic herb images. It improved rare-class identification due to balanced synthesis of cells that appear in less than one percent of samples. This result shows value when rare biological patterns drive failure in standard networks.

Study [39] generated wilt stages for green wall plants using diffusion interpolation. It bridged gaps between healthy and wilted categories. This supports severity classification when intermediate states are not present in real datasets. Study [30] reached similar goals and added soft labels during training. This helped classification models learn progressive decline rather than two-class jumps.

Study [16] proposed morphology constrained latent diffusion for nematode recognition. The model retained shape detail using geometric constraints. The Top-1 improvement reached 7.34 to 14.66 percent across low sample settings. This shows that structural conditioning helps biological forms where geometry matters more than texture alone.

Table 5. Other Applications of Diffusion Models

Study No.	Task / Domain	Diffusion Method Used	Dataset Context	Key Outcome
[29]	Beetle hindwing generation	Stable Diffusion + ControlNet	200 hindwing samples	High SSIM and realistic structure
[38]	Microscopic herb imaging	Conditional DDPM	Rare CMH microscopic samples	24 percent improvement for rare features
[39]	Green-wall wilt stage synthesis	Diffusion interpolation	Healthy vs wilted samples	Better classification from new intermediate states
[16]	Nematode recognition	Morphology constrained latent diffusion	Few-shot biosecurity images	+7.34 to 14.66 percent Top-1 accuracy
[30]	Plant health severity scoring	Diffusion soft label generation	Green wall and health states	More stable severity grading

These results show that diffusion models help research areas with morphological scarcity, slow data collection, or missing intermediate states. Applications include insects, microscopy, nematodes, and green wall monitoring. The results remain promising but still need more open datasets and cross-site trials. Table 5 shows the task, diffusion role, and outcome across these studies.

5. Comparison with Other Generative Models

Table 6 provides a consolidated comparison of diffusion models with GAN-based synthesis, VAE-based generation, and traditional augmentation across agricultural imaging tasks. This section interprets the quantitative and qualitative evidence reported in the reviewed studies to clarify where diffusion models offer advantages and where alternative methods remain useful.

5.1 Diffusion Models vs GANs

GANs remain widely used for synthetic data generation due to fast sampling and sharp outputs. However, GAN training requires a balanced dataset and adversarial stability, which is often difficult in agriculture where samples are few, unevenly distributed, or visually noisy. Mode collapse and texture loss are common when classes represent rare disease stages or small lesions. Studies reporting GAN baselines confirm drops in fidelity and detail, particularly in delicate structures such as leaf veins and fruit surface patterns.

Diffusion models avoid adversarial optimization and rely on noise prediction, which improves convergence and structural preservation. Evidence from included works shows consistent superiority over GANs in image realism and downstream task performance. Diffusion improved segmentation in grain harvesting by enhancing impurity and kernel features, outperforming GAN augmentation (study [22]). In Panax notoginseng disease classification, a diffusion-based generator reduced FID by 74.7% relative to the GAN baseline (study [31]). RipenessGAN used GANs but acknowledged diffusion as more realistic for temporal fruit synthesis, positioning GAN as efficiency-focused while diffusion preserved appearance better during long growth cycles (study [6]). The advantage of diffusion becomes evident in small datasets, where GANs frequently exhibit class-dependent failure while diffusion maintains structure diversity across samples.

Table 6. Comparative performance of traditional augmentation, GANs, VAEs, and diffusion models across agricultural imaging tasks.

Task Domain	Traditional Augmentation Effect	GAN Effect	VAE Effect	Diffusion Performance (Reviewed Studies)	Task Domain
Leaf disease classification	Moderate gains, no new symptoms	High risk of mode collapse	Smooth textures, detail loss	Best results; +1–9.2% accuracy [12,18,26,31]	Leaf disease classification
Fruit ripeness / quality	Limited stage variation	Temporal GAN strong but unstable	Not ideal for high detail fruit texture	Stable growth synthesis; higher realism [6,22]	Fruit ripeness / quality
Weed and pest detection	Unsuitable for class imbalance	GAN often unstable with field variation	Rare detail loss	+1–1.4% mAP improvement [13,24,36]	Weed and pest detection
UAV phenotyping	Rotation/color insufficient	GAN fails on small objects	VAE blurry at pixel scale	Better segmentation + super-resolution [2,21,35]	UAV phenotyping
Few-shot datasets	Minimal effect	Unreliable training	Blurry, low detail	+7–14% accuracy improvement [16,30]	Few-shot datasets

5.2 Diffusion Models vs Variational Autoencoders (VAEs)

VAEs generate smooth and coherent features but often blur high-frequency details. Agricultural images contain micro-textures such as fungal spots, chlorosis boundaries, nematode morphology, and pest structural traits. These details are essential for diagnosis, making VAE reconstructions insufficient for classification-driven augmentation.

Study [10] compared VAE, GAN, and diffusion architectures for scientific imaging and showed diffusion outperformed VAE in perceptual alignment and scientific validity. VAE outputs lacked discriminative lesion edges in leaf data and produced lower CLIPScore alignment with target structures. In nematode recognition (study [16]), latent diffusion preserved species-level morphology more accurately than VAE-style encodings, improving Top-1 accuracy by up to 14.6% in few-shot settings. VAEs remain useful for feature compression and latent embedding analysis, but for high-resolution augmentation and class balancing, diffusion provided better fidelity across all reviewed experiments.

5.3 Diffusion vs Traditional Augmentation

Conventional image augmentation (crop, rotate, flip, color jitter) increases dataset volume without creating new phenotypes. It does not model unseen disease progression or generate samples of rare classes, which limits generalization and increases bias toward majority classes. Multiple studies reported accuracy plateaus when only geometric transformations were used.

Diffusion augmentation introduced new visual patterns that traditional augmentation cannot replicate. For mushrooms, diffusion-based augmentation raised recall by 13.5% across 110 species (study [32]). In plant disease tasks, diffusion improved classification by 1–9% across multiple benchmarks (study [12,26,31]). In weed detection, diffusion-generated samples increased mAP by 1.2–1.4% on YOLO architectures (study [13,24]). In vineyard shoot detection (study [17]), text-guided diffusion raised mAP by up to 28.6%, far beyond what rotation-based augmentation achieved. Traditional augmentation remains beneficial as a baseline pre-processing step, but the reviewed evidence suggests diffusion has become the preferred augmentation method when dataset imbalance or rare symptom stages exist.

Diffusion consistently outperforms alternative methods when rare symptoms, limited datasets, or subtle morphological traits are present. GANs remain faster and useful for temporal modeling, VAEs for representation learning, and traditional augmentation as preprocessing, but diffusion emerges as the most practical generator for agriculture under realistic field data constraints.

6. Challenges in Applying Diffusion Models to Agricultural Imaging

Diffusion models improve data diversity and enhance classification performance across crop, fruit, and leaf datasets. However, several factors restrict their practical adoption in agricultural pipelines.

High computational cost remains the largest barrier. Diffusion sampling requires many inference steps, and training demands long GPU hours. Most studies employed single-task datasets with controlled environments [2,9,11,12,18,24,26]. Only few works tested scalability to multi-crop or multi-season datasets. Latent diffusion reduces cost, but real-time deployment is still unrealistic for edge devices in farms. Lightweight variants exist but lack benchmarking against field-grade data.

Training instability increases when real datasets are small. Agricultural datasets often contain 200–800 images per class. This leads to noise-amplified artifacts during synthesis. Works on nematodes [10,16] and mushrooms [32] highlight gains in low-sample regimes, but controlled tuning was required. Most studies used curated datasets and laboratory settings, not raw field images with blur, shadows, moisture or occlusion. Results from weed studies [9,13,24,33] show strong improvements, yet failure cases were not reported.

Generalization beyond local farm conditions also remains limited. Models trained on one region struggle when soil color, sun angle, leaf age or pest species differ. Only [13,17,25,33,34] tested cross-domain transfer. Vineyard [17] and green wall studies [30,39] address domain adaptation, but results are early. No study reports cross-country validation. Farm-to-farm robustness remains unclear.

Public datasets for diffusion-based agriculture are scarce. Only a few open releases exist such as beetle hindwing library [29], nematode dataset [10,16], potato dataset [26], green wall sets [30,39], and weed sets [9,13,24,33]. Many studies rely on private datasets. Lack of dataset access slows replication and prevents fair comparison.

Evaluation protocols are inconsistent. Metrics vary widely across the 27 studies. Some use FID and IS. Others report mAP, accuracy, recall or IoU. Few report perceptual quality or expert agronomic scoring. No benchmark exists for assessing synthetic realism in plant disease progression or fruit surface texture. Absence of standard metrics limits comparison and prevents estimating real utility across tasks.

Diffusion models therefore show strong promise but face technical, practical, and infrastructural barriers. More work is needed on model efficiency, domain transfer, multi-farm validation, open benchmarks, and unified evaluation standards.

7. Future Directions

Future work in diffusion-based agricultural imaging needs to address scale, modality, and deployment. Most studies use RGB images. Multimodal diffusion integrating hyperspectral, multispectral, thermal, and LiDAR data would strengthen disease detection under field variability. No study among the reviewed works used hyperspectral diffusion, despite clear value for early stress detection.

Foundation models trained on large agricultural corpora could reduce dependence on small datasets. Existing work fine-tunes Stable Diffusion or latent models for single crops or diseases. A unified pretrained agricultural diffusion backbone would support domain transfer across fruit, leaf, weed, and UAV imaging tasks.

On-device diffusion would reduce latency for field robotics. Current models run on desktop GPUs. Efficient variants using distillation or latent compression are needed for mobile sprayers, drones, and edge computing systems for real-time decisions.

A public large-scale generative benchmark is missing. Dataset standardization with fixed splits would enable reproducibility and fair comparison across future research.

Real-time synthetic augmentation pipelines integrated into monitoring systems will provide continuous learning under changing seasons and climates. This direction connects diffusion models to operational agriculture rather than laboratory settings.

8. Conclusion

Diffusion models have become a key method for synthetic image generation in agriculture. They address limits common to field datasets such as class imbalance, rare disease stages, inconsistent lighting, and low sample diversity. This review analyzed 27 Scopus-indexed studies published between 2020 and 2025 that applied DDPM, latent diffusion, ControlNet-based diffusion, or related variants for agricultural imaging tasks. The largest share of research focused on plant leaf disease classification and green wall plant health monitoring, followed by weed detection, nematode recognition, fruit disease assessment, and UAV-based crop phenotyping. Most studies reported accuracy gains when diffusion-based augmentation was used for training, with improvements ranging from small incremental boosts to notable increases above 10 percent in several cases.

Diffusion models surpassed GAN-based augmentation in handling fine texture details and rare phenotypes, particularly when dataset size was limited. Latent diffusion models reduced computational load and supported more flexible conditioning. Text-guided and ControlNet-guided approaches further enabled task-specific generation, such as unseen disease synthesis or controlled weed morphology. However, challenges remain. Training still requires non-trivial compute. Scaling to multi-farm environments and new crops is not established. Common datasets and evaluation protocols have not yet emerged, making direct comparison across studies difficult.

Future work should explore multimodal diffusion combining RGB, hyperspectral, thermal, and LiDAR inputs. Foundation-scale models trained on large agricultural corpora could support general downstream adaptation. Lightweight diffusion for mobile or edge deployment will improve real-field usability. Real-time augmentation pipelines integrated into phenotyping, pest detection, and harvesting systems would move diffusion from research to operations. Building shared benchmarks and open synthetic datasets will accelerate progress and improve reproducibility across research groups.

Writing Assistance Disclosure: This manuscript was drafted by the authors. Grammarly and ChatGPT were used occasionally for grammar checking and language refinement only. No content generation or analytical assistance was sought from AI tools.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

References:

1. Fenu, G.; Malloci, F.M. Evaluating Impacts between Laboratory and Field-Collected Datasets for Plant Disease Classification. *Agronomy* 2022, *12*, doi:10.3390/AGRONOMY12102359.
2. Yang, G.; Liu, C.; Li, G.; Chen, H.; Chen, K.; Wang, Y.; Hu, X. DiffKNet-TL: Maize Phenology Monitoring with Confidence-Aware Constrained Diffusion Model Based on UAV Platform. *Comput Electron Agric* 2025, *239*, doi:10.1016/j.compag.2025.110936.
3. Ayoub, S.; Gulzar, Y.; Reegu, F.A.; Turaev, S. Generating Image Captions Using Bahdanau Attention Mechanism and Transfer Learning. *Symmetry (Basel)* 2022, *14*, 2681.
4. Kumar, T.; Brennan, R.; Mileo, A.; Bendechache, M. Image Data Augmentation Approaches: A Comprehensive Survey and Future Directions. *IEEE Access* 2024, *12*, 187536–187571, doi:10.1109/ACCESS.2024.3470122.
5. Wang, Z.; Wang, P.; Liu, K.; Wang, P.; Fu, Y.; Lu, C.-T.; Aggarwal, C.C.; Pei, J.; Zhou, Y. A Comprehensive Survey on Data Augmentation. *IEEE Trans Knowl Data Eng* 2025, 1–20, doi:10.1109/TKDE.2025.3622600.
6. Kang, J.; Yoon, J.; Park, B.; Kim, J.; Jee, S.; Song, H.; Chung, H. RipenessGAN: Growth Day Embedding-Enhanced GAN for Stage-Wise Jujube Ripeness Data Generation. *Agronomy* 2025, *15*, doi:10.3390/agronomy15102409.
7. Espinoza, S.; Aguilera, C.; Rojas, L.; Campos, P.G. Analysis of Fruit Images With Deep Learning: A Systematic Literature Review and Future Directions. *IEEE Access* 2024, *12*, 3837–3859, doi:10.1109/ACCESS.2023.3345789.
8. Deng, B.; Lu, Y. Weed Image Augmentation by ControlNet-Added Stable Diffusion. *https://doi.org/10.1117/12.3014145* 2024, *13035*, 188–200, doi:10.1117/12.3014145.
9. Ambuj; Machavaram, R. Weed Image Classification with Generative AI Using Latent Denoising Diffusion Probabilistic Model and Wiener Filtering Approach. *Franklin Open* 2025, *12*, doi:10.1016/j.fraope.2025.100369.
10. Sordo, Z.; Chagnon, E.; Hu, Z.; Donatelli, J.J.; Andeer, P.F.; Nico, P.S.; Northen, T.R.; Ushizima, D.M.U. Synthetic Scientific Image Generation with VAE, GAN, and Diffusion Model Architectures. *J Imaging* 2025, *11*, doi:10.3390/jimaging11080252.
11. Yu, Y.; Wu, X.; Yu, P.; Wan, Q.; Dan, Y.; Xiao, Y.; Wang, Q. Location-Guided Lesions Representation Learning via Image

Generation for Assessing Plant Leaf Diseases Severity. *Plant Phenomics* 2025, *7*, doi:10.1016/j.plaphe.2025.100058.

12. Das, A.; Mahto, R.; Wang, W.; Jamali, A.; Singh, P.; Roy, S.K. A Generative Framework for Detection and Classification of Plant Leaf Disease Using Diffusion Network. *Appl Soft Comput* 2025, *177*, doi:10.1016/j.asoc.2025.113152.
13. Deng, B.; Lu, Y. Weed Image Augmentation by ControlNet-Added Stable Diffusion for Multi-Class Weed Detection. *Comput Electron Agric* 2025, *232*, doi:10.1016/j.compag.2025.110123.
14. Huang, J.; Wang, Z.; Xu, M.; Ma, L.; Wu, W.; Cao, J. Enhancing Few-Shot Plant Disease Classification with Diffusion Model. *Proceedings - 2024 China Automation Congress, CAC 2024* 2024, 2036–2040, doi:10.1109/CAC63892.2024.10864650.
15. Chen, D.; Qi, X.; Zheng, Y.; Lu, Y.; Huang, Y.; Li, Z. Synthetic Data Augmentation by Diffusion Probabilistic Models to Enhance Weed Recognition. *Comput Electron Agric* 2024, *216*, 108517, doi:10.1016/J.COMPAG.2023.108517.
16. Ouyang, X.; Zhuang, J.; Gu, J.; Ye, S. Few-Shot Data Augmentation by Morphology-Constrained Latent Diffusion for Enhanced Nematode Recognition. *Computers* 2025, *14*, doi:10.3390/computers14050198.
17. Hirahara, K.; Nakane, C.; Ebisawa, H.; Kuroda, T.; Iwaki, Y.; Utsumi, T.; Nomura, Y.; Koike, M.; Mineno, H. D4: Text-Guided Diffusion Model-Based Domain Adaptive Data Augmentation for Vineyard Shoot Detection. *Comput Electron Agric* 2025, *230*, doi:10.1016/j.compag.2024.109849.
18. Mori, N.; Naito, H.; Hosoi, F. Application of a Latent Diffusion Model to Plant Disease Detection by Generating Unseen Class Images. *AgriEngineering* 2024, *6*, 4901–4910, doi:10.3390/agriengineering6040279.
19. Ho, J.; Jain, A.; Abbeel, P. Denoising Diffusion Probabilistic Models. *Adv Neural Inf Process Syst* 2020, *33*, 6840–6851.
20. Song, J.; Meng, C.; Ermon, S. Denoising Diffusion Implicit Models. *ICLR 2021 - 9th International Conference on Learning Representations* 2020.
21. Li, J.; Zhu, C.; Yang, C.; Zheng, Q.; Wang, B.; Tu, J.; Zhang, Q.; Liu, S.; Wang, X.; Qiao, J. DM_CorrMatch: A Semi-Supervised Semantic Segmentation Framework for Rapeseed Flower Coverage Estimation Using UAV Imagery. *Plant Methods* 2025, *21*, doi:10.1186/s13007-025-01373-w.
22. Zhang, S.; Liu, L.; Li, G.; Du, Y.; Wu, X.; Song, Z.; Li, X. Diffusion Model-Based Image Generative Method for Quality Monitoring of Direct Grain Harvesting. *Comput Electron Agric* 2025, *233*, doi:10.1016/j.compag.2025.110130.
23. Rombach, R.; Blattmann, A.; Lorenz, D.; Esser, P.; Ommer, B. High-Resolution Image Synthesis With Latent Diffusion Models 2022, 10684–10695.
24. Deng, B.; Lu, Y. Image Prompt Adapter-Based Stable Diffusion for Enhanced Multi-Class Weed Generation and Detection. *AgriEngineering* 2025, *7*, doi:10.3390/agriengineering7110389.
25. Zhang, L.; Zhao, Y.; Zhou, C.; Zhang, J.; Yan, Y.; Chen, T.; Lv, C. JuDifformer: Multimodal Fusion Model with Transformer and Diffusion for Jujube Disease Detection. *Comput Electron Agric* 2025, *232*, doi:10.1016/j.compag.2025.110008.
26. Astuti, T.; Umar, A.N.; Wahyudi, R.; Rifai, Z. Enhancing Potato Leaf Disease Detection: Implementation of Convolutional Vision Transformers with Synthetic Datasets from Stable Diffusion. *International Journal on Informatics Visualization* 2024, *8*, 2054–2065, doi:10.62527/joiv.8.4.2167.
27. Ho, J.; Salimans, T. Classifier-Free Diffusion Guidance. 2022.
28. Zhang, L.; Rao, A.; Agrawala, M. Adding Conditional Control to Text-to-Image Diffusion Models 2023, 3836–3847.
29. Yang, Y.; Li, W.; Liu, R.; Wu, C.; Ren, J.; Shi, Y.; Ge, S. HindwingLib: A Library of Leaf Beetle Hindwings Generated by Stable Diffusion and ControlNet. *Sci Data* 2025, *12*, doi:10.1038/s41597-025-05010-y.
30. Lee, Y. Enhancing Plant Health Classification via Diffusion Model-Based Data Augmentation. *Multimed Syst* 2025, *31*, doi:10.1007/s00530-025-01745-1.
31. Wang, R.; Zhang, X.; Yang, Q.; Lei, L.; Liang, J.; Yang, L. Enhancing Panax Notoginseng Leaf Disease Classification with Inception-SSNet and Image Generation via Improved Diffusion Model. *Agronomy* 2024, *14*, doi:10.3390/agronomy14091982.
32. Du, M.; Wang, F.; Yan, W.; Guo, J.; Liu, L.; Lv, P.; He, Y.; Feng, X.; Wang, Y. Improving Food Safety: Synthetic Data Augmentation for Accurate Mushroom Species Identification in Complex Environments. *Applied Food Research* 2025, *5*, doi:10.1016/j.afres.2025.101039.
33. Modak, S.; Stein, A. Generative AI-Based Pipeline Architecture for Increasing Training Efficiency in Intelligent Weed Control Systems. *Journal of Systems Architecture* 2025, *167*, doi:10.1016/j.sysarc.2025.103464.
34. Cheng, Z.; Gu, X.; Zhang, Z.; Xu, Y.; Zhao, T.; Li, Y.; Sun, S.; Du, Y.; Cai, H. TAM-Net: A Deep Network Combining Tabular Diffusion Algorithm, Attention Mechanism, and Multi-Task Learning for Monitoring Crop Water Status from UAV Multi-Source Images. *European Journal of Agronomy* 2025, *170*, doi:10.1016/j.eja.2025.127778.
35. Wang, S.; Xu, S.; Lv, C.; Cai, F. RSSR-CDM: Remote Sensing Image Super-Resolution Based on an Improved Conditional Diffusion Model. *J Electron Imaging* 2025, *34*, doi:10.1117/1.JEI.34.4.043046.
36. Jasphin, E.T.J.; Joice, C.S. Weed Classification and Crop Health Monitoring in Microclimatic Conditions Using Thermal Image Analysis and Deep Learning Algorithms. *J Plant Growth Regul* 2025, *44*, 2247–2263, doi:10.1007/s00344-024-11542-1.
37. Wang, L.; Zhang, Z.; Yao, H. Agricrafter: Full-Growth-Cycle Crop Video Generation via Diffusion Models. *Comput Electron Agric* 2025, *239*, doi:10.1016/j.compag.2025.110946.
38. Li, W.; Yang, L.; Peng, G.; Pang, G.; Yu, Z.; Zhu, X. An Effective Microscopic Image Augmentation Approach. *Sci Rep* 2025, *15*, doi:10.1038/s41598-025-93954-x.
39. Yoon, M.; Lee, Y. Novel Augmentation Techniques Using Diffusion Models for Green Wall Plant Health Classification. *Comput Biol Med* 2025, *189*, doi:10.1016/j.compbiomed.2025.109899.