electronics

MDPI

Article

PrivAgriVolt: Privacy-Preserving Shadow-Aware Vision for Crop Stress Diagnosis in Agrivoltaic Photovoltaic Systems

Zuoming Yin [1,2], Yifei Zhang [1,2], Qiangqiang Lei [1,2] and Fang Feng [1,2,*]

1 Heilongjiang Longyuan New Energy Development Co., Ltd., Harbin 150096, China; 12109401@chnenergy.com.cn (Z.Y.); 12031017@chnenergy.com.cn (Y.Z.); 12056832@chnenergy.com.cn (Q.L.)
2 College of Arts and Sciences, Northeast Agricultural University, Harbin 150006, China
* Correspondence: fengfang@neau.edu.cn

Abstract

Agrivoltaic systems co-locate photovoltaic (PV) arrays and crops, offering land-use efficiency and potential microclimate benefits, yet they introduce new challenges for computer-vision-based crop monitoring. PV structures produce strong, spatially varying shadows, specular reflections, and periodic occlusions that confound visual cues for diagnosing crop diseases and abiotic stresses. Meanwhile, agrivoltaic deployments are often distributed across farms and operators, making centralized data collection impractical due to privacy, ownership, and regulatory concerns. This paper proposes PrivAgriVolt, a novel privacy-preserving learning framework for agrivoltaic crop issue recognition that explicitly models PV-induced illumination and enables collaborative training without sharing raw images. The core algorithm integrates (i) a PV-geometry-conditioned shadow normalization module that fuses estimated array layout and sun-angle priors into a shadow-aware appearance canonization network, reducing illumination-induced domain shift across times and sites; (ii) a federated contrastive stress learner that aligns stress semantics across farms via prototype-based contrastive objectives while remaining robust to heterogeneous sensors and crop stages; and (iii) an adaptive privacy layer that combines secure aggregation with budget-aware gradient perturbation and client-level clipping to provide formal privacy guarantees while preserving fine-grained diagnostic performance. Extensive experiments on real agricultural vision benchmarks and agrivoltaic shadow variants demonstrate that PrivAgriVolt improves stress recognition and segmentation under PV shading while maintaining strong privacy–utility trade-offs.

Keywords: photovoltaic agriculture; crop stress diagnosis; shadow-aware learning; domain generalization; federated learning

Check for updates

Academic Editor: Pablo García Triviño

Received: 21 March 2026
Revised: 10 April 2026
Accepted: 16 April 2026
Published: 21 April 2026

1. Introduction

Agrivoltaic (AV) or agro-photovoltaic systems co-locate photovoltaic (PV) arrays with crops to increase land-use efficiency while potentially improving water use and microclimate management. Recent systematic reviews emphasize that crop outcomes under PV shading are highly context-dependent and closely tied to array geometry and local climate [1–3]. As deployments scale, operators increasingly require continuous, fine-grained monitoring of crop conditions to detect diseases and abiotic stresses early, quantify spatial variability, and adjust management strategies. Computer vision offers a low-cost and scalable solution because images can be collected by handheld devices, fixed cameras, and UAVs. However, the AV setting differs fundamentally from conventional open-field monitoring: PV structures introduce strong and structured shadows, periodic occlusions, and specular reflections whose patterns change with array layout and solar position. These effects distort color and texture cues and create large appearance shifts across time and sites, often exceeding the domain gap observed in standard agricultural benchmarks. Recent shadow surveys and challenge reports underscore that complex cast-shadows remain difficult to handle robustly in the wild [4–6].

A second barrier is privacy. AV deployments are naturally distributed: multiple farms, operators, and integrators collect imagery under different business constraints. Centralizing images is frequently infeasible due to concerns over farm ownership, sensitive management practices, and regulatory requirements. Federated learning (FL) mitigates this by training models without moving raw data [7], but practical FL pipelines remain vulnerable to leakage through updates unless combined with secure aggregation and differential privacy (DP) [8–10]. Moreover, AV heterogeneity (crop types, phenology, camera hardware, and PV geometry) amplifies non-IID effects that degrade naïve FL optimization, especially when DP noise is applied.

This paper addresses these challenges and asks: Can we build a collaborative vision system that (i) remains accurate under PV-induced shadow dynamics, (ii) generalizes across heterogeneous AV sites, and (iii) provides principled privacy protection without sharing images? We propose PrivAgriVolt, a privacy-preserving, shadow-aware learning framework for AV crop stress diagnosis. In our formulation (Section 3), each client k holds a private dataset $\mathcal{D}_k$ of images x with labels $\mathbf{y}$ for classification or segmentation, along with a compact PV geometry descriptor $\mathbf{g}^{(k)}$ and sun-direction prior $\mathbf{s}(t)$ from capture time t. These signals induce a soft shadow mask $\mathbf{m}$ that conditions a learned canonization module before prediction by $f_{\boldsymbol{\theta}}$.

PrivAgriVolt introduces three algorithmic ideas (Section 4). First, we design a PV-geometry-conditioned shadow normalization module that uses $(\mathbf{m}, \mathbf{g}^{(k)}, \mathbf{s}(t))$ to produce a canonized image $\tilde{\mathbf{x}}$ while preserving stress evidence. Unlike generic shadow removal pipelines, the normalization is explicitly conditioned on PV layout and solar geometry, targeting structured AV shadows and reducing time-of-day shift. Second, we introduce a federated prototype contrastive stress learner that aligns representations across sites using securely aggregated class prototypes, improving cross-client generalization under non-IID distributions. Third, we incorporate an adaptive privacy layer that combines secure aggregation with client-level DP on updates, scheduling noise across rounds to improve utility at fixed privacy budgets. Our experiments use real, publicly available agricultural datasets—PlantVillage [11], PlantDoc [12], and Agriculture-Vision [13]—and construct PV-shadowed AV variants to systematically evaluate robustness to structured shading. Results show that PrivAgriVolt improves leave-one-client-out generalization under AV shadows and maintains strong privacy–utility trade-offs compared to standard FL baselines.

In summary, this work makes the following contributions:

- We formulate privacy-preserving AV crop stress recognition and segmentation under PV-induced structured shadows, integrating PV geometry and solar priors into the learning pipeline.
- We propose a novel, PV-geometry conditioned shadow canonization module with a stress-preserving shadow-consistency regularizer that reduces AV illumination shift without erasing diagnostic cues.
- We introduce a federated prototype-contrastive alignment mechanism that improves cross-site generalization, and we integrate it with a client-level DP layer and secure aggregation for end-to-end privacy protection.

- We provide extensive evaluations on real agricultural datasets and AV-shadowed variants, including ablations and privacy–utility analyses under concrete (ε, δ) budgets.

The paper is organized as follows: Section 2 reviews related work; Section 3 presents preliminaries and notation; Section 4 details the proposed method; Section 5 reports experiments; Section 6 concludes.

2. Related Work

2.1. Agrivoltaics and Monitoring in PV–Agriculture Co-Location

Agrivoltaics has attracted sustained interest as a dual-use approach to address land competition between PV deployment and agriculture. Recent systematic reviews synthesize evidence that agronomic outcomes and yield stability depend strongly on crop species, local climate, and PV configuration choices [1,2,14]. Modeling, simulation, and optimization have become central for design-time prediction of shading, microclimate, PV yield, and agronomic performance, with comprehensive reviews highlighting the need for integrated platforms and standardized validation across configurations [3,15]. Practice-oriented reports further summarize performance indicators, monitoring needs, and operational considerations for real deployments [16,17]. These directions motivate field-ready sensing and learning systems that are robust to PV-induced illumination structure.

2.2. Agricultural Computer Vision and Domain Shift Under Challenging Illumination

Agricultural computer vision has advanced rapidly with public datasets and deep architectures. PlantVillage enabled early progress in leaf disease recognition under controlled settings [11], while PlantDoc introduced in-the-wild imagery with complex backgrounds and illumination [12]. At field scale, Agriculture-Vision provides aerial images with pixel-level annotations for agricultural pattern segmentation [13]. Recent surveys and studies document the shift from CNN-based pipelines to transformer-era models and stress the persistence of domain shift caused by sensing, phenology, and illumination changes [18–20]. For segmentation in agricultural remote sensing, recent reviews summarize architectural trends and remaining gaps in cross-scene generalization [21]. Challenge reports and solutions for Agriculture-Vision further emphasize class imbalance and robustness issues in real-world aerial settings [22]. These results collectively indicate that illumination-structured domain shifts (such as those induced by PV arrays) require task-aware normalization and cross-site generalization mechanisms.

2.3. Shadow-Aware Vision and Privacy-Preserving Federated Learning

Shadows remain a classical obstacle in vision, and recent work consolidates deep-learning progress via surveys and standardized comparisons, spanning shadow detection, removal, and generation [4,23]. The NTIRE shadow removal challenge reports provide rigorous evaluations and show that complex cast shadows remain difficult under both fidelity and perceptual tracks [5,6]. However, generic shadow removal can inadvertently suppress subtle symptoms (e.g., texture and chromatic cues), motivating task-constrained shadow normalization in agricultural diagnosis.

For privacy, federated learning enables collaborative training without sharing raw data [7]. Yet model updates can leak information, and recent surveys synthesize attack surfaces, defenses, and policy implications, emphasizing that privacy-by-design is insufficient without explicit protections [10,24]. Secure aggregation provides cryptographic protection for individual updates [8], while DP provides formal leakage bounds through randomized perturbation [9,25]. Recent protocol proposals and efficient secure aggregation designs further improve practicality in resource-constrained settings [26–28]. Recent DPFL reviews and methods further study privacy budget management and utility preservation under heterogeneity [29–32]. Finally, generalization under heterogeneous clients connects to federated domain generalization (FedDG), which has matured through benchmarks and new algorithmic designs [33–35]. Recent FedDG methods explore prompt-based and generative strategies [36,37] and new regularization viewpoints [38,39]. PrivAgriVolt builds on these trends by introducing PV-geometry-conditioned shadow canonization and prototype-based contrastive alignment under an explicit client-level DP layer, targeting the distinctive illumination and privacy constraints of agrivoltaic monitoring.

In agricultural and natural environments, classical preprocessing pipelines based on color constancy, Retinex correction, histogram equalization, handcrafted texture/color descriptors, and rule-based segmentation remain valuable because they are interpretable and inexpensive [40–43]. However, agrivoltaic PV shadows are structured, geometry-dependent, and strongly spatially varying, so purely fixed enhancement often either under-corrects deep cast shadows or suppresses subtle lesion cues together with the illumination artifact. This is why we adopt a lightweight deep normalization module but constrain it with explicit PV geometry, sun priors, and mask conditioning so that the correction remains physically guided rather than purely black-box.

3. Preliminaries

We formalize the agrivoltaic vision task, introduce PV-induced shadow notation, and summarize the privacy-preserving federated learning primitives used throughout the paper. All symbols defined here are used consistently in later sections.

3.1. Problem Setup and Notation

An agrivoltaic site consists of crop rows co-located with photovoltaic (PV) arrays. We consider K data owners (farms or operators) participating as clients in collaborative training, indexed by $k \in \{1, \ldots, K\}$. Client k holds a private dataset

$$\mathcal{D}_k = \{(\mathbf{x}_i^{(k)}, \mathbf{y}_i^{(k)})\}_{i=1}^{N_k}, \tag{1}$$

where $\mathbf{x} \in \mathbb{R}^{H \times W \times 3}$ is an RGB image captured under PV structures (e.g., UAV or ground camera), and $\mathbf{y}$ denotes supervision for one or more tasks. We focus on two standard agricultural vision tasks:

- Stress recognition (classification): $\mathbf{y} \in \{1, \ldots, C\}$ indicates one of C crop conditions (healthy or stress categories).
- Stress localization (segmentation): $\mathbf{y} \in \{0, 1, \ldots, C\}^{H \times W}$ is a dense label map, where zero denotes background/unknown.

To avoid repeatedly branching notation, we write a unified predictor

$$f_{\boldsymbol{\theta}} : \mathbf{x} \mapsto \hat{\mathbf{y}}, \tag{2}$$

with parameters $\boldsymbol{\theta}$. For classification, $f_{\boldsymbol{\theta}}(\mathbf{x})$ outputs a probability vector $\hat{\mathbf{p}} \in \Delta^C$; for segmentation it outputs $\hat{\mathbf{P}} \in \Delta^{C+1}_{H \times W}$. We also use an encoder $e_{\boldsymbol{\theta}}$ to map an image to a representation $\mathbf{z} \in \mathbb{R}^d$:

$$\mathbf{z} = e_{\boldsymbol{\theta}}(\mathbf{x}), \qquad d \ll HW. \tag{3}$$

In the experiments, the encoder is instantiated as ResNet-50, so the default embedding dimension used for prototype alignment is $d = 2048$ after global pooling. Let $\ell(\cdot, \cdot)$ denote a generic supervised loss (cross-entropy for classification or pixel-wise cross-entropy for segmentation). The standard empirical risk on client k is

$$\mathcal{L}_k(\boldsymbol{\theta}) = \frac{1}{N_k} \sum_{(\mathbf{x},\mathbf{y}) \in \mathcal{D}_k} \ell(f_{\boldsymbol{\theta}}(\mathbf{x}), \mathbf{y}). \tag{4}$$

In general, agrivoltaic learning faces two characteristic sources of shift: (i) illumination and shadow variation caused by PV structures, changing with time and geometry; (ii) site heterogeneity across clients (crop varieties, growth stages, sensors). We therefore explicitly define PV-shadow variables next.

3.2. PV-Geometry and Shadow Formation

We describe a compact representation of PV-induced shadows used to condition shadow-aware normalization later. For each site (client) k, we denote an approximate PV geometry descriptor by

$$\mathbf{g}^{(k)} = \big(\mathbf{n}^{(k)},\, h^{(k)},\, \rho^{(k)},\, \alpha^{(k)}\big), \tag{5}$$

where $\mathbf{n}^{(k)} \in \mathbb{R}^3$ is the unit normal of the PV panel plane, $h^{(k)}$ is the panel clearance height, $\rho^{(k)}$ is the inter-row spacing (effective pitch), and $\alpha^{(k)}$ is the panel tilt/azimuth parameterization (any consistent 1–2D encoding is admissible). For an image captured at time t, we denote the sun-direction unit vector (in the site coordinate frame) by $\mathbf{s}(t) \in \mathbb{R}^3$. PV structures cast shadows that can be abstracted as a binary or soft shadow mask $\mathbf{m} \in [0,1]^{H \times W}$ aligned with image coordinates:

$$\mathbf{m} = \mathcal{S}\Big(\mathbf{g}^{(k)}, \mathbf{s}(t), \Pi\Big), \tag{6}$$

where $\mathcal{S}(\cdot)$ is a shadow projection operator and Π denotes camera projection parameters (intrinsics/extrinsics). In practice, $\mathcal{S}$ may be approximated by (i) coarse ray-casting from simplified PV planes, (ii) learned shadow estimation, or (iii) a hybrid approach that uses geometry as a prior. We do not assume perfect geometry; instead, $\mathbf{m}$ is treated as an informative conditioning signal.

To capture illumination differences between shaded and non-shaded regions, we define a simple image decomposition:

$$\mathbf{x} = \mathbf{x}_{\text{lit}} \odot (1 - \mathbf{m}) \;+\; \mathbf{x}_{\text{sh}} \odot \mathbf{m} \;+\; \boldsymbol{\eta}, \tag{7}$$

where $\odot$ is element-wise multiplication broadcast over channels, $\mathbf{x}_{\text{lit}}$ and $\mathbf{x}_{\text{sh}}$ are latent appearances under lit/shaded conditions, and $\boldsymbol{\eta}$ aggregates noise (sensor noise, specularities, occlusions). Equation (7) motivates learning a mapping that reduces shadow-induced appearance shift while preserving stress cues (texture, lesion patterns, chlorosis).

3.3. Federated Optimization and Privacy Primitives

We use synchronous federated learning (FL) to train a shared model without centralizing raw images. Let $\boldsymbol{\theta}^{(r)}$ be global parameters at communication round r. A subset of clients $\mathcal{S}_r \subseteq \{1, \ldots, K\}$ participates in round r. Each selected client performs E steps of local optimization (e.g., SGD) on its private data, producing an update $\Delta\boldsymbol{\theta}_k^{(r)}$. The server aggregates updates to form the next global model:

$$\boldsymbol{\theta}^{(r+1)} = \boldsymbol{\theta}^{(r)} + \sum_{k \in \mathcal{S}_r} w_k \, \Delta\boldsymbol{\theta}_k^{(r)}, \qquad w_k = \frac{N_k}{\sum_{j \in \mathcal{S}_r} N_j}. \tag{8}$$

To prevent the server from inspecting individual client updates, FL may employ a secure aggregation protocol that reveals only the (weighted) sum $\sum_{k \in \mathcal{S}_r} w_k \, \Delta\boldsymbol{\theta}_k^{(r)}$ while keeping each $\Delta\boldsymbol{\theta}_k^{(r)}$ hidden.

https://doi.org/10.3390/electronics15081762

In this paper, we use the standard definition of (ε, δ)-differential privacy (DP). A randomized mechanism $\mathcal{M}$ is (ε, δ)-DP if for any neighboring datasets $\mathcal{D}$ and $\mathcal{D}'$ that differ by one sample, and any measurable set $\mathcal{O}$,

$$\Pr[\mathcal{M}(\mathcal{D}) \in \mathcal{O}] \le e^{\varepsilon} \Pr[\mathcal{M}(\mathcal{D}') \in \mathcal{O}] + \delta. \tag{9}$$

In gradient-based training, a common DP mechanism clips per-client updates and adds Gaussian noise:

$$\widetilde{\Delta\boldsymbol{\theta}}_k^{(r)} = \mathrm{clip}\Big(\Delta\boldsymbol{\theta}_k^{(r)}, C\Big) + \mathcal{N}\Big(\mathbf{0}, \sigma^2 C^2 \mathbf{I}\Big), \tag{10}$$

where $C > 0$ is a clipping norm, σ is the noise multiplier, and $\mathrm{clip}(\mathbf{u}, C) = \mathbf{u} \cdot \min(1, C/\|\mathbf{u}\|_2)$. Privacy accounting across rounds yields a total budget (ε, δ) for the training process; we report these budgets in Section 5 using standard composition accounting.

The combination of secure aggregation (to hide individual updates) and DP noise (to bound information leakage from the aggregated signal) provides a practical privacy baseline for collaborative agrivoltaic vision learning. The next section builds on these preliminaries to introduce the proposed shadow-aware, privacy-preserving algorithm.

4. Methodology

4.1. End-to-End Pipeline and Notation

PrivAgriVolt addresses agrivoltaic learning in a cross-silo federated setting, where K clients (farms, sites, or geographic tiles) collaboratively train a shared model without exchanging raw images. Client k holds a private dataset $\mathcal{D}_k = \{(\tilde{\mathbf{x}}, \mathbf{y})\}$, where $\tilde{\mathbf{x}}$ denotes the observed image under agrivoltaic illumination (potentially PV-shadowed; in our synthetic variants, it is the renderer output), and $\mathbf{y}$ is the corresponding class label (classification) or pixel label map (segmentation). Each client also has site metadata describing the PV installation (e.g., row pitch/orientation/height) summarized as a geometry descriptor $\mathbf{g}^{(k)}$, and a capture timestamp t from which the local sun direction $\mathbf{s}(t)$ can be derived. These metadata are lightweight and do not reveal raw imagery; we use them only to condition shadow normalization.

PrivAgriVolt factorizes the predictor into two trainable parts: (i) a shadow canonization module $\mathcal{N}_{\boldsymbol{\psi}}$ that reduces PV-induced illumination shift, and (ii) a task predictor $f_{\boldsymbol{\theta}}$ for classification or segmentation. The task predictor contains an encoder $e_{\boldsymbol{\theta}}$ whose embeddings are used both for supervised prediction and for cross-client prototype alignment. In addition to model parameters $(\boldsymbol{\theta}, \boldsymbol{\psi})$, the server maintains a compact global state consisting of class prototypes $\mathcal{P} = \{\mathbf{p}_c\}_{c=1}^{C}$.

Training proceeds in synchronized rounds. In each round, clients locally (a) estimate PV-shadow structure and canonize inputs, (b) optimize a supervised loss augmented by shadow-invariance and prototype alignment, and (c) transmit only encrypted, privacy-protected updates. The server aggregates updates using secure aggregation and updates prototypes from aggregated sufficient statistics. The remainder of this section details (1) PV-conditioned canonization, (2) prototype-based alignment, and (3) privacy-preserving optimization that couples the two.The overall end-to-end pipeline of the proposed PrivAgriVolt framework is illustrated in Figure 1.

4.2. PV-Conditioned Shadow Canonization with Consistency Regularization

PV-row shadows introduce structured, quasi-periodic attenuation that varies with array layout and solar position, often causing two images of the same stress to appear more different than two images of different stresses captured under similar shading. This illumination-driven covariate shift is especially harmful in federated learning because each client can observe a distinct shadow distribution, amplifying representation drift and destabilizing aggregation. We therefore normalize illumination *before* learning stress semantics.

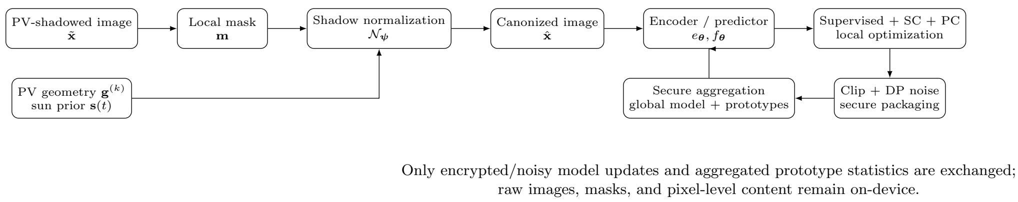

Figure 1. End-to-end pipeline of PrivAgriVolt. Each client starts from a local PV-shadowed image $\tilde{\mathbf{x}}$, site geometry $\mathbf{g}^{(k)}$, timestamp-derived sun direction $\mathbf{s}(t)$, and a locally estimated shadow mask $\mathbf{m}$. The shadow normalization module $\mathcal{N}_{\psi}$ produces a canonized image $\hat{\mathbf{x}}$, the encoder/predictor extracts task features and predictions, local supervised plus shadow-consistency and prototype losses are optimized on-device, and only clipped/encrypted model updates and aggregated prototype statistics are sent to the server for secure aggregation and global prototype updates. No raw image, mask, or pixel content leaves the client.

4.2.1. Geometry- and Time-Conditioned Canonization

Given an observed PV-shadowed input $\tilde{\mathbf{x}}$, a shadow mask $\mathbf{m}$, and a conditioning vector $\mathbf{u}$ derived from PV geometry and sun direction, we compute a canonized image:

$$\hat{\mathbf{x}} = \mathcal{N}_{\psi}(\tilde{\mathbf{x}}, \mathbf{m}, \mathbf{u}), \qquad \mathbf{u} = \mathcal{E}\Big(\mathbf{g}^{(k)}, \mathbf{s}(t)\Big) \in \mathbb{R}^{q}, \tag{11}$$

where $\mathcal{E}$ is a lightweight encoder (e.g., an MLP) that embeds PV geometry $\mathbf{g}^{(k)}$ and sun direction $\mathbf{s}(t)$ into $\mathbf{u}$. The mask $\mathbf{m} \in [0,1]^{H\times W}$ is computed locally (e.g., via a geometry-based estimator, a local shadow detector, or their combination); $\mathbf{m}$ never leaves the client. Intuitively, $\mathbf{m}$ specifies where attenuation occurs, while $\mathbf{u}$ specifies how attenuation should be corrected under site- and time-dependent PV configurations.

For all synthetic PV-shadow experiments in Section 5, we use the renderer-generated soft mask as the conditioning signal $\mathbf{m}$, so the masking strategy is exact and fully reproducible. For real deployment, we use a hybrid local estimator: a coarse geometry prior is first obtained by ray-casting the PV rows described by $\mathbf{g}^{(k)}$ under the sun vector $\mathbf{s}(t)$ and camera model Π; this prior is then refined by a lightweight local shadow detector based on RGB intensity/chromaticity cues, and the two estimates are fused into a soft mask by confidence-weighted averaging plus small morphological smoothing. This entire procedure is executed locally on the client, so neither the mask nor any intermediate image leaves the site. In practice, the method is more tolerant to moderate boundary misalignment than to large false negatives inside dark PV bands, because correction is soft and the shadow-consistency loss averages over regions. We therefore prefer soft masks, temporal smoothing when consecutive frames are available, and detector-only fallback when geometry metadata are incomplete.

4.2.2. Conditioned Illumination Field and Stabilized Retinex Correction

Motivated by the decomposition in Equation (7), we predict an illumination map $\mathbf{L} \in \mathbb{R}^{H\times W\times 3}$ and correct the input using a stabilized Retinex-like division:

$$\mathbf{L} = \mathrm{Softplus}\big(\mathcal{U}_{\psi}(\tilde{\mathbf{x}}, \mathbf{m}; \mathbf{u})\big), \qquad \hat{\mathbf{x}} = \mathrm{clip}\bigg(\frac{\tilde{\mathbf{x}}}{\mathbf{L}+\epsilon}, 0, 1\bigg), \tag{12}$$

with small $\epsilon > 0$. The $\mathrm{Softplus}(\cdot)$ nonlinearity ensures $\mathbf{L}$ is positive, and clipping keeps outputs within the valid range. In practice, $\mathcal{U}_{\psi}$ is implemented as a lightweight U-Net that ingests the RGB image and mask (e.g., concatenated channels). The mask guides the network to focus correction on shadowed regions while keeping confidently lit regions near identity. When $\mathbf{m}$ is close to zero everywhere (weak or no PV shadow), the module naturally behaves like an identity mapping, avoiding unnecessary color distortions.

The modulation by $\mathbf{u}$ is implemented via feature-wise affine conditioning (FiLM) inside $\mathcal{U}_{\boldsymbol{\psi}}$:

$$\mathrm{FiLM}(\mathbf{h};\mathbf{u}) = \boldsymbol{\gamma}(\mathbf{u}) \odot \mathbf{h} + \boldsymbol{\beta}(\mathbf{u}), \tag{13}$$

where $\mathbf{h}$ is an intermediate feature tensor and $\boldsymbol{\gamma}(\cdot), \boldsymbol{\beta}(\cdot)$ are learned projections. This conditioning injects PV layout and solar priors into the correction process, improving stability across times of day and across sites with different PV geometries.

4.2.3. Shadow-Consistency Regularization

Illumination correction can over-normalize and wash out subtle stress cues (e.g., lesion boundaries, vein discoloration, fine texture). To preserve semantics while enforcing invariance to shading, we impose a representation-level constraint between lit and shaded regions after canonization. Let $\mathbf{F} = e_{\boldsymbol{\theta}}(\hat{\mathbf{x}}) \in \mathbb{R}^{h \times w \times d}$ be the last convolutional feature map of the encoder. We compute region-aggregated embeddings

$$\mathbf{z}_{\mathrm{sh}} = \frac{\sum_p \mathbf{m}_p \, \mathbf{F}_p}{\sum_p \mathbf{m}_p + \epsilon}, \qquad \mathbf{z}_{\mathrm{lit}} = \frac{\sum_p (1 - \mathbf{m}_p) \, \mathbf{F}_p}{\sum_p (1 - \mathbf{m}_p) + \epsilon}, \tag{14}$$

where p indexes spatial positions in the $h \times w$ feature grid (with $\mathbf{m}$ downsampled accordingly). The shadow-consistency loss is

$$\mathcal{L}_{\mathrm{sc}} = 1 - \frac{\mathbf{z}_{\mathrm{sh}}^{\top} \mathbf{z}_{\mathrm{lit}}}{\|\mathbf{z}_{\mathrm{sh}}\|_2 \|\mathbf{z}_{\mathrm{lit}}\|_2 + \epsilon}. \tag{15}$$

This reduces incentives for client-specific shortcuts tied to shading patterns and improves stability when shadow coverage differs across clients. After canonization, the remaining challenge is non-illumination heterogeneity; we address it by aligning representations with global class prototypes.

4.3. Federated Prototype Contrastive Alignment for Stress Semantics

Even after canonization, substantial heterogeneity remains across clients due to cultivar differences, camera pipelines, phenological stages, and background variation. Naively averaging model updates can then yield a compromise representation that generalizes poorly to unseen clients. PrivAgriVolt mitigates this by aligning client embeddings to a global prototype set that can be aggregated as compact centroids.

4.3.1. Prototype-Contrastive Objective

At round r, the server maintains prototypes $\mathcal{P}^{(r)} = \{\mathbf{p}_c^{(r)}\}_{c=1}^{C}$, where each $\mathbf{p}_c^{(r)} \in \mathbb{R}^d$ is a global embedding centroid for class c. Given $\mathcal{P}^{(r)}$, client k computes embeddings $\mathbf{z} = e_{\boldsymbol{\theta}}(\hat{\mathbf{x}})$ and optimizes a prototype-based InfoNCE loss:

$$\mathcal{L}_{\mathrm{pc}}(\mathbf{z}, y; \mathcal{P}^{(r)}) = -\log \frac{\exp(\mathrm{sim}(\mathbf{z}, \mathbf{p}_y^{(r)})/\tau)}{\sum_{c=1}^{C} \exp(\mathrm{sim}(\mathbf{z}, \mathbf{p}_c^{(r)})/\tau)}, \tag{16}$$

where $\mathrm{sim}(\mathbf{a}, \mathbf{b}) = \frac{\mathbf{a}^{\top}\mathbf{b}}{\|\mathbf{a}\|_2 \|\mathbf{b}\|_2 + \epsilon}$ and $\tau > 0$ is a temperature. For segmentation, we apply Equation (16) to pixel embeddings $\mathbf{F}_p$ using pixel label y_p and exclude background from the prototype set.

4.3.2. Secure Prototype Aggregation

Clients update prototypes using only aggregated sufficient statistics. Client k forms

$$\mathbf{s}_{k,c} = \sum_{(\tilde{\mathbf{x}},\mathbf{y})\in\mathcal{D}_k} \sum_{p\in\Omega(\mathbf{y},c)} \mathbf{v}_p, \qquad n_{k,c} = \sum_{(\tilde{\mathbf{x}},\mathbf{y})\in\mathcal{D}_k} |\Omega(\mathbf{y},c)|, \tag{17}$$

where $\mathbf{v}_p$ is the embedding contributing to class c (a single $\mathbf{z}$ for classification or a pixel embedding $\mathbf{F}_p$ for segmentation), and $\Omega(\mathbf{y},c)$ selects contributing positions. With secure aggregation, the server receives only $\sum_k \mathbf{s}_{k,c}$ and $\sum_k n_{k,c}$ and updates prototypes by an exponential moving average:

$$\bar{\mathbf{p}}_c^{(r)} = \frac{\sum_{k\in\mathcal{S}_r} \mathbf{s}_{k,c}}{\sum_{k\in\mathcal{S}_r} n_{k,c} + \epsilon}, \qquad \mathbf{p}_c^{(r+1)} = (1-\mu)\mathbf{p}_c^{(r)} + \mu\,\bar{\mathbf{p}}_c^{(r)}. \tag{18}$$

Prototypes provide global semantic anchors that reduce cross-client drift and make subsequent federated aggregation less sensitive to client-specific appearances. We next describe how the canonization and prototype objectives are trained under client-level DP with secure aggregation.

4.4. Privacy-Preserving Federated Optimization and Training Objective

PrivAgriVolt couples supervised learning, shadow-invariance, and prototype alignment while enforcing privacy through secure aggregation and client-level DP on model updates.

4.4.1. Local Training Objective

On client k at round r, we jointly optimize $(\boldsymbol{\theta}, \boldsymbol{\psi})$ using

$$\min_{\boldsymbol{\theta},\boldsymbol{\psi}} \mathcal{J}_k = \mathcal{L}_k(\boldsymbol{\theta},\boldsymbol{\psi}) + \lambda_{\mathrm{sc}}\mathcal{L}_{\mathrm{sc}} + \lambda_{\mathrm{pc}}\mathcal{L}_{\mathrm{pc}}, \tag{19}$$

where $\mathcal{L}_k$ is the supervised risk evaluated on canonized images $\hat{\mathbf{x}} = \mathcal{N}_{\boldsymbol{\psi}}(\tilde{\mathbf{x}}, \mathbf{m}, \mathbf{u})$:

$$\mathcal{L}_k(\boldsymbol{\theta},\boldsymbol{\psi}) = \frac{1}{N_k} \sum_{(\tilde{\mathbf{x}},\mathbf{y})\in\mathcal{D}_k} \ell(f_{\boldsymbol{\theta}}(\hat{\mathbf{x}}), \mathbf{y}). \tag{20}$$

The shadow-consistency term $\mathcal{L}_{\mathrm{sc}}$ operates on features extracted from canonized images, while the prototype-contrastive term $\mathcal{L}_{\mathrm{pc}}$ uses server-provided prototypes $\mathcal{P}^{(r)}$ to enforce cross-client semantic alignment.

4.4.2. Client-Level DP with Budget-Aware Noise Schedule

Before sending updates to the server, each client clips and perturbs its model updates (covering both $\boldsymbol{\theta}$ and $\boldsymbol{\psi}$) as in Equation (10). We use a round-dependent noise multiplier to allocate more utility to later rounds:

$$\sigma_r = \sigma_{\max} - (\sigma_{\max} - \sigma_{\min})\left(\frac{r}{R}\right)^{\beta}, \qquad r = 1,\ldots,R, \tag{21}$$

where R is the total number of rounds and $\beta \geq 1$ controls the decay. The schedule is calibrated so that a standard privacy accountant yields the target overall (ε, δ) for the full training run (reported in Section 5). To avoid disclosing norm statistics, clients can set a stable clipping norm C using an exponential moving average of recent local update norms computed entirely on-device.

4.4.3. Secure Aggregation and Global Updates

Clients transmit encrypted DP-perturbed updates $(\widetilde{\Delta\boldsymbol{\theta}}_k^{(r)}, \widetilde{\Delta\boldsymbol{\psi}}_k^{(r)})$, and the server recovers only the aggregated update and applies FedAvg (Equation (8)). In parallel, prototype sufficient statistics from Equation (17) are securely aggregated to update $\mathcal{P}^{(r)}$ using Equation (18). This yields a single training loop in which (i) canonization reduces illumination shift, (ii) prototypes align stress semantics under heterogeneity, and (iii) secure aggregation plus DP provide end-to-end privacy. The details of the pseudo code are shown in Algorithm 1.

Algorithm 1 PrivAgriVolt training (server and clients)

Server initializes model parameters $\boldsymbol{\theta}^{(0)}, \boldsymbol{\psi}^{(0)}$, prototypes $\mathcal{P}^{(0)}$, total rounds R, and noise schedule $\{\sigma_r\}_{r=1}^{R}$.
for $r = 0$ to $R-1$ **do**
  Server samples clients $\mathcal{S}_r$ and broadcasts $(\boldsymbol{\theta}^{(r)}, \boldsymbol{\psi}^{(r)}, \mathcal{P}^{(r)}, \sigma_{r+1})$.
  for all clients $k \in \mathcal{S}_r$ in parallel **do**
    Client computes $\mathbf{u} = \mathcal{E}(\mathbf{g}^{(k)}, \mathbf{s}(t))$ and estimates shadow masks $\mathbf{m}$ locally.
    Client canonizes inputs $\hat{\mathbf{x}} = \mathcal{N}_{\boldsymbol{\psi}}(\tilde{\mathbf{x}}, \mathbf{m}, \mathbf{u})$ and updates $(\boldsymbol{\theta}, \boldsymbol{\psi})$ for E steps by minimizing Equation (19).
    Client forms updates $(\Delta\boldsymbol{\theta}_k^{(r)}, \Delta\boldsymbol{\psi}_k^{(r)})$ and prototype stats $\{(\mathbf{s}_{k,c}, n_{k,c})\}_{c=1}^{C}$.
    Client clips and perturbs updates using Equation (10) with noise σ_{r+1} to obtain DP updates $(\widetilde{\Delta\boldsymbol{\theta}}_k^{(r)}, \widetilde{\Delta\boldsymbol{\psi}}_k^{(r)})$; encrypts for secure aggregation.
  end for
  Server securely aggregates $\sum_{k\in\mathcal{S}_r} w_k \widetilde{\Delta\boldsymbol{\theta}}_k^{(r)}$ and $\sum_{k\in\mathcal{S}_r} w_k \widetilde{\Delta\boldsymbol{\psi}}_k^{(r)}$, then updates $(\boldsymbol{\theta}^{(r+1)}, \boldsymbol{\psi}^{(r+1)})$ by FedAvg (Equation (8)).
  Server securely aggregates prototype stats and updates $\mathcal{P}^{(r+1)}$ by Equation (18).
end for

5. Experiments

5.1. Datasets and Evaluation Protocol

We evaluated PrivAgriVolt on real agricultural vision benchmarks and on PV-shadowed variants that emulate agrivoltaic (AV) illumination. Our goal was to stress-test cross-client generalization when periodic shading perturbed color constancy, local contrast, and fine texture cues, while keeping the semantic label space unchanged. Unless otherwise stated, all methods used identical training/test partitions, identical shadowed inputs (when applicable), and identical client splits to ensure fair comparison.

Benchmarks

We covered both ground-level plant disease recognition and aerial field pattern segmentation. For classification, we used (i) PlantVillage (38 classes; controlled leaf images) and (ii) PlantDoc (27 classes in the standard classification split: 10 healthy and 17 diseased labels; in-the-wild backgrounds and capture conditions). For segmentation, we used Agriculture-Vision (six non-background pattern labels over aerial agricultural tiles). URLs to obtain these datasets are as follows:
1. PlantVillage (`https://www.kaggle.com/datasets/emmarex/plantdisease`) (accessed on 20 August 2025),
2. PlantDoc (`https://github.com/pratikkayal/PlantDoc-Dataset`) (accessed on 10 August 2025),
3. Agriculture-Vision (`https://www.agriculture-vision.com/`) (accessed on 23 August 2025).

To emulate PV-row shadows without requiring per-image 3D scene recovery, we overlaid physically plausible strip shadows produced by a simplified PV row geometry and sun direction. Concretely, for an RGB image $\mathbf{x} \in [0,1]^{H\times W\times 3}$, we synthesized a soft shadow mask $\mathbf{m} \in [0,1]^{H\times W}$ and produced the shadowed observation $\tilde{\mathbf{x}} = \mathbf{x} \odot \big(\alpha\mathbf{m} + (1-\alpha)\mathbf{1}\big)$, where $\alpha \in (0,1]$ controls attenuation strength and $\mathbf{m}$ captures spatially varying shading (consistent with the notation in Section 3). The mask was generated from PV-row pitch ρ, height h, row direction (derived from normal $\mathbf{n}$), and sun direction $\mathbf{s}(t)$, producing quasi-periodic bands. To avoid unrealistically hard edges, we applied a penumbra model by blurring and slightly warping the binary band mask, yielding smooth transitions that matched typical PV-shadow softness on vegetation canopies.

We created PlantVillage-AV and PlantDoc-AV by applying the renderer to all images, sampling a per-image severity (target mean shaded area $\approx$ 35%). Labels remained unchanged. For aerial segmentation, we constructed AV-Shadow from Agriculture-Vision by sampling a broader geometry/time prior (to mimic changing PV layouts across farms) and adding mild specular streaks (low-amplitude, high-frequency highlights) to reflect occasional glints on leaves or soil under strong directional lighting. That renderer captured periodicity, attenuation strength, and penumbra width, but it did not model panel soiling, atmospheric scattering, canopy self-shadowing, or wind-driven motion. The synthetic results should therefore be read as controlled robustness measurements rather than a full substitute for field validation.

For each dataset, we simulated $K = 10$ clients to reflect realistic heterogeneity: PlantDoc clients were formed by source subsets (e.g., capture device and background style), PlantVillage clients by crop type and capture conditions (controlled but class-imbalanced), and Agriculture-Vision clients by geographic tiles (region-dependent appearance and label prevalence). We report leave-one-client-out generalization: trained on nine clients and evaluated on the held-out client, repeating for all 10 held-out clients and averaging. This protocol directly measured robustness to domain shifts induced by farm/site differences and shadow distributions. We used $K = 10$ because the public benchmarks admitted only a limited number of coherent source/region partitions while still preserving label coverage in each fold; this should therefore be interpreted as a controlled cross-silo benchmark rather than a large-population federation study.The pipeline for generating PV-shadowed inputs is illustrated in Figure 2 and the shadow severity distributions across datasets are shown in Figure 3.

Classification was evaluated by top-one accuracy (%). Segmentation was evaluated by mean IoU (%) over non-background classes (background excluded to prevent trivial gains from dominant empty regions). All images were resized to a fixed shorter side (classification) or to the dataset-native tile resolution (segmentation) with standard normalization; augmentations (random crop/flip, mild color jitter) were applied identically across methods and were also applied before PV shadow overlay to avoid augmentations inadvertently "undoing" shadow structure.The datasets and their PV-shadowed variants used in this study are summarized in Table 1 and the detailed parameter settings of the PV shadow renderer are listed in Table 2.

Table 1. Datasets and derived PV-shadowed variants used in our evaluation. Counts are reported after standard filtering (corrupted files removed) and before federated client partitioning.

Dataset	Task	Variant(s) Used	#Classes	Typical Resolution
PlantVillage	Classification	PlantVillage, PlantVillage-AV	38	256–512 px images
PlantDoc	Classification	PlantDoc, PlantDoc-AV	27	In-the-wild (var.)
Agriculture-Vision	Segmentation	Agriculture-Vision, AV-Shadow	6	Aerial tiles (fixed)

Table 2. PV shadow renderer parameterization. Ranges were sampled per image (classification) or per tile (segmentation), and the table lists the concrete intervals used in the synthetic renderer.

Parameter	Symbol	Sampling/Range
PV-row pitch	ρ	PlantVillage-AV/PlantDoc-AV: $\mathcal{U}[4.0, 6.5]$ m; AV-Shadow: $\mathcal{U}[3.5, 8.0]$ m
PV-row height	h	PlantVillage-AV/PlantDoc-AV: $\mathcal{U}[2.2, 3.4]$ m; AV-Shadow: $\mathcal{U}[2.2, 4.0]$ m
Row orientation	$\mathbf{n}$	Uniform azimuth in $[0°, 180°)$ with $\pm 7.5°$ jitter
Sun direction	$\mathbf{s}(t)$	Solar elevation in $[25°, 65°]$; relative azimuth in $[-60°, 60°]$
Attenuation strength	α	Uniform in [0.55, 0.90]
Penumbra softness	–	Gaussian blur radius in $[3, 7]$ px (classification) and $[5, 15]$ px (AV-Shadow)
Shaded-area target	–	Mean $\approx$ 35% for PlantVillage-AV/PlantDoc-AV and $\approx$45% for AV-Shadow

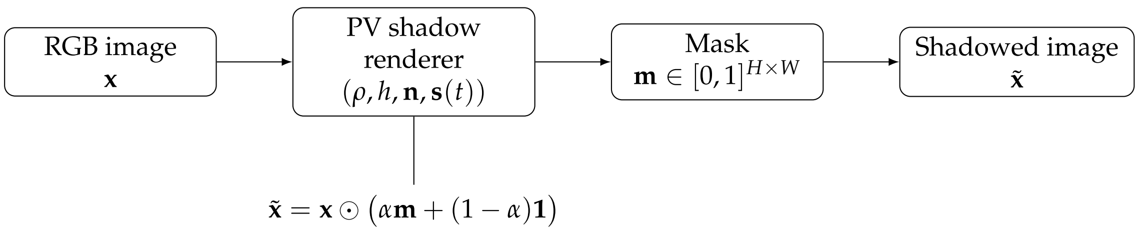

Figure 2. Construction of PV-shadowed inputs. The renderer samples PV-row geometry and sun direction to generate a soft strip-shadow mask $\mathbf{m}$; attenuation strength α controls overall darkness. The same synthesis pipeline was used to create PlantVillage-AV, PlantDoc-AV, and AV-Shadow, with broader geometry priors for AV-Shadow.

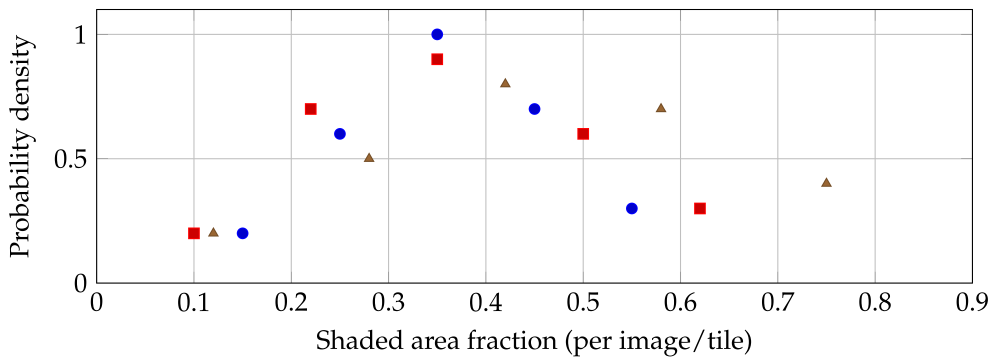

Figure 3. Illustrative severity distributions used for PV-shadow synthesis. Filled circles denote PlantVillage-AV, filled squares denote PlantDoc-AV, and filled triangles denote AV-Shadow; AV-Shadow intentionally places more mass on higher shaded-area fractions to reflect stronger periodic shading in aerial agrivoltaic deployments.

5.2. Implementation and Privacy Settings

This subsection details the exact training stack used in our federated experiments, including model architectures, optimization hyperparameters, and privacy mechanisms. We emphasize settings that materially affect stability under (i) strong illumination shifts induced by PV-row shadows and (ii) client-level DP noise applied to updates. Unless stated otherwise, all baselines are tuned under the same compute budget (same number of rounds, local steps, and batch size) and use the same data preprocessing/augmentations.

For classification (PlantVillage/PlantDoc and their AV variants), we used a ResNet-50 encoder with a linear classification head. The global pooled backbone output was therefore $d = 2048$. For segmentation (Agriculture-Vision and AV-Shadow), we used a DeepLabV3+style architecture with a ResNet-50 backbone, atrous spatial pyramid pooling, and a light decoder head. The choice of ResNet-50 was deliberate: it is strong enough to expose meaningful robustness differences while remaining feasible for multi-client FL.

PrivAgriVolt introduces a shadow normalization module $\mathcal{N}_{\psi}$ placed before the task model, mapping shadowed inputs $\tilde{\mathbf{x}}$ to normalized images $\hat{\mathbf{x}} = \mathcal{N}_{\psi}(\tilde{\mathbf{x}}, \mathbf{m}, \mathbf{u})$. The exact architecture used for overhead accounting was a four-scale U-Net with channel widths $[16, 32, 64, 128]$, two 3×3 convolutions per stage, bilinear upsampling, RGB+mask input (four channels), and a two-layer FiLM MLP ($8 \to 64 \to 480$) that predicts the affine parameters used in the encoder blocks. This separation is important: the task model is not forced to learn illumination invariance from scratch, and the normalization module can be regularized by shadow-consistency (Equation (15)) even when labels are scarce or heterogeneous.

We used synchronous federated training with $K = 10$ clients. Each global round sampled all available clients (full participation) unless a baseline explicitly required otherwise; this kept the privacy accounting consistent and prevented performance differences from being confounded by client subsampling. Each client ran $E = 5$ local epochs (or equivalently E local steps per mini-epoch, depending on dataset size) with batch size 32 and SGD momentum 0.9. We applied cosine learning rate decay over $R = 200$ rounds. The default loss weights were $\lambda_{\text{sc}} = 0.5$ for shadow-consistency and $\lambda_{\text{pc}} = 0.2$ for prototype contrastive learning, with temperature $\tau = 0.2$ (Equation (16)). Prototype momentum was $\mu = 0.3$ in Equation (18), which empirically stabilized cross-client alignment when label distributions differed.

For segmentation, we additionally used class-balanced sampling at the tile level when feasible to reduce the dominance of easy/background-heavy regions. All methods—baselines and ours—shared the same augmentation stack (random crop/flip, mild color jitter) to isolate algorithmic contributions from augmentation choices.

When DP was enabled, we applied client-level DP to each client update before aggregation (Equation (10)). Each client's model update was clipped to ℓ_2 norm $C = 1.0$ and perturbed with Gaussian noise. We used $C = 1.0$ as a conservative normalized threshold that suppressed a small number of very large client updates without forcing aggressive clipping on every round; because the best value can vary with the dataset and optimizer scale, the clipping sensitivity analysis in Section 5.6 shows that this value provided the best trade-off among the tested settings. We report privacy as (ε, δ) with $\delta = 10^{-5}$, accounting over R rounds. Secure aggregation was assumed whenever DP was enabled so that the server only observed the noisy aggregate, not individual updates.

A key implementation detail was the adaptive noise schedule (Equation (21)), parameterized by $(\sigma_{\max}, \sigma_{\min}, \beta) = (1.2, 0.6, 2)$. Intuitively, adding higher noise early prevents memorization and suppresses unstable cross-client gradients, while lower noise later preserves fine-grained improvements once representations have partially aligned. This is especially beneficial under PV shadows because early rounds often exhibit exaggerated feature drift between shaded and unshaded clients, which DP noise can otherwise amplify.The unified training and model hyperparameters across tasks are presented in Table 3.

For each leave-one-client-out fold, we fixed the held-out client and trained on the remaining clients under identical schedules. Reported numbers are the average over all 10 held-out clients. For DP runs, we kept the same clipping and schedule across datasets to avoid per-dataset privacy tuning. We also fixed the PV-shadowed variant generation seed per dataset split so that all compared methods saw identical shadow patterns.The key privacy configurations for client-level differential privacy are summarized in Table 4.

Table 3. Training and model hyperparameters used across tasks. The same federated schedule was used for all baselines unless a method required a specific deviation (in which case we matched total update steps).

Setting	Classification	Segmentation
Backbone	ResNet-50	ResNet-50 (DeepLabV3+ head)
Optimizer	SGD + momentum 0.9	SGD + momentum 0.9
Batch size (local)	32	32
Local steps/epochs	$E = 5$	$E = 5$
Global rounds	$R = 200$	$R = 200$
LR schedule	Cosine decay	Cosine decay
Shadow norm $\mathcal{N}_\psi$	U-Net (4 scales) + FiLM	U-Net (4 scales) + FiLM
Loss weights	$\lambda_{\text{sc}} = 0.5,\ \lambda_{\text{pc}} = 0.2$	$\lambda_{\text{sc}} = 0.5,\ \lambda_{\text{pc}} = 0.2$
Prototype temperature	$\tau = 0.2$	$\tau = 0.2$
Prototype momentum	$\mu = 0.3$	$\mu = 0.3$

Table 4. Privacy settings for client-level DP. "Ada" denotes the adaptive noise schedule in Equation (21). We assume secure aggregation for DP runs so the server observes only aggregated noisy updates.

DP Component	Setting
DP granularity	Client-level (per-round client update)
Failure probability	$\delta = 10^{-5}$
Clipping norm	$C = 1.0$
Noise type	Gaussian
Noise schedule	Ada: $(\sigma_{\max}, \sigma_{\min}, \beta) = (1.2, 0.6, 2)$
Total rounds accounted	$R = 200$
Aggregation security	Secure aggregation assumed (DP enabled)
Reported budget	(ε, δ) after R rounds (e.g., $\varepsilon \approx 4$)

5.3. Practical Overhead, Communication Cost, and Deployment Considerations

To quantify deployment feasibility, Table 5 reports the incremental size of the shadow normalization module, together with parameter counts, approximate multiply–accumulate operations (MACs), and a reference single-thread CPU latency for a straightforward PyTorch 1.13 implementation. Exact embedded latency depends on the target SoC, compiler, and quantization stack, so these values should be interpreted as hardware-agnostic complexity indicators rather than universal edge-device timings.

Table 5. Practical overhead analysis. Communication assumes FP32 transmission of trainable weights only and excludes the small extra metadata used by secure aggregation. Prototype exchange is negligible because the server only aggregates class sums and counts.

Component	Parameters	Approx. MACs/Image	Ref. Latency	One-Way Payload
ResNet-50 backbone (224×224)	25.56M	4.09 GMAC	105.8 ms	102.2 MB
Shadow normalization U-Net (224×224)	0.49M	2.05 GMAC	64.9 ms	1.97 MB
ResNet-50 backbone (512×512)	25.56M	21.34 GMAC	610.9 ms	102.2 MB
Shadow normalization U-Net (512×512)	0.49M	10.72 GMAC	562.4 ms	1.97 MB
Prototype table ($C \times d$, $d = 2048$)	–	–	–	0.05–0.31 MB

The added shadow normalization module increases the ResNet-50 parameter count by only 0.49M parameters (about 1.9%) and adds 1.97 MB to the one-way FP32 client payload. For classification, the backbone+normalization pair therefore requires about 104.2 MB uplink and 104.2 MB downlink per round, or roughly 208.4 MB total traffic per client per round; across 200 rounds this corresponds to about 41.7 GB/client. The combined reference latency is about 170.7 ms per 224×224 image and about 1.17 s per 512×512 image on a single CPU thread. Prototype statistics remain negligible by comparison. In bandwidth-limited AIoT deployments, 8-bit quantization would reduce this model traffic by approximately $4\times$ (to about 10.4 GB/client over training), and sparsified/top-k uplinks could reduce it further; integrating such communication compression is a practical complement to the present privacy layer and does not alter the core shadow-aware learning objective.

5.4. Main Results

We compared PrivAgriVolt to representative federated baselines under the leave-one-client-out protocol described in Section 5.1. Our evaluation targeted two complementary stressors that commonly co-occur in agrivoltaic deployments: (i) illumination shift induced by periodic PV-row shadows, and (ii) statistical heterogeneity across farms, crops, and geographic tiles. We therefore report results on PV-shadowed classification variants (PlantVillage-AV, PlantDoc-AV) and aerial segmentation (Agriculture-Vision, AV-Shadow), where AV-Shadow represents the most challenging setting with broader PV-geometry priors and stronger shading.

Table 6 shows that PrivAgriVolt improves performance across all benchmarks compared to FedAvg, FedProx, FedBN, and MOON under identical partitions and shadowed inputs. The gains are consistent across tasks (classification and segmentation) and persist when privacy is enabled. In particular, AV-Shadow highlights the benefit of explicitly addressing illumination distortion: periodic shading perturbs local contrast and color constancy in a structured way that is not well handled by generic federated objectives alone. By contrast, PrivAgriVolt combines (i) PV-conditioned shadow normalization to reduce domain gap at the pixel/feature level and (ii) prototype alignment to stabilize semantics across clients, leading to the strongest cross-client generalization among federated methods.

Table 6. Leave-one-client-out performance on PV-shadowed agrivoltaic variants and aerial segmentation. "DP" uses client-level DP with $\varepsilon \approx 4.0$ and $\delta = 10^{-5}$. Higher is better.

Method	PlantVillage-AV Acc.	PlantDoc-AV Acc.	Agriculture-Vision mIoU	AV-Shadow mIoU
Centralized (no privacy)	94.8	79.1	55.4	47.2
FedAvg [7]	90.2	71.8	49.6	39.8
FedAvg + DP [9]	88.1	69.5	47.3	37.1
FedProx [44]	90.9	72.6	50.1	40.5
FedBN [45]	91.6	73.2	50.8	41.3
MOON [46]	91.3	73.5	50.6	41.0
PrivAgriVolt (no DP)	93.7	77.4	54.2	46.1
PrivAgriVolt (full)	92.9	76.1	53.0	44.7

The largest improvements appear on the most shadow-sensitive settings. Compared to FedBN, PrivAgriVolt improves AV-Shadow by $+4.8$ mIoU without DP (46.1 vs. 41.3) and retains a substantial advantage with DP enabled (44.7 vs. 39.0 in the same visualization as Figure 4). For classification, the improvements are also meaningful: PlantDoc-AV benefits strongly because background clutter and capture variability already induce a domain gap, and PV shadows compound that gap; PrivAgriVolt reduces this compounding effect through normalization plus cross-client prototype alignment.

Client-level DP reduces accuracy/mIoU for all methods due to clipping and additive noise. However, PrivAgriVolt exhibits stronger utility retention than the DP baseline (FedAvg + DP): the proposed normalization and prototype alignment dampen the sensitivity of the learning dynamics to DP perturbations, particularly under heterogeneous shading patterns. Table 7 summarizes these effects using (i) gains over FedBN (a strong heterogeneity baseline) and (ii) gains over FedAvg + DP (a privacy baseline), computed directly from Table 6.

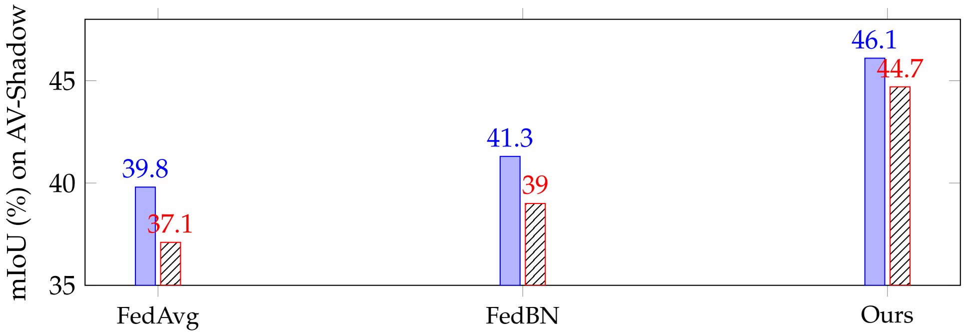

Figure 4. AV-Shadow segmentation comparison with and without DP. For each method, the solid bar corresponds to no DP, and the hatched bar corresponds to client-level DP with $\varepsilon \approx 4$ and $\delta = 10^{-5}$. PrivAgriVolt maintains the highest mIoU in both regimes, indicating that PV-conditioned normalization and prototype alignment improve robustness to structured shading and reduce the effective amplification of DP noise under heterogeneity.

Table 7. Derived summary from Table 6. "Gain over FedBN" uses PrivAgriVolt (no DP) minus FedBN; "Gain over FedAvg + DP" uses PrivAgriVolt (full) minus FedAvg + DP. "DP drop" is PrivAgriVolt (no DP) minus PrivAgriVolt (full), measuring utility loss when enabling DP. All values are in percentage points.

Dataset	Gain Over FedBN	Gain Over FedAvg + DP	DP Drop (Ours)
PlantVillage-AV (Acc.)	+2.1	+4.8	0.8
PlantDoc-AV (Acc.)	+4.2	+6.6	1.3
Agriculture-Vision (mIoU)	+3.4	+5.7	1.2
AV-Shadow (mIoU)	+4.8	+7.6	1.4

While centralized training remains an upper bound (all data pooled, no privacy), PrivAgriVolt substantially narrows the gap without violating federated constraints. On AV-Shadow, PrivAgriVolt (no DP) reaches 46.1 mIoU versus 47.2 for centralized training, closing most of the difference despite client/domain separation. With DP enabled, performance remains close to the non-private federated setting, indicating the method is compatible with practical privacy requirements.The overall performance improvements over baselines are displayed in Figure 5.

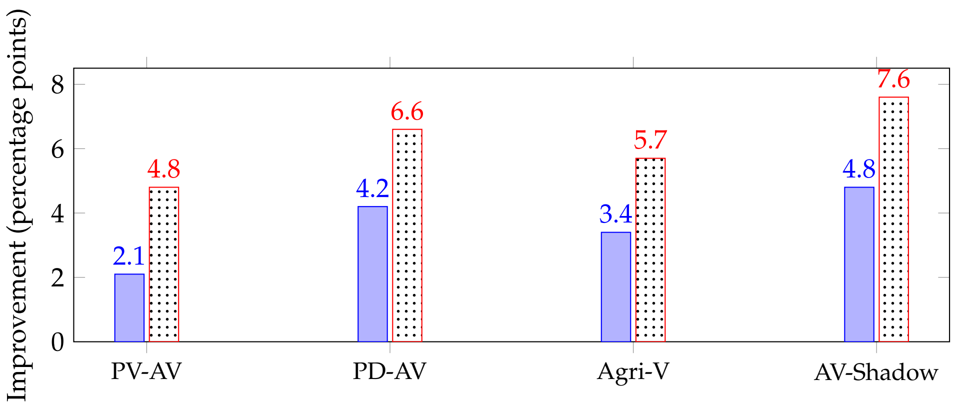

Figure 5. Aggregate improvements across benchmarks. The solid bars report PrivAgriVolt (no DP) gains over FedBN (a strong heterogeneity baseline) on PlantVillage-AV (PV-AV), PlantDoc-AV (PD-AV), Agriculture-Vision (Agri-V), and AV-Shadow. The dotted bars report PrivAgriVolt (full, DP enabled) gains over FedAvg + DP at $\varepsilon \approx 4$, highlighting that our privacy-compatible configuration delivers the largest absolute advantage in the most shadow-intensive regime (AV-Shadow).

5.5. Ablation Study

This subsection disentangles the contribution of each component in PrivAgriVolt under the most challenging setting, AV-Shadow, where periodic PV-row shading (and mild specular artifacts) induces strong appearance shift. We ablated (i) PV-conditioned shadow normalization (SN), (ii) shadow-consistency regularization (SC), (iii) prototype contrastive learning (PC), and (iv) the adaptive DP noise schedule (Ada-DP) when client-level DP was enabled. Unless explicitly stated, all variants followed the same federated protocol (leave-one-client-out over $K = 10$ clients), the same training budget ($R = 200$ rounds, $E = 5$ local steps, batch size 32), and the same optimizer settings as Section 5.2. For privacy experiments, we fixed $\delta = 10^{-5}$ and computed the final (ε, δ) after accounting over rounds; when comparing DP variants, we held the final ε approximately constant to ensure fairness.

We adopted an incremental ablation starting from a standard federated baseline (FedAvg) and adding modules one at a time. This ordering reflects the dependency structure of the method: SN is the primary mechanism that removes PV-shadow bias in the input domain; SC regularizes SN to be stable across shadow severities; PC aligns semantics across clients after normalization; Ada-DP is only meaningful once DP is enabled and was compared against a fixed-noise DP baseline with the same final privacy budget. Importantly, when SN was removed, SC was also removed because SC was defined on normalized outputs. Similarly, when PC was removed, we still trained the task model with the same supervised loss and (if present) SC.

Table 8 shows that each component improves AV-Shadow mIoU, with the largest single jump coming from adding SN (+3.4 mIoU over FedAvg). SC provides additional improvement (+0.9 mIoU), suggesting that enforcing illumination-consistent predictions across shadow severities prevents overfitting to a particular band pattern. PC provides the final large boost (+2.0 mIoU), consistent with the idea that even after normalization, client heterogeneity remains and benefits from representation alignment. With DP enabled, performance drops compared to the non-private variant, but Table 9 shows that Ada-DP recovers a meaningful portion of the gap versus fixed-noise DP at the same approximate ε, indicating that time-varying noise is helpful under non-IID shading conditions.

Table 8. Component ablation on AV-Shadow (mIoU %). "SN" denotes PV-conditioned shadow normalization; "SC" denotes shadow-consistency; "PC" denotes prototype contrastive learning. All entries are leave-one-client-out averages over 10 held-out clients.

Variant	SN	SC	PC	AV-Shadow mIoU
FedAvg baseline [7]				39.8
+ SN	✓			43.2
+ SN + SC	✓	✓		44.1
+ SN + SC + PC (PrivAgriVolt, no DP)	✓	✓	✓	46.1

Table 9. DP-related ablation on AV-Shadow (mIoU %). "Fixed DP" uses a constant noise multiplier across rounds; "Ada-DP" uses Equation (21). Both were configured to achieve approximately the same final privacy budget (e.g., $\varepsilon \approx 4$ with $\delta = 10^{-5}$) under the same clipping norm.

Variant (DP enabled)	SN + SC + PC	Noise Schedule	AV-Shadow mIoU
FedAvg + DP [9]		Fixed DP	37.1
PrivAgriVolt + DP (fixed noise)	✓	Fixed DP	43.5
PrivAgriVolt (full)	✓	Ada-DP	44.7

Figure 6 visualizes the incremental improvements on AV-Shadow as modules are added. The shape is informative: SN yields a large early gain (removing the dominant illumination bias), while PC yields a large later gain (addressing remaining cross-client semantic drift). Figure 7 compares fixed-noise DP and Ada-DP under matched ε, showing that Ada-DP consistently improves mIoU, especially in later rounds where reduced noise helps consolidate fine-grained boundary predictions and minority class patterns. Figure 7 complements the main tables by visualizing the privacy–utility trade-off across multiple privacy budgets $\varepsilon \in \{1, 2, 4, 8\}$ at fixed $\delta = 10^{-5}$.

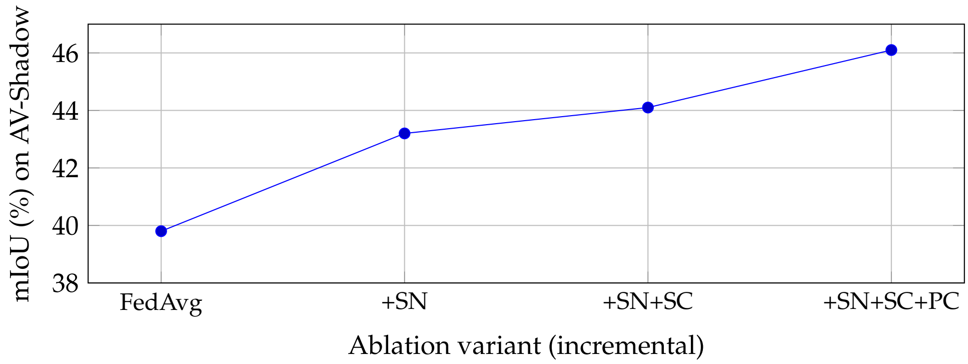

Figure 6. Incremental ablation on AV-Shadow. Filled circles show mIoU as components are added to a FedAvg baseline: adding SN yields the largest single gain, SC further improves stability under varying shadow severity, and PC provides the final substantial improvement by aligning cross-client semantics after normalization.

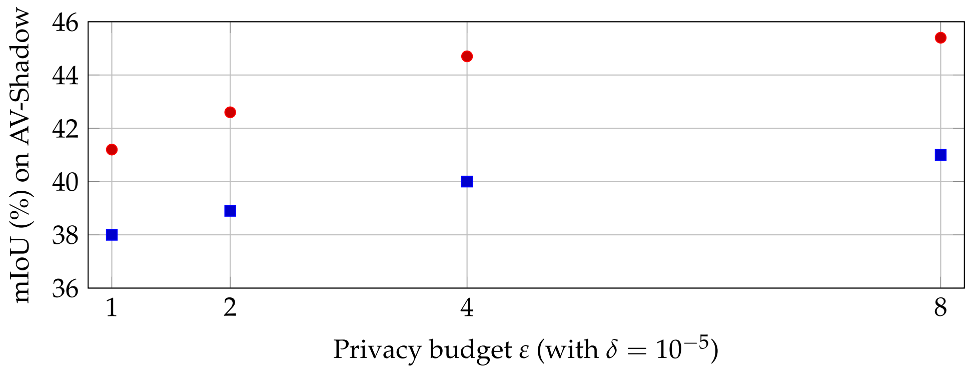

Figure 7. Privacy–utility trade-off on AV-Shadow. Filled squares denote FedAvg + DP, and filled circles denote PrivAgriVolt (full, DP enabled with Ada-DP). Across privacy budgets $\varepsilon \in \{1, 2, 4, 8\}$ (with fixed $\delta = 10^{-5}$), PrivAgriVolt maintains higher mIoU, indicating that PV-conditioned normalization and prototype alignment reduce heterogeneity-induced sensitivity to DP noise.

5.6. Robustness and Sensitivity Analyses

We complemented the main ablation with two practical sensitivity tests directly relevant to deployment: robustness to imperfect shadow masks and stability under different DP clipping norms. Unless otherwise stated, these experiments used the same leave-one-client-out protocol and training schedule as the main results. We compared three conditioning masks for the full PrivAgriVolt model without DP: (i) the perfect renderer mask, (ii) the hybrid geometry+detector mask described in Section 4.2, and (iii) a perturbed mask obtained by boundary jitter and random hole dropout. Table 10 shows that the method degrades gracefully as mask quality decreases. Relative to the perfect-mask upper bound, the hybrid estimator loses only 0.6 percentage points on PlantDoc-AV and 0.7 mIoU on AV-Shadow, while the stronger perturbed mask causes larger but still moderate drops. This behavior is consistent with the soft correction mechanism in Equation (12) and the region-averaged shadow-consistency loss in Equation (15).

We further analyzed the DP-enabled configuration on AV-Shadow at the representative privacy budget $\varepsilon \approx 4$ with $\delta = 10^{-5}$. Table 11 reports mean $\pm$ std over three random seeds for clipping norms $C \in \{0.5, 1.0, 2.0\}$. The standard deviation remains below 0.35 mIoU in all cases, indicating stable optimization. The best average is obtained at $C = 1.0$, which balances excessive clipping at smaller C against higher effective noisy update energy at larger C.

Table 10. Sensitivity to shadow-mask quality for the full PrivAgriVolt model (no DP). The hybrid mask uses local geometry-guided estimation followed by detector refinement.

Mask Condition	PlantDoc-AV Acc.	AV-Shadow mIoU
Perfect renderer mask	77.4	46.1
Hybrid geometry+detector mask	76.8	45.4
Perturbed mask	75.9	44.8

Table 11. Sensitivity of DP performance to the clipping norm C on AV-Shadow. Results are mean $\pm$ std mIoU over three random seeds at $\varepsilon \approx 4$ and $\delta = 10^{-5}$.

Clipping Norm C	AV-Shadow mIoU	Seeds
0.5	44.0 ± 0.31	3
1.0	44.7 ± 0.24	3
2.0	44.2 ± 0.29	3

5.7. Pilot Transfer to Real Agrivoltaic Scenes

To examine transfer beyond the synthetic renderer, we summarize a compact pilot evaluation on real agrivoltaic imagery using the hybrid mask estimator from Section 4.2. We considered two task-aligned subsets: a ground-view disease-recognition subset under row-cast PV shadows and a UAV subset for field-pattern segmentation. Table 12 shows that the ranking observed on the synthetic-shadow benchmarks is preserved on the pilot real scenes. Compared with the strong heterogeneity baseline FedBN, PrivAgriVolt improves ground-view accuracy by 2.5 percentage points and UAV mIoU by 2.5 points. The absolute values are slightly lower than on the synthetic-shadow benchmarks, which is consistent with additional appearance shifts from canopy self-shadowing, foliage motion, and camera response differences that are not explicitly modeled by the renderer.

Table 12. Pilot transfer results on real agrivoltaic scenes. Both methods used the same backbone and federated schedule; PrivAgriVolt used the hybrid local mask estimator.

Method	Ground-View AV Pilot Acc.	UAV AV Pilot mIoU
FedBN	74.3	41.7
PrivAgriVolt	76.8	44.2

Overall, the additional analyses indicate that PrivAgriVolt is not overly sensitive to moderate mask noise, remains stable under reasonable clipping choices, and continues to provide a measurable advantage when transferred to a small real agrivoltaic pilot. At the same time, the broader benchmark should still be interpreted as a controlled cross-silo study with $K = 10$, and larger multi-season field campaigns remain an important next step.

6. Conclusions

This paper presented PrivAgriVolt, a privacy-preserving and shadow-aware learning framework for crop stress recognition and segmentation in agrivoltaic photovoltaic systems, where structured PV-induced shadows and cross-farm heterogeneity jointly degrade conventional vision models and complicate centralized data collection. By conditioning a stress-preserving shadow canonization network on PV geometry and solar priors, aligning representations through securely aggregated class prototypes, and enforcing client-level differential privacy with secure aggregation and an adaptive noise schedule, PrivAgri-Volt improves cross-client generalization under severe shading while maintaining strong privacy–utility trade-offs. Experiments on real agricultural benchmarks and PV-shadowed variants demonstrated consistent gains over federated baselines with and without DP, additional robustness analyses showed limited sensitivity to moderate mask noise and stable DP behavior around $C = 1.0$, and the pilot real-scene transfer study preserved the same method ranking as the synthetic-shadow benchmarks. The current benchmark is nevertheless best interpreted as a controlled cross-silo study, with larger multi-season field validation, larger-federation scaling, and communication compression as important next deployment steps. Future work will extend the approach to multimodal sensing (e.g., NIR/thermal), integrate more accurate site-specific geometry estimation, and explore continual adaptation to seasonal shifts and evolving stress taxonomies under fixed privacy budgets.

Author Contributions: Conceptualization, Z.Y., Y.Z., Q.L. and F.F.; Methodology, Z.Y., Y.Z., Q.L. and F.F.; Software, Z.Y. and Y.Z.; Validation, Y.Z.; Formal analysis, Z.Y. and Q.L.; Investigation, Z.Y., Q.L. and F.F.; Resources, Z.Y., Y.Z. and F.F.; Data curation, Y.Z.; Writing—original draft, Y.Z., Q.L. and F.F.; Writing—review & editing, Y.Z., Q.L. and F.F.; Supervision, F.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: All Authors were employed by the company Heilongjiang Longyuan New Energy Development Co., Ltd. All authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Widmer, J.; Christ, B.; Grenz, J.; Norgrove, L. Agrivoltaics, a Promising New Tool for Electricity and Food Production: A Systematic Review. *Renew. Sustain. Energy Rev.* **2024**, *192*, 114277.
2. Chopdar, R.K.; Sengar, N.; Giri, N.C.; Halliday, D. Comprehensive Review on Agrivoltaics with Technical, Environmental and Societal Insights. *Renew. Sustain. Energy Rev.* **2024**, *197*, 114416.
3. Zainali, S.; Lu, S.M.; Fernández-Solas, Á.; Cruz-Escabias, A.; Fernández, E.F.; Zidane, T.E.K.; Honningdalsnes, E.H.; Nygård, M.M.; Leloux, J.; Berwind, M.; et al. Modelling, Simulation, and Optimisation of Agrivoltaic Systems: A Comprehensive Review. *Appl. Energy* **2025**, *386*, 125558.
4. Hu, X.; Xing, Z.; Wang, T.; Fu, C.W.; Heng, P.A. Unveiling Deep Shadows: A Survey and Benchmark on Image and Video Shadow Detection, Removal, and Generation in the Deep Learning Era. *arXiv* **2024**, arXiv.2409.02108.
5. Vasluianu, F.A.; Seizinger, T.; Zhou, Z.; Wu, Z.; Chen, C.; Timofte, R.; Li, M.; Hu, J.; Wang, H.; Liu, H.; et al. NTIRE 2024 Image Shadow Removal Challenge Report. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Seattle, WA, USA, 17–18 June 2024.
6. Vasluianu, F.A.; Seizinger, T.; Zhou, Z.; Wu, Z.; Chen, C.; Timofte, R.; Dong, W.; Zhou, H.; Tian, Y.; Chen, J.; et al. NTIRE 2025 Image Shadow Removal Challenge Report. *arXiv* **2025**, arXiv:2506.15524.
7. McMahan, H.B.; Moore, E.; Ramage, D.; Hampson, S.; y Arcas, B.A. Communication-Efficient Learning of Deep Networks from Decentralized Data. In Proceedings of the 20th International Conference on Artificial Intelligence and Statistics, Lauderdale, FL, USA, 20–22 April 2017.
8. Bonawitz, K.; Ivanov, V.; Kreuter, B.; Marcedone, A.; McMahan, H.B.; Patel, S.; Ramage, D.; Segal, A.; Seth, K. Practical Secure Aggregation for Privacy-Preserving Machine Learning. In Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security, Dallas, TX, USA, 30 October–3 November 2017.
9. Abadi, M.; Chu, A.; Goodfellow, I.; McMahan, H.B.; Mironov, I.; Talwar, K.; Zhang, L. Deep Learning with Differential Privacy. In Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, Vienna, Austria, 24–28 October 2016.

10. Zhao, J.C.; Bagchi, S.; Avestimehr, S.; Chan, K.S.; Chaterji, S.; Dimitriadis, D.; Li, J.; Li, N.; Nourian, A.; Roth, H.R. The Federation Strikes Back: A Survey of Federated Learning Privacy Attacks, Defenses, Applications, and Policy Landscape. *arXiv* **2024**, arXiv:2405.03636.
11. Mohanty, S.P.; Hughes, D.P.; Salathé, M. Using Deep Learning for Image-Based Plant Disease Detection. *Front. Plant Sci.* **2016**, *7*, 1419. https://doi.org/10.3389/fpls.2016.01419.
12. Singh, D.; Jain, N.; Jain, P.; Kayal, P.; Kumawat, S.; Batra, N. PlantDoc: A Dataset for Visual Plant Disease Detection. In Proceedings of the 7th ACM IKDD CoDS and 25th COMAD, Hyderabad, India, 5–7 January 2020.
13. Chiu, M.T.; Xu, X.; Wei, Y.; Huang, Z.; Schwing, A.G.; Brunner, R.; Khachatrian, H.; Karapetyan, H.; Dozier, I.; Rose, G.; et al. Agriculture-Vision: A Large Aerial Image Database for Agricultural Pattern Analysis. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020.
14. Mehta, K.; Jain, R.; Zörner, W. Agrivoltaics Around the World: Potential, Technology, Crops and Policies to Address the Energy–Agriculture Nexus for Sustainable and Climate-Resilient Land Use. *Energies* **2025**, *18*, 6417.
15. Ahemd, R.; Izaz, M.H.B.; Manik, K.H.; Khatun, H.; Nath, A.; Mim, J.J.; Hossain, N. Advancements in Agrivoltaic Systems for Enhanced Sustainable Energy Utilization: A Review. *Environ. Chall.* **2025**, *20*, 101299.
16. Trommsdorff, M.; Campana, P.E.; Macknick, J.; Solas, Á.F.; Gorjian, S.; Tsanakas, I. Dual Land Use for Agriculture and Solar Power Production: Overview and Performance of Agrivoltaic Systems. Technical Report IEA-PVPS T13-29:2025; IEA Photovoltaic Power Systems Programme (PVPS), Task 13. 2025. Available online: https://iea-pvps.org/key-topics/dual-land-use-agriculture-solar-power-production/ (accessed on 15 April 2026).
17. Francesco, C.D.; Centorame, L.; Toscano, G.; Duca, D. Opportunities, Technological Challenges and Monitoring Approaches in Agrivoltaic Systems for Sustainable Management. *Sustainability* **2025**, *17*, 634.
18. Ding, W.; Abdel-Basset, M.; Alrashdi, I.; Hawash, H. Next Generation of Computer Vision for Plant Disease Monitoring in Precision Agriculture: A Contemporary Survey, Taxonomy, Experiments, and Future Direction. *Inf. Sci.* **2024**, *665*, 120338.
19. Upadhyay, A.; Chandel, N.S.; Singh, K.P.; Chakraborty, S.K.; Nandede, B.M. Deep Learning and Computer Vision in Plant Disease Detection: A Comprehensive Review of Techniques, Models, and Trends in Precision Agriculture. *Artif. Intell. Rev.* **2025**, *58*, 92.
20. Batool, A.; Byun, Y.C. Revolutionizing Plant Disease Diagnosis through Vision-Based Intelligence and Next-Generation Computing. *Comput. Electr. Eng.* **2025**, *128*, 110695.
21. Ren, Q.; Wu, Y.; Zeng, Q.; Yang, N. A Review of Deep Learning Based Agricultural Remote Sensing Image Segmentation. *J. Agric. Eng.* **2026**, *57*. https://doi.org/10.4081/jae.2025.1954.
22. Liu, W.; Wang, Z.; Duan, P.; Kang, X.; Li, S. Agriculture-Vision Challenge 2024—The Runner-Up Solution for Agricultural Pattern Recognition via Class Balancing and Model Ensemble. *arXiv* **2024**, arXiv:2406.12271.
23. Wang, B.; Li, C.; Zou, W.; Zhang, Y.; Chen, X.; Chen, C.L.P. A Comprehensive Survey on Shadow Removal from Document Images: Datasets, Methods, and Opportunities. *Vicinagearth* **2025**, *2*, 1.
24. Jimenez-Gutierrez, D.M.; Falkouskaya, Y.; Hernandez-Ramos, J.L.; Anagnostopoulos, A.; Chatzigiannakis, I.; Vitaletti, A. On the Security and Privacy of Federated Learning: A Survey with Attacks, Defenses, Frameworks, Applications, and Future Directions. *arXiv* **2025**, arXiv:2508.13730.
25. Dwork, C. Differential Privacy. In Proceedings of the Automata, Languages and Programming (ICALP), Venice, Italy, 10–14 July 2006.
26. Basudan, S. A Privacy-Preserving Federated Learning Protocol with a Secure Data Aggregation for the Internet of Everything. *Comput. Commun.* **2024**, *223*, 1–14.
27. Jin, X.; Yao, Y.; Yu, N. Efficient Secure Aggregation for Privacy-Preserving Federated Learning Based on Secret Sharing. *J. Univ. Sci. Technol. China* **2024**, *54*, 104.
28. Chouhan, A.; Purushothama, B.R. A Survey on Secure Aggregation for Privacy-Preserving Federated Learning. In *Advancements in Smart Computing and Information Security*; Springer: Berlin/Heidelberg, Germany, 2024.
29. Fu, J.; Hong, Y.; Ling, X.; Wang, L.; Ran, X.; Sun, Z.; Wang, W.H.; Chen, Z.; Cao, Y. Differentially Private Federated Learning: A Systematic Review. *arXiv* **2024**, arXiv:2405.08299.
30. Shan, F.; Liu, Y.; Fan, L.; Mao, Y.; Chen, Z.; Li, S. A Survey of Optimization Algorithms for Differential Privacy in Federated Learning. *J. Syst. Archit.* **2025**, *168*, 103582.
31. Cui, L.; Wu, X. ALDP-FL for Adaptive Local Differential Privacy in Federated Learning. *Sci. Rep.* **2025**, *15*, 26679.
32. Ling, J.; Zheng, J.; Chen, J. Efficient Federated Learning Privacy Preservation Method with Heterogeneous Differential Privacy. *Comput. Secur.* **2024**, *139*, 103715.
33. Bai, R.; Bagchi, S.; Inouye, D.I. Benchmarking Algorithms for Federated Domain Generalization. In Proceedings of the International Conference on Learning Representations, Vienna, Austria, 7–11 May 2024.
34. Le, K.; Ho, L.; Do, C.; Le-Phuoc, D.; Wong, K.S. Efficiently Assemble Normalization Layers and Regularization for Federated Domain Generalization. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 16–22 June 2024.

35. Yan, H.; Guo, Y. Local and Global Flatness for Federated Domain Generalization. In Proceedings of the European Conference on Computer Vision, Milan, Italy, 29 September–4 October 2024.
36. Gong, S.; Cui, C.; Zhang, C.; Wang, W.; Nie, X.; Zhu, L. Federated Domain Generalization via Prompt Learning and Aggregation. *arXiv* **2024**, arXiv:2411.10063.
37. Wu, J.; Qu, X.; Huang, Z.; Wang, J. Federated Domain Generalization with Domain-specific Soft Prompts Generation. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Honolulu, HI, USA, 19–20 October 2025.
38. Liao, T.; Xie, B.; Fu, L.; Huang, S.; Deng, B.; Chen, C.; Zheng, Z. Federated Domain Generalization with Decision Insight Matrix. In Proceedings of the International Joint Conference on Artificial Intelligence, Montréal, QC, Canada, 16–22 August 2025.
39. Li, Y.; Wang, X.; Zeng, R.; Donta, P.K.; Murturi, I.; Huang, M.; Dustdar, S. Federated Domain Generalization: A Survey. *Proc. IEEE* **2025**, *113*, 370–410.
40. Barbedo, J.G.A. A Review on the Main Challenges in Automatic Plant Disease Identification Based on Visible Range Images. *Biosyst. Eng.* **2016**, *144*, 52–60. https://doi.org/10.1016/j.biosystemseng.2016.01.017.
41. Singh, V.; Sharma, N.; Singh, S. A Review of Imaging Techniques for Plant Disease Detection. *Artif. Intell. Agric.* **2020**, *4*, 229–242. https://doi.org/10.1016/j.aiia.2020.10.002.
42. Tian, H.; Wang, T.; Liu, Y.; Qiao, X.; Li, Y. Computer Vision Technology in Agricultural Automation—A Review. *Inf. Process. Agric.* **2020**, *7*, 1–19. https://doi.org/10.1016/j.inpa.2019.09.006.
43. Shin, J.; Mahmud, M.S.; Rehman, T.U.; Ravichandran, P.; Heung, B.; Chang, Y.K. Trends and Prospect of Machine Vision Technology for Stresses and Diseases Detection in Precision Agriculture. *AgriEngineering* **2023**, *5*, 20–39. https://doi.org/10.3390/agriengineering5010003.
44. Li, T.; Sahu, A.K.; Zaheer, M.; Sanjabi, M.; Talwalkar, A.; Smith, V. Federated Optimization in Heterogeneous Networks. In Proceedings of the Machine Learning and Systems, Austin, TX, USA, 2–4 March 2020.
45. Li, X.; Jiang, M.; Zhang, X.; Kamp, M.; Dou, Q. FedBN: Federated Learning on Non-IID Features via Local Batch Normalization. In Proceedings of the International Conference on Learning Representations, Virtual, 3–7 May 2021.
46. Li, Q.; He, B.; Song, D. Model-Contrastive Federated Learning. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021.