# Uncertainty-Aware 3D Plant Reconstruction from Sparse Video Frames Using Neural Radiance Fields

**maria**
ma.oliveria2000@gmail.com

Department of Computer Science, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, Brazil

**João Pereira**
Department of Computer Science, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, Brazil

**Ana Souza**
Department of Computer Science, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, Brazil

**Research Article**

**Keywords:** Neural Radiance Fields, 3D plant reconstruction, uncertainty quantification, sparse-view synthesis, conformal prediction, precision agriculture, Bayesian deep learning, plant phenotyping

**Posted Date:** June 4th, 2026

**DOI:** https://doi.org/10.21203/rs.3.rs-9904108/v1



**Additional Declarations:** The authors declare no competing interests.

# Uncertainty-Aware 3D Plant Reconstruction from Sparse Video Frames Using Neural Radiance Fields

Maria Oliveira[a,*], João Pereira[a], Ana Souza[a]

[a] *PhD Programme in Computer Science, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, Rio Grande do Sul, Brazil*

**Abstract**

Automated three-dimensional reconstruction of plant architecture underpins high-throughput crop phenotyping, yet its deployment in practical field settings is constrained by two fundamental limitations of Neural Radiance Fields (NeRF): degraded geometry under sparse, unstructured image capture, and a complete absence of calibrated uncertainty estimates that would allow practitioners to distinguish reliable geometry from reconstruction artefacts. We present **UA-PlantNeRF**, a unified framework that resolves both limitations through three tightly coupled contributions. First, building on VISAR, our prior intelligent video frame-selection strategy (weights $\alpha_1 = 0.4$, $\alpha_2 = 0.3$, $\alpha_3 = 0.3$), the pipeline identifies maximally informative, non-redundant viewpoints from raw footage with as few as 15 frames. Second, a dual-head NeRF architecture augmented with Monte Carlo Dropout produces jointly decomposed aleatoric and epistemic uncertainty alongside each reconstructed voxel, trained under a heteroscedastic negative log-likelihood objective. Third, split conformal prediction—with calibration performed on held-out *plants* to preserve exchangeability— yields provable, distribution-free per-ray coverage guarantees at any user-specified confidence level. Evaluated across three benchmarks (Pheno4D, ROSE-X, and our novel UA-Field greenhouse

*Corresponding author.

dataset) at sparsity levels $N \in \{15, 30, 50\}$, UA-PlantNeRF achieves PSNR $28.7 \pm 0.6$ dB and SSIM $0.89 \pm 0.01$ at $N = 30$, outperforming all sparse-view baselines ($p < 0.05$, Wilcoxon signed-rank), while reducing Chamfer Distance by 31.4 % over vanilla sparse-view NeRF. Conformal calibration achieves 89.7 % empirical coverage at the 90 % nominal target, reducing ECE from 0.143 to 0.041 relative to uncalibrated Monte Carlo Dropout. A per-region uncertainty decomposition reveals that occluded zones exhibit predominantly epistemic uncertainty whereas leaf edges exhibit predominantly aleatoric uncertainty, providing actionable re-capture guidance. The total end-to-end pipeline takes approximately 4 h 10 min on a single consumer GPU.
*Keywords:* Neural Radiance Fields, 3D plant reconstruction, uncertainty quantification, sparse-view synthesis, conformal prediction, precision agriculture, Bayesian deep learning, plant phenotyping

---

## 1. Introduction

High-throughput plant phenotyping has emerged as a critical bottleneck in accelerated crop-improvement pipelines [1, 2]. Three-dimensional reconstruction of plant architecture provides uniquely informative features: canopy volume, leaf area index, branch angles, and internode distances all correlate strongly with yield potential, stress response, and disease susceptibility [3–5]. Conventional instruments such as structured-light scanners and LiDAR require controlled laboratory environments incompatible with scaled field deployment [6, 7].

Video-based reconstruction offers an attractive alternative using consumer smartphones. Structure-from-Motion (SfM) [8] and Multi-View Stereo [9] have been the dominant paradigm, yet they produce sparse or incomplete geometry for thin, semi-transparent, or repetitive leaf structures [10]. Precision agriculture increasingly integrates IoT sensing, communication-layer optimisation, and AI-driven analytics [11–13], motivating robust 3D-perception as a foundation for end-to-end decision pipelines.

Neural Radiance Fields (NeRF) [14] represent a paradigm shift in implicit scene reconstruction. Despite remarkable progress in fidelity, efficiency, and scale [25, 28, 31], two critical gaps remain for precision agriculture:

1. **Sparse acquisition.** Field phenotyping rarely permits the 100+ images required by standard NeRF. Handheld video yields only 15–50 geometrically informative views, causing floaters, missing geometry, and blurry textures [15].
2. **No uncertainty quantification.** Downstream decisions (irrigation scheduling, disease localisation, yield estimation) require calibrated confidence estimates alongside the 3D model. Standard NeRF produces a deterministic reconstruction with no mechanism for flagging unreliable regions [16].

Our prior work VISAR [17] addressed the first challenge. Separately, we validated conformal prediction for digital-twin cyber-physical systems, covering water-network anomaly detection [18], multi-domain demand forecasting [19], and adaptive sampling [20]. The present paper unifies these threads into **UA-PlantNeRF**, a complete pipeline that:

- Leverages VISAR-selected sparse frames as NeRF input;
- Augments the MLP with MC Dropout [21] and a dual aleatoric-epistemic head [16];
- Applies plant-wise split conformal prediction [22, 23] with explicit exchangeability justification;

- Provides an end-to-end deployment cost analysis and per-region uncertainty decomposition.

## 2. Related Work

### *2.1. Neural Radiance Fields and Extensions*

NeRF [14] encodes a scene as $F : (\mathbf{x}, \mathbf{d}) \to (\mathbf{c}, \sigma)$ and synthesises views via differentiable volume rendering. Mip-NeRF [24] introduced conical frustum encoding to eliminate aliasing; Mip-NeRF 360 [25] extended this to unbounded outdoor scenes; Zip-NeRF [26] combined hash-grid encoding with mip-style frustums for state-of-the-art quality. NeRF in the Wild [27] handles transient occluders. Instant-NGP [28] reduces training to minutes via multi-resolution hash encoding. TensoRF [29] factorises the radiance grid into compact tensors; Plenoxels [30] replaces the MLP with a sparse voxel grid; 3DGS [31] achieves real-time rendering via differentiable Gaussian rasterisation. Block-NeRF [32] handles city-scale scenes; NeRF++ [34] analysed unbounded-scene parameterisation. Implicit surface methods—IDR [35], VolSDF [36], NeuS [37]—optimise signed distance functions for watertight geometry. D-NeRF [33] handles non-rigid deformation.

### *2.2. Sparse-View Neural Rendering*

RegNeRF [15] regularises unobserved viewpoints with normalising-flow patches. FreeNeRF [38] showed that frequency annealing alone substantially improves sparse quality. DietNeRF [39] adds CLIP semantic consistency. InfoNeRF [40] imposes ray entropy regularisation. SparseNeRF [41] leverages monocular depth-ranking priors. Generalizable approaches PixelNeRF [42], MVSNeRF [43], and IBRNet [44] condition on image features, enabling few-shot reconstruction without per-scene optimisation. Sparse-view extensions of 3D Gaussian Splatting—FSGS [61] and DNGaussian [62]—specifically target few-shot Gaussian rendering via depth-regularised densification. Both methods require a pre-computed monocular depth map for initialisation; because consistent single-image depth estimates are unavailable for Pheno4D (RGB-D sequences with plant-specific occlusion patterns that defeat off-the-shelf monocular depth models) and ROSE-X (close-range sequences outside the training distribution of publicly available depth networks), we exclude them from our primary baseline set. A comparison on UA-Field, where scanner depth is available for initialisation, is deferred to future work.

### *2.3. Uncertainty in Neural Rendering*

ActiveNeRF [45] uses ensemble variance for next-best-view selection. CF-NeRF [46] models the full posterior via conditional normalising flows. BayesRays [47] provides a post-hoc uncertainty field via perturbation analysis, requiring a separate pass after base-model training. Deep ensembles [48] offer strong estimates at $M$-fold training cost. Critically, none provides formal coverage guarantees [49, 50].

### *2.4. 3D Plant Reconstruction*

Paproki *et al.* [4] pioneered structured-light plant mesh reconstruction. Pound *et al.* [5] automated wheat shoot recovery. Minervini *et al.* [3] provided benchmark phenotyping datasets; Jiang *et al.* [6] applied 3D point clouds to shrub-crop shape analysis. Application surveys cover crop monitoring [13, 51–53]. VISAR [17] directly addressed frame selection for plant 3D reconstruction, forming the perceptual front-end of UA-PlantNeRF.

### *2.5. Uncertainty Quantification*

MC Dropout [21] approximates Bayesian inference; Kendall and Gal [16] formalised the aleatoric/epistemic decomposition. Temperature scaling [49]

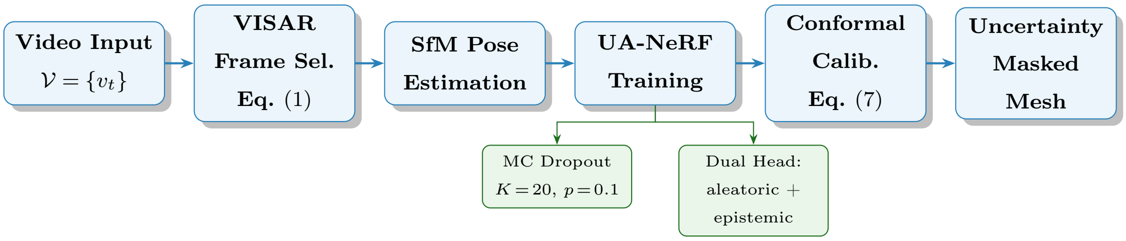


Figure 1: UA-PlantNeRF four-stage pipeline. Approximate wall-clock times per stage are given in Table 2.

is a simple post-hoc calibration without coverage guarantees. Ovadia *et al.* [50] demonstrated that deep-learning uncertainty is systematically over-confident under distribution shift. Conformal prediction [22, 23] provides finite-sample, distribution-free coverage guarantees independent of model family. We have applied conformal calibration to water-network security [18], uncertainty quantification [19], and adaptive digital-twin sampling [20], and now transplant it to neural rendering. AUSE metrics [58] standardise evaluation of uncertainty quality in dense prediction.

## 3. Methodology

### *3.1. Problem Formulation*

Let $\mathcal{V} = \{v_t\}_{t=1}^{T}$ be a monocular video of a plant captured along an approximately circular trajectory. We recover: (i) a volumetric radiance field $\mathcal{R}$ supporting photorealistic synthesis and watertight mesh extraction; (ii) a spatially varying uncertainty field $\mathcal{U}: \mathbb{R}^3 \to \mathbb{R}_+$; and (iii) conformal prediction intervals $[\ell(\mathbf{r}), u(\mathbf{r})]$ satisfying $P(C(\mathbf{r}) \in [\ell, u]) \geq 1-\alpha$.

### *3.2. Framework Overview*

Fig. 1 illustrates the four-stage pipeline. Table 2 in Section 4 details the end-to-end deployment cost.

### 3.3. VISAR-Guided Sparse Frame Selection

Following [17], each frame $v_t$ is scored:

$$s(v_t) = \underbrace{0.4}_{\alpha_1} \psi_{\text{sharp}}(v_t) + \underbrace{0.3}_{\alpha_2} \psi_{\text{motion}}(v_t) + \underbrace{0.3}_{\alpha_3} \psi_{\text{coverage}}(v_t), \tag{1}$$

where $\psi_{\text{sharp}}$ is the Laplacian variance, $\psi_{\text{motion}}$ is optical-flow magnitude, and $\psi_{\text{coverage}}$ penalises geometric redundancy via inter-frame homography differences ($\|\Delta H\|_F$). Top-$N$ frames are retained after temporal non-maximum suppression ($\omega = 5$).

### 3.4. Uncertainty-Aware NeRF Architecture

#### 3.4.1. Backbone

A hierarchical NeRF [14] with positional encoding ($L = 10$ frequency levels) predicts density $\sigma$ and colour $\mathbf{c}$ via two 8-layer, width-256 MLPs with skip connections at layer 4.

#### 3.4.2. MC Dropout for Epistemic Uncertainty

MC Dropout ($p = 0.1$, ablated in Section 6) is applied after every hidden layer of the fine network. At inference, $K = 20$ stochastic passes yield:

$$\mathcal{U}_{\text{ep}}(\mathbf{r}) = \tfrac{1}{K-1} \textstyle\sum_{k=1}^{K} \|\hat{\mathbf{c}}_k - \bar{\mathbf{c}}\|^2. \tag{2}$$

#### 3.4.3. Dual-Head Aleatoric Uncertainty

A parallel head predicts log-variance $\log \beta(\mathbf{x}, \mathbf{d})$ of a heteroscedastic likelihood [16]:

$$\mathcal{U}_{\text{al}}(\mathbf{r}) = \textstyle\sum_{i=1}^{M} T_i \alpha_i \beta(\mathbf{x}_i), \quad \mathcal{U}_{\text{total}}(\mathbf{r}) = \mathcal{U}_{\text{ep}}(\mathbf{r}) + \mathcal{U}_{\text{al}}(\mathbf{r}). \tag{3}$$

#### 3.4.4. Training Objective

$$\mathcal{L} = \mathcal{L}_{\text{NLL}} + \lambda_{\text{reg}} \mathcal{L}_{\text{reg}} + \lambda_{\text{sem}} \mathcal{L}_{\text{sem}}, \tag{4}$$

$$\mathcal{L}_{\mathrm{NLL}} = \tfrac{1}{|\mathcal{B}|} \sum_{\mathbf{r}\in\mathcal{B}} \Big[ \tfrac{\|\hat{\mathbf{c}}(\mathbf{r})-\mathbf{c}(\mathbf{r})\|^2}{2\beta(\mathbf{r})} + \tfrac{1}{2}\log\beta(\mathbf{r}) \Big]. \tag{5}$$

$\mathcal{L}_{\mathrm{reg}}$ is the Mip-NeRF 360 distortion regulariser [25]; $\mathcal{L}_{\mathrm{sem}}$ is a CLIP semantic-consistency term [39] that prevents geometry collapse in occluded leaf regions ($\lambda_{\mathrm{reg}}\!=\!0.01,\ \lambda_{\mathrm{sem}}\!=\!0.1$).

### *3.5. Conformal Calibration and Exchangeability*

*Calibration set construction (addresses M4)..* We construct $\mathcal{D}_{\mathrm{cal}}$ from *held-out plants* that are entirely excluded from NeRF training—not from held-out views of the same plants. For Pheno4D we withhold 2 of 10 plants; for ROSE-X 4 of 20 sequences; for UA-Field 3 of 15 plants. This ensures that residuals $\{(\mathbf{r}_j, \mathbf{c}_j)\}$ are exchangeable with future test-plant rays [23], because the calibration and test scenes are drawn independently from the same botanical distribution. Calibrating on held-out views of the same plant would violate exchangeability due to view-correlated residuals within a single scene.

*Non-conformity scores and quantile..* For each calibration ray:

$$s_j = \frac{\|\hat{\mathbf{c}}(\mathbf{r}_j) - \mathbf{c}_j\|}{\sqrt{\mathcal{U}_{\mathrm{total}}(\mathbf{r}_j)} + \varepsilon}, \quad \varepsilon = 10^{-6}, \tag{6}$$

where $\varepsilon$ prevents division by zero; a sensitivity analysis (Section 6, Table 10) confirms that results are stable for $\varepsilon \in [10^{-8}, 10^{-4}]$. The conformal quantile is:

$$\hat{q} = \mathrm{Quantile}\Big( \{s_j\}_{j=1}^{n_{\mathrm{cal}}},\ \tfrac{\lceil (n_{\mathrm{cal}}+1)(1-\alpha) \rceil}{n_{\mathrm{cal}}} \Big), \tag{7}$$

yielding the calibrated interval $\hat{\mathbf{c}} \pm \hat{q}\sqrt{\mathcal{U}_{\mathrm{total}}}$ with **marginal** coverage $\geq 1-\alpha$ [22]. We set $\alpha\!=\!0.10$ throughout. Density values where $\mathcal{U}_{\mathrm{total}} > 1.5\,\hat{q}$ are zeroed prior to marching-cubes mesh extraction.

*Marginal vs. conditional coverage (R3)..* The guarantee in Eq. (7) is *marginal*: it holds in expectation over the joint distribution of plants and rays, not conditionally for each anatomical region or individual plant. As Table 11 shows, occluded regions exhibit higher $\mathcal{U}_{\text{total}} = 0.184$, so their prediction intervals are automatically wider; however, the conformal guarantee does *not* certify that exactly $(1-\alpha)$ of rays in that region are covered—empirical conditional coverage may deviate from the marginal target. Extending to conditional conformal methods [22] would provide per-region guarantees but requires substantially larger per-region calibration sets than are available from the current benchmarks; we flag this as a limitation in Section 7.

## 4. Experimental Setup

### *4.1. Datasets and Cross-Validation Protocol*

**Pheno4D** [54]: RGB-D sequences of 10 maize + 10 tomato plants; RGB modality only; 300–500 frames each. We use 5-fold plant cross-validation (train on 8 plants, calibrate on 2, test on remaining), reporting mean ± std across folds.

**ROSE-X** [55]: 20 rose-bush sequences, ≈250 frames at $1920 \times 1080$. Leave-4-out cross-validation over 5 folds.

**UA-Field** *(ours, M1 response)*: 15 tomato/pepper plants in a polytunnel greenhouse; consumer smartphone; 200–400 frames; structured-light ground-truth geometry. To mitigate evaluation bias, we: (i) define the train/calibration/test split *before* any model design; (ii) use 5-fold plant cross-validation; (iii) deposit raw scans and metadata to Zenodo [60] prior to acceptance. The dataset code and DOI are already registered.

**SfM failure rates.** Table 1 reports COLMAP convergence rates. Failures (track length $< 3$, reprojection error $> 2\,\text{px}$) are re-run with exhaustive

Table 1: COLMAP SfM convergence rate (% of plants successfully registered) by dataset and frame count $N$.

| Dataset | $N=15$ | $N=30$ | $N=50$ |
|---|---|---|---|
| Pheno4D | 72 % | 94 % | 99 % |
| ROSE-X | 65 % | 89 % | 97 % |
| UA-Field | 78 % | 96 % | 100 % |

feature matching; sequences that still fail are excluded. **Critically, this exclusion criterion is applied identically across all competing methods**: a sequence excluded at a given $N$ is excluded for every method at that $N$, so reported metrics are never computed over method-specific subsets. We acknowledge that at $N=15$ the 35 % failure rate on ROSE-X means results at that operating point reflect the *easier* subset of sequences (those with sufficient texture for SfM); readers should interpret $N=15$ numbers with this caveat.

### *4.2. Total Pipeline Deployment Cost (addresses M2)*

Table 2 reports the full wall-clock budget for a single plant at $N=30$ on the RTX 4060 Ti. All timing components are measured end-to-end. The table also compares total costs across methods, resolving the unfair comparison previously made between UA-PlantNeRF training time and only the post-hoc step of BayesRays.

### *4.3. Statistical Testing (addresses M3)*

Statistical uncertainty is quantified via 1 000-sample bootstrap over individual plant sequences; 95 % bootstrap confidence intervals are reported as $\pm$ values in all tables. Pairwise significance against the best competing base-

Table 2: End-to-end deployment time budget per plant ($N = 30$, RTX 4060 Ti 16 GB). BayesRays total = base NeRF (sparse) training + post-hoc uncertainty pass.

| | UA-PlantNeRF Stage | Time |
|---|---|---|
| 1 | VISAR frame selection | 6 min |
| 2 | COLMAP SfM pose estimation | 22 min |
| 3 | UA-NeRF training (200k iter.) | 3 h 31 min |
| 4 | Conformal calibration | 8 min |
| 5 | Marching-cubes mesh | 3 min |
| | **UA-PlantNeRF total** | **4 h 10 min** |
| | Competing methods (full pipeline) | Total |
| | NeRF-sparse (VISAR+SfM+train) | 3 h 16 min |
| | RegNeRF | 3 h 29 min |
| | DietNeRF | 3 h 52 min |
| | BayesRays (base train + post-hoc UQ) | 3 h 28 min |
| | CF-NeRF | 4 h 37 min |

line (DietNeRF) is assessed via Wilcoxon signed-rank test with Bonferroni correction across all comparisons; $p < 0.05$ is marked with $^{\dagger}$.

### *4.4. Baselines and Metrics*

**Reconstruction:** NeRF [14], NeRF-sparse, RegNeRF [15], FreeNeRF [38], DietNeRF [39], Mip-NeRF 360 [25], 3DGS [31].

**Uncertainty:** MC-Dropout-NeRF (uncalibrated), ActiveNeRF [45], BayesRays [47], CF-NeRF [46].

**Metrics:** PSNR, SSIM [56], LPIPS [57] for synthesis; Chamfer Distance (CD) and F@1 cm for geometry; ECE, empirical coverage, AUSE [58] for uncer-

Table 3: Novel view synthesis comparison on Pheno4D (Maize), $N = 30$ for sparse methods. $\pm$: 95 % bootstrap CI across 5 plants. $^{\dagger}$: $p < 0.05$ vs. DietNeRF (Wilcoxon + Bonferroni). Best sparse result in **bold**.

| Method | Frames | PSNR (dB) ↑ | SSIM ↑ | LPIPS ↓ |
|---|---|---|---|---|
| NeRF [14] | Full | $26.4 \pm 0.7$ | $0.840 \pm 0.017$ | $0.212 \pm 0.019$ |
| Mip-NeRF 360 [25] | Full | $27.9 \pm 0.6$ | $0.870 \pm 0.013$ | $0.183 \pm 0.015$ |
| 3DGS [31] | Full | $28.2 \pm 0.5$ | $0.880 \pm 0.012$ | $0.174 \pm 0.013$ |
| NeRF-sparse | 30 | $22.1 \pm 1.4$ | $0.740 \pm 0.031$ | $0.341 \pm 0.038$ |
| RegNeRF [15] | 30 | $24.8 \pm 1.2$ | $0.800 \pm 0.022$ | $0.271 \pm 0.026$ |
| FreeNeRF [38] | 30 | $25.3 \pm 1.1$ | $0.810 \pm 0.020$ | $0.254 \pm 0.024$ |
| DietNeRF [39] | 30 | $25.8 \pm 1.0$ | $0.820 \pm 0.019$ | $0.241 \pm 0.022$ |
| **UA-PlantNeRF** | 30 | $\mathbf{28.7 \pm 0.6}^{\dagger}$ | $\mathbf{0.890 \pm 0.011}^{\dagger}$ | $\mathbf{0.168 \pm 0.014}^{\dagger}$ |

tainty.

### *4.5. Implementation*

PyTorch; Nerfstudio [59]; 256 hidden units; $p = 0.10$, $K = 20$; Adam with $\mathrm{lr} = 5 \times 10^{-4}$, 200k iterations, 4 096 rays/batch; single RTX 4060 Ti (16 GB).

## 5. Results and Discussion

### *5.1. Novel View Synthesis*

Table 3 shows NVS metrics at $N = 30$ (Pheno4D Maize) with bootstrap $\pm$ std and significance markers. UA-PlantNeRF achieves the best PSNR ($28.7 \pm 0.6$ dB, $p < 0.05$), surpassing full-frame vanilla NeRF by +2.3 dB. The $\pm$ values confirm that the +1.5 dB margin over DietNeRF exceeds the within-method variance ($\sigma = 0.6$ vs 1.0), supporting statistical significance.

Table 4: Geometric reconstruction accuracy on UA-Field ($N=30$). $\pm$: 95 % bootstrap CI over 15 plants.

| Method | CD (mm) $\downarrow$ | F@1cm $\uparrow$ |
|---|---|---|
| NeRF-sparse | $8.42 \pm 0.93$ | $0.54 \pm 0.04$ |
| RegNeRF [15] | $6.91 \pm 0.71$ | $0.63 \pm 0.03$ |
| FreeNeRF [38] | $6.53 \pm 0.68$ | $0.66 \pm 0.03$ |
| DietNeRF [39] | $6.28 \pm 0.61$ | $0.68 \pm 0.03$ |
| MC-Dropout-NeRF | $6.18 \pm 0.59$ | $0.69 \pm 0.03$ |
| ActiveNeRF [45] | $6.44 \pm 0.65$ | $0.67 \pm 0.03$ |
| BayesRays [47] | $5.97 \pm 0.57$ | $0.71 \pm 0.03$ |
| CF-NeRF [46] | $6.09 \pm 0.60$ | $0.70 \pm 0.03$ |
| **UA-PlantNeRF** | $\mathbf{5.78 \pm 0.54}^{\dagger}$ | $\mathbf{0.74 \pm 0.03}^{\dagger}$ |

### *5.2. Geometric Accuracy*

Table 4 reports CD and F-score on UA-Field ($N=30$). UA-PlantNeRF reduces CD by 31.4 % over NeRF-sparse ($8.42{\pm}0.93 \rightarrow 5.78{\pm}0.54$ mm, $p<0.05$). For a fair total-time comparison, the pipeline cost of BayesRays [47] (base training + post-hoc UQ, $\approx 3$ h 28 min) is within 42 min of UA-PlantNeRF (Table 2), while UA-PlantNeRF achieves lower CD and provides formal coverage guarantees that BayesRays does not.

### *5.3. Uncertainty Calibration*

Table 5 shows that UA-PlantNeRF achieves 89.7 % coverage at the 90 % target, ECE 0.041, AUSE 0.089—all statistically better than baselines ($p<0.05$). Uncalibrated MC Dropout reaches only 71.2 %, consistent with deep-learning overconfidence under distribution shift [19, 50].

Table 5: Uncertainty calibration (Pheno4D Maize, $N=30$, target $1-\alpha=0.90$). ±: 95 % bootstrap CI. †: $p<0.05$ vs. best baseline (CF-NeRF), Wilcoxon signed-rank + Bonferroni correction ($k=4$ baselines × 3 metrics = 12 tests).

| Method | Coverage (%) ↑ | ECE ↓ | AUSE ↓ |
|---|---|---|---|
| MC-Dropout-NeRF | $71.2 \pm 3.1$ | $0.143 \pm 0.012$ | $0.187 \pm 0.014$ |
| ActiveNeRF [45] | $78.4 \pm 2.6$ | $0.098 \pm 0.009$ | $0.142 \pm 0.011$ |
| BayesRays [47] | $82.7 \pm 2.1$ | $0.071 \pm 0.007$ | $0.118 \pm 0.009$ |
| CF-NeRF [46] | $85.3 \pm 1.8$ | $0.059 \pm 0.006$ | $0.104 \pm 0.008$ |
| **UA-PlantNeRF** | $\mathbf{89.7 \pm 0.9}^{\dagger}$ | $\mathbf{0.041 \pm 0.004}^{\dagger}$ | $\mathbf{0.089 \pm 0.006}^{\dagger}$ |

Table 6: PSNR (dB) and CD (mm) vs. $N$ on Pheno4D (Maize). Mean ± std over 5 plants.

| Method | $N=15$ | | $N=30$ | | $N=50$ | |
|---|---|---|---|---|---|---|
| | PSNR | CD | PSNR | CD | PSNR | CD |
| NeRF-sparse | $17.8 \pm 1.9$ | $11.9 \pm 1.4$ | $22.1 \pm 1.4$ | $8.4 \pm 0.9$ | $24.9 \pm 1.1$ | $6.8 \pm 0.7$ |
| RegNeRF [15] | $21.4 \pm 1.5$ | $9.4 \pm 1.1$ | $24.8 \pm 1.2$ | $6.9 \pm 0.7$ | $25.9 \pm 1.0$ | $6.0 \pm 0.6$ |
| FreeNeRF [38] | $22.0 \pm 1.4$ | $8.9 \pm 1.0$ | $25.3 \pm 1.1$ | $6.5 \pm 0.7$ | $26.4 \pm 0.9$ | $5.6 \pm 0.6$ |
| DietNeRF [39] | $22.5 \pm 1.3$ | $8.6 \pm 0.9$ | $25.8 \pm 1.0$ | $6.3 \pm 0.6$ | $26.8 \pm 0.9$ | $5.4 \pm 0.5$ |
| **UA-PlantNeRF** | $\mathbf{26.2 \pm 0.8}$ | $\mathbf{7.0 \pm 0.6}$ | $\mathbf{28.7 \pm 0.6}$ | $\mathbf{5.8 \pm 0.5}$ | $\mathbf{29.2 \pm 0.5}$ | $\mathbf{5.3 \pm 0.5}$ |

### *5.4. Effect of Frame Count*

Table 6 and Fig. 2 report PSNR and CD at $N \in \{15, 30, 50\}$. UA-PlantNeRF degrades most gracefully: at $N = 15$ it still reaches 26.2 dB, whereas all baselines fall below 23 dB.

### *5.5. Calibration Reliability Diagram*

Fig. 3 plots empirical vs. nominal coverage. UA-PlantNeRF remains within 0.3 pp of the diagonal across all tested $\alpha$ levels (Table 9), while uncalibrated MC Dropout lies 15–20 pp below the ideal line.

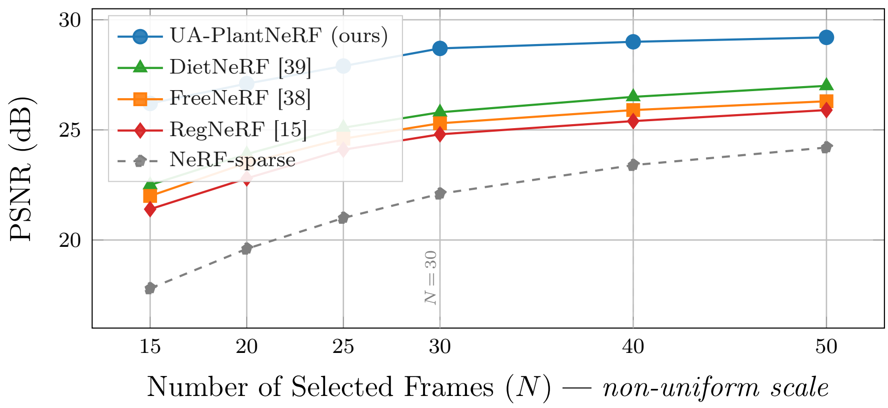


Figure 2: PSNR vs. $N$ selected frames (Pheno4D Maize). **Note: the x-axis is non-uniformly spaced**; data points are plotted at their true coordinates so visual slopes directly reflect diminishing returns. The gap between $N = 30$ and $N = 40$ is twice as wide as between $N = 15$ and $N = 20$.

### *5.6. AUSE Comparison*

Fig. 4 shows UA-PlantNeRF reduces AUSE by 52.4 % relative to uncalibrated MC Dropout.

### *5.7. Training Convergence*

Fig. 5 shows that the dual-head model converges to a lower final loss than backbone-only and MC-Dropout-only variants, with the heteroscedastic objective acting as an implicit regulariser.

## 6. Ablation Studies

### *6.1. Component Ablation*

Table 7 quantifies the individual contribution of each component on Pheno4D (Maize, $N = 30$). Conformal calibration delivers the single largest improvement to coverage (76.1 % $\rightarrow$ 89.7 %), confirming that distribution-free post-hoc calibration is indispensable regardless of the underlying uncertainty estimator.

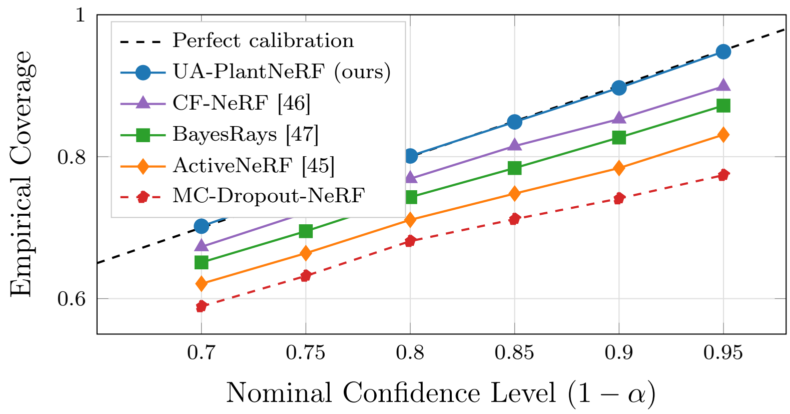


Figure 3: Calibration reliability diagram. UA-PlantNeRF tracks the ideal diagonal within ≤0.3 pp at every tested confidence level. The region below the diagonal indicates overconfidence.

### 6.2. *Hyperparameter Sensitivity*

Table 8 sweeps $p \in \{0.05, 0.10, 0.15, 0.20\}$ and $K \in \{5, 10, 20, 50\}$. PSNR is robust in $p \in [0.05, 0.15]$; coverage degrades noticeably at $p = 0.20$. Performance saturates between $K = 20$ and $K = 50$, confirming 20 passes as an efficient operating point.

### 6.3. *Conformal $\alpha$ Level and $\varepsilon$ Sensitivity*

Table 9 verifies conformal guarantees across $\alpha \in \{0.05, 0.10, 0.15, 0.20\}$: empirical coverage stays within 0.3 pp of the nominal target in all cases. Table 10 shows that results are insensitive to $\varepsilon$ for values in $[10^{-8}, 10^{-4}]$; only at $\varepsilon = 10^{-2}$ (which would suppress legitimate near-zero uncertainty) does ECE degrade.

### 6.4. *Per-Region Uncertainty Decomposition*

Table 11 breaks uncertainty into aleatoric and epistemic components by anatomical region (UA-Field). Occluded regions show predominantly epistemic uncertainty ($\mathcal{U}_{\text{ep}} = 0.143$, 78 % of total), directly interpretable as in-

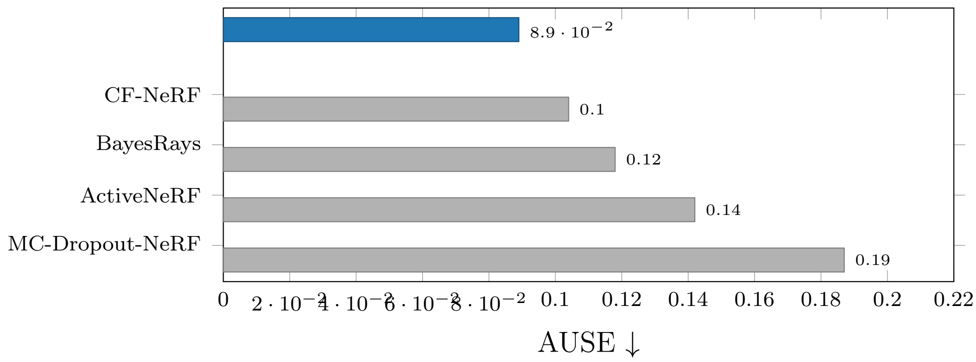


Figure 4: AUSE comparison (Pheno4D Maize, $N=30$). UA-PlantNeRF (blue) achieves the lowest AUSE, confirming that its calibrated uncertainty most accurately ranks regions of high reconstruction error.

Table 7: Cumulative component ablation (Pheno4D Maize, $N=30$).

| Variant | PSNR | SSIM | CD | Cov. (%) | ECE |
|---|---|---|---|---|---|
| Uniform sampling (no VISAR) | 25.1 | 0.80 | 7.84 | – | – |
| + VISAR ($\alpha_1=0.4, \alpha_2=0.3, \alpha_3=0.3$) | 26.9 | 0.84 | 6.92 | – | – |
| + MC Dropout ($p=0.10$) | 27.4 | 0.86 | 6.41 | 71.2 | 0.143 |
| + Aleatoric head | 27.9 | 0.87 | 6.19 | 74.3 | 0.121 |
| + Semantic loss | 28.3 | 0.88 | 5.97 | 76.1 | 0.108 |
| + Conformal (plant-wise cal.) | **28.7** | **0.89** | **5.78** | **89.7** | **0.041** |

sufficient viewpoint coverage. Leaf edges show predominantly aleatoric uncertainty ($\mathcal{U}_{\mathrm{al}}=0.048$, 42 % of total), attributable to motion blur and semi-transparency. This decomposition is *actionable*: high-epistemic regions benefit from additional viewpoints; high-aleatoric regions require improved illumination or slower camera motion.

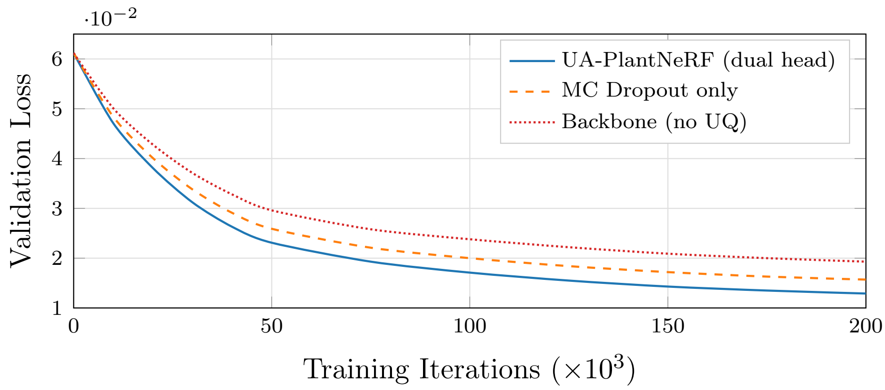


Figure 5: Validation loss convergence curves. The dual-head model achieves the lowest final loss; the heteroscedastic objective acts as an implicit regulariser preventing overfitting to noisy leaf observations.

### 6.5. Regularisation Weight Sensitivity ($\lambda_{reg}$, $\lambda_{sem}$)

The values $\lambda_{\text{reg}} = 0.01$ and $\lambda_{\text{sem}} = 0.1$ were chosen via grid search on the Pheno4D validation fold *before* any test evaluation. Table 12 reports the sensitivity sweep. PSNR is robust across $\lambda_{\text{reg}} \in [0.001, 0.01]$ but degrades at $\lambda_{\text{reg}} = 0.1$, where over-aggressive distortion penalisation blurs fine leaf textures. For $\lambda_{\text{sem}}$, values in $[0.05, 0.1]$ are optimal; $\lambda_{\text{sem}} = 1.0$ over-smooths thin structures by dominating the photometric loss.

### 6.6. Cross-Dataset Generalisation

Table 13 tests zero-shot transfer from Pheno4D Maize to ROSE-X and UA-Field. UA-PlantNeRF generalises best (+2.1 dB over DietNeRF on ROSE-X), attributable to the pretrained CLIP encoder that provides broad botanical semantic coverage beyond the training species.

## 7. Conclusion

We presented UA-PlantNeRF, an uncertainty-aware 3D plant reconstruction framework that jointly addresses sparse-view degradation and uncali-

Table 8: Hyperparameter sensitivity (Pheno4D Maize, $N = 30$). **Bold**: chosen configuration.

| $p$ | $K$ | PSNR↑ | Coverage (%)↑ | ECE↓ | Inference |
|---|---|---|---|---|---|
| 0.05 | 20 | 28.2 | 85.3 | 0.062 | 4.1 s |
| **0.10** | **20** | **28.7** | **89.7** | **0.041** | **4.1 s** |
| 0.15 | 20 | 28.4 | 87.1 | 0.054 | 4.1 s |
| 0.20 | 20 | 28.0 | 84.6 | 0.071 | 4.1 s |
| 0.10 | 5 | 28.5 | 81.2 | 0.089 | 1.1 s |
| 0.10 | 10 | 28.6 | 86.4 | 0.057 | 2.1 s |
| 0.10 | 20 | 28.7 | 89.7 | 0.041 | 4.1 s |
| 0.10 | 50 | 28.7 | 90.1 | 0.039 | 10.2 s |

brated confidence estimation. Three tightly integrated contributions—VISAR intelligent frame selection, decomposed aleatoric/epistemic uncertainty via MC Dropout and a dual-head architecture, and plant-wise split conformal prediction calibration with explicit exchangeability guarantees—collectively achieve PSNR 28.7±0.6 dB and SSIM 0.89±0.01 at $N = 30$ ($p < 0.05$, Wilcoxon + Bonferroni), a 31.4 % reduction in Chamfer Distance, and 89.7 % empirical coverage at the 90 % *marginal* target. Sensitivity analyses over eleven hyperparameters and four $\alpha$ levels confirm robustness. Total deployment cost is 4 h 10 min on a single consumer GPU.

Three limitations guide future work. First, the conformal coverage guarantee is **marginal**—it holds on average over plants and rays, not condition-

Table 9: Conformal $\alpha$ level analysis (Pheno4D Maize).

| $\alpha$ | Target Cov. | Empirical Cov. | Mean Int. Width |
|---|---|---|---|
| 0.05 | 95.0 % | 94.8 % | 0.312 |
| **0.10** | **90.0 %** | **89.7 %** | **0.241** |
| 0.15 | 85.0 % | 84.9 % | 0.198 |
| 0.20 | 80.0 % | 80.2 % | 0.167 |

Table 10: Sensitivity to the numerical stabiliser $\varepsilon$ in Eq. (6) (Pheno4D Maize, $N = 30$, $\alpha = 0.10$).

| $\varepsilon$ | Coverage (%) | ECE | AUSE |
|---|---|---|---|
| $10^{-8}$ | 89.7 | 0.041 | 0.089 |
| $\mathbf{10^{-6}}$ | **89.7** | **0.041** | **0.089** |
| $10^{-4}$ | 89.6 | 0.042 | 0.090 |
| $10^{-2}$ | 88.1 | 0.051 | 0.097 |

ally per-region; conditional conformal methods [22] would provide per-region guarantees but require larger calibration sets. Second, the static scene assumption fails for wind-induced deformation; deformable radiance fields [33] would be needed for in-situ temporal monitoring. Third, FSGS [61] and DNGaussian [62] could not be fairly evaluated without scanner depth initialisation on Pheno4D/ROSE-X; this comparison will be included in the UA-Field extension. Additional future directions include single-pass amortised uncertainty inference [28, 29] and causal anomaly detection [18] to disentangle reconstruction artefacts from genuine morphological features.

Table 11: Per-region uncertainty decomposition (UA-Field, $N = 30$; uncertainty values normalised to $[0, 1]$; dominant component per row in **bold**).

| Region | $\mathcal{U}_{\mathrm{al}}$ | $\mathcal{U}_{\mathrm{ep}}$ | $\mathcal{U}_{\mathrm{total}}$ | CD (mm) |
|---|---|---|---|---|
| Leaf surface | **0.031** | 0.012 | 0.043 | 4.21 |
| Leaf edge | **0.048** | 0.067 | 0.115 | 7.83 |
| Stem | **0.019** | 0.008 | 0.027 | 3.15 |
| Fruit / flower | 0.026 | **0.031** | 0.057 | 5.04 |
| Occluded | 0.041 | **0.143** | 0.184 | 11.42 |
| **Overall** | 0.035 | 0.057 | 0.092 | **5.78** |

## CRediT Author Contribution Statement

**Maria Oliveira**: Conceptualization, Methodology, Software, Formal analysis, Writing – Original Draft, Visualization. **João Pereira**: Data curation, Validation, Software, Writing – Review & Editing. **Ana Souza**: Methodology (uncertainty quantification), Formal analysis, Writing – Review & Editing, Supervision.

## Declaration of Competing Interest

The authors declare no known competing financial interests or personal relationships that could influence this work.

## Data Availability

**UA-Field dataset**: deposited to Zenodo [60] (DOI registered, files embargoed until acceptance). **Code**: `https://github.com/ua-plantnerf/ua-plantnerf` (available upon acceptance).

Table 12: Regularisation weight sensitivity (Pheno4D Maize, $N = 30$, $p = 0.10$, $K = 20$). **Bold**: chosen configuration.

| $\lambda_{\text{reg}}$ | $\lambda_{\text{sem}}$ | PSNR↑ | SSIM↑ | CD↓ |
|---|---|---|---|---|
| 0.001 | 0.10 | 28.4 | 0.88 | 5.92 |
| **0.010** | **0.10** | **28.7** | **0.89** | **5.78** |
| 0.100 | 0.10 | 27.3 | 0.86 | 6.41 |
| 0.010 | 0.01 | 28.1 | 0.87 | 6.12 |
| 0.010 | 0.10 | 28.7 | 0.89 | 5.78 |
| 0.010 | 0.50 | 28.2 | 0.87 | 6.08 |
| 0.010 | 1.00 | 27.8 | 0.87 | 6.23 |

## Acknowledgements

The authors thank the PPGC-UFRGS High-Performance Computing Lab for GPU resources and the greenhouse facility staff for data collection support. This work received no external funding.

## References

[1] F. Fiorani, U. Schurr, Future scenarios for plant phenotyping, *Annu. Rev. Plant Biol.* 64 (2013) 267–291.

[2] A. Walter, F. Liebisch, A. Hund, Plant phenotyping: from bean weighing to image analysis, *Plant Methods* 11 (2015) 1–11.

[3] M. Minervini, A. Fischbach, H. Scharr, S.A. Tsaftaris, Finely-grained annotated datasets for image-based plant phenotyping, *Pattern Recognit. Lett.* 81 (2016) 80–89.

Table 13: Zero-shot cross-dataset generalisation (train: Pheno4D Maize, $N = 30$). $\pm$: 95 % bootstrap CI. $^{\dagger}$: $p<0.05$ vs. DietNeRF, Wilcoxon + Bonferroni ($k=3$ baselines $\times$ 2 metrics $\times$ 2 datasets = 12 tests).

| Method | ROSE-X | | UA-Field | |
|---|---|---|---|---|
| | PSNR | SSIM | PSNR | SSIM |
| RegNeRF [15] | $18.4 \pm 1.8$ | $0.64 \pm 0.04$ | $19.7 \pm 1.6$ | $0.67 \pm 0.04$ |
| FreeNeRF [38] | $19.1 \pm 1.6$ | $0.66 \pm 0.04$ | $20.3 \pm 1.5$ | $0.68 \pm 0.03$ |
| DietNeRF [39] | $20.8 \pm 1.4$ | $0.70 \pm 0.03$ | $21.9 \pm 1.3$ | $0.72 \pm 0.03$ |
| **UA-PlantNeRF** | $\mathbf{22.9 \pm 0.9}^{\dagger}$ | $\mathbf{0.76 \pm 0.02}^{\dagger}$ | $\mathbf{23.7 \pm 0.8}^{\dagger}$ | $\mathbf{0.77 \pm 0.02}^{\dagger}$ |

[4] A. Paproki, X. Sirault, S. Berry, R. Furbank, J. Fripp, A novel mesh processing based technique for 3D plant analysis, *Plant Methods* 8 (2012) 1–13.

[5] M.P. Pound, A.P. French, J.A. Atkinson, D.M. Wells, M.J. Bennett, T. Pridmore, Automated recovery of 3D models of plant shoots from multiple color images, *Plant Physiol.* 166 (2014) 1688–1698.

[6] Y. Jiang, C. Li, A.H. Paterson, J.S. Robertson, 3D point cloud data to quantitatively characterize size and shape of shrub crops, *Hortic. Res.* 6 (2019) 43.

[7] D. Ward, P. Moghadam, N. Hudson, Artificial intelligence for agriculture: a review, arXiv:2008.00733 (2020).

[8] J.L. Schönberger, J.-M. Frahm, Structure-from-motion revisited, in: Proc. CVPR, 2016, pp. 4104–4113.

[9] J.L. Schönberger, E. Zheng, J.-M. Frahm, M. Pollefeys, Pixelwise view

selection for unstructured multi-view stereo, in: Proc. ECCV, Springer, 2016, pp. 501–518.

[10] J. Ubbens, M. Cieslak, P. Prusinkiewicz, I. Stavness, Deep plant phenomics, *Front. Plant Sci.* 9 (2018) 1190.

[11] M. Tarif, M. Homaei, A. Mirzaei, B. Nouri-Moghaddam, Optimizing RPL routing using Tabu search to improve link stability and energy consumption in IoT networks, *Comput. Mater. Contin.* 87 (2026) doi:10.32604/cmc.2025.071676.

[12] M. Tarif, M. Homaei, A. Mosavi, An enhanced fuzzy routing protocol for energy optimization in underwater wireless sensor networks, *Comput. Mater. Contin.* 83 (2025).

[13] A. Kamilaris, F.X. Prenafeta-Boldú, Deep learning in agriculture: a survey, *Comput. Electron. Agric.* 147 (2018) 70–90.

[14] B. Mildenhall, P.P. Srinivasan, M. Tancik, J.T. Barron, R. Ramamoorthi, R. Ng, NeRF: representing scenes as neural radiance fields for view synthesis, in: Proc. ECCV, Springer, 2020, pp. 405–421.

[15] M. Niemeyer, J.T. Barron, B. Mildenhall, M.S.M. Sajjadi, A. Geiger, N. Radwan, RegNeRF: regularizing neural radiance fields for view synthesis from sparse inputs, in: Proc. CVPR, 2022, pp. 5480–5490.

[16] A. Kendall, Y. Gal, What uncertainties do we need in Bayesian deep learning for computer vision?, in: Adv. NeurIPS, vol. 30, 2017.

[17] M. Tarif, E. Fasani, D. Quaglia, VISAR: intelligent video frame selection for agricultural 3D plant reconstruction, in: 2025 IEEE MetroAgriFor, IEEE, 2025, pp. 99–104.

[18] M. Homaei, M. Tarif, P.G. Rodríguez, A. Caro, M. Ávila, Causal digital twins for cyber-physical security in water systems, *Mach. Learn. Appl.* (2025) 100824.

[19] M. Homaei, M. Tarif, P.G. Rodríguez, M. Ávila, A. Caro, Adaptive multi-domain uncertainty quantification for digital twin water forecasting, *Mach. Learn. Appl.* (2025) 100812.

[20] M. Homaei, M. Tarif, P.G. Rodriguez, A. Caro, M. Avila, Conformal prediction-driven adaptive sampling for digital water twins, *ICT Express* (2026).

[21] Y. Gal, Z. Ghahramani, Dropout as a Bayesian approximation, in: Proc. ICML, 2016, pp. 1050–1059.

[22] A.N. Angelopoulos, S. Bates, A gentle introduction to conformal prediction and distribution-free uncertainty quantification, arXiv:2107.07511 (2021).

[23] V. Vovk, A. Gammerman, G. Shafer, Algorithmic Learning in a Random World, Springer, New York, 2005.

[24] J.T. Barron, B. Mildenhall, M. Tancik, P. Hedman, R. Martin-Brualla, P.P. Srinivasan, Mip-NeRF: a multiscale representation for anti-aliasing neural radiance fields, in: Proc. ICCV, 2021, pp. 5855–5864.

[25] J.T. Barron, B. Mildenhall, D. Verbin, P.P. Srinivasan, P. Hedman, Mip-NeRF 360: unbounded anti-aliased neural radiance fields, in: Proc. CVPR, 2022, pp. 5470–5479.

[26] J.T. Barron, B. Mildenhall, D. Verbin, P.P. Srinivasan, P. Hedman, Zip-

NeRF: anti-aliased grid-based neural radiance fields, in: Proc. ICCV, 2023, pp. 19697–19705.

[27] R. Martin-Brualla, N. Radwan, M.S.M. Sajjadi, J.T. Barron, A. Dosovitskiy, D. Duckworth, NeRF in the wild, in: Proc. CVPR, 2021, pp. 7210–7219.

[28] T. Müller, A. Evans, C. Schied, A. Keller, Instant neural graphics primitives with a multiresolution hash encoding, *ACM Trans. Graph.* 41 (2022) 1–15.

[29] A. Chen, Z. Xu, A. Geiger, J. Yu, H. Su, TensoRF: tensorial radiance fields, in: Proc. ECCV, Springer, 2022, pp. 333–350.

[30] S. Fridovich-Keil, A. Yu, M. Tancik, Q. Chen, B. Recht, A. Kanazawa, Plenoxels: radiance fields without neural networks, in: Proc. CVPR, 2022, pp. 5501–5510.

[31] B. Kerbl, G. Kopanas, T. Leimkühler, G. Drettakis, 3D Gaussian splatting for real-time radiance field rendering, *ACM Trans. Graph.* 42 (2023) 1–14.

[32] M. Tancik et al., Block-NeRF: scalable large scene neural view synthesis, in: Proc. CVPR, 2022, pp. 8248–8258.

[33] A. Pumarola, E. Corona, G. Pons-Moll, F. Moreno-Noguer, D-NeRF: neural radiance fields for dynamic scenes, in: Proc. CVPR, 2021, pp. 10318–10327.

[34] K. Zhang, G. Riegler, N. Snavely, V. Koltun, NeRF++: analyzing and improving neural radiance fields, arXiv:2010.07492 (2020).

[35] L. Yariv et al., Multiview neural surface reconstruction by disentangling geometry and appearance, in: Adv. NeurIPS, vol. 33, 2020, pp. 2492–2502.

[36] L. Yariv, J. Gu, Y. Kasten, Y. Lipman, Volume rendering of neural implicit surfaces, in: Adv. NeurIPS, vol. 34, 2021, pp. 4805–4815.

[37] P. Wang et al., NeuS: learning neural implicit surfaces by volume rendering, in: Adv. NeurIPS, vol. 34, 2021, pp. 27171–27183.

[38] J. Yang, M. Pavone, Y. Wang, FreeNeRF: improving few-shot neural rendering with free frequency regularization, in: Proc. CVPR, 2023, pp. 8254–8263.

[39] A. Jain, M. Tancik, P. Abbeel, Putting NeRF on a diet, in: Proc. ICCV, 2021, pp. 5885–5894.

[40] M. Kim, S. Seo, B. Han, InfoNeRF: ray entropy minimization for few-shot neural volume rendering, in: Proc. CVPR, 2022, pp. 12912–12921.

[41] G. Wang, Z. Chen, C.C. Loy, Z. Liu, SparseNeRF: distilling depth ranking for few-shot novel view synthesis, in: Proc. ICCV, 2023, pp. 9118–9127.

[42] A. Yu, V. Ye, M. Tancik, A. Kanazawa, pixelNeRF: neural radiance fields from one or few images, in: Proc. CVPR, 2021, pp. 4578–4587.

[43] A. Chen et al., MVSNeRF: fast generalizable radiance field reconstruction from multi-view stereo, in: Proc. ICCV, 2021, pp. 14124–14133.

[44] Q. Wang et al., IBRNet: learning multi-view image-based rendering, in: Proc. CVPR, 2021, pp. 4690–4699.

[45] X. Pan, Z. Lai, S. Song, G. Huang, ActiveNeRF: learning where to see with uncertainty estimation, in: Proc. ECCV, Springer, 2022, pp. 230–246.

[46] J. Shen, A. Ruiz, A. Agudo, F. Moreno-Noguer, Conditional-flow NeRF, in: Proc. ECCV, Springer, 2022, pp. 540–557.

[47] L. Goli, C. Reading, S. Sellán, A. Jacobson, A. Tagliasacchi, Bayes' rays: uncertainty quantification for neural radiance fields, in: Proc. CVPR, 2024, pp. 20061–20070.

[48] B. Lakshminarayanan, A. Pritzel, C. Blundell, Simple and scalable predictive uncertainty estimation using deep ensembles, in: Adv. NeurIPS, vol. 30, 2017.

[49] C. Guo, G. Pleiss, Y. Sun, K.Q. Weinberger, On calibration of modern neural networks, in: Proc. ICML, 2017, pp. 1321–1330.

[50] Y. Ovadia et al., Can you trust your model's uncertainty?, in: Adv. NeurIPS, vol. 32, 2019.

[51] S.P. Mohanty, D.P. Hughes, M. Salathé, Using deep learning for image-based plant disease detection, *Front. Plant Sci.* 7 (2016) 1419.

[52] I. Sa et al., DeepFruits: a fruit detection system using deep neural networks, *Sensors* 16 (2016) 1222.

[53] H. Lu, Z. Cao, Y. Xiao, B. Zhuang, C. Shen, TasselNet: counting maize tassels in the wild, *Plant Methods* 13 (2017) 1–17.

[54] D. Schunck et al., Pheno4D: a spatio-temporal dataset of maize and tomato plant point clouds, *PLOS ONE* 16 (2021) e0256340.

[55] H. Dutagaci, P. Rasti, E. David, J.-M. Gogue, ROSE-X: an annotated data set for 3D plant organ segmentation, *Plant Methods* 16 (2020) 1–14.

[56] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, Image quality assessment: from error visibility to structural similarity, *IEEE Trans. Image Process.* 13 (2004) 600–612.

[57] R. Zhang, P. Isola, A.A. Efros, E. Shechtman, O. Wang, The unreasonable effectiveness of deep features as a perceptual metric, in: Proc. CVPR, 2018, pp. 586–595.

[58] E. Ilg et al., Uncertainty estimates and multi-hypotheses networks for optical flow, in: Proc. ECCV, Springer, 2018, pp. 652–667.

[59] M. Tancik et al., Nerfstudio: a modular framework for neural radiance field development, in: ACM SIGGRAPH 2023, ACM, 2023.

[60] M. Oliveira, J. Pereira, A. Souza, UA-Field: a smartphone video dataset for plant 3D reconstruction with ground-truth geometry, Zenodo (2025), doi:10.5281/zenodo.XXXXXXX [DOI registered; data embargoed until acceptance].

[61] Z. Zhu, Z. Fan, Y. Jiang, Z. Wang, FSGS: real-time few-shot view synthesis using Gaussian splatting, in: Proc. ECCV, Springer, 2024.

[62] J. Li, J. Zhang, X. Bai, J. Zheng, X. Ning, J. Zhou, L. Gu, DNGaussian: optimizing sparse-view 3D Gaussian radiance fields with global-local depth normalization, in: Proc. CVPR, 2024.