

AngleCam V2: Predicting leaf inclination angles across taxa from daytime and nighttime photos

Luis Kremer¹, Jan Pisek², Ronny Richter^{3,4}, Julian Frey⁵, Daniel Lusk¹, Christiane Werner⁶,
Christian Wirth^{3,4}, and Teja Kattenborn¹

¹Chair of Sensor-based Geoinformatics (geosense), University of Freiburg, Germany

²Tartu Observatory, University of Tartu, Tõravere, Estonia

³German Centre for Integrative Biodiversity Research (iDiv), Halle-Jena-Leipzig, Leipzig,
Germany

⁴Systematic Botany and Functional Biodiversity, Institute of Biology, Leipzig University,
Leipzig, Germany

⁵Chair of Forest Growth and Dendroecology, University of Freiburg, Germany

⁶Chair of Ecosystem Physiology, University of Freiburg, Germany

2026

Correspondence: Luis Kremer, Chair of Sensor-based Geoinformatics (geosense), University of
Freiburg, Germany. Email: luis.kremer@geosense.uni-freiburg.de

Running Head: AngleCam V2: Leaf angles from photos

Acknowledgements

LK and TK received further funding from the German Research Foundation (DFG) under the projects
PANOPS (project number 504978936) and ECOSENSE (project number SFB 1537/1). TK received
funding from the Flexible Funds Program for junior scientists of the University of Leipzig (project
number 232201582). TK and CW gratefully acknowledge the support of iDiv funded by the German
Research Foundation (DFG-FZT 118, 202548816). JP was supported by the Estonian Research
Council Grant PRG 1405 and the Estonian Ministry of Education and Research, Centre of Excellence
for Sustainable Land Use (TK232).

Author Contributions

LK and TK conceived the ideas, designed the methodology, and led the analysis. TK, JP, RR, JF, LK, and DL collected the data. LK and TK led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

Data Availability Statement

All data are available on Zenodo at <https://doi.org/10.5281/zenodo.17086253>, organized into four directories: training/validation images with annotations (01_Training_Validation_Data/), multitemporal test imagery with TLS validation (02_Test_Data/), model weights and predictions (03_Model_Outputs/), and supplementary materials (04_Supplementary_Material/). The source code is available at <https://github.com/Luis-Kr/AngleCamV2>. The pretrained model is available at <https://doi.org/10.5281/zenodo.17101166>.

Conflicts of Interest

All authors declare that they have no conflicts of interest.

Abstract

Understanding how plants capture light and maintain their energy balance is crucial for predicting how ecosystems respond to environmental changes. By monitoring leaf inclination angle distributions (LIADs), we can gain insights into plant behaviour that directly influences ecosystem functioning. LIADs affect radiative transfer processes and reflectance signals, which are essential components of satellite-based vegetation monitoring. Despite their importance, scalable methods for continuously observing these dynamics across different plant species throughout day-night cycles are limited.

We present AngleCam V2, a deep learning model that estimates LIADs from both RGB and near-infrared (NIR) night-vision imagery. We compiled a dataset of over 4,500 images across 200 globally distributed species to facilitate generalization across taxa. To address the lack of labeled NIR data, a training strategy was developed that uses pseudo-NIR imagery derived from daytime RGB images. The model is based on a vision transformer architecture with mixed-modality training incorporating RGB, grayscale, and pseudo-NIR images.

AngleCam V2 achieved substantial improvements in generalization compared to AngleCam V1 ($R^2 = 0.62$ vs 0.12 on the same holdout dataset). Phylogenetic analysis across 100 genera revealed no systematic taxonomic bias in prediction errors. Testing against leaf angle dynamics from multitemporal terrestrial laser scanning confirmed the model's ability to track diurnal leaf movements ($R^2 = 0.61$ – 0.75) with a nighttime RMSE of 3.1° – 3.6° (compared to daytime RMSEs of 3.5° and 7.6° , respectively). Furthermore, the model successfully detected water limitation-induced changes over a 14-day monitoring period.

This method enables continuous monitoring of leaf angle dynamics using conventional cameras, enabling applications in ecosystem monitoring networks, plant stress detection, interpreting satellite vegetation signals, and citizen science platforms for global-scale understanding of plant structural responses.

Keywords: computer vision, deep learning, diurnal plant movements, leaf inclination angle distribution, plant phenotyping, remote sensing, terrestrial laser scanning, vision transformer

1 Introduction

Some plants hold their leaves horizontally like umbrellas, while others keep them upright like rows of solar panels. This difference in orientation, ranging from flat to erect, is described as the vertical leaf angle. Plants have evolved a wide variability of vertical leaf angle orientations across species, growth forms, and geographic regions (Puglielli et al., 2024).

Even within a single plant, leaf angles vary to optimize resource use. Upright upper leaves facilitate light transmission, while horizontal lower leaves maximize light absorption (Niinemets, 2010). This variation, which has been studied since Darwin (Darwin & Darwin, 1883), regulates light uptake, competition, microclimates, and canopy productivity (Hikosaka & Hirose, 1997; Mullen et al., 2006; Niinemets, 2010).

The vertical orientation of a leaf is defined by its surface angle, which ranges from 0° (entirely horizontal) to 90° (entirely vertical). The variability of these angles within a canopy is referred to as the leaf inclination angle distribution (LIAD). LIADs are not static, as plants adjust leaf angles through active and passive processes, including diurnal movements, phenological changes, and water-stress-induced drooping. These responses are influenced by light, temperature, competition, and plant water status (Apelt et al., 2017; Geldhof et al., 2021; Niinemets, 2010; Puglielli et al., 2017). Rapid changes in leaf angles can indicate stress, such as heat or photoinhibition, and regulate canopy energy balances (Ehleringer et al., 1987; Kattenborn et al., 2022; Leuzinger & Körner, 2007; Sastry et al., 2018; Van Zanten et al., 2010; Werner et al., 2001). LIAD variability also significantly influences canopy reflectance, making it critical for global vegetation monitoring and interpreting satellite-based signals (Braghiere et al., 2021; Dechant et al., 2020; Jablonski et al., 2025; Kattenborn et al., 2024).

Despite their importance, few scalable techniques exist for automatically tracking dynamic changes in LIADs (Yang et al., 2023). While inertial measurement units (IMUs) can track individual leaf movements (Geldhof et al., 2021), they are not scalable to entire canopies. 3D point cloud-based approaches using photogrammetry (Qi et al., 2019), terrestrial laser scanning (Bailey & Mahaffee, 2017; Murithi et al., 2025; Stovall et al., 2021; Zheng & Moskal, 2012), or stereo vision (Bernotas et al., 2019; Biskup et al., 2007; Müller-Linow et al., 2015) show promise but require specialized hardware and calibration, limiting their usability in field conditions.

Yet humans can easily perceive differences in leaf angles from real-life observations or photographs (Pisek et al., 2011; Zou et al., 2014). Automating this task using computer vision techniques led to the development of AngleCam (Kattenborn et al., 2022, 2024), a deep learning model designed to estimate LIADs from single, horizontally oriented RGB photographs or video frames. AngleCam was trained on approximately 2500 images and corresponding leaf inclination angle distributions, which were derived from annotating individual leaves in the image frames, and compared against independent TLS-derived LIADs across 25 plant species. The applicability and responsiveness of AngleCam were demonstrated through long-term time series in which predicted leaf inclination angle

distributions exhibited plausible dynamics in response to environmental drivers, including radiation, temperature, and soil humidity at multiple sites (Kattenborn et al., 2022, 2024). Despite these potentials, the original AngleCam (V1) remains limited to daylight observations and has restricted generalizability due to the limited diversity of its training dataset.

Here, we present AngleCam V2, which addresses these limitations through three major advancements: (1) we expanded the training and validation dataset to over 4500 images across 200 species, ensuring robust generalization across diverse growth forms and biomes. (2) we integrated a self-supervised vision transformer (DINOv2) to enhance feature extraction and robustness to varied scene conditions. (3) we developed a training strategy that combines grayscale and pseudo-NIR transformations to leverage existing RGB labels and minimize manual effort for NIR annotation. This approach enables the model to process both RGB and NIR night-vision imagery, supporting continuous 24-hour monitoring of leaf movements.

In this study, we validate the model on a holdout dataset and assess its generalization across 100 genera. Additionally, we test the approach in controlled indoor experiments, which include a comparison with multitemporal terrestrial laser scanning (TLS) and monitoring of a plant under water limitation. The enhanced generalization and temporal capabilities could make this approach a promising candidate for ecosystem monitoring networks and citizen science applications, with potential for contributing to global-scale understanding of leaf angle patterns and their environmental drivers.

2 Materials and Methods

2.1 AngleCam model development

2.1.1 RGB image data across taxa and geographic regions

The dataset for model training and validation grew from 2680 to 4795 images, enhancing taxonomic and morphological diversity with AngleCam V2 (Fig. 1). Unlike AngleCam V1, which relied solely on two datasets from 2021, AngleCam V2 includes these plus three additional collections (Tab. 1). The first is a diverse set from the Leipzig area, featuring 100 species from various environments. The second set consists of image time series taken within the project at the Canopy Crane research platform (Leipzig Canopy Crane [LCC]) of the German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig. This site covers a Leipzig floodplain forest, a structurally complex hardwood forest dominated by tree species such as European ash (*Fraxinus excelsior* L.), English oak (*Quercus robur* L.), Sycamore maple (*Acer pseudoplatanus* L.), European hornbeam (*Carpinus betulus* L.), and small-leaved lime (*Tilia cordata* Mill) (see Richter et al., 2022, for details). Together, both datasets provided a total of 102 different species. While *Tilia cordata* and *Acer pseudoplatanus* represented approximately 50% of the dataset, this reflects extensive temporal sampling from continuous monitoring at the Leipzig Canopy Crane, capturing a large variability of LIADs of these species across diverse environmental conditions and seasonal phenology rather than simple taxonomic overrepresentation (Fig. 1). These image series were captured using TLC-200 Pro timelapse cameras (Brinno Inc., Taipei, Taiwan) in high-dynamic range (HDR) mode (further details see Kattenborn et al., 2022).

With the new version of AngleCam, we have expanded the training and validation dataset to make the model more robust against various leaf shapes and illumination conditions. The entire dataset now includes three additional datasets in addition to the two from the first version:

- Image series taken within the field experiment MyDiv (Mycorrhiza in tree Diversity effects on ecosystem functioning; Ferlian et al., 2018).
- A globally sourced collection of over 600 tree and shrub samples representing more than 100 species from diverse climates and biomes.
- An independent test dataset with image series of three indoor plant species, including two validated against terrestrial laser scans (described in Section 2.3 & 2.4).

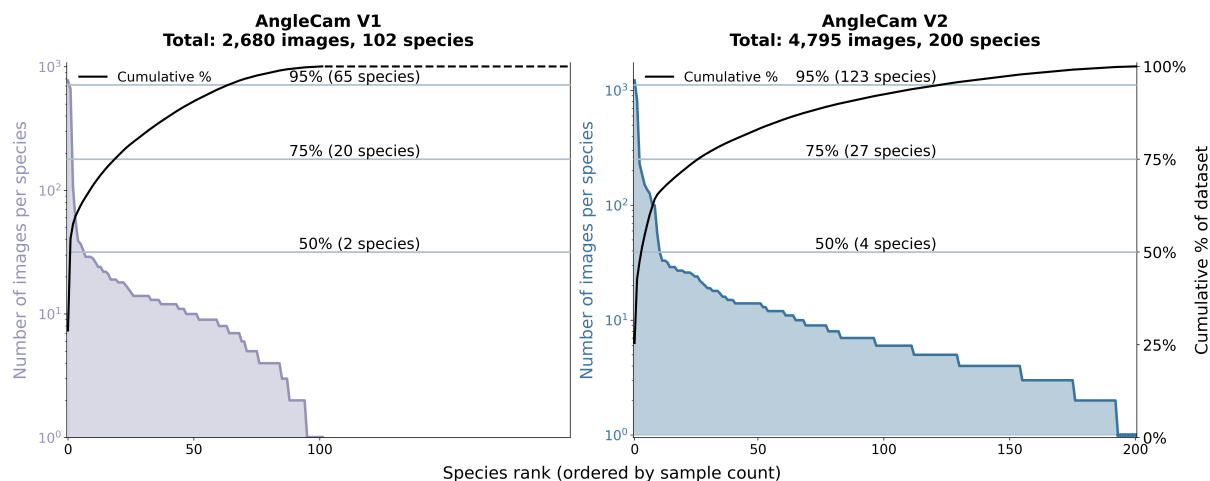


Figure 1: Species distribution in AngleCam (V1 & V2) training datasets. Rank-abundance curves showing the number of images per species (left y-axis, log scale) and cumulative dataset coverage (right y-axis) for Version 1 (left panel, 2680 images across 102 species) and Version 2 (right panel, 4795 images across 200 species). The curves illustrate how images are distributed across species in each dataset, with reference lines indicating the number of species required to achieve 50%, 75%, and 95% of total dataset coverage.

The MyDiv experiment, conducted at the Bad Lauchstädt Experimental Research Station of the Helmholtz Center for Environmental Research-UFZ in Saxony-Anhalt, Germany, is structured to investigate the effects of tree diversity on ecosystem functions (Ferlian et al., 2018). The data for this analysis specifically used measurements from monoculture plots of nine temperate deciduous tree species: *Acer pseudoplatanus* L., *Aesculus hippocastanum* L., *Fraxinus excelsior* L., *Sorbus aucuparia* L., *Betula pendula* R., *Carpinus betulus* L., *Fagus sylvatica* L., *Quercus petraea* (Matt.) Liebl., and *Tilia platyphyllos* Scop. For each of these nine species, two time-lapse cameras were deployed in the plot center of the respective monoculture plots. The cameras were positioned to capture a horizontal field of view of branches from two individuals located centrally within the plot. The installation height corresponded to the upper third of the tree crowns. Time series images were collected at 5-minute intervals during the primary vegetation period from June 6 to October 5, 2022 (further details see Kattenborn et al., 2024).

To improve the model's robustness across genera and scene conditions, we integrated a comprehensive dataset that includes 676 horizontal plant photographs representing over 100 distinct species across a broad taxonomic and ecological spectrum. The observations were collected from botanical gardens and arboreta worldwide, including locations in Europe (Czech Republic, Estonia, France, Italy, Netherlands, Portugal, Spain, Sweden, and the United Kingdom), as well as in the USA, Israel, and Australia. This collection covers a variety of biomes and climates, ranging from Mediterranean and temperate to subtropical regions. For several species and genera, such as *Eucalyptus*, *Quercus*, and

Betula, multiple samples from diverse locations were collected, resulting in a dataset that includes both intra- and interspecific variation under varied environmental conditions.

Table 1: Datasets used for AngleCam V1 and V2 model training and validation.

Dataset	Year(s)	Species Count	Image Count	Description	AngleCam Version
Area of Leipzig	2021	100	1237	Natural/semi-natural areas, parks, gardens, indoor plants	V1 & V2
Leipzig Canopy Crane (LCC)	2021	2	1431	Floodplain forest, a hardwood ecosystem	V1 & V2
MyDiv experiment	2022	9	1422	Tree monocultures	V2 only
Global Collection	2012–2022	>100	676	Botanical gardens, arboreta worldwide	V2 only
Indoor experiments	2024–2025	3	35	Controlled indoor plant monitoring	V2 only

2.1.2 Labeling RGB images with leaf inclination angle distributions

Deep learning models benefit from large, well-annotated training datasets. To generate leaf angle references efficiently, we used visual estimation based on RGB images. Previous studies obtained estimates of leaf inclination angle distributions by measuring the vertical angle of leaves that are oriented perpendicular to the camera’s line of sight (Pisek et al., 2011; Ryu et al., 2010). Multiple measurements of individual leaves can be summarized as a leaf inclination angle distribution. However, these strict geometric requirements limit the number of usable samples per image, the robustness of a leaf inclination angle distribution, and pose challenges for species with complex or curled leaf forms (Murithi et al., 2025). Here, we overcome these limitations by estimating the average leaf inclination of whole leaf surfaces through visual interpretation of their apparent angle from horizontal in RGB images. This process relied solely on visual estimation, without additional geometric cues. Only clearly visible leaves were annotated, while occluded leaves were excluded. For curled or non-planar leaves, annotators estimated the average inclination of the entire leaf surface. This approach was previously applied successfully for AngleCam V1 (Kattenborn et al., 2022).

We obtained a LIAD for each reference image by sampling the average leaf surface inclinations of 20 leaves. This sample size was chosen to obtain a representative intra-image leaf angle variability, as well as a large image dataset across species and scene conditions. We followed the approach from AngleCam V1, where this sampling strategy was validated by terrestrial laser scanning ($R^2 = 0.74$) (Kattenborn et al., 2022). For each image, we placed an evenly spaced grid of points and selected the leaf closest to each point for annotation. This method ensures spatial balance across the field of

view, reduces bias toward prominent leaves, and samples leaves throughout the vertical gradient of the visible canopy seen from the horizontal camera angle.

Although this approach enables efficient annotation, it may introduce observer-related uncertainty. To reduce sensitivity to label uncertainty, individual discrete angle measurements were not used directly as training targets. Instead, the 20 leaf samples were converted into probability distributions across the full 0° – 90° leaf angle range. The two-parameter beta distribution was employed, as it is well-suited to model continuous gradients of leaf inclination angle distribution shapes, ranging from strongly horizontal (planophile) to uniform, spherical, or strongly vertical (erectophile) forms (Goel & Strebel, 1984). Beta distribution fitting was performed in Python using the `scipy.stats` module (v1.13.0), with parameters (α, β) estimated by maximum likelihood.

Rather than directly predicting the (α, β) parameters of a beta distribution, the model was trained to predict the entire probability density function (PDF) across the 0° – 90° range at 2° intervals, since the underlying LIAD may differ from an idealized beta distribution in some cases (Kattenborn et al., 2022). To further mitigate the impact of potential annotation bias and enhance generalization, the fitted (α, β) values of the reference LIADs were augmented, generating 50 synthetic variants for each image by sampling (α, β) within $\pm 20\%$ of the standard deviation of the maximum likelihood estimates. This data augmentation increases the model's robustness to plausible measurement variations. Additionally, by training on over 3500 images, random interpretation errors are expected to be mitigated during optimization rather than systematically propagating into the learned signal (Rolnick et al., 2017). The robustness of this labeling approach was verified by evaluating the model against independent leaf angle measurements derived from terrestrial laser scanning (see Section 2.3).

2.1.3 Generating training data on pseudo-NIR images

While AngleCam V1 was trained exclusively on RGB images, we specifically adapted AngleCam V2 to also process NIR images acquired under nighttime conditions, enabling fully continuous monitoring of leaf angle dynamics.

Night-vision images are captured by cameras with infrared light-emitting diodes (NIR-LEDs) that emit near-infrared radiation (850–940 nm). The camera sensor records reflected NIR light, resulting in grayscale images that represent backscatter intensity (Sun et al., 2023). These images differ from daylight RGB images in several ways: the spectrum varies (e.g., vegetation appears bright in NIR), the effective range of the NIR LED is around 9 meters, and illumination decreases with the square of the distance from the light source (Pharr et al., 2023).

These fundamental differences posed a challenge, as a model trained only on RGB images would perform poorly on NIR images due to modality mismatch. Given that a large dataset of labeled RGB data was already available (Tab. 1), we decided not to invest the same workload to create an equally sized dataset of labeled NIR imagery. Instead, we developed an approach to create synthetic NIR imagery from the already extensively labeled RGB data.

Firstly, we randomly converted 50% of the RGB in the training dataset to grayscale using luminance weighting ($0.299R + 0.587G + 0.114B$) to reduce reliance on color. Secondly, we randomly added a distance-based dimming to 50% of the grayscale images (pseudo-NIR images) to mimic the limited range of the NIR illuminator and the corresponding brightness falloff with distance (Fig. 2). For this, we scaled the pixel intensities of each grayscale image according to the pixel-wise distances estimated from computer vision-derived depth maps (Depth Anything V2; see details below). The final transformation was defined as:

$$I_{\text{NIR}} = I_{\text{gray}} \cdot \left(\frac{d_0}{d + \varepsilon} \right)^2, \quad d_0 = 1, \varepsilon = 2.$$

where I_{gray} is the grayscale image, d is the per-pixel depth (in meters), and d_0 is a normalization distance. We set $\varepsilon = 2$ to ensure numerical stability for minimal depths and to slightly compress the distance attenuation. This yields pseudo-NIR images that appear brighter and less contrast-extreme (Fig. 2).

The depth maps used to simulate the NIR imagery were generated using Depth Anything V2, a state-of-the-art monocular depth estimation model (Yang et al., 2024). We used the large model variant, fine-tuned for metric depth estimation on the Virtual KITTI 2 dataset (Cabon et al., 2020). To align predicted depth values with the operational range of our NIR camera system, we constrained maximum depth predictions to 9 meters.

2.1.4 Model architecture and training

For AngleCam V2, we revised the architecture by replacing the EfficientNet-B7 backbone with the self-supervised DINOv2 ViT-S/14 model as the feature extractor (Oquab et al., 2023). Vision transformers offer a comprehensive feature representation, effectively recognizing structures and capturing depth cues through their global context processing (Dosovitskiy et al., 2020; Tolan et al., 2024).

The DINOv2 ViT-S/14 backbone outputs a 384-dimensional embedding, which we passed into a

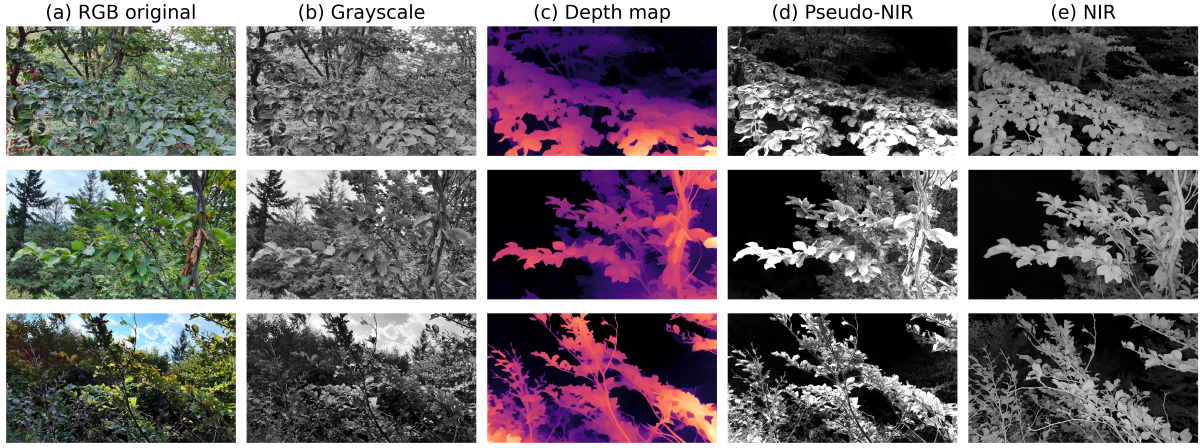


Figure 2: Pseudo-NIR image generation pipeline. The transformation from RGB to pseudo-NIR involves several steps: (a) RGB original image, (b) conversion to grayscale, (c) depth map estimation using Depth Anything V2 (closer objects to the camera appear brighter), and (d) pseudo-NIR image generation by scaling the intensity of the grayscale image with the corresponding depth map, simulating NIR-LED illumination falloff. (e) Actual NIR image captured at nighttime of the same scene; see Fig. A1 in the Appendix for further examples.

lightweight regression head: a fully connected layer ($384 \rightarrow 128$) with dropout (rate = 0.4), GELU activation (Hendrycks & Gimpel, 2016), another fully connected layer ($128 \rightarrow 43$), and a softmax output layer for a 43-bin probability distribution across 2° intervals from 0° to 90° .

Images were resized to 224×224 pixels via linear interpolation. The augmentation pipeline included *RandomResizedCrop* (scale = $[0.8, 1.0]$) and horizontal flipping (50% probability), while photometric augmentations adjusted brightness and contrast (variation factors = 0.3) and added Gaussian noise ($\sigma = 0.02$) for improved robustness to sensor noise and artifacts.

Instead of training separate models for RGB and NIR imagery, we developed AngleCam V2 using a mixed-modality approach. This design ensures that, regardless of whether NIR or RGB data are available, no model switching or dataset separation is required. During each epoch, 50% of images were randomly presented as original RGB, while the remaining 50% were non-RGB variants: true grayscale and pseudo-NIR images generated through distance-based dimming (Section 2.1.3; Fig. 2). All images were normalized using ImageNet statistics (mean = $[0.485, 0.456, 0.406]$, std = $[0.229, 0.224, 0.225]$).

The model was trained on an NVIDIA RTX A6000 GPU (48 GB RAM), with the final training run of 50 epochs completed in 24 minutes. Hyperparameters were determined through testing of different configurations, including batch sizes (8–64), optimizers (Adam, AdamW, and SGD), and learning rates (10^{-5} to 10^{-1}), alongside multiple regression head architectures with varying dropout rates (0–0.5) and weight decay settings (10^{-5} to 10^{-1}).

The final configuration used the AdamW optimizer with an initial learning rate of 1×10^{-4} , weight decay of 0.01, and gradient clipping (max norm = 3.0, L_2 norm) to prevent overfitting (Loshchilov & Hutter, 2017). Learning rate scheduling was managed with *ReduceLROnPlateau* (reduction factor 0.5, patience 5 epochs, minimum learning rate 1×10^{-6}). Training was performed with a batch size of 32 and Huber loss to reduce sensitivity to outliers (Sun et al., 2020). The final model was selected based on the lowest validation loss obtained through the 50 epochs.

Prior to splitting the dataset, three RGB images per genus were extracted to assess phylogenetic prediction errors (see next Section 2.2). The remaining labeled data were split 80/20 into training and validation sets, resulting in approximately 3600 images for training and 900 for validation. The random split resulted in some overlap of species and acquisition sites between training and validation sets. While fully independent validation datasets would be ideal, creating stratified splits across multiple dimensions (species, sites, temporal series, and leaf angles) would result in impractically small subsets. Importantly, images within shared species were captured across different seasons, phenological stages, and environmental conditions, providing substantial variability. Validation was performed exclusively on original RGB images since the pseudo-NIR and grayscale variants share the same labels. The model's capability to process NIR imagery was independently validated through controlled experiments using terrestrial laser scanning as reference data (Section 2.3).

2.2 Phylogenetic error analysis

Predictive models in ecology often exhibit systematic biases when applied across different taxonomic groups, particularly when training data is not balanced across taxonomic and phylogenetic lineages (Morales-Castilla et al., 2024). Such phylogenetic autocorrelation in prediction errors can indicate dataset gaps, model limitations, or evolutionary constraints that limit generalizability (Roberts et al., 2017). This consideration is relevant for our dataset, where temporal monitoring at specific sites resulted in uneven species representation, with approximately 50% of training images originating from four species. Given this taxonomic imbalance, we assessed whether the residuals of AngleCam V2 are phylogenetically autocorrelated.

For images where species identification was not available in the original metadata, we determined plant species using the Pl@ntNet online identification tool (PlantNet, 2025). Each identification was manually reviewed to assess plausibility at both species and genus levels. Due to increased uncertainty in species-level assignments from images, phylogenetic error analysis was performed at the genus level. To evaluate phylogenetic error in the residuals, three random images per genus were excluded

from the training dataset and used for analysis. Genera with fewer than three samples were not analyzed due to insufficient significance. Consequently, from a total of 129 genera, 100 genera with 300 samples were included in the analysis.

For each genus, we calculated the LIAD residuals as the mean absolute deviation between AngleCam V2 predictions and labeled LIAD references. The genus-level errors were then mapped onto a phylogenetic tree to visualize error patterns between taxonomic groups. We quantified the phylogenetic signal in prediction errors using Pagel's λ , which measures the degree to which closely related genera exhibit similar prediction errors (Pagel, 1999). A λ near 1 indicates a strong phylogenetic influence, while a λ near 0 suggests independence from phylogeny (Kamilar & Cooper, 2013). We estimated Pagel's λ with the function *phylosig* of the R-package *phytools* (v.2.4-4).

We quantified phylogenetic autocorrelation in model residuals using Moran's *I* statistic calculated on phylogenetic distances (Gittleman & Kot, 1990; Moran, 1950). Significant positive autocorrelation would indicate that prediction errors are clustered within particular clades, suggesting systematic biases that could limit model transferability to underrepresented taxonomic groups (Hawkins, 2012). We estimated Moran's *I* with the function *moran.idx* of the R-package *adephylo* (v.1.1-17).

2.3 Multitemporal model evaluation of AngleCam V2 with TLS-derived leaf inclination angle distributions

2.3.1 Experimental setup and data acquisition

We validated diurnal leaf inclination angle distributions obtained from AngleCam V2 against independent LIADs from terrestrial laser scanning (TLS). An indoor experiment was conducted with two species known for strong diurnal movements, *Maranta leuconeura* E.Morren and *Calathea ornata* (Linden) Körn, monitored from 17–20 January 2025.

The TLS scanner (Riegl VZ-400i; RIEGL Laser Measurement Systems GmbH, Horn, Austria) was positioned 1.5–2.0 m from the plants for unobstructed coverage, acquiring scans every 30 minutes. It operates with a near-infrared laser (1550 nm) and offers a measurement precision of 5 mm, with a field of view of 100° vertical and 360° horizontal.

Two Ubiquiti UniFi G5 Bullet cameras (Ubiquiti Inc., New York, NY, USA) were installed 1 m from each plant, capturing images every 30 minutes in both RGB and near-infrared modes (NIR range: 9 m) and synchronized with TLS acquisitions. Images were geometrically calibrated to correct for lens distortion.

2.3.2 TLS data processing and AngleCam comparison

The TLS point clouds were processed to extract LIADs for comparison with AngleCam V2 predictions. The pipeline focused solely on leaf surfaces by manually defining bounding boxes to isolate target plants, eliminating stems, petioles, and background objects.

Point clouds were centered around the plant base for a consistent coordinate system. Surface normals were estimated using principal component analysis (PCA) within a 1 cm radius, identifying the direction of least variance (Jolliffe, 2002; Rusu et al., 2008). Normals were oriented toward the scanner for consistent angle calculations, and the point clouds were subsampled with a voxel grid at a 1 cm resolution to reduce density variation.

For statistical analysis, TLS-derived LIADs were temporally aligned with AngleCam predictions based on acquisition timestamps. Quantitative comparisons were performed on average leaf angle values from LIAD distributions and the closest-in-time AngleCam predictions, calculating metrics like the coefficient of determination (R^2) and root mean square error (RMSE) for each matched time point.

2.4 Multitemporal model evaluation of AngleCam V2 under water limitation

We performed an additional indoor experiment to assess whether AngleCam V2 captures extended, multi-day trends under plant water limitation. We positioned an individual *Aglaonema commutatum* Schott in front of a window and recorded images as described above over 14 days (17–31 December 2024). Images were acquired every two minutes, resulting in 8783 images. The plant was last watered on December 12 and thus remained without irrigation for 19 days by the end of monitoring. While the experiment followed this known watering history, no concurrent measurements of soil moisture or plant water status were collected.

3 Results

3.1 Model performance on training and validation data

The model evaluation revealed distinct performance patterns across training and validation datasets, as well as a substantial improvement over the previous model version V1 (Fig. 3). For the training dataset, we found a strong correspondence between the predicted and reference average leaf angles ($R^2 = 0.75$, $RMSE = 7.37^\circ$, $n = 3591$). The regression relationship indicated a minor overestimation of lower shallow angles and an underestimation of steeper angles. The AngleCam V2 performance assessment on the validation dataset resulted in an $R^2 = 0.62$ and $RMSE = 9.32^\circ$ ($n = 899$, Fig. 3b).

Comparison with AngleCam V1 on the same validation dataset revealed clear performance improvements of AngleCam V2. AngleCam V1 achieved $R^2 = 0.12$ with $RMSE = 14.19^\circ$ (Fig. 3c). Its predictions exhibited greater dispersion and stronger systematic bias, tending to predict intermediate average leaf angles.

In addition to average angle performance, the model's ability to capture the entire distribution shape was evaluated by analyzing bin-wise prediction errors (predicted minus reference probability) across 43 leaf angle bins (Fig. 4). For the training set, the mean error remained centered around zero for all angle bins, with a stable percentile band and only a minor negative dip in the 10th percentile between 10° and 25° (Fig. 4a). In the validation set, AngleCam V2 similarly maintained a mean error near zero, with a slight positive deviation between 40° and 60° , and a more pronounced negative dip in the 10th percentile at lower angles (10° – 25°) (Fig. 4b). By contrast, AngleCam V1 exhibited a pronounced angle-dependent bias on the same validation data, with underprediction at low angles (10° – 20°), increasing overprediction at higher angles, and a widening distribution spread toward extreme orientations (Fig. 4c).

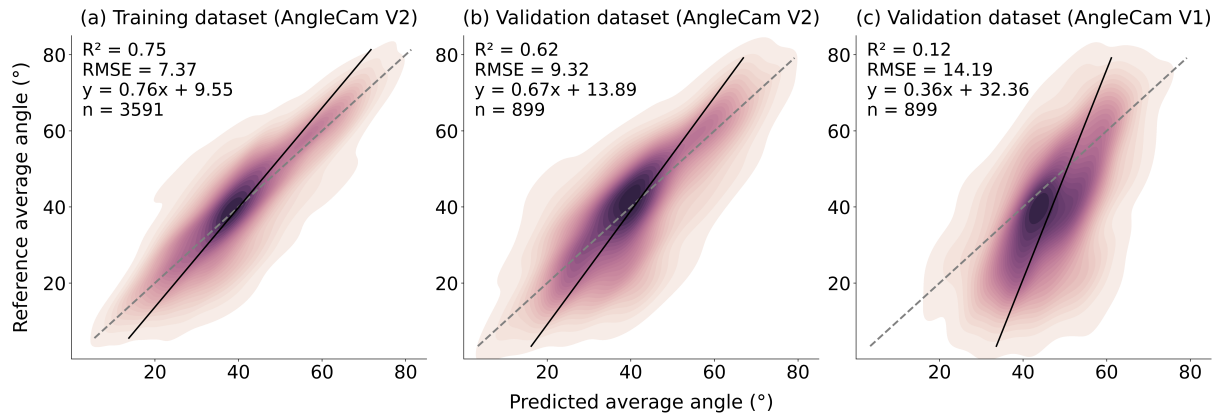


Figure 3: Comparison of AngleCam V1 and V2 performance across training and validation dataset. Kernel-density scatter plots show the joint distribution of predicted and reference average leaf angles ($0-90^\circ$) for (a) the training set of AngleCam V2, (b) the independent validation set of AngleCam V2, and (c) the same validation set evaluated with AngleCam V1. Colors represent point density (dark = high density), the dashed grey line marks the 1:1 relationship, and the solid black line is the ordinary-least-squares regression fit (equation inset). Insets additionally report the coefficient of determination (R^2), root-mean-square error (RMSE, degrees), and sample size (n).

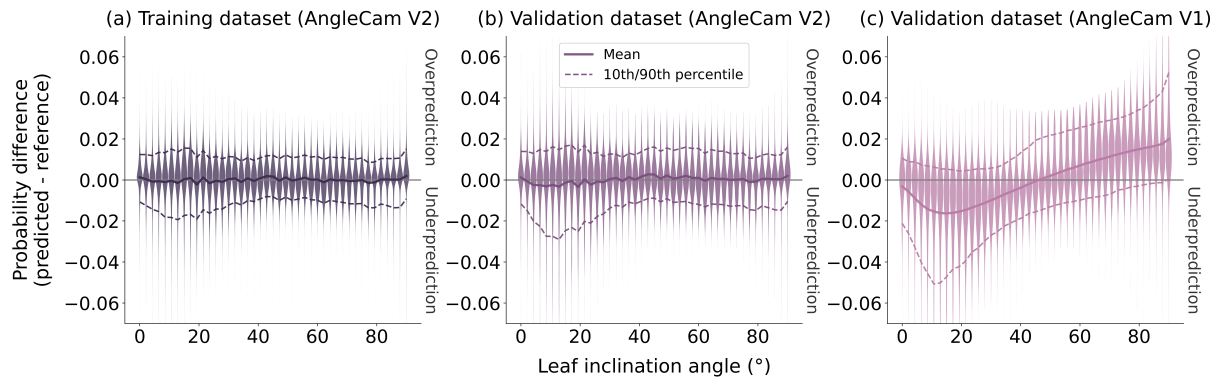


Figure 4: Bin-wise prediction errors for leaf inclination angle distributions are shown for: (a) the training dataset (AngleCam V2, $n = 3591$), (b) the validation dataset (AngleCam V2, $n = 899$), and (c) the validation dataset (AngleCam V1, $n = 899$). Each plot displays the error distribution (predicted minus reference probability) for every 2° angle bin. Solid lines show the mean error, while dashed lines represent the 10th and 90th percentiles. Positive values indicate overprediction and negative values underprediction.

3.2 Phylogenetic error assessment

We analyzed prediction errors across 100 genera to test whether model performance was systematically structured by phylogeny. Mean absolute errors ranged from around 1° to 15° , with the majority of genera falling between 4° and 10° (Fig. 5). The smallest errors were observed in genera such as *Pilea* and *Nothofagus*, while the largest occurred in *Arbutus* and *Olea* (see Table A1 in the Appendix for all genus-level MAE values). When mapped onto the phylogenetic tree, prediction errors showed no significant clustering. Pagel's λ was near zero ($\lambda = 7.41 \times 10^{-5}$), with no significant deviation from

the null hypothesis of $\lambda = 0$ (no phylogenetic structure, $p = 1.0$), indicating that no statistically significant phylogenetic structure was present in the residuals. Consistently, Moran's I showed weak positive autocorrelation ($I = 0.074$, $p = 0.23$), supporting the finding that model residuals were not significantly structured by evolutionary relatedness within the sampled 100 genera.

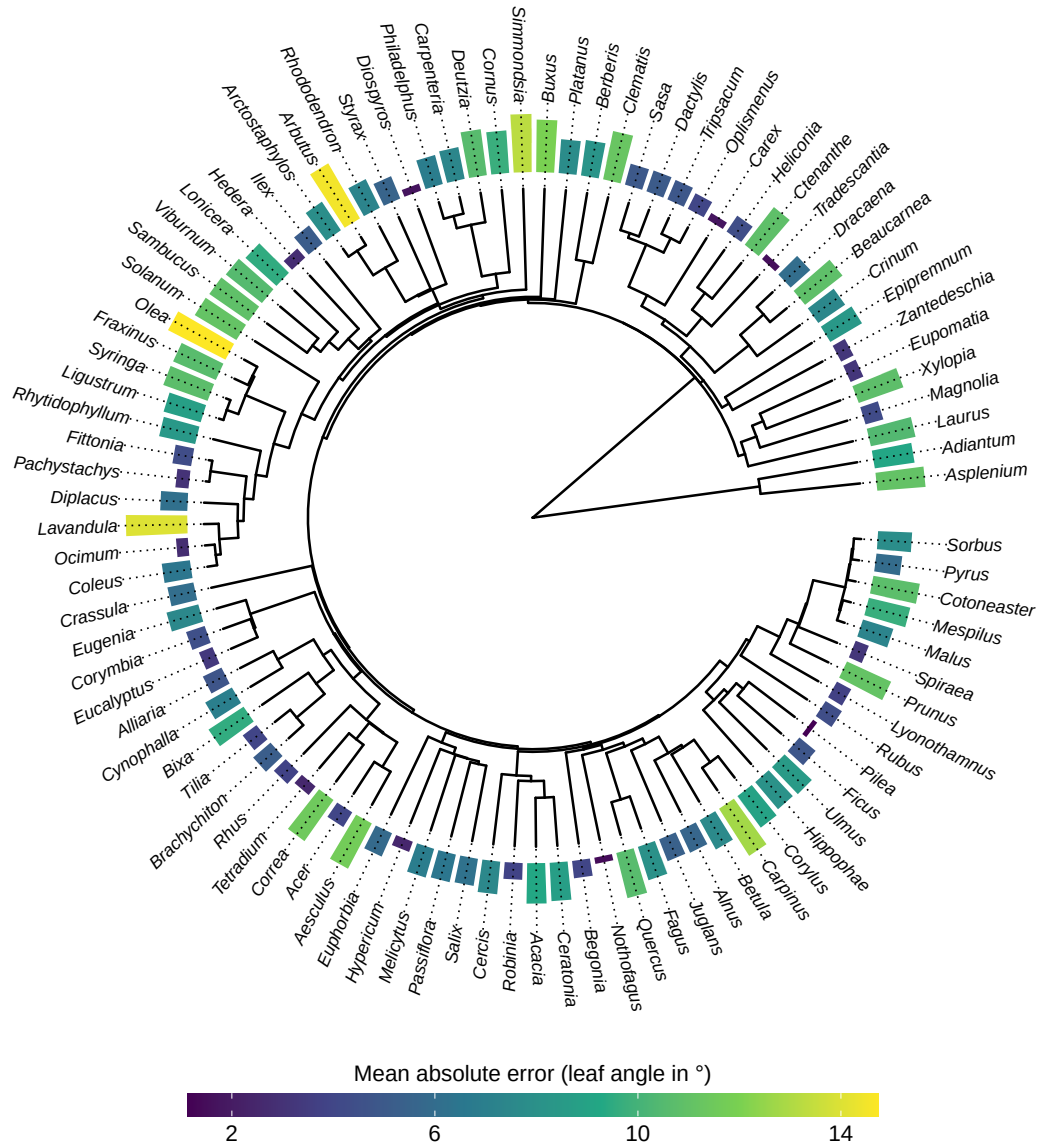


Figure 5: Phylogenetic distribution of AngleCam V2 prediction performance across 100 plant genera. Colored outer bars show residual per genus, computed for genera with at least three observations. Residuals show no significant clustering by clade. Pagel's λ is near zero ($\lambda = 7.41 \times 10^{-5}$, $p = 1.0$), indicating no phylogenetic structure. Moran's I shows weak autocorrelation ($I = 0.074$, $p = 0.23$). Together, these results indicate that there is no detectable phylogenetic autocorrelation in model errors.

3.3 Multitemporal model evaluation of AngleCam V2 with TLS-derived leaf inclination angle distributions

We validated AngleCam V2 predictions against independent TLS measurements using two indoor plant species (*Calathea ornata* and *Maranta leuconeura*) monitored continuously over nearly three days. Both species displayed pronounced diurnal leaf angle movements, which were captured by both the AngleCam and the TLS approach (Fig. 6).

For *Calathea ornata*, AngleCam predictions followed the TLS trends, but there was a systematic discrepancy throughout the monitoring period, most pronounced at lower average leaf angles (Fig. 6a). The range of average leaf angles obtained by the two methods was very similar, ranging from approximately 45° to 63° for AngleCam and 50° to 63° for TLS. The average leaf angles derived from the two methods yielded a high global correspondence ($R^2 = 0.61$). Disaggregated by modality, the Root Mean Square Error (RMSE) was 7.6° for daytime (RGB) and 3.6° for nighttime (NIR) imagery.

The results for *Maranta leuconeura* showed stronger agreement (Fig. 6b). AngleCam predictions closely matched TLS-based estimates across the entire time series, with average angles ranging from approximately 42° to 65° for both methods. The global correspondence was $R^2 = 0.75$. The model achieved close alignment across both modalities, with an RMSE of 3.5° for daytime (RGB) and 3.1° for nighttime (NIR) imagery.

3.4 Multitemporal model evaluation of AngleCam V2 under water limitation

We assessed AngleCam's sensitivity for tracking water limitation-induced changes in leaf angles by monitoring a non-watered individual of *Aglaonema commutatum* over an extended period of 14 days (Fig. 7a). The time series revealed a consistent increase in average leaf angle from approximately 43° to 48° (slope = 0.33°/day, $p < 0.001$, $n = 8783$), becoming especially evident after December 25. From this point onward, angles consistently exceeded the pre-December 25 baseline and showed an increased diurnal amplitude. The imagery (Fig. 7b) and corresponding leaf inclination angle distributions extracted at four time steps confirmed this trend, illustrating a gradual shift toward steeper angles (Fig. 7c). The overall shape of the distributions remained largely stable, with a slight tendency toward right skew, underscoring the robustness of AngleCam predictions under gradually changing physiological conditions.

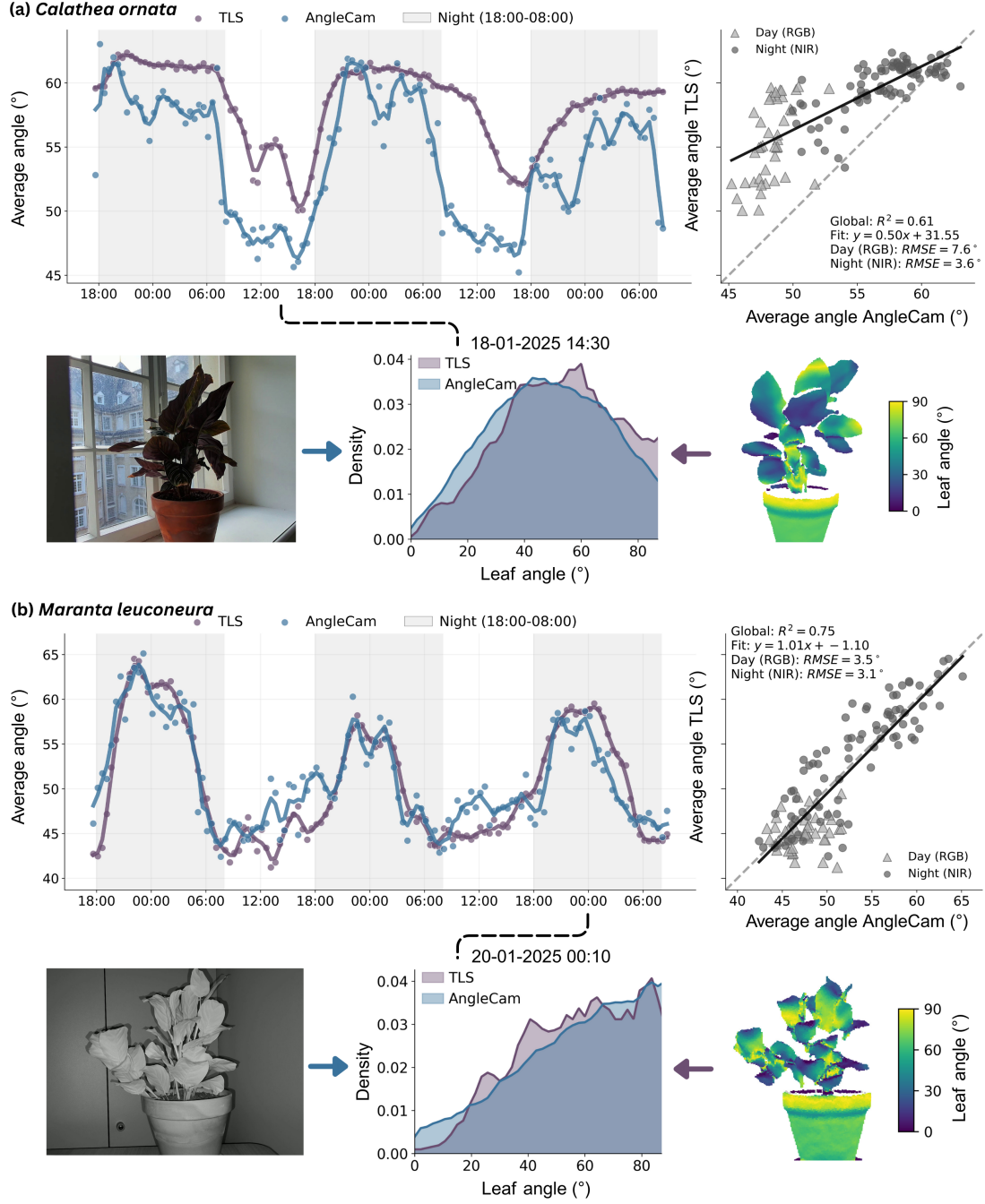


Figure 6: Multitemporal evaluation of AngleCam V2 predictions against TLS measurements for two plant species known for strong diurnal leaf movements. Time series show average leaf inclination angles over nearly three days for (a) *Calathea ornata* (Global $R^2 = 0.61$) and (b) *Maranta leuconeura* (Global $R^2 = 0.75$), with both RGB and near-infrared night-vision imagery captured every 30 minutes. Scatter plots (right panels) distinguish between daytime (RGB) and nighttime (NIR) performance, reporting separate Root Mean Square Errors ($RMSE$) for each modality. Distribution plots show example leaf inclination angle distributions at specific timestamps, with images illustrating plant appearance and 3D point cloud visualizations of the same setting. An interactive visualization can be accessed at [GitHub](#).

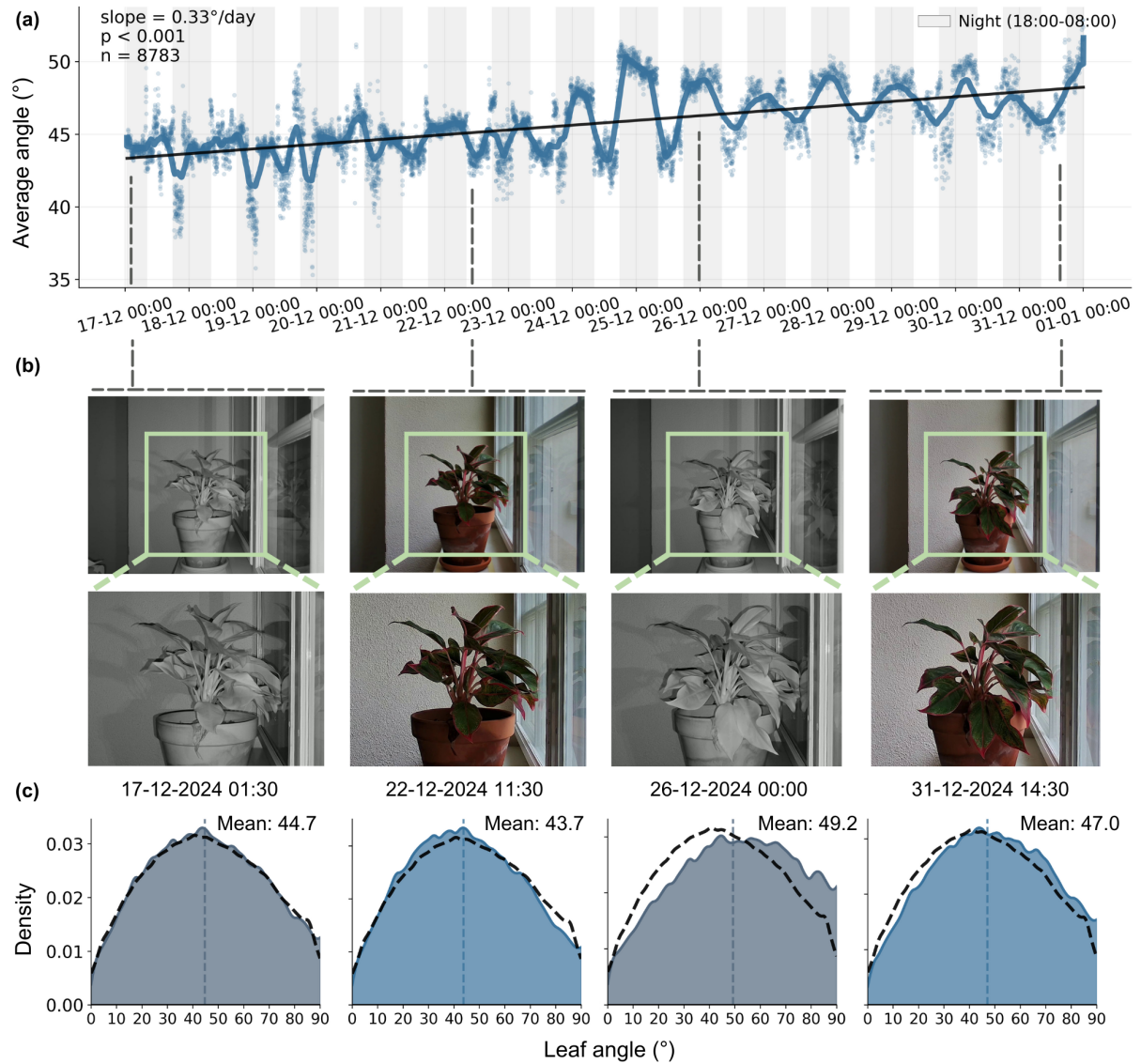


Figure 7: Tracking of leaf angle changes in *Aglaonema commutatum* induced by water limitation (14 days without irrigation). (a) Time series showing a gradual increase in average leaf angles from approximately 43° to 48° (slope = 0.33°/day, $p < 0.001$, $n = 8783$ measurements), with pronounced diurnal oscillations and increased amplitude compared to the pre-December 25 baseline. The black line shows a linear trend with smoothed daily patterns overlaid. (b) Sequential NIR and RGB imagery at four time points showing visual progression of water limitation effects (c) Corresponding leaf inclination angle distributions at the same time points, demonstrating the gradual shift toward steeper angles (dashed line shows average distribution of all 8783 samples) while maintaining overall distribution shape. The results demonstrate the potential of AngleCam for revealing gradual physiological changes and its application for long-term drought monitoring.

4 Discussion

4.1 Model performance on training and validation data

Our evaluation revealed that AngleCam V2 achieved stronger predictive performance compared to its predecessor (AngleCam V1). When evaluated on their respective original validation datasets, AngleCam V1 showed higher performance metrics (validation: $R^2 = 0.84$, $RMSE = 6.13^\circ$; Kattenborn et al. (2022)) compared to AngleCam V2 (validation: $R^2 = 0.62$, $RMSE = 9.32^\circ$). However, when both models were evaluated on the new, more globally representative dataset, AngleCam V2 demonstrated substantially superior generalization, with an R^2 of 0.62, compared to AngleCam V1's R^2 of 0.12. For this comparison, AngleCam V1 was applied as originally trained without any retraining on the expanded dataset. The increased performance can be attributed to the fact that AngleCam V2 was trained on an expanded dataset with considerably greater species diversity and scene heterogeneity than the first version. This expansion deliberately introduced more variation in leaf morphologies, growth forms, environmental conditions, cameras, and data acquisition settings, enhancing the model's ability to generalize across diverse ecological contexts and data acquisition scenarios.

The expanded dataset inevitably introduced additional label noise as well as variability in image features unrelated to actual leaf angle differences, both of which can bias a model (Belkin et al., 2019; Geman et al., 1992). While large training sets can help mitigate such effects (Rolnick et al., 2017), complex deep learning models may still memorize noise or spurious patterns instead of extracting meaningful signals. Given our dataset size and the deep transformer model (DINOv2), this risk would manifest as low bias but high variance if regularization were insufficient. To address this, we implemented stronger regularization strategies compared to AngleCam V1. A dropout rate of 0.4 in the regression head randomly deactivated 40% of units during each forward pass, facilitating the model to rely on robust features rather than memorizing training-specific patterns. Weight decay provided additional regularization by penalizing large parameter values, encouraging smoother angle approximations (Loshchilov & Hutter, 2017; Zhang et al., 2018). We also transitioned from mean squared error (MSE) to Huber loss, which treats outliers more robustly by being quadratic for small errors and linear for large errors (Sun et al., 2020). Together, these techniques facilitated the model to extract meaningful signals from the expanded dataset while ensuring the generalization of the model over the more variable scene conditions and taxa.

The validation results revealed a tendency where low leaf angles were overestimated and high angles were underestimated. We assume that this bias, also known as regression to the mean bias (Barnett et al., 2005), emerges from the training data distribution, which contains abundant samples between 35°-50° but relatively few extreme cases below 20° or above 60°. The bin-wise error analysis (Fig. 4) further clarifies this bias, where AngleCam V2 maintains a mean error near zero across the distribution range but shows an increased spread in the 10th percentile at lower angles (10°–25°) in the validation set. This pattern reflects greater uncertainty when the model encounters these less frequently represented orientations. Nevertheless, the stability of the mean error across all bins represents a major advancement over AngleCam V1, which exhibited a pronounced, systematic angle-dependent bias and increasing variance toward extreme orientations.

Consequently, AngleCam V2 achieves its highest performance for intermediate LIAD types, such as spherical or plagiophile, while demonstrating reduced accuracy for extreme planophile or erectophile canopy structures. Several approaches could address this imbalance, such as loss weighting to give more importance to rare, extreme angles during training, or a weighted training sampler that increases the selection frequency of underrepresented angle ranges. However, such interventions require careful implementation to avoid introducing artificial biases or unstable prediction behavior (Barnett et al., 2005). AngleCam development remains an ongoing process, and future versions will focus on expanding training diversity in extreme angle scenarios.

4.2 Phylogenetic error assessment

The residual analysis revealed no statistically significant phylogenetic signal in prediction errors (Fig. 5), indicating that AngleCam V2 learned general principles of leaf angle estimation that transfer across taxonomic and evolutionary lineages. The absence of phylogenetic autocorrelation suggests the model captures fundamental geometric relationships rather than taxon-specific visual features. Pagel's λ was near zero ($\lambda = 7.41 \times 10^{-5}$, $p = 1.0$), indicating that no statistically significant phylogenetic structure was detected in the prediction errors. Similarly, Moran's I showed only weak, non-significant autocorrelation ($I = 0.074$, $p = 0.23$), providing no evidence of systematic phylogenetic bias within the current dataset.

The substantial representation of *Tilia*, *Acer*, *Quercus*, and *Fagus* in our training dataset (see also Fig. 1, Tab. A1) results from continuous time-series monitoring, providing extensive environmental and phenological coverage across these genera. Notably, the phylogenetic analysis demonstrates that prediction errors are not systematically lower for such well-represented genera compared to those

with fewer training samples. Although more balanced taxonomic representation would be ideal for future datasets, the temporal richness within these dominant genera added training diversity across environmental conditions without compromising model generalization.

This supports applying AngleCam to new species, given that their broader taxonomic groups are reasonably represented in the training data. Still, generalization has its limits: performance may degrade for clades that are sparsely or not represented at all, and clade-specific biases may emerge in such cases. Overall, while the absence of a detectable phylogenetic signal suggests that within-genus variation in leaf morphology and growth dynamics is reasonably captured by our training set, we acknowledge that the limited sample size (100 genera) may mask modest clade-specific biases that could emerge with larger sample sizes or higher taxonomic resolution.

4.3 Evaluating AngleCam on representing leaf angle diurnal dynamics

Our controlled testing experiments with TLS demonstrated AngleCam V2's ability to track temporal leaf angle dynamics across day-night cycles. As demonstrated for *Maranta leuconeura* and *Calathea ornata*, the model captured natural circadian leaf movements despite never being explicitly trained on labeled NIR imagery. This capability emerged from our dual-modality training strategy, combining grayscale conversion with pseudo-NIR generation as a proxy for actual night-vision data. This workaround was required by the high labor intensity of manual leaf annotation. While our RGB dataset is extensive, replicating this sample size with manually labeled NIR imagery was not feasible within the scope of the current study.

As shown by the TLS comparison, this strategy proved effective for the tested species. For *Calathea ornata* (Global $R^2 = 0.61$), the model achieved an RMSE of 7.6° for daytime (RGB) and 3.6° for nighttime (NIR) imagery. For *Maranta leuconeura* (Global $R^2 = 0.75$), performance was even more consistent, with an RMSE of 3.5° for RGB and 3.1° for NIR imagery. Although these results indicate that the model can generalize to real night-vision environments, the pseudo-NIR pipeline remains an approximation with inherent physical limitations. Specifically, pseudo-NIR images retain daytime ambient shadows and may exhibit depth-based overexposure or less plausible illumination falloff compared to true NIR backscatter (a detailed characterization of these discrepancies and visual comparisons to real NIR imagery are provided in Appendix Section A.2). As a result, the synthetic labels do not serve as a perfect substitute for true NIR annotations. Nevertheless, this method effectively leverages the existing RGB dataset to enable 24-hour monitoring, avoiding the constraints of extensive manual NIR labeling. Future iterations incorporating large-scale, independently labeled

NIR datasets will be essential to further improve night-vision accuracy and to address sensor-specific discrepancies.

The systematic underestimation observed in *Calathea ornata* relative to TLS may result from sampling biases inherent to laser scanning, where vertical leaves intercept more pulses and can inflate angle estimates (Jiang et al., 2021). The use of 1 cm voxel downsampling and the spherical distribution of *Maranta leuconeura* minimized this effect. However, the fixed TLS position, which was necessary for continuous 30-minute sampling over several days, likely introduced some occlusion-related bias. Future studies could address this by employing multiple fixed scanners or, for shorter intensive campaigns, by conducting sequential multi-view scans.

Species-specific performance variations may also indicate current model generalization limits. The superior results for *Maranta leuconeura* likely reflect the presence of morphologically similar taxa in the training set (*Ctenanthe burle-marxii*, $n = 7$), which share comparable leaf architecture. In contrast, *Calathea ornata* lacked such representatives beyond the experimental samples themselves. This indicates that model performance can depend on exposure to similar species during training. When deploying AngleCam for experimental monitoring, particularly for species that are not adequately represented by morphologically similar taxa in the training dataset, we recommend acquiring at least 15-20 labeled images from the target scene for model calibration. We provide an open-source labeling tool, pre-trained model weights, and a retraining framework (GitHub), enabling a flexible adaptation of AngleCam to particular experimental conditions.

4.4 Application of AngleCam in future research

AngleCam V2 addresses limitations in low-cost and scalable leaf angle monitoring by enabling continuous day-night observations, opening up new possibilities for ecosystem research. The method's compatibility with standard time-lapse cameras, low hardware costs relative to laser scanning systems, and minimal field maintenance requirements make it a valuable asset for long-term monitoring networks (e.g., flux-towers, biodiversity experiments, PhenoCam). Integration with eddy covariance measurements could provide insights into how diurnal and seasonal leaf movements are coupled with carbon fluxes and ecosystem-atmosphere interactions (Baldocchi, 2020). Given that variable vertical leaf angle profiles impact light interception and canopy energy balance, continuous monitoring could enhance our understanding of ecosystem responses to environmental changes (Yang et al., 2023).

Despite recent advances, knowledge about factors driving leaf angle dynamics remains limited. While diurnal patterns and vertical LIAD gradients are documented (Ehleringer & Forseth, 1980; Kao

& Forseth, 1992; Liu et al., 2019; Pisek et al., 2022), the relative importance of light availability versus water stress in controlling these dynamics is poorly understood (Kattenborn et al., 2022; Yang et al., 2023). AngleCam's high temporal resolution and multi-height deployment capability can help to address this knowledge gap. Over the course of the 14-day experiment in this study, progressive soil drainage (with the plant remaining unwatered for a total of 19 days) revealed that the leaf inclination angle distributions of *Aglaonema commutatum* responded rapidly to water limitation. AngleCam's sensitivity to changes in plant condition enabled quantification of these dynamics, demonstrating its potential for tracking plant responses to climate extremes, such as atmospheric or soil drought. Interestingly, we observed two trends: a steady increase in average leaf angles and a progressive amplification of diurnal variation. These findings align with a previous study on five temperate tree species, which showed that leaf angle dynamics respond both in absolute values and in their temporal variability (Kattenborn et al., 2022, 2024). In the future, simultaneous monitoring of plant physiological parameters (e.g., sap flow, stem water potential), environmental conditions (e.g., temperature, light intensity), and ecosystem fluxes (e.g., carbon and water exchange) alongside leaf angle dynamics could help uncover the mechanistic drivers of these dynamics and support the use of leaf angle as an indicator of plant physiological status. However, given that species respond differently to environmental stresses, such applications would likely require species-specific training and calibration to establish reliable relationships between leaf angle changes and physiological conditions.

Past and current deployments of AngleCam reflect its potential for ecological and plant physiological research. Installations at the Leipzig Canopy Crane (Richter et al., 2021), the MyDiv experiment (Ferlian et al., 2018), and the ArboFun experiment (Kretz et al., 2025), all part of the German Centre for Integrative Biodiversity Research (iDiv), explore how structural, taxonomic, and functional diversity shape leaf angle dynamics for temperate tree species. In addition, AngleCam installations at the ECOSENSE site in Ettenheim (Werner et al., 2024) and at the Hartheim Research Station in Germany (Integrated Carbon Observation System (ICOS) flux-tower site) cover multiple individuals of temperate broad-leaved tree species. All of the above-mentioned installations capture leaf movements alongside comprehensive physiological and environmental measurements. Expanding such integrated monitoring to other research infrastructures and biomes could significantly enhance our understanding of leaf angle dynamics. Existing PhenoCam and eddy covariance flux tower networks (Brown et al., 2016; Richardson et al., 2018) offer ideal platforms for scaling up AngleCam installations and linking leaf angle variability to plant and ecosystem function at broader spatial and temporal scales. However, accurate LIAD estimation with AngleCam depends on a horizontal camera orientation. Successful

integration into these networks would require either identifying existing cameras with appropriate horizontal geometries or installing dedicated sensors that conform to the required acquisition geometry. Similarly, the distance to the canopy influences the effective spatial resolution. While the model training covered typical detection ranges (< 2 m), images must be taken close enough to clearly resolve individual leaves.

When deploying AngleCam for ecological monitoring, it is essential to consider the spatial representativeness of the imagery. A single horizontal photograph captures only the foliage within the camera's field of view, typically representing a local portion of the canopy at a distance of approximately 1 to 2 meters. As a result, AngleCam reports the LIAD of the visible leaf surfaces from that specific perspective. Because leaf orientation may display directional patterns or canopy asymmetries, such as those caused by phototropism or prevailing winds, LIAD estimates can vary depending on the azimuth from which photographs are taken. To achieve a representative characterization of the entire canopy structure, camera placement should account for these spatial variations. We recommend spatial replication, including image acquisition from multiple azimuths and different heights, to capture a comprehensive distribution of leaf angles.

Beyond ecological and physiological monitoring at local scales, understanding leaf angle dynamics is also critical for interpreting vegetation signals in satellite-based Earth observation. Leaf angles affect the scattering and absorption of light within plant canopies and thereby influence radiative transfer processes (Hase et al., 2022; Kattenborn et al., 2019). Temporal changes in vertical leaf angles can significantly alter reflectance signals and systematically confound widely used vegetation indices, such as kNDVI (Hase et al., 2022; Kattenborn et al., 2024) or fluorescence signals (Jablonski et al., 2025). These changes are often not random noise but reflect structured responses to environmental drivers like temperature or radiation (Kattenborn et al., 2022). Thus, unresolved leaf angle dynamics may obscure or mimic other physiological variability, such as vegetation productivity changes when observed from space. At the same time, this sensitivity to environmental cues means that remotely sensed variations in canopy reflectance may contain embedded signals of leaf angle movements. AngleCam can thus contribute to disentangling these effects by providing temporally resolved parameters in radiative transfer modeling or by acting as a covariate to interpret better satellite observations of vegetation stress (Hase et al., 2022; Jablonski et al., 2025; Kattenborn et al., 2024).

The model's demonstrated generalization across different scene conditions, growth forms, and taxonomic and phylogenetic lineages can unlock large-scale applications in concert with citizen science platforms. Particularly, citizen science apps for plant species recognition, such as iNaturalist or

599 PI@ntNet, acquire millions of geolocation plant images each year, many of which are acquired with
600 sufficient distance and camera perspectives compatible with AngleCam (Garcin et al., 2021; Mason
601 et al., 2025; Soltani et al., 2024). Previous studies have already highlighted the value of citizen science
602 data for uncovering plant functional traits or phenology across spatiotemporal scales (Mora et al.,
603 2024; Schiller et al., 2021; Wolf et al., 2022). Accordingly, citizen science data may also represent
604 an unprecedented data treasure to uncover global patterns in leaf inclination angle distributions at
605 the species and community level, biogeographic variations in leaf positioning strategies, or seasonal
606 adaptation patterns across biomes. This approach would substantially expand our understanding of
607 how leaf angles are organized within and across species worldwide, potentially uncovering previously
608 unknown global patterns in plant structural adaptation (Li & Fang, 2024).

5 Conclusions

We introduced AngleCam V2, an enhanced deep learning model for estimating leaf inclination angle distributions from RGB and night-vision NIR imagery. Expanding the training dataset to include diverse species and canopy structures, combined with grayscale and pseudo-NIR transformations that leverage extensive existing RGB labels, has improved generalization and temporal robustness, including compatibility with nighttime imagery.

Testing against terrestrial laser scanning (TLS) demonstrated the capability of AngleCam V2 to track diurnal leaf movements across day and night imagery. As shown in a water limitation experiment, the model also captured multi-day trends in leaf angle variability. Additionally, phylogenetic analysis across 100 genera revealed no statistically significant systematic bias in predictions, supporting the model's generalizability across a wide range of plant species.

A central feature of AngleCam is its broad applicability: it can be used with standard imagery from smartphones and time-lapse cameras to PhenoCam networks, provided cameras are oriented horizontally. The method requires no specialized hardware, making it easy to deploy in field, lab, or citizen science settings. This enables both high-frequency monitoring of leaf angle dynamics over time and flexible snapshot-based assessments using handheld devices. AngleCam V2 thus provides a scalable, accessible, and open-source tool for monitoring leaf angle dynamics across time, taxa, and environments.

6 References

References

- Apelt, F., Breuer, D., Olas, J. J., Annunziata, M. G., Flis, A., Nikoloski, Z., Kragler, F., & Stitt, M. (2017). Circadian, carbon, and light control of expansion growth and leaf movement. *Plant Physiology*, 174(3), 1949–1968.
- Bailey, B. N., & Mahaffee, W. F. (2017). Rapid measurement of the three-dimensional distribution of leaf orientation and the leaf angle probability density function using terrestrial lidar scanning. *Remote Sensing of Environment*, 194, 63–76.
- Baldocchi, D. D. (2020). How eddy covariance flux measurements have contributed to our understanding of global change biology. *Global change biology*, 26(1), 242–260.

- Barnett, A. G., Van Der Pols, J. C., & Dobson, A. J. (2005). Regression to the mean: What it is and how to deal with it. *International journal of epidemiology*, 34(1), 215–220.
- Belkin, M., Hsu, D., Ma, S., & Mandal, S. (2019). Reconciling modern machine-learning practice and the classical bias–variance trade-off. *Proceedings of the National Academy of Sciences*, 116(32), 15849–15854.
- Bernotas, G., Scorza, L. C., Hansen, M. F., Hales, I. J., Halliday, K. J., Smith, L. N., Smith, M. L., & McCormick, A. J. (2019). A photometric stereo-based 3d imaging system using computer vision and deep learning for tracking plant growth. *GigaScience*, 8(5), giz056.
- Biskup, B., Scharr, H., Schurr, U., & Rascher, U. (2007). A stereo imaging system for measuring structural parameters of plant canopies. *Plant, cell & environment*, 30(10), 1299–1308.
- Braghiere, R. K., Wang, Y., Doughty, R., Sousa, D., Magney, T., Widlowski, J.-L., Longo, M., Bloom, A. A., Worden, J., Gentine, P., et al. (2021). Accounting for canopy structure improves hyperspectral radiative transfer and sun-induced chlorophyll fluorescence representations in a new generation earth system model. *Remote Sensing of Environment*, 261, 112497.
- Brown, T. B., Hultine, K. R., Steltzer, H., Denny, E. G., Denslow, M. W., Granados, J., Henderson, S., Moore, D., Nagai, S., SanClements, M., et al. (2016). Using phenocams to monitor our changing earth: Toward a global phenocam network. *Frontiers in Ecology and the Environment*, 14(2), 84–93.
- Cabon, Y., Murray, N., & Humenberger, M. (2020). Virtual kitti 2. *arXiv preprint arXiv:2001.10773*.
- Darwin, C., & Darwin, F. (1883). *The power of movement in plants*. D. Appleton; Company.
- Dechant, B., Ryu, Y., Badgley, G., Zeng, Y., Berry, J. A., Zhang, Y., Goulas, Y., Li, Z., Zhang, Q., Kang, M., et al. (2020). Canopy structure explains the relationship between photosynthesis and sun-induced chlorophyll fluorescence in crops. *Remote Sensing of Environment*, 241, 111733.
- Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., et al. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*.
- Ehleringer, J., & Forseth, I. (1980). Solar tracking by plants. *Science*, 210(4474), 1094–1098.
- Ehleringer, J., Comstock, J., Tenhunen, J., Catarino, F., Lange, O., & Oechel, W. (1987). Plant response to stress.
- Ferlian, O., Cesarz, S., Craven, D., Hines, J., Barry, K. E., Bruelheide, H., Buscot, F., Haider, S., Heklau, H., Herrmann, S., et al. (2018). Mycorrhiza in tree diversity–ecosystem function

relationships: Conceptual framework and experimental implementation. *Ecosphere*, 9(5), e02226.

Garcin, C., Joly, A., Bonnet, P., Lombardo, J.-C., Affouard, A., Chouet, M., Servajean, M., Lorieul, T., & Salmon, J. (2021). PI@ ntnet-300k: A plant image dataset with high label ambiguity and a long-tailed distribution. *NeurIPS 2021-35th Conference on Neural Information Processing Systems*.

Geldhof, B., Pattyn, J., Eyland, D., Carpentier, S., & Van de Poel, B. (2021). A digital sensor to measure real-time leaf movements and detect abiotic stress in plants. *Plant Physiology*, 187(3), 1131–1148.

Geman, S., Bienenstock, E., & Doursat, R. (1992). Neural networks and the bias/variance dilemma. *Neural computation*, 4(1), 1–58.

Gittleman, J. L., & Kot, M. (1990). Adaptation: Statistics and a null model for estimating phylogenetic effects. *Systematic Zoology*, 39(3), 227–241.

Goel, N. S., & Strebel, D. E. (1984). Simple beta distribution representation of leaf orientation in vegetation canopies 1. *Agronomy Journal*, 76(5), 800–802.

Hase, N., Doktor, D., Rebmann, C., Dechant, B., Mollenhauer, H., & Cuntz, M. (2022). Identifying the main drivers of the seasonal decline of near-infrared reflectance of a temperate deciduous forest. *Agricultural and Forest Meteorology*, 313, 108746.

Hawkins, B. A. (2012). Eight (and a half) deadly sins of spatial analysis. *Journal of Biogeography*, 39(1), 1–9.

Hendrycks, D., & Gimpel, K. (2016). Gaussian error linear units (gelus). *arXiv preprint arXiv:1606.08415*.

Hikosaka, K., & Hirose, T. (1997). Leaf angle as a strategy for light competition: Optimal and evolutionarily stable light-extinction coefficient within a leaf canopy. *Ecoscience*, 4(4), 501–507.

Jablonski, A. D., Li, R., Kim, J., Lerda, M., Petras, C., & Yang, X. (2025). Spatial patterns of leaf angle distribution covary with canopy fluorescence yield, reflectance indices, and leaf chlorophyll content, in a mixed temperate forest. *Remote Sensing of Environment*, 330, 114996.

Jiang, H., Hu, R., Yan, G., Cheng, S., Li, F., Qi, J., Li, L., Xie, D., & Mu, X. (2021). Influencing factors in estimation of leaf angle distribution of an individual tree from terrestrial laser scanning data. *Remote Sensing*, 13(6), 1159.

Jolliffe, I. T. (2002). *Principal component analysis* (2nd). Springer-Verlag.

- Kamilar, J. M., & Cooper, N. (2013). Phylogenetic signal in primate behaviour, ecology and life history. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 368(1618), 20120341.
- Kao, W.-Y., & Forseth, I. (1992). Diurnal leaf movement, chlorophyll fluorescence and carbon assimilation in soybean grown under different nitrogen and water availabilities. *Plant, Cell & Environment*, 15(6), 703–710.
- Kattenborn, T., Fasnacht, F. E., & Schmidtlein, S. (2019). Differentiating plant functional types using reflectance: Which traits make the difference? *Remote Sensing in Ecology and Conservation*, 5(1), 5–19.
- Kattenborn, T., Richter, R., Guimarães-Steinicke, C., Feilhauer, H., & Wirth, C. (2022). Anglecam: Predicting the temporal variation of leaf angle distributions from image series with deep learning. *Methods in Ecology and Evolution*, 13(11), 2531–2545.
- Kattenborn, T., Wieneke, S., Montero, D., Mahecha, M. D., Richter, R., Guimarães-Steinicke, C., Wirth, C., Ferlian, O., Feilhauer, H., Sachsenmaier, L., et al. (2024). Temporal dynamics in vertical leaf angles can confound vegetation indices widely used in earth observations. *Communications Earth & Environment*, 5(1), 550.
- Kretz, L., Schnabel, F., Richter, R., Raabgrund, A., Kattge, J., Andraczek, K., Kahl, A., Kuenne, T., & Wirth, C. (2025). Functional traits explain growth response to successive hotter droughts across a wide set of common and future tree species in europe. *Plant Biology*.
- Leuzinger, S., & Körner, C. (2007). Tree species diversity affects canopy leaf temperatures in a mature temperate forest. *Agricultural and forest meteorology*, 146(1-2), 29–37.
- Li, S., & Fang, H. (2024). Mapping global leaf inclination angle (lia) based on field measurement data. *Earth System Science Data Discussions*, 2024, 1–29.
- Liu, J., Skidmore, A. K., Wang, T., Zhu, X., Premier, J., Heurich, M., Beudert, B., & Jones, S. (2019). Variation of leaf angle distribution quantified by terrestrial lidar in natural european beech forest. *ISPRS journal of photogrammetry and remote sensing*, 148, 208–220.
- Loshchilov, I., & Hutter, F. (2017). Decoupled weight decay regularization. *arXiv preprint arXiv:1711.05101*.
- Mason, B. M., Mesaglio, T., Barratt Heitmann, J., Chandler, M., Chowdhury, S., Gorta, S. B. Z., Grattarola, F., Groom, Q., Hitchcock, C., Hoskins, L., et al. (2025). Inaturalist accelerates biodiversity research. *BioScience*, biae104.

- Mora, K., Rzanny, M., Wäldchen, J., Feilhauer, H., Kattenborn, T., Kraemer, G., Mäder, P., Svidzinska, D., Wolf, S., & Mahecha, M. D. (2024). Macrophenological dynamics from citizen science plant occurrence data. *Methods in Ecology and Evolution*, 15(8), 1422–1437.
- Morales-Castilla, I., Davies, T., Legault, G., Buonaiuto, D., Chamberlain, C. J., Ettinger, A. K., Garner, M., Jones, F. A., Loughnan, D., Pearse, W. D., et al. (2024). Phylogenetic estimates of species-level phenology improve ecological forecasting. *Nature Climate Change*, 14(9), 989–995.
- Moran, P. A. (1950). Notes on continuous stochastic phenomena. *Biometrika*, 37(1/2), 17–23.
- Mullen, J. L., Weinig, C., & Hangarter, R. P. (2006). Shade avoidance and the regulation of leaf inclination in arabidopsis. *Plant, Cell & Environment*, 29(6), 1099–1106.
- Müller-Linow, M., Pinto-Espinosa, F., Scharr, H., & Rascher, U. (2015). The leaf angle distribution of natural plant populations: Assessing the canopy with a novel software tool. *Plant methods*, 11, 1–16.
- Murithi, C. M., Pisek, J., Schraik, D., Bailey, B. N., Liu, J., Stovall, A. E., Vicari, M. B., Zheng, G., & Skidmore, A. (2025). Intercomparison of methods for estimating leaf inclination angle distribution with terrestrial lidar for broadleaf tree species. *New Phytologist*.
- Niinemets, Ü. (2010). A review of light interception in plant stands from leaf to canopy in different plant functional types and in species with varying shade tolerance. *Ecological Research*, 25, 693–714.
- Oquab, M., Darcet, T., Moutakanni, T., Vo, H., Szafraniec, M., Khalidov, V., Fernandez, P., Haziza, D., Massa, F., El-Nouby, A., et al. (2023). Dinov2: Learning robust visual features without supervision. *arXiv preprint arXiv:2304.07193*.
- Pagel, M. (1999). Inferring the historical patterns of biological evolution. *Nature*, 401(6756), 877–884.
- Pharr, M., Jakob, W., & Humphreys, G. (2023). *Physically based rendering: From theory to implementation*. MIT Press.
- Pisek, J., Diaz-Pines, E., Matteucci, G., Noe, S., & Rebmann, C. (2022). On the leaf inclination angle distribution as a plant trait for the most abundant broadleaf tree species in europe. *Agricultural and Forest Meteorology*, 323, 109030.
- Pisek, J., Ryu, Y., & Alikas, K. (2011). Estimating leaf inclination and g-function from leveled digital camera photography in broadleaf canopies. *Trees*, 25, 919–924.
- PlantNet. (2025). PlantNet Plant Identification [Accessed: 2025-02-18]. <https://identify.plantnet.org/de>

- Puglielli, G., Bricca, A., Chelli, S., Petruzzellis, F., Acosta, A. T., Bacaro, G., Beccari, E., Bernardo, L., Bonari, G., Bolpagni, R., et al. (2024). Intraspecific variability of leaf form and function across habitat types. *Ecology Letters*, 27(3), e14396.
- Puglielli, G., Redondo-Gómez, S., Gratani, L., & Mateos-Naranjo, E. (2017). Highlighting the differential role of leaf paraheliotropism in two mediterranean cistus species under drought stress and well-watered conditions. *Journal of Plant Physiology*, 213, 199–208.
- Qi, J., Xie, D., Li, L., Zhang, W., Mu, X., & Yan, G. (2019). Estimating leaf angle distribution from smartphone photographs. *IEEE Geoscience and Remote Sensing Letters*, 16(8), 1190–1194.
- Richardson, A. D., Hufkens, K., Milliman, T., Aubrecht, D. M., Chen, M., Gray, J. M., Johnston, M. R., Keenan, T. F., Klosterman, S. T., Kosmala, M., et al. (2018). Tracking vegetation phenology across diverse north american biomes using phenocam imagery. *Scientific data*, 5(1), 1–24.
- Richter, R., Ballasus, H., Engelmann, R. A., Zielhofer, C., Sanaei, A., & Wirth, C. (2022). Tree species matter for forest microclimate regulation during the drought year 2018: Disentangling environmental drivers and biotic drivers. *Scientific Reports*, 12(1), 17559.
- Richter, R., Hutengs, C., Wirth, C., Bannehr, L., & Vohland, M. (2021). Detecting tree species effects on forest canopy temperatures with thermal remote sensing: The role of spatial resolution. *Remote Sensing*, 13(1), 135.
- Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillerá-Arroita, G., Hauenstein, S., Lahoz-Monfort, J. J., Schröder, B., Thuiller, W., et al. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, 40(8), 913–929.
- Rolnick, D., Veit, A., Belongie, S., & Shavit, N. (2017). Deep learning is robust to massive label noise. *arXiv preprint arXiv:1705.10694*.
- Rusu, R. B., Blodow, N., Marton, Z. C., & Beetz, M. (2008). Aligning point cloud views using persistent feature histograms. *2008 IEEE/RSJ international conference on intelligent robots and systems*, 3384–3391.
- Ryu, Y., Nilson, T., Kobayashi, H., Sonnentag, O., Law, B. E., & Baldocchi, D. D. (2010). On the correct estimation of effective leaf area index: Does it reveal information on clumping effects? *Agricultural and Forest Meteorology*, 150(3), 463–472.
- Sastry, A., Guha, A., & Barua, D. (2018). Leaf thermotolerance in dry tropical forest tree species: Relationships with leaf traits and effects of drought. *AoB Plants*, 10(1), plx070.

- Schiller, C., Schmidtlein, S., Boonman, C., Moreno-Martinez, A., & Kattenborn, T. (2021). Deep learning and citizen science enable automated plant trait predictions from photographs. *Scientific Reports*, *11*(1), 16395.
- Soltani, S., Ferlian, O., Eisenhauer, N., Feilhauer, H., & Kattenborn, T. (2024). From simple labels to semantic image segmentation: Leveraging citizen science plant photographs for tree species mapping in drone imagery. *Biogeosciences*, *21*(11), 2909–2935.
- Stovall, A. E., Masters, B., Fatoyinbo, L., & Yang, X. (2021). Tisleaf: Automatic leaf angle estimates from single-scan terrestrial laser scanning. *New Phytologist*, *232*(4), 1876–1892.
- Sun, Q., Zhou, W.-X., & Fan, J. (2020). Adaptive huber regression. *Journal of the American Statistical Association*, *115*(529), 254–265.
- Sun, Q., Wang, X., Yan, C., & Zhang, X. (2023). Vq-infratrans: A unified framework for rgb-ir translation with hybrid transformer. *Remote Sensing*, *15*(24), 5661.
- Tolan, J., Yang, H.-I., Nosarzewski, B., Couairon, G., Vo, H. V., Brandt, J., Spore, J., Majumdar, S., Haziza, D., Vamaraju, J., et al. (2024). Very high resolution canopy height maps from rgb imagery using self-supervised vision transformer and convolutional decoder trained on aerial lidar. *Remote Sensing of Environment*, *300*, 113888.
- Van Zanten, M., Pons, T., Janssen, J., Voesenek, L., & Peeters, A. (2010). On the relevance and control of leaf angle. *Critical Reviews in Plant Science*, *29*(5), 300–316.
- Werner, C., Ryel, R. J., Correia, O., & Beyschlag, W. (2001). Structural and functional variability within the canopy and its relevance for carbon gain and stress avoidance. *Acta Oecologica*, *22*(2), 129–138.
- Werner, C., Wallrabe, U., Christen, A., Comella, L. M., Dormann, C., Göritz, A., Grote, R., Haberstroh, S., Jouda, M., Kiese, R., et al. (2024). Ecosense-multi-scale quantification and modelling of spatio-temporal dynamics of ecosystem processes by smart autonomous sensor networks. *Research Ideas and Outcomes*, *10*, e129357.
- Wolf, S., Mahecha, M. D., Sabatini, F. M., Wirth, C., Bruelheide, H., Kattge, J., Moreno Martínez, Á., Mora, K., & Kattenborn, T. (2022). Citizen science plant observations encode global trait patterns. *Nature ecology & evolution*, *6*(12), 1850–1859.
- Yang, L., Kang, B., Huang, Z., Zhao, Z., Xu, X., Feng, J., & Zhao, H. (2024). Depth anything v2. *Advances in Neural Information Processing Systems*, *37*, 21875–21911.

- 825 Yang, X., Li, R., Jablonski, A., Stovall, A., Kim, J., Yi, K., Ma, Y., Beverly, D., Phillips, R., Novick,
826 K., et al. (2023). Leaf angle as a leaf and canopy trait: Rejuvenating its role in ecology with
827 new technology. *Ecology Letters*, 26(6), 1005–1020.
- 828 Zhang, G., Wang, C., Xu, B., & Grosse, R. (2018). Three mechanisms of weight decay regularization.
829 *arXiv preprint arXiv:1810.12281*.
- 830 Zheng, G., & Moskal, L. M. (2012). Leaf orientation retrieval from terrestrial laser scanning (tls)
831 data. *IEEE Transactions on Geoscience and Remote Sensing*, 50(10), 3970–3979.
- 832 Zou, X., Möttus, M., Tammeorg, P., Torres, C. L., Takala, T., Pisek, J., Mäkelä, P., Stoddard, F., &
833 Pellikka, P. (2014). Photographic measurement of leaf angles in field crops. *Agricultural and*
834 *Forest Meteorology*, 184, 137–146.

A Appendix

A.1 Genus representation across datasets

Table A1: Genus-level representation across training, validation, and phylogenetic analysis (PA) datasets. Training and validation counts show the number of images per genus used for model development. Phylogenetic analysis includes three samples per genus where available, with mean absolute error (MAE) values representing the mean absolute deviation between AngleCam V2 predictions and labeled leaf inclination angle distribution references in degrees. Empty MAE values (—) indicate genera not included in the phylogenetic analysis. Genera are ordered by phylogenetic data availability and total sample count.

Genus	Training	Validation	Total	PA	MAE (°)
<i>Tilia</i>	1014	249	1263	3	3.92
<i>Acer</i>	714	179	893	3	3.94
<i>Quercus</i>	198	60	258	3	10.61
<i>Fagus</i>	191	35	226	3	7.99
<i>Carpinus</i>	119	29	148	3	12.75
<i>Fraxinus</i>	107	31	138	3	10.62
<i>Sorbus</i>	93	31	124	3	7.76
<i>Eucalyptus</i>	86	20	106	3	3.32
<i>Rhododendron</i>	86	17	103	3	7.29
<i>Aesculus</i>	76	23	99	3	11.67
<i>Mespilus</i>	40	14	54	3	9.70
<i>Betula</i>	41	11	52	3	7.63
<i>Salix</i>	30	10	40	3	6.24
<i>Cynophalla</i>	28	8	36	3	6.98
<i>Prunus</i>	24	11	35	3	11.07
<i>Cornus</i>	26	7	33	3	9.71
<i>Sasa</i>	24	6	30	3	4.94
<i>Syringa</i>	25	5	30	3	10.73
<i>Crinum</i>	25	1	26	3	7.40
<i>Magnolia</i>	19	7	26	3	4.23

Continued on next page

Table A1 – continued from previous page

Genus	Training	Validation	Total	PA	MAE (°)
<i>Ocimum</i>	22	4	26	3	2.71
<i>Alnus</i>	18	6	24	3	5.61
<i>Rhytidophyllum</i>	18	6	24	3	8.41
<i>Olea</i>	18	5	23	3	14.72
<i>Pachystachys</i>	17	6	23	3	3.00
<i>Arbutus</i>	17	5	22	3	14.48
<i>Bixa</i>	17	4	21	3	9.54
<i>Cercis</i>	15	4	19	3	7.46
<i>Styrax</i>	14	4	18	3	5.51
<i>Deherainia</i>	14	2	16	3	7.49
<i>Hedera</i>	11	4	15	3	2.75
<i>Pyrus</i>	13	2	15	3	5.91
<i>Dracaena</i>	13	1	14	3	6.02
<i>Xylopi</i>	11	3	14	3	10.79
<i>Berberis</i>	12	1	13	3	8.16
<i>Coleus</i>	10	3	13	3	6.47
<i>Corylus</i>	7	5	12	3	8.94
<i>Laurus</i>	8	4	12	3	10.40
<i>Pilea</i>	10	2	12	3	1.14
<i>Ulmus</i>	10	2	12	3	8.53
<i>Alliaria</i>	10	1	11	3	4.72
<i>Clematis</i>	10	1	11	3	11.24
<i>Crassula</i>	10	1	11	3	6.03
<i>Ctenanthe</i>	7	4	11	3	10.85
<i>Euphorbia</i>	11	0	11	3	5.82
<i>Lavandula</i>	9	2	11	3	13.99
<i>Oplismenus</i>	11	0	11	3	3.96
<i>Rubus</i>	8	3	11	3	4.37

Continued on next page

Table A1 – continued from previous page

Genus	Training	Validation	Total	PA	MAE (°)
<i>Buxus</i>	8	1	9	3	11.96
<i>Goudaea</i>	7	2	9	3	1.43
<i>Passiflora</i>	8	1	9	3	6.46
<i>Epipremnum</i>	8	0	8	3	8.38
<i>Ficus</i>	5	3	8	3	4.78
<i>Heliconia</i>	6	2	8	3	4.36
<i>Hypericum</i>	7	0	7	3	2.48
<i>Urceolina</i>	5	2	7	3	10.58
<i>Viburnum</i>	6	1	7	3	10.51
<i>Adiantum</i>	5	1	6	3	9.16
<i>Begonia</i>	5	1	6	3	3.98
<i>Brachychiton</i>	6	0	6	3	5.22
<i>Callianthe</i>	5	1	6	3	4.83
<i>Diospyros</i>	6	0	6	3	1.67
<i>Ilex</i>	5	1	6	3	5.14
<i>Ligustrum</i>	5	1	6	3	8.79
<i>Platanus</i>	4	2	6	3	7.74
<i>Sambucus</i>	6	0	6	3	10.88
<i>Acacia</i>	4	1	5	3	9.22
<i>Asplenium</i>	4	1	5	3	11.03
<i>Carex</i>	2	3	5	3	1.73
<i>Eupomatia</i>	4	1	5	3	3.16
<i>Simmondsia</i>	3	2	5	3	13.25
<i>Fittonia</i>	4	0	4	3	4.32
<i>Hippophae</i>	2	2	4	3	8.07
<i>Malus</i>	4	0	4	3	7.25
<i>Microcitrus</i>	4	0	4	3	7.65
<i>Tripsacum</i>	3	1	4	3	4.81

Continued on next page

Table A1 – continued from previous page

Genus	Training	Validation	Total	PA	MAE (°)
<i>Zantedeschia</i>	4	0	4	3	3.19
<i>Corymbia</i>	2	1	3	3	4.47
<i>Eugenia</i>	1	2	3	3	7.49
<i>Melicytus</i>	3	0	3	3	6.83
<i>Rhus</i>	1	2	3	3	3.78
<i>Robinia</i>	3	0	3	3	3.82
<i>Tetradium</i>	3	0	3	3	2.52
<i>Arctostaphylos</i>	2	0	2	3	7.82
<i>Carpenteria</i>	2	0	2	3	7.39
<i>Ceratonia</i>	2	0	2	3	8.75
<i>Correa</i>	2	0	2	3	11.48
<i>Juglans</i>	2	0	2	3	5.36
<i>Lyonothamnus</i>	2	0	2	3	3.69
<i>Philadelphus</i>	2	0	2	3	6.85
<i>Solanum</i>	1	1	2	3	11.06
<i>Beaucarnea</i>	1	0	1	3	10.84
<i>Cotoneaster</i>	1	0	1	3	10.73
<i>Dactylis</i>	1	0	1	3	5.02
<i>Deutzia</i>	1	0	1	3	10.57
<i>Diplacus</i>	0	1	1	3	6.17
<i>Lonicera</i>	1	0	1	3	9.53
<i>Nothofagus</i>	1	0	1	3	1.44
<i>Spiraea</i>	1	0	1	3	3.21
<i>Tradescantia</i>	1	0	1	3	1.51
<i>Goeppertia</i>	19	1	20	—	—
<i>Maranta</i>	9	3	12	—	—
<i>Aglaonema</i>	2	1	3	—	—
<i>Chimonobambusa</i>	3	0	3	—	—

Continued on next page

Table A1 – continued from previous page

Genus	Training	Validation	Total	PA	MAE (°)
<i>Erica</i>	2	1	3	—	—
<i>Kumara</i>	1	2	3	—	—
<i>Lasiopetalum</i>	3	0	3	—	—
<i>Liriodendron</i>	1	2	3	—	—
<i>Morus</i>	3	0	3	—	—
<i>Reynoutria</i>	2	1	3	—	—
<i>Skimmia</i>	3	0	3	—	—
<i>Xiphidium</i>	3	0	3	—	—
<i>Buddleja</i>	2	0	2	—	—
<i>Castanea</i>	1	1	2	—	—
<i>Hosta</i>	1	1	2	—	—
<i>Miconia</i>	2	0	2	—	—
<i>Ostrya</i>	2	0	2	—	—
<i>Paliurus</i>	1	1	2	—	—
<i>Sedum</i>	2	0	2	—	—
<i>Selaginella</i>	2	0	2	—	—
<i>Spathiphyllum</i>	2	0	2	—	—
<i>Zelkova</i>	1	1	2	—	—
<i>Aria</i>	1	0	1	—	—
<i>Crataegus</i>	1	0	1	—	—
<i>Dimorphotheca</i>	1	0	1	—	—
<i>Doryanthes</i>	1	0	1	—	—
<i>Gustavia</i>	1	0	1	—	—
<i>Spyridium</i>	0	1	1	—	—
<i>Vaccinium</i>	1	0	1	—	—

A.2 Pseudo-NIR generation examples

Visual examples of the pseudo-NIR pipeline are compared to actual night-vision NIR imagery across ten forest scenes (Fig. A1). These scenes represent two tree species: European beech (*Fagus sylvatica*

L.) and European hornbeam (*Carpinus betulus* L.). Images were acquired in 2025 at the ECOSENSE site (Ettenheim, Germany) and the Hartheim Forest Research Site (Germany). The selection covers diverse canopy conditions, including clear skies (Fig. A1, rows 2, 3, 5, 6, 7), direct solar radiation (row 3), pronounced daytime shadows (rows 5, 6, 7), and rain (row 1). Phenological stages range from leaf onset (row 4) to leaf senescence (rows 2, 10).

The ten selected scenes show a good agreement in illumination falloff between the generated pseudo-NIR images (Fig. A1, column d) and the actual NIR imagery (Fig. A1, column e). Although grayscale images (column b) reduce the model's dependence on color-specific features, they do not necessarily reflect the night-vision environment. Columns (d) and (e) consistently show patterns in which foreground vegetation is highlighted by increased brightness, while distant backgrounds fade into darkness. This effect reduces background detail to levels characteristic of NIR imagery, whereas grayscale scenes (column b) often preserve unrealistic detail at greater distances. Additionally, the pseudo-NIR approach eliminates "blue sky" effects by darkening the background (Fig. A1, rows 2, 5), thereby aligning with the monochromatic response of night-vision sensors.

Despite its potential as a training proxy, the pseudo-NIR pipeline has several inherent limitations. Actual NIR images show high vegetation reflectance, making leaves appear bright and uniform. While the pseudo-NIR pipeline replicates this effect, leaves located very close to the camera become excessively bright, leading to localized overexposure. Since grayscale and ultimately pseudo-NIR images are derived from daytime RGB photos, they retain ambient shadows determined by the sun's position. In contrast, actual NIR imagery captured with NIR-LEDs typically exhibits minimal shadowing on surfaces directly facing the camera. This discrepancy becomes strongly evident under lens flare conditions (Fig. A1, row 3), where the pseudo-NIR image contains overexposure and light stripes that are absent in real night-vision scenes. Furthermore, the illumination falloff with increasing distance is often more pronounced in the pseudo-NIR pipeline than in actual NIR imagery. In certain cases, actual NIR imagery preserves more background detail than its synthetic counterpart (Fig. A1, rows 3, 7, 10). This difference is most pronounced in the deep background and less pronounced in images focused on vegetation near the camera (Fig. A1, rows 1, 5, 6, 9).

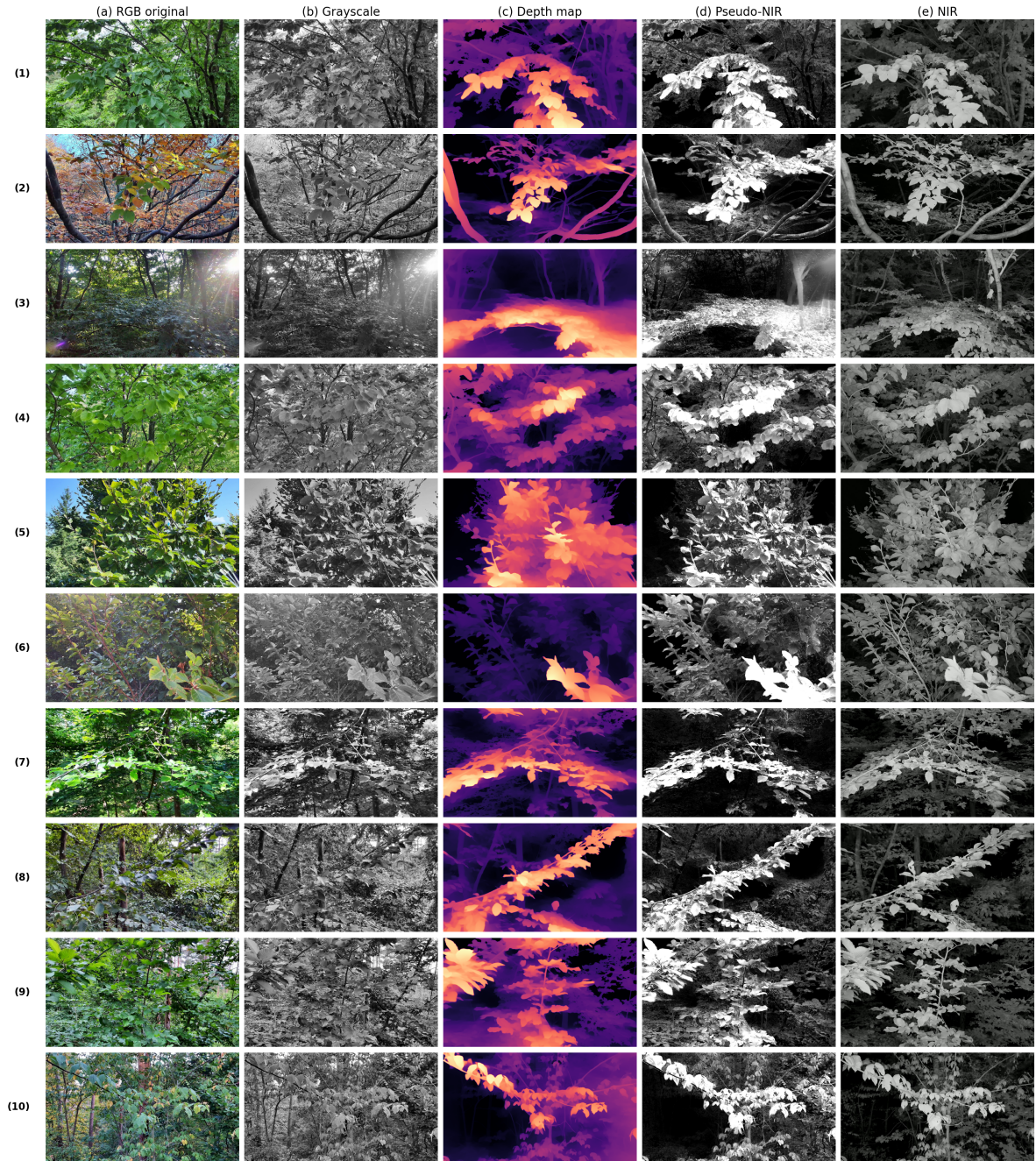


Figure A1: Visual examples of the pseudo-NIR image generation pipeline across ten different forest scenes (1–10). The columns represent the sequential transformation steps: (a) the original RGB photograph, (b) conversion to grayscale, (c) per-pixel depth estimation using Depth Anything V2, (d) the generated pseudo-NIR image simulating distance-based illumination falloff, and (e) the actual NIR night-vision image of the same scene captured during the nighttime period for reference.