

A Multi-Modal Dataset for Automated Phenological Stage Mapping in *Actinidia chinensis*

Isabel Pinheiro

`isabel.a.pinhoiro@inesctec.pt`

Institute for Systems and Computer Engineering, Technology and Science (INESC TEC)

<https://orcid.org/0000-0001-5387-0980>

Pedro Moura

Institute for Systems and Computer Engineering, Technology and Science (INESC TEC)

<https://orcid.org/0009-0001-3898-9841>

Leandro Rodrigues

Institute for Systems and Computer Engineering, Technology and Science (INESC TEC)

<https://orcid.org/0000-0003-4507-2481>

Germano Moreira

Institute for Systems and Computer Engineering, Technology and Science (INESC TEC)

<https://orcid.org/0000-0002-1382-8267>

Rui Manuel Coutinho

Institute for Systems and Computer Engineering, Technology and Science (INESC TEC)

<https://orcid.org/0000-0001-8821-2224>

Francisco Terra

Institute for Systems and Computer Engineering, Technology and Science (INESC TEC)

<https://orcid.org/0009-0000-2807-2506>

António Valente

School of Science and Technology, University of Trás-os-Montes e Alto Douro (UTAD)

<https://orcid.org/0000-0002-5798-1298>

Mário Cunha

Faculty of Sciences, University of Porto (FCUP) <https://orcid.org/0000-0002-8299-324X>

Filipe Neves dos Santos

Faculty of Engineering, University of Porto (FEUP) <https://orcid.org/0000-0002-8486-6113>

Data Note

Keywords:

Posted Date: March 13th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9097249/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: The authors declare no competing interests.

A Multi-Modal Dataset for Automated Phenological Stage Mapping in *Actinidia chinensis*

Isabel Pinheiro^{1,2*}, Pedro Moura^{1,2}, Leandro Rodrigues^{1,3}, Germano Moreira^{1,3},
Rui Manuel Coutinho^{1,4}, Francisco Terra¹, António Valente^{1,2},
Mário Cunha^{1,3}, Filipe Neves dos Santos^{1,4}

¹Institute for Systems and Computer Engineering, Technology and Science (INESC TEC), Porto, 4200-465, Portugal

²School of Science and Technology, University of Trás-os-Montes e Alto Douro (UTAD), Vila Real, 5000-801, Portugal

³Faculty of Sciences, University of Porto (FCUP), Porto, 4169-007, Portugal

⁴Faculty of Engineering, University of Porto (FEUP), Porto, 4169-007, Portugal

March 11, 2026

Abstract

Phenological monitoring of *Actinidia chinensis* is critical for optimising operational costs and yield prediction. However, current manual assessment methods are time-consuming, making them impractical for large-scale precision agriculture applications. Most existing phenological datasets focus exclusively on image data without spatial validation. The Multi-Modal *Actinidia chinensis* Phenology Dataset is composed of (i) 1 665 annotated images of phenological stages from bud to fruit set and (ii) georeferenced videos with systematic manual ground truth of spatial stage distributions. The dataset employs an adapted 17-class BBCH system that consolidates visually similar stages, excludes problematic categories, and introduces generic structural classes to address practical annotation difficulties. Additionally, the data is organised hierarchically across various plant structures, genders, and phenological stages. The annotated images offer versatility for a range of applications, including training data for computer vision models to detect phenological stages. Furthermore, the georeferenced videos facilitate the validation of automated counting algorithms. This combined approach enables plant-level detection accuracy and provides an illustrative methodology for spatial validation that users can extend to additional orchards, promoting the development and benchmarking of automated phenological monitoring systems for precision agriculture applications in kiwifruit production.

Background & Summary

Phenological monitoring in orchards relies on labour-intensive manual assessment methods that limit comprehensive evaluation. Fruit set rate estimation and assisted pollination timing decisions require manual counting of buds, flowers and fruits across small representative areas, with each square meter requiring approximately 10 minutes of assessment time. This sampling approach introduces significant limitations, as orchards exhibit spatial heterogeneity in flowering timing and fruit development that cannot be captured through localised observations of small representative areas, since orchard-wide generalisations are imprecise and potentially suboptimal for management interventions.

Existing phenological datasets employ various monitoring approaches spanning multiple spatial and temporal scales. Image-based datasets for crop phenology have been developed for species including raspberry [1], apple [2], cotton [3], and tomato [4], utilising computer vision (CV) techniques for stage classification and yield estimation. Temporal monitoring approaches integrate unmanned aerial vehicle imagery for repeated field-scale observations [5] and ground-based camera networks for continuous phenological tracking across diverse ecosystems [6, 7]. Continental-scale vegetation monitoring has been established through satellite remote sensing [8] and citizen science observation platforms [9]. Within *Actinidia* species, prior work has focused on flower detection and gender classification [10].

However, these existing approaches share a standard limitation whereby datasets focus exclusively on image data without spatial validation capabilities. At the same time, satellite-based methods lack the resolution to validate patterns at the individual plant level. For *Actinidia* specifically, no dataset currently provides both detailed phenological annotations and spatially-referenced validation data. The Multi-Modal *Actinidia chinensis* Phenology Dataset [11] is publicly available and addresses these gaps by providing georeferenced videos with manual ground truth counts, while maintaining a robust set of manually labelled images. While the image component captures morphological diversity across five orchards, the georeferenced video component demonstrates the spatial

*Email: isabel.a.pinheiro@inesctec.pt

validation methodology at a single orchard and serves as an illustrative example rather than a generalizable spatial ground truth. This dual-component approach enables assessment of phenological heterogeneity at the orchard level while maintaining accuracy at the individual plant level.

The Multi-Modal *Actinidia chinensis* Phenology Dataset [11] comprises two complementary components organised hierarchically across plant structures, gender, and phenological stages. The labelled images component provides 1 665 images with bounding box annotations for developing CV detection models. The georeferenced videos component provides spatial validation by systematically counting across defined sampling zones to benchmark automated counting algorithms. This modular implementation enables diverse applications, including phenological mapping for timing of agricultural operations such as thinning or assisted pollination, and fruit set assessment for yield prediction. Specifically for assisted pollination, phenological mapping enables the identification of receptive-flower distributions across the orchard, supporting targeted pollen-application strategies at optimal dates to maximise pollination efficiency, minimise pollen expenditure, and improve fruit quality.

Figure 1 illustrates the Multi-Modal *Actinidia chinensis* Phenology Dataset architecture comprising the Labelled Images Component and Georeferenced Videos Component. Both components employ a colour-coded schema to delineate the hierarchical organisation: (i) structure in blue, (ii) gender in yellow, and (iii) phenological stage in orange.

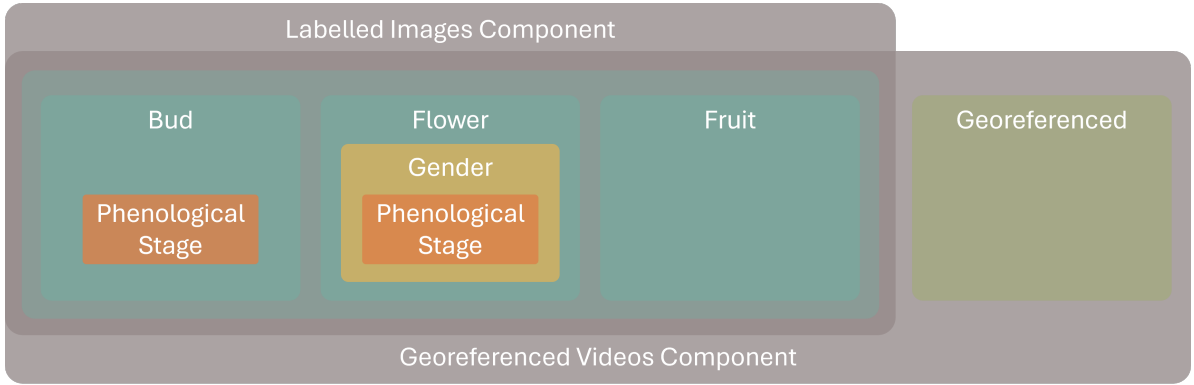


Figure 1: Multi-Modal *Actinidia chinensis* Phenology Dataset architecture illustrating the hierarchical organisation across two complementary components: (i) manually annotated images organised by plant structure, gender and phenological stage, and (ii) spatially referenced video recordings with systematic manual counts using the same structure.

This Data Descriptor details the creation and structure of the Multi-Modal *Actinidia chinensis* Phenology Dataset in three main sections. The Methods Section presents the data-acquisition methodology, including the phenological-stage-based classification framework and the annotation protocols for both dataset components. The Data Records Section details the dataset organisation, file specifications, and repository structure. The Technical Validation Section describes the validation procedures ensuring annotation quality and reliability.

Methods

Data Acquisition

The data acquisition procedures were designed to address the specific requirements of the dataset’s two components described above. The strategy for labelled images prioritised morphological and environmental diversity across multiple orchards and dates to capture morphological variability. At the same time, the georeferenced videos component focused on comprehensive spatial coverage within defined sampling zones to support validation of automated monitoring algorithms.

Labelled Images Component

The labelled images component provides morphological annotations for developing CV models. The images were collected to maximise morphological and environmental variability in spatial, temporal, and technical parameters.

The data were collected from five different *Actinidia chinensis* orchards in northern Portugal, covering variations in age, variety, cultivar, training system, and region. Table 1 characterises the five orchards used to acquire images.

Sampling across orchards with different ages, training systems, and regions captured morphological variability in flower arrangement and canopy structure, ensuring the dataset represents diverse production conditions. Regarding variety, the same variety does not present apparent differences in buds and flowers. However, the flowers are easily distinguished between varieties, since *Actinidia chinensis* var. *chinensis* has slightly smaller and more delicate flowers.

Table 1: Characteristics of the five *Actinidia chinensis* orchards used for data collection.

Orchard	Age (year)	Variety	Cultivar	Training system	Location (lat, long)	City
O1	12	<i>deliciosa</i>	"Hayward"	Pergola	41°19'18.14"N, 8°39'42.00"W	Vila do Conde
O2	30	<i>deliciosa</i>	"Hayward"	Pergola	41°7'35.88"N, 8°32'55.44"W	Gondomar
O3	15	<i>deliciosa</i>	"Bo.Erika"	Pergola	41°3'17.75"N, 8°35'16.99"W	Vila Nova de Gaia
O4	8	<i>deliciosa</i>	"Hayward"	T-bar	40°36'0.53"N, 8°32'40.29"W	Aveiro
O5	5	<i>chinensis</i>	"Dori"	T-bar	40°36'0.54"N, 8°32'34.61"W	Aveiro

The images were collected over nine days on 18 May 2023 and between 17 April 2024 and 5 June 2024. This period corresponds to the flowering phase, which is the most critical for guiding agronomic practices such as assisted pollination timing and fruit thinning decisions, capturing the complete sequence of reproductive development from bud emergence to fruit set. Three operators acquired images using smartphones (iPhone Xr (SP1), iPhone 15 Pro (SP2), or Xiaomi Redmi Note 12 (SP3)) with resolutions ranging from $1\,080 \times 1\,080$ to $6\,144 \times 6\,144$ pixels. The use of multiple smartphone models introduced variability in image characteristics (sensor specifications, colour rendering, and compression), enriching the dataset by reflecting real-world acquisition conditions while maintaining sufficient image quality for reliable morphological annotation across all devices. During acquisition, operators varied the distance from the vegetative wall and the viewing angle to ensure comprehensive morphological documentation from multiple perspectives. This intentional variability captures diverse acquisition scenarios representative of practical field conditions. Temporal metadata and precise georeferencing were not recorded, as the focus was on morphological diversity rather than spatiotemporal mapping. Table 2 shows the schedule of the nine image-acquisition sessions, listing the date, orchard, operator, and smartphone model for each.

Table 2: Characteristics of the nine image-acquisition sessions conducted across five *Actinidia chinensis* orchards (O1-O5). Each session involved an operator (OP1 - OP3) with a smartphone (iPhone Xr (SP1), iPhone 15 Pro (SP2) or Xiaomi Redmi Note 12 (SP3)).

Date	Orchard	Operator	Smartphone	Images
2023-05-18	O1	OP1	SP1	576
2024-04-17	O2	OP1	SP2	237
2024-04-21	O4	OP1	SP2	252
2024-04-23	O2	OP3	SP3	29
2024-04-25	O4	OP3	SP2	139
2024-04-25	O5	OP3	SP2	574
2024-05-03	O2	OP1	SP2	238
2024-05-10	O2	OP1	SP2	394
2024-05-23	O2	OP2	SP3	193
2024-05-23	O3	OP2	SP3	209
2024-06-05	O2	OP2	SP2	263

A total of 3 104 images were acquired, documenting the developmental stages from bud to fruit set. This collection includes 576 images from 2023 [12] and 2 528 images from 2024 [13], publicly available. From this set of images, 1 665 images were selected based on uniqueness calculations using the Fiftyone platform [14] to identify and remove duplicate or near-duplicate images, thereby avoiding accidental concept imbalance that can bias the model’s learning. The selected images were pre-processed and standardised to a resolution of $1\,024 \times 1\,024$ pixels to optimise compatibility with CV structures.

Georeferenced Videos Component

The georeferenced videos component provides spatial distribution data and manual ground truth validation for phenological mapping applications.

Data were collected at orchard O6, an *Actinidia chinensis* var. *chinensis* cv. "Hayward" site with 12-year-old plants trained on a pergola system, located in Vila do Conde (41°17'27.48"N, 8°38'57.57"W). Video acquisition occurred on a single date (28 May 2025) during peak reproductive activity, providing an illustrative demonstration of the spatial validation methodology rather than a comprehensive representation of spatial heterogeneity across multiple orchards or management conditions. The experimental design utilised four 5×5 meter sampling zones distributed across the orchard to capture within-orchard spatial heterogeneity in phenological development. Figure 2 illustrates the seven-wire pergola infrastructure within each sampling zone, creating six inter-wire sections that define the systematic sampling framework for spatial analysis.

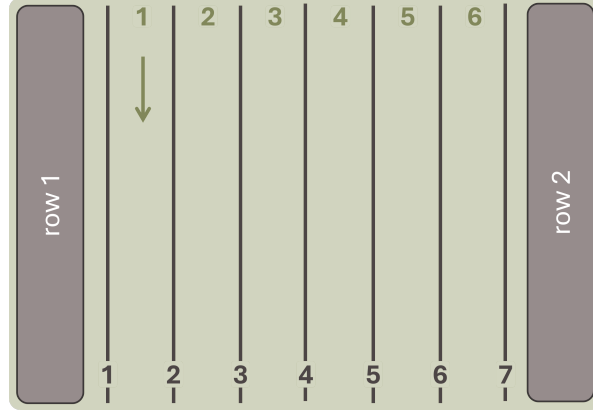


Figure 2: Schematic representation of the seven-wire pergola infrastructure within each sampling zone, creating six inter-wire sections that define the systematic sampling framework.

Video acquisition was conducted on 28 May 2025 during peak reproductive activity, using standardised collection protocols. Six videos were recorded per zone along predetermined transects parallel to the vegetative canopy. The camera was mounted on a tripod positioned approximately 1.80 metres from the vegetative wall, oriented parallel to the canopy, with the operator walking at approximately 0.25 metres per second between way-points recorded using the Geo++ RINEX Logger smartphone application [15], with GNSS data post-processed to achieve centimetre-level accuracy. This standardised setup minimised operator-specific biases arising from height or camera-angle variability. Videos were captured using Xiaomi Poco X6 Pro (SP4) at 3840×2160 pixel resolution and 30 fps, stored in MP4 format. Frame-level georeferencing was achieved by linearly interpolating between the start and end coordinates, assuming constant walking velocity and a straight-line trajectory. GPS Exchange Format (GPX) files were generated for each recording, containing calculated temporal-spatial metadata for individual frames.

Data Annotation

The Multi-Modal *Actinidia chinensis* Phenology Dataset follows a hierarchical classification framework that accommodates three levels of analysis (structure, gender, and phenological stage).

The first level distinguishes between bud, flower and fruit, three structures of *Actinidia chinensis*. The bud represents the initial stage of development, in which the floral organs are immature and not visible, and the flower develops. The flower represents the stage where the reproductive organs are visible, and individuals can be classified as female or male. When properly pollinated, the female flower develops into fruit, the final stage.

The second level distinguishes female and male flowers, addressing the dioecious nature of *Actinidia chinensis*. Recognising the gender distinction is crucial, as each gender has distinct functions. While female flowers develop into kiwifruit after pollination, which directly impacts yield, male flowers serve exclusively as pollen sources during the pollination phase. There are distinct morphological characteristics of the reproductive organs of flowers that allow for clear visual identification of the gender [16]. These distinct morphological features allow accurate gender classification during data collection and annotation. Figure 3 identifies the key morphological characteristics of the reproductive organs to differentiate the genders.



(a)



(b)

Figure 3: Key morphological characteristics to distinguishing gender in *Actinidia chinensis* flowers [17]. **(a)** Female flowers exhibit a star-shaped stigma (purple shadow) surrounded by the anthers that contain non-viable pollen (pink shadow). **(b)** Male flowers display numerous well-developed anthers with viable pollen (blue shadow) with reduced pistil development.

The third level is the most detailed classification level that implements standardised phenological staging according to the BBCH (Biologische Bundesanstalt, Bundessortenamt und CHemische Industrie) scale [18] adapted for *Actinidia chinensis* [16]. The BBCH scale is characterised by a decimal-based framework that matches plant development into ten principal growth stages, which are subdivided into ten secondary stages (not all numbers need to correspond to a stage of development), enabling precise documentation of phenological events [19]. For this work, the principal growth stages of interest are the inflorescence emergence and flowering stages (stage 5 and stage 6, respectively) [16]. Figure 4 shows the twelve secondary phenological stages, using the defined colour-coded schema to delineate the hierarchical organisation, while Table 3 provides detailed descriptions of the key morphological characteristics for distinguishing the respective secondary stages.



Figure 4: Examples of each phenological growth stage of *Actinidia chinensis* according to the BBCH scale from inflorescence bud swelling (stage 51) through fruit set (stage 69) [16].

Table 3: Criteria for distinguishing each phenological growth stage of *Actinidia chinensis* according to the BBCH scale from inflorescence bud swelling (stage 51) through fruit set (stage 69) [16].

BBCH	Description
51	Inflorescence emergence characterised by inflorescence bud swelling, more specifically closed buds without peduncles and visible greenish sepals, covered with trichomes.
53	Growing flower buds, specifically buds still closed with reddish peduncles elongating.
55	Sepals begin to separate, revealing a greenish-white corolla, and the reddish peduncles continue to elongate.
56	Continuation of the separation of the sepals, revealing the white corolla, which is longer than the calyx, with the peduncles elongating and thickening.
57	Balloon stage corolla with the white petals forming a hollow ball with one petal separating.
59	Several separate petals with no visible pistils, longer than the calyx.
60	First flowers open with a bell-shaped appearance.
61	Beginning of flowering with 10% of flowers open in the orchard.
65	Full flowering with 50% of flowers open in the orchard.
67	Petal fading begins with some pistils still fertile.
68	Most petals fallen or dry with non-functional and dry pistils.
69	End of flowering phase with fruit set visible.

The BBCH scale has limitations for dioecious species. Male flowers are not included, as they fulfil their function by releasing pollen without developing fruit, creating gaps in the phenological classification. Furthermore, some stages (stage 61 and stage 65) describe flowering patterns at the level of the orchard’s flower population, rather than individual flower characteristics.

The Multi-Modal *Actinidia chinensis* Phenology Dataset classes were defined according to the BBCH scale, and some adaptations were made to resolve the mentioned limitations.

- In the case of buds, stage 51 was excluded since these small buds without a peduncle were frequently occluded by the foliage. Stage 57 and stage 59 were combined into class bud_57, as they had minimal morphological differences, making identification difficult.
- In the case of female flowers, stage 61 and stage 65 at the population level were adapted to individual flower characteristics, with flower_female_61 redefined as fully open flowers.
- In the case of male flowers, female flower stages were adapted to include male flower development stages (flower_male_60, flower_male_61, flower_male_67).

The resulting system comprises 17 classes that combine specific BBCH stage categories with generic structural classes to accommodate cases where stage or gender determination is not feasible due to image quality, viewing angle, or developmental ambiguity. Figure 5 presents the 17 classes using the colour-coded hierarchical schema. Table 4 provides morphological criteria for each class and summarises the correspondence with standard BBCH stages.



Figure 5: Examples of each class, including specific classes for each phenological stage (from swelling of inflorescence buds (bud_53) to fruiting (flower_female_69) and generic classes for gender (flower_female and flower_male) and structure (bud, flower and fruit).

Table 4: Criteria for distinguishing each class and correspondence to standard BBCH stages. Male flower classes are dataset-specific extensions, and generic classes have no BBCH equivalent.

Class	BBCH	Description
bud_53	53	Closed bud with visible peduncles.
bud_55	55	White-greenish corolla visible through sepals.
bud_56	56	White corolla extends beyond sepal margins.
bud_57	57 + 59	Petal(s) separation with no visible stigma.
flower_female_60	60	Partially open petals with bell-shaped appearance and visible stigma.
flower_female_61	61 + 65	Fully open petals with receptive stigma.
flower_female_67	67	Petal fading with some pistils still fertile.
flower_female_68	68	Most petals fallen or dry with all pistils non-functional.
flower_female_69	69	Visible fruit swelling with advanced stigma senescence.
flower_male_60	new	Partially open petals with bell-shaped appearance.
flower_male_61	new	Fully open petals.
flower_male_67	new	Petal fading or dry.
flower_female	-	Generic class for female flowers when the phenological stage cannot be determined.
flower_male	-	Generic class for male flowers when the phenological stage cannot be determined.
bud	-	Generic class for buds when the phenological stage cannot be determined.
flower	-	Generic class for flower when the gender cannot be determined.
fruit	-	Generic class for fruit.

To ensure consistent classification across both dataset components, a structured annotation protocol was developed with explicit decision criteria for distinguishing between phenological stages. Figure 6 presents the complete decision flowchart guiding the classification process.

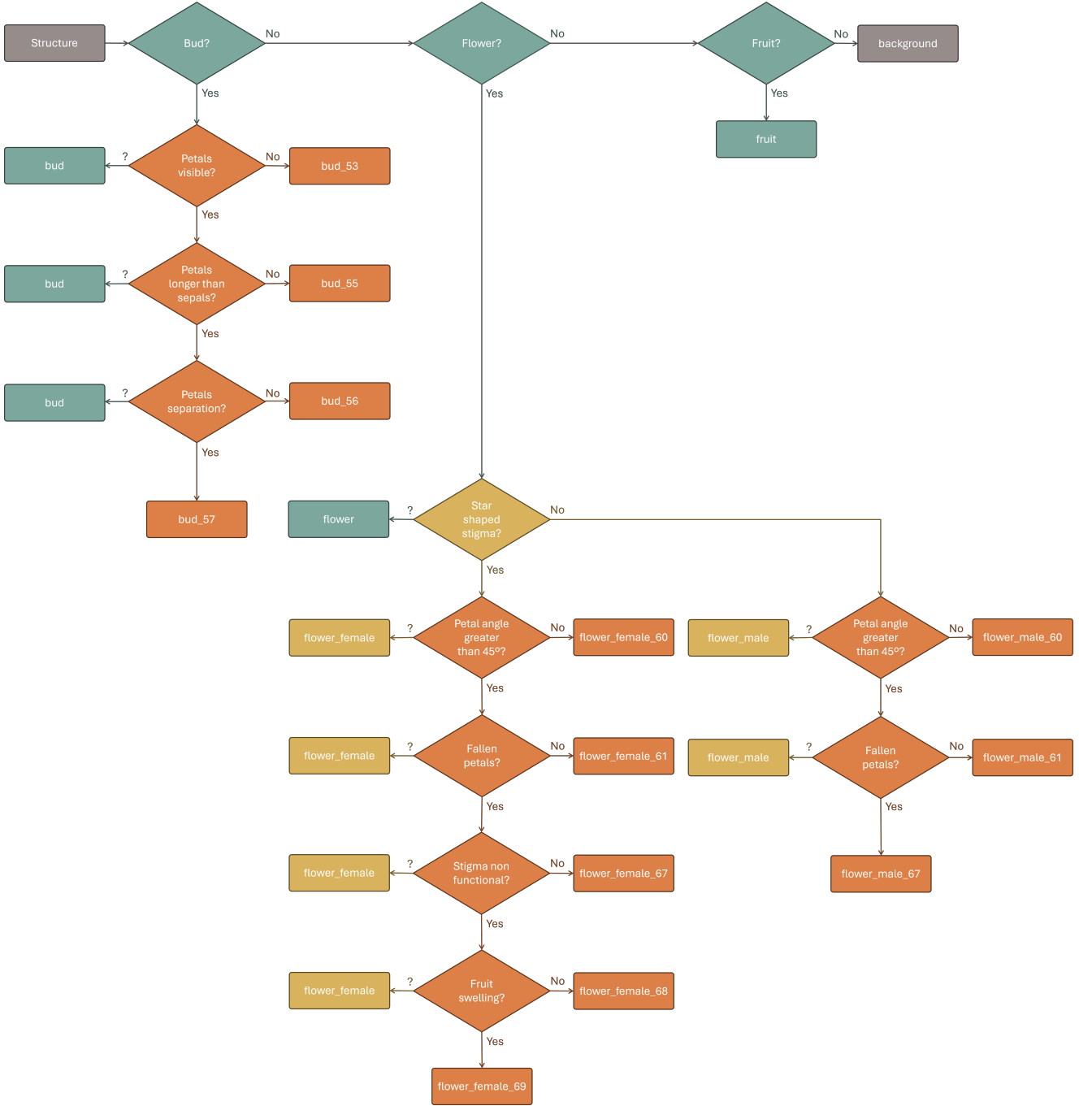


Figure 6: Annotation decision flowchart guiding classification from region of interest detection to final class assignment. Uncertain cases at any decision point are assigned to the appropriate generic class via "?" paths.

Starting from a structure, annotators first evaluate its type by sequentially assessing whether it corresponds to a bud, a flower, or a fruit. Structures that do not match any category are classified as background. The decision tree employed sequential morphological criteria specific to each structure type. For buds, annotators evaluated: (i) petal visibility through sepals, (ii) petal extension beyond sepal margins, and (iii) petal separation without visible stigma. For flowers, annotators first determined gender by identifying the star-shaped stigma characteristic of female flowers. For female flowers, annotators evaluated: (i) extent of petal opening, (ii) presence of petal fading, (iii) stigma functionality, and (iv) fruit development. Male flowers were assessed using the first two criteria, as they do not develop fruit. At each decision point, when annotators cannot confidently determine the answer, the structure is assigned to the appropriate generic class via the "?" paths, ensuring systematic handling of transitional or ambiguous instances.

The Multi-Modal *Actinidia chinensis* Phenology Dataset comprises two complementary components utilising these 17 classes. The first component consists of manually annotated images with bounding box labels for CV model training and validation. The second component includes georeferenced videos with systematic phenological stage counts across defined spatial areas, providing ground truth data for spatial distribution analysis and monitoring applications.

Labelled Images Component

Manual annotation was performed using the Computer Vision Annotation Tool (CVAT) [20] following the decision flowchart presented in Figure 6. Bounding boxes tightly enclosed each structure with minimal margins, excluding peduncles to focus on diagnostic morphology. Structures were annotated when at least 50% of defining characteristics were visible and unoccluded, permitting partial overlap between adjacent structures. The dataset includes images with single or multiple classes, as well as background images without target structures.

Annotation was performed by a primary annotator across all 1 665 images over three consecutive days to maintain temporal consistency. Independent validation was conducted by a trained secondary annotator on a randomly selected 10% subset (167 images), as detailed in Technical Validation.

Images were exported in Pascal Visual Object Classes (VOC) format, generating XML files with annotation metadata including bounding box coordinates, class assignments, and image dimensions. Table 5 presents the distribution of images and annotations across individual classes and hierarchical levels, with the colour scheme reflecting the hierarchical structure.

Table 5: Number of images and annotations for each class in the Multi-Modal *Actinidia chinensis* Phenology Dataset. Colour coding reflects hierarchical levels: structure (blue), gender (yellow), and phenological stage (orange).

Classes	3 Levels		2 Levels		1 Level		Total					
	images	annotations	images	annotations	images	annotations	images	annotations				
background	108	0	108	0	108	0	1 655	10 857				
bud_53	242	969	817	4 379	817	4 379						
bud_55	451	1 294										
bud_56	403	756										
bud_57	327	603										
bud	284	757										
flower_female_60	267	406	954	4 006	1245	6 363						
flower_female_61	481	931										
flower_female_67	415	1 299										
flower_female_68	154	355										
flower_female_69	171	489										
flower_female	330	526										
flower_male_60	156	274	295	1 724								
flower_male_61	260	802										
flower_male_67	113	221										
flower_male	195	427										
flower	319	633	319	633								
fruit	26	115	26	115	26	115						

Georeferenced Videos Component

Manual phenological assessment followed the decision criteria established in Figure 6. Unlike image-based annotation, field observers had direct access to structures from multiple viewing angles, enabling confident determination of both gender and phenological stage. Consequently, generic classes were unnecessary, and all structures were classified into the 13 stage-specific classes. For model validation workflows, users should evaluate stage-specific predictions against the 13-class ground truth and report generic class predictions separately as structures where a limited perspective prevented confident assignment.

Ground truth counting utilised a structured data matrix with individual plant shoots as rows and phenological stage classes as columns. Operators systematically inspected each sampling zone, cross-checking ambiguous structures from different angles and recording only clearly identifiable phenological features. Manual counts were digitised and compiled into standardised Excel files for each sampling zone, providing zone-level ground truth data corresponding to the collective set of six videos per zone.

Table 6 presents spatial distribution patterns and validation counts across sampling zones and phenological classes, with colour coding reflecting the hierarchical classification structure.

Table 6: Ground truth counts across sampling zones showing hierarchical aggregation levels for the georeferenced videos component. Colour coding reflects hierarchical levels: structure (blue), gender (yellow), and phenological stage (orange).

Class	Zones											
	A			B			C			D		
bud_53	53			37			12			14		
bud_55	120	269	269	45	114	114	37	102	102	46	113	113
bud_56	32			28			38			38		
bud_57	64			4			15			15		
flower_female_60	91			13			41			20		
flower_female_61	196			73			256			511		
flower_female_67	233	526		79	175		391	695		569	1 106	
flower_female_68	6			2			7			6		
flower_female_69	0		562	8		175	0		695	0		1 106
flower_male_60	24			0			0			0		
flower_male_61	0	36		0	0		0	0		0	0	
flower_male_67	12			0			0			0		
fruit	0	0	0	0	0	0	0	0	0	0	0	0
Total	831			289			797			1 219		

Data Record

The Multi-Modal *Actinidia chinensis* Phenology Dataset [11] is publicly available at Zenodo. The dataset comprises 1 665 annotated images and 24 georeferenced video segments, with corresponding ground truth validation data, structured to support hierarchical phenological classification and spatial mapping applications. Figure 7 illustrates the dataset architecture and file organisation.

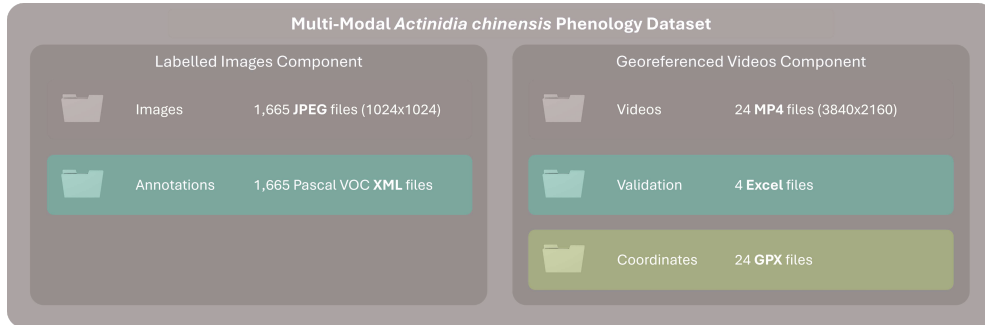


Figure 7: Multi-Modal *Actinidia chinensis* Phenology Dataset architecture showing file organisation and specifications for both components.

The labelled images component contains JPEG files standardised to $1\,024 \times 1\,024$ pixels with corresponding Pascal VOC XML annotations. Each XML file contains bounding box coordinates (xmin, ymin, xmax, ymax) and class labels from 17 categories encompassing stage-specific BBCH classifications and generic morphological groups.

The georeferenced videos component comprises 24 MP4 files distributed across four sampling zones (A-D) with six videos per zone. Each video has a corresponding GPX file containing track points with WGS84 coordinates, timestamps at 1-second intervals, and metadata linking to the source video. Frame-level spatial referencing is achieved by linearly interpolating between the recorded start and end positions.

Manual validation data are provided as Excel files, one per sampling zone, with plant shoots as rows and 13 stage-specific BBCH categories as columns. Generic morphological classes are excluded as field protocols enabled assignment to specific developmental stages.

Technical Validation

Labelled Images Component

Annotation quality was assessed through a systematic validation protocol. A primary annotator labelled all 1 665 images over three consecutive days to maintain temporal consistency, following the classification criteria established in Table 4 and the decision flowchart in Figure 6.

To assess inter-annotator reliability, a secondary annotator completed four hours of structured training and independently annotated a randomly selected 10% subset (167 images) without access to primary annotations. Inter-annotator agreement was quantified using Cohen’s Kappa coefficient ($\kappa = 0.8$) for class labels and mean Intersection over Union ($mIoU = 0.7$) for bounding box localisation [21].

The 10% validation confirms protocol consistency but does not assess systematic annotator bias. Fine-grained distinctions between consecutive stages require subjective judgment, and threshold decisions reflect primary annotator interpretation. Additionally, the dataset exhibits class imbalance reflecting natural phenological occurrence, with female flower stages more represented than male and late senescence stages underrepresented (Table 5). Users should consider class-weighted approaches for model training and account for potential bias toward dominant classes when interpreting performance metrics.

Georeferenced Videos Component

Manual phenological assessments were conducted in the field by operators following the classification criteria in Table 4 for 13 stage-specific classes. Generic classes were excluded as field conditions permitted detailed morphological assessment across multiple viewing angles, enabling reliable assignment to specific developmental stages. Manual counts are provided as Excel files containing plant-level phenological stage assignments for each sampling zone.

The spatial validation data represent a single orchard and serve as a methodological framework that users can replicate in additional orchards to extend spatial coverage. Class distribution is uneven, with male flowers absent from three of four sampling zones due to orchard spatial arrangement (Table 6). Users validating detection models against spatial counts should account for this uneven distribution across genders and stages.

Data Availability

The Multi-Modal *Actinidia chinensis* Phenology Dataset described in this Data Descriptor is publicly available at Zenodo: <https://doi.org/10.5281/zenodo.17371025>. This dataset comprises two components: (1) 1 665 JPEG images ($1\,024 \times 1\,024$ pixels) with corresponding Pascal VOC XML annotation files containing bounding box coordinates and phenological class labels, and (2) 24 MP4 video files ($3\,840 \times 2\,160$ pixels) with corresponding GPX coordinate files and Excel validation files containing manual ground truth counts.

Code Availability

Custom scripts for dataset preparation are publicly available under the MIT License at <https://github.com/INESCTEC/actinidia-dataset-preprocessing>. These scripts implement standard operations, including image resolution standardisation to $1\,024 \times 1\,024$ pixels (reproducible from original raw images) and GPX track generation for video georeferencing through linear interpolation (reproducible from original videos and coordinate logs). The scripts are provided for transparency and to enable users to apply similar preprocessing to their own data.

Image annotation was conducted using CVAT (<https://github.com/opencv/cvat>). Uniqueness calculations for duplicate detection were performed using the Fiftyone platform (<https://github.com/voxel51/fiftyone>).

Author Contributions

Conceptualisation: I.P., M.C., and F.N.d.S.; Methodology: I.P. and F.N.d.S.; Data collection: I.P., P.M., L.R., G.M., R.M.C. and F.T.; Data annotation: I.P. and P.M.; Software and validation: I.P., P.M., L.R. and G.M.; Formal analysis: I.P.; Resources: A.V, M.C., and F.N.d.S.; Writing—original draft preparation: I.P.; Writing—review and editing: P.M., L.R., G.M., R.M.C., F.T., A.V, M.C., and F.N.d.S.; Supervision: A.V, M.C., and F.N.d.S.. All authors have read and agreed to the published version of the manuscript.

Funding

This project has received funding from the European Union’s Horizon Europe research and innovation programme under the Grant Agreement 101182954.

Isabel Pinheiro was funded by the Portuguese Foundation for Science and Technology (FCT) and European Social Fund (ESF) through the scholarship SFRH/2025.02081.BD.

Germano Moreira was funded by the Portuguese Foundation for Science and Technology (FCT) and European Social Fund (ESF) through the scholarship SFRH/2022.09726.BD.

Leandro Rodrigues was funded by the Portuguese Foundation for Science and Technology (FCT) and European Social Fund (ESF) through the scholarship SFRH/2023.01424.BD.

Rui M. Coutinho was funded by the Portuguese Foundation for Science and Technology (FCT) and European Social Fund (ESF) through the scholarship SFRH/2022.09503.BD.

Competing Interests

The authors declare no competing interests.

References

- [1] Parham Jafary, Anna Bazangeya, Michelle Pham, Lesley G. Campbell, Sajad Saeedi, Kouros Zareinia, and Habiba Bougherara. Raspberry phenoset: A phenology-based dataset for automated growth detection and yield estimation, 2024.
- [2] Laura-Sophia von Hirschhausen, Jannes S. Magnusson, Mykyta Kovalenko, Fredrik Boye, Tanay Rawat, Peter Eisert, Anna Hilsmann, Sebastian Pretzsch, and Sebastian Bosse. Applegrowthvision: A large-scale stereo dataset for phenological analysis, fruit detection, and 3d reconstruction in apple orchards, 2025.
- [3] Vasileios Sitokonstantinou, Alkiviadis Koukos, Ilias Tsoumas, Nikolaos S Bartsotas, Charalampos Kontoes, and Vassilia Karathanassi. Fuzzy clustering for the within-season estimation of cotton phenology. *PLOS ONE*, 2023.
- [4] Karim Sabry, Jonas Anderegg, Norbert Kirchgessner, Helge Aasen, and Walter Kattenborn. Tomatodiff: On-plant tomato segmentation with denoising diffusion models. *arXiv preprint arXiv:2408.06296*, 2024.
- [5] Anne Katrine Mortensen, Mads Dyrmann, Henrik Karstoft, Rasmus Nyholm Jørgensen, and René Gislum. Growliflower: An image dataset for cauliflower growth stage classification. *Data in Brief*, 11:240–244, 2017.
- [6] A. D. Richardson, C. Schädel, O. Zimmerman, P. J. Giambuzzi, E. T. Cottrell, D. Watkins, T. Milliman, K. J. Pearson, M. P. Williams, M. B. Krassovski, and P. J. Hanson. Spruce vegetation phenology in experimental plots from phenocam imagery, 2015–2024, 2025.
- [7] Andrew D. Richardson, Koen Hufkens, Tom Milliman, Donald M. Aubrecht, Morgan E. Furze, Bijan Seyednasrollah, Miroslav B. Krassovski, Jennifer M. Latimer, W. Ral Nettles, and Robert R. Heidinger. Tracking vegetation phenology across diverse north american biomes using phenocam imagery. *Scientific Data*, 5:180028, 2018.
- [8] Mads S. Rasmussen, Rasmus Fensholt, Pietro Ceccato, and Simon R. Proud. A plant phenology index based on satellite-observed surface climate data from sentinel-2 and landsat-8. *Remote Sensing of Environment*, 247:111934, 2020.
- [9] Corey T. Callaghan, Jodi J. L. Rowley, William K. Cornwell, Alistair G. B. Poore, and Richard E. Major. A dataset of species occurrence records from citizen science platforms. *Scientific Data*, 6:180278, 2019.
- [10] I. Pinheiro, G. Moreira, S. Magalhães, A. Valente, M. Cunha, and F. Santos. Actinidia chinensis cv. 'hayward' flower dataset 2024, 2024. Data set.
- [11] Isabel Pinheiro. Multi-Modal Actinidia chinensis Phenology Dataset (Portugal, 2023-2024), 2025.
- [12] Isabel Pinheiro, Germano Moreira, Sandro Magalhães, António Valente, Mário Cunha, and Filipe Santos. Actinidia chinensis cv. 'hayward' flower dataset 2024, 2024. Available at <https://doi.org/10.5281/zenodo.13692223>.
- [13] I. Pinheiro, P. Moura, L. Rodrigues, G. Moreira, S. Magalhães, A. Valente, M. Cunha, and F. Santos. Actinidia chinensis bud and flower dataset (portugal, 2024), 2025. Available at <https://doi.org/10.5281/zenodo.17178632>.
- [14] B. E. Moore and J. J. Corso. Fiftyone. *GitHub*. Note: <https://github.com/voxel51/fiftyone>, 2020.
- [15] Farzaneh Zangenehnejad and Yang Gao. GNSS observation generation from smartphone android location API: Performance of existing apps, issues and improvement. *Sensors*, 23(2):777, 2023.
- [16] M.C. Salinero, P. Vela, and M.J. Sainz. Phenological growth stages of kiwifruit (actinidia deliciosa 'hayward'). *Scientia Horticulturae*, 121(1):27–31, 2009.
- [17] I. Pinheiro, G. Moreira, S. Magalhães, A. Valente, M. Cunha, and F. N. dos Santos. Deep learning based approach for actinidia flower detection and gender assessment. *Scientific Reports*, 14(1):24452, 2024.

- [18] Peter Lancashire, Hermann Bleiholder, T. Boom, P. Langeluddeke, R. Stauss, Elfriede Weber, and A. Witzemberger. A uniform decimal code for growth stages of crops and weeds. *Annals of Applied Biology*, 119:561 – 601, 02 2008.
- [19] Elisabeth Koch, Ekko Bruns, Frank-M Chmielewski, Claudio Defila, Wolfgang Lipa, and Annette Menzel. Guidelines for plant phenological observations. *World Climate Data and Monitoring Programme*, 01 2007.
- [20] CVAT.ai Corporation. Computer vision annotation tool (cvat). Software, 2023. Available at <https://doi.org/10.5281/zenodo.4009388>.
- [21] Jacob Cohen. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1):37–46, 1960.