Postharvest Biology and Technology 243 (2027) 114707

Contents lists available at ScienceDirect

Postharvest Biology and Technology

journal homepage: www.elsevier.com/locate/postharvbio

Applying convolutional neural networks to X-ray images to improve postharvest detection of fruit fly infestation

Francesca Scolari [a,*], Leani Serfontein [b], Wayne Kirkman [b], Sean Moore [b,c], Jakobus Hoffman [d], Lunga Bam [d], Andrea Moyano [a], Guy F. Sutton [c], Claudio Pezzoni [e], Grzegorz J. Możdżyński [e,f], Dario Pescini [e], Anna Sandionigi [e,g], Aruna Manrakhan [b,h,**]

[a] Institute of Molecular Genetics IGM CNR 'Luigi Luca Cavalli-Sforza', Pavia, Italy
[b] Citrus Research International, Mbombela, South Africa
[c] Centre for Biological Control, Department of Zoology and Entomology, Rhodes University, Makhanda, South Africa
[d] South African Nuclear Energy Corporation (NECSA), Pelindaba, South Africa
[e] Department of Statistics and Quantitative Methods, University of Milano-Bicocca, Milan, Italy
[f] Poznan University of Technology, Poznan, Poland
[g] Quantia Consulting Srl, Milan, Italy
[h] Department of Conservation Ecology and Entomology, Stellenbosch University, Stellenbosch, South Africa

ARTICLE INFO

Dataset link: Convolutional neural networks and X-ray analysis for fruit fly infestation detection

Keywords:
Mediterranean fruit fly
Ceratitis capitata
X-ray computed tomography
Deep learning
Postharvest pest detection
Citrus
Automated inspection

ABSTRACT

Early detection of fruit fly (Diptera: Tephritidae) infestation in traded fruit is important for maintaining phytosanitary security, particularly for quarantine pests. This study evaluated the integration of micro-focus X-ray computed tomography (μCT) and deep learning as a non-destructive, automated, and scalable solution for postharvest pest detection. The Mediterranean fruit fly (medfly), *Ceratitis capitata* Wiedemann, 1824, was used as a model species and Navel oranges, *Citrus sinensis* (Osbeck), served as the host commodity. Controlled infestation experiments were conducted using both artificially punctured fruit to promote oviposition and intact fruit to avoid confounding effects associated with mechanical damage. Oranges were exposed to varying adult medfly densities to generate different infestation levels. Following exposure, the infested oranges were subjected to varying incubation periods to simulate different infestation stages. High-resolution μCT scans were first assessed manually and subsequently analyzed using a staged deep learning pipeline incorporating convolutional neural networks (CNNs), 3D CNNs, CNN-Long Short-Term Memory (LSTM) hybrids and Transformer-based models. Manual μCT inspection achieved mean correct classification rates of 0.84 in punctured fruit and 0.69 in intact fruit, with detection performance increasing with infestation stage. Deep learning analyses were conducted on a curated punctured-fruit dataset using a slice-level classification and voting-based aggregation strategy. Among the evaluated transfer-learning architectures, DenseNet121 achieved the best overall performance, reaching an accuracy of 0.88, a precision of 0.86, a sensitivity of 1.00, and an AUC of 0.80; because the manual and automated analyses were performed on different sets of fruit, these accuracies are complementary rather than directly comparable.

Grad-CAM and attention visualizations confirmed that model predictions were driven by biologically meaningful tissue regions associated with larval damage. Although the deep learning workflow was developed and validated only on punctured fruit, the results demonstrate that μCT images contain sufficient information for automated detection of medfly infestation when infestation-associated damage is visible. These findings support the potential of combining μCT imaging and artificial intelligence (AI) for automated, reproducible and non-destructive postharvest pest detection and enhanced phytosanitary compliance in postharvest supply chains.

* Corresponding author.
** Corresponding author at: Citrus Research International, Mbombela, South Africa.
E-mail addresses: francesca.scolari@cnr.it (F. Scolari), leani@cri.co.za (L. Serfontein), waynek@cri.co.za (W. Kirkman), seanmoore@cri.co.za (S. Moore), jakobus.hoffman@necsa.co.za (J. Hoffman), lunga.bam@necsa.co.za (L. Bam), andreamyn19@gmail.com (A. Moyano), g.sutton@ru.ac.za (G.F. Sutton), c.pezzoni4@campus.unimib.it (C. Pezzoni), g.mozdzynski@campus.unimib.it (G.J. Możdżyński), dario.pescini@unimib.it (D. Pescini), anna.sandionigi@quantiaconsulting.com (A. Sandionigi), aruna@cri.co.za (A. Manrakhan).

https://doi.org/10.1016/j.postharvbio.2026.114707
Received 1 April 2026; Received in revised form 23 August 2026; Accepted 4 September 2026
Available online 12 September 2026
0925-5214/© 2026 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Postharvest inspection of fruit commodities forms an integral component in trade, ensuring compliance with phytosanitary requirements of exporting and importing countries. During inspection, samples from fruit consignments are checked for the presence of quarantine pests. At typical ports of entry, methods used to detect internal fruit infestation are destructive and are not always reliable (Adedeji et al., 2020). In several parts of the world, trade of fruit is affected by fruit fly (Diptera: Tephritidae) pests of quarantine importance (White and Elson-Harris, 1992). Early stages of fruit fly infestation are cryptic because eggs and young larvae develop beneath the fruit peel and typically produce few or no visible external symptoms until infestation is well established, by which point fruit are already compromised.

One of the most notorious fruit fly pests is the Mediterranean fruit fly (medfly), *Ceratitis capitata* Wiedemann, 1824 (Diptera: Tephritidae), a species with a wide geographical distribution but still absent from several major fruit growing regions of the world, including parts of North America and large parts of Asia (White and Elson-Harris, 1992). This highly polyphagous species infests more than 300 host plants (Liquido et al., 1991; Papadopoulos et al., 2001), which include many fruit species of commercial importance. Various *Citrus* species are recorded as hosts for medfly (De Meyer et al., 2002) whereas *Citrus limon* (L.) Burman F. cv Eureka has been demonstrated to be a non-host in commercial orchards in South Africa (Manrakhan et al., 2018). The wide geographic distribution of medfly, coupled with a high reproductive capacity, makes it a persistent challenge for fruit production and postharvest supply chains (Papadopoulos et al., 2024). Numerous pest control strategies have been deployed against *C. capitata* to mitigate its impact, including chemical treatments, biological control, mass trapping and the Sterile Insect Technique (SIT) (see (Giunti et al., 2023) for a review). While these approaches have been successful in suppressing pest populations, none can fully eliminate the risk of infested fruit entering postharvest channels. Consequently, the need for highly reliable detection at the inspection stage remains a critical gap in current management strategies. Addressing this challenge requires technological solutions that can be implemented at scale.

X-ray computed tomography (CT) has emerged as a promising tool for non-destructive fruit inspection, allowing the visualization of internal fruit structures in three dimensions (Du et al., 2019; Matsui et al., 2023; Tempelaere et al., 2023; Q. Wang et al., 2024). This imaging modality provides volumetric data that enable the detection of hidden defects such as bruises, rot, or insect damage, which would otherwise escape surface inspection (Arendse et al., 2016; Gao et al., 2025; He et al., 2024; Yu et al., 2024). Computed tomography imaging has a long history of application in agriculture, including the official inspection of grains for internal insect pests since the 1950s, and the evaluation of internal physiological disorders in produce such as apples, potatoes, and avocados (Karunakaran et al., 2004; Milner et al., 1950; Si and Sankaran, 2016; Stermer, 1972). Micro-focus X-ray computed tomography (μCT) is a refined version of this technique, capable of capturing micron-level details non-invasively (Yu et al., 2022). This method has been successfully applied to detect internal insect infestations in citrus and other crops, even in cases where fruit displayed no external symptoms (Alba-Tercedor and Ruano, 2024; Srivastava et al., 2020).

Despite its potential, the wide adoption of CT imaging for postharvest inspection has been limited by the need for manual or semi-automated image interpretation, which restricts throughput, standardization and scalability in commercial settings. Interpreting CT scans requires expertise, is time-consuming and is not easily standardized to meet the high demands of commercial inspection lines. This bottleneck underscores the importance of developing automated analysis pipelines that can translate the rich information content of CT data into fast, reliable, and reproducible detection outcomes suitable for real-world use. In addition to X-ray-based approaches, alternative non-destructive sensing technologies including visible/near-infrared (Vis/NIR) spectroscopy have been explored for the detection of citrus internal infestations, demonstrating promising classification performance when combined with multivariate and machine learning methods (Li et al., 2023). However, such approaches may be limited by the spatial heterogeneity of the infestation and the difficulties of capturing localized damage within the fruit.

Recent advances in artificial intelligence (AI) and deep learning provide an opportunity to address this limitation. Convolutional neural networks (CNNs), three-dimensional CNNs, hybrid deep learning models and attention-based architectures such as Transformers have demonstrated outstanding capabilities in high-dimensional pattern recognition tasks, including the classification of medical scans, industrial imaging and agricultural inspection (Ilesanmi et al., 2024; Kshatri and Singh, 2023; Ma et al., 2024; Wang et al., 2022; H. Wang et al., 2024; Zhang et al., 2024). When applied to CT images of citrus, such models have the potential to identify hidden infestation patterns, distinguish between healthy and infested tissues and even approximate the severity of infestation. Unlike classical feature-based approaches, deep learning methods can automatically extract relevant representations from raw image data, significantly enhancing accuracy and reproducibility. Synthetic data augmentation strategies, including conditional Generative Adversarial Networks (cGANs), have also been explored to address the scarcity of labelled datasets in imaging applications. By generating realistic synthetic training examples, these approaches may improve model robustness, although their effectiveness remains application dependent and requires careful validation (Mert, 2023).

A further difficulty, distinct from the scarcity of labelled data, is that annotation is available for the specimen as a whole, whereas the diagnostic evidence occupies only part of the reconstructed volume. Propagating the specimen-level label to every section introduces label noise, because many sections of an affected specimen contain no evidence of the condition. Weakly supervised formulations, and in particular multiple-instance learning, address this by treating the specimen as a set of sections and learning which of them carry the evidence, with supervision applied only at the specimen level (Ilse et al., 2018). Attention-based variants additionally indicate which sections determined each decision, which is informative when the expected signature is spatially localised.

The integration of μCT imaging and AI-driven analysis represents a significant step toward overcoming the limitations of traditional inspection methods and offering scalability, improved detection accuracy and rapid processing compatible with industrial workflows. By coupling volumetric imaging with advanced algorithms, it is increasingly possible to detect and characterize infestation patterns that elude conventional visual inspection, thereby expanding the potential applications of postharvest imaging technologies (He et al., 2024; Srivastava et al., 2020).

The limitations of existing pest control and inspection methods in preventing medfly and other tephritid pest infestations from entering markets highlight the urgent need for innovative technological solutions able to ensure crop security, maintain export market access and reduce the economic burden on producers and exporters (Ansah et al., 2023; Hadipour-Rokni et al., 2023).

The objective of the present study was to evaluate, for the first time, the potential of combining μCT with deep learning for the non-destructive detection of medfly infestation in citrus. Specifically, we aimed to assess the ability of μCT techniques initially optimized for the detection of another internal citrus pest to reveal fruit fly infestation of different stages and densities. We further sought to investigate whether coupling μCT imaging with AI models could provide an automated and reproducible framework for analyzing volumetric imaging data and supporting postharvest inspection. Our goal was to explore the feasibility, advantages and current constraints of this integrated approach, with a view to support the development of robust postharvest inspection systems for the citrus industry and beyond.

2. Materials and methods

2.1. Insects

Ceratitis capitata flies reared at the Citrus Research Centre, Citrus Research International (CRI), Mbombela, South Africa, were used in this study. The *C. capitata* culture used was established in April 2022 from approximately 600 wild flies reared from coffee berries collected in Burgershall, Hazyview, Mpumalanga Province, South Africa. Two-week-old males and females were used for the fruit infestation experiments.

2.2. Infestation experiments

Flies were placed into three 50 × 50 x 50 cm Perspex boxes at different densities: 50 (25 males and 25 females), 100 (50 males and 50 females) and 500 flies (250 males and 250 females), respectively, to obtain different fruit infestation levels (Fig. 1A). Flies had access to a mixture of sugar and yeast hydrolysate (3:1 ratio) and water *ad libitum*.

Navel oranges, *Citrus sinensis* (Osbeck) cv. Navel, bought from a local supermarket, were used in all tests. Fifteen fruit were placed in each Perspex box and exposed to flies for 24 h for infestation at four time intervals (days) before undergoing X-ray scanning for medfly larval detection: 10 (targeting third instar larvae), 7 (targeting second and third instars), 5 (targeting first and second instars) and 4 (targeting first instars) days, respectively. For each infestation time-point, six fruit were kept in a separate box and not exposed to fruit flies. These unexposed fruit served as uninfested controls for the corresponding incubation period.

To increase likelihood of oviposition, as previously demonstrated for *C. capitata* (Staub et al., 2008), a modified dissection needle was used to puncture three holes (6 mm deep) on three equidistant sides of each fruit (Fig. 1B) before placement in the Perspex boxes. The needle's penetration depth was controlled using a stopper to largely restrict it to the albedo layer. Fruit were punctured in equal distances around the middle of the fruit. The dissecting needle was rinsed in a 0.1% Sporekill® (Didecyldimethylammonium chloride) (ICA International Chemicals (Pty) Ltd, Stellenbosch, South Africa) solution (dilution with water) in between each fruit puncture. The control fruit were also punctured. Each puncture mark was delineated using a black permanent marker for clear identification. At 24 h after fly exposure, the fruit were removed from the Perspex boxes and placed into 20 L containers. Each fruit was wrapped in a paper bag. Containers with fruit (exposed and unexposed) were kept in a room at mean temperature (± standard error) of 24°C ± 2°C. For this experiment, there were three replicates using three different fly cohorts. For punctured fruit, each replicate consisted of 15 oranges per fly-density treatment (50, 100, and 500 flies) and six control oranges for each infestation stage. Across the three biological replicates, this resulted in 540 exposed punctured fruit and 72 punctured control fruit.

A second similar experiment was conducted using intact detached fruit (*i.e.*, without puncturing).

Only the highest fly-density treatment (500 flies) was evaluated for intact fruit because previous studies have shown that *C. capitata* females oviposit less readily in intact citrus than in substrates containing artificial openings (Papachristos and Papadopoulos, 2009). The use of the highest fly density was therefore intended to maximize the likelihood of successful oviposition and obtain sufficient numbers of infested fruit for meaningful evaluation. Inclusion of lower fly densities (50 and 100 flies) was expected to result in very low infestation rates and an insufficient number of positive fruit for analysis.

A total of 250 males and 250 females were placed into one Perspex box and the same experimental set-up described above was followed including incubation of fruit to generate the four infestation time-points. The mean temperature (± standard error) of the room where fruit were incubated was 20°C ± 4°C. Again, there were three replicates for the experiment on intact fruit using three different fly cohorts. Each replicate consisted of 15 exposed fruit and six control fruit.

2.3. Micro-focus X-ray CT Scanning

All fruit were transported in 20 L containers to the South African Nuclear Energy Corporation (NECSA) facility, Pelindaba, North-West Province, South Africa, for scanning using their MIXRAD facility (Nikon XTH 225 ST micro-focus X-ray tomography system) as described in (Hoffman and De Beer, 2012). Scans were performed at 100 kV with the beam operating at 200 μA. Each scan lasted 1 min 26 s following optimisation of the scan duration in a previous trial on *Thaumatotibia leucotreta* (Meyrick) (Lepidoptera: Tortricidae) in oranges (Kirkman and Moore, unpublished results). A total of 125 projections were used in the tomography system. This projection number was selected based on the minimum number of projections required to obtain detectable and observable infestation-associated damage while maintaining practical acquisition times. Detector binning was applied to improve image statistics by combining adjacent detector pixels, thereby increasing the number of detected photons contributing to each effective pixel and improving signal-to-noise ratio (SNR), although at the expense of spatial resolution. This reduction in resolution was considered acceptable because the resulting voxel size allowed visualisation of infestation-associated damage in the reconstructed volumes. The voxel volume of each tomogram was 0.0087 mm^3 (Hoffman and De Beer, 2012).

In CT imaging, projections refer to the individual X-ray images acquired at multiple rotational angles around the fruit and subsequently reconstructed into a three-dimensional tomographic volume.

In the punctured-fruit experiment, six exposed fruit were randomly selected from each fly-density x incubation-period combination in each biological replicate for μCT scanning, resulting in 216 scanned exposed fruit. All 72 punctured control fruit were also scanned, giving a total of 288 scanned punctured fruit across the three biological replicates.

In the trial with intact fruit, all fruit (treatment and control) for each



Fig. 1. Experimental setup for artificial infestation of punctured Navel oranges by medflies. A) The setup consisted of three Perspex boxes, each containing 15 fruit and a different fruit fly density (50, 100, or 500 flies, respectively). B) Photograph of the fruit puncturing procedure using a dissection needle fitted with a stopper to standardize puncture depth at 6 mm.

incubation period were scanned. Each scan corresponded to one individual fruit. During each scan, each fruit was removed from the paper bag. It took us two days to process and scan all the fruit used in the study. To control for continued larval developmental time in the fruit after scanning, each fruit was labelled and placed in individual bags under cool conditions (at approximately 4°C) until dissection to confirm presence of larvae and developmental stages.

All fruit were transported back to CRI for dissection. Of the fruit scanned, the images of one biological replicate and three individual scans from the remaining replicates were not available for computational analysis; the deep learning dataset therefore comprised 189 individually scanned punctured fruit (141 infested, 48 control), corresponding to two complete replicates less three scans. All three replicates contributed to the manual μCT analysis reported in Section 3.1.

2.4. Fruit dissection

Fruit were manually inspected and dissected over two days after scanning (Supplementary Figure S1). The presence of eggs or larvae was recorded for each fruit and the total number of eggs and larvae were recorded. The larvae were removed from the fruit, blanched and preserved in 70% ethanol. The larvae were then identified to instar based on the presence of preapical tooth and sclerotized mandibles (White and Elson-Harris, 1992). The number of larvae at each developmental stage (*i.e.*, first, second and third instars) were separately recorded.

2.5. Manual recognition of fruit fly infestation on X-ray μCT images and data analysis

Each scan generated a three-dimensional tomographic volume composed of 1024 sequential image slices. To facilitate manual inspection, every fifth slice was examined for black areas, resulting in 205 images per fruit for subsequent analysis. Every fifth image was selected to reduce redundancy between adjacent slices while maintaining sufficient spatial coverage of the fruit volume. Patterns of dark areas were then observed and categorised. Healthy fruit typically exhibited a central low-density region corresponding to the natural locular cavity (Supplementary Figure S2A), which was not considered indicative of infestation. In contrast, fruit fly-associated damage was characterised by additional clustered dark areas located in the pulp, often near the rind or distinct from the central cavity (Supplementary Figure S2B) and highlighted through image processing (Supplementary Figure S2C). These features were frequently aligned with visible discontinuities in the surrounding tissue (Supplementary Figure S2).

For each scan, the presence (1) or absence (0) of the fruit fly-associated damage was recorded. The association between fruit fly-associated damage observed manually in scans and presence of fruit fly (larvae or eggs detected during dissection) was determined using a Spearman correlation analysis. No eggs were detected in any dissected fruit. Consequently, the presence of *C. capitata* was defined exclusively by the presence of larvae and the earliest developmental stage represented in the dataset corresponded to first instar larvae.

From the *C. capitata*-associated damage data and the presence of larvae, which were both binary outcomes, sensitivity and specificity analyses were conducted for punctured and intact fruit exposed to flies. These analyses were based on a 2×2 contingency table using larval presence and scan-detected damage. Data were analyzed using XLSTAT (Addinsoft, New York, USA).

To model the effects of fly density and incubation time on fruit fly damage detection rates in X-ray μCT images, while accounting for repeated measurements of fruit originating from the same oviposition cages, data were analysed using generalized linear mixed models (GLMMs; Bolker et al., 2009) fitted with the glmmTMB package (Brooks et al., 2017) in R version 4.5.0 (R Core Team., 2025). The binary detection outcomes (presence or absence of X-ray detected damage) were modelled using a binomial error distribution and a logit link function. To resolve issues of complete separation (e.g., uninfested fruit including control containing zero larvae), the analysis was divided into three separate models. Firstly, fruit damage detection was evaluated across the full dataset to determine the probability of recording damage regardless of whether larvae were present or absent. Secondly, the false positive rate was evaluated on the uninfested fruit subset, where a false positive was defined as X-ray damage being recorded when larvae were absent. Lastly, the true positive rate was evaluated on the infested fruit subset, defining a true positive as X-ray damage being recorded for fruit exposed to fruit flies where fruit fly presence was confirmed by dissection.

For punctured fruit, fixed effects included fly density, incubation duration and their interaction. For intact fruit, fixed effects included fly exposure (unexposed control versus exposure to 500 flies), incubation duration and their interaction. In cases where low sample size caused convergence failure, the interaction term was omitted and an additive model was fitted. Oviposition cage was specified as a random intercept to account for non-independence of data collected over time (Bolker et al., 2009). In cases where low sample sizes caused model convergence failure, the multiplicative term was dropped, and only the additive model was fitted.

Model overdispersion was evaluated by computing dispersion ratios and *P*-values using simulated quantile residuals and Monte Carlo simulation using the 'performance' R package. Model diagnostics confirmed that extra-binomial dispersion was absent across all fitted models (Table S1). Wald chi-square tests were used to determine the significance of fixed effects using the 'car' R package (Fox and Weisberg, 2019). Post-hoc estimated marginal means and pairwise contrasts, either utilising Tukey or Dunnett adjustments for multiple comparisons, were computed using the 'emmeans' R package (Lenth, 2022).

Figures showing the numbers of different larval instars detected across fly densities and incubation periods were generated using Python 3.

2.6. Automatic recognition of fruit fly infestation on X-ray images and data analysis

2.6.1. Preprocessing of X-ray μCT images

To ensure the suitability of the X-ray μCT images for deep learning analysis and maximize the robustness of model training, a comprehensive preprocessing pipeline was implemented. Deep learning analyses were conducted exclusively on the punctured-fruit dataset because this experiment produced substantially higher infestation prevalence and more pronounced μCT-visible damage than the intact-fruit experiment, providing a sufficiently informative dataset for supervised model development.

Each fruit sample was represented by a stack of 72 vertical slices extracted from the reconstructed tomographic volume. The tomographic volume refers to the three-dimensional reconstruction of the fruit generated from the series of X-ray projections acquired during scanning, whereas the vertical slices correspond to two-dimensional cross-sectional images extracted sequentially through this volume and used as input for deep learning analysis.

All grayscale slices were resized to 128×128 pixels and normalized to a [0–1] range to standardize the input space across samples. Background removal was performed using Otsu's adaptive thresholding, followed by row-wise filtering to exclude non-informative slices, such as those containing only peel or air. Bounding boxes were subsequently applied to crop and centre the fruit region while preserving the original aspect ratio. To correct for minor positional asymmetries along the longitudinal axis, a thin black column was introduced at the fruit's central axis when necessary. This procedure reduced morphological variability and minimized imaging artefacts arising from interpolation and minor alignment errors.

All preprocessed slices belonging to the same fruit were grouped to enable both slice-level and volumetric analyses. The imaging repository

available for computational analysis contained more than 750,000 μCT slices derived from the punctured- and intact-fruit experiments. This repository was used during the exploratory stages of model development, including multiclass classification, volumetric architectures, sequential models and preliminary slice-level transfer-learning experiments. For final model development and evaluation, a curated dataset of 111 punctured fruit was assembled through manual review of the scans. Samples with corrupted or low quality images, inconsistent annotations, or other data-quality issues were excluded prior to model training. All slices originating from the same fruit were retained as a single specimen-level unit to support slice-level classification and subsequent fruit-level prediction through voting-based aggregation.

2.6.2. Deep learning

Deep learning analyses were conducted exclusively on the punctured-fruit cohort. The intact-fruit dataset was not used for model training or evaluation because of its low infestation prevalence, weak μCT signal and limited suitability for supervised learning, as discussed in the Results and discussion.

The deep learning workflow was developed iteratively through four successive task formulations: A) multiclass slice-level classification, in which individual μCT slices were assigned to one of the nine experimental classes; B) fruit-level binary classification, in which the complete stack of 72 slices was used to classify fruit as infested or non-infested; C) slice-level binary classification with fruit-level voting aggregation, in which individual slices were classified as infested or non-infested and the resulting predictions combined to generate a fruit-level classification; and D) fruit-level weakly supervised classification, in which the 72 sections of a fruit are treated as a set and supervision is applied only at the level of the whole fruit.

Formulations A and B served as exploratory stages that informed subsequent methodological refinements. Formulation C was retained as the pipeline of the original workflow and is reported in Sections 3.2.2–3.2.5. Formulation D was developed subsequently, to address the label noise inherent in slice-level supervision without recourse to the section-level curation that formulation C required; it is described in Section 2.6.3 and reported in Section 3.2.7.

The curated dataset of 111 punctured fruit was partitioned at the level of the individual fruit into a development subset of 77 fruit and a held-out test subset of 34 fruit (24 infested, 10 control), corresponding to the stated 70:30 proportion. All 72 sections of a given fruit were assigned to the same subset, so that no fruit contributed sections to more than one subset and no fruit in the test subset contributed sections to model development.

Within the development subset, a validation partition was defined during training by reserving the final 10% of the section-level training array (Keras validation_split = 0.10). Because the array was ordered by fruit and was not shuffled beforehand, this validation subset corresponded approximately to the sections of the final eight fruit in the ordered development array, although the boundary was defined at the section level rather than explicitly at the fruit level. It was used for hyper-parameter optimisation (Optuna, objective: validation AUC) and for early stopping (patience = 3 epochs on validation AUC). The architecture-specific voting threshold was selected by maximisation of the F1-score. The test subset was reserved for the final evaluation reported in the Results and discussion. Section 2.6.3 reports an analysis in which all three partitions are defined explicitly at the level of the fruit and their composition is given in full.

For formulation A, AlexNet and InceptionV3 (Krizhevsky et al., 2012; Szegedy et al., 2016) were trained for multiclass classification using the Adam optimizer (learning rate = 1×10^{-4}, weight decay = 1×10^{-5}) for 30 epochs. Performance was assessed using a weighted F1-score to account for class imbalance.

For formulation B, three architectures were evaluated: a 3D CNN treating the 72-slice stack of each fruit as a volumetric tensor, allowing the model to learn spatial dependencies across adjacent slices; a CNN–Long Short-Term Memory (LSTM) hybrid extracting spatial features from each slice with 2D convolutional layers and passing the resulting sequence to LSTM layers that modelled inter-slice dependencies; a pure RNN–LSTM model which omitted the convolutional stage and processed the slice sequence as a time series. Performance was assessed using the area under the receiver operating characteristic curve (AUC).

For formulation C, transfer learning was applied using ImageNet-pretrained backbones (EfficientNetB0, MobileNetV2, ResNet50, NASNetMobile, DenseNet121 and InceptionV3) (Rajendra and Wardhana, 2025). Individual slices were classified as infested or non-infested using binary cross-entropy loss and the Adam optimizer. Fruit-level predictions were obtained through a voting strategy in which a fruit was classified as infested when the number of positively classified slices exceeded an architecture-specific threshold.

Voting thresholds were selected by maximisation of the F1-score and subsequently applied unchanged to the test subset. Evaluation used this single stratified partition; no cross-validation was performed, and the metrics are accordingly reported as single-split point estimates with 95% confidence intervals. Section 2.6.3 describes a further analysis in which the threshold-selection procedure is specified in full and the accompanying code released (see Data statement).

Models were implemented in Python and trained using the Adam optimizer with binary cross-entropy loss. Early stopping (patience = 3 epochs) based on validation AUC was used to reduce overfitting. Training was configured to use GPU acceleration when available.

As an exploratory extension, slice sequences were converted into short image sequences and analysed using a Transformer-based architecture. Features were extracted using an ImageNet-pretrained DenseNet121 backbone and processed by a Transformer encoder. Because the available dataset was limited, an Auxiliary-Classifier Generative Adversarial Network (ACGAN) was used to generate additional training samples. The resulting multiclass experiment is reported separately in Section 3.2.6.

Model performance was evaluated using weighted F1-score, accuracy, precision, recall, confusion matrices, and receiver operating characteristic (ROC) curves with the corresponding area under the curve (AUC). Grad-CAM visualizations (Selvaraju et al., 2017) were generated for the final transfer-learning models to provide model interpretability and biological validation.

Proportion-based performance metrics are reported with Wilson score 95% confidence intervals, which are appropriate at the sample sizes involved; AUC and F1 are reported as point estimates for the benchmark analysis. For comparisons between formulations evaluated on the same held-out fruit, AUCs were compared as a paired quantity using percentile bootstrap confidence intervals obtained by resampling fruit, with each bootstrap replicate recomputing both AUCs on the same resample.

2.6.3. Formulation D: fruit-level weakly supervised classification

Formulations A and C supervise at the level of the individual section, which requires every section of an infested fruit to be labelled as infested. In most infested fruit only a few of the 72 sections intersect a larval gallery, so a substantial proportion of the training sections carry a positive label without containing any evidence of infestation. The section-level curation described in Section 2.6.1 was introduced to suppress this label noise. The same difficulty, expressed between fruit rather than between sections, is the reason the intact-fruit cohort could not support supervised learning (Section 3.3). Because that curation depends on visual judgement and could not be applied on an inspection line, we formulated the task so that the noise does not arise.

In a multiple-instance formulation the unit of supervision is the whole specimen. Each fruit is treated as a set of 72 sections carrying a single fruit-level label, and the model learns both which sections are informative and how to classify the fruit from them; no section is ever assigned a label of its own. This matches the structure of the problem, in

which infestation is a property of the fruit while its visible signature is confined to a few sections. We used the gated-attention formulation of Ilse et al. (2018), in which a learned weight is assigned to each section and the fruit-level representation is the attention-weighted mean of the section representations; the weights are interpretable as the contribution of each section to the decision.

An ImageNet-pretrained MobileNetV2 backbone mapped each section to a feature vector, from which the attention layer produced the fruit-level representation and a sigmoid output. Training proceeded in two stages: the attention layer was first trained with the backbone frozen (learning rate 1×10^{-3}), after which the last 20 non-normalisation layers of the backbone were unfrozen and training continued at a reduced learning rate (3×10^{-5}). Batch-normalisation layers were kept frozen throughout, the loss was weighted by inverse class frequency, and early stopping monitored validation AUC in both stages; whichever stage achieved the higher validation AUC was retained.

For deep-learning analysis, only image files that were successfully transferred and technically suitable for computational processing were included. Because no additional section-level curation was performed, this analysis used the resulting archive of 189 punctured-fruit scans (141 infested, 48 control; 13,608 sections). Unlike formulation C, which was developed on a manually curated subset of 111 punctured fruit to minimise section-level label noise, formulation D operated directly on the larger archive of technically suitable scans without section-level curation. Consequently, the sample sizes differ between the two formulations because they address different methodological objectives. Fruit were partitioned at the fruit level into training, validation and test subsets (105 / 27 / 57 fruit; 78 / 20 / 43 infested and 27 / 7 / 14 control), stratified by infestation status. The decision threshold was selected on the validation fruit by maximising balanced accuracy and then applied unchanged to the held-out test fruit. For evaluation on the fixed held-out test partition, two operating points were retained: one maximising balanced accuracy and one maximising F1-score, both selected exclusively on the validation fruit and subsequently applied unchanged to the test fruit. In addition, stratified group k-fold cross-validation (5 folds, grouped by fruit) was used to estimate performance variability. Within each cross-validation fold, the training portion was further divided into an inner training and validation subset (20% of the training fruit used for validation). The decision threshold was selected exclusively on this inner validation subset by maximising balanced accuracy and then applied unchanged to the corresponding held-out fold.

3. Results and discussion

3.1. Manual μCT inspection effectively reveals infestation patterns, particularly in punctured fruit

Manual evaluation of μCT scans showed a strong, positive, and statistically significant correlation between *C. capitata*-associated damage and larval presence in dissected punctured fruit (Spearman's $\rho = 0.645$; 95% CI: 0.565–0.714; $p < 0.0001$). The overall proportion of correctly classified punctured fruit was 0.84 (95% CI: 0.79–0.88). The sensitivity of the system was estimated at 0.83 (95% CI: 0.77–0.87), indicating a strong ability to correctly identify infested fruit. The specificity was 0.86 (95% CI: 0.77–0.92), reflecting the system's effectiveness in correctly identifying non-infested fruit.

Fruit fly damage detection rates were significantly higher in punctured fruit exposed to flies at all densities (50, 100, and 500 flies; all $P < 0.001$) than in unexposed fruit (Fig. 2A). However, detection rates in punctured fruit did not differ significantly among the three fly densities. Incubation duration was also a significant positive predictor of fruit fly damage detection in punctured fruit ($\chi^2 = 4.97$, df = 1, $P = 0.026$). The probability of detecting damage increased over time, rising from 15% on day 4–23% on day 7, and reaching 26% by day 10 (Fig. 2B). The interaction between fly density and incubation time was not statistically significant ($\chi^2 = 1.24$, df = 3, $P = 0.743$), indicating that the effect of incubation days on detection rates did not differ among fly densities. Fly density had no statistically significant effect on the probability of a false positives in punctured fruit ($\chi^2 = 6.44$, df = 3, $P = 0.092$) (Fig. 2C). In contrast, incubation duration was a statistically significant driver of false positive rates ($\chi^2 = 5.89$, df = 1, $P = 0.015$) (Fig. 2D). On average, probability of a false positive on an uninfested punctured fruit increased over time, increasing from 33% on day 4–51% on day 7, and reaching 69% by day 10 (Fig. 2D).

Furthermore, for fruit in which larvae were confirmed present by dissection (n = 178), fly density did not significantly correlate with fruit fly damage detection rates (true positive rate) ($\chi^2 = 4.81$, df = 2, $P = 0.090$; Fig. 2E). Similarly, the true positive rate was not significantly affected by incubation duration ($\chi^2 = 0.22$, df = 1, $P = 0.637$; Fig. 2F), and the interaction between fly density and incubation duration was not statistically significant ($\chi^2 = 2.91$, df = 2, $P = 0.234$; Fig. 2F). Larval abundance increased with adult fly density, with the highest larval counts generally observed in fruit exposed to the highest fly density (Fig. 3). The early (day 4) and early-mid stages (day 5) of infestation had predominantly second instar larvae, whilst the mid-late (day 7) and late stages (day 10) of infestation had predominantly third instar larvae (Fig. 4).

The increased detectability of infestation with advancing infestation stage observed in this study has also been previously recorded in punctured fruit of different fruit types, including citrus, exposed to *Bactrocera dorsalis* (Hendel) and X-ray scanned (Chuang et al., 2011). This high detectability observed in advanced infestation stages for fruit flies is likely due to the creation of larger feeding galleries by the third instar larvae within the fruit compared to those formed by earlier larval stages (Alba-Tercedor and Ruano, 2024). Similar trends have been reported in earlier X-ray imaging studies on *B. oleae* infestation, where detection accuracy increased markedly with the severity of internal damage, while early infestations were more difficult to identify due to the limited tissue alteration (Haff et al., 2016). Since darkened areas visible in X-ray images result from differences in tissue density between infested and non-infested areas within the fruit pulp (Chuang et al., 2011), larger galleries would increase the visibility of these contrasts in the images. For unexposed punctured fruit incubated for 10 days acting as a control for the latest stage of infestation, there was a high rate of false positives (Fig. 2D), likely due to advanced rot that may have been visually confounded with fruit fly–associated damage. Although fruit fly-induced damage may overlap with pathogen-associated decay within the fruit, this may not necessarily represent a limitation of the system. This source of error is likely to bias the system toward precautionary detection rather than allowing infested fruit to escape inspection. The frequency of false positives may be less problematic than false negatives because the fruit affected by rot or other internal damage would generally be rejected from export or marketing channels irrespective of the underlying cause. Consequently, where identification of the specific cause of tissue degradation is required, fruit classified as positive could subsequently be subjected to confirmatory dissection or diagnostic examination.

In exposed intact fruit, in contrast, detection performance was weakly correlated with true infestation (Spearman's $\rho = 0.317$; 95% CI: 0.195–0.549; $p < 0.0001$). The X-ray detection system was less sensitive but more specific when applied to intact fruit compared to punctured fruit, with images reviewed manually. Sensitivity was at 0.26 (95% CI: 0.18–0.35), indicating limited ability to detect infested fruit. Specificity was high at 0.96 (95% CI: 0.91–0.98), demonstrating strong performance in correctly identifying non-infested fruit. The mean proportion of correct classifications of medfly damage in infested intact fruit was 0.69 (95% CI: 0.63–0.74). However, the infestation rate of intact fruit (17.44%) were considerably lower than that of punctured fruit, which may partly explain the lower detection performance observed during manual image analysis.

In intact fruit, fly density had no statistically significant effect on

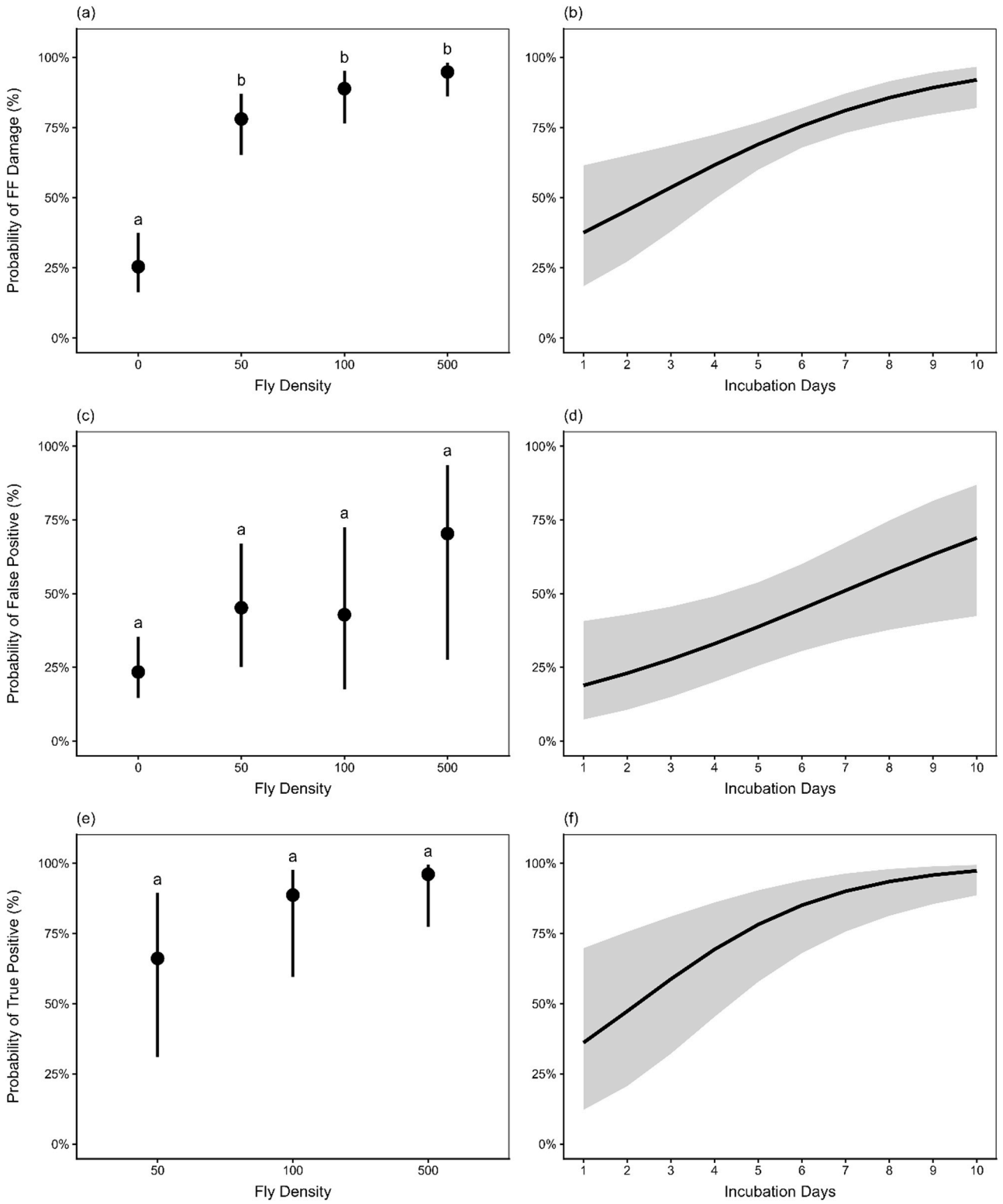


Fig. 2. X-ray detection performance on artificially punctured citrus fruit as a function of fly exposure density and incubation duration. *Top row*: Fruit damage detection probability across fly density treatments (A) and incubation days (B). *Middle row*: False positive probability (specificity) across fly density treatments (C) and incubation days (D). *Bottom row*: True positive probability (sensitivity) across fly density treatments (E) and incubation days (F). In panels A, C, and E, solid points and vertical bars represent estimated marginal means and 95% confidence intervals derived from binomial GLMMs. Density treatment means sharing a common letter are not statistically different based on post-hoc adjusted pairwise comparisons ($P > 0.05$). In panels B, D, and F, solid black lines and grey shaded ribbons denote mean expected regression curves and 95% confidence intervals across the 10-day incubation period.

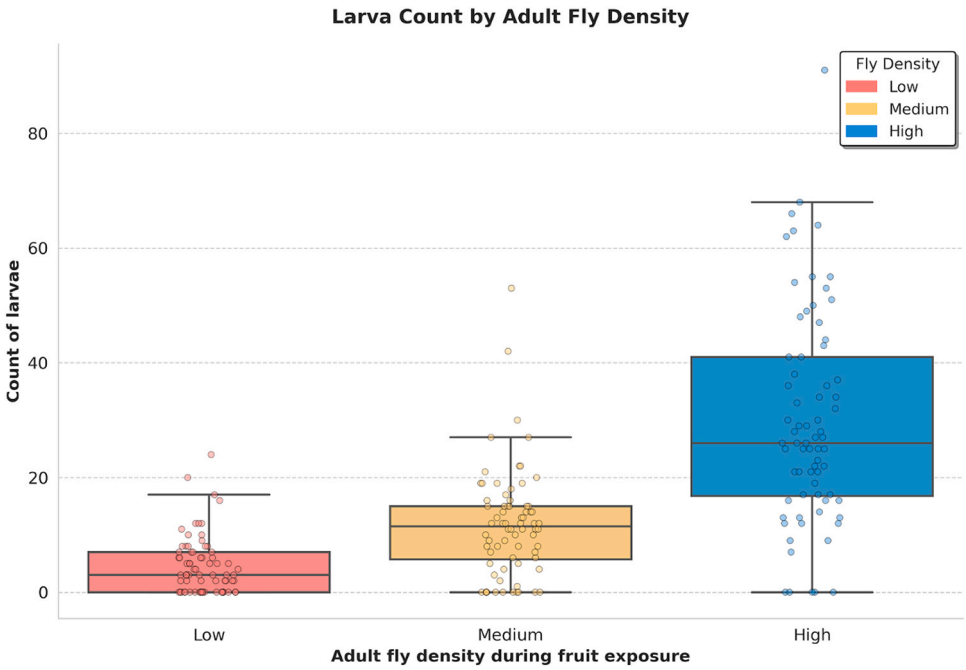


Fig. 3. Box-and-whisker plots showing *C. capitata* larval counts in punctured Navel oranges exposed to three fly-density treatments: low (50 flies), medium (100 flies) and high (500 flies). The horizontal line within each box represents the median value. The boxplots are presented for descriptive visualization of larval abundance only. Statistical inference regarding the effects of infestation stage and fly density on μCT detection performance was performed using GLMMs.

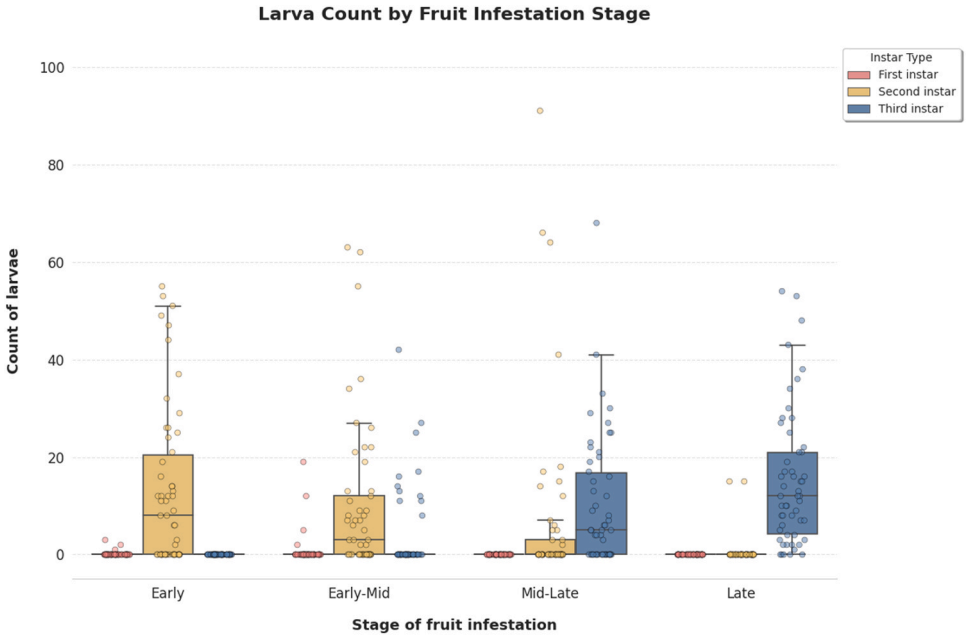


Fig. 4. Box-and-whisker plots showing counts of three developmental stages of *C. capitata* larvae (first, second and third instar) in punctured Navel oranges at four infestation stages corresponding to fruit incubation periods of 4 days (Early), 5 days (Early-Mid), 7 days (Mid-Late) and 10 days (Late). The horizontal line within each box represents the median value. The boxplots are presented for descriptive visualization of larval developmental-stage distributions only. Statistical inference regarding the effects of infestation stage and fly density on μCT detection performance was performed using GLMMs.

damage detection probabilities (Fig. 5A). Control fruit (0 flies) had a fruit damage detection rate of 35% (95% CI: 18–56%), which was similar to the 41% (95% CI: 24–61%) observed in infested fruit (500 flies) (Fig. 5A). Incubation duration was also not a statistically significant predictor of fruit damage detection rate in intact fruit ($X_2 = 0.23$, df $= 1$, $P = 0.631$). The probability of detecting damage on an unpunctured fruit remained relatively stable over time, holding at 38% on day 4, 37% on day 7, and 37% by day 10 (Fig. 5B). The interaction between fly density and incubation days was also not statistically significant ($X_2 = 0.54$, df $= 1$, $P = 0.463$). Fly density had no statistically significant effect on the probability of a false positive ($X_2 = 0.04$, df $= 1$, $P = 0.837$). The probability was comparable between control fruit (0 flies), which had a false positive rate of 35% (95% CI: 22–51%), and infested fruit (500 flies), which had a false positive rate of 33% (95% CI: 22–47%) (Fig. 5C). Similarly, incubation duration was not a statistically significant driver of false positive rates ($X_2 = 0.55$, df $= 1$, $P = 0.460$). On average, the

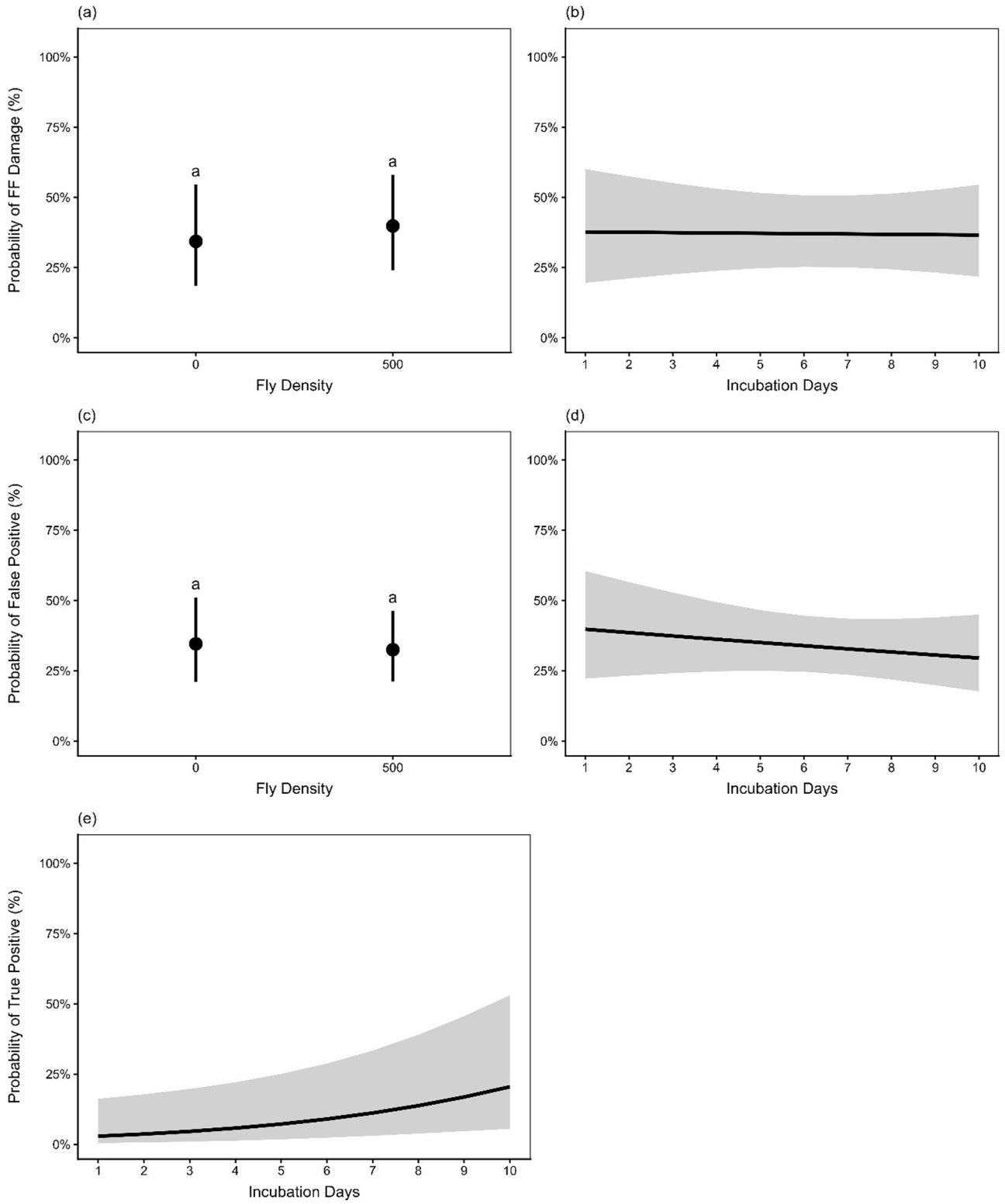


Fig. 5. X-ray detection performance on intact (unpunctured) fruit as a function of fly exposure density and incubation duration. *Top row*: Fruit damage detection probability across fly density treatments (A) and incubation days (B). *Middle row*: False-positive probability across fly density treatments (C) and incubation days (D). *Bottom row*: True-positive probability (sensitivity) across incubation days at 500 fly density (E). In panels A and C, solid points and vertical bars represent estimated marginal means and 95% confidence intervals derived from binomial GLMMs. Density treatment means sharing a common letter are not statistically different based on post-hoc adjusted pairwise comparisons ($P > 0.05$). In panels B, D, and E, solid black lines and grey shaded ribbons denote mean expected regression curves and 95% confidence intervals across the 10-day incubation period.

probability of a false positive on an uninfested intact fruit remained relatively steady over time, shifting slightly from 36% on day 4–33% on day 7, and reaching 30% by day 10 (Fig. 5D).

For confirmed infested intact fruit (n = 30), incubation duration was not a statistically significant predictor of true positive rates ($X_2 = 1.14$, df = 1, $P = 0.285$; Fig. 5E). In contrast to the punctured fruit, the early stage of infestation in the intact fruit had predominantly first instar larvae, whilst the early-mid stage of infestation in the intact fruit had predominantly second instar larvae (Fig. 6). There were both first instar and late instar larvae in the mid-late stage of infestation, whilst the late stage of infestation had predominantly third instar larvae (Fig. 6).

3.2. Deep learning enables automated μCT-based detection by identifying localized infestation signatures

3.2.1. Dataset characteristics and model development

The deep learning analysis was developed through a series of exploratory modelling stages that progressively refined both the classification task and the representation of the μCT data. These stages were designed to identify an appropriate strategy for automated detection of fruit fly infestation while accounting for the heterogeneous spatial distribution of infestation-related damage within individual fruit. The performance of the exploratory models is summarized in Table 1.

The complete imaging repository comprised 1360 μCT scans corresponding to 767,000 individual image slices derived from both punctured and intact fruit experiments. While this repository captured substantial variability in fruit anatomy, infestation severity and imaging conditions, it also highlighted methodological challenges associated with class imbalance, localized infestation signals and inconsistencies in image quality and annotation reliability. These observations informed the progressive refinement of the modelling workflow.

Multiclass slice-level classification (formulation A) was evaluated using AlexNet and InceptionV3. Both models achieved weighted F1-scores close to chance level (0.23 and 0.24, respectively; Table 1), indicating that individual slices did not contain sufficient information to reliably discriminate among the nine experimental classes. The combination of severe class imbalance and limited contextual information motivated reformulation of the problem as a binary infestation-detection task.

Fruit-level binary classification (formulation B) was subsequently evaluated using a 3D CNN, a CNN–LSTM hybrid and a pure RNN–LSTM architecture. Performance remained limited, with AUC values ranging from 0.50 to 0.61 (Table 1). These results suggested that assigning a single label to an entire fruit volume was not well suited to infestation damage that is often spatially localized within the fruit. Examination of misclassified samples further revealed the influence of annotation noise and variable image quality, motivating both a dataset-curation step and a shift toward slice-level classification combined with fruit-level aggregation.

An initial round of transfer learning under formulation C was performed on the uncurated dataset to evaluate the feasibility of slice-level binary classification. EfficientNetB0 achieved a per-image AUC of 0.29, whereas MobileNetV2 and ResNet50 achieved per-image AUCs of 0.64 and 0.40 respectively. Following slice aggregation, the corresponding fruit-level AUCs were 0.55 and 0.44 (Table 1). Although these results demonstrated the potential of transfer learning for infestation detection, performance was constrained by label noise and inconsistencies within the uncurated dataset.

A curated punctured-fruit dataset was therefore assembled through manual review of the scans, removing samples with corrupted or low-quality images, inconsistent annotations, or other data-quality issues. This curated dataset was used only for formulation C. Formulation D, described below, was intentionally evaluated on a larger archive without manual section-level curation.

Following curation and optimization of the voting threshold for each architecture, four ImageNet-pretrained backbones (MobileNetV2, NASNetMobile, DenseNet121, InceptionV3; Table 2) were retained for the final benchmark reported in Table 3 and Section 3.2.3. Collectively, the results in Table 2 illustrate the progression from multiclass classification and fruit-level volumetric models to a slice-level transfer-learning approach combined with fruit-level voting aggregation, which provided the most effective framework for automated μCT-based infestation detection.

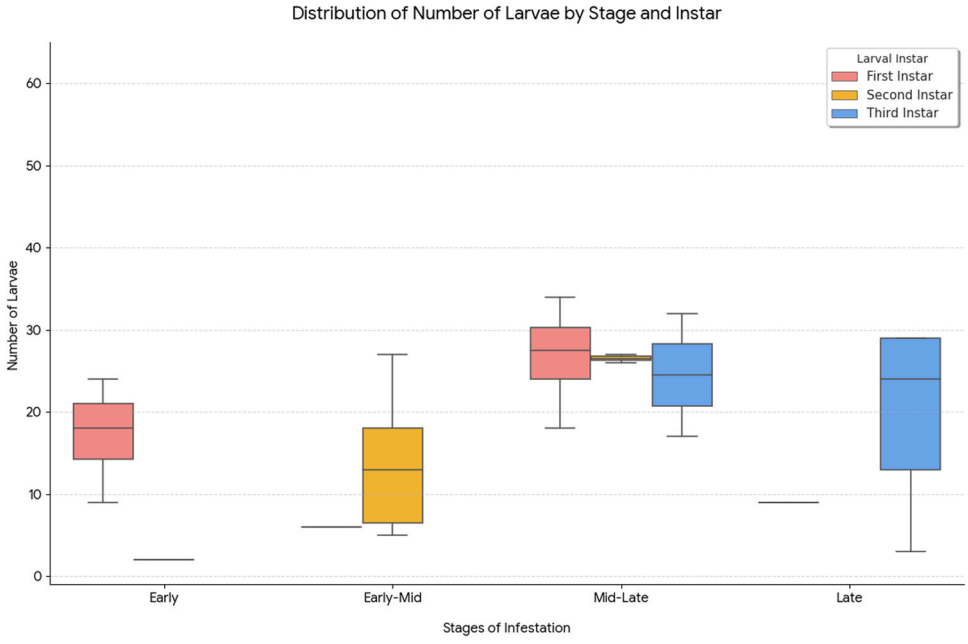


Fig. 6. Box-and-whisker plots showing counts of three developmental stages of *C. capitata* larvae (first, second and third instar) in intact Navel oranges at four infestation stages corresponding to fruit incubation periods of 4 days (Early), 5 days (Early-Mid), 7 days (Mid-Late) and 10 days (Late). The horizontal line within each box represents the median value. The boxplots are presented for descriptive visualization of larval developmental-stage distributions only; statistical inference regarding the effects of incubation period and fly exposure on μCT detection performance was performed using GLMMs.

Table 1
Performance of the exploratory deep learning architectures and task formulations evaluated during development of the automated μCT-based infestation-detection pipeline. Three of the four task formulations described in Section 2.6.2 are represented here: (A) multiclass slice-level classification of the nine experimental groups (control fruit plus combinations of infestation stage and fly density), evaluated using weighted F1-score; (B) fruit-level binary classification using volumetric and sequential architectures, evaluated using area under the receiver operating characteristic curve (AUC); and (C) an initial implementation of slice-level binary transfer learning on the uncurated dataset, evaluated using both per-image and per-fruit AUC following slice aggregation. The final column summarizes the principal observations that motivated subsequent refinement of the analytical pipeline. These exploratory experiments informed the development of the final benchmark reported in Table 2, which was conducted on the curated punctured-fruit dataset using slice-level classification and threshold-optimized voting aggregation; formulation D, developed subsequently to remove the dependence on section-level curation, is reported in Table 3.

Formulation	Model	Metric	Value	Rationale for subsequent pipeline refinement
A. Multiclass slice-level classification	AlexNet	Weighted F1	0.23	Limited discrimination among the nine experimental classes; motivated reformulation to a binary infestation-detection task
A. Multiclass slice-level classification	InceptionV3	Weighted F1	0.24	Performance close to chance level, further indicating limited discrimination among the nine experimental classes
B. Fruit-level binary classification	3D CNN	AUC	0.61	Moderate discrimination, indicating that fruit-level labels were insufficient to capture localized infestation patterns
B. Fruit-level binary classification	CNN-LSTM	AUC	0.57	Limited discrimination, with evidence of reduced generalization performance on the fruit-level dataset
B. Fruit-level binary classification	RNN-LSTM	AUC	0.50	Performance remained at chance level, suggesting that sequential modelling of slice order alone was insufficient to capture diagnostically relevant infestation patterns
C. Slice-level binary transfer learning (uncurated dataset)	EfficientNetB0	AUC (per-image)	0.29	Performance below chance level, indicating substantial label noise and motivating dataset curation
C. Slice-level binary transfer learning (uncurated dataset)	MobileNetV2	AUC (per-image) / (per-fruit)	0.64 / 0.55	Demonstrated the strongest performance among the initial transfer-learning models and was retained for evaluation following dataset curation and voting-threshold optimization
C. Slice-level binary transfer learning (uncurated dataset)	ResNet50	AUC (per-image) / (per-fruit)	0.40 / 0.44	Limited discrimination on the uncurated dataset and therefore not retained for the final benchmark

3.2.2. Slice-level classification and voting strategy

A key methodological improvement emerged from the observation that infestation-related damage was not uniformly distributed across all 72 slices of a given fruit. Instead, visible signs of infestation were often confined to a limited number of slices corresponding to localized regions of tissue damage. To better capture these spatially restricted patterns, a slice-level classification strategy was adopted, as described in Section 2.6.2 (Fig. 7; Supplementary Video S1). Each slice was independently classified as either infested or non-infested and the resulting predictions were subsequently aggregated using a voting threshold to generate a final fruit-level classification. This approach enabled the models to focus on localized infestation signatures while reducing the influence of non-informative slices that contained little or no evidence of infestation. By allowing fruit-level decisions to be based on the accumulation of multiple positive slice predictions rather than a single global assessment of the entire fruit volume, the voting strategy provided a more robust framework for detecting spatially heterogeneous infestation patterns.

3.2.3. Final model comparison

Following dataset curation and implementation of the slice-level classification with voting aggregation strategy described in Section 3.2.2, four ImageNet-pretrained architectures were retained for the final binary detection benchmark: MobileNetV2, NASNetMobile, DenseNet121, and InceptionV3 (Table 2). All models were trained and evaluated under an identical protocol, including the same fruit-level train:test partition, the same slice-level classification framework, and the same threshold-selection criterion, enabling direct comparison of model performance.

MobileNetV2 and NASNetMobile achieved identical performance, with an AUC of 0.70, an accuracy of 0.82, and an F1-score of 0.89 (Table 2). Both models achieved a sensitivity of 1.00 (95% CI 0.86–1.00) and a precision of 0.80, indicating that all infested fruit in the test set were correctly identified, although a small number of non-infested fruit were incorrectly classified as infested.

DenseNet121 achieved the strongest overall performance, reaching an accuracy of 0.88 (95% CI 0.73–0.95), an F1-score of 0.92 and a precision of 0.86. This model therefore achieved a sensitivity of 1.00 (95% CI 0.86–1.00) together with the highest overall classification accuracy among the evaluated architectures. The optimal voting threshold for DenseNet121 was also the highest among the four models (12 slices), indicating that reliable fruit-level predictions were obtained only when infestation-related signals were consistently detected across multiple slices.

InceptionV3 achieved the highest AUC (0.84) and the highest precision (0.91), together with the highest specificity (0.80), consistent with a more conservative decision profile but at the expense of a slightly lower recall (0.88). This performance profile indicates a more conservative classification behaviour, reducing false positives while increasing the likelihood of missed infestations relative to DenseNet121 (Table 2).

Overall, the results demonstrate that transfer learning combined with slice-level classification and voting aggregation can effectively

Table 2
Performance of the four ImageNet-pretrained transfer-learning architectures retained for the final binary infestation-detection benchmark. Models were evaluated on the curated punctured-fruit dataset using the slice-level classification and fruit-level voting aggregation strategy described in Section 3.2.2. All metrics were calculated under an identical evaluation protocol, including the same fruit-level train:test partition, the same slice-level classification framework, and voting thresholds selected independently for each architecture by maximisation of the F1-score. Area Under the Receiver Operating Characteristic Curve (AUC) measures the model's ability to distinguish between infested and non-infested fruit, with higher values indicating better discrimination. F1-score is the harmonic mean of Precision and Sensitivity and is particularly informative when class distributions are imbalanced. Sensitivity (recall, true positive rate) represents the proportion of infested fruit correctly identified by the model. Precision represents the proportion of fruit predicted as infested that were truly infested. Specificity represents the proportion of non-infested fruit correctly identified. Accuracy represents the proportion of all predictions that were correctly classified. Vote threshold indicates the minimum number of slice-level positive predictions (out of 72 slices per fruit) required for a fruit to be classified as infested. Sensitivity, precision, specificity and accuracy are accompanied by Wilson score 95% confidence intervals, which can be derived from the confusion matrix alone. AUC and F1 are reported as point estimates without dispersion, because interval estimates for these two quantities require resampling of the individual fruit-level predictions, which are not retained for this analysis; dispersion for both is reported in Table 3 for the analysis described in Section 2.6.3, where the per-fruit scores are available. Estimates derive from 34 test fruit, of which 10 were non-infested, and the intervals are correspondingly wide. The exploratory architectures and task formulations that preceded this final benchmark are summarized in Table 1 and discussed in Section 3.2.1.

Model	AUC	F1	Sensitivity (95% CI)	Precision (95% CI)	Specificity (95% CI)	Accuracy (95% CI)	Vote Threshold
MobileNetV2	0.70	0.89	1.00 (0.86–1.00)	0.80 (0.63–0.90)	0.40 (0.17–0.69)	0.82 (0.66–0.92)	5
NASNetMobile	0.70	0.89	1.00 (0.86–1.00)	0.80 (0.63–0.90)	0.40 (0.17–0.69)	0.82 (0.66–0.92)	6
DenseNet121	0.80	0.92	1.00 (0.86–1.00)	0.86 (0.69–0.94)	0.60 (0.31–0.83)	0.88 (0.73–0.95)	12
InceptionV3	0.84	0.89	0.88 (0.69–0.96)	0.91 (0.73–0.98)	0.80 (0.49–0.94)	0.85 (0.70–0.94)	7

Table 3
Fruit-level classification performance of formulation D. Performance of the weakly supervised, fruit-level formulation on the full set of 189 punctured-fruit scans available for deep-learning analysis. The first two rows report the held-out test partition at two decision thresholds, each selected on the validation fruit and then applied unchanged; they are based on the same continuous prediction scores and therefore share the same, threshold-independent AUC. The third row reports the mean and standard deviation across cross-validation folds grouped by fruit, and the fourth pools the out-of-fold predictions, giving one prediction per fruit over the entire archive. Proportions are reported with Wilson score 95% confidence intervals and AUC with a percentile bootstrap interval obtained by resampling fruit. Values in parentheses are 95% confidence intervals, whereas values following ± are standard deviations across folds; the two quantify different things and are not interchangeable. Balanced accuracy and F1 are given as point estimates in the single-split and pooled rows, since interval estimates for these two composite quantities would require a separate bootstrap procedure. The threshold was selected independently within each fold and varied over the range shown, which is itself an indication that the operating point is not identifiable at this sample size. No validation partition exists at the pooled level, so a fixed threshold of 0.50 was applied.Values in parentheses are Wilson score 95% confidence intervals for proportions and bootstrap 95% intervals for AUC. Values following ± are standard deviations across cross-validation folds.

Evaluation	Fruit (inf/ctrl)	Decision threshold	AUC	Sensitivity	Specificity	Precision	Accuracy	Balanced accuracy	F1
Held-out test set, threshold maximising F1	57 (43/14)	0.33	0.799 (0.673–0.907)	0.930 (0.814–0.976)	0.286 (0.117–0.546)	0.800 (0.670–0.888)	0.772 (0.648–0.862)	0.608	0.860
Held-out test set, threshold maximising balanced accuracy	57 (43/14)	0.81	0.799 (0.673–0.907)	0.488 (0.346–0.632)	1.000 (0.785–1.000)	1.000 (0.845–1.000)	0.614 (0.484–0.729)	0.744	0.656
Cross-validation, mean ± SD across folds	189 (141/48)	0.36–0.98[a]	0.783 ± 0.111	0.633 ± 0.155	0.769 ± 0.246	0.906 ± 0.090	0.667 ± 0.087	0.701 ± 0.086	0.731 ± 0.094
Pooled out-of-fold predictions	189 (141/48)	0.50[b]	0.752 (0.676–0.822)	0.730 (0.652–0.797)	0.521 (0.383–0.655	0.817 (0.741–0.875)	0.677 (0.608–0.740)	0.626	0.772

[a] The threshold was selected independently within each fold and varied over the range shown, which is itself an indication that the operating point is not identifiable at this sample size.
[b] No validation partition exists at the pooled level, so a fixed threshold of 0.50 was applied.

detect fruit fly infestation from μCT images. From an operational perspective, DenseNet121 represents the most suitable architecture among those evaluated because it missed no infested fruit while maintaining the best precision and accuracy of the models achieving that sensitivity. In phytosanitary inspection systems, where failing to detect an infested fruit is generally considered more costly than incorrectly flagging a non-infested fruit for secondary inspection, minimizing false negatives is typically the primary objective, although at the specificity achieved here (0.60) a substantial proportion of sound fruit would be routed to secondary inspection.

These estimates derive from 34 test fruit, of which 10 were controls, and the confidence intervals are correspondingly wide. The specificity estimates in particular should be treated as provisional: the three architectures reaching a sensitivity of 1.00 did so at specificities of 0.40–0.60, indicating a systematic preference for the infested class rather than balanced discrimination. No operating point can be recommended for deployment on the basis of a control set of this size.

3.2.4. Model interpretability and biological validation

To assess model interpretability, Grad-CAM (Gradient-weighted Class Activation Mapping) analyses were conducted on the final transfer-learning models. As shown in Fig. 8, the resulting activation maps consistently highlighted internal tissue regions associated with suspected infestation-related damage while assigning comparatively little importance to peripheral tissue or background areas. This correspondence between model attention and biologically meaningful image features suggest that the models learned relevant infestation-associated patterns rather than relying on spurious image artefacts. The Grad-CAM analyses therefore provide an additional level of biological validation for the automated classification framework and increase confidence in the interpretability of the model predictions.

Similar activation patterns were observed across other correctly

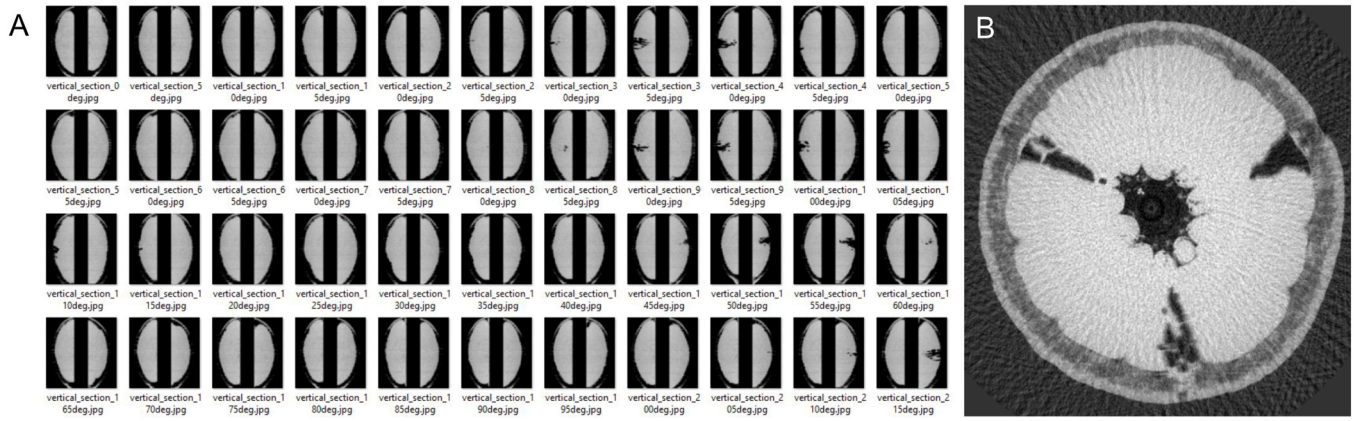


Fig. 7. Preprocessing pipeline and slice-level representation of μCT images used for deep learning analysis. (A) Example of the preprocessed dataset consisting of vertically extracted cross-sectional μCT slices obtained at multiple rotation angles from a single fruit volume. From the reconstructed tomographic volume comprising 1024 slices, 72 informative vertical slices were selected per fruit, resized (128 × 128 pixels), intensity-normalized to a [0–1] range, background-removed using Otsu's thresholding, and spatially aligned using bounding-box cropping and central-axis correction to ensure geometric consistency across samples. Each fruit was represented as a stack of standardized slices for slice-level classification and subsequent voting-based fruit-level prediction. (B) Representative original μCT slice prior to preprocessing, shown for comparison. The raw image includes background regions, intensity heterogeneity, and minor alignment variability that were corrected during preprocessing to reduce noise and improve model robustness.

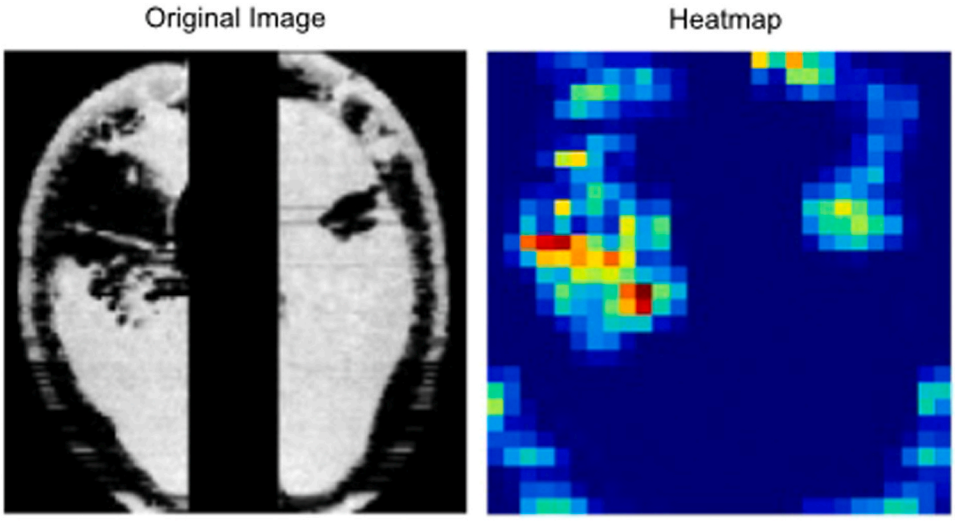


Fig. 8. Representative Grad-CAM visualization generated from the final transfer-learning model. The left panel shows the original μCT slice of an orange, while the right panel displays the corresponding Grad-CAM heatmap, highlighting the image regions that contributed most strongly to the classification decision. Warmer colors indicate higher activation and correspond to regions associated with suspected infestation-related tissue damage. Similar activation patterns were consistently observed across other correctly classified samples, indicating that the highlighted regions are representative of the model's general decision-making behaviour rather than unique to the example shown.

classified samples, suggesting that the highlighted regions were representative of the model's general decision-making behaviour rather than an isolated example. A complementary form of interpretability, obtained at the level of whole sections rather than within a section, is reported in Section 3.2.7.

3.2.5. *Impact of preprocessing on model performance*

The preprocessing pipeline described above played an important role in the development of the deep learning workflow. Standardization of image dimensions and slice orientation, background removal, exclusion of non-informative slices and correction of alignment inconsistencies reduced variability unrelated to infestation status and increased the consistency of the input data. These steps enabled the models to focus on biologically relevant image features associated with infestation rather than on imaging artefacts or positional variation among samples.

The improved performance of the final transfer-learning models following dataset curation and preprocessing suggests that careful preparation of μCT data is an important prerequisite for reliable automated detection. More broadly, these results highlight the importance of combining biological knowledge with appropriate image-processing procedures when developing deep learning approaches for agricultural and biological imaging applications.

3.2.6. *Multiclass extension with Transformer and ACGAN augmentation*

As a complementary exploration of the information content contained within the μCT data, the Transformer-based architecture described in Section 2.6.2 was trained on a nine-class formulation comprising control fruit and combinations of four incubation periods and two fly-density treatments. Because the original training dataset was too small for this multiclass task, the Auxiliary-Classifier GAN (ACGAN) described in Section 2.6.2 was used to augment the training data, increasing the number of available sequences to approximately 800.

Standard data augmentation procedures were subsequently applied to the training subset only. On the held-out test set, the model achieved an overall accuracy of 73.5% and a weighted F1-score of 0.746.

These results are reported separately from the binary detection benchmark presented in Table 2 because the underlying classification objective differs substantially. Whereas the models in Table 2 were optimized for the operationally relevant task of infested versus non-infested classification, the Transformer model was trained to discriminate among nine experimental classes. Direct comparison between the two approaches is therefore not appropriate.

Nevertheless, the results demonstrate that the μCT data contain information capable of supporting more detailed classification tasks beyond binary infestation detection. They also suggest that synthetic data augmentation using ACGANs may provide a useful strategy for addressing data scarcity in μCT-based agricultural imaging applications. Future studies based on larger annotated datasets will be required to determine whether these approaches can be reliably extended to the estimation of infestation stage and severity under operational conditions.

3.2.7. Formulation D: fruit-level supervision reproduces detection performance without section selection

Applied to the 189-fruit punctured-fruit archive available for deep-learning analysis, the weakly supervised model classified held-out fruit with an AUC of 0.80 (95% CI 0.67–0.91; 57 test fruit, 43 infested and 14 control). Cross-validation grouped by fruit gave a mean AUC of 0.78 ± 0.11 across five folds, with a mean balanced accuracy of 0.67 ± 0.09 and a mean specificity of 0.77 ± 0.25; pooling the out-of-fold predictions, which yields one prediction per fruit over all 189 fruit, gave an AUC of 0.75 (95% CI 0.68–0.82) (Table 3). The dispersion across folds is itself informative: with five folds the standard deviation is imprecisely estimated, and the marked variation in specificity reflects the small number of control fruit assigned to each fold.

For comparison, a systematic reproduction of formulation C on all 189 scans available for deep learning analysis, without section-level curation did not achieve discrimination distinguishable from chance: across four architectures and two hyper-parameter settings, fruit-level AUC ranged from 0.45 to 0.64 and every confidence interval included 0.5. These experiments extend the initial uncurated results summarised in Table 1 to the full set of architectures retained for the final benchmark. Compared on the same held-out fruit and with the same backbone (MobileNetV2), fruit-level supervision improved AUC by 0.32 (95% CI 0.16–0.50, $p < 0.001$), and the advantage held against all eight configurations of formulation C. Section-level validation AUC did not exceed 0.69 even with hyper-parameter search, indicating that the limitation lies in the assignment of the fruit-level label to every section rather than in model capacity or tuning.

Two features of the fruit-level result are informative. First, the learned attention weights were concentrated on a small subset of the 72 sections, with the highest-weighted sections corresponding to regions in which larval galleries were visible (Supplementary Figure S3). This indicates that the model was able to identify informative regions despite being trained only with fruit-level labels. The model therefore recovered, without supervision at the section level, the same selection that had previously been made by visual review. Second, an ablation in which the backbone was used as a fixed feature extractor reached an AUC of only 0.67, indicating that adapting the convolutional features to μCT contrast was necessary and that the aggregation scheme alone was not sufficient.

Detectability was ordered by infestation density (AUC 0.73, 0.74 and 0.79 for 50, 100 and 500 flies) and was lowest at the earliest incubation stage (0.71 at day 4, against 0.75–0.80 at days 5, 7 and 10). Both orderings are broadly consistent with the manual μCT analysis, in which incubation period significantly affected detection performance and fly density significantly affected the overall probability of detecting fruit-fly damage in punctured fruit (Section 3.1), and both were reproduced in an independent cross-validation run. Within this analysis, however, neither trend reached statistical significance (Spearman $p = 0.35$ for density; Mann–Whitney $p = 0.24$ for day 4 against later stages), the subgroups comprising 34–36 infested fruit against 48 controls. They are therefore reported as consistent tendencies rather than as independently established effects.

Specificity remained the limiting quantity. At the threshold selected on the validation fruit the model reached a specificity of 1.00 (95% CI 0.79–1.00) and a precision of 1.00, but a sensitivity of only 0.49; at the F1-optimized threshold (0.33), sensitivity increased to 0.93, while specificity fell to 0.29. Across cross-validation folds, the selected threshold ranged from 0.36 to 0.98, reflecting substantial variability in the operating point. Mean balanced accuracy across folds was 0.70 ± 0.09, while the held-out test set reached a balanced accuracy of 0.74 when the threshold was selected to maximise this metric (Table 3). Part of the residual false-positive rate is likely to reflect the same confounding documented for manual inspection in Section 3.1, where advanced rot in unexposed punctured fruit incubated for 10 days was visually confused with fruit fly damage. The archive contains 48 control fruit in total and 14 in the test partition; the number of non-infested specimens, rather than the modelling approach, is what currently limits the precision of these estimates and prevents any operating point from being recommended.

3.3. Biological, technological, and pest management implications

Overall, the results of this study demonstrate the potential of μCT imaging as a non-destructive tool for postharvest detection of fruit fly infestation in citrus fruit and highlight the value of combining μCT imaging with deep learning to support scalable, reproducible and interpretable inspection workflows. Rather than replacing expert assessment, automated analysis has the potential to complement manual inspection by increasing throughput and consistency while maintaining biological interpretability.

Controlled infestation experiments on both punctured and intact oranges, combined with high-resolution μCT imaging and biological validation through fruit dissection, confirmed that μCT can reveal internal larval activity and associated tissue damage that may not be detectable through conventional external inspection. From a biological perspective, manual μCT detection performance was influenced by infestation stage and larval density, although the magnitude and statistical significance of these effects differed between punctured and intact fruit and among the response variables considered (Section 3.1, Table S1); the automated analysis showed the same ordering, although its subgroups were too small to establish the effects independently (Section 3.2.7). The comparison between punctured and intact fruit should also be interpreted cautiously because the experiments were conducted under slightly different incubation temperatures, which may have influenced oviposition behavior and larval development rates. As larval development progresses, feeding galleries and tissue degradation became increasingly pronounced, generating stronger image contrast and facilitating detection. These patterns explain the higher detection rates observed at advanced infestation stages and highlight the continuing challenge of detecting early infestations, particularly in intact fruit where oviposition and larval activity remain sparse and highly localized. Because no eggs were detected during fruit dissection, the earliest biologically validated infestation stage represented in this study was the first instar larval stage.

In addition to biological factors influencing detectability, the selected μCT acquisition parameters also represent a compromise between imaging quality and practical throughput. Although the current scanning protocol was sufficient to detect infestation-associated damage in the examined samples, improvements in imaging conditions, such as acquiring more projections or using longer integration times, would improve angular sampling and signal-to-noise ratio respectively, and would enhance visualization of smaller or lower-contrast features

associated with very early infestation stages. Such improvements could potentially increase sensitivity towards first-instar infestations; however, they would require longer acquisition times and may reduce the suitability of μCT for high-throughput screening applications. Therefore, future optimization of imaging parameters should consider the balance between detection sensitivity, spatial resolution, and practical requirements for large-scale inspection.

The acquisition protocol was deliberately optimised for throughput: 125 projections acquired in 1 min 26 s per fruit, at a voxel volume of 0.0087 mm^3 . This entails two costs that act in the same direction for early detection. Angular undersampling limits the effective spatial resolution, so that the narrow tunnels produced by first-instar larvae approach the resolution limit; and the short integration time lowers the signal-to-noise ratio, reducing the contrast between a small air-filled cavity and the surrounding parenchyma. Both preferentially degrade the detection of small, early lesions, which is consistent with the lower performance observed at day 4 in Section 3.2.7. Denser angular sampling, longer integration, or iterative and learning-based reconstruction would be expected to improve sensitivity at the earliest stages, at the cost of scanning rate. On the evidence presented here, the next gains are more likely to come from improved image formation than from further refinement of the classifier.

A central limitation of the deep learning component is that model development and evaluation were conducted exclusively on the punctured-fruit cohort and not on the intact-fruit cohort that more closely represents commercial postharvest inspection conditions.

This restriction arose from three characteristics of the intact-fruit dataset. First, infestation prevalence was substantially lower in intact fruit than in punctured fruit, resulting in too few positive samples to support reliable supervised training and evaluation. Second, manual inspection of intact-fruit μCT scans achieved low sensitivity relative to dissection-based ground truth, indicating that infestation-related signals were only weakly represented in the images acquired under the scanning conditions used in this study. Third, the discrepancy between imaging data and biological ground truth would have introduced substantial label noise into any supervised learning framework because many biologically infested fruit contained little or no visually detectable infestation signal. The same consideration applies, at a finer scale, within the punctured cohort: a fruit-level label propagated to all 72 of its sections is noisy for the same reason. Formulation D addresses this directly by supervising at the level of the fruit, and its performance relative to formulation C on identical fruit (Section 3.2.7) quantifies how much of the earlier limitation was attributable to label noise rather than to the imaging.

Consequently, the deep learning results reported in this study should be interpreted as applying only to punctured fruit. Punctured fruit may represent preharvest surface injuries caused by mechanical damage in orchards, hail or feeding by other organisms that may create entry points similar to the artificial punctures used in this study and increase susceptibility to fruit fly infestation. The punctured-fruit model may therefore represent situations in which the natural barriers to oviposition have been compromised prior to harvest, resulting in increased susceptibility to infestation. For intact fruit, however, the principal limitation identified by this study appears to be the detectability of infestation-related signals within the μCT images rather than the analytical framework itself. Extending automated detection to intact commercial fruit will likely require larger and more comprehensively annotated datasets and/or modifications to imaging protocols that improve visualization of localized larval activity. We regard this as the principal direction for future research and the main qualification on the operational generalizability of the deep learning workflow reported here.

From a pest management perspective, the integration of non-destructive imaging and automated analysis could contribute to more effective phytosanitary inspection systems for fruit consignments. Technologies capable of detecting internal infestation without destructive sampling may increase inspection efficiency, support surveillance programs targeting quarantine pests such as *C. capitata* and permit higher sampling intensity while preserving commodity integrity. Although further work is required to improve sensitivity to very early infestations and to validate performance under operational conditions, the present results demonstrate how advanced imaging and AI-based analysis can contribute to the development of practical diagnostic tools for postharvest pest detection.

4. Conclusions

This study supports the feasibility of integrating μCT imaging with deep learning for the automated, non-destructive detection of medfly infestation in citrus fruit. Importantly, the objective of the deep learning component was not to surpass expert manual inspection in absolute accuracy, but to provide an automated and reproducible analysis framework capable of operating at throughput levels that are impractical for expert review alone.

The proposed workflow combines volumetric imaging with AI-based analysis to identify internal infestation signatures and translate complex μCT data into practical diagnostic outputs. The results demonstrate that infestation-related damage can be detected both through expert interpretation of μCT scans and through automated image analysis, particularly when infestation is sufficiently developed to generate visible internal tissue changes.

The same conclusion was reached by a formulation supervising at the level of the whole fruit, which requires no manual selection of informative sections, indicating that the detectability of internal infestation in μCT reconstructions is a property of the images rather than of a particular curation choice. This matters for the intended application: a procedure that depends on expert review of individual sections cannot be transferred to an inspection line, whereas fruit-level supervision can.

However, the deep learning workflow was developed and validated exclusively on punctured fruit, and additional work will be required to determine whether comparable performance can be achieved for intact fruit under commercial inspection conditions. Future studies should therefore focus on improving detection sensitivity in intact fruit, expanding annotated datasets, and evaluating system performance under operational phytosanitary inspection scenarios.

Present performance supports detection above chance but not deployment: across the full 189-fruit deep learning dataset, no decision threshold achieved acceptable sensitivity and specificity simultaneously, and with 48 non-infested specimens available the specificity of any operating point cannot be estimated with useful precision. Enlarging the control set is therefore a prerequisite for any operational assessment, and the evidence assembled here suggests that further gains will come primarily from improved image formation rather than from refinement of the classifier. With continued development of both imaging and analytical methodologies, μCT-based automated inspection retains the potential to support scalable and non-destructive phytosanitary monitoring within postharvest supply chains.

Data statement

The processed μCT images, the deep learning analysis code (organized in different formulations), and the infestation experiment raw data are available at https://osf.io/7qgzb/overview?view_only=8274dd2f09b44750820f913bddea0d56.

CRediT authorship contribution statement

Francesca Scolari: Writing – review & editing, Writing – original draft, Supervision, Methodology, Investigation, Funding acquisition, Conceptualization. **Leani Serfontein:** Writing – review & editing, Investigation, Data curation. **Wayne Kirkman:** Writing – review & editing, Funding acquisition, Conceptualization. **Sean Moore:** Writing –

review & editing, Conceptualization. **Jakobus Hoffman:** Writing – review & editing, Methodology, Investigation. **Lunga Bam:** Writing – review & editing, Methodology, Investigation. **Andrea Moyano:** Writing – review & editing, Investigation. **Guy F.Sutton:** Formal analysis, Visualization. **Claudio Pezzoni:** Writing – review & editing, Visualization, Investigation, Formal analysis, Data curation. **Grzegorz J. Możdżyński:** Writing – review & editing, Investigation, Formal analysis, Data curation. **Dario Pescini:** Writing – review & editing, Visualization, Investigation, Formal analysis, Data curation. **Anna Sandionigi:** Writing – review & editing, Visualization, Supervision, Methodology, Investigation, Formal analysis. **Aruna Manrakhan:** Writing – review & editing, Visualization, Supervision, Methodology, Investigation, Funding acquisition, Conceptualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We acknowledge the technical assistance provided at Citrus Research International by Glorious Shongwe for maintenance of the fly cultures, Phindile Rikhotso for preparation of infested fruit and dissection as well as Rooikie Beck for assistance during the X-ray scans. Funding for this research was provided by Citrus Research International under project 1360 and by the European Union's Horizon Europe Research and Innovation Programme REACT (Grant agreement 101059523).

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.postharvbio.2026.114707.

Data availability

OSF Link Provided

Convolutional neural networks and X-ray analysis for fruit fly infestation detection (OSF)

References

Adedeji, A.A., Ekramirad, N., Rady, A., Hamidisepehr, A., Donohue, K.D., Villanueva, R. T., Parrish, C.A., Li, M., 2020. Non-destructive technologies for detecting insect infestation in fruits and vegetables under postharvest conditions: a critical review. Foods 9, 927. https://doi.org/10.3390/foods9070927.

Alba-Tercedor, J., Ruano, F., 2024. Use of micro-computed tomography to monitor olive fruit damage caused by three insect pests, 14 Sci. Rep. 2024 14 (1), 21067. https://doi.org/10.1038/s41598-024-72029-3.

Ansah, F.A., Amo-Boateng, M., Siabi, E.K., Bordoh, P.K., 2023. Location of seed spoilage in mango fruit using X-ray imaging and convolutional neural networks. Sci. Afr. 20, e01649. https://doi.org/10.1016/J.SCIAF.2023.E01649.

Arendse, E., Fawole, O.A., Magwaza, L.S., Opara, U.L., 2016. Non-destructive characterization and volume estimation of pomegranate fruit external and internal morphological fractions using X-ray computed tomography. J. Food Eng. 186, 42–49. https://doi.org/10.1016/J.JFOODENG.2016.04.011.

Bolker, B.M., Brooks, M.E., Clark, C.J., Geange, S.W., Poulsen, J.R., Stevens, M.H.H., White, J.S.S., 2009. Generalized linear mixed models: a practical guide for ecology and evolution. Trends Ecol. & Evol. 24, 127–135. https://doi.org/10.1016/j.tree.2008.10.008.

Brooks, M.E., Kristensen, K., van, Benthem, K.J., Magnusson, A., Berg, C.W., Nielsen, A., Skaug, H.J., Mächler, M., Bolker, B.M., 2017. glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. R. J. 9, 378–400. https://doi.org/10.32614/RJ-2017-066.

Chuang, C.L., Ouyang, C.S., Lin, T., Te, Yang, M.M., Yang, E.C., Huang, T.W., Kuei, C.F., Luke, A., Jiang, J.A., 2011. Automatic X-ray quarantine scanner and pest infestation detector for agricultural products. Comput. Electron. Agric. 77, 41–59. https://doi.org/10.1016/j.compag.2011.03.007.

De Meyer, M., Copeland, R.S., Lux, S., Mansell, M., Wharton, R., White, I.M., Zenz, N., 2002. Annotated check list of host plants for afrotropical fruit flies (Diptera: Tephritidae) of the genus *Ceratitis*. Zool. Doc. K. Mus. voor Midden Afr. 27, 1–92.

Du, Z., Hu, Y., Ali Buttar, N., Mahmood, A., 2019. X-ray computed tomography for quality inspection of agricultural products: A review. Food Sci. Nutr. 7, 3146–3160. https://doi.org/10.1002/fsn3.1179.

Fox, J., Weisberg, S., 2019. An R Companion to Applied Regression, 3rd ed. Sage Publications.

Gao, J., Zhou, Y., Hui, Y., Zhang, Y., Wang, Q., Liu, J., Wang, X., Wang, H., Ding, Hao, Ding, Haiyang, 2025. Nondestructive detection and rapid segmentation of the development process of *Sitophilus zeamais* in single wheat kernels based on CT and MV-UNet. J. Econ. Entomol. 118 935947. https://doi.org/10.1093/JEE/TOAF015.

Giunti, G., Benelli, G., Campolo, O., Canale, A., Kapranas, A., Liedo, P., De Meyer, M., Nestel, D., Ruiu, L., Scolari, F., Wang, X., Papadopoulos, N.T., 2023. Management of the Mediterranean fruit fly, *Ceratitis capitata*: past, present, and future. Entomol. Gen. 43, 1241–1263. https://doi.org/10.1127/entomologia/2023/2136.

Hadipour-Rokni, R., Askari Asli-Ardeh, E., Jahanbakhshi, A., Esmaili paeen-Afrakoti, I., Sabzi, S., 2023. Intelligent detection of citrus fruit pests using machine vision system and convolutional neural network through transfer learning technique. Comput. Biol. Med. 155, 106611. https://doi.org/10.1016/J.COMPBIOMED.2023.106611.

Haff, R., Jackson, E., Moscetti, R., Massantini, R., 2016. Detection of fruit-fly infestation in olives using X-ray imaging: Algorithm development and prospects. Am. J. Agric. Sci. Technol. https://doi.org/10.7726/ajast.2016.1001.

He, J., Verlinde, S., Belien, T., Alhmedi, A., Verboven, P., Nicolai, B., 2024. Non-destructive larval infestation damage detection in pear fruits using deep learning and X-ray CT-generated radiographs. https://doi.org/10.58286/29252.

Hoffman, J.W., De Beer, F.C., 2012. Characteristics of the micro-focus X-ray tomography facility (MIXRAD) at Necsa in South Africa. 18th World Conf. Nondestruct. Test. 16–20.

Ilesanmi, A.E., Ilesanmi, T.O., Ajayi, B.O., 2024. Reviewing 3D convolutional neural network approaches for medical image segmentation. Heliyon 10, e27398. https://doi.org/10.1016/J.HELIYON.2024.E27398.

Ilse, M., Tomczak, J.M., Welling, M., 2018. Attention-based deep multiple instance learning. In: Proceedings of the 35th International Conference on Machine Learning, 80. PMLR, pp. 2127–2136.

Karunakaran, C., Jayas, D.S., White, N.D.G., 2004. Identification of wheat kernels damaged by the red flour beetle using X-ray images. Biosyst. Eng. 87, 267–274. https://doi.org/10.1016/j.biosystemseng.2003.12.002.

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. ImageNet classification with deep convolutional neural networks. Adv. Neural Inf. Process. Syst. 25, 1097–1105.

Kshatri, S.S., Singh, D., 2023. Convolutional neural network in medical image analysis: A review. Arch. Comput. Methods Eng. 30, 2793–2810. https://doi.org/10.1007/S11831-023-09898-W.

Lenth, R.V. (2022). *emmeans: Estimated Marginal Means, aka Least-Squares Means* (R package version 1.8.2). https://CRAN.R-project.org/package=emmeans.

Li, D., Long, J., Tang, Z., Han, L., Gong, Z., Wen, L., Peng, H., Wen, T., 2023. Detection and classification of citrus fruit infestation by *Bactrocera dorsalis* (Hendel) using a multi-path Vis/NIR spectroscopy system. Agriculture 13, 1642. https://doi.org/10.3390/agriculture13081642.

Liquido, N.J., Shinoda, L.A., Cunningham, R.T., 1991. Host plants of the Mediterranean fruit fly (Diptera: Tephritidae): An Annotated World Review. SPIE. https://doi.org/10.4182/CMLT2950.

Ma, Y., Yin, J., Huang, F., Li, Q., 2024. Surface defect inspection of industrial products with object detection deep networks: a systematic review, 57 Artif. Intell. Rev. 2024 57 (12), 333. https://doi.org/10.1007/S10462-024-10956-3.

Manrakhan, A., Daneel, J.-H., Beck, R., Theron, C.D., Weldon, C.W., Moore, S.D., Hattingh, V., 2018. Non-host status of commercial export grade lemon fruit (*Citrus limon* (L.) Burman f. cv. Eureka) for *Ceratitis capitata*, *Ceratitis rosa*, *Ceratitis quilicii* and *Bactrocera dorsalis* (Diptera: Tephritidae) in South Africa. Afr. Entomol. 26, 202–214. https://doi.org/10.4001/003.026.0202.

Matsui, T., Sugimori, H., Koseki, S., Koyama, K., 2023. Automated detection of internal fruit rot in Hass avocado via deep learning-based semantic segmentation of X-ray images. Postharvest Biol. Technol. 203, 112390. https://doi.org/10.1016/J.POSTHARVBIO.2023.112390.

Mert, A., 2023. Enhanced dataset synthesis using conditional generative adversarial networks. Biomed. Eng. Lett. 13, 41–48. https://doi.org/10.1007/s13534-022-00251-x.

Milner, M., Lee, M.R., Katz, R., 1950. Application of X-ray technique to the detection of internal insect infestation of grain. J. Econ. Entomol. 43, 933–935.

Papachristos, D.P., Papadopoulos, N.T., 2009. Are citrus species favorable hosts for the Mediterranean fruit fly? A demographic perspective. Entomol. Exp. Et. Appl. 132, 1–12. https://doi.org/10.1111/j.1570-7458.2009.00861.x.

Papadopoulos, N.T., De Meyer, M., Terblanche, J.S., Kriticos, D.J., 2024. Fruit flies: Challenges and opportunities to stem the tide of global invasions. Annu. Rev. Entomol. 69, 355–373. https://doi.org/10.1146/annurev-ento-022723-103200.

Papadopoulos, N.T., Katsoyannos, B.I., Carey, J.R., Kouloussis, N.A., 2001. Seasonal and Annual Occurrence of the Mediterranean Fruit Fly (Diptera: Tephritidae) in Northern Greece. Ann. Entomol. Soc. Am. 94, 41–50. https://doi.org/10.1603/0013-8746(2001)094[0041:SAAOOT]2.0.CO;2.

R Core Team, 2025. R: A language and environment for statistical computing. R. Found. Stat. Comput. Vienna Austria. https://www.R-project.org/.

Rajendra, D.F., Wardhana, A.K., 2025. Development of MobileNetV2 for CT-scan lung classification using transfer learning. J. Appl. Inform. Comput. 9, 2703–2710. https://doi.org/10.30871/jaic.v9i5.10282.

Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., Batra, D., 2017. Grad-CAM: Visual explanations from Deep Networks via Gradient-Based Localization, in:

2017. IEEE International Conference on Computer Vision (ICCV). IEEE, pp. 618–626. https://doi.org/10.1109/ICCV.2017.74.

Si, Y., Sankaran, S., 2016. Computed tomography imaging-based bitter pit evaluation in apples. Biosyst. Eng. 151, 9–16. https://doi.org/10.1016/j.biosystemseng.2016.08.008.

Srivastava, S., Mishra, G., Mishra, H.N., 2020. Application of an expert system of X- ray micro computed tomography imaging for identification of *Sitophilus oryzae* infestation in stored rice grains. Pest. Manag. Sci. 76, 952–960. https://doi.org/10.1002/PS.5603.

Staub, C.G., De Lima, F., Majer, J.D., 2008. Determination of host status of citrus fruits against the Mediterranean fruit fly, *Ceratitis capitata* (Wiedemann) (Diptera: Tephritidae). Aust. J. Entomol. 47, 184–187. https://doi.org/10.1111/j.1440-6055.2008.00646.x.

Stermer, R.A., 1972. Automated X-ray inspection of grain for insect infestation. Trans. ASAE 15, 1081–1085. https://doi.org/10.13031/2013.38075.

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z., 2016. Rethinking the Inception architecture for computer vision, in: 2016. IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, pp. 2818–2826. https://doi.org/10.1109/CVPR.2016.308.

Tempelaere, A., Minh Phan, H., Van De Looverbosch, T., Verboven, P., Nicolai, B., 2023. Non-destructive internal disorder segmentation in pear fruit by X-ray radiography and AI. Comput. Electron. Agric. 212, 108142. https://doi.org/10.1016/J.COMPAG.2023.108142.

Wang, C., Liu, S., Wang, Y., Xiong, J., Zhang, Z., Zhao, B., Luo, L., Lin, G., He, P., 2022. Application of Convolutional Neural Network-based detection methods in fresh fruit production: A comprehensive review. Front. Plant. Sci. 13, 868745. https://doi.org/10.3389/fpls.2022.868745.

Wang, H., Nguyen, T.H., Nguyen, T.N., Dang, M., 2024. PD-TR: End-to-end plant diseases detection using a transformer. Comput. Electron. Agric. 224, 109123. https://doi.org/10.1016/J.COMPAG.2024.109123.

Wang, Q., Xue, H., John Martin, J.J., Hou, M., Cao, H., Dong, Z., Li, J., Sun, C., 2024. Trends and applications of computed tomography in agricultural non-destructive testing. Agriculture 14, 2329. https://doi.org/10.3390/agriculture14122329.

White, I.M., Elson-Harris, M.M., 1992. Fruit. Flies Econ. Significance Their Identif. Bionomics. CAB Int. Wallingford Uk.

Yu, L., Liu, L., Yang, W., Wu, D., Wang, J., He, Q., Chen, Z.S., Liu, Q., 2022. A non-destructive coconut fruit and seed traits extraction method based on Micro-CT and deeplabV3+ model. Front. Plant. Sci. 13, 1069849. https://doi.org/10.3389/fpls.2022.1069849.

Yu, Y., Song, Z., Sun, M., Gao, S., Cheng, Y., Li, S., Chen, Y., 2024. Inline detection of citrus rind micro-wounds using contrast-enhanced X-ray imaging: a feasibility study. LWT 214, 117095. https://doi.org/10.1016/j.lwt.2024.117095.

Zhang, M., Zhang, Y., Liu, S., Han, Y., Cao, H., Qiao, B., 2024. Dual-attention transformer-based hybrid network for multi-modal medical image segmentation. Sci. Rep. 14, 25704. https://doi.org/10.1038/s41598-024-76234-y.