Applied Food Research 6 (2026) 102136

Contents lists available at ScienceDirect

Applied Food Research

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

Development and benchmarking of sensing technologies for *Tenebrio molitor* dead pupae identification

Francisco Oliveira [a,b,*], Vítor Tinoco [a], Leandro Rodrigues [a,c], Filipe N. Santos [a,d], Mário Cunha [a,c], Inês Vieira [e], Marisa V. Santos [e,f]

[a] INESC TEC - Institute for Systems and Computer Engineering, Technology and Science, Campus da FEUP, S/N, 4200-465, Porto, Portugal
[b] Department of Engineering, UTAD - University of Trás-os-Montes and Alto Douro, 5000-801, Vila Real, Portugal
[c] Department of Geosciences, Environment and Spatial Planning, Faculty of Sciences of the University of Porto, Rua do Campo Alegre S/N, 4169-007, Porto, Portugal
[d] Department of Electrical and Computer Engineering, Faculty of Engineering of the University of Porto, Campus da FEUP, S/N, 4200-465, Porto, Portugal
[e] Thunder Foods S.A., 2005-332, Santarém, Portugal
[f] Egas Moniz Center for Interdisciplinary Research (CiiEM), Egas Moniz School of Health & Science, Almada, Setúbal, Portugal

ARTICLE INFO

Keywords:
Edible insects
Automation
Industry
NIR
Thermography
Hyperspectral

ABSTRACT

The production of insects for both human and animal consumption has seen an interest increase in recent years. *Tenebrio molitor*, a beetle species whose larval stage was considered by the European Commission safe for human consumption, represents a promising candidate for large-scale production. Efficient production of *T. molitor* requires the separation of its different metamorphic stages to optimize productivity and reduce cannibalism. Additionally, dead pupae must be removed to prevent contamination and ensure the healthy development of the remaining individuals. Currently, the identification of dead pupae relies primarily on visual inspection based on colour changes, as dead individuals tend to exhibit surface melanization. However, this method becomes inefficient in large-scale operations. This study presents a benchmark of hyperspectral imaging (HSI) and thermal imaging as alternative sensing technologies for automated dead pupae identification. Hyperspectral data acquired in the near-infrared range (900–1700 nm) enabled accurate discrimination between dead and live pupae using both a PLS-DA and a Logistic Regression model, achieving F1-Scores above 90 % for both classes. Furthermore, key wavelengths (903 nm and 1259 nm) were identified and used to develop a normalized difference index (NDI) capable of distinguishing pupae health status using only two spectral bands. Thermal imaging revealed consistent temperature differences, with dead pupae presenting approximately 1–3 % lower temperatures than live pupae. The proven reliability and precision of the proposed sensing technologies validate their use in contributing to the development of scalable and reliable solutions for quality control in the insect farming industry.

1. Introduction

The human population increase, which is expected to surpass 9 billion individuals by 2050, requires adaptation on food production to assure democratized access to food products (Vågsholm et al., 2020). Furthermore, seeking sustainable production practices is essential to minimize the climatic and economic impact caused by the food industry, which often results in several wastes and non-utilized by-products (Gómez-García et al., 2021). In 2021, the European Commission, acknowledged the larval stage of *T. molitor* as a novel food, allowing it to be safely consumed by humans in frozen and dried forms (Bisht et al., 2025; EFSA Panel on Nutrition, Novel Foods and Food Allergens (NDA) et al., 2021). This novel food presents a valuable source of protein, unsaturated fatty acids, vitamins, and minerals, presenting health benefits combined with the advantage of its industrial production being cost-effective and sustainable (Muñoz-Seijas, Fernandes, & Domínguez, 2024; Ni et al., 2024; Sharma et al., 2024). Furthermore, *T. molitor* frass, which mainly consists of the insect's excrements, is rich in nutrients valuable to plants and can be used as an organic fertilizer (Poveda et al., 2019). The use of the by-products resulting from the rearing of these insects contributes to the

* Corresponding author at: INESC TEC - Institute for Systems and Computer Engineering, Technology and Science, Campus da FEUP, S/N, 4200-465, Porto, Portugal.
E-mail address: francisco.a.oliveira@inesctec.pt (F. Oliveira).

https://doi.org/10.1016/j.afres.2026.102136
Received 9 December 2025; Received in revised form 4 May 2026; Accepted 12 May 2026
Available online 19 May 2026

implementation of a sustainable food production system with minimal waste generation (Nuno Muñoz-Seijas, Fernandes et al., 2024). Moreover, agro-industrial by-products can be used as feeding substrates for *T. molitor* production (Gasco et al., 2020; Rumbos et al., 2022).

Considering the use of natural resources, insects tend to consume less water than other animals, and the feed requirement for rearing is low. *T. molitor* production has a water footprint of approximately one third of the value of beef production. Furthermore, its protein content per edible kilogramme is higher when compared to beef, which, consequently, leads to a smaller water footprint per gram of protein (Miglietta et al., 2015). This combined with the advantages previously presented related to by-product conversion makes industrial insect production capable of following a circular economy perspective (Madau et al., 2020).

The industrial insect production for human consumption, despite presenting advantages, still face a few difficulties which might compromise efficient production. Rearing *T. molitor* requires a consistently monitored production, in which it is crucial to separate the different metamorphosis stages of the insect to avoid cannibalism, optimize the beetles' breeding and easily harvest the larvae (Barrett et al., 2023). However, the monitoring solutions are still limited and in early development and do not consider the removal of dead pupae (Majewski et al., 2024).

Pupae are mostly motionless, only contracting their body when pinched as a defence mechanism, which makes them more likely to suffer cannibalism by larvae - estimated to affect between 15% and 25% of pupae mixed in trays with larvae (Ichikawa & Kurauchi, 2009). For an efficient industrial rearing of *T. molitor* the insects must be separated by life cycle stage to ensure optimal development. When conducting regular harvest of pupae from trays containing larvae, some might perish during the separation process. In order to avoid contamination of the full production batch by dead pupae, these must be removed and discarded. However, since pupae are naturally motionless, the identification of dead pupae becomes difficult to conduct successfully in a scale-up scenario, since this task is usually performed by observation of the colour change of the insects, which become darker when dead due to melanization (Noonin et al., 2011). Although this feature facilitates manual separation, its efficiency decreases with higher insect densities per tray, as the superficial colour change can be hidden by adjacent pupae and might not be visible without manipulation. Furthermore, human visual inspection performance has been reported to present error rates frequently in the range of 20–30 % depending on the conducted task (See, 2012). Moreover, despite some biological absorbers being possible to be characterized in the visible region (from near 400 nm to 700 nm), higher wavelengths might also provide useful data for physiological characterization (Gruensfelder et al., 2025).

This study aims to evaluate and compare different sensing methods for the efficient and accurate identification of dead *T. molitor* pupae, with a focus on overcoming the limitations of traditional visual inspection techniques. The specific objectives are:

- Investigate the limitations of colour-based identification of dead pupae;
- Explore and validate alternative sensing technologies, including spectral response and thermal imaging, for enhanced pupae detection;
- Develop and propose a Normalized Difference Index (NDI) for classifying pupal health status, enabling consistent and objective identification;
- Evaluate the efficiency of classifying *T. molitor* pupae in two classes (dead or alive) based on the variation of mean reflectance on the Near Infrared region.
- Assess the feasibility of implementing the proposed methods using cost-effective equipment, avoiding reliance on high-end hyperspectral imaging systems;
- Contribute to the development of a scalable and practical classification prototype for application in automated insect production systems.

2. Methodology

2.1. Tenebrio molitor pupae sample characterization

The pupae used on the experiments conducted were obtained from the Portuguese producer Thunder Foods, S.A.[1] The pupae measured approximately 1.5–1.7 cm in length and 0.4 cm in width.

The *T. molitor* samples were all reared under the same controlled conditions, with feeding substrate consisting of plant-derived materials, such as cereals (wheat bran), and vegetables (potato) as the water source. The production had a total duration of 14–18 weeks, with temperature and relative humidity ranging between 27 ± 1 °C and 50 ± 5%, respectively. The breeding process started with the oviposition by female adults (darkling beetles) in trays, egg hatching, and larvae growth. Larvae trays were maintained past the ideal harvest moment to allow some insects to reach the pupae stage. Afterwards, the *T. molitor* pupae were manually collected from the same batch and separated from the frass to be used in the tests, guaranteeing the homogeneity of the samples.

2.2. Data acquisition preparation

The methodology followed for data acquisition consisted on the preparation of a 3D printed polylactic acid (PLA) base structure in which *T. molitor* pupae were randomly placed and then captured with the three cameras. The use of this structure ensures a uniform background that could easily be differentiated from the pupae on the different sensors used. The environment of the laboratory was controlled to a constant temperature of 22 °C. However, the NIR camera used halogen lamps that increased the temperature of both the base structure and the analysed pupae. Due to this temperature increase, the data acquisition was, in each batch, firstly conducted with the thermal and RGB cameras and only afterwards with the hyperspectral camera. By doing this, we guaranteed the temperature of the samples were not heavily affected by the heat caused by the halogen lamps.

2.3. Sensing technologies

- **2D RGB dataset** - 20 red, green and blue (RGB) images were acquired using a smartphone (Samsung A54[2]) with image resolution of 3060 × 4080 px. This dataset was used as ground truth, since it is the most common method used to distinguish between live and dead pupae as stated before. Considering this, these RGB images were used during the labelling task of hyperspectral and thermal images as visual reference to aid the procedure.
- **Thermal images dataset** - 10 images were acquired with the thermal camera FLIR E8-Pro.[3] Further, the images were processed on the FLIR Thermal Studio[4] to compare the temperature range of live and dead samples. The data gathered with the thermal camera allowed for the comparison of the temperature between live and dead pupae on each batch. On each image, using the FLIR Thermal Studio software, the minimum and maximum temperature of each pupa was measured, as shown in Fig. 1. These measurements allowed to define an average temperature of each pupa in each image, and consequently assess if there were differences between live and dead pupae.

[1] Thunder Foods, S.A. - https://thunderfoods.pt/ (Accessed on 21/11/2025).

[2] Samsung A54 5G - https://www.samsung.com/pt/smartphones/galaxy-a/galaxy-a54-5g-green-128gb-sm-a546blgceub/ (Accessed on 21/11/2025).

[3] FLIR E8-Pro - https://www.flir.eu/products/e8-pro/ (Accessed on 21/11/2025).

[4] FLIR Thermal Studio Suite - https://www.flir.com/products/flir-thermal-studio-suite/ (Accessed on 21/11/2025).

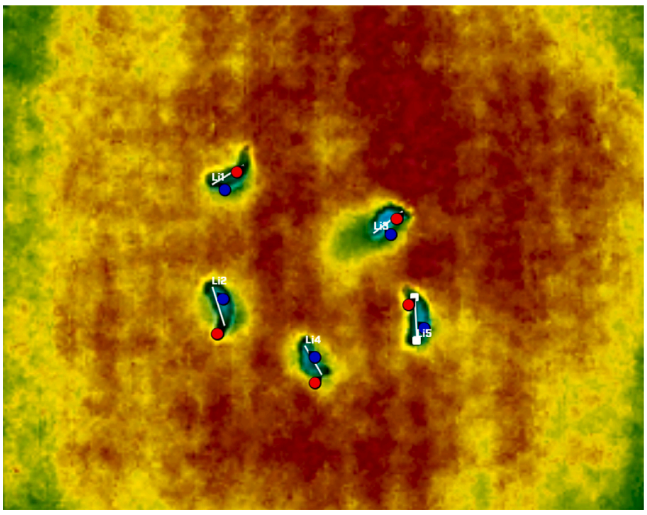


Fig. 1. Thermal image processed on the FLIR Thermal Studio software. Blue pointers correspond to coldest parts of the pupae and red pointers the warmest.

- **Hyperspectral dataset** - 26 Near-Infrared (NIR) readings were acquired using the Specim FX-17,[5] which is able to detect wavelengths between 900 nm and 1700 nm with a step of 3.57 nm and a spatial resolution of 640 px. These images were then processed on two different platforms separately to conduct different analysis.

2.4. Data processing

Initially, the software Breeze by Prediktera[6] was used to build models for pupae classification. This sotfware allows for analysis of hyperspectral images and also build and run sampling and classification models. Firstly, as a preprocessing method, a principal component analysis (PCA) (Abdi & Williams, 2010) was used to perform segmentation of the pupae on each frame, which allowed for each to be treated individually on further tests and consequently removing the background of the images. This analysis was performed by conducting training on 9 manually annotated images of the dataset. The output of the model resulted on the segmentation of the pupae on each acquired image, as can be seen in Fig. 2. Then, a partial least squares discriminant analysis (PLS-DA) (Barker & Rayens, 2003; Lee et al., 2018) was performed on 20 images to divide the pupae into two distinct classes, Dead (D) and Alive (A). Furthermore, since the model attributes a score to each wavelength band according to their relevance to distinguish between classes during training, these scores were further used to obtain the most relevant wavelengths during the training procedure. Afterwards, two of the most relevant wavelengths were selected to create an index that allowed to correlate the spectral response of the pupae with their health status.

The performance of the PLS-DA model was assessed using precision (Eq. (1)), recall (Eq. (2)) and F1-score (Eq. (3)), computed from the confusion matrix. These metrics were calculated based on the number of true positives (TP) false positives (FP), false negatives (FN), and true negatives (TN).

$$\text{Precision} = \frac{TP}{TP + FP} \tag{1}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{2}$$

[5] Specim FX-17 - https://www.specim.com/products/specim-fx17/ (Accessed on 21/11/2025).

[6] Prediktera - https://prediktera.com (Accessed on 21/11/2025).

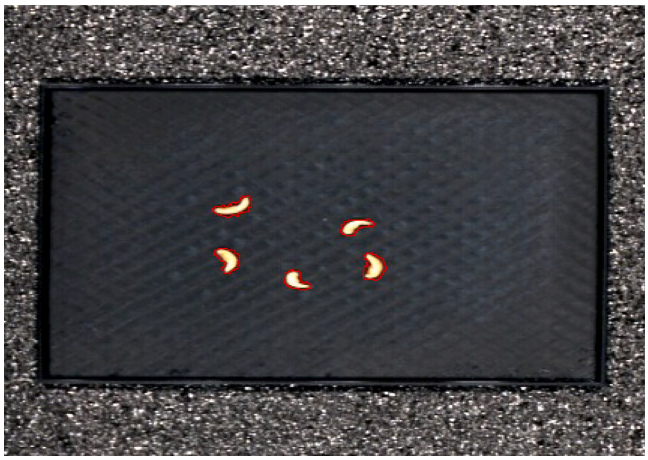


Fig. 2. Hyperspectral image with pupae segmentation performed by the PCA analysis (red contour).

$$\text{F1-Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{3}$$

The second processing method consisted on analysing the difference between alive and dead pupae using spectral slopes. Firstly, false RGB images were generated from the ENVI files to obtain segmentation masks of the pupae in each image, which are the brightest elements present on each frame due to the dark background generated by the black PLA tray. For each segmented pupa, the average reflectance on each wavelength across its correspondent pixels is obtained to determine the average reflectance for each sample. Afterwards, wavelength windows were defined based on the observed variations obtained in the reflectance plots. Within each window, a linear fit was applied to the mean spectrum, and the slope was calculated. In addition, the average reflectance in each window was also computed. Both the slope and mean values were stored on a feature vector for each pupa to be used on a class-balanced logistic regression model with a 5-fold stratified cross-validation to distinguish between alive and dead pupae.

The logistic regression model performance was assessed using the metrics mentioned on the previous method and also ROC-AUC (Fawcett, 2006). Furthermore, Balanced Accuracy Score was used to compute the best threshold for probability cutoff between the two classes.

3. Results

3.1. RGB images

As shown in Fig. 3(a), the dead pupae present a darker surface caused by melanization of the tissues (Noonin et al., 2011). On the different batches, it was possible to notice that this characteristic on the dead pupae was not always visible in its entirety. On a few samples, only a portion of the pupae was affected, as can be seen in Fig. 3(b).

This feature was more visible on the RGB images as a result of their higher resolution compared to the other cameras used in this experiment. Therefore, these images were considered during the hyperspectral and thermal data annotation procedure as references for the labelling task.

3.2. Hyperspectral images - PLS-DA & NDI

The application of the PCA model to perform the segmentation of the pupae in the acquired dataset resulted in a total of 230 pupae.

Analysing the average reflectance spectrum of the pupae, it was possible to notice that, overall, the spectral data presented positive

(a) (b)

Fig. 3. Pupae samples: (a) pupae batch containing two dead pupae; (b) pupa sample with half of the body presenting melanization.

Table 1
Confusion matrix of the PLS-DA model on a total of 230 pupae.

		Predicted	
		Dead	Alive
Actual	Dead	50 (86.2%)	7 (12.1%)
	Alive	0 (0.0%)	173 (100.0%)

Table 2
Metrics of the PLS-DA model.

	Class	
	Dead	Alive
Precision	100%	96.1%
Recall	86.2%	100%
F1-Score	92.6%	98%

peaks at around 1050 nm, 1235 nm and 1628 nm, and negative peaks at 1175 nm and 1428 nm, as can be seen in Fig. 4(a).

This characteristic was present in all the samples, being similar in all the acquired images; however, the magnitude of the values still exhibited some differences between images. Furthermore, when comparing live and dead pupae the same pattern persisted. The main difference was that the reflectance of the dead pupae tended to be lower in magnitude from wavelengths around 900 nm and 1100 nm, and afterwards higher to the upper limit of the camera of 1700 nm, as shown in Fig. 4(b). This data allows to understand that there are differences detected by spectral analysis between live and dead pupae, yet this is not clear by just visually comparing the reflectance of the samples. Therefore, the previously described PLS-DA model was employed. This approach constructs latent variables that classify each pixel based on the weighted contribution of multiple wavelengths, allowing each pixel to be distinguished between the two defined classes (Li et al., 2001).

The trained PLS-DA model achieved the performance metrics presented in Tables 1 and 2. As shown in these tables, the model demonstrated no difficulty in correctly classifying live pupae. Although the classification of dead pupae was generally accurate, a few samples were incorrectly identified as alive. Overall, the model provided reliable discrimination between live and dead *T. molitor* pupae, as shown in Fig. 5.

Hence, by analysing the weights attributed to each wavelength band by the PLS-DA, it is possible to determine which are the most relevant for classifying the samples into two classes based on their health status. In the literature, normalized difference indices (NDI) are highly used to explain relations between spectral data and physical parameters (Zanetti, 2023). Therefore, considering two of the most relevant wavelengths for the PLS-DA model, the 903 nm and the 1259 nm bands, a NDI was developed, as seen in Eq. (4). This allowed to assess whether these wavelengths were sufficient for comparison between live and dead pupae. The application of this index can be seen in Fig. 6, in which it is possible to observe that dead pupae present in the image correspond to lower values on this NDI compared to live pupae. Furthermore, the index presents negative values for portions of the dead pupae, which indicates that the reflectance for the 903 nm wavelength on those pixels is higher when compared to the 1259 nm wavelength. Considering this, it was possible to validate the NDI on misclassified pupae, such as the sample present on Fig. 3(b), which returns a negative NDI value of −0.0246, despite being considered alive by the PLS-DA prediction.

$$NDI = \frac{\rho_{1259} - \rho_{903}}{\rho_{1259} + \rho_{903}} \tag{4}$$

3.3. Hyperspectral images - Spectral slope analysis

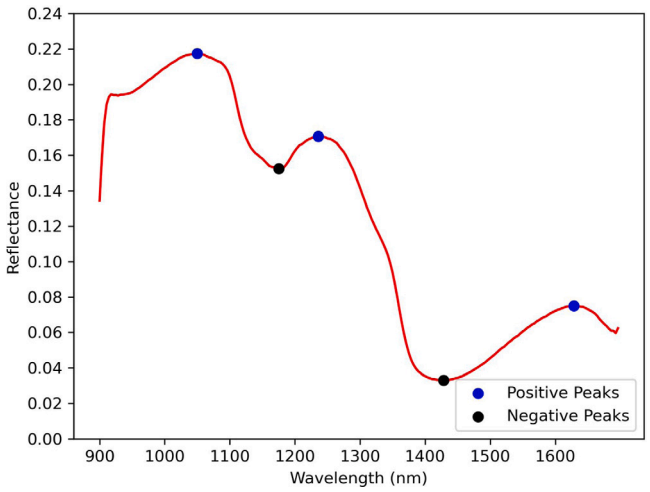
Applying the method of extracting the brightest elements on each of the 26 ENVI files resulted in a total of 341 pupae. The average reflectance per pupa was calculated on each image, as shown in the example of Fig. 7.

As demonstrated in Section 3.2, the average spectral signature of dead and live pupae present similar behaviour. With Spectral Slope analysis, this behaviour was analysed to determine if, despite similar, it could represent differences between the two classes of samples. Overall, it was possible to notice regions in which reflectance constantly presented abrupt changes in all pupae, as well as regions where the reflectance was approximately constant during several wavelengths. Therefore, the following regions were selected: 900–1050 nm, 1100–1200 nm, 1200–1300 nm, 1350–1450 nm and 1500–1650 nm.

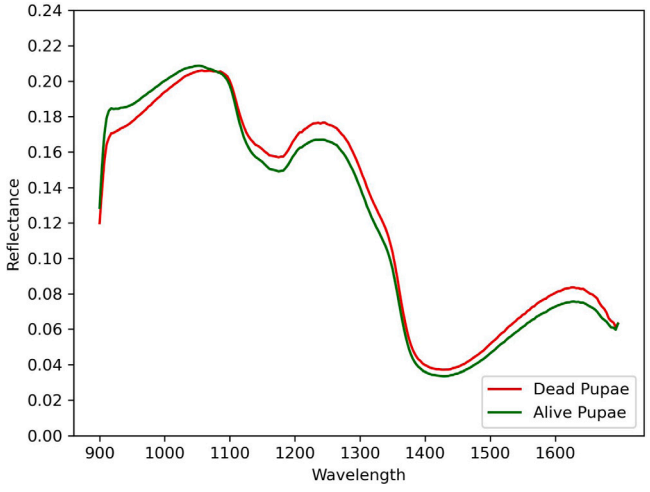
Using these windows, the slope and mean reflectance were computed for each individual pupa and used as input features on a class-balanced logistic regression classifier. Furthermore, a decision threshold for the predictions was selected to maximize balanced accuracy (Eq. (5)) and the reported metrics were computed at this threshold.

$$\text{Balanced Accuracy} = \frac{1}{2}\left(\frac{TP}{TP+FN} + \frac{TN}{TN+FP}\right) \tag{5}$$

The model achieved the performance metrics presented in Table 3.



(a)



(b)

Fig. 4. Average spectral signature of a pupae batch: (a) overall spectrum including both live and dead pupae; (b) comparison of live and dead pupae spectral signatures.

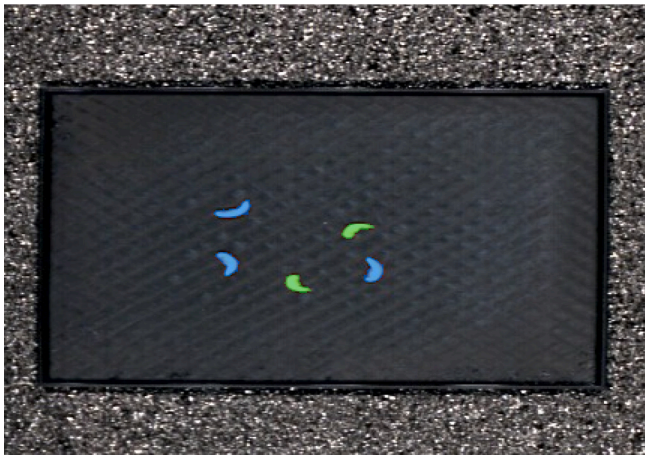


Fig. 5. Pupae classification performed by the PLS-DA model.

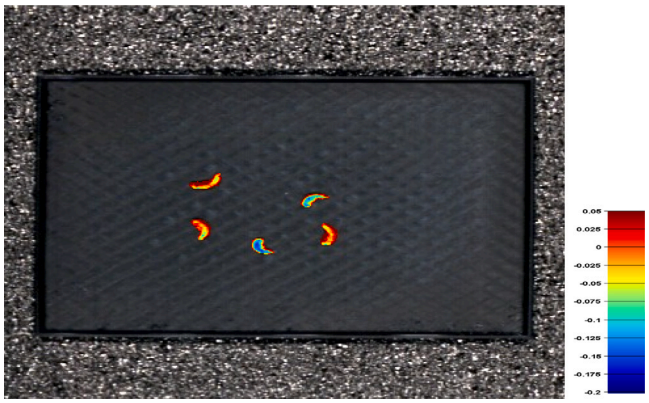


Fig. 6. NDI applied to pupae spectral signature.

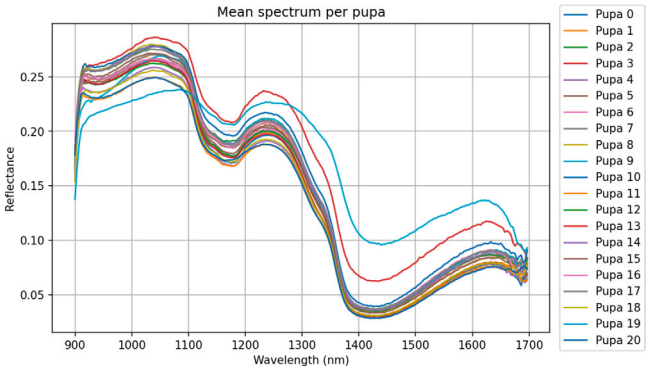


Fig. 7. Mean spectrum per pupa.

Table 3
Performance metrics of the Logistic Regression model.

Overall		
Metric	Value	
Balanced Accuracy	96.4%	
ROC-AUC	97.8%	
Class-wise		
Metric	Dead	Alive
Precision	92.5%	97.4%
Recall	94.2%	96.6%
F1-Score	93.3%	97.0%

3.4. Thermal images

The number of dead and live pupae varied across batches; however, live pupae generally exhibited higher temperatures than dead pupae. The temperature analysis of the previously presented batch as an example, containing five pupae, can be seen in Fig. 8(a). It is clear that the temperature differences between the two classes of pupae, despite noticeable, are still close in absolute values. Furthermore, to assess whether the body temperature of each pupa presented significant variations along its body, the minimum and maximum temperatures were measured on each pupa. It was observed that dead pupae exhibited a higher thermal gradient, while live pupae showed smaller temperature variations. These measurements are shown in Fig. 8(b), which corresponds to a batch containing a larger number of pupae, making the temperature variations of each individual easier to observe.

However, the absolute values obtained were not consistent along the batches, since, as previously stated, the halogen lamps on the data acquisition setup caused the overall temperature of the system to vary.

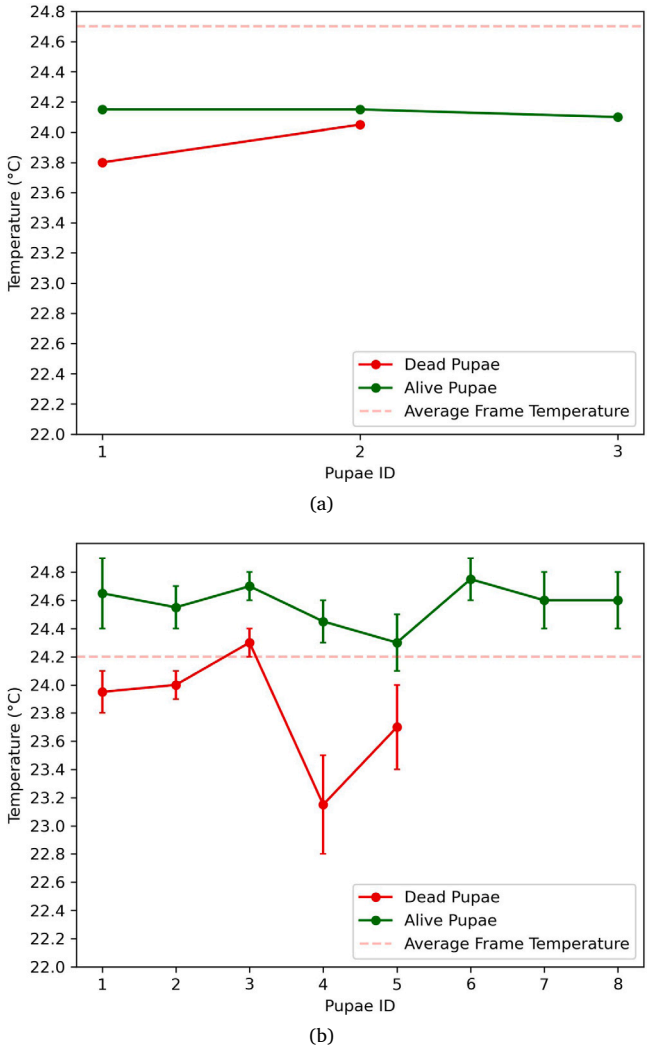


Fig. 8. Temperature analysis of pupae batches: (a) batch containing three live pupae and two dead pupae; (b) batch containing eight live pupae and five dead pupae.

Table 4
Temperature variations of the pupae along separate batches.

		Temperature (°C)	
		Minimum	Maximum
Class	Dead	22.20	25.55
	Alive	22.85	25.80

This caused the average temperature of the pupae to change up to 3 °C between the captured frames, as can be seen in Table 4. By analysing each batch individually, it was possible to understand how the temperature of the dead pupae differed from that of the live ones. Dead pupae had temperatures approximately 1% to 3% lower than live pupae, averaging around 1.8% lower.

4. Discussion

Overall, both the hyperspectral and thermal analyses provided results that enabled clear differentiation between dead and live pupae in each batch. However, while hyperspectral data is consistent between measurements providing precise information regarding the spectral signature of the pupae, allowing for the implementation of different classification models, thermal imaging is highly impacted by the environment of acquisition.

Table 5
Comparison between PLS-DA and Logistic Regression.

Method	Features	F1-Score (Dead)	F1-Score (Alive)
PLS-DA	Full Spectrum	92.6%	98%
Spectral Slopes (Logistic Regression)	Slopes + Mean Reflectance	93.3%	97%

Although the PLS-DA model accurately classified most pupae as dead or alive, some dead pupae were misclassified as alive. This may result from the progression of decay, which can lead to recently deceased pupae being identified as alive by the model. Furthermore, when analysing the calculated NDI value for these pupae, it was possible to observe that the reflectance at 903 nm was higher than at 1259 nm, resulting on a negative NDI value, similarly to pupae classified as dead by the model. The logistic regression model developed aims to minimize this risk, by adjusting the decision threshold accordingly to the maximum balanced accuracy obtained. However, despite the Recall when predicting dead pupae increasing considerably on the logistic regression compared to the PLS-DA (86.2% to 94.2%), Precision decreased from 100% to 92.5%. Consequently, the opposite behaviour happened on predicting live pupae. Despite these changes, this model, similarly to the PLS-DA, still presented more capability of detecting live pupae compared to dead ones. Table 5 shows a comparison between the two approaches side-by-side.

The PLS-DA model applied to the hyperspectral data allowed the overview of the most significant wavelengths contributing to the classification of the pupae into the two classes. As previously mentioned, the wavelengths of 903 nm and 1259 nm were considered significant by the model for that classification, which was further proven by the application of the developed NDI. This index, capable of differentiating live and dead pupae based only on two distinct wavelengths, supports the robustness of the weights previously attributed to them, proving sufficient for classifying the samples and, consequently, eliminating the need for a broader spectral range while guaranteeing a less complex system.

In previous works, such as the analysis developed by Kröncke et al. (2023), which analyses the spectral data of living mealworm larvae, the authors acknowledged that absorbance peaks near the 1200 nm can be indicators of fat content. Furthermore, Benes et al. (2022) reached a similar conclusion by analysing the content of insect powder mixtures containing *T. molitor*, assuming that absorbance in this spectral range is related to the presence of lipids. Although the lipid content of *T. molitor* pupae differs slightly from that of the larvae (Morales-Ramos et al., 2015), considering these previous works, it is possible to correlate with the behaviour found in the analysis of dead and live pupae. The spectral signature of live pupae presented a significantly lower reflectance between approximately 1200 nm and 1300 nm, and consequently higher absorbance, compared with the spectral signature of dead pupae. As dead pupae decompose, their fat content is expected to decrease, supporting previous findings regarding these wavelengths.

Furthermore, literature on vegetation indices supports that reflectance values around 900 nm can be used to estimate water content in plants (Li et al., 2021). In this study, differences in reflectance between 900 nm and 1100 nm were observed between dead and live pupae, supporting the assumption that their water content differs.

Simultaneously, logistic regression classification showed that by combining slope and mean reflectance across windows provided strong discrimination between dead and alive classes, as verified by the high balanced accuracy and ROC-AUC. This indicates that, on the NIR range, not only the reflectance on specific wavelengths can be used to classify the pupae as dead or alive but also the variation of reflectance on several windows.

Table 6
Comparison between methods for dead pupae identification.

Method	Strengths	Limitations
RGB/Manual	Only requires direct observation	Insufficient when melanization is occluded
Hyperspectral	High accuracy and provides physiological insight	Costly and complex
Thermal	Clear differences between dead and live pupae	Environment dependent

Although thermal analysis allowed comparison of temperature differences between dead and live pupae within each batch, it was found to be unreliable for implementation in a factory environment without ground-truth data. This limitation arises from the difficulty in defining consistent temperature ranges for each class. During data acquisition, variations in temperature around the samples increased the temperate gradient among pupae of the same class, sometimes causing overlap. Considering each batch individually, it was possible to notice that dead pupae presented lower temperature than live pupae. However, if the pupae were clustered and the temperature of the environment changed during data acquisition, this task would be highly compromised, resulting in incorrect classification.

Furthermore, hyperspectral analysis revealed a relationship between pupal decay and the developed NDI. Dead pupae with higher degree of deterioration, such as body perforations, corresponded to lower index values. However, in thermal analysis, this association could not be established, as differences between pupae at various stages of decomposition were not clearly distinguishable. Table 6 shows a summary of comparison between the methods addressed in this work.

5. Conclusion

The analysis of *T. molitor* pupae enabled the identification of characteristics that distinguish between dead and live pupae. Information on the health status of *T. molitor* pupae is valuable for the insect industry, as it facilitates the implementation of separation methods to remove dead pupae and prevent further contamination.

Among the methodologies tested in this study, hyperspectral data proved to be more consistent across pupae samples than thermal data, enabling a more direct assessment of their health status. Thermal imaging, however, confirmed the expected temperature differences between dead and live pupae, although the measurements were strongly influenced by ambient temperature. Both methods offer advantages over manual selection, as they do not rely solely on features observable in the visible spectrum, which may not always be easily detectable, as well as do not require an operator to classify each pupa.

However, the hyperspectral camera used in this study performs linear data acquisition, requiring a complex setup that enables sample movement in front of the camera lens during acquisition. This setup increases both system complexity and overall cost. Meanwhile, the thermal camera demands a more controlled environment to minimize the influence of the ambient temperature on measurements.

For future work, these methodologies are expected to be integrated into an autonomous sorting machine to enable the removal of dead pupae from production batches. Nevertheless, in industrial settings, pupae are more densely packed than in the datasets used for this study. This difference may affect the accuracy of the system in identifying dead pupae. To support further development, the feasibility demonstrated in this work of characterizing dead pupae using spectral and thermal sensors must be extended to more cluttered environments with higher pupae density. Additionally, improvements in real-time inference are essential to ensure effective performance of the sorting system.

CRediT authorship contribution statement

Francisco Oliveira: Writing – original draft, Methodology, Investigation, Data curation, Conceptualization. **Vítor Tinoco:** Writing – review & editing, Investigation, Data curation. **Leandro Rodrigues:** Writing – review & editing, Methodology, Investigation. **Filipe N. Santos:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization. **Mário Cunha:** Writing – review & editing, Validation, Supervision, Methodology. **Inês Vieira:** Writing – review & editing, Validation, Resources. **Marisa V. Santos:** Writing – review & editing, Validation, Resources.

Ethical statement

Ethical approval was not required for this study as it did not involve human participants, animals, or identifiable personal data.

Funding sources

Funding This work is co-financed by Component 5 - Capitalization and Business Innovation, integrated in the Resilience Dimension of the Recovery and Resilience Plan within the scope of the Recovery and Resilience Mechanism (MRR) of the European Union (EU), framed in the Next Generation EU, for the period 2021–2026, within project InsectERA, with reference 20.

Declaration of competing interest

The authors Inês Vieira and Marisa V. Santos are employed on the company that provided the pupae for the research developed during the course of this work.

Acknowledgements

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

Data availability

The authors do not have permission to share the data publicly. However, the data may be made available upon reasonable request, subject to approval and further analysis.

References

Abdi, H., & Williams, L. J. (2010). Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics*, *2*(4), 433–459. http://dx.doi.org/10.1002/wics.101.

Barker, M., & Rayens, W. (2003). Partial least squares for discrimination. *Journal of Chemometrics*, *17*(3), 166–173. http://dx.doi.org/10.1002/cem.785.

Barrett, M., Godfrey, R., Schnell, A., & Fischer, B. (2023). Farmed yellow mealworm (tenebrio molitor; coleoptera: Tenebrionidae) welfare: species-specific recommendations for a global industry. *Journal of Insects As Food and Feed*, *10*(6), 903–948. http://dx.doi.org/10.1163/23524588-20230104.

Benes, E., Biró, B., Fodor, M., & Gere, A. (2022). Analysis of wheat flour-insect powder mixtures based on their near infrared spectra. *Food Chemistry: X*, *13*, Article 100266. http://dx.doi.org/10.1016/j.fochx.2022.100266.

Bisht, A., Murugan, G., Hussain, M., & Abdi, G. (2025). Current advancements in alternative protein sources: Plant-based, cultured meat, and insect proteins for a sustainable future. *Food Bioscience*, *73*, Article 107708. http://dx.doi.org/10.1016/j.fbio.2025.107708.

EFSA Panel on Nutrition, Novel Foods and Food Allergens (NDA), Turck, D., Bohn, T., Castenmiller, J., De Henauw, S., Hirsch-Ernst, K. I., Maciuk, A., Mangelsdorf, I., McArdle, H. J., Naska, A., Pelaez, C., Pentieva, K., Siani, A., Thies, F., Tsabouri, S., Vinceti, M., Cubadda, F., Frenzel, T., Heinonen, M., ... Knutsen, H. K. (2021). Safety of frozen and dried formulations from whole yellow mealworm (tenebrio molitor larva) as a novel food pursuant to regulation (EU) 2015/2283. *EFSA Journal*, *19*(8), Article e06778. http://dx.doi.org/10.2903/j.efsa.2021.6778.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, *27*(8), 861–874. http://dx.doi.org/10.1016/j.patrec.2005.10.010.

Gasco, L., Biancarosa, I., & Liland, N. S. (2020). From waste to feed: A review of recent knowledge on insects as producers of protein and fat for animal feeds. *Current Opinion in Green and Sustainable Chemistry*, *23*, 67–79. http://dx.doi.org/10.1016/j.cogsc.2020.03.003.

Gómez-García, R., Campos, D. A., Aguilar, C. N., Madureira, A. R., & Pintado, M. (2021). Valorisation of food agro-industrial by-products: From the past to the present and perspectives. *Journal of Environmental Management*, *299*, Article 113571. http://dx.doi.org/10.1016/j.jenvman.2021.113571.

Gruensfelder, H. D. R., Shofu, F., Michie, M. S., Berezin, M. Y., Shmuylovich, L., & O'Brien, C. M. (2025). Characterization of biological absorption spectra spanning the visible to the short-wave infrared. *JoVE*, (215), Article e67403. http://dx.doi.org/10.3791/67403.

Ichikawa, T., & Kurauchi, T. (2009). Larval cannibalism and pupal defense against cannibalism in two species of tenebrionid beetles. *Zoological Science*, *26*(8), 525–529. http://dx.doi.org/10.2108/zsj.26.525.

Kröncke, N., Neumeister, M., & Benning, R. (2023). Near-infrared reflectance spectroscopy for quantitative analysis of fat and fatty acid content in living tenebrio molitor larvae to detect the influence of substrate on larval composition. *Insects*, *14*(2), http://dx.doi.org/10.3390/insects14020114.

Lee, L. C., Liong, C.-Y., & Jemain, A. A. (2018). Partial least squares-discriminant analysis (PLS-DA) for classification of high-dimensional (HD) data: a review of contemporary practice strategies and knowledge gaps. *Analyst*, *143*, 3526–3539. http://dx.doi.org/10.1039/C8AN00599K.

Li, B., Martin, E., & Morris, J. (2001). Latent variable selection in partial least squares modelling. *IFAC Proceedings Volumes*, *34*(25), 463–468. http://dx.doi.org/10.1016/S1474-6670(17)33867-3, 6th IFAC Symposium on Dynamics and Control of Process Systems 2001, Jejudo Island, Korea, 4-6 June 2001.

Li, H., Yang, W., Lei, J., She, J., & Zhou, X. (2021). Estimation of leaf water content from hyperspectral data of different plant species by using three new spectral absorption indices. *PLoS One*, *16*(3), 1–16. http://dx.doi.org/10.1371/journal.pone.0249351.

Madau, F. A., Arru, B., Furesi, R., & Pulina, P. (2020). Insect farming for feed and food production from a circular business model perspective. *Sustainability*, *12*(13), http://dx.doi.org/10.3390/su12135418.

Majewski, P., Lampa, P., Burduk, R., & Reiner, J. (2024). End-to-end solution for tenebrio molitor rearing monitoring with uncertainty estimation and domain shift detection. In *2024 IEEE/CVF conference on computer vision and pattern recognition workshops* (pp. 5498–5507). http://dx.doi.org/10.1109/CVPRW63382.2024.00559.

Miglietta, P. P., De Leo, F., Ruberti, M., & Massari, S. (2015). Mealworms for food: A water footprint perspective. *Water*, *7*(11), 6190–6203. http://dx.doi.org/10.3390/w7116190.

Morales-Ramos, J. A., Rojas, M. G., Shelby, K. S., & Coudron, T. A. (2015). Nutritional value of pupae versus larvae of tenebrio molitor (coleoptera: Tenebrionidae) as food for rearing podisus maculiventris (heteroptera: Pentatomidae). *Journal of Economic Entomology*, *109*(2), 564–571. http://dx.doi.org/10.1093/jee/tov338.

Muñoz-Seijas, N., Fernandes, H., Domínguez, J., & Salgado, J. (2024). Recent advances in biorefinery of tenebrio molitor adopting green technologies. *Food and Bioprocess Technology*, *18*, 1061–1078. http://dx.doi.org/10.1007/s11947-024-03510-0.

Muñoz-Seijas, N., Fernandes, H., Outeiriño, D., Morán-Aguilar, M. G., Domínguez, J. M., & Salgado, J. M. (2024). Potential use of frass from edible insect tenebrio molitor for proteases production by solid-state fermentation. *Food and Bioproducts Processing*, *144*, 146–155. http://dx.doi.org/10.1016/j.fbp.2024.01.002.

Ni, D., Nelis, J. L., Dawson, A. L., Bourne, N., Juliano, P., Colgrave, M. L., Juhász, A., & Bose, U. (2024). Application of near-infrared spectroscopy and chemometrics for the rapid detection of insect protein adulteration from a simulated matrix. *Food Control*, *159*, Article 110268. http://dx.doi.org/10.1016/j.foodcont.2023.110268.

Noonin, C., Jiravanichpaisal, P., Söderhäll, I., Merino, S., Tomás, J. M., & Söderhäll, K. (2011). Melanization and pathogenicity in the insect, tenebrio molitor, and the crustacean, pacifastacus leniusculus, by aeromonas hydrophila AH-3. *PLoS One*, *5*(12), 1–10. http://dx.doi.org/10.1371/journal.pone.0015728.

Poveda, J., Jiménez-Gómez, A., Saati-Santamaría, Z., Usategui-Martín, R., Rivas, R., & García-Fraile, P. (2019). Mealworm frass as a potential biofertilizer and abiotic stress tolerance-inductor in plants. *Applied Soil Ecology*, *142*, 110–122. http://dx.doi.org/10.1016/j.apsoil.2019.04.016.

Rumbos, C., Oonincx, D., Karapanagiotidis, I., Vrontaki, M., Gourgouta, M., Asimaki, A., Mente, E., & Athanassiou, C. (2022). Agricultural by-products from Greece as feed for yellow mealworm larvae: circular economy at a local level. *Journal of Insects As Food and Feed*, *8*(1), 9–22. http://dx.doi.org/10.3920/JIFF2021.0044.

See, J. E. (2012). *Visual inspection : a review of the literature*: *Technical report*, Sandia National Laboratories, http://dx.doi.org/10.2172/1055636.

Sharma, B., Yadav, D. K., Malakar, S., Singh, S., Sharma, M., Suri, S., & Sridhar, K. (2024). Insect proteins – production technologies, bio-functional, and food applications: A perspective. *Food Bioscience*, *61*, Article 104560. http://dx.doi.org/10.1016/j.fbio.2024.104560.

Vågsholm, I., Arzoomand, N. S., & Boqvist, S. (2020). Food security, safety, and sustainability—Getting the trade-offs right. *Frontiers in Sustainable Food Systems*, *4*, http://dx.doi.org/10.3389/fsufs.2020.00016.

Zanetti, M. (2023). A one-class classification model for burned-area detection based on mutual ordering of normalized differences. *IEEE Transactions on Geoscience and Remote Sensing*, *61*, 1–15. http://dx.doi.org/10.1109/TGRS.2023.3301056.