ICPA
17th INTERNATIONAL CONFERENCE ON PRECISION AGRICULTURE
ConBAP 2026
CONGRESSO BRASILEIRO DE AGRICULTURA DE PRECISÃO E DIGITAL
13-16 July 2026 | Porto Alegre, Brazil
ISPA
AsBraAP

Unified Detection and Weight Estimation of Small Fruits Using Multi-Task Vision Models in Precision Agriculture

Theoren Stroud [1], Shekhar Suman Borah[1], Prabha Sundaravadivel [1], Janie F. Ross [2], Blair J. Sampson [2], Patricia Knight [3], Siva P. Kumpatla [2]

[1]Center for Robotics and Intelligent Systems, The University of Texas at Tyler, Tyler, TX, USA

[2]Thad Cochran Southern Horticultural Laboratory, USDA-ARS, Poplarville, MS, USA.

[3]Coastal Research and Extension Center, Mississippi State University, Poplarville, MS, USA.

A paper from the Proceedings of the

17th International Conference on Precision Agriculture and 11th Brazilian Congress on Precision and Digital Agriculture

13-15 July 2026

Porto Alegre, Brazil

Abstract.

Fruit weight is an important trait for yield prediction, grading, and postharvest decisions, but it is usually measured with destructive and labor-intensive methods. We propose a branched YOLOv8 model that performs object detection and, through an additional head, regresses an informational vector (IV), defined as grape weight (grams) for real muscadine images and as rotation angle for a synthetic digit dataset. The regression head shares features with the YOLOv8 backbone and is trained with an extra mean squared error loss term, enabling multi-task learning while maintaining detection performance. On the synthetic digit benchmark, the model achieved 79.2% mAP50, 19.4% mAP50-95, and an IV mean squared error of 0.416 deg^2, compared with an angle variance of 675 deg^2. On the muscadine hyperspectral dataset, the final model reached 85.7% mAP50, 80.5% mAP50-95, and an IV mean squared error of 1.67 g^2, compared with a weight variance of about 2.829 g^2. These results show that the IV branch can recover rotation angles in the synthetic benchmark and predict muscadine grape weights with useful accuracy, suggesting that non-destructive image-based weight estimation is feasible. The model is designed to accept hyperspectral inputs through modified early layers, and its compact single-stage design makes it suitable for real-time edge deployment in precision agriculture applications.

Keywords.
Fruit phenotyping; muscadine grapes; YOLOv8; hyperspectral imaging; object detection; multi-task learning; weight estimation.

The authors are solely responsible for the content of this paper, which is not a refereed publication. Citation of this work should state that it is from the Proceedings of the 17th International Conference on Precision Agriculture and 11th Brazilian Congress on Precision and Digital Agriculture. EXAMPLE: Last Name, A. B. & Coauthor, C. D. (2026). Title of paper. In Proceedings of the 17th International Conference on Precision Agriculture and 11th Brazilian Congress on Precision and Digital Agriculture (unpaginated, online). Monticello, IL: International Society of Precision Agriculture and Brazilian Congress on Precision and Digital Agriculture.

Unified Detection and Weight Estimation of Small Fruits Using Multi-Task Vision Models in Precision Agriculture

1 Introduction

Fruit weight is an important phenotypic trait for yield estimation, grading, and postharvest decision-making. However, it is still commonly measured using manual and often destructive procedures that are labor-intensive and difficult to scale. To address these limitations, non-destructive sensing approaches based on computer vision and spectral imaging have received increasing attention for estimating fruit quality and phenotypic characteristics (Zhao et al., 2025; Sabouri et al., 2025). Recent studies have shown that fruit weight and other continuous quality attributes can be predicted from image-derived geometric features and spectral information. Machine learning and deep learning models have been successfully applied to estimate fruit weight from RGB imagery, while hyperspectral imaging (HSI) has been used to predict attributes such as firmness, soluble solids content, and maturity (Gidado et al., 2026; Lanke & Chandak, 2025). Despite these advances, most existing approaches estimate traits separately from object detection, which results in fragmented processing pipelines.

At the same time, YOLO-based object detectors have become widely adopted for fruit detection, counting, and yield estimation because of their high accuracy and real-time performance under field conditions (Badgujar et al., 2024; Vijayakumar & Vairavasundaram, 2024). However, these systems generally provide only object locations and class labels, with limited capability for estimating continuous phenotypic traits such as fruit weight. This limits their usefulness for detailed phenotyping and fruit grading applications.

Multi-task learning offers a promising solution by allowing related tasks to share a common feature representation. Previous studies have demonstrated the effectiveness of joint detection and regression frameworks for agricultural applications, including fruit size estimation and quality assessment (Albaaji et al., 2025; Khan et al., 2025). Nevertheless, limited research has explored the integration of a modern one-stage detector such as YOLOv8 with an auxiliary regression branch for simultaneous fruit detection and per-object weight estimation, particularly within a lightweight architecture suitable for edge deployment and future hyperspectral integration.

To address this gap, this study proposes a YOLOv8-based multi-task framework that combines object detection and continuous trait estimation through an auxiliary informational vector (IV) branch. The proposed architecture is evaluated using both a synthetic digit-rotation benchmark and a muscadine grape dataset, where the regression branch is used to predict rotation angle and grape weight, respectively. In addition, the framework is designed to support hyperspectral imagery through modifications to the early convolutional layers, enabling future integration of spectrally informative bands for enhanced quantitative trait estimation.

The main contributions of this work are as follows:

a) A YOLOv8-based multi-task framework is proposed that jointly performs object detection and scalar trait regression through an auxiliary informational vector branch.

b) A synthetic digit-rotation benchmark is introduced to evaluate the regression capability of the proposed architecture in a controlled multi-object environment.

c) The proposed framework is applied to non-destructive muscadine grape weight estimation, demonstrating simultaneous detection and quantitative trait prediction within a single model.

d) A hyperspectral extension of the framework is presented to support future integration of selected spectral bands within a unified lightweight detection and regression architecture.

2 Related Work

Deep object detection models have become a fundamental component of modern precision agriculture systems, particularly for fruit detection, counting, and yield estimation under challenging field conditions (Huang et al., 2023). YOLO-based architecture has been widely adopted because they provide high detection accuracy while maintaining real-time inference speeds. Improved variants of YOLOv4 and YOLOv5 have been successfully applied to citrus detection in orchard environments, demonstrating robust performance under varying illumination conditions, fruit occlusions, and complex backgrounds (Xu et al., 2023; Lawal et al., 2023). Similar approaches have also been developed for pear detection and multi-fruit counting applications, highlighting the suitability of YOLO architecture for deployment in automated harvesting and yield-monitoring systems (Liu et al., 2026; Zhao et al., 2024). Beyond detection accuracy, recent studies have emphasized the importance of evaluating model robustness and deployment reliability in agricultural vision systems. Annotation-free robustness metrics have been proposed to assess counting consistency and model stability across varying imaging conditions, providing complementary measures of practical performance without requiring extensive manual annotations (Manjunatha et al., 2026). Non-destructive estimation of fruit size and weight has been investigated using both traditional machine learning and deep learning approaches. Previous studies have used image-derived geometric features such as diameter, area, and perimeter to estimate fruit weight through regression models (Miranda et al., 2023). Deep-learning-based systems have further combined object detection and segmentation with regression techniques to estimate fruit dimensions and weight from image data. These approaches have demonstrated promising accuracy across various fruit species, indicating that visual information can provide reliable indicators of physical fruit characteristics (Charisis & Argyropoulos, 2024; Yang et al., 2024). However, most studies perform detection and regression as separate processing stages, increasing system complexity and computational requirements.

Hyperspectral imaging (HSI) has emerged as a powerful tool for non-destructive phenotyping because of its ability to capture detailed spectral information related to both external and internal fruit characteristics (Hong et al., 2026). Numerous studies have applied HSI and spectral regression techniques to predict quality attributes such as firmness, maturity, soluble solids content, and textural properties (Wu et al., 2026). Both traditional machine learning methods and deep learning architectures have demonstrated strong predictive performance using hyperspectral data, while band-selection strategies have shown that a limited number of informative wavelengths can often achieve performance comparable to full-spectrum models (Ahmad et al., 2025; Bai et al., 2025). These findings suggest that hyperspectral information can provide valuable complementary features for quantitative trait estimation. Recent advances have also explored multimodal agricultural AI frameworks that combine imaging, sensing, and intelligent analytics for large-scale phenotyping applications (Sundaravadivel et al., 2026), highlighting the broader trend toward extracting quantitative crop traits from image-based data. Multi-task learning has increasingly been adopted in computer vision to enable multiple related tasks to share a common feature representation within a single network. In agricultural applications, multi-task frameworks have been used to jointly perform fruit detection, segmentation, and size estimation. Ferrer-Ferrer et al. (2023) demonstrated that a single multitask deep neural network could jointly perform fruit detection and size estimation, illustrating the advantages of unified architectures that combine object detection with auxiliary prediction tasks to improve efficiency and reduce computational overhead.

Although previous studies have demonstrated the effectiveness of object detection, fruit trait regression, hyperspectral analysis, and multi-task learning individually, limited research has integrated these capabilities into a single lightweight framework. Few studies have explored the use of a modern one-stage detector such as YOLOv8 for simultaneous object detection and per-object quantitative trait estimation. The framework proposed in this work addresses this gap by incorporating an auxiliary informational vector branch into YOLOv8, enabling joint detection and regression while supporting future integration of hyperspectral imagery.

3. Materials and Methods

3.1 Proposed Multi-Task YOLOv8 Architecture

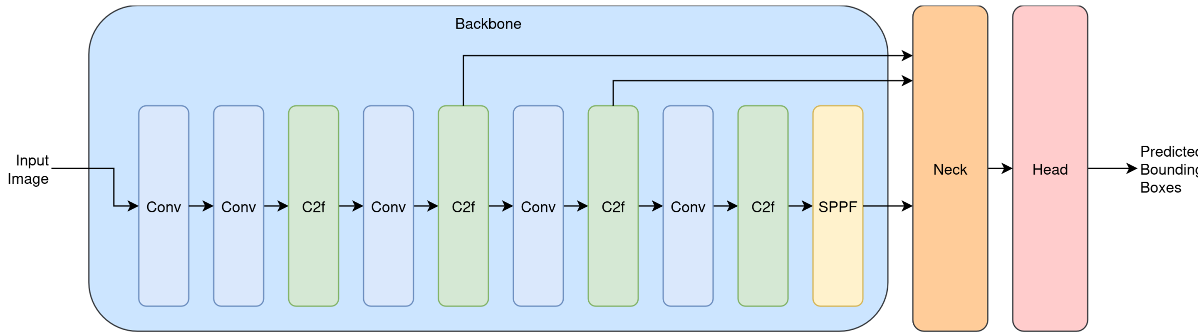


Figure 1: High level architecture of the YOLOv8 family of models with the backbone expanded.

The proposed framework extends the YOLOv8 object detection architecture by adding an auxiliary regression branch that predicts a secondary output vector, referred to as the informational vector (IV). This design enables object detection and image-level regression within a single multi-task framework. While the detection branch performs localization and classification, the IV branch predicts continuous quantitative traits associated with the input image. The YOLOv8 family was selected as the base architecture because it offers a good balance between detection accuracy and computational efficiency. As shown in Figure 1, the backbone network is responsible for feature extraction and serves as the attachment point for the proposed regression branch. The framework was implemented in PyTorch using modifications to the standard Ultralytics YOLOv8 architecture. The regression branch receives an intermediate three-dimensional feature tensor from the last convolutional layer of the YOLOv8 backbone and processes it through a series of layers to produce the IV output. This allows the regression task to use the same learned feature representations as object detection while adding only limited computational overhead.

To support hyperspectral imagery, the input layer of the YOLOv8 backbone was modified to accept an arbitrary number of channels. For the hyperspectral experiments, the original three-channel RGB input was expanded to 300 spectral channels. The first convolutional layer was expanded from three input channels to 300 input channels, and the output channel dimension of the first convolutional layer was increased by a factor of ten to provide a smoother transition from high-dimensional spectral input to later feature extraction layers. The expanded layers were initialized using bilinear upscaling of the pretrained convolutional tensors to preserve information from the pretrained model. To incorporate the regression objective during training, an additional informational loss term was introduced. This loss was computed as the mean squared error between the predicted IV and the corresponding ground-truth values. The informational loss was scaled by a configurable weighting factor and added to the standard YOLOv8 detection loss components, allowing both detection and regression objectives to be optimized jointly during training.

3.2 Datasets

Two datasets were used to evaluate the proposed multi-task framework. The first dataset was a synthetic benchmark designed to assess the regression capability of the informational vector (IV) branch independently of the hyperspectral fruit data. The dataset consisted of three-channel images containing strings of two to ten digits rotated by random angles ranging from −45° to 45°. Ground-truth annotations included bounding boxes and class labels for each digit, along with the global rotation angle applied to the image. During training, the object detection component was responsible for identifying digit locations and classes, while the IV branch was trained to predict the rotation angle. The second dataset consisted of hyperspectral images of muscadine grapes. The dataset contained 90 hyperspectral images, each comprising 300 spectral channels. Ground-truth annotations included grape bounding boxes, variety class labels (bronze, purple, and dark purple), and grape weights measured in grams. In this experiment, the object detection component was tasked with localizing and classifying grapes, while the IV branch was trained to estimate grape weight. The dataset was divided into approximately 80% training data and 20% validation data.

3.3 Training Procedure

Training was performed using the Ultralytics framework with custom modifications to support both the informational vector (IV) branch and hyperspectral image inputs. The proposed architecture was evaluated using the synthetic digit dataset and the hyperspectral muscadine grape dataset. For the synthetic digit experiment, YOLOv8-large was used as the base architecture together with the regression branch. The model was trained using the hyperparameters summarized in Table 1. During training, the detection component learned to identify digit classes and bounding boxes, while the IV branch learned to predict the global rotation angle of each image.

Table 1: Key hyperparameters during training of the synthetic digit dataset model.

Hyperparameter	Value
Initial Learning Rate (lr0)	7.14×10^{-4}
Final Learning Rate (lrf)	0.01
Momentum	0.9
Optimizer	AdamW
Batch Size	4
Box Loss Weight	7.5
Class Loss Weight	0.5
DFL Loss Weight	1.5
Informational Loss Weight	1.0

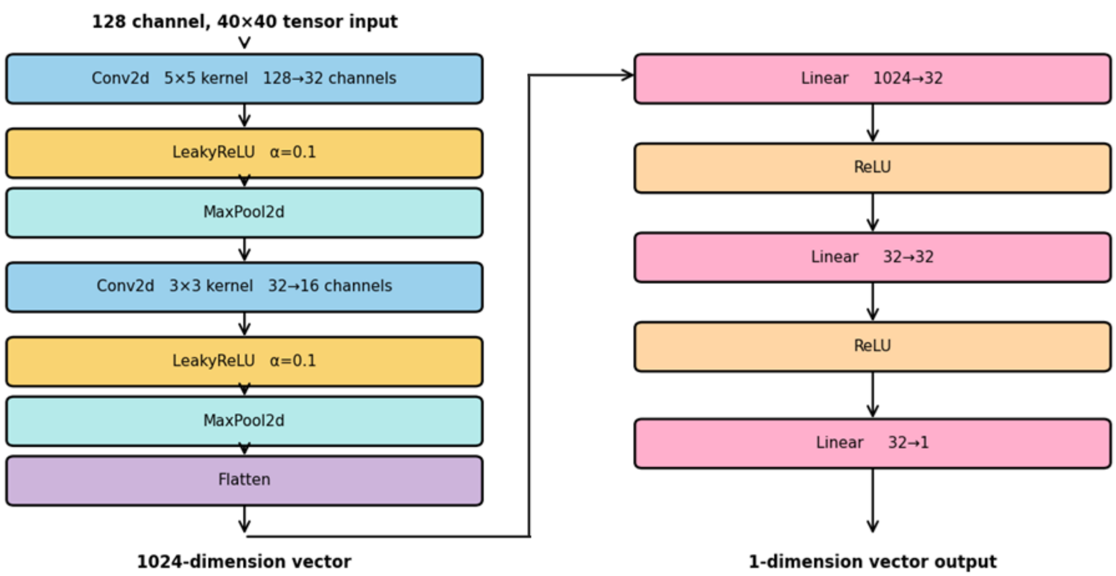


Figure 2: Branch model architecture used for hyperspectral data.

For the hyperspectral muscadine grape dataset, YOLOv8-nano was used as the base architecture together with the modified branch structure shown in Figure 2. Training was carried out in two stages. In the first stage, the model was trained for 300 epochs with the informational loss weight set to zero, so only the detection component was trained while the regression branch remained inactive. The hyperparameters used in this stage are summarized in Table 2.

The best-performing model from the first stage was then used to initialize a second 300-epoch training session. In this stage, the informational loss weight was set to 1.0, allowing joint optimization of detection and regression. The hyperparameters used for the final model are provided in Table 3. For the hyperspectral experiments, the default Ultralytics augmentation strategies were disabled. Instead, a custom augmentation pipeline was applied during data loading. The augmentations included image shifting, rotation, mirroring, brightness adjustment, and uniform noise.

Table 2: Key hyperparameters during training of the intermediate (no info loss) model.

Hyperparameter	Value
Initial Learning Rate (lr0)	3×10^{-6}
Final Learning Rate (lrf)	0.1
Momentum	0.937
Optimizer	SGD
Batch Size	2
Box Loss Weight	1.0
Class Loss Weight	200.0
DFL Loss Weight	0.5
Informational Loss Weight	0.0

Table 3: Key hyperparameters during training of the final model.

Hyperparameter	Value
Initial Learning Rate (lr0)	1×10^{-6}
Final Learning Rate (lrf)	0.1
Momentum	0.937
Optimizer	AdamW
Batch Size	2
Box Loss Weight	2.0
Class Loss Weight	100.0
DFL Loss Weight	1.0
Informational Loss Weight	1.0

4. Results and Discussion

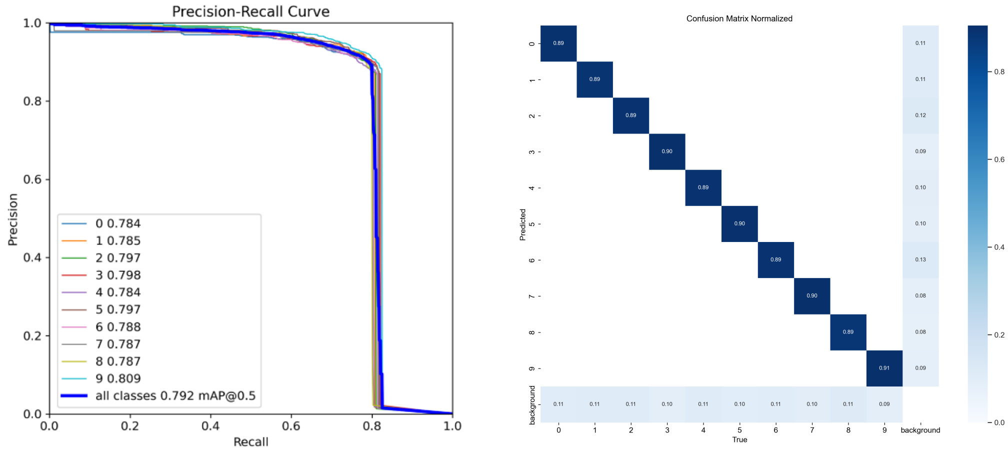
The proposed multi-task framework was evaluated using both a synthetic digit-rotation dataset and a hyperspectral muscadine grape dataset. The synthetic dataset was used to assess the regression capability of the informational vector (IV) branch in a controlled environment, while the muscadine dataset was used to evaluate the framework on a real-world fruit phenotyping task involving simultaneous detection, classification, and weight estimation.

4.1 Synthetic Digit-Rotation Experiment

The synthetic digit dataset was used to evaluate whether the proposed architecture could perform image-level regression while simultaneously carrying out object detection. Using YOLOv8-large as the base architecture, the model was trained to detect individual digits and predict the global rotation angle of each image. The results are summarized in Table 4, and the corresponding precision-recall curve and confusion matrix are shown in Figure 3. The model achieved a mAP50 of 79.2% and a mAP50-95 of 19.4%. Although the mAP50-95 value was relatively low, the confusion matrix showed a strong diagonal pattern, indicating that the model generally identified the correct digit classes. These results suggest that the lower mAP50-95 was mainly due to bounding-box localization rather than classification errors. For the regression task, the model achieved an informational mean squared error of 0.416 deg^2. This value is much lower than the variance of the rotation-angle distribution, which was 675 deg^2, showing that the regression branch was able to estimate the rotation angle accurately. Example predictions from the trained model are shown in Figure 4.

Table 4: Results of the best trained model.

mAP50	mAP50-95	Info MSE (deg^2)
79.2%	19.4%	0.416



(a) Precision-Recall Curve. **(b) Confusion Matrix**

Figure 3: Results of the digit identification model.

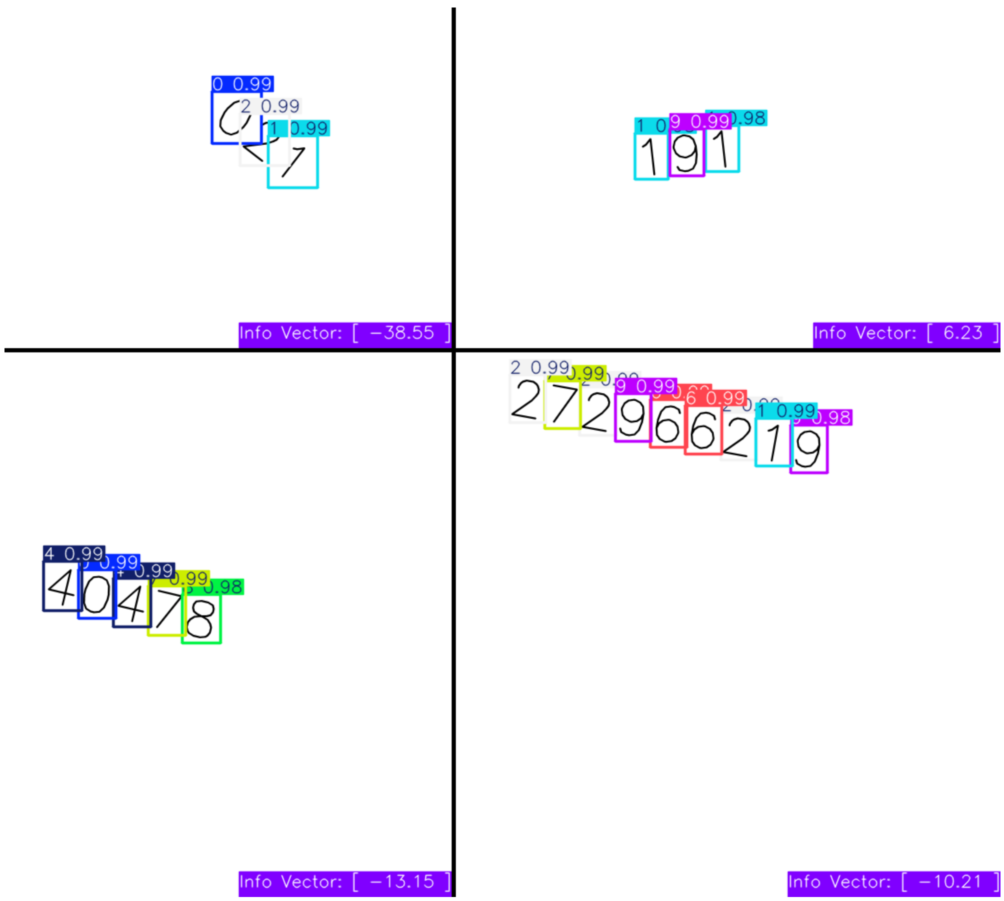


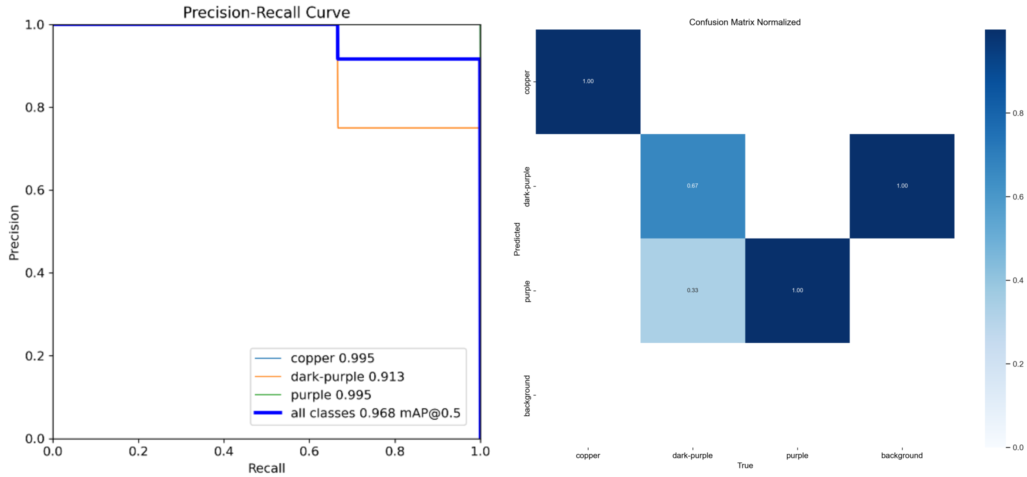
Figure 4: Example predictions of the number identifier model.

4.2 Hyperspectral Muscadine Grape Experiment

The proposed framework was further evaluated using a hyperspectral muscadine grape dataset consisting of 90 images with variety labels and grape weights. To assess the effect of the regression branch, training was performed in two stages. The first stage optimized only the object detection task, while the second stage jointly optimized both detection and regression objectives. The results of the intermediate model trained without informational loss are presented in Table 5 and Figure 5. This model achieved a mAP50 of 96.8% and a mAP50-95 of 96.3%, indicating strong detection and classification performance when the regression objective was excluded from training.

Table 5: Results of the best trained intermediate (no info loss) model.

mAP50	mAP50-95
96.8%	96.3%



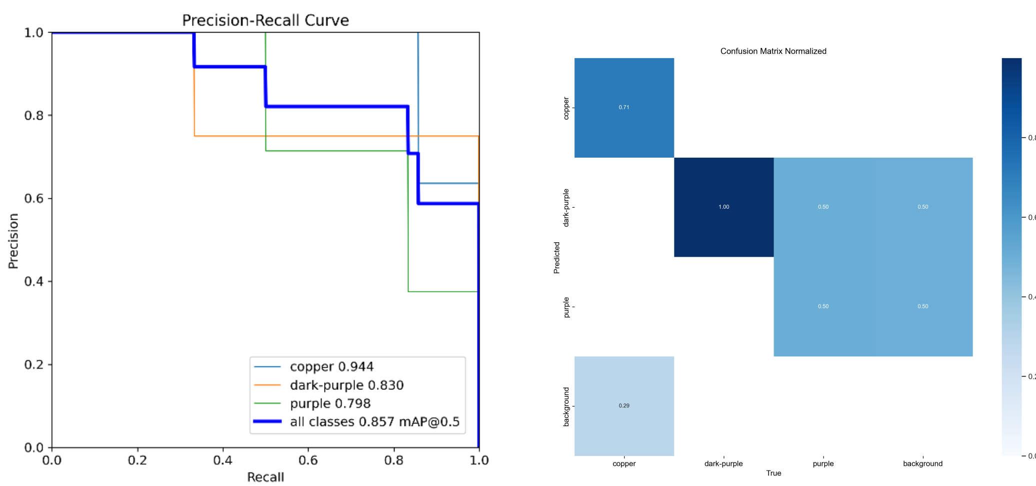
(a) **Precision-Recall Curve.** (b) **Confusion Matrix**

Figure 5: Results of the best trained intermediate (no info loss) model

The performance of the final multi-task model is summarized in Table 6 and Figure 6. After enabling the informational loss, the model achieved a mAP50 of 85.7% and a mAP50-95 of 80.5%. Although these values were lower than those of the intermediate model, the relatively small difference between mAP50 and mAP50-95 suggests that the predicted bounding boxes remained closely aligned with the ground-truth annotations. The confusion matrix indicates that the main source of error was class confusion between the purple and dark-purple grape varieties rather than inaccurate localization.

Table 6: Results of the best trained final model.

mAP50	mAP50-95	Info MSE (g^2)
85.7%	80.5%	1.67



(a) **Precision-Recall Curve.** (b) **Confusion Matrix.**

Figure 6: Results of the best trained final model

For the regression task, the final model achieved an informational mean squared error of 1.67 g². The grape-weight variance in the dataset was approximately 2.829 g². Because the prediction error was lower than the observed variance, the model was able to learn meaningful relationships between hyperspectral image features and grape weight. Example predictions from the final model are shown in Figure 7.

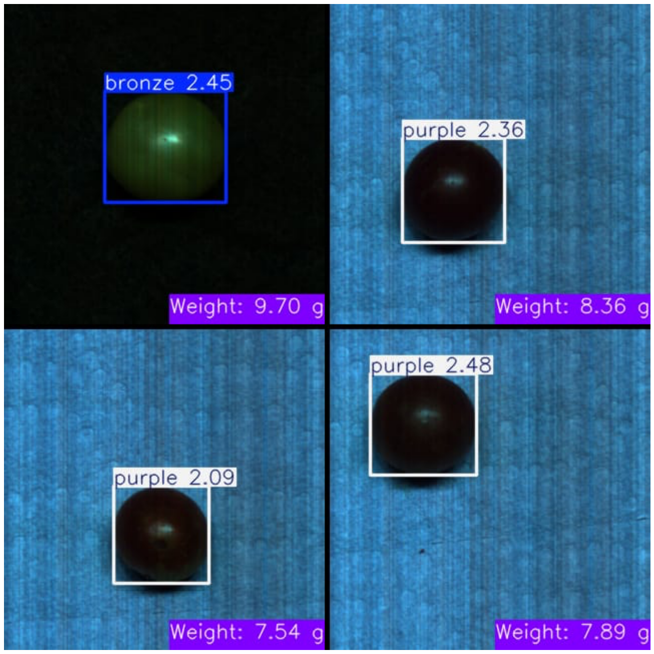


Figure 7: Example predictions of the best trained final model. The bottom-right image is a dark-purple grape, so it is a misclassification by the model. All other grapes were correctly classified.

The results show that the proposed architecture can perform object detection and quantitative trait estimation within a single framework. The synthetic experiment demonstrated the effectiveness of the IV branch for regression, while the muscadine grape experiment showed its usefulness for non-destructive fruit phenotyping. At the same time, the reduction in detection performance after adding the regression objective suggests a trade-off between detection and regression accuracy. Future work may benefit from exploring different loss weights, branch designs, and larger datasets, but the present results already show that the framework is practical for precision agriculture applications.

5. Limitations and Future Work

The proposed framework shows that a branched YOLOv8 model can support both object detection and quantitative trait regression in a single architecture. However, the results also show that adding the regression objective can reduce detection performance compared with a detection-only model. This suggests that the balance between the two tasks is important and may need further tuning. Another limitation is the use of relatively small datasets, especially for the muscadine grape experiment. With only 90 hyperspectral images, the model was evaluated on limited data, so the results should be interpreted with caution. The synthetic digit experiment is useful for testing the regression branch, but it does not fully reflect the complexity of real fruit phenotyping. The hyperspectral extension also requires modified input layers and custom handling of high-dimensional spectral data. While the model was adapted to accept 300-channel images, this setup may not generalize directly to other hyperspectral sensors or band configurations. In addition, the current experiments used a single fruit dataset, so broader validation across other fruit crops and imaging conditions is still needed. Future work should explore alternative loss weights, different branch designs, and other base architectures to improve the balance between detection and regression. Further testing with larger datasets and additional fruit species would also help assess the generality of the framework. Since the model is designed to remain lightweight, future studies may also examine its suitability for broader edge deployment and real-time field use.

6. Conclusion

This study proposed a lightweight multi-task YOLOv8 framework that combines object detection with continuous trait regression through an auxiliary informational vector (IV) branch. The model was evaluated using both a synthetic digit-rotation benchmark and a hyperspectral muscadine grape dataset, demonstrating that a single architecture can support object detection and image-level regression within a unified framework. The synthetic experiment demonstrated that the regression branch could accurately estimate rotation angle while maintaining digit detection performance. The muscadine experiment further showed that the framework could learn from hyperspectral imagery and predict grape weight in a non-destructive manner while simultaneously performing fruit detection and classification. These results demonstrate the feasibility of integrating quantitative trait estimation into a lightweight object detection framework for precision agriculture applications. The results also revealed a trade-off between detection and regression performance when the informational loss was incorporated into training, suggesting that further optimization of the multi-task learning process is needed. Future work will focus on improving the balance between detection and regression objectives, evaluating larger and more diverse datasets, and investigating the use of selected hyperspectral bands for enhanced trait estimation. With further refinement, the proposed framework may provide a useful foundation for field-deployable systems that combine detection and quantitative phenotyping within a single model.

Acknowledgement

The materials are based upon work supported by the USDA-ARS NACA 58-6062-6-006. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of USDA-ARS.

References

Ahmad, M., Distefano, S., Khan, A. M., Mazzara, M., Li, C., Li, H., Aryal, J., Ding, Y., Vivone, G., & Hong, D. (2025). A comprehensive survey for hyperspectral image classification: The evolution from conventional to transformers and Mamba models. *Neurocomputing, 644*, 130428. https://doi.org/10.1016/j.neucom.2025.130428

Albaaji, G. F., Chandra, V. C. S. S., & Sharafudeen, M. (2025). An intelligent multi-modal neural framework for accurate fruit grading localization and yield estimation. *Expert Systems with Applications, 268*, 126366. https://doi.org/10.1016/j.eswa.2024.126366

Badgujar, C. M., Poulose, A., & Gan, H. (2024). Agricultural object detection with You Only Look Once (YOLO) algorithm: A bibliometric and systematic literature review. *Computers and Electronics in Agriculture, 223*, 109090. https://doi.org/10.1016/j.compag.2024.109090

Bai, L., Chen, Z., Zhou, L., et al. (2025). HSI classification: A comprehensive review, dataset, and future trends. *Journal of the Indian Society of Remote Sensing, 53*, 3983–4004. https://doi.org/10.1007/s12524-025-02275-z

Charisis, C., & Argyropoulos, D. (2024). Deep learning-based instance segmentation architectures in agriculture: A review of the scopes and challenges. *Smart Agricultural Technology, 8*, 100448. https://doi.org/10.1016/j.atech.2024.100448

Ferrer-Ferrer, M., Ruiz-Hidalgo, J., Gregorio, E., Vilaplana, V., Morros, J.-R., & Gené-Mola, J. (2023). Simultaneous fruit detection and size estimation using multitask deep neural networks. *Biosystems Engineering, 233*, 63–75. https://doi.org/10.1016/j.biosystemseng.2023.07.010

Gidado, M. J., Gunny, A. A. N., Gopinath, S. C. B., Devi, M., & Ibrahim, A. M. (2026). Artificial intelligence and machine learning in postharvest fruit quality assessment: Current challenges, recent advances, and future prospects. *Journal of Stored Products Research, 116*, 102959. https://doi.org/10.1016/j.jspr.2026.102959

Hong, D., Li, C., Yokoya, N., et al. (2026). Hyperspectral imaging. *Nature Reviews Methods Primers, 6*, 19. https://doi.org/10.1038/s43586-026-00470-x

Huang, Y., Qian, Y., Wei, H., Lu, Y., Ling, B., & Qin, Y. (2023). A survey of deep learning-based object detection methods in crop counting. *Computers and Electronics in Agriculture, 215*, 108425. https://doi.org/10.1016/j.compag.2023.108425

Khan, Z., Shen, Y., & Liu, H. (2025). Object detection in agriculture: A comprehensive review of methods, applications, challenges, and future directions. *Agriculture, 15*(13), 1351. https://doi.org/10.3390/agriculture15131351

Lanke, N. P., & Chandak, M. B. (2025). Recent trends in deep learning and hyperspectral imaging for fruit quality analysis: An overview. *Journal of Optics*. https://doi.org/10.1007/s12596-025-02881-7

Lawal, O. M., Zhu, S., & Cheng, K. (2023). An improved YOLOv5s model using feature concatenation with attention mechanism for real-time fruit detection and counting. *Frontiers in Plant Science, 14*, 1153505. https://doi.org/10.3389/fpls.2023.1153505

Liu, S., Xue, J., Zhang, T., Lv, P., Zhao, T., Sun, H., Liu, R., & Chen, Y. (2026). Lightweight multi-classification pear fruit high-precision detection model in complex orchard scenes. *Engineering Applications of Artificial Intelligence, 175*, 114617. https://doi.org/10.1016/j.engappai.2026.114617

Manjunatha, H., Sundaravadivel, P., Borah, S. S., Tamil, L., Knight, P. R., Torbert, H. A., & Kumpatla, S. P. (2026). Beyond green: Domain-aware self-validated instance counting for Loropetalum and non-green ornamental species using annotation-free robustness metrics. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops* (pp. 9777–9786).

Miranda, J. C., Gené-Mola, J., Zude-Sasse, M., Tsoulias, N., Escolà, A., Arnó, J., Rosell-Polo, J. R., Sanz-Cortiella, R., Martínez-Casasnovas, J. A., & Gregorio, E. (2023). Fruit sizing using AI: A review of methods and challenges. *Postharvest Biology and Technology, 206*, 112587. https://doi.org/10.1016/j.postharvbio.2023.112587

Sabouri, A., Bakhshipour, A., Poorsalehi, M., et al. (2025). Machine learning techniques for non-destructive estimation of plum fruit weight. *Scientific Reports, 15*, 751. https://doi.org/10.1038/s41598-024-85051-2

Sundaravadivel, P., Manjunatha, H., Narasimhamurthy, K., Borah, S. S., Anand, A., Torbert, H. A., Knight, P., Tamil, L., & Kumpatla, S. P. (2026). Ros-AI: An LLM-enhanced scalable multimodal framework for UAV-based rose bloom analysis in precision agriculture. *Smart Agricultural Technology, 14*, 102267. https://doi.org/10.1016/j.atech.2026.102267

Vijayakumar, A., & Vairavasundaram, S. (2024). YOLO-based object detection models: A review and its applications. *Multimedia Tools and Applications, 83*, 83535–83574. https://doi.org/10.1007/s11042-024-18872-y

Wu, S., Huang, L., & Huang, R. (2026). Review of advances in hyper-spectral imaging-based non-destructive testing methods in food application. *Journal of Food Composition and Analysis, 150*, 108860. https://doi.org/10.1016/j.jfca.2026.108860

Xu, L., Wang, Y., Shi, X., Tang, Z., Chen, X., Wang, Y., Zou, Z., Huang, P., Liu, B., Yang, N., Lu, Z., He, Y., & Zhao, Y. (2023). Real-time and accurate detection of citrus in complex scenes based on HPL-YOLOv4. *Computers and Electronics in Agriculture, 205*, 107590. https://doi.org/10.1016/j.compag.2022.107590

Yang, X., Zhao, W., Wang, Y., et al. (2024). Lightweight and efficient deep learning models for fruit detection in orchards. *Scientific Reports, 14*, 26086. https://doi.org/10.1038/s41598-024-76662-w

Zhao, L., He, C., Liu, Z., Xu, Z., Sun, R., Chen, H., Li, W., Wang, X., & Chen, Y. (2025). PhenoCitrus: An automated platform to phenotyping morphological traits of citrus fruit. *Scientia Horticulturae, 354*, 114539. https://doi.org/10.1016/j.scienta.2025.114539

Zhao, P., Zhou, W., & Na, L. (2024). High-precision object detection network for automate pear picking. *Scientific Reports, 14*, 14965. https://doi.org/10.1038/s41598-024-65750-6