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

Mobile Edge AI for detection of grape clusters and disease symptoms in vineyards

Ana Elisandra Campos da Silva[1,2,3*], Eduarda Menezes da Silveira[1,2,3], Fernando Infantini Satte Alam Nogueira[1,2,3], Sandro da Silva Camargo[1,2,3], Eduardo Farinati Leite[1,2], Joao Francisco Valiati[1,2]

[1] Center for Embedded Devices and Research in Digital Agriculture (CEDRA), São Leopoldo, Brazil

[2] SENAI Innovation Institute for Sensing Systems (ISI-SIM), São Leopoldo, Brazil;

[3] Federal University of Pampa (Unipampa), Bagé, RS, Brazil

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.

Precision viticulture requires accessible technological solutions capable of supporting rapid disease diagnosis and production monitoring directly in the field. However, dependence on cloud computing, external servers, and specialized hardware still limits the adoption of computer vision systems by small and medium-sized producers. This work presents a mobile edge artificial intelligence approach for detecting grape clusters and leaves with disease symptoms using the You Only Look Once (YOLO) object detection architecture. The proposed solution decentralizes processing by transferring inference to Android smartphones, eliminating dependence on external computational infrastructure and internet connectivity.

Model training was performed using the public dataset proposed by Székely et al. (2024), composed of proximal RGB aerial images captured by unmanned aerial vehicles (UAVs) operating at approximately 3 m above the vineyard canopy. An initial subset containing 150 images was manually re-annotated to adapt the class structure to the objectives of the mobile application and to validate the training and inference pipeline. Subsequently, a curated expanded subset containing 6,700 images was employed, including variability in illumination conditions, canopy structures, occlusions, and phenological stages.

Experimental evaluation was conducted using precision, recall, mAP@50, and mAP@50–95 metrics. The best experimental scenario achieved 64.9% precision, 62.2% recall, 66.9% mAP@50, and 42.3% mAP@50–95. Class-level analysis revealed distinct behaviors between *the target categories. Grape clusters achieved higher detection rates, while unhealthy leaves presented greater challenges due to symptom variability and missed detections, highlighting the complexity of disease identification under field conditions. The best-performing model was integrated into an Android application developed in Kotlin for fully local object detection directly from the device camera. Practical tests performed on conventional mobile devices achieved processing rates between 0.3 and 0.5 frames per second (FPS), characterizing the current implementation as a functional proof of concept for validating the complete pipeline from UAV-based training to on-device inference. The obtained results indicate that the integration of UAV-trained object detection models with embedded mobile inference represents a promising strategy for expanding access to artificial intelligence technologies in precision viticulture.*

Keywords: *Precision Viticulture, YOLO, Object Detection, Mobile Application, On-device Inference*

Introduction

Precision viticulture has increasingly incorporated computer vision and artificial intelligence techniques to support crop monitoring, disease diagnosis, and yield estimation. Among these approaches, deep learning-based object detection models have demonstrated promising performance under field conditions (Santos et al., 2020; Sozzi et al., 2022). Recent benchmarks highlight that computer vision combined with generative and edge-driven models opens new technological opportunities for automation in field environments (Min et al., 2025).

Recent advances in embedded artificial intelligence have enabled the deployment of computer vision models directly on mobile devices, reducing dependence on remote servers and cloud infrastructures (Li et al., 2026; Zhao et al., 2026). This capability is particularly relevant in agricultural scenarios, where internet connectivity may be limited and access to high-performance computing resources is often restricted. By processing data locally, edge AI systems can provide faster responses, reduce communication costs, and increase operational reliability under field conditions. The shift toward decentralized intelligence is also aligned with modern Internet of Things (IoT) ecosystems and data-driven agricultural management strategies (Chen & Yang, 2021).

Among the available object detection frameworks, the YOLO (You Only Look Once) architecture has been widely adopted due to its favorable trade-off between computational efficiency and detection accuracy. These characteristics are especially important for edge computing environments, where processing power, memory, and energy consumption are constrained. In viticulture, YOLO-based approaches have been successfully employed for grape cluster detection, disease identification, and vineyard monitoring tasks, demonstrating robust performance under real-world field conditions (Sozzi et al., 2022).

In this context, this work proposes a mobile edge AI pipeline for detecting grape clusters and leaves with disease symptoms directly on smartphones. The proposed approach integrates UAV-based model training with fully local on-device inference, enabling decentralized field operation without cloud dependency.

Materials and Methods

Mobile Edge AI Architecture

This work proposes a mobile edge AI pipeline for vineyard monitoring using object detection models deployed directly on Android devices. The complete workflow comprises dataset preparation, annotation quality assessment, model training and validation, model selection, conversion to a mobile-compatible format, and local inference execution on smartphones. Figure 1 illustrates the proposed architecture, which is organized into two complementary stages: an offline training pipeline and an on-device inference pipeline.

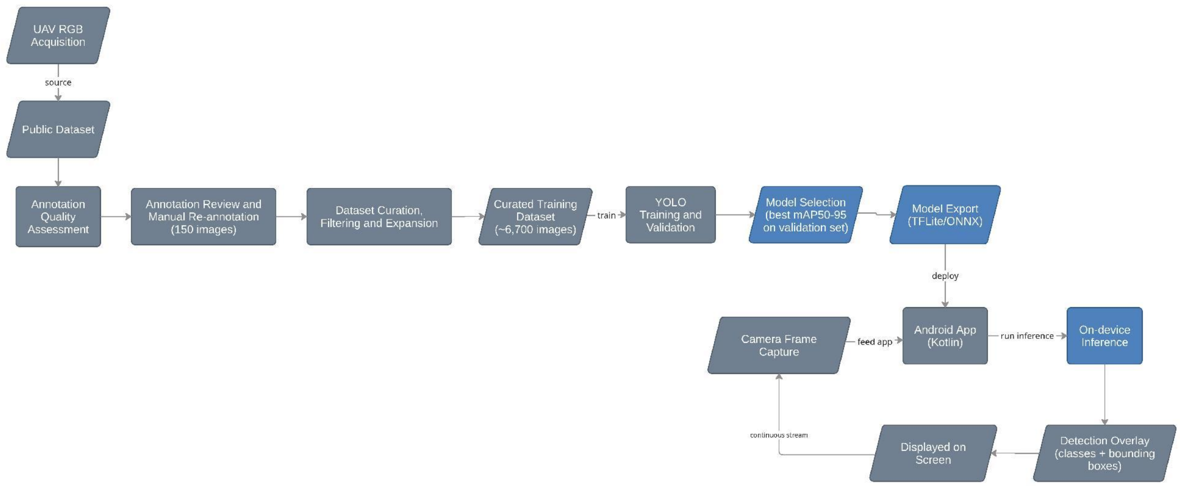

Fig. 1. Block diagram of the proposed mobile edge AI architecture, illustrating the offline training pipeline and the on-device inference loop for local Android deployment.

As shown in Fig. 1, the offline pipeline starts with UAV RGB image acquisition and the use of a public vineyard image dataset. The dataset was first submitted to an annotation quality assessment step, followed by manual review and re-annotation of an initial subset containing 150 images. After this stage, dataset curation and expansion were performed to obtain a curated training dataset containing approximately 6,700 images. This dataset was used for YOLO model training and validation, followed by model selection according to validation performance.

The selected model was then exported to a mobile-compatible format and deployed into an Android application developed in Kotlin. During operation, the application acquires frames from the smartphone camera, performs inference locally on the device processor, and displays the predicted classes and bounding boxes over the camera preview in a continuous inference loop.

Dataset Preparation and Re-annotation

The experiments were conducted using the public dataset proposed by Székely et al. (2024), composed of proximal RGB aerial images acquired by UAVs operating at approximately 3 m above the vineyard canopy. The dataset contains vineyard scenes acquired under different illumination conditions, canopy organizations, disease symptom manifestations, and phenological stages.

An initial subset containing 150 images was manually re-annotated after inconsistencies and incomplete annotations were identified in the original dataset. The re-annotation process aimed to improve annotation quality and adapt the class structure to the objectives of the mobile application, which focuses on detecting grape clusters (grape) and unhealthy leaves (unhealthy).

Following validation of the annotation strategy, a larger curated subset containing approximately 6,700 images was created. Images were selected through a manual curation process that removed samples presenting annotation inconsistencies, corrupted files, and unsuitable visual conditions. The final dataset preserved variability in illumination, canopy density, disease symptom distribution, and partial occlusions to improve model generalization.

Model Training and Evaluation

Object detection models were developed using the YOLO architecture. Training was performed using the curated dataset and evaluated through precision, recall, mAP@50, and mAP@50–95 metrics. Performance was analyzed using both macro-averaged metrics and class-specific metrics. Additionally, normalized confusion matrices were used to investigate detection behavior for grape clusters and unhealthy leaves separately.

The best-performing model was defined as the model achieving the highest mAP@50–95 on the validation dataset while maintaining balanced performance across both target classes. Lightweight YOLO variants were prioritized because the final application targets resource-constrained mobile devices and requires local execution without cloud support.

Mobile Deployment

After training, the selected model was exported to a mobile-compatible format and integrated into an Android application developed in Kotlin. The application acquires images through the smartphone camera and performs object detection locally on the device processor. All inference occurs on-device, eliminating dependence on cloud services or internet connectivity and enabling operation in remote agricultural environments. Practical tests were conducted on conventional Android devices to validate the complete training-to-deployment pipeline under real operating conditions.

Results and Discussion

The obtained results demonstrated the feasibility of integrating UAV-trained object detection models with mobile edge inference for vineyard monitoring applications. The best experimental configuration achieved the performance metrics presented in Table 1.

Metric	Result
Precision	64.9%
Recall	62.2%
mAP@50	66.9%
mAP@50–95	42.3%

Table 1. Object detection performance obtained in the best experimental scenario

The obtained detection metrics demonstrate the feasibility of applying YOLO-based object detection models for vineyard monitoring under challenging field conditions. Considering the variability present in the dataset, including illumination changes, canopy density variations, partial occlusions, and disease symptom diversity, the achieved precision, recall, and mAP values indicate that the proposed approach can successfully identify relevant vineyard targets. To better understand the model behavior, a class-level analysis was performed using the normalized confusion matrix presented in Fig. 2.

The grape class achieved a correct detection rate of approximately 74%, indicating that most grape clusters were successfully recognized by the model. Only a small fraction of grape instances were confused with the unhealthy class, while approximately 17% were missed and classified as background.

In contrast, the unhealthy class presented a lower detection rate of approximately 46%. More than half of the unhealthy instances were classified as background, indicating that disease symptoms remain the most challenging targets in the dataset. This behavior can be associated with the high variability of symptom appearance, differences in lighting conditions, partial occlusions, and the diversity of disease manifestations observed in vineyard environments. Additionally, isolated leaf regions often present lower visual quality and may be confused with grape clusters, shadows, dry leaves, naturally aged foliage, or background elements present in the vineyard scene. These factors increase the complexity of disease symptom identification when compared to grape cluster detection.

These results suggest that the proposed model is currently more effective for grape cluster detection than for disease symptom identification. Future work should focus on improving the robustness of unhealthy leaf detection through additional annotations, dataset expansion, and model optimization strategies.

The observed performance difference between the grape and unhealthy classes highlights distinct challenges associated with each detection task. While grape clusters tend to exhibit relatively consistent visual characteristics, disease symptoms may vary substantially according to severity, phenological stage, environmental conditions, and image acquisition circumstances. This observation reinforces the importance of dataset diversity and annotation quality for robust disease detection models.

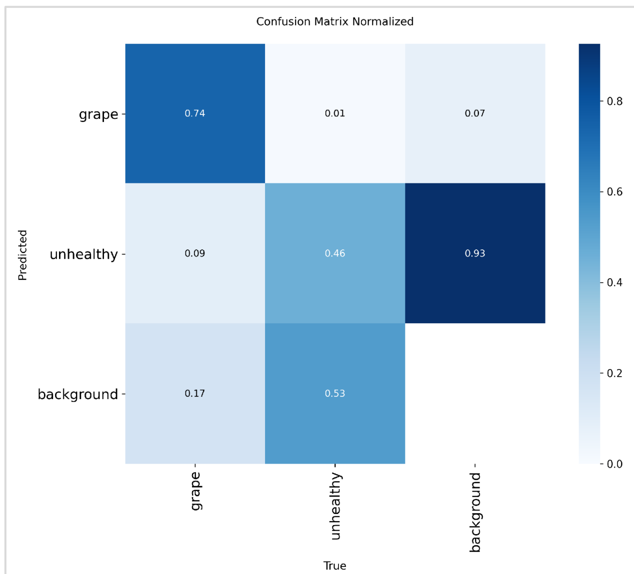

Fig. 2. Normalized confusion matrix for grape cluster and unhealthy leaf detection

To evaluate the practical feasibility of the proposed approach, the trained model was deployed on conventional smartphones and tested under real operating conditions. The objective of this stage was to validate the complete pipeline, including image acquisition, local inference execution, and results visualization directly on the mobile device.

Practical tests performed on Samsung Galaxy A36, Samsung Galaxy A56, and Motorola Edge 60 Pro achieved processing rates between 0.3 and 0.5 FPS. Although the current implementation does not yet support continuous real-time video processing, the obtained speed was considered sufficient for point-based field inspections performed by human operators, where image acquisition and analysis occur intermittently. Applications requiring continuous image acquisition or large-scale automated monitoring would require higher inference rates and additional optimization strategies. Examples of the developed application and on-device field tests are presented in Fig. 3.

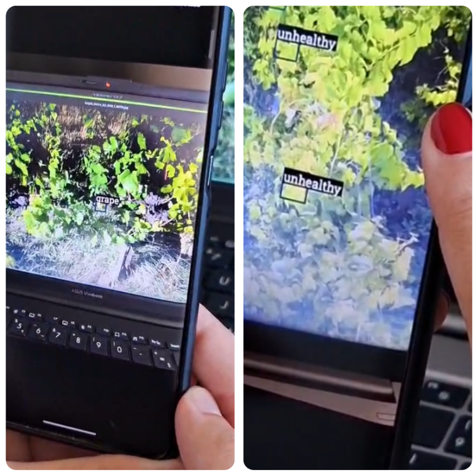

Fig. 3. Examples of on-device field tests performed using the developed mobile application.

The obtained performance may already support targeted vineyard inspections, disease verification in suspected plants, and localized grape cluster monitoring tasks where continuous video processing is not required.

The proposed approach contributes to expanding access to artificial intelligence technologies in precision viticulture, especially in low-connectivity agricultural environments.

Despite the promising results, some limitations remain. The model was trained and evaluated using images acquired under specific environmental conditions and vineyard configurations, which may limit its generalization to different grape varieties, climatic conditions, and production regions. Furthermore, variations in illumination, occlusions, and symptom appearance continue to affect disease detection performance. The current implementation also remains below real-time processing rates, requiring short pauses during field inspections. These limitations characterize the present system as a proof of concept and indicate opportunities for future improvements.

Conclusions

A mobile edge AI approach for detecting grape clusters and disease symptoms in vineyards was presented using YOLO-based object detection models integrated into a smartphone application. The obtained results demonstrated the technical feasibility of combining UAV-based model training with fully local on-device inference, enabling vineyard monitoring without dependence on permanent internet connectivity or cloud-based processing.

The proposed solution achieved satisfactory performance under challenging field conditions, reaching 64.9% precision, 62.2% recall, and 66.9% mAP@50. Class-level analysis revealed that grape cluster detection was considerably more robust than unhealthy leaf detection, indicating that disease symptom recognition remains the most challenging aspect of the proposed application.

The successful deployment of the model on conventional smartphones demonstrates the potential of embedded artificial intelligence to expand access to computer vision technologies in precision viticulture. However, the current implementation should be regarded as a proof of concept. Limitations related to inference speed, environmental variability, and disease symptom detection accuracy still require further investigation and optimization.

Future work will focus on improving unhealthy leaf detection performance, increasing inference speed, expanding validation across different vineyards and environmental conditions, and migrating the application to Flutter to improve cross-platform compatibility.

Acknowledgments

This work has been fully funded by the project FieldEye supported by the Center for Embedded Devices and Research in Digital Agriculture (CEDRA) of SENAI-RS, with financial resources from the PPI IoT/Manufatura 4.0 / PPI HardwareBR of the MCTI, grant number 053/2023, signed with EMBRAPII.

References

Chen, J., & Yang, A. (2021). Intelligent agriculture and its key technologies based on edge computing and data science. IEEE Access, 9, 153243-153255.

Codes-Alcaraz, A. M., Furnitto, N., Sottosanti, G., Failla, S., Puerto, H., Rocamora-Osorio, C., Freire-García, P., & Ramírez-Cuesta, J. M. (2025). Automatic Grape Cluster Detection Combining YOLO Model and Remote Sensing Imagery. *Remote Sensing*, *17*(2), Article 243. https://doi.org/10.3390/rs17020243

Hughes, D. P., & Salathé, M. (2015). An open access repository of images on plant health to enable the development of mobile disease diagnostics through machine learning and crowdsourcing. arXiv preprint arXiv:1511.08060.

Jeon, W., Kim, T., Choi, S., Yang, K., Kim, S.-Y., & Song, M. (2026). Bridging the Lab-to-Field gap in plant disease diagnosis through unsupervised domain adaptation enhanced by background recomposition. *Ecological Informatics*, *93*, Article 103579. https://doi.org/10.1016/j.ecoinf.2025.103579

Li, L., Liu, C. Z., Cao, K., Zhou, Z., & Alrowaily, M. (2026). Artificial intelligence of things for smart agriculture with optimal embedded edge intelligence. *Computers and Electrical Engineering*, *133*, Article 111039. https://doi.org/10.1016/j.compeleceng.2026.111039

Min, X., Ye, Y., Xiong, S., & Chen, X. (2025). Computer Vision Meets Generative Models in Agriculture: Technological Advances, Challenges and Opportunities. Applied Sciences, 15(14).

Moreira, G., dos Santos, F. N., & Cunha, M. (2026). A YOLO-based approach to grape berry detection and counting with ampelographic feature analysis for grapevine yield estimation. *Information Processing in Agriculture*. Advance online publication. https://doi.org/10.1016/j.inpa.2026.03.002

Santos, T. T., de Souza, L. L., dos Santos, A. A., & Avila, S. (2020). Grape detection, segmentation, and tracking using deep learning. Computers and Electronics in Agriculture, 170, 105247.

Sozzi, M., Cantalamessa, S., Cogato, A., Kayad, A., & Marinello, F. (2022). Automatic bunch detection in vineyards using YOLOv5. Agronomy, 12(3), 537.

Székely, D. E., Dobra, D., Dobre, A. E., Domşa, V., Drăghici, B. G., Ileni, T. A., et al. (2024). Bacterial-fungicidal vine disease detection with proximal aerial images. Journal of Plant Diseases and Protection, 131(4), 895–908.

Zhao, R., Deng, F., Que, H., Liu, M., Yue, X., & Mu, L. (2026). A Lightweight Real-Time Tomato Leaf Disease Detection System for Edge-Based Smart Agriculture. *Sensors*, *26*(11), Article 3474. https://doi.org/10.3390/s26113474