Baltic J. Modern Computing, Vol. 13 (2025), No. 2, pp. 486–501
https://doi.org/10.22364/bjmc.2025.13.2.10

Blackcurrant Leaf Analysis Using Instance Segmentation and Multi-label Classification

Reinis ODĪTIS, Ivo ODĪTIS, Kārlis FREIVALDS, Jānis BIČEVSKIS

University of Latvia, Raiņa bulvāris 19, Riga, Latvia

ro17020@students.lu.lv, ivo.oditis@lu.lv, karlis.freivalds@lu.lv, janis.bicevskis@lu.lv

ORCID: 0009-0003-5488-1608, ORCID: 0000-0003-2354-3780, ORCID: 0000-0003-2684-559X, ORCID: 0000-0001-5298-9859

Abstract. In agriculture, the health of plants can often be assessed by examining the appearance of their leaves. Traditionally, this evaluation is carried out by humans through visual observation. However, drone technology enables plant monitoring through aerial photography, removing the need for direct human presence. Furthermore, artificial intelligence offers the potential to replace human expertise in this process. In this study, the authors explore the use of machine learning methods to evaluate the condition of blackcurrants using visible light (RGB) images. The research reviews similar approaches where machine learning techniques have been applied to analyze plant leaves, aiming to identify various issues in a timely and efficient manner. Specifically, this study employs the YOLO model for leaf instance segmentation, followed by multi-label classification of the segmented leaf instances using the ResNet model. The study concludes that this method, while not perfect, provides sufficient accuracy to effectively identify field-level health issues and support targeted crop management strategies.

Keywords: Blackcurrant leaf analysis, Instance segmentation, Multi-label classification, Plant health monitoring, Drone-based agriculture

1 Introduction

In modern and precise agriculture, there is a continual search for new and cost-effective methods to diagnose plant conditions, enabling timely and accurate decisions regarding fertilization, spraying, and other interventions. One of the simplest methods for problem identification is visual plant observation. This can be done by walking through the fields and visually observing the plants; however, it is more convenient to perform this using automatically captured images. Field images can be obtained quickly and easily using many of the commercially available drones. These images provide farmers with sufficient visual information on plant condition without the need for physical field inspections. However, the process becomes more efficient when plant health information is obtained automatically.

Various comprehensive reviews (section 2) indicate that plant leaf images are widely utilized for assessing plant health through machine learning methods. While some studies explore the methodologies for image acquisition, the primary emphasis is placed on identifying potential issues and their nature using these techniques. These studies also encompass several shrub species, such as raspberries and quince. However, blackcurrants and the challenges (leaf identification in images and classification) associated with their cultivation remain unaddressed.

The publication examines the classification of blackcurrant leaves using current image analysis methods (section 3). To identify issues with blackcurrants, the process is divided into two stages: instance segmentation and classification. Images of blackcurrant bushes are segmented to isolate individual blackcurrant leaves using machine learning methods (YOLO model). The segmented images are then classified using the ResNet model to distinguish between healthy plants, diseased plants, and plants lacking sufficient nutrients.

The significance of this work is highlighted by the fact that the analysis of blackcurrant leaves and disease identification has not been addressed in previous studies. Additionally, there are no publicly available datasets of blackcurrant leaf images containing both healthy and damaged leaves with annotations. The authors could not ignore real-world challenges and worked with naturally obtained images of blackcurrant bushes, performing segmentation and classification of individual leaf instances. As a result, a solution prototype was developed that could be integrated into industrial applications.

2 Related works

Currently, the analysis of agricultural crops using photographs is being widely studied. Several comprehensive reviews of studies in this field of agriculture may be found (Debangshi (2021), Hafeez et al. (2023), Istiak et al. (2023), Anam et al. (2024)). While images from various multispectral cameras are increasingly being examined, much attention is still focused on what can be inferred about plants from visible light (RGB) images (Hafeez et al. (2023)). When evaluating plant leaves, various plant properties are assessed, such as plant health (Janani and Jebakumar (2023)), changes in plant biomass (Fei et al. (2023)), age (Bai et al. (2023)), and others.

The processing of plant leaf images typically involves two steps: isolating plant leaves from larger images and segmentation. In the following subsections, the authors review segmentation and multi-label classification methods used in other similar studies.

2.1 Instance segmentation

Instance segmentation is a crucial step in the classification process, as it facilitates the accurate identification of individual leaves within a single image. It is a computer vision technique combining object detection and semantic segmentation to identify and delineate individual object instances within an image precisely. Unlike semantic segmentation, which assigns a class label to each pixel without distinguishing separate instances, instance segmentation provides both class labels and unique boundaries for each instance of the same class. This method allows for more precise data collection for each leaf, ultimately leading to improved overall data quality for the entire image.

Following the review of studies (Gu et al. (2022)), two instance segmentation methods were examined: YOLO (You Only Look Once), particularly YOLOv8-seg, and R-CNN (Region-based Convolution Neural Networks). Both models are widely recognized for their effectiveness in segmentation tasks (Charisis and Argyropoulos (2024)). However, given the high volume of data involved in analyzing blackcurrant fields—where potentially thousands of leaves need to be segmented and classified—computation time becomes a critical factor. Given the findings from recent research (Sapkota et al. (2024)) indicating that YOLOv8 demonstrates superior inference speed and effectiveness in segmentation tasks compared to Mask R-CNN, and additional evidence (Khan et al. (2023)) from a study where YOLOv8 was tested on a maize disease detection dataset, demonstrating that it not only achieved high precision but also proved to be highly effective. This further supports that YOLOv8 is not only fast but also highly accurate, making it an excellent fit for the task.

2.2 Multi-label classification

Classification is a supervised machine learning task that involves predicting a categorical label for an input based on its features. In classification, a model is trained on labeled data, where each input is associated with a specific class. The goal is to learn patterns from the data that allow the model to assign correct labels to unseen examples.

As demonstrated in the works of other authors (Hosny et al. (2023), Elfatimi et al. (2024)), multi-label classification algorithms have proven to be highly effective for the identification of issues in plant leaves. In the context of multi-label classification, the task extends to scenarios where one object can be associated with more than one class. For instance, in the case of blackcurrant leaf health analysis, leaves can be associated with multiple diseases and may be nutrient deficient. Hosny et al. (2023) identifies several models applicable for multi-label classification of leaves, highlighting among them ResNet, VGG, and EfficientNet.

3 Proposed method

In this study, we propose a blackcurrant leaf analysis method using advanced deep learning techniques for instance segmentation and classification. The proposed approach utilizes instance segmentation to detect and isolate individual blackcurrant leaves accurately from complex backgrounds and overlapping foliage within an image, thereby ensuring precise delineation of leaf boundaries. Once segmentation is complete, each leaf instance undergoes further analysis via a multi-label classification model, which evaluates its health condition based on critical features. As discussed in section 2, YOLO models (YOLOv8n-seg, YOLOv9c-seg) are employed for instance segmentation, while ResNet architectures (ResNet-50, ResNet-101, ResNet-152) are utilized for multi-label classification (the model selection is justified in section 3.7). The proposed methodology encompasses the steps represented in figure 1.

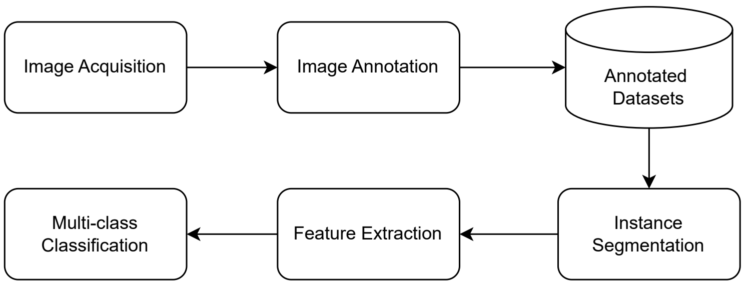

Fig. 1. Steps of leaf disease detection

3.1 Agricultural Context

Blackcurrants (*Ribes nigrum*) are valuable crops cultivated in temperate regions, thriving in well-drained soils with moderate moisture. However, they are highly susceptible to diseases such as powdery mildew, leaf spot, and rust, which manifest as discoloration, spotting, and necrosis, significantly reducing yield and quality. Early signs of these diseases are often identified through visual indicators on leaves, necessitating regular physical field monitoring. One of the most efficient ways to acquire such imagery is through drones equipped with automated missions, typically capturing images from a height of 3–5 meters to ensure sufficient detail for disease identification. The proposed method automates the identification of these diseases in images, associating leaf instances with disease classes. This information can be mapped onto aerial photos to provide a comprehensive view of field health. To be practically useful for farmers, the data should delineate the boundaries of affected areas, allowing targeted interventions. While instance-level metrics may not always achieve perfect accuracy, maintaining reliable field-level averages ensures the overall utility of the method for effective crop management.

3.2 Image acquisition

All images used in the datasets for training and validation, as described in subsection 3.4, were captured using a Nikon D3300 DSLR camera with a 24.2 MP DX-format CMOS sensor. The images were acquired from a local blackcurrant field, taken from various angles to ensure diversity in the dataset and enhance the robustness of the model training by simulating different perspectives of the leaf instances.

3.3 Image annotation

The acquired images were annotated using tailored methods to support the different stages of analysis. For instance segmentation, each image was manually annotated by outlining the boundaries of individual blackcurrant leaves using the Computer Vision Annotation Tool (CVAT). These annotations, provided in the form of point coordinates, were exported in the YOLOv8 segmentation 1.0 format. This process enables the model to accurately detect and isolate leaves from the background and overlapping objects, which is critical for precise instance segmentation.

For multi-label classification, each leaf instance was analyzed manually, and a label was assigned to describe its association with one or more of the predefined health status classes. The classification scheme, which includes the classes Healthy Leaves (HL), Nutrient Deficient Leaves (NDL), and Mycosphaerella ribis-Affected Leaves (MRL), was established in collaboration with a domain expert following an evaluation of local blackcurrant fields. The class labels were compiled in a structured file, which includes the image names and their corresponding class associations. This process ensures that both segmentation and classification tasks are fully supported by the annotations.

3.4 Datasets overview

This section provides an overview of the datasets employed in the study, which were designed to support different stages of the proposed method. Two distinct datasets were prepared: one for multi-label classification and the other for instance segmentation.

The multi-label classification dataset consists of 287 images, each containing a single blackcurrant leaf instance placed against a black background. These images are distributed across three subsets: 229 images for training, 49 for validation, and 49 for testing. The health status of each leaf instance was labeled based on the classification scheme developed in consultation with a domain expert, as outlined in the section 3.3. The dataset includes a single annotation file that lists image names and their associations with the three identified classes: Healthy Leaves (HL), Nutrient Deficient Leaves (NDL), and Mycosphaerella ribis-Affected Leaves (MRL), visually represented in figures 2, 3, and 4. A leaf instance may exhibit traits associated with multiple classes, thus requiring the model to handle overlapping features through a multi-label classification approach.

A single annotation file accompanies every subset, containing detailed information about the class associations for each image. The distribution of samples across classes in the multi-label classification dataset is presented in table 1. Because each instance can belong to multiple classes, the total number of labels in the table is greater than the total number of images.

The instance segmentation dataset consists of 87 images, each featuring blackcurrant leaves as the sole object type. It is divided into two subsets: a training set with 71 images and a validation set with 16 images. Compared to the classification dataset, the instance segmentation dataset is significantly smaller. This is because annotating this dataset is a complex and time-consuming process, requiring the precise outlining of each leaf instance. Each image was manually annotated to provide segmentation information in the form of points delineating each leaf. To analyze the potential impact and

Fig. 2. HL class example

Fig. 3. NDL class example

Fig. 4. MRL class example

Table 1. Distribution of class occurrences across training, validation, and test sets

Class	**Training (229 images)**	**Validation (49 images)**	**Test (49 images)**
Healthy Leaves (HL)	92	21	19
Nutrient Deficient Leaves (NDL)	85	19	16
Mycosphaerella ribis-Affected Leaves (MRL)	80	14	18
Total Labels	257	54	53

trends related to the dataset, the model was trained on different portions of the training set (1/4, 1/3, 1/2, 3/4, 1/1), with detailed results available in subsection 4.1. These detailed annotations enable the model to learn how to accurately identify and isolate individual leaves from complex backgrounds, ensuring reliable segmentation for further feature extraction and analysis.

3.5 Instance segmentation

For the instance segmentation task, authors selected YOLOv8n-seg and YOLOv9c-seg models, whose training was implemented using the Ultralytics package (LLC (2023)).

The ultralytics package provides an optimized pipeline for YOLO models, including configured setups for essential hyperparameters such as:

1. **Learning Rate (lr)**: Automatically initialized and dynamically adjusted in response to observed gradient behaviors.
2. **Optimizer**: The default Adam optimizer, selected to enhance gradient-based learning, was applied as configured within the Ultralytics framework for YOLO models.
3. **Batch Size**: Automatically determined according to available GPU memory, resulting in a batch size of 8 for this study.

For each model, the training was conducted in three separate configurations, each using images of a different size: 256x256 pixels for the first configuration, 512x512 pixels for the second, and 1024x1024 pixels for the third. This approach allowed the authors to assess how the models performed with varying resolutions and to understand the impact of image size on instance segmentation accuracy.

To determine the optimal number of epochs, the authors analyzed the model loss metrics by training the models with an image size of 256x256 over 100 epochs. The results are summarized in table 2.

Table 2. Initial and Final Loss Values: Box, Segmentation, Classification

Model	Loss metric	Initial value	Final value
YOLOv8n-seg	Box loss	2.3889	0.88616
YOLOv8n-seg	Segmentation loss	5.1496	1.1622
YOLOv8n-seg	Classification loss	3.4797	0.50577
YOLOv9c-seg	Box loss	2.0545	0.71893
YOLOv9c-seg	Segmentation loss	4.017	1.1462
YOLOv9c-seg	Classification loss	2.1793	0.50037

All metrics demonstrated substantial reductions within the initial 60–70 epochs, after which improvements plateaued, indicating diminishing returns. Consequently, the authors set the training limit at 70 epochs. The training loss metrics can be seen in figure 5.

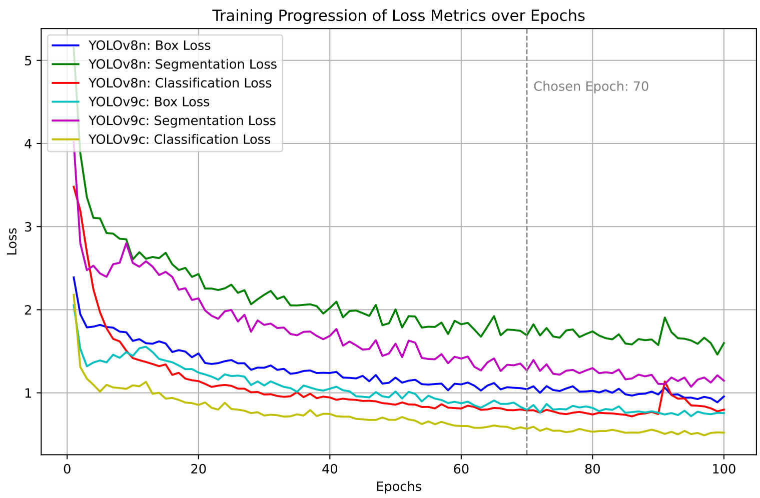

Fig. 5. Training progression of Box Loss, Segmentation Loss, and Classification Loss over 100 epochs

More detailed results, including additional insights and analysis of model performance, are covered in section 4.1.

3.6 Instance isolation

In the proposed method, instance isolation involves isolating the segmented leaf instances from the background to prepare them for classification. Following instance segmentation, the segmentation mask of each detected leaf is extracted and resized to match the original image's dimensions. Using bitwise operations, the segmentation mask is overlaid onto the original image, effectively isolating each leaf by removing unwanted background pixels. The bounding box of the mask is computed to define the region of interest, and the leaf is cropped accordingly. To standardize the input for the classification step, each cropped leaf instance is placed on a black background, ensuring a consistent visual format. These cropped instances, now devoid of any irrelevant features or noise, are then passed to the multi-label classification model to assess their health status.

3.7 Multi-label classification model selection

Based on the related researches (section 2.2), the authors selected three classification models for further investigation: ResNet, VGG and EfficientNet. The authors first conducted a comparative analysis of ResNet and VGG models. Following this initial evaluation, the results were further compared with EfficientNet, a more recent model designed to enhance the scalability of architectures like ResNet and similar convolutional neural networks. The comparison of solutions in this study is based on the analysis of existing research and performance reports, without conducting direct training of the models on the proposed dataset.

Based on the results from training on ImageNet, ResNet is a better choice than VGGNet for this particular task. Specifically, the research compared ResNet-152 and VGG-16, where ResNet-152 achieved a top-1 accuracy of 0.870 and a top-5 accuracy of 0.963, while VGG-16 showed a lower top-1 accuracy of 0.715 and a top-5 accuracy of 0.901 (Wani et al. (2020)).

In a comparison between ResNet and EfficientNet, the authors analyzed the study (Sinha and Patil (2024)) where a comparative analysis of CNN, EfficientNet, and ResNet was conducted for grape disease prediction. Both EfficientNet and ResNet demonstrated strong performance, with ResNet slightly outperforming EfficientNet. According to the study, ResNet achieved the highest accuracy of 98%, while EfficientNet closely followed with an accuracy of 97%. Both models were fine-tuned using transfer learning on a dataset containing high-resolution images of grape leaves affected by diseases such as black rot, leaf blight, and grapevine measles. Although EfficientNet is known for its efficiency in model scaling, the residual learning mechanism in ResNet provided a marginal advantage in this specific task, resulting in better overall classification performance. This study was chosen due to the nature of the task and its similarities to blackcurrant leaf analysis, making the findings highly relevant.

After comparing the models, ResNet has been selected for this study. The training will be conducted on ResNet-50, ResNet-101, and ResNet-152 to determine which architecture yields the best results for blackcurrant leaf analysis.

3.8 Multi-label classification model training

Classification is used to identify and categorize diseases in blackcurrant leaf instances based on distinguishing features. For this task, the authors trained and evaluated ResNet-50, ResNet-101, and ResNet-152.

The training process utilized PyTorch to implement a multi-label classification model based on ResNet architectures (ResNet-50, ResNet-101, and ResNet-152). The models were initialized with their default pre-trained weights provided by PyTorch to leverage feature representations learned from large-scale datasets. The final fully connected layer of each ResNet model was replaced to output predictions for three classes, making them suitable for multi-label classification. Input images were resized to 256x256 pixels and transformed into tensors for processing. The model was trained using the BCEWithLogitsLoss loss function , which combines sigmoid activation with binary cross-entropy, ensuring efficiency and numerical stability for multi-label tasks (Ansel et al. (2024)). The Adam optimizer was selected for its ability to adapt learning rates and incorporate momentum (Kingma and Ba (2017)). To assess performance, evaluation metrics such as accuracy, precision, recall, and specificity were computed.

To identify the optimal hyperparameters for model training, a grid search was conducted separately for each ResNet architecture—ResNet-50, ResNet-101, and ResNet-152—over three key hyperparameters: the number of epochs, learning rate, and batch size. The following ranges were tested: epochs in [20, 30, 50, 70, 100], learning rates in [0.01, 0.001, 0.0005, 0.0001], and batch sizes in [4, 8, 16, 32, 64]. Performance was evaluated using precision, accuracy, recall, specificity, and validation loss.

After completing the grid search, the best-performing hyperparameter configurations for each model were identified as follows: (1) ResNet-50: 70 epochs, a learning rate of 0.0001, and a batch size of 16; (2) ResNet-101: 70 epochs, a learning rate of 0.0001, and a batch size of 8; and (3) ResNet-152: 70 epochs, a learning rate of 0.0001, and a batch size of 16. These optimal configurations were then used to train and evaluate each model. Training was conducted on a dataset comprising 229 images, with validation performed on a separate set of 49 images. The final model evaluation was conducted using a test set of 49 images to assess overall performance. A detailed overview and result analysis can be found in section 4.2.

4 Results and discussion

4.1 Instance segmentation results

The instance segmentation results presented in table 3 highlight the performance of YOLOv8n-seg and YOLOv9c-seg across varying input image sizes.

The evaluation metrics include mean Average Precision (mAP) at IoU thresholds of 0.50 (mAP50) and 0.50–0.95 (mAP50–95), as well as precision for bounding box (B)

Table 3. Instance Segmentation results

Model	Image size	mAP50(B)	mAP50-95(B)	Pre(B)	mAP50(M)	mAP50-95(M)	Pre(M)
YOLOv8n-seg	256x 256	0.59164	0.38272	0.67455	0.58450	0.34303	0.74249
YOLOv8n-seg	512x 512	0.65656	0.49653	0.73628	0.65931	0.47840	0.76106
YOLOv8n-seg	1024x 1024	0.67211	0.52423	0.74736	0.66956	0.51023	0.75789
YOLOv9c-seg	256x 256	0.64149	0.46194	0.74662	0.65124	0.47340	0.72102
YOLOv9c-seg	512x 512	0.68866	0.54670	0.87470	0.69849	0.51147	0.79023
YOLOv9c-seg	**1024x 1024**	**0.72130**	**0.68643**	**0.88182**	**0.69930**	**0.53121**	**0.81230**

and mask (M) predictions. Higher mAP values indicate better performance of the model in accurately detecting and localizing objects. The formulas for these performance metrics are provided in table 4.

Table 4. Performance parameters

Indicator	Formula
mAP50	$\frac{1}{N}\sum_{i=1}^{N} \mathrm{AP}_i(IoU = 0.50)$
mAP50-95	$\frac{1}{10}\sum_{IoU=0.50}^{0.95} \frac{1}{N}\sum_{i=1}^{N} \mathrm{AP}_i(IoU)$
Accuracy (Acc)	$\left(\frac{\mathrm{TP+TN}}{\mathrm{TP+TN+FN+FP}} \times 100\right)\%$
Sensitivity (Sen)	$\left(\frac{\mathrm{TP}}{\mathrm{TP+FN}} \times 100\right)\%$
Specificity (Spe)	$\left(\frac{\mathrm{TN}}{\mathrm{FP+TN}} \times 100\right)\%$
Precision (Pre)	$\left(\frac{\mathrm{TP}}{\mathrm{TP+FP}} \times 100\right)\%$

Across all tested image sizes, YOLOv9c-seg consistently outperformed YOLOv8n-seg in both bounding box and mask segmentation metrics. At the highest resolution (1024 × 1024), YOLOv9c-seg achieved a bounding box mAP50 of 0.72130 and a mask mAP50 of 0.69930, compared to YOLOv8n-seg's corresponding scores of 0.67211 and 0.66956. Additionally, YOLOv9c-seg demonstrated better precision metrics, exceeding YOLOv8n-seg's box precision by 16.2% and mask precision by 5.4%.

Increasing the image resolution resulted in improved performance for both models, with mAP50 and mAP50–95 metrics rising consistently. For example, YOLOv9c-seg's bounding box mAP50–95 increased from 0.46194 at 256 × 256 to 0.68643 at 1024 × 1024, while YOLOv8n-seg showed a comparable increase from 0.38272 to 0.52423. This trend suggests that higher-resolution inputs provide richer feature details, enhancing segmentation performance.

In figures 6 and 7, the YOLOv9c-seg model with an image resolution of 1024 × 1024 was used to segment leaf instances. The results in the provided images align well with those presented in table 3, demonstrating strong detection and segmentation of leaves. The high precision (88.18% for bounding boxes and 81.23% for masks) ensures that most detected leaves are correctly classified, which is crucial for agricultural applications such as plant health monitoring.

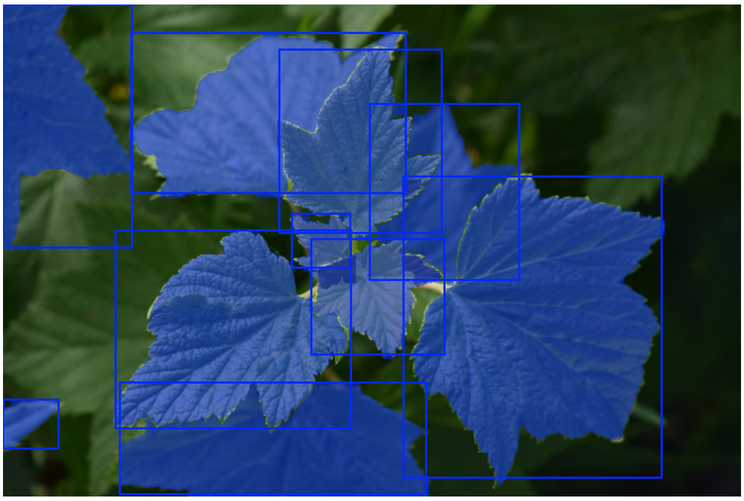

Fig. 6. Instance segmentation results using YOLOv9c-seg (1024 × 1024 resolution) on healthy leaves

However, the model encounters difficulties in recognizing damaged or diseased leaves, as seen in Figure 7, where some leaves with spots or holes show inaccuracies in both the bounding box and the segmentation mask. This limitation suggests that, in real-world scenarios, some unhealthy leaves might be missed, potentially delaying disease detection in crops. Additionally, overlapping detections in dense foliage indicate that the model may have difficulty distinguishing individual leaves in clustered environments, which could impact tasks such as automated pruning recommendations.

While the overall accuracy is promising, further improvements in fine-grained segmentation would enhance the model's ability to support precision agriculture by reliably identifying both healthy and unhealthy leaves.

Overall, the results demonstrate that YOLOv9c-seg is better suited for the given instance segmentation task, particularly at higher resolutions. Since YOLOv9c-seg achieved

Fig. 7. Instance segmentation results using YOLOv9c-seg (1024 × 1024 resolution) on leaves with visible damage

the best performance during training, the model was further trained using different portions of the training set (1/4, 1/3, 1/2, 3/4, 1/1) to analyze its robustness and adaptability. The training results are summarized in table 5 and visualized in figure 8.

Table 5. YOLOv9c-seg instance Segmentation results based on dataset fraction

Dataset fraction	mAP50(B)	mAP50-95(B)	Pre(B)	mAP50(M)	mAP50-95(M)	Pre(M)
1/4	0.49543	0.26224	0.72417	0.48971	0.23260	0.71650
1/3	0.51329	0.28320	0.74446	0.48492	0.25204	0.70009
1/2	0.53370	0.30844	0.77216	0.51555	0.26884	0.75563
3/4	0.59152	0.37815	0.78589	0.57172	0.33265	0.79359
1/1	0.72130	0.68643	0.88182	0.69930	0.53121	0.81230

The results indicate a clear trend of performance improvement as the dataset fraction increases. With only 1/4 of the training data, YOLOv9c-seg achieved a bounding box mAP50 of 0.49543 and a mask mAP50 of 0.48971, which gradually improved with larger dataset portions. Notably, at the full dataset, the model reached its highest performance, with a bounding box mAP50 of 0.72130 and a mask mAP50 of 0.69930. This demonstrates that increasing the dataset size leads to higher segmentation accu-

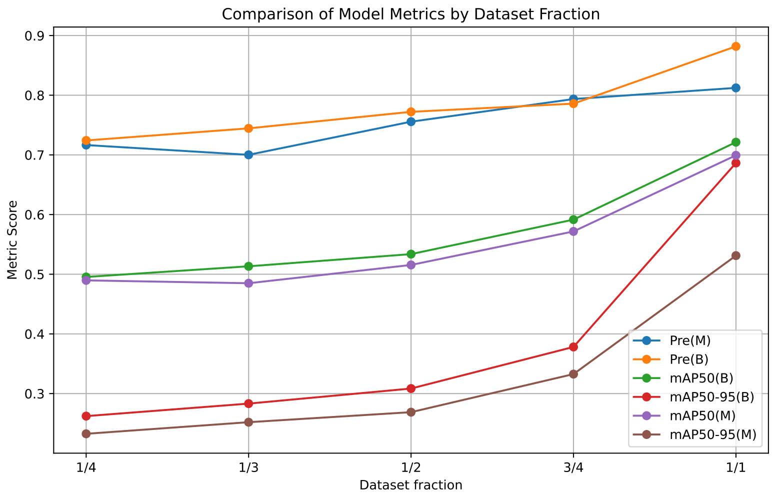

Fig. 8. Comparison of YOLOv9c-seg instance segmentation metrics by dataset fraction

racy, as no decline in performance was observed at any stage. Based on this trend, it is reasonable to assume that further expanding the dataset would yield even better results. However, since this study presents a conceptual rather than an industrial approach, and given that instance segmentation annotation is a time-consuming process, the dataset was not further extended.

4.2 Multi-label classification results

The performance of the multi-label classification models is summarized in table 6. The results were evaluated using accuracy, sensitivity, specificity, and precision, with all performance metrics defined in table 4. True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN) were used to compute these metrics, where TP and TN indicate correct predictions, while FP and FN represent misclassifications.

Table 6. Multi-label classification results

Model	Acc (%)	Sen (%)	Spe (%)	Pre (%)	Epochs	Batch size	Lr
ResNet-50	0.896	0.913	0.982	0.913	70	16	0.0001
ResNet-101	0.931	0.948	0.981	0.965	70	8	0.0001
ResNet-152	0.940	0.967	0.988	0.982	70	16	0.0001

ResNet-152 achieved the highest performance among all models, with accuracy, sensitivity, and specificity reaching 94.0%, 96.7%, and 98.8%, respectively. However, ResNet-101 also demonstrated strong results, attaining an accuracy of 93.1% and precision of 96.5%, indicating its effectiveness in distinguishing between the three classes. ResNet-50, while exhibiting slightly lower accuracy (89.6%), maintained competitive sensitivity and precision values. All models were trained for 70 epochs, with variations in batch size and learning rate having a limited impact on overall performance.

The results suggest that increasing model complexity does not yield substantial improvements, as even the smaller ResNet-50 model performed well. The similarity in metrics across all three architectures indicates that the dataset's characteristics, such as size and class distribution, may have a greater influence on performance than the choice of ResNet depth. This suggests that ResNet-50 provides a suitable trade-off between accuracy and computational efficiency, making it a practical choice for this classification task.

5 Conclusion

This study demonstrated the potential of a combined instance segmentation and multi-label classification approach for analyzing blackcurrant leaf health using RGB images. While the achieved accuracy was not the highest, it was sufficient to identify underlying issues in blackcurrant fields. YOLOv9c-seg excelled in instance segmentation, particularly at higher resolutions, enabling precise detection of individual leaves, and ResNet-50 provided a reliable balance between classification performance and computational efficiency. Instance segmentation training with different fractions of the dataset showed a tendency for the metrics to improve, with no decline even when using the full dataset, indicating that a larger dataset could further enhance overall performance. Importantly, even though the method may not yield perfect results for every individual instance, the aggregated data is effective for identifying patterns of health issues and marking their spatial distribution across the field. This capability supports targeted interventions and resource-efficient crop management. Future work should explore larger datasets and enhanced models to further improve accuracy and adaptability.

6 Acknowledgements

This work has been conducted within the research project "Competence Centre of Information and Communication Technologies" of The Recovery and Resilience Facility, contract No. 5.1.1.2.i.0/1/22/A/CFLA/008 signed between IT Competence Centre and Central Finance and Contracting Agency, Research No. 1.8 "Innovative Use of Drones in Horticulture."

References

Anam, I., Arafat, N., Hafiz, M. S., Jim, J. R., Kabir, M. M., Mridha, M. (2024). A systematic review of uav and ai integration for targeted disease detection, weed management, and pest

control in precision agriculture, *Smart Agricultural Technology* **9**, 100647.
https://doi.org/10.1016/j.atech.2024.100647

Ansel, J., Yang, E., He, H., Gimelshein, N., Jain, A., Voznesensky, M., Bao, B., Bell, P., Berard, D., Burovski, E., Chauhan, G., Chourdia, A., Constable, W., Desmaison, A., DeVito, Z., Ellison, E., Feng, W., Gong, J., Gschwind, M., Hirsh, B., Huang, S., Kalambarkar, K., Kirsch, L., Lazos, M., Lezcano, M., Liang, Y., Liang, J., Lu, Y., Luk, C., Maher, B., Pan, Y., Puhrsch, C., Reso, M., Saroufim, M., Siraichi, M. Y., Suk, H., Suo, M., Tillet, P., Wang, E., Wang, X., Wen, W., Zhang, S., Zhao, X., Zhou, K., Zou, R., Mathews, A., Chanan, G., Wu, P., Chintala, S. (2024). Pytorch 2: Faster machine learning through dynamic python bytecode transformation and graph compilation, *29th ACM International Conference on Architectural Support for Programming Languages and Operating Systems, Volume 2 (ASPLOS '24)*, ACM.
https://doi.org/10.1145/3620665.3640366

Bai, Y., Shi, L., Zha, Y., Liu, S., Nie, C., Xu, H., Yang, H., Shao, M., Yu, X., Cheng, M., Liu, Y., Lin, T., Cui, N., Wu, W., Jin, X. (2023). Estimating leaf age of maize seedlings using uav-based rgb and multispectral images, *Computers and Electronics in Agriculture* **215**, 108349.
https://doi.org/10.1016/j.compag.2023.108349

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

Debangshi, U. (2021). Drones-applications in agriculture, *Chronicle of Bioresource Management* **5**(Sep, 3), 115–120.

Elfatimi, E., Eryiğit, R., Elfatimi, L. (2024). Deep multi-scale convolutional neural networks for automated classification of multi-class leaf diseases in tomatoes, *Neural Computing and Applications* **36**(2), 803–822.
https://doi.org/10.1007/s00521-023-09062-2

Fei, S., Xiao, S., Li, Q., Shu, M., Zhai, W., Xiao, Y., Chen, Z., Yu, H., Ma, Y. (2023). Enhancing leaf area index and biomass estimation in maize with feature augmentation from unmanned aerial vehicle-based nadir and cross-circling oblique photography, *Computers and Electronics in Agriculture* **215**, 108462.
https://doi.org/10.1016/j.compag.2023.108462

Gu, W., Bai, S., Kong, L. (2022). A review on 2d instance segmentation based on deep neural networks, *Image and Vision Computing* **120**, 104401.
https://doi.org/10.1016/j.imavis.2022.104401

Hafeez, A., Husain, M. A., Singh, S., Chauhan, A., Khan, M. T., Kumar, N., Chauhan, A., Soni, S. (2023). Implementation of drone technology for farm monitoring & pesticide spraying: A review, *Information Processing in Agriculture* **10**(2), 192–203.
https://doi.org/10.1016/j.inpa.2022.02.002

Hosny, K. M., El-Hady, W. M., Samy, F. M., Vrochidou, E., Papakostas, G. A. (2023). Multi-class classification of plant leaf diseases using feature fusion of deep convolutional neural network and local binary pattern, *IEEE Access* **11**, 62307–62317.
https://doi.org/10.1109/CVPR.2016.90

Istiak, M. A., Syeed, M. M., Hossain, M. S., Uddin, M. F., Hasan, M., Khan, R. H., Azad, N. S. (2023). Adoption of unmanned aerial vehicle (uav) imagery in agricultural management: A systematic literature review, *Ecological Informatics* **78**, 102305.
https://doi.org/10.1016/j.ecoinf.2023.102305

Janani, M., Jebakumar, R. (2023). Detection and classification of groundnut leaf nutrient level extraction in rgb images, *Advances in Engineering Software* **175**, 103320.
https://doi.org/10.1016/j.advengsoft.2022.103320

Khan, F., Zafar, N., Tahir, M. N., Aqib, M., Waheed, H., Haroon, Z. (2023). A mobile-based system for maize plant leaf disease detection and classification using deep learning, *Frontiers in Plant Science* **14**.
https://doi.org/10.3389/fpls.2023.1079366

Kingma, D. P., Ba, J. (2017). Adam: A method for stochastic optimization.
https://doi.org/10.48550/arXiv.1412.6980

LLC, U. (2023). Ultralytics yolo: Real-time object detection and segmentation. Accessed: 2024-11-26.
https://ultralytics.com

Sapkota, R., Ahmed, D., Karkee, M. (2024). Comparing yolov8 and mask r-cnn for instance segmentation in complex orchard environments, *Artificial Intelligence in Agriculture* **13**, 84–99.
http://dx.doi.org/10.1016/j.aiia.2024.07.001

Sinha, S. V., Patil, B. M. (2024). Comparative analysis of cnn, efficientnet and resnet for grape disease prediction: A deep learning approach, *International Journal of Intelligent Systems and Applications in Engineering* **12**(3), 600–609.
https://ijisae.org/index.php/IJISAE/article/view/5291

Wani, M. A., Bhat, F. A., Afzal, S., Khan, A. I. (2020). *Basics of Supervised Deep Learning*, Springer Singapore, Singapore, pp. 13–29.
https://doi.org/10.1007/978-981-13-6794-6_2

Received November 28, 2024 , revised April 30, 2025, accepted June 13, 2025