An Instance Segmentation Model for Strawberry Diseases Based on YOLOv8

No Author Given

No Institute Given

Abstract. In the agriculture sector, the early and correct identification of plant diseases plays a key role in order to apply the appropriate treatments and reduce economic losses. In this way, the need arises for a rapid and low-cost identification of this type of diseases. In response of this problem, deep learning (DL) based approaches emerge as an alternative that allows accurate and fast detection without a high economic cost. With this motivation, the present study proposes a model based on the recently released YOLOv8 architecture to perform the strawberry disease instance segmentation. For this purpose, we make use of the Strawberry Disease Detection Dataset which comprises a total of 2500 images including a total of seven strawberry diseases. As a final result, we obtain a model that achieves a performance of 82.90% and 64.91% in the mAP50 and mAP50-95 metrics surpassing the results obtained by the unique work focused on instance segmentation on this particular dataset.

Keywords: Instance segmentation · Strawberry disease detection · YOLOv8 · Deep learning · Convolutional Neural Networks

1 Introduction

Nowadays, Artificial Intelligence (AI) has emerged as a promising tool in many areas, and the agricultural sector is no exception [7] [9] [8] [15]. Most of works AI focus in fruit detection [18] [2] [17] [6]. However, protecting plants against diseases is crucial for enhancing crop quality and reducing food production costs [4] [5] [4]. The traditional methods for identifying and diagnosing plant diseases and pests include inspection by a professional farmer or manual inspection in a laboratory, which is time-consuming and requires an expert level of knowledge [3]. For this reason, the need arises for a method of diagnosis of plant diseases that is efficient both in time and money.

The aforementioned challenges can be addressed by leveraging the potential of deep learning (DL) methods. DL techniques, specifically computer vision (CV) tasks are useful for solving problems related to classification, detection, segmentation, and object counting, among others taking as input digital images and videos [11]. In this way, these constitute a powerful tool that allows not only identify the type of plant disease but its exact location [3]. To meet this goal, one of the most appropriate CV tasks is image segmentation, specifically, instance segmentation. Instance segmentation involves assigning distinct labels to individual instances of objects belonging to the same class within an image [14]. This task goes beyond traditional object detection or semantic segmentation by providing pixel-level accuracy and distinguishing between multiple instances of the same object class.

In the agricultural field, strawberries have increased their popularity among countries due to different factors such as their ease of preparation and their nutrients [19]. However, these fruits are also recognized as a highly vulnerable crop [1]. For this reason, using DL-based methodology to perform the detection and segmentation of strawberry diseases results relevant. Currently, most of the works that address the identification of diseases in this fruit are focused on classification and detection. The work of Afzaal *et al.* [3], as of the date of elaboration of this article, is the only one that addresses this problem using an instance segmentation approach. Additionally, this work presents a public dataset called Strawberry Detection Dataset and the results obtained using Mask R-CNN. For this reason, this paper is the one that will be used as a reference for the present study. Different from this work, we will make use of YOLOv8 [16], a completely new architecture released in early 2023, to perform the instance segmentation task applied to strawberry diseases.

The rest of the article is organized as follows. Section 2 presents some recent works that address the identification of strawberry diseases using DL techniques. Section 3 includes the methodology to carry out this study including a description of the dataset used, the architecture, the evaluation metrics and details of the model training. Section 4 presents the results obtained and a discussion. Finally, conclusions of this work and future insights are presented in section 5.

2 Related Work

In recent years, different works have emerged focused on using DL techniques to identify strawberry diseases.

Several published works address this problem using a classification approach. In [13], Dinata *et al.* addressed the classification of six types of diseases in strawberry plants using a dataset composed of 4663 strawberry leaf disease images extracting their features. The authors reported an accuracy of 63,7%. Similarly, Xiao *et al.* [20] trained a ResNet50 model using two different datasets containing the original and feature images to classify three leaf diseases: leaf blight, gray mold, and powdery mildew. The authors obtained an accuracy of 98.06% and 99.60% for the original and feature dataset, respectively.

Although the works previously presented have obtained promising results, their datasets are not public, which means that it is not possible for other authors to use different approaches to face this task. In response, Azfaan *et al.* in [3] addressed the problem of a lack of available public datasets for the detection of plant diseases that allow for instance segmentation. In this way, the authors released the Strawberry Disease Detection Dataset, which comprises a total of

2500 images of seven different types of strawberry diseases. In addition, the authors proposed a model based on Mask R-CNN, and after applying different image augmentation techniques, they obtained a final mAP50 of 82.43% and a mAP50-95 of 59.94%.

Recently, there have been different works that make use of the Strawberry Disease Detection Dataset mainly focusing on object detection. In this manner, Cruz *et al.* in [12] proposed a strawberry disease detection model using the Strawberry Disease Detection Dataset based on YOLOv5 reporting an accuracy of 92%. Similarly, Chen et al. in [10] proposed an improved lightweight YOLOv5 model for the real-time detection of strawberry diseases using the same dataset. The authors introduced a ghost convolution (GhostConv) module into the base YOLOv5 network, which allowed reducing the parameter numbers and floating-point operations (FLOPs). They finally achieved mAP50 of 94.7% with 3.9 M parameters and 3.6 G FLOPs.

3 Methodology

3.1 Dataset

The dataset used in this study corresponds to the Strawberry Disease Detection Dataset, presented by Afzaal *et al.* at [3]. This dataset contains a total set of 2500 419×419 pixel images of seven different types of strawberry diseases, with their respective segmentation labels. The dataset is divided into 1450 images for training, 307 for validation, and 743 for testing. The classes or diseases included are *angular leafspot*, *anthracnose fruit rot*, *blossom blight*, *gray mold*, *leaf spot*, *powdery mildew fruit*, and *powdery mildew leaf*. Finally, it is important to mention that this is an unbalanced dataset, which means that not all classes have the same number of images. Figure 1 shows examples of each strawberry present in the dataset.

3.2 Model architecture

In this work, we will use the YOLOv8 architecture instead of the Mask R-CNN used in the reference paper [3]. Currently, there is no official published paper available for YOLOv8, so it is not possible to study its architecture in great detail. However, according to the official repository by Ultralytics in [16] and the Computer Vision community, YOLOv8 introduces new features than its predecessors to improve its performance, flexibility and efficiency. There are three main improvements: anchor-free detection, new convolutions and, mosaic augmentation. YOLOv8 is able to predict the center of an object in a direct way instead of using a known anchor box reducing the number of predicted boxes and therefore, the detection time. In addition, the new convolutions, which replace some old ones, let concatenate the bottleneck outputs reducing the tensor sizes without taking into account the channel dimensions. Last but not least, the training process is an important feature of its success. Therefore, YOLOv8 is

(a) Angular leafspot

(b) Anthracnose fruit rot

(c) Blossom blight

(d) Gray mold

(e) Leaf spot

(f) Powdery mildew fruit

(g) Powdery mildew leaf

Fig. 1: Classes in the Strawberry Disease Detection Dataset [3].

capable of producing some image variations on each epoch at run time. One of the techniques is mosaic augmentation which consists of the union of four different images to force the model to learn about objects in different positions with different occlusions and different background pixels. It has been shown that mosaic augmentation during whole training can hurt performance whence it is a good idea to turn off this augmentation for the last ten epochs.

Moreover, different from its previous versions, this state-of-the-art architecture supports different CV tasks such as object detection and tracking, instance segmentation, image classification and pose estimation. To fulfill the objective of this study, the models corresponding to instance segmentation will be used.

3.3 Evaluation metrics

The main evaluation metric in instance segmentation tasks is the mean average precision (mAP). The mAP represents the area under the precision-recall curve taking values that range between 0 and 1, and is calculated as the mean of the average precision (AP) metric over all classes:

$$\mathrm{mAP} = \frac{1}{N}\sum_{i=1}^{N} AP_i \tag{1}$$

where N represents the number of classes. In our case, as we have a total of 7 classes, then $N = 7$.

In addition, to obtain the mAP, it is necessary to find the Intersection Over Union (IoU) or Jaccard Index. The IoU takes values between 0 an 1 and measures the level of overlap between the ground truth masks and the mask predicted by the model. A high IoU indicates a good match between the prediction and the ground truth. The IoU can be calculated as:

$$\mathrm{IoU} = \frac{\text{Area of Intersection}}{\text{Area of Union}} \quad \text{where IoU} \in [0, 1] \tag{2}$$

Morover, the mAP can be calculated considering as correct predictions only those whose IoU is greater than a specific threshold. Similar to the reference paper, this particular study will take into account mAP at IoU $= 0.50$ and IoU $= 0.5 : 0.05 : 0.95$, or simply known as mAP, which performs the calculation of this metric by averaging the value obtained using different thresholds from 0.5 to 0.95 with a step of 0.05.

3.4 Training and Hyperparameter Tunning

The training was performed using the original YOLOv8 [16] implementation for instance segmentation. Five different YOLOv8 models are available, which differ mainly in their number of parameters. In this particular case, YOLOv8x, the largest model, was used since it led to better results after several experiments.

Also, to maximize the performance of the model, hyperparameter tunning was performed. Our YOLOv8x was trained using SGD optimizer, an image size of 640, an initial learning rate of 0.001, and a batch size of 32. Although the number of epochs was set to 200, early stopping was used to end training if there is not any improvement in the metrics after 15 epochs. This helps to avoid overfitting.

Finally, the best model is chosen based on a weighted combination of 2 metrics: mAP50 and mAP50-95. By default, in the YOLOv8 implementation used, a weight of 0.9 is given to mAP50-95, and 0.1 to mAP50, which means that the first metric has more relevance to choose the best model.

3.5 Hardware Acceleration

The experiments for this study were carried out using a workstation with 128 CPUs, 256 GB of RAM, and 4 NVIDIA A100 SXM4 of 40 GB of VRAM.

4 Results and Discussion

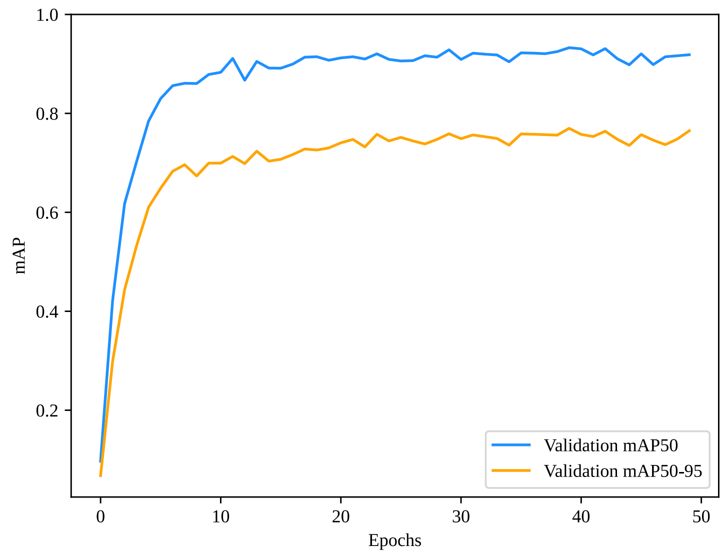
Figure 2a shows the evolution of the mAP50 and mAP50-95 on the validation set. It is possible to observe that at all the training epochs, the mAP50 is greater than the mAP50-95. This is an expected result since the mAP50-95 considers a wider range of IoU thresholds. As this threshold value increases, a higher overlap between the predictions and the ground truths is required to consider a detection as correct. An IoU threshold of 0.50 is relatively low; therefore, the model is more likely to give more predictions as correct, resulting in a higher value on this metric.

Figure 2b shows the evolution of the loss in the validation set throughout the training. Although a slight increase in loss is observed in the final epochs, the implemented early stopping ensures that overfitting does not occur. Finally, after a training of 50 epochs, a model was obtained which reached a mAP50 of 82.90% and a mAP50-95 of 64.51% on the test set. The values of this metric broken down by class can be seen in Table 1.

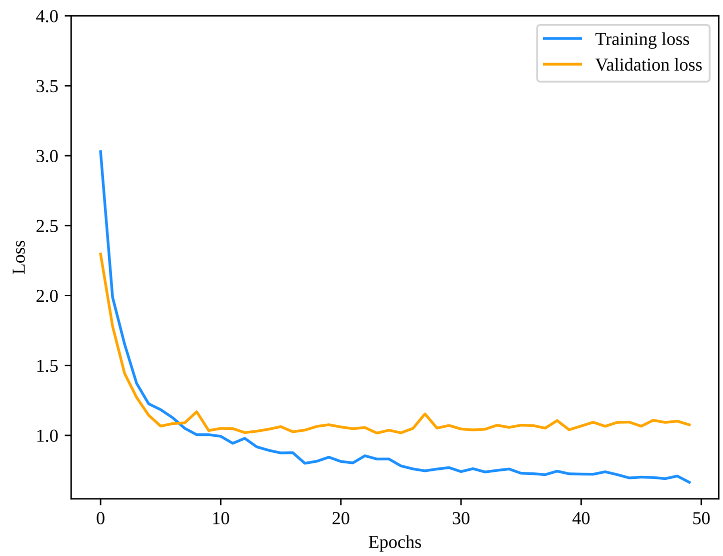
Table 1: Per-class AP on the test set.

Class	mAP50 (%)	mAP50-95 (%)
Angular Leafspot	78.20	53.26
Anthracnose Fruit Rot	79.82	40.43
Blossom Blight	83.37	70.33
Gray Mold	91.73	69.64
Leaf Spot	78.87	71.78
Powdery Mildew Fruit	76.45	60.60
Powdery Mildew Leaf	91.85	85.53

The normalized confusion matrix on the test data is shown in Fig. 3 in which the *background* class has been included. In this plot, the x-axis represents the



(a) mAP50 and mAP50-95 evolution on the validation set.



(b) mAP50 and mAP50-95 evolution on the validation set.

Fig. 2: Curves of training and validation accuracy and loss.

real class for each image, and the y-axis contains the predicted class. In general, we observe that the model tends to confuse several classes, but this is much more frequent with respect to the *background* class. In this manner, numerous instances of strawberry diseases are predicted as *background*, and vice-versa. This case is particularly noticeable in the classes *angular leafspot*, *anthracnose fruit rot*, *leafspot*, and *powdery mildew leaf*, in which we have 23%, 22%, 31%, and 25% false negatives, that is, cases in which instances of one of the diseases have been predicted as background. On the other hand, we also have high percentages of false positives in the classes *blossom blight*, *gray mold*, *leaf spot*, *powder mildew leaf* with 14%, 21%, 21%, and 24% respectively, since the model detected and predicted the background class as one of the diseases. Based on this, it is possible to affirm that the complex and varied background conditions of each class seem to generate the greatest confusion in the model. Finally, we see that the model also presented some slight confusion between the classes *angular leafspot* with *gray mold* and *leaf spot*; *anthracnose fruit rot* with *gray mold*; *leaf spot* with *gray mold*, and finally *powdery mildew fruit* with *gray mold*.

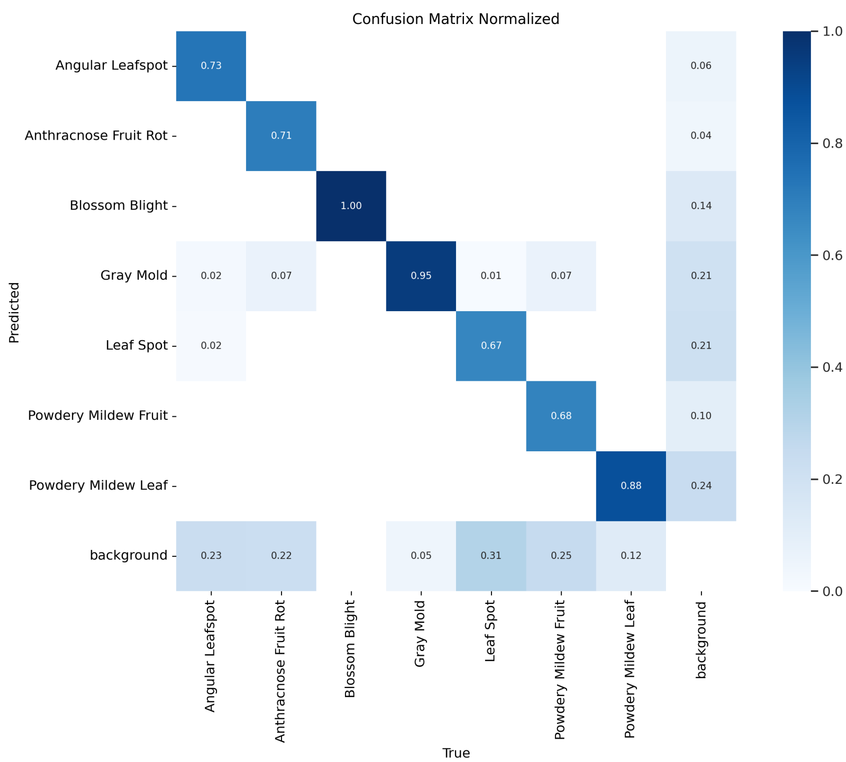


Fig. 3: Confusion matrix of the Strawberry Disease segmentation results

Compared with the reference paper, it is important to highlight that although our YOLOv8x presents similar problems when detecting various diseases as background, it better differentiates between instances of the different diseases and also significantly reduces the cases when the background is detected as a strawberry disease.

Figure 3 shows some predictions made by our model. Particularly, in the case of the example corresponding to the class *leaf spot*, it is shown that the network predicts only a small portion of the ground truth mask, taking the rest of the pixels as background, which corroborates what was observed in the confusion matrix for this particular class.

Finally, Table 2 presents a comparison of the mAP50 and mAP50-95 results obtained by our model and other different architectures on the Strawberry Disease Detection Dataset. It is important to mention that the results for both Mask R-CNN and YOLOACT are also reported in the work of Afzaal *et al.* [3]. We see that our model outperforms the results of Mask R-CNN with ResNet101 backbone by 0.57% and 7.63% in the mAP50 and mAP50-95 respectively.

Table 2: Comparison with other architectures

Network	mAP50 (%)	mAP50-95 (%)
Mask R-CNN (ResNet50) [3]	81.37	55.21
Mask R-CNN (ResNet101) [3]	82.43	59.94
YOLOACT (ResNet50) [3]	79.71	55.19
YOLOACT (ResNet50) [3]	79.39	55.81
YOLOv8	**82.90**	**64.51**

5 Conclusion and Future Work

In this article, we propose a deep learning-based model to detect and segment a total of seven strawberry diseases present in the Strawberry Disease Detection Dataset. Specifically, the recently revealed YOLOv8 was used, obtaining as a result a model that reaches a mAP50 and mAP50-95 of 82.90% and 64.91%, respectively. These results improve those obtained by architectures such as Mask-RCNN and YOLOACT. It's undeniable that the improvements introduced in the new YOLOv8 architecture have shown to be capable to enhance the training process and results with respect to other instance segmentation state-of-the-art models in this particular dataset. However, as more accurate models are presented, the demand for more powerful hardware infrastructure arises.

As future work, it would be highly relevant to develop work related to the acquisition of larger datasets with the purpose to improve the performance of DL models. If this is not possible, the use of different image augmentation techniques can be performed in order to find those that help to increase the performance of the architecture for this particular task. Additionally, it could be of great

(a) Original (b) Ground Truth (c) Predicted

Fig. 4: Predictions on the test set by the YOLOv8 network.

interest to deploy the model in a real-time system with an embedded system in order to test its performance in a real work environment.

References

1. Abbas, I., Liu, J., Amin, M., Tariq, A., Tunio, M.H.: Strawberry fungal leaf scorch disease identification in real-time strawberry field using deep learning architectures. Plants **10**(12), 2643 (2021)
2. Afonso, M., Fonteijn, H., Fiorentin, F.S., Lensink, D., Mooij, M., Faber, N., Polder, G., Wehrens, R.: Tomato fruit detection and counting in greenhouses using deep learning. Frontiers in plant science **11**, 1759 (2020)
3. Afzaal, U., Bhattarai, B., Pandeya, Y.R., Lee, J.: An instance segmentation model for strawberry diseases based on mask r-cnn. Sensors **21**(19), 6565 (2021)
4. Arsenovic, M., Karanovic, M., Sladojevic, S., Anderla, A., Stefanovic, D.: Solving current limitations of deep learning based approaches for plant disease detection. Symmetry **11**(7), 939 (2019)
5. Barbedo, J.G.: Factors influencing the use of deep learning for plant disease recognition. Biosystems engineering **172**, 84–91 (2018)
6. Bargoti, S., Underwood, J.: Deep fruit detection in orchards. In: 2017 IEEE International Conference on Robotics and Automation (ICRA). pp. 3626–3633. IEEE (2017)
7. Bauer, A., Bostrom, A.G., Ball, J., Applegate, C., Cheng, T., Laycock, S., Rojas, S.M., Kirwan, J., Zhou, J.: Combining computer vision and deep learning to enable ultra-scale aerial phenotyping and precision agriculture: A case study of lettuce production. Horticulture research **6**(1), 1–12 (2019)
8. Bosilj, P., Aptoula, E., Duckett, T., Cielniak, G.: Transfer learning between crop types for semantic segmentation of crops versus weeds in precision agriculture. Journal of Field Robotics **37**(1), 7–19 (2020)
9. Bosilj, P., Duckett, T., Cielniak, G.: Connected attribute morphology for unified vegetation segmentation and classification in precision agriculture. Computers in industry **98**, 226–240 (2018)
10. Chen, S., Liao, Y., Lin, F., Huang, B.: An improved lightweight yolov5 algorithm for detecting strawberry diseases. IEEE Access (2023)
11. Crespo, F., Crespo, A., Sierra-Martínez, L.M., Peluffo-Ordóñez, D.H., Morocho-Cayamcela, M.E.: A computer vision model to identify the incorrect use of face masks for covid-19 awareness. Applied Sciences **12**(14), 6924 (2022)
12. Cruz, M., Mafra, S., Teixeira, E., Figueiredo, F.: Smart strawberry farming using edge computing and iot. Sensors **22**(15), 5866 (2022)
13. Dinata, M.I., Mardi Susiki Nugroho, S., Rachmadi, R.F.: Classification of strawberry plant diseases with leaf image using cnn. In: 2021 International Conference on Artificial Intelligence and Computer Science Technology (ICAICST). pp. 68–72 (2021). `https://doi.org/10.1109/ICAICST53116.2021.9497830`
14. Hafiz, A.M., Bhat, G.M.: A survey on instance segmentation: state of the art. International journal of multimedia information retrieval **9**(3), 171–189 (2020)
15. Jiayu, Z., Shiwei, X., Zhemin, L., Wei, C., Dongjie, W.: Application of intelligence information fusion technology in agriculture monitoring and early-warning research. In: 2015 International Conference on Control, Automation and Robotics. pp. 114–117 (2015). `https://doi.org/10.1109/ICCAR.2015.7166013`

16. Jocher, G., Chaurasia, A., Qiu, J.: YOLO by Ultralytics (Jan 2023), `https://github.com/ultralytics/ultralytics`
17. Koirala, A., Walsh, K., Wang, Z., McCarthy, C.: Deep learning for real-time fruit detection and orchard fruit load estimation: Benchmarking of 'mangoyolo'. Precision Agriculture **20**(6), 1107–1135 (2019)
18. Sa, I., Ge, Z., Dayoub, F., Upcroft, B., Perez, T., McCool, C.: Deepfruits: A fruit detection system using deep neural networks. sensors **16**(8), 1222 (2016)
19. Simpson, D.: The economic importance of strawberry crops. The genomes of rosaceous berries and their wild relatives pp. 1–7 (2018)
20. Xiao, J.R., Chung, P.C., Wu, H.Y., Phan, Q.H., Yeh, J.L.A., Hou, M.T.K.: Detection of strawberry diseases using a convolutional neural network. Plants **10**(1), 31 (2020)