# ENHANCING REAL-TIME INSTANCE SEGMENTATION FOR PLANT DISEASE DETECTION WITH IMPROVED YOLOV8-SEG ALGORITHM

*Mohamed Ben Ammar**

Department of Information Systems, Faculty of Computing and Information Technology, Northern Border University, Rafha
Saudi Arabia

* Corresponding Author, e-mail: Mohammed.Ammar@nbu.edu.sa

**Abstract:** With widespread uses in areas as diverse as traffic analysis and medical imaging, picture segmentation is a basic problem in computer vision. Instance segmentation, which combines object recognition with segmentation, is a powerful tool for item identification and exact delineation. Using the Tomato Leaf disease dataset as an example, this research delves into the topic of segmentation training by capitalizing on the simplicity of enhanced YOLOv8-Seg models. Tomato leaf disease are the focus of this instance-segmentation dataset, which seeks to resolve the pressing problem of agricultural difficulties. One instance segmentation networks, YOLOv8n-Seg is presented and compared in this article for the purpose of Tomato leaf disease identification. The models are tested in difficult situations to see how well they can detect and separate garbage occurrences. Results show that enhanced YOLOv8-Seg is useful for agriculture by accurately segmenting instances of tomato leaf disease detection.

**Key words:** Tomato leaf disease detection, Instance Segmentation, YOLOv8-Seg, Object Detection.

## 1. INTRODUCTION

Precise delineation of tomatoes at different development stages and with ailments is crucial for activities such as automated harvesting, monitoring fruit health, and accurate assessment of size and quality [1]. The expeditious and accurate segmentation described herein confers benefits in implementations such as greenhouse management systems and machine vision [2]. Tomato instance segmentation presents several hurdles. Environmental factors like fluctuating lighting conditions, overlapping fruits, leaves blocking the tomatoes, and variations in viewing angles can significantly hinder the process [3]. Additionally, changes in the color and texture of the tomato surface itself can negatively impact the accuracy of the segmentation results.

Fruit recognition and segmentation have been a major area of research focus. Traditional methods rely on analysing specific features of the fruit, such as color, shape, and surface texture. For example, one approach involved distinguishing pears from the background based on colour variations [4]. In addition to edge, colour, magnitude, and orientation properties, multi-feature fusion strategies may also incorporate shape and colour attributes. In their study, Yin et al. [5] presented a technique for detecting mature tomatoes. This approach included reducing noise for fruits that were obstructed or overlapped, and then combining form and colour characteristics. Nevertheless, both the analysis of individual features and the combination of several features are hindered by constraints in terms of resilience and temporal effectiveness. Fluctuations in the colour and texture of fruit surfaces, especially under unregulated conditions, may greatly affect the accuracy of detecting and separating them.

It is essential in complex greenhouse environments to be able to accurately and quickly identify and separate individual tomato instances, especially those that have been damaged by illnesses. The quick selection of ripe fruits, the reduction of waste, and the simplification of the monitoring of tainted tomatoes in order to prevent the spread of illnesses are all made possible as a result of this. Object identification and instance segmentation are two areas in which deep learning has shown to be an effective technique due to its outstanding performance. A specialized branch of the Mask R-CNN architecture, which is commonly used in deep learning, is specifically designed to generate binary masks for the purpose of object detection and segmentation. This aspect of the architecture is very useful. In order to enable real-time object identification, this branch makes use of region proposal networks, which is a continuation of the work that was done by Fast R-CNN [6, 7].

This study presents an improved YOLOv8s-Seg technique that properly segments both healthy tomatoes and diseased tomatoes. The approach is based on considerable research conducted by several researchers. The following initiatives were part of the research. A dataset including information on leaf diseases is used to test the detection capability of YOLOv8 deep learning models. The subsequent parts of the paper are organized in the following manner: Section 2 present related work. Section 3 offers a detailed and complete account of the techniques and materials used, including full explanations and precise information. The assessment of the detection results of the proposed model is outlined in Section 4. The outcome is eventually given in Section 5.

## 2.RELATED WORK

Multiple experiments have shown the efficacy of Mask R-CNN in fruit segmentation. A dataset of 120 apple photos with substantial overlap was used by Jia et al. [8] to obtain 97.31% accuracy using an improved Mask R-CNN variation, demonstrating the technique's adaptability to difficult situations. Author in reference [10] introduced a fuzzy masking using R-CNN method to identify tomato ripeness. They achieved a 98.00% accuracy on a dataset of 100 pictures. For robust segmentation, they combined edge feature extraction with fuzzy logic in their technique. To distinguish between ripe and immature tomatoes, author in [11] used a typical Mask R-CNN with ResNet101 as the backbone. They achieved remarkable accuracy rates of 95% and 94%, respectively. These results demonstrate the capability of Mask R-CNN for a wide range of fruit segmentation tasks.

Reference [12] introduced an enhanced Mask RCNN models that incorporates the attention process to accurately segment fruits ripeness in many scenarios, including variations in lighting, occlusion, and overlapping. The test findings indicated an accuracy rate of 95.8% and a recall rate of 97.1%. Further examples of the high precision and robustness of the mask based RCNN algorithm in detecting objects and instance segmentation include tomato fruit segmentation [13], tomato fruit infection area detection [14], tomato maturity segmentation [15], and soil block segmentation [16]. Mask RCNN is a traditional instance segmentation model that follows a two-stage approach. Masks are produced using Mask RCNN using feature localization. After conducting pooling procedures on the area of interest, the identified features are then sent to the mask predictor. Nevertheless, carrying out these actions in a sequential manner might result in sluggish segmentation speed, a substantial model size, and a higher count of computational parameters.

## 3.MATERIALS AND METHODS

In this paper, object detection performances of YOLOv8 is tested on a dataset containing different types of leaf diseases. Figure 1 illustrates each step of detection process. Firstly, dataset was obtained from Roboflow platform and divided into training and validation parts for the training. Afterwards, training was performed with each detection model to get specific results in terms of some metrics discussed in Section IV. In this section, a brief information about YOLO8 will be given. Following this, dataset fed into the training phase, environment and hyperparameters of each model will be introduced.

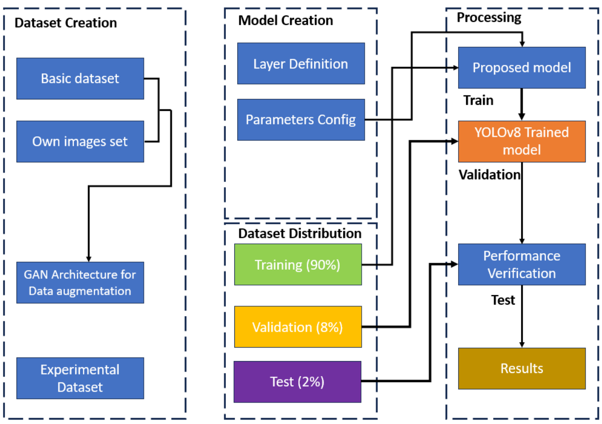


*Figure 1. Tomatoes disease detection proposed architecture.*

### 3.1. Dataset

The Tomato Leaf Diseases Dataset, which included 10,269 photos of tomato leaves afflicted by nine different disease classes, was utilized in this investigation. The classes present in this context include Early Blight, , Late Blight,Healthy, Leaf Miner, Leaf Mold, Mosaic Virus, Septoria, Spider Mites, and Yellow Leaf Curl Virus. The dataset is available to the public via Roboflow. The dataset is separated into three categories: 9236 photos for training (90%), 867 images for validation (8%), and 166 images for testing (2%). The dataset can be viewed at the following link: https://universe.roboflow.com/bryan-b56jm/tomato-leaf-disease-ssoha/dataset/63. The dataset was last accessed on 25 March 2024 [18].

A graph showing the total amount of annotations for all dataset classes can be seen in Figure 2 visually displays the distribution of bounding boxes, which provides information on the position and dimensions of the boxes. This visual representation improves understanding of the spatial configuration of bounding boxes, ensuring a diverse choice of item positions and dimensions for the enhanced model. By examining this graph, one may ascertain if the distribution of the bounding boxes is homogeneous or whether some sections of the dataset have higher significance. This study is crucial to ensure that the neural network can effectively identify objects in images, considering the variations in size and placement among all objects.

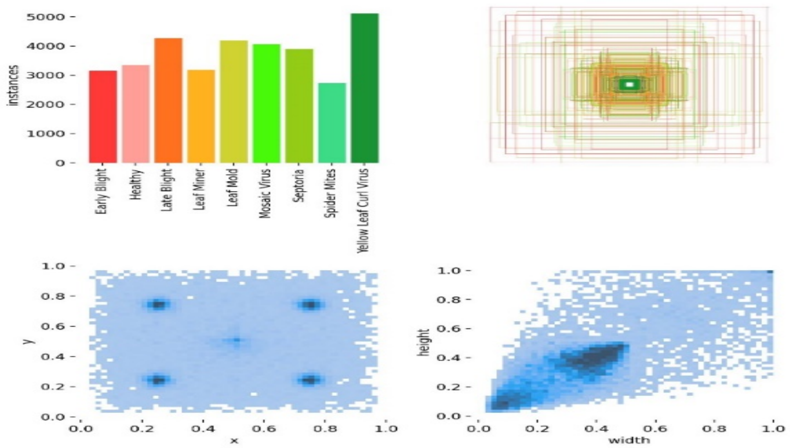


*Figure 2. Data visualization: Dataset's annotation distribution per class sizes.*

### 3.2. Improved tomato leaf disease detection based on YOLOv8-Seg

The YOLO model algorithm class is a specialized deep-learning models designed exclusively for the task of object detection. YOLOv8, produced by the same authors as YOLOv5, maintains a comparable overall structure while providing substantial improvements and optimizations. YOLOv8 demonstrates exceptional algorithmic performance, setting it apart from its predecessor.

With the introduction of the widely accepted decoupled head structure, YOLOv8's head section differs greatly from YOLOv5. The Task Aligned Assigner technique is made available in YOLOv8 for the purpose of allocating positive samples during the calculation of the loss function. Additionally, the distribution of focused loss is included. Throughout the training stage, YOLOv8 employs an approach that is provided in YOLOX, which is a strategy that involves deactivating mosaics augmentation in the ten most recent epochs. During the process of data augmentation, implementation of this adjustment has shown its effectiveness in enhancing accuracy. The YOLOv8s-Seg models are different versions of the YOLOv8 object recognition model that were made just for segmentation tasks. Real-time instance segmentation of objects is something that these models are able to do while still retaining a high segment mean average accuracy. These models are inspired by the YOLACT network. The topology of the YOLACT network is shown in a clear and straightforward manner in Figure 3.

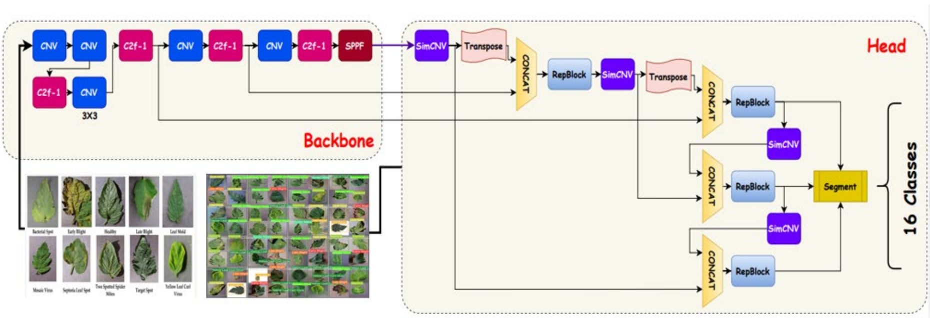


*Figure 3. Tomato disease leaf detection based on YOLOV8s model*

The head of YOLOv8-Seg is formed by the combination of several segments and certain parts of the neck. The Neck node is responsible for integrating the PANet and FPN feature fusion networks . YOLOv8-Seg, in contrast to YOLOv5 and YOLOv6, does away with the 1 × 1 convolution that occurs prior to upsampling as presented in Figure 4. Instead, it integrates feature maps from various layers of the backbone networks directly. The neck modules of YOLOv8-Seg are going to be optimized as part of this research in order to increase the overall performance of the network. In order to achieve a higher level of precision, we substituted the traditional convolutions encountered in the neck region with three-by-three Sim-Convolutions. Additionally, we included two one-by-one Sim-Convolutions prior to each upsampling step. Additionally, we took the C2f module and replaced it.

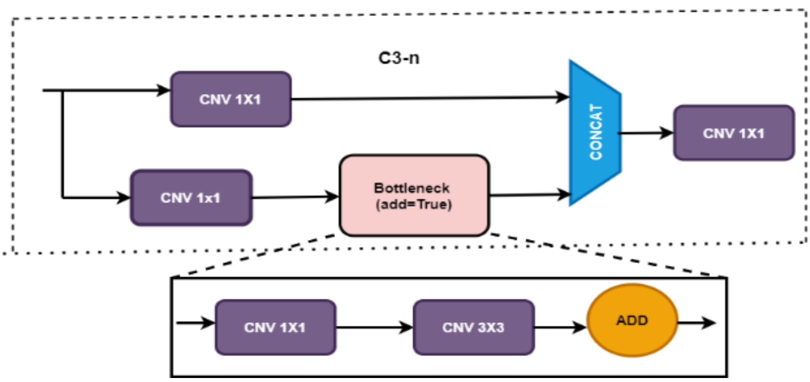


*Figure 4. C3 Architecture*

Enhanced YOLOv8s-Seg architecture is utilized by the tomato segmentation networks, as illustrated in Figure 5. This upgrade is comprised of three significant modifications: the C2f module is replaced with a RepBlock module; SimConv convolutions are included before to the upsampling phases of the neck module; and all remaining conventional convolutions are replaced with SimConv. Because of these modifications, the network is now better able to properly combine individual feature information.

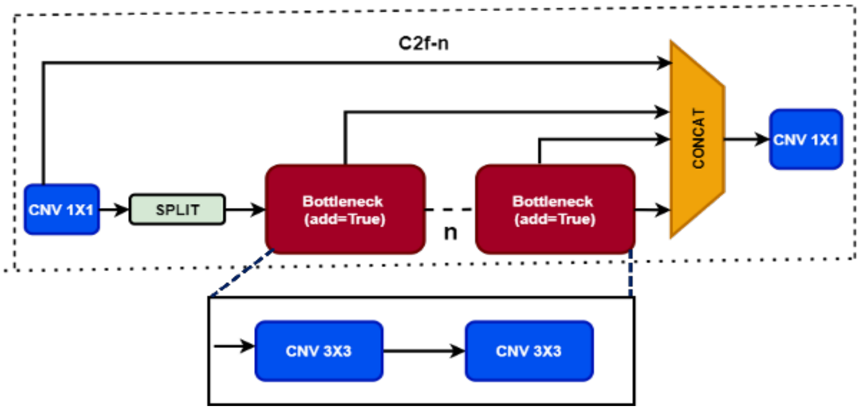


*Figure 5. C2f Architecture in the neck.*

### 3.3. Evaluation metrics

This study evaluates the performances of the proposal enhanced YOLOv8s-Seg model using four metrics: accuracy, recall, F1 score, and mAP@0.5. Accuracy, recall, and F1 score are employed to assess the effectiveness of tomato leaf disease detection, while mAP specifically measures the quality of the segmentation results. To determine accuracy, recall, F1 score, and mAP scores, the equations (1), (2), (3), and (4) are used in the calculation process. When these four factors have higher values, it indicates that the segmentation results are satisfactory.

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

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

$$F1 = \frac{2*Precision*Recall}{Precision+Recall} = \frac{2*TP}{2*TP+FP+FN} \tag{3}$$

$$mAP = \sum_{n=1}^{K} \frac{AP_i}{K} \tag{4}$$

Notes: TP denotes a positively predicted sample, FP denotes a negatively forecasted sample with an inaccurate positive prediction, and FN denotes a positively predicted sample with a missed positive prediction. AP is an abbreviation that stands for the average accuracy of categorization, and the effectiveness of the segmentation model improves as the AP score for the model grows. Here, K is the number of categories used for segmentation.

## 4.EXPERIMENTS AND RESULTS

### 4.1. Experimental Environment

This paper's models were trained using the Google Colab platform, which was employed. A graphics processing unit (GPU) that was equipped with CUDA capabilities and the use of the PyTorch environment were both essential components for the experimental setup. Through the use of parallel computing, the YOLOv8-Seg deep learning model was able to accomplish learning processes and object identification that were both accurate and efficient. PyTorch was used as the framework for the construction and training of the models, and the usage of a graphics processing unit (GPU) that supported CUDA was done in order to speed up the calculation. For the purpose of this experiment, the upgraded YOLOv8s-Seg model is run in an environment setup that is similar to the previous one, and the hyperparameter values are maintained throughout.The YOLOv8s-Seg model that has been enhanced went through training with a batch size of sixteen over the course of fifty epochs.

### 4.2. Results Analysis and Discussion

Training the suggested networks for a total of fifty epochs was a part of this article. Additionally, the training outcomes of the improved models were obtained by applying them to the Tomato leaf disease dataset. The numerous performance metrics for both the initial training set and the set used for validation are shown in the following table 1, which presents the data. A sustained reduction in both loss of training and validation loss indicates that the model's learning process is continuously improving. At the same time, there is a tendency toward an increase in accuracy, recall, mAP@0.5, and mAP@0.5:0.95, which indicates that the network's detection ability has been steadily improving.

*Table 1: Tomatoes leaf disease detection performance metrics*

| **Class** | **Images** | **Instances** | **Box(P)** | **R** | **mAP50** | **mAP50-95** |
|---|---|---|---|---|---|---|
| **all** | 843 | 2680 | 0.954 | 0.934 | 0.975 | 0.932 |
| **Early Blight** | 843 | 252 | 0.964 | 0.96 | 0.984 | 0.954 |
| **Healthy** | 843 | 277 | 0.902 | 0.903 | 0.951 | 0.886 |
| **Late Blight** | 843 | 344 | 0.976 | 0.953 | 0.987 | 0.952 |
| **Leaf Miner** | 843 | 259 | 0.991 | 1 | 0.995 | 0.98 |
| **Leaf Mold** | 843 | 338 | 0.953 | 0.903 | 0.97 | 0.928 |
| **Mosaic Virus** | 843 | 319 | 0.983 | 0.934 | 0.979 | 0.959 |
| **Septoria** | 843 | 324 | 0.973 | 0.901 | 0.981 | 0.943 |
| **Spider Mites** | 843 | 252 | 0.959 | 0.976 | 0.986 | 0.959 |
| **Yellow Leaf Curl Virus** | 843 | 315 | 0.885 | 0.873 | 0.938 | 0.827 |

The Improved YOLOv8s-Seg model concludes with a box mAP of 98.4% and a segmentation mAP of 93.8% in the final epoch. Notably, the Nano model outperforms slightly within the same epoch range. In particular, during the last epoch, the Nano model achieved a box mAP of 97.5% and a segmentation mask mAP of 95%. This comparison highlights the Nano model's gradual performance improvements, hinting at its potential for enhanced accuracy in object localization and segmentation, particularly in the context of Tomato leaf disease Instance Segmentation.

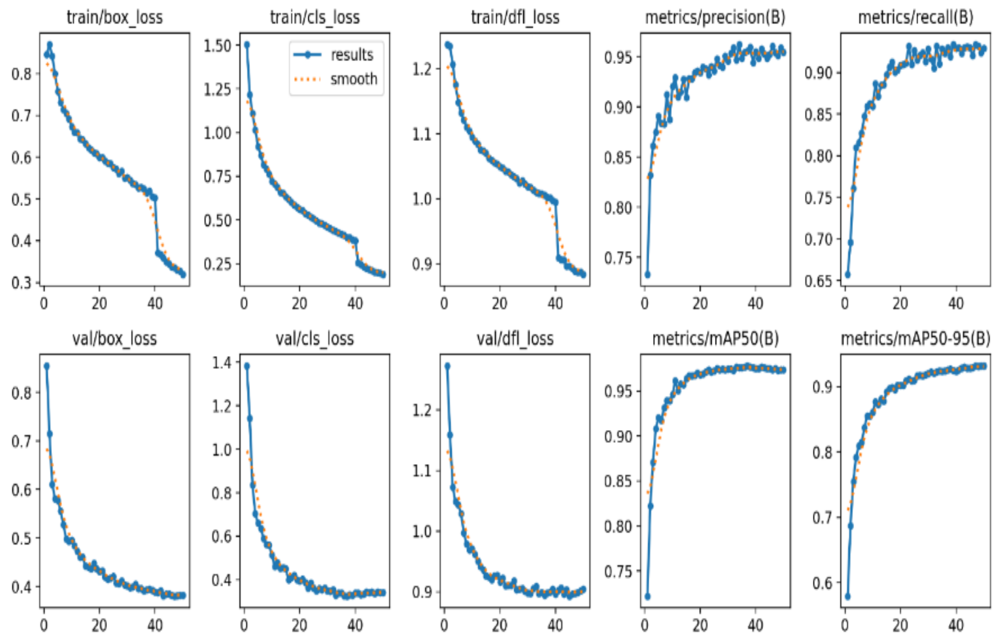


*Figure 6. Training Performance for improved YOLOv8s-Seg and YOLOv8n-Seg* Networks

The PR curves for Tomato leaf disease Instance Segmentation, utilizing the proposed improved YOLOv8s-Seg model, are depicted in Figure. 6. It can be observed that there is a noticeable acceleration in accuracy improvement with increasing recall. Figure 7 reinforces this observation, depicting Precision-Recall (PR) curves for our proposed model clustered in the upper right corner, signalling elevated levels of both recall and accuracy. The substantial area beneath these curves signifies robust model performance. Furthermore, the smoothness of the PR curves implies a consistent relationship between recall rate and accuracy, affirming the stability of our proposed model's performance.

The F1 curve, encapsulating the harmonic balance between precision and recall, provides a comprehensive evaluation of the model's performance in both Bounding Box and Segmentation Mask tasks, as shown in Figure. 7.

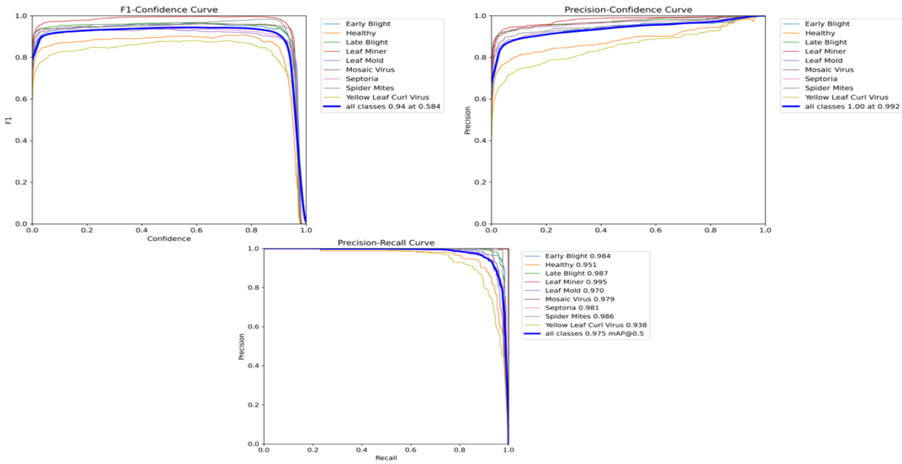


*Figure 7. PR curve comparison of Bounding box and Segmentation mask*

For the Bounding Box F1 curve, the visualization illustrates how the F1 score evolves across diverse confidence thresholds for positive class predictions. A higher F1 score indicates a superior equilibrium between precision and recall, signifying a more resilient model for object detection. Similarly, the Segmentation Mask F1 curve evaluates the F1 score under varying confidence thresholds, emphasizing the balance between racall and precision in the objects segmentation context. A higher F1 score in this curve reflects an enhanced equilibrium between precision and recall, demonstrating improved accuracy in outlining object boundaries.

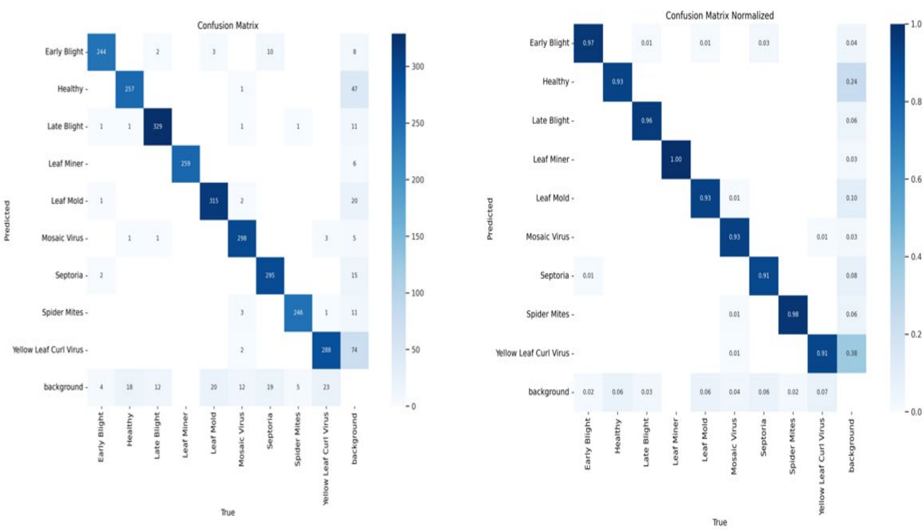


*Figure 8. Confusion Matrix of Tomato leaf disease instance segmentation based on Improved YOLOv8-Seg model.*

For tomato leaf disease instance segmentation, a confusion matrix (Figure 9) based on the enhanced YOLOv8-Seg model would usually show the model's performance as predicted by true negatives (TN), true positives (TP), false negatives (FN), and false positives (FP). Every each cell inside the matrix represents a specific pairing of anticipated and real categories. The suggested model's confusion matrix is shown in Figure. 8, which provides information on the enhanced YOLOv8-Seg model's prediction accuracy for Tomato leaf disease instance segmentation. The matrix offers a thorough depiction of the connections between forecasts.Detailing the model’s performance in terms of true positive, true negative, false positive, and false negative instances. This visual aid is instrumental in evaluating the effectiveness of the model in accurately detecting and segmenting tomato leaf instances within the dataset. The observed results indicate that the improved YOLOv8n-Seg achieved a notably higher accuracy rate for each category when compared to the improved YOLOv8s-Seg.

After the training phase is over, the model is evaluated using a separate testing dataset that contains photos that were not used during the training process. During this assessment, the model's objective is to identify and separate instances of garbage that are located tomato leaf disease. Each recognized garbage object's bounding box is used to graphically represent the findings.

## 5. CONCLUSION

This research proposes an enhanced YOLOv8s-Seg network for tomato disease identification and maturity assessment using instance segmentation. To improve feature fusion, the C2f module was updated using a RepBlock module. SimConv convolutions were incorporated before the upsampling stages in the feature fusion network, and all remaining standard convolutions were substituted with SimConv. The improved YOLOv8s-Seg achieved a segment mean average precision (mAP) of 97.5% at an intersection over union (IoU) threshold of 0.5 on the validation set. This represents a significant performance improvement compared to other segmentation models, surpassing YOLOv7-Seg by 2.8%, YOLOv5s-Seg by 4.2%, the original YOLOv8s-Seg by 3.4%, and Mask RCNN by 0.8%. The optimized YOLOv8s-Seg also achieved faster inference times compared to YOLOv8s-Seg and YOLOv5s-Seg (3.5 ms vs. 3.9 ms and 4.1 ms, respectively).

## ACKNOWLEDGEMENT

The authors extend their appreciation to the Deanship of Scientific Research at Northern Border University, Arar, KSA for funding this research work through the project number “NBU-FFR-2024-2439-*02*”.

## REFERENCES

[1].Lee, J., Nazki, H., Baek, J., Hong, Y., Lee, M. Artificial intelligence approach for tomato detection and mass estimation in precision agriculture. *Sustainability*, vol.12, no.21, 2020, 9138. https://doi.org/10.3390/su12219138

[2].Kaur, R., Singh, S. A comprehensive review of object detection with deep learning. *Digital Signal Processing*,132, 2023, 103812. https://doi.org/10.1016/j.dsp.2022.103812

[3].P. G. Shambharkar, S. Sharma. Plant disease detection and prevention using deep learning,. *2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS)*, Coimbatore, India, 2023, pp. 1832-1838, doi: 10.1109/ICACCS57279.2023.10112733.

[4].G. Lavanya, S. D. Pande. Enhancing Real-time object detection with YOLO algorithm. *EAI Endorsed Trans IoT*, vol. 10, Dec. 2023. https://doi.org/10.4108/eetiot.4541

[5].Stark, T., Ştefan, V., Wurm, M. et al. YOLO object detection models can locate and classify broad groups of flower-visiting arthropods in images. *Science Rep*. vol. 13, 16364, 2023. https://doi.org/10.1038/s41598-023-43482-3

[6].Yang, L., Zhong, J., Zhang, Y., Bai, S., Li, G., Yang, Y., Zhang, J. An improving faster-RCNN with multi-attention ResNet for small target detection in intelligent autonomous transport with 6G. *IEEE Transactions on Intelligent Transportation Systems*. 2022. 10.1109/TITS.2022.3193909

[7].R. Rubin, C. Jacob, S. M. Anzar, A. Panthakkan, Mask R-CNN with multi-backbones – a comparative analysis. *2022 5th International Conference on Signal Processing and Information Security (ICSPIS)*, Dubai, United Arab Emirates, 2022, pp. 55-60, doi: 10.1109/ICSPIS57063.2022.10002546.

[8].Jia, W., Tian, Y., Luo, R., Zhang, Z., Lian, J., Zheng, Y. Detection and segmentation of overlapped fruits based on optimized mask R-CNN application in apple harvesting robot. *Computers and Electronics in Agriculture*, 172, 105380, 2020. https://doi.org/10.1016/j.compag.2020.105380 .

[9].Bai, Y., Mao, S., Zhou, et al. Clustered tomato detection and picking point location using machine learning-aided image analysis for automatic robotic harvesting. *Precision Agric*, vol.24, 2023, pp.727–743. https://doi.org/10.1007/s11119-022-09972-6

[10]. Tang, Z., He, X., Zhou, G., Chen, A., Wang, Y., Li, L., Hu, Y. (2023). A precise image-based tomato leaf disease detection approach using PLPNet. *Plant Phenomics*, 5, 2-23. DOI: 10.34133/plantphenomics.0042

[11]. Jigar Patel , Hardik Talsania , Kirit Modi . ODNET: Optimized DenseNet for Indian food classification. *International Journal on Information Technologies and Security*, vol.15, no.4, 2023, pp. 27-36. https://doi.org/10.59035/FPBL3081

[12]. Arati Shahapurkar, S. F. Rodd. Efficient feature aware machine learning model for detecting fraudulent transaction in streaming environment. *International Journal on Information Technologies and Security*, vol.14, no.3, 2022, pp. 3-14. https://ijits-bg.com/2022.v14.i3.01

[13]. H. Yi, B. Liu, B. Zhao, E. Liu. Small object detection algorithm based on improved YOLOv8 for remote sensing, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 17, 2024, pp. 1734-1747. doi: 10.1109/JSTARS.2023.3339235.

[14]. A. Afdhal, K. Saddami, S. Sugiarto, Z. Fuadi, N. Nasaruddin, Real-time object detection performance of YOLOv8 models for self-driving cars in a mixed traffic environment. *2023 2nd International Conference on Computer System, Information Technology, and Electrical*

*Engineering (COSITE)*, Banda Aceh, Indonesia, 2023, pp. 260-265, doi: 10.1109/COSITE60233.2023.10249521.

[15]. L. Shen, B. Lang, Z. Song, "DS-YOLOv8-based object detection method for remote sensing images. *IEEE Access*, vol. 11, 2023, pp. 125122-125137. doi: 10.1109/ACCESS.2023.3330844

[16]. Gao, C., Zhang, Q., Tan, Z. et al. Applying optimized YOLOv8 for heritage conservation: enhanced object detection in Jiangnan traditional private gardens. *Herit Sci,* 12, 31, 2024. https://doi.org/10.1186/s40494-024-01144-1

[17]. Zhou, Y. A yolo-nl object detector for real-time detection. *Expert Systems with Applications*, 238, 2024, 122256.

[18]. Tomato-leaf-desease-ssoha_dataset: https://universe.roboflow.com/bryan-b56jm/tomato-leaf-disease-ssoha/dataset/63 (last visited 25-03-2024)

## *Information about the authors:*

**Mohamed Ben Ammar** – Dr. Mohamed Ben Ammar is an Assistant Professor in the Department of Information Systems at the Faculty of Computing and Information Technology, Northern Border University, in Rafha, Saudi Arabia. He holds a Ph.D. in Engineering of Information Systems from Sfax University's National Engineering School of Sfax (ENIS) in Tunisia. His expertise in affective computing, intelligent tutoring systems, sentiment analysis, and multimodal emotion recognition holds the potential to revolutionize how we interact with machines, from emotionally intelligent virtual assistants to personalized learning experiences.

**Manuscript received on 24 April 2024**