

PlantSeg: A Large-Scale In-the-wild Dataset for Plant Disease Segmentation

Tianqi Wei¹, Zhi Chen¹, Xin Yu¹, Scott Chapman², Paul Melloy³, and Zi Huang¹

¹School of Electrical Engineering and Computer Science, The University of Queensland, Brisbane, 4067, Australia

²School of Agriculture and Food Sustainability, The University of Queensland, Brisbane, 4067, Australia

³Agriculture and Food, CSIRO, Dutton Park, Queensland, 4102, Australia

ABSTRACT

Plant diseases pose significant threats to agriculture. It necessitates proper diagnosis and effective treatment to safeguard crop yields. To automate the diagnosis process, image segmentation is usually adopted for precisely identifying diseased regions, thereby advancing precision agriculture. Developing robust image segmentation models for plant diseases demands high-quality annotations across numerous images. However, existing plant disease datasets typically lack segmentation labels and are often confined to controlled laboratory settings, which do not adequately reflect the complexity of natural environments. Motivated by this fact, we established PlantSeg, a large-scale segmentation dataset for plant diseases. PlantSeg distinguishes itself from existing datasets in three key aspects. (1) Annotation type: Unlike the majority of existing datasets that only contain class labels or bounding boxes, each image in PlantSeg includes detailed and high-quality segmentation masks, associated with plant types and disease names. (2) Image source: Unlike typical datasets that contain images from laboratory settings, PlantSeg primarily comprises in-the-wild plant disease images. This choice enhances the practical applicability, as the trained models can be applied for integrated disease management. (3) Scale: PlantSeg is extensive, featuring 11,400 images with disease segmentation masks and an additional 8,000 healthy plant images categorized by plant type. Extensive technical experiments validate the high quality of PlantSeg's annotations. This dataset not only allows researchers to evaluate their image classification methods but also provides a critical foundation for developing and benchmarking advanced plant disease segmentation algorithms.

Background & Summary

Plant diseases are a serious threat to agricultural productivity and can significantly impact crop yields and quality¹. Globally, between 20% and 40% of all crops are lost due to plant diseases. According to The Food and Agriculture Organization of the United Nations², annual losses exceed 220 billion dollars due to plant diseases. Early and accurate plant disease detection and assessment is crucial for minimizing economic losses. Traditionally, manual diagnosis by plant pathologists is considered the most reliable method of assessment. However, diagnosticians are not always available to provide assessment in a timely manner, leading to potentially costly delays. Further, plant pathologists are often skilled at recognizing a limited number of plant diseases on a handful of hosts, thus multiple plant pathologists or taxonomists are usually required for a reliable diagnosis.

Arguably, one of the goals for precision agriculture³ includes improvements to agricultural systems enabling the automatic localization and segmentation of disease-affected plants and plant parts. Generic image segmentation methods^{4–7} have demonstrated outstanding performance on commonly used benchmark datasets, such as ADE20k⁸, Cityscapes⁹ and MSCOCO¹⁰. However, there is still a huge gap between the mainstream segmentation models and the common ones being used for plant disease segmentation. Most recent plant disease segmentation studies^{11–18} typically adopt obsolete deep learning models for segmenting narrow selections of host and pathogen relationships. In contrast, a more generalized approach to segmenting a wider variety of plant diseases sets a far more fine-grained and challenging task, to model the characteristics of different diseases.

The challenge of developing an advanced deep learning-based plant disease segmentation model is made more difficult due to the substantial number of annotated plant images required and the lack of publicly available high-quality datasets. Currently, the availability of plant disease datasets is limited, and most accessible datasets are not sufficiently labeled in terms of annotation type, image source, and scale. We provided the statistics of existing plant disease datasets in Table 1 and elaborate on their insufficiency as follows:

- **Annotation Type.** Most existing plant disease image datasets are designed for classification or object detection tasks. Classification involves identifying the global content in an image but does not provide any local information. Object detection can localize objects by drawing bounding boxes around them. There are only a few datasets available for image

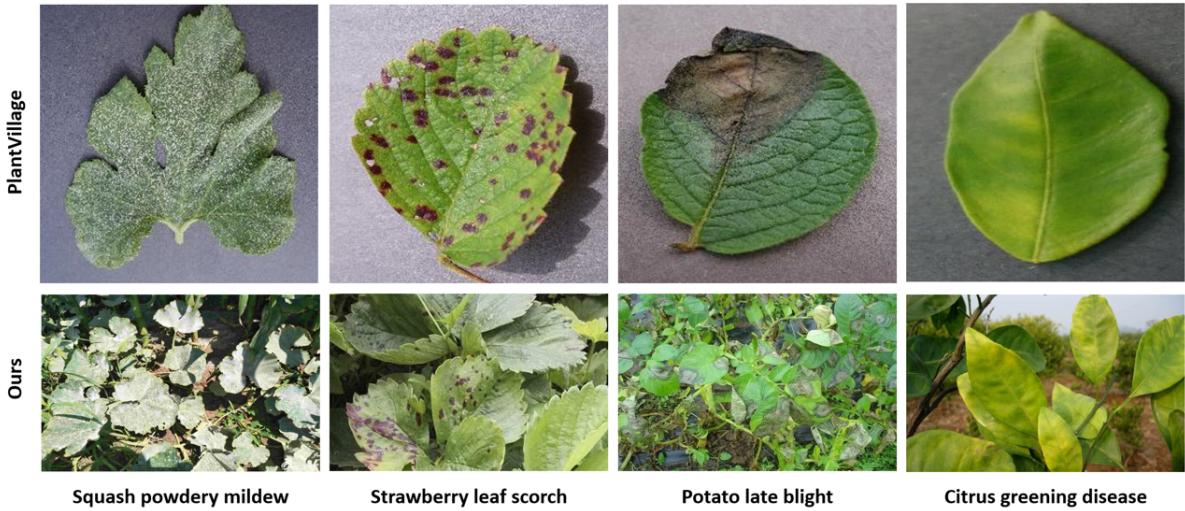


Figure 1. Examples of images of PlantVillage¹⁹ and our dataset. As collected in laboratory environments, each image in PlantVillage only contains one leaf and has a uniform background, while images of our dataset feature much more complex backgrounds, various viewpoints, and different lighting conditions.

Dataset Name	Year	#Images	#Classes	#Plants	In-the-wild	Bounding box	Segmentation mask	References
PlantVillage	2015	54,309	38	14	✗	✗	✗	Hughes, <i>et al.</i> ¹⁹
PlantDoc	2020	2,598	27	13	✓	✗	✗	Singh, <i>et al.</i> ²⁰
FieldPlant	2023	5,170	27	3	✓	✗	✗	Moupojou, <i>et al.</i> ²¹
PlantWild	2024	18,542	89	33	✓	✗	✗	Wei, <i>et al.</i> ²²
Tomato Disease	2020	4,178	4	1	✗	✓	✗	Zhang, <i>et al.</i> ²³
GLDD	2020	4,449	4	1	✗	✓	✗	Xie, <i>et al.</i> ¹⁵
APPLE & GRAPE	2021	1,150	6	2	✗	✓	✗	Savarimuthu, <i>et al.</i> ¹⁶
MSMSVDD	2022	1,000	5	3	✓	✓	✗	Li, <i>et al.</i> ¹⁴
LDSD	2021	588	1	N/A	✓	✗	✓	Fakhre, <i>et al.</i> ²⁴
NLB	2024	1,000	1	1	✓	✗	✓	Prashanth, <i>et al.</i> ²⁵
PlantSeg (ours)	2024	11,458	115	34	✓	✗	✓	N/A

Table 1. Summary of plant disease image datasets. Existing datasets

segmentation^{24,25}, which still require detailed and precise mask annotations. Accurately localizing the affected areas, pixel-wise, with segmentation masks provides improved granularity compared to purely predicting the bounding box with object detection^{14–16,23}. The affected areas are usually irregular shapes, the outputs from semantic segmentation can be directly used in automated systems for precision agriculture, such as quarantining affected areas within paddocks, variable fungicide application rates to minimize the spread through the paddock. These assessments might lead to the ability to use decision support tools and integrated disease management (IDM) on a sub-paddock scale.

- **Image Source.** Images in many existing datasets^{15, 16, 19, 23} are collected under controlled laboratory conditions. As illustrated in the upper row of Figure 1, such images lack the complex background information in field environments. As a result, algorithms trained on ‘staged’ datasets may not perform in real-world scenarios where plants are subjected to various environmental factors, such as varying lighting conditions, occlusions, and background noise. This makes lab-trained models unsuitable for segmentation tasks in the field where they are likely to be applied for integrated disease management.
- **Scale.** Existing datasets are often small in scale^{24,25}, which contain a limited number of images, categories, and a narrow host and pathogen focus. Consequently, algorithms trained on such datasets cannot be applied to detect diseases outside the scope for which they have been characterized as their lack of generalizability limits the practical application.

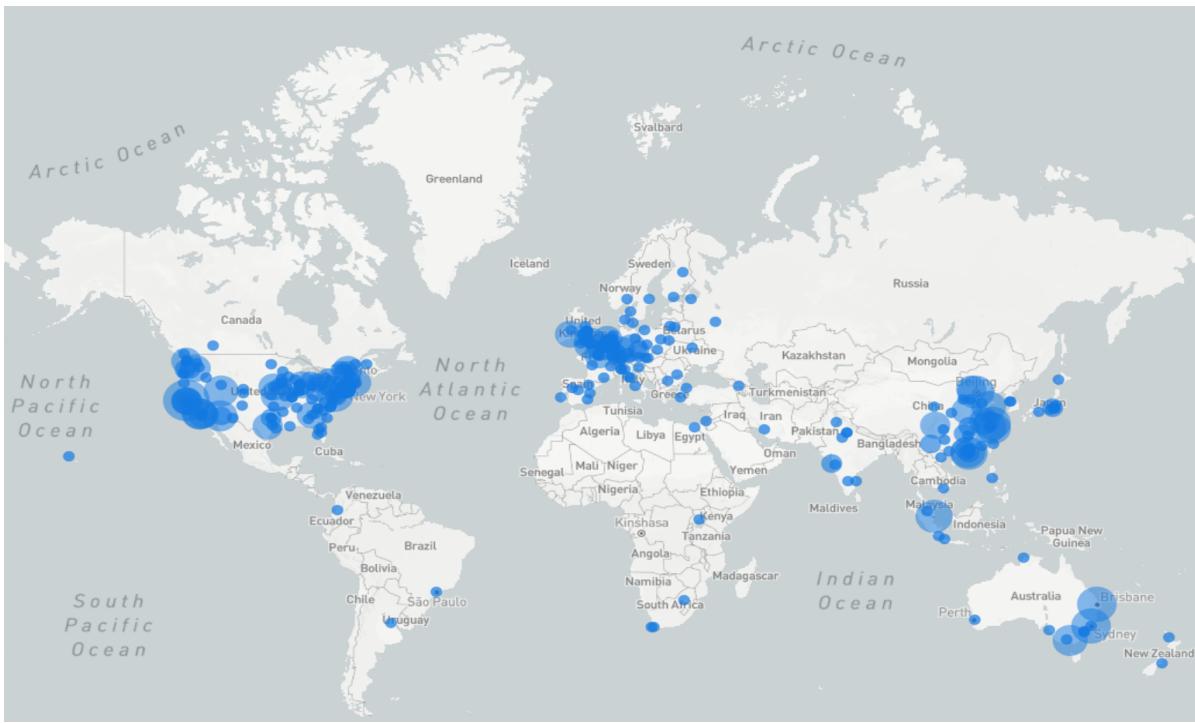


Figure 2. Locations of the source image acquired. The sizes of the plots represent the number of acquired images. The size of each circle demonstrates the number of images acquired from the address, and the color depth indicates the density of addresses within a nearby region.

To address these problems in plant disease segmentation research, we present PlantSeg, the largest dataset for plant disease segmentation in the wild. It has the most number of categories among all existing plant disease datasets. PlantSeg contains 11,458 images of 115 disease categories with corresponding segmentation annotations. The segmentation annotation is carried out by trained annotators and checked by expert pathologists to ensure accuracy. This paper showcases the characteristics of PlantSeg and benchmarks state-of-the-art segmentation models on plant disease segmentation. We believe our dataset can serve as a comprehensive benchmark for developing plant disease segmentation methods.

Methods

Image acquisition

To build our dataset, we carefully selected plants that hold substantial economic and nutritional significance. It consists of **profit crops** with high commercial value, **staple crops** that are essential for human consumption, as well as a diverse range of **fruits** and **vegetables**, which contribute significantly to agricultural production. Including a wide variety of plants, our dataset is both comprehensive and representative of the most important plants in the agriculture field. Consequently, we have identified 115 diseases across 34 plants for our dataset curation. For each plant disease, we utilized its name as the keyword to search relevant images from various internet sources, including Google Images, Bing Images, and Baidu Images. This comprehensive collection strategy can broaden the retrieval range and increase the diversity of results, as images were collected from websites all over the world. Therefore, the scale of our dataset was effectively expanded. The IP addresses of the image source websites are presented in Figure 2, and show our dataset incorporates images from diverse regions from around the world.

Image cleaning

The retrieved images were organized into folders corresponding to their respective disease names. Data quality was 'cleaned' by a review process before proceeding with segmentation annotation. The cleaning process involved our annotators carefully reviewing each image and removing any incorrect images in a class folder, retaining only accurate images of each class. Ambiguous images, or images difficult to classify, were also discarded. This process ensured the accuracy of every image through cross-validation by at least two annotators. Furthermore, in cases where there were discrepancies between the annotators' judgements, experts with extensive knowledge and experience were brought in to review and make final decisions.



Figure 3. Examples of images with annotated polygons on the disease-affected areas.

Metadata	Description
Name	The name of the image
Plants	The host plants of interest
Diseases	Disease type
Resolutions	The resolution of the image
Label files	The path of corresponding label file
Mask ratios	The proportion of annotated pixels to the total number of pixels.
URLs	Download link of the image if available
Training/Test split	Specify images of each disease type for Training/Test set

Table 2. The Metadata of PlantSeg.

Segmentation annotation

After the review process, we established a segmentation annotation standard to ensure consistent labelling of disease-affected areas in the images. For distinct and sizeable lesions, annotations were made with individual polygons. Overlapping lesions were annotated as a combined affected area. For diseases like rust and powdery mildew, which present as small, densely clustered symptoms on both leaves and fruits, we meticulously annotated the infected regions to accurately reflect the disease distribution. Additionally, any deformities in plant leaves or fruits caused by diseases were also annotated. Examples of annotated images are shown in Figure 3.

Under the guidance of two expert plant pathologists, a group of annotators participated in the segmentation annotation process. The annotators were trained on the annotation standard and then required to annotate 10 images for evaluation. The expert pathologists reviewed these annotations, and any annotator whose work was deemed unsatisfactory was asked to re-annotate the images. Annotators who consistently did not meet the re-annotation standards were disqualified from further annotation tasks. Conversely, if the annotation results met the standard, the annotators were approved to proceed to the subsequent stages of the project.

In total, 10 annotators were engaged in this detailed segmentation annotation work. We divide the images into various subsets. Images were subsetted by host plant and included a number of different disease. Each annotator was assigned a specific subset and used LabelMe (V5.5.0)²⁶ software to annotate the diseased parts. Following the image subset primary annotations, by the first annotator, the image subset was passed to another annotator for review, and correct any errors. The final annotations underwent a rigorous review by expert pathologists. Figure 4 demonstrates the whole workflow of the segmentation annotation process.

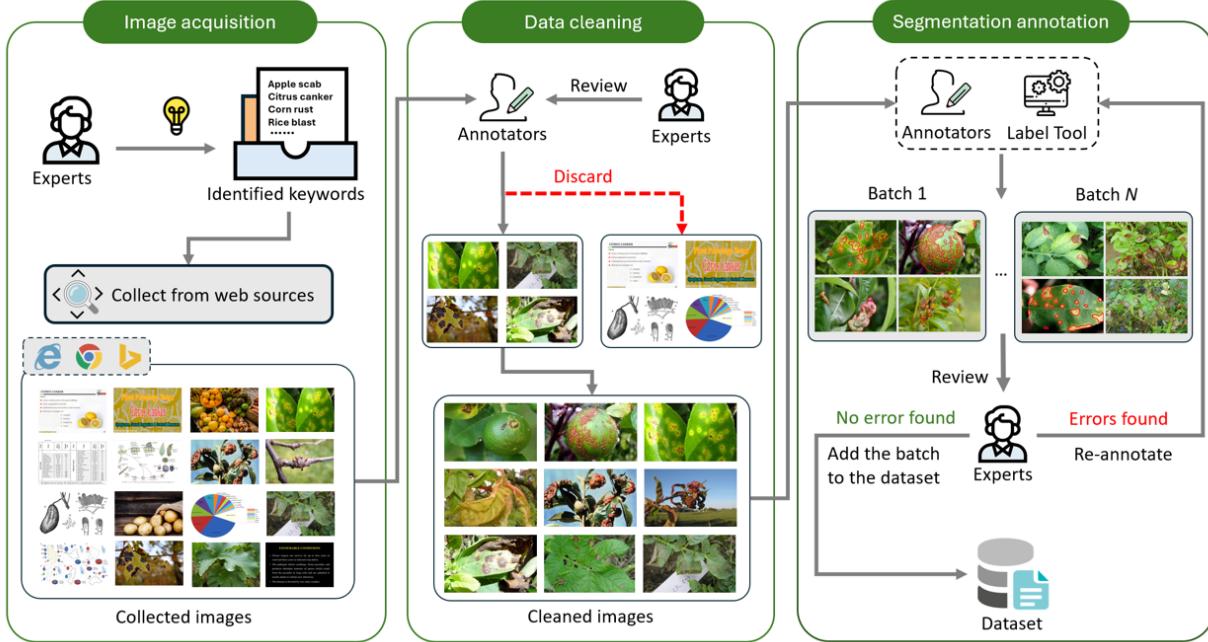


Figure 4. The curation process of the PlantSeg dataset involves three main steps: image acquisition, data cleaning, and annotation. In the image acquisition stage, images were collected from various internet sources using identified keywords and then stored according to their categories. During the data cleaning phase, incorrect images were identified and removed. For the segmentation annotation process, annotators utilized LabelMe²⁶ to annotate the cleaned images. These annotations were subsequently reviewed by experts and saved in JSON files.

PlantSeg metadata

In this section, we introduce the metadata of the PlantSeg dataset, which provides detailed information related to each image. An overview is presented in Table 2.

Plants and Diseases. 34 host plants of interest were determined according to the suggestions of experts. These were classified into three categories relating to their socioeconomic importance to humans. Profit crops, included commodities with high market demand and low nutritional value, *e.g.*, Coffee and Tobacco. Staple crops, refer to crops providing the majority of people's carbohydrate and protein requirements, such as wheat, corn, and potatoes. PlantSeg also contains a wide range of fruits and vegetables, such as apples, oranges, and tomatoes, which supplement nutritional needs and are vital for people's diets. Based on the plants of interest, 115 common disease categories were determined to be included in the dataset.

URLs. To aid in the reproducibility of this work, all downloaded images used to build the PlantSeg dataset were stored with URL links to the source websites. This allows researchers to verify the source of the images and ensure compliance with copyright regulations.

Label files. The label files share the same filenames as their corresponding images, differing only by their file extensions. These labels are stored as grayscale PNG images, where pixels representing diseased areas are annotated with specific class index values, while all other pixels are assigned a value of zero.

Image resolutions and Mask ratios. The image resolution indicates the image size including width and height, and the mask ratio denotes the proportion of labeled pixels to the total number of pixels in the image.

Training/Test sets of PlantSeg. PlantSeg is built to evaluate segmentation methods on plant disease images. We randomly selected 20% of the images from each disease as the test set, while the remaining images were used as the training set.

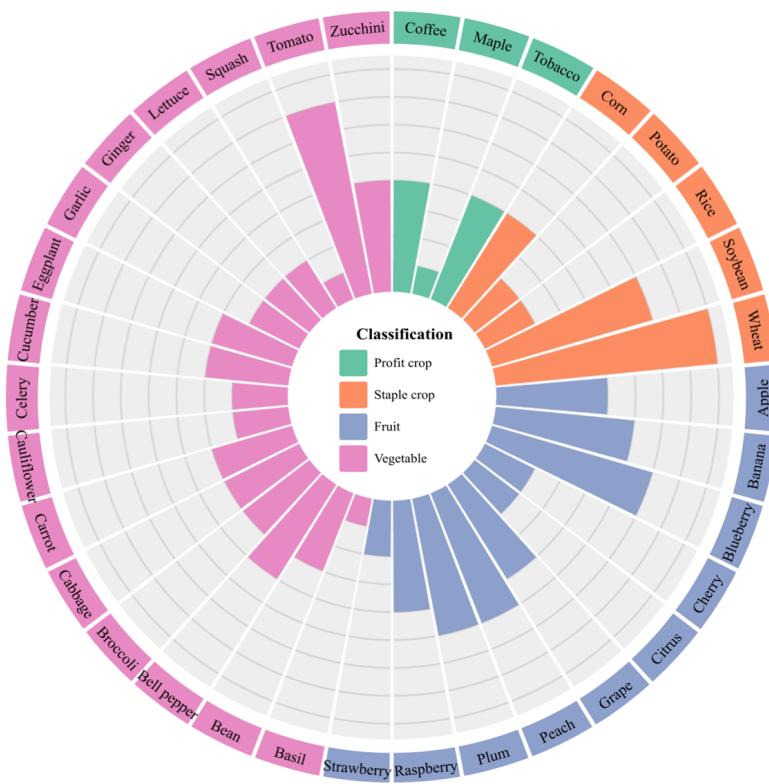


Figure 5. Disease distribution in PlantSeg according to plants and Socioeconomic classification. The height of the bars represents the number of diseases associated with each plant.

Data Records

The PlantSeg dataset can be downloaded through [https://doi.org/10.5281/zenodo.13293891²⁷](https://doi.org/10.5281/zenodo.13293891). The repository is covered under the CC BY-NC-ND 4.0 licence. Plant disease images are saved in JPEG format and are stored in the “images” folder, and the labels are saved in PNG format and are stored in the “annotations” folder. Each image and its corresponding label have the same file name except the file extension. In addition, the original label files generated by LabelMe²⁶ and saved in JSON format, are provided in the “json” folder. All images and labels are split into training and test sets with an 80/20 ratio. A PlantSeg-Meta.csv file is provided to store the meta-information presented in Table 2.

Technical Validation

Data property analysis

We conduct a comprehensive analysis of the PlantSeg dataset from multiple aspects, including distributions of disease type, image resolution, and segmentation mask ratio.

Plant and disease type distribution. The PlantSeg dataset includes 115 diseases across 34 plant hosts, which are categorized into four major socioeconomic groups: **profit crops**, **staple crops**, **fruits**, and **vegetables**. Figure 5 provides an overview of the distribution of plant hosts and disease types. Fifteen plant hosts from vegetables and ten fruit hosts constitute significant portions of the dataset, with 45 and 39 diseases respectively. In contrast, profit crops consist of only 9 diseases across 3 plant hosts, accounting for 7.8% of the total image database.

Image resolution distribution. We analyze image resolution distribution in PlantSeg and compare it with two widely used plant disease image datasets: PlantVillage¹⁹ and PlantDoc²⁰. PlantVillage consists of more than 50,000 images, all images are captured from plant hosts under controlled experimental conditions. PlantDoc contains images collected from field environments, and only includes about 2,600 images. The resolution distribution is presented in Figure 6. This scatter plot demonstrates that PlantSeg covers a wide range of image resolutions and reflects the variability typical of real-world conditions, compared to other databases. PlantDoc (red points) also exhibits considerable variability, albeit on a lower scale compared

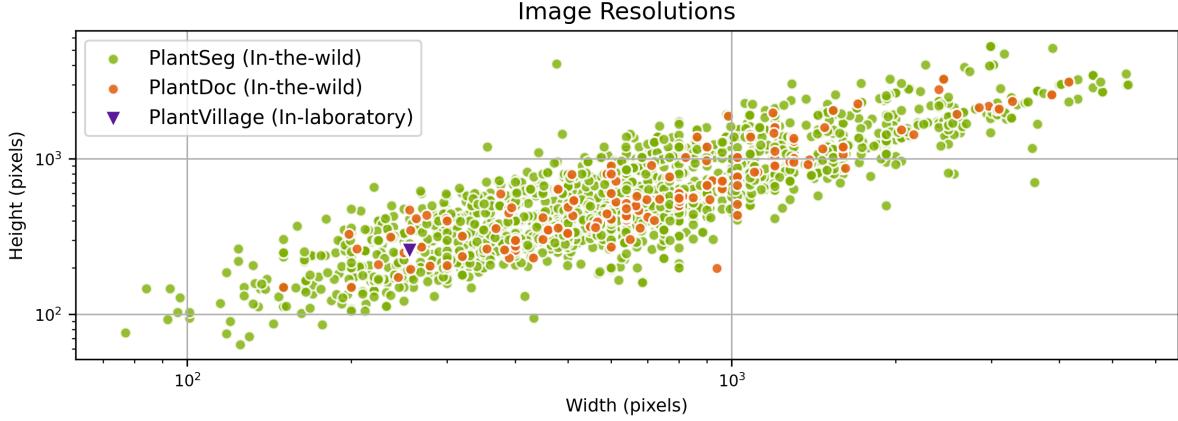


Figure 6. The resolution distributions of different datasets, including PlantVillage, PlantDoc and PlantSeg. Each green dot represents a single image in the PlantSeg database; Red, for each image in the PlantDoc database and inverted yellow triangles for the PlantVillage database.

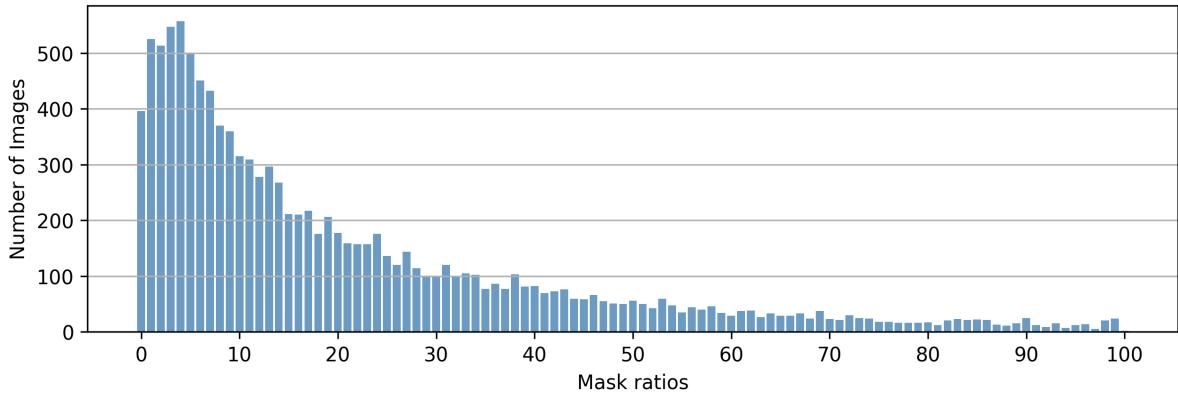


Figure 7. The horizontal axis represents the percentage of mask area relative to the entire image, while the vertical axis represents the number of corresponding images.

with PlantSeg. PlantVillage images were curated from uniform laboratory settings with a single resolution. All data points overlap in a single point on the plot. Figure 6 reveals the diversity and range of image resolutions in in-the-wild datasets compared to laboratory-collected data. It indicates the challenge when working on real-world data collection and the inherent variance in images. In addition, it also emphasizes the advantage of a dataset, such as ours, which focuses on field-based images with variable scale and diversity.

Segmentation mask ratio distribution. Figure 7 depicts the distribution of segmentation mask ratios in the PlantSeg dataset. A low ratio indicates the annotated disease area is relatively small, while a high ratio suggests a larger area. Overall the mask ratio of the PlantSeg image database exhibits significant variation. The box plot in Figure 8 shows considerable variation in mask ratio distribution among different diseases and within each disease type.

Evaluation on PlantSeg

To establish benchmarks for plant disease segmentation, we applied four segmentation methods and evaluated their performance on the PlantSeg dataset.

Baseline models. We employed four commonly used and state-of-the-art semantic segmentation methods as our baselines, including Side Adapter Network (SAN)²⁸, DeepLabv3⁴, DeepLabv3+⁵, and SegNeXt⁷. These methods were trained and evaluated on the same training and test sets respectively. They were implemented using PyTorch version 1.11. For DeepLabv3⁴, DeepLabv3+⁵, and SAN²⁸, we respectively leveraged different variants of ResNet²⁹ and Vision Transformer³⁰ as the backbones. All methods were trained using a Stochastic gradient descent (SGD) optimizer, with a learning rate of 0.001, a momentum of

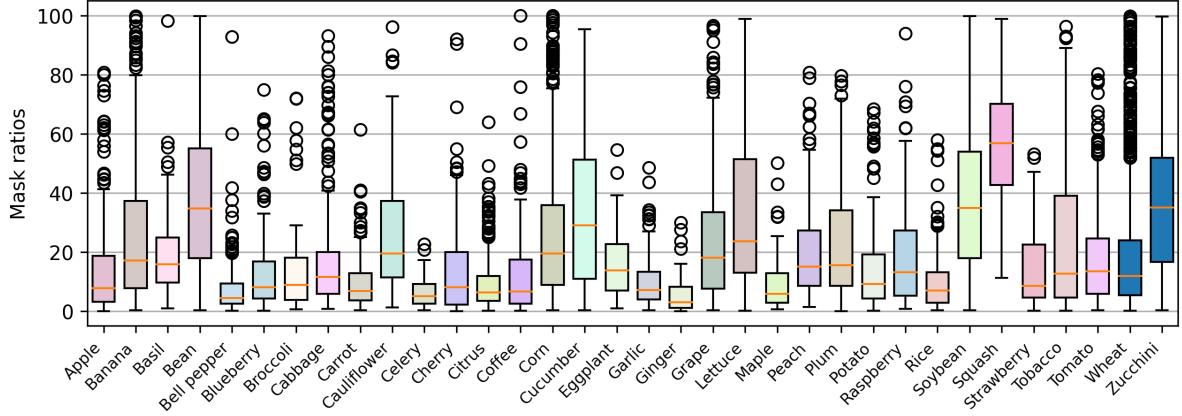


Figure 8. The boxplot shows the percentage distribution of mask area per image, which varies significantly among different plants.

Method	Backbone	MIoU	mAcc
DeepLabv3 ⁴	ResNet-50 ²⁹	17.24	37.95
DeepLabv3 ⁴	ResNet-101 ²⁹	20.72	40.63
DeepLabv3+ ⁵	ResNet-50 ²⁹	25.08	40.66
DeepLabv3+ ⁵	ResNet-101 ²⁹	27.18	42.29
SAN ²⁸	ViT-B/16 ³⁰	34.79	50.19
SAN ²⁸	ViT-L/14 ³⁰	36.91	52.81
SegNeXt ⁷	MSCAN-L	44.52	59.95

Table 3. Performance comparison of different methods on PlantSeg.

0.9, and a weight decay rate of 0.0005. We introduced cross-entropy loss to optimize the models. Each model was trained with a fixed batch size of 16.

Evaluation metrics. We introduced the Mean Intersection over Union (MIoU) and Mean Accuracy (mAcc) as our evaluation metrics. MIoU calculates the average Intersection over Union across all classes, offering insight into the model’s overall segmentation performance. Mean Accuracy measures the proportion of correctly classified pixels within each class and then averages the accuracies across all classes. These evaluation metrics are computed as:

$$\text{MIoU}_i = \frac{1}{N} \sum_{i=1}^N \frac{TP_i}{TP_i + FP_i + FN_i}, \quad \text{mAcc} = \frac{1}{N} \sum_{i=1}^N \frac{TP_i}{TP_i + FN_i}, \quad (1)$$

where N represents the number of classes. TP_i , FP_i , TN_i , FN_i are the number of true positive, false positive, false negative, and true negative pixels of the i -th class respectively.

Experiment results. The segmentation results of the selected baselines are summarized in Table 3. The findings show that DeepLabv3⁴ and DeepLabv3+⁵ with a deeper ResNet-101 backbone outperform those with a ResNet-50 backbone in both MIoU and mAcc scores. Similarly, SAN²⁸ with a ViT-L/14 backbone demonstrates superior performance compared to the ViT-B/16 backbone, with a 4.79% increase in MIoU and a 5.89% increase in mAcc. This suggests that using a larger backbone can benefit the methods and lead to substantial improvements in segmentation performance. Among all the baselines, SegNeXt⁷, which incorporates a large multi-scale convolutional backbone, achieves the highest performance, with the MIoU of 53.89% and the mAcc of 65.91%.

Figure 9 presents a series of visualizations comparing the ground truth masks with the predictions from various baselines, including DeepLabv3⁴ and SegNext⁷. The results from DeepLabv3 are particularly unsatisfactory, as it struggles to identify

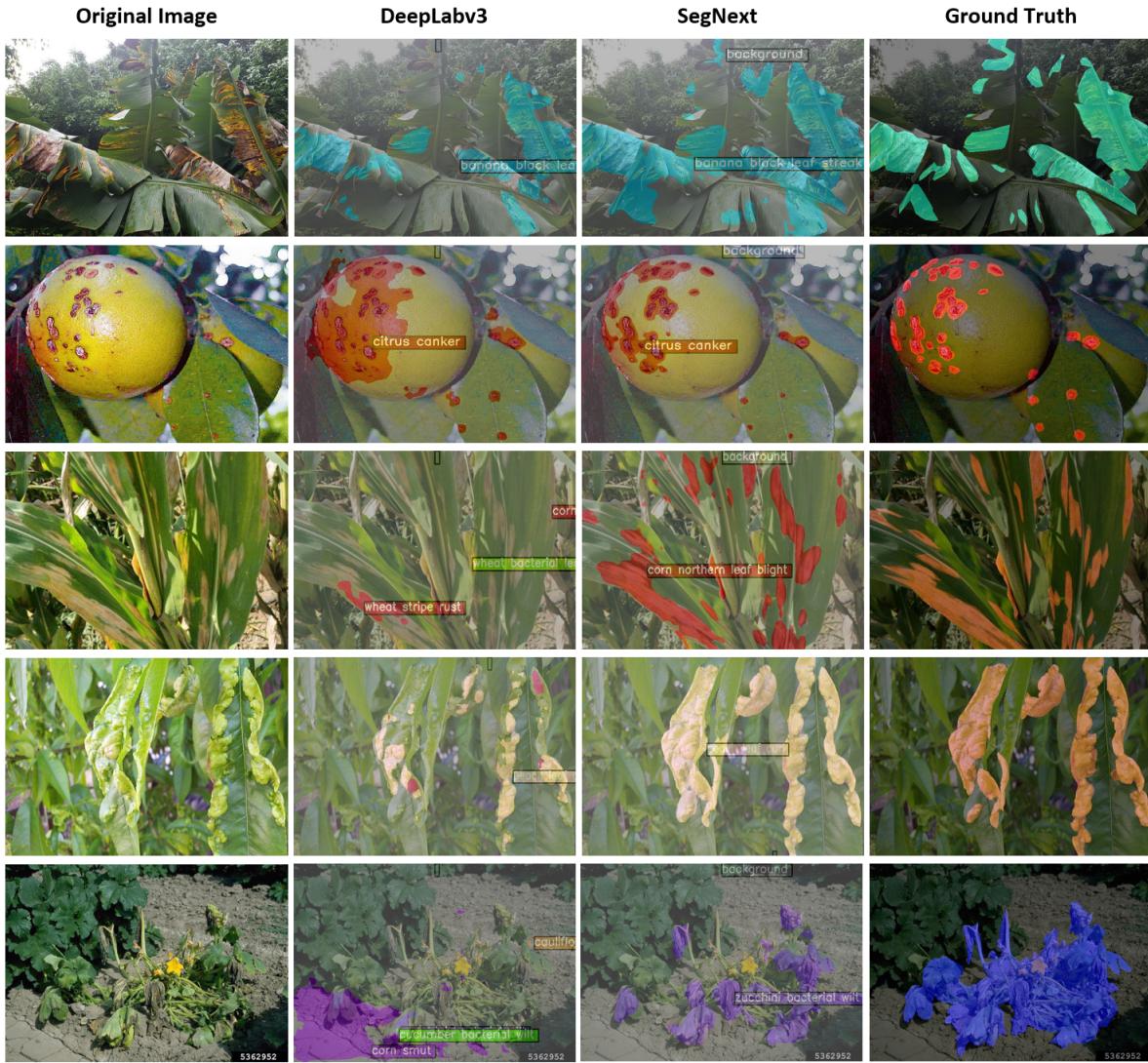


Figure 9. Visualization of some experimental results on the test set of PlantSeg. From left to right: image examples of PlantSeg, results of DeepLabv3⁴, results of SegNeXt⁷, and the ground truth annotations.

and locate the diseased areas accurately. The state-of-the-art SegNeXt delivers more accurate results than DeepLabv3, as it can effectively segment lesions and deformation on leaves and fruit according to the results shown in the 1st to 4th rows of Figure 9. However, in the cases of the 5th row, SegNeXt focuses on the wilted leaves but overlooks the collapsed stems. This suggests that segmentation becomes more challenging when the disease involves curling and deformation of stems.

Conclusion

PlantSeg offers a step forward towards automation for disease detection and quantification using ordinary proximal sensing devices such as RGB cameras. Multispectral and hyperspectral devices, which are currently used in research for disease detection and quantification in the field have a higher cost and are beyond the budgets for many in the agricultural industry³¹. Segmentation of images to define the area of affected plant parts can allow an automated method for quantification of diseases signs and symptoms at a stated point in time. This provides an unbiased method for estimating disease severity scores for researchers and industry decision-makers to make a judgment on the timing of integrated disease management practices. While laboratory-trained segmentation methods for disease symptom segmentation don't perform well in the field, they could still provide an automated disease method of laboratory experiments, lowering the requirements for highly skilled technicians with years of pathology experience.

Code availability

The codes for the baseline reproduction are presented in <https://github.com/tqwei05/PlantSeg>. The codes benefit from <https://github.com/open-mmlab/mmsegmentation>, which provides a benchmark toolbox for numerous segmentation methods.

References

1. Shoaib, M. *et al.* An advanced deep learning models-based plant disease detection: A review of recent research. *Front. Plant Sci.* **14**, 1158933 (2023).
2. Agrios, G. N. *Plant pathology* (2005).
3. Shafi, U. *et al.* Precision agriculture techniques and practices: From considerations to applications. *Sensors* **19**, 3796 (2019).
4. Chen, L.-C., Papandreou, G., Schroff, F. & Adam, H. Rethinking atrous convolution for semantic image segmentation. *arXiv preprint arXiv:1706.05587* (2017).
5. Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F. & Adam, H. Encoder-decoder with atrous separable convolution for semantic image segmentation. In *Proceedings of the European conference on computer vision (ECCV)*, 801–818 (2018).
6. Kirillov, A., Wu, Y., He, K. & Girshick, R. Pointrend: Image segmentation as rendering. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 9799–9808 (2020).
7. Guo, M.-H. *et al.* Segnext: Rethinking convolutional attention design for semantic segmentation. *Adv. Neural Inf. Process. Syst.* **35**, 1140–1156 (2022).
8. Zhou, B. *et al.* Semantic understanding of scenes through the ade20k dataset. *Int. J. Comput. Vis.* **127**, 302–321 (2019).
9. Cordts, M. *et al.* The cityscapes dataset for semantic urban scene understanding. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 3213–3223 (2016).
10. Lin, T.-Y. *et al.* Microsoft coco: Common objects in context. In *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6–12, 2014, Proceedings, Part V 13*, 740–755 (Springer, 2014).
11. Bhatti, M. A. *et al.* Advanced plant disease segmentation in precision agriculture using optimal dimensionality reduction with fuzzy c-means clustering and deep learning. *IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens.* (2024).
12. Jafar, A., Bibi, N., Naqvi, R. A., Sadeghi-Niaraki, A. & Jeong, D. Revolutionizing agriculture with artificial intelligence: plant disease detection methods, applications, and their limitations. *Front. Plant Sci.* **15**, 1356260 (2024).
13. Wang, D., Wang, J., Li, W. & Guan, P. T-cnn: Trilinear convolutional neural networks model for visual detection of plant diseases. *Comput. Electron. Agric.* **190**, 106468 (2021).
14. Li, J. *et al.* An improved yolov5-based vegetable disease detection method. *Comput. Electron. Agric.* **202**, 107345 (2022).
15. Xie, X. *et al.* A deep-learning-based real-time detector for grape leaf diseases using improved convolutional neural networks. *Front. plant science* **11**, 751 (2020).
16. Savarimuthu, N. *et al.* Investigation on object detection models for plant disease detection framework. In *2021 IEEE 6th international conference on computing, communication and automation (ICCCA)*, 214–218 (IEEE, 2021).
17. Wei, T., Chen, Z. & Yu, X. Snap and diagnose: An advanced multimodal retrieval system for identifying plant diseases in the wild. *arXiv preprint arXiv:2408.14723* (2024).
18. Shoaib, M. *et al.* Deep learning-based segmentation and classification of leaf images for detection of tomato plant disease. *Front. Plant Sci.* **13**, 1031748 (2022).
19. Hughes, D., Salathé, M. *et al.* An open access repository of images on plant health to enable the development of mobile disease diagnostics. *arXiv preprint arXiv:1511.08060* (2015).
20. Singh, D. *et al.* Plantdoc: A dataset for visual plant disease detection. In *Proceedings of the 7th ACM IKDD CoDS and 25th COMAD*, 249–253 (2020).
21. Moupojou, E. *et al.* Fieldplant: A dataset of field plant images for plant disease detection and classification with deep learning. *IEEE Access* **11**, 35398–35410 (2023).
22. Wei, T., Chen, Z., Huang, Z. & Yu, X. Benchmarking in-the-wild multimodal plant disease recognition and a versatile baseline. In *ACM International Conference of Multimedia* (2024).

23. Zhang, Y., Song, C. & Zhang, D. Deep learning-based object detection improvement for tomato disease. *IEEE access* **8**, 56607–56614 (2020).
24. Leaf disease segmentation dataset — kaggle.com. <https://www.kaggle.com/datasets/fakhrealam9537/leaf-disease-segmentation-dataset/data?select=data>. [Accessed 07-08-2024].
25. Prashanth, K., Harsha, J. S., Kumar, S. A. & Srilekha, J. Towards accurate disease segmentation in plant images: A comprehensive dataset creation and network evaluation. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 7086–7094 (2024).
26. GitHub - labelmeai/labelme: Image Polygonal Annotation with Python (polygon, rectangle, circle, line, point and image-level flag annotation). — github.com. <https://github.com/labelmeai/labelme>. [Accessed 07-08-2024].
27. Wei, T. A large-scale in-the-wild dataset for plant disease segmentation, [10.5281/ZENODO.13293891](https://doi.org/10.5281/ZENODO.13293891) (2024).
28. Xu, M., Zhang, Z., Wei, F., Hu, H. & Bai, X. Side adapter network for open-vocabulary semantic segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2945–2954 (2023).
29. He, K., Zhang, X., Ren, S. & Sun, J. Deep residual learning for image recognition. In *CVPR* (2016).
30. Dosovitskiy, A. *et al.* An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929* (2020).
31. Bock, C. H., Barbedo, J. G., Del Ponte, E. M., Bohnenkamp, D. & Mahlein, A.-K. From visual estimates to fully automated sensor-based measurements of plant disease severity: status and challenges for improving accuracy. *Phytopathol. Res.* **2**, 1–30 (2020).

Author contributions statement

Tianqi Wei designed the study, built the dataset, conducted experiments and wrote the manuscript. Zhi Chen designed the study, built the dataset and wrote the manuscript. Xin Yu designed the study and revised the manuscript. Scott Chapman and Paul Melloy supervised the annotation process, validated the data and reviewed the manuscript. Zi Huang administrated the project, offered resources and reviewed the manuscript.

Competing interests

The authors declare no competing interests.