

Reviewer 2:
Comments and Suggestions for Authors

This paper develops a deep learning approach to automatically detect and quantify the level of disease in isolated soybean leaves. It is based on YOLOv5 and produces bounding boxes of the infected areas of the leaves (cf. Figure 3) along with a quantification (one of 5 levels of degree of leaf damage). The method is explained and then tested to obtain quantitative measures of performance. As summarized in Fig. 6, the success rate of classification in those 5 classes is above 80%, a good result.

Response: We would like to thank the reviewer for his recommendations and suggestions to improve the manuscript. We have made the requested edits and modifications as suggested by the reviewer.

The training was done on images of collected leaves put flat on a uniform background, making segmentation quite easy and lowering image variation. The validation was done on similar images (with uniform backgrounds) so it is unclear how useful the authors' approach is in natural settings where the leaf is still on the plant and the background is uncontrolled.

Response: Thank you to the reviewer for mentioning this. We are planning to investigate this aspect on our future work in order to develop a model that can quantify the damage degree in natural settings where leaves may overlap. We are planning to plant several thousand of soybeans at the Horticulture Research Center (SIUC) during the summer 2024, use a drone to take pictures, and then use the developed approach to detect leaf damage. The scope of the work was introducing a novel DLM for accurate damage prediction and classification by training a high number of near-field soybean leaf images. The developed model successfully quantifies damage severity, distinguishing healthy/unhealthy leaves and offering a comprehensive solution in precision agriculture. The outcome of this research presents a robust DLM for soybean damage assessment, supporting informed agricultural decisions based on specific damage levels and enhancing crop management and productivity.

Also, they have 30 hyperparameters but give no details about how important those are. (See for example my request in Detailed points to see how performance is affected by threshold values.)

Response: We would like to thank the reviewer for catching this. Setting hyper-parameters in training deep learning models is challenging. Therefore, as we mentioned in the paper (lines 363-365), we used a genetic algorithm that has a built-in evolve function developed in YOLOv5. The evolve function optimized the hyperparameters used in the current study.

I recommend the authors have Supplementary Material to show to what extent the performances change under different runs of the hyperparameter optimizer unless it converges quite generally to the same local optimum.

Response: We used the YOLOv5's built-in 'evolve' function, which is based on a genetic algorithm for parameter optimization. This function converges to the best hyperparameter that we used in training the model. We have summarized the hyperparameters and their obtained values in Table 1.

Separate comments:

1: The English is somewhat clumsy but that is not an impediment to reading and understanding the text. (But do see what you can do to improve the sentences, e.g. using ChatGPT.)

Response: We have improved the reading of the manuscript.

2: You use 5 classes. Wouldn't it have been more natural to just output the proportion of the leaf that is damaged?

Response: We would like to thank the reviewer for this reflection. Unlike traditional programming, which relies on predefined rules and solutions to produce specific outputs. Machine learning (ML) and deep learning (DL) are data-driven approaches that learn from examples/data and generate models that can handle unseen data and scenarios. For instance, instead of writing a program to calculate the exact percentage of leaf damage based on fixed criteria, we trained an ML/DL model to learn from a large dataset of leaf images and labels. Then, we used the model to predict the confidence score of new leaf images belonging to one of the five classes we defined in Table 3.

Detailed points

- line 29: is "\$400/metric ton" referring to the price? (If so, rephrase as the reader is otherwise led to think you are referring to the mass produced but then the numbers do not make sense.) Suggestion: "In 2021/2022, of the production of US soybean, with an average cost of \$400 per metric ton, 21 million metric tons were used ..." Response: This has been updated (lines 29-30).
- line 42: billions of dollars? Response: This has been corrected.
- line 43: put SCN in full text the first time. Response: This has been updated.
- line 131: until the end of vegetative growth? Response: This has been updated.
- lines 138 to 141: duplicated sentence. Response: The duplicated sentence has been deleted.
- line 186: why "first"? In the formula, you want the sum of all a_i , no? Response: Thank you for catching this typo. It should be " a_i ". We corrected as follow: A_{T_i} is the area of defect i -th in the leaf.

- Figure 3: you might want to separate better the different sub-images to make clear you have 4 leaves per sub-image when that is the case.

Response: Figure 3 (updated figure 7) shows a Mosaic image which is automatically generated by YOLOv5, where four or more random images are combined into four tiles with a random ratio with a relative scale of the image and trained through a data loader.

- Figure 4: what is the input of XVIII Conv?

Response: The model Figure 4 (updated see Figure 8) removed the roman numbers that indicate levels in the YOLOv5 model to avoid ambiguity. Conv is an abbreviation for a convolutional neural network, which is used to process leaf images for feature extraction.

- Definition of GloU: you don't say what C is (and mention the definition of "/"). Also put in bold the other defined terms (beyond IoU and GloU).

Response: We updated the paper by defining "C" and we have explained the equation 2 and added the appropriate reference.

- Figure 5: there are only 2930 images so mention somewhere that there are multiple leaves per image. Do you have overlaps between leaves in the pictures or are the leaves always non overlapping, facilitating segmentation into the correct number of leaves?

Response: We have clarified that we have used a non-overlapping set of images.

- In table 2 (the benchmarks for the threshold IoU = 0.5) you should include data for other threshold values so we know how sensitive these performances are. Or better, give results as curves for the different metrics as a function of that threshold.

Response: Table 2 includes a performance metric mAP@0.95, which shows the mean precision for threshold values ranging from 0.5 to 0.95 with 0.05 increments.

- Check the first vs last name of some references and that you have not left out some of the associated authors, while there are other references that are duplicated.

Response: We have used endnote to insert all references.

Comments on the Quality of English Language

Should be improved (but not a major issue).