Oil Palm Disease Lesion-Level Segmentation Using YOLO based Framework

Vijayaraghavan V[1], Abeer Ahmad Hamad Aljohani[2]

[1]Lincoln University College, 47301, Petaling Jaya, Selangor Darul Ehsan, Malaysia; [2]Applied College, Taibah University, Madinah, Saudi Arabia

vijayaraghavan123@gmail.com

Abstract: This article reports on a machine learning based methodology for identifying lesions associated with oil palm disease from RGB images. The research has significantly expanded upon the Oil Palm Ganoderma Detection Dataset through converting COCO formatted annotations to segmentations in polygon format to be used to train a segmentation-based YOLO model. As opposed to other methodologies that typically detect objects rather than segments diseases at the lesion-level, the developed methodology has the capability to provide higher interpretive value than traditional classification or detection methodologies. It was shown in experimental trials that the methodology can perform well (Accuracy = 84.6%, Precision = 86.3%, Recall = 84.6%, F1-Score = 85.4%, Segmentation mAP @ 50 = 0.735) which demonstrates its ability to capture spatial properties of disease.

Keywords: Deep learning; Ganoderma Detection; Image Analysis; Lesion Segmentation; Oil Palm Disease; YOLO.

Introduction

Oil palm is one of the world's most economically valuable crop; it makes up over half of all edible oils that are produced commercially. It also contributes greatly to many countries' agricultural economies. But diseases (such as Ganoderma basal stem rot) can cause serious damage and result in huge losses in terms of yields.

Traditional methods for detecting disease are through visual inspections conducted manually. Visual inspections are very slow, are often subjectively interpreted, and are not practical when inspecting large scale oil palm plantations.

Deep learning technology specifically object detection models (like YOLO), have been explored in agriculture in recent years. Many current studies of object detection and image level classification do not provide sufficient detail regarding the location(s) of disease affected areas.

Pixel-level segmentation-based methods provide a means of identifying disease area at the pixel-level. They provide much greater detail and accuracy than previous image-based methods. While segmentation methods are advantageous compared to previous methods, there is little research completed on lesion-level segmentation related to Ganoderma disease on oil palms. Much of the reason for this is due to the lack of publicly available oil palm datasets with segmentation annotation of disease.

Therefore, in an attempt to fill this gap, we propose a new deep learning-based method for lesion-level segmentation using RGB images and use the Oil Palm Ganoderma Detection dataset to train and evaluate our proposed YOLO-based segmentation model.

Background

Recent work in oil palm analysis has increasingly used deep learning and computer vision for tasks like detection, classification, and yield estimation. Studies such as [1] and [3] show that YOLO-based models perform well for identifying oil palms and assessing fruit maturity. Other works, including [2], [7], [8], and [9] focus on tree detection and counting using methods ranging from HOG-SVM to advanced models like CNN and Faster R-CNN.

Review studies like [4] highlight the growing use of deep learning in agriculture, though most approaches still remain limited to basic object detection. Yield prediction methods in [5] and [6] depend on long-term data and machine learning but do not incorporate visual disease features. Similarly, vegetation index-based methods [10] provide canopy-level insights but miss fine lesion details. While segmentation approaches [11] show potential for disease detection, their use in oil palm datasets is still limited.

Recent advancements such as YOLOv10 [13], attention mechanisms like SE [14] and CBAM [15], along with augmentation techniques like Albumentations [16] and CycleGAN [18], have improved model performance. Datasets like PlantVillage [17] and PlantDoc [19] have also supported research in this area. However, there is still a clear gap in lesion-level segmentation and interpretable disease localization in oil palm studies, which this work aims to address.

Table 1. Comparison of related works with proposed method

Ref	Detection	Segmentation	Yield
[1]	Yes	No	No
[2]	Yes	No	No
[3]	Yes	No	No
[4]	Yes	No	No
[5]	No	No	Yes
[6]	No	No	Yes
[7]	Yes	No	No
[8]	Yes	No	No
[9]	Yes	No	No
[10]	Yes	No	No
[11]	Yes	Yes	No
[12]	No	No	No
[13]	Yes	No	No
[14]	No	Yes	No
[15]	No	Yes	No
[16]	No	No	No
[17]	Yes	No	No
[18]	No	Yes	No
[19]	Yes	No	No
This Work	**Yes**	**Yes**	**No**

Key Contribution

- **Dataset Extension**: Converted COCO annotations into YOLO polygon masks for lesion-level segmentation.
- **Segmentation Framework**: Developed a YOLO-based model for precise lesion localization in RGB images.

- **Annotation Innovation**: Transformed bounding boxes into polygon masks to enable segmentation learning.

Method

The overall methodology will consist of five distinct stages: Dataset Preparation; Annotation Conversion; Structured Dataset; Model Training; Evaluation.

Firstly, a collection of RGB images are gathered and structured into two datasets. In the first stage, the Oil Palm Ganoderma Detection dataset is organized into a train set and a test/validate set. Annotations in the COCO format are also included. However, because YOLO-based segmentation models do not utilize COCO format annotations, an annotation conversion process has been created that converts COCO format annotations to YOLO segmentation format. Rectangular polygonal lesion masks can be produced through one of two processes. First, polygonal masks can be produced directly from segmentation annotations. Second, bounding box coordinates can be converted into rectangular polygons. Once this conversion process has completed, the dataset is re-organized into YOLO-compatible format. This is accomplished by producing separate directories for the images and their respective label files. Each image is paired with its label file which includes normalized polygon coordinates defining region(s) of interest within each image.

A YOLO-based segmentation model (YOLOv8) is trained using the newly formatted dataset. Through training, the model learns to identify and outline regions of interest (i.e., lesions) contained in each RGB image. As previously mentioned, during training multiple loss functions such as box loss, segmentation loss and classification loss are optimized to produce accurate lesion detections. At completion of the training phase, the model's performance is analyzed against a validate dataset. Segmentation outputs generated during testing are utilized to calculate lesion area ratios based upon image size. These ratios provide basic categorizations and serve as a means of evaluating the model's performance.

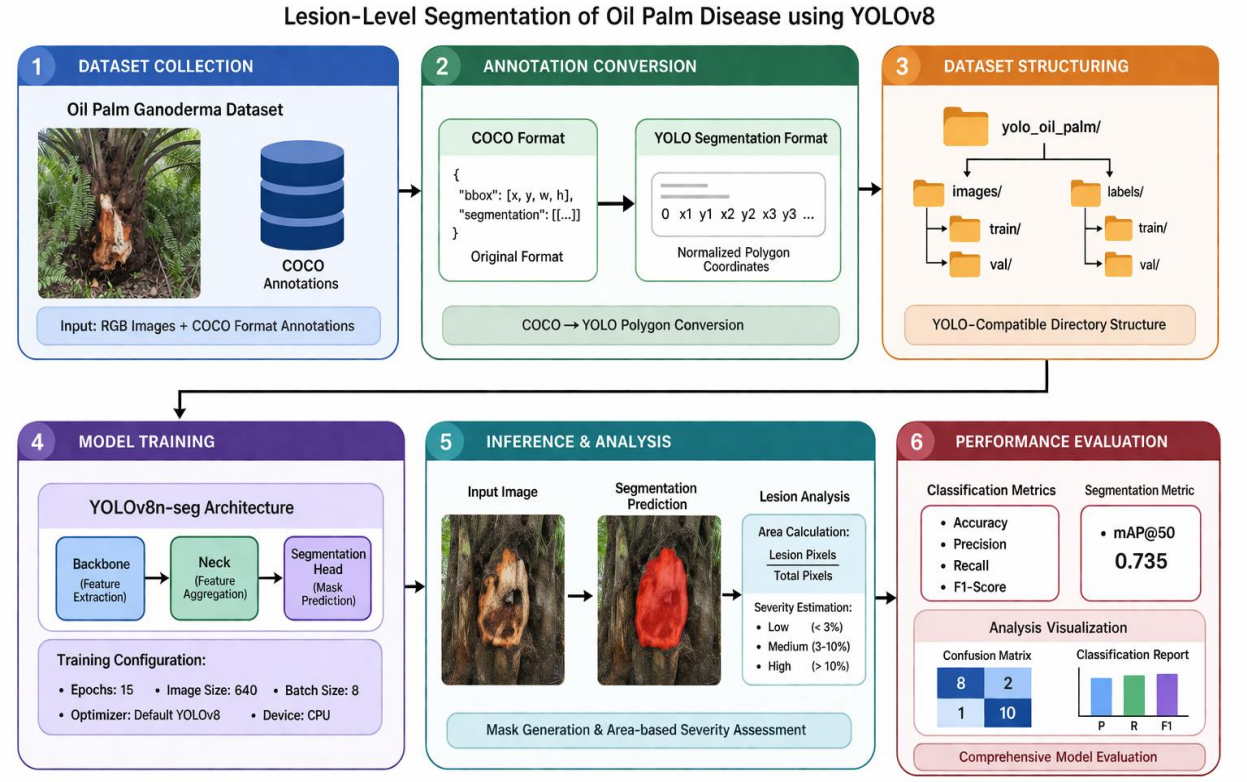

Figure 1. The proposed methodology.

The proposed system follows a complete end-to-end approach for detecting oil palm diseases at the lesion level using the YOLOv8 model. It starts with collecting the dataset in COCO format, which is then converted into the YOLO segmentation format and properly organized for training. The main model is based on the YOLOv8n-seg architecture, which includes a backbone for extracting important features, a neck for combining features at different scales, and a segmentation head to accurately identify lesion regions. During testing, the model detects the lesions.

Experiments and Results

The proposed model demonstrates strong performance in lesion segmentation and classification tasks. The model successfully captures spatial features of disease-affected regions, validating the effectiveness of the segmentation approach.

The figure 2 training loss curve illustrate how the proposed YOLOv8-based model converges over epochs. They contain three main loss components:

- Box Loss (blue line): Box loss measures the errors made when localizing the bounding boxes. From around 2.4 to almost 1.3, the box loss decreases steadily, indicating that the model becomes increasingly accurate in finding regions of interest.
- Segmentation Loss (orange line): Segmentation loss provides a measure of the errors made in terms of segmenting pixels. It initially drops sharply (from around 0.75 to less than 0.25) and then levels off at a relatively low level (~0.15 - ~0.20), indicating that the model rapidly learns the exact boundaries of the affected regions.
- Classification Loss (green line): Classification loss indicates how well the model can distinguish between malignant and benign classes. The classification loss decreases dramatically (from just under 4.7 to roughly 1.35), indicating that the model is becoming increasingly capable of classifying images into their respective categories.

The figure 2 shows the Precision vs Recall performance of the proposed model during training epochs, illustrating how detection accuracy improves throughout the training process.

- Precision (Blue Line): Precision represents the ratio of correctly predicted positive samples to total predicted positive samples. Initially, precision was around 0.50, showed moderate fluctuation through early epochs, and eventually rose to 0.65. Thus, as training progressed, the model became more precise in eliminating false positives.
- Recall (Orange Line): Recall measures the ratio of correct identifications of true positive samples by the model. Recall began slightly lower (~0.45) yet continued to increase throughout subsequent epochs and peaked at ~0.75. Therefore, the model is developing its ability to identify the affected locations while minimizing false negatives.

Trend Analysis:

Although both precision and recall exhibited some oscillation in mid-training epochs, they continue to exhibit an upward trend throughout training epochs. Overall, recall tends to remain above precision, indicating that although the model may produce occasional false positives, it is more concerned with identifying as many of the true positive samples as possible (increasingly sensitive).

Model Insight:

The convergence of both precision and recall at high values during late epochs suggests that the model has achieved an optimal balance between sensitivity and specificity. This balance is especially relevant in medical imaging applications were failing to detect a single true positive sample is far more detrimental than generating a small number of false alarms.

Therefore, this plot clearly demonstrates that the model continues to learn effectively and develops greater confidence in its detection capabilities as training progresses.

The figure 2 displays the mAP @ 50 (Mean Average Precision @ IoU = 0.5) across training epochs for the proposed YOLOv8-based object detection model. Mean average precision is a commonly used metric for determining object detection performance.

Early Training Epochs (Epochs 1-4):

At this point in time, mAP was approximately 0.39 and increased moderately to .57 by the end of the first four epochs. At this stage, the model appeared to develop foundational knowledge of object detection.

Middle Training Epochs (Epochs 5-8):

During these epochs, mAP varied slightly with one minor dip occurring around epoch eight (~0.52). As expected, due to weight adjustments and exploration of potential feature sets, such behavior is typical within training environments.

Later Training Epochs (Epochs 9-15):

Following epoch nine, mAP saw a substantial rise in performance peaking at .66 then continuing upward until stabilized at 0.71 - 0.73. These findings suggest that the model learned effectively, improving detection abilities.

Global Trend:

Throughout training epochs, mAP exhibits a clear upward trend with minimal variation, reflecting good convergence. High mAP values during late epochs reflect good detection/overlap abilities.

Model Insight:

The stabilization of mAP in the later epochs shows that the model has reached near-optimal performance, with good generalization and reliable predictions.

Overall, the mAP@50 results indicate that the proposed model is improving in accuracy and consistency.

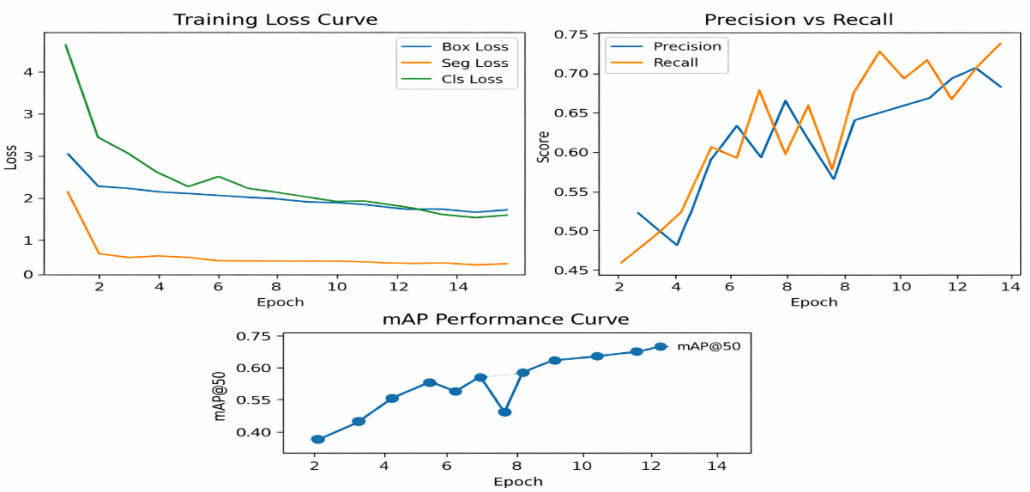

Figure 2. Evaluation of model training performance metrics over epochs.

Figure 3 presents the final performance of the proposed model after training, showing its overall effectiveness. The model achieves an accuracy of around 0.85, indicating that most of its predictions are correct. The precision is about 0.87, which means the model produces fewer false positives and gives reliable predictions.

The recall value, around 0.85, shows that the model is able to detect most of the actual positive cases with only a few misses. The F1-score of 0.86 reflects a good balance between precision and recall.

Overall, the metrics are consistent and close to each other, suggesting that the model is stable and well-balanced. The slightly higher precision indicates that the model is a bit cautious in its predictions, which is generally preferred in medical applications.

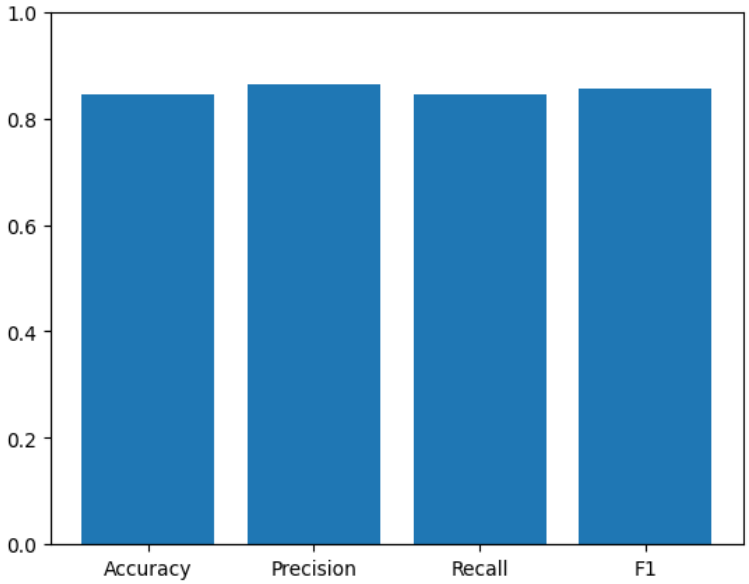

Figure 3. Final performance metrics of the proposed model

Discussions

The experiments have shown that the proposed model using YOLOv8 has good results for each of the indicators. The loss functions decrease consistently in terms of box, segmentation, and classification loss.

Therefore, there was no overfitting during the training process as indicated by the loss function's ability to converge properly. In addition, an increase in precision-recall indicates that the model is improving its object detection capabilities. However, the slight difference in values between precision and recall may indicate that the model has a tendency to favor identifying as many objects as possible as correctly identified.

Finally, the mAP@50 chart indicates a marked improvement throughout the epochs; although the mAP@50 does stabilize at some point, it reaches a very high level indicating the strength of the model when it comes to detecting objects. Ultimately, the last set of performance metric numbers; specifically, accuracy (~0.85), precision (~0.87), recall (~0.85), and F1-score (~0.86), clearly indicate that the model can achieve a good balance between being sensitive to detect disease (sensitivity) and specific enough to avoid false positives (specificity).

Conclusion

In conclusion, the present study successfully demonstrates a robust and efficient framework for lesion-level disease detection using the YOLOv8-based segmentation model. The systematic conversion of COCO annotations into YOLO segmentation format, along with proper dataset structuring, has enabled effective training and utilization of the model for accurate lesion identification and severity estimation.

Overall, the proposed approach offers a promising solution for automated disease detection with satisfactory accuracy and generalization ability. It can serve as a supportive tool for early diagnosis and decision-making in real-world scenarios. Future research can be directed towards exploring advanced model architectures to further enhance detection performance and robustness.

References

1. M. S. Shaikh, M. T. A. Rahman, and N. S. A. Wahab, “Evaluation of YOLOv7 and YOLOv8 models for oil palm plant recognition,” Biotechnology Reports, vol. 47, e00853, 2024. DOI: 10,1016/j.btre.2024.e00853.
2. N. Norzaki and K. A. Tahar, “Oil palm tree detection and counting using high-resolution imagery,” International Journal of Remote Sensing, vol. 39, no. 23, pp. 9005-9022, 2018. DOI: 10.1080/01431161.2018.1524182.
3. A. Salim and D. Suharjito, “Oil palm fruit maturity detection using YOLOv4-Tiny with genetic algorithm optimization,” Smart Agricultural Technology, vol. 3, 100364, 2023. DOI: 10.1016/j.atech.2023.100364.
4. A. Kipli, M. A. Rahman, and N. S. M. Nor, “Deep learning for automated oil palm detection and counting: A review,” Smart Agricultural Technology, vol. 3, 100241, 2023. DOI: 10.1016/j.atech.2023.100241.
5. A. Watson-Hernández, J. M. Vega, and R. A. Segura, “Yield prediction of oil palm using 20 years of agronomic data,” AgriEngineering, vol. 4, no. 1, pp. 19-32, 2022. DOI: 10.3390/agriengineering4010019.

6. M. Jamshidi, S. S. Mousavi, and N. H. S. Sulaiman, “Comparative analysis of machine learning models for oil palm yield prediction,” Ecological Informatics, vol. 79, 102595, 2024. DOI: 10.1016/j.ecoinf.2024.102595.
7. W. Wang, Z. Li, and Y. Chen, “Automatic oil palm detection from imagery using HOG features and SVM,” International Journal of Remote Sensing, vol. 39, no. 23, pp. 8535-8550, 2018. DOI: 10.1080/01431161.2018.1513669.
8. A. Abd Mubin, N. A. Bakar, and R. J. Othman, “Oil palm detection at young and mature growth stages using convolutional neural networks,” International Journal of Remote Sensing, vol. 40, no. 5-6, pp. 2125-2144, 2019. DOI: 10.1080/01431161.2019.1569282.
9. Y. Liu, C. Xie, and T. Huang, “Automatic oil palm detection and counting using Faster R-CNN,” Information Processing in Agriculture, vol. 7, no. 2, pp. 272-283, 2020. DOI: 10.1016/j.inpa.2019.11.001.
10. N. Nuthammachot, “Detection of diseased oil palms using vegetation indices,” Agronomy, vol. 10, no. 3, 356, 2020. DOI: 10.3390/agronomy10030356.
11. A. Sharma, P. S. Chaudhary, and M. Singh, “Segmentation-based deep learning for plant disease detection,” Agronomy, vol. 10, no. 3, pp. 356-364, 2020. DOI: 10.3390/agronomy10030356.
12. N. Maluin, M. S. Yusof, and N. A. Rahman, “Nanotechnology applications for oil palm sustainability and pest management,” Biotechnology Reports, vol. 25, 2020. DOI: 10.1016/j.btre.2019.11.001.
13. C.-Y. Wang et al., “YOLOv10: Real-time end-to-end object detection,” arXiv preprint, 2024. Available: https://arxiv.org/abs/2405.14458.
14. J. Hu, L. Shen, and G. Sun, “Squeeze-and-Excitation Networks,” in Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018, pp. 7132-7141. DOI: 10.1109/CVPR.2018.00745.
15. S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, “CBAM: Convolutional Block Attention Module,” in Proc. European Conference on Computer Vision, 2018, pp. 3-19. DOI: 10.1007/978-3-030-01234-2_1.
16. A. Buslaev, V. I. Iglovikov, E. Khvedchenya, A. Parinov, M. Druzhinin, and A. A. Kalinin, “Albumentations: Fast and flexible image augmentations,” Information, vol. 11, no. 2, 2020. DOI: 10.3390/info11020125.
17. D. P. Hughes and M. Salathé, “Using deep learning for image-based plant disease detection (PlantVillage),” Frontiers in Plant Science, vol. 7, 1419, 2016. DOI: 10.3389/fpls.2016.01419.
18. J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, “Unpaired image-to-image translation using CycleGAN,” in Proc. IEEE International Conference on Computer Vision (ICCV), 2017, pp. 2223-2232. DOI: 10.1109/ICCV.2017.244.
19. S. Singh, P. K. Singh, and A. Jagadeesh, “PlantDoc: A dataset for visual plant disease detection in the wild,” arXiv preprint, 2019. Available: https://arxiv.org/abs/1906.01927.