Automated Detection of Pineapple Plants in Drone-Captured Aerial Imagery for Precision Agriculture

Matheus Arroyo de Lima*, Thiago Mantovani Martins*,
Vitor Matheus Soares Siqueira De Almeida†, Rafael Buosi Jardim†,
Victor Hugo Zani‡, Luiza Hoehl Loureiro Alves Barbosa§
*Barão de Mauá University Center, Ribeirão Preto, Brazil
Emails: matheusarroyodelima@gmail.com, thiagomantovanimartins@outlook.com
†REGES University, Ribeirão Preto, Brazil
Emails: matheus2famas@gmail.com, rbjardim2005@gmail.com
‡Ribeirão Preto Medical School, University of São Paulo, Ribeirão Preto, Brazil
Email: victor_zani@usp.br
§Altamap, Ribeirão Preto, Brazil
Email: luiza.hoehl@gmail.com

***Abstract*—Pineapple harvesting remains largely manual due to scattered planting patterns and complex fruit structure. This study presents a method for detecting pineapple plants in large orthomosaic images using the Slicing Aided Hyper Inference (SAHI) technique combined with the YOLOv8 segmentation model. SAHI divides large images into smaller patches, enabling accurate detection. A dataset of 867 training and 97 validation images from two orthomosaics was used, with the model achieving 93% precision and 88% accuracy. Despite high precision, challenges with false negatives suggest future improvements. This approach shows promise for automating pineapple harvesting and improving agricultural efficiency.**

***Index Terms*—Pineapple detection, deep learning, YOLOv8, SAHI, agricultural automation, orthomosaic images, precision agriculture.**

I. Introduction

Pineapple, a tropical and subtropical fruit known for its distinct flavor, ranks fourth in global fruit production. The international pineapple market is robust, with Costa Rica leading exports, shipping about 3 million tons in 2023, which represents approximately 50% of global exports [1]. Brazil is also a major exporter, with exports totaling around 150,000 tons in the same year [2]. Additionally, pineapple consumption has shown a growing trend, driven by its popularity in international markets and the demand for derived products such as juices and preserves. In 2023, global pineapple consumption was estimated at about 24 million tons, reflecting a stable and increasing demand for this nutritious fruit [3].

Despite Brazil's role in pineapple production, it faces challenges in maintaining competitiveness in the global market. Alves [4] analyzed Brazil's competitive performance from 1997 to 2006, revealing that while the state of Ceará demonstrated a competitive advantage in pineapple exports, Brazil as a whole did not achieve the same level of performance. The competitiveness of Brazilian pineapple exports has been influenced by a variety of factors, including domestic policies, market access, and global demand.

On a global scale, Indonesia is recognized as a major exporter of canned pineapple, maintaining a strong competitive advantage in the international market. Wiranthi and Mubarok [5] demonstrated that Indonesia's exports of canned pineapple positioned the country as a "rising star" in key markets such as the United States, Spain, and China. Their study highlighted the importance of improving the quality and quantity of production to maintain this competitive edge.

Technological advancements in agriculture have also played a crucial role in improving the efficiency and profitability of pineapple farming. For example, Kleemann et al. [6] examined the adoption of organic certification by pineapple farmers in Ghana, demonstrating that organic-certified farming achieved a significantly higher return on investment (ROI) compared to GlobalGAP-certified farming, primarily due to the price premium in organic markets. This trend towards organic and certified farming has opened new market opportunities for pineapple producers in developing countries.

The development of mechanized harvesting and fruit detection technologies has been essential to enhancing productivity and reducing labor dependency in pineapple farming. The use of deep learning models, such as YOLOv5, has facilitated the detection of ripening pineapples, enabling farmers to optimize harvest timing and improve operational efficiency [7]. Moreover, Li et al. [8] developed a deep learning-based approach to classify pineapple quality, combining VGG16, LSTM, and GAN models to accurately predict maturity and improve grading efficiency.

In summary, the global pineapple market continues to expand, driven by increasing demand and technological innovations in agricultural practices. However, challenges remain, particularly in maintaining competitiveness and optimizing production processes in a highly competitive export market. This study aims to build on these advancements by exploring novel techniques in pineapple plant detection, leveraging the Slicing Aided Hyper Inference (SAHI) technique and YOLOv8 segmentation model to improve detection accuracy and facilitate large-scale agricultural automation.

II. Related Works

Recent advances in object detection and machine learning techniques have significantly contributed to the automation of pineapple plant detection, counting, classification, and harvesting. Several studies have explored different methods and models to tackle the challenges of detecting pineapples in complex environments.

Lai et al. (2023) proposed an improved version of YOLOv7 to accurately detect pineapples and classify their maturity levels in complex field environments. The model incorporated the SimAM attention mechanism and the soft-NMS algorithm to improve feature extraction and handle occlusion and overlaps in the images. The model achieved a mean average precision (mAP) of 95.82% and a recall of 89.83%, outperforming the original YOLOv7. This method highlights the potential for improving the vision systems in robotic harvesting [9].

Sharma et al. (2022) developed a deep learning approach based on YOLOv5 to detect the ripening stages of pineapples. The model was designed to assist in optimizing harvest timing and reducing human resource dependency in large pineapple fields. The YOLOv5 model achieved over 95% accuracy in detecting ripening, demonstrating faster training times and superior performance compared to previous versions like YOLOv4. This approach offers promising applications for real-time fruit detection in agricultural settings [7].

Wan Nurazwin et al. (2022) developed an automated method to detect and count pineapple crowns using machine learning classifiers such as ANN, SVM, and KNN. The study used high-resolution aerial images captured by UAVs, pre-processed and segmented to extract features like color, shape, and texture. The best performance was achieved using the ANN-GDX classifier, with an accuracy of 94.4%. This method showed potential for automating fruit detection and yield estimation in large-scale fields [10].

Li et al. (2023) proposed a deep learning model that integrates VGG16, LSTM, and GAN techniques for pineapple quality classification and maturity prediction. The model was trained on a dataset of 1,431 MD2 pineapple images and demonstrated high precision, recall, and mAP scores. This approach provides a reliable solution for grading pineapples in commercial settings, aiding in optimizing harvest times and maximizing crop yield [8].

Anh et al. (2020) introduced a robotic system for the autonomous harvesting of pineapples, integrating machine vision and robotic manipulators. The YOLOv3 model was used to detect and recognize pineapples, achieving a mAP of 90.82%. The robotic system, equipped with custom end-effectors, demonstrated a success rate of 95.55% for pineapple harvesting and an average harvest time of 12 seconds per fruit. This study underscores the potential of combining machine vision and robotics for efficient fruit harvesting in agricultural fields [11].

Rahutomo et al. (2019) implemented an AI-based web application for counting pineapple objects over large areas using aerial images. The application was developed using Python and Flask, with AI models trained using Keras-RetinaNet. This system achieved accurate object counting, helping optimize the use of resources such as water, fertilizers, and packaging materials. The researchers highlighted the benefits of integrating AI in agricultural applications [12].

Hobbs et al. (2021) developed a deep learning-based density-estimation model for counting pineapple inflorescences in large fields. Using a U-net backbone, the model achieved a mean absolute error (MAE) of 11.5 and a mean absolute percentage deviation (MAPD) of 6.37%. The model's efficiency in processing large-scale fields, with over 1.6 million flowering plants, was a key advantage. This approach improves field management by detecting vegetative and failed forcing areas, optimizing the harvesting process [13].

Wan et al. (2023) used a combination of Faster R-CNN and Mask R-CNN to detect and segment mature pineapple fruits in UAV-captured imagery. Their study focused on predicting yield by detecting fruits that were either covered with plastic hats or nets, which serve as sun protection. Their approach achieved high precision, with an F1-score of 0.849 and an overall classification accuracy of 97.46%. However, this study differs from ours in its focus on fruit detection rather than plant counting, as well as its reliance on zonal statistics for yield estimation. While their approach demonstrates the effectiveness of using multiple detection models, the method requires specific conditions (e.g., the use of protective coverings) that are not applicable to our dataset of open-field pineapple plants [16].

These studies demonstrate the effectiveness of machine learning and deep learning approaches in solving key challenges in pineapple detection, classification, and yield estimation. Each method contributes to the broader goal of improving agricultural efficiency and accuracy, particularly through the automation of traditionally manual tasks.

III. Methodology

A. Dataset Preparation

The dataset utilized in this study was acquired by the authors through aerial images captured with a DJI Mavic 2 Pro drone at an altitude of 70 meters. The images were taken in pineapple fields located in the region of Pital, Costa Rica (Latitude: 10.526293°, Longitude: -84.182759°). These orthomosaics serve as the primary source for training and validating the YOLOv8 segmentation model in detecting individual pineapple plants. The original images have a resolution of 27,199 x 26,808 pixels, offering a comprehensive view of the surveyed area.

The dataset used in this study is proprietary and was provided by the agricultural mapping company Altamap. Any inquiries regarding access to this dataset should be directed to Altamap.

To create the training and validation datasets, one of the orthomosaic images was used exclusively for training, while the other was reserved for validation. This approach ensures that the model is validated on a separate image, preventing overfitting and providing a robust evaluation of its performance on unseen data.

Due to the large size of these orthomosaic images, directly training and performing inference on the entire image would be computationally expensive and could result in inaccuracies when detecting small objects, such as individual pineapple plants. To address this, the images were divided into smaller patches using the **Slicing Aided Hyper Inference (SAHI)** technique. The SAHI method was employed to divide the large orthomosaic images into smaller patches of 640 x 640 pixels, with a 20% overlap between adjacent patches to ensure that no part of a pineapple plant was truncated at the edges of a patch.

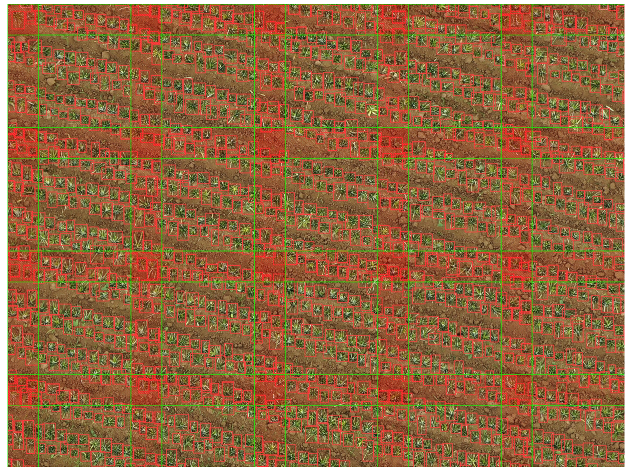

Fig. 1. Grid over a section of the original orthomosaic image, showing where the image is sliced into smaller patches, including annotations of the detected pineapple plants.

B. Slicing Aided Hyper Inference (SAHI)

SAHI is a technique designed to improve the efficiency and accuracy of object detection models when working with large-scale images. Traditional object detection methods often struggle with large images, especially when objects are small in relation to the image size. SAHI mitigates this by slicing the large images into smaller, manageable patches, allowing the model to process each patch individually [14].

This method ensures that the detection model, such as YOLOv8, can focus on smaller sections of the image where objects, such as pineapple plants, are more likely to be detected. After processing, the results from each slice are merged back into the original large image, maintaining spatial consistency and improving detection accuracy.

C. Slicing Process

The SAHI technique was employed to divide the large orthomosaic images into smaller patches of 640 x 640 pixels, with a 20% overlap between adjacent patches in both the x and y directions. This overlap ensures that no part of a pineapple plant is truncated at the edges of a patch, which could result in missed or incomplete detections. Through this process, a total of 867 smaller images were generated for training, while 97 images were reserved for validation.

Figure 1 illustrates how the original orthomosaic image is divided into smaller patches using a grid. The grid ensures that each 640x640 image contains complete portions of the pineapple plants, preventing truncation at the edges due to the 20% overlap.

During the slicing process, some patches were discarded as they did not contain any areas with pineapple plants, thus ensuring that the dataset used for model training and validation focuses solely on relevant regions. The rationale behind this decision is that the model learns to distinguish between background and the object of interest from the images that contain the pineapple plants. Including patches without any plants could lead to dataset imbalance, as a significant portion of the area in the orthomosaic does not contain any objects, potentially diluting the model's ability to focus on the relevant detection task.

Future work may explore the impact of incorporating such images without plants into the training process to assess their effect on the model's robustness. However, this was not applied in the current work to maintain the dataset's focus on meaningful data and avoid overfitting the model to non-informative background areas.

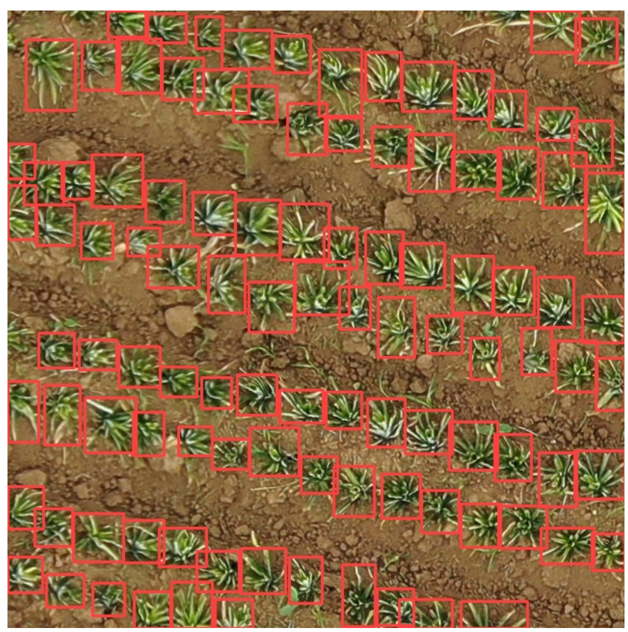

Fig. 2. Example of a 640x640 sliced patch from the original image, showing the annotated pineapple plants.

Figure 2 presents an example of a 640x640 pixel image slice generated during the process, with annotations of the detected pineapple plants. These smaller images form the basis for training the YOLOv8x segmentation model.

D. Model Training and Hyperparameters

For training, the **YOLOv8x segmentation** model was selected due to its state-of-the-art performance in object detection tasks [15], particularly for handling dense and small objects in large-scale images. The YOLOv8x variant, known for its larger architecture and ability to process high-resolution images, was ideally suited for detecting pineapple plants in the orthomosaic patches used in this study.

The model was trained on the sliced patches, each of size 640x640 pixels, for a total of **100 epochs**. The Adam optimizer was used, which is a popular choice in deep learning due to its adaptive learning rate and robust performance across a wide variety of tasks. A batch size of 8 was employed, which was the maximum allowable size on the GPU used for training—a Tesla V100 with 16GB of VRAM. This batch size provides a good balance between memory efficiency and computational speed, ensuring stable and effective learning throughout the training process.

Regarding the loss functions, YOLOv8x applies two standard loss components for object detection tasks. The loss for bounding boxes was computed using Complete Intersection over Union (CIoU) loss, which optimizes the localization of the bounding boxes by considering the distance between the predicted and ground truth boxes, along with their size and aspect ratio. For classification, the Cross Entropy loss was used, which is well-suited for multi-class classification tasks by comparing the predicted class probabilities with the ground truth labels.

While the results obtained in this study are promising, it is important to note that they are based on a single training run. A limitation of this work is the absence of cross-validation, which would have provided more robust insights into the model's generalization performance. Due to the relatively small size of the dataset and time constraints, cross-validation was not performed. Future work will address this by employing cross-validation techniques to better assess the model's ability to generalize across different data distributions.

E. Data Augmentation and Preprocessing

No explicit normalization was applied to the dataset beyond the inherent preprocessing steps of the YOLOv8 framework. However, data augmentation techniques were employed to enhance the diversity of the training data. These included random horizontal flipping, scaling, and color jittering, which are common augmentation techniques used to improve the model's robustness against variations in lighting and plant orientation.

These augmentations aimed to simulate real-world variability, such as changes in sunlight, camera angles, and slight variations in the plants' appearance across different sections of the field. This helped prevent overfitting to the limited training data and improved the model's generalization.

F. Training and Validation Curves

To illustrate the model's learning process, we include a plot of the training and validation loss curves (Figure 3). These curves show the evolution of the loss over the course of 100 epochs. The graph indicates that the model's loss steadily decreased on both the training and validation datasets, though with some variance in the training set on final epochs, suggesting that further training would cause overfitting. The final model's performance metrics were obtained at the epoch with the lowest validation loss, ensuring that the model did not overfit to the training data.

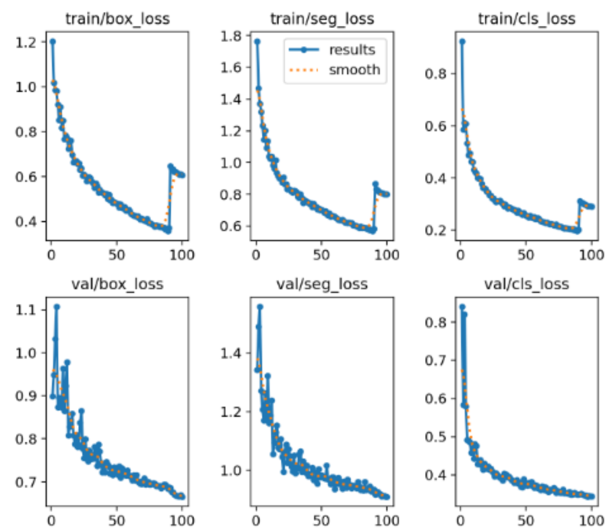

Fig. 3. Training and validation loss curves over 100 epochs, showing the evolution of the YOLOv8 model's performance.

G. Merging Detections

After inference on all slices, the detected bounding boxes were merged to form the final detections on the original large image. The SAHI technique ensures that overlapping detections are properly merged, taking into account the overlap in the slicing process. The merging process involves calculating the IoU of overlapping boxes and combining them into a single detection when the IoU exceeds a predefined threshold. This merging is crucial to avoid duplicate detections and to ensure that the final output accurately reflects the detected objects in the full-resolution image.

H. Model Validation

The validation of the model was performed using the same slicing technique applied to the validation set, consisting of 97 images. The performance was assessed using multiple key metrics to evaluate how well the YOLOv8x segmentation model detected pineapple plants. The following metrics were used:

- **True Positives (TP)**: The number of correctly detected pineapple plants that match the ground truth labels.

- **False Positives (FP)**: The number of incorrect detections where the model mistakenly identified an object as a pineapple plant.
- **False Negatives (FN)**: The number of missed detections where the model failed to identify a pineapple plant that was present in the ground truth.
- **Accuracy**: The percentage of correctly identified pineapple plants out of the total number of plants in the validation set. It provides an overall measure of the model's performance, reflecting both true positives and true negatives.
- **Precision**: The ratio of true positive detections to the sum of true positives and false positives. Precision measures the model's ability to avoid false positive detections and indicates how reliable the detected plants are.

For the validation, a **confidence threshold** of 0.3 was applied, meaning that the model only considered a detection valid if its confidence score was above this threshold. Additionally, an IoU threshold of **0.5** was used to assess the overlap between predicted bounding boxes and the ground truth. If the IoU between the predicted and true bounding boxes was below 0.5, the detection was classified as a false positive.

IV. Results and Discussion

The results from the validation set, after applying the SAHI technique for slicing and merging the detections, demonstrate the effectiveness of the YOLOv8x segmentation model in detecting pineapple plants from large orthomosaic images. The model's performance was evaluated using the key metrics outlined in the previous sections. The final metrics obtained from the validation set are as follows:

TABLE I
Performance Metrics of the YOLOv8x Model on the Validation Set

Metric	Value
True Positives (TP)	30,029
False Positives (FP)	2,223
False Negatives (FN)	1,766
Accuracy	88%
Precision	93%

These results indicate that the model achieves a high precision, meaning that most of the detected pineapple plants are correct, with relatively few false positives. The accuracy, at 88%, reflects the model's ability to correctly identify pineapple plants out of the total number of plants, though some false negatives suggest that a portion of the plants were missed during detection.

A. Performance Evaluation

The high precision of 93% suggests that the model performs well in terms of avoiding false positives, which is critical in agricultural applications where overestimating the number of plants could lead to incorrect yield predictions or inefficient resource allocation. However, the presence of 1,766 false negatives indicates that there are still challenges in ensuring that all pineapple plants in the validation set are detected. This could be due to factors such as occlusion, plant clustering, or the resolution of the patches used during inference.

B. Model as an Estimation Tool

Given the results, the YOLOv8x segmentation model demonstrates the potential to be used as a tool for estimating the number of pineapple plants in large agricultural fields. While it does not achieve perfect detection accuracy, the model can still be effectively applied in scenarios where a general estimate is sufficient, such as yield prediction, field management, and resource allocation. The high precision ensures that the plants detected are mostly correct, reducing the likelihood of overestimation.

However, due to the number of false negatives, the model may underestimate the total number of plants. This suggests that while it can provide valuable insights for large-scale field management, further refinement would be needed for tasks that require near-perfect detection, such as automated harvesting or precise yield estimation.

C. Comparison with Previous Methods

While many studies in the field of agricultural automation focus on object detection in aerial imagery, we did not find any works that specifically address the task of detecting pineapple plants using orthomosaic images. Many existing methods leverage geospatial factors for detection and classification, which differs from our approach that focuses solely on deep learning-based detection without additional geospatial inputs. For instance, works such as [16] utilize object detection models, such as Faster R-CNN, but they focus on detecting and segmenting mature pineapple fruits under specific conditions (e.g., covered with plastic hats or nets). Although their approach shares some similarities with our work, the primary goal and dataset differ significantly, making direct comparisons difficult.

However, we acknowledge that our results, particularly regarding false negatives, indicate that further improvements are necessary to achieve the same levels of precision found in related works. This highlights the need for continued optimization and the exploration of more advanced segmentation and detection techniques in future research.

D. Challenges and Future Improvements

Future research should focus on optimizing the use of drones for capturing orthomosaic images by investigating the maximum altitude at which images can be captured without significantly compromising detection accuracy. This would help improve the efficiency of drone flights, allowing for broader coverage in less time while maintaining reliable results.

Additionally, it is important to explore the impact of plant maturity on detection accuracy. As pineapple plants grow, their leaves can become denser and overlap, potentially complicating the counting process. Understanding the most advanced growth stage at which accurate plant detection is still possible will be crucial for ensuring the model's robustness across different field conditions.

Finally, given that each pineapple plant produces a single fruit, the ability to count plants accurately provides a direct method for estimating total crop production. This makes plant detection not only useful for field management but also a powerful tool for yield estimation, supporting more efficient planning and resource allocation in agricultural operations.

V. Conclusion

In conclusion, the results of this study show that the YOLOv8x segmentation model, when combined with the SAHI technique, can effectively detect pineapple plants in high-resolution orthomosaic images with high precision. The model holds potential as an estimation tool in agricultural settings, where exact detection is less critical, but general trends and plant densities are needed for decision-making. Further improvements could enhance the model's performance, particularly in reducing false negatives, making it a more robust solution for precision agriculture.

References

[1] U.S. Department of Agriculture (USDA), Pineapple Production and Export Data, 2023.

[2] ABAPA - Associação Brasileira dos Produtores de Abacaxi, Relatório de Exportação de Abacaxi 2023, 2023.

[3] FAO - Food and Agriculture Organization. World Programme for the Census of Agriculture. 2023.

[4] L. M. M. Alves, Uma análise da competitividade das exportações de fruticultura cearense e brasileira: O caso do abacaxi e da melancia, Dissertação (Mestrado) - Universidade Federal do Ceará, Centro de Ciências Agrárias, Fortaleza-CE, 2009.

[5] P. E. Wiranthi and F. Mubarok, Competitiveness and the Factors Affecting Export of the Indonesia Canned Pineapple in the World and the Destination Countries, KnE Life Sciences, vol. 2, no. 6, pp. 339-352, 2017. DOI: 10.18502/kls.v2i6.1056.

[6] L. Kleemann, A. Abdulai, and M. Buss, Certification and Access to Export Markets: Adoption and Return on Investment of Organic-Certified Pineapple Farming in Ghana, World Development, vol. 64, pp. 79-92, 2014. DOI: https://doi.org/10.1016/j.worlddev.2014.05.005.

[7] A. K. Sharma, H. H. C. Nguyen, T. X. Bui, S. Bhardwa, and D. V. Thang, An Approach to Ripening of Pineapple Fruit with Model Yolo v5, in *2022 IEEE 7th International Conference for Convergence in Technology (I2CT)*, 2022, pp. 1-5. DOI: 10.1109/I2CT54291.2022.9824067.

[8] Z. Li, Y.-W. Chong, M. N. Ab Wahab, G.-K. Lim, and R. Dawood, Classification and Prediction of Pineapple Quality using Deep Learning, in *2023 4th International Conference on Big Data Analytics and Practices (IBDAP)*, 2023, pp. 1-6. DOI: 10.1109/IBDAP58581.2023.10271948.

[9] Y. Lai, R. Ma, Y. Chen, T. Wan, R. Jiao, and H. He, A Pineapple Target Detection Method in a Field Environment Based on Improved YOLOv7, *Applied Sciences*, vol. 13, no. 4, Article 2691, 2023. [Online]. Available: https://www.mdpi.com/2076-3417/13/4/2691. DOI: 10.3390/app13042691.

[10] R. Wan Nurazwin Syazwani, H. Muhammad Asraf, M.A. Megat Syahirul Amin, and K.A. Nur Dalila, Automated image identification, detection and fruit counting of top-view pineapple crown using machine learning, *Alexandria Engineering Journal*, vol. 61, no. 2, pp. 1265-1276, 2022. DOI: https://doi.org/10.1016/j.aej.2021.06.053.

[11] N. P. T. Anh, S. Hoang, D. Van Tai, and B. L. C. Quoc, Developing Robotic System for Harvesting Pineapples, in *2020 International Conference on Advanced Mechatronic Systems (ICAMechS)*, 2020, pp. 39-44. DOI: 10.1109/ICAMechS49982.2020.9310079.

[12] R. Rahutomo, A. S. Perbangsa, Y. Lie, T. W. Cenggoro, and B. Pardamean, Artificial Intelligence Model Implementation in Web-Based Application for Pineapple Object Counting, in *2019 International Conference on Information Management and Technology (ICIMTech)*, vol. 1, pp. 525-530. DOI: 10.1109/ICIMTech.2019.8843741.

[13] J. Hobbs, P. Prakash, R. Paull, H. Hovhannisyan, B. Markowicz, and G. Rose, Large-Scale Counting and Localization of Pineapple Inflorescence Through Deep Density-Estimation, *Frontiers in Plant Science*, vol. 11, Article 599705, Jan. 2021. DOI: 10.3389/fpls.2020.599705.

[14] Akyon, F. C., Altinuc, S. O., Temizel, A., Slicing Aided Hyper Inference and Fine-tuning for Small Object Detection, *2022 IEEE International Conference on Image Processing (ICIP)*, 2022, pp. 966-970, doi: 10.1109/ICIP46576.2022.9897990.

[15] G. Jocher, A. Chaurasia, and J. Qiu, Ultralytics YOLO, version 8.0.0, Jan. 2023. [Online]. Available: https://github.com/ultralytics/ultralytics. [Accessed: Sep. 08, 2024].

[16] Z. Wan, H. Wang, H. Li, L. Zhao, and Z. Ma, Pineapples Detection and Segmentation Based on Faster and Mask R-CNN in UAV Imagery, *Remote Sensing*, vol. 15, no. 3, pp. 739-755, 2023. DOI: 10.3390/rs15030739.