Deep Learning for Agricultural Crime Prevention: YOLOv8-x for Real-Time Durian Theft Detection in Low-Light Conditions

I Nyoman Eddy Indrayana[1], Gde Brahupadhya Subiksa[2], Putu Indah Ciptayani[3], I Wayan Suasnawa[4], and I Putu Sutawinaya[5]

[1,2,3,4] Information Technology Department, Politeknik Negeri Bali, Bali, Indonesia
[5] Electrical Engineering Department, Politeknik Negeri Bali, Bali, Indonesia
eddyindrayana@pnb.ac.id.

Abstract. Nocturnal durian theft poses a significant challenge for farmers, leading to substantial economic losses. This research proposes a deep learning approach for night-time durian theft detection, leveraging the capabilities of the YOLOv8 object detection network. A unique dataset was collected in a nocturnal environment, simulating actual theft scenarios, including actors carrying tools like sickles and sacks, with their faces obscured by cloth and hats. This presented complex detection challenges. To optimize detection performance under these demanding conditions, various YOLOv8 variants (n, s, m, l, and x) were extensively evaluated. Experimental results consistently show that YOLOv8-x achieved the best detection performance, with the highest mean Average Precision (mAP) compared to other variants. These findings highlight the potential of YOLOv8-x as an effective and robust solution for preventing nocturnal durian theft, contributing to enhanced agricultural security and mitigating losses for farmers. This study paves the way for developing computer vision-based early warning systems to protect agricultural assets.

Keywords: Deep Learning, Durian Theft Detection, Yolov8

1 Introduction

Durian, sometimes regarded as the "King of Fruits", possesses considerable economic importance in Southeast Asian nations, especially in Indonesia, where its cultivation significantly bolsters local agricultural economies (Barakat et al., 2023). The elevated demand and premium market price of durian, however, make its orchards particularly susceptible to theft, especially during the peak harvest season. Nocturnal theft poses a significant challenge, as reduced visibility at night provides optimal concealment for offenders, resulting in considerable financial losses for farmers and impacting regional agricultural stability (Indrayana et al., 2024). Conventional security systems, including human monitoring and traditional fencing, often fall short for extensive orchards, as they are labor-intensive and prone to human errors or supervision issues in challenging, low-light conditions (Patel et al., 2022). The intrinsic limitations of these approaches

A. A. N. G. Sapteka et al. (eds.), *Proceedings of the International Conference on Sustainable Green Tourism Applied Science - Engineering Applied Science 2025 (ICOSTAS-EAS 2025)*, Advances in Engineering Research 280,
https://doi.org/10.2991/978-94-6463-878-3_16

underscore the urgent need for sophisticated, automated solutions to protect agricultural assets.

The swift advancement of computer vision and deep learning has transformed several areas, including surveillance and security technology (Sharma et al., 2023). Object recognition, a fundamental function of deep learning, has shown exceptional skill in accurately detecting and localizing particular objects within intricate visual input streams. Contemporary object recognition models, especially those utilizing Convolutional Neural Networks (CNNs), are capable of discerning complex patterns and traits that signify suspicious actions, even under suboptimal environmental conditions (Dai et al., 2021; Kim et al., 2017). Architectures such as the You Only Look Once (YOLO) (Indrayana et al., 2024; Lee & Hwang, 2022; Redmon & Farhadi, 2017) series have become leaders in real-time object detection (Redmon et al., 2016) owing to their single-pass prediction efficiency and remarkable accuracy (Han et al., 2021; Li et al., 2022) rendering them exceptionally appropriate for practical security applications where prompt alarms are essential.

Notwithstanding significant progress in deep learning-based object recognition, a critical research need remains in agricultural crime prevention, particularly with nocturnal theft under adverse visual circumstances. The majority of current studies in agricultural monitoring predominantly concentrate on daylight activities, including crop health evaluation and pest identification. Research specifically focused on human intruder identification in low-light conditions, particularly when individuals employ concealment methods such as facial coverings or carry implements, is limited. The absence of specialized solutions underscores a critical demand for a strong, efficient, and highly precise automated system designed to address the distinct issues of nocturnal agricultural theft, therefore enhancing or supplanting conventional, less effective security methods.

This study seeks to fill a significant gap by creating and thoroughly assessing a deep learning system for real-time detection of durian theft at night, employing the advanced YOLOv8 object identification framework (Safaldin et al., 2024; Sohan et al., 2024). Our primary purpose is to systematically analyze the performance of multiple YOLOv8 variants, namely, n (Nano), s (Small), m (Medium), l (Large), and x (Extra-Large), to determine which version attains the highest mean Average Precision (mAP) and optimal Recall. The primary objective is to ascertain the most efficient YOLOv8 setup for accurately detecting masked individuals wielding particular instruments in difficult, low-light orchard settings. This project aims to illustrate the feasibility and enhanced effectiveness of sophisticated deep learning models in improving agricultural security. The originality of this research is evident in numerous critical aspects. First, it entails the development of a distinctive, bespoke dataset meticulously designed to simulate authentic nocturnal durian theft situations, incorporating individuals with obscured identities and equipped with diverse theft-related implements, thereby directly confronting the inadequacies of publicly accessible datasets for this specialized purpose. Second, it presents a thorough comparison study of all popular YOLOv8 variations (n, s, m, l, x), providing empirical data on their specific performance trade-offs in a demanding low-light, obscured-feature identification task. This study identifies the optimal YOLOv8 variant, establishing a vital foundation for the development of effective, high-performance, computer vision-based early warning systems aimed at safeguarding high-value agricultural products from nocturnal criminal activities, thus significantly contributing to smart farming security solutions.

2 Methodology

2.1 Research steps

The initial phase involves compiling a specialized dataset by gathering photos and videos depicting nocturnal durian-stealing activity. The data is subsequently tagged with bounding boxes to indicate the presence of thieves, whose identities may include individuals such as the ill, campers, or those using face masks. The dataset is subsequently partitioned into three segments: training data (train), validation data (val), and testing data (test), to ensure an effective and objective training and assessment procedure. The researchers selected many variations of the YOLOv8 model: YOLOv8n (Bai et al., 2023), YOLOv8s (Tahir et al., 2024), YOLOv8m, YOLOv8l, and YOLOv8x. Each version is subsequently trained by configuring hyperparameters, including the number of epochs, batch size, and learning rate. The training process is conducted individually for each variant to evaluate the performance of each model. Post-training, the model undergoes evaluation using the validation and test datasets by computing essential metrics, such as mean Average Precision (mAP@0.5 and mAP@0.5:0.95), precision, and recall. The evaluation outcomes of each version are then thoroughly compared to determine the optimal option. This research is crucial to ensure the system developed achieves maximum performance and reliably detects durian theft during nighttime.

2.2 Yolov8 Variants

In the domain of computer vision, particularly within deep learning-based object recognition, model efficiency and accuracy are two critical metrics that frequently exhibit an inverse relationship. Ultralytics, the creator of the YOLOv8 series, offers many model variants: Nano (n), Small (s), Medium (m), Large (l), and Extra-Large (x) that adeptly manage this trade-off. The primary distinction among these variations is not due to entirely disparate core architectures, but rather the methodical scaling of network components via two essential parameters in their YAML configuration files: depth_multiple and width_multiple. A comprehensive comprehension of how these characteristics influence the network's internal architecture is essential for selecting the appropriate model for your individual application requirements, such as the nocturnal durian theft detection you are investigating. Essentially, depth_multiple governs the network's depth, which correlates directly with the frequency of critical functional block repeats, such as the C2f module, a significant innovation in YOLOv8 (Zhu et al., 2024). An increased depth_multiple value results in a greater number of recurrent layers in the network, enabling the model to extract more intricate and hierarchical features from the input data. Conversely, width_multiple governs the "breadth" of the network, as seen by the quantity of channels (or filters) in each convolutional layer.

Augmenting the width multiple enables each layer to handle more intricate and varied visual information, leading to enhanced and more distinct feature representations. The astute integration of these two scaling factors is fundamental to the diverse performance profiles of each YOLOv8 variation.

The YOLOv8 variants (n, s, m, l, x) are crafted with varying architectural scales through depth_multiple (network depth) and width_multiple (number of channels) to achieve an optimal balance between speed and accuracy. YOLOv8n(0.33,0.25) represents the smallest model with 3.2M parameters and 8.7 GFLOPs (Mas et al., 2024), making it particularly suitable for devices with limited resources. YOLOv8s(0.33,0.50) demonstrates enhanced width, striking an effective balance with 11.2M parameters and 28.6 GFLOPs. YOLOv8m(0.67,0.75) exhibits a significant increase in both depth and width, featuring 25.9M parameters and 78.9 GFLOPs, which contributes to enhanced accuracy. YOLOv8l (1.00,1.00) denotes the complete baseline architecture, comprising 43.7 million parameters and 165.2 GFLOPs, making it appropriate for achieving high accuracy. Ultimately, YOLOv8x (1.00/1.33, 1.25) stands out as the largest and most precise model (68.2M parameters, 257.8 GFLOPs) (Mas et al., 2024), featuring an increased number of channels, which positions it as the ideal option for applications that necessitate the highest level of accuracy, even with considerable computational requirements.

3 Result and Discussion

3.1 Result

A thorough assessment of several YOLOv8 model variations (n, s, m, l, and x) is performed utilizing a meticulously curated and annotated dataset pertaining to nocturnal durian stealing. The efficacy of each model is evaluated using conventional object detection metrics, encompassing mean Average Precision (mAP) at an Intersection over Union (IoU) threshold of 0.5 (mAP@0.5) and mAP over an IoU range from 0.5 to 0.95 (mAP@0.5:0.95), in addition to Precision and Recall. All models are trained and assessed on the identical hardware platform to guarantee equitable comparison. The outcomes of identifying durian thieves during nighttime utilising YOLOv8 variants n, s, m, l, and x are presented in Figure 1, Figure 2, Figure 3, Figure 4, and Figure 5, respectively. Table 1 presents a summary of precision, recall, mAP50, and mAP50:95.

Figure 1. Detection Results Using Yolov8n

Figure 2. Detection Results Using Yolov8s

Figure 3. Detection Results Using Yolov8m

Figure 4. Detection Results Using Yolov8l

Figure 5. Detection Results Using Yolov8x

Table 1. Performance Results of the Yolov8 Variant Model (n,s,m,l,x) in Detecting Durian Thieves at Night

Model	Precision	Recall	mAP50	mAP50:95
Yolov8n	0.898	0.935	0.909	0.579
Yolov8s	0.952	0.938	0.977	0.587
Yolov8m	0.976	0.937	0.982	0.597
Yolov8l	0.978	0.941	0.984	0.612
Yolov8x	0.981	0.944	0.988	0.624

The experimental results shown in Table 1 provide compelling evidence of the correlation between the increase in YOLOv8 model capacity and the enhancement in object detection performance within the challenging context of nighttime durian theft scenarios. There is a distinct upward trend in all evaluation metrics, including Precision, Recall, mAP@0.5, and mAP@0.5:0.95, ranging from the smallest variant (YOLOv8n) to the largest (YOLOv8x). Notably, YOLOv8x stands out as the model exhibiting exceptional performance, attaining the highest metrics for mAP@0.5 (0.988), mAP@0.5:0.95 (0.624), Precision (0.984), and Recall (0.944). This ongoing enhancement demonstrates that more extensive and intricate networks possess enhanced abilities to extract and process strong features from frequently noisy and detail-restricted images, which are common traits of the low-light conditions central to this investigation.

3.2 Discussion

The enhanced performance of YOLOv8x is due to its more extensive architecture, marked by increased depth_multiple and width_multiple in comparison to other variants. The enhanced capacity of this model enables it to acquire more distinct feature representations, which is vital in low-light scenarios where the visual information is limited. The enhanced capacity of the model to create intricate filters and feature hierarchies is essential for differentiating between target objects (thief, carrying tools) and background noise. Moreover, the variability in the appearance of objects, including actors donning face coverings (such as clothing and hats) and wielding different tools (like sickles and tongs), necessitates robust generalization abilities from the model. YOLOv8x, featuring a more sophisticated architecture, demonstrates a superior capacity to handle these variations, effectively identifying objects even amidst the obscuration and image degradation typical of nighttime settings. While smaller variants, such as YOLOv8n and YOLOv8s, exhibit impressive computational efficiency and satisfactory accuracy for simpler tasks, their performance reveals constraints when addressing the stringent accuracy requirements essential for vital security applications, including theft prevention. For instance, while YOLOv8n demonstrates a commendable mAP@0.5 of 0.902, its mAP@0.5:0.95 accuracy of 0.579 significantly trails that of YOLOv8x, which stands at 0.624. This highlights the challenges associated with producing highly accurate bounding boxes at elevated IoU thresholds. The most notable enhancement in mAP@0.5:0.95 performance develops progressively from variants 'n' to 'x', underscoring the model's ability to attain greater spatial precision and resilience against environmental noise. The exceptionally high Precision and Recall demonstrated by YOLOv8l and YOLOv8x reflect remarkably low rates of false positives and false negatives, which is an essential characteristic for a dependable early warning system.

The impressive performance of YOLOv8x presents exciting possibilities for real-world applications; however, one must take into account its significant computational demands. The model possesses a significantly greater number of parameters and GLOPs (floating-point operations) compared to other variants, indicating a need for more robust computational resources (e.g., high-end GPUs) and possibly extended inference times. The identified limitations could hinder implementation on edge devices with limited resources, such as battery-operated, standalone security cameras situated in remote areas. Additionally, although this dataset aims to mirror real-world situations, it may not completely capture the variations found in actual conditions, such as severe weather events like rain or fog, diverse forms of artificial lighting, or advanced burglary techniques.

4 Conclusion

This research has successfully developed and thoroughly assessed a deep learning-based system for detecting durian theft, which operates efficiently in low-light conditions —a significant challenge in agricultural security. A rigorous study of several YOLOv8 architectural variations (n, s, m, l, and x) revealed a distinct association between model capacity and detection performance. The testing results indicate that YOLOv8x markedly surpasses other versions, attaining the greatest mean Average Precision (mAP@0.5) of 0.988 and a Recall of 0.944, along with enhanced performance on the mAP@0.5:0.95 and Precision measures. This verifies that deeper and broader networks possess the capability to extract strong and discriminative characteristics from complex and changeable pictures, characteristic of nighttime surveillance settings with hidden objects.

The selection of YOLOv8x as the superior variation underscores the significance of enhanced model ability to tackle specific issues, including low visibility, offender concealment, and nuanced object detection in agricultural crime prevention. When the system detects a potential thief, it can be used to automatically trigger an alarm, send a notification to the owner, or activate other security responses such as turning on a spotlight that points at the durian thief. This achievement demonstrates the viability of utilizing sophisticated deep learning for security applications in agriculture, providing an automated solution that significantly surpasses conventional approaches. This research primarily contributes by demonstrating the exceptional performance of YOLOv8x in challenging situations and by creating a distinctive dataset that accurately simulates nighttime burglary scenarios, serving as a valuable resource for future investigations in this field. However, the study acknowledges certain limitations. The superior YOLOv8x model has high computational requirements, which could be a challenge for implementation on low-resource edge devices. Furthermore, the dataset, while comprehensive, may not encompass all potential real-world variables, such as extreme weather conditions or different theft tactics.

Acknowledgment

This research was financially supported by DIPA Politeknik Negeri Bali under the grant number SP DIPA-139.03.2.693476/2025, with the third revision on April 8, 2025.

References

Bai, R., Shen, F., Wang, M., Lu, J., & Zhang, Z. (2023). Improving detection capabilities of YOLOv8-n for small objects in remote sensing imagery: Towards better precision with simplified model complexity. *Research Square*, 0–9. https://doi.org/10.21203/rs.3.rs-3085871/v1.

Barakat, M., Chung, G. C., Lee, I. E., Pang, W. L., & Chan, K. Y. (2023). Detection and sizing of durian using zero-shot deep learning models. *International Journal of Technology*, *14*(6). https://doi.org/10.14716/ijtech.v14i6.6640.

Dai, X., Hu, J., Zhang, H., Shitu, A., Luo, C., Osman, A., Sfarra, S., & Duan, Y. (2021). Multi-task faster R-CNN for nighttime pedestrian detection and distance estimation. *Infrared Physics and Technology*, *115*. https://doi.org/10.1016/j.infrared.2021.103694.

Han, X., Chang, J., & Wang, K. (2021). Real-time object detection based on YOLO-v2 for tiny vehicle objects. *Procedia Computer Science*, *183*. https://doi.org/10.1016/j.procs.2021.02.031.

Indrayana, I. N. E., Subiksa, G. B., Prihatini, P. M., Suasnawa, I. W., & Ciptayani, P. I. (2024). Yolo model for durian theft detection in night vision. *Proceedings of the International Conference on Sustainable Green Tourism Applied Science - Engineering Applied Science 2024 (ICoSTAS-EAS 2024)*, *2024*, 339–348. https://doi.org/10.2991/978-94-6463-587-4_39.

Indrayana, I. N. E., Sudarma, M., Putra, I. K. G. D., & Sudana, A. A. K. O. (2024). Improve nighttime highway vehicles and pedestrian detection using Yolov8+CLAHE. *2024 Ninth International Conference on Informatics and Computing (ICIC)*, 1–6. https://doi.org/10.1109/ICIC64337.2024.10956682.

Kim, J. H., Hong, H. G., & Park, K. R. (2017). Convolutional neural network-based human detection in nighttime images using visible light camera sensors. *Sensors (Switzerland)*, *17*(5). https://doi.org/10.3390/s17051065.

Lee, J., & Hwang, K. (2022). YOLO with adaptive frame control for real-time object detection applications. *Multimedia Tools and Applications*, *81*(25). https://doi.org/10.1007/s11042-021-11480-0.

Li, C., Li, L., Jiang, H., Weng, K., Geng, Y., Li, L., Ke, Z., Li, Q., Cheng, M., Nie, W., Li, Y., Zhang, B., Liang, Y., Zhou, L., Xu, X., Chu, X., Wei, X., & Wei, X. (2022). *YOLOv6: A single-stage object detection framework for industrial applications*. http://arxiv.org/abs/2209.02976.

Mas, M. S., Saidah, S., & Ibrahim, N. (2024). Detection and counting of wheat ear using YOLOv8. *International Journal of Electrical and Computer Engineering*, *14*(5), 5813–5823. https://doi.org/10.11591/ijece.v14i5.pp5813-5823.

Patel, H., Prajapati, K., Sarvaiya, A., Upla, K., Raja, K., Ramachandra, R., & Busch, C. (2022). Depthwise convolution for compact object detector in nighttime images. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, *2022-June*. https://doi.org/10.1109/CVPRW56347.2022.00053.

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, *2016-December*. https://doi.org/10.1109/CVPR.2016.91.

Redmon, J., & Farhadi, A. (2017). YOLO9000: Better, faster, stronger. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, *2017-January*. https://doi.org/10.1109/CVPR.2017.690.

Safaldin, M., Zaghden, N., & Mejdoub, M. (2024). An improved YOLOv8 to detect moving objects. *IEEE Access*, *12*(April), 59782–59806. https://doi.org/10.1109/ACCESS.2024.3393835.

Sharma, K., Rawat, S. S., Parashar, D., Sharma, S., Roy, S., & Sahoo, S. (2023). State-of-the-Art analysis of multiple object detection techniques using deep learning. *International Journal* *of Advanced Computer Science and Applications*, *14*(6). https://doi.org/10.14569/IJACSA.2023.0140657

Sohan, M., Sai Ram, T., & Rami Reddy, C. V. (2024). *A Review on YOLOv8 and Its Advancements*. https://doi.org/10.1007/978-981-99-7962-2_39

Tahir, N. U. A., Long, Z., Zhang, Z., Asim, M., & ELAffendi, M. (2024). PVswin-YOLOv8s: UAV-based pedestrian and vehicle detection for traffic management in smart cities using improved YOLOv8. *Drones*, *8*(3). https://doi.org/10.3390/drones8030084

Zhu, J., Hu, T., Zheng, L., Zhou, N., Ge, H., & Hong, Z. (2024). YOLOv8-C2f-Faster-EMA: An improved underwater trash detection model based on YOLOv8. *Sensors*, *24*(8). https://doi.org/10.3390/s24082483