Automated Mangrove Tree Counting Using Hybrid Trunk Detection Model

Wong Yu Ting1, a), Tee Connie1, 2, b) and Michael Kah Ong Goh1, 2, c)

1Faculty of Information Science and Technology, Multimedia University, Jalan Ayer Keroh Lama, 75450 Bukit Beruang, Melaka, Malaysia

2Centre of Image and Vision Computing, CoE for Artificial Intelligence, Multimedia University, Persiaran Multimedia, 63100 Cyberjaya, Malaysia.

b) Corresponding author: tee.connie@mmu.edu.my

a) 1211104759@student.mmu.edu.my

c)michael.goh@mmu.edu.my

**Abstract.** Mangrove trees are one of the Malaysia’s significant assets, with the country ranking highly in mangrove forest conservation. These trees thrive along coastal wetlands and are essential for maintaining ecological balance, yet they face significant threats from deforestation and climate change. This study presents an innovative approach that combines deep learning and computer vision techniques for detecting and counting mangrove trees with its trunk from the side view for better accuracy. The dataset was manually captured in the Straits of Malacca and annotated using Roboflow. Experiments reveal that relying solely on deep learning, specifically YOLO in this study, leads to challenges such as duplicate detections of overlapping trunks and misclassification of objects with similar colours. The Hybrid Trunk Detection (HTD) Model is introduced to address limitations of YOLO, leading to improved detection and counting accuracy. Experimental results demonstrate that the HTD Model effectively reduces false positives and achieves significant improvement in F1 score and mAP50 over the baseline results. The findings highlight the potential of AI-driven approaches in ecological conservation.

# INTRODUCTION

Mangroves thrive in salty oxygen poor soil and are suited for coastal intertidal zones [1]. They are essential to preserving ecological balance and sustaining human livelihoods. They can be natural nursery and constitute a basis for coastal food web. Roots of mangrove trees can reduce the motion of tidal waters, protect coastlines from erosion, storms, and flooding. Significant amounts of carbon from the atmosphere are also absorbed and stored by mangroves. Unfortunately, deforestation, climate change, and human activities have been the cause of threats to the population of mangroves. The need to monitor, manage and conserve the ecosystem is urgent. Traditional approach of monitoring mangroves can consume a huge amount of time and manpower and has potential risk of errors, highlighting the need for innovative approaches. This study presents a Hybrid Trunk Detection (HTD) Model for detecting and counting mangrove trees with its trunk from the side view for better accuracy.

Over the years, researchers have explored both conventional and deep learning methods for tree detection. Conventional approaches include [2], who developed an automated system for olive tree detection using RGB satellite images, combining Laplacian of Gaussian filtering, improved K-Means clustering, and Random Forest classification, achieving highest accuracy of 97.52%. Similarly, [3] applied Gray-Level Co-occurrence Matrix (GLCM), Wavelet Transform, and Template Matching, for oil palm trees detection using drone imagery, achieving the highest accuracy of 82.07%.

[4] utilized Connected Component Labeling (CCL) and morphological operations to detect citrus trees, reporting 95.9% precision and 97.9% recall%. [5] used a conventional vision-based method involving uneven illumination correction, filtering, and a circular Hough transform to detect olive trees with 94% true positive rate and 95.1% precision. [6] combined spectral and spatial information using Delaunay triangulation for olive tree canopy detection, achieving a 74% F1 score.

In deep learning approaches, [7] employs YOLOv5 variants (YOLOv5s, YOLOv5m, and YOLOv5x) for cherry trees detection from drone images, with YOLOv5s achieving an F1 score of 98%. Training used techniques like mosaic data augmentation and Stochastic Gradient Descent optimization.[8] introduced TrunkNet, a two-stage deep learning model for urban tree trunk detection significantly outperforming PoolNet and U2-Net with a 96.6% F-measure. [9] compared YOLOv5, SSD, and UNET for tree detection using satellite imagery, with UNET achieving the highest accuracy of 81%.

Recent studies include [10], who presented an Object Based Image Analysis-Convolution Neural Network (OBIA-CNN) framework for olive tree detection from UAV imagery, achieving 99% F-score. [11] developed HR-SFANet, for urban tree detection, outperforming PyCrown and DeepForest with a 73.5% F-score. [12] advanced oil palm detection by enhancing RetinaNet with deformable convolutions and IoU-aware detection, achieving F1 scores of 94.7% and 90.2% for single- and dual-species detection, respectively.

## PROPOSED SOLUTION

The proposed solution begins with capturing an input image of mangrove trees. The image is passed through a trained YOLOv8s model to detect tree trunks. The predicted bounding boxes are then processed using a custom object filtering and colour space enhancement technique. This stage aims to refine YOLO outputs by confirming true tree trunks and removing false positives based on HSV and contour features. Finally, the number of valid detections is counted to determine the total number of trees in the image. Figure 1 shows the flow diagram that illustrates flow of detection and counting with the proposed Hybrid Trunk Detection (HTD) Model.

|  |
| --- |
|  |

**FIGURE 1.** Proposed hybrid trunk detection (HTD) model workflow

### Data Collection

The dataset used in this study consisted of a total of 610 manually collected images taken in horizontal view, with different angles from the side view of mangrove trees using mobile phones, at the Straits of Melaka, Malaysia, right in front of The Shore Hotel and Residences shown in Figure 2. Image was captured with the resolution 4608 × 3456 and 4608 × 2592, using Vivo iQOO Z7 5G and Vivo T1 5G as the devices. Data collection is conducted for a few days at the same location, as changes and happening of natural phenomena such as rising tide provide different presentations and reflection to the image captured. The data collected includes different complexities in daylight setting for the model to train and increase the ability to fit with real world scenarios. The images were then annotated utilising Roboflow, and split into train, validation, and test set in 7:2:1 ratio.

|  |  |  |
| --- | --- | --- |
|  |  |  |

**FIGURE 2.** Sample data collected under different daylight settings

### The Baseline YOLO Architecture

This study utilises YOLO (You Only Look Once) as the baseline object detection module. YOLO treats object detection as a single regression problem, enables the algorithm to predict bounding boxes and class probabilities for objects directly from image in a single forward pass of network. Several versions of YOLO, such as YOLOv5s, YOLOv8s, and YOLOv11s were chosen for this mangrove tree counting project. YOLOv5s is a lightweight version optimised for resource efficiency and fast inference speed. YOLOv8 is the model with significant improvements over YOLOv5, introducing architectural changes and path aggregation, where the small (s) variance model offers a balance of speed and accuracy. YOLOv11 introduced more efficient architecture and advanced attention mechanisms, designed to enhance small object detection and improve accuracy while maintaining the real-time inference speed that YOLO is known for. YOLO produces confidence scores for each predicted bounding box and calculates a loss function to optimize its predictions. The confidence score combines the probabilities of abjectness score and class probability, represents the likelihood that an object is present in the box.

### Hybrid Trunk Detection (HTD) Model

The YOLO performs satisfactorily in detecting tree trunks under normal conditions. However, it exhibits limitations in challenging scenarios, particularly under low-light conditions or when other object share similar colours and textures with tree trunks. Therefore, this study proposes a Hybrid Trunk Detection (HTD) Model designed to enhance mangrove tree trunk detection and overcome the limitations of YOLO. The HTD Model integrates the YOLOv8s model with a custom colour based processing verification step. This hybrid approach leverages YOLOv8’s robust object localization capabilities, followed by a domain-specific computer vision filter to confirm the presence of actual tree trunks within each predicted bounding box.

The YOLOv8s model was first used to detect potential tree trunks in each image. The model outputs bounding boxes along with their corresponding confidence scores. These predictions, however, may include false positives due to visual similarities between trunks and other background elements such as shadows, roots, etc.

To reduce false positives, colour based processing was applied to each YOLO-detected bounding box. Specifically, additional non-maximum suppression was applied to the images after YOLO prediction to remove recurring detection and drawing of bounding boxes on the same region. Then the region of interest (ROI) inside each predicted box was extracted and then converted to HSV (Hue, Saturation, Value) colour space, and a binary mask was generated by thresholding the HSV transformed region, highlighting pixels falling within the specified hue, saturation, and value range, which are areas likely to contain tree trunk features (refer to Equation (1)),

where *I* represent input image, each *di* is a candidate detection predicted by YOLO. For each detected window, *x*, the region within the window is cropped (see Equation (2)):

The cropped RGB region is then converted into HSV colour scheme using Equation (3),

where *R*, *G* and *B* denote the red, green, and blue components respectively. is the maximum among these three values, represents the intensity or brightness of the dominant colour channel, and is computed as in Equation (4),

The saturation, *S*, and value, *V*, are computed as in Equations (5) and (6),

Pixel-wise thresholding is performed after converting the image into HSV (see Equation (7)). Each single pixel is kept if:

A binary mask is produced, and the contours are extracted from the binary mask. If a sufficiently large contour was detected, the prediction was considered a valid tree trunk. This filtering step ensures that only detections containing the visual characteristics of tree trunks are retained. The verification function *V(di)* is given by Equation (8).

A contour is considered valid if it exceeds a minimum area (threshold). The final verified trunk set, *Y*, is computed as in Equation (9),

## EXPERIMENTAL RESULTS

### Comparison among different YOLO models

All models were trained with 70 epoch, batch size 16 with 640 image size. Table 1 shows the validation result of each YOLO model. From the results in Table 1, YOLOv8s outperforms the other models by achieving the highest precision at 0.862, recall at 0.767, mAP50 at 0.813 and mAP50-95 at 0.627. Thus, YOLOv8s model is adopted for the subsequent tests.

**TABLE 1.** Results of different YOLO models.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Metrics** | **YOLOv5s (val)** | **YOLOv8s (val)** | **YOLOv11s (val)** | **YOLOv8s (test)** |
| Precision | 0.834 | 0.862 | 0.800 | 0.815 |
| Recall | 0.749 | 0.767 | 0.754 | 0.814 |
| F1 score | 0.789 | 0.813 | 0.776 | 0.814 |
| mAP50 | 0.842 | 0.857 | 0.830 | 0.847 |
| mAP50-95 | 0.613 | 0.627 | 0.620 | - |

### Evaluation of Hybrid Trunk Detection (HTD) Model

While the YOLO model demonstrates strengths in tree trunk detection, it also exhibits specific limitations. The YOLO model, although generally accurate, tends to overcount in cases of partial occlusion or overlapping trunks with similar textures. To address these challenges, the proposed HTD Model leverages the precision of YOLO for initial detection and utilizing the colour-based processing method as a validation step to enhance reliability and reduce false positives, aiming to improve the robustness of detection in complex mangrove environments.

To evaluate the performance of the proposed approach, extensive experiments under various combinations of Non-Maximum Suppression (NMS) and Intersection of Union (IoU) threshold values were conducted. Table 2 presents the results of the proposed HTD model across NMS thresholds ranging from 0.1 to 0.6 and IoU thresholds between 0.1 to 0.5.

From the results shown in Table 2, it demonstrates that the proposed method significantly enhances detection performance compared to the baseline YOLOv8s model. The best overall performance was achieved with a NMS threshold of 0.6 and an IoU threshold of 0.1. where the model attained a precision of 0.7937, recall of 0.9091, F1 score of .8475, and the highest mAP of 0.8851, outperformed the baseline YOLOv8s test results. This indicates excellent object localization and detection coverage. A closely competitive configuration was observed under an NMS threshold of 0.5, which achieved a slightly higher F1 score of 0.8553 and a very strong mAP50 of 0.8797 at the same IoU threshold. Although the mAP50 was marginally lower than that of the NMS 0.6 configuration, the higher F1 score reflects better balance between precision and recall.

### Visual Inspection of the Hybrid Trunk Detection (HTD) Model

Figure 3 illustrates the result of the tree trunk detection in every step of the HTD Model. The YOLO model tends to mis-detect tree trunks in certain challenging scenarios such as mistakenly interprets a single trunk as two or more separate trunks when leaves partially cover a trunk, and struggle to distinguish trunks under low lighting conditions when overlapping trunks or other object have similar colours and textures. The HTD Model incorporates a colour-based processing filter to verify and refine YOLO’s detections, reducing false positives and enhances overall reliability of mangrove tree trunk detection.

**TABLE 2.** Results of proposed approach under different thresholds setting.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **NMS Threshold** | **Metrics** | **HTD Model (IoU threshold =0.1)** | **HTD Model (IoU threshold =0.2)** | **HTD Model (IoU threshold =0.3)** | **HTD Model (IoU threshold =0.4)** | **HTD Model (IoU threshold =0.5)** |
| NMS=0.2 | Precision | 0.8673 | 0.8531 | 0.8294 | 0.8104 | 0.7725 |
| Recall | 0.8318 | 0.8182 | 0.7955 | 0.7773 | 0.7409 |
| F1 score | 0.8492 | 0.8353 | 0.8121 | 0.7935 | 0.7564 |
| mAP50 | 0.8132 | 0.8061 | 0.7889 | 0.7663 | 0.7155 |
| NMS=0.3 | Precision | 0.8559 | 0.8423 | 0.8243 | 0.8108 | 0.7748 |
| Recall | 0.8636 | 0.8500 | 0.8318 | 0.8182 | 0.7818 |
| F1 score | 0.8597 | 0.8462 | 0.8281 | 0.8145 | 0.7783 |
| mAP50 | 0.8434 | 0.8408 | 0.8278 | 0.8004 | 0.7525 |
| NMS=0.4 | Precision | 0.8362 | 0.8233 | 0.8103 | 0.8017 | 0.7672 |
| Recall | 0.8818 | 0.8682 | 0.8545 | 0.8455 | 0.8091 |
| F1 score | 0.8584 | 0.8451 | 0.8319 | 0.8230 | 0.7876 |
| mAP50 | 0.8591 | 0.8567 | 0.8429 | 0.8198 | 0.7756 |
| NMS=0.5 | Precision | 0.8148 | 0.8025 | 0.7984 | 0.7860 | 0.7531 |
| Recall | 0.9000 | 0.8864 | 0.8818 | 0.8682 | 0.8318 |
| F1 score | 0.8553 | 0.8423 | 0.8380 | 0.8251 | 0.7905 |
| mAP50 | 0.8797 | 0.8720 | 0.8594 | 0.8383 | 0.7932 |
| NMS=0.6 | Precision | 0.7937 | 0.7817 | 0.7778 | 0.7659 | 0.7381 |
| Recall | 0.9091 | 0.8955 | 0.8909 | 0.8773 | 0.8455 |
| F1 score | 0.8475 | 0.8347 | 0.8305 | 0.8178 | 0.7881 |
| mAP50 | 0.8851 | 0.8775 | 0.8691 | 0.8480 | 0.8046 |

# CONCLUSION

In conclusion, this study proposed and evaluated a Hybrid Trunk Detection Model to detect and count mangrove trees. The proposed approach capitalizes on YOLO’s detection capabilities and leverages the rule-based filtering of the computer vision method to validate and refine detections. This strategy effectively reduces false positives and enhances detection reliability in visually complex natural environments. Experimental results show that the proposed approach provides a more accurate estimation of tree trunk count compared to either method used in isolation. This work contributes an approach that supports ecological monitoring, offering robust and scalable tree trunk detection and counting in complex mangrove environments. Future work may focus on extending this framework to support ecological studies, such as analysing tree growth patterns, monitoring regeneration and reproduction.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Sample 1** | **Sample 2** | **Sample 3** |
| **Original Image** |  |  |  |
| **YOLO** | **Number of trunks detected: 11** | **Number of trunks detected: 4** | **Number of trunks detected: 7** |
| **Colour-based Processing** |  |  |  |
| **HTD Model** | **Number of trunks detected: 7** | **Number of trunks detected: 3** | **Number of trunks detected: 5** |

**FIGURE 3.** Visual comparison of trunk detection results across different stages

# References

1. S. Srikanth, S.K.Y. Lum, and Z. Chen, “Mangrove root: adaptations and ecological importance,” Trees **30**(2), 451–465 (2016).
2. M. Waleed, T.-W. Um, A. Khan, and U. Khan, “Automatic Detection System of Olive Trees Using Improved K-Means Algorithm,” Remote Sensing **12**(5), 760 (2020).
3. S.A. Hashim, S. Daliman, I.N.M. Rodi, N.A. Aziz, N.A. Amaludin, and A.E. Rak, “Analysis of Oil Palm Tree Recognition using Drone-Based Remote Sensing Images,” IOP Conf. Ser.: Earth Environ. Sci. **596**(1), 012070 (2020).
4. C. Donmez, O. Villi, S. Berberoglu, and A. Cilek, “Computer vision-based citrus tree detection in a cultivated environment using UAV imagery,” Computers and Electronics in Agriculture **187**, 106273 (2021).
5. O.A.A. Al-Tikriti, B.S.F. Al-Saffar, A.M. Bozdoğan, and S. Arıca, “Detection and Counting of Olive Trees in Images Taken by an Unmanned Aerial Vehicle,” in *2022 4th International Conference on Current Research in Engineering and Science Applications (ICCRESA)*, (IEEE, Baghdad, Iraq, 2022), pp. 75–79.
6. S. Illana Rico, D.M. Martínez Gila, P. Cano Marchal, and J. Gómez Ortega, “Automatic Detection of Olive Tree Canopies for Groves with Thick Plant Cover on the Ground,” Sensors **22**(16), 6219 (2022).
7. T. Ozer, C. Akdogan, E. Cengiz, M.M. Kelek, K. Yildirim, Y. Oguz, and H. Akkoc, “Cherry Tree Detection with Deep Learning,” in *2022 Innovations in Intelligent Systems and Applications Conference (ASYU)*, (IEEE, Antalya, Turkey, 2022), pp. 1–4.
8. R. Li, G. Sun, S. Wang, T. Tan, and F. Xu, “Tree trunk detection in urban scenes using a multiscale attention-based deep learning method,” Ecological Informatics **77**, 102215 (2023).
9. U. E Hani, S. Munir, S. Younis, T. Saeed, and H. Younis, “Automatic Tree Counting from Satellite Imagery Using YOLO V5, SSD and UNET Models: A case study of a campus in Islamabad, Pakistan,” in *2023 3rd International Conference on Artificial Intelligence (ICAI)*, (IEEE, Islamabad, Pakistan, 2023), pp. 88–94.
10. O. Ameslek, H. Zahir, H. Latifi, and E.M. Bachaoui, “Combining OBIA, CNN, and UAV imagery for automated detection and mapping of individual olive trees,” Smart Agricultural Technology **9**, 100546 (2024).
11. J. Ventura, C. Pawlak, M. Honsberger, C. Gonsalves, J. Rice, N.L.R. Love, S. Han, V. Nguyen, K. Sugano, J. Doremus, G.A. Fricker, J. Yost, and M. Ritter, “Individual tree detection in large-scale urban environments using high-resolution multispectral imagery,” International Journal of Applied Earth Observation and Geoinformation **130**, 103848 (2024).
12. S.S. Lee, L.G. Lim, S. Palaiahnakote, J.X. Cheong, S.S.M. Lock, and M.N.B. Ayub, “Oil palm tree detection in UAV imagery using an enhanced RetinaNet,” Computers and Electronics in Agriculture **227**, 109530 (2024).