Mango Fruit Detection from Aerial Image

Marawan Ashraf Fawzy Eldeib

Mohd Haris Lye Abdullah

Abstract

This research investigates the use of deep learning models to detect mango fruits from aerial images captured by drones, focusing on Faster R-CNN and YOLO variants. The study evaluates these models using datasets from the University of Sydney, CQUniversity, and a locally prepared dataset. Among the models, YOLOv8 demonstrated the highest performance, achieving 98.5% mAP on the Sydney dataset and 80.9% mAP on the local dataset. Known for its real-time capability, YOLOv8 also shows strong potential for applications requiring rapid detection. The MangoVision GUI, designed with bilingual support and features such as GPS coordinate extraction, integrates these models into a user-friendly tool for precise mango detection. This study highlights the potential of AI-driven solutions to enhance agricultural practices, making YOLOv8 a top choice for precision farming.

Problem Statement & Objectives

Problem Statement:
Manual mango detection in large orchards is labour-intensive, time-consuming, and prone to errors. Additionally, environmental challenges like dense foliage, occlusion, and varying drone altitudes and angles further complicate accurate detection.

Objectives:

- Evaluate public datasets on mango fruit detection. Collect and label a mango fruit detection dataset from aerial images captured by a drone.
- Design and train an object detection model aimed at detecting mango fruits from aerial images.
- Evaluate the performance of the model based on the prepared dataset.

Literature Review

- The integration of AI with drone technology is transforming agriculture, enabling precise and efficient crop monitoring. The latest YOLO models, YOLOv8 and YOLOv10, excel in real-time detection, offering both high speed and accuracy. Earlier studies, such as those by Alzubaidi et al. (2021) and Ranjan & Machavaram (2022), demonstrated the effectiveness of YOLOv2 and YOLOv5 for mango detection, highlighting the advancements that can be brought by these newer models.
- Bargoti & Underwood (2017) employed Faster R-CNN with VGG-16 backbone for fruit detection, achieving high accuracy. Pham's (2023) use of ResNet-50 with the Detectron2 framework has further improved Faster R-CNN efficiency, though its two-stage process remains more computationally demanding compared to the faster, one-stage YOLO models.

Methodology

This study systematically integrates advanced deep learning models for mango detection from aerial images, utilizing a locally collected dataset alongside public datasets from CQUniversity and the University of Sydney. The methodology begins with data collection, followed by preprocessing steps such as annotation, resizing, and augmentation to enhance model performance. The data is then split into training, validation, and test sets to rigorously assess model accuracy. The models evaluated include YOLOv8, YOLOv10, and Faster R-CNN with VGG-16 and ResNet-50 backbones, as well as Detectron2 with ResNet-50. Each model undergoes hyperparameter tuning, focusing on key factors such as learning rates, optimizers, and early stopping, with performance metrics like mean Average Precision (mAP) and F1 Score guiding the evaluation. These models are integrated into MangoVision, a bilingual graphical user interface (GUI) that includes GPS coordinate extraction, providing a practical and user-friendly tool for precise mango detection. This comprehensive approach ensures that the study not only evaluates state-of-the-art models but also applies these advancements to real-world agricultural challenges.

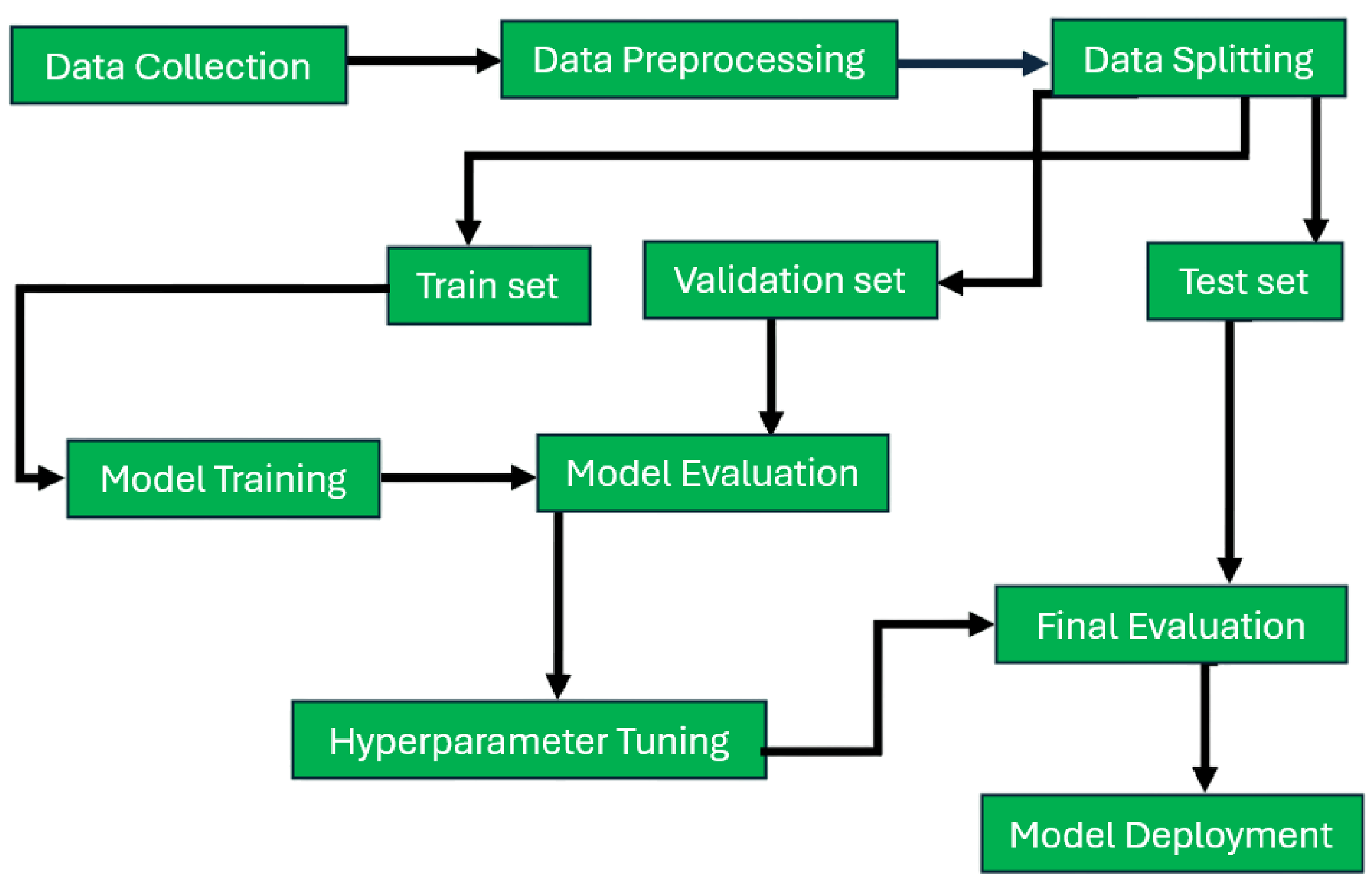

Figure 1: Methodology Flowchart.

Testing & Evaluation

The implementation phase involved training and testing the YOLOv8, YOLOv10, and Faster R-CNN models with VGG-16 and ResNet-50 backbones using datasets from the University of Sydney, CQUniversity, and a locally collected dataset. YOLOv8 emerged as the top performer, achieving a mean Average Precision (mAP) of 98.5% on the Sydney dataset, surpassing YOLOv10 and Detectron2 ResNet-50 by 4.6% and 5.5%, respectively, and showing a 10.9% improvement over the previous Faster R-CNN (VGG-16) study. On the local dataset, YOLOv8 achieved a strong mAP of 80.9%, further demonstrating its robustness across different environments.

Model	Backbone	mAP@0.5
Faster R-CNN (Previous Study)	VGG-16	0.89
YOLOv8n	YOLOv8	0.985
Faster R-CNN	ResNet-50 with Detectron2	0.9302

Table 1: Models Comparison Table with Previous Study on Sydney Dataset.

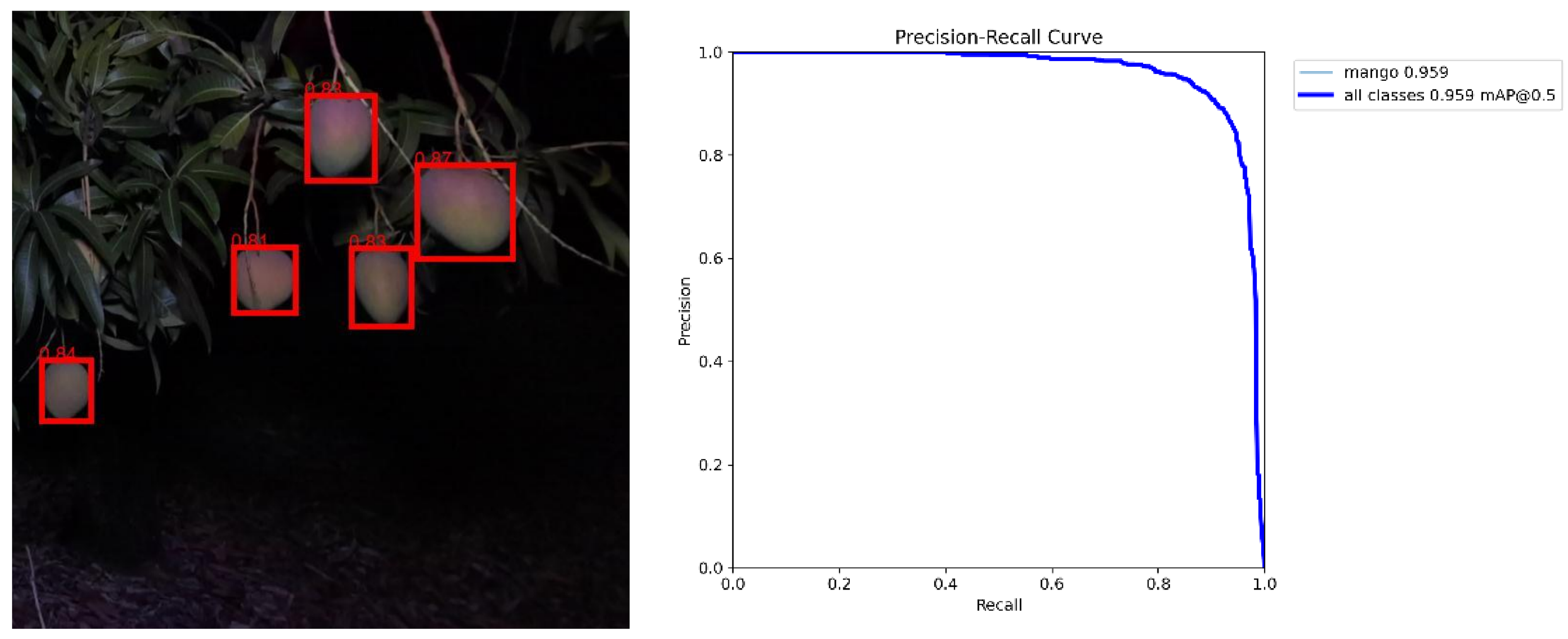

Figure 2: Example of Mango Detection and Precision-Recall Curve for our Best Model

The Precision-Recall Curve was particularly significant in evaluating model performance, offering a clear view of how well the models balanced precision and recall critical for ensuring accurate detection of mangoes. The results are visualized through MangoVision, where the GUI's bilingual support and GPS extraction features make it accessible and practical for real-world agricultural applications.

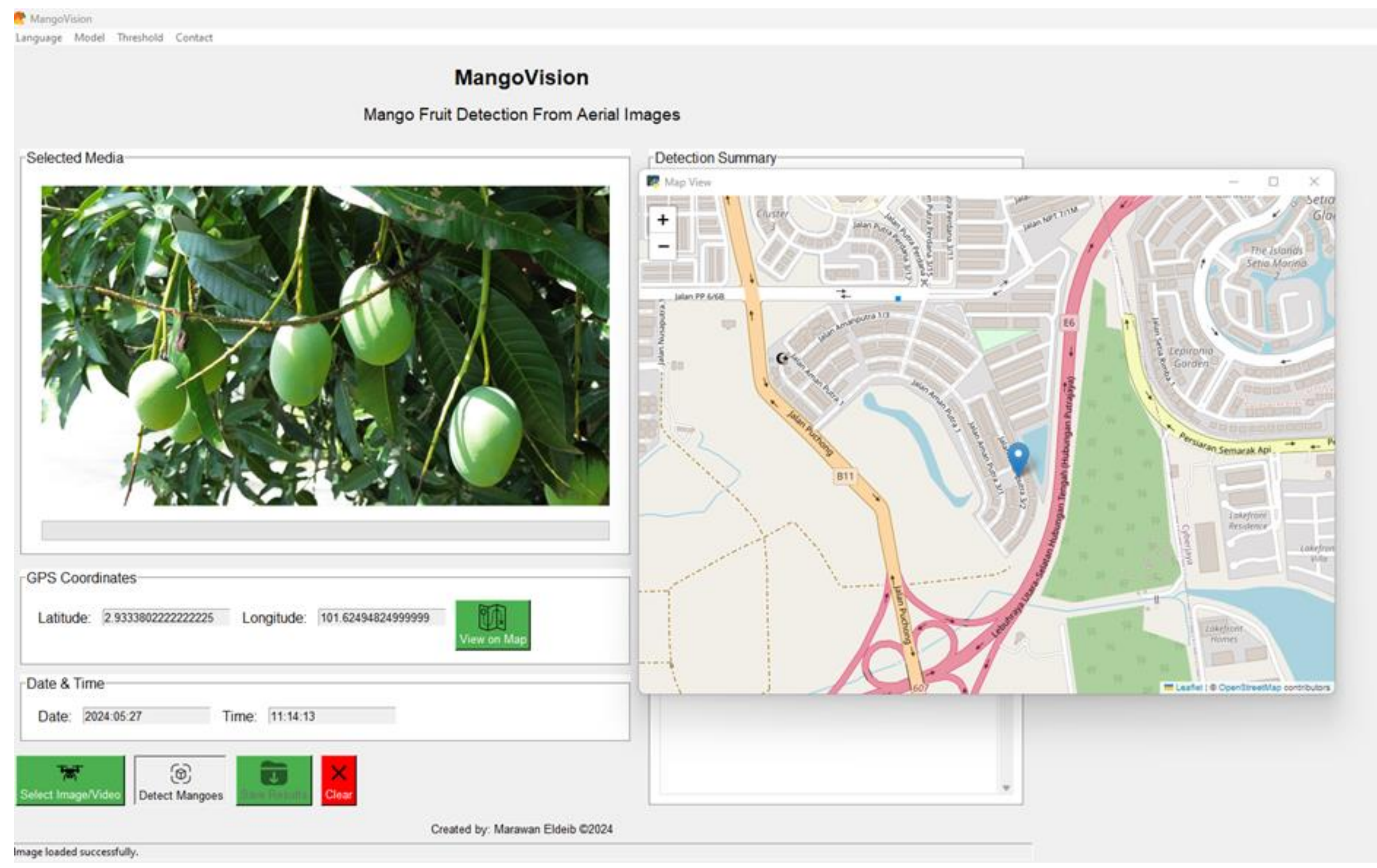

Figure 3: MangoVision GUI - GPS Coordinate Extraction and Map Visualization.

Conclusion

This project successfully applied deep learning models, specifically YOLO and Faster R-CNN, to detect mango fruits from aerial images. Among these models, YOLOv8 stood out, achieving a mAP of 98.5% on the Sydney dataset, surpassing both YOLOv10 and Faster R-CNN with Detectron2. The integration of YOLO models into the MangoVision GUI, which includes image and video detection, bilingual support, and GPS extraction, offers a smart, user-friendly solution for precision agriculture. This research highlights the role of advanced AI in driving smarter, more precise agricultural monitoring, paving the way for further innovations in farming technology.

FACULTY OF ENGINEERING

MMU MULTIMEDIA UNIVERSITY