

การจำแนกความสุกของปาล์มน้ำมันสดโดยใช้การเรียนรู้เชิงลึก

นงนุช สังข์ทอง¹, วีระ สอ้ง²

บทคัดย่อ

จากปัญหาของการจำแนกระดับความสุกของผลปาล์มน้ำมันสด ทำให้ผู้วิจัยสนใจศึกษาปัญหาและวิธีการที่ใช้ในการจำแนกระดับความสุกของผลปาล์มน้ำมันสดจากงานวิจัยในอดีต เพื่อนำเสนอวิธีการจำแนกระดับความสุกของผลปาล์มน้ำมันสดจากภาพถ่ายผลปาล์มน้ำมันโดยใช้การเรียนรู้เชิงลึก โดยผู้วิจัยเลือกใช้แบบจำลองการเรียนรู้เชิงลึกสำหรับการจำแนกระดับความสุกของผลปาล์มน้ำมันสด ที่สามารถพัฒนาไปสู่การใช้งานแบบเรียลไทม์ หรือใช้งานบนโทรศัพท์มือถือได้ เพื่อเป็นแนวทางในการประยุกต์วิธีการจำแนกระดับความสุกของผลปาล์มน้ำมันสด ให้สามารถใช้งานได้ในสถานการณ์จริงในอนาคต โดยเลือกแบบจำลองการเรียนรู้เชิงลึกที่นำหลักการของโครงข่ายประสาทเทียมแบบคอนโวลูชัน ได้แก่ แบบจำลอง YOLOv8n (nano) และทำการปรับค่าพารามิเตอร์การเรียนรู้ต่าง ๆ เพื่อให้ได้แบบจำลองการจำแนกระดับความสุกของผลปาล์มน้ำมันสดที่สามารถจำแนกได้อย่างแม่นยำที่สุด จากผลการดำเนินงานวิจัยพบว่า การจำแนกระดับความสุกของผลปาล์มน้ำมันสด ด้วยแบบจำลองการเรียนรู้เชิงลึกที่ชื่อว่า YOLOv8n (nano) ที่มีการปรับปรุงค่าพารามิเตอร์การเรียนรู้ ดังต่อไปนี้ Optimizer = Adam, Learning rate = 0.0001 และ Batch size = 32 แสดงค่าความแม่นยำสูงที่สุด จากการฝึกสอน (Training) และตรวจสอบ (Validate) ด้วยชุดข้อมูลของการฝึกสอนและชุดข้อมูลประเมินผล หลังจากนั้นทำการทดสอบแบบจำลองด้วยชุดข้อมูลทดสอบ (Testing dataset) ที่แสดงค่าความถูกต้อง (Precision) 98.8%, ค่าของการทำนายถูกต้อง (Recall) 99.7%, ค่าความเที่ยงตรงเฉลี่ยที่ตรวจพบวัตถุจริงอย่างน้อย 50% (mAP50) 99.5% และ ค่าความเที่ยงตรงเฉลี่ยที่ตรวจพบวัตถุจริงในช่วง 50% ถึง 95% (mAP50-95) 99.5%

คำสำคัญ : การจำแนกความสุก ปาล์มน้ำมัน การวิเคราะห์ภาพ การเรียนรู้เชิงลึก

¹ หลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิทยาการข้อมูล คณะวิทยาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ กรุงเทพฯ 10110

² คณะวิทยาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ กรุงเทพฯ 10110

* Corresponding author: Tel.: 061-4545644 E-mail address: nongnoot.por@gs.wu.ac.th

Ripeness Classification of Fresh Oil Palm Using Deep Learning

Nongnoot Sungthong^{1*}, Vera Sa-ing²

Abstract

The challenge of accurately ripeness level classifying from fresh oil palm fruit, this research investigated the problems and methods that were used in ripeness level classification. The aim of our research is to propose a method for classifying the ripeness level of oil palm fruit from images using deep learning. This research applies a deep learning model for classifying the ripeness level of fresh oil palm fruit and aiming for real-time application or mobile usage. The proposed model, Yolov8n (nano), adapts principles from convolutional neural networks that was tuned the learning parameters consist of the Adam optimizer, a learning rate of 0.0001, and a batch size of 32, result in the most accurate ripeness level classification model. The research demonstrates the Yolov8n (nano) improved model with optimized parameters to achieve the highest accuracy. Training and validation datasets represent precision of 98.8%, recall of 99.7%, mean average precision at 50% (mAP50) of 99.5%, and mAP50-95 of 99.5%. These findings confirm the model's effectiveness in accurately classifying the ripeness level of oil palm fruit from images.

Keywords: Ripeness classification, Oil palm, Image analysis, Deep learning

¹ Data Science, Faculty of Science, Srinakharinwirot University, Bangkok, 10110, Thailand

² Faculty of Science, Srinakharinwirot University, Bangkok, 10110, Thailand

* Corresponding author: Tel.: 061-4545644 E-mail address: nongnoot.por@g.swu.ac.th

บทนำ

ปาล์มน้ำมัน มีชื่อวิทยาศาสตร์ว่า “*Elaeis guineensis*” มีถิ่นกำเนิดมาจากแถบแอฟริกา สำหรับประเทศไทย พบว่าได้มีผู้นำเข้ามาปลูก ในปี พ.ศ. 2472 ที่จังหวัดสงขลา และจังหวัดจันทบุรี โดยมีการปลูกเป็นการค้าครั้งแรกที่ ตำบลปริก อำเภอสะเดา จังหวัดสงขลา [1] จากการสำรวจการปลูกปาล์มในประเทศไทย พบว่า ปัจจุบันพื้นที่ปลูกปาล์มน้ำมันหลักของประเทศส่วนใหญ่อยู่ในภาคใต้ คิดเป็นอัตราส่วน 87.69% ของพื้นที่ปลูกทั่วประเทศ โดยเฉพาะในจังหวัดสุราษฎร์ธานี กระบี่ และชุมพร (อัตราส่วนรวมกันเกือบ 58%) โดยเกษตรกรมีความสนใจในการปลูกปาล์มน้ำมันเพิ่มขึ้น ตามยุทธศาสตร์ของแผนพลังงานทดแทนและพลังงานทางเลือกของประเทศ ทำให้ในปี 2565 มีพื้นที่ให้ผลผลิตทั่วประเทศ ประมาณ 6.15 ล้านไร่ เพิ่มขึ้นจากในปี 2564 ที่มีพื้นที่ให้ผลผลิตทั่วประเทศประมาณ 6.03 ล้านไร่ หรือขยายเพิ่มขึ้นตัว 1.99% [2]

ปาล์มน้ำมันที่ปลูกสำหรับการค้าโดยทั่วไป มี 3 ลักษณะ ได้แก่ แบบที่ 1 ผลดิบมีสีเขียวและจะเปลี่ยนเป็นสีส้มเมื่อผลสุก ลักษณะนี้เรียกว่า *Virescens* แบบที่ 2 ผลดิบมีสีดำปลายผลและมีสีข้างที่ขั้วผล และจะเปลี่ยนเป็นสีแดงเมื่อผลสุก ลักษณะนี้เรียกว่า *Nigrescens* แบบที่ 3 เมื่อผลสุกจะมีสีผิวเปลือกสีเหลืองซีด (แบบนี้พบน้อยมาก) เรียกว่า *Albescens* [3]

โดยปัจจุบันในขั้นตอนการเก็บเกี่ยวหรือการคัดแยกคุณภาพผลปาล์มที่โรงงานสกัดน้ำมันปาล์มจะดำเนินการจำแนกระดับความสุกของผลปาล์มน้ำมันสด ด้วยการตรวจสอบด้วยสายตา ซึ่งอาจส่งผลให้เกิดความผิดพลาดได้ง่าย ด้วยปัจจัยต่างๆ เช่น เกษตรกรที่มีประสบการณ์น้อย สภาพแสง ระยะความสูงของต้นปาล์ม เป็นต้น จึงส่งผลให้เกิดความเสียหายและสูญเสียรายได้ เพราะหากเก็บเกี่ยวผลปาล์มน้ำมันสดที่ไม่มีคุณภาพ เช่น เกษตรกรเก็บเกี่ยวทะลายปาล์มน้ำมันที่ไม่สุกหรือดิบไปขาย อาจทำให้โรงงานสกัดน้ำมันปาล์มปฏิเสธการรับซื้อได้ เพราะการเก็บเกี่ยวในระยะที่ถูกต้องจะทำให้ได้ทะลายปาล์มที่มีคุณภาพ มีเปอร์เซ็นต์น้ำมันสูง ซึ่งจะมีผลต่อราคาทะลายปาล์มด้วย เนื่องด้วยระดับความสุกส่งผลต่อคุณภาพน้ำมันปาล์มที่สกัดได้ ดังนั้น เกษตรกรควรต้องคำนึงและให้ความสำคัญในขั้นตอนการเก็บเกี่ยวผลปาล์มน้ำมันสดเป็นอย่างยิ่ง

จากการทบทวนงานวิจัยทั้งในอดีตและปัจจุบัน ที่เกี่ยวกับการจำแนกระดับความสุกของปาล์มน้ำมันสด พบว่า มีการนำเสนอวิธีการที่ต่างกันหลายวิธี และวิธีที่ได้รับความนิยมมากที่สุดคือการใช้เทคนิคการมองเห็นด้วยคอมพิวเตอร์ (Computer vision) และหนึ่งในอัลกอริทึมที่ใช้ในเทคนิค Computer vision คือโครงข่ายประสาทเทียมแบบคอนโวลูชัน Convolution Neural Network (CNN) โดยจากการทบทวนงานวิจัยพบว่าเทคนิค Computer vision ได้พิสูจน์ให้เห็นว่ามีประสิทธิภาพในการจำแนกความสุกของผลปาล์มน้ำมันสดและมีความแม่นยำสูงเมื่อเทียบกับเทคนิคอื่น ๆ ทั้งนี้ แบบจำลองที่นิยมใช้และมีความแม่นยำในการจำแนกสูง ได้แก่ แบบจำลองการตรวจจับวัตถุ (Object Detection Model) ซึ่งเป็นแบบจำลองที่สามารถปรับปรุงการตรวจจับและจำแนกระดับความสุกของผลปาล์มน้ำมันสด ณ สถานที่จริงแบบเรียลไทม์ได้ในอนาคต ที่สามารถช่วยให้เกษตรกรหรือผู้ที่เกี่ยวข้องสามารถเก็บเกี่ยวผลปาล์มน้ำมันสด ได้อย่างถูกต้องในเวลาที่เหมาะสมอีกด้วย [4,5,6,7,8,9,10,11,12,13,14]

ทฤษฎีที่เกี่ยวข้อง

1. แบบจำลองการตรวจจับวัตถุ (Object Detection Model) [15]

โครงสร้างที่ใช้ในการตรวจจับวัตถุสามารถแบ่งเป็น 2 ประเภทคือ Region proposal networks และ Unified networks แบบจำลองที่ใช้ระบบ Region proposal networks เรียกว่า Multi-stage หรือ Two-stage models ส่วน Unified models เรียกว่า Single-stage model โดยแบบจำลอง Two-stage มีการทำงาน 2 ขั้นตอน คือ 1) สร้างพื้นที่ที่น่าสนใจ Region of Interest (ROI) และ 2) การจำแนกคลาส ทั้งนี้ การตรวจจับวัตถุแบบ Two-stage มีความแม่นยำสูง แต่มีการประมวลผลช้า และการตรวจจับวัตถุแบบ Single-stage มีการประมวลผลที่เร็วกว่า เนื่องจากมีขั้นตอนการทำงานที่น้อยกว่าในเรื่องของการคำนวณ โดยแบบจำลองการตรวจจับวัตถุแบบ Two-stage ได้แก่ Mask R-CNN, R-CNN, Fast R-CNN และ Faster R-CNN และแบบจำลองตรวจจับวัตถุแบบ Single-stage ได้แก่ YOLO และ SSD ทั้งนี้ ผลลัพธ์จากการตรวจจับวัตถุจะมีกล่องสี่เหลี่ยมรอบวัตถุพร้อมกับระดับความมั่นใจในการตรวจจับ

1.1. Two-Stage Detectors

- **Region-Based Convolutional Neural Network (R-CNN)**

R-CNN เป็นหนึ่งในแบบจำลองตรวจจับวัตถุที่นิยมใช้กันมาก โดยจัดอยู่ในประเภทแบบจำลอง 2 ขั้นตอน (Two-stage object detectors) นำเสนอโดย Girshick et al., (2014) [16] โดย R-CNN เป็นแบบจำลองตรวจจับวัตถุแรกที่ใช้ CNN แบบแบ่งพื้นที่ (Region-based) โดยหลักการ R-CNN ใช้ Selective Search ในการสร้าง ROI หรือบริเวณที่น่าจะมีวัตถุ ประมาณ 2,000 บริเวณ จากนั้นป้อนข้อมูลเหล่านี้เข้าสู่ CNN เพื่อสร้าง Feature Vector ขนาด 4,096 มิติ จากนั้นนำ Feature Vector ที่ได้ไปใช้กับ Support Vector Machine (SVM) เพื่อจำแนกประเภทของวัตถุ นอกจากนี้ ยังวาดกรอบล้อมรอบวัตถุที่ตรวจพบอีกด้วย

- **Fast R-CNN**

หลังจากระบบ R-CNN ถูกเผยแพร่เพียงปีเดียว Girshick (2015) [17] ก็ปล่อย R-CNN เวอร์ชันที่เร็วกว่าออกมาชื่อว่า Fast R-CNN ปัญหาของ R-CNN คือใช้เวลาประมวลผลนาน เนื่องจากใช้การ Convolution กับแต่ละ ROI ที่คาดว่าจะมีวัตถุ ทำให้ช้า Fast R-CNN จึงนำแนวคิดของ ROI มาลดเวลาการทำงาน โดยใช้ชั้น ROI pooling เพื่อลดขนาด Feature map จากแต่ละ ROI ให้เป็นขนาดเดียวกัน ไม่ว่าขนาดของ ROI เดิมจะต่างกัน จากนั้นใช้เทคนิค Max pooling แบ่งพื้นที่ ROI เป็นช่องตารางเล็กๆ แล้วเลือกค่าพิกเซลที่มีค่าสูงสุดในแต่ละช่อง ผลลัพธ์คือ Fast R-CNN ทำงานเร็วกว่า R-CNN เล็กน้อย (mAP 66.9% vs 66.0%) โดยทดสอบบนชุดข้อมูล PASCAL VOC 2007 สำหรับทั้งการจำแนกประเภทวัตถุและปรับกรอบวัตถุ สรุปแล้ว Fast R-CNN พัฒนาต่อยอดจาก R-CNN โดยเน้นประมวลผลเฉพาะบริเวณที่น่าจะมีวัตถุ ผสานเทคนิคหลายอย่างให้ทำงานเร็วขึ้นและแม่นยำขึ้น

- **Faster R-CNN**

Faster R-CNN พัฒนาต่อยอดจาก R-CNN โดยนำเสนอแนวคิด Region Proposal Network (RPN) ที่ Ren et al., (2015) [18] ได้นำเสนอในปี 2016 โดย RPN จะเข้ามาแทนที่ Selective Search ซึ่งเป็นอัลกอริทึมที่ช้าใน Fast R-CNN RPN เป็นเครือข่าย CNN เต็มรูปแบบที่ทำหน้าที่คาดการณ์บริเวณที่น่าจะมีวัตถุ (Region Proposal) ขั้นตอนการทำงานเริ่มจากการสร้าง Anchor Box ซึ่งเป็นกล่องสี่เหลี่ยมรอบๆ บริเวณที่คาดว่าจะมีวัตถุ จากนั้นใช้ฟังก์ชันการสูญเสีย (Loss Function) เพื่อคำนวณหาความน่าจะเป็นของข้อผิดพลาด สุดท้าย เครือข่ายหลัก (Backbone Network) จะสร้าง Feature map และ RPN จะคาดการณ์ชุดของ Region Proposal เหล่านี้จะถูกส่งต่อไปยังชั้นถัดไป นั่นคือ Roi Pooling Layer ชั้นนี้จะแปลง Feature ที่ได้จาก CNN ที่ผ่าน

การปรับแต่ง ให้เป็น Feature map ขนาดคงที่ สุดท้าย ชั้นการจำแนก (Classification Layer) จะทำนายประเภทของวัตถุ ส่วน Bounding Box Regression จะสร้างกล่องสี่เหลี่ยมล้อมรอบวัตถุเพื่อระบุตำแหน่ง

- **Mask R-CNN**

ต่อยอดจาก Faster R-CNN คือ Mask R-CNN ที่ He et al., (2017) [19] นำเสนอ โดยมุ่งเน้นที่การแบ่งส่วนวัตถุ (Instance segmentation) จากภาพ Mask R-CNN เป็นแบบจำลองต่อยอดจาก Faster R-CNN ที่นอกเหนือจากการระบุประเภทและกรอบวัตถุแล้ว ยังสามารถสร้าง Mask ของวัตถุนั้นด้วย การแบ่งส่วนวัตถุอย่างแม่นยำเป็นสิ่งสำคัญสำหรับงานนี้ ดังนั้น Mask R-CNN จึงผสมผสานสองแง่มุมสำคัญของ Computer vision นั่นคือ การตรวจจับวัตถุ (Object detection) เพื่อจำแนกและระบุตำแหน่งวัตถุในภาพ และการแบ่งส่วนความหมาย (Semantic segmentation) เพื่อจำแนกและกำหนดแต่ละพิกเซลให้เป็นหมวดหมู่ที่กำหนดไว้ Mask R-CNN นำเสนอ RoIAlign layer ที่ช่วยแก้ปัญหาความไม่ตรงกันของตำแหน่ง

1.2. Single-Stage Detectors

- **You Only Look Once (YOLO)**

YOLO เป็นอีกหนึ่งแบบจำลองตรวจจับวัตถุประเภท Single-stage และนิยมใช้สำหรับงานตรวจจับวัตถุแบบเรียลไทม์ โดย แบบจำลองนี้มีความโดดเด่นที่การทำนายกรอบวัตถุ (Bounding box) และประเภทของวัตถุพร้อมกันภายในครั้งเดียว เรียกได้ว่าเป็นเครือข่ายแบบ Unified และมีความเร็วสูงเมื่อเทียบกับ Faster R-CNN โดย YOLO ใช้เครือข่ายประสาทเทียมชนิด Convolutional neural network (CNN) แบบเดี่ยว ทั้งนี้ CNN ที่ใช้ใน YOLO เดิมสร้างจากโมเดล GoogLeNet และรุ่นอัปเดตเรียกว่า DarkNet ซึ่งอิงจาก VGG หลักการของ YOLO จะแบ่งภาพที่ป้อนเข้าเป็นตารางกริด โดยแต่ละเซลล์จะทำนายประเภทของวัตถุและกรอบวัตถุโดยตรง ที่มีการกำหนดกรอบวัตถุจำนวนมากก่อนที่จะรวมเป็นผลลัพธ์สุดท้าย [20]

- **Single-Shot Detector (SSD)**

SSD เป็นอีกหนึ่งแบบจำลองตรวจจับวัตถุประเภท Single-stage เหมือนกับ YOLO ที่ Liu et al., (2016) [21] พัฒนา โดยใช้การประมวลผลเพียงครั้งเดียว (single shot) เพื่อตรวจจับวัตถุหลายๆ อย่างภายในภาพ ทั้งการระบุตำแหน่งและประเภทของวัตถุทำพร้อมกันในรอบเดียว

จากข้อมูลข้างต้น สรุปความสามารถในแต่ละหลักการได้ ดังตารางที่ 1 แสดงคุณสมบัติและข้อจำกัดของแบบจำลองตรวจจับวัตถุที่เป็นที่นิยมต่าง ๆ

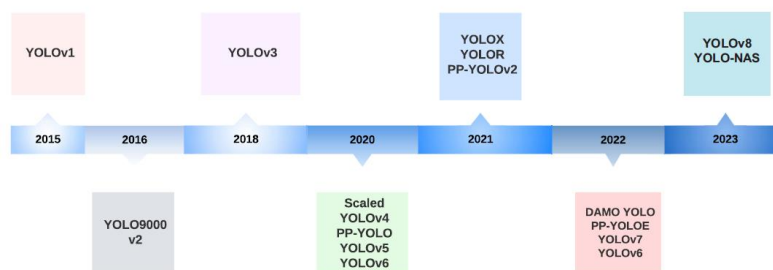
ตารางที่ 1 เปรียบเทียบแบบจำลองตรวจจับวัตถุแบบ Single-stage และ Two-stage [15]

แบบจำลอง	คุณสมบัติ	ข้อจำกัด
R-CNN	ใช้ Selective Search algorithm ค้นหา แยกประเภท และระบุตำแหน่งวัตถุบนภาพด้วยกรอบสี่เหลี่ยม (bounding box)	ค้นหาตำแหน่งวัตถุ (Region proposals) ในภาพได้ถึง 2,000 ตำแหน่ง แต่ใช้เวลานานมากในการฝึกแบบจำลอง จึงไม่สามารถใช้งานจริงแบบ Real-time เพราะการทดสอบซ้ำมาก
Fast R-CNN	มีการพัฒนาเลเยอร์ ROI pooling ที่กำหนดพื้นที่ ROI ขนาดคงที่ ผสมผสานแบบจำลอง 3 แบบจำลอง จาก R-CNN และใช้ Softmax แทน SVM ในการจำแนก	อัลกอริทึม Selective Search ที่ใช้หา Region Proposal ทำให้การประมวลผลช้าและใช้ระยะเวลานาน

แบบจำลอง	คุณสมบัติ	ข้อจำกัด
	ประเภท ทำให้มีความแม่นยำและรวดเร็ว R-CNN เดิม	
Faster R-CNN	ใช้ "Region Proposal Network (RPN)" แทน Selective Search ทำให้การตรวจจับวัตถุแม่นยำ รวดเร็ว และมีประสิทธิภาพดีกว่า Fast R-CNN มาก	การหา Region Proposal ใช้เวลานาน ต้องผ่านหลายรอบกว่าจะเสร็จสิ้นกระบวนการทั้งหมด
YOLO	เมื่อเปรียบเทียบกับอัลกอริทึมตระกูล R-CNN แล้ว YOLO มีความเร็วอย่างมาก โดยใช้ CNN เพียงตัวเดียวทั้งสำหรับการระบุตำแหน่ง (localization) และจำแนกประเภท (classification) เหมาะสำหรับการตรวจจับวัตถุแบบเรียลไทม์ และมีประสิทธิภาพในการแสดงลักษณะของวัตถุที่ดีกว่า	มีข้อจำกัดในการตรวจจับวัตถุขนาดเล็กที่อยู่รวมกัน โดยมีข้อผิดพลาดหลักเกิดจากการระบุตำแหน่งที่ไม่ถูกต้อง และมีปัญหาเกี่ยวกับการแสดงลักษณะของวัตถุที่มีอัตราส่วนภาพไม่ปกติ
SSD	สามารถฝึกแบบ End-to-end ได้ โดยใช้ CNN filters ขนาดเล็กสำหรับการคาดการณ์ประเภทของวัตถุ	มีความแม่นยำมากกว่า YOLO แต่ช้ากว่าเล็กน้อย แต่เร็วกว่า Faster R-CNN และค่าความแม่นยำด้อยกว่า
Mask R-CNN	ใช้สำหรับ Instance Segmentation และมีการใช้ RoIAlign layer เพื่อเพิ่มความแม่นยำในการแบ่งส่วนวัตถุ	การจำแนกประเภทวัตถุขึ้นอยู่กับกระบวนการแบ่งส่วนวัตถุที่ถูกต้อง

2. You Only Look Once (YOLO) [22]

อัลกอริทึม YOLO หรือ "You Only Look Once" เปรียบเสมือนผู้บุกเบิกที่พลิกโฉมหน้าวงการตรวจจับวัตถุ ด้วยวิธีการอันเป็นเอกลักษณ์ ซึ่งต่างจากวิธีดั้งเดิมที่ยุ่งยากและซับซ้อน แบ่งเป็นหลายขั้นตอน เริ่มต้นด้วยการหาตำแหน่งวัตถุ (Region proposals) และตามด้วยการจำแนกประเภท แต่ YOLO ทำงานด้วยความรวดเร็วเพียงรอบเดียว (Single forward pass) โดยแบ่งภาพเป็นตาราง แล้วคาดการณ์ทั้งกล่อง (Bounding boxes) และประเภทของวัตถุพร้อมกัน ภายในแต่ละช่อง ส่งผลให้ทำงานได้อย่างรวดเร็วและมีความแม่นยำสูง ซึ่ง YOLO จะมีการพัฒนาดังแสดงในภาพประกอบที่ 1



ภาพที่ 1 A timeline of YOLO versions. [23]

ตารางที่ 2 ตารางแสดงรายละเอียดอัลกอริทึม YOLO ในแต่ละเวอร์ชัน [22]

เวอร์ชัน	รายละเอียด
YOLOv1	YOLOv1 เปิดตัวครั้งแรก ถือเป็นการปฏิวัติวงการตรวจจับวัตถุ ด้วยแนวคิดใหม่ของแบบจำลองการตรวจจับแบบ Single - stage ซึ่งแตกต่างจากวิธีเดิมที่ต้องผ่านหลายขั้นตอน YOLOv1 ไม่ต้องใช้กลไกการเสนอพื้นที่ (Region proposal) ทำให้สามารถวิเคราะห์ภาพทั้งหมดได้ในคราวเดียว ได้อย่างมีประสิทธิภาพ
YOLOv2 และ YOLOv3	YOLOv2 และ YOLOv3 พัฒนาต่อยอดจาก YOLOv1 ด้วยเทคนิคที่ล้ำสมัย เช่น โครงสร้าง Feature Pyramid Networks (FPN) การฝึกแบบจำลองแบบหลายขนาด และกล่อง Anchor boxes ส่งผลให้สามารถตรวจจับวัตถุได้แม่นยำและหลากหลายขึ้น โดย FPN ช่วยให้แบบจำลองเรียนรู้รายละเอียดวัตถุได้หลายขนาด การฝึกโมเดลแบบหลายขนาดช่วยให้โมเดลตรวจจับวัตถุได้หลากหลายขนาด และกล่อง Anchor boxes ช่วยให้แบบจำลองคาดการณ์ตำแหน่งวัตถุได้แม่นยำขึ้น
YOLOv4	ผู้พัฒนาได้ทุ่มเทอย่างจริงจังในการพัฒนา YOLOv4 เพื่อนำการเปลี่ยนแปลงที่ล้ำสมัยมาสู่ YOLO โดยได้มีการนำเสนอการเปลี่ยนแปลงในสถาปัตยกรรม Darknet53 ซึ่งกลายเป็น CSPDarknet53-PANet-SPP เทคนิคต่างๆ เช่น bag of specials and freebies อัลกอริทึมวิวัฒนาการทางพันธุกรรม โมดูล Attention และอื่นๆ
YOLOv5	YOLOv5 แม้จะพัฒนาต่อยอดจาก YOLOv4 ด้วยการเปลี่ยนแปลงเพียงเล็กน้อยและยังคงเทคนิคส่วนใหญ่ไว้ แต่สิ่งที่ทำให้ YOLOv5 โดดเด่นคือการใช้งาน PyTorch ทำให้การใช้งานและฝึกแบบจำลองได้อย่างง่ายดายยิ่งขึ้น โดย Ultralytics ได้ออกแบบ YOLOv5 ให้มีการใช้งานที่ง่าย การติดตั้งที่สะดวก และมีคำแนะนำการใช้งานอย่างชัดเจน
YOLO-R	ใน YOLO-R ผู้พัฒนาได้ทดลองใช้วิธีการใหม่และมุ่งพัฒนาแบบจำลองเพื่อรองรับการเรียนรู้แบบหลายงาน (Multi-task learning) เป้าหมายหลักคือเพื่อขยายขีดความสามารถของโมเดล YOLO ให้สามารถทำงานได้หลากหลายมากขึ้น โดยไม่ต้องพึ่งพาแบบจำลองแยกต่างหากสำหรับงานแต่ละประเภท
YOLO-X	YOLO-X เป็นแบบจำลองตรวจจับวัตถุที่พัฒนาต่อยอดจาก YOLOv3 โดยนำเทคนิคใหม่ ๆ มาผสมรวมเข้าด้วยกัน เช่น Anchor-free, Decoupled heads, Label assignment และ strong augmentations ช่วยให้ YOLO-X มีประสิทธิภาพและความแม่นยำในการตรวจจับวัตถุที่ดีขึ้น
YOLOv6 และ YOLOv7	ใน YOLOv7 และ YOLOv6 ผู้เขียนได้ทำการทดลองกับโครงสร้างของแบบจำลอง โดยใน YOLOv7 มีการนำเสนอแนวคิดใหม่เกี่ยวกับการปรับพารามิเตอร์ใน YOLO และการปรับขนาดของแบบจำลองเป็นครั้งแรก ในขณะเดียวกัน YOLOv6 มีการเพิ่มเทคนิคของการกลั่น (Distillation) และการควอนไทเซชัน (Quantization) เข้าไปในแบบจำลองด้วย โดยแบบจำลอง YOLOv6 เหมาะสำหรับใช้งานแบบ Real-time บนอุปกรณ์ที่มีทรัพยากรจำกัด ต้องการความเร็ว และ YOLOv7 เหมาะกับงานที่ต้องการความแม่นยำสูง บนอุปกรณ์ที่มีทรัพยากรสูง
YOLOv8	YOLOv8 นั้นมีประสิทธิภาพที่ดีกว่า เมื่อเทียบกับแบบจำลองรุ่นก่อนๆ อย่าง YOLOv3 หรือ YOLOv5 อาจดูเหมือนว่าแบบจำลองรุ่นก่อนๆ นั้นออกแบบมาซับซ้อนเกินความจำเป็น เพราะ YOLOv8 นั้นพัฒนาต่อยอดจาก YOLOv5 โดยมีการเปลี่ยนแปลงโครงสร้างเพียงเล็กน้อย แต่กลับมีประสิทธิภาพที่เหนือกว่าสาเหตุหลักๆ ที่ทำให้ YOLOv8 มีประสิทธิภาพดีกว่า เนื่องจากการเปลี่ยนแปลง 2 ประการ ได้แก่ 1) การแทนที่ CSPLayer ด้วยโมดูล C2f: CSPLayer นั้นเป็นเลเยอร์หลักใน YOLOv5 ที่ทำหน้าที่เพิ่มประสิทธิภาพและลดขนาดของแบบจำลอง แต่โมดูล C2f นั้นมีประสิทธิภาพที่เหนือกว่า CSPLayer จึงทำให้ YOLOv8 ทำงานได้เร็วขึ้นและแม่นยำยิ่งขึ้น 2) การเพิ่มฟังก์ชันการสูญเสียที่ดีกว่า ฟังก์ชันการสูญเสียมีหน้าที่ในการคำนวณความแตกต่างระหว่างผลลัพธ์ที่คาดหวังกับผลลัพธ์ที่ได้จริงของแบบจำลอง การเพิ่มฟังก์ชันการสูญเสียที่ดีกว่า ช่วยให้

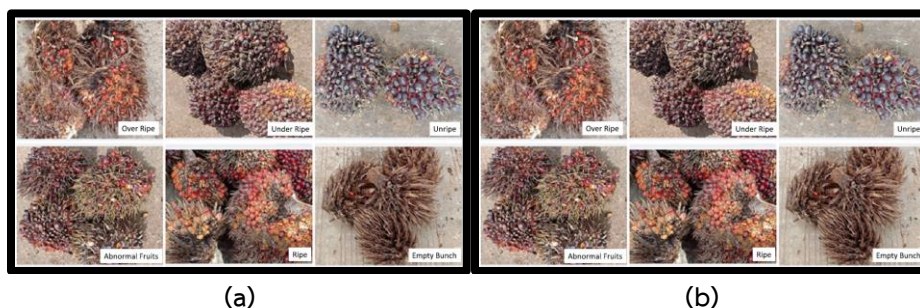
เวอร์ชัน	รายละเอียด
	โมเดลเรียนรู้ได้ดีขึ้น และสามารถตรวจจับวัตถุขนาดเล็กได้แม่นยำยิ่งขึ้น โดยสรุป YOLOv8 นั้นเป็นแบบจำลองตรวจจับวัตถุที่มีประสิทธิภาพสูง ใช้งานง่าย และมีขนาดเล็ก เหมาะสำหรับการใช้งานบนอุปกรณ์ที่มีทรัพยากรจำกัด เช่น โทรศัพท์มือถือ หรือกล้องวงจรปิด
YOLOv-NAS	YOLO-NAS เป็นแบบจำลองตรวจจับวัตถุที่มีจุดเด่นคือ Quantization Blocks เป็นเทคโนโลยีที่ล้ำสมัยที่ช่วยเพิ่มความเร็วและความแม่นยำ ทั้งนี้ เทคโนโลยี Quantization Blocks ช่วยให้ YOLO-NAS ประหยัดพลังงาน

วิธีดำเนินการ

ในการวิจัยครั้งนี้ ผู้วิจัยมุ่งเน้นศึกษาวิธีการจำแนกความสุกของปาล์มน้ำมันสด เพื่อช่วยให้สามารถจำแนกความสุกของปาล์มน้ำมันสดได้อย่างถูกต้องแม่นยำ ทั้งนี้ เพื่อป้องกันความผิดพลาดในขั้นตอนเก็บเกี่ยว หรือขั้นตอนการคัดแยกผลปาล์ม ณ โรงงานสกัดน้ำมันปาล์ม เป็นต้น โดยผู้วิจัยเลือกใช้อัลกอริทึมที่สามารถพัฒนาไปสู่การใช้งานแบบเรียลไทม์ หรือใช้งานบนโทรศัพท์มือถือได้ เพื่อนำไปสู่การประยุกต์ใช้งานได้ในสถานการณ์จริงในอนาคต ผู้วิจัยได้เลือกใช้หลักการโครงข่ายประสาทเทียมแบบคอนโวลูชัน (CNN) ที่มีพื้นฐานจากวิธีการเรียนรู้เชิงลึก มีประสิทธิภาพและสามารถนำมาประยุกต์ใช้ในงานวิจัยนี้ได้ ทั้งนี้ ผู้วิจัยได้ออกแบบขั้นตอนในการดำเนินงานวิจัยโดยแบ่งออกเป็น 4 ขั้นตอน ดังต่อไปนี้

ขั้นตอนที่ 1 : การเก็บรวบรวมข้อมูล (Data Acquisition)

ข้อมูลที่ใช้ในงานวิจัยนี้ เป็นข้อมูลที่อยู่ในรูปแบบของข้อมูลรูปภาพผลปาล์มน้ำมันสด โดยผู้วิจัยได้นำชุดข้อมูลดังกล่าวมาจากเว็บไซต์ Science Data Bank [24] ซึ่งเป็นชุดข้อมูลที่ถูกรวบรวมและจัดเก็บไว้เพื่อให้ผู้ที่สนใจนำข้อมูลไปใช้ประโยชน์ในการพัฒนาแบบจำลองที่สามารถใช้สร้างแอปพลิเคชันบนสมาร์ตโฟน หรือแอปพลิเคชันที่ฝังอยู่ในหุ่นยนต์หรืออุปกรณ์อื่น ๆ ได้ เพื่อช่วยให้การจำแนกระดับความสุกของผลปาล์มน้ำมันสดมีความแม่นยำและถูกต้อง โดยชุดข้อมูลต้นฉบับมีระดับความสุกของปาล์มน้ำมัน จำนวน 6 หมวดหมู่ ได้แก่ ดิบ (Unripe) กึ่งสุกกึ่งดิบ (Under-ripe) สุกพอดี (Ripe) สุกเกินไป (Over-ripe) ทะลายเปล่า (Empty bunch) และทะลายผิดปกติ (Abnormal FFB) โดยชุดข้อมูลเป็นรูปภาพต่อเนื่องที่ถูกแยกเฟรมมาจากวิดีโอของทะลายปาล์มน้ำมันสด มีจำนวนทั้งหมด 4,160 ภาพ มีความละเอียดเท่ากับ 416x416 และเป็นรูปภาพที่ผ่านการกำหนดป้ายชื่อประเภทระดับความสุก หรือการทำ Labelling เรียบร้อยแล้ว โดยข้อมูลภาพจะมี 2 ลักษณะ ได้แก่ ภาพของทะลายปาล์มน้ำมันสดที่มีระดับความสุกหมวดหมู่เดียวกัน ดังตัวอย่างภาพประกอบที่ 2(a) และจากภาพที่ 2(b) แสดงภาพของกองทะลายปาล์มน้ำมันสดที่มีระดับความสุกแตกต่างกัน และข้อมูลของรูปภาพที่ได้รับมาทั้งหมด ดังแสดงในตารางที่ 3



ภาพที่ 2 ตัวอย่างภาพกองทะลายปาล์มน้ำมันสด (a) ที่มีระดับความสุกหมวดหมู่เดียวกัน (b) ที่มีระดับความสุกแตกต่างกัน [25]

ตารางที่ 3 แสดงการกระจายของรูปภาพ และ Object สำหรับแต่ละหมวดหมู่ [25]

Category of FFB	Objects	Image
Unripe	2,913	1,130
Under-ripe	2,575	1,289
Ripe	2,973	1,880
Over-ripe	2,641	1,162
Empty bunch	857	473
Abnormal FFB	2,599	1,237
Total	14,558	7,171

ชุดข้อมูลต้นฉบับได้ถูกแบ่งออกเป็น ชุดข้อมูลสำหรับฝึกอบรมแบบจำลอง (Training dataset) ข้อมูลสำหรับตรวจสอบความถูกต้องแบบจำลอง (Validation dataset) และชุดข้อมูลสำหรับการทดสอบแบบจำลอง (Testing dataset) มีอัตราส่วน 70:20:10 ดังตารางที่ 4

ตารางที่ 4 อัตราส่วนการแบ่งชุดข้อมูลต้นฉบับ [25]

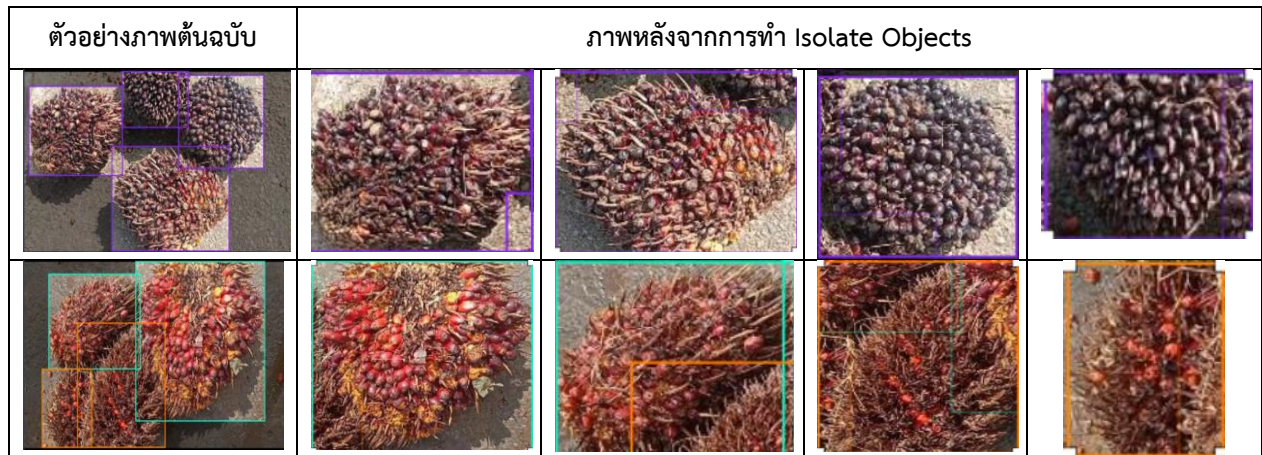
ชุดข้อมูล	จำนวน (ภาพ)	อัตราส่วน (%)
Training dataset	2,908	70
Validation dataset	835	20
Testing dataset	417	10
Total	4,160	100

ขั้นตอนที่ 2 : การเตรียมข้อมูล (Data Preparation)

ผู้วิจัยใช้ภาพกองทะลายปาล์มน้ำมันสดที่ผ่านการทำ Labelling เรียบร้อยแล้ว ซึ่งเป็นข้อมูลภาพต้นฉบับ มาทำ Data preparation ด้วยเครื่องมือ Roboflow โดย Roboflow เป็นเครื่องมือที่ช่วยในการจัดการและปรับปรุงข้อมูลภาพ (Image data) สำหรับการฝึกอบรมแบบจำลองทางด้าน Computer vision โดย Roboflow มีฟีเจอร์ต่าง ๆ ที่ช่วยให้ง่ายต่อการเตรียมข้อมูล สำหรับการฝึกอบรมแบบจำลอง โดยมีขั้นตอน ดังนี้

- การเตรียมข้อมูล (Data Preprocessing)

เพื่อให้การฝึกอบรมแบบจำลองมีความแม่นยำและมีประสิทธิภาพ ผู้วิจัยได้ใช้วิธีการทำ Isolate objects เพื่อตัดภาพทะลายปาล์มน้ำมันออกเป็น 1 ทะลาย ต่อ 1 ภาพ จากเดิมเป็นรูปภาพกองทะลายปาล์มน้ำมันสด ซึ่งมีปาล์มน้ำมันหลายทะลาย ใน 1 ภาพ ดังภาพที่ 3 และทำการ Modify classes โดยเลือกใช้หมวดหมู่ระดับความสุกของปาล์มน้ำมัน สำหรับการดำเนินการวิจัย เพียง 3 หมวดหมู่ จากทั้งหมด 6 หมวดหมู่ ได้แก่ ไม่สุกหรือดิบ (Unripe) สุกพอดี (Ripe) และ สุกเกินไป (Over-ripe) ทั้งนี้ สาเหตุที่เลือกใช้ 3 หมวดหมู่ดังกล่าว เนื่องจากเป็นหมวดหมู่ที่เกิดความผิดพลาดในการจำแนกระดับความสุกด้วยสายตาของมนุษย์ได้ง่าย



ภาพที่ 3 แสดงตัวอย่างการทำ Objects Isolation

- **การเพิ่มปริมาณข้อมูลภาพสอน (Data Augmentations)**

การเพิ่มปริมาณข้อมูลภาพสอน (Data augmentations) เป็นการสร้างข้อมูลรูปภาพให้มีความหลากหลายและมีจำนวนเพิ่มขึ้น ผู้วิจัยได้เลือกใช้วิธีการต่าง ๆ ได้แก่ การกลับด้านรูปภาพ (Flip) การตัดบางส่วนของรูปภาพ (Crop) การหมุนรูปภาพ (Rotation) การปรับแสงสว่าง (Brightness) และการเพิ่มสิ่งรบกวนลงในรูปภาพ (Noise)

- **การแบ่งข้อมูล (Split Data)**

ผลจากการทำ Data preparation ทำให้ชุดข้อมูลที่ใช้สำหรับดำเนินงานวิจัยมีรูปภาพทั้งหมดจำนวน 14,477 ภาพ โดยชุดข้อมูลจะถูกแบ่งออกเป็น 3 ชุด ได้แก่ ชุดข้อมูลสำหรับฝึกอบรมแบบจำลอง (Training dataset) จำนวน 11,900 ภาพ ชุดข้อมูลสำหรับตรวจสอบความถูกต้องแบบจำลอง (Validation dataset) จำนวน 1,782 ภาพ และชุดข้อมูลสำหรับการทดสอบแบบจำลอง (Testing dataset) จำนวน 795 ภาพ ดังตารางที่ 5

ตารางที่ 5 แสดงจำนวนชุดข้อมูลสำหรับดำเนินงานวิจัย

Class	Train	Validation	Test	Total
Unripe	4,010	642	266	4,918
Ripe	4,284	585	246	5,115
Over-ripe	3,606	555	283	4,444
Total	11,900	1,782	795	14,477

ขั้นตอนที่ 3 : การฝึกฝนแบบจำลอง (Training and Validation Model)

สำหรับขั้นตอนการฝึกฝนแบบจำลอง ผู้วิจัยได้เลือกใช้แบบจำลองที่สามารถพัฒนาไปสู่การใช้งานแบบเรียลไทม์ได้ในอนาคต โดยเลือกใช้แบบจำลอง You Only Look Once หรือ YOLO ที่มีการระบุตำแหน่งของวัตถุด้วยกรอบสี่เหลี่ยม (Bounding box) เพื่อระบุพิกัดของวัตถุในภาพ พร้อมทั้งสามารถระบุชนิดหรือประเภทของวัตถุต่าง ๆ ในภาพได้ ทั้งนี้ ผู้วิจัยได้เลือกใช้แบบจำลอง YOLOv8 ซึ่งเป็นเวอร์ชันใหม่ล่าสุดที่ถูกพัฒนาขึ้นมาในปี พ.ศ. 2566 โดย YOLOv8 มีแบบจำลองต่างๆ ได้แก่ YOLOv8n (nano), YOLOv8s (small), YOLOv8m (medium), YOLOv8l (large), YOLOv8x ดังตารางที่ 6 แสดงค่า mAP และค่า FPS ของ YOLOv8 แต่ละแบบจำลอง ทั้งนี้ ผู้วิจัยเลือกใช้แบบจำลองที่มีขนาดเล็กอย่าง YOLOv8n (nano) เนื่องจากมีจำนวน

พารามิเตอร์น้อย ประมวลผลเร็ว และใช้พื้นที่หน่วยความจำในการประมวลผลน้อย และมีความแม่นยำสูง ด้วยเหตุผลดังกล่าวข้างต้น จึงทำให้แบบจำลองขนาดเล็กอย่าง YOLOv8n เหมาะสำหรับแอปพลิเคชันที่ต้องใช้การประมวลผลแบบเรียลไทม์ และเหมาะสมกับอุปกรณ์ที่มีทรัพยากรจำกัด โดยแบบจำลองดังกล่าวได้รับการฝึกฝนด้วยชุดข้อมูล Common Objects in Context หรือ COCO โดย COCO เป็นชุดข้อมูลภาพขนาดใหญ่สำหรับการฝึกอบรมแบบจำลองการรู้จำวัตถุ และการตรวจจับวัตถุ ชุดข้อมูลประกอบด้วยภาพมากกว่า 330,000 ภาพ แบ่งออกเป็น 80 คลาส ของวัตถุ ดังนั้นเพื่อปรับปรุงประสิทธิภาพของแบบจำลองดังกล่าว เพื่อใช้ในการจำแนกระดับความสุกของผลปาล์มน้ำมันสด ผู้วิจัยจึงต้องฝึกฝนแบบจำลองด้วยชุดข้อมูลที่ได้เตรียมไว้สำหรับงานวิจัยนี้ เพื่อให้แบบจำลองสามารถทำนายระดับความสุกของผลปาล์มน้ำมันสดได้ตามที่กำหนด

ตารางที่ 6 แสดงค่าความสามารถของแบบจำลอง YOLOv8 [26]

Model	Size (pixel)	mAP ^{val} 50-95	Speed	Speed	Params (M)	FLOPs (B)
			CPU ONNX (ms)	A100 TensorRT (ms)		
YOLOv8n	640	37.3	80.4	0.99	3.2	8.7
YOLOv8s	640	44.9	128.4	1.20	11.2	28.6
YOLOv8m	640	50.2	234.7	1.83	25.9	78.9
YOLOv8l	640	52.9	375.2	2.39	43.7	165.2
YOLOv8x	640	53.9	479.1	3.53	68.2	257.8

*mAPval values are for single-model single-scale on COCO val2017 dataset. Reproduce by yolo val detect data=coco.yaml device=0

*Speed averaged over COCO val images using an Amazon EC2 P4d instance. Reproduce by yolo val detect data=coco128.yaml batch=1 device=0|cpu

ขั้นตอนที่ 4 : การประเมินผลการทดลอง (Model Evaluation)

การประเมินประสิทธิภาพของแบบจำลองตรวจจับวัตถุ แบ่งออกเป็น 2 ส่วน ดังนี้

- การประเมินตำแหน่งที่สามารถตรวจจับวัตถุได้

การประเมินตำแหน่งที่สามารถตรวจจับวัตถุได้ วัดได้จากตัวชี้วัดความถูกต้องของการตรวจจับวัตถุ IoU (Intersection over Union) โดยการวัดความซ้อนทับระหว่าง Bounding box ที่แบบจำลองทำนาย และ Bounding box จริงของวัตถุที่อยู่ในภาพ ซึ่งมีความสำคัญในการประเมินคุณภาพของการตรวจจับวัตถุ โดยค่า IoU เป็นค่าที่อยู่ในช่วง [0, 1] ที่ใช้เป็นเกณฑ์ตัดสินใจว่าการตรวจจับวัตถุนั้นถูกต้องหรือไม่ โดยทั่วไปแล้ว ค่า IoU ที่มากกว่าหรือเท่ากับค่าที่กำหนด (เช่น 0.5) จะถือว่าเป็น Detection ถูกต้อง (True Positive) ในขณะที่ค่า IoU ที่น้อยกว่าค่าที่กำหนดจะถือว่าเป็น Detection ไม่ถูกต้อง (False Positive) หรือถ้าไม่มีการตรวจจับวัตถุเลยจะเป็น False Negative ดังสมการที่ 1

$$IoU = \frac{\text{Intersection Area}}{\text{Union Area}} \quad (1)$$

- การประเมินความสามารถในการจำแนกวัตถุ

การประเมินความสามารถในการจำแนกวัตถุ สามารถประเมินได้จากตัวชี้วัดต่างๆ ดังนี้

1) Confusion Matrix ใช้ในการประเมินผลการทำนาย หรือ Prediction ที่ทำนายจากแบบจำลองที่สร้างขึ้น เพื่อแสดงผลลัพธ์ของการทำนายว่าตรงกับค่าจริงมากน้อยเพียงใด ดังแสดงในตารางที่ 7 ที่ประกอบด้วยค่าต่างๆ ดังนี้ 1) True Positive (TP): แบบจำลองทำนายถูก ว่าข้อมูลเป็นคลาส Positive และข้อมูลนั้นเป็นคลาส Positive 2) False Positive (FP): แบบจำลองทำนายผิด ว่าข้อมูลเป็นคลาส Positive แต่ข้อมูลนั้นเป็นคลาส Negative 3) True Negative (TN): แบบจำลองทำนายถูก ว่าข้อมูลเป็นคลาส Negative และข้อมูลนั้นเป็นคลาส Negative 4) False Negative (FN): แบบจำลองทำนายผิด ว่าข้อมูลเป็นคลาส Negative แต่ข้อมูลนั้นเป็นคลาส Positive

ตารางที่ 7 ตารางแสดงค่า Confusion Matrix [27]

	Actual Positive (1)	Actual Negative (0)
Predicted Positive (1)	True Positive	False Positive
Predicted Negative (0)	False Negative	True Negative

2) Precision เป็นหนึ่งในตัวชี้วัดความแม่นยำของแบบจำลองในงานจำแนกวัตถุ (Object Classification) โดยค่า Precision จะวัดความแม่นยำของการทำนายว่าเป็นคลาสบวก (Positive class) โดยการนับจำนวน True Positives (TP) และ False Positives (FP) ทั้งนี้ ค่า Precision มีค่าอยู่ในช่วง [0, 1] โดยค่าที่มากขึ้นแสดงว่าการทำนายว่าเป็นคลาสบวกของแบบจำลองมีความแม่นยำมากขึ้น ดังสมการที่ 2

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \quad (2)$$

3) Recall เป็นตัวชี้วัดที่ใช้วัดสัดส่วนของค่า Positive จริงทั้งหมดที่ทำนายถูกต้อง คือ True Positives หาด้วยทั้งหมดของ Ground truths (ค่าจริงทั้งหมด) เพื่อประเมินว่าแบบจำลองสามารถระบุข้อมูลที่เป็ค่าบวก (Positive) ได้ทั้งหมดมากเพียงใด มีค่าอยู่ในช่วง [0, 1] โดยค่าที่มากขึ้นแสดงว่าแบบจำลองมีความสามารถในการตรวจจับคลาสบวกมีประสิทธิภาพมากขึ้น ดังสมการที่ 3

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \quad (3)$$

4) Average Precision (AP) เป็นหนึ่งในตัวชี้วัดที่ใช้ในการประเมินความแม่นยำของแบบจำลองในงานตรวจจับวัตถุ (Object detection) คำนีใช้ในการวัดประสิทธิภาพของโมเดลในการตรวจจับและจำแนกวัตถุที่มีหลายหมวดหมู่ คำนวณจาก Precision-Recall Curve ซึ่งเป็นกราฟที่แสดงความสัมพันธ์ระหว่าง Precision และ Recall ที่เปลี่ยนแปลงตามการปรับค่าความเชื่อมั่น (Confidence threshold) ของแบบจำลอง มีค่าอยู่ในช่วง [0, 1] โดยค่าที่สูงกว่าแสดงถึงประสิทธิภาพที่ดีขึ้นในการตรวจจับและจำแนกวัตถุ ดังสมการที่ 4

$$AP = \int_0^1 p(r)dr \quad (4)$$

5) Mean Average Precision (mAP) เป็นตัวชี้วัดที่ใช้ในการประเมินประสิทธิภาพของโมเดลในงานตรวจจับวัตถุ (Object detection) ที่มีหลายหมวดหมู่ โดยการคำนวณค่า Average Precision (AP) ของแต่ละคลาสแยกต่างหาก แล้วนำมาหาค่าเฉลี่ย ซึ่งสะท้อนถึงประสิทธิภาพโดยรวมของแบบจำลองในการตรวจจับวัตถุจากคลาสต่าง ๆ ดังสมการที่ 5 ตัวอย่างเช่น ถ้าแบบจำลองสามารถตรวจจับคลาส "คน" ได้ด้วย AP = 0.8 และคลาส "รถ" ได้ด้วย AP = 0.7 ดังนั้น ค่า mAP ของแบบจำลองจะเท่ากับ $(0.8 + 0.7) / 2 = 0.75$

$$mAP = \frac{1}{n} * \sum(AP) \quad (5)$$

ผลการวิจัยและอภิปรายผลการวิจัย

งานวิจัยนี้ได้ดำเนินการโดยใช้ภาษา Python และ ultralytics libraries โดยใช้ GPU T4 ด้วย Google Colab โดยได้ทำการฝึกฝนแบบจำลอง YOLOv8n (Pretrained weight) ซึ่งเป็น Weight ที่ได้จากการฝึกฝนด้วยชุดข้อมูล COCO test-dev 2017 โดยมีการกำหนดค่าพารามิเตอร์ต่าง ๆ สำหรับปรับแต่งแบบจำลอง

การปรับแต่งแบบจำลอง YOLOv8n

ในงานวิจัยนี้ได้ทำการปรับแต่งแบบจำลองโดยการกำหนดค่าพารามิเตอร์ต่าง ๆ เพื่อหาแบบจำลอง YOLOv8n ที่ดีที่สุดสำหรับการจำแนกระดับความสูงของปาล์มน้ำมันสด

- **Optimizer**

เป็นอัลกอริทึมที่ใช้ในกระบวนการฝึก (Training) แบบจำลอง เป้าหมายหลักของ Optimizer คือการปรับค่าพารามิเตอร์เพื่อลดค่าฟังก์ชันความสูญเสีย (Loss function) หรือเพื่อให้แบบจำลองทำนายผลลัพธ์ที่ถูกต้องมากขึ้น นอกจากนี้ยังมีผลต่อความเร็วและประสิทธิภาพในการฝึกแบบจำลองอีกด้วย โดยในงานวิจัยนี้ใช้ Optimizer ที่แตกต่างกัน 2 ชนิดเพื่อเปรียบเทียบกัน ได้แก่

- 1) Stochastic Gradient Descent (SGD)
- 2) Adaptive Moment Estimation (Adam)

- **Learning rate**

Learning rate คือค่าที่กำหนดขนาดของการเปลี่ยนแปลงพารามิเตอร์ (Weight) ของแบบจำลองในการเรียนรู้แต่ละรอบ เปรียบเสมือนความเร็วในการปรับตัวของแบบจำลอง หน้าที่หลักของ Learning rate คือการควบคุมขนาดของการเปลี่ยนแปลงพารามิเตอร์ของแบบจำลองตามผลลัพธ์ที่ได้จากการฝึกฝน หาก Learning rate มีค่าสูง แบบจำลองก็จะปรับตัวเร็วขึ้น แต่เสี่ยงที่จะกระโดดข้ามจุดที่เป็นค่าตอบที่ดีที่สุด (Global minimum) ได้ ในทางกลับกัน ถ้า Learning rate มีค่าต่ำ แบบจำลองจะปรับตัวช้าลง ใช้เวลานานในการฝึกฝนนาน และอาจติดอยู่ในหลุมค่าตอบที่ไม่ดี (Local minimum) ได้ โดยในงานวิจัยนี้ได้ทดลองใช้ Learning rate ด้วยค่าที่แตกต่างกัน ได้แก่ 0.01 0.001 และ 0.0001

- **Batch Size**

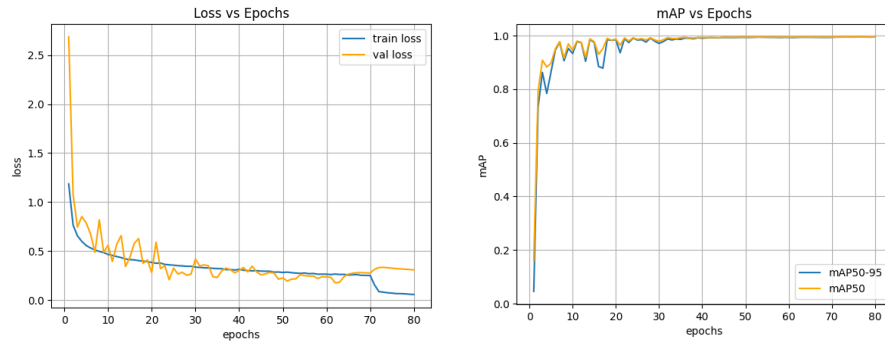
Batch size คือ จำนวนตัวอย่างข้อมูล (รูปภาพ) ที่แบบจำลองจะใช้ในการปรับพารามิเตอร์แต่ละรอบของการเรียนรู้ ทั้งนี้ ขนาดของ Batch size ส่งผลต่อการเรียนรู้ของแบบจำลองในด้านต่าง ๆ ได้แก่ ความเร็วในการเรียนรู้ ความเสถียรของการเรียนรู้ การใช้หน่วยความจำ เป็นต้น โดยในงานวิจัยนี้ได้เลือกใช้ Batch size จำนวน 4 ขนาด ได้แก่ 16 32 64 และ 128

ในงานวิจัยนี้ได้ทำการปรับแต่งแบบจำลอง YOLOv8n ด้วยการกำหนดค่าพารามิเตอร์ต่าง ๆ ได้แก่ Optimizer จำนวน 2 ชนิด Learning rate จำนวน 3 ค่า และ Batch Size จำนวน 4 ขนาด และ กำหนดจำนวนรอบในการฝึกฝนแบบจำลองเท่ากับ 80 Epochs ดังนั้น ทำให้ได้แบบจำลองเพื่อทำการเปรียบเทียบประสิทธิภาพทั้งหมดจำนวน 24 แบบจำลอง ดังแสดงผลลัพธ์ในตารางที่ 8

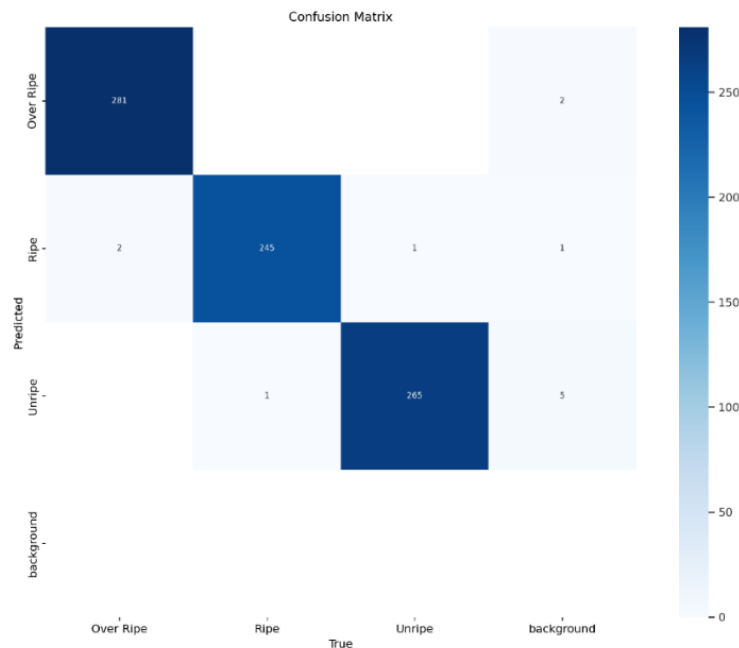
ตารางที่ 8 แสดงผลการทดสอบประสิทธิภาพแบบจำลอง ด้วยการกำหนดค่าพารามิเตอร์ต่าง ๆ

Parameter			Result				
Optimizer	Learning rate	Batch size	Precision	Recall	mAP50	mAP50-95	Training time (hrs)
SGD	0.01	16	0.995	0.997	0.995	0.994	1.643
SGD	0.01	32	0.995	0.997	0.995	0.995	1.229
SGD	0.01	64	0.997	0.996	0.995	0.995	1.000
SGD	0.01	128	0.997	0.997	0.995	0.995	0.942
SGD	0.001	16	0.995	0.995	0.993	0.993	1.730
SGD	0.001	32	0.995	0.995	0.994	0.994	1.122
SGD	0.001	64	0.996	0.995	0.995	0.995	1.074
SGD	0.001	128	0.994	0.995	0.994	0.994	0.953
SGD	0.0001	16	0.960	0.966	0.988	0.988	1.664
SGD	0.0001	32	0.948	0.967	0.986	0.986	1.225
SGD	0.0001	64	0.956	0.940	0.983	0.983	1.093
SGD	0.0001	128	0.944	0.937	0.975	0.975	1.082
Adam	0.01	16	0.972	0.973	0.991	0.984	1.757
Adam	0.01	32	0.928	0.945	0.979	0.968	1.261
Adam	0.01	64	0.962	0.943	0.984	0.969	1.638
Adam	0.01	128	0.942	0.946	0.986	0.979	1.478
Adam	0.001	16	0.993	0.994	0.995	0.995	1.739
Adam	0.001	32	0.998	0.996	0.995	0.995	1.189
Adam	0.001	64	0.995	0.996	0.995	0.995	1.649
Adam	0.001	128	0.997	0.996	0.995	0.995	1.478
Adam	0.0001	16	0.993	0.997	0.994	0.994	1.827
Adam	0.0001	32	0.998	0.997	0.995	0.995	1.900
Adam	0.0001	64	0.990	0.994	0.995	0.995	1.601
Adam	0.0001	128	0.992	0.990	0.994	0.994	1.431

จากตารางที่ 8 การประเมินประสิทธิภาพแบบจำลองด้วยชุดข้อมูล Testing dataset จะเห็นว่า YOLOv8n มีประสิทธิภาพดีที่สุด ด้วยค่าพารามิเตอร์ต่าง ๆ ดังนี้ Adam Optimizer, Learning rate = 0.0001, Batch size = 32 ดังภาพที่ 4 แสดงกราฟค่าความสูญเสียและค่าความแม่นยำระหว่างการฝึกอบรมแบบจำลอง ทั้งนี้ ผลลัพธ์จากการประเมินแบบจำลองมีค่า Precision = 0.998, Recall = 0.997, mAP50 = 0.995 และ mAP50-95 = 0.995



ภาพที่ 4 กราฟแสดงค่าความสูญเสีย (loss) และค่าความแม่นยำ (mAP) ของ YOLOv8n ที่มีประสิทธิภาพดีที่สุด



ภาพที่ 5 ภาพแสดง Confusion Matrix ของแบบจำลอง YOLOv8n ที่มีประสิทธิภาพดีที่สุด ประเมินด้วยชุดข้อมูล Testing dataset

จาก Confusion Matrix ที่แสดงในรูปภาพที่ 5 สามารถสรุปผลได้ดังนี้

1) คลาสที่สุกเกินไป (Ove-ripe) มีจำนวนรูปภาพสำหรับการทดสอบทั้งหมดจำนวน 283 ภาพ มีค่าความแม่นยำในการทำนายของคลาสนี้มีค่าเท่ากับ 99.29% โดยทำนายถูกต้องจำนวน 281 ภาพ และทำนายผิดพลาดเป็นคลาสสุกพอดี (Ripe) จำนวน 2 ภาพ

2) คลาสที่สุกพอดี (Ripe) มีจำนวนรูปภาพสำหรับการทดสอบทั้งหมดจำนวน 246 ภาพ มีค่าความแม่นยำในการทำนายของคลาสนี้มีค่าเท่ากับ 99.59% โดยทำนายถูกต้องจำนวน 245 ภาพ และทำนายผิดพลาดเป็นคลาสดิบ (Unripe) จำนวน 1 ภาพ

3) คลาสที่ไม่สุกหรือดิบ (Unripe) มีจำนวนรูปภาพสำหรับการทดสอบทั้งหมดจำนวน 266 ภาพ มีค่าความแม่นยำในการทำนายของคลาสนี้มีค่าเท่ากับ 99.62% โดยทำนายถูกต้องจำนวน 265 ภาพ และทำนายผิดพลาดเป็นคลาสสุกพอดี (Ripe) จำนวน 1 ภาพ

สรุปผลการวิจัย

ระดับความสุกของผลปาล์มน้ำมัน มีผลโดยตรงต่อคุณภาพของน้ำมันปาล์ม เพราะผลปาล์มน้ำมันที่มีระดับความสุกพอดี เมื่อถูกสกัดแล้วจะทำให้ได้น้ำมันปาล์มที่มีคุณภาพสูง ดังนั้น หากเกิดความผิดพลาดในขั้นตอนการจำแนกระดับความสุกของปาล์ม น้ำมันสดและเก็บเกี่ยวไม่ถูกต้อง ส่งผลให้เกิดความเสียหายและสูญเสียรายได้ต่อเกษตรกรได้ งานวิจัยนี้จึงนำเสนอวิธีการจำแนก ระดับความสุกของปาล์มน้ำมันสด โดยใช้หลักการประสาทเทียมแบบคอนโวลูชันของ YOLOv8n (nano) ซึ่งชุดข้อมูลต้นฉบับที่ใช้สำหรับการวิจัย ได้รับมาจากเว็บไซต์ Science Data Bank ทั้งนี้ หมวดหมู่ระดับความสุกของปาล์มน้ำมัน สำหรับการดำเนินการวิจัยมี 3 หมวดหมู่ ได้แก่ ไม่สุกหรือดิบ (Unripe) สุกพอดี (Ripe) และสุกเกินไป (Over-ripe) ในขั้นตอนการเตรียมข้อมูล ผู้วิจัยได้นำข้อมูลภาพต้นฉบับมาทำ Data preparation ด้วยเครื่องมือ Roboflow ทั้งนี้ เพื่อให้การฝึกอบรมแบบจำลองมีความแม่นยำและมีประสิทธิภาพ ผู้วิจัยได้ใช้วิธีการทำ Isolate objects และทำการเพิ่มปริมาณข้อมูลภาพสอน (Data augmentations) ซึ่งในงานวิจัยนี้ได้ทำการปรับแต่งแบบจำลองโดยการปรับค่าพารามิเตอร์ต่าง ๆ ได้แก่ Optimizer, Learning rate และ Batch size จากผลลัพธ์แสดงให้เห็นว่าแบบจำลองที่มีความแม่นยำสูงที่สุดสำหรับการจำแนกระดับความสุกของปาล์มน้ำมันสด ด้วยแบบจำลอง YOLOv8n (nano) ที่มีค่าพารามิเตอร์ต่าง ๆ ดังนี้ Adam Optimizer, Learning rate = 0.0001, Batch size = 32 ทั้งนี้ ผลลัพธ์จากการประเมินแบบจำลองมีค่า Precision = 0.997, Recall = 0.997, mAP50 = 0.995 และ mAP50-95 = 0.995

กิตติกรรมประกาศ

กลุ่มผู้เขียนขอขอบคุณ หลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิทยาการคอมพิวเตอร์ มหาวิทยาลัยศรีนครินทรวิโรฒ ประเทศไทย สำหรับการสนับสนุนทางวิชาการและความรู้อันทรงคุณค่าที่มีส่วนสนับสนุนในการดำเนินการจนนำไปสู่ความสำเร็จของการทำวิจัยนี้ นอกจากนี้ กลุ่มผู้เขียนขอขอบคุณบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประเทศไทย ที่ได้มอบทุนสนับสนุนการตีพิมพ์ผลงานของเรา

เอกสารอ้างอิง

- [1] Chumphon Provincial Agriculture and Cooperatives office.(2021). Data for Agricultural Product Development Planning, 2021, (In thai). Chumphon Province (Palm Oil). Retrieved from <https://www.opsmoac.go.th/chumphon-performance-preview-431391791827>
- [2] Kritchaya, C., Nisakon, R., Kantita, J., (2023). The Situation of Production and Marketing of Palm Oil and Palm Oil in Thailand. (In thai). (1-6). Oil Palm and Palm Oil Research and Innovation Institute Journal, Prince of Songkla University
- [3] Teerapong, J., (2019). Agricultural Handbook for Efficient Palm Oil Production. (In thai). (4th edition). Research and Development Center for Efficient Palm Oil Production, Faculty of Natural Resources, Prince of Songkla University
- [4] Ibrahim, Z., Sabri, N., & Isa, D. (2018). Palm Oil Fresh Fruit Bunch Ripeness Grading Recognition Using Convolutional Neural Network. Journal of Telecommunication, Electronic and Computer Engineering (JTEC), 10(3-2), 109-113. <https://jtec.utem.edu.my/jtec/article/view/4720>
- [5] Septiarini, A., Hatta, H. R., Hamdani, H., Oktavia, A., Kasim, A. A., & Suyanto, S. (2020, 3-4 Nov. 2020). Maturity Grading of Oil Palm Fresh Fruit Bunches Based on a Machine Learning Approach. 2020 Fifth International Conference on Informatics and Computing (ICIC).
- [6] Herman, H., Cenggoro, T. W., Susanto, A., & Pardamean, B. (2021, 19-20 Aug. 2021). Deep Learning for Oil Palm Fruit Ripeness Classification with DenseNet. 2021 International Conference on Information Management and Technology (ICIMTech).
- [7] Ashari, S., Yanris, G. J., & Purnama, I. (2022). Oil Palm Fruit Ripeness Detection using Deep Learning. Sinkron : jurnal dan penelitian teknik informatika, 7(2), 649-656. <https://doi.org/10.33395/sinkron.v7i2.11420>
- [8] Khamis, N., Selamat, H., Ghazalli, S., Saleh, N. I. M., & Yusoff, N. (2022, 4-7 May 2022). Comparison of Palm Oil Fresh Fruit Bunches (FFB) Ripeness Classification Technique using Deep Learning Method. 2022 13th Asian Control Conference (ASCC)
- [9] Lai, J. W., Ramli, H. R., Ismail, L. I., & Hasan, W. Z. W. (2022). Real-Time Detection of Ripe Oil Palm Fresh Fruit Bunch Based on YOLOv4. IEEE Access, 10, 95763-95770. <https://doi.org/10.1109/ACCESS.2022.3204762>

- [10] Chen, Z., Elsaid, S., Sim, D. Y. Y., Maul, T., & Liao, I. Y. (2022, 16-18 Dec. 2022). Detection of Oil Palm Fresh Fruit Bunches (FFBS) with Computer Vision Models. 2022 International Conference on Mechanical, Automation and Electrical Engineering (CMAEE)
- [11] Mohtar, L., Muhammad, H., & Ab Rahman, M. (2022). Comparison CNN and Mobilenet_v2 Model for Oil Palm FBB Ripeness Classification. *Evolution in Electrical and Electronic Engineering*, 3(2), 146-154.
<https://publisher.uthm.edu.my/periodicals/index.php/eeee/article/view/8591>
- [12] Dambul, K. D., Choo, K. Y., & Mansour, M. Y. M. A. (2022). Object Detection Algorithms for Ripeness Classification of Oil Palm Fresh Fruit Bunch. *International Journal of Technology*, 13(6), 291-319.
<https://doi.org/https://doi.org/10.14716/ijtech.v13i6.5932>
- [13] Azman, H., & Suriani, N. S. (2023). Grading Oil Palm Fruit Bunch using Convolution Neural Network. *Evolution in Electrical and Electronic Engineering*, 4(1), 185-194.
<https://penerbit.uthm.edu.my/periodicals/index.php/eeee/article/view/10589>
- [14] Lai, J. W., Ramli, H. R., Ismail, L. I., & Wan Hasan, W. Z. (2023). Oil Palm Fresh Fruit Bunch Ripeness Detection Methods: A Systematic Review. *Agriculture*, 13(1)
- [15] Patel, S., & Patel, A. (2020). Object Detection with Convolutional Neural Networks. In (pp. 529-539).
https://doi.org/10.1007/978-981-15-7106-0_52
- [16] Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. *Proceedings of the IEEE conference on computer vision and pattern recognition*
- [17] Girshick, R. (2015). Fast r-cnn. *Proceedings of the IEEE international conference on computer vision* In (pp. 1440-1448)
- [18] Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in neural information processing systems*, 28.
- [19] He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask r-cnn. *Proceedings of the IEEE International Conference on Computer*
- [20] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. *Proceedings of the IEEE conference on computer vision and pattern recognition*
- [21] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., & Berg, A. C. (2016, 2016//). SSD: Single Shot MultiBox Detector. *Computer Vision – ECCV 2016, Cham*

- [22] Soumyadip., (2023). Mastering All YOLO Models from YOLOv1 to YOLO-NAS: Papers Explained (2024) Retrieved from <https://learnopencv.com/mastering-all-YOLO-models/>
- [23] Terven, J., & Cordova-Esparza, D. (2023). A comprehensive review of YOLO: From YOLOv1 to YOLOv8 and beyond. arXiv preprint arXiv:2304.00501
- [24] Suharjito, (2023) Annotated Video Dataset of Oil Palm Fruit Bunch Piles for Ripeness Grading. Retrieved from <https://www.scidb.cn/en/detail?dataSetId=7f2dfbed88c5484195d4217814b21542&version=V2>
- [25] Suharjito, Junior, F. A., Koeswandy, Y. P., Debi, Nurhayati, P. W., Asrol, M., & Marimin. (2023). Annotated Datasets of Oil Palm Fruit Bunch Piles for Ripeness Grading Using Deep Learning. Scientific Data, 10(1), 72. <https://doi.org/10.1038/s41597-023-01958-x>
- [26] Ultralytics Inc, (2024). Ultralytics YOLOv8 Docs. Retrieved from <https://docs.ultralytics.com/>
- [27] Elsevier B.V., (2024) Confusion Matrix. Retrieved from <https://www.sciencedirect.com/topics/engineering/confusion-matrix>