

การจำแนกขนาดเม็ดมะม่วงหิมพานต์ด้วยเทคนิคการเรียนรู้แบบผสมผสาน

Cashew nut size classification using hybrid learning techniques

ภวัต ฉิมเล็ก¹, ธงรบ อักษร¹, เสกสรรค์ ศิวาลัย¹ และสุธาสินี จิตต์อนันต์²

Pawat Chimlek¹, Thongrob Auxsorn¹, Sakesan Sivilai^{1*} and Sutasinee Jitanan²

Received: 28 February 2025 ; **Revised:** 15 May 2025 ; **Accepted:** 1 July 2025

บทคัดย่อ

ที่มาของปัญหาการวิจัยนี้เกิดจากข้อจำกัดของวิธีการคัดแยกขนาดเม็ดมะม่วงหิมพานต์แบบดั้งเดิม ซึ่งส่งผลให้มีต้นทุนแรงงานสูงและคุณภาพผลิตภัณฑ์ไม่สม่ำเสมอ อันส่งผลกระทบต่อความสามารถในการแข่งขันของอุตสาหกรรมเม็ดมะม่วงหิมพานต์ไทยในตลาดโลก วัตถุประสงค์ของการวิจัยคือการพัฒนากระบวนการคัดแยกขนาดเม็ดมะม่วงหิมพานต์ที่มีความแม่นยำสูงด้วยการผสมผสานเทคโนโลยีการเรียนรู้เชิงลึก (deep learning) และการเรียนรู้ของเครื่อง (machine Learning) วิธีการศึกษาได้นำเสนอระบบการคัดขนาดเม็ดมะม่วงหิมพานต์แบบผสมผสาน โดยประยุกต์ใช้การเรียนรู้เชิงลึกที่ใช้เทคนิค YOLOv5 ในการตรวจจับเม็ดมะม่วงหิมพานต์จากภาพ โดยได้ทดลองเพื่อหาพารามิเตอร์ที่เหมาะสมและมีประสิทธิภาพสูง และใช้การเรียนรู้ของเครื่องโดยใช้ SVM ในการคัดขนาดซึ่งวิเคราะห์จากลักษณะโครงสร้างหลักที่ชัดเจน ได้แก่ ความกว้าง ความยาว และพื้นที่ โดยได้ทดลองกับเคอร์เนล (kernel) 4 รูปแบบของ SVM ได้แก่ linear, RBF, polynomial และ sigmoid ผลการวิจัยแสดงให้เห็นว่าโมเดลตรวจจับมีประสิทธิภาพที่ดีที่สุดเมื่อใช้ Epochs เท่ากับ 400 และ batch size เท่ากับ 64 โดยมีค่า Precision เท่ากับ 0.9817 และ Recall เท่ากับ 0.9706 ขณะที่โมเดลคัดขนาดมีประสิทธิภาพสมบูรณ์ด้วยค่า Precision และ Recall ใกล้เคียง 1 ในทุกเคอร์เนล แม้ว่าประสิทธิภาพโดยรวมของระบบยังมีข้อจำกัดเนื่องจากขั้นตอนการตรวจจับที่ยังมีข้อผิดพลาดในบางกรณี การศึกษานี้สามารถสรุปได้ว่าระบบสามารถช่วยเพิ่มความแม่นยำในการคัดขนาดเม็ดมะม่วงหิมพานต์และทำให้สามารถช่วยเพิ่มความสามารถในการแข่งขันของอุตสาหกรรมเม็ดมะม่วงหิมพานต์ไทยในตลาดโลก โดยช่วยลดต้นทุนแรงงานและยกระดับคุณภาพผลิตภัณฑ์ให้มีความสม่ำเสมอมากขึ้น การพัฒนาในอนาคตควรมุ่งเน้นการปรับปรุงขั้นตอนการตรวจจับเพื่อเพิ่มประสิทธิภาพของระบบในภาพรวม

คำสำคัญ: การคัดแยกเม็ดมะม่วงหิมพานต์, การเรียนรู้เชิงลึก, การเรียนรู้ของเครื่อง, การจำแนกตามขนาด

Abstract

The research problem originated from the limitations of traditional cashew nut size classification methods, which result in high labor costs and inconsistent product quality, affecting the competitiveness of Thailand's cashew nut industry in the global market. The research objective was to develop a high-precision cashew nut size classification system by integrating deep learning and machine learning techniques. The proposed methodology involves a hybrid cashew nut size classification system. First, a deep learning technique utilizing YOLOv5 was applied to detect cashew nuts from images, with experiments conducted to determine the optimal and most efficient parameters. Subsequently, machine learning using a Support Vector Machine (SVM) was employed for size classification based on dominant morphological features, including width, length, and area. The SVM model was evaluated using four kernel types: linear, radial basis function (RBF), polynomial, and sigmoid. The results indicated that the detection model performed optimally when using 400 epochs and a batch size of 64, achieving a precision of 0.9817 and a recall of 0.9706. Furthermore, the size classification model achieved perfect performance, yielding precision and recall values of 1.0 across all tested

¹ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏพิบูลสงคราม

² ภาควิชาวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ, คณะวิทยาศาสตร์ มหาวิทยาลัยนเรศวร

¹ Faculty of Science and Technology, Pibulsongkram Rajabhat University

² Department of Computer Science and Information Technology, Faculty of Science, Naresuan University

* Corresponding author; E-mail: sakesan@psru.ac.th

kernels. However, the overall system performance still had minor limitations due to occasional errors in the detection step. In conclusion, this system can improve the accuracy of cashew nut size classification and enhance the competitiveness of Thailand's cashew nut industry in the global market by reducing labor costs and improving product quality consistency. Future developments should focus on refining the detection process to further enhance the overall system efficiency.

Keywords: Cashew nut classification, deep learning, machine learning, size classification, YOLOv5

บทนำ

อุตสาหกรรมเม็ดมะม่วงหิมพานต์ของประเทศไทยกำลังเผชิญความท้าทายที่สำคัญ แม้จะเป็นผู้ส่งออกอันดับที่ 16 ของโลก โดยมีปริมาณการส่งออก 24,923 ตันในปี 2565 (World Population Review, n.d.) แต่อุตสาหกรรมนี้มีแนวโน้มถดถอยอย่างต่อเนื่องในช่วงทศวรรษที่ผ่านมา สาเหตุหลักเกิดจากปัญหาด้านการควบคุมคุณภาพและกระบวนการผลิต โดยเฉพาะกระบวนการคัดขนาดซึ่งเป็นปัจจัยสำคัญที่ส่งผลต่อคุณภาพผลิตภัณฑ์และความสามารถในการแข่งขันในตลาดโลก เม็ดมะม่วงหิมพานต์เป็นพืชเศรษฐกิจสำคัญที่มีการเพาะปลูกอย่างแพร่หลายในจังหวัดอุดรธานี ชลบุรี และระนอง อย่างไรก็ตาม กระบวนการคัดแยกขนาดแบบดั้งเดิมยังคงต้องพึ่งพาแรงงานคนและเครื่องมือกลอย่างตะแกรงร่อนหลายชั้น ซึ่งมีข้อจำกัดด้านความแม่นยำ ประสิทธิภาพ และการรักษาคุณภาพของผลิตภัณฑ์

งานวิจัยก่อนหน้าได้นำเสนอแนวทางทางเทคนิคหลายรูปแบบสำหรับการจำแนกเม็ดมะม่วงหิมพานต์แบบอัตโนมัติ ซึ่งสามารถจำแนกออกเป็น 3 แนวทางหลัก ได้แก่ การประมวลผลภาพ (Image Processing) การเรียนรู้ของเครื่อง (Machine Learning) และการเรียนรู้เชิงลึก (Deep Learning) ในด้านการประมวลผลภาพ งานวิจัยของ Sunoj *et al.* (2018) ได้พัฒนาวิธีการที่ใช้ความเข้มของระดับสีเทาและมิติของเงาด้วยเทคนิคการแบ่งส่วนภาพ เพื่อแยกเม็ดมะม่วงหิมพานต์ออกจากพื้นหลัง โดยสามารถจำแนกเม็ดมะม่วงหิมพานต์ออกเป็น 2 กลุ่ม คือเม็ดที่สมบูรณ์ทั้งเม็ดและเม็ดที่แตก แม้ว่าวิธีนี้จะให้ความแม่นยำสูงสุดในสภาพแวดล้อมที่มีการควบคุม แต่ยังมีข้อจำกัดกับสัญญาณรบกวนที่เกิดจากแสงที่ไม่สม่ำเสมอซึ่งส่งผลต่อการวิเคราะห์ระดับสีเทาและโครงสร้างของเม็ดมะม่วงหิมพานต์ (Sunoj *et al.*, 2018)

การเรียนรู้ของเครื่องเป็นอีกแนวทางที่ได้รับการพัฒนา อัลกอริทึม Random Forest (Baitu *et al.*, 2023) ได้ถูกพัฒนาขึ้นให้สามารถจำแนกเม็ดสีขาวและเม็ดที่มีรอยไหม้ได้อย่างแม่นยำถึง 99.8% อย่างไรก็ตาม วิธีนี้ถูกจำกัดให้ใช้กับการจำแนกประเภทแบบทวิภาค (Binary Classification) เท่านั้น ในขณะที่ระบบคัดเกรดอัตโนมัติที่ใช้ Support Vector

Machine (SVM) สามารถจำแนกเม็ดมะม่วงหิมพานต์ตามสีและขนาดได้ด้วยความแม่นยำ 95% แสดงให้เห็นถึงประสิทธิภาพของ SVM ในการพัฒนาระบบคัดเกรดที่แม่นยำและเชื่อถือได้ ส่วนการเรียนรู้เชิงลึกมีบทบาทสำคัญในการเพิ่มความแม่นยำในการจำแนกประเภท โดยเฉพาะการใช้โครงข่ายประสาทเทียมแบบคอนโวลูชัน (CNN) ในการตรวจจับเม็ดมะม่วงหิมพานต์ที่ไม่ได้คุณภาพ งานวิจัยหลายงานได้ประยุกต์ใช้สถาปัตยกรรมเครือข่ายประสาทเทียมเชิงลึกแบบ DCNN (Kumari *et al.*, 2021) ซึ่งให้ความแม่นยำในการจำแนกประเภทเม็ดมะม่วงหิมพานต์ตั้งแต่ 95.1% ถึง 98.4% แม้ว่าประสิทธิภาพสูง แต่ยังคงมีความท้าทายด้านความซับซ้อนของโมเดลและปัญหาการเกิด Overfitting

งานวิจัยปัจจุบันมุ่งไปที่การใช้สถาปัตยกรรม YOLOv8 ที่สามารถจำแนกประเภทเม็ดมะม่วงหิมพานต์ (Pham *et al.*, 2024) โดยมีอัตราความผิดพลาดเพียง 2.96% ระบบนี้ใช้กล้องต้นทุนต่ำร่วมกับสายพานลำเลียง และฝึกโมเดลด้วยภาพตัวอย่าง 1,500 ภาพ และมีงานวิจัยที่ประยุกต์ใช้ YOLOv8 ในการตรวจจับวัชพืชหลายชนิดในไร่ฝ้ายของสหรัฐอเมริกา โดยเปรียบเทียบกับโมเดลอื่น ๆ เช่น YOLOv5 และ Faster R-CNN ผลการทดลองแสดงให้เห็นว่า YOLOv8 มีความแม่นยำสูง (Zha *et al.*, 2025)

อย่างไรก็ตาม จากการทบทวนวรรณกรรมพบว่า งานวิจัยส่วนใหญ่มุ่งเน้นไปที่การคัดแยกคุณภาพของเม็ดมะม่วงหิมพานต์ ในขณะที่ยังขาดงานวิจัยที่เกี่ยวข้องกับการคัดแยกขนาดซึ่งเป็นอีกปัจจัยหนึ่งที่สำคัญในการกำหนดราคา โดยงานวิจัยส่วนใหญ่จะเกี่ยวข้องกับการคัดขนาดของผลผลิตทางการเกษตรอื่น เช่น การพัฒนาเทคโนโลยีระบบอัตโนมัติสำหรับคัดแยกและคัดเกรดผลไม้ โดยใช้เทคนิคการประมวลผลภาพเพื่อวิเคราะห์ลักษณะของผลไม้ เช่น สี รูปร่าง และขนาด ควบคู่กับอัลกอริทึมการเรียนรู้ของเครื่องที่สามารถปรับปรุงความแม่นยำได้อย่างต่อเนื่องเมื่อมีข้อมูลมากขึ้น ข้อจำกัดหลักของงานวิจัยนี้อยู่ที่สภาพแวดล้อมที่หลากหลายอาจส่งผลต่อความแม่นยำ

นอกจากนี้ยังมียานวิจัยที่เกี่ยวข้องกับการคัดขนาดแอปเปิ้ลโดยใช้เทคโนโลยีการรู้จำภาพและเทคนิค Laplacian

(Zhang *et al.*, 2024) เพื่อหาขอบของแอปเปิ้ลและคำนวณหาเส้นผ่านศูนย์กลาง ผลการทดลองพบว่าการใช้เส้นผ่านศูนย์กลางของแอปเปิ้ลมีความถูกต้อง 95% ส่วนการใช้สี่เหลี่ยมภายนอกมีความถูกต้อง 82.5% รวมถึงมีงานวิจัยที่นำเสนอการวัดขนาดแอปเปิ้ลบนต้นโดยใช้กล้องความลึกและ YOLOv5 (Song *et al.*, 2022) ที่มีความถูกต้องเท่ากับ 92% แต่ยังมีข้อจำกัดที่อาจเกิดจากการเปลี่ยนแปลงของแสง

ความท้าทายสำคัญในการพัฒนาระบบคัดแยกขนาดเม็ดมะม่วงหิมพานต์อยู่ที่การพัฒนาโมเดลให้สามารถรองรับปัจจัยสภาพแวดล้อมที่มีหลากหลาย และการบูรณาการคุณลักษณะหลายประเภทเพื่อเพิ่มความแม่นยำและประสิทธิภาพ ดังนั้นงานวิจัยนี้จึงมุ่งเน้นการพัฒนาแบบคัดแยกเม็ดมะม่วงหิมพานต์ตามขนาดที่มีความแม่นยำสูง โดยอาศัยเทคโนโลยี YOLO เพื่อแก้ไขข้อจำกัดของวิธีการแบบดั้งเดิม ลดต้นทุนแรงงาน และยกระดับคุณภาพผลิตภัณฑ์ ซึ่งจะช่วยเสริมสร้างความสามารถในการแข่งขันของอุตสาหกรรมเม็ดมะม่วงหิมพานต์ไทยในตลาดโลกได้อย่างยั่งยืน

วิธีดำเนินการวิจัย

งานวิจัยนี้เสนอวิธีการแบบผสมที่รวมการเรียนรู้เชิงลึก (Deep Learning) กับการเรียนรู้ของเครื่อง (Machine Learning) โดยใช้การเรียนรู้เชิงลึกในการตรวจจับเม็ดมะม่วงหิมพานต์ และการเรียนรู้ของเครื่องในการจำแนกขนาดเม็ดมะม่วงหิมพานต์ วิธีการแบบผสมนี้ใช้จุดแข็งของทั้งการเรียนรู้เชิงลึกและการเรียนรู้ของเครื่องแบบดั้งเดิม ทำให้ระบบมีประสิทธิภาพมากขึ้นในการประมวลผลเม็ดมะม่วงหิมพานต์ ซึ่งการเรียนรู้เชิงลึกลดข้อจำกัดเรื่องสภาพแวดล้อมที่หลากหลายที่ส่งผลต่อความแม่นยำ ในการตรวจจับและการแบ่งส่วนเม็ดมะม่วงหิมพานต์ ในขณะที่การเรียนรู้ของเครื่องแบบดั้งเดิมจะทำการจำแนกขนาดได้อย่างแม่นยำโดยใช้อัลกอริทึม

ที่เป็นที่ยอมรับและคุณลักษณะเฉพาะด้านการจำแนกขนาดของเม็ดมะม่วงหิมพานต์

ผู้วิจัยใช้เม็ดมะม่วงหิมพานต์ที่ผลิตในประเทศไทย โดยเฉพาะจากจังหวัดอุดรธานี ซึ่งจัดกลุ่มตามขนาดมาตรฐาน ได้แก่ W180: ขนาดใหญ่ที่สุด (180 เม็ดต่อปอนด์), W240: ขนาดใหญ่รองลงมา (240 เม็ดต่อปอนด์), W320: ขนาดกลาง (320 เม็ดต่อปอนด์) และ W450: ขนาดเล็กที่สุด (450 เม็ดต่อปอนด์) โดยมีละเอียดดัง Table 1

Table 1 Details of cashew nuts by size

Size	Width (mm)	Height (mm)
W180	20	34
W240	17	27
W320	15	23
W450	13	22

1. วิธีการเก็บรวบรวมข้อมูล

เม็ดมะม่วงหิมพานต์ถูกจัดเป็น 4 ขนาด โดยใช้เครื่องวัดความกว้างและความยาวของแต่ละเม็ดด้วยเวอร์เนียคาลิเปอร์เพื่อยืนยันความถูกต้องของขนาดก่อนที่จะใช้ในการถ่ายภาพ ภาพจะถูกถ่ายด้วยกล้องมือถือดิจิทัลที่มีความละเอียด 1024*1024 พิกเซลภายใต้สภาพแสงที่เหมาะสมและในสภาพแวดล้อมที่ควบคุมเพื่อให้ได้ภาพคุณภาพสูงสำหรับการประมวลผล เม็ดมะม่วงหิมพานต์ถูกจัดเรียงตามขนาดที่กำหนดบนกระดาษสีขาว โดยกล้องจะวางอยู่ที่ความสูง 30 เซนติเมตรจากพื้นที่จัดไว้

การศึกษานี้ได้ถ่ายภาพ โดยมีรายละเอียดแสดงดัง Table 2 โดยในภาพที่ถ่ายมีทั้งเม็ดที่ติดกันรวมอยู่ด้วยเพื่อให้ใกล้เคียงกับสภาพแวดล้อมจริง

Table 2 Image Acquisition Details of Cashew Nuts

Size	Number of Cashew nuts image	Train set	Validation set	Test set	Total images	Total cashew nuts
W180	9	26	4	8	38	342
W240	10	26	4	8	38	380
W320	11	26	4	7	37	407
W450	12	27	3	7	37	444
Total images		105	15	30	150	
Total cashew nuts		1,104	156	313		1573

จากภาพที่ถ่ายได้มีการสร้างภาพเพิ่มให้มีความหลากหลายของแสงสว่างของภาพเพื่อให้ได้โมเดลตรวจจับเม็ดมะม่วงหิมพานต์ที่มีประสิทธิภาพกับภาพที่แสดงที่แตกต่างกันด้วย โดยได้มีการปรับเพิ่มลดแสงให้ภาพแบบสุ่มภาพละ 5 ระดับ ดังนั้นจะมีภาพทั้งหมด 750 ภาพ และมีจำนวนเม็ดทั้งหมด 7865 เม็ด

ตัวอย่างภาพที่แสดงใน Figure 1 และ Figure 2 โดยชุดข้อมูลนี้ถูกแบ่งออกเป็นสามชุด ได้แก่ ชุดฝึกอบรม (training set), ชุดทดสอบ (validation set) และชุดทดสอบจริง (testing set) ในอัตราส่วน 70:20:10 ตามลำดับ

Dataset จะถูกแบ่งออกเป็น 2 รูปแบบคือ Dataset 1 คือ Dataset ของโมเดลตรวจจับเม็ดมะม่วงหิมพานต์ ซึ่งจะเป็นการตีกรอบรอบเม็ดมะม่วงหิมพานต์ และสร้างป้ายกำกับ (Label) เป็น binary class คือเป็นเม็ดมะม่วงหิมพานต์หรือไม่ใช่ ส่วน Dataset 2 คือ Dataset ของโมเดลคัดแยกขนาดเม็ดมะม่วงหิมพานต์ ซึ่งจะเป็นการตีกรอบรอบเม็ดมะม่วงหิมพานต์ และสร้าง สร้างป้ายกำกับเป็น 4 คลาส (Class) คือ 4 ขนาดมาตรฐานของเม็ดมะม่วงหิมพานต์ แสดงดัง Figure 1

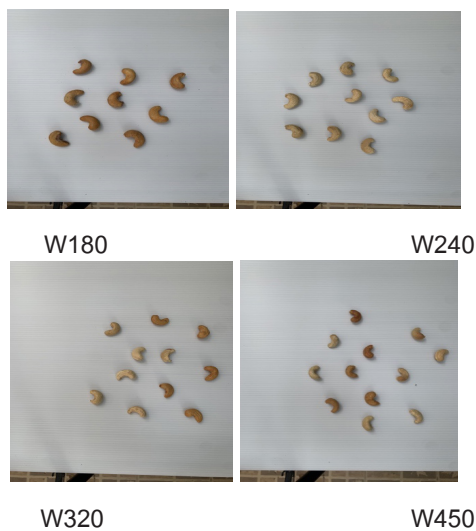


Figure 1 Sample images in dataset



Figure 2 Examples of Images with Brightness Variations

วิธีการนี้ช่วยให้การเก็บข้อมูลภาพเป็นมาตรฐานและมีคุณภาพสูงสำหรับการประมวลผลและวิเคราะห์ในขั้นตอนถัดไป การให้ความสำคัญในการตรวจสอบการวัดและการถ่ายภาพในสภาวะที่ควบคุมได้ช่วยรักษาความสอดคล้องและความน่าเชื่อถือของข้อมูลสำหรับกระบวนการฝึกเครื่องเรียนรู้

2. วิธีการคัดขนาดเม็ดมะม่วงหิมพานต์

งานวิจัยนี้ออกแบบระบบคัดขนาดเม็ดมะม่วงหิมพานต์แบบผสมผสานโดยการรวม Deep Learning และ Machine Learning เข้าด้วยกัน โดยใช้ Deep Learning ในการพัฒนาโมเดลเพื่อจับภาพเม็ดมะม่วงหิมพานต์แต่ละเม็ด เนื่องจากสามารถตรวจจับวัตถุได้อย่างมีประสิทธิภาพในสภาพแสงที่หลากหลายและพื้นหลังที่แตกต่างกัน ขณะเดียวกัน Machine Learning ถูกนำมาใช้ในการสร้างโมเดลสำหรับการคัดขนาด เนื่องจากกระบวนการนี้ต้องการขอบเขตขนาดที่ชัดเจนโดยอ้างอิงจากคุณลักษณะโครงสร้างหลักของเม็ดมะม่วงหิมพานต์ เช่น ความกว้าง ความสูง หรือพื้นที่ โดยคุณลักษณะเหล่านี้มีความสำคัญสำหรับการพัฒนาโมเดลการคัดแยกที่มีประสิทธิภาพสูงและ Machine Learning เหมาะสมกว่า Deep Learning สำหรับงานนี้ เพราะคุณลักษณะที่เลือกใช้ตรงกับคุณลักษณะที่ใช้ในกระบวนการคัดแยกแบบดั้งเดิมที่อิงจากมนุษย์ กระบวนการทำงานของระบบนี้แสดงใน Figure 3

2.1 โมเดลตรวจจับเม็ดมะม่วงหิมพานต์ (Cashew Nut Detection Model)

การศึกษานี้จะประยุกต์ใช้ YOLOv8 ในตรวจหาเม็ดมะม่วงหิมพานต์ เพื่อให้ได้ขอบเขตของเม็ดมะม่วงหิมพานต์ เพื่อนำไปใช้สำหรับดึงลักษณะสำคัญที่ใช้สำหรับการคัดแยกขนาดขั้นตอนต่อไป ได้แก่ ความกว้าง ความสูง และพื้นที่ของเม็ดมะม่วงหิมพานต์ โดยจะนำภาพใน Dataset1 ที่มีป้ายกำกับสำหรับระบุตำแหน่ง ขอบเขตของเม็ดมะม่วงหิมพานต์ในแต่ละภาพ

การสร้างโมเดลตรวจจับเม็ดมะม่วงหิมพานต์โดยใช้ YOLOv8 ซึ่งจะฝึกโมเดลเพื่อหาพารามิเตอร์ที่เหมาะสมเพื่อให้โมเดลที่สามารถตรวจหาเม็ดมะม่วงหิมพานต์ได้อย่างแม่นยำ

2.2 โมเดลคัดแยกขนาดเม็ดมะม่วงหิมพานต์ (Cashew Nut Size Classification Model)

ขั้นตอนนี้จะนำผลลัพธ์โมเดลตรวจจับเม็ดมะม่วงหิมพานต์ที่เป็นกรอบสี่เหลี่ยมรอบเม็ดมะม่วงหิมพานต์ในแต่ละภาพมาสร้างโมเดลสำหรับการคัดแยกขนาดของเม็ดมะม่วงหิมพานต์ โดยการสร้างโมเดลคัดแยกขนาดเม็ดมะม่วง

หิมพานต์ได้ใช้ Support Vector Machine (SVM) เนื่องจาก SVM มีความสามารถในการสร้างขอบเขตที่ชัดเจนระหว่างกลุ่มข้อมูลที่แตกต่างกัน โดยข้อมูลนำเข้าที่ใช้ในกระบวนการนี้ประกอบด้วยคุณลักษณะของเม็ดมะม่วงหิมพานต์ ได้แก่ ความกว้าง ความสูง และพื้นที่ ซึ่งได้จาก Dataset 2 ซึ่งทำ

ป้ายกำกับเป็นกรอบสี่เหลี่ยมล้อมรอบเม็ดมะม่วงหิมพานต์ในแต่ละขนาด 4 ตามขนาดมาตรฐาน ได้แก่ W180, W240, W320, และ W450 โดยเม็ดมะม่วงหิมพานต์แต่ละขนาดก่อนที่จะนำมาถ่ายรูปเก็บได้ถูกวัดขนาดที่แน่นอนด้วยเวอร์เนียร์เพื่อให้ได้ขนาดที่ตรงตามช่วงที่กำหนดให้ขนาดมาตรฐาน

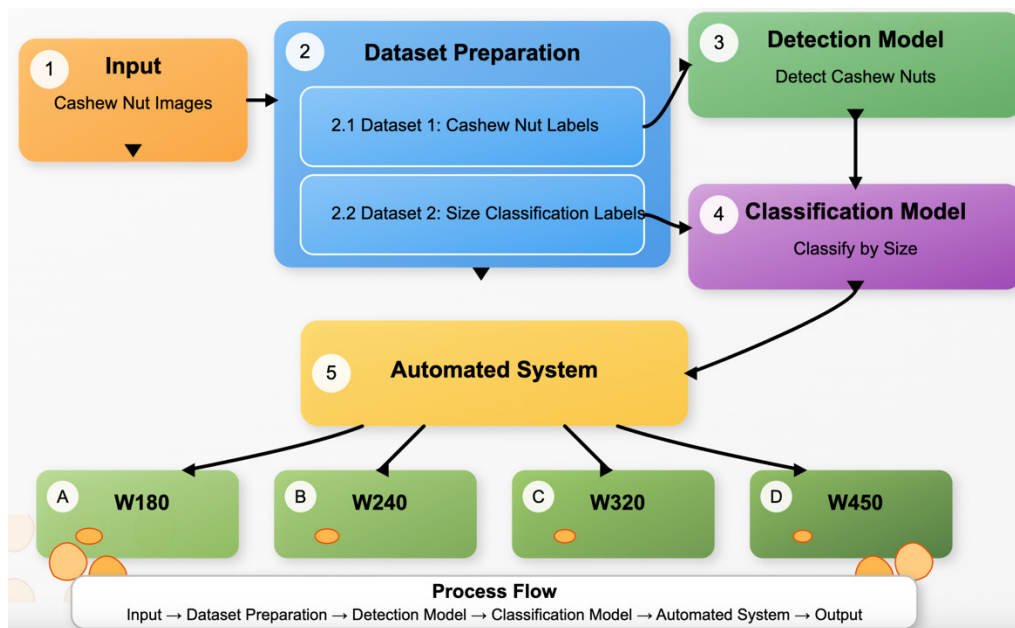


Figure 3 Cashew nut size classification using hybrid learning techniques

3. เกณฑ์การวัดประสิทธิภาพ

3.1 การวัดประสิทธิภาพของโมเดลตรวจจับเม็ดมะม่วงหิมพานต์

ค่า Mean Average Precision (mAP) จะถูกใช้เพื่อตัวชี้วัดที่ใช้ประเมินประสิทธิภาพของแบบจำลองการตรวจจับเม็ดมะม่วงหิมพานต์ โดยจะพิจารณาความแม่นยำในการจำแนกตำแหน่งของเม็ดมะม่วงหิมพานต์ที่พบในภาพหรือไม่พบ โดยการคำนวณ mAP จะใช้ค่า Precision และ Recall ที่ได้จากการประเมินโมเดลในแต่ละ Threshold ของค่า Intersection over Union (IoU) ซึ่งแสดงถึงความแม่นยำในการทำนายตำแหน่งของวัตถุในภาพ โดย mAP จะรวมค่า Average Precision (AP) จากหลายๆ คลาส หรือหลายๆ Threshold มาเป็นค่าเดียวที่สามารถใช้เปรียบเทียบประสิทธิภาพของโมเดลได้ (LearnOpenCV, 2021)

สมการของ Precision และ Recall สำหรับการตรวจจับเม็ดมะม่วงหิมพานต์ สามารถคำนวณดังสมการ 1 และ 2

$$Precision = \frac{True\ Positives(TP)}{True\ Positives(TP) + False\ Positives(FP)} \quad (1)$$

$$Recall = \frac{True\ Positives(TP)}{True\ Positives(TP) + False\ Negatives(FN)} \quad (2)$$

โดยที่ True Positive (TP) คือจำนวนที่ตรวจจับเม็ดมะม่วงหิมพานต์ได้ถูกต้อง

False Positive (FP) คือ จำนวนที่ตรวจจับเม็ดมะม่วงหิมพานต์ผิด

False Negative (FN) คือ จำนวนที่ตรวจจับเม็ดมะม่วงหิมพานต์ไม่พบ

การคำนวณ AP สำหรับแต่ละคลาสจะเป็นการคำนวณค่าเฉลี่ยของ Precision ที่หลากหลายระดับของ Recall และ mAP จะถูกคำนวณโดยการหาค่าเฉลี่ยของ AP สำหรับแต่ละคลาส ในงานวิจัยนี้คือคลาสเม็ดมะม่วงหิมพานต์และไม่ใช่เม็ดมะม่วงหิมพานต์ ตามสมการ 3

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \quad (3)$$

โดยที่ N คือจำนวนของคลาส และ AP_i คือค่า Average Precision ของแต่ละคลาส ซึ่งในการศึกษาจะใช้ mAP50 ซึ่งจะใช้ค่า IoU ที่ 0.50 โดยที่ค่า IoU ที่มากกว่าหรือเท่ากับ 0.50 จะถือว่าการตรวจจับนั้นถูกต้อง นอกจากนี้จะใช้ mAP50-95 ที่คำนวณจากค่า IoU ตั้งแต่ 0.50 ถึง 0.95 โดยเพิ่มขึ้นทีละ 0.05 ค่า IoU ที่ใช้คำนวณ ได้แก่ 0.50, 0.55, 0.60, ..., 0.95 ค่าที่ได้จะเป็นค่าเฉลี่ยของ AP ในช่วงนี้ mAP50-95 มีความแม่นยำในการระบุตำแหน่งวัตถุที่สูงกว่า mAP50 เพราะใช้ IoU หลายค่า ทำให้สะท้อนประสิทธิภาพของโมเดลได้ดีกว่า

นอกจากนี้จะใช้ค่า Fitness เป็นตัวชี้วัดที่ใช้ประเมินคุณภาพของโมเดลที่ฝึกมา และช่วยเลือก Weights ที่ดีที่สุดจากแต่ละ epoch ในระหว่างการฝึก ทำให้ได้โมเดลที่มีความแม่นยำสูงสุด โดยที่โมเดลที่ดีจะมีค่า Fitness สูงกว่า 0.5 ขึ้นไป โดยสามารถคำนวณด้วย สมการ 4

$$Fitness = 0.9 \times mAP50-95 + 0.1 \times Recall \quad (4)$$

3.2 การวัดประสิทธิภาพของโมเดลคัดขนาดเม็ดมะม่วงหิมพานต์

ในการจำแนกขนาดเม็ดมะม่วงหิมพานต์ที่มี 4 ขนาด จะถูกวัดประสิทธิภาพด้วยค่า Accuracy, Precision, Recall และ F1-Score โดยที่ Accuracy คืออัตราส่วนของจำนวนตัวอย่างที่โมเดลทำนายถูกต้องต่อจำนวนตัวอย่างทั้งหมด ดังสมการ 5

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (5)$$

โดยที่ TP (True Positive) คือจำนวนครั้งที่โมเดลทำนายขนาดถูกต้อง

TN (True Negative) คือจำนวนครั้งที่โมเดลทำนายว่าไม่ใช่ขนาดนั้นและถูกต้อง

FP (False Positive) คือจำนวนครั้งที่โมเดลทำนายผิดว่าเป็นขนาดหนึ่งในขณะที่คำตอบคือไม่ใช่ขนาดนั้น

FN (False Negative) คือจำนวนครั้งที่โมเดลผิดพลาดไม่สามารถจำแนกขนาดที่ถูกต้องได้

Precision และ Recall เป็นตัวชี้วัดที่สำคัญเมื่อพิจารณาการจำแนกแต่ละขนาดของเม็ดมะม่วงหิมพานต์ โดย Precision จะคำนวณจากสัดส่วนของตัวอย่างที่โมเดลทำนายว่าเป็นขนาดใดขนาดหนึ่งแล้วถูกต้องเทียบกับจำนวนทั้งหมด

ที่โมเดลทำนายว่าเป็นขนาดนั้น ซึ่งแสดงถึงความแม่นยำของโมเดลในแต่ละขนาด โดยที่ Recall จะวัดว่าโมเดลสามารถจับตัวอย่างทั้งหมดของขนาดนั้นๆ ออกมาได้มากน้อยเพียงใด ซึ่งค่าสูงหมายถึงโมเดลสามารถจับตัวอย่างที่เป็นขนาดนั้นๆ ได้ดีโดยมีความผิดพลาด (False Negative) น้อย ซึ่งชี้ให้เห็นความแม่นยำในการคัดแยกเม็ดมะม่วงหิมพานต์แต่ละขนาดได้ครบถ้วน หาก Precision สูงแสดงว่าโมเดลให้ผลลัพธ์มีความแม่นยำ แต่ถ้า Recall สูงหมายถึงโมเดลสามารถจับตัวอย่างทั้งหมดของขนาดนั้นได้ครบถ้วน

F1-Score เป็นค่าเฉลี่ยแบบถ่วงน้ำหนักระหว่าง Precision และ Recall เพื่อลดปัญหาการพึ่งค่าใดค่าหนึ่งมากเกินไป โดยคำนวณจากสมการ 6

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (6)$$

ค่าที่สูงหมายถึงโมเดลมีความสมดุลระหว่าง Precision และ Recall ทำให้สามารถจำแนกขนาดของเม็ดมะม่วงหิมพานต์ได้ทั้งแม่นยำและครบถ้วน ส่วนค่า F1-Score มีประโยชน์อย่างมากในกรณีที่มีความไม่สมดุลของข้อมูลระหว่างแต่ละขนาดของเม็ดมะม่วงหิมพานต์ เช่น หากขนาดหนึ่งมีตัวอย่างมากกว่าอีกขนาดหนึ่งมากเกินไป ซึ่งค่า Accuracy ที่สูงแต่ไม่ได้สะท้อนถึงประสิทธิภาพจริงของโมเดล ดังนั้นค่า F1-Score จึงช่วยให้สามารถประเมินผลได้แม่นยำยิ่งขึ้นสำหรับแต่ละขนาด และสามารถเปรียบเทียบประสิทธิภาพของโมเดลได้ดียิ่งขึ้น

ผลการทดลองและอภิปรายผล

1. ผลการทดลองโมเดลตรวจจับเม็ดมะม่วงหิมพานต์

การทดลองสำหรับโมเดลตรวจจับเม็ดมะม่วงหิมพานต์เพื่อให้ได้บริเวณที่เป็นเม็ดมะม่วงหิมพานต์ โดยใช้ Dataset 1 และมีปรับพารามิเตอร์สำหรับการสร้างโมเดล ได้แก่ Epochs(E) จำนวนในการฝึก (100,200,300,400,500) และ Batch size (B) ขนาดของชุดข้อมูลที่ถูกใช้ในแต่ละรอบการฝึกอบรม คือ 64 ผลการทดลองแสดงดัง Table 3

จากผลการทดลองแสดงให้เห็นว่าการเพิ่ม Epochs จาก 100 เป็น 500 ภายใต้ Batch size 64 ช่วยเพิ่มประสิทธิภาพของโมเดล โดยค่า mAP50 เพิ่มจาก 0.9842 เป็น 0.9875 และค่า mAP50-95 สูงสุดที่ epoch 300 ก่อนลดลงเล็กน้อย นอกจากนี้ Precision และ Recall ที่ดีขึ้นส่งผลให้ F1 Score เพิ่มจาก 0.9662 เป็น 0.9950 แสดงให้เห็นประสิทธิภาพในการตรวจจับเม็ดมะม่วงหิมพานต์ได้อย่างแม่นยำและครอบคลุม ซึ่งหากพิจารณาทั้งด้านประสิทธิภาพ ความเสี่ยงของ overfitting และเวลาในการฝึก โมเดลที่เหมาะสมที่สุดคือ

โมเดลที่ฝึกด้วย Epoch เท่ากับ 400 เพราะมีค่า Precision, Recall, mAP50, mAP50-95 และ F1 Score อยู่ในระดับสูงมากอย่างสม่ำเสมอ โดยเฉพาะ F1 Score ที่สูงถึง 0.9761 ซึ่งแสดงถึงความสมดุลที่ดีระหว่าง Precision และ Recall ขณะเดียวกันก็ยังควบคุมความเสี่ยงจาก overfitting ได้ดีกว่าโมเดลที่ Epoch เท่ากับ 500 โดยแสดงผลลัพธ์ของโมเดลตรวจจับเม็ดมะม่วงหิมพานต์ดัง Figure 4 และ Figure 5

ดังนั้นโมเดลตรวจจับเม็ดมะม่วงหิมพานต์จาก Batch Size 64 และ Epochs 400 จะถูกนำไปใช้ในการตรวจจับเม็ดมะม่วงหิมพานต์เพื่อให้ได้กรอบของเม็ดมะม่วงหิมพานต์และนำไปใช้สำหรับคัดขนาดเม็ดมะม่วงหิมพานต์ในขั้นตอนถัดไป รวมถึงพัฒนาเพื่อการประยุกต์ใช้งานในระบบอัตโนมัติสำหรับการแยกขนาดเม็ดมะม่วงหิมพานต์ในอนาคต

Table 3 Comparative performance results of cashew nut detection models based on epochs and batch size Annotation: Epochs(E), Batch size (B), Precision(P), Recall(R)

E	B	P	R	mAP 50	mAP 50-95	F1 Score
100	64	0.9657	0.9667	0.9842	0.8595	0.9662
200	64	0.9798	0.9529	0.9858	0.8626	0.9662
300	64	0.9686	0.9671	0.9874	0.8694	0.9678
400	64	0.9817	0.9706	0.9875	0.8673	0.9761
500	64	0.9998	0.9902	0.9950	0.8634	0.9950



Figure 4 Experimental results of the cashew nut detection model with Batch Size 64 and 400 Epochs

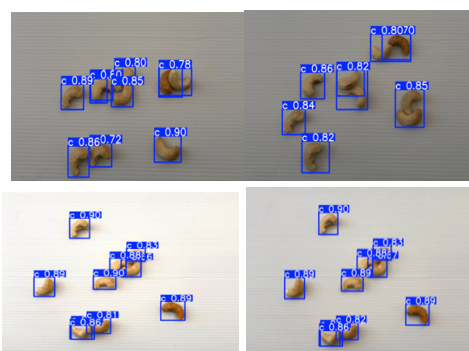


Figure 5 Example detection results of overlapping cashew nuts

2. ผลการทดลองโมเดลคัดขนาดเม็ดมะม่วงหิมพานต์

ในกระบวนการคัดแยกขนาดเม็ดมะม่วงหิมพานต์ด้วย SVM โดยใช้ Dataset 2 ซึ่งประกอบด้วย 4 ขนาดมาตรฐานมะม่วงหิมพานต์ โดยนำความกว้าง ความยาว และพื้นที่ของกรอบสี่เหลี่ยมล้อมรอบเม็ดมะม่วงหิมพานต์เป็นลักษณะสำคัญ (Feature) สำหรับสร้างโมเดลคัดแยกขนาดเม็ดมะม่วงหิมพานต์โดยใช้ SVM ซึ่งได้ทดลองกับเคอร์เนล (Kernel) 3 รูปแบบของ SVM ได้แก่ Linear, RBF และ Polynomial โดยแต่ละ Kernel ได้ทดลอง Hyperparameter Tuning แบบ Grid Search เพื่อให้ได้พารามิเตอร์ที่ดีที่สุดสำหรับแต่ละ Kernel ซึ่ง Grid search จะเหมาะสมกับการทดลองที่มีพารามิเตอร์น้อยและต้องการให้ได้ความแม่นยำสูง โดยในการทดลองได้มีการปรับพารามิเตอร์ที่ใช้ในแต่ละ Kernel ได้แก่ C คือค่าควบคุมขนาดของ Margin ประกอบด้วยค่า [0.1, 1, 10, 100, 200], gamma คือควบคุมอิทธิพลของตัวอย่างข้อมูล ซึ่งมีค่า [1, 0.1, 0.01, 0.001] และ class_weight จะปรับค่าอัตราโน้มนำให้เหมาะกับจำนวนข้อมูลในแต่ละคลาส ซึ่งจะทดลองโดยใช้ 10-fold Cross-Validation เพื่อหาโมเดลที่ดีที่สุด ซึ่งผลการทดลองแสดงดัง Table 4

Table 4 Performance of Cashew Nut Sizing Models

Kernel	Precision	Recall	F1-Score
Linear	0.99	1.00	1.00
RBF	0.99	0.99	0.99
Polynomial	1.00	0.98	0.99

ผลลัพธ์จากตารางแสดงให้เห็นว่า โมเดล SVM ที่ใช้ Kernel ต่าง ๆ ได้แก่ Linear, RBF และ Polynomial มีค่า Precision, Recall, F1-Score และ Accuracy เท่ากับ 1.00 ทั้งหมด แสดงให้เห็นว่าโมเดลสามารถจำแนกข้อมูลได้ใกล้เคียง 100% ซึ่งชี้ให้เห็นว่าชุดข้อมูลของขนาดเม็ดมะม่วงหิมพานต์อาจมีโครงสร้างที่แยกกันอย่างชัดเจน ทำให้ SVM สามารถสร้างเส้นแบ่งที่เหมาะสมได้ทุก Kernel อย่างมีประสิทธิภาพนอกจากนี้ การที่ค่า F1-Score ใกล้เคียง 1.00 สะท้อนว่าโมเดลไม่มีปัญหาความไม่สมดุลของข้อมูล (Imbalanced Data) และสามารถดึงคุณลักษณะสำคัญออกมาได้ดี และแสดงถึงความสามารถของโมเดลในการคัดแยกขนาดเม็ดมะม่วงหิมพานต์ได้อย่างแม่นยำและครอบคลุมในทุกขนาด โดยแสดงตัวอย่างผลการทดลองดัง Figure 6 ซึ่งทดลองแบบมีหลายขนาดและมีวัตถุอื่นๆ



Figure 6 Results of cashew nut size classification model

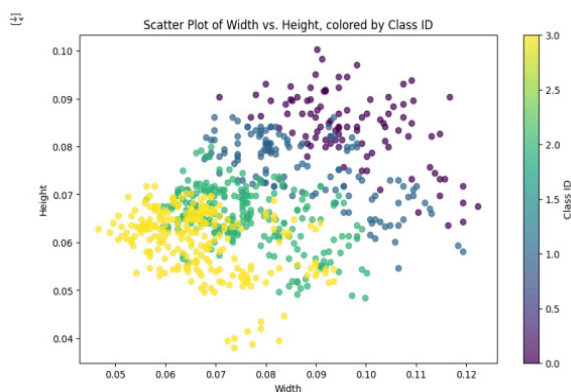


Figure 7 Scatter plot showing the relationship between width and height of cashew nuts of each size

จากกราฟ Scatter Plot ใน Figure 7 แสดงความสัมพันธ์ระหว่างความกว้างและความยาวของกรอบที่ล้อมรอบเม็ดมะม่วงหิมพานต์แต่ละเม็ด โดยมีการแบ่งสีตามขนาด 4 ขนาด จะเห็นว่าความกว้างและความยาวของแต่ละขนาดสามารถแบ่งแยกกันค่อนข้างชัดเจน ซึ่งทำให้เห็นว่าการใช้ SVM ในการแยกขนาดเม็ดมะม่วงหิมพานต์ได้ดี แม้จะมีบางส่วนที่ซ้อนทับกันบ้าง

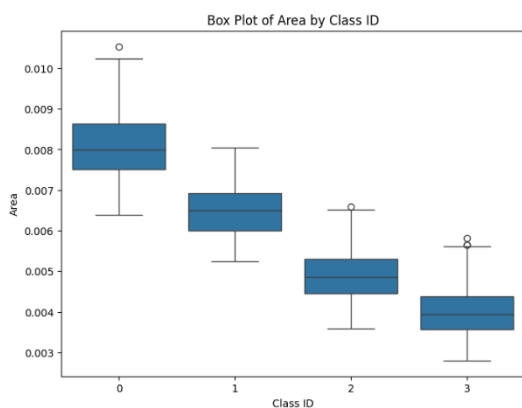


Figure 8 Box plot showing the area range of cashew nuts by size

จากการวิเคราะห์พื้นที่ของกรอบเม็ดมะม่วงหิมพานต์ โดยใช้ Box Plot ใน Figure 8 ที่แสดงการกระจายตัวของพื้นที่ของเม็ดมะม่วงหิมพานต์แต่ละขนาดตาม Class ID (0-3) แสดงให้เห็นความแตกต่างที่ชัดเจนระหว่างแต่ละขนาด ซึ่งการแบ่งแยกที่ชัดเจนนี้บ่งชี้ว่าค่าพื้นที่เป็นตัวแปรที่ดีในการจำแนกขนาดเม็ดมะม่วงหิมพานต์

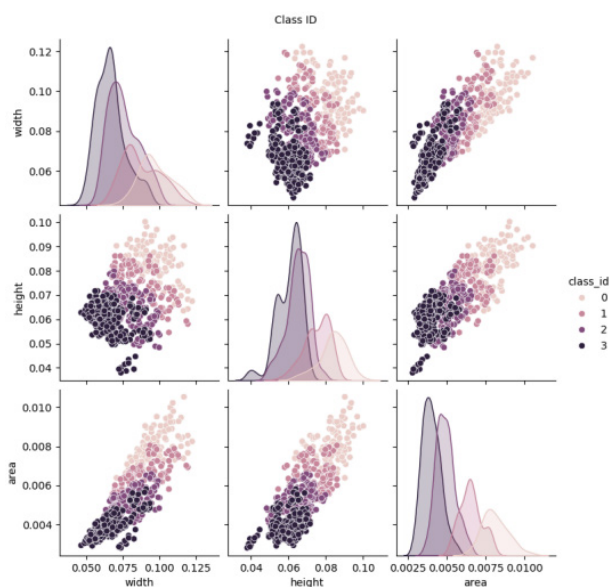


Figure 9 Graph showing width, length, and area of cashew nuts of each size

จากกราฟ Pair Plot ใน Figure 9 แสดงความสัมพันธ์ระหว่างตัวแปรความกว้าง ความยาวและพื้นที่ของเม็ดมะม่วงหิมพานต์โดยแบ่งตาม class_id ทั้ง 4 ขนาด (0-3) ตามขนาดของเม็ดมะม่วงหิมพานต์จากเล็กไปใหญ่ พบว่าข้อมูลมีการแบ่งกลุ่มที่ค่อนข้างชัดเจน โดยเฉพาะเมื่อพิจารณาความสัมพันธ์ระหว่างความกว้างกับความยาวที่แสดงการจัดกลุ่มได้ดี ส่วนกราฟความหนาแน่น (Density plot) บนแนวทแยงมุมแสดงการกระจายตัวของแต่ละตัวแปร ซึ่งมีการซ้อนทับกันบางส่วนระหว่างคลาส แต่ยังสามารถแยกแยะได้ นอกจากนี้ยังพบความสัมพันธ์เชิงบวกระหว่างตัวแปรทั้งสาม คือเมื่อค่าหนึ่งเพิ่มขึ้น อีกค่าหนึ่งมักจะเพิ่มขึ้นด้วย ลักษณะการกระจายตัวเช่นนี้บ่งชี้ว่าชุดข้อมูลนี้น่าจะเหมาะสมสำหรับการใช้ในการสร้างโมเดลคัดแยกขนาดเม็ดมะม่วงหิมพานต์ เนื่องจากมีการแบ่งกลุ่มที่ค่อนข้างชัดเจน

เมื่อพิจารณานำผลการทดลองโมเดลตรวจจับเม็ดมะม่วงหิมพานต์มีค่า Precision 0.9817 และ Recall สูงถึง 0.9706 บ่งบอกว่าโมเดลสามารถตรวจจับเม็ดมะม่วงหิมพานต์ส่วนใหญ่ได้ดีมีโอกาสพลาดน้อย อย่างไรก็ตาม เมื่อส่งผลการตรวจจับเม็ดมะม่วงหิมพานต์ไปยัง โมเดลคัดแยกขนาดเม็ด

มะม่วงหิมพานต์พบว่าค่า Precision และ Recall ใกล้เคียงกับ 1 หมายความว่าเมื่อโมเดลได้รับผลลัพธ์ของเม็ดมะม่วงหิมพานต์ที่ตรวจจับมาแล้วจะสามารถจำแนกขนาดได้อย่างแม่นยำ แต่ประสิทธิภาพโดยรวมของระบบอาจถูกจำกัดด้วยขั้นตอนการตรวจจับซึ่งยังมีการตรวจจับผิดพลาด

สรุปผลการทดลองและข้อเสนอแนะ

การศึกษานี้ได้ออกแบบระบบภาคตัดขนาดเม็ดมะม่วงหิมพานต์แบบผสมผสานโดยใช้ Deep Learning และ Machine Learning ร่วมกัน โดยใช้ Deep Learning ในการพัฒนาโมเดลตรวจจับเม็ดมะม่วงหิมพานต์ในภาพ ซึ่งสามารถระบุตำแหน่งของเม็ดมะม่วงหิมพานต์ในสภาพแสงและเงาที่หลากหลายได้อย่างมีประสิทธิภาพ ในขณะที่ Machine Learning ใช้ในการสร้างโมเดลสำหรับคัดขนาด เนื่องจากกระบวนการนี้ต้องการขอบเขตขนาดที่ชัดเจนโดยอาศัยลักษณะโครงสร้างหลัก เช่น ความกว้าง ความสูง หรือพื้นที่ ซึ่ง Machine Learning เหมาะสมกว่าการใช้ Deep Learning สำหรับงานนี้ เนื่องจากลักษณะเหล่านี้ตรงกับลักษณะที่ใช้ในการคัดขนาดโดยมนุษย์ ผลการทดลองของโมเดลตรวจจับเม็ดมะม่วงหิมพานต์แสดงให้เห็นสามารถตรวจจับเม็ดมะม่วงหิมพานต์ได้ส่วนใหญ่โดยมีโอกาสผิดพลาดน้อย ในขณะที่โมเดลจัดประเภทขนาดมีประสิทธิภาพสูงเมื่อได้รับผลการตรวจจับจากโมเดลแล้ว ระบบสามารถจัดประเภทขนาดได้อย่างถูกต้องทั้งหมด อย่างไรก็ตาม ประสิทธิภาพโดยรวมของระบบอาจได้รับผลกระทบจากขั้นตอนการตรวจจับที่ยังมีข้อผิดพลาดบ้างในบางกรณี ดังนั้นการศึกษในอนาคตจะพิจารณาปรับปรุงขั้นตอนการตรวจจับและพัฒนาโมเดลให้มีความแม่นยำมากยิ่งขึ้นในสภาพแวดล้อมที่หลากหลายมากยิ่งขึ้น เพื่อเพิ่มประสิทธิภาพของระบบในภาพรวม

เอกสารอ้างอิง

- Baitu, G. P., Gadalla, O. A. A., & Öztekin, Y. B. (2023). Traditional machine learning-based classification of cashew kernels using colour features. *Journal of Tekirdag Agricultural Faculty*, 20(1), 115–124. <https://doi.org/10.33462/jotaf.1100782>
- Kumari, A. S. S. S. R., Anjaneyulu, K. S. R., & Krishna, K. S. R. (2021). Classification of first quality fancy cashew kernels using four deep convolutional neural network models. *Materials Today: Proceedings*, 44, 3423–3429. <https://doi.org/10.1016/j.matpr.2020.11.897>

LearnOpenCV. (2021). *Mean average precision (mAP) in object detection*. <https://learnopencv.com/mean-average-precision-map-object-detection-model-evaluation-metric/>

- Pham, V.-N., Do Ba, Q.-H., Tran Le, D.-A., Nguyen, Q.-M., Do Van, D., & Nguyen, L. (2024). A low-cost deep-learning-based system for grading cashew nuts. *Computers*, 13(3), Article 71. <https://doi.org/10.3390/computers13030071>
- Song, X., Li, Z., Zhang, X., & Liu, J. (2022). Recognition and counting of apples in a dynamic state using a 3D camera and deep learning algorithms for robotic harvesting systems. *Sensors*, 22(19), Article 7012. <https://doi.org/10.3390/s22197012>
- Sunoj, S., Igathinathane, C., & Jenicka, S. (2018). Cashews whole and splits classification using a novel machine vision approach. *Computers and Electronics in Agriculture*, 145, 19. <https://doi.org/10.1016/j.compag.2017.12.034>
- World Population Review. (n.d.). *Cashew production by country*. Retrieved February 27, 2025, from <https://worldpopulationreview.com/country-rankings/cashew-production-by-country>
- Zha, Y., Yang, L., Liu, Y., Ma, Y., Li, D., & Wang, H. (2025). Weed recognition in cotton fields based on YOLOv8. *Intelligent Agriculture*, 1(1), Article 100019. <https://doi.org/10.1016/j.inag.2025.100019>
- Zhang, Y., Zhang, Y., & Zhang, Y. (2024). Apple fruit edge detection model using a rough set and convolutional neural network. *Sensors*, 24(7), Article 2283. <https://doi.org/10.3390/s24072283>