

การใช้โครงข่ายประสาทเทียมแบบคอนโวลูชันชนิด YOLOv8 ร่วมกับระบบแจ้งเตือนแบบเรียลไทม์เพื่อตรวจจับเศษโลหะในถุงพลาสติกในระหว่างกระบวนการผลิต

Using YOLOv8-Based Convolutional Neural Network with a Real-Time Notification System to Detect Metal Debris in Clear Plastic Pouches During Production

ดำรงศักดิ์ กิจเดช*, ทแก้ว เยี่ยมสวัสดิ์, สมิทธิ์ ชั่งสมบุญ, วัฒนา กลสิกุล และ วีระพันธ์ ดั่งทองสุข

สาขาวิชาวิศวกรรมเครื่องกล คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเอเชียอาคเนย์

19/1 ถนนเพชรเกษม แขวงหนองค้างพลู เขตหนองแขม กรุงเทพฯ 10160

*ติดต่อ: dumrongsakk@sau.ac.th, 02-8074500-27 ต่อ 404, 302

บทคัดย่อ

ปัจจุบันปัญหาในกระบวนการผลิตในโรงงานเริ่มมีการนำปัญญาประดิษฐ์ใช้งานร่วมกับกล้องในหลากหลายงาน เนื่องจากการใช้ปัญญาประดิษฐ์มีความยืดหยุ่นมากกว่าการใช้กระบวนการทางภาพ งานวิจัยนี้ต้องการตรวจจับเศษโลหะที่ติดมากับถุงพลาสติกใส ซึ่งมาจากกระบวนการเป่า อย่างไรก็ตาม ถุงพลาสติกใสทำให้มีการสะท้อนแสงที่ไม่แน่นอน และยังมี การพิมพ์ข้อความบนถุงไว้อยู่แล้วด้วย ซึ่งเศษโลหะมีรูปร่างไม่แน่นอนขนาดประมาณ 2 – 4 มิลลิเมตร เวลาในการตรวจจับ ถูกจำกัดที่ประมาณ 0.6 วินาที ระหว่างกระบวนการตัดถุงพลาสติก งานวิจัยนี้จึงได้นำเสนอการใช้โครงข่ายประสาทเทียมแบบคอนโวลูชันชนิด YOLOv8 ร่วมกับระบบแจ้งเตือนแบบเรียลไทม์เพื่อตรวจจับเศษโลหะในถุงพลาสติกในระหว่างกระบวนการผลิต หลักการทำงานคือกล้องจะทำการรับภาพและส่งไปยังคอมพิวเตอร์ เพื่อตรวจจับโลหะและส่งเป็นสัญญาณไปยังเสาสัญญาณไฟ การทดลองเป็นการตรวจจับโลหะโดยใช้ภาพที่มีอยู่แล้ว ผลที่ได้ค่าความถูกต้องอยู่ที่ 81.25 เปอร์เซ็นต์

คำหลัก: โครงข่ายประสาทเทียม, คอนโวลูชัน, ถุงพลาสติก, ตรวจจับโลหะ, ปัญญาประดิษฐ์

Abstract

In recent years, artificial intelligence integrated with vision systems has been increasingly used in manufacturing processes, as artificial intelligence is more flexible than image processing. This study focuses on the detection of metal debris attached to clear plastic pouches during the bag blowing process. However, clear plastic pouches exhibit unpredictable light reflections and contain printed text. The irregularly shaped metal debris has an approximate size of 2 – 4 millimeters. The detection process is limited to approximately 0.6 seconds during the cutting process. this research proposes a YOLOv8-based convolutional neural network with a real-time notification system to detect metal debris in clear plastic pouches during production. The working principle of the system involves capturing images from a

camera and transmitting them to a computer for metal debris is detected, a signal sent to activate an indicator light. The experimental setup was designed to evaluate the system by comparing existing images. The results indicated that the accuracy was 82 percent.

Keywords: Neural Networks, Convolution, Plastic Pouches, Metal Debris Detection, AI

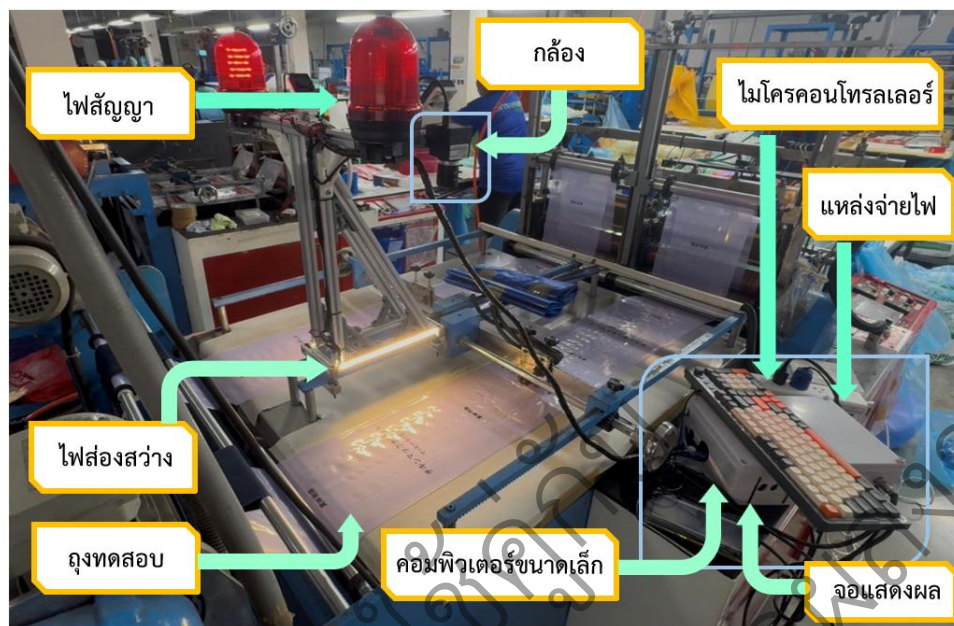
1. บทนำ

โจทย์ปัญหาในอุตสาหกรรมมีจำนวนมากที่มีความซับซ้อนมาก จำเป็นต้องใช้ปัญญาประดิษฐ์เข้ามาช่วยในการแก้ไขปัญหา ซึ่งในงานวิจัยนี้ก็เช่น ทางโรงงานได้ถูกลูกค้าสั่งคืน (Reject) กระจกพลาสติกที่สั่งทำสำหรับการหีบห่อสินค้า เนื่องจากมีเศษโลหะติดมาที่กระจกพลาสติกใส ซึ่งแหล่งที่มาเกิดจากการสึกหลอของเฟืองที่อยู่ในกระบวนการเป่า เศษโลหะนี้มีขนาดอยู่ระหว่าง 2-4 มิลลิเมตร ส่วนกระจกมีขนาดเท่ากับกระดาษ A4 งานวิจัยนี้จึงใช้กระจกที่ถูกตีกลับมาเป็นตัวอย่างในการทดลอง โดยจากเหตุผลข้างต้นการใช้โครงข่ายประสาทเทียมแบบคอนโวลูชันชนิด YOLOv8 ร่วมกับระบบแจ้งเตือนแบบเรียลไทม์เพื่อตรวจจับเศษโลหะในกระจกพลาสติกใสระหว่างกระบวนการผลิต จึงเป็นการตอบโจทย์ ในงานวิจัยนี้ทำการทบทวนวรรณกรรมส่วนของการหาปัญญาประดิษฐ์ที่เหมาะสมกับข้อจำกัดของงานวิจัยนี้

ปัญญาประดิษฐ์มีหลายชนิด จึงต้องศึกษาว่าปัญญาประดิษฐ์ชนิดไหนเหมาะสมที่สุดกับงานวิจัยนี้ โดยจากการศึกษาพบว่าปัญญาประดิษฐ์ที่ใช้ในงานตรวจหาวัตถุอยู่ 2 แบบหลักๆ คือ Mask R-CNN (Region-based Convolutional Neural Network) และ YOLO (You Only Look Once) โดยมีผลดังนี้ Mask R-CNN ถูกใช้ในงานที่หลากหลาย เช่นการตรวจจับใบหน้า [1] การตรวจจับ dPCR (digital polymerase chain reaction) อันเนื่องมาจากโครงสร้างที่ไม่เป็นระเบียบของไมโครอาร์เรย์สำหรับการใช้งานทางการแพทย์ [2] การตรวจจับป้ายทะเบียนรถที่มีมุมเอียงมากผิดปกติ [3] การตรวจจับพื้นที่ว่างในลานจอดรถ [4], การระบุทิศทางของวัตถุที่มีรูปทรงสมมาตร [5] และ การตรวจจับก๊อกน้ำทองเหลือง

ที่มีรูปร่างของครีที่ไม่แน่นอนและมีการสะท้อนแสง [6] ขณะที่ YOLO มีการใช้อย่างแพร่หลายเช่นกันโดยมีการพัฒนาอย่างต่อเนื่องโดยมีตั้งแต่ YOLOv1 – YOLOv11 ตัวอย่างงานวิจัย เช่น การนำ YOLOv5 มาใช้ในการคัดแยกชนิดแคคตัสและระบุตำแหน่งเพื่อการรดน้ำต้นไม้ในปริมาณที่เหมาะสม [7] และยังมีงานนำ YOLOv5 มาใช้ในการทำนายท่าทางของคนเพื่อใช้ในการควบคุมหุ่นยนต์อีกด้วย [8] ขณะที่ [9] ใช้ YOLOv8 ในการตรวจหาโรคในแคคตัสเพื่อช่วยในการฉีดยากำจัดเพลี้ยแป้งได้ก่อนลุกลาม [10] เป็นการเปรียบเทียบระหว่าง YOLO V3 กับ YOLO V5 ซึ่งผลที่ได้คือ YOLO V5 มีความเร็วที่มากกว่าในส่วนของการฝึกข้อมูล ส่วนของความเร็วในการตรวจจับวัตถุและระบุตำแหน่งทำให้เร็วกว่าเมื่อใช้โครงสร้างของโครงข่ายประสาทเทียมขนาดเล็กสุด ในขณะที่ใช้โครงข่ายประสาทเทียมขนาดใหญ่จะได้ความเร็วในการระบุตำแหน่งวัตถุที่เท่ากัน และ B. Benjdira et al. [11] ทำการเปรียบเทียบ Fast R-CNN กับ YOLO V3 ซึ่ง YOLO V3 มีประสิทธิภาพที่มากกว่าสรุปคือ Mask R-CNN มีความเร็วที่ต่ำกว่าแต่มีความยืดหยุ่นในการปรับแก้ไขหรือประยุกต์ใช้งานมากกว่า YOLO โดยงานวิจัยนี้เพียงแค่นำไปใช้เท่านั้นและต้องการความเร็วที่มากจึงมุ่งเน้นไปที่ YOLO ซึ่ง YOLO ที่ใช้เป็น YOLOv8 เนื่องจากข้อจำกัดในส่วนของฮาร์ดแวร์

ระเบียบวิธีวิจัยถูกแบ่งเป็น 4 คือ อภิปรายการทดลอง การพัฒนาระบบควบคุม การเตรียมข้อมูลและการฝึกข้อมูล และขั้นตอนการทดลอง โดยมีรายละเอียดดังต่อไปนี้



รูปที่ 1 ส่วนประกอบอุปกรณ์ที่ใช้ในการทดลอง

2. ระเบียบวิธีวิจัย

ระเบียบวิธีวิจัยถูกแบ่งเป็น 4 คือ อภิปรายการทดลอง การพัฒนาระบบควบคุม การเตรียมข้อมูลและการฝึกข้อมูล และขั้นตอนการทดลอง โดยมีรายละเอียดดังต่อไปนี้

2.1 อุปกรณ์การทดลอง

อุปกรณ์ที่ใช้ในการทดลองมีส่วนประกอบหลักดังนี้คือ กล้อง Webcam Logitech C525 ความละเอียด 720p อัตราภาพ 30 fps ใช้ในการเก็บรวบรวมภาพสำหรับฝึกปัญญาประดิษฐ์ชนิด YOLOv8 และสำหรับตรวจจับโรคแบบเรียลไทม์ต่อไป ส่วนต่อมาเป็นบอร์ดคอมพิวเตอร์ขนาดเล็กใช้เป็น Jetson Orin nano reComputer J3011 ทำหน้าที่เป็นปัญญาประดิษฐ์ ส่วนถัดมาบอร์ดไมโครคอนโทรลเลอร์ใช้เป็นบอร์ด ESP32 ทำหน้าที่เป็นตัวรับสัญญาณคำสั่งจาก Jetson Orin nano ผ่านช่อง USB เพื่อเปิดปิดไฟสัญญาณ โดยมีโซลิตสเตรียเป็นตัวกลางอีกที โดยระบบไฟใช้ไฟกระแสตรง 24 โวลต์ จอแสดงผลแบบจอสัมผัส Mellow Fly HDMI7 ขนาดจอ 86 x 154 มิลลิเมตร ความละเอียดหน้าจอ 1024 x 600 พิกเซล ใช้ในการแสดงผลภาพการตรวจหาเชื้อโลหะแบบเวลาจริง ไฟสัญญาณมีหลอดแสง 3 โหมด

และโหมดเสียง 3 โหมด ฐานกว้าง 30 เซนติเมตร สูง 40 เซนติเมตร เพื่อให้มีระยะห่างพอสำหรับเก็บภาพทั้งชุด โดยที่ตำแหน่งของส่วนประกอบทั้งหมดแสดงดังรูปที่ 1

2.2 การพัฒนาระบบควบคุม

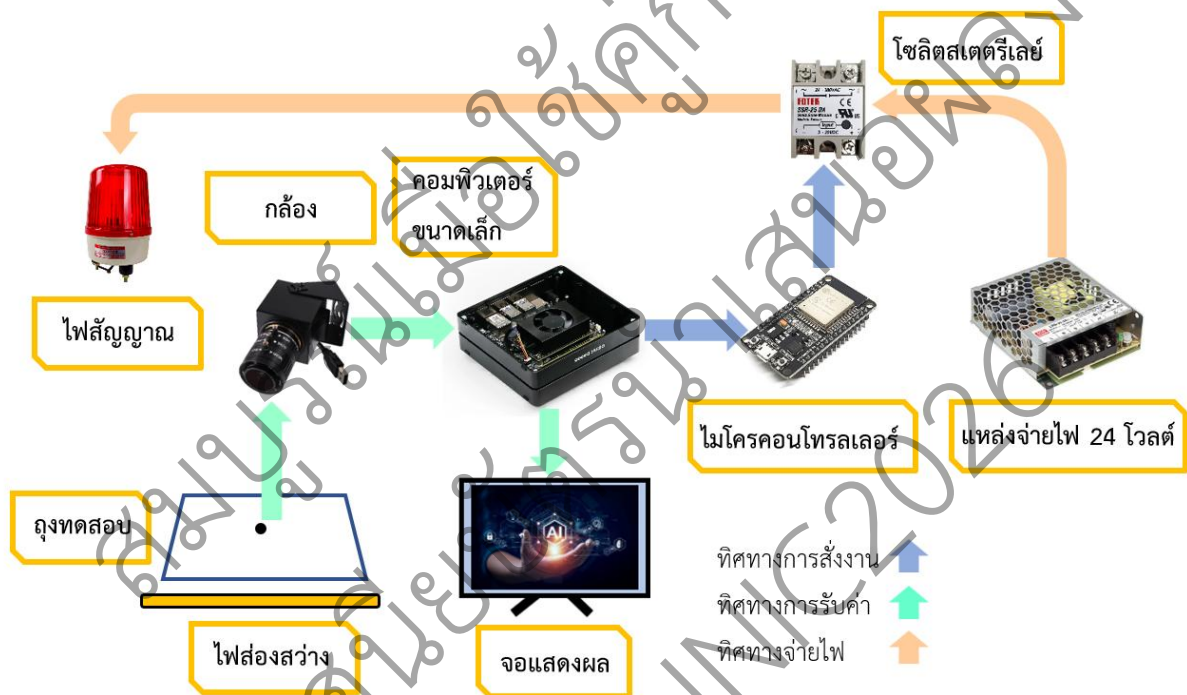
ภาพรวมของระบบแจ้งเตือนแบบเรียลไทม์เพื่อตรวจจับเชื้อโลหะแสดงในรูปที่ 2 ระบบการทำงานเริ่มจากกล้องจะทำการถ่ายภาพเมื่อภาพ 2 ภาพ ไม่มีการเปลี่ยนแปลง เพื่อแก้ปัญหาภาพเบลอ จากนั้นภาพจะถูกส่งไปยัง YOLOv8 ที่อยู่ใน Jetson Orin nano เพื่อทำการตรวจหาเชื้อโลหะบนถุงขนาด A4 ซึ่งแสงปกติของโรงงานไม่เพียงพอและเกิดแสงสะท้อนที่บดบังการมองเห็นเชื้อโลหะได้ง่ายจึงต้องมีการติดตั้งไฟแสงธรรมชาติเพิ่มเติมบริเวณข้างถุงเพื่อหลีกเลี่ยงการสะท้อนโดยตรง จากนั้นภาพจะถูกแสดงจอแสดงผลเพื่อให้สามารถตรวจสอบความถูกต้องได้ หากมีการตรวจพบเชื้อโลหะ YOLOv8 จะทำการตีกรอบบริเวณที่มีเชื้อโลหะและส่งสัญญาณผ่าน USB ไปยังบอร์ด ESP32 เพื่อสั่งเปิดไซเรนอีกทีหนึ่ง อย่างไรก็ตามตัวไซเรนใช้ไฟกระแสตรง 24 โวลต์ จึงจำเป็นต้องส่งสัญญาณไปที่โซลิตสเตรียอีกทีหนึ่ง โดยมีเงื่อนไขว่าหากมีการพบเชื้อโลหะ 1 ชิ้น จะทำการเปิดไซเรน 3 วินาที แต่หากมี

มากกว่านั้นระบบจะทำการบวกเวลาเพิ่มไปเรื่อย ๆ เพื่อแก้ปัญหาการตรวจพบติดกันโดยที่ไซเรนยังไม่หยุดทำงาน

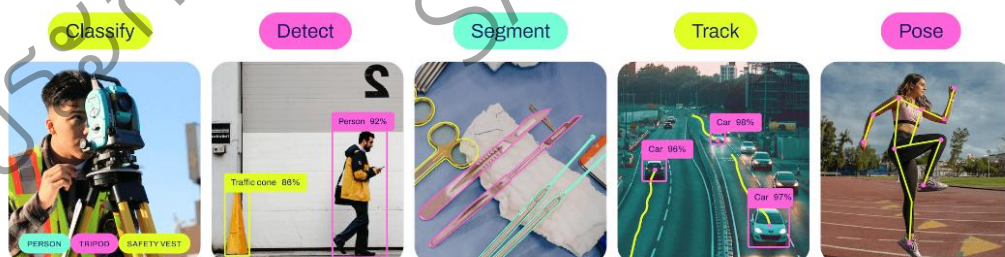
2.3 การเตรียมข้อมูล (Dataset) และการฝึกข้อมูล (Data Training)

งานวิจัยนี้ใช้ปัญญาประดิษฐ์ชนิด YOLOv8 (You Only Look Once Version 8) การตรวจหาเศษโลหะบนถุงขนาด A4 โดยที่ YOLOv8 มีพื้นฐานมาจากเทคนิคโครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolution Neural Networks) ร่วมกับการเรียนรู้เชิงลึก (Deep Learning) โดย Ultralytics [12] แสดงให้

เห็นว่า YOLOv8 ถูกแบ่งออกเป็น 5 แบบหลัก ได้แก่ (1) การจำแนกประเภทวัตถุ (Classify), (2) การตรวจจับตำแหน่งของวัตถุ (Detect), (3) การแยกขอบเขตของวัตถุ (Segment), (4) การติดตามการเคลื่อนไหวของวัตถุ (Track) และ (5) การระบุท่าทางของวัตถุ (Pose) ดังแสดงในรูปที่ 3 โดยงานวิจัยนี้เลือกใช้ Segment ซึ่งมีความเหมาะสมกับงานมากที่สุด



รูปที่ 2 การเชื่อมต่อของระบบและหลักการทำงาน



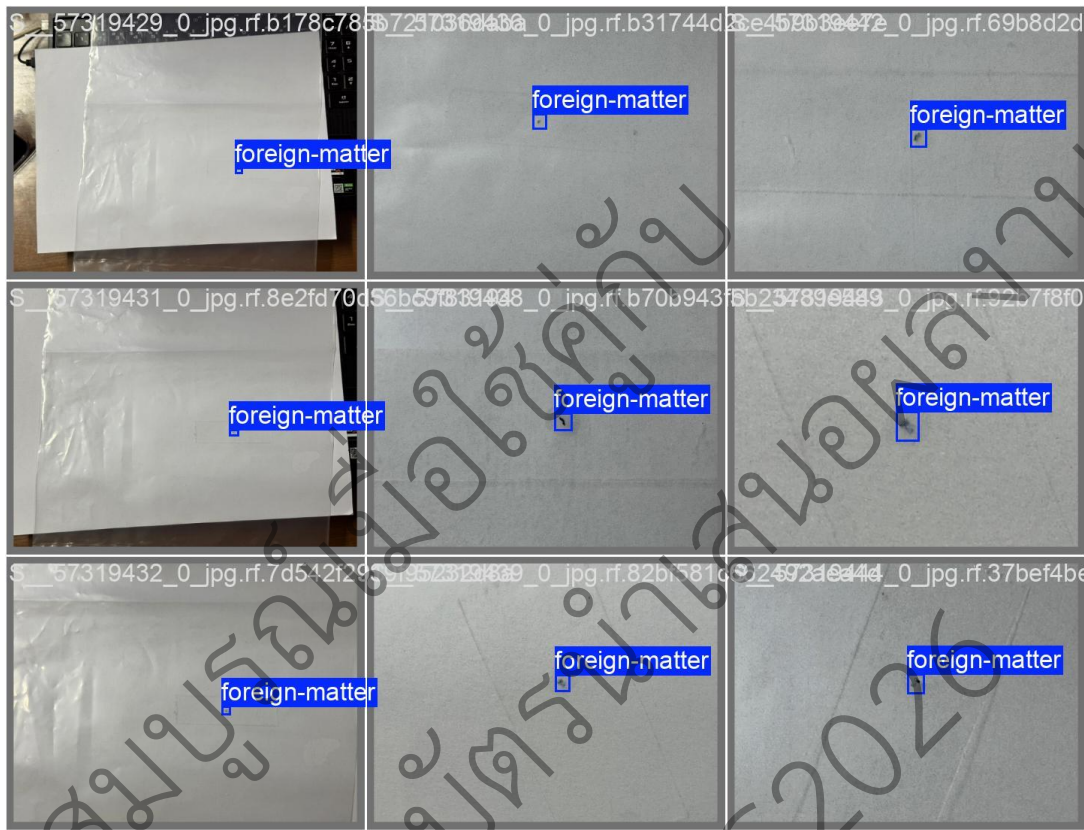
รูปที่ 3 รูปแบบในการทำปัญญาประดิษฐ์ของ YOLOv8 (Ultralytics [20])

ปัจจุบัน YOLO เวอร์ชันล่าสุด คือ YOLOv11 อย่างไรก็ตามเนื่องจากระบบปฏิบัติการบนบอร์ด Jetson

Orin Nano ไม่รองรับการทำงานของ YOLOv11 บนการ์ดแสดงผล Nvidia ในปัจจุบัน รองรับเพียงแค่

YOLOv8 โครงข่ายประสาทเทียม (Neural Networks) คือการจำลองโครงข่ายประสาทในสมองของมนุษย์ซึ่งช่วยให้สามารถตัดสินใจเรื่องที่มีความซับซ้อนได้ เช่น พื้นหลังที่มีการเปลี่ยนแปลงตลอดเวลา แสงและเงาที่ไม่แน่นอน

เป็นต้น YOLOv8 มีความเร็วในการประมวลผลสูง ส่วนของความแม่นยำขึ้นอยู่กับความซับซ้อนของโครงข่ายประสาทเทียม จำนวนของภาพและความหลากหลายของภาพที่นำมาฝึกปัญญาประดิษฐ์



รูปที่ 4 แสดงตัวอย่าง Dataset ที่ถูกใช้ในการทดลองนี้

ภาพที่ใช้ในการฝึกปัญญาประดิษฐ์จะใช้ภาพที่ได้จากกล้อง Webcam ถูกลดความละเอียดลงเหลือ 640 x 640 pixel โดยภาพที่ใช้ในการฝึกมีรูปร่างและตำแหน่งของเศษโลหะที่แตกต่างกันโดยมีภาพเริ่มต้นอยู่ที่ 42 ภาพ จากนั้นภาพถูกทำให้เพิ่มขึ้นเป็น 210 ภาพเพื่อใช้เป็นชุดข้อมูล (Dataset) สำหรับการฝึกปัญญาประดิษฐ์โดยใช้กระบวนการ Augmentation หมุนภาพ 90 องศา (Rotate) เอียงภาพ (Shear) และเบลอภาพ (Blur) จากนั้นทำฉลาก (Label) ให้กับ Dataset ทั้งหมดถูกทำบนเว็บไซต์ Roboflow ตัวอย่างภาพที่ถูกทำ Label แล้วถูกแสดงดังรูปที่ 4 ค่าตัวแปรต่าง ๆ ที่ใช้ในการฝึกปัญญาประดิษฐ์มีดังนี้ จำนวนชั้นในการฝึก (Layer)

ทั้งหมด 409 เลเยอร์, มีพารามิเตอร์ทั้งหมด 20,057,634 พารามิเตอร์, Optimizer เป็น Adam, อัตราการเรียนรู้ (Learning Rate) อยู่ที่ 0.001, Momentum อยู่ที่ 0.9, Batch size ใช้เป็น 60% ของ CUDA Memory จำนวนรอบในการฝึก (Epoch) อยู่ที่ 150 รอบ

การฝึกปัญญาประดิษฐ์จำเป็นต้องใช้คอมพิวเตอร์ที่มีการ์ดจอที่มีความสามารถในการประมวลผลที่สูงมากสำหรับการการฝึกปัญญาประดิษฐ์ โดยงานวิจัยนี้ทำการฝึกบน Google Collab ซึ่งเป็นการประมวลผลแบบออนไลน์ โดยใช้การ์ดจอ NVIDIA Tesla T4 ใช้ Python เวอร์ชัน 3.12.11 ในการเขียนโปรแกรมฝึกและใช้ Pytorch เวอร์ชัน 2.8.0 เป็น Back Bone ส่วนโปรแกรม

ที่ใช้ความคุมการทำงานการจ้งเป็น cu126 และเวลาที่ใช้ในการฝึกข้อมูลทั้งหมด 0.82 ชั่วโมง หลังจากทำการฝึกจะได้แบบจำลอง (Model) สำหรับการตรวจหาเศษโลหะและระบุตำแหน่งโดยการติเส้นขอบวัตถุ

2.4 ขั้นตอนในการทดลอง

การทดลองเป็นการหาค่าความถูกต้องของปัญญาประดิษฐ์ผ่านผลลัพธ์การฝึก (Training) และการประเมินผล (Validation) บนภาพที่ถ่ายมาแล้วเพียงอย่างเดียวจำนวน 16 ภาพ

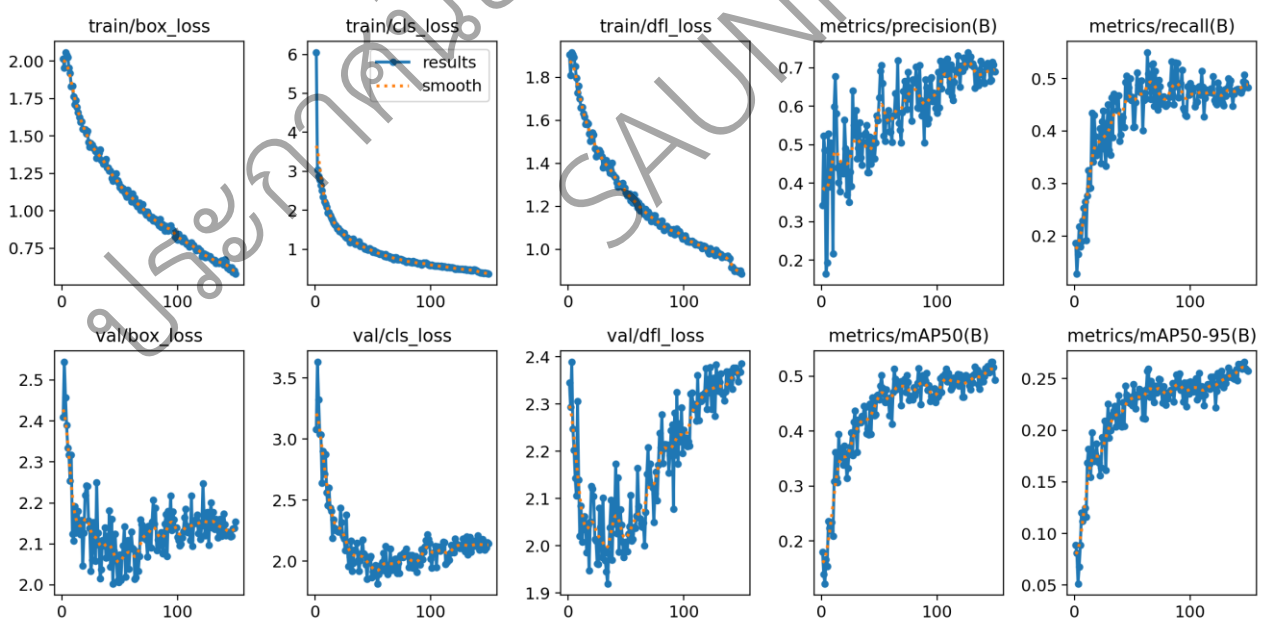
3. ผลการทดลอง

ผลการฝึกปัญญาประดิษฐ์ถูกแสดงดังรูปที่ 5 ซึ่งแสดงให้เห็นสามารถเรียนรู้จากข้อมูลได้อย่างมีประสิทธิภาพ โดยพิจารณาจากค่าความสูญเสียของข้อมูลฝึก (training loss) นั่นคือ box loss, classification loss และ distribution focal loss (DFL) มีแนวโน้มลดลงอย่างต่อเนื่องตลอดกระบวนการฝึกสอน แสดงให้เห็นว่าพารามิเตอร์ของแบบจำลองได้รับการปรับค่าอย่างเหมาะสมและแบบจำลองมีการลู่เข้า (convergence) ที่ดี อย่างไรก็ตาม เมื่อพิจารณาค่าความสูญเสียของข้อมูลประเมินผล (validation loss) พบว่าแม้ค่า box loss และ classification loss จะลดลงในช่วงแรก ยังคงมี

แนวโน้มคงที่หรือเพิ่มขึ้นเล็กน้อยในช่วงท้ายของการฝึก นอกจากนี้ค่า validation DFL ยังมีแนวโน้มเพิ่มขึ้นอย่างชัดเจน แสดงให้เห็นว่าเริ่มมีการจดจำเฉลยแทนที่จะเรียนรู้ (overfitting) ในระดับเล็กน้อย

ค่าความแม่นยำ (precision) มีแนวโน้มเพิ่มขึ้นจนมีค่าประมาณ 0.7 แสดงให้เห็นว่าแบบจำลองมีความแม่นยำในการทำนายวัตถุที่ตรวจพบค่อนข้างดี ในขณะที่ค่า recall อยู่ประมาณ 0.5 ซึ่งบ่งชี้ว่ายังมีวัตถุบางส่วนที่แบบจำลองไม่สามารถตรวจจับได้อยู่ ส่งผลทำให้เกิดความไม่สมดุลระหว่าง precision และ recall โดยโมเดลมีแนวโน้มทำนายแบบอนุรักษ์นิยม (conservative prediction) กล่าวคือ ปัญญาประดิษฐ์เลือกทำนายเฉพาะกรณีที่มีมั่นใจสูงเท่านั้น โดยจะไม่ทำนายหากไม่แน่ใจ

ส่วนของค่าความแม่นยำเฉลี่ย (mean Average Precision: mAP) พบว่า mAP@0.5 (mAP50) มีค่าประมาณ 0.5 และ mAP@0.5:0.95 (mAP50-95) อยู่ที่ประมาณ 0.25 ซึ่งอยู่ในระดับปานกลาง แสดงให้เห็นว่าแบบจำลองสามารถตรวจจับวัตถุได้ในระดับที่ยอมรับได้ แต่ยังมีข้อจำกัดในด้านความแม่นยำของตำแหน่งอยู่บ้าง



รูปที่ 5 ผลการฝึกข้อมูลและประเมินผลข้อมูล



รูปที่ 6 แสดงผลการตรวจหาเศษโลหะและระบุตำแหน่ง

ส่วนรูปที่ 6 แสดงให้เห็นว่าแบบจำลองยังมีความแม่นยำต่ำอยู่ ค่าต่ำสุด 30% ค่าสูงสุด 70% โดยมีค่าเฉลี่ยอยู่ที่ 61.87%

4. สรุป

แบบจำลองมีความสามารถในการเรียนรู้และให้ผลลัพธ์ในระดับที่น่าพอใจ อย่างไรก็ตาม เนื่องจากทางโรงงานให้ข้อมูลตัวอย่างและงบประมาณมาค่อนข้างจำกัด ซึ่งหากไม่มีข้อจำกัดดังกล่าวข้างต้น จะทำให้สามารถปรับปรุงประสิทธิภาพได้ โดยการเพิ่มจำนวนข้อมูล การใช้เทคนิคการเพิ่มข้อมูล (data augmentation) และการปรับแต่งพารามิเตอร์ของโมเดล เพื่อเพิ่มค่า recall ลดปัญหา overfitting และยกระดับค่า mAP ให้สูงขึ้นได้

หากกำหนดค่าความมั่นใจที่ 60% เพื่อหลีกเลี่ยงการตรวจจับเศษโลหะผิดพลาด จะทำให้ตรวจจับภาพที่มีอยู่จริงไม่พบ ดังนั้นจึงสามารถตรวจจับได้ 81.25%

5. กิตติกรรมประกาศ

ขอขอบคุณสาขาวิชาวิศวกรรมเครื่องกล คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเอเชียอาคเนย์ ที่เอื้อเฟื้อสถานที่ในการทดลอง ขอขอบคุณบริษัท อินทราพรแพค จำกัด ที่เอื้อเฟื้อสถานที่ในการทดลองและตัวอย่างที่ใช้ในการทดลอง

6. เอกสารอ้างอิง

[1] Lin, K., et al. (2020). Face detection and segmentation based on improved Mask R-

- CNN. Discrete Dynamics in Nature and Society, 2020, 1–11.
- [2] Hu, Z., et al. (2019). A novel method based on a Mask R-CNN model for processing dPCR images. *Analytical Methods*, 11(27), 3410–3418.
- [3] Lin, C.-H., & Li, Y. (2019). A license plate recognition system for severe tilt angles using Mask R-CNN. In *2019 International Conference on Advanced Mechatronic Systems (ICAMEchS)* (pp. 229–234). IEEE.
- [4] Agrawal, T., & Urolagin, S. (2020). Multi-angle parking detection system using Mask R-CNN. In *Proceedings of the 2020 2nd International Conference on Big Data Engineering and Technology* (pp. 76–80).
- [5] He, P., Zuo, L., Zhang, C., & Zhang, Z. (2019). A value recognition algorithm for pointer meter based on improved Mask R-CNN. In *2019 9th International Conference on Information Science and Technology (ICIST)* (pp. 108–113). IEEE.
- [6] Kijdech, D., & Vongbunyong, S. (2024). Manipulation of a complex object using dual-arm robot with Mask R-CNN and grasping strategy. *Journal of Intelligent & Robotic Systems*, 110, 103. <https://doi.org/10.1007/s10846-024-02132-0>
- [7] Kijdech D. and Vongbunyong S. (2021). Artificial Intelligence in Localization and Classification of Cactus for Automatic Watering Works, The 8 th SAU National Interdisciplinary Conference 2021, 4-5 June 2021.
- [8] Kijdech, D., & Duangthongsuk, W. (2022). Using the convolution neural network to predict human body languages for two wheel drive robot control. *SAU Journal of Science & Technology*, 8(1). 28-39.
- [9] Kijdech, D., & Duangthongsuk, W. (2025). Using the convolution neural network (YOLOv8) with Cartesian robot for automatic disease detection and watering. *RMUTP Research Journal Sciences and Technology*, 19(2), 101-118.
- [10] Kuznetsova A., Maleva T. and Soloviev V., (2020). “Detecting Apples in Orchards Using YOLOv3 and YOLOv5 in General and Close-Up Images,” *Advances in Neural Networks – ISNN*, pp. 233-243.
- [11] Benjdira B., Khursheed T., Koubaa A., Ammar A. and Ouni K., (2019). “Car Detection using Unmanned Aerial Vehicles: Comparison between Faster R-CNN and YOLOv3 ,” *Proceedings of the 1 st International Conference on Unmanned Vehicle Systems (UVS)*, Muscat, Oman, 5-7 February.
- [12] Ultralytics, "YOLOv8 Docs – Modes: Detect, Segment, Classify, Pose, Track," [Online]. Available: <https://docs.ultralytics.com/modes/>. [Accessed: Jun. 6, 2025].