

การใช้โครงข่ายประสาทเทียมแบบคอนโวลูชันชนิด YOLOv8 ร่วมกับระบบแจ้งเตือนแบบเรียลไทม์เพื่อตรวจจับเศษโลหะในถุงพลาสติกใสระหว่างกระบวนการผลิต

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

ดำรงศศักดิ์ ทิวเดช, ทนถวัลย์ เยี่ยมสวัสดี, สมินทร์ ชิงสมบูรณ์, วัฒนา กสิกุล และ วีระพันธ์ ด้วงทองสุข

บทคัดย่อ

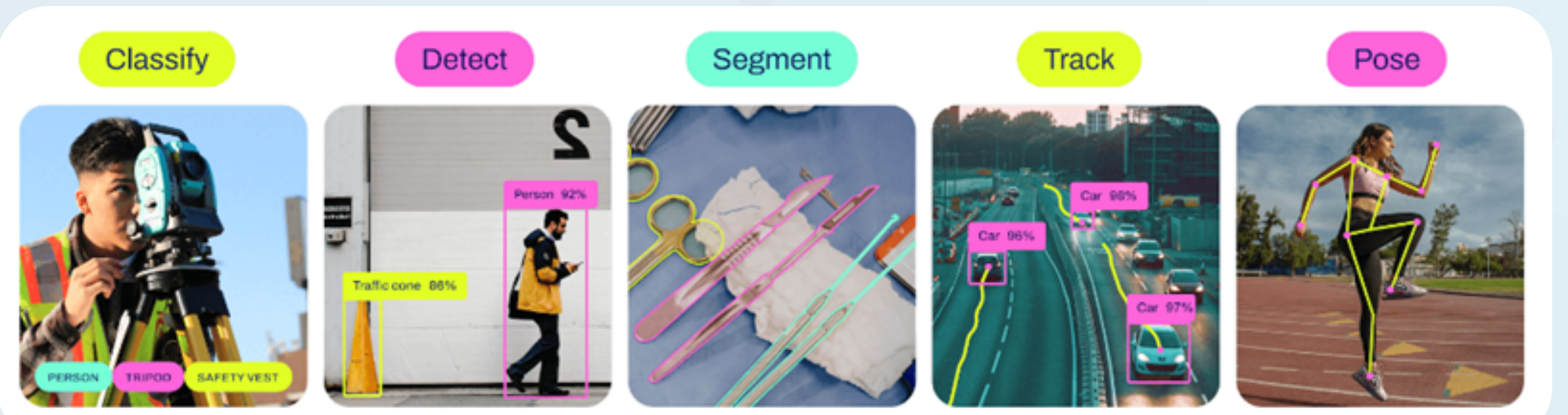
ปัจจุบันปัญหาในกระบวนการผลิตในโรงงานเริ่มมีการนำปัญญาประดิษฐ์เข้ามาช่วยร่วมกับกล้องในหลากหลายงาน เนื่องจากการใช้ปัญญาประดิษฐ์มีความยืดหยุ่นมากกว่าการใช้กระบวนการทางภาพ งานวิจัยนี้ต้องการตรวจจับเศษโลหะที่ติดมากับถุงพลาสติกใส ซึ่งมาจากกระบวนการเป่า อย่างไรก็ตาม ถุงพลาสติกใสทำให้มีการสะท้อนแสงที่ไม่แน่นอน และยังมีกรรมพันธุ์ผิวของถุงพลาสติกใสด้วย ซึ่งเศษโลหะมีรูปร่างไม่แน่นอนขนาดประมาณ 2 - 4 มิลลิเมตร เวลาในการตรวจจับถูกจำกัดที่ประมาณ 0.6 วินาที ระหว่างกระบวนการตัดถุงพลาสติก งานวิจัยนี้จึงได้นำเสนอการใช้โครงข่ายประสาทเทียมแบบคอนโวลูชันชนิด YOLOv8 ร่วมกับระบบแจ้งเตือนแบบเรียลไทม์เพื่อตรวจจับเศษโลหะในถุงพลาสติกใสระหว่างกระบวนการผลิต หลักการทำงานคือกล้องจะทำการรับภาพและส่งไปยังคอมพิวเตอร์ เพื่อตรวจจับโลหะและส่งเป็นสัญญาณไปยังเสาสัญญาณไฟ การทดลองเป็นการตรวจจับโลหะโดยใช้ภาพที่มีอยู่แล้ว ผลที่ได้ค่าความถูกต้องอยู่ที่ 81.25 เปอร์เซ็นต์

บทนำ

โจทย์ปัญหาในอุตสาหกรรมมีจำนวนมากที่มีความซับซ้อนมาก จำเป็นต้องใช้ปัญญาประดิษฐ์เข้ามาช่วยในการแก้ไขปัญหานั้น ซึ่งในงานวิจัยนี้ก็คือ ทางโรงงานได้ถูกลูกค้าสั่งคืน (Reject) ถุงพลาสติกใสที่สั่งทำสำหรับการหีบห่อสินค้า เนื่องจากมีเศษโลหะติดมากับถุงพลาสติกใส ซึ่งแหล่งที่มาเกิดจากการสีกหล่อของเฟืองที่อยู่ในกระบวนการเป่า เศษโลหะนี้มีขนาดอยู่ระหว่าง 2-4 มิลลิเมตร ส่วนถุงมีขนาดเท่ากับกระดาษ A4 งานวิจัยนี้จึงใช้โครงข่ายประสาทเทียมแบบคอนโวลูชันชนิด YOLOv8 ร่วมกับระบบแจ้งเตือนแบบเรียลไทม์เพื่อตรวจจับเศษโลหะในถุงพลาสติกใสระหว่างกระบวนการผลิต จึงเป็นการตอบโจทย์ ในงานวิจัยนี้ทำการทบทวนวรรณกรรมส่วนของการหาปัญญาประดิษฐ์ที่เหมาะสมกับข้อจำกัดของงานวิจัยนี้

ระเบียบวิธีวิจัย

ระเบียบวิธีวิจัยถูกแบ่งเป็น 4 คือ อุลตรนิกการทดลอง การพัฒนาระบบควบคุม การเตรียมข้อมูลและการฝึกข้อมูล และขั้นตอนการทดลอง โดยมีรายละเอียดดังต่อไปนี้



รูปที่ 3 รูปแบบในการทำปัญญาประดิษฐ์ของ YOLOv8 (Ultralytics [20])

อุปกรณ์การทดลอง

อุปกรณ์ที่ใช้ในการทดลองมีส่วนประกอบหลักดังนี้คือ กล้อง Webcam Logitech C525 ความละเอียด 720p อัตราภาพ 30 fps ใช้ในการเก็บรวบรวมภาพสำหรับฝึกปัญญาประดิษฐ์ชนิด YOLOv8 และสำหรับตรวจจับโรคแบบเรียลไทม์ต่อไป ส่วนต่อมาคือบอร์ดคอมพิวเตอร์ขนาดเล็กใช้เป็นที่ Jetson Orin nano reComputer J3011 ทำหน้าที่เป็นปัญญาประดิษฐ์ ส่วนตัวบอร์ดไมโครคอนโทรลเลอร์ใช้เป็นที่บอร์ด ESP32 ทำหน้าที่เป็นตัวรับสัญญาณคำสั่งจาก Jetson Orin nano ผ่านช่อง USB เพื่อเปิดปิดไฟสัญญาณ โดยมีชุดเสตดรีเลย์เป็นตัวกลางอีกที โดยระบบไฟใช้ไฟกระแสตรง 24 โวลต์ จอแสดงผลแบบจอสัมผัส Mellow Fly HDMI7 ขนาดจอ 86 x 154 มิลลิเมตร ความละเอียดหน้าจอ 1024 x 600 พิกเซล ใช้ในการแสดงผลภาพการตรวจหาเศษโลหะแบบเวลาจริง ไฟสัญญาณมีโหมดแสง 3 โหมด และโหมดเสียง 3 โหมดฐานกว้าง 30 เซนติเมตร สูง 40 เซนติเมตร เพื่อให้มีระยะห่างพอสำหรับเก็บภาพทุกทั้งถุง โดยที่ตำแหน่งของส่วนประกอบทั้งหมดแสดงดังรูปที่ 1

การพัฒนาระบบควบคุม

ภาพรวมของระบบแจ้งเตือนแบบเรียลไทม์เพื่อตรวจจับเศษโลหะแสดงในรูปที่ 2 ระบบการทำงานเริ่มจากกล้องจะทำการถ่ายภาพเมื่อภาพ 2 ภาพ ไม่มีการเปลี่ยนแปลง เพื่อแก้ปัญหาภาพเบลอ จากนั้นภาพจะถูกส่งไปยัง YOLOv8 ที่อยู่ใน Jetson Orin nano เพื่อทำการตรวจหาเศษโลหะบนถุงขนาด A4 ซึ่งแสงปกติของโรงงานไม่เพียงพอและเกิดแสงสะท้อนที่ดับ การมองเห็นเศษโลหะได้ดียิ่งขึ้นจึงต้องมีการติดตั้งไฟแสงธรรมชาติเพิ่มเติมบริเวณข้างถุงเพื่อหลีกเลี่ยงการสะท้อนโดยตรง จากนั้นภาพจะถูกแสดงจอแสดงผลเพื่อให้สามารถตรวจสอบความถูกต้องได้ หากมีการตรวจพบเศษโลหะ YOLOv8 จะทำการติกรอบบริเวณที่มีเศษโลหะและส่งสัญญาณผ่าน USB ไปยังบอร์ด ESP32 เพื่อส่งเปิดไฟระบบอีกทีหนึ่ง อย่างไรก็ตามตัวชิประบบใช้ไฟกระแสตรง 24 โวลต์ จึงจำเป็นต้องส่งสัญญาณไปที่ชุดเสตดรีเลย์อีกทีหนึ่ง โดยมีเงื่อนไขว่าหากมีการพบเศษโลหะ 1 ชิ้น จะทำการเปิดไฟระบบ 3 วินาที แต่หากมีมากกว่านั้นระบบจะทำการกวาดเวลาเพิ่มไปเรื่อย ๆ เพื่อแก้ปัญหาการตรวจพบติดกันโดยที่ชิประบบยังไม่หยุดทำงาน

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] แสดงให้เห็นว่า YOLOv8 ถูกแบ่งออกเป็น 5 แบบหลัก

ปัจจุบัน YOLO เวอร์ชันล่าสุด คือ YOLOv11 อย่างไรก็ตามเนื่องจากกระบวนการปฏิบัติการบนบอร์ด Jetson Orin Nano ไม่รองรับการทำงานของ YOLOv11 บนการ์ดแสดงผล Nvidia ในปัจจุบันรองรับเพียงแค่ YOLOv8

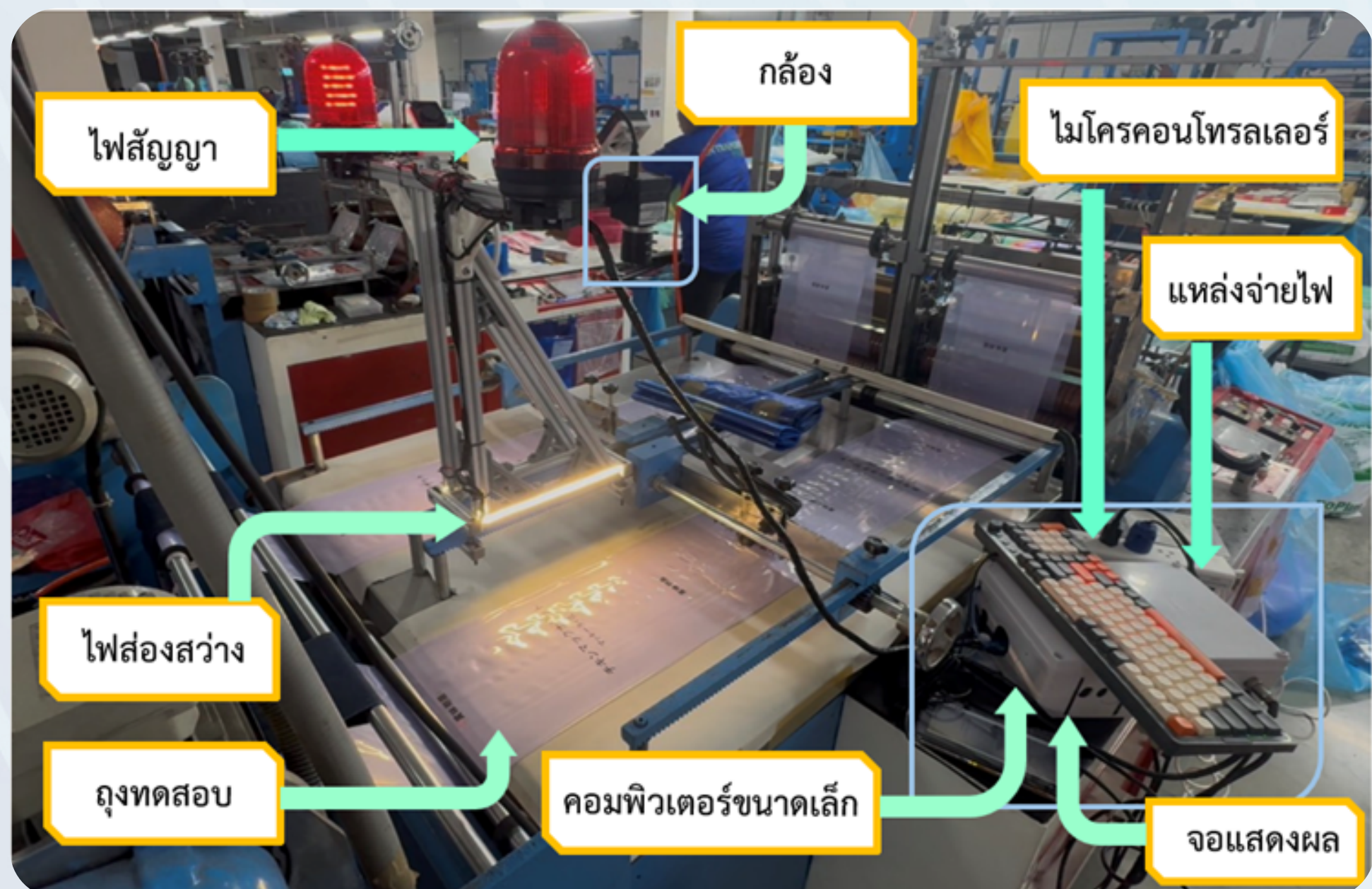
ภาพที่ใช้ในการฝึกปัญญาประดิษฐ์จะใช้ภาพที่ได้จากกล้อง Webcam ถูกลดความละเอียดลงเหลือ 640 x 640 pixel โดยภาพที่ใช้ในการฝึกมีรูปร่างและตำแหน่งของเศษโลหะที่แตกต่างกันโดยมีภาพเริ่มต้นอยู่ที่ 42 ภาพ จากนั้นภาพถูกทำให้เพิ่มขึ้นเป็น 210 ภาพ เพื่อใช้เป็นชุดข้อมูล (Dataset) สำหรับการฝึกปัญญาประดิษฐ์ โดยใช้กระบวนการ Augmentation หมุนภาพ 90 องศา (Rotate) เอียงภาพ (Shear) และเบลอภาพ (Blur) จากนั้นทำเลเบล (Label) ให้กับ Dataset ทั้งหมดถูกทำบนเว็บไซต์ Roboflow ตัวอย่างภาพที่ถูกทำ Label แล้ว ถูกแสดงดังรูปที่ 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

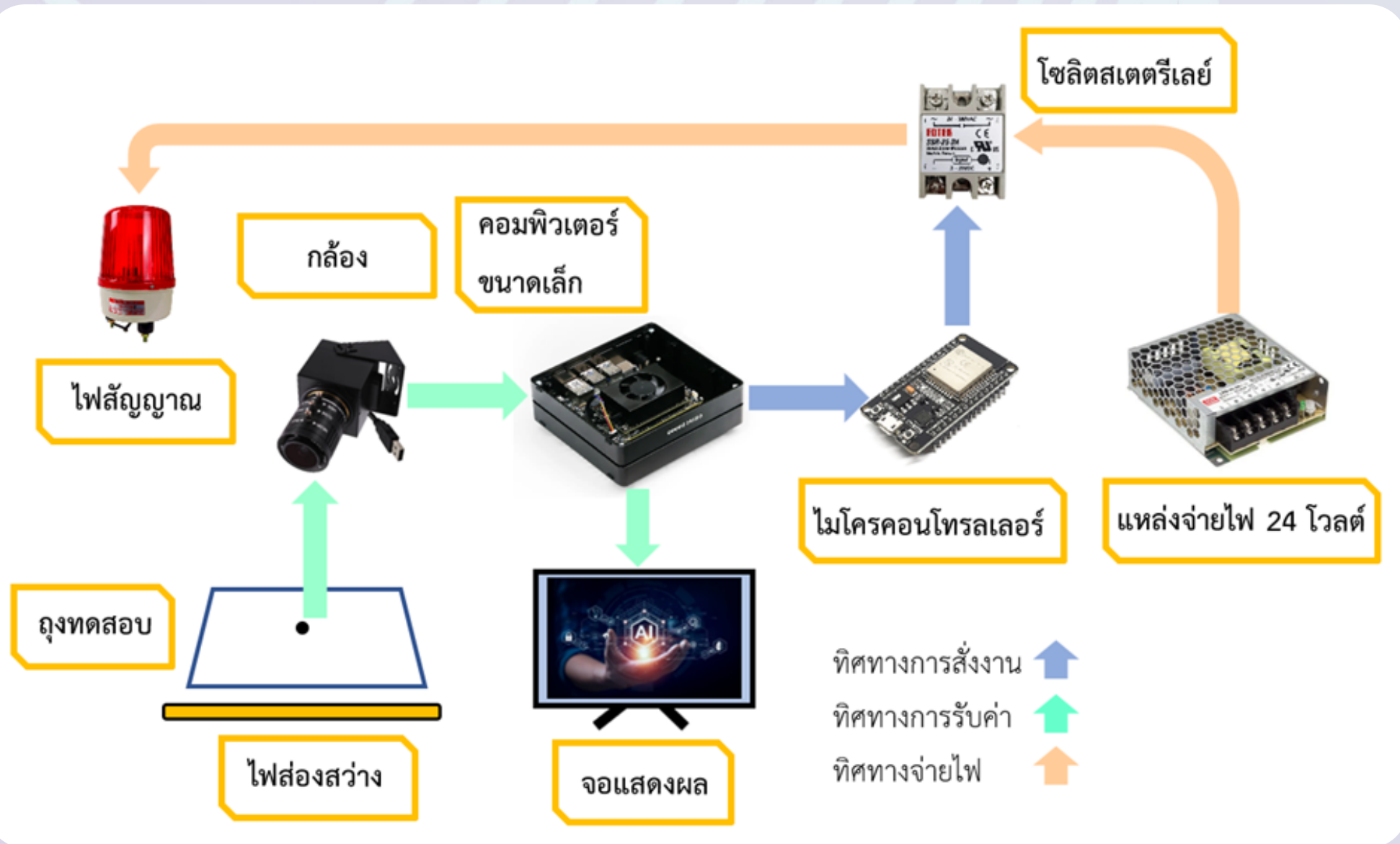
การฝึกปัญญาประดิษฐ์ งานวิจัยนี้ทำการฝึกบน Google Collab ซึ่งเป็นการประมวลผลแบบออนไลน์ โดยใช้การ์ดจอ NVIDIA Tesla T4 ใช้ Python เวอร์ชัน 3.12.11 ในการเขียนโปรแกรมฝึกและใช้ Pytorch เวอร์ชัน 2.8.0 เป็น Back Bone ส่วนโปรแกรมที่ใช้ความคุมการทำงานการดอเป็น cv2 และเวลาที่ใช้ในการฝึกข้อมูลทั้งหมด 0.82 ชั่วโมง หลังจากทำการฝึกจะได้แบบจำลอง (Model)

เอกสารอ้างอิง

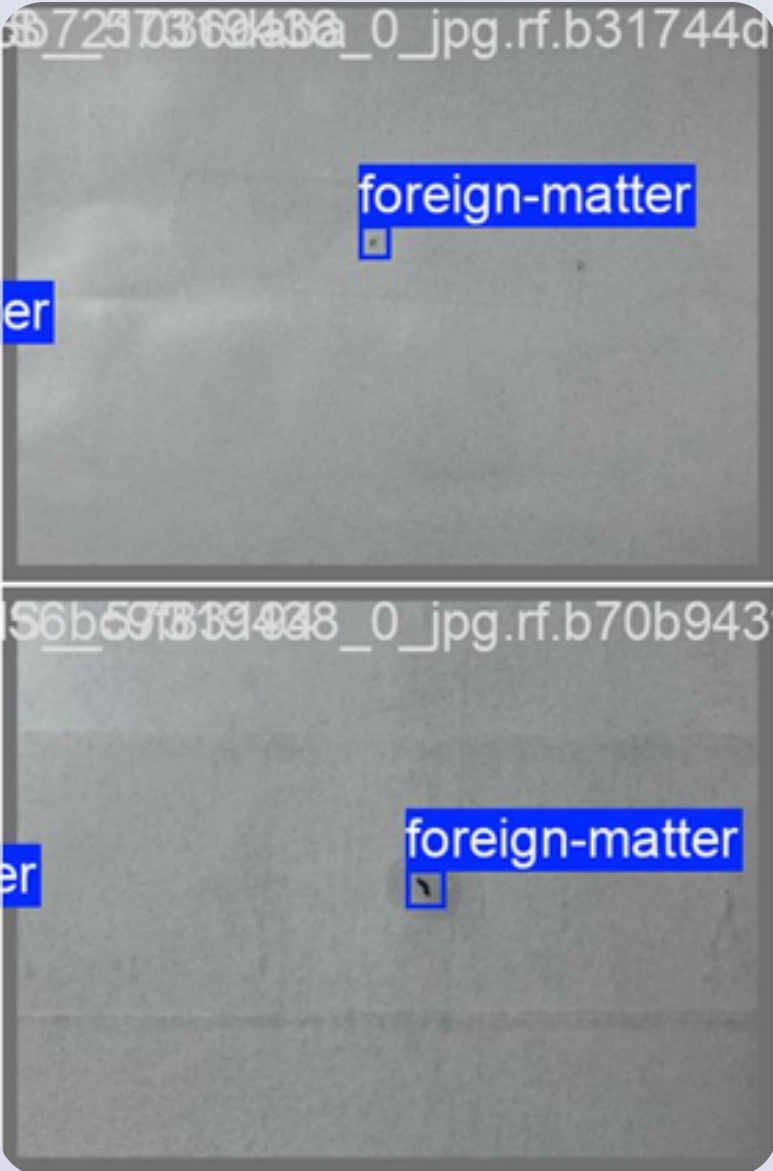
- [6] Kijdech, D., & Vongbunyong, S. (2024). Manipulation of a complex object using dual-arm robot with Mask R-CNN and grasping strategy. Journal of Intelligent & Robotic Systems, 110, 103. <https://doi.org/10.1007/s10846-024-02132-0>
- [7] Kijdech D. and Vongbunyong S. (2021). Artificial Intelligence in Localization and Classification of Cactus for Automatic Watering Works, The 8th 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 1st 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].



รูปที่ 1 ส่วนประกอบอุปกรณ์ที่ใช้ในการทดลอง



รูปที่ 2 การเชื่อมต่อของระบบและหลักการทำงาน

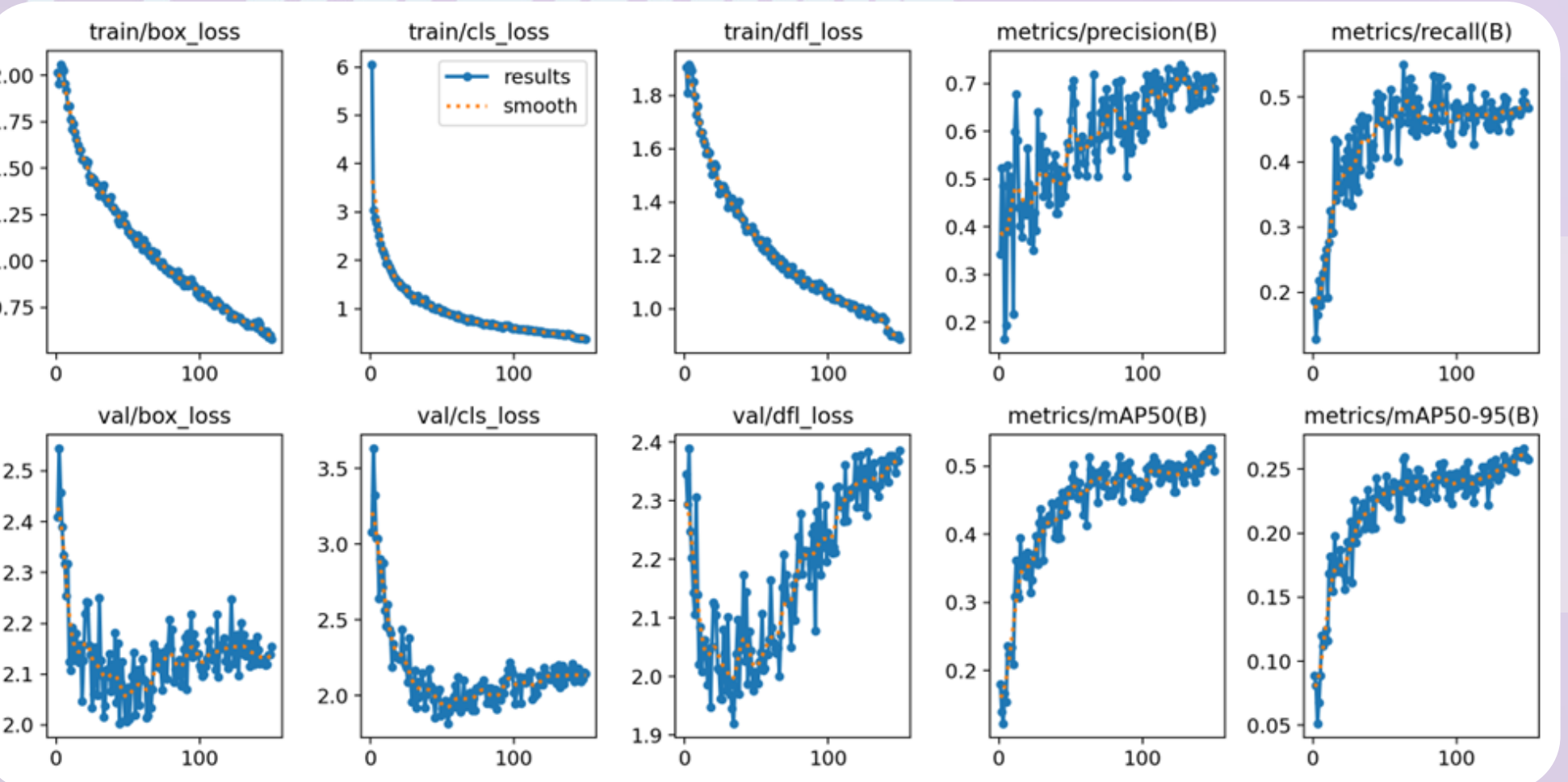


รูปที่ 4 แสดงตัวอย่าง Dataset

ผลการทดลอง

รูปที่ 5 แสดงให้เห็นสามารถเรียนรู้จากข้อมูลได้อย่างมีประสิทธิภาพ จากค่าความสูญเสียของข้อมูลฝึก (training loss) นั่นคือ box loss, classification loss และ distribution focal loss (DFL) มีแนวโน้มลดลงอย่างต่อเนื่องตลอดกระบวนการฝึกสอน แสดงให้เห็นว่าพารามิเตอร์ได้รับการปรับค่าอย่างเหมาะสมและแบบจำลองมีการลู่เข้า (convergence) ที่ดี อย่างไรก็ตาม ค่าความสูญเสียของข้อมูลประเมินผล (validation loss) พบว่าแม้ค่า box loss และ classification loss จะลดลงในช่วงแรก ยังคงมีแนวโน้มคงที่หรือเพิ่มขึ้นเล็กน้อยในช่วงท้าย validation DFL ยังมีแนวโน้มเพิ่มขึ้นอย่างชัดเจน แสดงให้เห็นว่าเริ่มมีการจดจำเลเบลแทนที่จะเรียนรู้ (overfitting) ในระดับน้อย ค่าความแม่นยำ (precision) มีแนวโน้มเพิ่มขึ้นจนมีค่าประมาณ 0.7 แสดงให้เห็นว่าแบบจำลองมีความแม่นยำในการทำนายวัตถุที่ตรวจพบค่อนข้างดี ในขณะที่ค่า recall อยู่ที่ประมาณ 0.5 ซึ่งบ่งชี้ว่ายังมีวัตถุบางส่วนที่แบบจำลองไม่สามารถตรวจจับได้อยู่ ส่งผลทำให้เกิดความไม่สมดุลระหว่าง precision และ recall กล่าวคือ ปัญญาประดิษฐ์เลือกทำนายเฉพาะกรณีที่มั่นใจสูงเท่านั้น 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 ส่วนรูปที่ 6 แสดงให้เห็นว่าแบบจำลองยังมีความมั่นใจต่ำอยู่ ค่าต่ำสุด 30% ค่าสูงสุด 70% โดยมีค่าเฉลี่ยอยู่ที่ 61.87%



รูปที่ 5 ผลการฝึกข้อมูลและประเมินผลข้อมูล



รูปที่ 6 แสดงผลการตรวจหาเศษโลหะ

สรุป

แบบจำลองมีความสามารถในการเรียนรู้และให้ผลลัพธ์ในระดับที่น่าพอใจ อย่างไรก็ตาม เนื่องจากทางโรงงานให้ข้อมูลตัวอย่างและงบประมาณค่อนข้างจำกัด ซึ่งหากไม่มีข้อจำกัดดังกล่าวข้างต้น จะทำให้สามารถปรับปรุงประสิทธิภาพได้ โดยการเพิ่มจำนวนข้อมูล การใช้เทคนิคการเพิ่มข้อมูล (data augmentation) และการปรับแต่งพารามิเตอร์ของโมเดล เพื่อเพิ่มค่า recall ลดปัญหา overfitting และยกระดับค่า mAP ให้สูงขึ้นได้ หากกำหนดค่าความมั่นใจที่ 60% เพื่อหลีกเลี่ยงการตรวจจับเศษโลหะผิดพลาด จะทำให้ตรวจจับภาพที่มีอยู่จริงไม่พบ ดังนั้นจึงสามารถตรวจจับได้ 81.25%