การตรวจจับการสวมหมวกนิรภัยของผู้ขับขี่รถจักรยานยนต์ ภายในพื้นที่โรงพยาบาลศิริราช ด้วยวิธีการเรียนรู้เชิงลึก

ธวัชชัย แซ่ลี้[1*], วราภรณ์ วิยานนท์[2]

บทคัดย่อ

ปัจจุบันการสัญจรภายในพื้นที่โรงพยาบาลศิริราชด้วยรถจักรยานยนต์ ยังคงพบการไม่สวมใส่หมวกนิรภัย ซึ่งอาจจะนำมาสู่การบาดเจ็บรุนแรงบริเวณศีรษะของผู้ขับขี่และผู้ซ้อนในกรณีที่เกิดอุบัติเหตุ ในอดีตการตรวจจับการสวมใส่หมวกนิรภัยนั้นจะทำด้วยวิธีการสังเกตการณ์ด้วยตัวบุคคล (Manual system) ทำให้เกิดความผิดพลาดขึ้นได้ (Human error) อันเนื่องมาจากสายตาและสมองของมนุษย์มีข้อจำกัดในการมองเห็นและการจดจำ ความผิดพลาดของการตรวจจับจะส่งผลให้ไม่เป็นไปตามนโยบายเรื่องกำหนดมาตรการให้คณะแพทยศาสตร์ศิริราชพยาบาล เป็นเขตสวมหมวกนิรภัย พ.ศ. 2563 งานนี้วิจัยนี้จึงได้นำเสนอการตรวจจับการสวมใส่หมวกนิรภัยของผู้ขับขี่และผู้นั่งซ้อนท้ายที่สัญจรภายในโรงพยาบาลศิริราช ด้วยการเรียนรู้จากชุดข้อมูลที่เก็บรวบรวมจากกล้องวงจรปิด (Custom data) ร่วมกับโมเดลการเรียนรู้เชิงลึก YOLOv8 ซึ่งเป็นโมเดลที่ใช้สำหรับงานตรวจจับวัตถุ (Objective Detection) เพื่อมาทดแทนวิธีการสังเกตการณ์ด้วยตัวบุคคล ทำให้ได้ผลลัพธ์คือโมเดล YOLOv8l และ YOLOv8x มีประสิทธิภาพสูงที่สุดจากทุกโมเดลโดย YOLOv8l มีค่าเฉลี่ยความแม่นยำ (Precision) อยู่ที่ 0.868, ค่าเฉลี่ยความถูกต้อง (Recall) อยู่ที่ 0.790, ค่าเฉลี่ย mAP50 อยู่ที่ 0.859 และค่าเฉลี่ย F1-Score อยู่ที่ 0.827 ส่วน YOLOv8x มีค่าเฉลี่ยความแม่นยำ (Precision) อยู่ที่ 0.842, ค่าเฉลี่ยความถูกต้อง (Recall) อยู่ที่ 0.811, ค่าเฉลี่ย mAP50 อยู่ที่ 0.858 และค่าเฉลี่ย F1-Score อยู่ที่ 0.826

คำสำคัญ : การตรวจจับหมวกนิรภัย, Object Detection, Deep learning, YOLOv8

---

[1] หลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิทยาการข้อมูล คณะวิทยาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ กรุงเทพฯ 10110

[2] คณะวิทยาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ กรุงเทพฯ 10110

* Corresponding author: Tel.: 085-3429451 E-mail address: thawatchai.sae@g.swu.ac.th

Helmet Detection for Motorcycle Riders and Passengers In Siriraj Hospital Area using Deep Learning

Thawatchai Saelee[1*], Waraporn Viyanon[2]

Abstract

Siriraj Hospital has a large number of patients, some of whom use motorcycles without helmets, leading to serious injuries when accidents occur. To address this issue, the hospital implemented a policy in 2020 requiring everyone in the Siriraj Faculty of Medicine to wear a helmet. However, visual inspection by staff has limitations and is prone to errors, making it difficult to enforce the policy. This research presents a system for detecting helmet use by riders and passengers within Siriraj Hospital. The system uses the YOLOv8 deep learning model for object detection, trained on a custom dataset collected from CCTV cameras. This approach outperforms manual observation and achieves high accuracy, with a precision of 0.868, recall of 0.790, mAP50 of 0.859, and F1-Score of 0.827 for YOLOv8l and a precision of 0.842, recall of 0.811, mAP50 of 0.858, and F1-Score of 0.826 for YOLOv8x The system has the potential to improve helmet compliance and reduce motorcycle-related injuries at Siriraj Hospital.

Keywords : Helmet detection, Object Detection, Deep Learning, YOLOv8

---

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

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

* Corresponding author: Tel.: 085-3429451 E-mail address: thawatchai.sae@g.swu.ac.th

บทนำ

คณะแพทยศาสตร์ศิริราชพยาบาล เป็นสถานที่ให้บริการทางการแพทย์และการศึกษา โดยมีบุคลากร นักศึกษา และประชาชนทั่วไป เข้ามาใช้บริการเป็นจำนวนมาก พบว่าปัญหาความปลอดภัยด้านการจราจรเป็นปัญหาที่สำคัญ ทุกปีมีผู้เสียชีวิตบนท้องถนนทั่วโลกเกือบ 1.2 ล้านคน ร้อยละ 70 ถึง 90 ของผู้เสียชีวิตในประเทศไทยนั้นเป็นผู้ขับขี่รถจักรยานยนต์ โดยการบาดเจ็บที่ศีรษะเป็นสาเหตุหลักของการเสียชีวิตและพิการ การสวมหมวกนิรภัยจะช่วยลดโอกาสการเสียชีวิตเนื่องมาจากการบาดเจ็บที่ศีรษะได้ 43% สำหรับผู้ขับขี่และ 58% สำหรับผู้นั่งซ้อนท้าย [1]

จากปัญหาดังกล่าว จึงได้มีการจัดตั้งโครงการศิริราชรณรงค์สวมหมวกนิรภัยขึ้น ซึ่งดำเนินการตั้งแต่ ปี พ.ศ.2562 จนถึงปัจจุบัน ทั้งนี้ได้มีการกำหนดมาตรการให้คณะแพทยศาสตร์ศิริราชพยาบาล เป็นเขตสวมหมวกนิรภัยและปฏิบัติตามกฎจราจร รวมถึงมีบทลงโทษในกรณีพบผู้ขับขี่หรือผู้นั่งซ้อนท้ายรถจักรยานยนต์ที่ไม่สวมหมวกนิรภัย และกำหนดให้คณะทำงานผู้รับผิดชอบนำนโยบายไปสู่การปฏิบัติจริง ปัญหาที่พบจากการปฏิบัติ คือ การตรวจจับการสวมหมวกนิรภัย โดยใช้ระบบสังเกตการณ์ด้วยตัวบุคคล (Manual System) นั้นทำให้เกิดความผิดพลาดในการตรวจจับได้ เนื่องมาจากการมองเห็นและจดจำของมนุษย์นั้นมีข้อจำกัด ทางผู้วิจัยจึงได้ศึกษากระบวนการที่นำเทคโนโลยีคอมพิวเตอร์ในการระบุและตรวจจับวัตถุ (Object Detection) โดยใช้การเรียนรู้เชิงลึก (Deep learning) มาพัฒนาการตรวจจับการสวมใส่หมวกนิรภัยของผู้ขับขี่และผู้นั่งซ้อนท้าย ทดแทนการสังเกตการณ์ด้วยตัวบุคคล เพื่อเป็นการลดข้อจำกัดและลดการใช้งานทรัพยากรบุคคล โดยผู้วิจัยได้เลือกใช้โมเดลการเรียนเชิงลึก YOLOv8 มาทำการฝึกกับชุดข้อมูลรูปภาพ จำนวน 3,000 รูป โดยทำการเก็บรวบรวมจากกล้องวงจรปิดทั้ง 4 บริเวณ ภายในโรงพยาบาลศิริราชเพื่อนำมาพัฒนาการตรวจจับการสวมหมวกนิรภัย

งานวิจัยที่เกี่ยวข้อง

ทางผู้วิจัยได้ทำการศึกษาเรื่องการนำเทคโนโลยีคอมพิวเตอร์ในการระบุและตรวจจับวัตถุ (Object Detection) มาประยุกต์ใช้ งานพบว่าในงานด้านการตรวจจับวัตถุนั้น สามารถแบ่งได้เป็น 2 ประเภท ได้แก่ One-stage detectors และ Two-stage detectors ซึ่งจะมีความแตกต่างกันอย่างชัดเจนในวิธีการทำงานและประสิทธิภาพของการตรวจจับ โดย One-stage detectors นั้นจะมีความเร็วและประสิทธิภาพในการตรวจจับสูง อันเนื่องมาจากการที่ไม่ต้องผ่านขั้นตอนของการสร้าง Region Proposals ก่อน ซึ่งช่วยลดความซับซ้อนของกระบวนการทำงานได้ และสามารถประมวลผลภาพได้ในเวลาเรียลไทม์ ส่วน Two-stage detectors มีความแม่นยำสูงกว่าในการตรวจจับวัตถุที่มีขนาดเล็กหรือตำแหน่งที่ซับซ้อน แต่ใช้ระยะเวลาในการทำงานที่นานกว่าแบบ One-stage detectors และมีขั้นตอนการทำงานที่ซับซ้อนโดยโมเดลในกลุ่ม One-stage detectors ตัวอย่างเช่น Yolo (You Only Look Once), SSD และโมเดลในกลุ่ม Two-stage detectors ตัวอย่างเช่น R-CNN, Fast-RCNN, Faster-RCNN [2] จากความต้องการและลักษณะของงานตรวจจับการสวมใส่หมวกนิรภัยที่ต้องการตรวจจับวัตถุอย่างรวดเร็วนั้น ทางผู้วิจัยจึงได้เลือกใช้การตรวจจับแบบ One-stage detectors และในการวัดประสิทธิภาพของโมเดลต่างๆ ในงานตรวจจับวัตถุนั้นสามารถทำ

ได้โดยการวัด Frames Per Second (FPS) หรือจำนวนเฟรมที่โมเดลสามารถประมวลผลได้ในหนึ่งวินาที จำนวน FPS ที่สูงมักจะแสดงถึงความเร็วในการทำงานของโมเดล

Model	Frame Per Second (FPS)
YOLOv8	40 - 155
SSD	22 - 46
Faster R-CNN	5 - 7

ตารางที่ 1 แสดงผลการเปรียบเทียบ Frame Per Second (FPS)

จากตารางที่ 1 Yolov8 จะมีความเร็วในการประมวลผลจำนวนเฟรมในหนึ่งวินาที ที่สูงกว่า SDD และ Faster R-CNN จากความเร็วในการประมวล FPS [3] นี้ สามารถทำให้ได้ผลลัพธ์ในเวลาเรียลไทม์หรือใกล้เคียงและมีประสิทธิภาพในการทำงานที่เร็ว ดังนั้นทางผู้วิจัยจึงได้เลือกใช้โมเดล YOLO (You Only Look Once)

ผู้วิจัยทำการทบทวนวรรณกรรมที่เกี่ยวข้องกับการตรวจจับวัตถุ โดยในประเทศจีนศึกษาการพัฒนาระบบตรวจจับหมวกนิรภัยของผู้ขับขี่รถจักรยานยนต์ โดยใช้อัลกอริทึม YOLOv5 มีการเปรียบเทียบโมเดลแบบ Two-stage ได้แก่ Faster R-CNN, Cascade R-CNN และ Libra R-CNN กับแบบ One-Stage ได้แก่ SSD, RetinaNet และ FCOS ซึ่งผลลัพธ์ที่ได้ โมเดล YOLOv5 ที่เป็นแบบ One-Stage มีค่า mAP มากที่สุดเมื่อเปรียบเทียบกับทุกโมเดล โดยมีค่า mAP อยู่ที่ 97.7% และ F1-Score อยู่ที่ 92.7% [4] ต่อมาในประเทศพม่า ก็ได้ทำการศึกษาพัฒนาระบบตรวจจับหมวกนิรภัยเช่นกัน แต่จะนำวิธี CNN-base multi-task (MTL) มาใช้ โดยงานวิจัยนี้ได้เลือกใช้โมเดล YOLOv2 และ RetinaNet มาทำการเปรียบเทียบประสิทธิภาพ โดยผลลัพธ์ที่ได้คือ โมเดล YOLOv2 มีค่า F1-Score อยู่ที่ 60.0% และโมเดล RetinaNet มีค่า F1-Score อยู่ที่ 67.3% [5] จะสังเกตได้ว่า จากผลลัพธ์ที่ได้เปรียบเทียบระหว่างการใช้ YOLOv5 ในงานวิจัยในประเทศจีนและการใช้ YOLOv2 และ RetinaNet ในงานวิจัยในประเทศพม่า จะเห็นได้ว่า YOLOv5 มีประสิทธิภาพสูงกว่าทั้ง YOLOv2 และ RetinaNet ในการตรวจจับการสวมหมวกนิรภัยของผู้ขับขี่รถจักรยานยนต์อันเนื่องมาจากโมเดล YOLOv5 นั้นได้รับการพัฒนาปรับปรุงมาจากเวอร์ชั่นก่อนหน้า

หลังจากนั้นทางผู้วิจัยได้ทบทวนวรรณกรรมที่เกี่ยวข้องกับการนำโมเดล YOLO ไปประยุกต์ใช้ร่วมกับ CNN โดยมีงานวิจัย [6] ได้นำ YOLO และ CNN มาทำงานร่วมกันในการตรวจจับหมวกนิรภัยจากชุดข้อมูลภาพ ชุดข้อมูลรูปภาพที่ใช้ในการฝึกมีทั้งหมด 30,000 รูปภาพ โดยแบ่งเป็น 13,000 รูปที่เป็นรูปภาพผู้ขับขี่รถจักรยานยนต์ที่ไม่สวมหมวกนิรภัย และ 17,000 รูปที่เป็นรูปภาพผู้ขับขี่จักรยานยนต์ที่สวมหมวกนิรภัย โดยนำชุดข้อมูลมาทำการฝึกกับ CNN เพื่อใช้ในการแยกคุณลักษณะที่จำเป็นจากรูปภาพ หลังนั้นนำคุณลักษณะที่ได้ไปทำการตรวจจับวัตถุและจำแนกโดยใช้ YOLO ได้ผลลัพธ์ที่มีความแม่นยำอยู่ที่ 94.29% และงานวิจัย [7] ก็ได้เลือกใช้ YOLO และ CNN มาทำงานร่วมกัน แต่จะแตกต่างกันในส่วนของการประยุกต์ใช้งาน โดย YOLO จะถูกนำมาใช้ในการตรวจจับวัตถุในรูปภาพ ได้แก่ ผู้ขับขี่รถจักรยานยนต์, ผู้ซ้อนรถจักรยานยนต์ และรถจักรยานยนต์ ส่วน CNN จะใช้ในการตรวจจับการสวมใส่หมวกนิรภัยของผู้ขับขี่รถจักรยานยนต์และผู้นั่งซ้อนท้าย โดยผลลัพธ์ในการตรวจจับผู้ขับขี่รถจักรยานยนต์ ผู้นั่งซ้อนท้ายรถจักรยานยนต์ และรถจักรยานยนต์ นั้นมีความแม่นยำอยู่ที่ 91.08% และในตรวจจับการสวมหมวกนิรภัยมีความแม่นยำอยู่ที่ 96.23% จะสังเกตได้ว่าทั้งงานวิจัยที่ [6] และ [7] นำเสนอโมเดล YOLO เป็นวิธีการหลักในการตรวจจับวัตถุในรูปภาพ

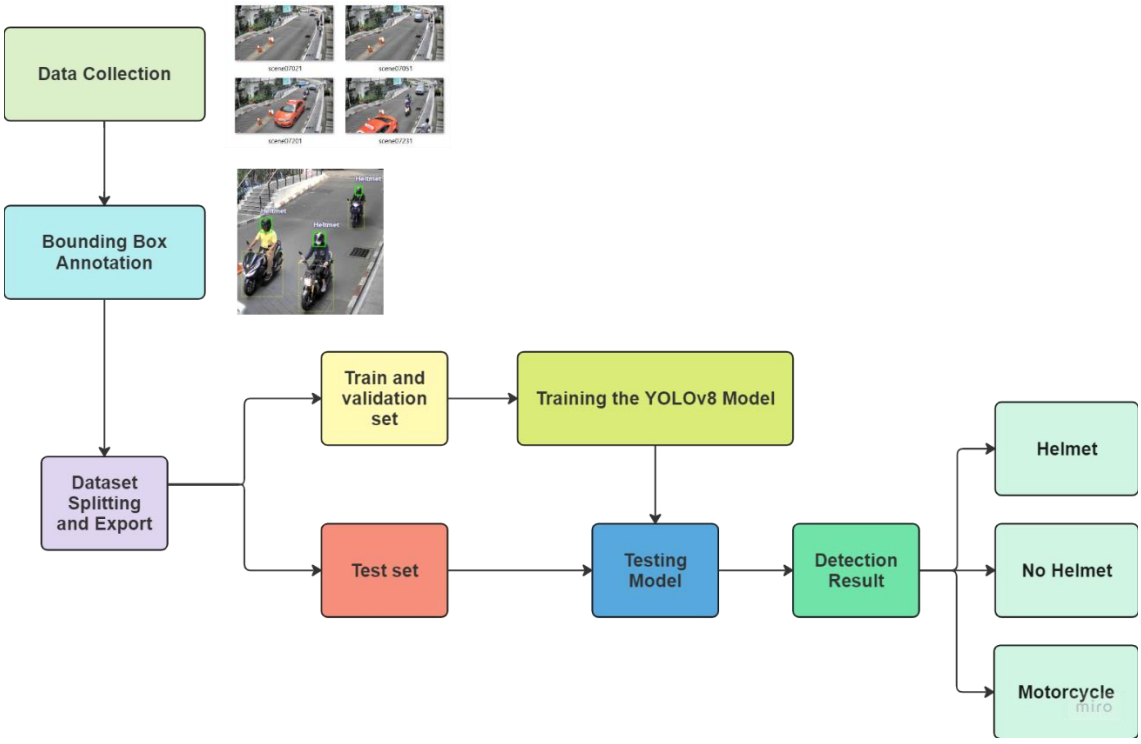
และใช้ CNN เพื่อประมวลผลเพิ่มเติมหรือการแยกแยะคุณลักษณะของวัตถุก่อนที่ YOLO จะทำการตรวจจับ นั้นมีประสิทธิภาพในการตรวจจับที่ค่อนข้างสูง

ผู้วิจัยมีความสนใจที่จะใช้ Google Colab เป็นเครื่องมือในการพัฒนาการตรวจจับการสวมหมวกนิรภัย จึงได้ทบทวนวรรณกรรมที่เกี่ยวข้องกับการใช้ Google Colab ในการพัฒนา พบว่างานวิจัย [8] ได้ใช้ Google Colab เป็นเครื่องมือในการพัฒนาการตรวจจับการสวมหมวกนิรภัยกับชุดข้อมูลรูปภาพที่ใช้ในการฝึกโมเดลมาจากเว็บไซต์โอเพนซอร์ส มีจำนวนทั้งหมด 2,121 รูปภาพ ทำการฝึกกับ 3 โมเดล ได้แก่ YOLOv3, SSD300 และ Caffe Model โดยผลลัพธ์พบว่าโมเดล YOLOv3 มีความแม่นยำมากที่สุดจากทั้ง 3 โมเดล โดยมีความแม่นยำอยู่ที่ 90.13%

จากการทบทวนวรรณกรรมที่เกี่ยวข้องการตรวจจับวัตถุและความเหมาะสมกับการตรวจจับการสวมหมวกนิรภัยนั้น ทางผู้วิจัยจึงได้เลือกโมเดล Yolo (You Only Look Once) ที่เป็นการตรวจจับวัตถุแบบ One-stage detectors มาเป็นโมเดลในการฝึกกับชุดข้อมูลที่ผู้วิจัยได้ทำการเก็บรวบรวมข้อมูลมา

วิธีดำเนินการ

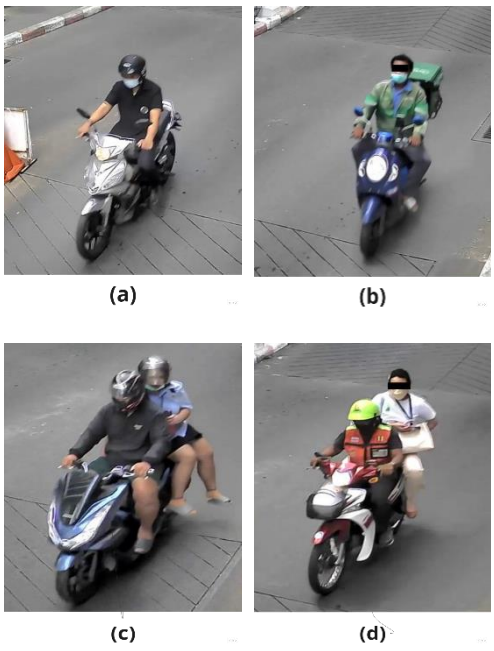
ในส่วนนี้ผู้วิจัยได้ออกแบบแผนผังขั้นตอนในการพัฒนาการตรวจจับการสวมหมวกนิรภัย จากรูปที่ 1



รูปที่ 1 แสดงแผนผังการพัฒนาการตรวจจับการสวมหมวกนิรภัย

ขั้นตอนที่ 1 : Data Collection

ทำการดาวห์โหลดไฟล์ข้อมูลจากกล้องวงจรปิดทั้ง 4 บริเวณ ได้แก่ ประตู 1, ประตู 5, ประตู 7 และประตู 8 ช่วงเดือน พ.ย. 2566 - ม.ค. 2567 ในช่วงเวลา 7.00 น. ถึง 9.00 น. และ 16.00 ถึง 17.00 น. ซึ่งไฟล์ข้อมูลจากกล้องวงจรจะอยู่ในรูปแบบของไฟล์วิดีโอ ซึ่งไม่สามารถนำไปฝึกโมเดลได้จะต้องทำการแปลงไฟล์วิดีโอให้เป็นไฟล์ภาพก่อนจึงจะสามารถนำไปใช้ฝึกโมเดลได้ โดยผู้วิจัยได้เลือกใช้โปรแกรม VLC ในการแปลงไฟล์วิดีโอเป็นไฟล์ภาพที่มีขนาด 2560 x 1440 เมื่อได้ข้อมูลไฟล์ภาพแล้ว ทำการคัดเลือกภาพที่สามารถนำไปใช้ในการฝึกโมเดล เพราะเนื่องจากข้อมูลภาพที่ได้มาจากการแปลงข้อมูลทั้งหมดนั้นจะมีทั้งภาพที่สามารถใช้ในการฝึกโมเดลได้และภาพที่ไม่สามารถใช้การฝึกโมเดลได้ โดยผู้วิจัยมีวิธีการในการคัดเลือกภาพ ดังนี้ ข้อมูลรูปภาพต้องมีวัตถุที่ผู้วิจัยต้องการจะตรวจจับ ได้แก่ รถจักรยานยนต์ การสวมหมวกนิรภัย และการไม่สวมหมวกนิรภัย, ข้อมูลรูปภาพสามารถระบุผู้ขับขี่รถจักรยานยนต์ได้ว่าสวมหมวกนิรภัยหรือไม่สวมหมวกนิรภัย, ข้อมูลรูปภาพสามารถระบุผู้ซ้อนท้ายรถจักรยานยนต์ได้ว่าสวมหมวกนิรภัยหรือไม่สวมหมวกนิรภัย, ข้อมูลรูปภาพในแต่ละรูปจะไม่มีวัตถุที่ผู้วิจัยต้องการตรวจจับที่ซ้ำกัน เพื่อป้องกันการรั่วไหลของข้อมูล (Data Leakage) ในขั้นตอนของการแบ่งชุดข้อมูล (Data Splitting) ลักษณะของภาพที่สามารถใช้ในการฝึกโมเดล จะมีลักษณะเป็นภาพที่มีผู้ขับขี่รถจักรยานยต์และผู้นั่งซ้อยท้ายที่สวมหมวกนิรภัยหรือไม่สวมหมวกนิรภัย ส่วนภาพที่ไม่สามารถใช้ในการฝึกโมเดลจะเป็นภาพที่ไม่มีวัตถุที่ผู้วิจัยต้องการจะตรวจจับ เช่น ภาพของถนนที่ไม่มีการสัญจร, ภาพของถนนที่มีแต่รถยนต์สัญจร เป็นต้น ในที่นี้ผู้วิจัยได้ข้อมูลภาพที่ใช้ในการฝึกและทดสอบโมเดลมาทั้งสิ้นจำนวน 3,000 รูป ดังรูปที่ 2

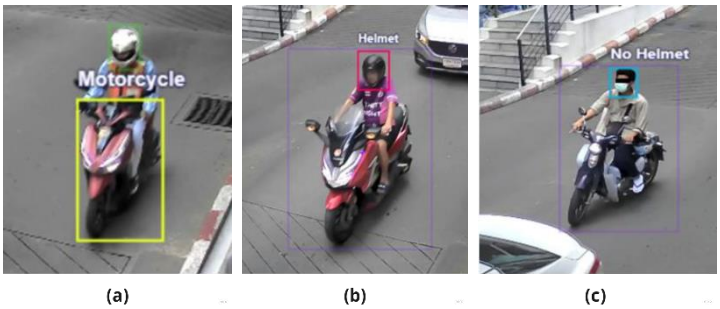


รูปที่ 2 แสดงตัวอย่างข้อมูลภาพที่ใช้ในการฝึกและทดสอบโมเดล

(a) แสดงภาพผู้ขับขี่รถจักรยานยนต์ที่สวมหมวกนิรภัย, (b) แสดงภาพผู้ขับขี่รถจักรยานยนต์ที่ไม่สวมหมวกนิรภัย, (c) แสดงภาพผู้ขับขี่และผู้นั่งซ้อนท้ายรถจักรยานยนต์ที่สวมหมวกนิรภัย, (d) แสดงภาพผู้ขับขี่รถจักรยานยนต์ที่สวมหมวกนิรภัยและผู้นั่งซ้อนท้ายที่ไม่สวมหมวกนิรภัย

ขั้นตอนที่ 2 : Bounding Box Annotation

หลังจากที่ได้ไฟล์ข้อมูลรูปภาพที่ผ่านการคัดกรองแล้วทำการสร้าง Bounding Box ผ่านแพลตฟอร์ม Roboflow [9] ซึ่งเป็นขั้นตอนที่สำคัญสำหรับการเรียนรู้แบบมีผู้สอนโดยกำหนดคลาสทั้งหมด 3 คลาส ได้แก่ คลาส Motorcycle สำหรับภาพที่มีผู้ขับขี่รถจักรยานยนต์, คลาส Helmet สำหรับภาพที่มีการสวมหมวกนิรภัยของทั้งผู้ขับขี่รถจักรยานยนต์และผู้นั่งซ้อนท้ายและคลาส No Helmet สำหรับภาพที่ไม่มีการสวมหมวกนิรภัยของทั้งผู้ขับขี่รถจักรยานยนต์และผู้นั่งซ้อนท้าย ดังรูปที่ 3

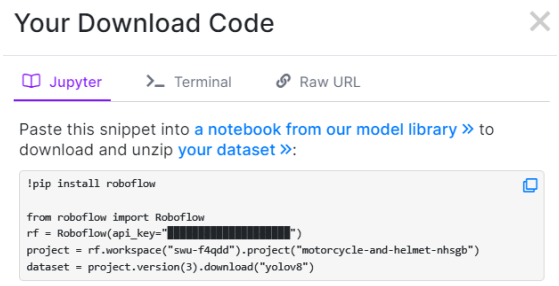


รูปที่ 3 แสดงตัวอย่าง Bounding Box ทั้ง 3 คลาส

(a) แสดงภาพ Bounding Box ของคลาส Motorcycle สำหรับภาพที่มีผู้ขับขี่รถจักรยานยนต์, (b) แสดงภาพ Bounding Box ของคลาส Helmet สำหรับภาพที่มีการสวมใส่หมวกนิรภัย และ (c) แสดงภาพ Bounding Box ของคลาส No Helmet สำหรับภาพที่ไม่มีการสวมใส่หมวกนิรภัย

ขั้นตอนที่ 3 : Dataset Splitting and Export

หลังจากกำหนด Bounding Box ให้กับข้อมูลรูปภาพทั้งหมดแล้ว ทางผู้วิจัยได้ทำการแบ่งชุดข้อมูล (Dataset Split) ให้ได้ชุดข้อมูลสำหรับใช้ในการฝึกและทดสอบโมเดล ในอัตราส่วน 70% สำหรับฝึกโมเดลและ 30% สำหรับทดสอบโมเดล โดยรูปภาพที่ใช้สำหรับการฝึกโมเดลมีจำนวน 2,105 รูป, รูปภาพที่ใช้ในการตรวจสอบโมเดลมีจำนวน 600 รูป และรูปภาพที่ใช้ในการทดสอบโมเดลมีจำนวน 295 รูป หลังจากนั้นทำการ Export ชุดข้อมูลโดยทางผู้วิจัยได้กำหนดรูปแบบให้เหมาะสมสำหรับโมเดลโดยการเลือกรูปแบบที่เป็น YOLOv8 และทำการ Export ชุดข้อมูลออกมาเป็นลักษณะ API KEY เพื่อเตรียมไว้ใช้ในขั้นตอนฝึกโมเดลใน Google Colab ต่อไป ดังรูปที่ 4



รูปที่ 4 แสดงตัวอย่าง API KEY ของชุดข้อมูลที่เตรียมไว้ใช้ในขั้นตอนการฝึกโมเดล

ขั้นตอนที่ 4 : Training the YOLOv8 model

โดยก่อนทำการทดลอง ผู้วิจัยได้ศึกษาข้อมูลของ Pretrained Detect models ของ YOLOv8 จากเว็บไซต์ Ultralytics [10] ดังตารางที่ 2 โดย Pretrained Detect models จะประกอบไปด้วย YOLOv8n, YOLOv8s, YOLOv8m, YOLOv8l และ YOLOv8x โดทุกรุ่นจะมีความแตกต่างกันในด้านของประสิทธิภาพ, ความเร็วในการประมวลผล, จำนวนพารามิเตอร์ ขึ้นอยู่กับลักษณะงานและความต้องการของผู้เลือกใช้

Model	Size (pixels)	mAP^{val} 50-95	Speed CPU ONNX (ms)	Speed CPU ONNX (ms)	Params (M)	FLOPs (B)
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

ตารางที่ 2 แสดงตัวอย่างประสิทธิภาพ, ความเร็วในการประมวลผล, จำนวนพารามิเตอร์ ที่แตกต่างกันของ YOLOv8

หลังจากนั้นทางผู้วิจัยได้ทดลองเลือกใช้ Pretrained Detect models ได้แก่ YOLOv8n, YOLOv8s, YOLOv8, YOLOv8l และ YOLOv8x ร่วมกับ Epoch จำนวน 100 รอบ เพื่อดูความแตกต่างในการฝึกโมเดล พบว่าระยะเวลาในการฝึกโมเดลของทุกรุ่นนั้นมีความแตกต่างกันดังตารางที่ 3

Model	Training time (hours)	Epoch
YOLOv8n	0.471	100
YOLOv8s	0.541	100
YOLOv8m	1.024	100
YOLOv8l	1.201	100
YOLOv8x	2.470	100

ตารางที่ 3 แสดงระยะเวลาการฝึก Pretrained Detect models แบบต่างๆ

ผู้วิจัยได้กำหนดการตั้งค่าต่างๆ ได้แก่ Task, Mode, Model, Data, Epochs และ Imgsz ผู้วิจัยจึงได้กำหนกการตั้งค่าดังนี้ Task ผู้วิจัยกำหนดเป็น Detect เพื่อใช้สำหรับงานที่ต้องการระบุตำแหน่งของวัตถุหรือพื้นที่ที่น่าสนใจในภาพหรือวิดีโอ, Mode ผู้วิจัยกำหนดเป็น Train เพื่อใช้สำหรับการฝึกแบบโมเดล YOLOv8 ร่วมกับชุดข้อมูลที่กำหนดเอง, Model ผู้วิจัยกำหนดเป็น yolo8n.pt, yolo8s.pt, yolo8m.pt, yolo8l.pt และ yolo8x.pt, Data ผู้วิจัยกำหนดเป็นตำแหน่งของชุดข้อมูลที่ใช้ในการฝึกโมเดลซึ่งจะถูกระบุไว้ในไฟล์ data.yaml, Epoch ผู้วิจัยกำหนดเป็น 100 รอบที่จะทำการฝึก, Imgsz ผู้วิจัยกำหนดเป็น 640 ซึ่งเป็นขนาดของภาพที่ใช้ในการฝึก

ขั้นตอนที่ 5 : Testing model

เมื่อทำการฝึกโมเดล YOLOv8 ทุกรุ่นเสร็จสิ้น ก็จะได้ไฟล์โมเดลที่มีชื่อไฟล์ว่า best.pt สำหรับโมเดลในแต่ละรุ่น โดยไฟล์นี้จะเป็นไฟล์โมเดลที่ได้ทำการฝึกกับชุดข้อมูลรูปภาพที่ได้เก็บรวบรวมมา ซึ่งสามารถใช้เป็น Testing Model ในการตรวจจับกับชุดข้อมูลทดสอบได้

ผลการวิจัยและอภิปรายผลการวิจัย

ในส่วนนี้ผู้วิจัยจะนำเสนอรายละเอียดของผลลัพธ์และอภิปรายประเด็นจากผลลัพธ์ที่น่าสนใจ โดยผู้วิจัยเลือกตัวชี้วัดที่ใช้ในการประเมินผลลัพธ์ ได้แก่ Precision, Recall, mAP50, F1-Score และผลลัพธ์รูปภาพจากชุดข้อมูลทดสอบ

รายละเอียดการทดลอง

ในการทดลองหลังจากที่สร้างกรอบวัตถุ (Bouding box) ให้กับชุดข้อมูลแล้วนั้น ในขั้นตอนของการทำ Preprocessing ผู้วิจัยได้ทำการปรับเปลี่ยนขนาดของภาพ (Resize) ทุกภาพในชุดข้อมูลจาก 2560*1440 ให้มีขนาด 640*640 โดยขนาดภาพ 640*640 นั้นจะเป็นขนาดที่เหมาะสมสำหรับโมเดลในการประมวลผล เพราะถ้าขนาดของภาพใหญ่เกินไปอาจทำให้การตรวจจับวัตถุไม่แม่นยำ เนื่องจากโมเดลอาจไม่สามารถจดจำรายละเอียดของวัตถุได้ดีพอในภาพที่มีขนาดใหญ่ เครื่องมือที่ใช้ในการฝึกและทดสอบโมเดลนั้นทางผู้วิจัยเลือกใช้เป็น Google Colab ที่เลือก Runtime type เป็น Python 3 และกำหนด Hardware accelerator เป็น NVIDIA Tesla V100 SXM2 16 GB และทำการการตั้งค่าพารามิเตอร์ที่สำคัญในการฝึกโมเดลตามตารางที่ 4

Parameter	Value
imgsz	640
batch	16
epochs	100
optimizer	AdamW
Learning rate	0.01

ตารางที่ 4 ตั้งค่าพารามิเตอร์ที่สำคัญในการฝึกโมเดล

ผลการทดลอง

จากการที่ได้ทดลองฝึกโมเดล YOLOv8n, YOLOv8s, YOLOv8m, YOLOv8l และ YOLOv8x ร่วมกับชุดข้อมูลที่ได้เก็บรวบรวมมาเองนั้น ผู้วิจัยได้ประเมินผลลัพธ์โดยรวม (ALL) ของโมเดลแต่ละรุ่นได้ผลลัพธ์ตามตารางที่ 5 จะพบว่าระยะเวลาในการฝึกโมเดลนั้น YOLOv8x ใช้ระยะเวลาในการฝึกนานที่สุดเมื่อเปรียบเทียบกับโมเดลในรุ่นอื่นๆ ซึ่งเหตุผลมาจากจำนวนพารามิเตอร์ของโมเดล YOLOv8x นั้นมีจำนวนพารามิเตอร์ที่มากที่สุด ส่วน YOLOv8n และ YOLOv8s ใช้ระยะเวลาในการฝึกที่น้อยเมื่อเปรียบเทียบกับโมเดลในรุ่นอื่นๆ ในส่วนของค่าความแม่นยำ (Precision) ของ YOLOv8l มีความแม่นยำสูงที่สุดอยู่ที่ 0.868, ค่าความถูกต้อง (Recall) ของ YOLOv8x มีความถูกต้องสูงที่สุดอยู่ที่ 0.811 และค่าเฉลี่ยความแม่นยำ (mAP50) ของ YOLOv8l และ YOLOv8x มีค่าที่ใกล้เคียงกัน จะเห็นได้ว่า YOLOv8l มีความแม่นยำและประสิทธิภาพในการตรวจจับที่ครอบคลุมสูง แต่ใช้ระยะเวลาในการฝึกที่น้อยกว่า YOLOv8x ส่วน YOLOv8m เป็นตัวเลือกที่ดีในเรื่องของความสมดุลระหว่างระยะเวลาในฝึกและประสิทธิภาพของโมเดล โดยมีค่า Recall, mAP50 และ F1-Score ที่ใกล้เคียงกับ YOLOv8l แต่มีใช้ระยะเวลาฝึกที่น้อยกว่า

Model	Training time (hours)	Precision	Recall	mAP50	F1-Score
YOLOv8n	0.471	0.799	0.751	0.803	0.774
YOLOv8s	0.541	0.837	0.783	0.839	0.809
YOLOv8m	1.024	0.855	0.791	0.854	0.822
YOLOv8l	1.201	0.868	0.790	0.859	0.827
YOLOv8x	2.470	0.842	0.811	0.858	0.826

ตารางที่ 5 ประเมินผลลัพธ์โดยรวมของโมเดลแต่ละรุ่น

ผลลัพธ์ในคลาสสวมหมวกนิรภัย (Helmet) ในตารางที่ 6 จะพบว่า ค่าความแม่นยำ (Precision) ของ YOLOv8l มีความแม่นยำสูงที่สุด อยู่ที่ 0.855 ส่วนค่าความถูกต้อง (Recall) ของ YOLOv8x มีความถูกต้องสูงที่สุด อยู่ที่ 0.832 และค่า mAP50 ของทั้ง YOLOv8l และ YOLOv8x นั้นมีค่าสูงที่สุดและมีค่าเท่ากันอยู่ที่ 0.870 จากการวิเคราะห์ผลลัพธ์ในคลาสสวมหมวกนิรภัย ผู้วิจัยพบว่าถ้าหากต้องการความแม่นยำของการตรวจจับการสวมหมวกนิรภัยที่สูง สามารถเลือกใช้โมเดล YOLO ได้ทุกรุ่นยกเว้น YOLOv8n เนื่องจากค่าความแม่นยำ (Precision) ของ YOLOv8n นั้นมีค่าไม่สูงมากเมื่อเทียบกับรุ่นอื่น แต่ถ้าต้องการความถูกต้องในการตรวจจับการสวมหมวกนิรภัย (Recall) สามารถเลือกใช้โมเดล YOLOv8l และ YOLOv8x ได้เพราะมีค่าสูงถึง 80% ขึ้นไป

Model	Class	Precision	Recall	mAP50
YOLOv8n	Helmet	0.772	0.743	0.802
YOLOv8s	Helmet	0.816	0.798	0.839
YOLOv8m	Helmet	0.844	0.789	0.865
YOLOv8l	Helmet	0.855	0.809	0.870
YOLOv8x	Helmet	0.829	0.832	0.870

ตารางที่ 6 ผลลัพธ์คลาสสวมหมวกนิรภัย (Helmet) ของโมเดลแต่ละรุ่น

ผลลัพธ์ในคลาสรถจักรยานยนต์ ในตารางที่ 7 จะพบว่า ค่าความแม่นยำ (Precision) ของ YOLOv8s มีความแม่นยำสูงที่สุด อยู่ที่ 0.899, ค่าความถูกต้อง (Recall) ของทั้ง YOLOv8s, YOLOv8m และ YOLOv8x มีความถูกต้องที่เท่ากันและมีค่าสูงที่สุด อยู่ที่ 0.940 และค่า mAP50 ของ YOLOv8l มีค่าสูงที่สุด อยู่ที่ 0.967 จากการวิเคราะห์ผลลัพธ์ในคลาสรถจักรยานยนต์ ค่าความถูกต้อง (Recall) และค่า mAP50 ของทุกโมเดลทุ่นรุ่นนั้นมีค่าที่สูง อันเนื่องมาจากโมเดลทุกรุ่นที่ผู้วิจัยนำมาทดลองนั้นเป็น Pretrained detect models ที่ได้ผ่านการฝึกกับชุดข้อมูล COCO ซึ่งภายในชุดข้อมูลมีคลาสรถจักรยานยนต์ด้วยเช่นกัน

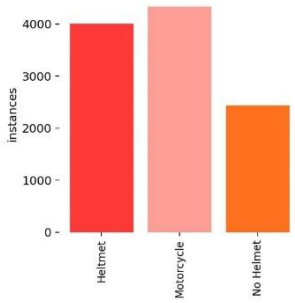
Model	Class	Precision	Recall	mAP50
YOLOv8n	Motorcycle	0.878	0.937	0.959
YOLOv8s	Motorcycle	0.899	0.940	0.962
YOLOv8m	Motorcycle	0.883	0.940	0.966
YOLOv8l	Motorcycle	0.893	0.925	0.967
YOLOv8x	Motorcycle	0.874	0.940	0.965

ตารางที่ 7 ผลลัพธ์คลาสรถจักรยานตร์ (Motorcycle) ของโมเดลแต่ละรุ่น

ผลลัพธ์ในคลาสไม่สวมหมวกนิรภัย (No Helmet) ในตารางที่ 8 จะพบว่า ค่าความแม่นยำ (Precision) ของ YOLOv8l มีความแม่นยำสูงที่สุด อยู่ที่ 0.856, ค่าความถูกต้อง (Recall) ของ YOLOv8x มีความถูกต้องที่สูงที่สุด อยู่ที่ 0.660 และค่า mAP50 ของ YOLOv8l มีค่าสูงที่สุด อยู่ที่ 0.740 จากการวิเคราะห์ผลลัพธ์ในคลาสไม่สวมหมวกนิรภัย ในส่วนของค่าความถูกต้อง (Recall) ของโมเดลทุกรุ่นนั้นให้ค่าผลลัพธ์ที่ยังไม่สูงมาก เนื่องมาจากปริมาณของข้อมูลของคลาสไม่สวมหมวกนิรภัย ที่ใช้ในการฝึกโมเดลนั้นมีปริมาณที่น้อยกว่าเมื่อเปรียบเทียบกับคลาสอื่นในรูปที่ 5

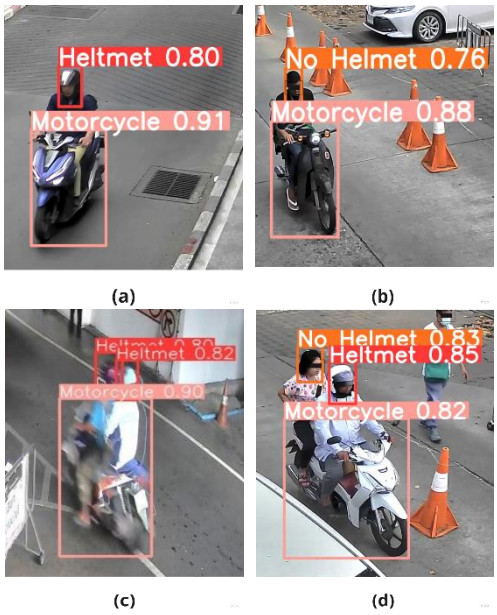
Model	Class	Precision	Recall	mAP50
YOLOv8n	No Helmet	0.747	0.572	0.648
YOLOv8s	No Helmet	0.797	0.611	0.714
YOLOv8m	No Helmet	0.837	0.644	0.730
YOLOv8l	No Helmet	0.856	0.634	0.740
YOLOv8x	No Helmet	0.823	0.660	0.738

ตารางที่ 8 ผลลัพธ์คลาสไม่สวมหมวกนิรภัย (No Helmet) ของโมเดลแต่ละรุ่น



รูปที่ 5 แสดงการเปรียบเทียบปริมาณข้อมูลทุกคลาสที่ใช้ในการฝึกโมเดล

ผลลัพธ์รูปภาพจากชุดข้อมูลทดสอบ



รูปที่ 6 แสดงตัวอย่างผลลัพธ์รูปภาพจากชุดข้อมูลทดสอบ

(a) แสดงตัวอย่างผลลัพธ์ที่โมเดลตรวจจับรถจักรยานยนต์และผู้ขับขี่ที่สวมหมวกนิรภัย, (b) แสดงตัวอย่างผลลัพธ์ที่โมเดลตรวจจับรถจักรยานยนต์และผู้ขับขี่ที่ไม่สวมหมวกนิรภัย, (c) แสดงตัวอย่างผลลัพธ์ที่โมเดลตรวจจับรถจักรยานยนต์ ผู้ขับขี่และผู้นั่งซ้อนท้ายที่สวมหมวกนิรภัย, (d) แสดงตัวอย่างผลลัพธ์ที่โมเดลตรวจจับรถจักรยานยนต์ ผู้ขับขี่ที่สวมหมวกนิรภัยและผู้นั่งซ้อนท้ายที่ไม่สวมหมวกนิรภัย

รูปที่ 6 จะแสดงให้เห็นว่าโมเดล YOLOv8 สามารถที่จะตรวจจับวัตถุในชุดข้อมูลทดสอบได้อย่างมีประสิทธิภาพ แต่จะพบความผิดพลาดในบางกรณี เช่น การไม่สามารถตรวจจับการสวมหมวกนิรภัยของเด็กที่นั่งอยู่บริเวณด้านหน้าของผู้ขับขี่ได้ เนื่องจากตำแหน่งที่นั่งของเด็กและขนาดศีรษะที่เล็กนั้นส่งผลให้โมเดลตรวจจับได้ผิดพลาด รวมไปถึงยังพบการไม่ตรวจจับวัตถุในการทดสอบกับข้อมูลวิดีโอ เพราะภายในข้อมูลวิดีโอที่เป็นภาพเคลื่อนไหวรถจักรยานยนต์บางคันจะขับขี่ด้วยความรวดเร็วทำให้โมเดลประมวลผลในการตรวจจับได้ไม่ทัน

สรุปผลการวิจัย

จากปัญหาการสัญจรด้วยรถจักรยานยนต์ภายในพื้นที่โรงพยาบาลศิริราช ที่ผู้ขับขี่และผู้ซ้อนรถจักรยานยนต์ นั้นมีการสวมใส่และไม่สวมใส่หมวกนิรภัยขณะสัญจร ทำให้ในบางครั้งการไม่ใส่หมวกนิรภัยนั้นนำมาซึ่งอุบัติเหตุที่ร้ายแรงต่อร่างกายได้ งานวิจัยนี้จึงสร้างขึ้นมาเพื่อเป็นระบบตรวจจับการสวมหมวกนิรภัยของผู้ขับขี่รถจักรยานยนต์ภายในพื้นที่โรงพยาบาลศิริราชทดแทนการตรวจจับที่ใช้ระบบสังเกตการณ์ด้วยบุคคล ทางผู้วิจัยจึงได้ทำการสร้างระบบตรวจจับการสวมหมวกนิรภัยของผู้ขับขี่

รถจักรยานยนต์ ที่นำเทคโนโลยีคอมพิวเตอร์ในการระบุและตรวจจับวัตถุ (Object Detection) โดยใช้การเรียนรู้เชิงลึก (Deep learning) คือ โมเดล YOLOv8 ซึ่งเป็นโมเดลที่สามารถใช้ในการระบุหรือตรวจจับวัตถุ ร่วมกับข้อมูลไฟล์ภาพที่รวบรวมจากกล้องวงจรปิด ทั้ง 4 ประตูภายในพื้นที่โรงพยาบาลศิริราช ประกอบไปด้วย ประตู 1, ประตู 5, ประตู 7 และประตู 8 ได้จำนวนรูปภาพที่ใช้ในการฝึกและทดสอบ จำนวน 3,000 รูป ผลลัพธ์ของการตรวจจับการสวมหมวกนิรภัยของผู้ขับขี่รถจักรยานยนต์ของทุกคลาส พบว่าโมเดล YOLOv8l และ YOLOv8x มีประสิทธิภาพสูงที่สุดจากทุกโมเดลโดย YOLOv8l และ YOLOv8x มีประสิทธิภาพสูงที่สุดจากทุกโมเดลโดย YOLOv8l มีค่าเฉลี่ยความแม่นยำ (Precision) อยู่ที่ 0.868, ค่าเฉลี่ยความถูกต้อง (Recall) อยู่ที่ 0.790, ค่าเฉลี่ย mAP50 อยู่ที่ 0.859 และค่าเฉลี่ย F1-Score อยู่ที่ 0.827 ส่วน YOLOv8x มีค่าเฉลี่ยความแม่นยำ (Precision) อยู่ที่ 0.842, ค่าเฉลี่ยความถูกต้อง (Recall) อยู่ที่ 0.811, ค่าเฉลี่ย mAP50 อยู่ที่ 0.858 และค่าเฉลี่ย F1-Score อยู่ที่ 0.826

กิตติกรรมประกาศ

การจัดทำวิจัยฉบับบนี้สำเร็จลุล่วงไปได้ด้วยการให้ความช่วยเหลือแนะนำของ ผศ.ดร.วราภรณ์ วิยานนท์ ซึ่งเป็นอาจารย์ที่ปรึกษาที่ได้กรุณาให้คำแนะนำข้อคิดเห็น ตรวจสอบ และแก้ไขร่างสารนิพนธ์มาโดยตลอด จึงขอขอบคุณมา ณ ที่นี้

ขอขอบคุณ คณาจารย์ทุกท่านในภาควิชาวิทยาการข้อมูล มหาวิทยาลัยศรีนครินทรวิโรฒ ที่แนะนำ ความรู้และแนวทางการทำสารนิพนธ์นี้

เอกสารอ้างอิง

[1] สวมหมวกนิรภัยลดโอกาสเสียชีวิต. (2023). NSTDA Retrieved 4 กุมภาพันธ์ 2566 from https://www.nstda.or.th/home/knowledge_post/government-areas-100percent-helmet-wearing-04/

[2] Rizzoli, A. (2564). One and two stage detectors. Vlabs. Retrieved 2 พฤศจิกายน 2566 from https://www.v7labs.com/blog/object-detection-guide

[3] YOLOv8 Performance Comparison. (2566). Keylabs. Retrieved 2 พฤศจิกายน 2566 from https://keylabs.ai/blog/yolov8-vs-ssd-choosing-the-right-object-detection-model/

[4] Jia W, Xu S, Liang Z, Zhao Y, Min H, Li S, et al. Real-time automatic helmet detection of motorcyclists in urban traffic using improved YOLOv5 detector. IET Image Processing. 2021;15(14):3623-37.

[5] Lin, H., Deng, J. D., Albers, D., & Siebert, F. W. (2020). Helmet use detection of tracked motorcycles using cnn-based multi-task learning. IEEE Access, 8, 162073-162084.

[6] Sanchana, M. A., & Eliyas, S. (2023). Automated Motorcycle Helmet Detection Using The Combination of YOLO AND CNN. 2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE),

[7] Dasgupta, M., Bandyopadhyay, O., & Chatterji, S. (2019). Automated helmet detection for multiple motorcycle riders using CNN. 2019 IEEE Conference on Information and Communication Technology,

[8] Sivaraj, J., Adithya, R. S., Alexander, A., Vishnudeep, M., Mohammed, S., Farhanudin, P., & Anusha, T. (2021). Helmet Violation Detection Application for Road Safety.

[9] Use Roboflow Annotate. (2563). Roboflow Retrieved 2 พฤศจิกายน 2566 from https://docs.roboflow.com/annotate/use-roboflow-annotate

[10] glenn-jocher, L.-q., AyushExel. (2566). Models. ultralytics. Retrieved 2 พฤศจิกายน 2566 from https://docs.ultralytics.com/tasks/detect/