ระบบตรวจจับเรือในแม่น้ำเจ้าพระยาด้วยเทคนิค YOLOv7

Boat Detection System in the Chao Phraya River Using YOLOv7 Techniques

สราวุฒิ สายทวี[1]* สมชาติ จิริวิภากร[2]

[1]สาขาวิชาวิศกรรมป้องกันประเทศ, คณะวิศวกรรมศาสตร์, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ลาดกระบัง, กรุงเทพมหานคร 10520

[2]สาขาวิชาวิศกรรมไฟฟ้า, คณะวิศวกรรมศาสตร์, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ลาดกระบัง, กรุงเทพมหานคร 10520

Sarawut Saitawee[1]* Somchat Jiriwibhakorn[2]

[1] Engineering Program in Defence Engineering, School of Engineering, King Mongkut's Institute of Technology Ladkrabang, Lat Krabang, Bangkok, 10520, Thailand

[2] Department of Electrical Engineering, School of Engineering, King Mongkut's Institute of Technology Ladkrabang, Lat Krabang, Bangkok, 10520, Thailand

Corresponding author Email: s.saitawee@gmail.com[1] somchat.ji@kmitl.ac.th[2]

(Received: October 2, 2024; Revised: June 24, 2025; Accepted: June 30, 2025)

บทคัดย่อ

งานวิจัยนี้มุ่งพัฒนาระบบตรวจจับเรือในแม่น้ำเจ้าพระยาโดยใช้เทคนิค YOLOv7 เพื่อแก้ไขปัญหาการจราจรทางน้ำที่หนาแน่นและอันตราย โดยเปรียบเทียบประสิทธิภาพของ 3 อัลกอริทึม ได้แก่ YOLO (You Only Look Once), SSD (Single Shot MultiBox Detector) และ Faster R-CNN (Region-based Convolutional Neural Network) บนชุดข้อมูลภาพถ่ายจากกล้องที่มีความละเอียด 608x608 พิกเซล จำนวน 1,000 ภาพ แบ่งเป็นภาพที่มีเรือ 750 ภาพ และภาพไม่มีเรือ 250 ภาพ ประเมินประสิทธิภาพของโมเดลโดยใช้ confusion matrix ผลการทดสอบพบว่าอัลกอริทึม YOLOv7 ให้ค่า F1-Score สูงสุด (95.0%) รองลงมาคือ Faster R-CNN (93.0%) และ SSD (82.0%) เพราะฉะนั้น อัลกอริทึม YOLOv7 มีประสิทธิภาพในการตรวจจับเรือในแม่น้ำเจ้าพระยาสูงที่สุดในการทดลองนี้

คำสำคัญ: ระบบตรวจจับเรือ แม่น้ำเจ้าพระยา YOLOv7 จราจรทางน้ำ ประสิทธิภาพ

ABSTRACT

This research aims to develop a system for detecting boats in the Chao Phraya River using the YOLOv7 technique to address the issues of heavy and hazardous water traffic. The performance of three algorithms YOLO (You Only Look Once), SSD (Single Shot MultiBox Detector), and Faster R-CNN (Region-based Convolutional Neural Network) was compared using a dataset of 1,000 images from cameras with a resolution of 608x608 pixels, consisting of 750 images with boats and 250 images without boats. The models' performance was evaluated using a confusion matrix. The test results showed that the YOLOv7 algorithm achieved the highest F1-Score (95.0%), followed by Faster R-CNN (93.0%) and

SSD (82.0%). Therefore, the YOLOv7 algorithm was the most effective in detecting boats in the Chao Phraya River in this experiment.

Keyword: Boat Detection, Chao Phraya River, YOLOv7, water traffic, performance.

1. บทนำ

แม่น้ำเจ้าพระยาเป็นแม่น้ำสายหลักของประเทศไทย เป็นแหล่งคมนาคมขนส่งที่สำคัญของประเทศ และเป็นแหล่งทรัพยากรธรรมชาติที่สำคัญของประเทศเช่นกัน อย่างไรก็ตามแม่น้ำเจ้าพระยายังเป็นเส้นทางลักลอบขนสินค้าผิดกฎหมายที่สำคัญของประเทศอีกด้วย สินค้าผิดกฎหมายเหล่านี้มักเป็นสินค้าหนีภาษี เช่น น้ำมันเถื่อน บุหรี่หนีภาษี สุราหนีภาษี ยาผิดกฎหมาย และสินค้าอื่น ๆ ที่ผิดกฎหมาย การลักลอบขนสินค้าผิดกฎหมายเหล่านี้ก่อให้เกิดความเสียหายต่อประเทศทั้งทางเศรษฐกิจและสังคม ปัญหาเรื่องเรือผิดกฎหมายลักลอบเข้ามาทางแม่น้ำเจ้าพระยาโดยเรือผิดกฎหมายเหล่านี้มักลักลอบขนสินค้าที่ผิดกฎหมาย และเรือเหล่านี้มักลักลอบเข้ามาทางแม่น้ำเจ้าพระยาจากต่างประเทศหรือจากพื้นที่ชายแดน จากนั้นจึงนำสินค้าผิดกฎหมายไปจำหน่ายยังพื้นที่ต่าง ๆ ในประเทศไทย ก่อให้เกิดผลกระทบต่อประเทศหลายประการ เช่น รายได้ของภาครัฐลดลง เนื่องจากไม่ได้รับภาษีจากสินค้าผิดกฎหมาย เกิดการแข่งขันไม่เป็นธรรมกับผู้ประกอบการที่ปฏิบัติตามกฎหมาย ส่งผลกระทบต่อสุขภาพของประชาชน เนื่องจากสินค้าผิดกฎหมายบางชนิดอาจก่อให้เกิดอันตราย กระทบต่อความมั่นคงของประเทศ เนื่องจากสินค้าผิดกฎหมายบางชนิดอาจเกี่ยวข้องกับการก่อการร้ายหรืออาชญากรรมอื่น ๆ [1-3]

1.1 ความเป็นมาและความสำคัญของปัญหา

ปัญหาเรือผิดกฎหมายลักลอบเข้ามาทางแม่น้ำเจ้าพระยามีมาอย่างยาวนาน สาเหตุหลักมาจากการที่แม่น้ำเจ้าพระยาเป็นเส้นทางน้ำสายหลักที่เชื่อมต่อระหว่างประเทศไทยกับประเทศเพื่อนบ้านในภูมิภาคเอเชียตะวันออกเฉียงใต้ มีความยาวกว่า 372 กิโลเมตร และมีความกว้างตั้งแต่ 200 ถึง 1,000 เมตร ประกอบกับสภาพภูมิประเทศที่เป็นแม่น้ำลำคลองมากมาย ทำให้ยากต่อการควบคุมและตรวจตรา นอกจากนี้ ยังมีปัจจัยอื่น ๆ ที่เอื้อต่อการลักลอบนำเข้าสินค้าผิดกฎหมายทางแม่น้ำเจ้าพระยา เช่น ช่องโหว่ในกฎหมายและมาตรการทางศุลกากร ขาดความร่วมมือจากหน่วยงานที่เกี่ยวข้อง และความยากลำบากในการพิสูจน์ทราบสินค้าผิดกฎหมาย

ระบบการตรวจจับเรือขนาดเล็ก ซึ่งมีความสำคัญต่อเทคโนโลยีเรืออัตโนมัติและความปลอดภัยในการเดินเรือ วิธีการตรวจจับแบบเดิมที่อาศัยเรดาร์บนเรือ ไม่สามารถตรวจจับเรือขนาดเล็กหรือเรือที่อยู่ใกล้ได้อย่างมีประสิทธิภาพ ดังนั้น บทความเสนอวิธีการใหม่ที่ใช้ deep learning ร่วมกับกล้องวีดิโอ เพื่อตรวจจับเรือขนาดเล็กโดยอาศัยปัญญาประดิษฐ์[4]

การใช้กล้องในการตรวจจับด้วยการนำภาพดิจิทัลมาประมวลผลและแก้ไขปัญหานี้ เป็นระบบตรวจจับเรือด้วยเทคนิค Deep Learning ซึ่งระบบประกอบด้วยโมดูลหลัก 2 โมดูล ดังนี้ โมดูลการตรวจจับวัตถุ (Object Detection) โมดูลนี้ใช้เทคนิค Deep Learning เช่น YOLOv5, SSD หรือ Faster R-CNN เพื่อตรวจจับวัตถุที่ปรากฏในภาพดิจิทัล โมดูลนี้จะระบุตำแหน่งและขนาดของวัตถุที่เป็นเรือ และโมดูลการจำแนกวัตถุ (Object Classification) โมดูลนี้ใช้เทคนิค Deep Learning เช่น ResNet หรือ Inception เพื่อจำแนกประเภทของวัตถุที่ตรวจจับได้ โมดูลนี้จะจำแนกวัตถุที่เป็นเรือออกจากวัตถุอื่น ๆ เช่น ต้นไม้ อาคาร หรือทุ่น แต่ในงานวิจัยนี้จะเน้นไปที่โมดูลการตรวจจับเรือเท่านั้น เพื่อหาอัลกอริทึมที่เหมาะสมในการนำไปพัฒนาแก้ปัญหาได้ต่อไป

1.2 วรรณกรรมที่เกี่ยวข้อง

YOLO, SSD และ Faster R-CNN ล้วนเป็นอัลกอริทึมการตรวจจับวัตถุ (Object Detection) ที่ได้รับความนิยม แต่ละอัลกอริทึมมีจุดเด่นและจุดด้อยที่แตกต่างกัน เหมาะกับการใช้งานที่แตกต่างกันไป สำหรับความเหมือนคือ ทั้ง

สามอัลกอริทึมใช้ Convolutional Neural Networks (CNN) สำหรับการดึงคุณลักษณะ (feature extraction) จากภาพ แต่ละอัลกอริทึมสามารถตรวจจับวัตถุได้หลายประเภทพร้อมกัน (multi-class object detection) และแต่ละอัลกอริทึมสามารถระบุขอบเขต (bounding box) ของวัตถุที่ตรวจจับพบได้ ในเรื่องของความแตกต่างในเรื่องวิธีการโดย Faster R-CNN จะประกอบด้วยสองขั้นตอน ขั้นตอนแรกสร้าง Region Proposals และขั้นตอนที่สองทำการจำแนกวัตถุและระบุขอบเขต SSD จะเป็นอัลกอริทึมแบบ One-stage ทำการดึงคุณลักษณะและทำนายวัตถุในขั้นตอนเดียว และYOLO จะเป็นอัลกอริทึมแบบ One-stage คล้ายกับ SSD แต่มีโครงสร้างเครือข่ายที่เรียบง่ายกว่า ทำให้มีความเร็วสูงขึ้น นอกจากนั้นยังมีในเรื่องของ ความเร็ว ความแม่นยำ และความเหมาะกับงาน ที่มีความแตกต่างกันบ้าง

โดยสรุปแล้ว Faster R-CNN เหมาะกับงานที่ต้องการความแม่นยำสูง [5] SSD เหมาะกับงานที่ต้องการความสมดุลระหว่างความเร็วและความแม่นยำ [6] และ YOLO เหมาะกับงานที่ต้องการความเร็วสูง [7] ทั้งนี้ก็ขึ้นอยู่กับการนำไปใช้งานให้เหมาะสมกับความต้องการ ประเภทของงาน ความต้องการด้านความเร็วและความแม่นยำและทรัพยากรการประมวลผลที่มีอยู่

ในงานวิจัยที่ผ่านมามีงานวิจัยหลายตัวที่ได้ศึกษาข้อมูลเพิ่มเติมเกี่ยวกับการพัฒนาเทคโนโลยีการตรวจจับวัตถุ อาทิเช่น Ren, S. และคณะ[8] นำเสนออัลกอริทึม Faster R-CNN ใหม่สำหรับการตรวจจับวัตถุแบบเรียลไทม์ ผลการทดลอง Faster R-CNN บนชุดข้อมูล COCO 2015 ผลลัพธ์แสดงให้เห็นว่า Faster R-CNN ทำงานได้เร็วกว่าอัลกอริทึมตรวจจับวัตถุแบบเรียลไทม์อื่นๆ ที่มีความแม่นยำใกล้เคียงกัน สามารถตรวจจับวัตถุได้ 7 fps (frames per second) บน CPU Intel i7-4790K ในขณะที่ Faster RCNN (VGG16) ทำงานได้เพียง 0.3 fps บน CPU เดียวกัน นอกจากนี้ Faster R-CNN ยังประสบความสำเร็จในการตรวจจับวัตถุ 80 ประเภท บนชุดข้อมูล COCO 2015 Faster R-CNN บรรลุค่า mAP (mean average precision) ที่ 33.2% ซึ่งใกล้เคียงกับ Faster RCNN (VGG16) (31.2%) Dai, J. และคณะ[9] นำเสนอวิธีการตรวจจับวัตถุ R-FCN ซึ่งใช้ Fully Convolutional Networks (FCN) การทดลองในงานวิจัยนี้ได้เปรียบเทียบประสิทธิภาพของ R-FCN กับวิธีการอื่นๆ บนชุดข้อมูล PASCAL VOC โดยพบว่า R-FCN มีผลลัพธ์ที่แข็งแกร่งและมีความเร็วในการทดสอบที่ดีกว่า นอกจากนี้ การใช้ Position-Sensitive Score Maps ยังช่วยลดปัญหาความต่างของตำแหน่งในการตรวจจับวัตถุอย่างมีประสิทธิภาพ

Cai, Z. และคณะ[10] นำเสนอสถาปัตยกรรม cascaded R-CNN สำหรับการตรวจจับวัตถุ เพื่อแก้ปัญหาสองปัจจัยหลักที่ทำให้ประสิทธิภาพของตัวตรวจจับวัตถุลดลง ปัจจัยแรกคือการโอเวอร์ฟิตติ้ง (Overfitting) ระหว่างการฝึกอบรม เกิดจากการขาดตัวอย่างเชิงบวก ปัจจัยที่สองคือความไม่ตรงกันระหว่างเกณฑ์ IoU (Intersection over Union) ที่ใช้ในเวลาฝึกอบรมและเวลาอนุมาน ซึ่งสถาปัตยกรรมนี้ประกอบด้วยตัวตรวจจับหลายขั้นตอน แต่ละขั้นตอนได้รับการฝึกฝนด้วยเกณฑ์ IoU ที่สูงขึ้น การออกแบบแบบเรียงซ้อนช่วยปรับปรุงประสิทธิภาพโดยลดการโอเวอร์ฟิตติ้งและทำให้แน่ใจว่าการฝึกอบรมและการอนุมานตรงกันมากขึ้น ในการทดลอง cascaded R-CNN ทำงานได้ดีกว่าตัวตรวจจับอื่นๆ บนชุดข้อมูล COCO Fu, C. และคณะ[11] เสนอวิธีการตรวจจับวัตถุแบบ Single-Shot Multibox Detector (SSD) ซึ่งใช้ Single Deep Neural Network (DNN) โดย SSD ทำงานโดย discretize output space ของ bounding box เป็นชุดของ default boxes หลายขนาด และ aspect ratios บนแต่ละ feature map location ในขั้นตอนการอนุมาน DNN จะสร้างคะแนนสำหรับการมีอยู่ของวัตถุแต่ละประเภทใน default box แต่ละกล่อง และปรับ bounding box ให้ดีขึ้นเพื่อให้ตรงกับรูปร่างของวัตถุ SSD ผสานรวมการคาดการณ์จาก feature maps หลายความละเอียดเพื่อจัดการวัตถุที่มีขนาดหลากหลาย ข้อดีของ SSD คือ ความ

เรียบง่ายเมื่อเทียบกับวิธีการที่ต้องใช้ Region Proposal Network (RPN) ทดสอบ SSD บนชุดข้อมูล PASCAL VOC, MS COCO และ ILSVRC ผลการทดลองแสดงให้เห็นว่า SSD มีประสิทธิภาพใกล้เคียงกับวิธีการที่ใช้ RPN Meng, S. และคณะ[12] นำเสนอ Two-Stage SSD (TS-SSD) ซึ่งเป็นสถาปัตยกรรมการตรวจจับวัตถุแบบสองขั้นตอนที่ผสานรวมข้อดีของ Single-Stage SSD (SSD) และ Two-Stage detectors TS-SSD ประกอบด้วยสองขั้นตอน คือขั้นตอนแรกการค้นหา Region ใช้ Single-Shot MultiBox Detector (SSD) เพื่อค้นหา Region ของ Interest (RoI) ขั้นตอนถัดไปคือการจำแนกประเภท ใช้ Region-based Convolutional Neural Network (R-CNN) เพื่อจำแนกวัตถุภายใน ผลการทดลองทดสอบ TS-SSD บนชุดข้อมูล PASCAL VOC, MS COCO และ ImageNet แสดงให้เห็นว่า TS-SSD มีประสิทธิภาพและความแม่นยำสูงกว่า SSD แบบดั้งเดิม TS-SSD เป็นสถาปัตยกรรมการตรวจจับวัตถุที่มีประสิทธิภาพและความแม่นยำสูง เหมาะสำหรับการใช้งานที่ต้องการทั้งความเร็วและความแม่นยำ อย่างไรก็ตาม TS-SSD มีความซับซ้อนสูงกว่า SSD แบบดั้งเดิม

Redmon, J. และคณะ[13] เสนอ You Only Look Once (YOLO) ซึ่งเป็นวิธีการตรวจจับวัตถุแบบ Single-Stage YOLO ประกอบด้วย Convolutional Neural Network (CNN) เพียงชุดเดียว ทำหน้าที่ทั้งค้นหา Region ของ Interest (RoI) และจำแนกวัตถุภายใน RoI ทำงานโดยแบ่งภาพออกเป็นชุดของ Grid Cells แต่ละ Grid Cell จะทำหน้าที่คาดการณ์ Bounding Box ของวัตถุที่ปรากฏใน Grid Cell นั้น ทดสอบ YOLO บนชุดข้อมูล PASCAL VOC, COCO และ ImageNet ผลการทดลองแสดงให้เห็นว่า YOLO มีประสิทธิภาพและความเร็วในการตรวจจับวัตถุสูงกว่าวิธีการอื่น ๆ และเหมาะสำหรับการใช้งานแบบ Real-Time Ge, Z. และคณะ[14] นำเสนอ YOLOX ในการตรวจจับวัตถุโดยผสมผสานเทคนิคการ stitching แบบหลายระดับเข้ากับ YOLOv5 ทดสอบบนชุดข้อมูล COCO 2017 ผลที่ได้คือ YOLOX สามารถบรรลุความแม่นยำ mAP (mean Average Precision) 56.5% บน COCO 2017 ซึ่งสูงกว่า YOLOv5 4.4% ผลสรุปคือ YOLOX เป็นโมเดลการตรวจจับวัตถุแบบ Single-Stage ที่มีความแม่นยำสูง scalable และประสิทธิภาพดี

Hu, H. และคณะ [15] ได้นำเสนอการตรวจจับวัตถุเบื้องต้นด้วยโมเดล YOLOv7 ที่ได้รับการปรับปรุงในการตรวจจับวัตถุในภาพหรือวิดีโอ เฟรม โมเดลนี้ได้รับการปรับแต่งด้วยชุดข้อมูลเรือเพื่อเพิ่มประสิทธิภาพการตรวจจับ และในการติดตามวัตถุได้ใช้ตัวกรอง Kalman ช่วยคาดการณ์ตำแหน่งและสถานะของวัตถุในเฟรมถัดไป ซึ่งช่วยให้สามารถติดตามวัตถุได้อย่างราบรื่นแม้ว่าวัตถุจะบดบังหรือเคลื่อนที่เร็ว ผลการทดลองแสดงให้เห็นว่าวิธีการที่เสนอนี้มีประสิทธิภาพและแม่นยำในการตรวจจับเรือแบบเรียลไทม์ วิธีนี้สามารถตรวจจับเรือได้หลากหลายประเภท แม้ว่าจะอยู่ในสภาพแสงที่หลากหลายและสภาพพื้นผิวที่ซับซ้อน Xie, S.และคณะ [16] นำเสนอโมเดล YOLOv7 ที่ได้รับการปรับปรุงเพื่อสกัดคุณลักษณะจากภาพหรือวิดีโอ เฟรม โมเดลนี้ได้รับการปรับแต่งด้วยชุดข้อมูลเรือเพื่อเพิ่มประสิทธิภาพการสกัดคุณลักษณะ โดยมุ่งเน้นไปที่บริเวณของภาพหรือวิดีโอ เฟรมที่น่าจะเป็นเรือ กลไกนี้ช่วยเพิ่มประสิทธิภาพการตรวจจับโดยมุ่งเน้นไปที่บริเวณที่เกี่ยวข้องมากที่สุด โดยโมเดลการจำแนกประเภทใช้เพื่อจำแนกวัตถุที่ตรวจจับได้ว่าเป็นเรือหรือไม่ โมเดลนี้ได้รับการฝึกฝนด้วยชุดข้อมูลเรือเพื่อเพิ่มประสิทธิภาพการจำแนกประเภท ผลการทดลองแสดงให้เห็นว่าวิธีการที่เสนอนี้มีประสิทธิภาพและแม่นยำในการตรวจจับเรือ Wang, Y. และคณะ [17] เสนอวิธีการตรวจจับเรือแบบเรียลไทม์โดยใช้ YOLOv7 และเครือข่ายน้ำหนักเบา เริ่มต้นด้วยการปรับแต่ง YOLOv7 ด้วยการลดขนาดโมเดลและเพิ่มโมดูลคอขวดแบบขยาย โมเดลที่ปรับแต่งแล้วนี้เรียกว่า YOLOLite-v7 จากนั้น YOLOLite-v7 นำไปฝึกอบรมบนชุดข้อมูลเรือขนาดใหญ่ ผลการทดลองแสดงให้เห็นว่า YOLOLite-v7 สามารถตรวจจับเรือได้

อย่างแม่นยำและรวดเร็ว โดยมีค่า mAP (mean Average Precision) อยู่ที่ 92.5% ที่ความเร็ว 50 FPS

จากการศึกษาวรรณกรรมข้างต้นพบว่าเทคโนโลยีการตรวจจับวัตถุมีการพัฒนาอย่างรวดเร็ว มีการนำเสนอวิธีการใหม่ ๆ อยู่เสมอ วิธีการเหล่านี้มีทั้งแบบ Single-Stage และ Two-Stage แต่ละวิธีการมีข้อดีและข้อเสียแตกต่างกันไป ขึ้นอยู่กับความต้องการใช้งาน โดยวิธีการตรวจจับวัตถุแบบ Single-Stage มีข้อดีคือความเร็วในการตรวจจับวัตถุสูง และมีความซับซ้อนต่ำ แต่มีข้อเสียคือความแม่นยำในการตรวจจับวัตถุอาจจะต่ำกว่าวิธีการแบบ Two-Stage ส่วนวิธีการตรวจจับวัตถุแบบ Two-Stage มีข้อดีคือ ความแม่นยำในการตรวจจับวัตถุสูง แต่มีข้อเสียคือ ความเร็วในการตรวจจับวัตถุต่ำ และความซับซ้อนสูง

Faster R-CNN เป็นวิธีการแบบ Two-Stage ที่มีความแม่นยำสูง แต่ความเร็วในการตรวจจับวัตถุค่อนข้างต่ำ SSD เป็นวิธีการแบบ Single-Stage ที่มีความเร็วในการตรวจจับวัตถุสูง SSD มีประสิทธิภาพใกล้เคียงกับวิธีการแบบ Two-Stage ที่ใช้ RPN TS-SSD เป็นวิธีการแบบ Two-Stage ที่ผสานข้อดีของ SSD และ Two-Stage detectors เข้าด้วยกัน TS-SSD มีประสิทธิภาพและความแม่นยำสูงกว่า SSD แบบดั้งเดิม YOLO เป็นวิธีการแบบ Single-Stage ที่มีความเร็วในการตรวจจับวัตถุสูง YOLO มีประสิทธิภาพและความเร็วในการตรวจจับวัตถุสูงกว่าวิธีการอื่น ๆ บนชุดข้อมูล PASCAL VOC, COCO และ ImageNet YOLOv7 เป็นวิธีการแบบ Single-Stage ที่มีความแม่นยำสูง YOLOv7 สามารถตรวจจับเรือได้หลากหลายประเภท แม้ว่าจะอยู่ในสภาพแสงที่หลากหลายและสภาพพื้นผิวที่ซับซ้อน

ได้เลือกอัลกอริทึม YOLOv7, SSD และ Faster R-CNN เพราะทั้งสามเป็น อัลกอริทึมที่มีชื่อเสียงและได้รับการยอมรับอย่างแพร่หลาย ในงานด้าน object detection และครอบคลุม ลักษณะเด่นต่างกัน ได้แก่

- YOLOv7 เป็นอัลกอริทึมแบบ one-stage ที่เน้นความเร็วและมีความแม่นยำสูงในเวอร์ชันล่าสุด จึงเหมาะสำหรับงาน real-time

- SSD เป็น one-stage เช่นกัน แต่มีสมดุลระหว่างความเร็วและความแม่นยำ

- Faster R-CNN เป็น two-stage ที่เน้นความแม่นยำมากกว่า แต่ประมวลผลช้ากว่า

ดังนั้น การเลือกทั้ง 3 อัลกอริทึมนี้มาเปรียบเทียบครอบคลุมหลากหลายมุมมอง ทั้งความเร็ว ความแม่นยำ และความซับซ้อนของโมเดล เพื่อหาวิธีที่ เหมาะสมกับงานตรวจจับเรือในแม่น้ำเจ้าพระยา ที่ต้องเผชิญกับฉากหลังซับซ้อนและมีวัตถุขนาดเล็ก

แนวโน้มในอนาคตเทคโนโลยีการตรวจจับวัตถุมีแนวโน้มที่จะพัฒนาอย่างต่อเนื่อง ในอนาคต วิธีการตรวจจับวัตถุน่าจะมีความแม่นยำและรวดเร็วมากขึ้น สามารถตรวจจับวัตถุได้หลากหลายประเภท ในสภาพแสงที่หลากหลายและสภาพพื้นผิวที่ซับซ้อน และสามารถใช้งานได้กับอุปกรณ์ที่มีทรัพยากรจำกัด

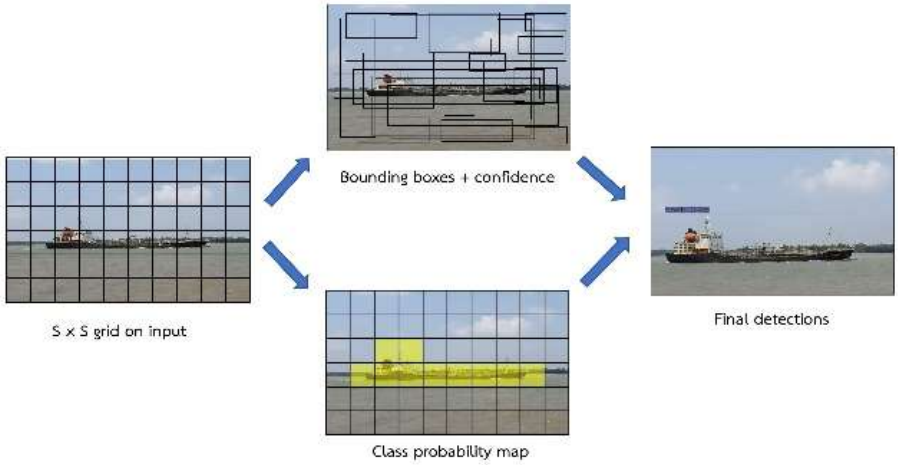
1.3 วัตถุประสงค์ของงานวิจัย

งานวิจัยนี้มุ่งพัฒนาระบบตรวจจับเรือในแม่น้ำเจ้าพระยาโดยใช้เทคนิค Deep Learning เพื่อแก้ไขปัญหาการจราจรทางน้ำที่หนาแน่นและอันตราย วัตถุประสงค์หลักของงานวิจัยนี้มีดังนี้

1) พัฒนาระบบตรวจจับเรือในแม่น้ำเจ้าพระยาโดยใช้เทคนิค Deep Learning เน้นการทำงานของอัลกอริทึม YOLOv7

2) เปรียบเทียบประสิทธิภาพของอัลกอริทึม Deep Learning 3 ชนิด ประกอบไปด้วยอัลกอริทึม YOLOv7, SSD และ Faster R-CNN ในการตรวจจับเรือจากภาพถ่าย การวัดประสิทธิภาพของโมเดลจะประเมินโดยใช้ confusion matrix

3) ค้นหาอัลกอริทึม Deep Learning ที่เหมาะสมที่สุดสำหรับการตรวจจับเรือในแม่น้ำเจ้าพระยา งานวิจัยนี้มุ่งค้นหาอัลกอริทึม Deep Learning ที่มีความแม่นยำสูง ใช้เวลาประมวลผลน้อย และเหมาะสมกับการใช้งานจริง ตามที่แสดงดังรูปที่ 1



รูปที่ 1 หลักการทำงานของโครงสร้างเครือข่ายในอัลกอริทึม YOLOv7

2. ขอบเขตงานวิจัย

1) งานวิจัยมุ่งเน้นไปที่การพัฒนาระบบตรวจจับเรือในแม่น้ำเจ้าพระยาโดยใช้เทคนิค Deep Learning

2) เปรียบเทียบประสิทธิภาพของ 3 อัลกอริทึม Deep Learning ได้แก่ YOLOv7 (You Only Look Once), SSD (Single Shot MultiBox Detector) และ Faster R-CNN (Region-based Convolutional Neural Network) บนชุดข้อมูลภาพถ่ายจากกล้องที่มีความละเอียด 608x608 พิกเซล

3) ชุดข้อมูลประกอบด้วย ภาพที่มีเรือ 750 ภาพ ประกอบไปด้วยภาพเรือ เรือสินค้า, เรือบรรทุกแก๊ส - น้ำมัน, เรือประมงขนาดเล็ก, เรือข้ามฝาก และเรือลากจูงอย่างล่ะเท่า ๆ กัน โดยใช้ภาพเรือด้านข้างแนวนอนของเรือเท่านั้น และภาพที่ไม่มีเรือ 250 ภาพ ภาพเรือที่ได้จากเรือที่วิ่งผ่านบริเวณร่องน้ำเจ้าพระยา ซึ่งครอบคลุมพื้นที่แม่น้ำเจ้าพระยาขอบเขตของพื้นที่ท่าเสือซ้อนเล็บ รร.นายเรือ จังหวัดสมุทรปราการ ตามแนวชายฝั่งที่มีความกว้างประมาณ 800 เมตรเท่านั้น

4) ใช้ชุดข้อมูล COCO 2015 (Common Objects in Context) เป็นชุดข้อมูลสำหรับใช้ฝึกโมเดล

5) ภาพที่ใช้มีแสงสว่างเพียงพอ ไม่มีเงามาบดบัง

6) วัดประสิทธิภาพด้วยค่า (Confusion Matrix)

7) คอมพิวเตอร์โน๊ตบุ๊ค ระบบปฏิบัติการ Windows 10 โปรเซสเซอร์ AMD Ryzen 7 รุ่น 4800H ความเร็วสัญญาณนาฬิกา 2.90 GHz กราฟิก Radeon Graphics

8) ขนาด input 608x608 พิกเซล การทำ preprocessing มีการดำเนินการ เช่น Resize ภาพ ให้เข้ากับ input size ของ YOLOv7 แปลงค่าสี เป็น RGB หรือ BGR ตามที่โมเดลรองรับ และ normalize ค่า pixel (เป็นมาตรฐานทั่วไปใน deep learning)

3. ทฤษฎีที่นำเสนอ

ในหัวข้อนี้จะกล่าวถึงคำอธิบายเกี่ยวกับเทคนิคการเรียนรู้เชิงลึก (Deep Learning) และวิธีการทำงาน ประโยชน์และความสามารถของ Deep Learning ในการประมวลผลภาพและตรวจจับวัตถุ ภาพรวมของปัญหาการตรวจจับวัตถุในภาพและความสำคัญของการพัฒนาอัลกอริทึมในการแก้ไขปัญหานี้ อธิบายถึงความท้าทายที่เกี่ยวข้องกับการตรวจจับเรือในแม่น้ำเจ้าพระยา อัลกอริทึม YOLO (You Only Look Once) ในส่วนของหลักการทำงานของ YOLO และเหตุผลที่ทำให้มันมีประสิทธิภาพสูงในงานตรวจจับวัตถุ ข้อดีและข้อเสียของ YOLO เมื่อเปรียบเทียบกับอัลกอริทึมอื่น ๆ อัลกอริทึม SSD (Single Shot MultiBox Detector) อธิบายเกี่ยวกับการทำงานของ SSD และความสามารถในการตรวจจับ

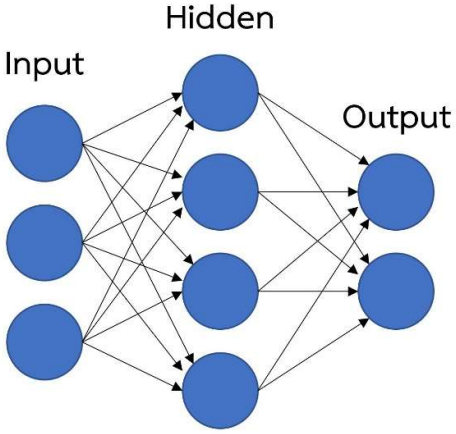
วัตถุ จุดเด่นและจุดด้อยของ SSD อัลกอริทึม Faster R-CNN (Region-based Convolutional Neural Network) หลักการทำงานของ Faster R-CNN และวิธีการที่ใช้ในการตรวจจับวัตถุในภาพ เปรียบเทียบประสิทธิภาพของ Faster R-CNN กับ YOLO และ SSD ทฤษฎี Confusion Matrix การประเมินประสิทธิภาพของโมเดลโดยใช้ Confusion Matrix การคำนวณค่า F1-Score และความหมายของมันในบริบทของการตรวจจับวัตถุ

3.1 Deep Learning

Deep Learning [18,19] หรือ การเรียนรู้เชิงลึกเป็นเทคนิคการเรียนรู้ของเครื่องประเภทหนึ่ง โดยใช้โครงข่ายประสาทเทียม (Neural Networks) ที่มีโครงสร้างซับซ้อนหลายชั้น เพื่อเรียนรู้จากข้อมูลจำนวนมาก Deep Learning ประสบความสำเร็จอย่างมากในงานประมวลผลภาษาธรรมชาติ การจดจำภาพ การแปลภาษา และงานอื่น ๆ อีกมากมาย

หลักการพื้นฐานของ Deep Learning ประกอบไปด้วย 1) โครงข่ายประสาทเทียม (Neural Networks) โครงสร้างที่เลียนแบบการทำงานของสมองมนุษย์ ประกอบด้วยหน่วยประมวลผลขนาดเล็กที่เรียกว่า นิวรอน เชื่อมต่อกันเป็นชั้น ๆ นิวรอนแต่ละตัวทำหน้าที่รับข้อมูล ประมวลผล และส่งผลลัพธ์ไปยังนิวรอนถัดไป 2) การเรียนรู้ Deep Learning เรียนรู้จากข้อมูลจำนวนมาก โดยปรับค่าพารามิเตอร์ของโครงข่ายประสาทเทียมให้เหมาะสม เพื่อให้สามารถจำลองรูปแบบของข้อมูลได้อย่างถูกต้อง 3) ฟังก์ชันการกระตุ้น (Activation Functions) กำหนดว่านิวรอนจะส่งผลลัพธ์อะไร ขึ้นอยู่กับข้อมูลที่รับเข้ามา 4) ฟังก์ชันการสูญเสีย (Loss Functions) วัดความผิดพลาดของโมเดล ช่วยให้นำโมเดลไปปรับปรุงให้ดีขึ้น และ5) อัลกอริทึมการปรับแต่ง (Optimization Algorithms) ปรับค่าพารามิเตอร์ของโมเดล เพื่อลดค่าฟังก์ชันการสูญเสีย

Deep Learning ประสบความสำเร็จอย่างมากในงานตรวจจับวัตถุ โดยมีอัลกอริทึมยอดนิยมหลายตัว เช่น YOLO (You Only Look Once) ที่มีความเร็วในการทำงานสูง SSD (Single Shot MultiBox Detector) ที่มีความเร็วในการทำงานปานกลาง มีความแม่นยำสูง และ Faster R-CNN (Region-based Convolutional Neural Network) ที่มีความแม่นยำสูง แต่มีความเร็วในการทำงานช้า ดังแสดงโครงข่ายประสาทเทียมซึ่งเป็นโครงสร้างพื้นฐานของ Deep Learning ดังรูปที่ 2



รูปที่ 2 โครงสร้างพื้นฐานของ Deep Learning

จากรูปที่ 2 โครงสร้างพื้นฐานของ Deep Learning ประกอบด้วยหน่วยประมวลผลขนาดเล็กที่เรียกว่า นิวรอน เชื่อมต่อกันเป็นชั้น ๆ นิวรอนแต่ละตัวทำหน้าที่รับข้อมูล ประมวลผล และส่งผลลัพธ์ไปยังนิวรอนถัดไป โครงข่ายประสาทเทียมเรียนรู้จากข้อมูลจำนวนมาก โดยปรับค่าพารามิเตอร์ของการเชื่อมต่อระหว่างนิวรอน เพื่อให้สามารถจำลองรูปแบบของข้อมูลได้อย่างถูกต้อง

3.2 ทฤษฎี YOLOv7

YOLOv7[20,21] YOLO (You Only Look Once) เป็นสถาปัตยกรรมที่ใช้สำหรับการตรวจจับวัตถุ (Object Detection) ในภาพ YOLOv7 เป็นเวอร์ชั่นล่าสุดของ YOLO ที่ได้รับการปรับปรุงจาก YOLO รุ่นก่อน ๆ โดยเน้นการเพิ่มประสิทธิภาพและความแม่นยำในการตรวจจับวัตถุ หลักการทำงานของ YOLOv7 ประกอบด้วยขั้นตอนหลัก ๆ ดังนี้

ขั้นที่ 1 การแบ่งภาพเป็นกริด ภาพถูกแบ่งเป็นกริดของเซลล์ที่มีขนาดเท่ากัน (เช่น 13x13, 19x19 ฯลฯ)

ขั้นที่ 2 การพยากรณ์กรอบเขตแดน (Bounding Boxes) และความน่าจะเป็น ในแต่ละเซลล์ของกริด YOLO จะพยากรณ์กรอบเขตแดนหลายกรอบพร้อมกับความน่าจะเป็นที่กรอบนั้นจะครอบคลุมวัตถุที่สนใจ

ขั้นที่ 3 การประมวลผลกรอบเขตแดน โดยใช้ Non-Maximum Suppression (NMS) เพื่อลดกรอบเขตแดนที่ซ้อนทับกันและเก็บเฉพาะกรอบที่มีคะแนนสูงสุด

ขั้นที่ 4 การพยากรณ์คลาส YOLO จะพยากรณ์ความน่าจะเป็นที่กรอบเขตแดนจะเป็นคลาสวัตถุใด ๆ จากชุดข้อมูลที่ใช้ฝึก

จุดเด่นของ YOLOv7 จะประกอบไปด้วยด้านต่าง ๆ ดังนี้ ความเร็ว YOLOv7 มีการออกแบบให้ทำงานได้รวดเร็ว ซึ่งสามารถประมวลผลภาพในเวลาจริง (real-time) ได้ ความแม่นยำYOLOv7 มีการปรับปรุงให้มีความแม่นยำสูงขึ้นเมื่อเทียบกับเวอร์ชันก่อน ๆ ประสิทธิภาพในการจัดการวัตถุเล็ก YOLOv7 มีการปรับปรุงการตรวจจับวัตถุเล็ก ๆ และการจัดการหลายวัตถุในภาพ และความง่ายในการปรับใช้ YOLOv7 ได้รับการออกแบบให้ใช้งานง่ายและสามารถปรับใช้ในแอปพลิเคชันที่หลากหลายได้อย่างง่ายดาย

สถาปัตยกรรมของ YOLOv7 มีการปรับปรุงสถาปัตยกรรมจาก YOLO รุ่นก่อน ๆ โดยมีการใช้ Backbone ใหม่ที่เรียกว่า CSPDarknet ซึ่งมีการรวม Residual Blocks และ Cross Stage Partial Connections เพื่อเพิ่มความสามารถในการตรวจจับวัตถุ นอกจากนี้ยังมีการใช้ PANet และ SPP (Spatial Pyramid Pooling) เพื่อเพิ่มประสิทธิภาพในการจัดการหลายขนาดของวัตถุ ประกอบด้วยโมดูลหลักดังต่อไปนี้ Backbone โครงข่าย CNN แบบ Backbone ใหม่ มีประสิทธิภาพและความแม่นยำสูง Neck โมดูล Neck ประกอบด้วยโมดูล CSP และ PANet Head โมดูล Head ประกอบด้วยโมดูล Focus Module และการประมวลผลหลัง NMS และ Adaptive IoU Thresholding

Backbone ของ YOLOv7 ใช้โครงข่าย CNN แบบใหม่ที่มีประสิทธิภาพและความแม่นยำสูง โครงข่ายนี้ประกอบด้วยโมดูล Cross-Stage Partial connections (CSP) หลายโมดูล โมดูล CSP ช่วยให้โมเดลสามารถเรียนรู้คุณสมบัติของวัตถุจากหลายระดับ

Neck ของ YOLOv7 ประกอบด้วยโมดูล CSP และ PANet โมดูล CSP ช่วยให้โมเดลสามารถรวมคุณสมบัติจากหลายระดับ โมดูล PANet ช่วยให้โมเดลสามารถดึงคุณสมบัติที่เกี่ยวข้องกับวัตถุ

Head ของ YOLOv7 ประกอบด้วยโมดูล Focus Module โมดูล Focus Module ช่วยให้โมเดลสามารถทำนายคลาสและตำแหน่งของวัตถุ

YOLOv7 ใช้เทคนิคการประมวลผลหลังสองแบบเพื่อปรับปรุงประสิทธิภาพคือ Non-Maximum Suppression (NMS) ช่วยให้โมเดลสามารถลบการตรวจจับวัตถุที่ซ้อนทับกัน และAdaptive IoU Thresholding ที่ช่วยให้โมเดลสามารถปรับค่าเกณฑ์ IoU ให้เหมาะสมกับขนาดของวัตถุดังแสดงดังรูปที่ 3

รูปที่ 3 การตรวจจับภาพเรือด้วย YOLOv7

3.3 ขั้นตอนการตรวจจับภาพเรือด้วย YOLOv7

วิธีการตรวจจับภาพเรือด้วย YOLOv7[20] มีขั้นตอนดังนี้

ขั้นที่ 1 เตรียมข้อมูล ดาวน์โหลดโมเดล YOLOv7 เลือกโมเดลที่เหมาะสมกับความต้องการ เช่น โมเดลที่ผ่าน

การฝึกอบรมบนชุดข้อมูลเรือ เตรียมชุดข้อมูลภาพเรือ ภาพเรือที่ใช้จะมีหลากหลายประเภท มุมมอง แสง และสภาพแวดล้อม มีป้ายกำกับคลาสที่ถูกต้อง

ขั้นที่ 2 ประมวลผลภาพ โหลดภาพเรือที่ต้องการตรวจจับ แปลงขนาดภาพให้เป็นขนาดที่โมเดล YOLOv7 รองรับ และแปลงค่าสีของภาพให้เป็นรูปแบบที่โมเดล YOLOv7 รองรับ

ขั้นที่ 3 ตรวจจับวัตถุ ป้อนภาพที่ประมวลผลแล้วไปยังโมเดล YOLOv7 ประมวลผลภาพโดยโมเดล YOLOv7 จะประมวลผลภาพและทำนายวัตถุ ดึงผลลัพธ์จากโมเดล YOLOv7

ขั้นที่ 4 วิเคราะห์ผลลัพธ์ กรองผลลัพธ์ที่ไม่น่าเชื่อถือ เช่น คะแนนความมั่นใจต่ำ วาดกรอบรอบวัตถุ วาดกรอบรอบวัตถุที่ตรวจจับได้ และแสดงผลลัพธ์ เช่น คลาสของวัตถุ คะแนนความมั่นใจ ตำแหน่งของวัตถุดังแสดงดังรูปที่ 4 การตรวจจับภาพเรือด้วย YOLOv7

รูปที่ 4 การตรวจจับภาพเรือด้วย YOLOv7

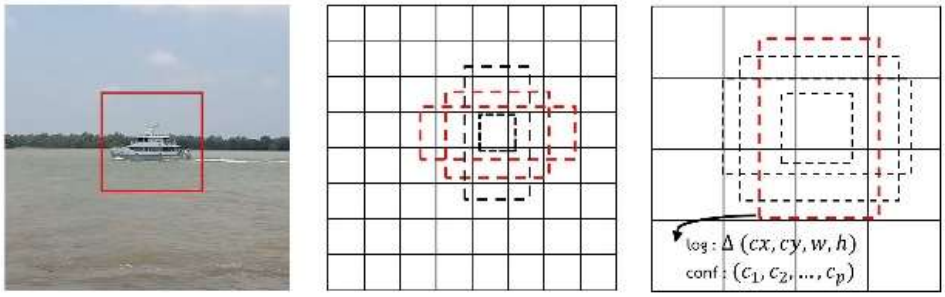
3.4 ขั้นตอนการตรวจจับภาพเรือด้วย SSD

SSD (Single Shot MultiBox Detector) [22] เป็นอัลกอริทึม Deep Learning ที่มีประสิทธิภาพและความเร็วสูง เหมาะสำหรับการตรวจจับวัตถุในภาพและวิดีโอ SSD ถูกนำไปใช้งานจริงอย่างแพร่หลาย เช่น การตรวจจับภาพเรือในแม่น้ำ ขั้นตอนการตรวจจับภาพเรือด้วย SSD มีดังนี้

ขั้นที่ 1 เตรียมชุดข้อมูล รวบรวมภาพถ่ายหรือวิดีโอที่มีภาพเรือในแม่น้ำเจ้าพระยา กำหนด Bounding Box ครอบเรือในแต่ละภาพ แบ่งชุดข้อมูลออกเป็น 3 ชุด คือ ชุดฝึกอบรม (Training Set) ชุดตรวจสอบ (Validation Set) และชุดทดสอบ (Test Set)

ขั้นที่ 2 ฝึกโมเดล SSD ดาวน์โหลดโมเดล SSD ปรับแต่งโมเดล SSD ให้เหมาะกับชุดข้อมูล ฝึกโมเดล SSD บนชุดข้อมูลฝึกอบรม ประเมินประสิทธิภาพของโมเดลบนชุดข้อมูลตรวจสอบ ปรับแต่งโมเดล SSD เพิ่มเติมหากจำเป็น

ขั้นที่ 3 ตรวจจับภาพเรือ โหลดโมเดล SSD ที่ได้รับการฝึกฝน ป้อนภาพถ่ายหรือวิดีโอให้กับโมเดล SSD โมเดล SSD จะค้นหารูปทรง (Bounding Box) ของเรือในภาพ และแสดงผลลัพธ์บนภาพถ่ายหรือวิดีโอดังแสดงดังรูปที่ 5



รูปที่ 5 การตรวจจับภาพเรือด้วย SSD

ในภาพเรือ SSD Framework จะแบ่งภาพออกเป็นกริด แต่ละเซลล์ของกริดจะทำนายวัตถุที่ปรากฏในเซลล์นั้น โมเดลจะสร้าง Default boxes หลายขนาดและรูปร่าง Default boxes เหล่านี้จะถูกประมวลผลด้วย CNN เพื่อดึงคุณลักษณะของภาพ จากนั้นโมเดลจะใช้คุณลักษณะเหล่านี้เพื่อทำนายวัตถุในภาพ โมเดลจะทำนายทั้งตำแหน่ง (Localization loss) และประเภท (Confidence loss) ของวัตถุ สุดท้ายโมเดลจะเลือก Default boxes ที่ดีที่สุด โดย Default boxes ที่ดีที่สุดนี้จะล้อมรอบเรือในภาพได้อย่างแม่นยำ

3.5 ขั้นตอนการตรวจจับภาพเรือด้วย Faster R-CNN

Faster R-CNN (Region-based Convolutional Neural Network) [23] เป็นโมเดลการตรวจจับวัตถุ (Object Detection) ยอดนิยมที่ใช้ Convolutional Neural Network (CNN) ในการระบุและจัดประเภทวัตถุในภาพ โมเดลนี้ทำงานได้รวดเร็วและแม่นยำ และถูกนำไปใช้กับงานตรวจจับวัตถุหลากหลายประเภท รวมถึงการตรวจจับภาพเรือ ขั้นตอนโดยละเอียดของการตรวจจับภาพเรือด้วย Faster R-CNN มีดังนี้

ขั้นที่ 1 เตรียมข้อมูล โดยรวบรวมภาพเรือ หารูปภาพเรือจำนวนมากที่มีความหลากหลายในแง่ของมุมมองแสงสว่างและประเภทของเรือ สร้างป้ายกำกับ (Annotation) กำหนดกรอบสี่เหลี่ยม (Bounding Boxes) รอบเรือในแต่ละภาพ ป้ายกำกับเหล่านี้จะใช้ในการฝึกโมเดล

ขั้นที่ 2 ฝึกโมเดล Faster R-CNN เตรียมโมเดล เลือกโมเดล Faster R-CNN ที่ผ่านการฝึกอบรมแล้ว หรือฝึกโมเดลใหม่จากจุดเริ่มต้น ป้อนข้อมูล ป้อนภาพเรือและป้ายกำกับที่สร้างขึ้นในขั้นตอนที่ 1 ไปยังโมเดล ฝึกโมเดลโดยปรับโมเดลให้เรียนรู้ที่จะระบุและจัดประเภทเรือในภาพโดยใช้ฟังก์ชันการสูญเสีย

ขั้นที่ 3 ตรวจจับภาพเรือ ป้อนภาพ ป้อนภาพใหม่ที่ต้องการตรวจจับเรือไปยังโมเดล ใช้โมเดลจะทำนายกรอบสี่เหลี่ยม (Bounding Boxes) รอบเรือที่ตรวจพบในภาพและคะแนนความมั่นใจสำหรับแต่ละการตรวจจับ

ขั้นที่ 4 วิเคราะห์ผลลัพธ์ ประเมินผลลัพธ์ วิเคราะห์ความถูกต้องของการตรวจจับโดยเปรียบเทียบผลลัพธ์กับป้ายกำกับที่แท้จริง (Ground Truth) ปรับโมเดล หากจำเป็น ปรับโมเดลเพิ่มเติมเพื่อปรับปรุงความถูกต้องของการตรวจจับ

สามารถแสดงขั้นตอนการทำงานของทั้ง 3 อัลกอริทึมได้ดังรูปที่ 6

รูปที่ 6 ขั้นตอนการทำงานของ YOLOv7, Faster R-CNN, และ SSD

จากรูปที่ 6 สามารถอธิบายแต่ละขั้นตอนได้ดังนี้

1) รับภาพถ่ายจากกล้องใช้กล้องจับภาพบริเวณแม่น้ำเจ้าพระยาความละเอียดของภาพ 608x608 พิกเซล จำนวนภาพทั้งหมด 1,000 ภาพภาพที่มีเรือ 750 ภาพ ภาพที่ไม่มีเรือ 250 ภาพภาพถ่ายต้องอยู่ในสภาพแสงปกติ ไม่มีการบดบังเรือ เช่น เงา

2) การเตรียมข้อมูล (Data Preparation) การเตรียมข้อมูลสำหรับระบบตรวจจับเรือ จะใช้ COCO Dataset เป็นฐานข้อมูลหลัก โดยเริ่มต้นด้วยการคัดกรองเฉพาะรูปภาพและ Annotation ที่มีวัตถุประเภท "เรือ" เท่านั้น เพื่อให้สอดคล้องกับการตรวจจับแบบ Binary Class (มีเพียง Class "เรือ" เดียว)

3) การประมวลผลล่วงหน้า (Preprocessing) Resize ปรับขนาดภาพเป็น 608x608 พิกเซล Color Normalization แปลงค่าสีของภาพ (เช่น RGB ⟶ BGR) Format Conversion แปลงภาพเป็น Tensor หรือ array ให้พร้อมป้อนเข้าโมเดล YOLOv7 ต้องการ input ที่มีขนาดสม่ำเสมอและ preprocessing ที่เหมาะสมเพื่อให้ผลลัพธ์แม่นยำ

4) การเลือกและฝึกอัลกอริทึมตรวจจับ (Model Training) ใช้โมเดล Deep Learning จำนวน 3 โมเดล ได้แก่ YOLOv7, Faster R-CNN, และ SSD

5) การตรวจจับวัตถุป้อนภาพเข้าโมเดล ได้ output เป็น Bounding Box รอบเรือ

6) การประเมินผล (Evaluation) ใช้ Confusion Matrix เพื่อวัดผลTrue Positive (TP) False Positive (FP) False Negative (FN) True Negative (TN) คำนวณค่า Precision, Recall และF1-Score

7) การแสดงผลลัพธ์และเปรียบเทียบแสดงภาพผลลัพธ์การตรวจจับ พร้อม bounding boxวิเคราะห์ความแม่นยำและความเร็วของแต่ละโมเดลเลือกโมเดลที่เหมาะสมที่สุดกับบริบทจริง

3.6 การวัดประสิทธิภาพด้วย confusion matrix

Confusion Matrix [24] เป็นเครื่องมือที่ใช้ประเมินประสิทธิภาพของโมเดลการจำแนกประเภท ช่วยให้เข้าใจว่าโมเดลทำนายผลผิดพลาดอย่างไร โดยแบ่งผลลัพธ์ออกเป็นสี่ประเภท

True Positive (TP) กรณีที่โมเดลทำนายถูกต้องว่าเป็นกลุ่มนั้น ๆ จำนวนภาพที่มีเรือจริง ๆ ที่โมเดลระบุว่ามีเรือ

False Positive (FP) กรณีที่โมเดลทำนายผิดพลาดว่าเป็นกลุ่มนั้น ๆ ทั้งที่จริงไม่ใช่ จำนวนภาพที่ไม่มีเรือจริง ๆ แต่โมเดลระบุว่ามีเรือ

True Negative (TN) กรณีที่โมเดลทำนายถูกต้องว่าไม่ใช่กลุ่มนั้น ๆ จำนวนภาพที่ไม่มีเรือจริง ๆ และโมเดลระบุว่าไม่มีเรือ

False Negative (FN) กรณีที่โมเดลทำนายผิดพลาดว่าไม่ใช่กลุ่มนั้น ๆ ทั้งที่จริงเป็น จำนวนภาพที่มีเรือจริง ๆ แต่โมเดลระบุว่าไม่มีเรือ

ค่า TP, FP, FN, และ TN เมื่อได้ค่ามาแล้วสามารถนำไปคำนวณประเมินประสิทธิภาพของโมเดลการจำแนกประเภทต่าง ๆ ได้ดังนี้

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

$$Recall = \frac{TP}{(TP + FN)} \quad (2)$$

$$F1\ score = 2 \times \frac{(Precision \times Recall)}{(Precision + Reca\ )} \quad (3)$$

โดยที่

Precision (ความแม่นยำเชิงบวก) ในบริบทของงานวิจัยนี้ หมายถึง สัดส่วนของภาพที่โมเดลระบุว่ามีเรือและเป็นจริงที่มีเรือ แสดงประสิทธิภาพของโมเดลในการระบุภาพที่มีเรือได้ถูกต้อง

Recall (ความไว) ในบริบทของงานวิจัยนี้ หมายถึง สัดส่วนของภาพที่มีเรือจริง ๆ ที่โมเดลระบุว่ามีเรือ แสดงประสิทธิภาพของโมเดลในการค้นหาภาพที่มีเรือทั้งหมด

F1-score คะแนน F1-score เป็นค่าเฉลี่ยฮาร์โมนิกของ Precision และ Recall แสดงประสิทธิภาพโดยรวมของโมเดล

4. ผลการวิจัย

งานวิจัยนี้มุ่งพัฒนาระบบตรวจจับเรือในแม่น้ำเจ้าพระยาโดยใช้เทคนิค YOLOv7 เปรียบเทียบประสิทธิภาพของอัลกอริทึมตรวจจับเรือ 2 ชนิด ได้แก่ SSD (Single Shot MultiBox Detector) และ Faster R-CNN (Region-based Convolutional Neural Network) บนชุดข้อมูลภาพถ่ายความละเอียด 608x608 พิกเซล จำนวน 1,000 ภาพ แบ่งเป็นภาพที่มีเรือ 750 ภาพ และภาพไม่มีเรือ 250 ภาพ ดังแสดงผลการทดลองตารางที่ 1

ตารางที่ 1 การประเมินประสิทธิภาพการตรวจจับเรือในแม่น้ำเจ้าพระยา

อัลกอริทึม	YOLO	Faster R-CNN	SSD
Precision	98.0%	97.0%	92.0%
Recall	92.0%	89.0%	74.0%
F1-Score	95.0%	93.0%	82.0%

รูปที่ 7 การตรวจจับภาพเรือด้วย YOLOv7

รูปที่ 8 การตรวจจับภาพเรือด้วย YOLOv7

รูปที่ 9 การตรวจจับภาพเรือด้วย Faster R-CNN

รูปที่ 10 การตรวจจับภาพเรือด้วย Faster R-CNN

รูปที่ 11 การตรวจจับภาพเรือด้วย SSD

รูปที่ 12 การตรวจจับภาพเรือด้วย SSD

5. สรุป

จากตารางการทดลองที่ 1 ในหัวข้อผลการวิจัยโดยการเปรียบเทียบประสิทธิภาพของ 3 อัลกอริทึม ได้แก่ YOLO (You Only Look Once), SSD (Single Shot MultiBox Detector) และ Faster R-CNN (Region-based Convolutional Neural Network) บนชุดข้อมูลภาพถ่ายจากกล้องที่มีความละเอียด 608x608 พิกเซล จำนวน 1,000 ภาพ แบ่งเป็นภาพที่มีเรือ 750 ภาพ และภาพไม่มีเรือ 250 ภาพ ประเมินประสิทธิภาพของโมเดลโดยใช้ confusion matrix สามารถสรุปผลการทดลองได้ดังนี้

5.1 ประสิทธิภาพในการตรวจจับ

อัลกอริทึม YOLOv7 มีประสิทธิภาพในการตรวจจับเรือในแม่น้ำเจ้าพระยาสูงสุด โดยมีค่า F1-Score สูงสุด YOLOv7 มีค่า F1-Score 95.0% ซึ่งสูงกว่า Faster R-CNN (93.0%) และ SSD (82.0%) มีค่า Precision และ Recall สูง YOLOv7 มีค่า Precision 98.0% และ Recall 92.0%

5.2 โมเดลที่เหมาะสม

จากผลการทดลองซึ่งบ่งชี้ว่าโมเดลอัลกอริทึม YOLOv7 สามารถระบุเรือได้อย่างถูกต้อง และเหมาะสำหรับการใช้งานจริงในการตรวจจับเรือเพื่อแก้ไขปัญหาการจราจรทางน้ำที่หนาแน่นและอันตรายในสภาวะแวดล้อมของงานวิจัยนี้

5.3 การอภิปรายผล

จุดเด่นของงานวิจัยนี้ คือความแม่นยำสูงของอัลกอริทึม YOLOv7 ที่ให้ค่า F1-Score สูงสุด (95.0%) แสดงให้เห็นว่าโมเดลสามารถตรวจจับเรือได้อย่างถูกต้องแม่นยำ เหมาะสำหรับการใช้งานจริง เมื่อเปรียบเทียบประสิทธิภาพของ 3 อัลกอริทึมบนชุดข้อมูลเดียวกัน ทำให้เลือกอัลกอริทึมที่เหมาะสมกับงานในสภาวะแวดล้อมนี้ได้ ผลการวิจัยแสดงให้เห็นถึงศักยภาพของ YOLOv7 ในการใช้งานจริงสำหรับการตรวจจับเรือในแม่น้ำเจ้าพระยา แต่ข้อจำกัดของงานวิจัยคือขนาดชุดข้อมูล โดยชุดข้อมูลที่ใช้ในการทดสอบมีขนาดค่อนข้างเล็ก (1,000 ภาพ) และความหลากหลายของข้อมูลที่จำกัดเพียงเรือที่ผ่านมาในช่วงที่เก็บข้อมูลเท่านั้น และในเรื่องการทดสอบงานวิจัยนี้ทดสอบอัลกอริทึมบนภาพถ่ายจากกล้องเท่านั้น ยังไม่มีการทดสอบบนวิดีโอหรือสภาพแสงที่แตกต่างกันซึ่งก็อาจได้ผลที่แตกต่างกันได้

YOLOv7 มีความสามารถโดดเด่นด้าน multi-scale feature extraction และ backbone ที่ทันสมัย (CSPDarknet + PANet + SPP) ซึ่งช่วยในการ

- จับลักษณะของเรือขนาดเล็กและหลากหลายชนิดได้ดี
- มีโมดูล Focus และ NMS ที่แม่นยำในการคัดกรอง bounding boxes
- รองรับ input ที่มีความละเอียด 608x608 พิกเซล ได้ดี ทำให้การจับตำแหน่งวัตถุแม่นยำ

ดังนั้นเมื่อจับคู่กับชุดข้อมูลที่มีลักษณะภาพค่อนข้างคงที่และเรือมีขนาดต่างกัน YOLOv7 จึงสามารถคาดการณ์ bounding box ได้แม่นยำ ส่งผลให้ precision สูงถึง 98%

5.4 ข้อแนะนำสำหรับงานวิจัยในอนาคต

ควรทดสอบ YOLOv7 กับชุดข้อมูลที่ใหญ่ขึ้นและหลากหลายมากขึ้น

ควรเปรียบเทียบ YOLOv7 กับอัลกอริทึมการตรวจจับวัตถุอื่น ๆ มากขึ้น

ควรพัฒนาโมเดล YOLOv7 เพิ่มเติมเพื่อให้สามารถตรวจจับเรือในสภาพต่าง ๆ ได้อย่างแม่นยำมากขึ้น

การทดลองในงานวิจัยนี้ ไม่ได้เน้นการแยกประเภท (Classification) ของเรือ โดยเน้นที่การตรวจจับว่ามีเรือหรือไม่เท่านั้น แต่อย่างไรก็ตาม หากต้องการนำไปแยกประเภทเรือในอนาคต สามารถต่อยอดด้วยโมเดล classification (เช่น ResNet, Inception) ร่วมกับ YOLO ได้ไม่ยาก

ในงานวิจัยนี้ไม่มีการเปรียบเทียบกับ YOLO เวอร์ชันอื่น ๆ แต่ในการเปรียบเทียบ YOLOv7 กับ YOLO เวอร์ชันอื่น (เช่น YOLOv5 หรือ YOLOX) จะช่วยยืนยันในการตรวจจับเรือในแม่น้ำเจ้าพระยาได้อย่างชัดเจน

ยิ่งขึ้น และยังทำให้เห็นความแตกต่างด้านประสิทธิภาพ ทำให้งานวิจัยมีความสมบูรณ์ น่าเชื่อถือทางวิชาการมากขึ้น อีกทั้งยังให้ข้อมูลเชิงลึกอันมีค่าเกี่ยวกับจุดแข็งและจุดอ่อนของแต่ละเวอร์ชัน ซึ่งจะเป็นประโยชน์สำหรับการพัฒนาเทคนิคการตรวจจับวัตถุในอนาคต

6. กิตติกรรมประกาศ

ข้าพเจ้าขอกราบขอบพระคุณ รศ. ดร. สมชาติ จิริวิภากร อาจารย์ที่ปรึกษา เป็นอย่างสูง ที่ได้ให้คำปรึกษา แนะนำแนวทาง และช่วยแก้ไขปัญหาต่าง ๆ ตลอดระยะเวลาการทำวิจัย นอกจากนี้ ข้าพเจ้ายังขอขอบคุณ น.ท. ผศ. ดร. พิศณุ คูมีชัย อาจารย์กองวิชาวิศวกรรมศาสตร์ ฝ่ายศึกษาโรงเรียนนายเรือ ที่เอื้อเฟื้ออุปกรณ์และเครื่องมือในการวิจัย แนะนำการเขียนโปรแกรม และให้คำปรึกษาจนงานวิจัยนี้สำเร็จลุล่วงไปด้วยดี ความสำเร็จของงานวิจัยนี้เกิดขึ้นได้ด้วยการสนับสนุนที่สำคัญจากอาจารย์ทั้งสองท่าน ข้าพเจ้าหวังว่างานวิจัยนี้จะเป็นประโยชน์ต่อผู้สนใจ และสามารถนำไปพัฒนาต่อยอดเพื่อสร้างผลงานใหม่ ๆ ต่อไปได้

7. เอกสารอ้างอิง

[1] Royal Customs Department, *Statistics on Smuggled Goods Seizure*, Bangkok, Thailand: Royal Customs Department, 2022.

[2] Marine Police Department, *Statistics on Illegal Vessel Seizure*, Bangkok, Thailand: Marine Police Department, 2022.

[3] Royal Thai Police, *Annual Report 2021*, Bangkok, Thailand: Royal Thai Police, 2021.

[4] Z. Chen, D. Chen, Y. Zhang, X. Cheng, M. Zhang and C. Wu, “Deep learning for autonomous ship-oriented small ship detection,” *Safety Sci.*, vol. 130, Art. no. 104812, 2020.

[5] P. Arumugam and N. Natarajan, “A survey of face detection, extraction and recognition,” *Int. J. Comput. Appl.*, vol. 163, no. 7, pp. 1–6, 2017.

[6] W. Liu *et al*., “SSD: Single Shot MultiBox Detector,” *arXiv preprint*, arXiv: 1512.02325, 2016.

[7] J. Redmon, S. Divakar, A. Sivaram and R. Girshick, “You Only Look Once: Unified, real-time object detection,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 9, pp. 1656–1667, 2016.

[8] S. Ren, K. He, R. Girshick and J. Sun, “Faster R-CNN: Towards real-time object detection with region proposal networks,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, 2016.

[9] J. Dai, Y. Li, K. He and J. Sun, “R-FCN: Fully convolutional networks for object detection,” *arXiv preprint*, arXiv: 1605.06409, 2016.

[10] Z. Cai, N. Vasconcelos and M. Li, “Cascade R-CNN: Delving into high-resolution feature representations for object detection,” *arXiv preprint*, arXiv: 1712.00726, 2018.

[11] C. Fu, W. Liu, D. Anguelov and C. Szegedy, “SSD: Single-Shot MultiBox Detector,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 9, pp. 2163–2177, 2017.

[12] S. Meng, P. Sun and L. Lin, “Two-Stage SSD for real-time object detection,” *arXiv preprint*, arXiv: 1805.14173, 2018.

[13] J. Redmon, S. Divakar, A. Sivaram and R. Girshick, “You Only Look Once: Unified, real-time object detection,” *IEEE Trans.*

Pattern Anal. Mach. Intell., vol. 38, no. 9, pp. 1656–1667, 2016.

[14] Z. Ge, S. Deng and X. Li, "YOLOX: Scalable and efficient object detection with cross-scale stitching," *arXiv preprint*, arXiv: 2304.14979, 2024.

[15] H. Hu, W. Jiang, Z. Qin and J. Tang, "A real-time ship detection method based on improved YOLOv7 and Kalman filter," *Sensors*, vol. 22, no. 1, p. 142, 2022.

[16] S. Xie, M. Li, Q. Zhang and Z. Shao, "A ship detection method based on YOLOv7 and improved attention mechanism," *J. Coast. Res.*, vol. 149, pp. 1–8, 2022.

[17] Y. Wang, C. Zhang and Z. Xu, "A real-time ship detection method based on YOLOv7 and lightweight network," *Appl. Sci.*, vol. 12, no. 1, p. 362, 2022.

[18] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1929–1958, 2014.

[19] Y. LeCun, Y. Bengio and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[20] C. Y. Wang, A. Bochkovskiy and H. Y. M. Liao, "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors," *arXiv preprint*, arXiv:2207.02696, 2022.

[21] A. Bochkovskiy, C. Y. Wang and H. Y. M. Liao, "YOLOv4: Optimal speed and accuracy of object detection," *arXiv preprint*, arXiv:2004.10934, 2020.

[22] W. Liu *et al*., "SSD: Single-shot MultiBox Detector," *arXiv preprint*, arXiv: 1603.03991, 2016.

[23] S. Ren *et al*., "Faster R-CNN: Towards real-time object detection with region-based convolutional neural networks," *arXiv preprint*, arXiv: 1506.01497, 2015.

[24] J. Li, X. Zhang and H. Wang, "A survey of deep learning-based ship detection methods," *Remote Sens.*, vol. 13, no. 11, pp. 1–24, 2021.