Science and Engineering Connect, ISSN 3027-7914 (Online)
Vol. 47 No. 1 pp. 62-77 (2024)

บทความวิจัย (Research Article)

ระบบตรวจจับยานพาหนะที่ละเมิดทางม้าลายด้วยการจอดรถผิดกฎหมาย
Detection System for Vehicles Violating Pedestrian Crosswalks through Illegal Parking

วัชรวิทย์ ลาพินี, ธนัตถ์กรณ์ ชื่นบรรลือสุข, วริศรา บุญยวง, วิศัลย์ศยา ทิพย์โรจน์, ภูมิพิชญ์ เอี๊ยบทวี, ฐิติรัตน์ ศิริบวรรัตนกุล*
Watcharavit Lapinee, Thanatkorn Chuenbanluesuk, Varisara Boonyuang, Wisansaya Tipparot, Poompich Eiabtawee, Thitirat Siriborvornratanakul*
คณะสถิติประยุกต์ สถาบันบัณฑิตพัฒนบริหารศาสตร์ กรุงเทพฯ ประเทศไทย
Graduate School of Applied Statistics, National Institute of Development Administration, Bangkok, Thailand
*Corresponding author E-mail: thitirat@as.nida.ac.th

Received 12 July 2022; Revised 19 March 2024; Accepted 21 March 2024

บทคัดย่อ

ความเป็นมาและวัตถุประสงค์ : การบังคับใช้กฎหมายอย่างจริงจังและมีประสิทธิภาพมีบทบาทสำคัญที่จะช่วยให้ผู้ขับขี่ตระหนักและรับผิดชอบต่อผู้คนบนท้องถนนมากขึ้น นำไปสู่การลดลงของอุบัติเหตุและความสูญเสียที่จะเกิดขึ้นต่อชีวิตและทรัพย์สิน งานวิจัยนี้จึงทำการพัฒนาระบบอัตโนมัติสำหรับตรวจจับยานพาหนะที่ทำผิดกฎจราจรบนทางม้าลาย โดยระบบจะตรวจจับยานพาหนะและนับจำนวนผู้ขับขี่ที่จอดพาหนะทับทางม้าลายในขณะที่รถจอดติดไฟแดง

วิธีดำเนินการวิจัย : ข้อมูลที่ใช้ในงานวิจัยเก็บจากกล้องวงจรปิดของเว็บไซต์สำนักการจราจรและขนส่งจำนวน 12 จุด รวมทั้งหมด 334 รูป ข้อมูลดังกล่าวถูกแบ่งออกเป็น 2 ส่วน ได้แก่ ข้อมูลสำหรับฝึกสอน 301 รูป และสำหรับทดสอบ 33 รูป เทคนิคที่ใช้คือ YOLOv5 เพื่อตรวจจับวัตถุ อันได้แก่ ยานพาหนะและทางม้าลาย นอกจากนี้ เพื่อให้การตรวจจับและการนับจำนวนเป็นไปอย่างถูกต้องในวิดีโอที่เป็นภาพต่อเนื่องจากกล้องวงจรปิด จึงใช้เทคนิคติดตามวัตถุในภาพ ได้แก่ วิธี DeepSORT และวิธี Manual ประกอบกันไปด้วย

ผลการวิจัย : การใช้เทคนิค YOLOv5 เพื่อตรวจจับรถยนต์ รถจักรยานยนต์ และทางม้าลาย ได้ค่า Average Precision เท่ากับ 0.837 นอกจากนี้ ในส่วนของการติดตามวัตถุนั้น ผู้วิจัยพบว่าวิธี DeepSORT ให้ผลลัพธ์การติดตามวัตถุที่ถูกตรวจจับโดย YOLOv5 ดีกว่าวิธี Manual โดยสามารถติดตามและนับจำนวนยานพาหนะที่จอดทับทางม้าลายเกินกว่าระยะเวลาที่กำหนดได้อย่างมีประสิทธิภาพ

สรุป : งานวิจัยนี้เสนอการใช้ขั้นตอนวิธีอัตโนมัติในการวิเคราะห์ภาพต่อเนื่องจากวิดีโอกล้องวงจรปิดเพื่อตรวจจับยานพาหนะที่ทำผิดกฎจราจรโดยการจอดทับทางม้าลายระหว่างรถติดไฟแดง เทคนิคตรวจจับวัตถุYOLOv5

สามารถตรวจจับยานพาหนะและทางม้าลายในภาพได้ดี แต่ยังจำเป็นต้องเสริมด้วยเทคนิคการติดตามวัตถุในภาพ เพื่อไม่ให้เกิดการตรวจจับพาหนะซ้ำซ้อนกันในภาพต่อเนื่องของกล้องวงจรปิด

การนำไปใช้ประโยชน์ในเชิงปฏิบัติ : ระบบที่นำเสนอในงานวิจัยนี้สามารถนำไปต่อยอดเพื่อตรวจจับผู้ขับขี่ที่จอดพาหนะทับทางม้าลายในขณะที่รถจอดติดไฟแดงได้โดยอัตโนมัติ ทั้งนี้ อาศัยข้อมูลภาพจากกล้องวงจรปิดที่มีการติดตั้งอยู่แล้วและเพิ่มเติมขั้นตอนวิธีการวิเคราะห์ภาพเสริมเข้าไปเท่านั้น

คำสำคัญ : การตรวจจับทางม้าลาย, การตรวจจับยานพาหนะ, การติดตามตำแหน่งของวัตถุ

Abstract

Background and Objectives: Strict and effective enforcement of laws plays a crucial role in increasing driver awareness and responsibility towards people on the roads. This could in turn lead to a reduction in accidents and losses to life and properties. The present research therefore developed an automated system for detecting vehicles violating traffic rules on pedestrian crosswalks. The system detects vehicles and counts the number of drivers parking their vehicles on pedestrian crosswalks while waiting at red lights.

Methodology: Data for the study were collected from 12 surveillance camera sites of the Department of Traffic and Transportation, totaling 334 images. The data were divided into two parts: training data consisting of 301 images and testing data of 33 images. The YOLOv5 technique was employed to detect such objects as vehicles and pedestrian crosswalks. Additionally, to ensure accurate detection and counting in continuous video images from surveillance cameras, two alternatives viz. DeepSORT object tracking and manual methods were experimented.

Main Results: Using the YOLOv5 technique to detect cars, motorcycles, and pedestrian crosswalks resulted in an Average Precision score of 0.837. In terms of object tracking, the DeepSORT method outperformed the manual method in effectively tracking and counting vehicles parked over pedestrian crosswalks beyond the designated time frame.

Conclusions: The present research proposes the use of automated vision analysis from continuous video surveillance to detect vehicles violating traffic laws by parking over pedestrian crosswalks while waiting at red lights. The study reveals that while the YOLOv5

technique effectively detects vehicles and pedestrian crosswalks in images, it requires an additional object tracking technique to prevent duplicate vehicle detection in continuous camera images.

Practical Application: The system investigated in the present research can be further developed to automatically detect drivers parking over pedestrian crosswalks while waiting at red lights. This can be accomplished using existing surveillance camera image data, with simple addition of an enhanced image analysis technique.

Keywords: Crosswalk Detection, Vehicle Detection, Object Tracking

Introduction

อุบัติเหตุรถชนคนเดินเท้าเป็นหนึ่งในเหตุการณ์ที่พบได้บ่อย อีกทั้งลักษณะของการชนมักทำให้เกิดความรุนแรงสูง จากสถิติอุบัติเหตุที่เกิดขึ้นบนถนนของกรมทางหลวงในช่วง พ.ศ. 2556-2560 พบว่ามีผู้เสียชีวิตเฉลี่ยสูงถึง 55 รายต่ออุบัติเหตุ 100 ครั้ง โดยสาเหตุที่ทำให้เกิดอุบัติเหตุเกิดจากการขับรถเร็วเกินอัตรากำหนดการฝ่าฝืนสัญญาณไฟและเครื่องหมายจราจร [1] ข้อมูลจากกระทรวงสาธารณสุขใน พ.ศ. 2559 - 2564 พบว่าสถิติคนเดินเท้าประสบเหตุถึง 224,068 คน โดยร้อยละ 42.26 มีรถจักรยานยนต์เป็นคู่กรณีที่ทำให้เกิดอุบัติเหตุชนคนข้ามถนนหรือเกิดการบาดเจ็บกับคนเดินเท้า [2]

ในประเทศที่ให้ความสำคัญกับคนเดินเท้าและคนข้ามถนน อาทิ ประเทศญี่ปุ่น จะมีกฎหมาย [3] ที่บังคับให้รถยนต์ซึ่งขับขี่ผ่านมาจะต้องหยุดรถทันทีเพียงแค่คนที่ต้องการข้ามถนนยืนอยู่ในจุดที่รอข้ามถนน แต่สำหรับในประเทศไทยนั้น แม้ว่าสัญญาณไฟจราจรสำหรับคนข้ามจะเปลี่ยนเป็นสีเขียวแล้ว แต่ผู้ขับขี่จำนวนหนึ่งก็ยังไม่หยุดรถหรือแม้แต่มีการจอดรถขวางทับเส้นทางม้าลายเอาไว้ สำหรับพื้นที่กรุงเทพมหานครนั้นมีจำนวนทางม้าลายมากถึง 1,112 จุด หลายจุดเป็นพื้นที่ “จับ-ปรับ” เช่น แยกอโศกมนตรี แต่กลับพบการกระทำผิดมากกว่า 25,000 ครั้ง ในเดือนมกราคม พ.ศ. 2565 [4]

เพื่อแก้ปัญหาดังกล่าว ผู้วิจัยเชื่อว่าการบังคับใช้กฎหมายอย่างจริงจังและมีประสิทธิภาพจะสามารถเข้ามามีบทบาทสำคัญช่วยให้ผู้ขับขี่มีความตระหนักและความรับผิดชอบต่อผู้คนบนท้องถนนมากขึ้น นำไปสู่การลดลงของอุบัติเหตุและลดความสูญเสียที่จะเกิดขึ้นต่อชีวิตและทรัพย์สิน ในปัจจุบันมีการนำเสนอข่าวโดยระบุว่ากรุงเทพมหานครนำเทคโนโลยีปัญญาประดิษฐ์มาใช้ตรวจจับผู้ที่ฝ่าฝืนกฎจราจรบริเวณทางม้าลายแยกอโศก [4] เพื่อตรวจจับผู้กระทำความผิดบนพื้นที่สีแดงเพียงเท่านั้น ซึ่งข่าวดังกล่าวยังไม่มีการตีพิมพ์เผยแพร่ทางวิชาการหรือมีการเผยแพร่วิธีการที่ใช้อย่างเป็นรูปธรรม ในการนี้คณะผู้วิจัยจึงมีเป้าหมายในการพัฒนาระบบต้นแบบสำหรับตรวจจับผู้กระทำผิดกฎจราจรและนับจำนวนผู้กระทำผิด โดยในเบื้องต้นจะมุ่งเน้นไปที่การตรวจจับยานพาหนะที่มีการจอดทับทางม้าลายในขณะที่รถจอดติดไฟแดง ทั้งนี้โดยอาศัยข้อมูลจากกล้องวงจรปิดที่มีอยู่แล้ว

ตามแยกต่าง ๆ ไม่จำกัดเฉพาะแต่ทางม้าลายที่เป็นสีขาวสลับแดงแต่ยังรวมถึงทางม้าลายที่มีสีขาวสลับดำได้อีกด้วย เพื่อให้ผลลัพธ์ที่ได้สามารถนำไปปรับใช้กับทางม้าลายบริเวณแยกไฟแดงอื่น ๆ ในกรุงเทพมหานคร

Related Works

จากการทบทวนวรรณกรรมพบว่าการตรวจจับการละเมิดทางม้าลายเพื่อตรวจจับผู้ที่ทำผิดกฎจราจรนั้นประกอบด้วย 4 ส่วน ได้แก่ 1. การตรวจจับทางม้าลาย 2. การตรวจจับยานพาหนะ 3. การติดตามตำแหน่งของวัตถุ และ 4. การตรวจจับการจอดรถทับทางม้าลาย ในส่วนของการประเมินและวัดผลระบบลักษณะนี้ได้ถูกสรุปไว้ในข้อ 5

1. Pedestrian crosswalk detection (การตรวจจับทางม้าลาย) การตรวจจับทางม้าลายเป็นการระบุว่าพื้นที่บริเวณหนึ่ง ๆ เป็นทางม้าลายสำหรับคนข้ามถนนหรือไม่ ซึ่งมักต้องใช้การวิเคราะห์รูปภาพเข้ามาเกี่ยวข้อง เทคนิคที่พบในส่วนนี้มักมีอยู่ 2 แบบ ได้แก่ การตรวจจับทางม้าลายโดยใช้รูปจากกล้องวงจรปิดและการตรวจจับทางม้าลายโดยใช้กล้องหน้ารถ งานของ Ibadov และคณะ [5] เสนอวิธีตรวจจับทางม้าลายโดยใช้รูปจากกล้องวงจรปิด โดยนำรูปมาทำเป็นรูปขาว-ดำ (Binary mask) ทำให้สามารถแยกทางม้าลายออกมาได้ ในงานของ Zhang และคณะ [6] ศึกษาการตรวจจับทางม้าลายจากภาพของกล้องหน้ารถในสภาพอากาศที่แตกต่างกัน โดยมีการจำลองรูปภาพเสมือนของหมอกหนาเพื่อให้โมเดลสามารถตรวจจับทางม้าลายในสภาพอากาศที่ผิดปกติได้ดีขึ้น ผลการทดลองพบว่าการใช้เทคนิค CDNet มีความแม่นยำและความเร็วในการตรวจจับที่สูงกว่าการใช้ YOLO v5 ในกรณีที่พบว่าทางม้าลายไม่ชัดเจน ซีดจาง หรือเส้นขอบของทางม้าลายขาดหาย ทำให้ระบบอัตโนมัติไม่อาจแน่ใจได้ว่าภาพที่ปรากฏเป็นทางม้าลายหรือไม่นั้น งานของ Malbog [7] ทดลองตรวจจับทางม้าลายด้วยเทคนิค Mask R-CNN และได้ความแม่นยำถึง 97% ในขณะที่งานของ Romić และคณะ [8] ได้ทำการศึกษาลักษณะรูปแบบของทางม้าลายและทำนายทางม้าลายจากรูปภาพกว่า 300 รูป โดยใช้วิธี Column และ Row Structure Analysis พบว่ามี precision และ recall สูงถึง 98.7% ซึ่งจะช่วยให้ทำนายว่าเป็นทางม้าลายได้ดีขึ้น

จากการทบทวนวรรณกรรม ผู้วิจัยพบว่าสำหรับการตรวจจับทางม้าลายภายใต้สภาพแวดล้อมจริงที่มีตัวแปรที่ไม่อาจควบคุมได้หลายตัว อาทิ สภาพแสง หมอกควัน ความสมบูรณ์ของลายเส้นทางม้าลาย เทคนิคซึ่งเป็นที่นิยมคือเทคนิคการตรวจจับวัตถุ (Object detection) ที่ใช้สถาปัตยกรรมของการเรียนรู้เชิงลึก (Deep learning) ซึ่งถูกใช้อยู่ทั้งใน CDNet, YOLO, และ Mask R-CNN ด้วยเหตุนี้งานวิจัยชิ้นนี้จึงจะเลือกนำเทคนิคดังกล่าวมาใช้ในงานด้วย โดยเป้าหมายคือการใช้เทคนิคดังกล่าวในการพยากรณ์หา Bounding box กรอบสี่เหลี่ยมที่ทำหน้าที่ระบุว่าพื้นที่บริเวณใดในภาพคือพื้นที่ของทางม้าลาย

2. Vehicle detection (การตรวจจับยานพาหนะ) การตรวจจับยานพาหนะเป็นการระบุว่าในพื้นที่หนึ่ง ๆ นั้นมีรถยนต์ รถจักรยานยนต์ หรือรถโดยสารปรากฏอยู่หรือไม่ จากการทบทวนวรรณกรรมผู้วิจัยพบว่างานในอดีตมีการนำเทคนิคการเรียนรู้เชิงลึกมาช่วยจำแนกประเภทของยานพาหนะที่อยู่บนท้องถนนอย่างแพร่หลาย แม้ว่ายานพาหนะที่ปรากฏจะมีทั้งยี่ห้อ ประเภทของยานพาหนะ และสีที่แตกต่างกัน อาทิ การศึกษาของ Tonge และคณะ [9] ที่จำแนกรถยนต์และรถจักรยานยนต์ด้วยเทคนิค YOLO (you only look

once) การศึกษาของ Ibadov และคณะ [5] ที่ตรวจจับรถยนต์ด้วยเทคนิค Faster R-CNN การศึกษาของ Ibadov และคณะ [10] และการศึกษาของ Prasanthi และคณะ [11] ที่ศึกษาการตรวจจับยานพาหนะด้วย YOLOv3 และพบว่าสามารถประมวลผลได้เร็วกว่า R-CNN และ Fast R-CNN อีกทั้งมีค่าเฉลี่ยของความแม่นยำอยู่ที่ 85% และมีการใช้ ORB ในการเก็บข้อมูลทิศทางของยานพาหนะและนับจำนวนยานพาหนะได้อีกด้วย นอกจากนี้งานของ Navastara และคณะ [12] ได้ทดลองจำแนกยานพาหนะจากภาพกล้องวงจรปิด โดยใช้ Gaussian Mixture Model (GMM) ตรวจจับยานพาหนะและจำแนกประเภทของยานพาหนะด้วยการใช้ Hierarchical Multi-SVM ผลที่ได้มีความแม่นยำ 80.28% และ F-score 82.76% ในขณะที่ Pan และ Wang [13] ทำการตรวจจับยานพาหนะจากวิดีโอของกล้องวงจรปิดที่ถูกติดตั้งอยู่ตามถนน Freeway ของประเทศไต้หวัน โดยใช้ YOLOv3 ปัญหาที่พบได้แก่ ความแตกต่างกันของวัตถุที่ต้องตรวจจับ มุมกล้อง พื้นที่ยานพาหนะและความเร็วของยานพาหนะ รวมถึงช่วงเวลาที่มีผลต่อแสงในรูป นอกจากนี้รูปที่ได้จากกล้องวงจรปิดก็มีความคมชัดต่ำ ผลลัพธ์ของการศึกษาพบว่าความแม่นยำในช่วงกลางวันคือ 44% ในขณะที่ช่วงเวลากลางคืนมีความแม่นยำลดลงเหลือ 41%

นอกจากนี้ Zhu และคณะ [14] ได้เสนอวิธีการหาตำแหน่งของยานพาหนะในมุมมอง 3 มิติจากรูปภาพของกล้องจราจร โดยทำการแปลงรูปภาพให้เป็นรูปภาพในมุมมอง Bird’s eye view ด้วยวิธี Planar homography และใช้ YOLOv3 ในการตรวจจับยานพาหนะจากรูปภาพในมุมมองดังกล่าวซึ่งจะมีความบิดเบี้ยวของรูปภาพมากกว่ารูปปกติ ผลการทดลองพบว่า เมื่อเปรียบเทียบประสิทธิภาพกับโมเดลอื่น ๆ บนชุดข้อมูล BrnoCompSpeed ผลลัพธ์ของงานนี้ได้ความแม่นยำที่ดีขึ้น ในส่วนของการศึกษาโดย Shepelev และคณะ [15] ได้ศึกษาลักษณะเฉพาะของสี่แยก 20 แยกที่สำคัญของเมืองเชเลียบินสค์ โดยใช้ภาพจากกล้องวิดีโอจราจรมาวิเคราะห์ปริมาณยานพาหนะแบบทันทีทันใดด้วยเทคนิค R-CNN, YOLOv3, และ SORT ในการตรวจจับวัตถุ แบ่งประเภทของยานพาหนะ และติดตามวัตถุในวิดีโอ โดยมีจุดประสงค์เพื่อจัดระเบียบการจราจรและสัญญาณไฟจราจรของเมือง ผลการทดลองพบว่าระบบดังกล่าวสามารถช่วยลดระยะเวลาในการรอสัญญาณไฟจราจรสำหรับคนข้ามถนนได้ถึง 30% เมื่อเทียบกับระบบเดิม

จากการทบทวนวรรณกรรมจะเห็นว่า งานการตรวจจับยานพาหนะบนท้องถนนนั้นนิยมใช้แบบจำลองการเรียนรู้เชิงลึกในกลุ่มของ YOLO, R-CNN และ Fast R-CNN โดยปัญหาที่พบจากงานของ Pan และ Wang [13] คือ ความแตกต่างของวัตถุที่ต้องตรวจจับ มุมกล้อง พื้นที่ ยานพาหนะและความเร็วของยานพาหนะ รวมถึงช่วงเวลาที่มีผลต่อแสงในรูป และความคมชัดของกล้องวงจรปิด เนื่องจากรูปที่คณะผู้วิจัยจะนำมาใช้ในงานวิจัยนี้ก็เป็นรูปจากกล้องวงจรปิดเช่นกัน จึงมีความเป็นไปได้สูงที่จะพบปัญหาเรื่องความคมชัดของภาพจากกล้องวงจรปิดและความเร็วในการตรวจจับรูป ด้วยเหตุนี้คณะผู้วิจัยจึงจะคัดเลือกรูปจากกล้องวงจรปิดมาหลายภาพ (ให้ระบบส่งรูปเข้ามารูปละ 1 วินาที) และเลือกใช้เทคนิค YOLO ในการตรวจจับยานพาหนะเนื่องจากเป็นเทคนิคที่มีชื่อเสียงในด้านความเร็วของการประมวลผล ผลลัพธ์ที่ต้องการคือระบบที่จะพยากรณ์กรอบสี่เหลี่ยม Bounding box ที่ระบุว่าพื้นที่ใดในภาพบ้างที่เป็นภาพของยานพาหนะ

3. Object tracking (การติดตามตำแหน่งของวัตถุ) การติดตามตำแหน่งของวัตถุเป็นการระบุว่าวัตถุที่

สามารถตรวจจับได้ในภาพต่อเนื่อง (วิดีโอ) นั้น เป็นวัตถุชิ้นเดิมในแต่ละเฟรมของภาพหรือไม่ อาทิ การศึกษาของ Hou และคณะ [16] ทดลองติดตามตำแหน่งของรถยนต์ด้วยเทคนิค DeepSORT เช่นเดียวกับการศึกษาของ Parico และ Ahamed [17] ที่ตรวจจับลูกแพร์ด้วย YOLOv4 และนับจำนวนจากการติดตามตำแหน่งด้วยเทคนิค DeepSORT นอกจากนี้ยังมีงานของ Santos และคณะ [18] ที่เสนอวิธีการนับจำนวนรถยนต์บนท้องถนนในประเทศบราซิล โดยอาศัยการจำแนกรถยนต์ด้วย YOLOv3 และนับจำนวนของรถยนต์ที่สามารถจำแนกได้ด้วยการติดตามตำแหน่งของรถยนต์จากเทคนิค DeepSORT

4. Detection of parking vehicles over pedestrian crosswalks (การตรวจจับการจอดรถทับทางม้าลาย) ในเชิงของการวิเคราะห์ภาพนั้น การตรวจจับการจอดรถทับทางม้าลายสามารถอาศัยการพิจารณาจากตำแหน่งของทางม้าลายเปรียบเทียบกับตำแหน่งของยานพาหนะว่ามีการซ้อนทับกันอย่างมีนัยสำคัญหรือไม่ อาทิ การศึกษาของ Tonge และคณะ [9] ที่ตรวจจับยานพาหนะที่ทำผิดกฎจราจรโดยใช้รูปที่มาจากกล้องวงจรปิดซึ่งถูกถ่ายไว้ขณะไฟแดงทุก ๆ 15 วินาที (เพื่อป้องกันไม่ให้รูปที่ถ่ายออกมาซ้ำกัน) โดยหลังจากที่ตรวจจับยานพาหนะและทางม้าลายแยกกันได้แล้ว ก็นำ Bounding box ที่สร้างล้อมรอบยานพาหนะและทางม้าลายมาพิจารณาว่ามีการซ้อนทับกันหรือไม่

5. Evaluation (การประเมินผล) จากการทบทวนวรรณกรรมพบว่าวิธีการประเมินผลในงานการตรวจจับวัตถุ (Object detection) นิยมใช้ค่า Average Precision (mAP) ซึ่งประกอบด้วย 4 ตัวชี้วัดย่อย ได้แก่ Intersection over Union (IoU), confusion matrix, precision และ recall ดังรายละเอียดต่อไปนี้ Intersection over Union (IoU) เป็นการตรวจสอบระดับการซ้อนทับกันระหว่าง Bounding box ที่เป็นผลลัพธ์ของการทำนายและ Ground truth box ที่เป็นคำเฉลยที่ถูกต้องดังสมการที่ 1

$$\text{Iou} = \frac{\text{Area of Overlap}}{\text{Area of Union}} \tag{1}$$

Confusion Matrix เป็นการประเมินผลของการจำแนกประเภทของวัตถุ (Classification) โดยคำนวณจากค่า IoU และผลการจำแนกประเภทของการตรวจจับวัตถุ ซึ่งผลการจำแนกประเภทนั้นสามารถแบ่งออกได้เป็น 4 กรณี คือ (1) กรณี True Positives (TP) หรือกรณีที่โมเดลทำนายแล้วว่าได้กรณี positives และทำนายผลลัพธ์ตรงกับความเป็นจริง (2) กรณี True Negatives (TN) หรือกรณีที่โมเดลทำนายแล้วว่าได้กรณี negatives และทำนายผลลัพธ์ตรงกับความเป็นจริง (3) กรณี False Positives (FP) หรือกรณีที่โมเดลทำนายแล้วว่าได้กรณี positives แต่ทำนายผลลัพธ์ไม่ตรงกับความเป็นจริง (Type I Error) และ (4) กรณี False Negatives (TN) หรือกรณีที่โมเดลทำนายแล้วว่าได้กรณี negatives แต่ทำนายผลลัพธ์ไม่ตรงกับความเป็นจริง (Type II Error)

Precision คือ ค่าความแม่นยำ คำนวณโดยใช้สูตรดังสมการที่ 2

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{2}$$

Recall คือ ค่าความไวคำนวณโดยใช้สูตรดังสมการที่ 3

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (3)$$

นอกจากนี้ยังจะมีการใช้การคำนวณพื้นที่ใต้กราฟระหว่าง precision - recall และหาค่า mAP โดยใช้สูตรรายละเอียดดังสมการที่ 4 ด้วย เมื่อ mAP หมายถึง ผลลัพธ์จากการคำนวณค่า Average Precision, AP หมายถึง ค่าเฉลี่ยของ precision ในแต่ละประเภท และ N หมายถึง จำนวนประเภทของวัตถุ

$$\text{mAP} = \frac{1}{N} \sum_{i=1}^{N} \text{AP}i \quad (4)$$

จากการทบทวนวรรณกรรมในหัวข้อนี้คณะผู้วิจัยได้ข้อสรุปว่า สำหรับการสร้างระบบตรวจจับยานพาหนะที่มีการละเมิดทางม้าลายและนับจำนวนผู้กระทำความผิดในขณะที่รถจอดติดไฟแดงโดยใช้ข้อมูลจากกล้องวงจรปิดที่มีอยู่แล้วตามสี่แยกต่าง ๆ นั้น คณะผู้วิจัยจะนำรูปจากกล้องวงจรปิดมาประมวลผลทุก ๆ 1 วินาทีซึ่งเป็นความถี่ที่ไม่มากเกินไปสำหรับการประมวลผลแต่ก็ไม่น้อยเกินไปจนอาจทำให้พลาดจังหวะสำคัญในวิดีโอ ในส่วนของการตรวจจับวัตถุและสร้าง Bounding box นั้น ผู้วิจัยเลือกใช้เทคนิค YOLO เนื่องจากเป็นวิธีที่สามารถประมวลผลได้รวดเร็วอีกทั้งมีประสิทธิภาพที่ดีในการตรวจจับทั้งในส่วนของทางม้าลายและในส่วนของยานพาหนะ (อ้างอิงจากการทบทวนวรรณกรรม)

เพื่อเพิ่มความต่อเนื่องและความแม่นยำในการตรวจจับและนับวัตถุในภาพต่อเนื่องอย่างวิดีโอ เทคนิคการติดตามของ DeepSORT [19] จะถูกนำมาใช้ร่วมด้วยในงานนี้ และในขั้นตอนสุดท้ายของการตรวจสอบว่ามีการจอดรถทับทางม้าลายหรือไม่ คณะผู้วิจัยจะพิจารณาจากเกณฑ์ที่ว่าหากผู้ขับขี่มีการจอดรถทับทางม้าลายเกิน 5 วินาทีให้ระบบทำการเก็บข้อมูลและนับเป็นผู้กระทำผิดกฎหมายทางม้าลาย โดยการพัฒนาระบบอัตโนมัติลักษณะนี้จะช่วยให้การส่งข้อมูลให้เจ้าหน้าที่ตำรวจเป็นไปได้อย่างรวดเร็วและทันทีทันดี สามารถอำนวยความสะดวกให้กับเจ้าหน้าที่ตำรวจ และเสริมความปลอดภัยให้กับคนเดินเท้า ช่วยลดอุบัติเหตุรถชนคนเดินเท้าได้ในระยะยาว

Research Methodology

การเก็บรวบรวมข้อมูล

คณะผู้วิจัยทำการเก็บรูปภาพจากกล้องวงจรปิดบนเว็บไซต์ของสำนักการจราจรและขนส่ง (เข้าถึงได้จาก http://www.bmatraffic.com/index.aspx) โดยเลือกเก็บรูปภาพจากเฉพาะกล้องวงจรปิดที่สามารถมองเห็นทางม้าลายได้อย่างชัดเจนในช่วงเวลากลางวัน รวมทั้งหมด 12 จุด ได้แก่ แยกพระรามที่ 3, แยกรัชดา-ราชพฤกษ์, แยกประชาอุทิศ-ประดิษฐ์มนูญธรรม, แยกเดโช, แยกตรอกจันทร์, แยกอุภัย, แยกวัดเบญฯ, แยกนครไชยศรี, แยกราชวัตร, แยกสุรศักดิ์, แยกบ้านแขก, และ แยกราชวิถี

ในการเก็บรูปภาพนั้นคณะผู้วิจัยทำการเก็บ 1 รูปต่อความยาววิดีโอ 1 วินาที รูปที่เก็บมามีขนาด 352 x 288 พิกเซล จำนวนรูปภาพทั้งหมดที่เก็บมาคือ 334 รูป ซึ่งเป็นจำนวนเบื้องต้นที่เพียงพอสำหรับใช้ในการทดสอบระบบต้นแบบของงานวิจัยนี้ เนื่องจากการพัฒนาระบบต้นแบบนี้อาศัยการฝึกสอนต่อยอดจากแบบจำลองการเรียนรู้เชิงลึกที่เคยถูกฝึกสอนไว้แล้ว (Pre-trained model) การฝึกสอนต่อยอดจึงไม่จำเป็นต้องใช้จำนวนข้อมูลฝึกสอนปริมาณมากเท่ากับการฝึกสอนแบบจำลองการเรียนรู้เชิงลึกตั้งแต่ต้น (Train from scratch) ตามปกติ

การกำกับข้อมูล

เมื่อได้รูปภาพจากกล้องวงจรปิดแล้ว คณะผู้วิจัยจึงนำภาพทั้งหมดมาผ่านโปรแกรม LabelImg (https://github.com/HumanSignal/labelImg) ซึ่งเป็นโปรแกรมตัวช่วยสำหรับให้คณะผู้วิจัยสามารถกำหนดกรอบ Bounding box สี่เหลี่ยมที่แสดงถึงขอบเขตของวัตถุที่ต้องการตรวจจับในภาพได้ (ดังแสดงใน Figure 1) อีกทั้งสามารถระบุได้ด้วยว่าแต่ละกรอบนั้นคือวัตถุเป้าหมายชิ้นใด โดยสำหรับงานชิ้นนี้กำหนดให้มีวัตถุเป้าหมาย 3 ประเภท ได้แก่ รถยนต์ (Class 0), รถจักรยานยนต์ (Class 1) และทางม้าลาย (Class 2) ดังแสดงใน Figure 2

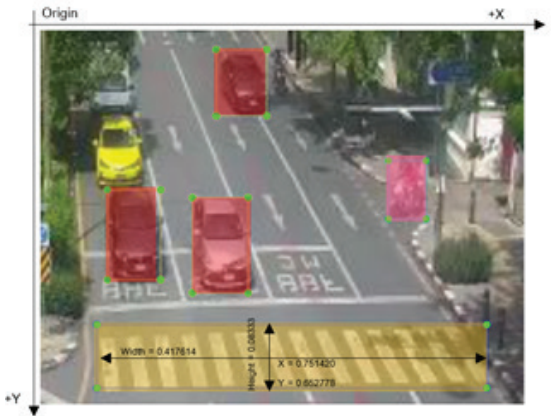

Figure 1 Object localization on images

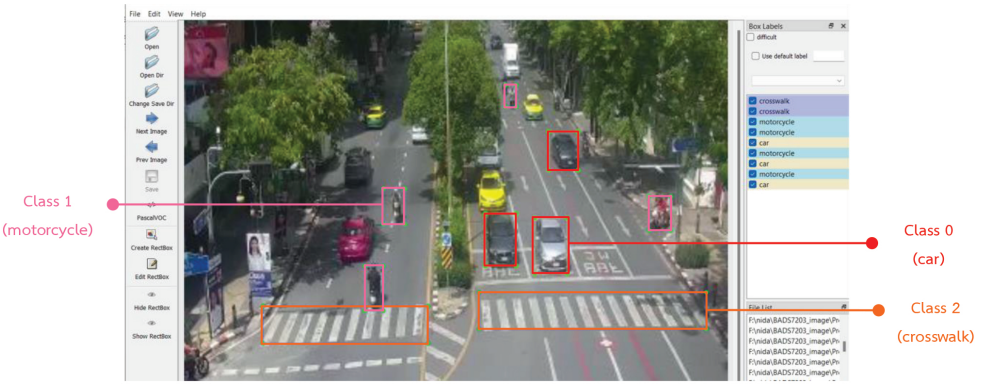

Figure 2 Object classification on images

ข้อมูลตำแหน่งและประเภทของ Bounding box ทั้งหมดที่คณะผู้วิจัยสร้างขึ้นด้วยโปรแกรม LabelImg จะถูกบันทึกให้อยู่ในรูปแบบที่พร้อมสำหรับการใช้งานกับแบบจำลอง YOLOv5 (https://github.com/ultralytics/yolov5) โดยบันทึกเป็นตัวเลข 5 ตัวต่อ Bounding box หนึ่งอัน ได้แก่ class, centerX, centerY, width, height

การตรวจจับวัตถุและการประเมินผล

แม้แบบจำลอง YOLOv5 จะมีความสามารถในการตรวจจับวัตถุหลายชนิดอยู่เดิมแล้ว แต่เพื่อให้แบบจำลองสามารถตรวจจับวัตถุในบริบทเฉพาะของทางม้าลายและยานพาหนะแบบของประเทศไทยได้ดีขึ้น ผู้วิจัยจึงนำรูปภาพพร้อมคำเฉลยจากหัวข้อการกำกับข้อมูล จำนวน 334 รูปมาทำการแบ่งออกเป็น 2 ส่วน ได้แก่ ข้อมูลสำหรับฝึกสอนจำนวน 301 รูป (90%) และสำหรับทดสอบจำนวน 33 รูป (10%) โดยการแบ่งส่วนนี้ผู้วิจัยเน้นให้มีข้อมูลฝึกสอนเป็นจำนวนมากเพื่อเลี่ยงปัญหา Overfit ทั้งนี้เทคนิคอื่นอย่าง Early stopping ไม่ได้ถูกใช้เพื่อแก้ปัญหา Overfit ในงานนี้เนื่องจากการทำ Early stopping ในแบบจำลองการเรียนรู้เชิงลึกมีความเสี่ยงต่อปรากฏการณ์ Double descent [20] จากนั้นคณะผู้วิจัยจึงนำข้อมูลฝึกสอนมาใช้ในการฝึกสอนแบบจำลอง YOLOv5 ต่อยอดจากเดิม และทำการวัดประสิทธิภาพในการตรวจจับวัตถุเป้าหมายของงานวิจัยนี้ด้วยข้อมูลส่วนทดสอบอีกทีหนึ่ง ทั้งนี้เพื่อยืนยันถึงประสิทธิภาพของการทำงานส่วนตรวจจับวัตถุที่ถูกฝึกสอนเพิ่มเติมนี้ก่อนจะถูกนำไปใช้ต่อยอดในขั้นถัดไป

ในงานวิจัยนี้การฝึกสอนแบบจำลองถูกทำโดยแพลตฟอร์มและไลบรารีที่เกี่ยวข้องดังต่อไปนี้ Google Colab และ torch 1.11.0+cu113 (Tesla T4) โดยใช้ค่า Image size = 416 , bitch size 16 , ทำการสร้างโมเดลจำนวน 100, 150, 200 epochs และใช้ Pre-trained จาก YOLOv5s จาก https://github.com/ultralytics/yolov5 ในส่วนของการประเมินผลแบบจำลองที่ได้ในการตรวจจับวัตถุ คณะผู้วิจัยเลือกใช้ค่า Average Precision (mAP) ดังผลลัพธ์ที่แสดงใน Table 1 ในส่วนของการนำแบบจำลองไปใช้ตรวจจับวัตถุบนรูปภาพชุดทดสอบนั้น คณะผู้วิจัยกำหนดค่า confidence มากกว่า 0.6 ดัง Figure 3 ทั้งนี้ค่าต่าง ๆ ที่ใช้ในการทดลองเป็นค่าที่ให้ผลลัพธ์ดีที่สุดหรือสมดุลเหมาะสมที่สุด ซึ่งผู้วิจัยพบจากการทดลองปรับจูนซ้ำ ๆ ในระหว่างการทดลอง

Table 1 Experimental results of YOLOv5 models trained with different epochs. Note that values in this table are summaries of model evaluation metrics calculated by the method proposed by YOLOv5

Epoch	mAP	Train box loss	Validation box loss	Time
100 epochs	0.823	0.03093	0.03393	6m 1s
150 epochs	0.826	0.02753	0.03364	8m 58s
200 epochs	0.837	0.0254	0.03299	11m 57s

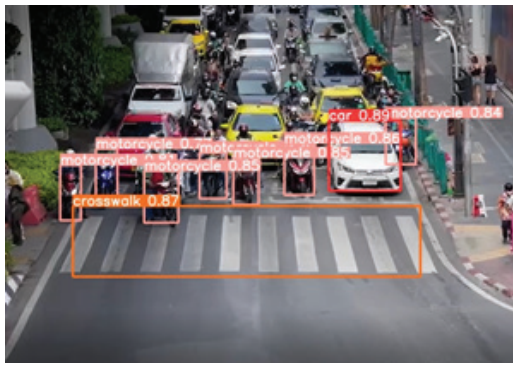

Figure 3 Vehicle and pedestrian crosswalk detection with YOLOv5, where the largest orange bounding box indicates the pedestrian crosswalk and other colored boxes represent all detected vehicles by YOLOv5

การตรวจจับการจอดรถทับทางม้าลาย

1. การติดตามวัตถุด้วยเทคนิค DeepSORT เมื่อเสร็จสิ้นการตรวจจับหารถยนต์ รถจักรยานยนต์ และทางม้าลายในภาพแล้ว คณะผู้วิจัยทำการติดตามตำแหน่งของยานพาหนะต่อด้วยเทคนิค DeepSORT [19] เพื่อตรวจจับว่ายานพาหนะที่ปรากฏอยู่ในรูปภาพนั้น เป็นยานพาหนะคันเดิมที่จอดอยู่นิ่ง ๆ หรือไม่ หากพบว่ายานพาหนะไม่มีการเคลื่อนไหวจึงค่อยจะทำการตรวจสอบต่อว่า ถ้า Bounding box ของรถยนต์และ/หรือรถจักรยานยนต์ซ้อนทับกับ Bounding box ของทางม้าลายนานเป็นเวลาเกินกว่า 5 วินาที ระบบจะทำการบันทึกว่ายานพาหนะคันดังกล่าวกำลังจอดทับทางม้าลายอยู่ในขณะที่กำลังติดไฟแดง อีกทั้งจะทำการนับจำนวนยานพาหนะที่จอดทับทางม้าลายอยู่ ณ ขณะหนึ่ง ๆ แสดงดัง Figure 4

Figure 4 Pedestrian crosswalk violation detection and offender counting using DeepSORT, where the largest orange bounding box indicates the pedestrian crosswalk and red shaded boxes indicate vehicles violating the pedestrian crosswalk

2. การติดตามวัตถุแบบ Manual การติดตามวัตถุแบบ Manual เป็นอีกหนึ่งวิธีที่คณะผู้วิจัยทำการทดลองเพื่อติดตามตำแหน่งยานพาหนะ โดยผู้วิจัยทำการเขียนโปรแกรมเพื่อเปรียบเทียบตำแหน่งของ Bounding box ในแต่ละเฟรมด้วยตนเองว่ามีพื้นที่ทับซ้อนกันเป็นร้อยละเท่าใดเมื่อเทียบกับ Bounding box ในเฟรมก่อนหน้า ทั้งนี้โดยจะสนใจเฉพาะยานพาหนะที่มี Bounding box ทับซ้อนกันอยู่กับ Bounding box ของทางม้าลายเท่านั้นดัง Figure 5

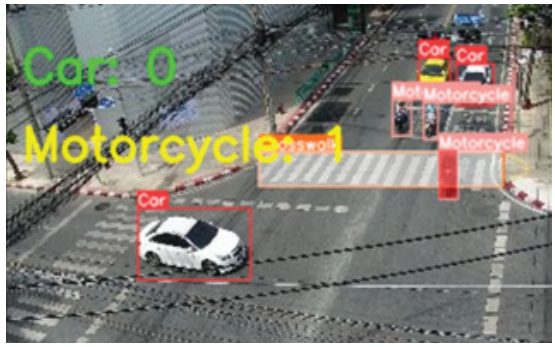

Figure 5 Pedestrian crosswalk violation detection and offender counting using the manual method, where the largest orange bounding box indicates the pedestrian crosswalk and red shaded boxes indicate vehicles violating the pedestrian crosswalk

3. การทดลองเปรียบเทียบ เพื่อเปรียบเทียบผลการติดตามวัตถุของเทคนิค DeepSORT และเทคนิค Manual คณะผู้วิจัยเลือกใช้ตัวชี้วัดชื่อ Multiple Object Tracking Accuracy (MOTA) ซึ่งสามารถถูกคำนวณโดยใช้สูตรดังสมการที่ 5 เมื่อ MOTA คือผลลัพธ์จากการคำนวณค่าการทำนายการตรวจจับวัตถุเทียบกับค่าตรวจจับวัตถุจริง, t คือเวลา ณ จุด ๆ หนึ่ง, FN คือผลการทำนายว่า negatives แต่ทำนายผิด, IDS คือ Mismatch error, FP คือผลการทำนายว่า positive แต่ทำนายผิด, และ GT คือ Ground truth object count

$$\text{MOTA} = 1 - \frac{\sum_t FN_t + FP_t + IDS_t}{\sum_t GT_t} \quad (5)$$

ค่า MOTA นี้เป็นการตรวจสอบว่าการตรวจจับติดตามวัตถุมีความแม่นยำหรือไม่ โดยการให้ทั้ง 2 เทคนิคทำนาย Bounding box ของตนออกมา ดังตัวอย่างใน Figure 6 จากนั้นจึงนำผลการทดลองไปเปรียบเทียบกับ Bounding box จริง

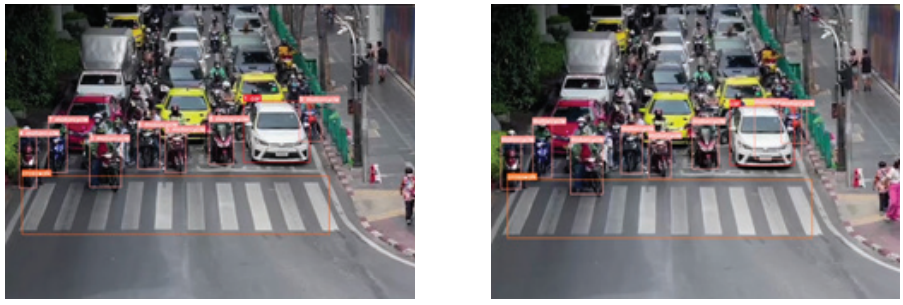

Figure 6 Examples of predicted bounding boxes obtained from DeepSORT tracking (a) and manual tracking (b)

Results

สำหรับการตรวจจับทางม้าลายและยานพาหนะโดยใช้เทคนิค YOLOv5 ผลการศึกษาพบว่า mAP มีค่า 0.837 , 0.826, 0.823 เมื่อทำการฝึกสอนต่อยอดเป็นจำนวน 200 epochs, 150 epochs และ 100 epochs ตามลำดับ ผลของ Validation box loss (ถูกคำนวณอยู่ภายใน YOLOv5) มีค่าอยู่ระหว่าง 0.03299 - 0.03393 และระยะเวลาที่ใช้ในการฝึกสอนจะเพิ่มขึ้นตามจำนวน epoch

ในงานวิจัยนี้ผู้วิจัยเลือกแบบจำลอง YOLOv5 ที่เป็นผลลัพธ์ของการฝึกสอนต่อยอดเป็นจำนวน 200 epochs มาเป็นแบบจำลองที่ใช้ในการตรวจจับวัตถุ เนื่องจากผลการประเมิน (ด้วยตัวชี้วัดของ YOLOv5) โดยภาพรวมพบว่า mAP 0.5-0.95 มีค่า 0.455 , Precision มีค่า 0.806, recall มีค่า 0.796 ทั้งนี้หากพิจารณาในชุดข้อมูลฝึกสอนจำนวน 301 รูปจะพบว่า box loss มีค่า 0.0254 , cls loss มีค่า 0.0017493 และ obj loss มีค่า 0.02421 อีกทั้งสำหรับในชุดข้อมูลทดสอบจำนวน 33 รูปพบว่า box loss มีค่า 0.03299 , cls loss มีค่า 0.0027196 และ obj loss มีค่า 0.03557 ดัง Figure 7 ทั้งนี้ผลการวัดค่าพื้นที่ใต้กราฟระหว่าง precision - recall ของรถยนต์มีค่า 0.820 , ของรถจักรยานยนต์มีค่า 0.848 และของทางม้าลายมีค่า 0.844 ดัง Figure 8

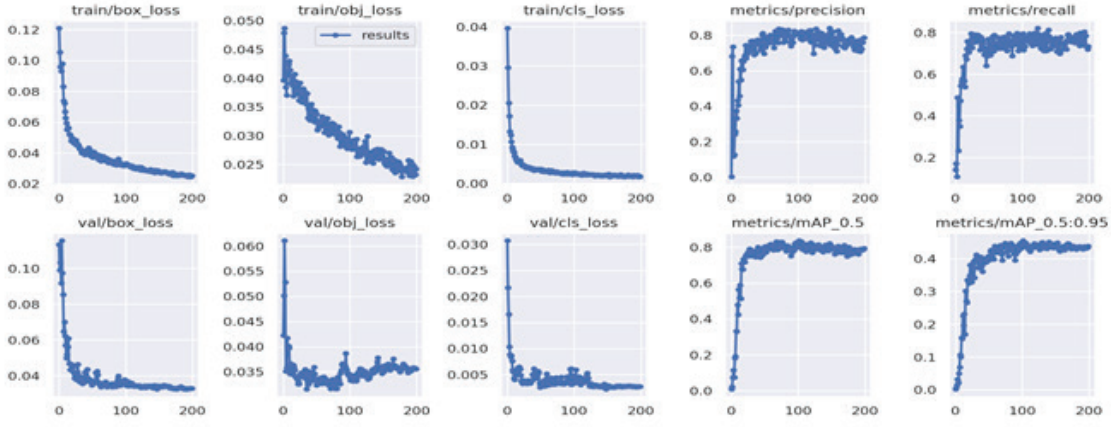

Figure 7 Evaluation of the YOLOv5 model trained with 200 epochs

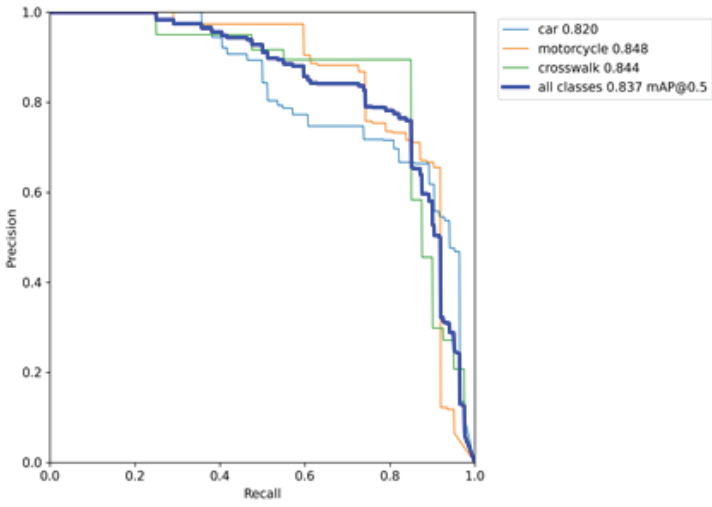

Figure 8 Area under the precision-recall curve

ในส่วนผลการทดลองเปรียบเทียบการติดตามวัตถุด้วยเทคนิค DeepSORT และ Manual นั้นแสดงดัง Table 2 จาก Table 2 พบว่าการใช้เทคนิค DeepSORT ได้ค่า MOTA = 1 อีกทั้งยังสามารถติดตามได้ว่าวัตถุดังกล่าวเป็นวัตถุเดิมหรือไม่ ในขณะที่เทคนิค Manual แม้จะได้ค่าใกล้เคียงกันคือ MOTA = 0.89 แต่กลับไม่สามารถติดตามได้ว่าวัตถุดังกล่าวเป็นยานพาหนะเดิมทำให้เกิดการนับยานพาหนะซ้ำหรือการตรวจจับระยะเวลาที่จอดทับทางม้าลายผิดพลาดไปจากความเป็นจริงได้ ผู้วิจัยจึงสรุปว่าในงานนี้การใช้เทคนิค DeepSORT สามารถติดตามวัตถุที่ถูกตรวจจับโดย YOLOv5 ได้ดีกว่าเทคนิค Manual

Table 2 Experimental comparison of object tracking using DeepSORT and manual techniques

	DeepSORT	Manual
FN	0	0
FP	0	1
IDS	0	0
GT	9	9
MOTA	1	0.89
Tracking ID	Yes	No

Conclusion

งานวิจัยนี้นำเสนอระบบต้นแบบสำหรับตรวจจับการละเมิดทางม้าลายในขณะที่รถจอดติดไฟแดง โดยมีการตรวจจับรถยนต์ รถจักรยานยนต์ และทางม้าลายด้วยเทคนิค YOLOv5 ซึ่งมีประสิทธิภาพของค่า mAP 0.837 และค่าพื้นที่ใต้กราฟระหว่าง precision - recall ของรถยนต์ = 0.820 , รถจักรยานยนต์ = 0.848 และทางม้าลาย = 0.844 เมื่อนำผลการตรวจจับวัตถุมาต่อยอดด้วยเทคนิคการติดตามวัตถุ 2 เทคนิคพบว่าเทคนิค DeepSORT วัดผลได้ค่า MOTA = 1 อีกทั้งสามารถตรวจสอบได้ว่ายานพาหนะคือคันเดิมและมีการซ้อนทับกับทางม้าลายเกินระยะเวลาที่กำหนดหรือไม่ แต่สำหรับเทคนิคที่สองคือเทคนิค Manual ซึ่งมีค่า MOTA = 0.89 นั้นแม้จะใช้ตรวจจับการจอดทับทางม้าลายได้ แต่ด้วยความผิดพลาดที่ไม่สามารถตรวจจับยานพาหนะคันเดิมที่อยู่ในต่างเฟรมได้อย่างแม่นยำจึงถือว่ามีประสิทธิภาพที่ด้อยกว่าเทคนิค DeepSORT

Future Works

1. ระบบต้นแบบปัจจุบันยังมีข้อจำกัด อาทิ สามารถตรวจจับผู้กระทำความผิดในช่วงเวลากลางวันเท่านั้น และบางครั้งมีการเข้าใจผิดว่าคนที่เดินข้ามทางม้าลายเป็นรถจักรยานยนต์ เพื่อพัฒนาให้ระบบมีความแม่นยำในการตรวจจับที่ดีขึ้น ในอนาคตควรเพิ่มจำนวนรูปภาพในการฝึกสอนต่อยอดแบบจำลองตรวจจับวัตถุ (YOLOv5) โดยให้ภาพที่ใช้ฝึกสอนครอบคลุมทั้งในช่วงเวลากลางวันและช่วงเวลากลางคืน รวมทั้งควรเพิ่มการตรวจจับคนในการฝึกสอนด้วย เพื่อให้ระบบสามารถแยกคนข้ามทางม้าลายออกจากรถจักรยานยนต์ได้ดีขึ้น

2. เนื่องจากงานวิจัยนี้ใช้เฉพาะรูปภาพจากกล้องวงจรปิดที่เห็นมุมด้านหน้าแยกไฟแดงเท่านั้น หากต้องการที่จะพัฒนาระบบให้สามารถตรวจจับป้ายทะเบียนของยานพาหนะที่กระทำความผิดได้แบบอัตโนมัติด้วย อาจต้องพิจารณาเปลี่ยนหรือเพิ่มมุมกล้องในการจับภาพให้เป็นมุมที่สามารถมองเห็นป้ายทะเบียนของยานพาหนะที่กระทำความผิดได้ชัดเจนด้วย อีกหนึ่งข้อจำกัดของการใช้มุมกล้องด้านหน้าเพียงมุมเดียวและไม่สามารถมองเห็นสัญญาณไฟจราจรปัจจุบันได้แบบทันทีทันใด คือ ระบบปัจจุบันไม่สามารถแยกแยะได้ว่ายานพาหนะจอดทับทางม้าลายเนื่องจากติดไฟแดงอยู่ หรือเนื่องจากรถติดยาวมาจากด้านหน้าทำให้จอดทับทางม้าลายโดยบังเอิญ

3. งานวิจัยนี้สามารถนำไปประยุกต์เพื่อตรวจจับการกระทำความผิดในรูปแบบอื่นได้ เช่น รถจักรยานยนต์ขับขี่บนทางเท้า รถยนต์จอดในที่ห้ามจอด

References

1. ThaiRoads Foundation, 2019, Key Facts on Road Safety Situations in Thailand 2016-2017 [Online], Available: http://trso. thairoads.org/system/files/rs_book_59-60_ final_resize_0. pdf. [1 June 2022] (In Thai)

2. Department of Disease Control, 2022, Pedestrians and the Crosswalk Safety [Online], Available: https://ddc.moph.go.th/ dip/news.php?news=23028&deptcode=dip. [1 June 2022] (In Thai)
3. National Police Agency, "Pedestrians First at Road Crossings," National Police Agency [Online], Available: https://www.npa.go.jp/english/bureau/traffic/document/06_ PedestriansFirstatRoadCrossings.pdf. [11 March 2024]
4. Bangkok Biz News, 2022, Bangkok Raises 'Asoke Montri Intersection' as a Safe Crosswalk after using AI Cameras to Catch and Penalize Offenders [Online], Available: https:// www.bangkokbiznews. com/news/984596. [1 June 2022] (In Thai)
5. Ibadov, S., Ibadov, R., Kalmukov, B. and Krutov, V., 2017, "Algorithm for Detecting Violations of Traffic Rules based on Computer Vision Approaches," *MATEC Web of Conferences*, 132 (05005), 4 p. https://doi.org/10.1051/matecconf/201713205005
6. Zhang, Z.D., Tan, M.L., Lan, Z.C., Liu, H.C, Pei, L. and Yu, W.X., 2022, "CDNet: A Real-time and Robust Crosswalk Detection Network on Jetson Nano based on YOLOv5," *Neural Computing and Applications*, 34 (13), pp. 10719-10730.
7. Malbog, M.A., 2019, "MASK R-CNN for Pedestrian Crosswalk Detection and Instance Segmentation," *IEEE International Conference on Engineering Technologies and Applied Sciences (ICETAS)*, 20-21 December 2019, Kuala Lumpur, Malaysia, pp. 1-5.
8. Romić, K., Galić, I., Leventić, H. and Habijan, M., 2021, "Pedestrian Crosswalk Detection Using a Column and Row Structure Analysis in Assistance Systems for the Visually Impaired," *Acta Polytechnica Hungarica*, 18 (7), pp. 25-45.
9. Tonge, A., Chandak, S., Khiste, R., Khan, U. and Bewoor, L.A., 2020, "Traffic Rules Violation Detection using Deep Learning," *International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, 5-7 November 2020, Coimbatore, India, pp. 1250-1257.
10. Ibadov, S.R., Kalmykov, B.Y., Ibadov, R.R. and Sizyakin, R.A., 2019, "Method of Automated Detection of Traffic Violation with a Convolutional Neural Network," *EPJ Web of Conferences*, 224 (04004), 7 p. https://doi.org/ 10.1051/epjconf/201922404004
11. Prasanthi, K., Chowdhary, S.V., Prasanthi, T.L. and Neelima, N., 2022, "Analysis on Applications of Machine Learning Based Image Processing for Vehicle Detection and Counting System," *International Journal of Mechanical Engineering*, 7 (1), 5 p.

12. Navastara, D.A., Maulana, M.F., Suciati, N. and Sarena, S.T., 2021, "Vehicle Classification Based on CCTV Video Recording Using Histogram of Oriented Gradients, Local Binary Patterns, and Hierarchical Multi-SVM," *IOP Conference Series: Materials Science and Engineering*, 1077 (1), 012068. https://doi.org/ 10.1088/1757-899X/1077/1/012068
13. Pan, S. and Wang, S., 2021, "Identifying Vehicles Dynamically on Freeway CCTV Images through the YOLO Deep Learning Model," *Sensors and Materials*, 33 (5), pp. 1517–1530.
14. Zhu, M., Zhang, S., Zhong, Y., Lu, P., Peng, H. and Lenneman, J.K., 2021, "Monocular 3D Vehicle Detection Using Uncalibrated Traffic Cameras through Homography," *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 27 September– 1 October 2021, Prague, Czech Republic, pp. 3814-3821.
15. Shepelev, V., Aliukov, S., Glushkov, A. and Shabiev, S., 2020, "Identification of Distinguishing Characteristics of Intersections based on Statistical Analysis and Data from Video Cameras," *Journal of Big Data*, 7 (46), 23 p. https://doi.org/10.1186/s40537-020-00324-7
16. Hou, X., Wang, Y. and Chau, L., 2019, "Vehicle Tracking Using Deep SORT with Low Confidence Track Filtering," *IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, 18-21 September 2019, Taipei, Taiwan, pp. 1-6.
17. Parico, A.I.B. and Ahamed, T., 2021, "Real Time Pear Fruit Detection and Counting Using YOLOv4 Models and Deep SORT," *Sensors*, 21 (14), 32 p. https://doi.org/10.3390/s21144803
18. Santos, A.M., Bastos-Filho, C.J.A., Maciel, A.M.A. and Lima, E., 2020, "Counting Vehicle with High-Precision in Brazilian Roads Using YOLOv3 and Deep SORT," *SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, 7-10 November 2020, Porto de Galinhas, Brazil, pp. 69-76.
19. Wojke, N., Bewley, A. and Paulus, D., 2017, "Simple Online and Realtime Tracking with a Deep Association Metric," *IEEE International Conference on Image Processing (ICIP)*, 17-20 September 2017, Beijing, China, pp. 3645-3649.
20. Lafon, M. and Thomas, A., 2024, "Understanding the Double Descent Phenomenon in Deep Learning," *arXiv*, 2403 (10459), 21 p. https://doi.org/10.48550/arXiv.2403.10459