

การติดตามการเคลื่อนที่ของยานพาหนะบนเครื่องคอมพิวเตอร์สมรรถนะต่ำ
โดยใช้โมเดลการเรียนรู้เชิงลึกและการติดตามแบบหลายวัตถุ

VEHICLE TRACKING ON A LOW-PERFORMANCE COMPUTER
ENVIRONMENT USING DEEP LEARNING AND MULTI-TRACKING
MODELS

วิษณุ โกเมศ¹
ดร.ชนภัทร มั่งคะจิตร²

บทคัดย่อ

การติดตามการเคลื่อนที่ของยานพาหนะบนท้องถนนถือเป็นงานสำคัญในการบังคับใช้กฎหมาย ซึ่งส่งผลให้ลดการเกิดอุบัติเหตุได้เป็นอย่างดี จากความก้าวหน้าของเทคโนโลยีด้านปัญญาประดิษฐ์สำหรับงานทางด้านคอมพิวเตอร์วิทัศน์ช่วยให้การพัฒนาการตรวจจับวัตถุและการติดตามการเคลื่อนที่ของวัตถุสามารถทำได้ง่ายและมีประสิทธิภาพ อย่างไรก็ตามสภาพแวดล้อมการทำงานจริงบนท้องถนนไม่เอื้อต่อการติดตั้งคอมพิวเตอร์สมรรถนะสูง ดังนั้นในงานวิจัยนี้จึงได้นำเสนอแนวทางการติดตามการเคลื่อนที่ของยานพาหนะบนเครื่องสมรรถนะต่ำ (คอมพิวเตอร์อุตสาหกรรม) ซึ่งมีความทนทานต่อสภาพแวดล้อมบนท้องถนนได้เป็นอย่างดี โดยงานวิจัยนี้ได้ทำการปรับแต่งแบบจำลอง YOLOv5 ให้มีความแม่นยำในการตรวจจับยานพาหนะ 3 ประเภทหลักบนท้องถนน ได้แก่ รถยนต์, รถจักรยานยนต์และรถบรรทุก หลังจากนั้นจึงนำยานพาหนะที่ตรวจจับได้ไปทำงานร่วมกับอัลกอริทึม DeepSORT เพื่อติดตามการเคลื่อนที่ของยานพาหนะเหล่านั้น โดยมีการปรับปรุงในส่วน of IOU Matching และ Matching Cascade เพื่อให้สามารถทำงานได้รวดเร็วและแม่นยำ ผลการทดลองการตรวจจับการติดตามการเคลื่อนที่ของยานพาหนะบนท้องถนนพบว่าให้ความแม่นยำในการตรวจจับและติดตามวัตถุถึง 89% และให้ความเร็วในการประมวลผลถึง 10 fps บนเครื่องคอมพิวเตอร์อุตสาหกรรม (CPU I7 Gen 6) ซึ่งเพียงพอต่อการนำไปประยุกต์ใช้ตรวจจับการกระทำความผิดกฎหมายทางถนน

¹ นักศึกษาหลักสูตรวิศวกรรมศาสตรมหาบัณฑิต สาขาวิศวกรรมข้อมูลขนาดใหญ่

² ที่ปรึกษาสารนิพนธ์หลัก

ABSTRACT

Vehicle tracking is a crucial task in traffic enforcement resulting in a decreasing number of road accidents. Regarding the advances of artificial intelligence technology in computer vision tasks, the implementation of object detection and multi-object tracking tasks can be performed efficiently and conveniently. Unfortunately, installing high-performance computing equipment is not permitted in a real working environment. In this research, we proposed an efficient approach to track vehicles on the roads using low-performance computers (i.e. industrial computers) in tough environments. Firstly, a pre-trained YoloV5 model was re-trained on the custom datasets to precisely detect three types of vehicles; cars, motorcycles, and trucks. Then, the DeepSORT algorithm was applied to track the movement of the detected vehicles. To track vehicles efficiently, we modified the IOU matching and matching cascade modules of DeepSORT. The experimental results in the real working environment showed that our approach produced the satisfied accuracy (0.89) and fast processing time (10 fps) on the industrial computer (CPU i7 Gen 6). As a consequence, our approach is capable of fulfilling the various tasks of traffic enforcement.

บทนำ

โดยปัจจุบันเทคโนโลยีในเรื่องการตรวจจับวัตถุ ด้วย Deep learning มีความพัฒนาไปอย่างก้าวกระโดดทำให้ปัจจุบันมีความแม่นยำสูงอีกทั้งยังมีการประมวลผลที่เร็วขึ้นทำให้สามารถนำไปประยุกต์ใช้ในงานต่างๆได้มากมาย ในงานวิจัยชิ้นนี้ก็ได้ประยุกต์นำเอาเทคโนโลยีการตรวจจับวัตถุ หรือ Object Detect ด้วย Deep learning มาใช้สำหรับตรวจจับยานพาหนะบนท้องถนน หรือการติดตามบุคคล เป็นต้น โดยได้นำเอา DeepSORT เฟรมเวิร์ก ซึ่งเป็นที่นิยมในงานการติดตามวัตถุมาใช้สำหรับ Tracking ในเฟรมภาพจากกล้อง CCTV เพื่อตรวจจับและติดตามวัตถุ

ในงานวิจัยนี้ได้นำเสนอวิธีการติดตามแบบหลายวัตถุ ในเฟรมภาพ โดยไม่ต้องทำการ Pre-training ก่อน โดยจะเริ่มประมวลผลจากภาพในเฟรมต่อเฟรม โดย Hungarian Assignment กับ Association metric เพื่อใช้ในการคำนวณหาความสัมพันธ์ระหว่างวัตถุ ทำให้การติดตามวัตถุมีความแม่นยำ อีกทั้งยังสามารถ Tracking object ได้โดยซ้อนทับกันบางส่วนได้

วัตถุประสงค์

1. เพื่อใช้ตรวจจับและติดตามแบบหลายวัตถุใน Stream ภาพบนเครื่องคอมพิวเตอร์ที่มีสมรรถนะต่ำ
2. นำไปประยุกต์ใช้กับ Applications ในงานต่างๆที่ต้องการ การติดตามวัตถุแบบ Realtime โดยที่สามารถติดตามวัตถุได้แม้วัตถุหายไปชั่วคราว และปรากฏขึ้นอีกครั้ง

ขอบเขตของการวิจัย

1. ตรวจจับวัตถุ และติดตามแบบหลายวัตถุได้ บนเครื่องคอมพิวเตอร์ที่มีสมรรถนะต่ำ
2. สามารถตรวจจับวัตถุ และนำเสนอแบบ ID ของวัตถุนั้นๆ
3. การทดลองในงานวิจัยนี้ ทำบริเวณเขตกรุงเทพมหานคร มุมมองถ่ายด้านหลังยานพาหนะช่วง คอสะพาน ทดสอบในสภาพการจราจรแบบปกติ รถไม่ติด ช่วงเวลากลางวันและกลางคืน สภาพอากาศไม่มีฝนตก

ประโยชน์ที่คาดว่าจะได้รับ

1. สามารถตรวจจับและติดตามแบบหลายวัตถุในเฟรมของวิดีโอได้
2. สามารถนำไปใช้ในงานที่เกี่ยวข้องกับการติดตามวัตถุ หรือการนับยานพาหนะโดยใช้กล้อง CCTV หรือ Streaming ได้
3. นำเสนอเฟรมวีร์กตรวจจับและติดตามแบบหลายวัตถุ บนเครื่องคอมพิวเตอร์ที่มีสมรรถนะต่ำ

แนวคิด ทฤษฎี และงานวิจัยที่เกี่ยวข้อง

งานวิจัยที่เกี่ยวข้องกับการตรวจจับและติดตามแบบหลายวัตถุ โดยการเรียนรู้เชิงลึกนั้น ปัจจุบันที่ได้รับความนิยมอยู่หลายตัว ซึ่งแต่ละตัวมี ข้อดี ข้อเสีย แตกต่างกันไป ผู้วิจัยได้ศึกษางานวิจัยที่ได้รับการตีพิมพ์และสรุปได้ดังนี้

DeepSORT: Simple online and realtime tracking with a deep association metric (Wojke,2017) เป็นงานวิจัยนำเสนออัลกอริทึมการตรวจจับและติดตามวัตถุ (Object tracking) ที่มีประสิทธิภาพจากการวิเคราะห์ลำดับของภาพ หรือเฟรมของภาพ โดยตรวจสอบการเปลี่ยนแปลงเชิงพื้นที่และเวลาของวัตถุในระหว่างลำดับของเฟรม โดยไม่ต้องทำการ Pre-train เพื่อรู้จำใดๆก่อน โดยจะเริ่มประมวลผลจากภาพในเฟรม เป็นการ Tracking object ใน แบบ Realtime โดยใช้ Hungarian assignment เพื่อใช้ในการคำนวณหาความสัมพันธ์ระหว่างวัตถุ และประมวลผล โดยผลลัพธ์การติดตามวัตถุค่อนข้างแม่นยำ อีกทั้งยังสามารถ Tracking object ได้แม้ ซ้อนทับกันบางส่วน แต่ถ้าวัตถุโดนซ้อนทับเกินกว่า 70% จะทำให้เกิด Loss ในการ

ติดตามได้ ทั้งยังมีข้อจำกัดเรื่องความ การเปลี่ยนระหว่างเฟรมที่จะเกิดระยะความใกล้เคียงของวัตถุ ความยุ่งเหยิงของพื้นหลัง มีผลต่อค่า Loss ทั้งสิ้น

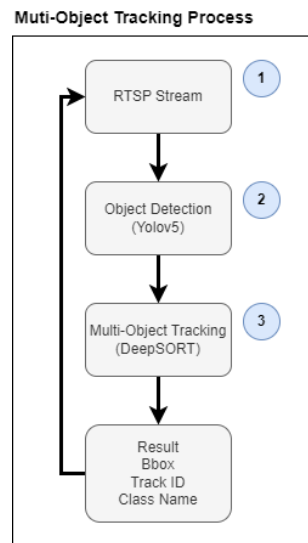
StrongSORT: Make DeepSORT Great Again (Yunhao Du,2021) เป็นงานวิจัยที่นำเสนอการปรับปรุง DeepSORT สำหรับแก้ไขการดักจับวัตถุ และเพิ่มอัลกอริทึม Appearance-free link model (AFLink) และ Gaussian-smoothed interpolation (GSI) โดย AFLink เสนอการเชื่อมโยงการ Track (Associate short track) โดยการเชื่อมโยง Global link model และ GSI เป็นอัลกอริทึมการช่วยจัดการ Loss ของวัตถุ ภายในระบบโดยจะมาแทน Kalman Filter เดิมของ DeepSORT โดยที่กล่าวมาจะทำให้การตรวจจับและความแม่นยำมากขึ้นแต่ใช้เวลาในการประมวลผลมากยิ่งขึ้น

Zero-Shot Object Tracking (Solawetz & Roboflow,2021) Roboflowซึ่งเป็น Website ให้บริการด้าน AI สำหรับนักพัฒนาในการ Labeling data และ Training model ที่เป็นที่นิยม โดยนำเสนอการติดตามวัตถุแบบ Zero-Shot โดยได้นำ DeepSORT มาประยุกต์ใช้ร่วมกับ CLIP (Contrastive Language-Image Pre-Training) โดยนำส่วน Image Encoder ของ CLIP มาเป็น Input และ Parameter ใน DeepSORT แทนที่ CNN และ IOU โดยจะทำให้เฟรมเวิร์ก จำวัตถุได้แม่นยำมากยิ่งขึ้นและการ Tracking มีความทนทานต่อ Noise และยังทำงานได้ดีเมื่อมีการซ้อนทับกันของวัตถุ แต่ข้อเสียใช้เวลาในการประมวลผลเพื่อนำ Latent Feature จากส่วน Image Encoder มาก โดยถ้าเป็น CPU จะทำให้ไม่สามารถใช้งานแบบ Realtime ได้เลย

ระเบียบวิธีวิจัย

งานวิจัยนี้นำเสนอการติดตามแบบหลายวัตถุ เป็นส่วนหนึ่งของงานด้านคอมพิวเตอร์วิทัศน์ เพื่อติดตามวัตถุในแบบ Realtime และใช้เครื่องประมวลผลที่มีประสิทธิภาพระดับกลาง โดยทำงานบน CPU เป็นหลัก โดยใช้การเรียนรู้เชิงลึกและโมเดลการติดตามวัตถุ โดยเครื่องประมวลผลที่ทดสอบมีคุณสมบัติดังนี้ CPU Cor i7 Gen6, Ram 16 GB, HDD 1 TB, OS ubuntu 18.04

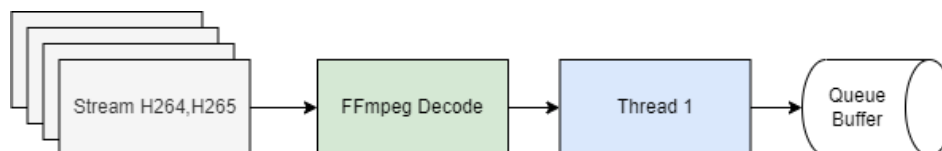
ภาพรวมของระบบและโมดูลประกอบไปด้วยส่วนหลัก 3 ส่วน 1) ส่วนรับภาพ RTSP Stream 2) ส่วนดักจับวัตถุหรือการตรวจจับวัตถุ (Object Detection) 3) ส่วนการติดตามวัตถุ (Multi-Object Tracking)



ภาพที่ 1 ภาพรวมโมดูลการติดตามวัตถุ

1. การรับภาพจาก Stream (RTSP Stream)

- Libraries FFmpeg ที่ใช้ในการถอดรหัสภาพ FFmpeg เป็นโปรแกรมจัดการไฟล์มีเดีย FFmpeg เป็นซอฟต์แวร์ที่เปิดให้นักพัฒนาสามารถนำไปใช้งานได้ฟรี ไม่มีค่าใช้จ่าย (Open Source) สามารถใช้งานได้ทั้ง 3 ระบบปฏิบัติการหลัก คือ Linux, MacOS และ ระบบปฏิบัติการ Window ใช้เป็นตัวถอดรหัสภาพและจัดการ Streaming
- ถอดรหัส H264,H265 ที่ 15 fps สร้างโมดูลการรับภาพและจัดการ Stream จาก CCTV โดยทำการเขียนโปรแกรมจัดการข้อมูลภาพที่ส่งมาในรูปแบบ Stream เพื่อแปลงข้อมูลดิบมาเป็นข้อมูลภาพเพื่อนำไปประมวลผลใน Process ถัดไป



ภาพที่ 2 การ Decoder ภาพเพื่อนำไปประมวลผล

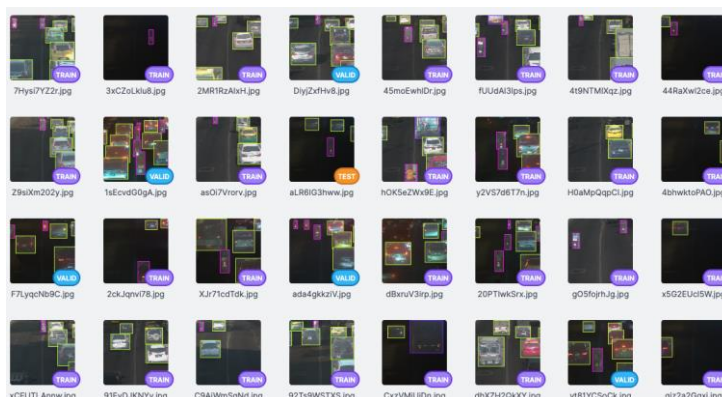
2. การตรวจจับวัตถุ (Object Detection)

โมเดล YOLOv5 นำมาใช้ในการตรวจจับวัตถุภาพในภาพ โดยทดลองจาก Custom Dataset ที่ Snapshot ภาพจากกล้อง CCTV ในมุมมองของการตรวจจับยานพาหนะ

แบ่งข้อมูลจากการ Snapshot ภาพจากกล้อง CCTV มุมมองกว้าง ทำการแบ่งข้อมูลออกเป็น 3 ส่วนสำหรับนำไปเรียนรู้อย่างนี้

- (1) Training Set: 3,375 รูป สำหรับนำไปเข้าโมเดลการเรียนรู้
- (2) Validation Set: 998 รูป สำหรับนำไปเป็นข้อมูล Validation ในโมเดลการเรียนรู้
- (3) Testing Set: 500 รูป สำหรับทดสอบโมเดล
- (4) Resize รูปเป็น 416x416 Pixel สำหรับนำเข้า โมเดล YOLOv5s

ทำการ Custom YOLO ด้วยการ Label ข้อมูลที่สนใจเฉพาะ มอเตอร์ไซด์ รถยนต์ 4 ล้อ และ รถบรรทุก โดยทั้งหมดมี 3 Class m=มอเตอร์ไซด์, c=รถยนต์ 4 ล้อ, t=รถบรรทุก โดยการเทรนโมเดลด้วย YOLOv5s และให้ IOU threshold for Non-max suppression = 0.5



ภาพที่ 3 การเตรียมข้อมูลสำหรับเทรน YOLO

โมเดล YOLO ที่ใช้ในการเทรน โดยตัว YOLO เป็นสถาปัตยกรรมที่ออกแบบไว้สำหรับทำ Image Detection ได้อย่างรวดเร็วและมีประสิทธิภาพ โดยมีหลายโมเดลสำหรับการนำไปใช้งานโดยแต่ละโมเดลจะมีขนาดที่ต่างกัน โดยยิ่งโมเดลใหญ่ก็จะทำให้การประมวลผลนานขึ้น แต่ก็มีคามแม่นยำมากยิ่งขึ้น โดยในงานวิจัยนี้เลือกใช้ YOLOv5s เพราะมีความเร็ว และมีความแม่นยำในระดับที่น่าพอใจที่สุด

				
Nano YOLOv5n	Small YOLOv5s	Medium YOLOv5m	Large YOLOv5l	XLarge YOLOv5x
4 MB _{FP16} 6.3 ms _{V100} 28.4 mAP _{COCO}	14 MB _{FP16} 6.4 ms _{V100} 37.2 mAP _{COCO}	41 MB _{FP16} 8.2 ms _{V100} 45.2 mAP _{COCO}	89 MB _{FP16} 10.1 ms _{V100} 48.8 mAP _{COCO}	166 MB _{FP16} 12.1 ms _{V100} 50.7 mAP _{COCO}

ภาพที่ 4 YOLOv5 โมเดลต่างๆ

YOLO จะให้ผลลัพธ์เป็น Boundary box ในภาพ โดยจะนำเอาไปใช้ต่อในการเข้า โมเดลการตรวจจับวัตถุต่อไป ตัวอย่าง ผลลัพธ์ของ YOLO [xmin, ymin, xmax, ymax, confidence, class, name]



ภาพที่ 5 ผลลัพธ์การ ตรวจจับวัตถุ ด้วย YOLOv5s

3. การติดตามแบบหลายวัตถุ (Multi-Object Tracking)

ในงานวิจัยนี้ได้ทำการศึกษาในส่วน DeepSORT (Simple Online and Realtime Tracking with a Deep Association Metric) ซึ่งเป็นอัลกอริทึมในการติดตามแบบหลายๆวัตถุ (Multi-Object tracking) ที่มีประสิทธิภาพจากการวิเคราะห์ลำดับของภาพ ในเฟรมภาพ

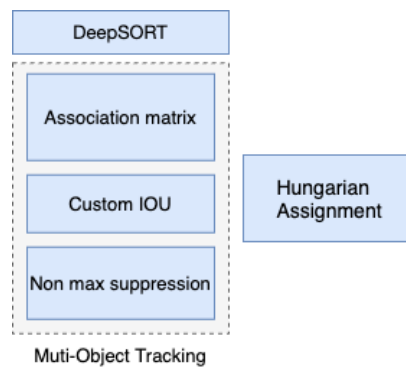
เมื่อทำการตรวจจับ Bbox (Bounding box coordinate of the object) จาก YOLOv5O นั้น จะทำการส่ง Bbox, id ของ Class และชื่อ Class เข้าไปในโมเดลการตรวจจับ DeepSORT ซึ่งประกอบไปด้วยส่วนหลัก ๆ ดังนี้

Detection มาจาก YOLOv5 สำหรับสร้างเป็นข้อมูลเริ่มต้นและการ update, delete ข้อมูลการ tracking ของวัตถุโดยส่งมาเป็นลำดับเรื่อย ๆ

Estimation การประมาณค่าตำแหน่งถัดไปโดยใช้แบบจำลอง Kalman filter เพื่อเชื่อมโยงกับตำแหน่งและขอบเขตของวัตถุที่ตรวจพบ และจะถูก Update ข้อมูลและค่าต่างๆในระบบ

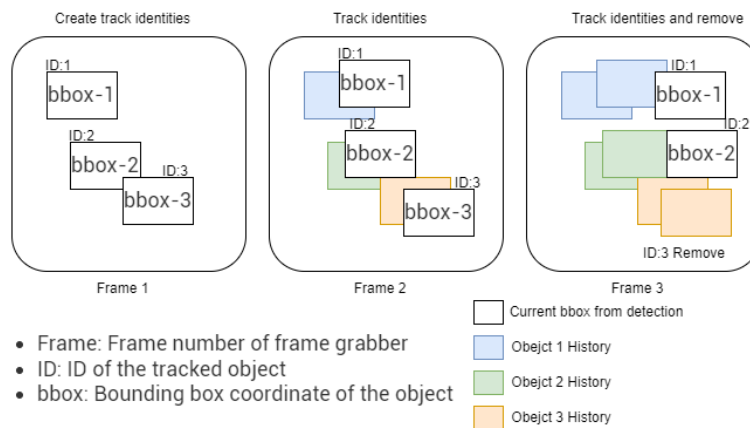
Data association จะคำนวณการหาความสัมพันธ์ระหว่างข้อมูลในระบบด้วย cost matrix และ intersection-over-union (IOU) ในการ Update ID ในระบบ Tracking โดยใช้ Hungarian assignment ระหว่างข้อมูลในระบบเป็นการเชื่อมโยงหาความสัมพันธ์ระหว่างข้อมูลในระบบโดยจะ Update สถานการณ์ Track ใน Matrix

Creation and Deletion of Track Identities สร้างและลบในการจัดการการติดตามวัตถุด้วย ID เป็นส่วนในการหาค่า Status ต่างๆที่เป็นเงื่อนไขในการ Tracking



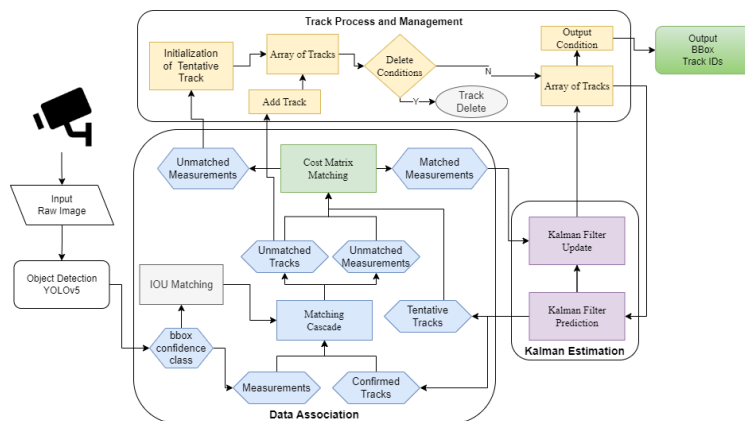
ภาพที่ 6 Hungarian Assignment (DeepSORT)

การติดตามวัตถุ ด้วย SORT ภายในจะมี Kalman filter และ Hungarian Assignment ที่ใช้ในการติดตามวัตถุที่ยังอยู่ในเฟรมภาพ (online) และ ในกรอบ IOU ระหว่างเฟรม โดยจะทำการ Update Bbox ในทุกๆ เฟรม เพื่อหา Cost Matrix ที่น้อยที่สุดสำหรับ Update ใน Track Array



ภาพที่ 7 ตัวอย่างการ Tracking ด้วย DeepSORT

การประมวลผลในระบบ เมื่อรับภาพจากกล้องจะส่งไปตรวจจับ Object Detection ด้วย YOLOv5 จะได้ตำแหน่งของวัตถุในภาพ เป็น Bbox, Class, Confidence แยกเป็น 2 ทาง 1.ไปหา IOU Matching 2.ไป Measurement โดยใช้ตัวกรอง Kalman เพื่อทำนาย Bbox สำหรับนำไปเทียบกับข้อมูล ID ในระบบ Matching Cascade โดยจะทราบ ว่าเป็นวัตถุในระบบหรือไม่ ถ้าเป็นให้ทำการ Update ถ้าไม่พบจะเข้า Process การเพิ่มข้อมูลการ Track ใหม่ โดยทุกๆรอบการทำงานจะให้ผลลัพธ์เป็นตำแหน่งของ Bbox, Track ID



ภาพที่ 8 DeepSORT Framework

การ Custom DeepSORT เพื่อให้ทำงานกับ CPU ได้ดีนั้น จะทำการนำส่วนการแกะ Feature ใน Bbox แต่ละ Object ออก เพราะในส่วนนี้จะประมวลผลหนัก โดยเปลี่ยนมาใช้ค่า IOU และ Aspect ratio ของ IOU แต่ละตัวแทน โดยจะนำไปเข้า Matching Cascade และนำไปหา Cost Matrix ต่อไป

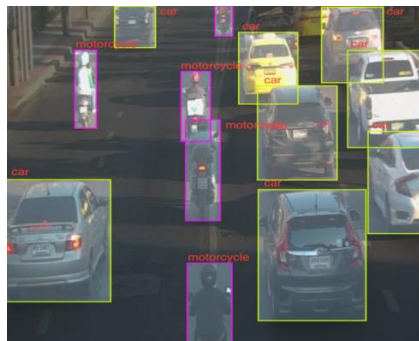
4. การทดสอบโมเดล (Model Evaluation)

ในการวัดประสิทธิภาพจะทำการบันทึก วิดีโอมุมมองการ Tracking ยานพาหนะโดยสุมบันทึก วิดีโอ มาทำการทดสอบ 3 นาที โดยทำการนับ ID ที่ Tracking ด้วย DeepSORT ในสภาพแวดล้อมจากกล้อง CCTV ดังภาพที่ 9



ภาพที่ 9 มุมมองภาพสำหรับนำไปใช้ทดสอบการ Tracking

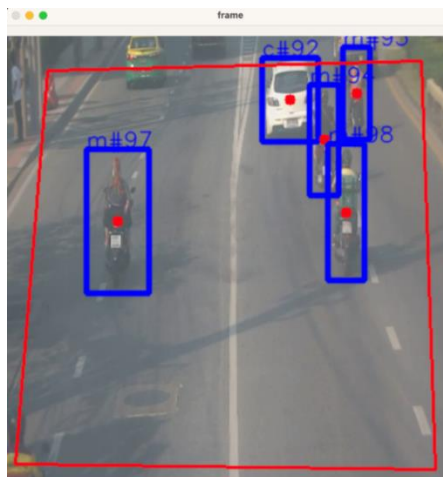
สามารถนำการติดตามวัตถุมาใช้ประโยชน์ในหลาย ๆ ด้านโดยยกตัวอย่างการตรวจนับรถ จากกล้อง CCTV บนท้องถนน เพื่อนำไปเป็นข้อมูลด้านงานจราจรหรืออื่น ๆ ที่เกี่ยวข้อง ดังภาพที่ 9



ภาพที่ 10 ตรวจจับยานพาหนะด้วย YOLOv5 แบบ Custom Dataset

5. การนำมาใช้งาน

สามารถนำการติดตามวัตถุมาใช้ประโยชน์ในหลาย ๆ ด้านโดยยกตัวอย่างการตรวจนับรถจากกล้อง CCTV บนท้องถนน เพื่อนำไปเป็นข้อมูลด้านงานจราจรหรืออื่น ๆ ที่เกี่ยวข้อง ดังภาพที่ 9



ภาพที่ 11 การนับรถด้วยการ Tracking Object แบบติดตามด้วย ID

โดยเมื่อผ่านกระบวนการ Tracking ทำให้รู้ตำแหน่งของยานพาหนะ หรือ Object นั้น ๆ แบบเฟรมต่อเฟรม ทำให้การนำไปใช้งานหรือออกแบบ Applications ต่าง ๆ นั้นนำไปประยุกต์ใช้ได้หลากหลาย จากภาพที่ 8 จะเห็นได้ว่าสามารถอ้างอิงโดยใช้ ID ในการติดตามวัตถุนั้น ๆ จนกว่าจะออกจากกรอบสีแดง หรือ Region of interest (ROI) ในภาพที่ 11

ผลการวิจัย

การหา Mean Average Precision ด้วยการใช้ Non-Maximum Suppression

ตารางที่ 1 การวัดผลจากการตรวจจับ Object Detection (Custom) ที่ได้จากโมเดล

YOLOv5s Model	Tracking
Precision	0.94302
Recall	0.83672
Mean Average Precision (mAP)	0.91722

เก็บข้อมูลภาพจากกล้อง CCTV มุมมองการตรวจนับยานพาหนะ จากรูปที่ 7 เป็น วิดีโอ ไฟล์ 3 ไฟล์สำหรับการทดสอบความแม่นยำในการ Tricking ดังนี้

- ไฟล์ No.1 ช่วงเวลาบันทึก 10.30-10.32 น. เวลา 2 นาที
- ไฟล์ No.2 ช่วงเวลาบันทึก 16.00-16.02 น. เวลา 1 นาที
- ไฟล์ No.3 ช่วงเวลาบันทึก 20.30-20.32 น. เวลา 2 นาที

ตารางที่ 2 ผลของการ Tracking

Ground truth			Tracking			
No.	Motorcycle	Car	Motorcycle	Car	Loss	Accuracy
1.	28	41	27	39	3	0.95
2.	8	20	7	20	1	0.96
3.	22	18	16	15	9	0.77

ผลการ Tracking โดยการสุ่มจากการบันทึก วิดีโอ จะมีค่าความแม่นยำ Accuracy ที่ 0.89 หรือ 89% ในสภาพแวดล้อม รถวิ่ง จากมุมมองด้านหลัง

สรุปและอภิปรายผลการวิจัย

การติดตามแบบหลายวัตถุ นั้น ได้ทำการ Custom Dataset (มุมมองการติดตามรถ ดังภาพที่ 9) ทั้งกลางวัน และกลางคืน สำหรับนำไปเทรนนิ่งในโมเดลการตรวจจับวัตถุ ด้วย YOLOv5 ใน 3 Class ประกอบไปด้วย 1. Car (c), 2. Motorcycle (m), 3. Truck (t) มีค่าความแม่นยำ mAP ที่ 0.91722 แสดงให้เห็นว่าประสิทธิภาพในการตรวจจับวัตถุ เพื่อนำผลลัพธ์ไปเข้า โมเดล Custom DeepSORT ต่อไป

ทำการ Custom DeepSORT โดยตัดส่วนที่เป็นการหา Feature ในภาพออก จากนั้นเปลี่ยนเป็น ค่า IOU และ Aspect ratio ของ IOU แต่ละตัวแทนใน ID ที่อยู่ใน Array of Track มา Matching Cascade และนำไปหา Cost Matrix และ Update Array of Track โดยมีค่าความแม่นยำ สำหรับ Dataset ไฟล์วิดีโอจากตารางที่ 2 ที่ 89%

ในการนำไปใช้งานนั้น สร้าง Application การติดตามยานพาหนะ เพื่อนับยานพาหนะหรือนำไปใช้ประโยชน์ในด้านอื่น ๆ ต่อไป ดังภาพที่ 11

ข้อเสนอแนะ

ในการศึกษาด้านการตรวจจับวัตถุและการติดตามแบบหลายวัตถุนั้นโดยใช้การเรียนรู้เชิงลึกและโมเดล DeepSORT สามารถเพิ่มประสิทธิภาพของโมเดลให้มี fps เพิ่มขึ้นไปอีกโดยการแปลงโมเดลการตรวจจับ Object Tracking ให้อยู่ในรูปแบบ ONNX เพื่อเป็นส่วนในการช่วยประมวลผลของ OpenVINO โมเดลจะทำให้การ Tracking มี ประสิทธิภาพสูงมากยิ่งขึ้น โดย OpenVINO เป็นชุดเครื่องมือฟรีที่ช่วยอำนวยความสะดวกในการเพิ่มประสิทธิภาพของโมเดลการเรียนรู้เชิงลึกจากเฟรมเวิร์ก และการปรับใช้โดยใช้ Engine บนฮาร์ดแวร์ของ Intel

บรรณานุกรม

Wojke, Bewley & Paulus. (2017). DeepSORT: Simple online and realtime tracking with a deep association metric. IEEE International Conference on Image Processing (ICIP)

Yunhao Du, Y. S., Bo Yang & Yanyun Zhao, (2022). StrongSORT: Make DeepSORT Great Again. arXiv:2202.13514 [cs.CV] Feb 2022

Bansal, K. S., G. S., Chellappa & Divakaran. (2018). Zero-Shot Object Detection. arXiv:1804.04340 [cs.CV]

Radford, J. W. K., C. H., A. R., G. G., S. A., G. S., A. A., P. M., J. C., Krueger & Sutskever. (2021). Learning Transferable Visual Models from Natural Language Supervision. arXiv:2103.00020 [cs.CV]

Redmon, S. D, Girshick & Farhadi. (2016). You Only Look Once: Unified, Real-Time Object Detection. arXiv:1506.02640 [cs.CV]

Voigtlaender, M. K., A. O., J. L., B. B. G. S., Geiger & Leibe. MOTs: Multi-Object Tracking and Segmentation. arXiv:1902.03604 [cs.CV]