

ระบบตรวจจับบุคคลและยานพาหนะแบบเรียลไทม์โดยใช้ดีพีแอลเอ็นอิงสำหรับมหาวิทยาลัยอัจฉริยะ

ภาณุพงศ์ แซ่โก

TNII

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

บัณฑิตศึกษา สาขาวิชาเทคโนโลยีสารสนเทศ

สถาบันเทคโนโลยีไทย-ญี่ปุ่น

ปีการศึกษา 2568

REAL-TIME PERSON AND VEHICLE DETECTION SYSTEM USING DEEP LEARNING FOR  
SMART UNIVERSITY

Panupong Saego

TNII

A Thesis Submitted in Partial Fulfillment of Requirements  
for the Degree of Master of Science Program in Information Technology  
Graduate Studies

Thai-Nichi Institute of Technology

Academic Year 2025

หัวข้อวิทยานิพนธ์

ระบบตรวจจับบุคคลและยานพาหนะแบบเรียลไทม์

โดยใช้ดีพเลิร์นนิงสำหรับมหาวิทยาลัยอัสสัมชัญ

โดย

ภาณุพงศ์ แซ่โก

สาขาวิชา

เทคโนโลยีสารสนเทศ

อาจารย์ที่ปรึกษาวิทยานิพนธ์

ดร.ศรายุทธ นนท์ศิริ

บัณฑิตศึกษา สถาบันเทคโนโลยีไทย-ญี่ปุ่น อนุมัติให้หัวข้อวิทยานิพนธ์ฉบับนี้เป็น ส่วนหนึ่ง  
ของการศึกษาตามหลักสูตรปริญญาโทบัณฑิต

.....รองอธิการบดีฝ่ายวิชาการ

(รองศาสตราจารย์ ดร.วรากร ศรีเชวงทรัพย์)

วันที่.....เดือน.....พ.ศ.....

คณะกรรมการสอบวิทยานิพนธ์

.....ประธานกรรมการ

(รองศาสตราจารย์ ดร.อรรณพ หมั่นสกุล)

.....กรรมการ

(ดร.ประมุข บุญเสียง)

.....กรรมการ

(ว่าที่ร้อยตรี ดร.พิชิตชัย คำอินทร์)

.....อาจารย์ที่ปรึกษาวิทยานิพนธ์

(ดร.ศรายุทธ นนท์ศิริ)

งานพุงค์ แซ่โก: ระบบตรวจจับบุคคลและยานพาหนะแบบเรียลไทม์โดยใช้ดีพีเลอร์นิง  
สำหรับมหาวิทยาลัยอัจฉริยะ. อาจารย์ที่ปรึกษา: ดร.ศรายุทธ นนท์ศิริ, 43 หน้า.

ความซับซ้อนที่เพิ่มขึ้นของการบริหารจัดการฝูงชนในพื้นที่สาธารณะ รวมถึงสถานการณ์  
การแพร่ระบาดของโควิด 19 ต้องอาศัยแนวทางที่เป็นนวัตกรรมมากยิ่งขึ้น มหาวิทยาลัยซึ่งเป็นศูนย์  
รวมกิจกรรมของผู้คนเผชิญความท้าทาย โดยเฉพาะในบริเวณโรงอาหารและลานจอดรถ งานวิจัยนี้  
ประยุกต์ใช้อัลกอริทึมการตรวจจับวัตถุที่ล้ำสมัยกับการบริหารจัดการฝูงชนในมหาวิทยาลัย โดยใช้  
YOLO เพื่อให้สามารถตรวจจับบุคคลในโรงอาหารและยานพาหนะในลานจอดรถได้อย่างแม่นยำ  
ระบบที่นำเสนอมีเป้าหมายเพื่อยกระดับความปลอดภัยภายในมหาวิทยาลัย เพิ่มประสิทธิภาพการ  
จัดสรรทรัพยากร และเสริมประสบการณ์โดยรวมของนักศึกษาผ่านระบบตรวจจับบุคคลและ  
ยานพาหนะ

อัลกอริทึมที่ใช้ในการประเมินประกอบด้วย YOLOv11n (nano), YOLOv11s (small),  
YOLOv11m (medium) และมีการเปรียบเทียบกับเวอร์ชันก่อนหน้า ได้แก่ YOLOv8 และ  
YOLOv10 ผลการทดลองแสดงให้เห็นว่า YOLOv8m ให้ความแม่นยำสูงสุดที่ 89.5% ในพื้นที่โรง  
อาหารบริเวณที่ 1 ขณะที่ YOLOv11s ทำได้ 85% ในพื้นที่โรงอาหารบริเวณที่ 2 สำหรับการตรวจจับ  
ยานพาหนะ โมเดลทั้งหมดให้ความแม่นยำ 90%

บัณฑิตศึกษา

สาขาวิชา เทคโนโลยีสารสนเทศ

ปีการศึกษา 2568

ลายมือชื่อนักศึกษา .....

ลายมือชื่ออาจารย์ที่ปรึกษา .....

PANUPONG SAEGO: REAL-TIME PERSON AND VEHICLE DETECTION SYSTEM USING DEEP LEARNING FOR SMART UNIVERSITY. ADVISOR: DR. SARAYUT NONSIRI, 43 PP.

The increasing complexity of crowd management in public spaces, including COVID-19 pandemic, needs innovative solutions. Universities, as hubs of human activity, face unique challenges in managing large crowds, particularly in cafeterias and parking lots. This research applied the cutting-edge object detection algorithm to crowd management in university using YOLO, by accurately detecting individuals in cafeterias and vehicles in parking lots. The proposed system aims to improve campus safety, optimize resource allocation, and enhance the overall student experience with human and vehicle detect systems. The different algorithms were evaluated such as YOLOv11n (nano), YOLOv11s (small), YOLOv11m (medium) and be compared with previous versions, including YOLOv8 and YOLOv10. The results shown YOLOv8m achieved the highest accuracy of 89.5% in 1st area of cafeterias, while YOLOv11s reached 85% in 2nd area of cafeterias. For vehicle detection, all models achieved an accuracy of 90%.

Graduate Studies

Field of Study Information Technology

Academic Year 2025

Student's Signature.....

Advisor's Signature.....

## กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้ด้วยความกรุณาและการสนับสนุนจากหลายฝ่ายเป็นอย่างยิ่ง โดยเฉพาะอย่างยิ่ง ดร.ศรายุทธ นนท์ศิริ อาจารย์ที่ปรึกษา ซึ่งได้มอบความรู้ คำแนะนำ แนวทางในการทำวิจัย ตลอดจนความช่วยเหลือในทุกขั้นตอนของการดำเนินงาน ผู้วิจัยขอกราบขอบพระคุณมา ณ โอกาสนี้

ขอกราบขอบพระคุณ รองศาสตราจารย์ ดร.อรรณพ หมั่นสกุล ดร.ประมุข บุญเสียง และ ว่าที่ร้อยตรี ดร.พิชิตชัย คำอินทร์ ในฐานะประธานและคณะกรรมการสอบวิทยานิพนธ์ ที่ได้ให้ข้อเสนอแนะ อันเป็นประโยชน์ต่อการปรับปรุงเนื้อหาและคุณภาพของวิทยานิพนธ์ฉบับนี้ให้สมบูรณ์ยิ่งขึ้น

ขอแสดงความขอบพระคุณต่อ คณาจารย์ทุกท่านของสถาบันเทคโนโลยีไทย-ญี่ปุ่น ที่ได้ถ่ายทอดองค์ความรู้แก่ผู้วิจัย และ บุคลากรของคณะเทคโนโลยีสารสนเทศ ที่ให้คำปรึกษาเกี่ยวกับเอกสาร ตลอดจนความช่วยเหลือในกระบวนการต่าง ๆ ที่เกี่ยวข้องกับการดำเนินงานของบัณฑิตวิทยาลัย

ขอขอบคุณ ครอบครัวอันเป็นที่รัก และ เพื่อน ๆ สำหรับกำลังใจ คำแนะนำ และแรงสนับสนุนที่มีคุณค่า ทำให้ผู้วิจัยสามารถดำเนินการวิจัยจนแล้วเสร็จสมบูรณ์ตามเป้าหมาย

ท้ายที่สุด ผู้วิจัยหวังเป็นอย่างยิ่งว่า ผลงานวิจัยฉบับนี้จะเป็นประโยชน์ต่อผู้ที่สนใจ และสามารถนำไปประยุกต์ใช้ต่อยอดในด้านต่าง ๆ ที่เกี่ยวข้องได้อย่างมีประสิทธิภาพและเกิดประโยชน์สูงสุดในอนาคต

ภาณุพงศ์ แซ่โก

## สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ง
บทคัดย่อภาษาอังกฤษ.....	จ
กิตติกรรมประกาศ.....	ฉ
สารบัญ.....	ช
สารบัญตาราง.....	ฎ
สารบัญรูป.....	ฉ
บทที่	
1 บทนำ.....	1
1.1 ความเป็นมาและความสำคัญของปัญหา.....	1
1.2 วัตถุประสงค์.....	3
1.2.1 เพื่อพัฒนาระบบในการตรวจจับบุคคลที่เข้ามาใช้โรงอาหาร.....	3
1.2.2 เพื่อพัฒนาระบบในการตรวจจับยานพาหนะที่เข้ามาจอดในที่จอดรถ.....	3
1.2.3 เพื่อแสดงผลลัพธ์ที่ได้จากการตรวจจับในรูปแบบแดชบอร์ด.....	3
1.3 ขอบเขตการดำเนินงาน.....	3
1.3.1 สร้างระบบตรวจจับบุคคลบริเวณโรงอาหาร และ ตรวจจับรถยนต์บริเวณ หน้าทางเข้าสถาบันเทคโนโลยีไทย-ญี่ปุ่น.....	3
1.3.2 ใช้อัลกอริทึม YOLO ในการตรวจจับวัตถุ.....	3
1.4 ประโยชน์ที่คาดว่าจะได้รับ.....	3
1.4.1 สามารถตรวจจับและเก็บข้อมูลจำนวนผู้เข้าใช้ บริการโรงอาหารผ่านภาพวิดีโอได้.....	3
1.4.2 สามารถตรวจจับและเก็บข้อมูลจำนวนยานพาหนะ ในที่จอดรถผ่านภาพวิดีโอ.....	3
1.4.3 สามารถแสดงผลลัพธ์ที่ได้จากการตรวจจับในรูปแบบแดชบอร์ดได้.....	3
2 หลักการ ทฤษฎีและงานวิจัยที่เกี่ยวข้อง.....	4
2.1 ปัญญาประดิษฐ์.....	4

## สารบัญ (ต่อ)

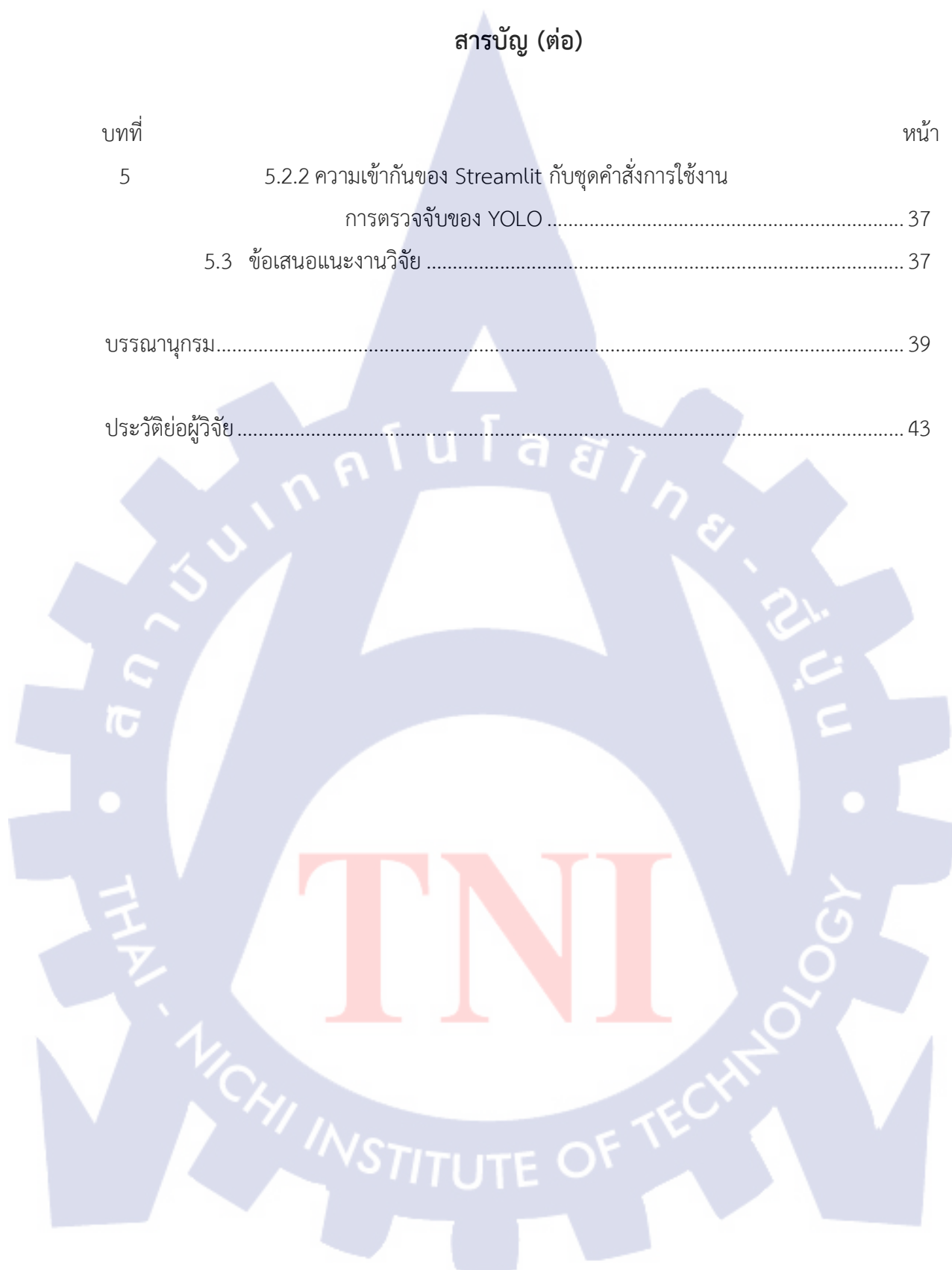
บทที่		หน้า
2	2.2 การเรียนรู้ของเครื่อง (MACHINE LEARNING).....	5
	2.3 การเรียนรู้เชิงลึก (DEEP LEARNING).....	6
	2.4 วิธีโครงข่ายประสาทแบบคอนโวลูชัน (CONVOLUTIONAL NEURAL NETWORKS).....	6
	2.4.1 ระดับชั้นของคอนโวลูชัน (Convolutional layer).....	6
	2.4.2 Max pooling .....	7
	2.4.3 Average pooling .....	8
	2.4.4 ANN for classification .....	8
	2.5 การตรวจจับวัตถุ (OBJECT DETECTION).....	9
	2.6 YOLO.....	9
	2.7 YOLOV5.....	10
	2.8 YOLOV6.....	10
	2.9 YOLOV7.....	11
	2.10 YOLOV8.....	11
	2.11 YOLOV9.....	11
	2.11.1 Programmable Gradient Information (PGI) .....	11
	2.11.2 Generalized Efficient Layer Aggregation Network (GELAN).....	12
	2.12 YOLOv10.....	12
	2.12.1 NMS-Free Training .....	12
	2.12.2 Enhanced Model Capabilities .....	12
	2.13 YOLO11.....	14
	2.14 งานวิจัยที่เกี่ยวข้อง .....	14
3	วิธีดำเนินการวิจัย .....	18
	3.1 การเตรียมข้อมูล .....	18
	3.1.1 วิดีโอบริเวณโรงอาหารที่หนึ่งความยาว 3 นาที.....	18
	3.1.2 วิดีโอบริเวณโรงอาหารที่สองความยาว 3 นาที .....	19
	3.1.3 วิดีโอบริเวณหน้าต่างเข้าสถาบันความยาว 3 นาที 30 วินาที.....	20

## สารบัญ (ต่อ)

บทที่		หน้า
3	3.2 เครื่องมือที่นำมาใช้.....	20
	3.3 การตรวจจับวัตถุด้วยโมเดล YOLO11.....	20
	3.3.1 การนำโมเดลของ YOLO11 มาใช้ในการตรวจจับ.....	20
	3.3.2 การเก็บข้อมูลที่ได้จากการตรวจจับ .....	21
	3.4 การประเมินประสิทธิภาพในการตรวจจับ.....	23
	3.4.1 Precision .....	23
	3.4.2 Recall .....	23
	3.4.3 F1-score .....	24
	3.4.4 Intersection over Union (IoU).....	24
	3.4.5 mAP .....	24
	3.5 การสร้างเว็บไซต์สำหรับแสดงผลการตรวจจับ.....	25
	3.5.1 เลือกประเภทของวัตถุที่ต้องการตรวจจับระหว่างบุคคลและยานพาหนะ. 26	
	3.5.2 กรณีเลือกยานพาหนะจะมีวิธีให้เลือกเพิ่มระหว่างการตรวจจับแบบ	
	1 ทิศทาง หรือ 2 ทิศทาง.....	26
	3.5.3 กรณีเลือกการตรวจจับแบบ 2 ทิศทาง จะมีตัวเลือกเพิ่มเติมคือ	
	ตรวจจับแบบแนวนอนหรือแนวตั้ง.....	27
	3.5.4 เลือกโมเดลในการตรวจจับ.....	27
	3.5.5 เลือกช่วงเวลาในการจัดเก็บและสรุปข้อมูล.....	28
	3.5.6 เลือกประเภทของการทำงาน.....	28
4	ผลการวิจัย.....	32
5	บทสรุปและข้อเสนอแนะ .....	36
	5.1 สรุปและอภิปรายผลการวิจัย.....	36
	5.2 ปัญหาและอุปสรรคในการทำวิจัย .....	37
	5.2.1 ปัญหาและปัจจัยด้านอุปกรณ์และการหาข้อมูล .....	37

สารบัญ (ต่อ)

บทที่		หน้า
5	5.2.2 ความเข้ากันของ Streamlit กับชุดคำสั่งการใช้งาน การตรวจจับของ YOLO .....	37
5.3	ข้อเสนอแผนงานวิจัย .....	37
	บรรณานุกรม.....	39
	ประวัติย่อผู้วิจัย.....	43



## สารบัญตาราง

ตาราง		หน้า
2.1	แสดงงานวิจัยที่เกี่ยวข้อง.....	16
4.1	แสดงความแม่นยำในการตรวจจับในแต่ละสถานที่.....	33
4.2	แสดงค่า precision ในการตรวจจับในแต่ละสถานที่.....	33
4.3	แสดงค่า recall ในการตรวจจับในแต่ละสถานที่.....	34
4.4	แสดงค่า F1-score ในการตรวจจับในแต่ละสถานที่.....	34
4.5	แสดงความสามารถในการประมวลผลโดยเฉลี่ยในแต่ละโมเดล .....	35



## สารบัญรูป

รูป	หน้า
2.1 แสดงการเปลี่ยนตำแหน่งของฟิลเตอร์.....	7
2.2 การทำพูลลิงแบบค่ามากที่สุด.....	8
2.3 การทำพูลลิงแบบหาค่าเฉลี่ย.....	8
2.4 แสดงกระบวนการแบ่งกลุ่มของ ANN.....	9
2.5 แสดงการวางกริดของ YOLO .....	10
2.6 แสดงลำดับการทำงานของ Large-kernel convolution [23] .....	13
2.7 แสดงลำดับการทำงานของ Partial self-attention (PSA) [23] .....	13
3.1 ภาพบริเวณโรงอาหารที่หนึ่ง .....	18
3.2 ภาพบริเวณโรงอาหารที่สอง .....	19
3.3 ภาพบริเวณหน้าต่างเข้าสถาบัน .....	20
3.4 พื้นที่การตรวจจับบริเวณโรงอาหารที่หนึ่ง.....	22
3.5 พื้นที่การตรวจจับบริเวณโรงอาหารที่สอง .....	22
3.6 พื้นที่การตรวจจับบริเวณหน้าต่างเข้าสถาบัน.....	23
3.7 หน้าจอหลักของเว็บไซต์ .....	25
3.8 เมนูการเลือกประเภทของวัตถุ.....	26
3.9 เมนูการเลือกประเภททิศทางการตรวจจับยานพาหนะ .....	26
3.10 เมนูการเลือกประเภทระนาบในการตรวจจับยานพาหนะ.....	27
3.11 เมนูการเลือกโมเดลในการตรวจจับ .....	27
3.12 เมนูการเลือกช่วงเวลาในการเก็บข้อมูล.....	28
3.13 เมนูการเลือกประเภทของงาน.....	28
3.14 หน้าจอการกำหนดพื้นที่ในการตรวจจับของยานพาหนะแบบ 1 ทิศทาง .....	29
3.15 หน้าจอการกำหนดพื้นที่ในการตรวจจับทั่วไป .....	30
3.16 หน้าจอแสดงผลการตรวจจับและแดชบอร์ด.....	31

## บทที่ 1

### บทนำ

#### 1.1 ความเป็นมาและความสำคัญของปัญหา

การรวมตัวของฝูงชนจำนวนมากในที่สาธารณะก่อให้เกิดความท้าทายต่อการดูแลและกลายเป็นข้อกังวลที่สำคัญ การจัดการฝูงชนไม่เพียงพอก่อให้เกิดผลกระทบต่อความปลอดภัยต่อสาธารณะ ประชาชนอาจได้รับบาดเจ็บจากความแออัดอย่างการกระแทกไปมา การเหยียบเท้าผู้อื่น การขาดอากาศหายใจจนเป็นลมหรือเสียชีวิต นอกจากการบาดเจ็บที่เกิดกับร่างกายโดยตรง การรวมตัวของฝูงชนจำนวนมากก่อให้เกิดผลกระทบต่อสภาพจิตใจ [1] นอกจากนี้ด้วยสถานการณ์การแพร่ระบาดของเชื้อไวรัสโคโรนา 2019 ทำให้ความต้องการศึกษาและเข้าใจพฤติกรรมของฝูงชนเพิ่มมากขึ้น ทั้งนี้จากการพัฒนาอย่างต่อเนื่องของเทคโนโลยีและนวัตกรรม การนำมาปรับใช้ให้เกิดประโยชน์สูงสุดเป็นสิ่งต้องคึกหรือผู้เกี่ยวข้องต่างให้ความสำคัญ มหาวิทยาลัยซึ่งเปรียบเสมือนหัวใจสำคัญในการขับเคลื่อนสังคม จำเป็นต้องควบคุมและดูแลกลุ่มคนจำนวนมาก และมีบทบาทสำคัญอยู่สามประการ คือ การจัดการศึกษาระดับสูงสุด การวิจัยทางวิชาการและการเผยแพร่ผลการวิจัย และการบริการสังคม มหาวิทยาลัยอัจฉริยะ (Smart University) เป็นเป้าหมายหนึ่งที่มหาวิทยาลัยชั้นนำต่างมุ่งหวังที่จะเป็นเพื่อยกระดับการเรียนรู้ของนักศึกษาให้สามารถกลายเป็นบุคคลที่เป็นประโยชน์แก่สังคมต่อไป ซึ่งเป็นการบูรณาการของระบบทางกายภาพ ดิจิทัล และระบบสังคมในการสร้างสภาพแวดล้อมทางสังคมที่ยั่งยืน [2] โดยใจความสำคัญหนึ่งที่ทำให้มหาวิทยาลัยอัจฉริยะแตกต่างจากมหาวิทยาลัยแบบดั้งเดิม คือ การเปลี่ยนแปลงทางดิจิทัล (Digital Transformation) ซึ่งไม่ได้หมายถึงการแปลงข้อมูลทางกายภาพให้เป็นข้อมูลดิจิทัลเพียงอย่างเดียว แต่เป็นการนำเอาเทคโนโลยีและกลยุทธ์การทำงานแบบดิจิทัลเข้ามาใช้ในการวางรากฐาน กำหนดเป้าหมาย ไปจนถึงปรับใช้กับขั้นตอนการทำงานและการขับเคลื่อนวัฒนธรรมองค์กร โดยเทคโนโลยีอย่างการเรียนรู้ผ่านสื่ออิเล็กทรอนิกส์ (E-learning) ช่วยให้ผู้เรียนสามารถเรียนรู้จากที่ไหนก็ได้ผ่านอินเทอร์เน็ต อินเทอร์เน็ตในทุกสิ่ง (Internet of Things) มีการนำเครือข่ายอุปกรณ์ทางกายภาพที่สามารถถ่ายโอนข้อมูลระหว่างกันโดยไม่ต้องมีมนุษย์ในการควบคุมหรือออกคำสั่ง เช่น ระบบเซ็นเซอร์ที่สามารถตรวจสอบ อุณหภูมิ การเคลื่อนไหว แสง และเสียงได้ เครือข่ายอินเทอร์เน็ตและซอฟต์แวร์สำหรับประมวลผลและวิเคราะห์ข้อมูล คลาวด์คอมพิวติ้ง (Cloud Computing) บริการที่ช่วยประมวลผลและจัดเก็บข้อมูลผ่านระบบออนไลน์และข้อมูลขนาดใหญ่ (Big Data) ข้อมูลจำนวนมากที่นำมาประมวลผลเพื่อหาความสัมพันธ์ระหว่าง คณาจารย์ ผู้เรียน และสภาพแวดล้อมของมหาวิทยาลัย ผ่านปัญญาประดิษฐ์ (Artificial Intelligence) เพื่อเข้าใจข้อมูลเชิงลึกหรือรูปแบบพฤติกรรมนำไปสู่การปรับปรุงรูปแบบการสอนของ

คณาจารย์และหลักสูตรให้ตรงกับความต้องการของผู้เรียน และทันสมัย รวมไปถึงการปรับปรุงสภาพแวดล้อมของมหาวิทยาลัยในการสร้างความรู้สึกรักเชิงบวกแก่ผู้ที่เข้ามาใช้พื้นที่ [3], [4]

ปัญญาประดิษฐ์ได้รับการพัฒนาอย่างต่อเนื่องและส่งผลต่อชีวิตประจำวัน ตัวอย่างเช่น โทรศัพท์มือถือที่นำปัญญาประดิษฐ์มาปรับปรุงคุณภาพและคุณสมบัติของกล้อง เช่น โหมดแนวตั้ง โหมดกลางคืน ความละเอียดขั้นสูง และการจดจำใบหน้า การสร้างรถยนต์ไร้คนขับที่ใช้ปัญญาประดิษฐ์ผ่านการประมวลผลและวิเคราะห์ข้อมูลเส้นทาง สภาพแวดล้อมผ่านเซ็นเซอร์ในยานพาหนะ เช่น กล้อง อุปกรณ์ตรวจจับคลื่นแสง เรดาร์ก่อนทำการตัดสินใจเคลื่อนที่ ระบบแนะนำสิ่งที่น่าสนใจในสื่อสังคมออนไลน์โดยอาศัยข้อมูลพฤติกรรมการใช้งานและใช้ปัญญาประดิษฐ์หาความสัมพันธ์เพื่อแนะนำเนื้อหาที่ใกล้เคียงกับพฤติกรรมหรือลักษณะนิสัยของผู้ใช้งานมากที่สุด การเรียนรู้ของเครื่อง (Machine learning) เป็นแนวทางที่ถูกนำมาใช้อย่างแพร่หลายในการพัฒนาระบบปัญญาประดิษฐ์ เพื่อช่วยให้ระบบเกิดความสามารถในการรู้จำคล้ายคลึงกับมนุษย์ การสอนระบบให้รับรู้ถึงรูปแบบที่เกิดขึ้นตามการป้อนข้อมูลลงไป ทั้งนี้ความสามารถในการเรียนรู้ของเครื่องจะขึ้นอยู่กับความซับซ้อนของข้อมูล จำนวนข้อมูลที่นำมาใช้ และรูปแบบที่นำมาใช้ในการเรียนรู้ การเรียนรู้เชิงลึก (Deep learning) เป็นวิธีที่ได้รับความนิยมในการสอนระบบให้เข้าใจรูปแบบของข้อมูลที่นำมาฝึกสอนผ่านระบบโครงข่ายประสาท (Neural network) ที่ได้รับแรงบันดาลใจจากสมองมนุษย์ ส่งผลให้คอมพิวเตอร์เกิดการเรียนรู้ข้อผิดพลาดและปรับปรุง อย่างไรก็ตามวัตถุประสงค์ที่แตกต่างกันทำให้เกิดระบบโครงข่ายประสาทรูปแบบใหม่ที่แตกต่างกัน เช่น โครงข่ายประสาทแบบคอนโวลูชัน (Convolutional neural network) ซึ่งมักถูกนำมาสร้างระบบตรวจจับวัตถุและระบุสิ่งที่ตรวจพบ [5]

ระบบตรวจจับบุคคลถูกพัฒนาอย่างต่อเนื่องและมีความแม่นยำสูง [6] ทำให้ถูกนำมาใช้อย่างแพร่หลาย อย่างการตรวจจับบุคคลเพื่อตรวจหาพฤติกรรมและความต้องการของมนุษย์ [7] การตรวจจับบุคคลที่น่าสงสัยที่อาจก่อเหตุร้ายและทำการแจ้งเจ้าหน้าที่รักษาความปลอดภัยให้พิจารณาเพิ่มเติมเพื่อแก้ไขสถานการณ์ได้อย่างทันท่วงที [8] การตรวจหาผู้มีความเสี่ยงต่อการติดเชื้อ COVID-19 [9] การใช้อากาศยานไร้คนขับตรวจสอบความหนาแน่นของฝูงชน อย่างไรก็ตามปัญหาหนึ่งที่น่าสนใจ คือ การตรวจจับผู้ใช้บริการร้านอาหารซึ่งโดยทั่วไป ต้องรองรับผู้ใช้บริการเป็นจำนวนมากส่งผลให้เกิดความท้าทายในการบริหารจัดการ หากมีการบริหารจัดการที่ผิดพลาดจะทำให้เกิดความวุ่นวายในร้านอาหาร การขาดแคลนอุปกรณ์ จำนวนพนักงานทำความสะอาดไม่สอดคล้องกับจำนวนผู้ใช้บริการในแต่ละช่วงเวลา ส่งผลให้ผู้ใช้บริการร้านอาหารลดน้อยลงและเป็นภาพลักษณ์ที่ไม่ดีแก่สถาบัน ทั้งนี้หากฝูงชนมีความแออัดสูงจะกระทบต่อความแม่นยำและรวดเร็วในการตรวจจับ การสร้างรูปแบบการตรวจจับเพื่อเพิ่มความแม่นยำและรวดเร็วจึงเป็นความท้าทายอย่างหนึ่ง [10]

การตรวจจับยานพาหนะถูกพัฒนาและมีความแม่นยำสูงโดยอาศัยภาพจากหลายมุมมอง [11] ซึ่งสามารถนำมาตรวจจับเพื่อวิเคราะห์สภาพจราจรเพื่อสร้างระบบจัดการเส้นทางสำหรับลดความแออัดบนท้องถนน [12] การเพิ่มเงื่อนไขการตรวจจับมากขึ้นเพื่อนำมาใช้ระบุสิ่งที่ต้องการของยานพาหนะ เช่น ขนาดของยานพาหนะ ประเภทของยานพาหนะ ยี่ห้อ และ สี รวมไปถึงการตรวจจับป้ายทะเบียนรถเพื่อเก็บข้อมูลตัวอักษร [13] ความท้าทายของการตรวจจับยานพาหนะที่เข้ามาจอดในที่จอดรถกลางแจ้ง แม้เซ็นเซอร์ตรวจจับวัตถุจะสามารถนำมาใช้ตรวจจับยานพาหนะที่เข้ามาจอดในที่จอดรถใต้ การตรวจจับยานพาหนะที่เข้ามาจอดผ่านกล้องวงจรปิดจะช่วยลดค่าใช้จ่ายและเก็บข้อมูลยานพาหนะเพิ่มเติมได้ [14]

YOLO (You only look one) เป็นเครื่องมือหนึ่งที่ถูกนำมาใช้ตรวจจับวัตถุ ซึ่งถูกพัฒนาอย่างต่อเนื่องให้มีความแม่นยำในการตรวจจับวัตถุสูง โดยงานวิจัยนี้มุ่งเน้นในการตรวจจับบุคคลที่เข้ามาใช้ร้านอาหารและยานพาหนะที่เข้ามาจอดในที่จอดรถของสถาบันการศึกษาด้วยเครื่องมือ YOLO เนื่องจาก YOLO มีการพัฒนาและปรับปรุงอย่างต่อเนื่อง ทำให้เกิด YOLO เวอร์ชันใหม่ที่ถูกนำเสนอออกมา โดยเวอร์ชันหลักที่นำมาใช้ในงานวิจัยนี้ คือ YOLO11 ซึ่งเป็นเวอร์ชันล่าสุด ณ ตอนนี้และมีคุณสมบัติใหม่อย่างการตรวจจับเส้นทางการเดินทาง ตรวจจับการแสดงท่าทางของมนุษย์ เป็นต้น นอกจากนี้ยังมีการปรับปรุงประสิทธิภาพและความยืดหยุ่นให้แม่นยำ รวดเร็วและง่ายต่อการใช้งาน

## 1.2 วัตถุประสงค์

- 1.2.1 เพื่อพัฒนาระบบในการตรวจจับบุคคลที่เข้ามาใช้ร้านอาหาร
- 1.2.2 เพื่อพัฒนาระบบในการตรวจจับยานพาหนะที่เข้ามาจอดในที่จอดรถ
- 1.2.3 เพื่อแสดงผลลัพธ์ที่ได้จากการตรวจจับในรูปแบบแดชบอร์ด

## 1.3 ขอบเขตการดำเนินงาน

- 1.3.1 สร้างระบบตรวจจับบุคคลบริเวณร้านอาหาร และ ตรวจจับรถยนต์บริเวณหน้าทางเข้าสถาบันเทคโนโลยีไทย-ญี่ปุ่น
- 1.3.2 ใช้อัลกอริทึม YOLO ในการตรวจจับวัตถุ

## 1.4 ประโยชน์ที่คาดว่าจะได้รับ

- 1.4.1 สามารถตรวจจับและเก็บข้อมูลจำนวนผู้ใช้บริการร้านอาหารผ่านภาพวิดีโอได้
- 1.4.2 สามารถตรวจจับและเก็บข้อมูลจำนวนยานพาหนะในที่จอดรถผ่านภาพวิดีโอ
- 1.4.3 สามารถแสดงผลลัพธ์ที่ได้จากการตรวจจับในรูปแบบแดชบอร์ดได้

## บทที่ 2

### หลักการ ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ในงานวิจัยครั้งนี้มีความเกี่ยวข้องกับการตรวจจับผู้คนและยานพาหนะสำหรับนำข้อมูลที่ตรวจจับได้มาวิเคราะห์รูปแบบที่เป็นประโยชน์ในการจัดการฝูงชนและที่จอดรถ ทั้งนี้เพื่อให้เข้าใจถึงที่มาที่ไปและกระบวนการทำงาน ผู้วิจัยได้ศึกษาเอกสารที่เกี่ยวข้องกับปัญญาประดิษฐ์ การเรียนรู้ของเครื่อง การเรียนรู้เชิงลึก การตรวจจับวัตถุ วิธีโครงข่ายประสาทแบบคอนโวลูชัน และ YOLO เวอร์ชันต่าง ๆ รวมถึงศึกษางานวิจัยที่เกี่ยวข้องกับการตรวจจับวัตถุด้วย YOLO เพื่อใช้เป็นข้อมูลอ้างอิงในการทำวิจัย โดยมีการนำเสนอตามหัวข้อต่อไปนี้

#### 2.1 ปัญญาประดิษฐ์

การสร้างระบบซอฟต์แวร์ที่สามารถดำเนินการในระบบดิจิทัลหรือระบบทางกายภาพได้ด้วยตัวเอง [15] เช่น รถยนต์ที่สามารถขับเคลื่อนเองได้หรือระบบที่สามารถโต้ตอบกับมนุษย์ได้อย่างชาญฉลาด ทั้งนี้เนื่องจากคำว่า ความฉลาด เป็นคำที่ยากต่อการนิยาม การนิยามในส่วนของปัญญาประดิษฐ์จึงมีการนิยามหลายรูปแบบ หนึ่งในนั้นคือการนำพื้นฐานของพฤติกรรมหรือวิธีคิดแบบมนุษย์มาใช้ เช่น the Turing test ที่ถูกนำเสนอโดย Alan Turing เมื่อปี 1950 ซึ่งเป็นการทดสอบที่ต้องการระบุการกระทำที่ถูกสร้างขึ้นโดยระบบหรือหุ่นยนต์จะสามารถถูกแยกออกกว่ามาจากมนุษย์หรือซอฟต์แวร์ได้หรือไม่ อย่างไรก็ตามระบบปัญญาประดิษฐ์ส่วนใหญ่ในปัจจุบันมุ่งเน้นที่การสร้างตัวแทนที่คิดหรือกระทำอย่างมีเหตุผล ในการสร้างระบบที่คิดอย่างมีเหตุผลมักใช้ระบบการเสนอเหตุผลและระบบการให้เหตุผลตามตรรกะ

แนวคิดพื้นฐานของปัญญาประดิษฐ์ คือ คอมพิวเตอร์สามารถใช้ภาษาและแนวคิดเชิงนามธรรม หรือปรับปรุงพฤติกรรมของตนเองอย่างง่ายได้ ในช่วงทศวรรษ 1980 ปัญญาประดิษฐ์มีการเติบโตที่รวดเร็ว โดยคำถามเกี่ยวกับการเป็นตัวแทนความรู้และการประมวลผลความรู้นั้นมีบทบาทสำคัญ ซึ่งในบริบทนี้ระบบที่มีความชำนาญเริ่มได้รับความนิยมนมากขึ้น ระบบดังกล่าวใช้คลังความรู้จำนวนมากจากข้อมูลข้อเท็จจริงและกฎเกณฑ์ เพื่อสรุปผลและหาแนวทางในการแก้ไขปัญหา แม้ว่าในตอนแรกจะประสบความสำเร็จที่ดีกับระบบที่มีความเชี่ยวชาญ แต่ความสำเร็จเหล่านี้ได้ลดลงจนกระทั่งถึงช่วงทศวรรษ 1990 เมื่อวิธีการทางคณิตศาสตร์และความน่าจะเป็นเข้ามามีบทบาทมากขึ้น และทำให้เกิดการตื่นตัวในวงการปัญญาประดิษฐ์อีกครั้ง โดยวิธีที่โดดเด่นวิธีหนึ่ง คือ วิธีโครงข่ายแบบเบย์ ระบบที่เกิดจากเทคนิคนี้มีประสิทธิภาพมากกว่าระบบที่ใช้เทคนิคเชิงสัญลักษณ์เป็นอย่างมาก ในช่วงนี้เป็นจุดเริ่มต้นของเทคนิคการเรียนรู้ของเครื่องตามแนวคิดความน่าจะเป็นและ

คณิตศาสตร์ ตัวอย่างเช่น ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machines) ที่ปฏิวัติการใช้การเรียนรู้ของเครื่อง ซึ่งเป็นแนวทางที่มีประสิทธิภาพที่สุดวิธีหนึ่งในการแก้ปัญหาการจำแนกประเภท นำไปสู่แนวทางในการจดจำรูปแบบและการประมวลผลภาพ อัลกอริทึมการจดจำใบหน้า และการรู้จำคำพูดที่ถูกนำไปใช้ในผลิตภัณฑ์สำหรับชีวิตประจำวัน เช่น กล้องถ่ายรูป โทรศัพท์มือถือ การจดจำใบหน้าโดยอัตโนมัติของกล้องถ่ายรูปหรือโทรศัพท์มือถือ และโทรศัพท์มือถือที่สามารถควบคุมด้วยเสียง วิธีการเหล่านี้กับการนำมาใช้ในรถยนต์ เช่น การควบคุมอุปกรณ์ในรถยนต์ได้ด้วยคำพูด ความสามารถของระบบนำทางในการวางแผนเส้นทางที่สั้นที่สุด และนำทางไปยังจุดหมายปลายทางอย่างมีประสิทธิภาพตามแผนการที่กำหนด บทบาทสำคัญต่อวงการเกมคอมพิวเตอร์ในการจำลองระบบอัจฉริยะที่สามารถสร้างสภาพแวดล้อมเสมือนจริงได้อย่างมีประสิทธิภาพ กระบวนทัศน์ในด้านวิทยาการหุ่นยนต์ เป็นต้น วิธีความน่าจะเป็นมีผลกระทบอย่างมีนัยสำคัญต่อการเคลื่อนที่ของหุ่นยนต์ การสร้างระบบเคลื่อนที่ได้อัตโนมัติในสภาพแวดล้อมต่าง ๆ การเรียนรู้ของเครื่องและหุ่นยนต์มีแนวโน้มสำคัญในการนำมาใช้ร่วมกับข้อมูลขนาดใหญ่ การเรียนรู้เชิงลึก การนำทางและการจัดการอัตโนมัติ

## 2.2 การเรียนรู้ของเครื่อง (Machine Learning)

การเรียนรู้ของเครื่องเป็นสาขาที่มุ่งเน้นไปที่การพัฒนาอัลกอริทึมเพื่อปรับปรุงโมเดลโดยอิงตามข้อมูลจำนวนมากที่ถูกใช้ในการเรียนรู้และไม่จำเป็นต้องเขียนโปรแกรมโดยตรง บทบาทสำคัญของการเรียนรู้ของเครื่องเกี่ยวข้องกับการจำแนกประเภท วัตถุประสงค์เพื่อให้ระบบเรียนรู้ฟังก์ชันที่สามารถเชื่อมคุณสมบัติหรือคุณลักษณะของกรณีเหล่านั้นกับคลาสที่เกี่ยวข้องได้อย่างแม่นยำ และทำการลดข้อผิดพลาดในการจำแนกประเภทให้เหลือน้อยที่สุด อย่างไรก็ตามโมเดลที่ได้รับการเรียนรู้ที่มีความแม่นยำสูงมากเกินไป นอกจากจะต้องใช้คอมพิวเตอร์ที่มีประสิทธิภาพสูงและเวลาในการประมวลผล โมเดลจะมีความสามารถในการทำนายความแม่นยำกับชุดข้อมูลใหม่ที่อยู่นอกเหนือจากชุดข้อมูลเดิมได้น้อยลง ในกรณีนี้เรียกว่าการเกิด Overfitting

รูปแบบการเรียนรู้ของเครื่องสามารถแบ่งออกได้สองรูปแบบหลัก คือ การเรียนรู้แบบมีผู้สอน (Supervised learning) โดยจะเกี่ยวข้องกับการจำแนกประเภทอย่างวิธีแผนผังการตัดสินใจ (Decision trees) การสุ่มแบบฟอเรสต์ (Random forest) ลำดับพจนานุกรม (Lexicographic order) ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machines) และ Boosting การเรียนรู้อีกรูปแบบหนึ่ง คือ การเรียนรู้แบบไม่มีผู้สอน (Unsupervised learning) มีเป้าหมายเพื่อหาฟังก์ชันที่สามารถอธิบายข้อมูลที่ได้รับมาได้อย่างแม่นยำโดยอาศัยการจัดกลุ่ม (Clustering)

## 2.3 การเรียนรู้เชิงลึก (Deep Learning)

การเรียนรู้เชิงลึกเป็นวิธีการแก้ปัญหาที่มีความแม่นยำและรวดเร็วสำหรับการเรียนรู้ของเครื่องโดยการเรียนรู้จากชุดข้อมูลขนาดใหญ่ วิธีที่แพร่หลายที่สุดในการเรียนรู้เชิงลึกคือ โครงข่ายประสาทเทียมเชิงลึก วิธีโครงข่ายประสาทเทียมสามารถย้อนกลับไปถึงงานของ Warren McCulloch และ Walter Pitts ในปี 1943 เมื่อพวกเขาพยายามเลียนแบบการทำงานของเซลล์ประสาทในสมอง โดยใช้วงจรอิเล็กทรอนิกส์ทำให้เกิดสิ่งที่เรียกว่าโครงข่ายประสาทเทียม (Artificial neural networks หรือ ANN)

หลักการดำเนินการของโครงข่ายประสาทเทียม คือ การสร้างเครือข่ายที่ประกอบด้วยชั้นของโหนดที่เชื่อมต่อถึงกัน ในโครงสร้างของเครือข่าย ระดับชั้นแรกสุดทำหน้าที่เป็นระดับชั้นในการนำเข้าข้อมูล ในขณะที่ระดับชั้นสุดท้ายทำหน้าที่เป็นระดับชั้นของผลลัพธ์ ผลลัพธ์ที่สร้างในแต่ละโหนดภายในระดับชั้นจะถูกส่งไปยังโหนดในระดับชั้นถัดไป และการส่งข้อมูลนี้จะถูกส่งไปพร้อมกับผลรวมของค่าถ่วงน้ำหนัก โครงข่ายประสาทเทียมมีบทบาทสำคัญในการเรียนรู้ของเครื่อง จนกระทั่งถูกแทนที่ด้วยวิธีของความน่าจะเป็น เช่น เครือข่ายแบบเบย์เซียน ชัฟฟอร์ตเวกเตอร์แมชชีน หรือ กระบวนการแบบเกาส์เซียน โครงข่ายประสาทเทียมเชิงลึก วิธีการเหล่านี้มีส่วนสำคัญในด้านการเรียนรู้ของเครื่องและนำมาใช้สร้างโมเดลในการทำนาย

## 2.4 วิธีโครงข่ายประสาทแบบคอนโวลูชัน (Convolutional Neural Networks)

วิธีโครงข่ายประสาทแบบคอนโวลูชันหรือเรียกโดยย่อในภาษาอังกฤษว่า CNN มีความโดดเด่นในฐานะโมเดลโครงข่ายประสาทเทียมเฉพาะที่ออกแบบมาสำหรับประมวลผลข้อมูลภาพสองมิติ และสามารถนำมาปรับใช้กับข้อมูลมิติเดียวและสามมิติได้ด้วยเช่นกัน โครงข่ายประสาทเทียมประเภทนี้ยังสามารถใช้งานได้ดีกับข้อมูลที่เป็นข้อมูลตามลำดับ (Sequential Data)

### 2.4.1 ระดับชั้นของคอนโวลูชัน (Convolutional layer)

เกิดจากการคูณระหว่างอาร์เรย์ของข้อมูลนำเข้า (Input) กับอาร์เรย์ของค่าน้ำหนัก (weights) แบบสองมิติ โดยอาร์เรย์ของค่าน้ำหนักนี้ถูกเรียกว่าตัวกรอง (filter) หรือเคอร์เนล (kernel) โดยทั่วไปตัวกรองที่ใช้ในโครงข่ายประสาทเทียมจะมีขนาดเล็กกว่าเมื่อเทียบกับข้อมูลนำเข้า การคูณระหว่างส่วนของข้อมูลนำเข้าที่ถูกกรองให้มีขนาดเท่ากับตัวกรองกับตัวกรองเองนั้นเรียกว่าจุดที่เกิดผลิตภัณฑ์ (dot product) การดำเนินการนี้มักเรียกว่าผลคูณสเกลาร์เนื่องจากคุณสมบัติของผลลัพธ์ที่ได้เป็นค่าสเกลาร์เดียวเสมอ การใช้ตัวกรองที่มีขนาดเล็กลงโดยเจตนาถือเป็นกลยุทธ์อย่างหนึ่ง เนื่องจากทำให้ตัวกรองเดียวกันซึ่งในที่นี้คือชุดของค่าน้ำหนักสามารถนำไปใช้ได้หลายครั้งใน

ตำแหน่งต่าง ๆ ภายในพื้นที่ของข้อมูลนำเข้า เพื่ออธิบายอย่างละเอียดตัวกรองจะถูกย้ายอย่างเป็นระบบผ่านพื้นที่ของข้อมูลนำเข้า โดยตรวจสอบแต่ละส่วนของข้อมูลนำเข้าที่ถูกกรองให้มีขนาดเท่ากับตัวกรองที่ละรายการ จากซ้ายไปขวาและบนลงล่าง กระบวนการนี้ช่วยให้เครือข่ายสามารถหารูปแบบและคุณลักษณะเฉพาะต่าง ๆ ภายในข้อมูลนำเข้าผ่านการใช้ตัวกรองเดียวกันซ้ำ ๆ ได้

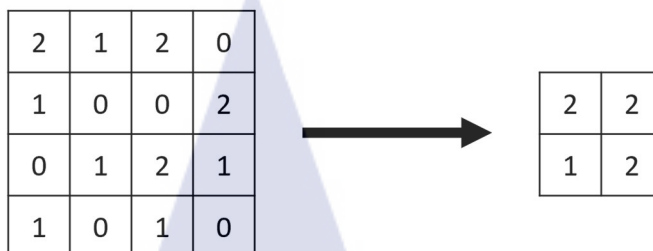
อย่างไรก็ตามเนื่องจากตัวกรองถูกย้ายอย่างเป็นระบบและนำไปใช้กับตำแหน่งต่าง ๆ ทั้งหมดอาร์เรย์ข้อมูลนำเข้าเพื่อนำมาสร้างอาร์เรย์สองมิติ อาร์เรย์นี้ประกอบด้วยค่าที่ถูกนำออก (Output) หลายค่า ซึ่งแต่ละค่าแสดงถึงผลลัพธ์ของการกรองข้อมูลนำเข้าในตำแหน่งที่แตกต่างกัน ผลลัพธ์ของอาร์เรย์สองมิติที่เกิดขึ้นจากกระบวนการนี้เรียกว่าแผนผังคุณลักษณะ (feature map) โดยเราสามารถนำแต่ละค่าในแผนผังคุณลักษณะมาคำนวณทางสถิติที่ไม่เป็นเชิงเส้น (nonlinearity) อย่าง Rectified Linear Unit (ReLU)

2	1	2	0	2	1	2	0	2	1	2	0	2	1	2	0
1	0	0	2	1	0	0	2	1	0	0	2	1	0	0	2
0	1	2	1	0	1	2	1	0	1	2	1	0	1	2	1
1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0
Step 1				Step 2				Step 3				Step 4			

รูปที่ 2.1 แสดงการเปลี่ยนตำแหน่งของฟิลเตอร์

#### 2.4.2 Max pooling

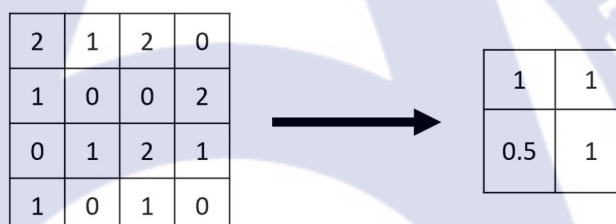
เป็นกระบวนการที่เกิดขึ้นหลังจากขั้นตอน Convolutional layer โดยทำการลดขนาดของแผนผังคุณลักษณะลงจากการแบ่งส่วนของข้อมูลนำเข้าให้มีพื้นที่ที่เล็กลงและเลือกค่าน้ำหนักที่มากที่สุดจากแต่ละพื้นที่ย่อย ซึ่งช่วยให้ได้ผลลัพธ์ของมิติที่เล็กลง เพิ่มความสะดวกและรวดเร็วมากยิ่งขึ้น



รูปที่ 2.2 การทำพูลลิ่งแบบค่ามากที่สุด

#### 2.4.3 Average pooling

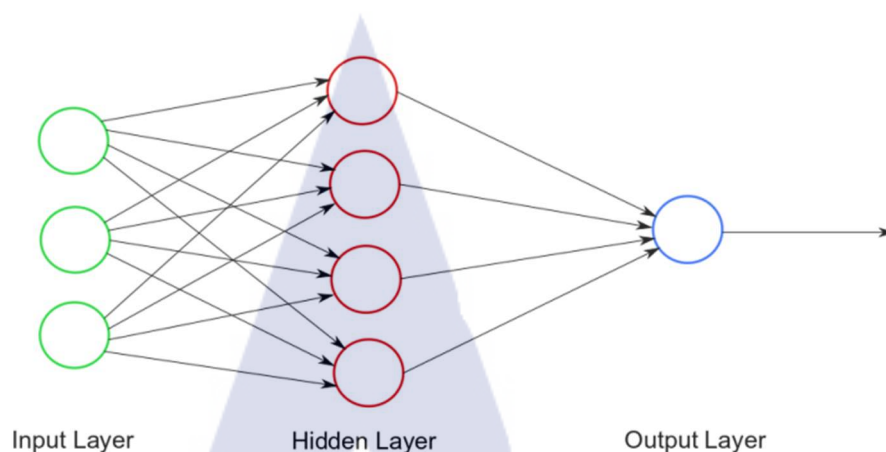
อีกกระบวนการหนึ่งในการพูลลิ่งเพื่อลดขนาดของแผนผังคุณลักษณะลงจากการแบ่งส่วนของข้อมูลนำเข้าให้มีพื้นที่เล็กลง เกิดขึ้นหลังจากขั้นตอน Convolutional layer เช่นเดียวกับ Max pooling แต่จะทำการหาค่าเฉลี่ยของค่าน้ำหนักทั้งหมดจากแต่ละพื้นที่ย่อย



รูปที่ 2.3 การทำพูลลิ่งแบบหาค่าเฉลี่ย

#### 2.4.4 ANN for classification

เป็นรูปแบบการประมวลผลของ ANN โดยประกอบด้วยโหนดที่เชื่อมต่อกันที่เรียกว่าเซลล์ประสาท ที่ถูกจัดเรียงเป็นชั้น ๆ โดยพื้นฐานแล้วจะแบ่งเป็น 3 ระดับ คือ Input layer ในการรับข้อมูลดิบ Hidden layers สำหรับประมวลผลและแปลงข้อมูลผ่านฟังก์ชันการประมวลผล ในระดับชั้นนี้จะทำการแยกคุณลักษณะที่ซับซ้อนจากข้อมูล Output layer ระดับชั้นที่จะสร้างผลลัพธ์สุดท้ายในการแบ่งกลุ่ม



รูปที่ 2.4 แสดงกระบวนการแบ่งกลุ่มของ ANN

## 2.5 การตรวจจับวัตถุ (Object Detection)

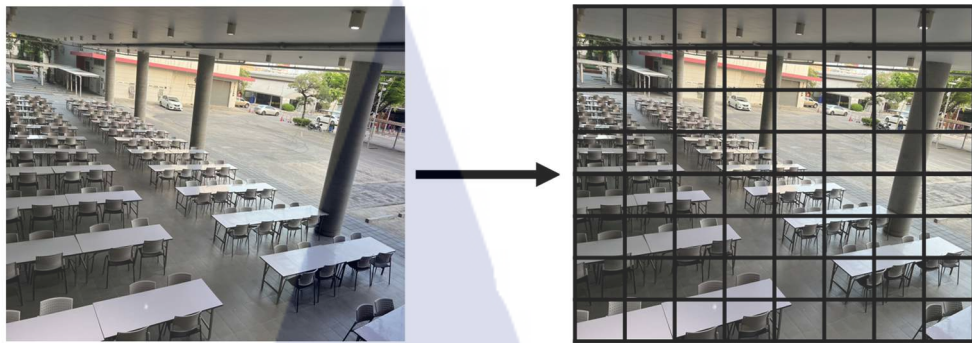
การตรวจจับวัตถุผสมผสานสองกระบวนการของคอมพิวเตอร์วิทัศน์ (Computer Vision) เข้าด้วยกัน ได้แก่ การจำแนกรูปภาพ (Image classification) และการระบุตำแหน่งวัตถุ (object localization) การตรวจจับวัตถุจะทำการวางกรอบล้อมรอบวัตถุแต่ละชิ้นในรูปภาพและระบุประเภทให้กับวัตถุเหล่านั้น [16]

## 2.6 YOLO

แบบจำลอง YOLO (You Only Look Once) ถูกนำเสนอครั้งแรกในบทความ "You Only Look Once: Unified, Real-Time Object Detection." เขียนโดย J. Redmon และคณะ ในปี 2015 [17]

YOLO มีความโดดเด่นในการใช้โครงข่ายประสาทเทียมที่ได้รับการฝึกอบรม และทำการเรียนรู้จากข้อมูลภาพจำนวนมาก สามารถทำการทำนายตำแหน่งและระบุประเภทของวัตถุได้อย่างแม่นยำ นอกจากนี้ยังสามารถประมวลผลแบบเรียลไทม์ที่ 155 เฟรมต่อวินาที

แบบจำลองนี้ดำเนินการโดยการแบ่งภาพที่ป้อนเข้าออกเป็นตารางของเซลล์ โดยแต่ละเซลล์จะได้รับมอบหมายให้ทำนายกรอบขอบเขตหากมีจุดศูนย์กลางของกรอบขอบเขตอยู่บริเวณวัตถุเป้าหมาย ในแต่ละเซลล์จะทำการทำนายกรอบขอบเขตเพื่อระบุตำแหน่งวัตถุ รวมถึงพิกัด  $x$  และ  $y$  ความกว้าง ความสูง ค่าความมั่นใจ และการจัดกลุ่ม



รูปที่ 2.5 แสดงการวางกริดของ YOLO

## 2.7 YOLOv5

YOLOv5 [18] ได้พัฒนาต่อยอดจาก YOLOv4 ประมาณสองเดือนในปี 2020 โดย Glen Jocher ผู้ก่อตั้งและ CEO ของ Ultralytics YOLOv5 ใช้เทคนิคที่ถูกปรับปรุงหลายอย่างจาก YOLOv4 แต่พัฒนาบน Pytorch แทน Darknet นอกจากนี้ YOLOv5 มีการเพิ่มอัลกอริทึม AutoAnchor ของ Ultralytics เครื่องมือสำหรับทดสอบ (Pre-training tool) นี้ เครื่องมือนี้จะตรวจสอบและปรับ anchor box ให้เหมาะสมกับชุดข้อมูลและการตั้งค่าการฝึก เช่น ขนาดภาพ เริ่มแรก AutoAnchor จะใช้ฟังก์ชัน k-means กับ label ของชุดข้อมูลเพื่อสร้างเงื่อนไขเริ่มต้นสำหรับอัลกอริทึม Genetic Evolution (GE) หลังจากนั้น อัลกอริทึม GE จะถูกพัฒนา anchor box ต่อไปอีก 1000 รุ่น เป็นค่าเริ่มต้น จากการใช้ฟังก์ชัน CloU loss [19] และ ค่า Recall ที่ดีที่สุดเป็น ฟังก์ชันที่เหมาะสม

## 2.8 YOLOv6

YOLOv6 [20] ถูกเผยแพร่ใน ArXiv เมื่อเดือนกันยายน 2022 โดยแผนก Meituan Vision AI การออกแบบเครือข่ายประกอบด้วยส่วนหลักที่มีประสิทธิภาพ โดยใช้บล็อก RepVGG หรือ CSPStackRep ที่สามารถแยกส่วนบริเวณสี่เหลี่ยมที่มีประสิทธิภาพจาก hybrid-channel strategy ในการตรวจจับใบหน้ามนุษย์ นอกจากนี้มีการนำเทคนิคการเพิ่มประสิทธิภาพโดยใช้ post-training quantization และ channel-wise distillation ส่งผลให้มีการตรวจจับที่เร็วและแม่นยำมากยิ่งขึ้น โดยรวมแล้ว YOLOv6 มีประสิทธิภาพการทำงานเหนือกว่าโมเดลก่อนหน้านี้ในด้านความแม่นยำและความเร็วอย่าง YOLOv5 YOLOX และ PP-YOLOE

## 2.9 YOLOv7

YOLOv7 [21] ถูกเผยแพร่บน ArXiv ในเดือนกรกฎาคม 2022 โดยทีมผู้สร้างเดียวกันกับ YOLOv4 และ YOLOR ในเวลานั้น YOLOv7 มีความเร็วและความแม่นยำเหนือกว่าตัวตรวจจับวัตถุอื่น ๆ ในเวลานั้นที่ช่วงความเร็ว 5 FPS ถึง 160 FPS เช่นเดียวกับ YOLOv4 ที่ถูกฝึกโดยใช้ชุดข้อมูล MS COCO เพียงอย่างเดียว โดยไม่ใช้ชั้นเครือข่ายประสาทเทียมเบื้องต้นที่ผ่านการฝึกมาล่วงหน้า YOLOv7 ได้เสนอการเปลี่ยนแปลงทางสถาปัตยกรรม และชุดเทคนิคที่เรียกว่า "bag-of-freebies" ซึ่งช่วยเพิ่มความแม่นยำโดยไม่ส่งผลต่อความเร็วในการประมวลผล แต่อาจส่งผลต่อระยะเวลาในการฝึก

## 2.10 YOLOv8

Ultralytics YOLOv8 เป็นโมเดลที่สร้างขึ้นบนความสำเร็จของ YOLO รุ่นก่อนหน้าและนำเสนอคุณสมบัติและการปรับปรุงใหม่ ที่เพิ่มประสิทธิภาพและความยืดหยุ่นในการใช้งาน YOLOv8 ถูกออกแบบให้มีความรวดเร็ว แม่นยำ และสามารถนำมาใช้งานได้ง่ายขึ้น โดยสามารถนำมาใช้งานได้ทั้งหมดสี่วิธี คือ การตรวจจับวัตถุ (object detection) การติดตามวัตถุ (tracking) การจำแนกประเภทข้อมูลภาพ (Image classification) และการประมาณการณที่วางท่าของมนุษย์ (Pose estimation)

## 2.11 YOLOv9

YOLOv9 เป็นจุดเปลี่ยนสำคัญทางเทคโนโลยีในการนำเสนอ Programmable Gradient Information (PGI) และ Generalized Efficient Layer Aggregation Network (GELAN) [22] ที่ช่วยให้โมเดลมีความแม่นยำมากยิ่งขึ้น และง่ายต่อการนำมาประยุกต์ใช้ นอกจากนี้ตัวโมเดลยังสร้างมาตรฐานใหม่ในการตรวจจับวัตถุผ่านชุดข้อมูล MS COCO ตัวโมเดล YOLOv9 เป็นโปรเจกต์ที่ถูกพัฒนาผ่านทีม open source โดยใช้ Ultralytics YOLOv5 เป็นต้นแบบในการพัฒนา

### 2.11.1 Programmable Gradient Information (PGI)

ถูกพัฒนาขึ้นบน auxiliary branch ทำให้ไม่มีการเสียทรัพยากรเพิ่มเติม อีกทั้ง PGI สามารถเลือกฟังก์ชันความสูญเสียที่เหมาะสมกับเป้าหมายได้อย่างอิสระ จึงสามารถแก้ไขข้อจำกัดที่พบใน mask modeling ได้ นอกจากนี้ กลไก PGI ที่นำเสนอสามารถนำไปประยุกต์ใช้กับโครงข่ายประสาทเทียมเชิงลึกที่มีขนาดต่าง ๆ ได้อย่างยืดหยุ่น และมีความเป็นสากลมากกว่ากลไก deep supervision ซึ่งโดยทั่วไปเหมาะสำหรับโครงข่ายประสาทเทียมเชิงลึกที่มีโครงสร้างซับซ้อนเท่านั้น

### 2.11.2 Generalized Efficient Layer Aggregation Network (GELAN)

การออกแบบของ GELAN คำนึงถึงจำนวนพารามิเตอร์ ความซับซ้อนในการคำนวณ ความแม่นยำ และความเร็วในการอนุมานไปพร้อมกัน การออกแบบนี้ช่วยให้ผู้ใช้สามารถเลือกบล็อกการคำนวณที่เหมาะสมกับ different inference devices ได้อย่างอิสระ

## 2.12 YOLOv10

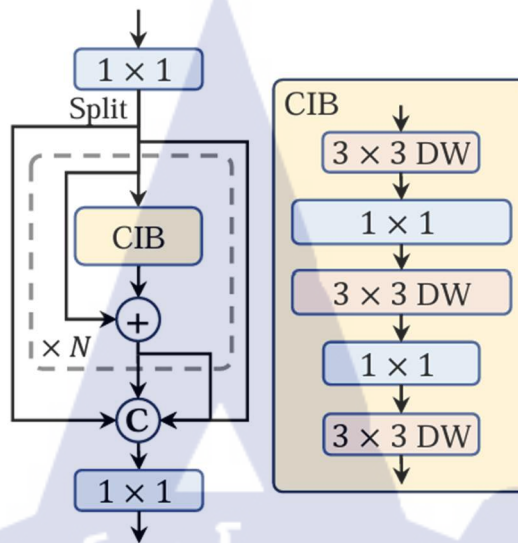
ในเวอร์ชันที่ 10 ของ YOLO [23] มีการนำ NMS-Free Training มาปรับปรุงให้สามารถลดภาระในการคำนวณลงได้เป็นอย่างมากพร้อมกับยังคงความแม่นยำไว้ได้ และการปรับปรุงโครงสร้างของโมเดลด้วย large-kernel convolutions และ partial self-attention modules ที่ปรับปรุงความสามารถในการประมวลผลโดยไม่สูญเสียทรัพยากรในการประมวลผลอย่างมีนัยสำคัญ

### 2.12.1 NMS-Free Training

โดยทั่วไป YOLO จำเป็นต้องใช้ non-maximum suppression (NMS) สำหรับการประมวลผลแบบ one-to-many assignment ซึ่งทำให้เกิดการประเมินผลลัพธ์ที่ไม่มีประสิทธิภาพ ในขณะที่อีกวิธีหนึ่งที่น่ามาใช้อย่าง one-to-one matching ที่ช่วยลดความซับซ้อนในการทำงาน แต่จำเป็นต้องเพิ่มทรัพยากรในการประเมินผลลัพธ์และกระทบต่อการประมวลผล เพื่อแก้ปัญหาดังกล่าวจึงมีการนำ one-to-many assignment กับ one-to-one matching มารวมกันเรียกว่า Dual label assignments ซึ่งช่วยแก้ข้อด้อยของทั้งสองกลยุทธ์ได้

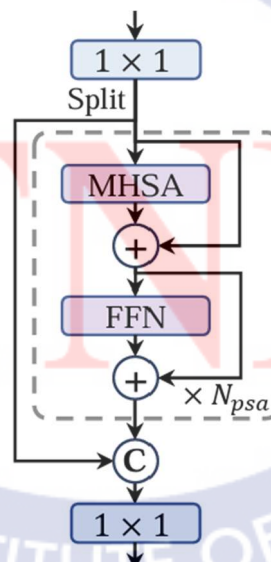
### 2.12.2 Enhanced Model Capabilities

การนำ Large-kernel convolution มาใช้ในการขยายขอบเขตการรับรู้ของโมเดลในส่วนของโครงสร้าง compact inverted block (CIB) โดยการเพิ่มขนาดเคอร์เนลจาก  $3 \times 3$  เป็น  $7 \times 7$  พร้อมเทคนิคการปรับโครงสร้างพารามิเตอร์โดยไม่กระทบต่อการประเมินผลลัพธ์ การออกแบบนี้เหมาะกับโมเดลขนาดเล็ก เพราะโมเดลขนาดใหญ่มีขอบเขตการรับรู้ที่กว้างอยู่แล้ว



รูปที่ 2.6 แสดงลำดับการทำงานของ Large-kernel convolution [23]

Partial self-attention (PSA) ทำการแบ่งกระบวนการออกเป็นสองส่วน คือ multi-head self-attention (MHSA) และ feed-forward network (FFN) โดยวิธีการดังกล่าวช่วยลดทรัพยากรในการประมวลผล และเพิ่มประสิทธิภาพในการประมวลผล



รูปที่ 2.7 แสดงลำดับการทำงานของ Partial self-attention (PSA) [23]

### 2.13 YOLO11

YOLO11 ได้รับการปรับปรุงโครงสร้างโมเดลและวิธีในการฝึกฝนในระดับ Backbone และ Neck ซึ่งเป็นโครงสร้างหลักในการตรวจจับวัตถุของ YOLO ทำให้ตัวโมเดลมีความแม่นยำมากยิ่งขึ้น ในขณะที่ใช้พารามิเตอร์ในการฝึกฝนน้อยลงกว่า 22% โดยเฉพาะโมเดล YOLO11m ด้วยการใช้ทรัพยากรประมวลผลที่ลดลง ช่วยให้สามารถนำโมเดลไปติดตั้งและใช้งานกับอุปกรณ์ที่มีประสิทธิภาพจำกัดได้หลากหลายมากขึ้น สนับสนุนการตรวจจับวัตถุในมุมมองที่หลากหลายมากขึ้น

### 2.14 งานวิจัยที่เกี่ยวข้อง

YOLO ถูกนำมาใช้ตรวจจับฝูงชนและยานพาหนะได้อย่างแม่นยำพร้อมกับมีการปรับปรุงเวอร์ชันใหม่ออกมาอย่างต่อเนื่อง อย่างไรก็ตามภาพที่ได้จากสภาพอากาศที่มีหมอกหนาแน่น ส่งผลให้การตรวจจับภาพเป็นไปได้ยากและมีความแม่นยำลดลง จึงมีการปรับแต่งโมเดล YOLO ให้สามารถตรวจจับวัตถุจากภาพที่เต็มไปด้วยหมอกได้แม่นยำมากขึ้น และเรียกโมเดลที่ได้รับการปรับแต่งนี้ว่า Image-Adaptive YOLO (IA-YOLO) โดยการนำ Differentiable Image Processing (DIP) มาใช้ปรับปรุง YOLO ให้สามารถตรวจจับวัตถุในสภาพอากาศที่ยากต่อการมองเห็น และใช้ Convolutional Neural Network (CNN) สำหรับทำนายค่าพารามิเตอร์ที่เกี่ยวข้องกับสภาพอากาศ เพื่อให้แน่ใจว่า CNN จะสามารถนำ DIP ไปใช้เพิ่มความแม่นยำในการตรวจจับภาพได้อย่างแม่นยำมากยิ่งขึ้นในแต่ละสภาพอากาศ [24] จากงานวิจัยนี้พบว่า IA-YOLO สามารถตรวจจับวัตถุทั้งในสภาพอากาศปกติและสภาพอากาศที่ยากต่อการมองเห็นได้ดียิ่งขึ้นโดยเฉพาะกับภาพที่มีหมอกและมีแสงน้อย นอกจากนี้การที่กล้องโดรนบังคับไปชั่วขณะหนึ่งจนหลุดโฟกัสหรือสภาพอากาศที่ทำให้ยากต่อการมองเห็นจนภาพไม่ชัดเจน ทำให้มีความต้องการแก้ปัญหาดังกล่าวจากการนำ YOLOv5 มาทำการปรับปรุงด้วย Flip-Mosaic algorithm และนำมาใช้ร่วมกับชุดข้อมูลที่มีคุณลักษณะยานพาหนะหลากหลายในแต่ละสถานการณ์ [25] จากงานวิจัยนี้พบว่า Flip-Mosaic algorithm สามารถเพิ่มความแม่นยำในการตรวจจับยานพาหนะได้มากขึ้นและลดอัตราการตรวจจับที่ล้มเหลวลง

อากาศยานไร้คนขับ Unmanned aerial vehicle (UAV) เป็นมุมมองในการตรวจจับที่มีมุมมองที่หลากหลายและความเร็วภาพที่แปรเปลี่ยนไปตามความเร็วของอากาศยานไร้คนขับ ในช่วงของสถานการณ์การแพร่ระบาดของเชื้อไวรัสโคโรนา 2019 จึงมีการนำอากาศยานไร้คนขับมาใช้ตรวจจับเพื่อสำรวจพื้นที่แออัดของฝูงชน [26] ในงานวิจัยนี้นำ YOLOv3 มาใช้ในการตรวจจับพร้อมกับกำหนดเส้นทางการบินสำรวจที่ชัดเจน โดยมีความแม่นยำถึง 90% ในการตรวจจับผู้คนรวมไปถึงความแออัดของฝูงชน นอกจากนี้ในอีกงานวิจัยหนึ่งมีการเพิ่มความสามารถในการตรวจจับของอากาศยานไร้คนขับพร้อมกับภาพตรวจจับอินฟราเรด Thermal infrared (TIR) ทำให้ได้มุมมองที่แตกต่างจาก

ภาพโดยทั่วไป โดย YOLOv5 ได้ถูกนำมาใช้ในการตรวจจับ [27] และพบว่ามีความแม่นยำเฉลี่ยในการตรวจจับอยู่ที่ร้อยละ 88.69 และตรวจจับได้ด้วยความเร็วภาพสูงสุด 50 เฟรมต่อวินาที

จากผลกระทบการแพร่ระบาดของเชื้อไวรัสโคโรนา 2019 ทำให้มีความสนใจในการจัดการฝูงชนมากขึ้น อย่างบริเวณห้างสรรพสินค้าซึ่งนอกจากจะใช้ในการป้องกันการแพร่ระบาดแล้ว ยังมีประโยชน์ต่อการบริหารจัดการฝูงชนให้เกิดความคล่องตัวในการเดินทาง และศึกษาพฤติกรรมของฝูงชนเพื่อนำไปปรับปรุงการบริการของห้างสรรพสินค้าต่อไป [28] ในการวิจัยนี้ได้นำเสนอระบบตรวจจับฝูงชนโดยการนำ YOLOv3 มาใช้ในการตรวจจับและแยกแยะผู้คนที่พร้อมกันอัลกอริทึมในการติดตาม DeepSORT เพื่อวางแผนแบ่งในการนับจำนวนผู้คน นอกจากนี้มีการนำ YOLOv3 ที่ได้รับการฝึกมาใช้ร่วมกับ TensorFlow เพื่อให้เกิดการประมวลผลแบบเรียลไทม์ได้ดียิ่งขึ้น โดยมีความแม่นยำอยู่ที่ร้อยละ 91.07

เนื่องจากการตรวจจับผู้คนจำนวนมากแบบเรียลไทม์จำเป็นต้องใช้อุปกรณ์ที่มีความสามารถในการประมวลผลที่สูง ข้อจำกัดด้านประสิทธิภาพของอุปกรณ์ทำให้มีความพยายามสร้างอัลกอริทึมในการตรวจจับวัตถุให้กับอุปกรณ์ที่มีความสามารถในการประมวลผลที่จำกัด การปรับปรุงอัลกอริทึมในการตรวจจับวัตถุให้มีการใช้ทรัพยากรในการประมวลผลลดลงในชื่อ PG-YOLO ซึ่งเกิดจากการปรับแต่งโมเดลดั้งเดิมอย่าง YOLOv5 [29] โดยผลลัพธ์คือ PG-YOLO เป็นอัลกอริทึมที่มีการประมวลผลลดลงและรวดเร็วมากขึ้นเมื่อเทียบกับ YOLOv5 ซึ่งสามารถประมวลผลข้อมูลได้มากกว่าเดิม 9 เท่า มีความแม่นยำในการทำนายที่ต่ำกว่าโมเดลเดิมเพียง 0.1% แต่ใช้เวลาในการประมวลผลลดลง 32.7%

อัลกอริทึมที่ถูกพัฒนาขึ้นมาเพื่อแก้ไขปัญหาที่มีลักษณะคล้ายกันนี้ คือ hybrid YOLOv4 ที่นำ YOLOv4 มาผ่านกระบวนการ Training phase pruning และ Convolutional Block Attention Module (CBAM) [30] ซึ่งช่วยลดความต้องการของหน่วยความจำในการประมวลผลโดยมีความแม่นยำในการตรวจจับที่เพิ่มขึ้น 33% มีค่า mAP ที่ 92.1%

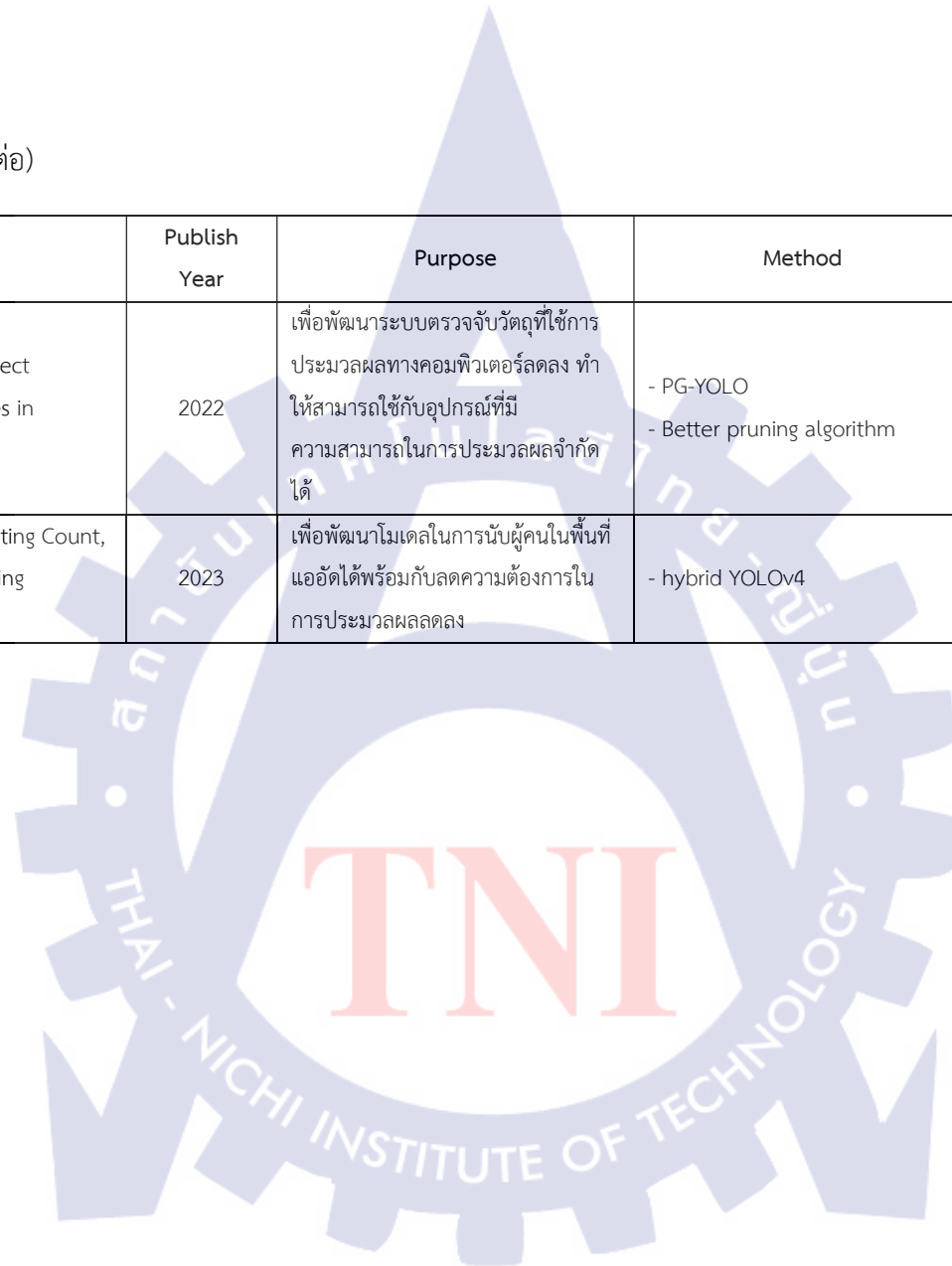
จากงานวิจัยทั้งหมดนี้พบว่า YOLO ถูกนำมาใช้ในการตรวจจับผู้คนและยานพาหนะให้กับอากาศยานไร้คนขับ พร้อมกับความพยายามในการเพิ่มความแม่นยำในการตรวจจับกรณีภาพจากกล้องวงจรปิดไม่ชัดเจนจากหมอกหรือถูกบดบังทัศนวิสัยกะทันหันจนกล้องหลุดโฟกัส การสร้างระบบการตรวจจับสำหรับบริหารจัดการฝูงชน และการปรับอัลกอริทึมให้สามารถใช้ตรวจจับฝูงชนกับอุปกรณ์ที่มีความสามารถในการประมวลผลที่จำกัดได้ อย่างไรก็ตาม YOLO มีการพัฒนาและปรับปรุงอย่างรวดเร็วทำให้งานวิจัยมีการใช้ YOLO ในหลากหลายเวอร์ชัน โดยในงานวิจัยนี้จะมุ่งเน้นไปที่การพัฒนาาระบบตรวจจับพร้อมๆกับสร้างเว็บไซต์ที่สามารถแสดงผลการตรวจจับวัตถุจากภาพวิดีโอและสรุปข้อมูลที่ได้ในรูปแบบแดชบอร์ดจากการตรวจจับบุคคลในบริเวณโรงอาหารและยานพาหนะในที่จอดรถ เพื่อนำมาวิเคราะห์ในการบริหารจัดการฝูงชนและที่จอดรถต่อไป

ตารางที่ 2.1 แสดงงานวิจัยที่เกี่ยวข้อง

No.	Topic	Publish Year	Purpose	Method	Result
1	Image-Adaptive YOLO for Object Detection in Adverse Weather Conditions [24]	2022	เพื่อตรวจจับวัตถุจากภาพที่มีคุณภาพต่ำจากสภาพอากาศที่ยากต่อการมองเห็น	- IA-YOLO	- mAP 37.08% สำหรับสภาพแวดล้อมที่มีหมอกหนา
2	Real-Time Vehicle Detection Based on Improved YOLO v5 [25]	2022	เพื่อลดความล้มเหลวในการตรวจจับยานพาหนะที่เกิดจากการบดบังโดยฉับพลัน	- Improved YOLOv5	- mAP 90.5%
3	Developing smart COVID-19 social distancing surveillance drone using YOLO Implemented in Robot Operating System Simulation Environment [26]	2020	เพื่อให้โดรนสามารถระบุตำแหน่งและนำทางฝูงชนในพื้นที่แออัดได้	- YOLOv3	- Accuracy 90%
4	Object detection from UAV thermal infrared images and videos using YOLO models [27]	2022	เพื่อให้อากาศยานไร้คนขับสามารถตรวจจับวัตถุผ่านภาพอินฟราเรดได้	- YOLOv5	- mAP 88.69% at 50 FPS
5	Real-Time Human Detection and Counting System Using Deep Learning Computer Vision Techniques [28]	2022	เพื่อสร้างระบบตรวจจับผู้คนที่แสดงหน้าจอกการทำงานที่เหมาะสมแก่ผู้ใช้และมีระบบในการจัดการ	- YOLOv3 - DeepSORT tracking algorithm	- Accuracy 91.07%

ตารางที่ 2.1 แสดงงานวิจัยที่เกี่ยวข้อง (ต่อ)

No.	Topic	Publish Year	Purpose	Method	Result
6	PG-YOLO: A Novel Lightweight Object Detection Method for Edge Devices in Industrial Internet of Things [29]	2022	เพื่อพัฒนาระบบตรวจจับวัตถุที่ใช้การประมวลผลทางคอมพิวเตอร์ลดลง ทำให้สามารถใช้กับอุปกรณ์ที่มีความสามารถในการประมวลผลจำกัดได้	- PG-YOLO - Better pruning algorithm	- Accuracy 93.4%
7	Realtime Crowd Monitoring Estimating Count, Speed and Direction of People Using Hybridized YOLOv4 [30]	2023	เพื่อพัฒนาโมเดลในการนับผู้คนในพื้นที่แออัดได้พร้อมกับลดความต้องการในการประมวลผลลดลง	- hybrid YOLOv4	- Accuracy has increased 33% - mAP 92.1%



### บทที่ 3 วิธีดำเนินการวิจัย

ในงานวิจัยนี้จะมุ่งเน้นไปที่การพัฒนาระบบตรวจจับผู้คนและยานพาหนะโดยใช้ YOLO11 พร้อมกับการสร้างเว็บไซต์ที่สามารถแสดงผลการตรวจจับวัตถุจากภาพวิดีโอพร้อมกับสรุปข้อมูลที่ได้ในรูปแบบแดชบอร์ด เพื่อนำข้อมูลไปใช้ในการวิเคราะห์รูปแบบในการจัดการฝูงชนและที่จอดรถต่อไป ซึ่งในหัวข้อนี้จะกล่าวถึง การเตรียมข้อมูล เครื่องมือที่นำมาใช้ การตรวจจับวัตถุด้วยโมเดล YOLO11 การประเมินประสิทธิภาพในการตรวจจับ และการสร้างเว็บไซต์สำหรับแสดงผลการตรวจจับ

#### 3.1 การเตรียมข้อมูล

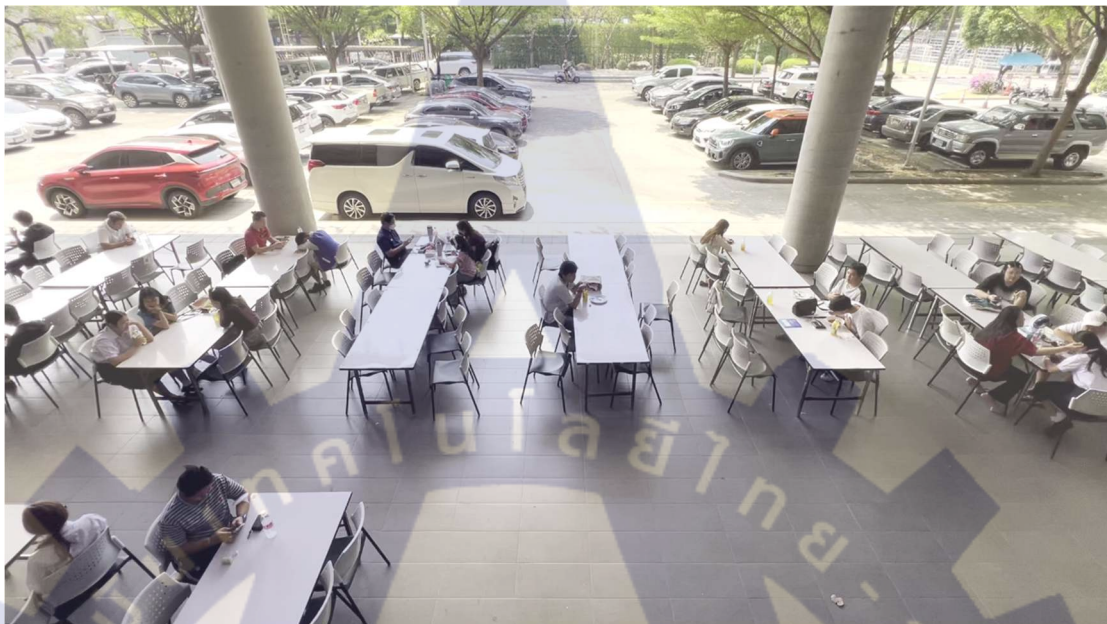
ข้อมูลภาพวิดีโอจากกล้องมือถือบริเวณโรงอาหาร 2 ตัวอย่าง และหน้าทางเข้าสถาบันเทคโนโลยีไทย-ญี่ปุ่น 1 ตัวอย่าง โดยมีรายละเอียด ดังนี้

##### 3.1.1 วิดีโอบริเวณโรงอาหารที่หนึ่งความยาว 3 นาที



รูปที่ 3.1 ภาพบริเวณโรงอาหารที่หนึ่ง

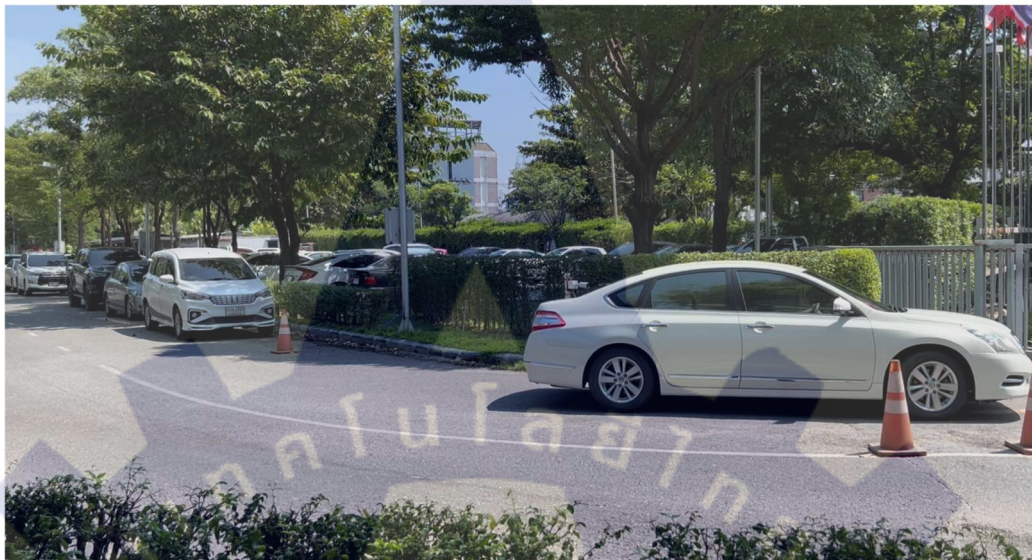
### 3.1.2 วิดีโอบริเวณโรงอาหารที่สองความยาว 3 นาที



รูปที่ 3.2 ภาพบริเวณโรงอาหารที่สอง



### 3.1.3 วิดีโอบริเวณหน้าทางเข้าสถาบันความยาว 3 นาที 30 วินาที



รูปที่ 3.3 ภาพบริเวณหน้าทางเข้าสถาบัน

## 3.2 เครื่องมือที่นำมาใช้

การพัฒนาาระบบตรวจจับและเว็บไซต์ในการวิจัยนี้ ถูกพัฒนาผ่านระบบปฏิบัติการ Windows 11 Home Single Language ประมวลผลด้วย AMD Ryzen 7 7800X3D 8-Core Processor RAM 32 GB. NVIDIA GeForce RTX 4070Ti. ภาษาโปรแกรมที่ใช้ คือ Python programming language version 3.10.15 torch version 2.5.1 torchvision version 0.20.1 with CUDA version 12.7 ผ่านโปรแกรมแก้ไขซอร์สโค้ดวิซวลสตูดิโอโค้ด (Visual Studio Code) โดยใช้ Anaconda ในการช่วยจัดการ package และ Python environment

## 3.3 การตรวจจับวัตถุด้วยโมเดล YOLO11

การนำโมเดลสำหรับตรวจจับด้วย YOLO11 มาใช้จะมีขั้นตอนในการดำเนินการอยู่ 2 ขั้นตอนหลัก ๆ คือ

### 3.3.1 การนำโมเดลของ YOLO11 มาใช้ในการตรวจจับ

ด้วยโมเดลดั้งเดิมของ YOLO11 สามารถตรวจจับผู้คนและยานพาหนะได้อย่างแม่นยำ โดยโมเดลดั้งเดิมที่ได้รับการฝึกมาก่อนแล้วแบ่งเป็น YOLO11n YOLO11s YOLO11m

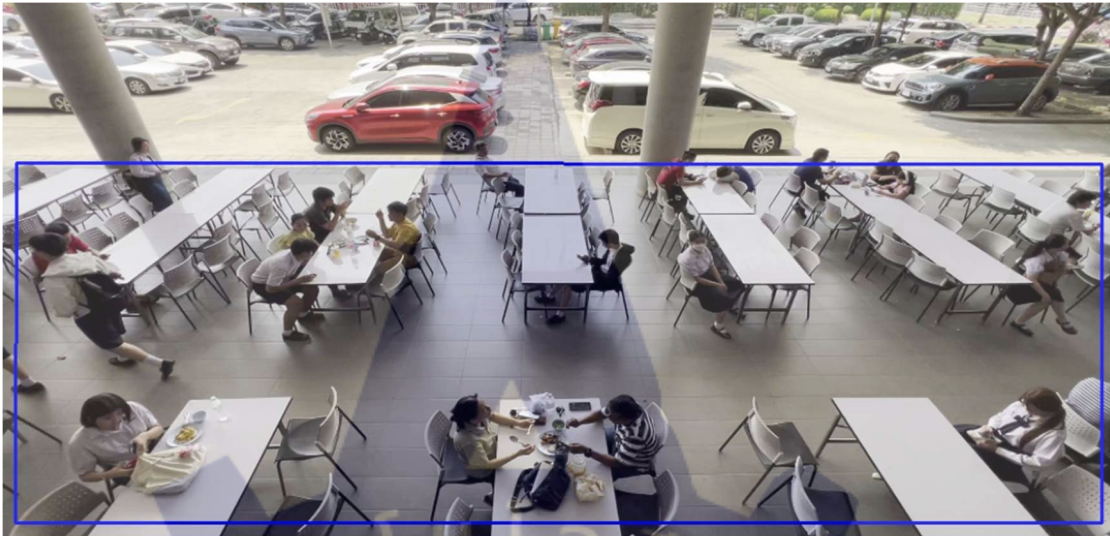
YOLO11l YOLO11x ซึ่งมีความเร็วและความแม่นยำแตกต่างกันตามจำนวนพารามิเตอร์ และได้รับการฝึกผ่านชุดข้อมูล COCO ซึ่งเป็นชุดข้อมูลขนาดใหญ่ที่ถูกออกแบบสำหรับการทำงานวิจัยและทดสอบประสิทธิภาพของคอมพิวเตอร์วิทัศน์ มีการรวบรวมรูปภาพกว่า 330,000 รูป แบ่งประเภทของวัตถุได้ 80 ประเภท

### 3.3.2 การเก็บข้อมูลที่ได้จากการตรวจจับ

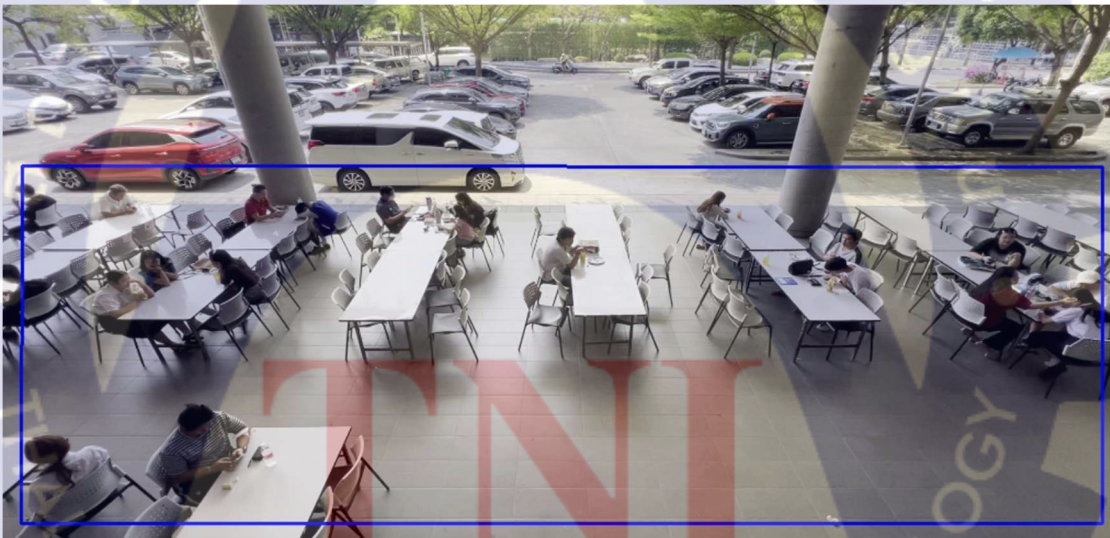
ในการศึกษานี้ มีการเก็บข้อมูลเกี่ยวกับบุคคลที่เดินทางเข้าออกบริเวณโรงอาหารและยานพาหนะที่เคลื่อนที่ภายในพื้นที่จอดรถ โดยอาศัยการกำหนดเส้นแบ่งเพื่อใช้ในการนับวัตถุที่เคลื่อนที่ผ่าน ระบบที่พัฒนาขึ้นสามารถบันทึกข้อมูลที่สำคัญ ได้แก่ จำนวนวัตถุที่ตรวจจับได้ทั้งหมด จำนวนวัตถุที่เคลื่อนที่ผ่านเส้นแบ่ง เวลาที่วัตถุเคลื่อนที่ผ่านเส้นแบ่ง ทิศทางของวัตถุขณะผ่านเส้นแบ่ง และความแม่นยำของการตรวจจับ

ข้อมูลที่รวบรวมได้จากกระบวนการดังกล่าวจะถูกนำไปวิเคราะห์ผ่านระบบบนเว็บไซต์ในลำดับถัดไป โดยมีการกำหนดพื้นที่ตรวจจับ 3 รูปแบบ ได้แก่ พื้นที่ตรวจจับที่ใช้สำหรับนับจำนวนวัตถุที่เคลื่อนที่เข้าสู่รัศมีของพื้นที่ตรวจจับ พื้นที่ตรวจจับที่ใช้สำหรับนับจำนวนวัตถุพิจารณาตามทิศทางวัตถุ พื้นที่ตรวจจับแบบหลายพื้นที่ช่วยคำนวณจำนวนวัตถุที่เข้าสู่แต่ละพื้นที่ได้อย่างอิสระ และ วิเคราะห์ระยะห่างระหว่างวัตถุภายในพื้นที่เดียวกัน

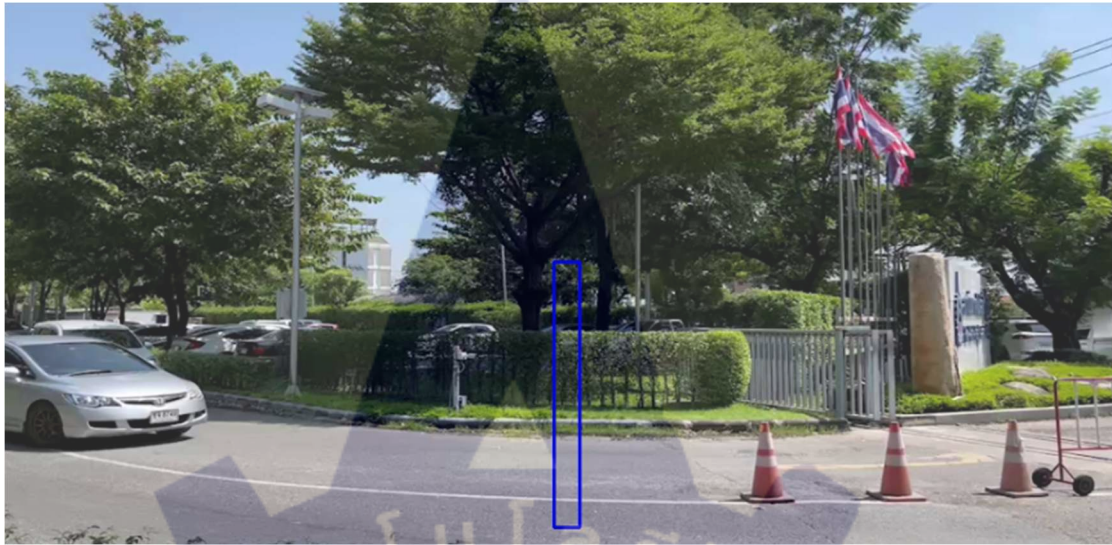
การกำหนดรูปแบบพื้นที่ตรวจจับถูกออกแบบให้เหมาะสมกับวัตถุประสงค์ของการศึกษา โดยมีโครงสร้างพื้นที่ตรวจจับแตกต่างกันไปตามสภาพแวดล้อมของสถานที่ ดังรูป 3.4 ถึง 3.6



รูปที่ 3.4 พื้นที่การตรวจจับบริเวณร้านอาหารที่หนึ่ง



รูปที่ 3.5 พื้นที่การตรวจจับบริเวณร้านอาหารที่สอง



รูปที่ 3.6 พื้นที่การตรวจจับบริเวณหน้าทางเข้าสถาบัน

### 3.4 การประเมินประสิทธิภาพในการตรวจจับ

#### 3.4.1 Precision

บ่งบอกถึงความเที่ยงตรงในแต่ละประเภทของกลุ่มข้อมูลที่สนใจว่ามีความแม่นยำของผลลัพธ์ของแบบจำลองได้ถูกต้อง เมื่อเทียบกับกลุ่มข้อมูลประเภทอื่นที่ถูกนำผลลัพธ์เพิ่มเข้ามาในประเภทของกลุ่มที่สนใจ

$$Precision = \frac{true\ positives}{true\ positives + false\ positives}$$

#### 3.4.2 Recall

บ่งบอกถึงความไวในการตรวจจับข้อมูลในแต่ละประเภทของกลุ่มข้อมูล ว่ากลุ่มข้อมูลที่สนใจมีความสามารถในการตรวจจับข้อมูลได้ถูกต้อง เมื่อเทียบกับข้อมูลของประเภทเดียวกันที่ถูกทำนายไปเป็นกลุ่มข้อมูลอื่น

$$Recall = \frac{true\ positives}{true\ positives + false\ negatives}$$

### 3.4.3 F1-score

F1-score เป็นตัวชี้วัดประสิทธิภาพที่เหมาะสมมากกว่าค่าความแม่นยำในการประเมินประสิทธิภาพของแบบจำลองบนชุดข้อมูลที่ค่ามีการแบ่งกลุ่มที่แตกต่างกันมากเกินไปหรือที่เรียกว่าชุดข้อมูลที่ไม่สมดุล (imbalanced dataset)

$$F1 = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$$

### 3.4.4 Intersection over Union (IoU)

IoU คือ ตัวชี้วัดการประเมินตำแหน่งของวัตถุในการตรวจจับพื้นที่ทับซ้อนระหว่างกรอบขอบเขตข้อมูลจริง (ground truth bounding box) กับกรอบขอบเขตที่ได้จากการทำนาย (predicted bounding box) หาด้วยพื้นที่รวมของทั้งสองขอบเขต

$$IoU = \frac{\text{Area of Intersection}}{\text{Area of Union}}$$

IoU มีค่าอยู่ระหว่าง 0 ถึง 1 โดยค่า 1 จะบ่งชี้ถึงการทับซ้อนอย่างสมบูรณ์ ในขณะที่ค่า 0 จะหมายถึงไม่มีการทับซ้อนกัน

### 3.4.5 mAP

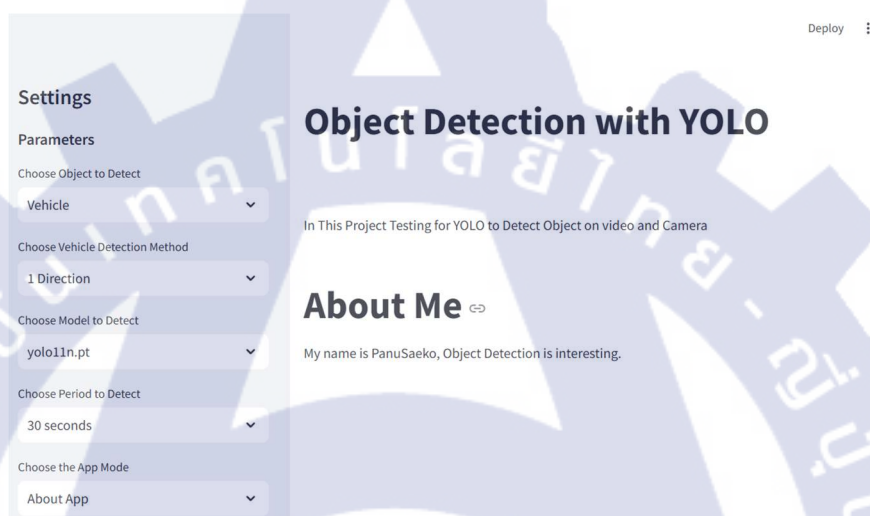
ค่าเฉลี่ยของความเที่ยงตรง (mean Average Precision หรือ mAP) เป็นการคำนวณที่เกี่ยวข้องกับกรอบขอบเขต (bounding boxes) และคุณภาพของการทำนายกรอบขอบเขต จะถูกกำหนดด้วยความสอดคล้องกับกรอบขอบเขตจริง การประเมินตำแหน่งซึ่งแสดงเป็นอัตราส่วนของพื้นที่ ๆ เป็นจุดตัดต่อพื้นที่ยูเนียนเรียกว่า จุดตัดเหนือยูเนียน (Intersection over Union หรือ IoU) การทำนายกล่องขอบเขตที่สมบูรณ์จะได้ค่า IoU เป็น 1 ซึ่งโดยทั่วไปจะพิจารณาการทำนายกล่องขอบเขตว่าถูกต้อง หากค่า IoU เกิน 0.5 ซึ่งบ่งชี้ว่ามีการทับซ้อนกันที่ 50% หรือมากกว่าเมื่อเทียบกับค่าจริง

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i$$

mAP คำนวณโดยการหาความแม่นยำเฉลี่ย (AP) สำหรับแต่ละคลาส จากนั้นจึงหาค่าเฉลี่ยจากจำนวนคลาส

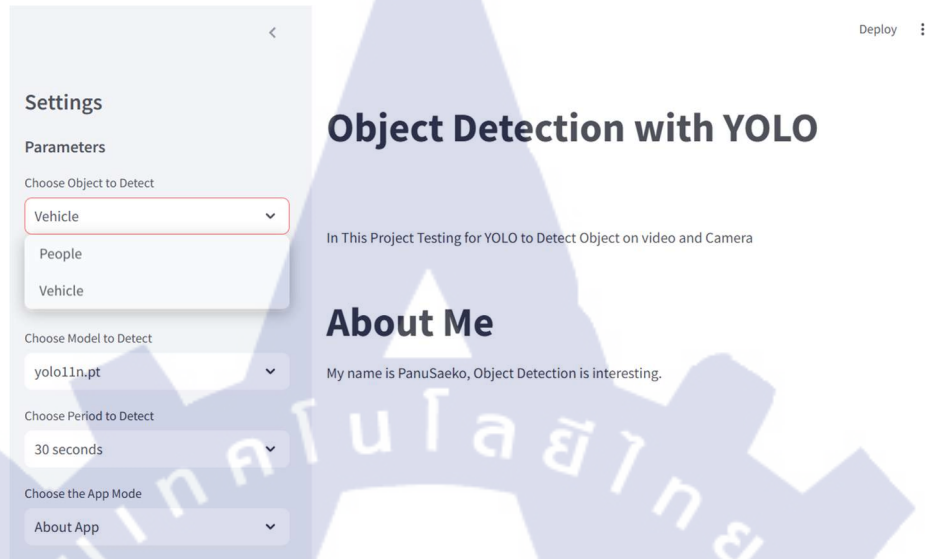
### 3.5 การสร้างเว็บไซต์สำหรับแสดงผลการตรวจจับ

เพื่อให้ง่ายต่อการใช้งานและการแสดงผล จึงได้ทำการสร้างระบบการตรวจจับบนเว็บไซต์ผ่าน Streamlit เวอร์ชัน 1.40.1 โดยตัวเว็บไซต์มีลำดับและฟังก์ชันการใช้งานดังนี้



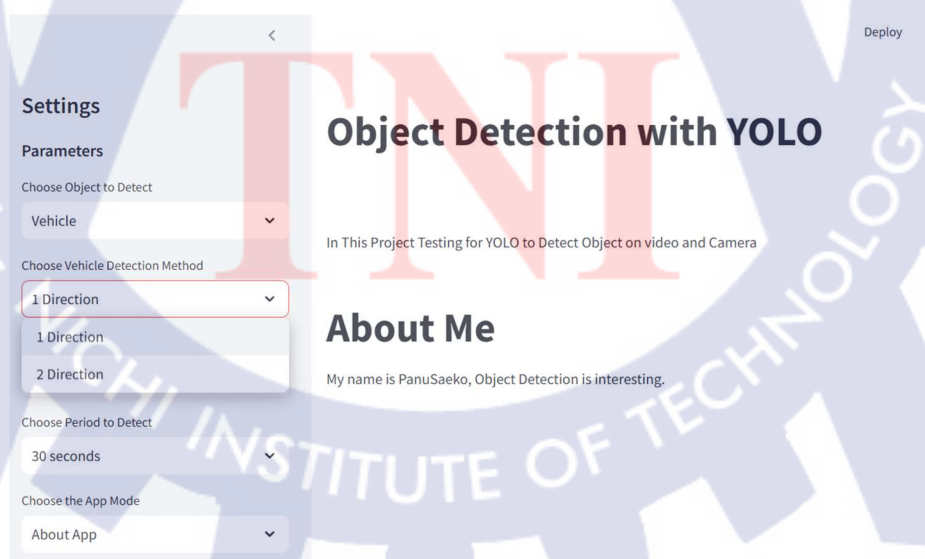
รูปที่ 3.7 หน้าจอหลักของเว็บไซต์

### 3.5.1 เลือกประเภทของวัตถุที่ต้องการตรวจจับระหว่างบุคคลและยานพาหนะ



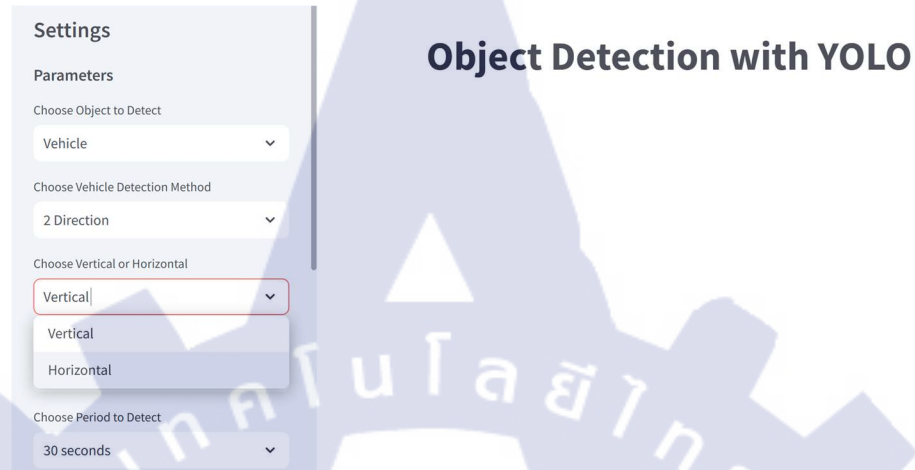
รูปที่ 3.8 เมนูการเลือกประเภทของวัตถุ

### 3.5.2 กรณีเลือกยานพาหนะจะมีวิธีให้เลือกเพิ่มระหว่างการตรวจจับแบบ 1 ทิศทาง หรือ 2 ทิศทาง



รูปที่ 3.9 เมนูการเลือกประเภททิศทางในการตรวจจับยานพาหนะ

3.5.3 กรณีเลือกการตรวจจับแบบ 2 ทิศทาง จะมีตัวเลือกเพิ่มเติมคือ ตรวจจับแบบแนวอนหรือแนวตั้ง



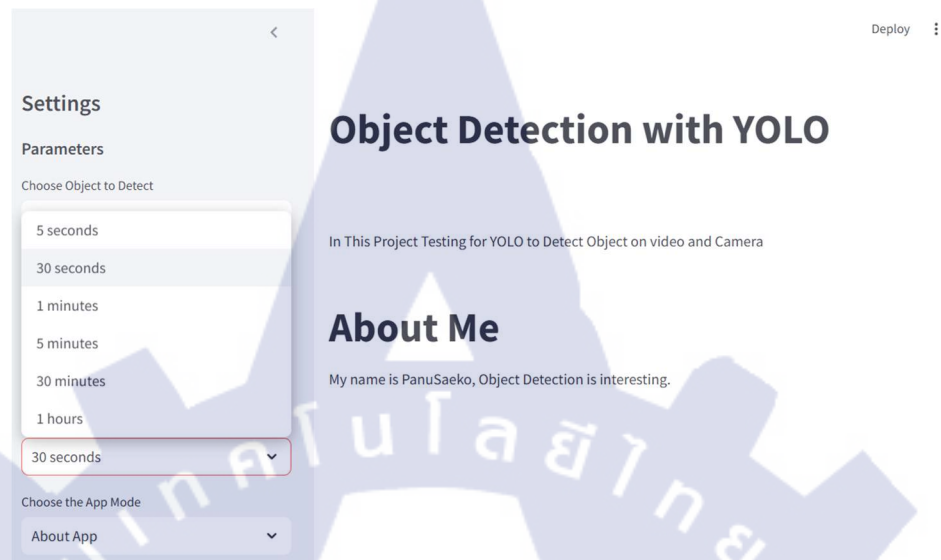
รูปที่ 3.10 เมนูการเลือกประเภทระนาบในการตรวจจับยานพาหนะ

3.5.4 เลือกโมเดลในการตรวจจับ



รูปที่ 3.11 เมนูการเลือกโมเดลในการตรวจจับ

### 3.5.5 เลือกช่วงเวลาในการจัดเก็บและสรุปข้อมูล



รูปที่ 3.12 เมนูการเลือกช่วงเวลาในการเก็บข้อมูล

### 3.5.6 เลือกประเภทของการทำงาน



รูปที่ 3.13 เมนูการเลือกประเภทของงาน

### 3.5.6.1 Draw Detecting Area

ใช้สำหรับกำหนดพื้นที่ในการตรวจจับ โดยในกรณีของยานพาหนะแบบ 1 ทิศทาง จะสามารถกำหนดพื้นที่ได้หลายช่อง โดยเมื่อกำหนดได้แล้วจะสามารถบันทึกการตั้งค่าเก็บไว้ได้

The screenshot displays the 'Object Detection with YOLO' web application interface. On the left is a 'Settings' panel with the following parameters:

- Parameters:**
  - Choose Object to Detect: Vehicle
  - Choose Vehicle Detection Method: 1 Direction
  - Choose Model to Detect: yolo11n.pt
  - Choose Period to Detect: 30 seconds
  - Choose the App Mode: Draw Detecting Area
- Upload a Video:**
  - Drag and drop file here (Limit 200MB per file - MP4, AVI, MOV, ASF, WEBM, MPEG4)
  - Browse files button
  - File: 752895556.790287.mp4 (49.9MB)
  - Input Video player showing a video frame with a duration of 0:00 / 3:39.

The main area is titled 'Object Detection with YOLO' and 'Region Control'. It shows a 'Current Frame' image of a road with two blue bounding boxes around vehicles. Below the image are 'Control points for Region' sliders:

- Top left Point X 3: 540 (range 0 to 1028)
- Top left Point Y 3: 239 (range 0 to 500)
- Bottom Right Point X 3: 604 (range 0 to 1028)
- Bottom Right Point Y 3: 477 (range 0 to 500)

Control buttons include: Save Region 3 (checked), Delete Region 2, and Export All Regions. Below is the 'Imported Regions Data' section with an 'Upload Region Data (JSON)' field and a 'Browse files' button.

รูปที่ 3.14 หน้าจอการกำหนดพื้นที่ในการตรวจจับของยานพาหนะแบบ 1 ทิศทาง

## Object Detection with YOLO

### Region Control

**Settings**

**Parameters**

Choose Object to Detect  
Vehicle

Choose Vehicle Detection Method  
2 Direction

Choose Vertical or Horizontal  
Horizontal

Choose Model to Detect  
yolo11n.pt

Choose Period to Detect  
5 seconds

Choose the App Mode  
Draw Detecting Area

---


Upload a Video

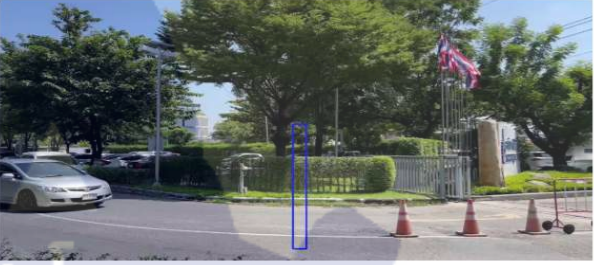
Drag and drop file here  
Limit: 200MB per file • MP4, AVI, MOV, ASF, WEBM, MPEG4

Browse files

752895556.790287.mp4  
49.9MB

Input Video





Current Frame

Control points for Region 1:

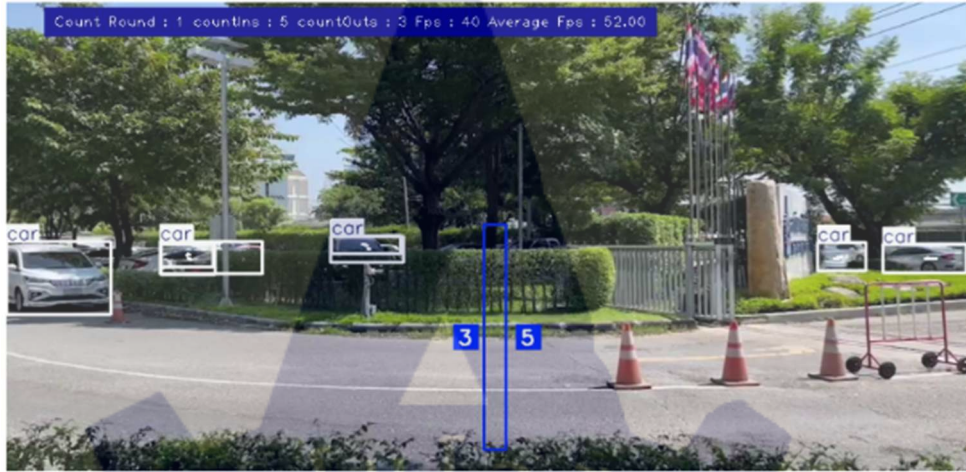
Top left Point X	504	1828
Top left Point Y	239	508
Bottom left Point X	584	1828
Bottom left Point Y	477	508
Bottom Right Point X	527	1828
Bottom Right Point Y	477	508
Top Right Point X	527	1828
Top Right Point Y	239	508

รูปที่ 3.15 หน้าจอการกำหนดพื้นที่ในการตรวจจับทั่วไป

#### 3.5.6.2 Run on video and Camera

ในหน้า Run on video และ Run on Camera จะมีลักษณะการทำงานที่คล้ายกัน โดยจะแสดงผลการตรวจจับแบบ real time พร้อมกับเก็บข้อมูลเป็นช่วง ๆ เพื่อแสดงบนแดชบอร์ด และบันทึกลงไฟล์ .csv

# Object Detection with YOLO



Count Left To Right

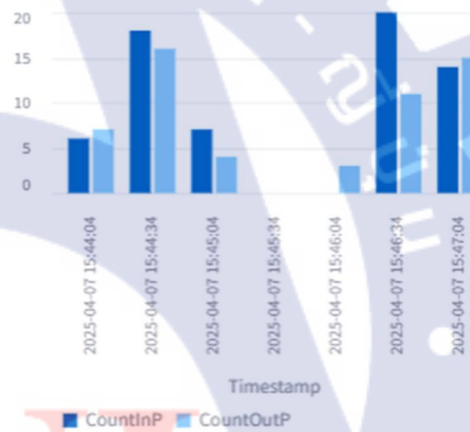
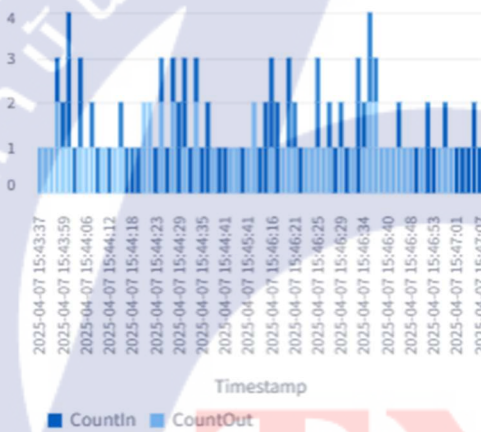
Count Right To Left

**3.0**

**5.0**

CountEachTime

CountPeriodTime



	Timestamp	CountInP	CountOutP	Total CountInP	Total CountOutP
0	2025-04-07 15:44:04	6	7	6	7
1	2025-04-07 15:44:34	18	16	24	23
2	2025-04-07 15:45:04	7	4	31	27
3	2025-04-07 15:45:34	0	0	31	27
4	2025-04-07 15:46:04	0	3	31	30
5	2025-04-07 15:46:34	20	11	51	41
6	2025-04-07 15:47:04	14	15	65	56

รูปที่ 3.16 หน้าจอแสดงผลการตรวจจับและแดชบอร์ด

## บทที่ 4 ผลการวิจัย

จากการพัฒนาเว็บไซต์ที่สร้างด้วย Streamlit พบว่าตัวเว็บไซต์สามารถใช้ตรวจจับวัตถุตามที่ต้องการได้ รวมไปถึงการแสดงผลการตรวจจับผ่านแดชบอร์ด และบันทึกข้อมูลการตรวจจับเป็นไฟล์ csv ได้ แสดงให้เห็นถึงความเป็นไปได้ในการนำไปต่อยอดให้ตัวเว็บไซต์มีฟังก์ชันการทำงานที่มีความหลากหลายและตอบโจทย์ในการนำไปใช้ในแต่ละองค์กรได้

ค่าความแม่นยำ precision recall และ F1-score ของแต่ละโมเดลในสภาพแวดล้อมที่ต่างกันแสดงไว้ในตารางที่ 4.1 ถึง 4.4 โดยในพื้นที่ส่วนแรกของโรงอาหาร YOLOv8m ทำได้ดีที่สุดด้วยความแม่นยำ 0.895 และมีค่า precision 1.0 และ recall 0.895 ส่งผลให้ได้ค่า F1-score 0.944 ซึ่งสูงที่สุด สำหรับพื้นที่ส่วนที่สองของโรงอาหาร โมเดลที่ให้ความแม่นยำสูงสุดคือ YOLOv11s โดยมีความแม่นยำ 0.850 ค่า precision 0.850 และ recall 1.0 ส่งผลให้ได้ค่า F1-score 0.919 สูงสุด โมเดลนี้มีประสิทธิภาพมากที่สุดในการแยกแยะบริเวณที่มีบุคคลหนาแน่นหรืออยู่ในระยะไกล

ข้อจำกัดสำคัญของการตรวจจับบุคคลในแต่ละพื้นที่คือการที่ผู้คนนั่งอยู่ใกล้กันมาก ทำให้โมเดลขนาดเล็กมีความไม่สม่ำเสมอในการตรวจจับ และมักเกิดการตรวจจับซ้ำของบุคคลที่อยู่ใกล้กันตรงข้ามกับโมเดลขนาดใหญ่ที่ช่วยเพิ่มความแม่นยำในการตรวจจับบริเวณชายขอบของวัตถุ อย่างไรก็ตาม ความท้าทายอย่างการมีบุคคลซ้อนทับกัน หรือมีบุคคลอยู่ด้านหลังผู้อื่นยังส่งผลกระทบต่อ การตรวจจับ โมเดลขนาดเล็กมีแนวโน้มเกิดการตรวจจับซ้ำ ขณะที่โมเดลขนาดใหญ่มีโอกาสพลาดการตรวจจับบุคคลที่มีลักษณะไม่ชัดเจน

ในงานตรวจจับยานพาหนะบริเวณทางเข้าสถาบัน โมเดลทุกตัวให้ระดับความแม่นยำและค่า F1-score ใกล้เคียงกัน เนื่องจากพื้นที่ดังกล่าวมียานพาหนะจำนวนไม่มาก และตำแหน่งบันทึกวิดีโออยู่ใกล้กับถนนที่มียานพาหนะผ่านบ่อย ทำให้ความแม่นยำในการตรวจจับของทุกโมเดลอยู่ในระดับสูง ทั้งนี้ ความท้าทายหนึ่งคือการเลือกตำแหน่งบันทึกที่เหมาะสม หากมียานพาหนะจากทั้งสองทิศทางมาถึงพร้อมกัน ยานพาหนะคันหนึ่งอาจบังการมองเห็นของอีกคัน ทำให้ไม่สามารถตรวจจับได้

ความแม่นยำในการตรวจจับมีความสำคัญอย่างยิ่ง และโดยทั่วไปโมเดลที่มีขนาดใหญ่กว่าจะให้ความแม่นยำสูงกว่า แต่ต้องแลกมากับการใช้ทรัพยากรคอมพิวเตอร์มากขึ้นและเวลาประมวลผลที่ยาวนานขึ้น งานวิจัยนี้ประเมินสมรรถนะการประมวลผลทั้งบน CPU และ GPU โดยพบว่า YOLOv8n ให้ความเร็วสูงสุดเนื่องจากเป็นโมเดลที่มีขนาดเล็ก นอกจากนี้ยังพบว่า YOLOv8 ประมวลผลได้เร็วกว่า YOLOv11 เล็กน้อย ทั้งนี้เนื่องจากการปรับปรุงโมเดลของ YOLO อย่างมีนัยสำคัญเพื่อยกระดับ

ความแม่นยำในการตรวจจับ ซึ่งอาจส่งผลกระทบต่อสมรรถนะระหว่างการทำงาน ผลลัพธ์สรุปแสดงไว้ในตารางที่ 4.5

ตารางที่ 4.1 แสดงความแม่นยำในการตรวจจับในแต่ละสถานที่

โมเดล	ร้านอาหารที่ 1	ร้านอาหารที่ 2	ทางเข้าสถาบัน
YOLOv8n	0.640	0.316	0.900
YOLOv8s	0.819	0.667	0.900
YOLOv8m	<b>0.895</b>	0.750	0.900
YOLOv10n	0.250	0.167	0.900
YOLOv10s	0.550	0.500	0.900
YOLOv10m	0.422	0.579	0.900
YOLO11n	0.550	0.278	0.900
YOLO11s	0.773	<b>0.850</b>	0.900
YOLO11m	0.810	0.723	0.900

ตารางที่ 4.2 แสดงค่า precision ในการตรวจจับในแต่ละสถานที่

โมเดล	ร้านอาหารที่ 1	ร้านอาหารที่ 2	ทางเข้าสถาบัน
YOLOv8n	0.842	0.857	1.000
YOLOv8s	0.900	0.889	1.000
YOLOv8m	<b>1.000</b>	0.882	1.000
YOLOv10n	<b>1.000</b>	<b>1.000</b>	1.000
YOLOv10s	0.917	<b>1.000</b>	1.000
YOLOv10m	<b>1.000</b>	<b>1.000</b>	1.000
YOLO11n	0.917	<b>1.000</b>	1.000
YOLO11s	0.850	0.850	1.000
YOLO11m	0.810	0.929	1.000

ตารางที่ 4.3 แสดงค่า recall ในการตรวจจับในแต่ละสถานที่

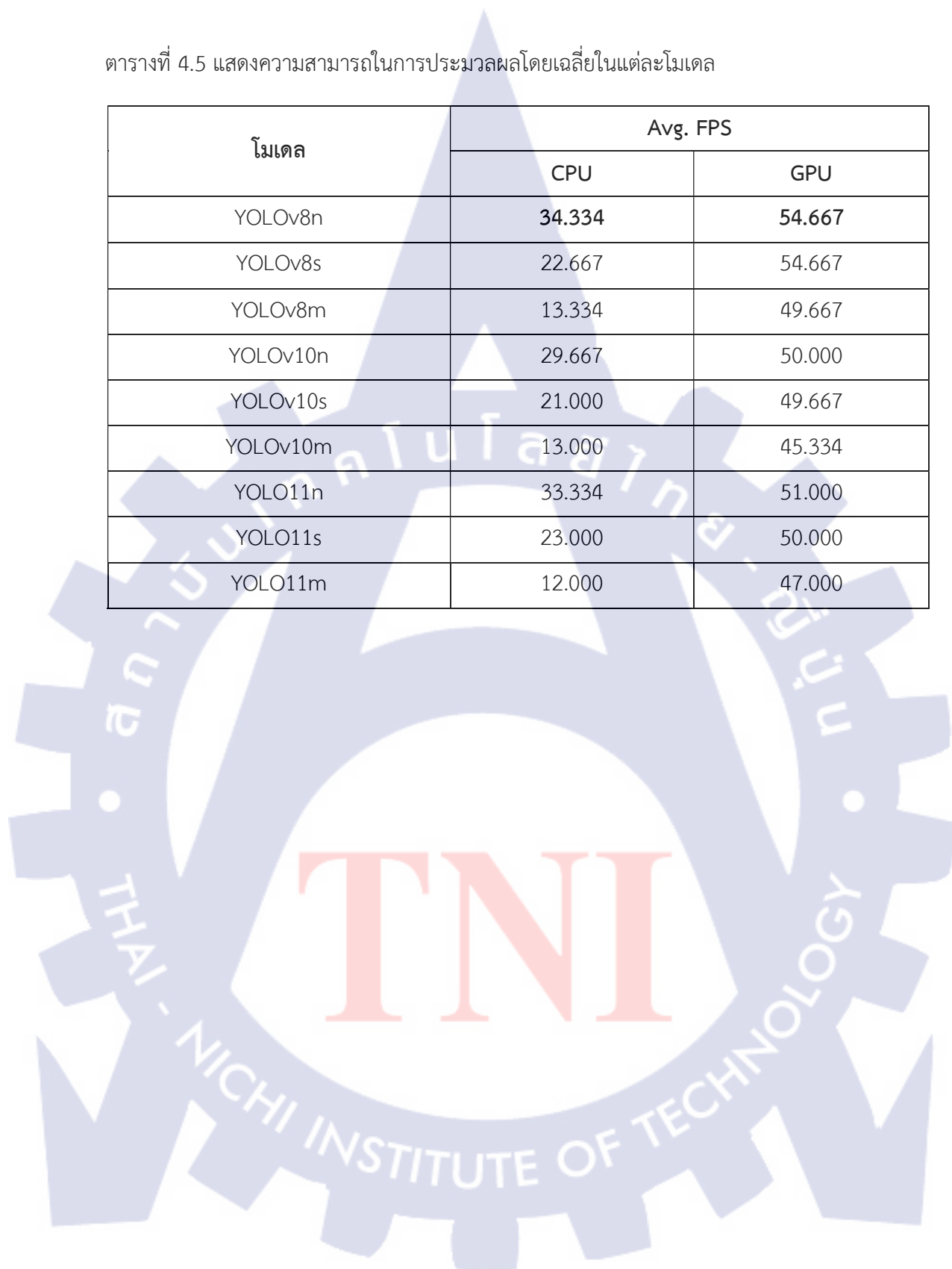
โมเดล	ร้านอาหารที่ 1	ร้านอาหารที่ 2	ทางเข้าสถาบัน
YOLOv8n	0.727	0.333	0.900
YOLOv8s	0.900	0.727	0.900
YOLOv8m	0.895	0.833	0.900
YOLOv10n	0.250	0.167	0.900
YOLOv10s	0.579	0.500	0.900
YOLOv10m	0.421	0.579	0.900
YOLO11n	0.579	0.278	0.900
YOLO11s	0.895	<b>1.000</b>	0.900
YOLO11m	<b>1.000</b>	0.765	0.900

ตารางที่ 4.4 แสดงค่า F1-score ในการตรวจจับในแต่ละสถานที่

โมเดล	ร้านอาหารที่ 1	ร้านอาหารที่ 2	ทางเข้าสถาบัน
YOLOv8n	0.781	0.480	0.947
YOLOv8s	0.900	0.800	0.947
YOLOv8m	<b>0.944</b>	0.857	0.947
YOLOv10n	0.400	0.286	0.947
YOLOv10s	0.710	0.667	0.947
YOLOv10m	0.593	0.733	0.947
YOLO11n	0.710	0.435	0.947
YOLO11s	0.872	<b>0.919</b>	0.947
YOLO11m	0.895	0.839	0.947

ตารางที่ 4.5 แสดงความสามารถในการประมวลผลโดยเฉลี่ยในแต่ละโมเดล

โมเดล	Avg. FPS	
	CPU	GPU
YOLOv8n	<b>34.334</b>	<b>54.667</b>
YOLOv8s	22.667	54.667
YOLOv8m	13.334	49.667
YOLOv10n	29.667	50.000
YOLOv10s	21.000	49.667
YOLOv10m	13.000	45.334
YOLO11n	33.334	51.000
YOLO11s	23.000	50.000
YOLO11m	12.000	47.000



## บทที่ 5

### บทสรุปและข้อเสนอแนะ

การศึกษาวิจัยการสร้างระบบตรวจจับบุคคลที่เข้ามาใช้โรงอาหารและยานพาหนะในลานจอดรถของสถาบันการศึกษาโดยใช้ YOLO สำหรับนำไปใช้ในการจัดการฝูงชน โดยอ้างอิงจากจุดประสงค์ที่กล่าวในบทที่ 1 สามารถสรุปได้ตามหัวข้อต่อไปนี้

- 5.1 สรุปผลการวิจัย
- 5.2 ปัญหาและอุปสรรคในการทำวิจัย
- 5.3 ข้อเสนอแนะงานวิจัย

#### 5.1 สรุปและอภิปรายผลการวิจัย

จากการพัฒนาระบบตรวจจับบุคคลบริเวณโรงอาหารและยานพาหนะที่เข้ามาใช้งานที่จอดรถ ในงานวิจัยนี้ได้นำ Streamlit มาใช้ในการพัฒนาหน้าเว็บไซต์ที่มีความยืดหยุ่นและสามารถนำมาประยุกต์ใช้งานได้สะดวกเนื่องจาก Streamlit มีรูปแบบการจัดวางโครงสร้างเว็บไซต์ที่สำเร็จรูป ทำให้สามารถออกแบบและจัดการหน้าเว็บไซต์ได้ง่ายและรวดเร็ว ทำให้สามารถมุ่งเน้นไปที่การพัฒนา ระบบการตรวจจับเป็นหลักสำหรับนำไปใช้งานร่วมกับเว็บไซต์ที่เตรียมเอาไว้ ตัวระบบการตรวจจับได้ทำการนำ YOLO ที่มีการปรับปรุงและพัฒนาโมเดลอย่างต่อเนื่อง และไม่กระทบต่อผู้ที่นำเวอร์ชันก่อนมาใช้มากนักหลังการอัปเดต การตรวจจับวัตถุโดยทั่วไปที่ไม่ได้เจาะจงในรายละเอียดหรือมีเงื่อนไขเฉพาะสามารถนำ pre train model ของตัว YOLO มาใช้งานได้เลย โดยไม่จำเป็นต้องนำข้อมูลมาฝึกฝนใหม่ด้วยตัวเอง จากวัตถุประสงค์ในงานวิจัยนี้ที่ต้องการตรวจจับตัวบุคคลและยานพาหนะเป็นหลัก จึงไม่จำเป็นต้องฝึกฝนโมเดลใหม่ ทำให้ลดระยะเวลาในการพัฒนาระบบได้เป็นอย่างมาก ด้วยเหตุนี้จึงสามารถมุ่งเน้นไปที่การปรับแต่งการตรวจจับวัตถุให้เป็นไปตามวัตถุประสงค์ที่ระบุไว้ก่อนหน้านี้ได้ เนื่องจากเงื่อนไขในการนับผลการตรวจจับวัตถุแต่ละชนิดไม่เหมือนกัน ความท้าทายในการตั้งเงื่อนไขการนับจำนวนผลลัพธ์จึงเป็นประเด็นสำคัญในงานวิจัยนี้ นอกจากนี้การนำผลลัพธ์ที่ได้มาแสดงผลบนหน้าเว็บไซต์แบบ Real time เป็นข้อดีอย่างหนึ่งของ Streamlit โดยสรุปในงานวิจัยนี้สามารถนำโมเดลการตรวจจับมาทำงานร่วมกับ Streamlit ได้ประสบผลสำเร็จ โดยเว็บไซต์สามารถเลือกชนิดของวัตถุ โมเดล เงื่อนไขในการนับวัตถุ ประเภทข้อมูล และสามารถกำหนดพื้นที่ ๆ ต้องการตรวจจับได้อย่างอิสระ พร้อมกับแสดงผลลัพธ์ในการตรวจจับเบื้องต้นผ่านแดชบอร์ดบนหน้าเว็บไซต์เดียวกันได้ และบันทึกข้อมูลลงไฟล์ csv แบบ real time ทำให้สามารถนำไปต่อยอดในการวิเคราะห์ข้อมูลสำหรับนำไปพัฒนาและบริหารจัดการฝูงชนต่อไป อย่างไรก็ตามตัวเว็บไซต์ที่มี

การทำงานแบบ real time ทำให้ต้องใช้เครื่องประมวลผลที่มีประสิทธิภาพสูงระดับหนึ่ง จึงไม่เหมาะในการนำไปใช้กับอุปกรณ์ที่มีความสามารถในการประมวลผลจำกัด และข้อจำกัดในการวางมุกกล้องเพื่อให้ตอบสนองต่อการใช้งานตามสภาพแวดล้อมที่แตกต่างและยากต่อการคาดเดาพฤติกรรมได้ การเลือกชนิดของโมเดลให้เหมาะสมจึงมีส่วนสำคัญต่อเป้าหมายและการจัดสรรทรัพยากร ทั้งนี้ตัวระบบยังมีส่วนที่สามารถพัฒนาและต่อยอดทั้งในส่วนของเว็บไซต์และตัวโมเดลการตรวจจับเพื่อนำไปใช้ในองค์กรได้อย่างมีประสิทธิภาพมากยิ่งขึ้นต่อไป

## 5.2 ปัญหาและอุปสรรคในการทำวิจัย

### 5.2.1 ปัญหาและปัจจัยด้านอุปกรณ์และการหาข้อมูล

ในการใช้ระบบการตรวจจับผ่าน Streamlit ด้วยโมเดลขนาดใหญ่จำเป็นต้องใช้คอมพิวเตอร์ประสิทธิภาพสูงระดับหนึ่งในการใช้งาน ปัญหาการรวบรวมข้อมูลภาพวิดีโอเนื่องจากจำเป็นต้องบันทึกในช่วงที่มีปริมาณผู้คนเยอะระดับหนึ่งสำหรับวิดีโอบริเวณโรงอาหาร และปริมาณยานพาหนะสำหรับที่จอดรถ แต่เวลาดังกล่าวมักไม่เอื้อต่อผู้วิจัยให้เดินทางมาเก็บข้อมูลได้

### 5.2.2 ความเข้ากันของ Streamlit กับชุดคำสั่งการใช้งานการตรวจจับของ YOLO

การพัฒนาเว็บไซต์เพื่อให้ชุดคำสั่งบนหน้าเว็บไซต์สามารถทำงานร่วมกับชุดคำสั่งในการตรวจจับของ YOLO มีความท้าทายในระดับหนึ่ง เนื่องจากโค้ดในการทำงานจะแตกต่างจากการทำงานแบบแยกออกจากกัน รวมไปถึงการสร้างแดชบอร์ดที่มีความซับซ้อนระดับหนึ่งในการส่งข้อมูลไปมาของทั้งสองระบบ

## 5.3 ข้อเสนอแนะงานวิจัย

ตัวระบบยังสามารถพัฒนาต่อยอดทั้งในส่วนของ Streamlit และระบบการตรวจจับสำหรับ Streamlit ตัวเว็บไซต์ยังมีลูกเล่นอีกมากในการพัฒนาหน้าเว็บไซต์ให้หน้าตื่นตา และสะดวกในการใช้งาน

ในส่วนของระบบตรวจจับวัตถุ YOLO มีการพัฒนาและปรับปรุงอยู่เสมอ ทำให้มีลูกเล่นใหม่ๆ ที่สามารถนำมาเสริมให้กับตัวระบบ รวมไปถึงการกำหนดเงื่อนไขในการนับผลการตรวจจับวัตถุ เพื่อให้สอดคล้องกับความต้องการทั้งต่อชุมชนและองค์กร นอกจากนี้การนำโมเดลที่ได้ทำการฝึกฝนด้วยตัวเองหรือมีการปรับแต่งช่วยให้ระบบสามารถทำการตรวจจับวัตถุที่มีเงื่อนไขเฉพาะทางได้



บรรณานุกรม

TNI

## บรรณานุกรม

- [1] A. E. Berlonghi, “Understanding and planning for different spectator crowds,” *Safety Science*, vol. 18, no. 4, pp. 239–247, February 1995, doi: 10.1016/0925-7535(94)00033-Y.
- [2] ISO/IEC JTC 1, *Smart Cities—Preliminary Report 2014*, Geneva, Switzerland: ISO/IEC, 2014.
- [3] V. L. Uskov et al., “Smart universities: Concepts, systems and technologies,” *Smart Innovation, Systems and Technologies*, vol. 70, pp. 47-188, May 2017, doi: 10.1007/978-3-319-59454-5.
- [4] J. Anttila and K. Jussila, “Universities and smart cities: the challenges to high quality,” *Total Quality Management & Business Excellence*, vol. 29, no. 9–10, pp. 1058–1073, July 2018, doi: 10.1080/14783363.2018.1486552.
- [5] L. Tan et al., “Comparison of retinaNet, SSD, and YOLO v3 for real-time pill identification,” *BMC Med Inform Decis Mak*, vol. 21, no. 1, pp. 324, December 2021, doi: 10.1186/s12911-021-01691-8.
- [6] A. Duraj and D. Duczynski, “Nested binary classifier as an outlier detection method in human activity recognition systems,” *Entropy*, vol. 25, no. 8, pp. 1121, July 2023, doi: 10.3390/e25081121.
- [7] J.-J. Lin et al., “Machine learning for human motion intention detection,” *Sensors*, vol. 23, no. 16, pp. 7203, August 2023, doi: 10.3390/s23167203.
- [8] R. Thomas et al., “Detection and localization of suspicious objects including human intruder,” *The 24th Topical Conference on Radio Frequency Power In Plasmas*, Annapolis, USA, August 18, 2023, pp. 020012, doi: 10.1063/5.0165274.

- [9] F. Styrzynski et al., “Machine learning successfully detects patients with COVID-19 prior to PCR results and predicts their survival based on standard laboratory parameters in an observational study,” *Infectious Diseases and Therapy*, vol. 12, no. 1, pp. 111–129, January 2023, doi: 10.1007/s40121-022-00707-8.
- [10] M. Yang et al., “Cluster-based crowd movement behavior detection,” in *2018 Digital Image Computing: Techniques and Applications (DICTA)*, Canberra, Australia, December 2018, pp. 1–8, doi: 10.1109/DICTA.2018.8615809.
- [11] Y. Song et al., “Vehicle target detection method for wide-area SAR images based on coarse-grained judgment and fine-grained detection,” *Remote Sensing*, vol. 15, no. 13, pp. 3242, June 2023, doi: 10.3390/rs15133242.
- [12] A. Elbery et al., “IoT-Based crowd management framework for departure control and navigation,” *IEEE Transactions on Vehicular Technology*, vol. 70, no. 1, pp. 95–106, January 2021, doi: 10.1109/TVT.2020.3048336.
- [13] P. Hidayatullah et al., “License plate detection and recognition for Indonesian cars,” *International Journal on Electrical Engineering and Informatics*, vol. 8, no. 2, pp. 331–346, June 2016, doi: 10.15676/ijeei.2016.8.2.7.
- [14] D. P. Carrasco et al., “T-YOLO: Tiny vehicle detection based on YOLO and multi-scale convolutional neural networks,” *IEEE Access*, vol. 11, pp. 22430–22440, March 2023, doi: 10.1109/ACCESS.2021.3137638.
- [15] S. Voeneky et al., “*The Cambridge Handbook of Responsible Artificial Intelligence: Interdisciplinary Perspectives*,” Cambridge, U.K.: Cambridge University Press, 2022, doi: 10.1017/9781009207898.
- [16] J. Brownlee, *Deep Learning for Computer Vision: Image Classification Object Detection, and Face Recognition in Python*, San Francisco, USA: Machine Learning Mastery, 2019.
- [17] J. Redmon et al., “You Only Look Once: *Unified, Real-Time Object Detection*,” arXiv, May 9, 2016, doi: 10.48550/arXiv.1506.02640.

- [18] J. Terven and D. C. Esparza, "A Comprehensive Review of YOLO: From YOLOv1 and Beyond," arXiv, August 7, 2023, [Online]. Available: <http://arxiv.org/abs/2304.00501>. [Accessed: September 17, 2023].
- [19] Z. Zheng et al., "Distance-IoU Loss: Faster and Better Learning for Bounding Box Regression," arXiv, November 19, 2019, doi: 10.48550/arXiv.1911.08287.
- [20] C. Li et al., "YOLOv6: A Single-Stage Object Detection Framework for Industrial Applications," arXiv, September 7, 2022, doi: 10.48550/arXiv.2209.02976.
- [21] C. Y. Wang et al., "YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors," arXiv, July 6, 2022, doi: 10.48550/arXiv.2207.02696.
- [22] Wang et al., "YOLOv9: Learning What You Want to Learn Using Programmable Gradient Information," arXiv, February 29, 2024, [Online]. Available: <https://doi.org/10.48550/arXiv.2402.13616>. [Accessed: February 29, 2024].
- [23] Wang et al., "YOLOv10: Real-Time End-to-End Object Detection," arXiv, October 30, 2024, [Online]. Available: <https://doi.org/10.48550/arXiv.2405.14458>.
- [24] W. Liu et al., "Image-adaptive YOLO for object detection in adverse weather conditions," *AAAI*, vol. 36, no. 2, pp. 1792–1800, June 2022, doi: 10.1609/aaai.v36i2.20072.
- [25] Y. Zhang et al., "Real-time vehicle detection based on improved YOLO v5," *Sustainability*, vol. 14, no. 19, pp. 12274, September 2022, doi: 10.3390/su141912274.
- [26] P. Somaldo et al., "Developing smart COVID-19 social distancing surveillance drone using YOLO implemented in robot operating system simulation environment," in *2020 IEEE 8th R10 Humanitarian Technology Conference (R10-HTC)*, Chiba University of Commerce, Japan, December 3, 2020, pp. 1–6, doi: 10.1109/R10-HTC49770.2020.9357040.

- [27] C. Jiang et al., "Object detection from UAV thermal infrared images and videos using YOLO models," *International Journal of Applied Earth Observation and Geoinformation*, vol. 112, Art. no. 102912, no page, August 2022, doi: 10.1016/j.jag.2022.102912.
- [28] H. Mokayed et al., "Real-time human detection and counting system using deep learning computer vision techniques," *Artificial Intelligence and Applications*, vol. 1, no. 4, pp. 205-213, October 2022, doi: 10.47852/bonviewAIA2202391.
- [29] C. Dong et al., "PG-YOLO: A novel lightweight object detection method for edge devices in industrial internet of things," *IEEE Access*, vol. 10, pp. 123736–123745, November 2022, doi: 10.1109/ACCESS.2022.3223997.
- [30] M. H. K. Khel et al., "Realtime crowd monitoring—estimating count, speed and direction of people using hybridized YOLOv4," *IEEE Access*, vol. 11, pp. 56368–56379, May 2023, doi: 10.1109/ACCESS.2023.3272481.



## ประวัติย่อผู้วิจัย



ชื่อ - สกุล

นายภานุพงศ์ แซ่โก

วัน เดือน ปีเกิด

13 ธันวาคม 2534

ที่อยู่ปัจจุบัน

6/153 ตำบล สำโรงเหนือ อำเภอบางพลี จังหวัด สมุทรปราการ 10270

โทรศัพท์มือถือ : 061-994-6599

E-mail: pangya\_saeko@hotmail.com

## ประวัติการศึกษา

พ.ศ. 2568

วิทยาศาสตร์มหาบัณฑิต สาขาเทคโนโลยีสารสนเทศ

สถาบันเทคโนโลยีไทย-ญี่ปุ่น

พ.ศ. 2556

วิทยาศาสตร สาขาสถิติประยุกต์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

## ประวัติการทำงาน

พ.ศ. 2565 - ปัจจุบัน

ผู้ช่วยวิทยากร

บริษัท นายน์เอ็กซ์เพิร์ท จำกัด

พ.ศ. 2558 - 2565

ประสานงานผู้ค้าต่างประเทศ

Growthstairlifts.Ltd

พ.ศ. 2556 - 2558

เจ้าหน้าที่ฝ่ายการค้าและการประสานงานระหว่างทีม  
บริษัทหลักทรัพย์จัดการกองทุน กรุงศรี จำกัด