การวัดค่าระดับน้ำโดยใช้วิธีการตรวจจับวัตถุและการเรียนรู้เชิงลึก

WATER LEVEL MONITORING USING OBJECT DETECTION AND DEEP LEARNING

เปมิกา คุณาวิริยะสิริ
PEMIKA CUNAVIRIYASIRI

การค้นคว้าอิสระนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาวิทยาการข้อมูลและการวิเคราะห์
ศูนย์วิเคราะห์ข้อมูลดิจิทัลอัจฉริยะพระจอมเกล้าลาดกระบัง
คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
พ.ศ. 2566
KMITL-2023-SC-M-017-044

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

WATER LEVEL MONITORING USING OBJECT DETECTION AND DEEP LEARNING

PEMIKA CUNAVIRIYASIRI

AN INDEPENDENT STUDY SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENT FOR THE DEGREE OF MASTER OF SCIENCE
IN DATA SCIENCE AND ANALYTICS
KMITL DIGITAL ANALYTICS AND INTELLIGENCE CENTER SCHOOL OF SCIENCE
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG
2023

KMITL-2023-SC-M-017-044

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

COPYRIGHT 2023
SCHOOL OF SCIENCE
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อการค้นคว้าอิสระ การวัดค่าระดับน้ำโดยใช้วิธีการตรวจจับวัตถุและการเรียนรู้เชิงลึก
ชื่อนักศึกษา เปมิกา คุณาวิริยะสิริ
รหัสประจำตัว 64605075
ปริญญา วิทยาศาสตรมหาบัณฑิต (วิทยาการข้อมูลและการวิเคราะห์) ศูนย์วิเคราะห์ข้อมูลดิจิทัลอัจฉริยะพระจอมเกล้าลาดกระบัง
พ.ศ. 2566
อาจารย์ที่ปรึกษาการค้นคว้าอิสระ ผู้ช่วยศาสตราจารย์ ดร.บุษยมาส พิมพ์พรรณชาติ

บทคัดย่อ

งานวิจัยนี้มีจุดประสงค์เพื่อศึกษาและเปรียบเทียบการอ่านค่าวัดระดับน้ำจากรูปภาพจากกล้องวงจรปิด (Closed-circuit Television : CCTV) โดยได้มีการสืบค้นรูปภาพไม้วัดระดับน้ำที่มีการใช้งานจริงตามสถานที่ต่าง ๆ มากกว่า 200 รูปภาพจากเว็บไซต์ต่าง ๆ มาใช้ในงานวิจัยครั้งนี้ ได้มีการใช้การจัดแต่งรูปภาพรูปแบบต่าง ๆ เช่น การปรับความสว่างของภาพ และ การเพิ่มสัญญาณรบกวนให้แก่ภาพ เพื่อทำให้มีข้อมูลที่เพียงพอต่อการเรียนรู้และสามารถสื่อถึงแต่ละสถานการณ์จริงที่เกิดขึ้นได้ งานวิจัยนี้ได้มีการนำตัวแบบที่ผ่านการเรียนรู้มาแล้วชื่อว่า YOLOv8m มาใช้กับชุดข้อมูลที่ได้จัดเตรียมไว้สำหรับขั้นตอนของการเรียนรู้ โดยมีจุดประสงค์เพื่อระบุตำแหน่งของไม้วัดระดับน้ำที่อยู่ในรูปภาพ นอกจากนี้ งานวิจัยได้มีการใช้ตัวแบบที่ผ่านการเรียนรู้มาแล้วอีกครั้งหนึ่งในการอ่านค่าและระบุตำแหน่งของตัวเลขที่อยู่ในภาพ เพื่ออ่านค่าว่าระดับน้ำในรูปภาพอยู่ตรงกับตัวเลขใด และระดับน้ำควรอ่านค่าออกมาเป็นเท่าใด โดยการเรียนรู้ของตัวแบบที่สอง ได้มีการใช้รูปภาพจาก Roboflow Universe ที่เป็นชุดข้อมูลของรูปภาพตัวเลขที่มีลักษณะคล้ายกับตัวเลขที่อยู่บนไม้วัดระดับน้ำมากกว่า 600 รูปภาพ มาใช้สำหรับการเรียนรู้ของตัวแบบที่สองในการอ่านค่าและระบุตำแหน่งของตัวเลขจากไม้วัดระดับน้ำ ทั้งสองการเรียนรู้ได้มีการทดลองโดยการแบ่งข้อมูลสำหรับการเรียนรู้ การทดสอบ และการตรวจสอบความถูกต้องออกเป็นทั้งหมด 4 รูปแบบ และนำผลลัพธ์ที่ได้จากทั้งสองตัวแบบมาเปรียบเทียบประสิทธิภาพโดยใช้ค่าความเที่ยง ค่าเรียกคืน และค่ากลางของค่าเฉลี่ยความเที่ยง เพื่อหาการแบ่งส่วนของข้อมูลที่ทำให้ตัวแบบมีประสิทธิภาพที่สุด หลังจากนั้น ได้มีการพัฒนาหน้าต่างสำหรับการแสดงผล โดยหน้าต่างการแสดงผลประกอบด้วย ภาพจากกล้องวงจรปิด กราฟเส้นแสดงระดับน้ำในปัจจุบันและอดีตรายชั่วโมง ค่าระดับน้ำ สถานที่ และสถานะการเปิดและปิดของประตูระบายน้ำ

คำสำคัญ : การตรวจจับวัตถุ การจัดแต่งรูปภาพ การติดป้ายกำกับ ค่าเฉลี่ยความเที่ยง และไม้วัดระดับน้ำ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Independent Study Title	Water Level Monitoring Using Object Detection and Deep Learning
Student Name	Pemika Cunaviriyasiri
Student ID	64605075
Degree	Master of Science (Data Science and Analytics)
	KMITL-Digital Analytics and Intelligence Center
Year	2023
Independent Study Advisor	Asst.Prof.Dr. Busayamas Pimpunchat

Abstract

The objective of this research is to study and observe the detection of water level measurements from an image that was snapped by CCTV cameras. More than 200 images of water level scale that are actually used in different locations have been used in this research. Augmentations techniques, such as adjusting image brightness and adding noise to an image, were being applied to provide enough images to represent each situation as it really happened. In this research, a pre-trained model, YOLOv8m, was applied to the dataset provided for the learning process. The purpose of the first model is to detect the position and get the location of the water scale in the image. Another pre-trained model, YOLOv8m, was also used to read and locate the number from the image obtained by the first model to get the water level from the water scale in the image. Both models have been trained with different types of partitioning for train, validation, and test. The results were compared by using precision, recall, and mean average precision (mAP). Last, a user interface which shows images from CCTV camera, line graph which represents water level hourly, current water level, location, and the door status of the water gate has been implemented to show to use case of this research.

Keywords : Object Detection, Image Augmentation, Image Annotation, Mean Average Precision, Water Level Scale

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

งานวิจัยเรื่อง การวัดค่าระดับน้ำโดยใช้วิธีตรวจจับวัตถุและการเรียนรู้ สามารถดำเนินการจนประสบความสำเร็จลุล่วงไปด้วยดี เนื่องจากได้รับความอนุเคราะห์และสนับสนุนเป็นอย่างดียิ่งจาก ผศ.ดร.บุษยมาส พิมพ์พรรณชาติ อาจารย์ที่ปรึกษาการค้นคว้าอิสระที่ได้กรุณาให้คำปรึกษา ให้ความรู้ ข้อคิดเห็น ข้อแนะนำ และปรับปรุงแก้ไขข้อบกพร่องต่าง ๆ ตั้งแต่เริ่มดำเนินการจนกระทั่งดำเนินการเสร็จสมบูรณ์ ผู้วิจัยขอกราบขอบพระคุณเป็นอย่างสูงไว้ ณ โอกาสนี้

ขอขอบคุณอาจารย์และเพื่อนคณะวิทยาศาสตร์สาขาวิทยาการข้อมูลและการวิเคราะห์ทุกท่านที่ให้ความรู้ ถ่ายทอดประสบการณ์ทางวิชาการที่มีประโยชน์ และช่วยเหลือในการทำการวิจัยในครั้งนี้

สุดท้ายนี้ ผู้วิจัยหวังว่างานวิจัยฉบับนี้จะเป็นประโยชน์สำหรับหน่วยงานที่เกี่ยวข้อง และผู้ที่สนใจศึกษาต่อไป

เปมิกา คุณาวิริยะสิริ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

	หน้า
บทคัดย่อภาษาไทย	ก
บทคัดย่อภาษาอังกฤษ	ข
กิตติกรรมประกาศ	ค
สารบัญ	ง
สารบัญตาราง	ช
สารบัญรูป	ซ
บทที่ 1 บทนำ	1
1.1 ความเป็นมาและความสำคัญของปัญหา	2
1.2 วัตถุประสงค์ของงานวิจัย	2
1.3 ขอบเขตของงานวิจัย	2
1.4 ประโยชน์ที่คาดว่าจะได้รับ	2
1.5 ขั้นตอนการดำเนินงาน	2
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง	4
2.1 การประมวลผลภาพ	4
2.1.1 การไล่ระดับสีเทา	4
2.1.2 การทำให้รูปภาพไม่ชัด	5
2.1.3 การทำการหาจุดเริ่มเปลี่ยน	6
2.2 การอ่านข้อความจากรูปภาพ	8
2.3 ตัวแบบที่ผ่านการเรียนรู้มาแล้ว	8
2.4 การตรวจจับวัตถุแบบไร้ที่ยึดเหนี่ยว	11
2.4.1 ส่วนหลัง	12
2.4.2 คุณลักษระพีระมิด	12
2.4.3 ส่วนหัว	12
2.5 การจัดแต่งรูปภาพ	13
2.5.1 การหมุนรูปภาพ	13
2.5.2 การเลื่อนรูปภาพ	13
2.5.3 การพลิกรูปภาพ	13
2.5.4 การเพิ่มความหยาบของรูปภาพ	13
2.5.5 การจัดแต่งรูปภาพแบบโมเสค	13
2.6 ความแม่นยำ RF100	14
2.7 โครงข่ายประสาทเทียมแบบคอนโวลูชัน	14
2.7.1 ชั้นแบบคอนโวลูชัน	15
2.7.2 ชั้นแบบพูลลิง	16
2.7.3 ชั้นที่เชื่อมกันอย่างสมบูรณ์	16
2.7.4 ตัวอย่างงานที่มีการนำโครงข่ายปราสาทเทียมแบบคอนโวลูชันไปใช้	17
2.7.5 ตัวแบบจากโครงข่ายประสาทเทียม	18

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

หน้า

2.8 มาตรวัดประสิทธิภาพของตัวแบบตรวจจับวัตถุ 22
2.8.1 ค่าความเที่ยง 22
2.8.2 ค่าเรียกคืน 23
2.8.3 ค่าคะแนน F1 23
2.8.4 Intersection Over Union (IoU) 23
2.8.5 ค่าเฉลี่ยความเที่ยง 23
2.8.6 ค่าเฉลี่ยความเที่ยง COCO 24
2.9 งานวิจัยที่เกี่ยวข้อง 24
บทที่ 3 วิธีการดำเนินงานวิจัย 27
3.1 รูปภาพที่ใช้ในตัวแบบที่ 1 28
3.2 ติดป้ายกำกับให้กับไม้วัดระดับน้ำ 29
3.3 จัดแต่งรูปภาพจากภาพที่มีอยู่ให้กับตัวแบบที่ 1 30
3.4 การเรียนรู้ของตัวแบบที่ 1 31
3.5 รูปภาพที่ใช้ในตัวแบบที่ 2 32
3.6 ติดป้ายกำกับให้กับข้อมูลตัวเลข 33
3.7 การเรียนรู้ของตัวแบบที่ 2 35
3.8 การหาอัตราส่วนของรูปภาพ 37
3.9 สร้างหน้าต่างสำหรับการใช้งาน 38
3.10 การทำงานของระบบ 39
บทที่ 4 ผลการวิจัยและการอภิปรายผล 41
4.1 ผลการวิเคราะห์การประเมินประสิทธิภาพของตัวแบบที่ 1 41
4.1.1 การประเมินประสิทธิภาพของการทดสอบที่ 1 ตัวแบบที่ 1 41
4.1.2 การประเมินประสิทธิภาพของการทดสอบที่ 2 ตัวแบบที่ 1 43
4.1.3 การประเมินประสิทธิภาพของการทดสอบที่ 3 ตัวแบบที่ 1 45
4.1.4 การประเมินประสิทธิภาพของการทดสอบที่ 4 ตัวแบบที่ 1 46
4.1.5 ผลการวิเคราะห์ประสิทธิภาพของตัวแบบที่ 1 48
4.2 ผลการวิเคราะห์การประเมินประสิทธิภาพของตัวแบบที่ 2 49
4.2.1 การประเมินประสิทธิภาพของการทดสอบที่ 1 ตัวแบบที่ 2 49
4.2.2 การประเมินประสิทธิภาพของการทดสอบที่ 2 ตัวแบบที่ 2 51
4.2.3 การประเมินประสิทธิภาพของการทดสอบที่ 3 ตัวแบบที่ 2 53
4.2.4 การประเมินประสิทธิภาพของการทดสอบที่ 4 ตัวแบบที่ 2 55
4.2.5 ผลการวิเคราะห์ประสิทธิภาพของตัวแบบที่ 2 57
4.3 การวัดสัดส่วนเพื่อหาค่าระดับน้ำ 59
บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ 60
5.1 สรุปผลการวิจัย 60
5.2 ข้อเสนอแนะ 60

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

	หน้า
เอกสารอ้างอิง	61
ประวัติผู้เขียน	63

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตารางที่	หน้า
2.1 รายละเอียดประสิทธิภาพของทำงานของตัวแบบ YOLOv8 รูปแบบต่าง ๆ	11
2.2 ประสิทธิภาพของงานวิจัย	26
3.1 การแบ่งข้อมูลสำหรับการเรียนรู้ของตัวแบบที่ 1	32
3.2 การแบ่งข้อมูลสำหรับการเรียนรู้ของตัวแบบที่ 2	35
3.3 รายละเอียดการติดป้ายกำกับให้กับชุดข้อมูลที่สองในการทดสอบที่ 1	36
3.4 รายละเอียดการติดป้ายกำกับให้กับชุดข้อมูลที่สองในการทดสอบที่ 2	36
3.5 รายละเอียดการติดป้ายกำกับให้กับชุดข้อมูลที่สองในการทดสอบที่ 3	37
3.6 รายละเอียดการติดป้ายกำกับให้กับชุดข้อมูลที่สองในการทดสอบที่ 4	37
3.7 โครงสร้างของการจัดเก็บข้อมูลระดับน้ำลงบนฐานข้อมูล	38
4.1 ผลลัพธ์ของตัวแบบที่ 1	401
4.2 ผลลัพธ์ของตัวแบบที่ 2	49

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป

รูปที่	หน้า
2.1 ความแตกต่างของภาพ RGB และภาพที่มีระดับสีเทา	4
2.2 สถาปัตยกรรมของตัวแบบ YOLOv8	10
2.3 สถาปัตยกรรมของตัวแบบ Fully Convolutional One-stage	12
2.4 แผนภาพรูปกล่องของการประเมินค่า mAP ของทั้ง 3 ตัวแบบ	14
2.5 การตรวจจับวัตถุในระบบอัตโนมัติของยานพาหนะ	17
2.6 การตรวจจับใบหน้า	18
3.1 ขั้นตอนการทำงานของตัวแบบที่ 1	27
3.2 ขั้นตอนการทำงานของตัวแบบที่ 2	28
3.3 รูปภาพตัวอย่างของชุดข้อมูลนำมาใช้ในตัวแบบที่ 1	28
3.4 ตัวอย่างผลลัพธ์ที่ต้องการจากตัวแบบที่ 1	29
3.5 ตำแหน่งของจุดสี่จุดล้อมรอบไม้วัดระดับน้ำ	29
3.6 การปรับแสงของรูปภาพสำหรับการเรียนรู้	30
3.7 การเพิ่มสัญญาณรบกวนให้กับรูปภาพสำหรับการเรียนรู้	31
3.8 รูปภาพตัวอย่างของชุดข้อมูลที่นำมาใช้ในตัวแบบที่ 2	33
3.9 ตัวอย่างการติดป้ายกำกับให้กับข้อมูลตัวเลข	34
3.10 ตำแหน่งของจุดสี่จุดล้อมรอบตัวเลข	34
3.11 การทำงานของระบบการอ่านค่าระดับน้ำจากกล้องวงจรปิด (CCTV)	39
3.12 หน้าต่างสำหรับการมอนิเตอร์ระดับน้ำ	39
4.1 ผลลัพธ์ของเมทริกซ์ความสับสนของการทดสอบที่ 1 ตัวแบบที่ 1	42
4.2 เส้นโค้งความเชื่อมั่นคะแนน F1 ของการทดสอบที่ 1 ตัวแบบที่ 1	42
4.3 เส้นโค้งค่าความเที่ยงและค่าเรียกคืนของการทดสอบที่ 1 ตัวแบบที่ 1	43
4.4 ผลลัพธ์ของเมทริกซ์ความสับสนของการทดสอบที่ 2 ตัวแบบที่ 1	43
4.5 เส้นโค้งความเชื่อมั่นคะแนน F1 ของการทดสอบที่ 2 ตัวแบบที่ 1	44
4.6 เส้นโค้งค่าความเที่ยงและค่าเรียกคืนของการทดสอบที่ 2 ตัวแบบที่ 1	44
4.7 ผลลัพธ์ของเมทริกซ์ความสับสนของการทดสอบที่ 3 ตัวแบบที่ 1	45
4.8 เส้นโค้งความเชื่อมั่นคะแนน F1 ของการทดสอบที่ 3 ตัวแบบที่ 1	45
4.9 เส้นโค้งค่าความเที่ยงและค่าเรียกคืนของการทดสอบที่ 3 ตัวแบบที่ 1	46
4.10 ผลลัพธ์ของเมทริกซ์ความสับสนของการทดสอบที่ 4 ตัวแบบที่ 1	46
4.11 เส้นโค้งความเชื่อมั่นคะแนน F1 ของการทดสอบที่ 4 ตัวแบบที่ 1	47
4.12 เส้นโค้งค่าความเที่ยงและค่าเรียกคืนของการทดสอบที่ 4 ตัวแบบที่ 1	47
4.13 การตรวจจับวัตถุของการทดสอบที่ 3 ของตัวแบบที่ 1	48
4.14 ผลลัพธ์ของตัวแบบที่ 1	48
4.15 ผลลัพธ์ของเมทริกซ์ความสับสนของการทดสอบที่ 1 ตัวแบบที่ 2	50
4.16 เส้นโค้งความเชื่อมั่นคะแนน F1 ของการทดสอบที่ 1 ตัวแบบที่ 2	50
4.17 เส้นโค้งค่าความเที่ยงและค่าเรียกคืนของการทดสอบที่ 1 ตัวแบบที่ 2	51

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป (ต่อ)

รูปที่	หน้า
4.18 ผลลัพธ์ของเมทริกซ์ความสับสนของการทดสอบที่ 2 ตัวแบบที่ 2	52
4.19 เส้นโค้งความเชื่อมั่นคะแนน F1 ของการทดสอบที่ 2 ตัวแบบที่ 2	52
4.20 เส้นโค้งค่าความเที่ยงและค่าเรียกคืนของการทดสอบที่ 2 ตัวแบบที่ 2	53
4.21 ผลลัพธ์ของเมทริกซ์ความสับสนของการทดสอบที่ 3 ตัวแบบที่ 2	54
4.22 เส้นโค้งความเชื่อมั่นคะแนน F1 ของการทดสอบที่ 3 ตัวแบบที่ 2	54
4.23 เส้นโค้งค่าความเที่ยงและค่าเรียกคืนของการทดสอบที่ 3 ตัวแบบที่ 2	55
4.24 ผลลัพธ์ของเมทริกซ์ความสับสนของการทดสอบที่ 4 ตัวแบบที่ 2	56
4.25 เส้นโค้งความเชื่อมั่นคะแนน F1 ของการทดสอบที่ 4 ตัวแบบที่ 2	56
4.26 เส้นโค้งค่าความเที่ยงและค่าเรียกคืนของการทดสอบที่ 4 ตัวแบบที่ 2	57
4.27 การตรวจจับวัตถุของการทดสอบที่ 1 ของตัวแบบที่ 2	58
4.28 ผลลัพธ์ของตัวแบบที่ 2	58
4.29 การหาความสูงของรูปภาพ	59
4.30 การวัดมาตราส่วนในการอ่านค่าระดับน้ำ	59

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1
บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ปัญหาอุทกภัยยังคงเป็นปัญหาสำคัญที่ประเทศไทยประสบพบเจอและยังคงหลีกเลี่ยงไม่ได้ ในปี พ.ศ. 2565 ฤดูฝนเริ่มต้นเร็วกว่าปกติอันเนื่องมาจากปรากฏการณ์ลานีญา หรือเป็นช่วงที่อุณหภูมิพื้นผิวน้ำทะเลเพิ่มสูงขึ้น ส่งผลให้หลายพื้นที่มีปริมาณน้ำฝนสะสมสูง

ข่าวจาก BBC News ไทย (2022) ได้กล่าวว่าเมื่อปี พ.ศ. 2565 มีพายุ 3 ลูก ได้แก่ พายุ “มู่หลาน” และพายุ “หมาอ๊อน” ที่เคลื่อนตัวมาแล้วอ่อนกำลังลงก่อนถึงไทย จึงส่งผลกระทบไม่มาก แต่ล่าสุดพายุ “โนรู” พัดเข้าไทยโดยตรงเมื่อปลายเดือนกันยายนที่ผ่านมา ส่งผลให้มีฝนตกหนักถึงหนักมากแทบทั้งประเทศ ไม่ว่าจะเป็น ภาคเหนือ ภาคอีสาน ภาคกลาง กรุงเทพมหานครและปริมณฑล ภาคตะวันออก และภาคใต้

กระบวนการจัดสรรน้ำจึงเป็นกระบวนการที่สำคัญสำหรับการรองรับปัญหาอุทกภัย เครื่องมืออุปกรณ์และเทคนิคการควบคุมการส่งน้ำในปัจจุบันยังไม่ได้รับการพัฒนาเท่าที่ควร งานวิจัยของ Joe Hughes (2023) ได้กล่าวว่าการควบคุมน้ำในคลองยังคงเป็นการควบคุมด้วยมือ ซึ่งต้องใช้พนักงานส่งน้ำจำนวนมาก และปัญหาที่ทุกโครงการชลประทานประสบอยู่คือขาดอัตรากำลังด้านงานส่งน้ำ ถึงแม้ว่าจะมีการนำเอาประตูน้ำอัตโนมัติแบบกลไฮโดร (Hydro-mechanical) มาทดลองใช้ แต่การใช้งานยังจำกัดเฉพาะบางคลองในบางโครงการเท่านั้น การใช้เครื่องรับรู้ในการวัดปริมาณน้ำในแม่น้ำมีข้อจำกัดหลากหลายจากปัจจัยภายนอก เช่น ความชื้นในอากาศ อุณภูมิ ความดันบรรยากาศ และสารเคมีบางชนิด ปัจจัยเหล่านี้ส่งผลเสียต่อเครื่องรับรู้ ทำให้มีการวัดที่ผิดพลาดเกิดขึ้น จึงทำให้การวัดระดับน้ำด้วยเครื่องรับรู้นั้นยังไม่ใช่ทางออกที่เหมาะสมสำหรับทุกสถานการณ์

งานวิจัยของ ธริศร์ (2018) ได้กล่าวว่าในช่วง 30 ปีที่ผ่านมา การประมวลผลจากภาพดิจิทัลมีการพัฒนาอย่างรวดเร็วและมีการนำมาประยุกต์ใช้กันอย่างแพร่หลาย คอมพิวเตอร์ได้เข้ามามีบทบาทในการประมวลผลแทนระบบประมวลผลของมนุษย์ การประมวลผลด้วยภาพมีประสิทธิภาพสูงและมีการประมวลผลแบบอัตโนมัติ

Krizhevsky ได้มีการสร้าง AlexNet จากทฤษฎีการเรียนรู้เชิงลึกและคว้าชัยชนะในการแข่งขันการประมวลจากภาพประเภท ILSVRC AlexNet มีข้อได้เปรียบทั้งด้านการจำแนกประเภทและด้านความรวดเร็วในการประมวลผล หลังจากนั้นไม่นาน ได้มีการพัฒนา CNN หรือขั้นตอนวิธีการเรียนรู้เชิงลึก เพื่อนำเสนอวิธีการใหม่สำหรับการประมวลผลจากภาพ ผู้ชนะในการแข่งขัน ILSVRC ลำดับต่อมา ได้มีการใช้ขั้นตอนวิธีการจำแนกประเภทจากภาพที่เป็นการเรียนรู้เชิงลึกทั้งหมด ไม่ว่าจะเป็น ZFNet VGG-Net GoogleNet หรือ ResNet

จึงเป็นที่มาของงานวิจัยในครั้งนี้ งานวิจัยนี้ต้องการพัฒนาเครื่องมือการอ่านค่าระดับน้ำด้วยการประยุกต์ใช้กล้องวงจรปิด (CCTV) ที่มีอยู่ โดยทำการวิเคราะห์ภาพถ่ายที่ได้รับโดยตรงจากกล้องวงจรปิด ซึ่งสามารถวิเคราะห์ได้ง่ายและสามารถตรวจสอบความถูกต้องกับภาพที่แสดงในเวลานั้นได้โดยไม่ต้องตรวจสอบด้วยวิธีอื่น และสามารถนำไปปรับใช้กับกล้องวงจรปิดมาตรฐานที่มีการใช้งานอยู่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.2 วัตถุประสงค์ของงานวิจัย

1) เพื่อนำเทคโนโลยีที่ทันสมัยที่มีอยู่ในปัจจุบันมาประยุกต์ใช้กับปัญหาด้านอัตรากำลังคนและความถูกต้องแม่นยำของการวัดค่าระดับน้ำ
2) เพื่อศึกษาการใช้การประมวลผลจากภาพและการเรียนรู้เชิงลึกในการวัดระดับน้ำจากภาพถ่ายที่ได้มาจากกล้องวงจรปิด (CCTV)

1.3 ขอบเขตของงานวิจัย

1) รูปภาพที่ใช้ในการวิจัย ได้แก่ รูปภาพจากกล้องวงจรปิด (CCTV) และภาพถ่ายที่ถ่ายให้เห็นไม้วัดระดับน้ำที่มีการใช้งานอยู่ในแต่ละพื้นที่ ซึ่งได้มีการสืบค้นข้อมูลรูปภาพจาก Google และชุดข้อมูลรูปภาพจากเว็บไซต์ https://universe.roboflow.com/
2) รูปภาพตัวเลขที่มีลักษณะคล้ายกับตัวเลขที่ปรากฏอยู่บนไม้วัดระดับน้ำ ซึ่งได้มีการสืบค้นข้อมูลรูปภาพจาก Google และชุดข้อมูลรูปภาพจากเว็บไซต์ https://universe.roboflow.com/
3) การใช้การตรวจจับวัตถุในการหาตำแหน่งของไม้วัดระดับน้ำ รวมถึงการใช้การตรวจจับวัตถุในการหาตำแหน่งและอ่านค่าตัวเลขที่อยู่บนไม้วัดระดับน้ำเพื่อให้ได้ค่าตัวเลขของระดับน้ำ

1.4 ประโยชน์ที่คาดว่าจะได้รับ

1) ผลการวิจัยในครั้งนี้ทำให้สามารถอ่านค่าระดับน้ำจากรูปภาพที่ได้มาจากกล้องวงจรปิด (CCTV)
2) ผลการวิจัยในครั้งทำให้เกิดทางเลือกเพิ่มขึ้นสำหรับการอ่านค่าระดับน้ำแบบอัตโนมัติ โดยการประยุกต์ใช้ผลการวิจัยกับอุปกรณ์ที่มีอยู่ และสามารถลดปัญหาในเรื่องของอัตรากำลังคนและความถูกต้องแม่นยำ

1.5 ขั้นตอนการดำเนินงาน

1) ค้นหารูปภาพของไม้วัดระดับน้ำ โดยรูปภาพจะต้องประกอบด้วยไม้วัดระดับน้ำรูปแบบต่าง ๆ และมีตัวเลขที่แสดงถึงค่าระดับน้ำสำหรับการนำไปใช้ในขั้นตอนการเรียนรู้ของตัวแบบ
2) ติดป้ายกำกับชุดข้อมูลที่มีอยู่ทั้งหมด เพื่อให้ตัวแบบสามารถตรวจจับวัตถุในภาพตามจุดประสงค์ของงานวิจัย โดยนำกรอบสี่เหลี่ยมมาครอบบริเวณที่เป็นไม้วัดระดับน้ำ เพื่อหาจุดของกรอบสี่เหลี่ยมที่จะนำไปใช้สำหรับการเรียนรู้ของตัวแบบ
3) นำรูปภาพที่ได้จากข้อที่ 2 มาใช้ในขั้นตอนการเรียนรู้ของตัวแบบ เพื่อให้ตัวแบบสามารถระบุตำแหน่งของวัตถุที่ต้องการได้
4) ค้นหารูปภาพของตัวเลขที่มีลักษณะคล้ายคลึงกับตัวเลขที่ปรากฏบนไม้วัดระดับน้ำ
5) ติดป้ายกำกับชุดข้อมูลรูปภาพที่ได้จากข้อที่ 4 เพื่อให้ตัวแบบสามารถอ่านค่าตัวเลขที่อยู่บนไม้วัดระดับน้ำ และระบุตำแหน่งของตัวเลขได้
6) นำรูปภาพที่ได้จากข้อที่ 2 และข้อที่ 4 มาสอนตัวแบบที่ 2 เพื่อให้ตัวแบบสามารถอ่านตัวเลขที่ปรากฏบนไม้วัดระดับน้ำและหาตำแหน่งของตัวเลขที่อยู่ติดกับระดับน้ำมากที่สุดได้
7) นำรูปภาพที่ได้จากข้อ 2 และตำแหน่งของตัวเลขที่ได้จากข้อที่ 6 มาคำนวณความใกล้ระหว่างตัวเลขกับขอบของรูปภาพด้านล่างเพื่อหาค่าทศนิยมของระดับน้ำให้ใกล้เคียงกับความเป็นจริงมากที่สุด โดยสามารถอ่านค่าระดับน้ำเป็นทศนิยม 2 ตำแหน่ง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

8) สร้างหน้าต่างสำหรับการใช้งานจริง ซึ่งหน้าต่างการแสดงผลประกอบด้วย กราฟแสดงข้อมูลระดับน้ำย้อนหลังรายชั่วโมง ภาพจากกล้องวงจรปิด ค่าระดับน้ำที่ได้จากตัวแบบ สถานที่ที่กล้องวงจรปิดติดตั้งอยู่ และสถานะการเปิดปิดของประตูระบายน้ำ ณ เวลานั้น ๆ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2
ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ในการดำเนินการวิจัยครั้งนี้ ผู้วิจัยได้ทำการศึกษาค้นคว้าเอกสารเกี่ยวกับทฤษฎีและงานวิจัยที่เกี่ยวข้องกับการตรวจจับวัตถุและการอ่านค่าระดับน้ำจากกล้องวงจรปิด (CCTV) ด้วยวิธีการเรียนรู้เชิงลึก ดังรายละเอียดต่อไปนี้

1. การประมวลผลภาพ
2. การอ่านข้อความจากรูปภาพ
3. ตัวแบบที่ผ่านการเรียนรู้มาแล้ว (YOLOv8)
4. การตรวจจับวัตถุแบบไร้ที่ยึดเหนี่ยว
5. การจัดแต่งรูปภาพ
6. ความแม่นยำ RF100
7. โครงข่ายประสาทเทียมแบบคอนโวลูชัน
8. มาตรวัดประสิทธิภาพของตัวแบบการตรวจจับวัตถุ
9. งานวิจัยที่เกี่ยวข้อง

2.1 การประมวลผลภาพ

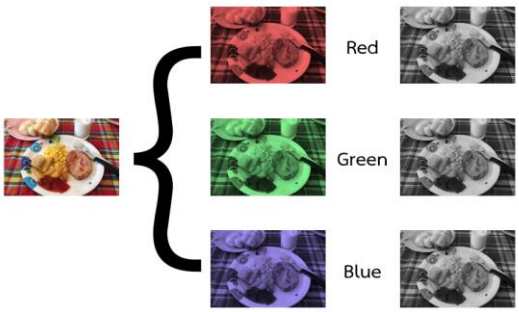
2.1.1 การไล่ระดับสีเทา

การไล่ระดับสีเทาเป็นกระบวนการในการแปลงรูปภาพจากภาพที่มีลักษณะของสีอื่น เช่น RGB, CMYK และ HSV ให้เป็นเฉดสีเทา ในตัวแบบ RGB นั้น ช่วงของทั้งสามองค์ประกอบ R, G และ B คือ [0, 225] จึงทำให้ภาพ 1 ภาพมีช่วงการเปลี่ยนสีถึง 16.58 ล้านรูปแบบ ภาพที่ผ่านการไล่ระดับสีเทาจะถูกแสดงด้วยค่าระดับสีเทาเพียงค่าเดียวเท่านั้น ดังนั้นภาพที่ผ่านการไล่ระดับสีเทาจะมีส่วนช่วยในการปรับปรุงประสิทธิภาพก่อนนำรูปภาพไปใช้กับขั้นตอนวิธีอื่น ๆ ในงานวิจัยครั้งนี้เป็นอย่างมาก

ความสำคัญของการไล่ระดับสีเทามีดังนี้

2.1.1.1 ลดมิติของข้อมูล

ในภาพ RGB จะมีช่องสีทั้งหมด 3 ช่องจึงทำให้ภาพมี 3 มิติ ในขณะที่ภาพระดับสีเทาจะมีเพียงมิติเดียวเท่านั้น ดังรูปที่ 2.1



รูปที่ 2.1 ความแตกต่างของภาพ RGB และภาพที่มีระดับสีเทา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.1.1.2 ลดความซับซ้อนของตัวแบบ

หากนำภาพ RGB ที่มีขนาด 10x10x3 พิกเซลมาใช้กับการเรียนรู้ของตัวแบบด้วยโครงข่ายประสาทเทียม จะทำให้มีข้อมูลในการป้อนทั้งหมด 300 โหนด กลับกัน หากใช้โครงข่ายประสาทเทียมกับภาพที่มีระดับสีเทา จะทำให้มีข้อมูลในการป้อนเพียง 100 โหนดเท่านั้น

2.1.1.3 เพื่อใช้ในการทำขั้นตอนวิธีอื่น

ขั้นตอนวิธีที่นำมาใช้ประมวลผลรูปภาพส่วนมาก ถูกสร้างให้ทำงานเฉพาะกับรูปภาพที่มีระดับสีเทาเท่านั้น เช่น ฟังก์ชันการตรวจจับมุมแคนนี่ (Canny Edge) ซึ่งเป็นไลบรารีของ OpenCV ที่รองรับเฉพาะรูปภาพที่มีระดับสีเทาเพียงลักษณะเดียว

ขั้นตอนวิธีการไล่ระดับสีเทาของไลบรารี OpenCV มีการเฉลี่ยถ่วงน้ำหนักขององค์ประกอบสี R, G, และ B ของภาพสีในการไล่ระดับสีเทา สูตรของค่าเฉลี่ยถ่วงน้ำหนักตั้งต้นของขั้นตอนวิธีนี้เป็นไปดังสมการที่ 2.1

$$Y = 0.2999\ R + 0.587\ G + 0.114\ B \tag{2.1}$$

2.1.2 การทำให้รูปภาพไม่ชัด

การทำให้รูปภาพไม่ชัด (Image Blurring) เป็นขั้นตอนก่อนการประมวลผลที่สำคัญที่สุดในการประมวลผลภาพและการมองเห็นของคอมพิวเตอร์ทั้งหมด การทำให้รูปภาพไม่ชัดสามารถเพิ่มประสิทธิภาพให้กับเทคนิคต่าง ๆ ได้ เช่น การทำให้รูปภาพไม่ชัดก่อนการตรวจจับขอบของวัตถุที่อยู่ในภาพสามารถลดจำนวนเนื้อหาที่มีความถี่สูง และลดรายละเอียดของภาพลงได้ การลดรายละเอียดของภาพลงจะทำให้สามารถหาวัตถุที่ต้องการได้ง่ายขึ้น และจะทำให้ตัวแบบสามารถโฟกัสไปที่วัตถุที่มีโครงสร้างขนาดใหญ่ในภาพได้อีกด้วย

การทำให้รูปภาพไม่ชัดแบบเกาส์เซียน (Gaussian Blur) ใช้ค่าเฉลี่ยถ่วงน้ำหนัก โดยพิกเซลที่อยู่ตรงกลางของรูปภาพจะมีน้ำหนักมากกว่าพิกเซลที่อยู่ตรงขอบ การทำให้รูปภาพไม่ชัดแบบเกาส์เซียนใช้วิธีลดสิ่งรบกวนด้วยวิธีการแจกแจงแบบเกาส์ (Gaussian Distribution) หรือเป็นการแจกแจงความน่าจะเป็นของค่าตัวแปรสุ่มที่เป็นค่าแบบต่อเนื่อง โดยที่ค่าของตัวแปรสุ่มมีแนวโน้มที่จะมีค่าอยู่ใกล้กับค่า ๆ หนึ่ง กราฟแสดงค่าฟังก์ชันความหนาแน่น (probability density function) จะเป็นรูปคล้ายระฆังคว่ำ หรือเรียกว่าสมการเกาส์เซียน (Guassian function) ดังสมการที่ 2.2

ผลลัพธ์ของวิธีนี้อาจจะไม่ทำให้รูปภาพไม่ชัดเท่าวิธีอื่น แต่ภาพจะไม่ชัดในลักษณะที่เป็นธรรมชาติมากกว่าวิธีอื่น การทำให้รูปภาพไม่ชัดแบบเกาส์เซียนใช้เคอร์เนลของ M X N โดยที่ทั้ง M และ N เป็นจำนวนเต็มคี่

อย่างไรก็ตาม เนื่องจากวิธีนี้เป็นการถ่วงน้ำหนักของพิกเซลตามระยะห่างจากพิกเซลที่อยู่ตรงกลาง เราจึงต้องใช้สมการเพื่อสร้างเคอร์เนลโดยสมการของฟังก์ชันเกาส์เซียนทิศทางเดียวเป็นไปดังสมการที่ 2.2

$$\mathrm{G(x)} = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{2.2}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดย x แทนตัวแปรสุ่ม พารามิเตอร์ μ แสดงค่ามัชฌิม และ σ^2 คือค่าความแปรปรวน (variance) ซึ่งเป็นค่าที่ใช้บอกปริมาณการกระจายของการแจกแจง การแจกแจงปรกติที่มีค่า $\mu = 0$ และ $\sigma^2 = 1$ จะถูกเรียกว่า การแจกแจงปกติมาตรฐาน

จากนั้น จึงขยายสมการนี้เป็นสองทิศทาง ทิศทางหนึ่งสำหรับแกน x และอีกทิศทางหนึ่งสำหรับแกน y ดังสมการที่ 2.3

$$G(x, y) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (2.3)$$

การแจกแจงปกติเป็นการแจกแจงที่เด่นที่สุดในทางวิชาความน่าจะเป็นและสถิติศาสตร์ ซึ่งหนึ่งในเหตุผลมาจากทฤษฎีจำกัดกลาง (central limit theorem) ที่กล่าวว่า ภายใต้สภาพทั่ว ๆ ไปแล้ว ค่าเฉลี่ยจากการสุ่มค่าของตัวแปรสุ่มอิสระจากการแจกแจงใด ๆ ที่มีค่าเฉลี่ยและค่าความแปรปรวนจำกัด ถ้าจำนวนการสุ่มนั้นใหญ่พอ ค่าเฉลี่ยนั้นจะมีการแจกแจงประมาณได้เป็นการแจกแจงปกติ

ลักษณะที่สำคัญของการแจกแจงปกติมีดังนี้

- f(x) > 0 ทุกค่าของ x
- f(x) จะลดลงเรื่อย ๆ หากค่า x ห่างจาก μ เพิ่มขึ้นเรื่อย ๆ
- f(x) สมมาตรที่ μ คือ f(μ + x) = f(μ - x) ทุกค่า x
- เมื่อ x = μ แล้ว f(x) จะมีค่าสูงสุด และ μ มีค่าเท่ากับมัธยฐานกับฐานนิยม
- ถ้า σ ลดลง ส่วนโค้งจะแคบลง
- พื้นที่ใต้ส่วนโค้งระหว่าง
  - $\mu - \sigma$ กับ $\mu + \sigma = 0.68$
  - $\mu - 2\sigma$ กับ $\mu + 2\sigma = 0.95$
  - $\mu - 3\sigma$ กับ $\mu + 3\sigma = 0.99$

2.1.3 การหาจุดเริ่มเปลี่ยน

การหาจุดเริ่มเปลี่ยนกับรูปภาพ (Image Thresholding) เป็นการแบ่งส่วนภาพอย่างง่าย และเป็นวิธีสร้างภาพทวิภาคจากภาพที่มีสีเทา โดยทั่วไปจะทำเพื่อแยกวัตถุ หรือพิกเซลที่เป็นจุดสนใจด้านหน้าออกจากพิกเซลที่เป็นพื้นหลังเพื่อช่วยในการประมวลผลภาพ

หลักการทำงานของการหาจุดเริ่มเปลี่ยนคือการใช้ค่าคงที่ค่าหนึ่งในการเปรียบเทียบค่าของพิกเซลในแต่ละพื้นที่ หากค่าของพิกเซลในพื้นที่นั้นมีค่าน้อยกว่าค่าคงที่ที่กำหนด จะเปลี่ยนค่าพิกเซลของพื้นที่นั้นให้เป็น 0 แต่หากค่าของพิกเซลในพื้นที่นั้นมีค่ามากกว่าค่าคงที่ที่กำหนด จะเปลี่ยนค่าพิกเซลของพื้นที่นั้นเป็น 255 สมการฟังชั่นก์ของการหาจุดเริ่มเปลี่ยนเป็นดังสมการที่ 2.4 และ 2.5

$$dst(x, y) = 0, if\ f(x, y) \geq src(x, y) \quad (2.4)$$

$$dst(x, y) = \max value, if\ f(x, y) < src(x, y) \quad (2.5)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดย f(x,y) คือ ภาพระดับเทาของตำแหน่ง x,y
src(x,y) คือ ค่าจุดเริ่มเปลี่ยน
dst(x,y) คือ ภาพขาวดำสองระดับที่ได้ของตำแหน่ง x,y

ลักษณะของภาพขาวดำสองระดับขึ้นอยู่กับค่าจุดเริ่มเปลี่ยน ดังนั้นหากค่าจุดเริ่มเปลี่ยนมีค่าสูง ภาพที่ได้จะเป็นสีดำมาก แต่หากค่าจุดเริ่มเปลี่ยนมีค่าต่ำ ภาพที่ได้จะเป็นสีขาวมาก ซึ่งการกำหนดค่าจุดเริ่มเปลี่ยนที่เหมาะสมจะช่วยให้ภาพขาวดำสองระดับมีความถูกต้องใกล้เคียงมากที่สุด โดยทั่วไปจะกำหนดค่าจุดเริ่มเปลี่ยนด้วยการเลือกค่าพิกเซล 2 วิธี ได้แก่

วิธีที่ 1 เลือกค่าพิกเซลสูงกว่าค่าจุดเริ่มเปลี่ยน ถ้าค่าพิกเซลมากกว่าจุดเริ่มเปลี่ยน (วัตถุมีความสว่างมากกว่าพื้นหลัง) พิกเซลภาพจะมีค่า 255 (สีขาว) และพิกเซลพื้นหลังมีค่า 0 (สีดำ) เรียกว่า พิกเซลสูงกว่าจุดเริ่มเปลี่ยน

วิธีที่ 2 เลือกค่าพิกเซลต่ำกว่าจุดเริ่มเปลี่ยน ถ้าค่าพิกเซลต่ำกว่าจุดเริ่มเปลี่ยน พิกเซลภาพจะมีค่า 0 (สีดำ) และพิกเซลพื้นหลังมีค่า 255 (สีขาว)

โดยทั่วไปแล้ว ค่าพิกเซลของวัตถุจะมีค่าเท่ากับ 1 (สีขาว) และค่าพิกเซลของพื้นหลังจะมีค่าเท่ากับ 0 (สีดำ) ภาพทวิภาคจะถูกสร้างขึ้นโดยพิกเซลสีขาวหรือสีดำ

การนำทฤษฎีการหาจุดเริ่มเปลี่ยนมาใช้จะเป็นการหาค่าความเข้มข้นให้ค่าที่มีความแตกต่างระหว่างตัววัตถุและพื้นหลัง เช่น กรณีของรูปภาพเอกสารที่มีการเขียนตัวอักษรเพื่อทำดัชนีจะมีความเข้มข้นของตัวอักษรเป็น 0 (สีดำ) และควาเข้มข้นของพื้นหลังเป้น 255 (สีขาว) เป็นต้น ซึ่งในทฤษฎีการหาจุดเริ่มเปลี่ยน ยังสามารถแบ่งการหาค่าได้โดยใช้ การหาจุดเริ่มเปลี่ยนแบบครอบคลุม (Global Thresholding) และการหาจุดเริ่มเปลี่ยนแบบปรับตัว (Adaptive Thresholding)

การหาจุดเริ่มเปลี่ยนแบบครอบคลุม เป็นการทำให้ค่าจุดเริ่มเปลี่ยนเป็นค่าคงที่หนึ่งค่า โดยใช้ฮิสโทแกรมจากรูปภาพที่นำมาทดสอบ ซึ่งค่าจุดเริ่มเปลี่ยนที่ใช้ควรเลือกจากค่าฮิสโทแกรมที่อยู่ที่จุดสูงสุด (Peaks)

การหาจุดเริ่มเปลี่ยนแบบปรับตัวเป็นการหาค่าจุดเริ่มเปลี่ยนที่มีได้หลายค่า โดยจะทำการหาค่าจุดเริ่มเปลี่ยนจากค่าความเข้มของสีในบริเวณกรอบสี่เหลี่ยมของภาพที่กำหนด โดยมีเงื่อนไขว่ากรอบสี่เหลี่ยมจะต้องมีขนาดที่เหมาะสม แต่ถ้าหากรูปภาพไม่มีลักษณะเป็นขอบสูงชันพอค่าจุดเริ่มเปลี่ยนจะมีผลกระทบต่อตำแหน่งขอบเขตและขนาดของวัตถุจริง ทำให้การวัดขนาดและการคำนวณหาพื้นที่คลาดเคลื่อนได้ มักนิยมนำมาใช้สำหรับการทำให้ภาพสีหรือภาพระดับสีเทาที่มีแสงสว่างไม่สม่ำเสมอ (Nonuniform Illumination) อยู่ในรูปแบบภาพขาวดำ (Binary Image) เทคนิควิธีการหาค่าจุดเริ่มเปลี่ยนแบบปรับตัวทำได้โดยการสร้างหน้าต่าง (Window) มาบวกกันแล้วหารด้วยจำนวนช่องทั่งหมดของหน้าต่างทั้งหมด จะได้ค่าจุดเริ่มเปลี่ยนที่อยู่ภายในหน้าต่างนั้น จากนั้นทำการหาค่าจุดเริ่มเปลี่ยนเช่นนี้ไปเรื่อย ๆ กับบริเวณที่ไม่ซ้ำกันจนกระทั่งได้มีการกำหนดค่าจุดเริ่มเปลี่ยนครบในทุก ๆ พิกเซล ถ้าค่าระดับสีเทาของพิกเซลนั้นมีค่ามากกว่าค่าจุดเริ่มเปลี่ยนของพิกเซลนั้นจะกำหนดให้เป็นสีขาว แต่ถ้าค่าระดับสีเทาของพิกเซลนั้นน้อยกว่าจุดเริ่มเปลี่ยนของพิกเซลนั้นจะกำหนดให้เป็นสีดำ ทำเช่นนี้จนครบทุกพิกเซลจะได้ผลลัพธ์เป็นภาพขาวดำ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2 การอ่านข้อความจากรูปภาพ

การอ่านข้อความจากรูปภาพ เป็นการเปลี่ยนจากรูปภาพ 2 มิติที่เป็นตัวอักษร ให้กลายเป็นข้อความที่พิมพ์ด้วยเครื่องหรือข้อความที่เขียนด้วยลายมือซึ่งสามารถอ่านโดยเครื่องได้ การอ่านข้อความจากรูปภาพเป็นกระบวนการที่ประกอบไปด้วยหลายกระบวนการย่อยเพื่อให้สามารถทำงานได้อย่างแม่นยำที่สุด โดยกระบวนย่อยได้แก่

- การประมวลผลภาพล่วงหน้า
- การแปลงข้อความให้เป็นภาษาถิ่น
- การแบ่งส่วนตัวหนังสือ
- การแบ่งแยกตัวอักขระ
- การประมวลผลย้อนหลัง

กระบวนการย่อยในรายการด้านบนอาจแตกต่างกัน แต่ขั้นตอนเหล่านี้เป็นขั้นตอนที่จำเป็นในการอ่านข้อความจากรูปภาพอัตโนมัติ ในซอฟต์แวร์ของการอ่านข้อความจากรูปภาพมีเป้าหมายหลักคือการระบุคำที่ไม่ซ้ำกันทั้งหมดโดยใช้ภาษาต่าง ๆ จากข้อความที่เขียนด้วยลายมือ

เป็นเวลาเกือบสองทศวรรษที่ระบบการรู้จำอักขระด้วแสงถูกนำมาใช้อย่างแพร่หลายเพื่อให้ป้อนข้อความอัตโนมัติลงในระบบคอมพิวเตอร์ แต่ตลอดเวลานี้ ระบบการอ่านข้อความจากรูปภาพออนไลน์ทั่วไป (เช่น การอ่านข้อความจากรูปภาพแบบโซน) ไม่สามารถอ่านแบบอักษรและรูปแบบหน้ากระดาษได้มากนัก การพิมพ์แบบเว้นระยะตามสัดส่วน (ซึ่งรวมถึงสำเนาเรียงพิมพ์แทบทั้งหมด) ฟอนต์เครื่องพิมพ์เลเซอร์ และแม้แต่ฟอนต์เครื่องพิมพ์ดีดที่ไม่ได้สัดส่วนจำนวนมาก ยังคงอยู่เกินความสามารถของระบบเหล่านี้ และด้วยเหตุนี้ การอ่านข้อความจากรูปภาพ แบบเดิมจึงไม่เคยประสบความสำเร็จในการทำงานอย่างสมบูรณ์ได้ในการแปลงเอกสารทั้งหมดให้อยู่ในรูปแบบดิจิทัล

เครื่องประมวลผลการอ่านข้อความจากรูปภาพรุ่นต่อไปจัดการกับปัญหาเหล่านี้ที่กล่าวถึงข้างต้นได้ดีมากโดยใช้การวิจัยล่าสุดในด้านการเรียนรู้เชิงลึก ด้วยการใช้ประโยชน์จากการผสมผสานระหว่างตัวแบบเชิงลึกและชุดข้อมูลขนาดใหญ่ที่เปิดเผยต่อสาธารณะ แบบจำลองจึงมีความแม่นยำล้ำสมัยในขอบเขตงานที่กำหนด ทุกวันนี้ ยังเป็นไปได้ที่จะสร้างข้อมูลสังเคราะห์ด้วยฟอนต์ต่าง ๆ โดยใช้โครงข่ายฝ่ายตรงข้ามเชิงสร้างสรรค์และแนวทางเชิงสร้างสรรค์อื่น ๆ อีกเล็กน้อย

การรู้จำอักขระด้วยแสงยังคงเป็นปัญหาที่ท้าทายเมื่อข้อความอยู่ในสภาพแวดล้อมที่ไม่มีข้อจำกัด เช่น ฉากธรรมชาติ เนื่องจากความผิดเพี้ยนทางเรขาคณิต พื้นหลังที่ซับซ้อน และแบบอักษรที่หลากหลาย เทคโนโลยีนี้ยังคงมีศักยภาพมหาศาลเนื่องจากกรณีการใช้งานที่หลากหลายของการอ่านข้อความจากรูปภาพ ที่ใช้การเรียนรู้เชิงลึก เช่น การอ่านทะเบียนรถยนต์ ใบแจ้งหนี้ดิจิทัล เมนูดิจิทัล และบัตรประจำตัวประชนดิจิทัล เป็นต้น

2.3 ตัวแบบที่ผ่านการเรียนรู้มาแล้ว

YOLO หรือ You Only Look Once คือตัวแบบการตรวจจับวัตถุที่มีความโดดเด่นเรื่องความเร็วและความถูกต้องในการตรวจจับวัตถุ หลักการทำงานของ YOLO คือการนำสี่เหลี่ยมมาล้อมรอบวัตถุที่อยู่ในรูปภาพ โดยหาจุดกึ่งกลางของแต่ละวัตถุ และทำนายว่าสิ่งที่อยู่ในกรอบสี่เหลี่ยมนั้นคืออะไร YOLO มีตัวแบบที่ได้ผ่านการเรียนรู้แล้ว 80 classes โดยตัวแบบจะสามารถบอกถึงความน่าจะเป็นในการทำนายวัตถุนั้น ๆ ได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ปัจจุบัน YOLO ได้พัฒนามาถึงรุ่นที่ 8 โดยทีม Ultralytics หรือทีมที่ทำงานวิจัยด้านปัญญาประดิษฐ์ ทางทีมมีเป้าหมายในการพัฒนาตัวแบบให้ตัวแบบสามารถเข้าถึงง่าย มีประสิทธิภาพ และง่ายต่อการนำไปใช้ YOLOv8 มีการพัฒนาตัวแบบเพิ่มขึ้นโดยมีรายละเอียด ดังนี้

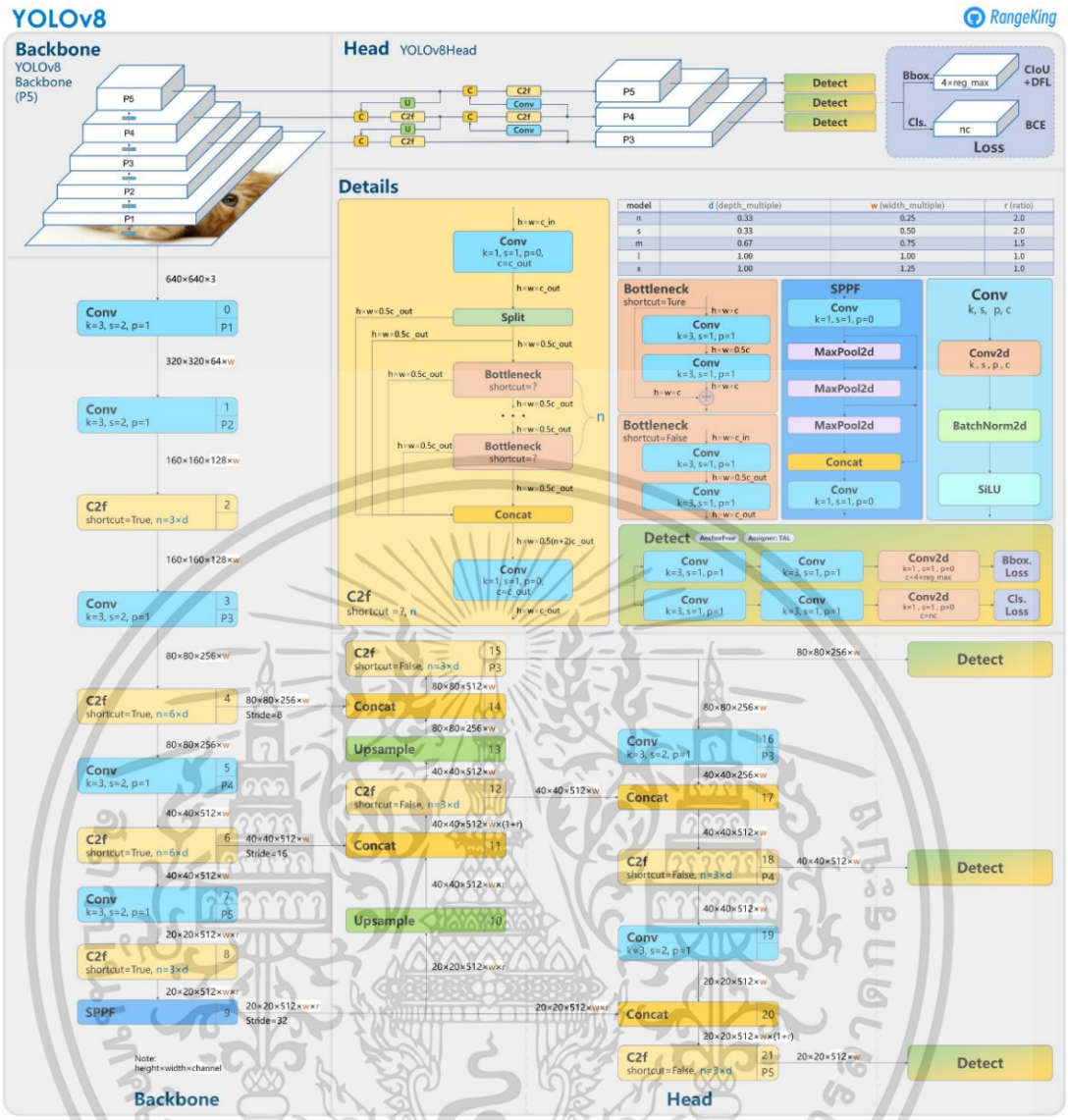
- การพัฒนาความแม่นยำในการตรวจจับวัตถุให้ดีขึ้นโดยผสมเทคนิคใหม่และการเพิ่มประสิทธิภาพให้กับตัวแบบ
- มีความเร็วในการคำนวณมากกว่าตัวแบบในการตรวจจับวัตถุรุ่นอื่น ๆ ในขณะที่ยังคงความแม่นยำที่ดี
- YOLOv8 รองรับแกนหลักชนิดต่าง ๆ เช่น EfficientNet, ResNet และ CSPDarknet ทำให้มีความยืดหยุ่นต่อผู้ใช้งานในการเลือกแกนหลักที่ดีที่สุดและเหมาะสมกับการใช้งานนั้น ๆ
- YOLOv8 ใช้การเรียนรู้แบบปรับตัว (Adaptive Learning) หรือการปรับให้ตัวแบบเหมาะสมกับระดับความสามารถของผู้เรียน เพื่อปรับอัตราการเรียนรู้ให้เหมาะสมและปรับสมดุลของฟังก์ชันการสูญเสียระหว่างการเรียนรู้ ซึ่งนำไปสู่ประสิทธิภาพของตัวแบบที่ดีขึ้น มีการใช้เทคนิคการจัดแต่งรูปภาพให้ได้รูปภาพใหม่ขั้นสูง เช่น MixUp และ CutMix เพื่อปรับปรุงความทนทานและลักษณะของตัวแบบ
- สามารถปรับแต่งสถาปัตยกรรมของตัวแบบได้อย่างยืดหยุ่น ช่วยให้ผู้ใช้สามารถปรับเปลี่ยนโครงสร้างและพารามิเตอร์ของตัวแบบให้เหมาะสมกับความต้องการได้อย่างง่ายดาย
- YOLOv8 ประกอบด้วยตัวแบบที่ผ่านการเรียนรู้ไว้ล่วงหน้า (Pretrained Model) เพื่อการใช้งานที่ง่ายดายและถ่ายโอนการเรียนรู้ในชุดข้อมูลที่หลากหลาย

สถาปัตยกรรมของ YOLOv8 ถูกพัฒนาขึ้นจากตัวแบบ YOLO ก่อนหน้า YOLOv8 ใช้โครงข่ายประสาทเทียมเต็มรูปแบบ (Convolutional Neural Network : CNN) ที่สามารถแบ่งออกเป็นสองส่วน ได้แก่ ส่วนแกนหลักและส่วนหัว

- ส่วนแกนหลักถูกพัฒนาต่อยอดจากสถาปัตยกรรมของ CSPDarknet53 สถาปัตยกรรมนี้ประกอบไปด้วยโครงข่ายประสาทเทียม 53 ชั้น และใช้เทคนิคที่เรียกว่าการเชื่อมต่อแบบไขว้กันบางส่วน (cross-stage partial connections) เพื่อปรับปรุงการไหลของข้อมูลระหว่างชั้นต่าง ๆ ของโครงข่าย
- ส่วนหัวของตัวแบบประกอบด้วยโครงข่ายประสาทเทียมหลายชั้นตามด้วยชั้นที่ต่อกันอย่างสมบูรณ์ ชั้นเหล่านี้มีหน้าที่ในการคาดเดาตำแหน่งของกล่องสี่เหลี่ยมล้อมรอบวัตถุ (Bounding Box), คะแนนความเป็นวัตถุ และ ค่าความเชื่อมั่นในการคาดเดาวัตถุ คุณสมบัติหลักของ YOLOv8 อีกอย่างหนึ่งคือการใช้กลไกการพึ่งพาตนเอง (self-attention) ในส่วนหัวของโครงข่าย กลไกนี้ ทำให้ตัวแบบสามารถให้ความสนใจไปที่ส่วนต่าง ๆ ของภาพและปรับความสำคัญของคุณสมบัติต่าง ๆ ตามความเกี่ยวข้องกับงาน

คุณสมบัติที่สำคัญของ YOLOv8 อีกอย่างหนึ่งคือความสามารถในการตรวจจับวัตถุที่มีหลากหลายขนาด ตัวแบบนี้ใช้เครือข่ายคุณลักษณะพีระมิด (Feature Pyramid Network) เพื่อตรวจจับวัตถุที่มีขนาดและมาตราส่วนที่ต่างกันภายในภาพ เครือข่ายคุณลักษณะพีระมิดประกอบไปด้วยชั้นหลายชั้นในการตรวจจับวัตถุที่มีขนาดต่างกัน ทำให้ตัวแบบสามารถตรวจจับวัตถุขนาดใหญ่และขนาดเล็กภายในภาพได้ ดังรูปที่ 2.2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.2 สถาปัตยกรรมของตัวแบบ YOLOv8
ที่มา: Solawetz at el. (2023)

รายละเอียดของตัวแบบ YOLOv8 รูปแบบต่าง ๆ ที่เปิดให้ใช้งานมีรายละเอียดดังตารางที่ 2.1 โดย Size คือ ขนาดของรูปภาพ

mAP^{val} คือ ค่ากลางของค่าเฉลี่ยความเที่ยง

Speed คือ ความเร็วในการประมวลผล

Params คือ จำนวนของพารามิเตอร์

FLOPs คือ ค่าที่ใช้ในการวัดประสิทธิภาพมาตรฐานที่ใช้โดยรายการซูเปอร์คอมพิวเตอร์ TOP500 โดยจะจัดอันดับคอมพิวเตอร์ตามการดำเนินการ 64 บิตต่อวินาทีโดยใช้ไลบรารี่ที่มีชื่อว่า High Performance LINPACK ซึ่งถูกพัฒนาโดยกลุ่มวิจัยของศาสตราจารย์ Jack Dongarra ที่ University of Tennessee, Knoxville โดยค่า FLOPs สามารถคำนวณได้โดยใช้สมการที่ 2.6

$$\text{Performance in GFlops} = A \times B \times C \times D \quad (2.6)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดย A คือ ความเร็วของ CPU ในหน่วย GHz
B คือ จำนวน CPU cores
C คือ ค่าคำสั่งของ CPU ในแต่ละรอบ
D คือ จำนวน CPU ต่อหนึ่งหน่วยประมวลผล

ตารางที่ 2.1 รายละเอียดประสิทธิภาพของทำงานของตัวแบบ YOLOv8 รูปแบบต่าง ๆ

Model	size (pixels)	mAP^{val} 50-95	Speed CPU ONNX (ms)	Speed A100 TensorRT (ms)	params (M)	FLOPs (B)
YOLOv8n	640	37.3	80.4	0.99	3.2	8.7
YOLOv8s	640	44.9	128.4	1.20	11.2	28.6
YOLOv8m	640	50.2	234.7	1.83	25.9	78.9
YOLOv8l	640	52.9	375.2	2.39	43.7	165.2
YOLOv8x	640	53.9	479.1	3.53	68.2	257.8

YOLOv8 ได้มีการเพิ่มตัวแบบที่ผ่านการเรียนรู้แล้ว (pre-trained model) เพื่อเป็นประโยชน์ต่อการประมวลผลภาพรูปแบบต่าง ๆ โดยการประมวลผลภาพที่เพิ่มขึ้นมาในครั้งนี้ ได้แก่

1. การแบ่งแยกภาพออกเป็นส่วนย่อย (Instance Segmentation) เป็นการตรวจจับวัตถุภายในภาพแต่จะแตกต่างจากการตรวจจับวัตถุโดยที่การตรวจจับวัตถุในการเรียนรู้จะไม่เป็นสี่เหลี่ยม แต่จะเป็นจุดล้อมรอบวัตถุนั้น ๆ โดยตัวแบบนี้ผ่านการเรียนรู้โดยใช้ชุดข้อมูลรูปภาพจากชุดข้อมูล COCO โดยมีรายละเอียดของรูปภาพที่ 640 พิกเซล
2. การจำแนกประเภทข้อมูลภาพ (Image Classification) เป็นการมองรูปภาพโดยรวมแล้วสอนให้ตัวแบบเรียนรู้ว่ารูปภาพที่มีอยู่ต้องการจะสื่ออะไร เช่น รูปภาพของทุ่งข้าว หากรูปภาพประกอบด้วยต้นข้าวที่มีสีเขียว อาจหมายถึง ต้นข้าวยังไม่พร้อมสำหรับการเก็บเกี่ยว หากรูปภาพของทุ่งข้าวเปลี่ยนเป็นสีเหลืองแล้ว อาจหมายถึง ต้นข้าวที่อยู่ในรูปภาพนั้นพร้อมสำหรับการเก็บเกี่ยวแล้ว โดยตัวแบบนี้ผ่านการเรียนรู้โดยใช้ชุดข้อมูลรูปภาพจากชุดข้อมูล ImageNet โดยมีรายละเอียดของรูปภาพที่ 224 พิกเซล
3. การตรวจจับวัตถุ (Object Detection) เป็นการระบุว่ารูปภาพประกอบไปด้วยวัตถุอะไรบ้าง และตำแหน่งของวัตถุในรูปภาพอยู่ที่ใด โดยผลลัพธ์ของตัวแบบนี้จะเป็นตำแหน่งของจุด 4 จุด ได้แก่ x1, x2, y1 และ y2 หรือจุดเริ่มต้นและสิ้นสุดของแกน x และแกน y ที่ล้อมรอบวัตถุแต่ละวัตถุเอาไว้ โดยตัวแบบนี้ผ่านการเรียนรู้โดยใช้ชุดข้อมูลรูปภาพจากชุดข้อมูล COCO โดยมีรายละเอียดของรูปภาพที่ 640 พิกเซล

2.4 การตรวจจับวัตถุแบบไร้ที่ยึดเหนี่ยว

การตรวจจับวัตถุแบบไร้ที่ยึดเหนี่ยว (Fully Convolutional One-stage Object Detection : FCOS) เป็นเครื่องมือตรวจจับวัตถุแบบไร้ที่ยึดเหนี่ยว วิธีนี้ช่วยแก้ปัญหาการตรวจจับ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

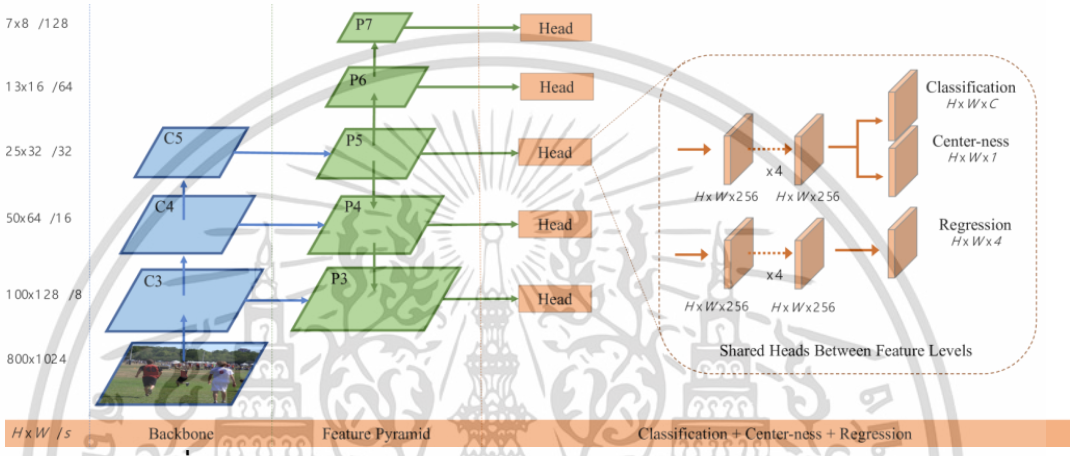
วัตถุในรูปแบบการคาดคะเนต่อพิกเซลคล้ายกับการแบ่งแยกภาพออกเป็นส่วนย่อย การศึกษาเกี่ยวกับการตรวจจับวัตถุแบบไร้ที่ยึดเหนี่ยวส่วนใหญ่ใช้ FCOS เป็นพื้นฐานของโครงสร้างสถาปัตยกรรม

FCOS ได้รับแรงบันดาลใจจาก FCN หรือ Fully Convolutional Networks สำหรับการแบ่งวัตถุในภาพออกเป็นส่วนย่อย โดยใช้การคาดคะเนพิกเซลที่คล้ายคลึงกันสำหรับการตรวจจับวัตถุ โดยสถาปัตยกรรมของ FCOS ประกอบไปด้วย 3 ส่วน ดังรูปที่ 2.3 ดังนี้

2.4.1 ส่วนหลัง

2.4.2 คุณลักษระพีระมิด

2.4.3 ส่วนหัว



รูปที่ 2.3 สถาปัตยกรรมของตัวแบบ Fully Convolutional One-stage

ที่มา: Lang at el. (2019)

ส่วนหัวของโครงข่ายสามารถแบ่งออกได้ 2 กรณี ดังนี้

2.4.3.1 การทำนายระดับความเชื่อมั่นและอีกส่วนทำนายความเป็นศูนย์กลางของวัตถุในการจำแนกประเภท

2.4.3.2 การทำนายขอบเขตโดยการใช้สมการถดถอย

รูปร่างผลลัพธ์ของคุณลักษณะที่แสดงในรูปที่ 2.3 ใช้สำหรับรูปภาพนำเข้าที่มีขนาด 800 x 1024 พิกเซล ระยะส่งออก (อัตราสุ่มตัวอย่างต่อขนาดภาพนำเข้า) ที่ C3/P3, C4/P4, C5/P5, P6 และ P7 คือ 8, 16, 32, 64 และ 128 ตามละดับ เมื่อระยะเพิ่มขึ้น ความละเอียดของรูปภาพที่ส่งออกมาจะน้อยลง ตัวอย่างเช่น ความกว้างและความสูงของภาพส่งออกที่ระดับ C3/P3 คือ 100 x 128 และระดับ C4/P4 คือ 50 x 64 อย่างไรก็ตาม ในทุกระดับจะมีความแตกต่างกัน ดังนี้

- ความลึกของการจำแนกประเภทคือ C โดย C คือจำนวนของประเภทวัตถุทั้งหมด
- ความลึกของการหาศูนย์กลางคือ 1 เพราะการทำนายศูนย์กลางของขอบเขตนั้น ใช้เพียงหนึ่งมาตราส่วนก็เพียงพอต่อการแสดงถึงคุณภาพที่ดีในการหาศูนย์กลาง
- ความลึกของการถดถอยคือ 4 เนื่องจากต้องใช้ตัวเลขสี่ตัวในการแสดงกรอบขอบเขตในระนาบ 2 มิติ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.5 การจัดแต่งรูปภาพ

การจัดแต่งรูปภาพ (Image Augmentation) คือการนำข้อมูลรูปภาพที่มีอยู่มาสร้างภาพขึ้นมาใหม่เพื่อเพิ่มจำนวนข้อมูลในขั้นตอนของการเรียนรู้ รูปภาพที่ได้มาใหม่จะถูกสร้างขึ้นจากรูปภาพที่มีอยู่แล้ว เราจึงไม่ต้องหารูปภาพเหล่านี้เอง จึงเป็นวิธีที่ดีเมื่อเราต้องการข้อมูลในการเรียนรู้เพิ่มขึ้น รูปภาพที่ได้จากการจัดแต่งภาพจะมีการติดป้ายกำกับเช่นเดียวกับรูปภาพตั้งต้นเสมอ การจัดแต่งรูปภาพที่มักถูกนำมาใช้มีดังนี้

2.5.1 การหมุนรูปภาพ

การหมุนรูปภาพ (Image Rotation) เป็นวิธีที่นิยมวิธีหนึ่ง ไม่ว่าเราจะหมุนภาพอย่างไร ข้อมูลที่ได้จากรูปภาพก็ยังคงเหมือนเดิมเสมอ วิธีนี้มันถูกนำมาใช้เพื่อเพิ่มจำนวนข้อมูลในการเรียนรู้ของตัวแบบ โดยนำรูปภาพตั้งต้นมาหมุนให้มีองศาที่แตกต่างกันออกไป

2.5.2 การเลื่อนรูปภาพ

การเลื่อนรูปภาพ (Image Shifting) ทำให้เราสามารถย้ายตำแหน่งของวัตถุที่อยู่ภายในรูปภาพได้ การเลื่อนรูปภาพทำให้รูปภาพที่เรามีอยู่มีความหลากหลายมากยิ่งขึ้น และสามารถทำให้ตัวแบบของเรามีประสิทธิภาพมากขึ้น การเลื่อนรูปภาพเป็นการย้ายตำแหน่งของทุกวัตถุในภาพจากตำแหน่งหนึ่งไปยังตำแหน่งหนึ่งของรูปภาพใหม่ หากวัตถุอยู่ในตำแหน่ง x, y ในรูปภาพตั้งต้น วัตถุนี้จะถูกย้ายไปที่ตำแหน่ง X, Y ในรูปภาพใหม่ ดังสมการต่อไปนี้

$$X = x + dx \tag{2.7}$$

$$Y = y + dy \tag{2.8}$$

โดย dx และ dy คือการเปลี่ยนแปลงพร้อมกับทิศทางที่วัตถุเลื่อนไป

2.5.3 การพลิกรูปภาพ

การพลิกรูปภาพ (Image Flipping) เป็นส่วนที่เพิ่มเติมขึ้นมาของการหมุนรูปภาพ วิธีนี้ทำให้เราสามารถพลิกรูปภาพไปทางซ้าย ขวา บน และ ล่างได้

2.5.4 การเพิ่มความหยาบของรูปภาพ

วิธีนี้คือการเพิ่มสัญญาณรบกวนให้กับรูปภาพ (Image Noising) โดยมักนำไปใช้กับตัวแบบที่เรียนรู้การแยกจุดรบกวนในรูปภาพออกจากรูปภาพ วิธีนี้จะทำให้ตัวแบบของเรามีความยืดหยุ่นมากขึ้นเมื่อเกิดการเปลี่ยนแปลงภายในภาพ

2.5.5 การจัดแต่งรูปภาพแบบโมเสค

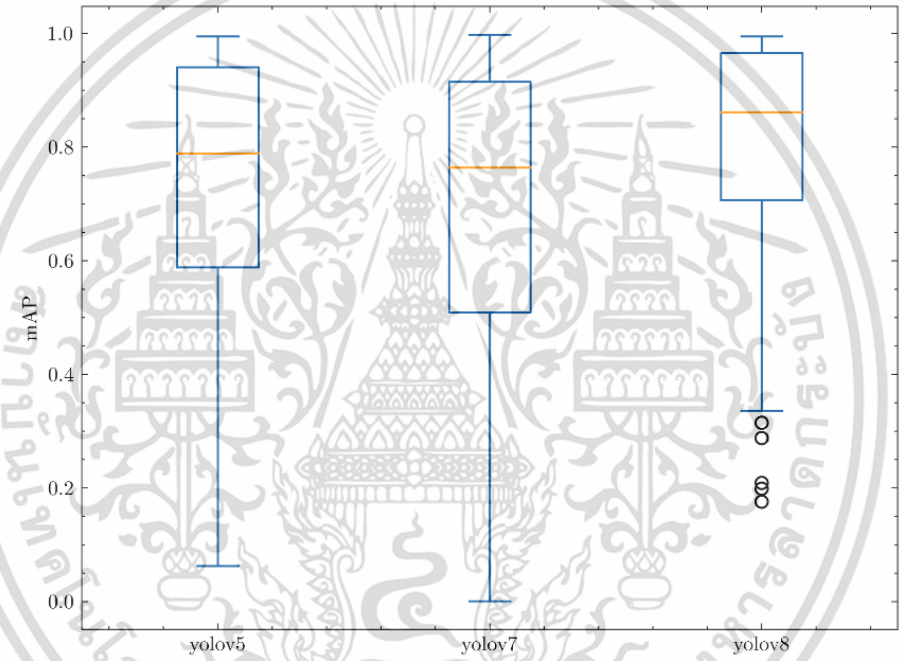
การจัดแต่งรูปภาพแบบโมเสค (Mosaic Augmentation) ถูกนำมาใช้ครั้งแรกใน YOLOv4 และถูกพัฒนามาจากการจัดแต่งรูปภาพแบบ CutMix การจัดแต่งภาพด้วยวิธีนี้คือการนำรูปภาพ 4 รูปมาต่อกันให้เป็นรูปเดียว โดยจะทำการปรับขนาดของรูปภาพ 4 รูป และนำรูปภาพที่ได้มาต่อกัน จากนั้นจะสุ่มตัดส่วนใดส่วนหนึ่งของรูปภาพที่ได้ออกมาและนำไปสู่รูปภาพแบบโมเสค

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.6 ความแม่นยำ RF100

ทีม Roboflow คือทีมที่ทำการค้นคว้าเกี่ยวกับปัญญาประดิษฐ์ โดยทีมได้รวบรวมข้อมูลรูปภาพไว้มากกว่า 10,000 ชุดข้อมูลบนแพลทฟอร์มที่เรียกว่า Roboflow Universe โดยคัดเลือก 100 ชุดข้อมูลตัวอย่างที่ไว้ใช้สำหรับการทดสอบตัวแบบต่าง ๆ โดยเรียกข้อมูล 100 ชุดนี้ว่า RF100 เกณฑ์ที่ใช้ในการรวบรวมข้อมูลได้รับการสนับสนุนจากบริษัท Intel และเกณฑ์นี้ถูกนำมาใช้เพื่อสามารถตอบได้ว่า ตัวแบบจะทำงานได้ดีเพียงใดหากนำมาใช้กับชุดข้อมูลที่กำหนดขึ้นมาเอง

YOLOv8 ได้ถูกนำมาประเมินโดยใช้ RF100 พร้อมกับ YOLOv5 และ YOLOv7 โดย mAP ที่ 0.5 ของทั้งสามตัวแบบเป็นไปตามรูปที่ 2.3 โดยทางทีมได้ทำการทดสอบตัวแบบ 100 ครั้ง แผนภาพรูปกล่อง (Box Plot) ในรูปที่ 2.4 สามารถบอกได้ว่า YOLOv8 มีค่านอกเกณฑ์ที่น้อยกว่า YOLOv5 และ YOLOv7 และมีค่าเฉลี่ยความแม่นยำมากกว่าเมื่อวัดด้วยเกณฑ์มารตฐานของ RF100



รูปที่ 2.4 แผนภาพรูปกล่องของการประเมินค่า mAP ของทั้ง 3 ตัวแบบ
ที่มา: Solawetz at el. (2023)

2.7 โครงข่ายประสาทเทียมแบบคอนโวลูชัน

โครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional Layer: CNN) เป็นโครงข่ายประสาทเทียมหนึ่งในกลุ่ม bio-inspired โดยที่ CNN จะเชี่ยวชาญในการประมวลผลข้อมูลที่มีลักษณะเป็นตาราง เช่น รูปภาพ ภาพดิจิทัลที่มีลักษณะเป็นทวิภาค โดยจะประกอบด้วยชุดพิกเซลที่ถูกจัดเรียงในลักษณะที่คล้ายกับตาราง ซึ่งแต่ละพิกเซลจะมีค่าค่าหนึ่งซึ่งแสดงถึงความสว่างและสีของพิกเซลนั้น ๆ

สมองของมนุษย์ประมวลผลข้อมูลจำนวนมหาศาลในวินาทีที่เราเห็นรูปภาพรูปหนึ่ง เซลล์ประสาทแต่ละเซลล์ทำงานในการรับสัญญาณของตัวเองและเชื่อมต่อกับเซลล์ประสาทอื่น ๆ ในลักษณะที่ทำให้เราสามารถมองเห็นทุกอย่างเป็นภาพเดียวกันได้ เช่นเดียวกับเซลล์ประสาทของเราที่ทำงานของตัวเองเพื่อตอบสนองต่อสิ่งเร้าทางชีวภาพ เซลล์ประสาทของ CNN ก็ประมวลผลข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เฉพาะในส่วนที่ตนเองรับผิดชอบเช่นกัน ชั้นต่าง ๆ ของ CNN จะถูกจัดเรียงให้สามารถตรวจจับรูปแบบที่เรียบง่ายได้ก่อน เช่น เส้น เส้นโค้ง เป็นต้น และรูปแบบที่ซับซ้อนมาขึ้น เช่น ใบหน้าและวัตถุจะจัดเรียงในส่วนถัด ๆ ไป เมื่อเราใช้งาน CNN เราสามารถทำให้คอมพิวเตอร์สามารถมองเห็นเช่นเดียวกับมนุษย์ได้

โดยทั่วไปโครงข่ายประสาทเทียมแบบคอนโวลูชันจะประกอบไปด้วย 3 ชั้นได้แก่ ชั้นคอนโวลูชัน (Convolutional Layer), ชั้นพูลลิง (Pooling Layer) และ ชั้นที่เชื่อมต่อกันอย่างสมบูรณ์ (Fully Connected Layer)

2.7.1 ชั้นคอนโวลูชัน (Convolutional Layer)

ชั้นนี้เป็นองค์ประกอบหลักในการทำงานของ CNN หรือเป็นส่วนหลักในการคำนวณของโครงข่ายทั้งหมด ชั้นนี้ทำการคูณสองเมทริกซ์เข้าด้วยกัน โดยที่เมทริกซ์แรกคือชุดของพารามิเตอร์ที่เรียนรู้ได้ หรือมีชื่อว่าเคอร์เนล (kernel) และเมทริกซ์ที่สองเป็นส่วนจำกัดของการรับสัญญาณ เคอร์เนลมีขนาดเล็กกว่าในเชิงพื้นที่แต่มีความลึกมากกว่า ซึ่งหมายความว่า หากรูปภาพประกอบด้วยช่องสัญญาณสามช่อง (RGB) ความสูงและความกว้างของเคอร์เนลจะมีขนาดเล็กกว่าในเชิงพื้นที่แต่ความลึกจะขยายออกไปถึงสามช่องสัญญาณ

ระหว่างการส่งต่อข้อมูล เคอร์เนลจะเลื่อนไปตามความสูงและความกว้างของรูปภาพ ซึ่งทำให้เกิดการแสดงรูปภาพของพื้นที่นั้น การเลื่อนไปตามความสูงและความกว้างของรูปภาพทำให้เกิดรูปภาพสองมิติที่เรียกว่า Activation Map ที่ให้การตอบสนองของเคอร์เนลในแต่ละตำแหน่งของรูปภาพ การเลื่อนของเคอร์เนลเรียกว่าสไตรด์ (stride)

หากรูปภาพนำเข้ามีขนาด W x W x D และเคอร์เนลมีขนาด F สไตรด์มีขนาด S และจำนวนการเสริมเต็ม (padding) คือ P ขนาดรูปภาพของผลลัพธ์สามารถคำนวณได้จากสมการที่ 2.9

$$W_{out} = \frac{W-F+2P}{S} + 1 \qquad (2.9)$$

การคิดค้นคอนโวลูชันเกิดจากแนวคิดสำคัญสามประการ โดยแนวคิดสามประการนี้เป็นแรงผลักดันให้นักวิจัยศึกษาเกี่ยวกับการมองเห็นด้วยคอมพิวเตอร์ แนวคิดสามประการมีรายละเอียด ดังต่อไปนี้

ชั้นของโครงข่ายประสาทเทียมขนาดเล็กใช้การคูณเมทริกซ์ด้วยเมทริกซ์ของพารามิเตอร์ที่อธิบายการโต้ตอบของข้อมูลนำเข้าและข้อมูลที่ส่งออก ซึ่งหมายความว่าทุกหน่วยของข้อมูลส่งออกได้มีการโต้ตอบกับข้อมูลนำเข้า ในขณะที่โครงข่ายประสาทเทียบแบบคอนโวลูชันมีการปฏิสัมพันธ์กันแบบกระจัดกระจาย (Sparse Interaction) โดยการทำให้เคอร์เนลมีขนาดเล็กกว่าข้อมูลนำเข้า เช่น รูปภาพสามารถมีพิกเซลถึงหนึ่งล้านหรือหนึ่งพันพิกเซลได้ แต่ในขณะที่ประมวลผลโดยใช้เคอร์เนล เราสามารถตรวจจับข้อมูลที่มีขนาดสิบหรือหนึ่งร้อยพิกเซลได้ ซึ่งหมายความว่าเราสามารถจัดเก็บพารามิเตอร์ได้น้อยลง วิธีนี้ไม่เพียงแต่ลดขนาดของหน่วยความจำของตัวแบบ แต่ยังปรับปรุงประสิทธิภาพทางสถิติของตัวแบบให้ดีขึ้นได้อีกด้วย

หากการคำนวณคุณลักษณะหนึ่งที่จุด (x1, y1) มีประโยชน์ การคำนวนคุณลักษณะที่จุด (x2, y2) ก็ควรมีประโยชน์ด้วยเช่นกัน ดังนั้น สำหรับชิ้นส่วนสองมิติ เซลล์ประสาทควรถูกจำกัดให้ใช้น้ำหนักชุดเดียวกัน ในโครงข่ายประสาทเทียบแบบดั้งเดิม แต่ละองค์ประกอบน้ำหนักของเมทริกซ์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จะถูกนำมาใช้เพียงครั้งเดียวและไม่นำกลับมาใช้อีก ในขณะที่โครงข่ายประสาทเทียมแบบคอนโวลูชันมีพารามิเตอร์ที่ใช้ร่วมกัน เช่น ในการนำไปสู่ผลลัพธ์ น้ำหนักที่ใช้กับข้อมูลนำเข้าจะเหมือนกับน้ำหนักที่นำไปใช้กับส่วนอื่น

เนื่องจากมีการแบ่งปันพารามิเตอร์ ชั้นของโครงข่ายประสาทเทียมคอนโวลูชันจะมีคุณสมบัติเทียบเท่ากับการแปลงค่า หากเราเปลี่ยนข้อมูลในทางใดทางหนึ่ง ข้อมูลส่งออกก็จะถูกเปลี่ยนไปในลักษณะเดียวกัน

2.7.2 ชั้นพูลลิง

ชั้นในส่วนนี้จะทำการแทนที่ข้อมูลส่งออกของโครงข่ายในบางตำแหน่งโดยรับสถิติสรุปของข้อมูลนำออกในบริเวณที่อยู่ใกล้เคียง วิธีนี้เป็นการช่วยลดขนาดเชิงพื้นที่โดยทำหน้าที่เป็นตัวแทนของพื้นที่ทั้งหมดแทน ซึ่งจะสามารถลดปริมาณการคำนวณและน้ำหนักที่ต้องการได้ การดำเนินการในส่วนนี้จะถูกประมวลผลในทุก ๆ ส่วนของการแทนค่าแต่ละตัว

ฟังก์ชันพูลลิงมีหลายหลายฟังก์ชัน ไม่ว่าจะเป็น ค่าเฉลี่ยสี่เหลี่ยมของพื้นที่ใกล้เคียง นอร์ม L2 สี่เหลี่ยมของพื้นที่ใกล้เคียง หรือ ค่าเฉลี่ยถ่วงน้ำหนักตามระยะห่างจากพิกเซลที่อยู่ ณ จุดกึ่งกลาง อย่างไรก็ตาม ฟังก์ชันที่ได้รับความนิยมที่สุดคือพูลลิงที่มีค่ามากที่สุด (Max Pooling) ซึ่งเป็นการรายงานข้อมูลส่งออกที่สูงที่สุดของพื้นที่ใกล้เคียง

หากรูปภาพนำเข้ามีขนาด W x W x D และเคอร์เนลพูลลิงมีขนาด F และสไตรด์มีขนาด S ขนาดรูปภาพของผลลัพธ์สามารถคำนวณได้จากสมการที่ 2.10

$$W_{out} = \frac{W-F}{S} + 1 \tag{2.10}$$

โดยผลลัพธ์ที่ได้คือข้ามูลนำออกขนาด W_{out} x W_{out} x D ในทุกกรณี พูลลิงจะให้ค่าคงที่ซึ่งจะทำให้วัตถุถูกจดจำได้โดยไม่คำนึงว่าวัตถุนั้นจะปรากฏที่ตำแหน่งใดในเฟรม

2.7.3 ชั้นที่เชื่อมกันอย่างสมบูรณ์

เซลล์ประสาทในชั้นนี้มีการเชื่อมต่อกันแบบเต็มรูปแบบกับเซลล์ประสาททั้งหมดในชั้นก่อนหน้าและชั้นถัดไป ดังที่เราเห็นปกติใน FCNN นี่คือเหตุผลที่ไม่ว่าข้อมูลจะมีความไม่เท่าเทียมแค่ไหน ก็สามารถคำนวณได้ตามปกติโดยใช้การคูณเมทริกซ์ ชั้น FC ช่วยในการเชื่อมระหว่างข้อมูลนำเข้าและข้อมูลส่งออก

เนื่องจากคอนโวลูชันเป็นการดำเนินการเชิงเส้นและนิยามของข้อมูลรูปภาพไกลจากการเป็นเชิงเส้นเป็นอย่างมาก ชั้นที่ไม่ใช่เชิงเส้นจึงมักจะถูกวางหลังชั้นแบบคอนโวลูชันเพื่อทำให้ Activation Map รู้จับกับการดำเนินการเชิงเส้น การดำเนินการแบบไม่ใช่เชิงเส้นมีหลายประเภทและประเภทที่ได้รับความนิยม ได้แก่

2.7.3.1 ซิกมอยด์ (Sigmoid)

ซิกมอยด์ที่ไม่ใช่เชิงเส้นมีรูปแบบทางคณิตศาสตร์ดังสมการที่ 2.11

$$\sigma(\kappa) = 1/(1 + e^{-}\kappa) \tag{2.11}$$

วิธีนี้จะนำตัวเลขที่เป็นค่าจริงมาบีบให้อยู่ในช่วงระหว่าง 0 ถึง 1 ข้อเสียอย่างหนึ่งของซิกมอยด์คือเมื่อการกระตุ้นอยู่ที่ปลายฝั่งใดฝั่งนึง การไล่ระดับจะเกือบกลายเป็น 0 โดยทันที

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หากการไล่ระดับในพื้นที่ใกล้เคียงมีขนาดเล็กมาก การแพร่กระจายแบบกลับหลัง (back propagation) จะทำการฆ่าการไล่ระดับนั้นทิ้งอย่างมีประสิทธิภาพ หากข้อมูลที่เข้ามาในเซลล์ประสาทมีค่าเป็นบวก ผลลัพธ์ที่ได้จากซิกมอยด์จะเป็นค่าบวกทั้งหมดหรือลบทั้งหมด ส่งผลให้เกิดการอัปเดตการไล่ระดับแบบซิกแซกของค่าน้ำหนักที่มีอยู่

2.7.3.2 แทน (Tanh)

ค่านี้จะบีบตัวเลขจำนวนจริงให้อยู่ในช่วง [-1,1] การกระตุ้นทำงานเหมือนซิกมอยด์ แต่ผลลัพธ์จะอิงค่า 0 ให้อยู่ตรงกลาง ไม่เหมือนกับเซลล์ประสาทของซิกมอยด์

2.7.3.3 ReLU

Rectified Linear Unit หรือ ReLU ได้รับความนิยมอย่างมากในช่วงไม่กี่ปีที่ผ่านมา และคำนวณโดยใช้สมการที่ 2.12

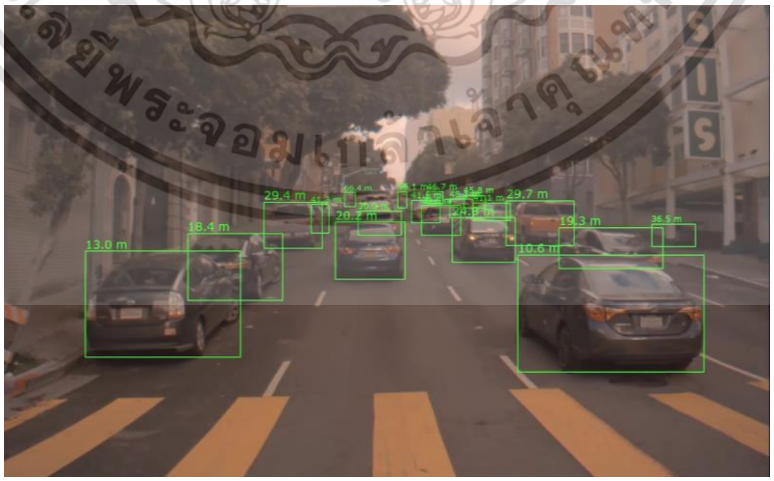
$$f(\kappa) = \max(0, \kappa) \tag{2.12}$$

หรือสามารถเรียกได้ว่าการกระตุ้นโดยใช้ 0 เป็นเกณฑ์ เมื่อเปรียบเทียบกับ sigmoid และ tan แล้ว ReLU มีความน่าเชื่อถือกว่า และเร่งการบรรจบกันให้เร็วกว่าได้ถึงหกเท่า แต่ข้อเสียของ ReLU คือ ReLU อาจเกิดการผิดพลาดระหว่างการเรียนรู้ได้ เนื่องจากการไล่ระดับขนาดใหญ่ที่ไหลผ่านถูกอัปเดตโดยที่เซลล์ประสาทไม่ได้รับการอัปเดตเพิ่มเติม อย่างไรก็ตาม เราสามารถแก้ไขปัญหานี้โดยกำหนดอัตราการเรียนรู้ให้เหมาะสมกับการทำงาน

2.7.4 ตัวอย่างงานที่มีการนำโครงข่ายปราสาทเทียมคอนโวลูชันไปใช้

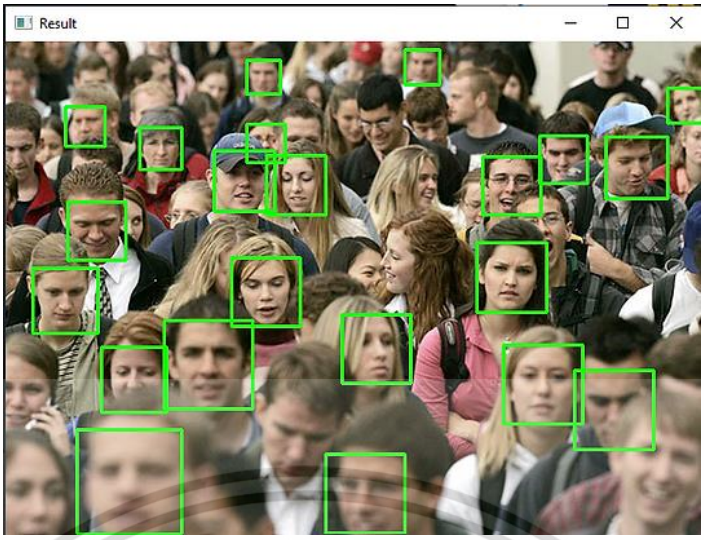
2.7.4.1 การตรวจจับวัตถุ

โครงข่ายประสาทเทียมแบบคอนโวลูชันทำให้ปัจจุบันมีตัวแบบที่ซับซ้อนมากขึ้น เช่น R-CNN, Fast R-CNN และ Faster R-CNN ซึ่งเป็นส่วนหลักที่นำไปใช้ในตัวแบบการตรวจจับวัตถุจำนวนมากที่ถูกนำไปใช้งานในระบบอัตโนมัติของยานพาหนะดังรูปที่ 2.4 และการตรวจจับใบหน้าดังรูปที่ 2.5 เป็นต้น



รูปที่ 2.5 การตรวจจับวัตถุในระบบอัตโนมัติของยานพาหนะ
ที่มา: Ritika (2022)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.6 การตรวจจับใบหน้า
ที่มา: Ultimate Python (2021)

2.7.4.2 การแบ่งส่วนวัตถุในรูปภาพ

ในปี พ.ศ. 2558 กลุ่มนักวิจัยจากฮ่องกงได้พัฒนา Deep Parsing Network ที่ใช้ CNN เพื่อรวมข้อมูลที่สมบูร์เข้ากับตัวแบบการแบ่งส่วนภาพ นักวิจัยจาก UC Berkeley ได้สร้างโครงข่ายคอนโวลูชันซึ่งได้รับการพัฒนาตามการแบ่งกลุ่มรูปภาพที่ทันสมัย

2.7.4.3 คำบรรยายภาพ

CNN ถูกนำมาใช้กับโครงข่ายประสาทแบบที่เกิดซ้ำเพื่อเขียนคำบรรณยายให้กับรูปภาพและวิดีโอ วิธีนี้สามารถนำไปใช้กับการทำงานหลากหลายรูปแบบ เช่น การตรวจจับกิจกรรมหรือการบรรยายวิดีโอและรูปภาพสำหรับผู้พิการทางสายตา YouTube ได้มีการใช้ตัวแบบนี้กับวิดีโอจำนวนมากที่ถูกนำเข้าไปยังแพลตฟอร์มอย่างสม่ำเสมอ

2.7.5 ตัวแบบจากโครงข่ายประสาทเทียม

2.7.5.1 EfficientNet

EfficientNet เป็นโครงข่ายประสาทเทียมคอนโวลูชันอย่างหนึ่งซึ่งถูกพัฒนาโดยทีม Google Brain โดยจุดกำเนิดของ ConvNet คือความต้องการที่จะเพิ่มประสิทธิภาพและความแม่นยำให้กับตัวแบบ ซึ่งจากงานวิจัยหลาย ๆ แหล่ง การเพิ่มมิติของเครือข่ายไม่ว่าจะเป็นความกว้าง ความลึก หรือความละเอียดทางใดทางหนึ่งจะสามารถช่วยเพิ่มความแม่นยำให้กับตัวแบบ ที่ผ่านมาตัวแบบส่วนใหญ่จะเน้นไปกับการเพิ่มมิติให้กับตัวแบบ ทางทีม Google Brain ผู้คิดค้น EfficientNet ได้ค้นพบวิธีในการเพิ่มประสิทธิภาพและความแม่นยำให้กับตัวแบบ ซึ่งวิธีนี้เรียกว่าการปรับมาตราส่วนเชิงประกอบ (Compound Scaling)

การปรับมาตราส่วนเชิงประกอบเป็นการฉีกกฎการเพิ่มความแม่นยำให้กับตัวแบบเมื่อเปรียบเทียบกับตัวแบบที่ผ่านมาจะเพิ่มแค่มิติเพียงอย่างเดียว วิธีที่การปรับมาตราส่วนเชิงประกอบใช้คือการเพิ่มทั้ง ความกว้าง ความลึก และความละเอียดด้วยค่าคงที่ค่าเดียวกัน ซึ่งการเพิ่มทั้งสามอย่างนี้จะทำให้ตัวแบบนั้นสามารถสกัดคุณลักษณะได้อย่างแม่นยำมากยิ่งขึ้น ดังนั้น EfficientNet จึงใช้วิธีการปรับมาตราส่วนเชิงประกอบเพื่อเพิ่มความแม่นยำให้กับตัวแบบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดย Efficient Net แต่ละรุ่นตั้งแต่ B0 ถึง B7 เป็นการเพิ่มค่า Ø (constant ratio) ขึ้นตามลำดับรุ่นของตัวแบบ

2.7.5.2 ResNet

ResNet หรือ Deep Residual Network เป็นตัวแบบที่ได้รับความนิยมเป็นอย่างสูงสำหรับงานวิจัยประเภทการเรียนรู้เชิงลึก (Deep Learning) และมักนำมาใช้ในงานวิจัยใหม่ ๆ ที่ทำเกี่ยวกับคอมพิวเตอร์วิทัศน์ (Computer Vision) เกือบทุกงานมักจะนำตัวแบบนี้มาเป็นตัวแบบที่ใช้ในการสร้างคุณลักษณะ ตัวแบบนี้ได้รับการนำเสนอครั้งแรกในงานวิจัยที่มีชื่อว่า Deep Residual Learning for Image Recognition และได้รับการอ้างอิงในฐานข้อมูล Scopus แล้วมากกว่า 3,700 ครั้ง ตัวแบบนี้ได้นำเสนอวิธีการแก้ปัญหาเรื่อง Vanishing Gradient ซึ่งมักเกิดขึ้นกับโครงข่ายที่มีความลึกค่อนข้างมาก ด้วยการใส่ทางลัด (shortcut) ลงไปในโครงข่าย

หากเรามีผลลัพธ์จากชั้นคอนโวลูชันอันแรกอยู่ในแถวลำดับ (array) x และต่อมาเรามีชั้นคอนโวลูชันสองอันเพิ่มเติมตามมา โดยปกติแล้ว x จะเริ่มเลือนหายและเราสามารถติดตามได้ยากกว่าการเปลี่ยนค่าพารามิเตอร์ในชั้นแรกเพราะจะถูกกลืนไปกับชั้นที่อยู่ถัดมา ปัญหานี้ถูกแก้โดยการส่งผ่านค่า x ข้ามชั้นอื่น ๆ ขึ้นมา แล้วนำไปบวกกับผลลัพธ์ที่ได้จากชั้นที่สาม โดยชั้นที่นำมาบวกกันจำเป็นต้องมีขนาดเท่ากัน

ทางทีมวิจัยจาก Cornell ได้ทดลองเปลี่ยนความลึกของ ResNet ดู และพบว่า ResNet นั้นมีพฤติกรรมการรวมตัวของเครือข่ายย่อย ๆ ในตัวเอง โดยทางคณะได้สุ่มไม่เปลี่ยนพารามิเตอร์ระหว่างขั้นตอนการเรียนรู้ของตัวแบบเป็นบล็อกหรือหลายบล็อก โดยวิธีการเรียนรู้แบบนี้ทำให้ตัวแบบเรียนรู้ได้ดีกว่าเดิม การทดลองยังพบว่า ถ้าเราลองลดชั้นของ ResNet ที่ได้ทำการเรียนรู้แล้ว โดยเอาบางบล็อกออก ผลลัพธ์แทบจะไม่ลดลงเลย ต่างจาก VGG-19 ที่ลดลงอย่างมาก

2.7.5.3 Coarse2Fine (C2F)

Coarses2Fine ได้มีการถูกเสนอในงานวิจัย Dong. at el. (2015) โดยเริ่มต้นจากปัญหาการดึงภาพขนาดใหญ่ จึงได้มีการเสนอวิธีการฟิวชั่นสองชั้นซึ่ง ได้มีการใช้ประโยชน์จากฐานข้อมูลรูปภาพอันดับต้น ๆ หลายแหล่งจากภาพที่มีลักษณะหยาบไปจนถึงภาพที่มีรายละเอียดสูง โดยมีจุดประสงค์เพื่อลดการใช้หน่วยความจำและความซับซ้อนในการคำนวณ โดยไม่สูญเสียความจำในการดึงข้อมูล C2F โดดเด่นด้วยขั้นตอนที่แบ่งเป็นชั้นที่ประกอบด้วยการกรองและการกลั่น C2F ประกอบด้วย 3 องค์ประกอบได้แก่

การกรองสิ่งรบกวน ด้วยการแสดงภาพแบบองค์รวม รูปภาพที่มีสัญญาณรบกวนจึงถูกกรองออกจากฐานข้อมูล ดังนั้น ภาพที่นำมาใช้ในการเปรียบเทียบจึงมีจำนวนลดลงเป็นอย่างมาก หากมีภาพสีทั้งหมด N ภาพในฐานข้อมูล $\{x_n\}_{n=1}^{N}$ การใช้คุณลักษณะแบบองค์รวม อย่างเช่น HSV รูปภาพจากฐานข้อมูลจะถูกแทนที่ด้วยแบบรูปเชิงเส้นทิศทาง (vector) P-dimensional histogram $V_n = \{h_1, h_2, \dots, h_p\}$ และกำหนดให้ชุดรูปภาพเป็น $\{V_1, V_2, \dots, V_p\}$ เพื่อเป็นการลดผลกระทบของรอบที่มีค่ามาก จึงทำ l1-normalized และการยกกำลังกับข้อมูลในค่าของ HSV ดังสมการที่ 2.13 และ 2.14

$$h_i = \frac{\mathrm{h}_i}{\sqrt{\sum_{i=1}^{P} h_i}} \tag{2.13}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$h_i = |h_i|^{\alpha} \tag{2.14}$$

เมื่อ α เป็นค่าสัมประสิทธิ์และมีค่าเป็น 0.5

เรานำระยะห่าง cosine ระหว่างแบบรูปเชิงเส้นทิศทางของการค้นหาและแบบรูปเชิงเส้นทิศทางของรูปภาพจากฐานข้อมูลมาเป็นการวัดหาความคล้ายเคียง เพื่อใช้ในการเลือกรูปภาพจากชุดข้อมูล ดังสมการที่ 2.15

$$s_{q,d}^{G} = \cos(q,d) = \frac{\sum_{i=1}^{P} V_{q,i} \cdot V_{d,i}}{\sqrt{\sum_{i=1}^{P} V_{q,i}^2}\sqrt{\sum_{i=1}^{P} V_{d,i}^2}} \tag{2.15}$$

เมื่อ q คือ การค้นหา

D คือ รูปภาพจากฐานข้อมูล และ

$s_{q,i}^{G}$ เป็นตัวแทนของค่าความเหมือนที่ได้จากคุณลักษณะภายนอก

ความเหมือนของ cosine ระหว่างการค้นหาและชุดข้อมูลรูปภาพเป็นไปดังสมการที่ 2.16

$$s_{q}^{G} = (s_{q,1}, s_{q,2}, \dots, s_{q,N})^{G} \tag{2.16}$$

หากคะแนน cosine มีค่ามากเพียงใด นั้นหมายความว่าชุดรูปภาพมีความคล้ายกับการค้นหามากเท่านั้น รูปภาพจากฐานข้อมูลมีความหลายหลายเป็นอย่างมาก ดังนั้น จึงสันนิษฐานว่ารูปภาพจากฐานข้อมูลส่วนใหญ่มีความแตกต่างกันอย่างมีนัยสำคัญ และสรุปได้ว่ารูปภาพตัวแทนแบบองค์รวมสามารถใช้เพื่อกรองสัญญาณรบกวนออกจากรูปภาพได้

สำหรับแต่ละการค้นหาตัวแทน K ที่ใช้ข้อมูลร่วมกับคำค้นหาจะถูกเลือกแบบที่สามารถปรับเปลี่ยนได้จากฐานข้อมูลเดิม การที่ใช้คำค้นหาต่างกันจึงทำให้มีรูปภาพตัวแทนที่ต่างกันตามไปด้วย

การปรับน้ำหนัก (Adaptive Weighting) ในการค้นหาครั้งหนึ่ง จะสามารถหาความคล้ายคลึงของรูปภาพจากคะแนนที่องค์ประกอบรวมทำได้โดยวัดคะแนนจากรูปภาพที่มีลักษณะใกล้เคียงกัน หลังจากขั้นตอนการกรองที่กล่าวถึงข้างต้นดำเนินการเสร็จสิ้น รูปภาพของตัวแทน K จะมีความคล้ายกับการค้นหามากกว่ารูปภาพจากฐานข้อมูลเดิม ความคล้ายคลึงนี้จะแตกต่างกันออกไปขึ้นอยู่กับปัจจัยที่ได้กล่าวไว้ข้างต้น การออกแบบการปรับน้ำหนักให้กับรูปภาพตัวแทนเพื่อประเมินระดับความเหมือนกับการค้นหาจึงเป็นสิ่งที่สำคัญ

เมื่อพิจารณาถึงตัวแทนแบบองค์รวมเช่น HSV ฮิสโทแกรมแล้ว ค่า cosine ที่ถูกคำนวณก่อนหน้าจึงทำถูกนำมาเป็นพื้นฐานของน้ำหนักในการเรียนรู้ของการกระจายคุณลักษณะสีโดยรวมในรูปภาพ โดยคะแนนความเหมือนของการค้นหากับตัวแทน K ถูกแทนด้วย $(s_{q,1}, s_{q,2}, \dots, s_{q,K})^{G}$ ค่า cosine ที่มากขึ้นสอดคล้องกับความสำคัญที่สูงขึ้นของรูปภาพตัวแทน เพื่อเป็นการลดความไม่เท่าเทียมของค่าคะแนนของตัวแทนที่มีคะแนนมากกว่า ค่าความเหมือน cosine จึงถูกนำไปทำ min-max และนำไปทำให้เป็นบรรทัดฐานโดยให้อยู่ในช่วง [0,1] ดังสมการที่ 2.17

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$s_{q,i}^{G} = \left(\frac{s_{q,i} - \min(s_{q,i})}{\max(s_{q,i}) - \min(s_{q,i})}\right)^{G} \quad (2.17)$$

หลังจากนั้น คะแนนที่ผ่านการทำให้เป็นบรรทัดฐาน จะถูกนำไปใช้กับการปรับน้ำหนัก สำหรับแต่ละรูปภาพของตัวแทน K ดังสมการที่ 2.18 และ 2.19

$$s_{q,i}^{G} = \left(\frac{s_{q,i}}{\sum_{i=1}^{K} s_{q,i}}\right)^{G} \quad (2.18)$$

$$w(q,i) = s_{q,i}^{G}, i = 1,2,\dots,K \quad (2.19)$$

ผลการทดลองแสดงให้เห็นว่าการปรับน้ำหนักได้ช่วยเพิ่มความแม่นยำในการดึงรูปภาพจากทั้งชุดข้อมูลวันหยุดและชุดข้อมูล Ukbench อย่างมีนัยสำคัญ

การปรับแต่งรูปภาพ การดึงข้อมูลที่แม่นยำนั้นดำเนินการผ่านคุณลักษณะเฉพาะ ซึ่งได้นำน้ำหนักมาปรับค่าด้วยค่าที่คำนวณไว้ล่วงหน้า ขั้นตอนแรกคือการสกัด SIFT ออกจากรูปภาพโดย Hessian-affine detector $I = x_1, x_2, \dots, x_M \in R^{DxM}$ โดยรูปภาพประกอบด้วย M ตัว และ D มิติ

หลังจากนั้น ตัวแบบที่ได้รับการเรียบรู้บนชุดข้อมูล Filckr60k ได้ใช้ขั้นตอนวิธีการจัดกลุ่ม Approximate K-Means (AKM) ในการกำหนดปริมาณคุณลักษณะ SIFT คุณลักษณะ X_i ของรูปภาพถูกกำหนดให้คำที่ใกล้เคียงมากที่สุดโดยแสดงด้วย ID ของคำที่เกี่ยวข้อง ดังสมการที่ 2.20

$$q(x_i) = \arg min \|x_i - \mu_j\|^2, \mu_j \in C \quad (2.20)$$

เพื่อเพิ่มความถูกต้องให้กับคำที่ใช้ คุณลักษณะ db-bit binary signatures ถูกสร้างขึ้นสำหรับแต่ละคุณลักษณะ SIFT $b_{(x)} = [b_1(x), b_2(x), \dots, b_{d_h}(x)]$ ซึ่งทำการเข้ารหัสตำแหน่งของ SIFT ตัวบอกภายในเซลล์ Voronoi ระยะห่างระหว่างตัวบอก x และ y ที่อยู่ในเซลล์เดียวกันนั้นถูกประมาณด้วยระยะแฮมมิ่ง (Hamming Distance) ของแต่ละ Binary Signature ดังสมการที่ 2.21

$$h(b(x), b(y)) = \sum_{i=1}^{d_b} |b_i(x) = b_i(y)| \quad (2.21)$$

ตัวบอก SIFT ถูกแสดงด้วย q(x), b(x) และดัชนี Standard Inverted จะถูกสร้างขึ้น จากนั้นกำหนดให้ฟังก์ชันสำหรับการจับคู่เป็นดังสมการที่ 2.22

$$f(x,y)\begin{cases} 1, & if\ q(x) = q(y) and\ h\big(b(x), b(y)\big) \le h_t, \\ 0, & otherwises \end{cases} \quad (2.22)$$

โดย h คือระยะห่างแฮมมิ่ง และ
h_t คือค่าคงที่ของแฮมมิ่ง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ค่าคะแนนความเหมือนระหว่างการค้นหาและรูปภาพตัวแทนเป็นไปดังสมการที่ 2.23

$$s_{q,d}^{L} = \sum_{j=1}^{M} tf\ idf(q(x))^2\ if\ q(x) = q(y) and\ h\big(b(x), b(y)\big) \leq\ h_t \quad (2.23)$$

โดย M คือจำนวนคุณลักษณะในการค้นหารูปภาพ q
X และ y คือคุณลักษณะของการค้นหารูปภาพจากฐานข้อมูล และ
L แทนการแสดงคะแนนที่ถูกคำนวณมาจากคุณลักษณะที่อยู่ใกล้เคียง

พิจารณาจากองศาของความเหมือนที่ได้จากการวัดการแทนแบบองค์รวมพบว่าการเรียงลำดับของรูปภาพตัวแทน K ได้มาจากคุณลักษณะที่ได้รับการปรับปรุงโดยใช้การปรับน้ำหนักตามที่ได้กล่าวมาข้างต้น คะแนนความเหมือนท้ายสุดของการค้นหาและรูปภาพตัวแทนเป็นไปตามสมการที่ 2.24

$$s_{q,d}^{F} = s_{q,d}^{L} \times w_{q,d} \quad (2.24)$$

โดย $s_{q,d}^{F}$ คือคะแนนความเหมือนท้ายสุดของการค้นหาและรูปภาพ candidate
D คือ ชุดข้อมูลรูปภาพ และ
$w_{q,d}$ คือค่าน้ำหนักที่ได้จากวิธีการถ่วงน้ำหนัก

งานวิจัยในครั้งนี้ได้เสนอการทำงานฟิวชั่นแบบสองชั้นโดยใช้ประโยชน์จากข้อมูลที่อยู่ใกล้เคียงและทำการเรียงข้อมูลที่อยู่ในฐานข้อมูลจากรูปที่หยาบที่สุดไปจนถึงรูปที่ละเอียดที่สุด โดยคำนึงถึงประสิทธิภาพและการใช้พื้นที่ในการเก็บข้อมูลที่มีอยู่น้อย งานวิจัยนี้ได้ใช้ HSV ฮิสโทแกรม เพื่อค้นหารูปภาพจากฐานข้อมูลและกรองรูปภาพที่มีสัญญาณรบกวนออก จากนั้นจึงทำการค้นหารูปภาพที่เหมาะสมจะมาเป็นตัวแทน โดยให้น้ำหนักแบบปรับตัวที่ได้เรียนรู้ผ่านคะแนนความเหมือนแบบทั่วไปเพื่อปรับปรุงการดึงข้อมูลให้แม่นยำขึ้น

มีการทดลองวิธีนี้กับชุดข้อมูลมาตรฐานสองชุด ได้แก่ ชุดข้อมูลวันหยุด และชุดข้อมูล Ukbench งานวิจัยนี้ได้แสดงให้เห็นว่าวิธีนี้มีประสิทธิภาพดีกว่าวิธีการฟิวชั่นรุ่นล่าสุดในแง่ของปริมาณการใช้พื้นที่ในการเก็บข้อมูล ความซับซ้อนในการคำนวณ และความแม่นยำที่สามารถแข่งขันกับเทคโนโลยีล้ำสมัยได้ โดยความแม่นยำของ C2F กับชุดข้อมูลวันหยุดนั้นสูงกว่าวิธีการฟิวชั่นแบบ Query Adaptive ถึง 3.31%

2.8 มาตรวัดประสิทธิภาพของตัวแบบตรวจจับวัตถุ

2.8.1 ค่าความเที่ยง

ค่าความแม่นยำคืออัตราส่วนระหว่างค่าบวกจริง (True Positives) และการทำนายที่เป็นบวก (Positives) ทั้งหมด ยกตัวอย่างเช่น การระบุได้ถูกต้องว่าผู้ป่วยเป็นโรคหัวใจจากผู้ป่วยทั้งหมดที่เป็นโรคหัวใจอยู่จริง โดยมีสมการดังสมการที่ 2.25

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$Precision = \frac{True\ Positive\ (TP)}{True\ Positive\ (TP)+False\ Positive\ (FP)} \quad (2.25)$$

2.8.2 ค่าเรียกคืน

ค่าเรียกคืนคือการวัดว่าตัวแบบของเรานั้นสามารถระบุค่าบวกจริง (True Positive) ได้อย่างถูกต้องเท่าใด ยกตัวอย่างเช่น สำหรับผู้ป่วยทุกรายที่เป็นโรคหัวใจจริง ค่าเรียกคืนสามารถบอกได้ว่ามีกี่รายที่ตัวแบบระบุได้ว่าเป็นโรคหัวใจอย่างถูกต้อง โดยมีสมการดังสมการที่ 2.26

$$Recall = \frac{True\ Positive\ (TP)}{True\ Positive\ (TP)+False\ Negative\ (FN)} \quad (2.26)$$

2.8.3 ค่าคะแนน F1

ในกรณีที่เราพิจารณาว่าค่าเรียกคืนที่สูงมีความสำคัญมากกว่าค่าความเที่ยงที่สูง เช่น หากต้องการตรวจหาผู้ป่วยที่เป็นโรคหัวใจให้ได้มากที่สุด หรืออีกตัวอย่างหนึ่ง เช่น การจำแนกว่าลูกค้าธนาคารนั้นเป็นผู้ที่มีความเป็นไปได้ในการไม่ชำระหนี้หรือไม่ ตัวแบบนี้ควรมีความเที่ยงที่สูง เนื่องจากธนาคารไม่ต้องการเสียลูกค้าที่ชำระหนี้ตามเวลา ค่าความเที่ยงที่สูงจึงมีความสำคัญกว่าในการพิจารณาตัวแบบ

นอกจากนี้ยังมีกรณีที่ค่าความเที่ยงและค่าเรียกคืนนั้นมีความสำคัญเท่ากัน ยกตัวอย่างเช่น หากแพทย์แจ้งว่าผู้ป่วยที่เป็นโรคอื่นและไม่ได้เป็นโรคหัวใจมีความสำคัญเท่ากับผู้ป่วยที่เป็นโรคหัวใจเนื่องจากการตัดสินใจว่าผู้ป่วยเป็นโรคใดโรคหนึ่งจะสามารถทำให้ผู้ป่วยคนนี้อยู่ต่อเพื่อรักษาโรคอื่น ๆ ได้ หากจุดประสงค์ของเราเป็นไปดังตัวอย่างสุดท้าย เราจะต้องพิจารณาทั้งค่าความเที่ยงที่สูงและค่าเรียกคืนที่สูงด้วยเช่นกัน ดังนั้น จึงมีตัววัดประสิทธิภาพอีกตัวหนึ่งที่เรียกว่าค่าคะแนน F1 หรือเป็นการหาค่าเฉลี่ยแบบฮาร์มอนิก (Harmonic Mean) ของค่าความเที่ยงและค่าเรียกคืน โดยมีสมการที่ 2.27

$$F1\ Score = 2 \cdot \frac{Precision \cdot Recall}{Precision+Recall} \quad (2.27)$$

2.8.4 Intersection Over Union (IoU)

Intersection Over Union คือการวัดการทับซ้อนระหว่าง 2 ขอบเขต เราใช้ค่านี้เพื่อวัดว่าขอบเขตที่ตัวแบบคาดการณ์นั้นทับซ้อนกับขอบเขตความจริงบนรูปภาพมากน้อยเพียงใด ในบางชุดข้อมูล เรามีการกำหนดค่า IoU ไว้ล่วงหน้า เช่น 0.5 ในการจำแนกประเภท เพื่อวัดว่าการคาดคะเนนั้นเป็นค่าบวกจริงหรือค่าบวกเท็จ

2.8.5 ค่าเฉลี่ยความเที่ยง

ค่าเฉลี่ยความเที่ยง (Average Precision) คือการหาพื้นที่ภายใต้เส้นโค้ง precision-recall ดังสมการที่ 2.28

$$AP = \int_0^1 p(r)dx \quad (2.28)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ค่าความเที่ยงและค่าเรียกคืนอยู่ระหว่าง 0 และ 1 เสมอ จึงทำให้ค่าเฉลี่ยความเที่ยง (Average Precision : AP) มีค่าระหว่าง 0 และ 1 เช่นกัน ก่อนที่จะทำการคิดค่า AP สำหรับการตรวจจับวัตถุ เรามักจะทำกให้ค่าเรียบเพื่อหารูปแบบก่อนเสมอ

2.8.6 ค่าเฉลี่ยความเที่ยงของชุดข้อมูล COCO (COCO mAP)

งานวิจัยหลายแหล่งมักจะให้ผลลัพธ์โดยใช้ชุดข้อมูล COCO เท่านั้น ใน mAP ของ COCO interlated AP จำนวน 101 จุด ถูกนำมาใช้ในการคำนวณ สำหรับ COCO นั้น AP คือ ค่าเฉลี่ยความเที่ยงของ IoU หลายรายการ เช่น AP@[.5:.95] หมายถึง AP เฉลี่ยสำหรับ IoU ตั้งแต่ 0.5 จนถึง 0.95 โดยมีระยะห่างที่ 0.05

mAP หรือ mean average precision คือค่ากลางเฉลี่ยของค่า AP ในบางงานวิจัยได้มีการนำค่า AP ที่ได้จากแต่ละ class มาหาค่าเฉลี่ย แต่ในบางงานวิจัยกำหนดให้ AP และ mAP นั้นมีความหมายเหมือนกัน ในงานวิจัยของ COCO ได้กำหนดให้ AP และ mAP ไม่มีความแตกต่างกัน และถือว่าเป็นค่าเดียวกัน

2.9 งานวิจัยที่เกี่ยวข้อง

เนื่องจากไม้วัดระดับในแต่ละประเทศมีลักษณะที่ไม่เหมือนกัน จึงมีงานวิจัยเกี่ยวกับอ่านค่าระดับน้ำที่แตกต่างกันออกไปทั้งงานวิจัยของต่างประเทศและงานวิจัยในประเทศ ถึงแม้จะมีงานวิจัยที่น่าสนใจแต่มีบางส่วนของขั้นตอนการทดลองเท่านั้นที่สามารถนำเทคนิคมาประยุกต์ใช้ในงานวิจัยในครั้งนี้ได้

ธริศร์ และ อุรินทร์ (พ.ศ. 2562) ได้มีการพัฒนาระบบอ่านค่าระดับน้ำด้วยการวิเคราะห์ค่าจุดเริ่มเปลี่ยนของงภาพระดับสีและเทคนิคการวิเคราะห์การถดถอย โดยการนำภาพที่ได้รับมาจากกล้อง CCTV มาแปลงภาพสีให้เป็นเป็นภาพระดับสีเทา และปรับตำแหน่งของแกน X ที่จะใช้วิเคราะห์ให้อยู่ในแนวตั้งฉาก จากนั้น ทำการประมวลผลเพื่อหาจุดเริ่มเปลี่ยนโดยวิเคราะห์ค่าจุดเริ่มเปลี่ยนของทุก ๆ พิกเซลในแนวเกน Y ของจำแหน่งพิกเซลในแนวแกน X เช่น หากแกนกลางของเกจวัดระดับน้ำอยู่ตรงตำแหน่งพิกเซลที่ 200 ตามแนวแกน X ตำแหน่งพิกเซลตั้งแต่ x_{200}, y_1 ถึง x_{200}, y_{270} ก็จะถูกนำมาประมวลผล ซึ่งจะทำให้ได้ค่าจุดเริ่มเปลี่ยนจำนวน 270 ค่า นำค่าจุดเริ่มเปลี่ยนไปสร้างกราฟในแต่ละพิกเซลตามแนวแกน X ผลลัพธ์ที่ได้คือระดับเส้นแนวนอนของระดับน้ำในภาพ ผู้วิจัยได้นำค่าพิกเซลที่ได้จากการประมวลผลมาสร้างสมการหรือแบบบจำลองทางคณิตศาสตร์สำหรับการคำนวณหาค่าระดับน้ำจากการแปลงค่าพิกเซลได้ โดยใช้เทคนิคการวิเคราะห์ถดถอยอย่างง่ายหรือแบบเชิงเส้น (Linear-Regression Analysis Technique) สำหรับสถานีคลอง 13 สถานีมหาสวัสดิ์ และสถานีประตูระบายน้ำจุฬาลงกรณ์ ตามลำดับ ดังสมการที่ 2.29, 2.30 และ 2.31

$$Y = (X - 397.39) / -177.86 \quad (2.29)$$

$$Y = (X - 490) / -160 \quad (2.30)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$Y = (X - 334.35) / -100.31 \quad (2.31)$$

เมื่อ X คือ ค่าพิกเซล และ Y คือ ค่าระดับน้ำ ค่าสัมประสิทธิ์สหสัมพันธ์ของสถานีที่ 1-3 มีค่าเท่ากับ 0.987, 0.963 และ 0.945 ตามลำดับ และค่าสัมประสิทธิ์ของการตัดสินในหรือค่า R-Square ของสถานีที่ 1-3 มีค่าเท่ากับ 0.0974, 0.923 และ 0.928 ตามลำดับ จึงสามารถสรุปได้ว่าระบบอ่านค่าระดับน้ำที่พัฒนาขึ้นโดยใช้การประมวลผลด้วยสมการที่กล่าวข้างต้น เป็นระบบที่มีประสิทธิภาพโดยรวมอยู่ในเกณฑ์ที่ดี สามารถใช้อ่านค่าระดับน้ำแทนการอ่านด้วยสายตาได้

Dou et al. (2023) ได้มีการใช้การประมวลผลภาพและโครงข่ายประสาทเทียมคอนโวลูชันในการอ่านค่าระดับน้ำ โดยวิธีที่ใช้คือการประมวลผลภาพและโครงข่ายประสาทเทียมคอนโวลูชันเพื่อหาค่าระดับน้ำจากรูปภาพของไม้วัดระดับน้ำ โดยรูปภาพนำเข้าคือรูปภาพจากกล้องที่ถ่ายจากสถานที่จริงและได้มีการใช้วิธีการประมวลผลภาพต่าง ๆ ในการจัดการกับรูปภาพที่ได้จากกล้อง ตัวอักษรและพิกัดของตำแหน่งไม้วัดระดับน้ำถูกนำไปเปรียบเทียบและแยกออกเป็นส่วน และใช้ CNN ในการตรวจจับตัวเลขที่ปรากฏบนไม้วัดระดับน้ำ จำนวนเส้นที่ใช้ในการวัดอัตราส่วนถูกคำนวนมากจากการสำรวจจุดพิกเซลของภาพทวิภาคจากพิกัดของตัวเลขตามการเปลี่ยนแปลงของค่าพิกเซล และท้ายสุด ค่าของระดับน้ำถูกคำนวณโดยใช้ความสัมพันธ์ทางคณิตศาสตร์ระหว่างค่าของตัวเลขและจำนวนเส้นบนไม้วัดระดับน้ำที่ปรากฏอยู่

งานวิจัยที่ได้มีการนำตัวแบบ YOLO ไปใช้คืองานวิจัยของ Wang. at el. (2018) ได้กล่าวว่าในการตรวจจับวัตถุที่อยู่ใต้น้ำโดยการนำตัวแบบ YOLO มาปรับให้มีประสิทธิภาพมากขึ้น และงานวิจัยของ Chengwenyang at el. (2015) ได้กล่าวว่า ในการตรวจจับวัตถุที่อยู่บนผิวน้ำเพื่อช่วยในการทำความสะอาดแวดล้อมของพื้นผิวน้ำ โดยงานวิจัยได้มีการปรับโครงสร้างและสถาปัตยกรรมบางส่วนของ YOLO ให้ดีขึ้น งานวิจัยการตรวจจับวัตถุใต้น้ำได้ใช้ YOLO มาตอบข้อจำกัดของรูปภาพนำเข้าที่มีลักษณะที่ไม่ชัด มีความผิดเพี้ยนของสี และมีสัญญาณรบกวนที่เกิดขึ้นได้สำหรับการถ่ายภาพใต้น้ำ งานวิจัยประกอบไปด้วย 2 ส่วน ได้แก่ การสร้างโครงข่ายเพื่อเสริมคุณภาพให้กับรูปภาพ และอีกส่วนหนึ่งคือการปรับปรุง YOLO ให้มีประสิทธิภาพขึ้น การเพิ่มประสิทธิภาพในส่วนของรูปภาพ ได้มีการใช้ทฤษฎีของง Retinex เพื่อลบสิ่งที่ไม่จำเป็นออกจากภาพถ่ายที่ได้จากใต้น้ำ โครงข่ายของ YOLO ที่ได้รับการปรับปรุงแล้ว สามารถแยกคุณลักษณะในหลาย ๆ มาตราส่วนและรวมทุกอย่างเข้าด้วยกันได้ นอกจากนี้ ตัวแบบนี้ยังแยกการจำแนกภาพและการถดถอยออกจากกัน ซึ่งทำให้สองขั้นตอนนี้มีความแตกต่างกันอย่างชัดเจน

เนื่องจาก YOLOv5 มีปัญหาความแม่นยำที่ไม่สูงสำหรับการตรวจจับวัตถุบนผิวน้ำที่มีขนาดเล็ก งานวิจัยจึงเน้นไปในการจรวจจับวัตถุบนพื้นผิวน้ำจากกล้องที่อยู่บนเรือไร้คนขับ งานวิจัยนี้ได้มีการเพิ่มประสิทธิภาพให้กับ YOLO ด้วยการเพิ่มชั้นสำหรับการตรวจจับวัตถุเข้ามาแทนที่โครงสร้าง C3 ที่มีอยู่ในโครงข่ายของ YOLOv5 เพื่อเพิ่มโครงสร้างที่เข้ารหัสในโครงข่ายเพิ่มเติม และเพิ่มการทำงาน CBAM Attention Mechanism จากการทดลองที่เกิดขึ้น ตัวแบบ YOLOv5 ที่ได้รับการปรับปรุงแล้วไม่ได้ดีกว่า YOLOv5 แค่ค่า P-R curve และ F1 Score เพียงอย่างเดียว แต่ยังดีกว่าอีกสาม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นตอนที่ได้มีการเปรียบเทียบในงานวิจัยนี้ในเรื่องของความแม่นยำและความเร็วอีกด้วย จึงสามารถสรุปได้ว่าขั้นตอนวิธีนี้สามารถนำไปใช้งานได้จริงและมีประสิทธิภาพเหนือกว่าตัวแบบอื่น ๆ ในการตรวจจับวัตถุจากภาพถ่ายจากเรือที่ไม่มีคนขับ

งานวิจัยของ Manasa (2019) ได้มีการใช้ตัวแบบที่ผ่านการเรียนรู้แล้ว (YOLOv5) ในการอ่านเลขทะเบียนรถยนต์ โดยได้นำรูปภาพที่ใช้ในการเรียนรู้มาจากกล้องวงจรปิดที่ติดตั้งอยู่บริเวณทางเข้าและทางออก มีการบันทึกภาพทั้งกลางวันและกลางคืน และรถยนต์แต่ละคันเคลื่อนที่ในความเร็วที่ไม่เท่ากัน มีการเก็บรูปภาพทั้งหมด 600 รูปภาพในขั้นตอนของการเรียนรู้

ได้มีการนำตัวแบบที่ผ่านการเรียนรู้มาแล้วมาใช้ทั้งหมดสองครั้ง โดยครั้งแรกคือการตรวจหาตำแหน่งของป้ายทะเบียนรถยนต์ และครั้งที่สองคือการหาว่าตัวเลขแต่ละตำแหน่งคือตัวเลขใดบ้าง ประสิทธิภาพของงานวิจัยเป็นไปดังตารางที่ 2.2

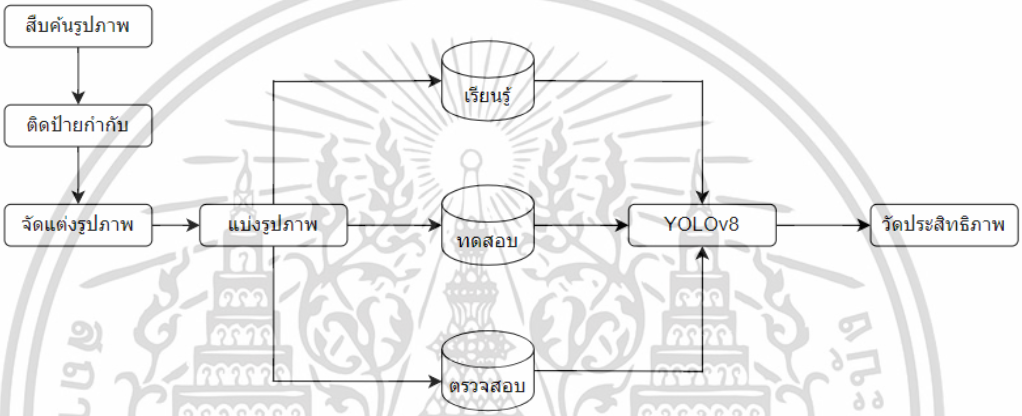
ตารางที่ 2.2 ประสิทธิภาพของงานวิจัย

Epoch’s	TP	FP	FN
1000	384	273	227
2000	544	76	67
3000	545	69	66
4000	526	77	85
5000	548	69	63
6000	550	62	61
7000	549	65	62
8000	549	70	62
9000	546	73	65
10000	548	56	63

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3
วิธีการดำเนินงานวิจัย

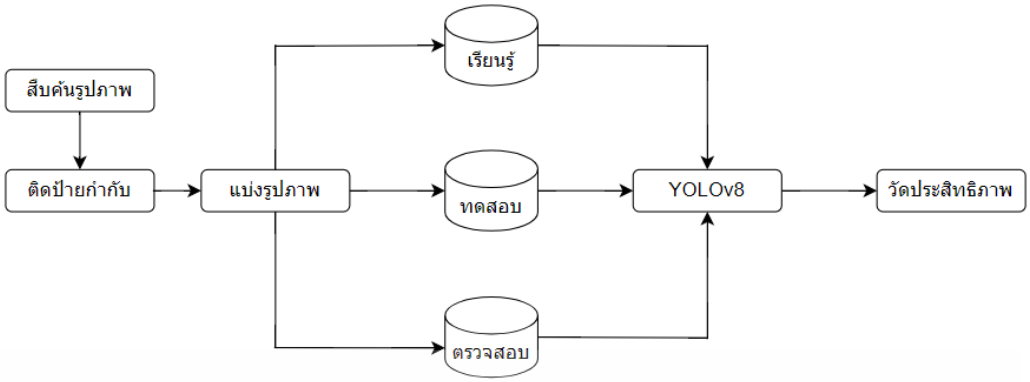
การวิจัยในครั้งนี้ เป็นการสอนให้ตัวแบบเรียนรู้เพื่ออ่านค่าระดับน้ำจากรูปภาพได้ เพื่อเป็นการใช้เทคโนโลยีและตัวแบบที่ทันสมัยมาประยุกต์ใช้กับปัญหาที่เกิดขึ้นจริงในสังคม โดยมีการใช้ตัวแบบที่ผ่านการเรียนรู้มาแล้วทั้งหมดสองครั้ง สำหรับการตรวจจับไม้วัดระดับน้ำในรูปภาพในตัวแบบที่ 1 และสำหรับการหาตำแหน่งและอ่านค่าระดับน้ำในตัวแบบที่ 2 โดยมีขั้นตอนการทำงานของตัวแบบที่ 1 และตัวแบบที่ 2 ดังรูปที่ 3.1 และรูปที่ 3.2



รูปที่ 3.1 ขั้นตอนการทำงานของตัวแบบที่ 1

จากรูปที่ 3.1 จะพบว่างานวิจัยนี้ได้มีการสืบค้นข้อมูลรูปภาพจากเว็บไซต์ Google และ Roboflow Universe เพื่อใช้เป็นรูปภาพสำหรับการเรียนรู้ของตัวแบบ จากนั้น ได้มีการนำรูปภาพที่ได้มาติดป้ายกำกับสำหรับการเรียนรู้ เพื่อระบุตำแหน่งของไม้วัดระดับน้ำจากรูปภาพ โดยการติดป้ายกำกับนั้น ได้มีการทำภาพเว็บไซต์ Roboflow หลังจากการติดป้ายกำกับแล้ว งานวิจัยนำรูปภาพในส่วนที่ใช้สำหรับการเรียนรู้ไปจัดแต่งเพื่อให้ได้รูปภาพที่สื่อถึงสถานการณ์จริงเพิ่มขึ้น โดยวิธีที่ใช้คือการปรับความสว่างของรูปภาพและการเพิ่มสัญญาณรบกวนให้กับรูปภาพ หลังจากนั้น รูปภาพทั้งหมดจะถูกแบ่งออกเป็นรูปภาพสำหรับการเรียนรู้ การทดสอบ และการตรวจสอบ หลังจากนั้น นำรูปภาพที่แบ่งออกทั้งหมดเข้าสู่ตัวแบบที่ผ่านการเรียนรู้แล้วที่มีชื่อว่า YOLOv8m ท้ายสุด งานวิจัยได้นำผลลัพธ์ที่ได้มาทำการเปรียบเทียบเพื่อหาการแบ่งสัดส่วนที่ดีที่สุดเพื่อให้ได้ผลลัพธ์ที่มีประสิทธิภาพมากที่สุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

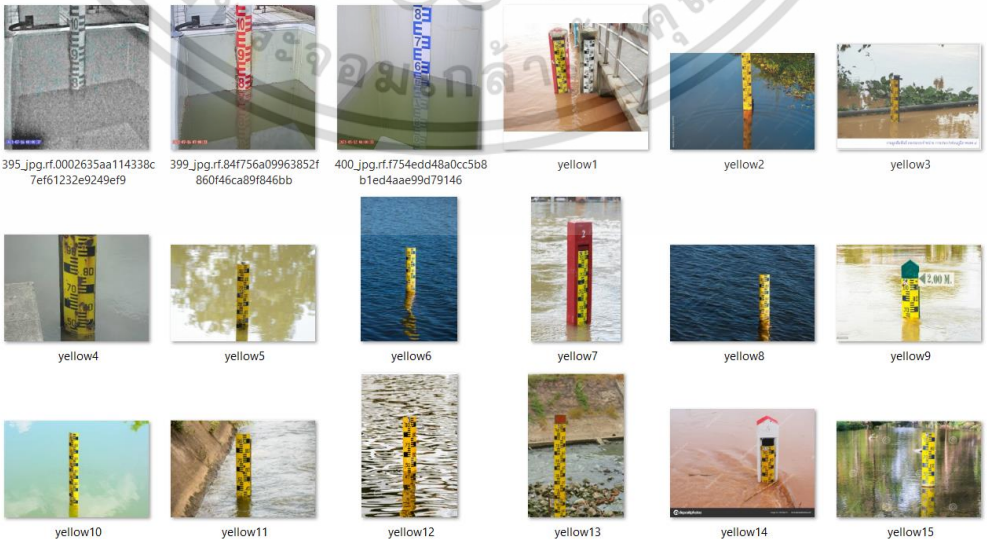


รูปที่ 3.2 ขั้นตอนการทำงานของตัวแบบที่ 2

จากรูปที่ 3.2 จะพบว่างานวิจัยนี้ได้มีการสืบค้นข้อมูลรูปภาพจากเว็บไซต์ Google และ Roboflow Universe เพื่อใช้เป็นรูปภาพสำหรับการเรียนรู้ของตัวแบบ จากนั้น ได้มีการนำรูปภาพที่ได้มาติดป้ายกำกับสำหรับการเรียนรู้ เพื่อหาตำแหน่งและค่าของตัวเลขที่อยู่บนไม้วัดระดับน้ำจากรูปภาพ ตัวแบบที่ 2 ไม่ได้มีการจัดแต่งรูปภาพเพิ่มเติมเหมือนกับตัวแบบที่ 1 เนื่องจากรูปภาพของตัวเลขมีมากเพียงพอสำหรับการทดสอบแล้ว นำรูปภาพที่ติดป้ายกำกับแล้วแบ่งออกเป็นรูปภาพสำหรับการเรียนรู้ การทดสอบ และการตรวจสอบเช่นเดียวกับตัวแบบที่ 1 และนำรูปภาพที่ได้แบ่งออกเป็นส่วนต่าง ๆ เข้าสู่ตัวแบบที่ผ่านการเรียนรู้แล้วที่มีชื่อว่า YOLOv8m นำผลลัพธ์ที่ได้มาทำการเปรียบเทียบเพื่อหาการแบ่งสัดส่วนที่ดีที่สุดเพื่อให้ได้ผลลัพธ์ที่มีประสิทธิภาพมากที่สุด

3.1 รูปภาพที่ใช้ในตัวแบบที่ 1

รูปภาพที่ใช้ในงานวิจัยครั้งนี้มาจากการสืบค้นรูปภาพจากอินเทอร์เน็ต โดยแหล่งที่มาของรูปภาพคือ Google และ RoboFlow รูปภาพที่ใช้เป็นรูปภาพของไม้วัดระดับน้ำ โดยในภาพต้องประกอบไปด้วยไม้วัดระดับน้ำ และตัวเลขบนไม้วัดระดับน้ำที่สามารถเห็นได้อย่างชัดเจนในระดับหนึ่ง รูปภาพที่ใช้ในการเรียนรู้มีทั้งหมด 146 รูปภาพ ดังแสดงในรูปที่ 3.3



รูปที่ 3.3 รูปภาพตัวอย่างของชุดข้อมูลนำมาใช้ในตัวแบบที่ 1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 3.3 จะพบว่ารูปภาพทั้งหมดประกอบไปด้วยไม้วัดระดับน้ำที่อยู่ตามพื้นที่ต่าง ๆ รูปภาพที่สืบค้นมาจาก Google มีหลากหลายขนาด และอยู่ในสภาพแวดล้อมที่แตกต่างกันออกไป รูปภาพที่สืบค้นมาจาก Roboflow จะมีความหลากหลายน้อยกว่า เนื่องจากเป็นภาพที่ได้มาจากกล้องตัวเดียวกัน และมีการจำลองสถานการณ์เพื่อวัตถุประสงค์ที่ใกล้เคียงกับงานวิจัยในครั้งนี้

ในส่วนของการทดลองใช้งานจริง ได้นำรูปภาพมาจากกล้องวงจรปิดจริง ซึ่งได้มีการใช้งานอยู่ในปัจจุบัน และไม่ได้ใช้สำหรับการเรียนรู้ การทดสอบ และการตรวจสอบมาก่อน เพื่อใช้ในการวัดความแม่นยำของตัวแบบกับสถานการณ์จริงอีกครั้งหนึ่ง

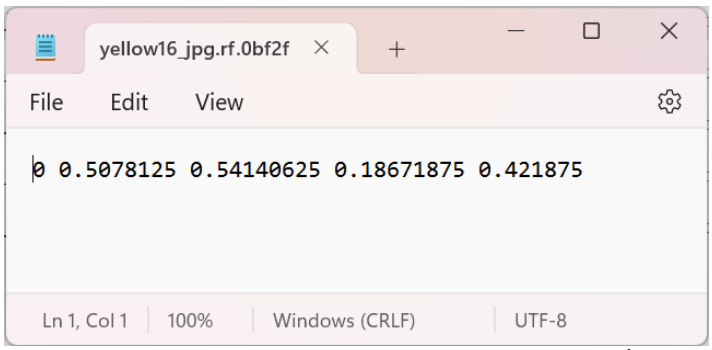
3.2 ติดป้ายกำกับให้กับไม้วัดระดับน้ำ

ติดป้ายกำกับให้กับวัตถุที่อยู่ในรูปภาพ โดยกำหนดให้ตำแหน่งของจุดสี่เหลี่ยมสี่จุดครอบคลุมพื้นที่ของวัตถุที่ต้องการเพื่อนำตำแหน่งของวัตถุไปใช้ในการเรียนรู้ของตัวแบบต่อไป โดยส่วนที่ต้องการคือตำแหน่งของไม้วัดระดับน้ำในส่วนที่อยู่เหนือน้ำขึ้นมาดังแสดงในรูปที่ 3.4

รูปที่ 3.4 ตัวอย่างผลลัพธ์ที่ต้องการจากตัวแบบแรก

จากรูปที่ 3.4 จะพบว่าผลลัพธ์ที่ต้องการจากการติดป้ายกำกับคือส่วนของไม้วัดระดับน้ำบริเวณที่เหนือน้ำขึ้นมาจนถึงขอบด้านบนสุดของไม้วัดระดับน้ำ ซึ่งสามารถอ่านค่าตัวเลขที่ปรากฏอยู่บนไม้วัดระดับน้ำได้อย่างชัดเจนในระดับหนึ่ง

เนื่องจากมีวัตถุเพียงอย่างเดียวที่ต้องการตรวจจับ ชุดข้อมูลสำหรับการติดป้ายกำกับที่ได้ทำการแบ่งออกเป็น การเรียน การทดสอบ และการตรวจสอบ จึงถูกเก็บในลักษณะของจุดสี่จุดล้อมรอบไม้วัดระดับน้ำเพียงอย่างละ 1 ค่าเท่านั้น ซึ่งข้อมูลการติดป้ายกำกับของรูปที่ 3.4 เป็นไปดังรูปที่ 3.5



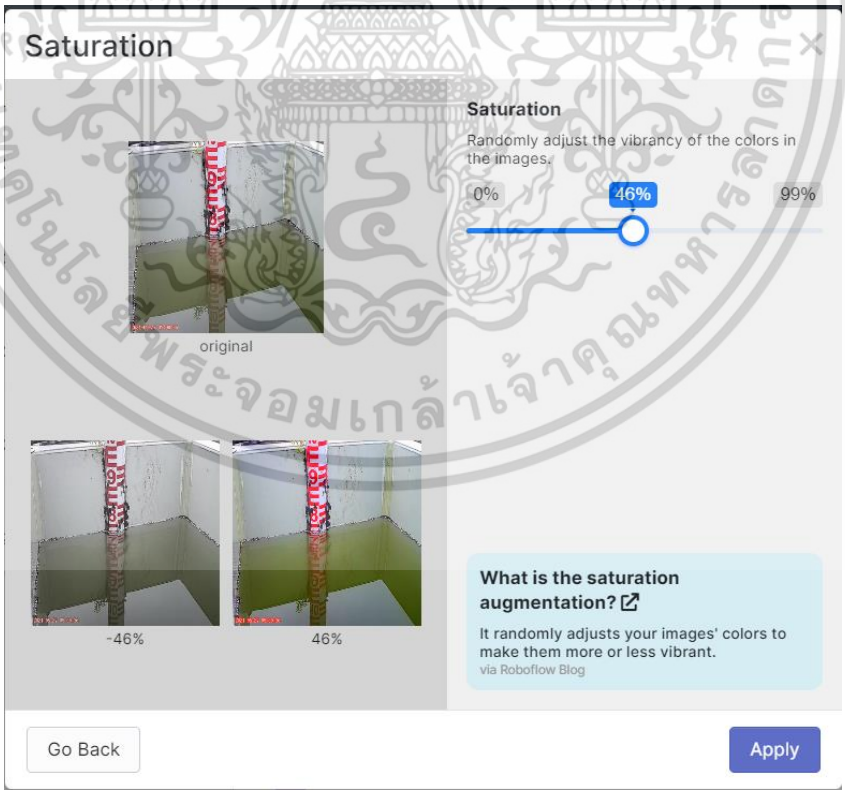
รูปที่ 3.5 ตำแหน่งของจุดสี่จุดล้อมรอบไม้วัดระดับน้ำ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 3.5 จะพบว่าตัวเลขที่ได้จากการติดป้ายกำกับมีทั้งหมด 5 ตัวเลข เลข 0 ในตำแหน่งแรกเป็นการระบุว่าวัตถุเป้าหมายคือ class ที่ 1 ซึ่งในตัวแบบที่ 1 มีเพียง class เดียวนั้นนั่นคือไม้วัดระดับน้ำ เลขในตำแหน่งแรกจึงเป็นเลข 0 เสมอ เลขในตำแหน่งที่ 2, 3, 4 และ 5 คือ x1, x2, y1 และ y2 ตามลำดับ ซึ่ง x1 และ x2 คือจุดเริ่มต้นและจุดสิ้นสุดของแกน x y1 และ y2 คือจุดเริ่มต้นและจุดสิ้นสุดของแกน y โดยจุดที่ x=0 และ y=0 คือมุมบนด้านซ้ายของรูปภาพ หาก x เข้าใกล้ 0 หมายความว่าตำแหน่งของไม้วัดระดับน้ำอยู่ทางซ้ายของรูปภาพ และหาก y เข้าใกล้ 0 หมายความว่าตำแหน่งของไม้วัดระดับน้ำอยู่ทางด้านบนของรูปภาพ

3.3 จัดแต่งรูปภาพจากภาพที่มีอยู่ให้กับตัวแบบที่ 1

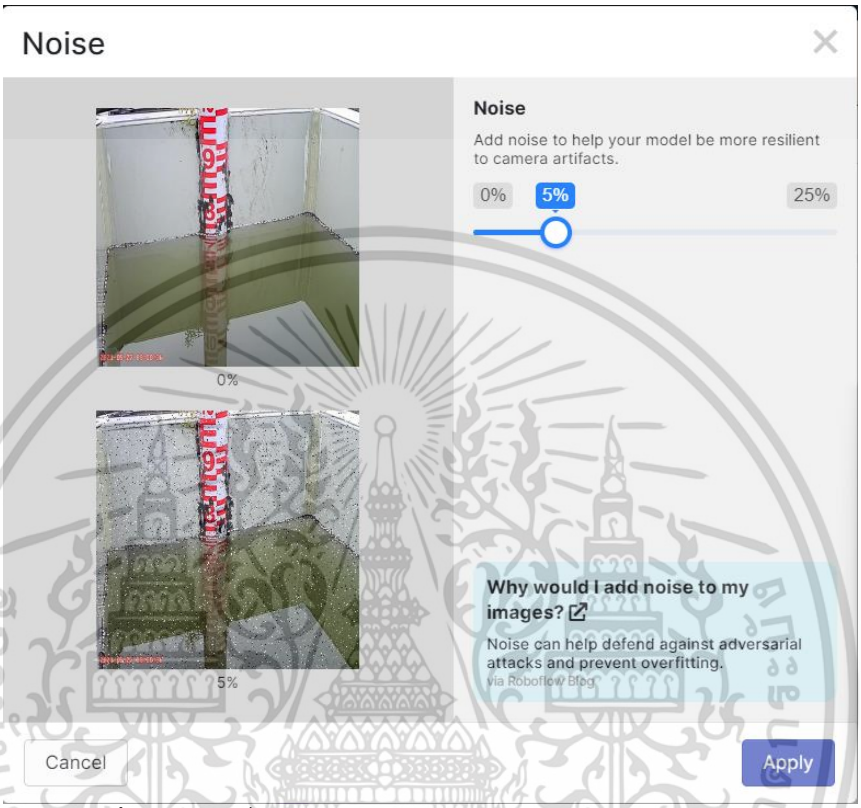
เนื่องจากภาพที่มีอยู่มีปริมาณน้อย จึงเพิ่มขั้นตอนการจัดแต่งรูปภาพให้กับงานวิจัยในครั้งนี้ โดยวิธีที่เลือกใช้คือ การปรับแสงของภาพให้สื่อถึงช่วงเช้า กลางวัน และกลางคืน และการเพิ่มสัญญาณรบกวนให้กับรูปภาพเพื่อแสดงถึงสัญญาณรบกวนเวลาได้รับภาพจากกล้องวงจรปิดในสถานการณ์จริงให้กับรูปภาพที่ได้แบ่งไว้สำหรับการเรียนรู้ ดังแสดงในรูปที่ 3.6 และ รูปที่ 3.7 ตามลำดับ ส่วนของภาพที่ใช้ในการทดสอบนั้น ไม่ได้มีการจัดแต่งรูปภาพเพิ่มเติม และได้มีการปรับรูปภาพทั้งหมดให้มีขนาด 640 x 640 พิกเซล เนื่องจากเป็นขนาดรูปภาพนำเข้าที่ตัวแบบต้องการ โดยขั้นตอนทั้งหมดนี้ทำให้เรามีรูปภาพทั้งหมด 276 รูปภาพสำหรับการทดสอบที่ 1 และ 2 และ 312 รูปภาพ สำหรับการทดสอบที่ 3 และ 4 ชุดข้อมูลแรกมีจำนวน class เพียง 1 class และรูปภาพแต่ละรูปประกอบไปด้วยไม้วัดระดับน้ำเพียง 1 ไม้เท่านั้น ดังตารางที่ 3.1



รูปที่ 3.6 การปรับแสงของรูปภาพสำหรับการเรียนรู้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 3.6 จะพบว่าได้มีการปรับแสงของรูปภาพในส่วนของการเรียนรู้โดยกำหนดให้ความสว่างของภาพเพิ่มขึ้นและลดลง 46% ซึ่งรูปภาพที่ได้หลังจากการปรับแสงจะสามารถสื่อถึงเวลากลางวันและกลางคืนได้ การปรับแสงของรูปภาพจะได้รูปภาพสำหรับการเรียนรู้เพิ่มขึ้นถึงสองเท่าจากที่มีอยู่เดิม



รูปที่ 3.7 การเพิ่มสัญญาณรบกวนให้กับรูปภาพสำหรับการเรียนรู้

จากรูปที่ 3.7 จะพบว่าได้มีการเพิ่มสัญญาณรบกวนให้กับรูปภาพ 5% โดยรูปภาพหลังจากการปรับแต่งจะมีจุดสัญญาณรบกวนปรากฏขึ้นแต่ยังคงความชัดเจนในการมองเห็นไม้วัดระดับน้ำไว้คงเดิม โดยขั้นตอนการเพิ่มสัญญาณรบกวนให้กับรูปภาพจะทำให้ได้ภาพสำหรับการเรียนรู้เพิ่มหนึ่งเท่าจากที่มีอยู่เดิม

3.4 การเรียนรู้ของตัวแบบที่ 1

นำรูปภาพที่ผ่านการติดป้ายกำกับแล้วมาใช้สำหรับการเรียนรู้ของตัวแบบ โดยใช้ตัวแบบที่ผ่านการเรียนรู้แล้วที่มีชื่อว่า YOLOv8m ทำการแบ่งข้อมูลสำหรับการเรียนรู้ การทดสอบหามาตรวัดประสิทธิภาพ และการตรวจสอบหลังจากที่ได้ตัวแบบที่ดีสุดมาแล้วว่าตัวแบบสามารถทำงานได้ดีเพียงใดกับข้อมูลที่ไม่เคยเห็นมาก่อน ได้มีการแบ่งข้อมูลออกเป็นสามส่วนดังตารางที่ 3.1 เพื่อให้ได้การแบ่งส่วนที่ทำให้ตัวแบบมีประสิทธิภาพมากที่สุด และทำการกำหนดพารามิเตอร์ โดยกำหนดให้ epoch = 100, Image Size = 640 และ Patience = 100

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.1 การแบ่งข้อมูลสำหรับการเรียนรู้ของตัวแบบที่ 1

รอบการทดสอบ	การเรียนรู้	การทดสอบ	การตรวจสอบ	รวม
การทดสอบที่ 1	195 รูปภาพ 70%	41 รูปภาพ 15%	40 รูปภาพ 15%	276 รูปภาพ 100%
การทดสอบที่ 2	195 รูปภาพ 70%	55 รูปภาพ 20%	26 รูปภาพ 10%	276 รูปภาพ 100%
การทดสอบที่ 3	249 รูปภาพ 80%	32 รูปภาพ 10%	31 รูปภาพ 10%	312 รูปภาพ 100%
การทดสอบที่ 4	249 รูปภาพ 80%	47 รูปภาพ 15%	16 รูปภาพ 5%	312 รูปภาพ 100%

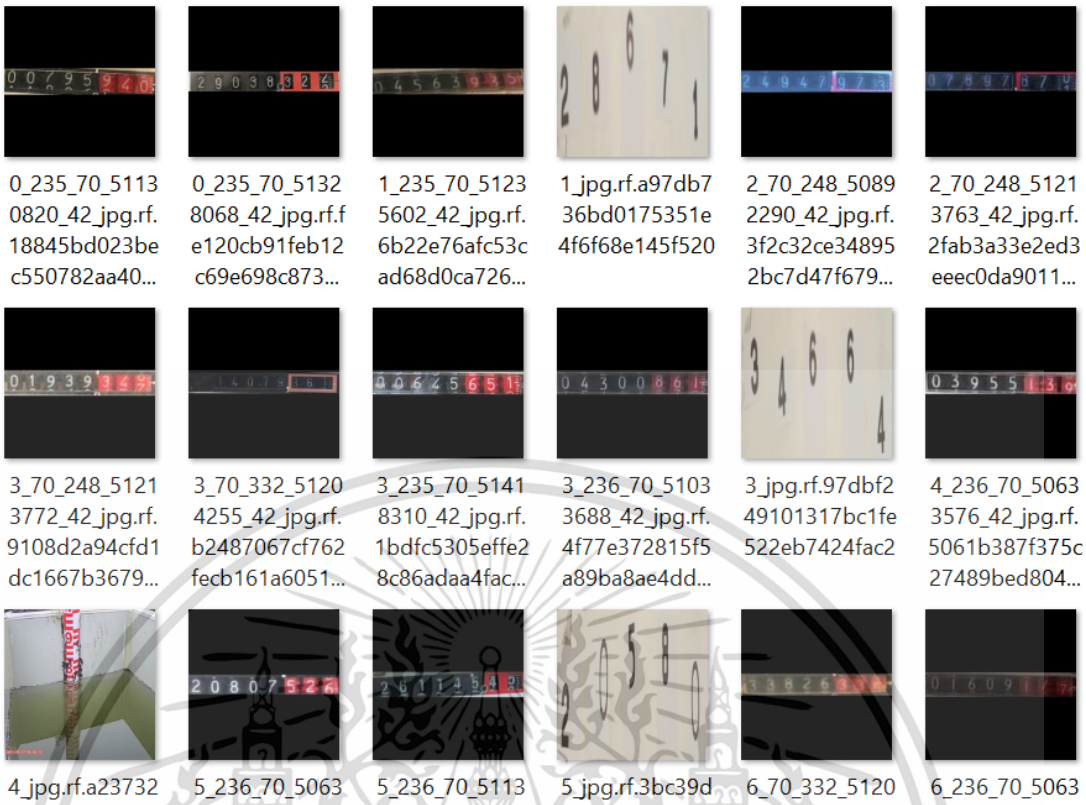
จากตารางที่ 3.1 จะพบว่าการทดสอบที่ 1 และ 2 มีรูปภาพทั้งหมด 276 รูปภาพ และการทดสอบที่ 3 และ 4 มีรูปภาพทั้งหมด 312 รูปภาพ สาเหตุที่จำนวนรูปภาพไม่เท่ากันเพราะว่าได้มีการจัดแต่งรูปภาพให้กับรูปภาพที่ใช้สำหรับการเรียนรู้เท่านั้น โดยจำนวนรูปภาพในส่วนของการเรียนรู้ การทดสอบ และการทดสอบในแต่ละรอบของการทดสอบจะไม่เท่ากันดังตารางที่ 3.1

3.5 รูปภาพที่ใช้ในตัวแบบที่ 2

ในส่วนของการเรียนรู้การอ่านค่าตัวเลขที่อยู่ในรูปภาพ ผู้วิจัยได้ทำการหารูปภาพเพิ่มเติม เนื่องจากรูปภาพที่มีอยู่มีตัวเลขไม่ครบถ้วนเท่าที่ต้องการ จึงสืบค้นข้อมูลเพิ่มเติมจาก Roboflow Universe โดยใช้หลักเกณฑ์ในการหารูปภาพของตัวเลขที่มีลักษณะใกล้เคียงกับตัวเลขที่ปรากฏบนไม้วัดระดับน้ำมากที่สุด โดยชุดข้อมูลที่เลือกนำมาใช้เป็นภาพของตัวเลขที่ถูกพิมพ์ออกมาโดยใช้ฟอนต์ที่ต่างกันออกไป และมีลักษณะใกล้เคียงกับตัวเลขที่ปรากฏบนไม้วัดระดับน้ำ หลังจากนั้น ได้ทำการรวบรวมรูปทั้งหมดที่มีอยู่เพื่อนำมาใช้ในขั้นตอนการเรียนรู้ของตัวแบบให้สามารถอ่านค่าของระดับน้ำได้

นำรูปภาพจากตัวแบบที่ 1 เพิ่มเข้าไปกับรูปภาพในส่วนแรกเพื่อทำให้ชุดข้อมูลสอดคล้องกับสถานการณ์จริงมากยิ่งขึ้น เนื่องจากต้องการให้ตัวแบบสามารถอ่านตัวเลขได้อย่างแม่นยำไม่ว่ารูปภาพจะมีความสว่างไม่เท่ากัน หรือมีสัญญาณรบกวนจากสภาพแวดล้อม เนื่องจากชุดข้อมูลที่มีอยู่มีความหลากหลายอยู่แล้ว จึงไม่ได้มีการจัดแต่งรูปภาพเพิ่มเติม และได้มีการปรับรูปภาพให้มีขนาด 640 x 640 พิกเซลเช่นเดียวกับชุดข้อมูลแรก โดยรูปภาพที่สืบค้นมาเพิ่มและรูปภาพที่มีอยู่เดิมมีทั้งหมด 639 รูปภาพ ดังรูปที่ 3.8

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



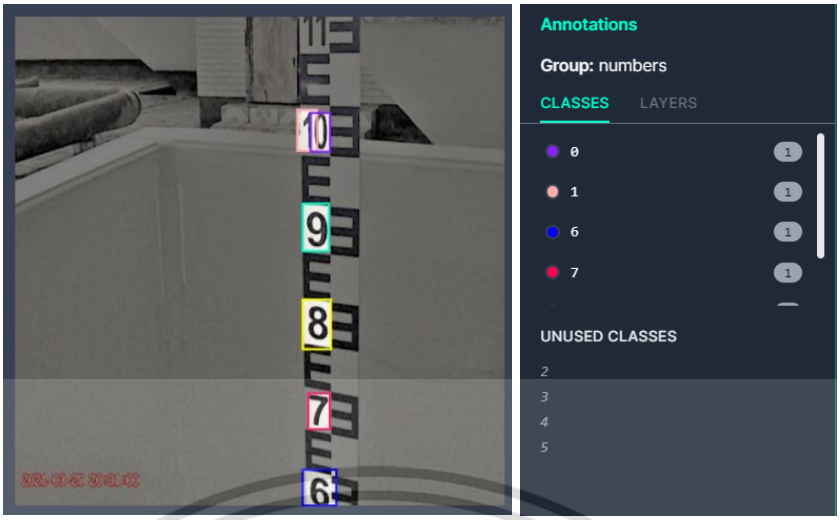
รูปที่ 3.8 รูปภาพตัวอย่างของชุดข้อมูลที่ใช้ในตัวแบบที่ 2

จากรูปที่ 3.8 จะพบว่ารูปภาพที่ใช้สำหรับการเรียนรู้มีลักษณะของตัวเลขคล้ายคลึงกับตัวเลขที่ปรากฏอยู่บนไม้วัดระดับน้ำ โดยตัวเลขที่อยู่ในรูปภาพมาจากหลายแหล่ง ได้แก่ ตัวเลขบนไม้วัดระดับน้ำ ตัวเลขที่ถูกพิมพ์ลงบนกระดาษ และตัวเลขจากมิเตอร์ไฟฟ้า โดยสีของตัวเลขมีทั้งหมด 3 สี ได้แก่ ขาว ดำ และแดง ลักษณะของตัวเลขจะมีความแตกต่างกันออกไปที่ความสูงและความกว้าง ซึ่งเหมาะกับการนำมาใช้เป็นรูปภาพสำหรับการเรียนรู้ในงานวิจัยในครั้งนี้ เนื่องจากกล้องวงจรปิดที่ถ่ายให้เห็นไม้วัดระดับน้ำนั้นมักถูกติดตั้งไว้ในมุมมองที่แตกต่างกันออกไป ทำให้การมองเห็นตัวเลขบนไม้วัดระดับน้ำไม่ได้มองเห็นจากด้านที่ขนานกับกล้องวงจรปิดเพียงมุมเดียวเท่านั้น

3.6 ติดป้ายกำกับให้กับข้อมูลตัวเลข

ติดป้ายกำกับให้กับวัตถุที่อยู่ในรูปภาพ โดยกำหนดให้ตำแหน่งของจุดสี่เหลี่ยมสี่จุดที่ครอบคลุมพื้นที่ของวัตถุที่ต้องการเพื่อนำตำแหน่งของวัตถุไปใช้ในการสอนตัวแบบต่อไป โดยส่วนที่ต้องการคือตำแหน่งของตัวเลข ทำการติดป้ายกำกับว่าตัวเลขที่ปรากฏอยู่นั้นเป็นตัวเลขใด เพื่อให้ตัวแบบสามารถอ่านค่าตัวเลขที่ปรากฏบนรูปภาพได้ ดังรูปที่ 3.9

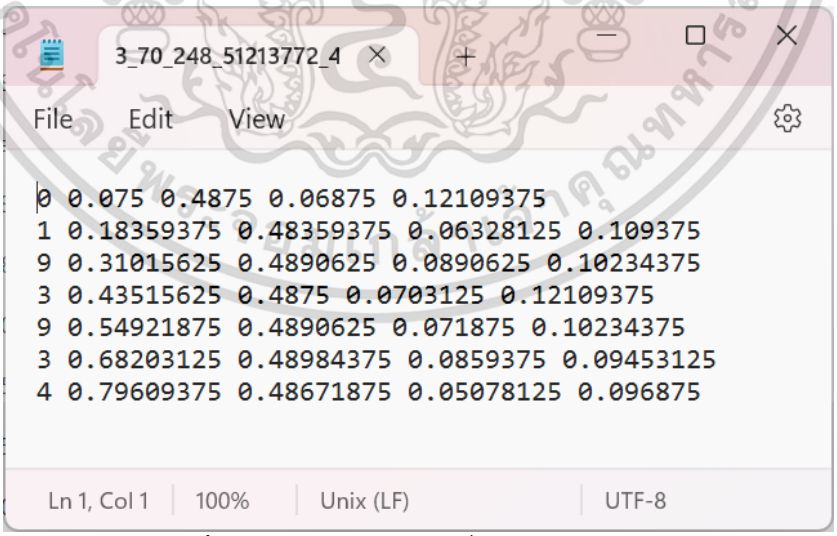
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.9 ตัวอย่างการติดป้ายกำกับให้กับข้อมูลตัวเลข

จากรูปที่ 3.9 จะพบว่ารูปภาพ 1 ภาพประกอบไปด้วยตัวเลขหลายตัวเลข การติดป้ายกำกับให้กับตัวเลขจะมองตัวเลขทีละหลักเท่านั้น จึงทำให้รูปที่ 3.9 มีตัวเลขอยู่ทั้งหมด 6 ตัวเลข ได้แก่ เลข 0, 1, 6, 7, 8 และ 9 ซึ่งการติดป้ายกำกับจะแบ่งเลข 10 ออกเป็น เลข 1 และเลข 0 ส่วนด้านขวาของรูปที่ 3.9 แสดงจำนวนตัวเลขที่พบในรูปภาพ และแสดงตัวเลขที่ไม่พบในรูปภาพ

เนื่องจากมีวัตถุที่ต้องการจำนวน 10 วัตถุได้แก่ เลข 1 ถึงเลข 9 ชุดข้อมูลสำหรับการติดป้ายกำกับที่ได้ทำการแบ่งออกเป็น การเรียน การทดสอบ และการตรวจสอบ จึงถูกเก็บในลักษณะของจุดสี่จุดล้อมรอบไม้วัดระดับน้ำ โดยรูปภาพ 1 รูปสามารถประกอบไปด้วยหลายวัตถุได้ดังรูปที่ 3.10 โดยเลขในตำแหน่งแรกในรูปภาพเป็นค่าของตัวเลขที่ระบุได้ และเลขในตำแหน่งที่ 2, 3, 4 และ 5 คือ x1, x2, y1 และ y2 ตามลำดับ



รูปที่ 3.10 ตำแหน่งของจุดสี่จุดล้อมรอบตัวเลข

จากรูปที่ 3.10 จะพบว่าเลขที่อยู่ในตำแหน่งแรกคือค่าจริงของตัวเลขที่ปรากฏ เลขในตำแหน่งที่ 2, 3, 4 และ 5 คือ x1, x2, y1 และ y2 ตามลำดับ ซึ่ง x1 และ x2 คือจุดเริ่มต้นและจุดสิ้นสุดของแกน x y1 และ y2 คือจุดเริ่มต้นและจุดสิ้นสุดของแกน y โดยจุดที่ x=0 และ y=0 คือ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

มุมบนด้านซ้ายของรูปภาพ หาก x เข้าใกล้ 0 หมายความว่าตำแหน่งของตัวเลขอยู่ทางซ้ายของรูปภาพ และหาก y เข้าใกล้ 0 หมายความว่าตำแหน่งของตัวเลขอยู่ทางด้านบนของรูปภาพ เช่นเดียวกับขั้นตอนการติดป้ายกำกับของตัวแบบที่ 1

จากรูปที่ 3.10 สามารถตีความได้ว่ารูปภาพประกอบไปด้วยตัวเลขทั้งหมด 7 ตัวเลข เลข 0, 1 และ 4 มีอย่างละ 1 ตัว และตัวเลข 3 และ 9 มีอย่างละ 2 ตัว

3.7 การเรียนรู้ของตัวแบบที่ 2

ได้มีการทดลองอ่านค่าตัวเลขจากรูปที่ได้ 2 วิธี ได้แก่ การใช้ไลบรารี่ บน python ที่มีชื่อว่า pytesseract และการนำรูปภาพที่ผ่านการติดป้ายกำกับให้ YOLOv8m เรียนรู้ พบว่า YOLOv8m สามารถอ่านค่าตัวเลขได้ดีและแม่นยำกว่า pytesseract เป็นอย่างมาก จึงเลือกใช้ตัวแบบนี้ โดยแบ่งข้อมูลสำหรับการเรียนรู้ การทดสอบหามาตรวัดประสิทธิภาพและการทดสอบหลังจากที่ได้ตัวแบบที่ดีที่สุดมาแล้วว่าตัวแบบสามารถทำงานได้ดีแค่ไหนกับข้อมูลที่ไม่เคยเห็นมาก่อน ได้มีการแบ่งข้อมูลสำหรับส่วนต่าง ๆ ดังตารางที่ 3.2 เพื่อให้ได้การแบ่งส่วนที่ทำให้ตัวแบบมีประสิทธิภาพมากที่สุด และทำการกำหนดพารามิเตอร์โดยกำหนดให้ epoch = 150, Image Size = 640 และ Patience = 100

ตารางที่ 3.2 การแบ่งข้อมูลสำหรับการเรียนรู้ของตัวแบบที่ 2

รอบการทดสอบ	การเรียนรู้	การทดสอบ	การตรวจสอบ	รวม
การทดสอบที่ 1	447 รูปภาพ 70%	96 รูปภาพ 15%	96 รูปภาพ 15%	639 รูปภาพ 100%
การทดสอบที่ 2	447 รูปภาพ 70%	127 รูปภาพ 20%	65 รูปภาพ 10%	639 รูปภาพ 100%
การทดสอบที่ 3	511 รูปภาพ 80%	64 รูปภาพ 10%	64 รูปภาพ 10%	639 รูปภาพ 100%
การทดสอบที่ 4	511 รูปภาพ 80%	95 รูปภาพ 15%	33 รูปภาพ 5%	639 รูปภาพ 100%

ชุดข้อมูลที่สองนั้นมีจำนวน class ทั้งหมด 10 class ได้แก่ 1, 2, 3, 4, 5, 6, 7, 8, 9 และ 0 ไม่เหมือนกับตัวแบบที่ 1 ที่มีอยู่เพียง 1 class เท่านั้น ภาพที่ใช้ในการสอน 1 ภาพนั้นจะประกอบไปด้วยตัวเลขหลายตัวเช่นเดียวกับตัวเลขที่ปรากฏบนไม้วัดระดับน้ำ จำนวน class ในการติดป้ายกำกับให้กับชุดข้อมูลรูปภาพในแต่ละการทดสอบมีรายละเอียดดังตารางที่ 3.3 ถึงตารางที่ 3.6

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.3 รายละเอียดการติดป้ายกำกับให้กับชุดข้อมูลที่สองในการทดสอบที่ 1

ตัวเลข	การเรียนรู้ (70%)	การทดสอบ (15%)	การตรวจสอบ (15%)
0	613	97	158
1	270	66	62
2	276	41	68
3	255	51	68
4	235	45	60
5	279	51	59
6	281	48	61
7	250	50	63
8	307	58	70
9	301	69	72

ตารางที่ 3.4 รายละเอียดการติดป้ายกำกับให้กับชุดข้อมูลที่สองในการทดสอบที่ 2

ตัวเลข	การเรียนรู้ (70%)	การทดสอบ (20%)	การตรวจสอบ (10%)
0	613	172	83
1	270	85	43
2	276	71	38
3	255	79	40
4	235	73	32
5	279	71	39
6	281	69	40
7	250	71	42
8	307	85	43
9	301	93	48

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.5 รายละเอียดการติดป้ายกำกับให้กับชุดข้อมูลที่สองในการทดสอบที่ 3

ตัวเลข	การเรียนรู้ (80%)	การทดสอบ (10%)	การตรวจสอบ (10%)
0	716	67	85
1	317	43	38
2	328	21	36
3	306	23	45
4	285	23	32
5	337	18	34
6	324	25	41
7	293	29	41
8	359	35	41
9	359	39	44

ตารางที่ 3.6 รายละเอียดการติดป้ายกำกับให้กับชุดข้อมูลที่สองในการทดสอบที่ 4

ตัวเลข	การเรียนรู้ (80%)	การทดสอบ (15%)	การตรวจสอบ (5%)
0	716	117	35
1	317	64	17
2	328	45	12
3	306	53	15
4	285	46	9
5	337	43	9
6	324	54	12
7	293	54	16
8	359	58	18
9	359	63	20

3.8 การหาอัตราส่วนของรูปภาพ

หลังจากได้ x1, x2, y1 และ y2 ของตัวเลขจากตัวแบบที่ 2 แล้ว ทำการหาตัวเลขที่มีตำแหน่งต่ำที่สุด หรือตัวเลขที่มีค่า y2 มากที่สุดเพื่อใช้เป็นค่าของตัวเลขหน้าจุดทศนิยม หลังจากนั้น ทำการหาค่าหลังจุดทศนิยมของตัวเลขเพื่อประเมินค่าว่าระดับน้ำอยู่ในช่วงใด โดยสมการที่ใช้เป็นดังสมการที่ 3.1

$$y2 - y1 / h - y2 \qquad (3.1)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.9 สร้างหน้าต่างสำหรับการใช้งาน

สร้างหน้าต่างสำหรับการมอนิเตอร์ โดยใช้ library ใน python ที่ชื่อว่า tkinter ในการแสดงผลลัพธ์ ดังรูปที่ โดยหน้าต่างในการมอนิเตอร์ประกอบด้วย

1. กราฟเส้นแสดงข้อมูลของระดับน้ำย้อนหลัง 10 ชั่วโมง
2. ภาพที่ได้จากกล้องวงจรปิดที่ได้มีการดึงภาพมาแสดง ซึ่งในตัวอย่างได้นำวิดีโอจากกล้องจรปิดมาใช้สำหรับการแสดงผล
3. ค่าระดับน้ำในปัจจุบันที่ผ่านขั้นตอนการอ่านค่าระดับน้ำที่ได้กล่าวมาในงานวิจัย
4. สถานะของการปิดและเปิดของประตู สำหรับการแจ้งเตือนหากระดับน้ำมีค่ามากเกินกว่าที่กำหนด แต่ประตูปล่อยน้ำยังคงเปิดอยู่

ข้อมูลระดับน้ำจะถูกบันทึกลงฐานข้อมูลที่เตรียมไว้ทุก ๆ 1 ชั่วโมง โดยโครงสร้างของฐานข้อมูลมีรายละเอียดดังตารางที่ 3.7

ตารางที่ 3.7 โครงสร้างของการจัดเก็บข้อมูลระดับน้ำลงบนฐานข้อมูล

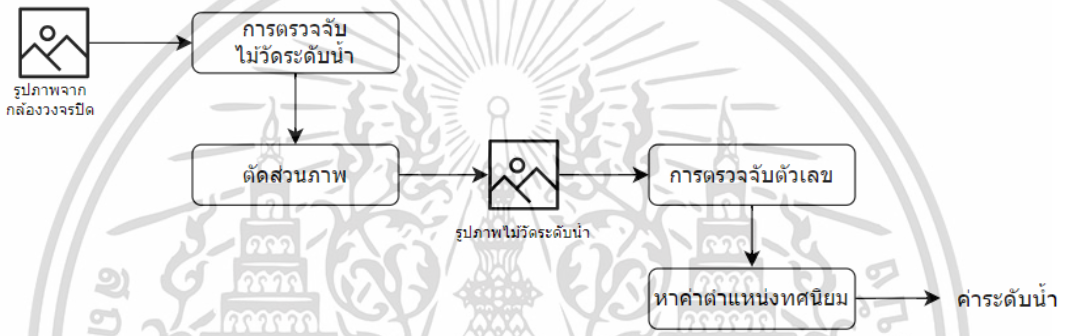
ชื่อหัวตาราง	ชนิดของข้อมูล	รายละเอียดข้อมูล	ตัวอย่างข้อมูล
ID	text	รหัสในการจัดเก็บข้อมูล	1
location	text	สถานที่ติดตั้งไม้วัดระดับน้ำ	Wat Sri Bunruang Water Gate
latitude	numeric	ละติจูด	13.7619139
longitude	numeric	ลองจิจูด	100.6198322
time_stamp	timestamp	การประทับเวลา	29/4/2023 15:00:00
water_level	numeric	ค่าระดับน้ำ	3.5
door_status	text	สถานการณ์ปิดเปิดของประตู (open, close)	open
street_view	text	ลิงก์ในการเข้าถึง Google Street View เพื่อแสดงให้เห็นภาพในแม่น้ำสำหรับการนำไปวิเคราะห์ภาพถ่ายการเปลี่ยนแปลงของระดับน้ำในอนาคต	https://www.google.com/maps/@13.7619139,100.6198322,446m/data=!3m1!1e3

จากตารางที่ 3.7 จะพบว่าระบบได้มีการกำหนดให้เก็บค่าระดับน้ำทุก ๆ 1 ชั่วโมง และทำการจัดเก็บรายละเอียดอื่น ๆ เพิ่มเติม เช่น สถานที่ที่มีการติดตั้งไม้วัดระดับน้ำ เวลาการบันทึกข้อมูล สถานการณ์เปิดและปิดประตูระบายน้ำ และลิงก์ไปที่ Google Street View เพื่อให้มีข้อมูลเพียงพอสำหรับการทำนายค่าระดับน้ำในอนาคต โดยข้อมูลรูปภาพจาก Google Street View จะทำให้เห็นถึงระดับน้ำของแม่น้ำในคลองได้จริง จึงสามารถนำข้อมูลในส่วนนี้ไปใช้ในการทำนายค่าระดับน้ำหรืองานวิจัยที่เกี่ยวข้องต่อไป

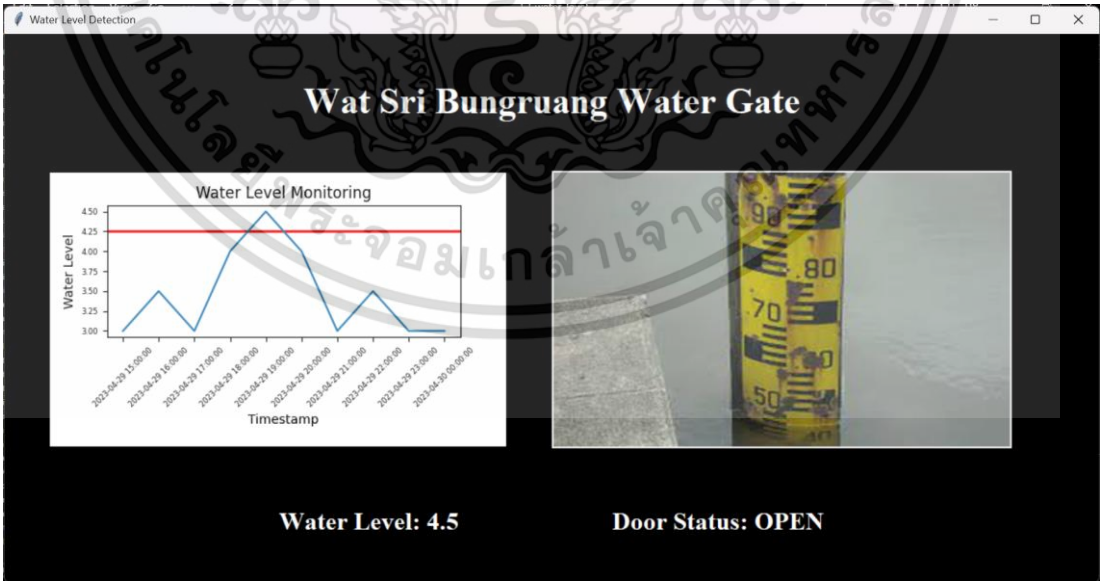
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.10 การทำงานของระบบ

เมื่อมีรูปภาพนำเข้าจากกล้องวงจรปิด (CCTV) เกิดขึ้น ระบบจะกำหนดให้รูปภาพเป็นรูปภาพนำเข้าของตัวแบบที่ 1 เพื่อให้ตัวแบบที่ 1 ตรวจจับตำแหน่งของไม้วัดระดับน้ำที่ปรากฏในรูปภาพ หลังจากนั้น ระบบจะนำผลลัพธ์ที่ได้จากตัวแบบที่ 1 ซึ่งเป็นการตัดภาพในส่วนของไม้วัดระดับน้ำไปใช้เป็นรูปภาพนำเข้าของตัวแบบที่ 2 เพื่อให้ตัวแบบที่ 2 ตรวจจับตัวเลขที่อยู่ในตำแหน่งที่ใกล้เคียงกับระดับน้ำมากที่สุด จากนั้นระบบจะทำการอ่านค่าตัวเลขนั้นและนำขนาดของตัวเลขที่ได้และระยะห่างของตัวเลขกับขอบด้านล่างสุดของรูปภาพมาเปรียบเทียบมาตราส่วนเพื่อหาตัวเลขหลังจุดทศนิยม พร้อมกับบันทึกค่าระดับน้ำสุดท้ายที่ได้ลงบนฐานข้อมูล การทำงานของระบบทั้งหมดเป็นไปดังรูปที่ 3.12 และการแสดงผลของการอ่านค่าระดับน้ำจากกล้องวงจรปิด เป็นไปดังรูปที่ 3.13



รูปที่ 3.11 การทำงานของระบบการอ่านค่าระดับน้ำจากกล้องวงจรปิด (CCTV)



รูปที่ 3.12 หน้าต่างสำหรับการมอนิเตอร์ระดับน้ำ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 3.12 จะพบว่าหน้าต่างสำหรับการแสดงผลประกอบไปด้วยชื่อของสถานที่ที่มีการติดตั้งไม้วัดระดับน้ำและกล้องวงจรปิด รูปภาพทางด้านขวาคือรูปภาพจากกล้องวงจรปิด เมื่อมีการอ่านค่าระดับน้ำเกิดขึ้น ค่าระดับน้ำจะประกฏตรงมุมซ้ายล่างของรูปภาพ หลังจากนั้น ระบบจะบันทึกค่าระดับน้ำลงไปในฐานข้อมูล ระบบจะดึงข้อมูลค่าระดับน้ำย้อนหลัง 10 ชั่วโมงมาแสดงบนกราฟเส้นทางด้านซ้ายของรูปภาพ ซึ่งผู้ใช้งานสามารถกำหนดระดับน้ำที่แต่ละจุดสามารถรับได้ โดยระดับน้ำสูงสุดที่สามารถรับได้จะถูกแสดงด้วยเส้นสีแดงในรูปที่ 3.12 หากระดับน้ำมีค่ามากกว่าเกณฑ์ ระบบจะส่งการแจ้งเตือนไปที่ผู้ดูแล ณ จุดที่มีการติดตั้งไม้วัดระดับน้ำเพื่อดำเนินการตามขั้นตอนต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4
ผลการวิจัยและการอภิปรายผล

งานวิจัยนี้มีวัตถุประสงค์เพื่ออ่านค่าระดับน้ำจากรูปภาพจากกล้องวงจรปิด (CCTV) มีการสืบค้นรูปภาพไม้วัดระดับน้ำจากอินเทอร์เน็ตเพื่อนำมาเป็นรูปภาพนำเข้าของงานวิจัย ได้มีการทำการติดป้ายกำกับรูปภาพเพื่อใช้สำหรับการเรียนรู้ และใช้ตัวแบบที่ผ่านการเรียนรู้มาแล้ว (YOLOv8) ในการทดสอบ 2 ตัวแบบ ตัวแบบที่หนึ่งใช้สำหรับการตรวจจับตำแหน่งของไม้วัดระดับน้ำ และตัวแบบที่สองใช้สำหรับการอ่านค่าตัวเลขที่ปรากฏบนไม้วัดระดับน้ำ ได้มีการทดสอบโดยการแบ่งข้อมูลสำหรับการเรียนรู้ การทดสอบ และการตรวจสอบออกเป็นหลากหลายรูปแบบเพื่อให้ได้ตัวแบบที่มีประสิทธิในการอ่านค่าระดับน้ำมากที่สุด โดยเปรียบเทียบประสิทธิภาพด้วยค่าความเที่ยง ค่าเรียกคืน และค่าเฉลี่ยความเที่ยง โดยผลการวิเคราะห์ได้เสนอตามลำดับ ดังนี้

4.1 ผลการวิเคราะห์การประเมินประสิทธิภาพของตัวแบบที่ 1

จากการทดสอบการทำงานของตัวแบบที่ 1 พบว่า การแบ่งข้อมูลสำหรับการเรียนรู้ การทดสอบ และการตรวจสอบไม่เท่ากันทำให้ได้ผลลัพธ์ที่แตกต่างกันออกไป ดังตารางที่ 4.1

ตารางที่ 4.1 ผลลัพธ์ของตัวแบบที่ 1

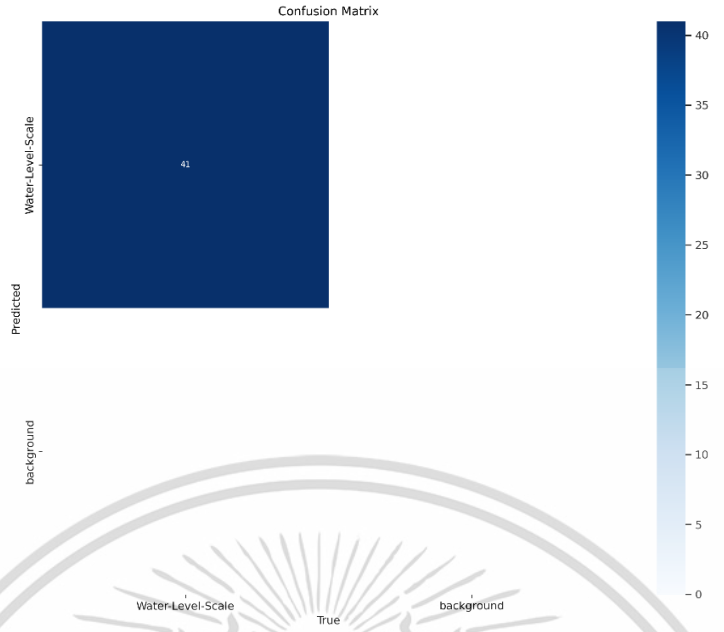
การทดสอบ	mAP50	mAP50-95	ค่าความเที่ยง	ค่าเรียกคืน
การทดสอบที่ 1 (70:15:15)	99.3	83.8	97.3	**100**
การทดสอบที่ 2 (70:20:10)	95.8	78.9	89.7	95.3
การทดสอบที่ 3 (80:10:10)	**99.4**	**86.8**	96.9	97.1
การทดสอบที่ 4 (80:15:5)	**99.4**	86.1	**97.8**	**100**

จากตารางที่ 4.1 จะพบว่า การทดสอบที่ 1 มีค่าเรียกคืนสูงที่สุด การทดสอบที่ 3 มีค่าเฉลี่ยความเที่ยงเมื่อ IoU เท่ากับ 50-95 และค่าเฉลี่ยควาทเที่ยงเมื่อ IoU เท่ากับ 50 มากที่สุด การทดสอบที่ 4 มีค่าเฉลี่ยความเที่ยงเมื่อ IoU เท่ากับ 50 ค่าความเที่ยง และค่าเรียกคืนมากที่สุด

4.1.1 การประเมินประสิทธิภาพของการทดสอบที่ 1 ตัวแบบที่ 1

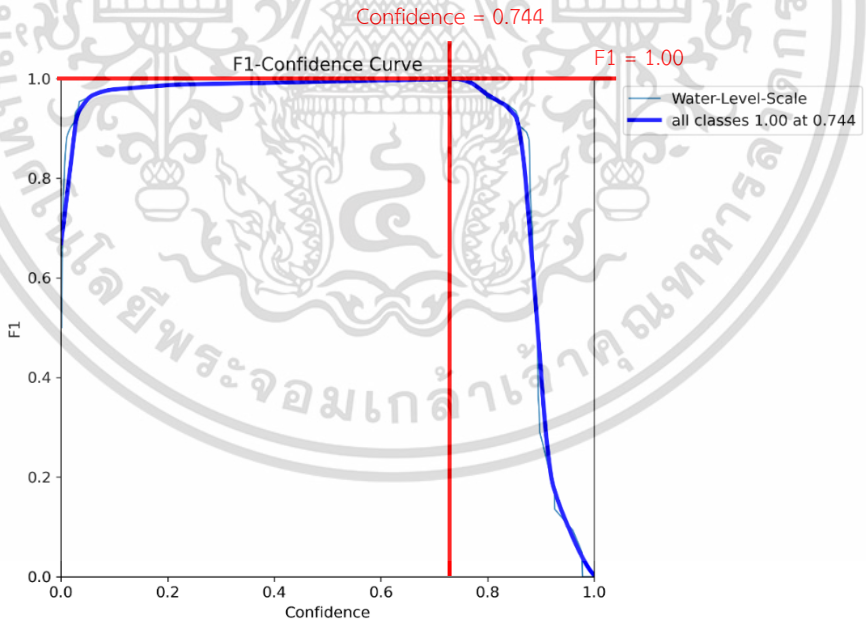
การแบ่งข้อมูลสำหรับการทดสอบที่ 1 ได้มีการแบ่งข้อมูลออกเป็น 70:15:15 ทำให้ได้ mAP50-95 อยู่ที่ 83.8 ดังตารางที่ 4.1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.1 ผลลัพธ์ของเมทริกซ์ความสับสนของการทดสอบที่ 1 ตัวแบบที่ 1

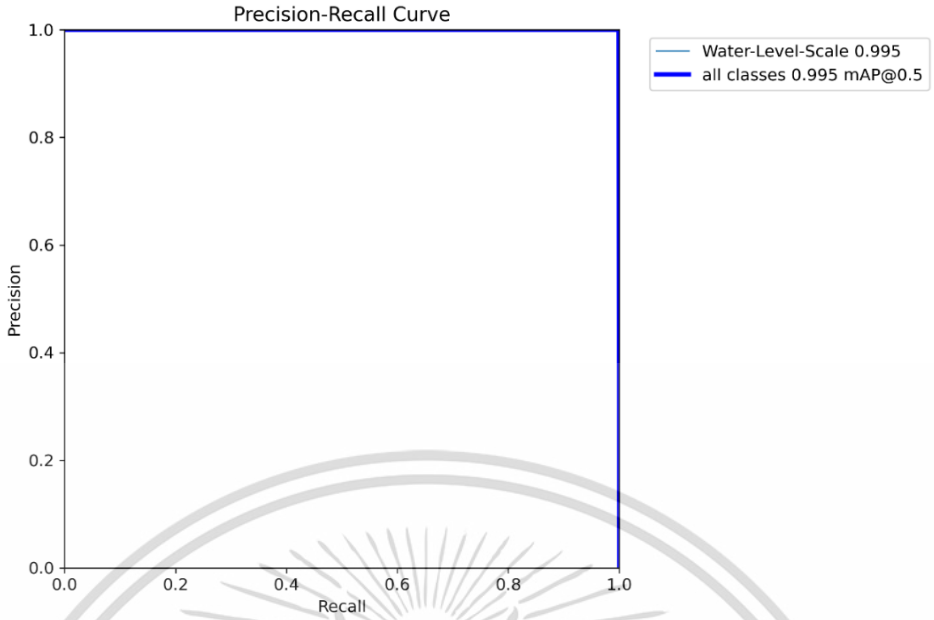
จากรูปที่ 4.1 เมทริกซ์ความสับสนแสดงให้เห็นว่าตัวแบบสามารถทำนายรูปภาพที่ใช้สำหรับการทดสอบได้ถูกต้องทั้ง 41 รูป



รูปที่ 4.2 เส้นโค้งความเชื่อมั่นคะแนน F1 ของการทดสอบที่ 1 ตัวแบบที่ 1

จากรูปที่ 4.2 จะพบว่าการทดสอบที่ 1 ของตัวแบบที่ 1 มีค่าคะแนน F1 เป็น 100 เมื่อค่าความเชื่อมั่นเท่ากับ 0.744 และค่าความเชื่อมั่นที่ 0.744 ทำให้ค่าเรียกคืนและค่าความเที่ยงเหมาะสมที่สุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

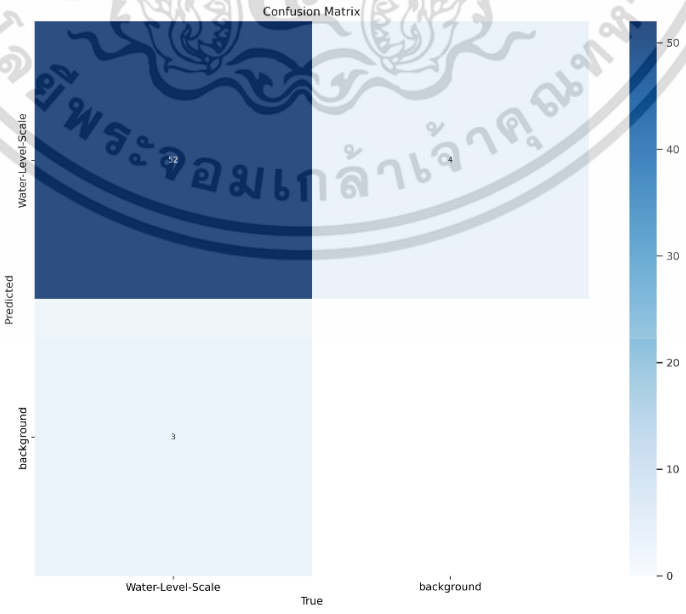


รูปที่ 4.3 เส้นโค้งค่าความเที่ยงและค่าเรียกคืนของการทดสอบที่ 1 ตัวแบบที่ 1

จากรูปที่ 4.3 จะพบว่าเส้นโค้งของค่าความเที่ยงและค่าเรียกคืนเท่ากับ 0.995 เมื่อค่าเฉลี่ยความเที่ยงที่ IoU เท่ากับ 0.5

4.1.2 การประเมินประสิทธิภาพของการทดสอบที่ 2 ตัวแบบที่ 1

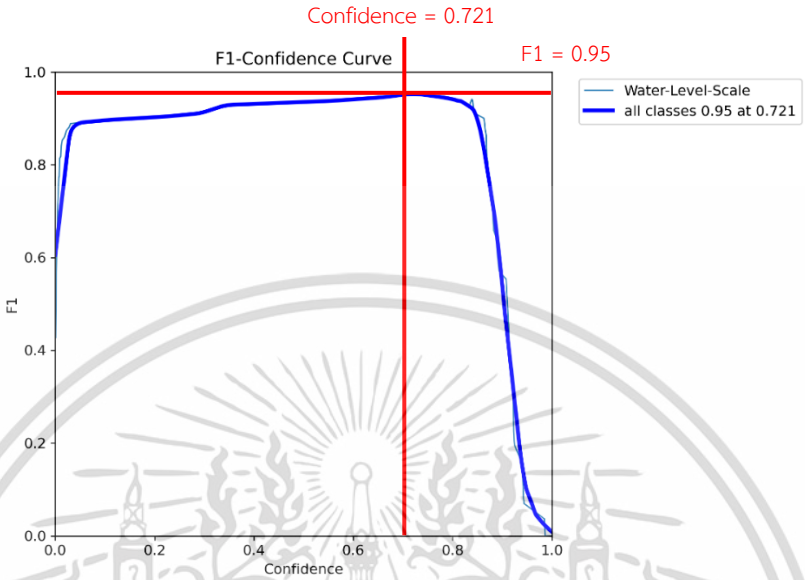
ข้อมูลสำหรับการทดสอบที่ 2 ได้แบ่งออกเป็น 70:20:10 ทำให้ได้ mAP50-95 อยู่ที่ 78.9 ดังตารางที่ 4.1



รูปที่ 4.4 ผลลัพธ์ของเมทริกซ์ความสับสนของการทดสอบที่ 2 ตัวแบบที่ 1

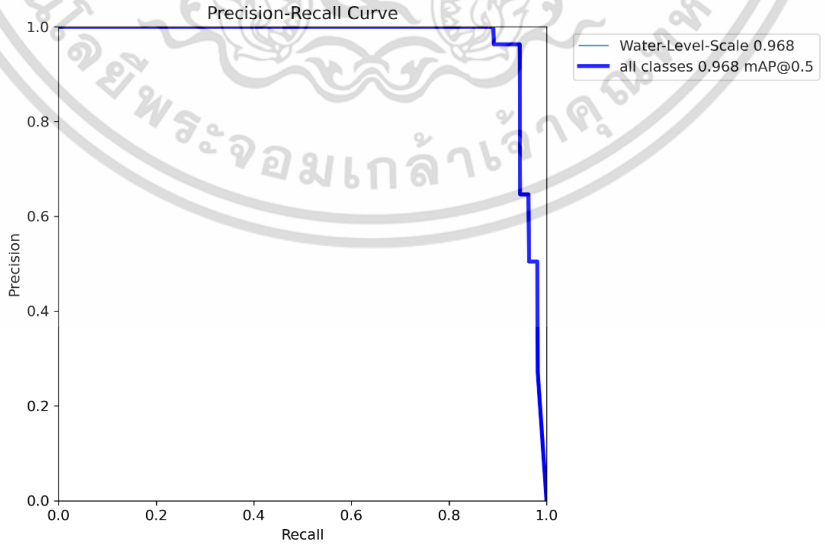
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 4.4 เมทริกซ์ความสับสนแสดงให้เห็นว่าตัวแบบสามารถทำนายรูปภาพที่ใช้สำหรับการทดสอบได้ถูกต้องทั้งหมด 52 รูป มีรูปพื้นหลังที่ทำนายว่าเป็นไม้วัดระดับน้ำทั้งหมด 4 รูป และมีรูปไม้ระดับที่ทำนายว่าเป็นภาพพื้นหลังทั้งหมด 3 รูป



รูปที่ 4.5 เส้นโค้งความเชื่อมั่นคะแนน F1 ของการทดสอบที่ 2 ตัวแบบที่ 1

จากรูปที่ 4.5 การทดสอบที่ 2 ของตัวแบบที่ 1 มีค่าคะแนน F1 เป็น 0.95 เมื่อค่าความเชื่อมั่นเท่ากับ 0.721 และค่าความเชื่อมั่นที่ 0.721 เป็นค่าที่ทำให้ค่าความเที่ยงกับค่าเรียกคืนเหมาะสมที่สุด



รูปที่ 4.6 เส้นโค้งค่าความเที่ยงและค่าเรียกคืนของการทดสอบที่ 2 ตัวแบบที่ 1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

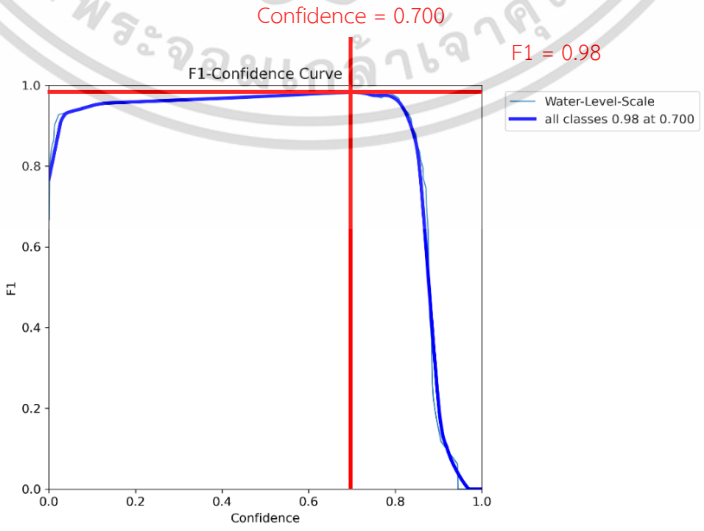
จากรูปที่ 4.6 จะพบว่าเส้นโค้งของค่าความเที่ยงและค่าเรียกคืนเท่ากับ 0.968 เมื่อค่าเฉลี่ยความเที่ยงที่ IoU เท่ากับ 0.5

4.1.3 การประเมินประสิทธิภาพของการทดสอบที่ 3 ตัวแบบที่ 1

การแบ่งข้อมูลสำหรับการทดสอบที่ 3 แบ่งออกเป็น 80:10:10 ทำให้ได้ mAP50-95 อยู่ที่ 86.8 ดังตารางที่ 4.1

รูปที่ 4.7 ผลลัพธ์ของเมทริกซ์ความสับสนของการทดสอบที่ 3 ตัวแบบที่ 1

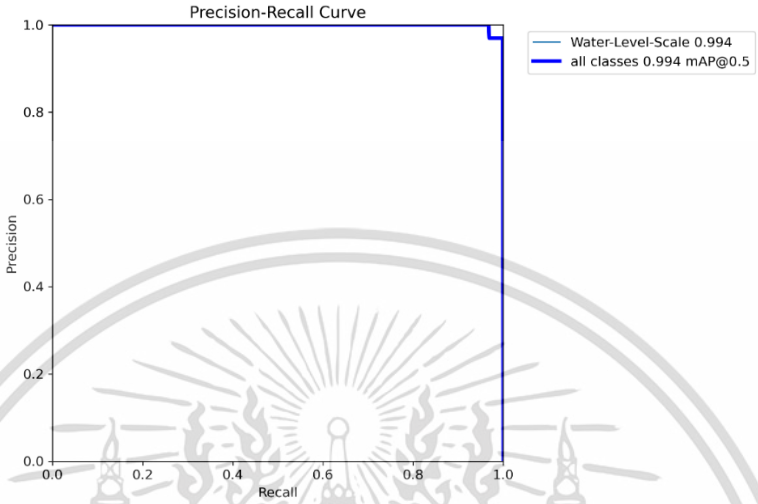
จากรูปที่ 4.7 เมทริกซ์ความสับสนแสดงให้เห็นว่าตัวแบบสามารถทำนายรูปภาพที่ใช้สำหรับการทดสอบได้ถูกต้องทั้งหมด 32 รูป มีรูปพื้นหลังที่ทำนายว่าเป็นไม้วัดระดับน้ำทั้งหมด 2 รูป



รูปที่ 4.8 เส้นโค้งความเชื่อมั่นคะแนน F1 ของการทดสอบที่ 3 ตัวแบบที่ 1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 4.8 จะพบว่าการทดสอบที่ 3 ของตัวแบบที่ 1 มีค่าคะแนน F1 เป็น 0.98 เมื่อค่าความเชื่อมั่นเท่ากับ 0.700 และค่าความเชื่อมั่นเท่ากับ 0.700 เป็นค่าที่ทำให้ค่าความเที่ยงและค่าเรียกคืนเหมาะสมที่สุด



รูปที่ 4.9 เส้นโค้งค่าความเที่ยงและค่าเรียกคืนของการทดสอบที่ 3 ตัวแบบที่ 1

จากรูปที่ 4.9 จะพบว่าเส้นโค้งของค่าความเที่ยงและค่าเรียกคืนเท่ากับ 0.994 เมื่อค่าเฉลี่ยความเที่ยงที่ IoU เท่ากับ 0.5

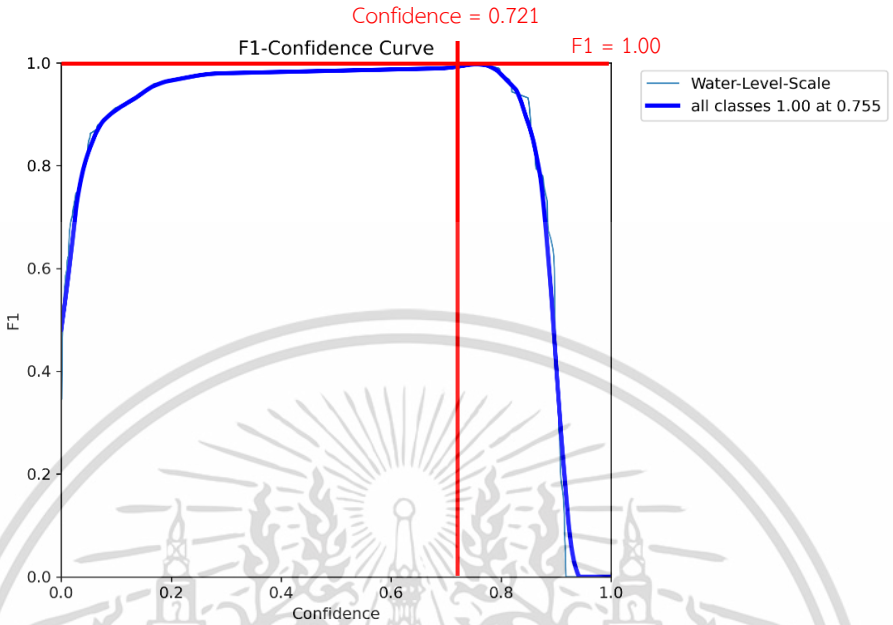
4.1.4 การประเมินประสิทธิภาพของการทดสอบที่ 4 ตัวแบบที่ 1

การแบ่งข้อมูลสำหรับการทดสอบที่ 4 แบ่งออกเป็น 80:15:5 ทำให้ได้ mAP50-95 อยู่ที่ 86.1 ดังตารางที่ 4.1

รูปที่ 4.10 ผลลัพธ์ของเมทริกซ์ความสับสนของการทดสอบที่ 4 ตัวแบบที่ 1

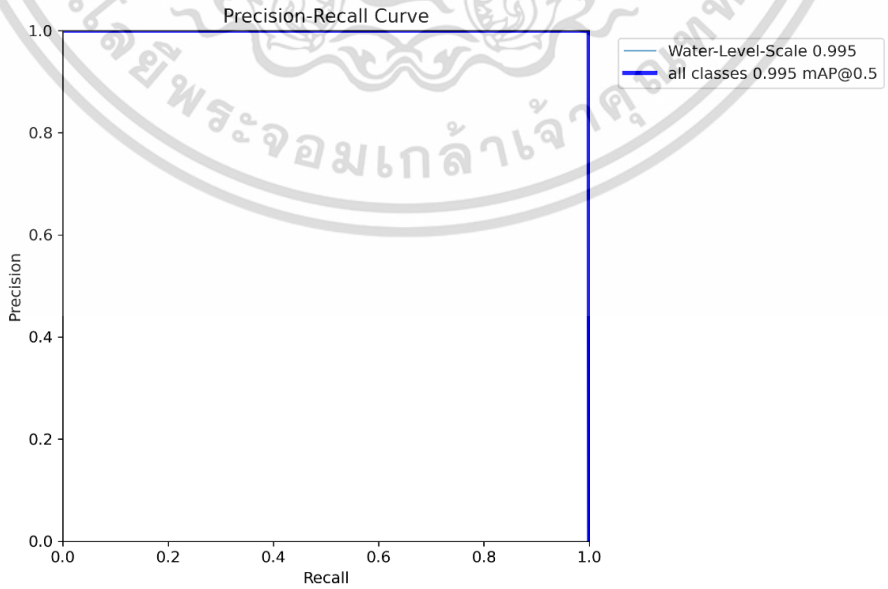
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 4.10 เมทริกซ์ความสับสนแสดงให้เห็นว่าตัวแบบสามารถทำนายรูปภาพที่ใช้สำหรับการทดสอบได้ถูกต้องทั้งหมด 47 รูป มีรูปพื้นหลังที่ทำนายว่าเป็นไม้วัดระดับน้ำทั้งหมด 2 รูป



รูปที่ 4.11 เส้นโค้งความเชื่อมั่นคะแนน F1 ของการทดสอบที่ 4 ตัวแบบที่ 1

จากรูปที่ 4.11 จะพบว่าการทดสอบที่ 4 ของตัวแบบที่ 1 มีค่าคะแนน F1 เป็น 1.00 เมื่อค่าความเชื่อมั่นเท่ากับ 0.755 และค่าความเชื่อมั่นเท่ากับ 0.755 ทำให้ค่าความเที่ยงและค่าเรียกคืนมีความเหมาะสมที่สุด



รูปที่ 4.12 เส้นโค้งค่าความเที่ยงและค่าเรียกคืนของการทดสอบที่ 4 ตัวแบบที่ 1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 4.12 จะพบว่าเส้นโค้งของค่าความเที่ยงและค่าเรียกคืนเท่ากับ 0.955 เมื่อค่าเฉลี่ยความเที่ยงที่ IoU เท่ากับ 0.5

4.1.5 ผลการวิเคราะห์ประสิทธิภาพของตัวแบบที่ 1

ค่าเฉลี่ยความเที่ยงเมื่อ IoU เท่ากับ 0.5 ค่าความเที่ยง และค่าเรียกคืนของการทดสอบที่ 4 หรือการแบ่งข้อมูลออกเป็น 80:15:5 มีค่าเท่ากับ 99.4, 97.8 และ 100 ตามลำดับ ซึ่งเป็นการทดสอบที่มีประสิทธิภาพดีที่สุดเมื่อเทียบกับการทดสอบอื่น ๆ จึงเลือกใช้ตัวแบบที่ได้จากการทดสอบที่ 4

เมื่อทดลองนำตัวแบบมาใช้กับรูปภาพที่ไม่เคยใช้ในการเรียนรู้ การทดสอบ หรือการตรวจสอบมาก่อน พบว่า ตัวแบบสามารถหาตำแหน่งของวัตถุได้ ดังรูปที่ 4.13 และแสดงผลลัพธ์ดังรูปที่ 4.14

รูปที่ 4.13 การตรวจจับวัตถุของการทดสอบที่ 3 ของตัวแบบที่ 1

จากรูปที่ 4.13 จะพบว่าตัวแบบสามารถหาตำแหน่งของไม้วัดระดับน้ำได้อย่างถูกต้อง โดยสามารถนำกรอบสี่เหลี่ยมสีน้ำเงินมาล้อมรอบวัตถุที่ต้องการได้


```
{'data': [{'box': {'x1': 248.0670928955078,
    'x2': 322.58367919921875,
    'y1': 1.149383544921875,
    'y2': 309.06280517578125},
   'class': 0,
   'confidence': 0.8203278183937073,
   'name': 'Water-Level-Scale'}],
 'message': 'Inference complete.',
 'success': True}
```


รูปที่ 4.14 ผลลัพธ์ของตัวแบบที่ 1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 4.14 จะพบว่า x1 และ x2 คือจุดเริ่มต้นและสิ้นสุดของแกน x ที่มีวัตถุอยู่ตรงกลาง y1 และ y2 คือจุดเริ่มต้นและสิ้นสุดของแกน y ที่มีวัตถุอยู่ตรงกลางเช่นเดียวกัน class คือเป้าหมายของงานวิจัย หรือไม้วัดระดับน้ำ ค่า confidence คือค่าความเชื่อมั่นที่ตัวแบบสามารถทำนายได้ว่ากรอบ 4 เหลี่ยมนี่ล้อมรอบไม้วัดระดับน้ำได้อย่างถูกต้องเพียงใด

4.2 ผลการวิเคราะห์การประเมินประสิทธิภาพของตัวแบบที่ 2

จากการทดสอบการทำงานของตัวแบบที่ 2 พบว่า การแบ่งข้อมูลสำหรับการทดสอบต่าง ๆ ทำให้ได้ค่าประสิทธิที่แตกต่างกันออกไป ดังตารางที่ 4.2

ตารางที่ 4.2 ผลลัพธ์ของตัวแบบที่ 2

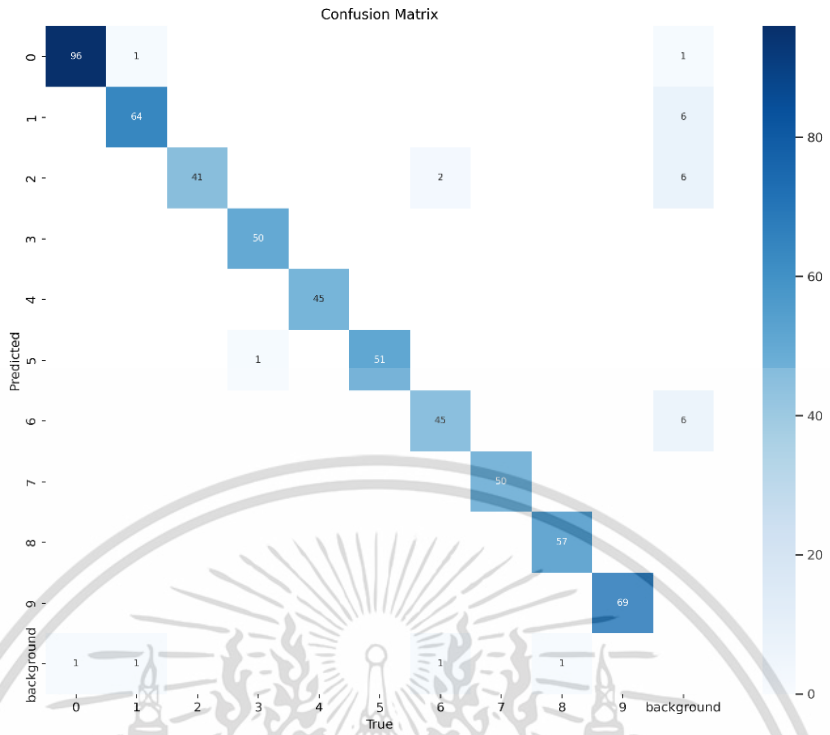
การทดสอบ	mAP50	mAP50-95	ค่าความเที่ยง	ค่าเรียกคืน
การทดสอบที่ 1 (70:15:15)	**98.9**	**67.4**	**98.2**	**98.7**
การทดสอบที่ 2 (70:20:10)	**98.9**	66.5	97.0	96.7
การทดสอบที่ 3 (80:10:10)	98.3	62.3	94.7	97.0
การทดสอบที่ 4 (80:15:5)	98.2	65.6	97.0	95.7

จากตารางที่ 4.2 จะพบว่า การทดสอบที่ 1 มีค่าเฉลี่ยความเที่ยงเมื่อ IoU เท่ากับ 50, ค่าเฉลี่ยความเที่ยงเมื่อ IoU เท่ากับ 50-95, ค่าความเที่ยง และค่าเรียกคืนมากที่สุด การทดสอบที่ 2 มีค่าเฉลี่ยความเที่ยงเมื่อ IoU เท่ากับ 50 มากที่สุด

4.2.1 การประเมินประสิทธิภาพของการทดสอบที่ 1 ตัวแบบที่ 2

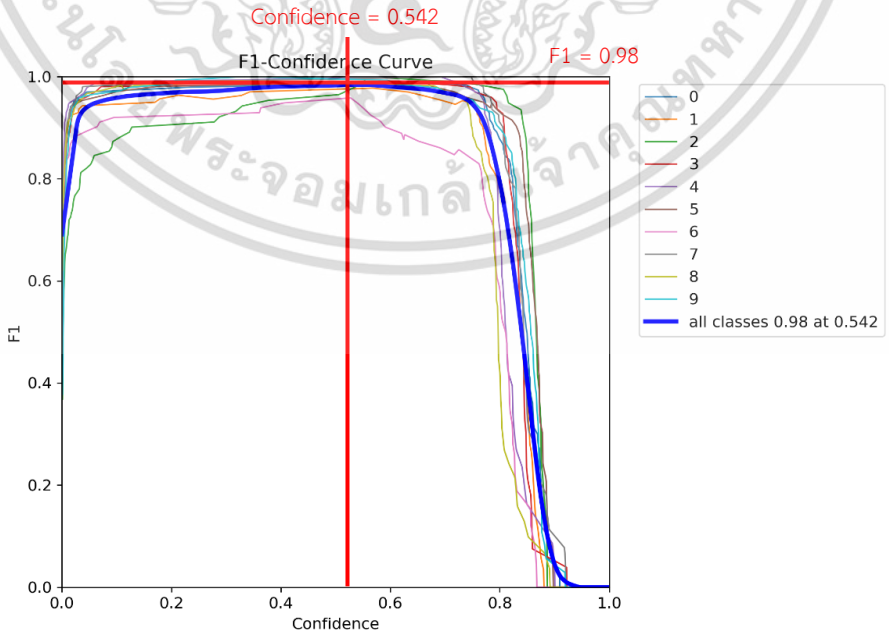
การแบ่งข้อมูลสำหรับการทดสอบที่ 1 แบ่งออกเป็น 70:15:15 ทำให้ได้ mAP50-95 อยู่ที่ 67.4 ดังตารางที่ 4.2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.15 ผลลัพธ์ของเมทริกซ์ความสับสนของการทดสอบที่ 1 ตัวแบบที่ 2

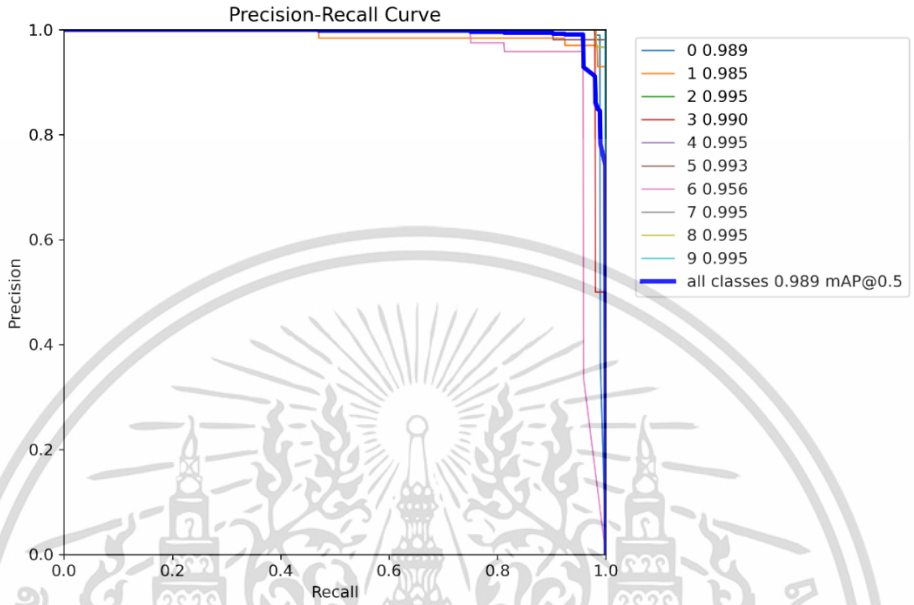
จากรูปที่ 4.15 เมทริกซ์ความสับสนแสดงให้เห็นว่า ตัวเลข 1, 2 และ 6 เป็นตัวเลขที่ทำนายผิดพลาดหรือเป็นตัวเลขที่ตัวแบบเกิดความสับสนระหว่างตัวเลขดังกล่าวกับรูปภาพพื้นหลัง โดยทำนายว่าภาพพื้นหลังเป็นเลข 1 ถึง 6 ครั้ง ทำนายว่าภาพพื้นหลังเป็นเลข 2 ถึง 8 ครั้ง และทำนายว่าภาพพื้นหลังเป็นเลข 6 ถึง 6 ครั้ง



รูปที่ 4.16 เส้นโค้งความเชื่อมั่นคะแนน F1 ของการทดสอบที่ 1 ตัวแบบที่ 2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 4.16 จะพบว่าตัวเลขที่ทำนายได้ทั้งหมดของการทดสอบที่ 1 มีค่าคะแนน F1 เป็น 0.98 เมื่อค่าความเชื่อมั่นเท่ากับ 0.542 และค่าความเชื่อมั่นเท่ากับ 0.542 เป็นค่าที่ทำให้ค่าความเที่ยงและค่าเรียกคืนมีความเหมาะสมที่สุด



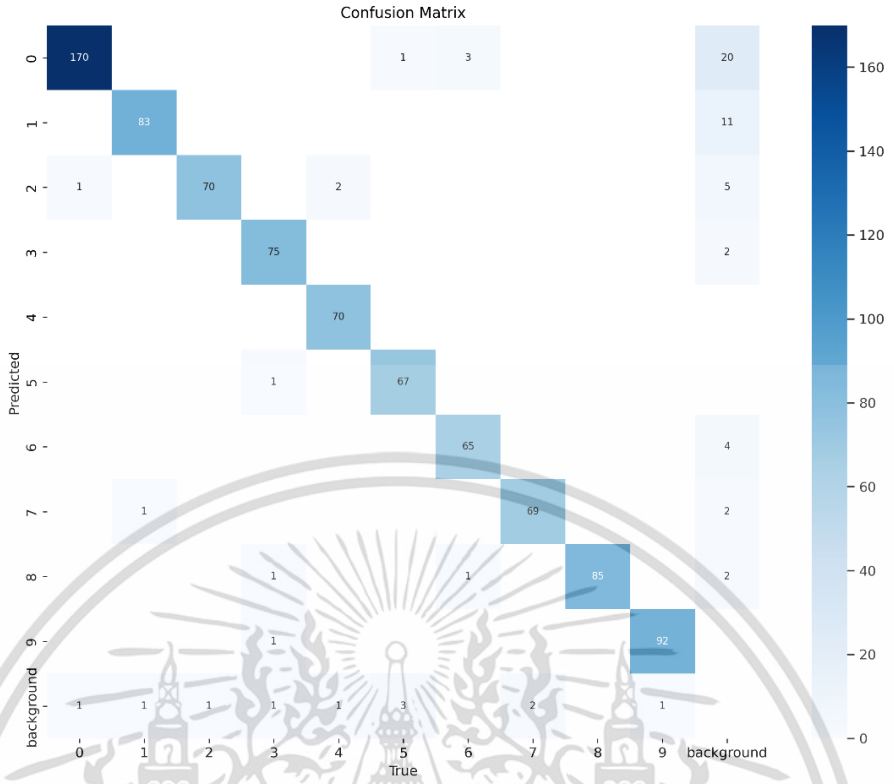
รูปที่ 4.17 เส้นโค้งค่าความเที่ยงและค่าเรียกคืนของการทดสอบที่ 1 ตัวแบบที่ 2

จากรูปที่ 4.17 จะพบว่าเส้นโค้งของค่าความเที่ยงและค่าเรียกคืนเท่ากับ 0.989 เมื่อค่าเฉลี่ยความเที่ยงที่ IoU เท่ากับ 0.5

4.2.2 การประเมินประสิทธิภาพของการทดสอบที่ 2 ตัวแบบที่ 2

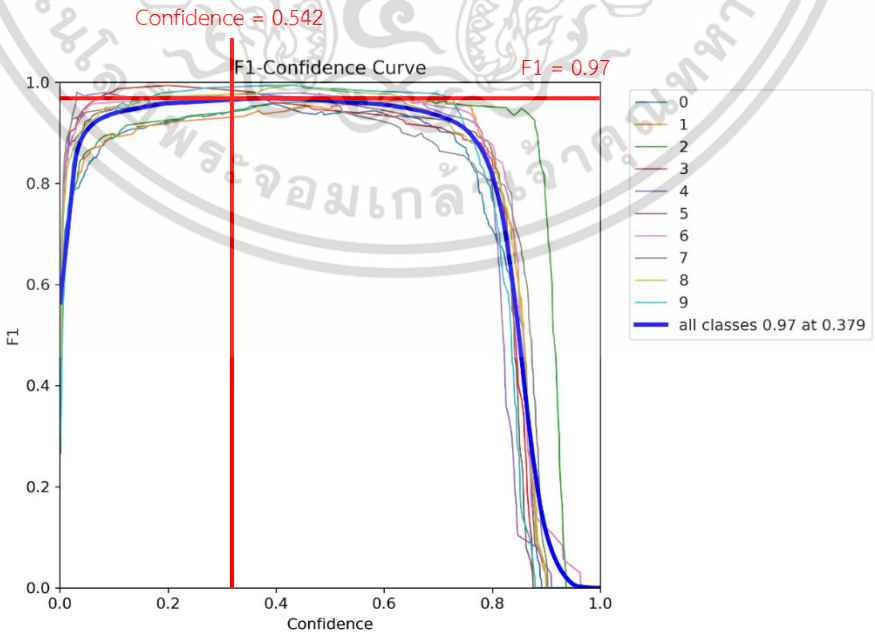
การแบ่งข้อมูลสำหรับการทดสอบที่ 2 แบ่งออกเป็น 70:20:10 ทำให้ได้ mAP50-95 อยู่ที่ 66.5 ดังตารางที่ 4.2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.18 ผลลัพธ์ของเมทริกซ์ความสับสนของการทดสอบที่ 2 ตัวแบบที่ 2

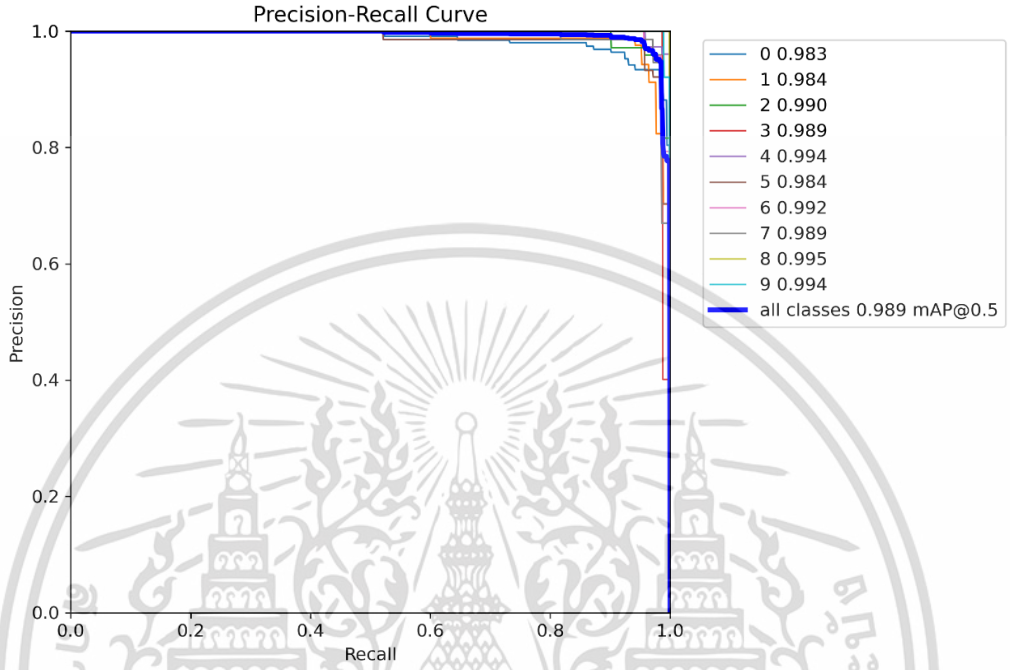
จากรูปที่ 4.18 เมทริกซ์ความสับสนแสดงให้เห็นว่าตัวเลข 0 และ 1 เป็นตัวเลขที่ทำนายผิดพลาดหรือเป็นตัวเลขที่ตัวแบบเกิดความสับสนระหว่างตัวเลขดังกล่าวกับรูปภาพพื้นหลัง ซึ่งตัวแบบมองภาพพื้นหลังเป็นเลข 0 ถึง 20 ครั้ง และมองภาพพื้นหลังเป็น 1 ถึง 11 ครั้ง



รูปที่ 4.19 เส้นโค้งความเชื่อมั่นคะแนน F1ของการทดสอบที่ 2 ตัวแบบที่ 2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 4.19 จะพบว่าตัวเลขที่ทำนายได้ทั้งหมดของการทดสอบที่ 2 มีค่าคะแนน F1 เป็น 0.97 เมื่อค่าความเชื่อมั่นเท่ากับ 0.379 และค่าความเชื่อมั่นเท่ากับ 0.379 เป็นค่าที่ทำให้ค่าความเที่ยงและค่าเรียกคืนมีความเหมาะสมที่สุด



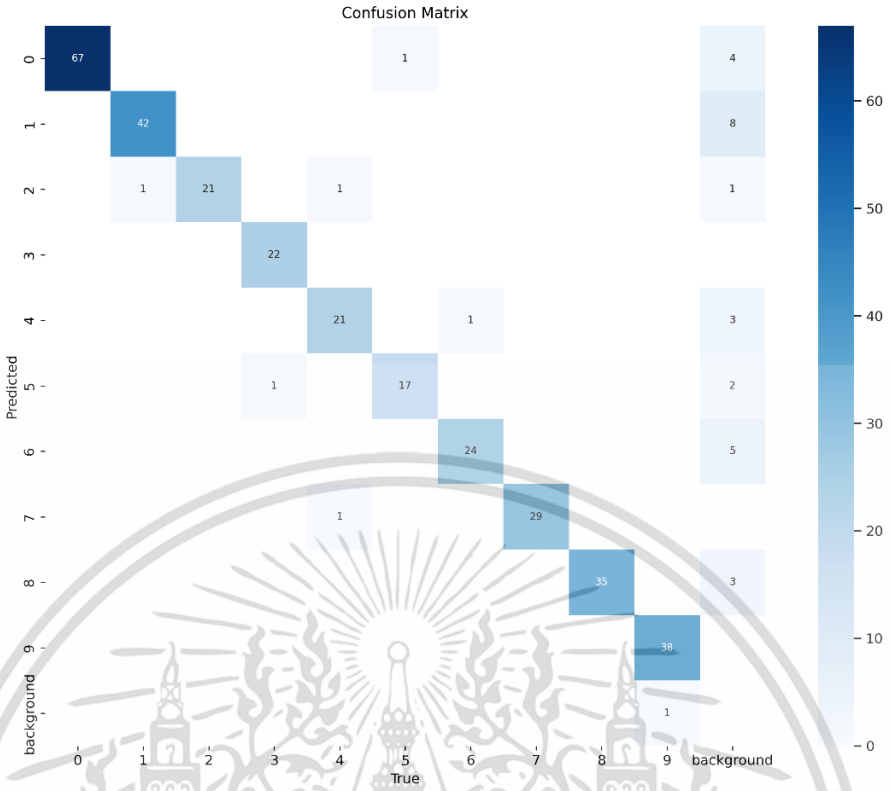
รูปที่ 4.20 เส้นโค้งค่าความเที่ยงและค่าเรียกคืนของการทดสอบที่ 2 ตัวแบบที่ 2

จากรูปที่ 4.20 จะพบว่า เส้นโค้งของค่าความเที่ยงและค่าเรียกคืนเท่ากับ 0.989 เมื่อค่าเฉลี่ยความเที่ยงที่ IoU เท่ากับ 0.5 เหมือนกับการทดสอบที่ 1

4.2.3 การประเมินประสิทธิภาพของการทดสอบที่ 3 ตัวแบบที่ 2

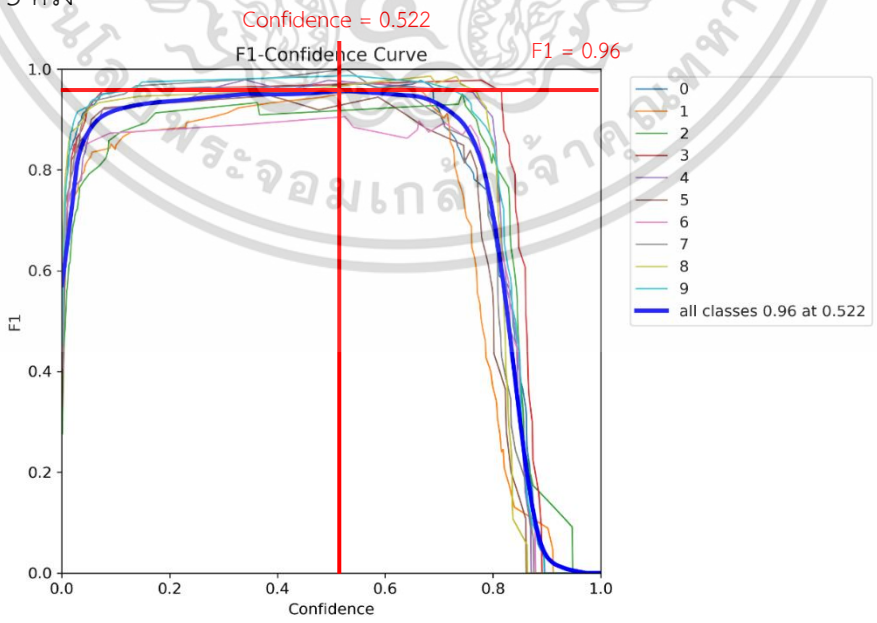
การแบ่งข้อมูลสำหรับการทดสอบที่ 3 แบ่งออกเป็น 80:10:10 ทำให้ได้ mAP50-95 อยู่ที่ 66.5 ดังตารางที่ 4.2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.21 ผลลัพธ์ของเมทริกซ์ความสับสนของการทดสอบที่ 3 ตัวแบบที่ 2

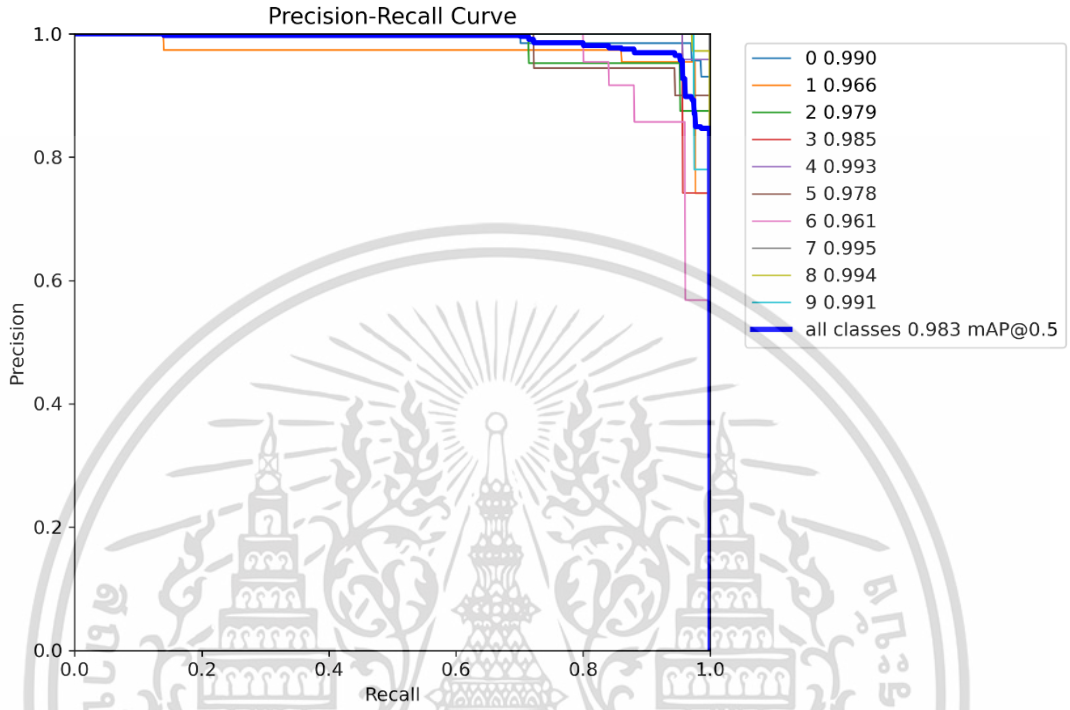
จากรูปที่ 4.21 จะพบว่าเมทริกซ์ความสับสนแสดงให้เห็นว่าตัวเลข 0, 1 และ 6 เป็นตัวเลขที่ทำนายผิดพลาดหรือเป็นตัวเลขที่ตัวแบบเกิดความสับสนระหว่างตัวเลขดังกล่าวกับรูปภาพพื้นหลัง ซึ่งตัวแบบมองภาพพื้นหลังเป็นเลข 0 ถึง 4 ครั้ง มองภาพพื้นหลังเป็น 1 ถึง 8 ครั้ง และมองภาพพื้นหลังเป็น 6 ถึง 5 ครั้ง



รูปที่ 4.22 เส้นโค้งความเชื่อมั่นคะแนน F1ของการทดสอบที่ 3 ตัวแบบที่ 2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 4.22 จะพบว่าตัวเลขที่ทำนายได้ทั้งหมดของการทดสอบที่ 3 มีค่าคะแนน F1 เป็น 0.96 เมื่อค่าความเชื่อมั่นเท่ากับ 0.522 และค่าความเชื่อมั่นเท่ากับ 0.522 เป็นค่าที่ทำให้ค่าความเที่ยงและค่าเรียกคืนมีความเหมาะสมที่สุด



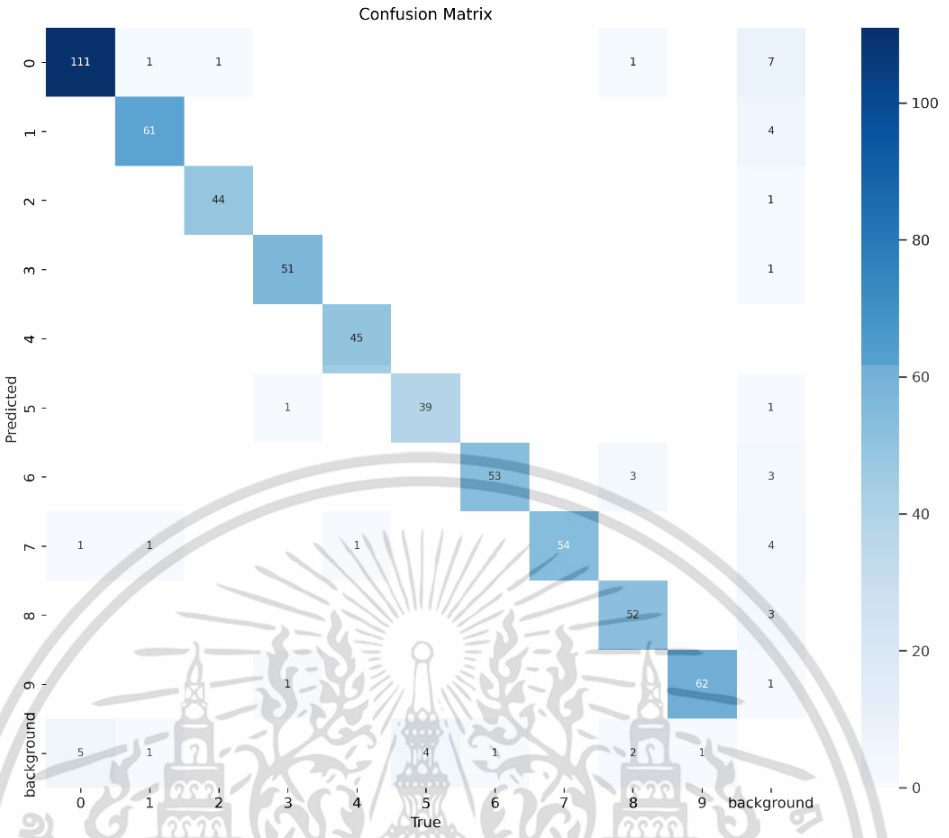
รูปที่ 4.23 เส้นโค้งค่าความเที่ยงและค่าเรียกคืนของการทดสอบที่ 3 ตัวแบบที่ 2

จากรูปที่ 4.23 จะพบว่า เส้นโค้งของค่าความเที่ยงและค่าเรียกคืนเท่ากับ 0.983 เมื่อค่าเฉลี่ยความเที่ยงที่ IoU เท่ากับ 0.5

4.2.4 การประเมินประสิทธิภาพของการทดสอบที่ 4 ตัวแบบที่ 2

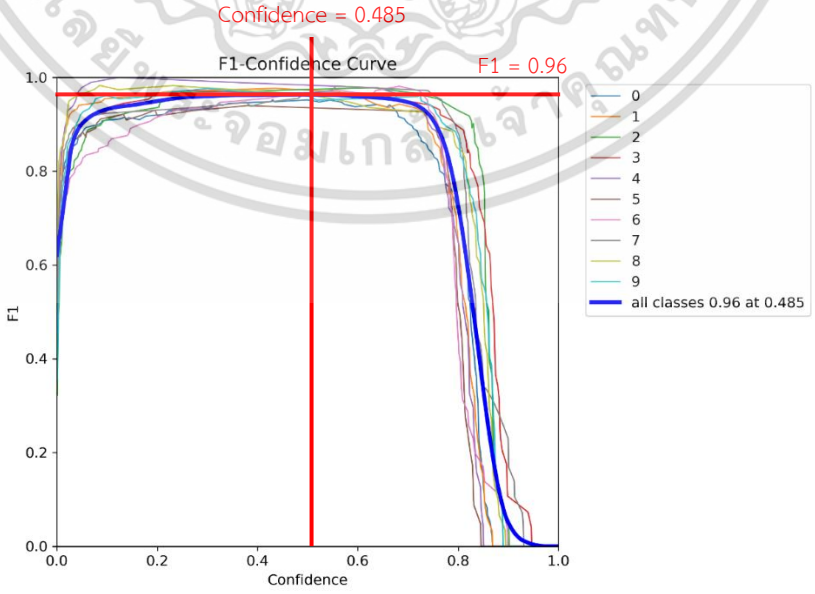
การแบ่งข้อมูลสำหรับการทดสอบที่ 4 แบ่งออกเป็น 80:15:5 ทำให้ได้ mAP50-95 อยู่ที่ 66.5 ดังตารางที่ 4.2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.24 ผลลัพธ์ของเมทริกซ์ความสับสนของการทดสอบที่ 4 ตัวแบบที่ 2

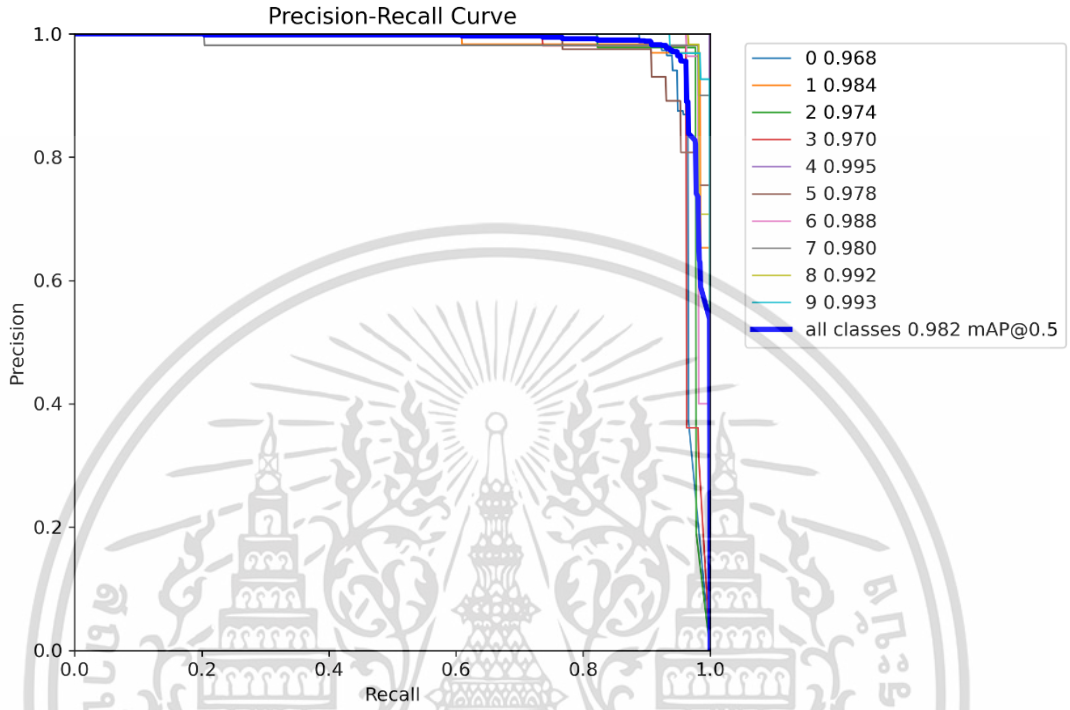
จากรูปที่ 4.24 จะพบว่าเมทริกซ์ความสับสนแสดงให้เห็นว่าตัวเลข 0 เป็นตัวเลขที่ทำนายผิดพลาดหรือเป็นตัวเลขที่ตัวแบบเกิดความสับสนระหว่างตัวเลขดังกล่าวกับรูปภาพพื้นหลัง ซึ่งตัวแบบมองภาพพื้นหลังเป็นเลข 0 ถึง 7 ครั้ง และมองเลข 0 เป็นภาพพื้นหลังถึง 5 ครั้ง



รูปที่ 4.25 เส้นโค้งความเชื่อมั่นคะแนน F1ของการทดสอบที่ 4 ตัวแบบที่ 2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 4.25 จะพบว่าตัวเลขที่ทำนายได้ทั้งหมดของการทดสอบที่ 4 มีค่าคะแนน F1 เป็น 0.96 เมื่อค่าความเชื่อมั่นเท่ากับ 0.485 และค่าความเชื่อมั่นเท่ากับ 0.485 เป็นค่าที่ทำให้ค่าความเที่ยงและค่าเรียกคืนมีความเหมาะสมที่สุด



รูปที่ 4.26 เส้นโค้งค่าความเที่ยงและค่าเรียกคืนของการทดสอบที่ 4 ตัวแบบที่ 2

จากรูปที่ 4.26 จะพบว่า เส้นโค้งของค่าความเที่ยงและค่าเรียกคืนเท่ากับ 0.982 เมื่อค่าเฉลี่ยความเที่ยงที่ IoU เท่ากับ 0.5

4.2.5 ผลการวิเคราะห์ประสิทธิภาพของตัวแบบที่ 2

จากการเปรียบเทียบประสิทธิภาพของการทดสอบทั้ง 4 การทดสอบพบว่า การทดสอบที่ 1 หรือการแบ่งข้อมูลออกเป็น 70:15:15 ทำให้ตัวแบบมีค่าวัดประสิทธิภาพทั้ง 4 ค่ามากที่สุด จึงเลือกใช้ตัวแบบจากการทดสอบที่ 1 ในการทำนายค่าระดับน้ำ

เมื่อนำตัวแบบจากการทดสอบที่ 1 มาใช้กับข้อมูลที่ไม่เคยใช้ในการเรียนรู้ การทดสอบ หรือการตรวจสอบมาก่อน พบว่า ตัวแบบสามารถตรวจจับและระบุตำแหน่งของตัวเลขได้ดังรูปที่ 4.7

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 4.27 การตรวจจับวัตถุของการทดสอบที่ 1 ของตัวแบบที่ 2

จากรูปที่ 4.27 จะพบว่าตัวแบบสามารถตรวจหาตัวเลขได้ 2 ตัวเลข ได้แก่เลข 8 และเลข 9 และตัวแบบสามารถนำสี่เหลี่ยมสีน้ำเงินไปล้อมรอบตัวเลขได้อย่างถูกต้อง


```
{'data': [{'box': {'x1': 8.340680122375488,
    'x2': 48.76849365234375,
    'y1': 88.90127563476562,
    'y2': 157.4007568359375},
   'class': 9,
   'confidence': 0.6971051692962646,
   'name': '9'},
  {'box': {'x1': 11.10673999786377,
    'x2': 46.45735168457031,
    'y1': 218.2248077392578,
    'y2': 280.234619140625},
   'class': 8,
   'confidence': 0.272097110748291,
   'name': '8'}],
 'message': 'Inference complete.',
 'success': True}
```

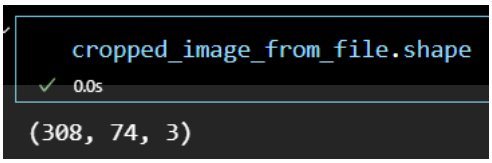

รูปที่ 4.28 ผลลัพธ์ของตัวแบบที่ 2

จากรูปที่ 4.28 จะพบว่าตัวแบบแสดงผลลัพธ์ของตัวเลขที่อยู่ในรูปภาพได้ โดย x1 และ x2 คือจุดเริ่มต้นและสิ้นสุดของแกน x ที่มีตัวเลขอยู่ตรงกลาง y1 และ y2 คือจุดเริ่มต้นและสิ้นสุดของแกน y ที่มีตัวเลขอยู่ตรงกลาง class คือเป้าหมายของงานวิจัย หรือค่าตัวเลขต่าง ๆ ที่เราต้องการอ่าน ค่า confidence คือค่าความเชื่อมั่นที่ตัวแบบสามารถทำนายได้ว่ากรอบ 4 เหลี่ยมสามารถล้อมรอบตัวเลขได้ถูกต้องมากน้อยเพียงใด โดยตัวแบบสามารถทำนายได้ว่า รูปภาพนำเข้าที่มาจากตัวแบบที่ 1 ประกอบไปด้วยเลข 9 และเลข 8

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.3 การวัดสัดส่วนเพื่อหาค่าระดับน้ำ

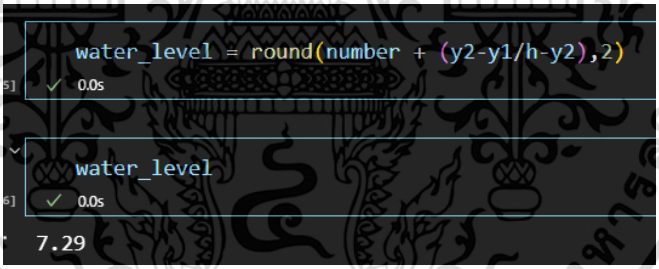
เมื่อได้ตำแหน่งของตัวเลขมาแล้ว ตัวแบบจะนำเลขที่มีค่า y2 มากที่สุด หรือหมายถึงเลข 8 ในรูปที่ 4.5 มาใช้ในการหาความสูงของตัวเลข โดยการนำ y2 – y1 (280 – 218 = 62) เมื่อได้ค่ามาแล้ว นำความสูงของรูปภาพมาลบกับ y2 เพื่อหาความชิดของตัวเลขกับขอบของรูปภาพ



รูปที่ 4.29 การหาความสูงของรูปภาพ

จากรูปที่ 4.29 จะพบว่าขนาดของรูปภาพคือ 308x74 โดยเลข 3 ในตำแหน่งสุดท้ายคือมิติของรูปภาพ เลขที่นำไปใช้แทนความสูงของรูปภาพคือเลขในตำแหน่งแรก หรือในที่นี่คือ 308

นำ y2 – y1 หารด้วยความสูงของรูปภาพ – y2 (280 – 218 / 308 – 280 = -0.71) จะได้อัตราส่วนของไม้วัดระดับน้ำในส่วนที่อยู่ใต้ตัวเลขจนถึงพื้นผิวน้ำ หลังจากนั้น นำค่าที่ได้มาบวกกับตัวเลขที่อยู่ใกล้ระดับน้ำมากที่สุด ผลลัพธ์ของการใช้สมการกับรูปที่ 4.27 เป็นดังรูปที่ 4.30



รูปที่ 4.30 การวัดมาตราส่วนในการอ่านค่าระดับน้ำ

จากรูปที่ 4.30 จะพบว่าเมื่อนำเลขที่อยู่ใกล้พื้นผิวน้ำที่สุดมาบวกกับตัวเลขที่ได้จากสมการแล้ว ค่าที่ได้คือ 7.29 ซึ่งสอดคล้องกับความเป็นจริง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5
สรุปผลการวิจัยและข้อเสนอแนะ

5.1 สรุปผลการวิจัย

งานวิจัยนี้ได้มีการนำตัวแบบที่ผ่านการเรียนรู้มาแล้ว (YOLOv8) มาใช้สำหรับการเรียนรู้ โดยชุดข้อมูลที่นำมาใช้ในตัวแบบแรกคือรูปภาพของไม้วัดระดับน้ำที่มีการจัดแต่งรูปภาพเพิ่มเติมให้สามารถสื่อถึงสถานการณ์จริงได้ โดยผลลัพธ์ที่ต้องการคือตำแหน่งของกรอบสี่เหลี่ยมที่ล้อมรอบไม้วัดระดับน้ำ ชุดข้อมูลที่นำมาใช้ในตัวแบบที่สองคือรูปภาพของตัวเลขที่มีลักษณะใกล้เคียงกับตัวเลขที่ปรากฏบนไม้วัดระดับน้ำ โดยผลลัพธ์ที่ต้องการคือตำแหน่งของตัวเลขที่อยู่ใกล้กับระดับน้ำมากที่สุดและต้องการทราบว่าตัวเลขที่ปรากฏเป็นเลขใด จากนั้น ได้มีการหาค่าตัวเลขหลังจุดทศนิยมเพื่อทำให้สามารถอ่านค่าระดับน้ำได้อย่างแม่นยำมากขึ้น งานวิจัยนี้ได้มีการพัฒนาหน้าต่างสำหรับการแสดงรายละเอียดต่าง ๆ ในการอ่านค่าระดับน้ำจากกล้องวงจรปิด (CCTV)

จากการทดสอบของตัวแบบที่ 1 พบว่า การทดสอบที่ 4 มีประสิทธิภาพที่ดีที่สุด โดยมีค่าในการวัดประสิทธิภาพสูงกว่าการทดสอบอื่นในเรื่องของค่าเฉลี่ยความเที่ยงเมื่อ IoU เท่ากับ 50 ค่าความเที่ยง และค่าเรียกคืน งานวิจัยจึงได้มีการเลือกใช้การทดสอบที่ 4 มาใช้สำหรับการตรวจจับไม้วัดระดับน้ำจากภาพถ่ายจากกล้องวงจรปิด

จากการทดสอบของจัวแบบที่ 2 พบว่า การทดสอบที่ 3 มีประสิทธิภาพที่ดีที่สุด โดยมีค่าในการวัดประสิทธิภาพสูงกว่าการทดสอบอื่นในทุกด้าน ได้แก่ ค่าเฉลี่ยความเที่ยงเมื่อ IoU เท่ากับ 50 ค่าเฉลี่ยความเที่ยงเมื่อ IoU เท่ากับ 50-95 ค่าความเที่ยง และค่าเรียกคืน งานวิจัยจึงได้มีการเลือกใช้การทดสอบที่ 3 มาใช้สำหรับการตรวจหาตัวเลขที่ใกล้เคียงกับระดับน้ำมากที่สุดจากรูปภาพที่เป็นผลลัพธ์ของตัวแบบที่ 1

หลังจากได้ตัวเลขบนไม้วัดระดับน้ำที่อยู่ใกล้กับระดับน้ำมากที่สุดแล้ว งานวิจัยได้มีการคำนวณหาอัตราส่วนเพื่อหาค่าในตำแหน่งทศนิยมเพื่อให้ได้ผลลัพธ์ที่มีความละเอียดมากยิ่งขึ้น และได้มีการพัฒนาหน้าต่างสำหรับการแสดงผลเพื่อให้เห็นตัวอย่างการใช้งานจริงของงานวิจัยในครั้งนี้

5.2 ข้อเสนอแนะ

งานวิจัยนี้เป็นการนำตัวแบบที่ผ่านการเรียนรู้แล้วมาประยุกต์ใช้กับการอ่านค่าระดับน้ำจากกล้องวงจรปิด เนื่องจากมีข้อมูลรูปภาพตั้งต้นที่ค่อนข้างจำกัด จึงมีรูปภาพในการเรียนรู้ไม่มากนัก หากผู้อื่นที่ต้องการศึกษาในหัวข้อใกล้เคียง สามารถนำรูปภาพจากกล้องวงจรปิด ณ สถานที่จริงมาใช้ในงานวิจัยได้ จะทำให้มีรูปภาพสำหรับการเรียนรู้มีจำนวนมากขึ้น และสามารถเพิ่มประสิทธิภาพให้กับตัวแบบได้ การเพิ่มจำนวนรอบในกรณีที่มีชุดข้อมูลรูปภาพมากขึ้น อาจทำให้ตัวแบบมีประสิทธิภาพเพิ่มขึ้นอีกด้วย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เอกสารอ้างอิง

[1] Adrain Rosebrock. 2021. **OpenCV Smoothing and Blurring.** [Online]. Available : https://pyimagesearch.com/2021/04/28/opencv-smoothing-and-blurring/

[2] Akshit Mehra. 2023. **Understanding YOLOv8 Architecture, Applications & Features.** [Online]. Available : https://www.labellerr.com/blog/understanding-yolov8-architecture-applications-features/

[3] BBC News ไทย. 2022. **เทียบข้อมูลน้ำท่วม 2554 กับ 2565 หลังรัฐยืนยันไม่ซ้ำรอย**. [Online]. Available: https://www.bbc.com/thai/articles/cx802845nxno

[4] Casella George and Berger Rogoer L. 2001. Statistical Inference (2nd ed.). Duxbury.

[5] Gang Dou, Rensheng Chen, Chuntan Han, Zhangwen Liu and Junfeng Liu. Research on Water-Level Recognition Method Based on Image Processing and Convolutional Neural Networks. Water 2022, 14, 1890. https://doi.org/10.3390/w14121890

[6] Ie Dong, Gaipeng Kong, Wenpu Dong, Liang Zhen, and Qi Tian. "Coarse2Fine: Two-layer Fusion for Image Retrieval." Journal of Latex Class Files, Vol. 14, No. 8, 2015

[7] Jacob Solawetz and Francesco. 2023. **What is YOLOv8? The Ultimate Guide**. [Online]. Available: https://blog.roboflow.com/whats-new-in-yolov8/

[8] Joe Hughes. 2023. **IoT & Ultrasonic Sensors: Revolutionising River Level Monitoring**. [Online]. Available: https://www.manxtechgroup.com/river-level-monitoring-withiot/#:~:text=The%20three%20most%20common%20sensors,ultrasonic%2C%20radar%20and%20depth%20sensors

[9] Jonathan Hui. 2018. **mAP (mean Average Precision) for Object Detection**. [Online]. Available: https://jonathan-hui.medium.com/map-mean-average-precision-for-object-detection-45c121a31173

[10] Tharis Thimthong and Urin Soteyome. 2562. "Development of an Automatic Water Level Measurement System from CCTV Pictures Using Gray Scale – Thresholding and Regression Analysis Techniques (Pilot Phase)." Information Technology Journal. 15(1). Graduate School of Public Administration, Burapha University. Irrigation Development Institute, Royal Irrigation Department. Vol. 15, No. 1, pp. 41-49, 2019

[11] Nadimpalli Lakshmi Manasa. 2019. Thai Digit Recognition on License Plates using YoloV3. Degree of Master of Engineering in Microelectronics and Embedded Systems.

[12] opencv dev team. 2019. **Miscellaneous Image Transformations.** [Online]. Available : https://docs.opencv.org/2.4/modules/imgproc/doc/miscellaneous_transformations.html#void%20cvtColor%28InputArray%20src,%20OutputArray%20dst,%20int%20code,%20int%20dstCn%29

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เอกสารอ้างอิง (ต่อ)

[13] Peratham Wiriyathammabhum. 2020. **Deep Residual Learning for Image Recognition (ResNet)**. [Online]. Available: https://peratham.medium.com/deep-residual-learning-for-image-recognition-resnet-daa83c22300

[14] Ritika. 2022. **Understanding technology behind self-driving cars: Stereo Vision**. [Online]. Available: https://www.labellerr.com/blog/understanding-technology-behind-self-driving-cars-stereo-vision/

[15] Stevan Lang, Fabrizio Ventola, Kristian Kersting. 2021. A One-Stage Anchor-Free Approach for Oriented Object Detection. Computer Science. Computer Vision and Pattern Recognition. Cornell University

[16] Sourabn_Sinha. 2023. **Python | Grayscaling of Images using OpenCV.** [Online]. Available : https://www.geeksforgeeks.org/python-grayscaling-of-images-using-opencv/

[17] Ultimate Python. 2021. **ตรวจจับใบหน้า ใน 11 บรรทัด Python Face Detection**. [Online] Available: https://www.ultimatepython.co/post/face-detection-opencv-python

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประวัติผู้เขียน

ชื่อ	นางสาว เปมิกา คุณาวิริยะสิริ
วัน เดือน ปีเกิด	26 สิงหาคม พ.ศ. 2542
ที่อยู่ปัจจุบัน	1213/75 ลาดพร้าว 95 แขวงพลับพลา เขตวังทองหลาง กทม. 10310
ประวัติการศึกษา	(2021) วิศวกรรมศาสตรบัณฑิต สาขาวิศวกรรมคอมพิวเตอร์ เกรดเฉลี่ย 3.45 (สถาบันเทคโนโลยีนานาชาติสิรินธร มหาวิทยาลัยธรรมศาสตร์)
ทุนการศึกษาที่ได้รับ	ทุนนวัตกรรมดิจิทัลในสาขาที่ขาดแคลนของสำนักงาน ก.พ.
ผลงานทางวิชาการ	ไม่มี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้