



## การคัดแยกขยะรีไซเคิลจากภาพด้วยเทคนิคการเรียนรู้เชิงลึก

โดย

นางสาวพิชญาภา สารทองคง

โครงการพิเศษนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร  
วิทยาศาสตรบัณฑิต  
สาขาวิชาวิทยาการคอมพิวเตอร์  
คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์  
ปีการศึกษา 2567  
ลิขสิทธิ์ของมหาวิทยาลัยธรรมศาสตร์

การคัดแยกขยะรีไซเคิลจากภาพด้วยเทคนิคการเรียนรู้เชิงลึก

โดย

นางสาวพิชญาภา สระทองคต

โครงการพิเศษนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

วิทยาศาสตรบัณฑิต

สาขาวิชาวิทยาการคอมพิวเตอร์

คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์

ปีการศึกษา 2567

ลิขสิทธิ์ของมหาวิทยาลัยธรรมศาสตร์

RECYCLABLE WASTE SEGREGATION FROM IMAGES USING DEEP  
LEARNING TECHNIQUES

BY

MISS PHITCHAYAPHA SRATHONGKHOD

A FINAL-YEAR PROJECT REPORT SUBMITTED IN PARTIAL FULFILLMENT  
OF THE REQUIREMENTS FOR THE DEGREE OF BACHELOR OF SCIENCE  
COMPUTER SCIENCE  
FACULTY OF SCIENCE AND TECHNOLOGY  
THAMMASAT UNIVERSITY  
ACADEMIC YEAR 2024  
COPYRIGHT OF THAMMASAT UNIVERSITY

(4)

มหาวิทยาลัยธรรมศาสตร์  
คณะวิทยาศาสตร์และเทคโนโลยี

รายงานโครงการพิเศษ

ของ

นางสาวพิชญา สารทองคต

เรื่อง

การคัดแยกขยะรีไซเคิลจากภาพด้วยเทคนิคการเรียนรู้เชิงลึก

ได้รับการตรวจสอบและอนุมัติ ให้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร  
หลักสูตรวิทยาศาสตรบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์  
เมื่อ วันที่ 30 พฤษภาคม พ.ศ. 2568

อาจารย์ที่ปรึกษา

(ผศ. ดร. วิรัตน์ จาริวงศ์เพบูลย์)

กรรมการสอบโครงการพิเศษ

(ผศ. ดร. เสาวลักษณ์ วรรณนาภา)

กรรมการสอบโครงการพิเศษ

๖๒ ๗๑๘ ๗๗๔ ๗๗๔  
(อ. ดร. นవฤกษ์ ชลารักษ์)

(5)

มหาวิทยาลัยธรรมศาสตร์  
คณะวิทยาศาสตร์และเทคโนโลยี

รายงานโครงการพิเศษ

ของ

นางสาวพิชญา สารทองคต

เรื่อง

การคัดแยกขยะรีไซเคิลจากภาพด้วยเทคนิคการเรียนรู้เชิงลึก

ได้รับการตรวจสอบและอนุมัติ ให้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร  
หลักสูตรวิทยาศาสตรบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์  
เมื่อ วันที่ 30 พฤษภาคม พ.ศ. 2568

อาจารย์ที่ปรึกษา

(ผศ. ดร. วิรัตน์ จาริวงศ์เพบูล์)

กรรมการสอบโครงการพิเศษ

(ผศ. ดร. เสาวลักษณ์ วรรธนาภา)

กรรมการสอบโครงการพิเศษ

๖๒๗๑๘๗ ๗๗๗๗๗๗

(อ. ดร. นฤกษ์ ชลาธักษ์)

(6)

หัวข้อโครงการพิเศษ	การคัดแยกขยะรีไซเคิลจากภาพด้วยเทคนิคการเรียนรู้เชิงลึก
ชื่อผู้เขียน	นางสาวพิชญาภา สารทองคต
ชื่อปริญญา	วิทยาศาสตรบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์
สาขาวิชา/คณะ/มหาวิทยาลัย	สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์
อาจารย์ที่ปรึกษาโครงการพิเศษ	ผศ. ดร. วิรัตน์ จารีวงศ์เพบูลร์
ปีการศึกษา	2567

## บทคัดย่อ

ในปัจจุบันปัญหาการจัดการขยะรีไซเคิลยังคงเป็นประเด็นสำคัญที่ส่งผลกระทบต่อสิ่งแวดล้อมและคุณภาพชีวิต การคัดแยกขยะอย่างถูกต้องถือเป็นวิธีการสำคัญในการเพิ่มประสิทธิภาพการนำวัสดุกลับมาใช้ใหม่ แต่กระบวนการดังกล่าวยังคงต้องพึ่งพาความรู้และความใส่ใจของมนุษย์ ทำให้เกิดความผิดพลาดในการคัดแยกบ่อยครั้ง ผู้พัฒนาโครงการจึงเล็งเห็นถึงความสำคัญของการพัฒนาโมเดลที่ช่วยสนับสนุนการคัดแยกขยะให้มีความสะดวกและแม่นยำมากยิ่งขึ้น

โครงการนี้ได้นำเทคนิคการเรียนรู้เชิงลึก (Deep Learning) มาประยุกต์ใช้ในการพัฒนาโมเดลสำหรับจำแนกประเภทของขยะจากภาพถ่ายโดยอัตโนมัติ เพื่อเปรียบเทียบประสิทธิภาพของโมเดล จึงได้ทดลองใช้โมเดล You Only Look Once (YOLO) และ Convolutional Neural Network (CNN) ด้วยชุดข้อมูลเดียวกันจำนวน 1,200 ภาพ ผลการทดลองพบว่า YOLO ให้ค่า F1-Score เท่ากับ 87 ในขณะที่ CNN ให้ค่า F1-Score เท่ากับ 60 แต่เมื่อ YOLO จะให้ผลลัพธ์ที่แม่นยำกว่า แต่ CNN มีความหมายสมในด้านการนำไปใช้งานจริงมากกว่า เนื่องจากพัฒนาได้ง่าย ใช้ทรัพยากรน้อย และไม่ต้องเตรียมข้อมูล annotation

ภายหลังจากการเลือกใช้โมเดล CNN ได้มีการเพิ่มปริมาณข้อมูลฝึกเป็นจำนวนทั้งสิ้น 8,031 ภาพ และทำการฝึกโมเดลใหม่อีกครั้ง ผลการประเมินโมเดลด้วยชุดทดสอบแสดงค่า Accuracy เท่ากับ 78% ซึ่งสะท้อนถึงศักยภาพของโมเดลในการจำแนกประเภทขยะได้อย่างมีประสิทธิภาพ และสามารถนำไปประยุกต์ใช้เพื่อสนับสนุนกระบวนการคัดแยกขยะในชีวิตประจำวันได้อย่างเป็นรูปธรรม

**คำสำคัญ:** การคัดแยกขยะรีไซเคิล, การเรียนรู้เชิงลึก, การจำแนกขยะ

Thesis Title	RECYCLABLE WASTE SEGREGATION FROM IMAGES USING DEEP LEARNING TECHNIQUES
Author	MISS PHITCHAYAPHA SRATHONGKHOD
Degree	Bachelor of Science
Major Field/Faculty/University	Computer Science Faculty of Science and Technology Thammasat University
Project Advisor	Asst. Prof. Ph.D. Wirat Jareevongpiboon
Academic Years	2024

## ABSTRACT

Recyclable waste management is a critical issue that impacts both the environment and public well-being. Proper waste separation significantly enhances the efficiency of material reuse. However, this process still depends on human awareness and knowledge, which often leads to classification errors. This project aims to develop an intelligent model to assist with more accurate and convenient waste sorting.

Deep learning techniques were applied to create an image classification model that automatically identifies types of waste. Two models—You Only Look Once (YOLO) and Convolutional Neural Network (CNN)—were trained and compared using the same dataset of 1,200 images. YOLO achieved an F1-Score of 87, while CNN achieved an F1-Score of 60. Although YOLO was more accurate, CNN was selected for this project due to its simplicity, lower resource requirements, and ease of deployment.

After selecting CNN, the dataset was expanded to 8,031 images. The model was retrained and achieved an accuracy of 78% on the test set. These results demonstrate that the model is effective for supporting waste classification and can be practically applied in real-world scenarios to improve recycling efficiency.

**Keywords:** Waste recycling, Deep learning, Waste classification

## กิตติกรรมประกาศ

ในการศึกษาโครงการนการจำแนกขยะรีไซเคิลจากภาพด้วยเทคนิคการเรียนรู้ของเครื่อง สำเร็จลุล่วงได้ด้วยความกรุณา และความช่วยเหลือจาก ผศ.ดร.วิรัตน์ จาเรืองศิริพูลย์ อาจารย์ที่ปรึกษา ที่ได้กรุณาให้คำปรึกษาแนะนำ ตรวจสอบ และแก้ไขข้อบกพร่อง ทุกขั้นตอนของการจัดทำโครงการ ผู้พัฒนาโครงการขอขอบพระคุณเป็นอย่างสูง

สุดท้ายนี้ ขอกราบขอบพระคุณ บิดา มารดา คุณย่า คุณอา วงศ์SEVENTEEN และเพื่อนนักศึกษา ตลอดจนผู้ที่มีส่วนเกี่ยวข้องทุกท่านที่ไม่ได้กล่าวนามไว้ ณ ที่นี่ ที่ได้ให้คำแนะนำ คำปรึกษา กำลังใจและส่วนช่วยเหลือให้โครงการพิเศษฉบับนี้สำเร็จลุล่วงได้ด้วยดี

นางสาวพิชญาภา สารทองคต

<b>สารบัญ</b>	
หน้า	
บทคัดย่อ	7
ABSTRACT	9
กิตติกรรมประกาศ	10
สารบัญ	11
สารบัญตาราง	14
สารบัญภาพ	15
รายการสัญลักษณ์และคำย่อ	17
บทที่ 1 บทนำ	1
1.1 ความเป็นมาและความสำคัญของโครงการ	1
1.2 วัตถุประสงค์	3
1.3 ขอบเขตของโครงการ	3
1.4 ประโยชน์ของโครงการ	4
1.5 ข้อจำกัดของโครงการ	4
บทที่ 2 วรรณกรรมและงานวิจัยที่เกี่ยวข้อง	5
2.1 แนวคิดทฤษฎีที่เกี่ยวข้อง	5
2.1.1 ขยายรีเซ็ต	5
2.1.2 การประมวลผลภาพ (image processing)	7
2.1.3 การเรียนรู้ของเครื่อง (Machine Learning)	9
2.1.4 การเรียนรู้เชิงลึก (Deep Learning)	11

2.2 เทคโนโลยีที่เกี่ยวข้อง	13
2.2.1 TensorFlow	13
2.2.2 Data Augmentation	14
2.2.3 Convolutional Neural Networks (CNN)	15
2.2.4 Mask R-CNN (Mask Region-based Convolutional Neural Network)	18
2.2.5 YOLO (You Only Look Once)	19
2.2.6 Roboflow	21
2.2.7 Google colab	22
2.2.8 Python	24
2.2.9 Gradio	25
2.3 การวัดประสิทธิภาพและการประเมินผล	26
2.4 งานวิจัยที่เกี่ยวข้อง	30
2.4.1 การจำแนกพาราเซอร์องค์วายการเรียนรู้เชิงลึก	30
2.4.2 การจำแนกภาพหมวดแบบเซ็ตเปิดด้วยโครงข่ายประสาทเทียมแบบคอนโวลูชัน	30
2.4.3 การพัฒนาและประยุกต์ใช้เทคนิค CNN ในการจำแนก ขยายชาญฝั่งทะเล	32
2.4.4 เครื่องคัดแยกขยะรีไซเคิลด้วยการประมวลผลภาพ	33
2.4.5 สรุปและเปรียบเทียบความแตกต่างของระบบ	34
บทที่ 3 วิธีการวิจัย	36
3.1 ภาพรวมของโครงงาน	36
3.1.1 ผู้ใช้งาน (User)	36
3.1.2 การตรวจสอบข้อมูลนำเข้า (Input Validation)	36
3.1.3 การเตรียมข้อมูล (Data Preprocessing)	36
3.1.4 การจำแนกประเภทโดยโมเดล (Classification by Model)	37
3.1.5 หมวดหมู่การจำแนกประเภท (Classification Categories)	37
3.2 การทดลองเบื้องต้นเพื่อเลือกแนวทางการพัฒนาโมเดล	37
3.2.1 การทดลองด้วย YOLO (You Only Look Once)	38

3.2.2 การทดลองด้วย CNN (Convolutional Neural Networks)	43
3.3 ประดิษฐ์ป่าสนใจและสิ่งที่ท้าทาย	54
3.4 ผลลัพธ์ที่คาดหวัง	56
3.5 ระบบต้นแบบและผลลัพธ์เบื้องต้น	56
บทที่ 4 ผลการดำเนินงาน	59
4.1 การจัดเตรียมhaarดเวย์และซอฟต์แวร์	59
4.1.1 haarดเวย์ที่ใช้ในการพัฒนา	59
4.1.2 ซอฟต์แวร์ที่ใช้ในการพัฒนาโปรแกรม	60
4.2 แผนการดำเนินงาน	60
4.3 ผลการทดลอง	62
4.3.1 ผลลัพธ์เบื้องต้นของโมเดล YOLO	62
4.3.2 ผลลัพธ์เบื้องต้นของโมเดล CNN	69
4.3.3 การเปรียบเทียบและวิเคราะห์ความเหมาะสมของโมเดล CNN และ YOLO สำหรับการจำแนกประเภทขยะรีไซเคิล	78
4.4 การพัฒนาและประเมินโมเดล CNN สำหรับการจำแนกประเภทขยะรีไซเคิล	85
4.4.1 ผลการฝึกโมเดล	85
4.4.2 การประเมินประสิทธิภาพของโมเดลด้วยชุดข้อมูลสอบ (test set)	88
4.4.3 การทดสอบโมเดลกับข้อมูลภายนอกผ่าน Gradio Interface	95
บทที่ 5 สรุป	100
5.1 สรุปผลการดำเนินงาน	100
5.2 ข้อเสนอแนะ	101
รายการอ้างอิง	102

## สารบัญตาราง

	หน้า
ตารางที่ 2.1 สรุปและเปรียบเทียบความแตกต่างของระบบ	34
ตารางที่ 3.1 จำนวนข้อมูล	39
ตารางที่ 3.2 จำนวนข้อมูลฝึกอบรม	45
ตารางที่ 3.3 จำนวนข้อมูลทดสอบ	45
ตารางที่ 4.1 ข้อมูลไฮาร์ดแวร์ที่ใช้ในการพัฒนา	59
ตารางที่ 4.2 การดำเนินงานที่ผ่านมาจนถึงปัจจุบัน	60
ตารางที่ 4.3 ผลการประเมินประสิทธิภาพของโมเดล YOLO สำหรับการจำแนกประเภทยีรีไซเคิล รายคลาส	62
ตารางที่ 4.4 ผลการทดสอบโมเดล YOLO กับภาพถ่ายที่มีขยะหลายประเภทรวมอยู่ในภาพเดียว	68
ตารางที่ 4.5 ผลการประเมินประสิทธิภาพของโมเดล CNN สำหรับการจำแนกประเภทยีรีไซเคิล รายคลาส	69
ตารางที่ 4.6 ผลการทดสอบโมเดล CNN กับข้อมูลนอก Dataset	78
ตารางที่ 4.7 การเปรียบเทียบประสิทธิภาพของโมเดลด้วยค่าซีวัด	79
ตารางที่ 4.8 การเปรียบเทียบประสิทธิภาพของโมเดลกับข้อมูลนอก Dataset	79
ตารางที่ 4.9 การเปรียบเทียบคุณสมบัติและผลลัพธ์ของโมเดล CNN และ YOLO ในงานจำแนก ประเภทยีรีไซเคิล	83
ตารางที่ 4.10 แสดงจำนวนภาพขยะก่อน-หลังการเพิ่มข้อมูล	85
ตารางที่ 4.11 ค่าความแม่นยำในแต่ละประเภทหลังจากทดสอบโมเดล CNN ด้วยชุดข้อมูลใหญ่	88

## สารบัญภาพ

	หน้า
ภาพที่ 1.1 ภาพถั่งขยะประเภทต่างๆ	1
ภาพที่ 2.1 สัญลักษณ์รีไซเคิล	5
ภาพที่ 2.2 กระดาษ	5
ภาพที่ 2.3 พลาสติก	6
ภาพที่ 2.4 โลหะ	6
ภาพที่ 2.5 แก้ว	7
ภาพที่ 2.6 ระบบตรวจรácทางคำตอป	8
ภาพที่ 2.7 ระบบตรวจจับใบหน้าในกล้องดิจิตอล	9
ภาพที่ 2.8 โครงสร้างของโครงข่ายประสาทเทียม	12
ภาพที่ 2.9 โครงสร้างของ CNN	15
ภาพที่ 2.10 Convolution layer	16
ภาพที่ 2.11 Pooling layer	17
ภาพที่ 2.12 Fully Connected layer	18
ภาพที่ 2.13 โครงสร้างของ Mask R-CNN	18
ภาพที่ 2.14 ตัวอย่างการใช้ Mask R-CNN	19
ภาพที่ 2.15 ตัวอย่างองค์ประกอบสถาปัตยกรรมโมเดลตรวจจับวัตถุ YOLO	20
ภาพที่ 2.16 ตัวอย่างการใช้ YOLO ในภาคการจราจร	21
ภาพที่ 2.17 Roboflow	21
ภาพที่ 2.18 Google colab	23
ภาพที่ 2.19 Python	24
ภาพที่ 2.20 ตัวอย่างการใช้งาน Gradio	25
ภาพที่ 2.21 สูตรการคำนวณหาค่า Intersection over union	29
ภาพที่ 3.1 สถาปัตยกรรมระบบ	36
ภาพที่ 3.2 ตัวอย่างการวิเคราะห์ขอบเขตและความต้องการของ YOLO	38
ภาพที่ 3.3 การรวมข้อมูลจาก Kaggle	38
ภาพที่ 3.4 การทำ Annotation	39
ภาพที่ 3.5 โค้ดการฝึกโมเดล YOLO	40
ภาพที่ 3.6 ตัวอย่างการวิเคราะห์ขอบเขตและความต้องการของ CNN	43

ภาพที่ 3.7 การจัดเก็บภาพของขยะไว้บน Google Drive	44
ภาพที่ 3.8 การแบ่งไฟล์เดอร์ข้อมูลฝึกอบรมและข้อมูลทดสอบ	44
ภาพที่ 3.9 โค้ดการทำ Data Augmentation	46
ภาพที่ 3.10 การทำ rotation_range=20	46
ภาพที่ 3.11 การทำ width_shift_range=0.1	47
ภาพที่ 3.12 การทำ height_shift_range=0.1	47
ภาพที่ 3.13 การทำ zoom_range=0.2	47
ภาพที่ 3.14 การทำ horizontal_flip=True	48
ภาพที่ 3.15 การสร้างโมเดล CNN	50
ภาพที่ 3.16 การฝึกสอนโมเดล CNN	51
ภาพที่ 3.17 หน้าหลักของระบบ	56
ภาพที่ 3.18 หน้าการแสดงผลข้อมูล	57
ภาพที่ 4.1 Confusion Matrix ของ YOLO	64
ภาพที่ 4.2 ภาพที่มีการลากกรอบจากข้อมูลจริง ของ YOLO	65
ภาพที่ 4.3 ภาพที่ไม่ได้ YOLO ทำนาย	66
ภาพที่ 4.4 Confusion Matrix ของ CNN	71
ภาพที่ 4.5 ตัวอย่างการทำนายที่ผิดพลาดของโมเดลจากข้อมูลทดสอบ CNN	72
ภาพที่ 4.6 กราฟความแม่นยำ (Accuracy) ต่อ Epoch	86
ภาพที่ 4.7 กราฟค่าความสูญเสีย (Loss) ต่อ Epoch	87
ภาพที่ 4.8 ตาราง Confusion Matrix ของโมเดล CNN ด้วยชุดข้อมูลใหญ่	90
ภาพที่ 4.9 ตัวอย่างภาพที่แก้วุกทำนายเป็นพลาสติก (1)	91
ภาพที่ 4.10 ตัวอย่างภาพที่แก้วุกทำนายเป็นพลาสติก (2)	92
ภาพที่ 4.11 ตัวอย่างภาพที่โลหะถูกทำนายเป็นกระดาษ (1)	92
ภาพที่ 4.12 ตัวอย่างภาพที่โลหะถูกทำนายเป็นกระดาษ (2)	93
ภาพที่ 4.13 ตัวอย่างภาพที่พลาสติกถูกทำนายเป็นกระดาษ (1)	93
ภาพที่ 4.14 ตัวอย่างภาพที่พลาสติกถูกทำนายเป็นกระดาษ (2)	94
ภาพที่ 4.15 ตัวอย่างภาพที่กระดาษถูกทำนายเป็นพลาสติก (1)	95
ภาพที่ 4.16 ตัวอย่างภาพที่กระดาษถูกทำนายเป็นพลาสติก (2)	95
ภาพที่ 4.17 ตัวอย่าง Gradio Interface ที่ใช้ทดสอบโมเดลกับภาพภายนอก	96
ภาพที่ 4.18 ตัวอย่างภาพที่ไม่สามารถทำนายได้ถูกต้อง	97
ภาพที่ 4.19 ตัวอย่างภาพที่ไม่ได้ทำนายผิด	98

## รายการสัญลักษณ์และคำย่อ

สัญลักษณ์/คำย่อ

CNN

YOLO

คำเต็ม/คำจำกัดความ

Convolutional Neural Network

You Only Look Once

## บทที่ 1

### บทนำ

#### 1.1 ความเป็นมาและความสำคัญของโครงการ

ในปัจจุบันถังขยะหลากหลายสีถูกจัดวางไว้ในแต่ละสถานที่อย่างแพร่หลายไม่ว่าจะเป็นมหาวิทยาลัย โรงพยาบาล ห้างสรรพสินค้า หรือพื้นที่อื่นๆ ต่างก็มีถังขยะหลายสีวางเรียงรายกัน โดยหลักๆแล้ว สีของถังขยะที่พบเห็นส่วนใหญ่ ได้แก่ สีเขียว, สีเหลือง, สีแดง และ สีน้ำเงิน ซึ่งแสดงถึงการแยกประเภทขยะตามลักษณะการจัดการ ดังนี้ 1.ขยะอินทรีย์ 2.ขยะรีไซเคิล 3.ขยะอันตราย 4.ขยะทั่วไป การมีถังขยะเหล่านี้อยู่ในชุมชนและพื้นที่สาธารณะต่างๆ สะท้อนให้เห็นถึงความตระหนักรและความพยายามของสังคมในการส่งเสริมการแยกประเภทของขยะเพื่อการจัดการทรัพยากรและสิ่งแวดล้อมที่มีประสิทธิภาพและยั่งยืนมากยิ่งขึ้น



ภาพที่ 1.1 ภาพถังขยะประเภทต่างๆ

แต่กลับกัน แม้ว่าถังขยะจะถูกจัดวางอย่างชัดเจนตามสีและประเภทมากเพียงใด แต่การแยกขยะยังคงเป็นความท้าทายสำหรับหลายคน ซึ่งอาจเกิดได้จากข้อจำกัดในการเข้าใจประเภทของขยะ ความไม่ทราบประเภทของขยะ หรือความยากในการจำแนกขยะบางประเภท หากขยะได้รับการจัดการอย่างไม่ถูกต้องจะเกิดปัญหาที่สำคัญคือ ขยะมูลฝอยจะเพิ่มมากยิ่งขึ้น ทั้งนี้ ข้อมูลจากการควบคุมมลพิษในปี 2566 มีขยะมูลฝอยในประเทศไทยจำนวน 26.95 ล้านตัน โดยปริมาณขยะที่เกิดขึ้นทั้งหมด มีมากถึง 27.7% ที่ไม่สามารถนำไปใช้ประโยชน์หรือกำจัดได้อย่างถูกต้อง เนื่องจาก การคัดแยกขยะที่ไม่ถูกวิธี (กรมควบคุมมลพิษ, 2566)

การจัดการกับขยะรีไซเคิลก็เป็นสิ่งที่ต้องให้ความสนใจเช่นกัน หากขยะประเภทนี้ไม่ได้รับการแยกอย่างถูกต้อง ก็จะต้องทิ้งไปในกับขยะอื่นๆ ผลสุดท้ายก็จะไปอยู่ที่โรงงานขยะ ซึ่งมีขยะจำนวนมาก มหาศาล ขยะรีไซเคิลที่ปนเปื้อนจะถูกจัดการเหมือนขยะมูลฝอยทั่วไป ซึ่งอาจนำไปสู่การกำจัดด้วยการเผาหรือฝังกลบ ส่งผลให้วัสดุรีไซเคิลสูญเสียคุณค่าและเพิ่มปริมาณขยะที่ส่งผลกระทบต่อสิ่งแวดล้อม ขยะรีไซเคิลที่ถูกทิ้งผิดวิธีไม่เพียงแต่ทำให้สูญเสียโอกาสในการนำวัสดุกลับมาใช้ใหม่ แต่ยังสร้างภาระให้กับโรงงานขยะ ซึ่งต้องจัดการกับขยะที่ปนกันอย่างซับซ้อน ทำให้กระบวนการแยก และรีไซเคิลมีความยากลำบากมากขึ้น ส่งผลให้ประสิทธิภาพของระบบการจัดการลดลง และอาจต้องพึ่งพาวิธีการกำจัด เช่นการฝังกลบหรือเผา ซึ่งเพิ่มผลกระทบต่อสิ่งแวดล้อมในระยะยาว

อีกหนึ่งปัญหาสำคัญที่ต้องให้ความสนใจคือสภาพการทำงานของพนักงานในโรงงานขยะ ซึ่งต้องเผชิญกับความท้าทายและความเสี่ยงจากการจัดการขยะที่ไม่ได้รับการแยกประเภทอย่างเหมาะสม ขยะรีไซเคิลที่ปนเปื้อนหรือขยะมูลฝอยที่มีสารพิษและเชื้อโรค อาจสร้างความเสี่ยงต่อสุขภาพของพนักงาน เช่น การติดเชื้อจากเชื้อโรคในขยะชีวภาพ การได้รับสารเคมีที่เป็นอันตราย หรือการบาดเจ็บจากการวัตถุมีคม นอกจากนี้ ขยะเหล่านี้ยังสร้างปัญหาด้านความปลอดภัย โดยเฉพาะในขั้นตอนการคัดแยกที่ต้องสัมผัสกับขยะโดยตรง นอกจากนี้ ลักษณะงานที่ซ้ำซากและต้องใช้แรงงานหนัก ยังส่งผลต่อความเนื่องอ่อนล้าทั้งทางร่างกายและจิตใจ ทำให้พนักงานในโรงงานขยะต้องทำงานในสภาพที่ไม่เหมาะสมสมต่อคุณภาพชีวิต

ปัญหาดังกล่าวสะท้อนให้เห็นถึงความสำคัญของการจัดการขยะรีไซเคิลอีกด้วย ไม่ว่าจะเป็นการสูญเสียโอกาสในการนำวัสดุกลับมาใช้ใหม่ การเพิ่มปริมาณขยะที่ต้องกำจัดด้วยวิธีการเผาหรือฝังกลบ และการเพิ่มภาระให้กับโรงงานขยะในการจัดการกับขยะที่ปนกัน

แม้ว่าจะมีงานวิจัยและโครงการหลายชิ้นที่พัฒนาและทดลองใช้เทคโนโลยีในการคัดแยกขยะ เช่น การใช้ระบบการประมวลผลภาพ (computer vision) และการเรียนรู้ของเครื่อง (machine learning) เพื่อเพิ่มประสิทธิภาพในการจำแนกประเภทตั้งแต่ต้นมักรับผลกระทบในหลายมิติ ไม่ว่าจะเป็นการสูญเสียโอกาสในการนำวัสดุกลับมาใช้ใหม่ การเพิ่มปริมาณขยะที่ต้องกำจัดด้วยวิธีการเผาหรือฝังกลบ และการเพิ่มภาระให้กับโรงงานขยะในการจัดการกับขยะที่ปนกัน

ทางผู้จัดทำจึงได้เลือกให้เป็นถึงความสำคัญในการพัฒนาและปรับปรุงการคัดแยกขยะรีไซเคิลจากภาพในโรงงานขยะ โดยหนึ่งในแนวทางที่มีความเหมาะสมคือการใช้เทคนิค Deep Learning เช่น Convolutional Neural Network (CNN) และ You Only Look Once (YOLO) ซึ่งเป็นอัลกอริทึมที่มีประสิทธิภาพในการประมวลผลภาพ โดย CNN หรือ โครงข่ายประสาทแบบคอนволูชัน ซึ่งเป็น

โมเดล Deep Learning ที่ได้รับความนิยมในการประมวลผลภาพ โดยทำงานจากการมองภาพเป็นชิ้นส่วนเล็กๆ แล้วค่อยๆ วิเคราะห์คุณสมบัติที่สำคัญ เช่น ขอบ รูปร่าง หรือสี ของวัตถุในภาพ เมื่อโมเดลได้เรียนรู้ลักษณะเหล่านี้แล้ว ก็สามารถจำแนกประเภทของวัตถุในภาพได้ ขณะที่ YOLO เป็นอัลกอริทึมที่สามารถตรวจจับและจำแนกวัตถุในภาพได้อย่างรวดเร็วแบบเรียลไทม์ โดยระบุตำแหน่ง และประเภทของวัตถุพร้อมกันในขั้นตอนเดียว ซึ่งหมายความว่าระบบต้องการความเร็วในการประมวลผล โดยการประยุกต์ใช้ CNN และ YOLO มาใช้ในการคัดแยกขยะรีไซเคิลนั้นสามารถเพิ่มความแม่นยำในการจำแนกประเภทของขยะ และช่วยลดข้อผิดพลาดที่อาจเกิดจากแรงงานคน ทำให้กระบวนการคัดแยกขยะในโรงงานมีความรวดเร็วและปลอดภัยมากยิ่งขึ้น พร้อมทั้งสร้างสภาพแวดล้อมการทำงานที่ปลอดภัยและมีประสิทธิภาพสำหรับพนักงาน โดยยังคงรักษาความสามารถในการแยกขยะรีไซเคิลได้อย่างเต็มที่ นอกจากนี้ ผู้พัฒนาได้คาดหวังว่าจะเป็นต้นแบบเพื่อพัฒนาต่ออยู่ด้วยกับโรงงานขยะหลายแห่งนำไปใช้ เพื่อเสริมสร้างการจัดการขยะที่ยั่งยืนและมีประสิทธิภาพในระดับอุตสาหกรรม

## 1.2 วัตถุประสงค์

โครงการนี้มีเป้าหมายเพื่อสร้างและปรับเปลี่ยนโมเดลการคัดแยกขยะรีไซเคิลจากภาพด้วยเทคนิคการเรียนรู้เชิงลึก ทำให้ผู้ใช้งานสามารถแยกขยะได้อย่างแม่นยำ และเพื่อให้บรรลุเป้าหมายดังกล่าว จึงกำหนดวัตถุประสงค์ของโครงการดังต่อไปนี้

1. เพื่อศึกษาลักษณะของขยะรีไซเคิลประเภทต่างๆ
2. เพื่อศึกษาวิธีการจำแนกขยะที่ถูกต้อง
3. เพื่อสร้างโมเดลสำหรับการจำแนกประเภทของขยะจากภาพถ่ายและปรับเปลี่ยน
4. เพื่อศึกษาโมเดลที่เกี่ยวกับการจำแนกภาพ ได้แก่ YOLO (You Only Look Once) และ CNN (Convolutional Neural Network)

## 1.3 ขอบเขตของโครงการ

เพื่อให้โครงการนี้บรรลุได้ตามวัตถุประสงค์ ทางผู้พัฒนาจึงได้กำหนดขอบเขตของงานดังนี้

1. ประเภทของขยะที่ใช้ในการจำแนก ได้แก่

1.1 กระดาษ

1.2 พลาสติก

1.3 โลหะ

#### 1.4 แก้ว

2. โดยจะทำการเก็บรวบรวมข้อมูลและภาพถ่ายของขยะแต่ละประเภทจาก Kaggle เพื่อนำมาฝึกสอนและสร้างโมเดลการจำแนกขยะจากการพัฒนาด้วยอัลกอริทึม YOLO และ CNN
3. ภาพของขยะที่นำมาใช้ในการฝึกสอนและสร้างโมเดลการจำแนกขยะจากการพัฒนาด้วยอัลกอริทึม YOLO และ CNN เป็นภาพถ่ายของขยะหนึ่งประเภทที่แสดงลักษณะเฉพาะ เช่น รูปร่าง สี หรือพื้นผิวของขยะนั้น จำนวน 1200 รูป และได้ขยายจำนวนภาพเป็น 8031 รูปเพื่อใช้ในการทดลองเพิ่มเติม
4. ภาพของขยะถูกปรับขนาดเป็น (224, 224) พิกเซล เนื่องจากขนาดนี้ช่วยให้โมเดลสามารถประมวลผลข้อมูลได้อย่างมีประสิทธิภาพ

#### 1.4 ประโยชน์ของโครงงาน

1. เป็นเครื่องมือที่ช่วยในการจำแนกขยะรีไซเคิลได้อย่างแม่นยำ
2. ช่วยเพิ่มประสิทธิภาพในการจำแนกขยะ
3. สามารถช่วยลดการพึ่งพาแรงงานมนุษย์ในการคัดแยก ลดความเสี่ยงในการสัมผัสถูกขยะที่อาจเป็นอันตรายหรือมีสารพิษ
4. เป็นต้นแบบที่สามารถนำไปประยุกต์ใช้ในโรงงานคัดแยกขยะ

#### 1.5 ข้อจำกัดของโครงงาน

1. ไม่เดลการคัดแยกสามารถคัดแยกขยะรีไซเคิลได้เพียง 4 ประเภทเท่านั้น คือ แก้ว พลาสติก กระดาษ และโลหะ
2. ไม่เดลสามารถประมวลผลภาพของขยะรีไซเคิลได้เพียง 1 ชิ้นต่อ 1 ภาพเท่านั้น ไม่สามารถประมวลผลภาพขยะหลายชิ้นในภาพเดียวกันได้ เพราะอาจส่งผล กระทบต่อความแม่นยำ
3. ประสิทธิภาพการประมวลผลอาจลดลงหากภาพที่ได้รับมีความลละเอียดต่ำหรือคุณภาพไม่ดี เช่น ภาพที่เบลอมากๆ หรือภาพวัวๆ ที่ไม่ชัดเจน

## บทที่ 2

### วรรณกรรมและงานวิจัยที่เกี่ยวข้อง

#### 2.1 แนวคิดทฤษฎีที่เกี่ยวข้อง

##### 2.1.1 ขยายรีไซเคิล

การรีไซเคิล คือ การนำวัสดุที่ไม่ใช้แล้วหรือของเสียกลับมาผ่านกระบวนการปรับปรุงหรือเปลี่ยนแปลงคุณภาพ เพื่อให้สามารถใช้งานได้ใหม่ในรูปแบบเดิมหรือใกล้เคียงเดิม หรือเปลี่ยนเป็นวัตถุดิบ หรือผลิตภัณฑ์ใหม่ที่มีประโยชน์อีกรั้ง (Recycle Engineering, 2024)



ภาพที่ 2.1 สัญลักษณ์รีไซเคิล

ขยายรีไซเคิลแต่ละประเภท มีดังนี้

##### 2.1.1.1 กระดาษ

กระดาษถือเป็นองค์ประกอบสำคัญของขยะที่เกิดจากบ้านเรือนและสำนักงานเนื่องจากการใช้งานกระดาษมีความเกี่ยวข้องกับชีวิตประจำวันของเรา เช่น หนังสือพิมพ์ กระดาษคอมพิวเตอร์ ถุงกระดาษ และลังกระดาษ กระดาษส่วนใหญ่สามารถนำกลับมาใช้การรีไซเคิลได้



ภาพที่ 2.2 กระดาษ

### 2.1.1.2 พลาสติก

ขยะพลาสติกเพิ่มขึ้นอย่างรวดเร็ว เนื่องจากสินค้าและบรรจุภัณฑ์ส่วนใหญ่ทำจากพลาสติก เช่น ขวดนม กระป๋องยา และขวดน้ำดื่ม พลาสติกเป็นที่นิยมอย่างมากในปัจจุบัน เพราะมีคุณสมบัติที่เหมาะสม เช่น น้ำหนักเบา สีสันสวยงาม ไม่เป็นสนิม ทนทาน และมีหลากหลายประเภท อย่างไรก็ตาม ขยะพลาสติกเหล่านี้หากถูกกำจัดด้วยการเผา จะก่อให้เกิดมลพิษต่อสิ่งแวดล้อม และหากนำไปฝังกลดจะเปลืองพื้นที่ วิธีที่ดีที่สุดและเหมาะสมที่สุดในการจัดการขยะพลาสติกคือการนำกลับมาใช้เชิงลูปเพื่อนำไปซีลใหม่



ภาพที่ 2.3 พลาสติก

### 2.1.1.3 โลหะ

โลหะหลากหลายชนิดสามารถนำกลับมาใช้เชิงลูปได้โดยการหลอมและเปลี่ยนรูปให้กล้ายเป็นผลิตภัณฑ์ใหม่ โดยโลหะแบ่งออกเป็น 3 กลุ่มใหญ่คือ

- โลหะประเภทเหล็ก
- โลหะประเภทอลูมิเนียม
- โลหะมีค่า



ภาพที่ 2.4 โลหะ

#### 2.1.1.4 แก้ว

ขวดแก้วสามารถแบ่งออกเป็นสองประเภทหลัก คือ ขวดแก้วแบบสมบูรณ์และขวดแก้วแบบไม่สมบูรณ์ ขวดแก้วแบบสมบูรณ์หมายถึงขวดที่ยังไม่แตกหรือเสียหาย เช่น ขวดน้ำปลาและขวดเครื่องดื่ม ขวดประเภทนี้จะถูกคัดแยกและส่งกลับไปยังโรงงานเพื่อนำไปทำความสะอาดและใช้ซ้ำอีกครั้ง ส่วนขวดแก้วแบบไม่สมบูรณ์ หรือขวดที่แตกหรือเสียหาย จะถูกแยกประเภทตามสี เช่น ขวดแก้วสีชาและขวดแก้วสีใส เพื่อนำเข้าสู่กระบวนการหลอมเศษแก้วในโรงงานก่อนเข้าสู่กระบวนการผลิตใหม่ในขั้นตอนต่อไป



ภาพที่ 2.5 แก้ว

#### 2.1.2 การประมวลผลภาพ (image processing)

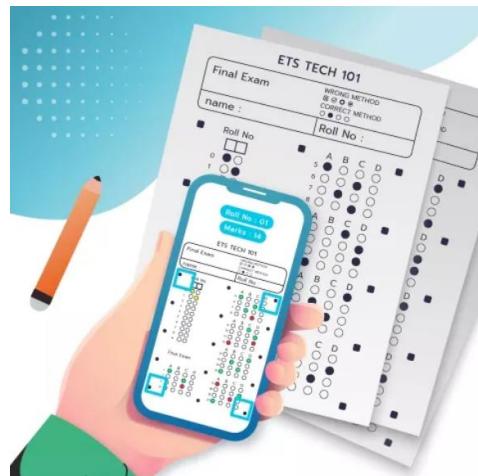
กระบวนการจัดการและวิเคราะห์รูปภาพในรูปแบบดิจิทัลโดยใช้คอมพิวเตอร์ คือ การนำภาพถ่ายหรือภาพที่มีอยู่มาประมวลผลเพื่อแปลงข้อมูลที่อยู่ในภาพให้กลายเป็นข้อมูลเชิงดิจิทัลที่สามารถนำไปใช้งานได้ กระบวนการนี้ครอบคลุมทั้งการดึงข้อมูลในเชิงคุณภาพ เช่น การระบุสีลักษณะพื้นผิว หรือรูปแบบของวัตถุ และข้อมูลในเชิงปริมาณ เช่น ขนาด ความสูง ความกว้าง และรูปร่างของวัตถุในภาพ

กระบวนการดังกล่าวประกอบด้วยขั้นตอนสำคัญคือ การปรับปรุงความคมชัดของภาพและลดหรือกำจัดสัญญาณรบกวนที่อาจส่งผลกระทบต่อความแม่นยำในการประมวลผลภาพ ต่อมาคือ การแยกส่วนของวัตถุเป้าหมายออกจากภาพพื้นหลังอย่างชัดเจน เพื่อนำภาพวัตถุที่ได้มาทำการวิเคราะห์ในเชิงปริมาณ เช่น การวัดขนาด รูปร่าง หรือการศึกษาพิสัยทางการเคลื่อนที่ของวัตถุในภาพ หลังจากนั้น ข้อมูลเชิงปริมาณที่ได้สามารถนำไปใช้ในการวิเคราะห์เพิ่มเติมหรือพัฒนาเป็นระบบอัตโนมัติ เพื่อนำไปประยุกต์ใช้ในงานด้านต่าง ๆ อย่างมีประสิทธิภาพ เช่น การวิจัย การพัฒนาเทคโนโลยี หรือการปรับปรุงกระบวนการในภาคอุตสาหกรรม

ตัวอย่างการประยุกต์ใช้ Image Processing เช่น

### 1. ระบบตรวจสอบกระดาษคำตอบ

ทำงานโดยการนำภาพกระดาษคำตอบที่ถูกต้องมาเปรียบเทียบกับภาพกระดาษคำตอบที่ตรวจสอบ เพื่อวิเคราะห์ว่าตำแหน่งของคำตอบในทั้งสองภาพตรงกันหรือไม่ หากตำแหน่งตรงกัน ระบบจะคำนวณคะแนนให้ตามคำตอบที่ถูกต้อง



ภาพที่ 2.6 ระบบตรวจสอบกระดาษคำตอบ

ที่มา : [ออนไลน์]

<https://techintegration.ets.kmutt.ac.th/content/tech-review/evalbee>

### 2. ระบบตรวจจับใบหน้าในกล้องดิจิทัล

ทำงานโดยการนำทำงานโดยการตรวจสอบส่วนต่าง ๆ ของภาพเพื่อหาลักษณะที่คล้ายกับใบหน้า เมื่อระบบพบตำแหน่งที่มีลักษณะคล้ายใบหน้า กล้องจะทำการโฟกัสไปที่ตำแหน่งนั้นเพื่อให้ภาพมีความคมชัดมากขึ้น โดยจะใช้ข้อมูลเชิงปริมาณ เช่น ระยะห่างระหว่างคิ้ว มุกปาก จมูก โนนกแก้ม และโครงหน้า เพื่อปรับการโฟกัสให้ตรงกับตำแหน่งใบหน้าที่ตรวจจับได้



ภาพที่ 2.7 ระบบตรวจจับใบหน้าในกล้องดิจิตอล

ที่มา : [ออนไลน์] <https://www.falco.co.th/news-update/face-scan/biometrics-face-scan/>

### 2.1.3 การเรียนรู้ของเครื่อง (Machine Learning)

Machine Learning (การเรียนรู้ของเครื่อง) คือ กระบวนการที่ทำให้คอมพิวเตอร์สามารถเรียนรู้และพัฒนาประสิทธิภาพการทำงานของตนเองจากข้อมูลและสภาพแวดล้อมที่ได้รับการเรียนรู้ โดยไม่จำเป็นต้องได้รับการโปรแกรมหรือคำสั่งจากมนุษย์อย่างต่อเนื่อง ระบบที่ใช้เทคนิค Machine Learning จะสามารถปรับตัวและพัฒนาไปตามข้อมูลใหม่ๆ ที่มีการเปลี่ยนแปลง โดยไม่ต้องมีการเขียนโปรแกรมเพิ่มเติมจากมนุษย์ในแต่ละครั้งที่ข้อมูลใหม่เกิดขึ้น กล่าวอีกนัยหนึ่ง คอมพิวเตอร์สามารถตีความข้อมูลและทำการตัดสินใจหรือทำนายผลลัพธ์ได้ด้วยตนเองตามบริบทที่ได้รับการฝึกฝน ทำให้มีความยืดหยุ่นและความสามารถในการรับมือกับข้อมูลรูปแบบใหม่ๆ โดยไม่ต้องพึ่งพาการปรับแต่งจากผู้ใช้งานโดยตรง

โดยหลักการของ Machine Learning จะแบ่งออกเป็น 3 ประเภท ดังนี้

## 1. Supervised Learning

การเรียนรู้แบบมีผู้สอนคือการทำให้คอมพิวเตอร์สามารถหาคำตอบของปัญหาได้ด้วยตัวเองหลังจากที่มันได้รับการฝึกฝนจากชุดข้อมูลตัวอย่าง เมื่อคอมพิวเตอร์ได้รับข้อมูล (Input) เช่น รูปภาพของขยะ ที่เบื้องต้นคอมพิวเตอร์จะไม่สามารถรู้ได้ว่าเป็นขยะประเภทไหน เราจึงต้องสอนให้คอมพิวเตอร์รู้จักระบบท่องขยะนั้นๆ โดยการให้ข้อมูลที่มีการกำหนดป้ายกำกับ (Label) ชัดเจน เช่น ขยะพลาสติก ขยะกระดาษ หรือขยะอาหาร เป็นต้น

## 2. Unsupervised Learning

การเรียนรู้แบบไม่มีผู้สอนคือการที่คอมพิวเตอร์สามารถเรียนรู้และวิเคราะห์ความสัมพันธ์ของข้อมูลได้โดยไม่ต้องมีการกำหนดป้ายกำกับและไม่ต้องมีการฝึกสอน กล่าวคือ เรายังไม่ได้ให้ข้อมูล (Input) กับคอมพิวเตอร์ แต่ไม่บอกว่าข้อมูลนั้นคืออะไร คอมพิวเตอร์จะต้องเรียนรู้จากลักษณะของข้อมูลเอง โดยการค้นหาความสัมพันธ์หรือโครงสร้างที่ซ่อนอยู่ในข้อมูล ตัวอย่างเช่น เมื่อเราป้อนข้อมูลที่เป็นภาพของขยะหลายประเภทเข้าไป แต่ไม่ได้บอกว่าขยะแต่ละชิ้นคือขยะประเภทไหน คอมพิวเตอร์จะทำการวิเคราะห์ข้อมูล (Feature Extraction) เพื่อหาลักษณะร่วมกันของขยะแต่ละประเภท เช่น ขยะที่มีลักษณะคล้ายกัน อาจจะมีสี รูปร่าง หรือวัสดุที่คล้ายกัน จากนั้นคอมพิวเตอร์จะทำการ จัดกลุ่ม (Clustering) ข้อมูลเหล่านั้นออกเป็นกลุ่มๆ เช่น ขยะพลาสติก, ขยะกระดาษ, ขยะอาหาร หรือขยะอื่นๆ ที่มีลักษณะคล้ายกัน โดยไม่จำเป็นต้องบอกล่วงหน้าว่าแต่ละกลุ่มคืออะไร

## 3. Reinforcement Learning

การเรียนรู้แบบเสริมกำลังเป็นวิธีการเรียนรู้ที่คอมพิวเตอร์หรือระบบ (เรียกว่า Agent) เรียนรู้การทำงานหรือการตัดสินใจจากการปฏิสัมพันธ์ (Interaction) กับสิ่งแวดล้อม (Environment) โดยการลองผิดลองถูก ซึ่งทำให้ Agent สามารถปรับปรุงการตัดสินใจและการกระทำการของตนเองให้ดีขึ้นเรื่อยๆ เพื่อให้ได้ผลลัพธ์ที่ดีที่สุดหรือบรรลุเป้าหมายที่กำหนด วิธีการนี้ทำงานโดยที่คอมพิวเตอร์จะทำการเลือกกระทำการอย่างในสถานการณ์หรือระบบจำลอง เช่น การเล่นเกม หรือการควบคุมหุ่นยนต์ โดยการเลือกกระทำการที่นำไปสู่รางวัลหรือผลลัพธ์ที่ต้องการ เมื่อ Agent ทำการที่ถูกต้องจะได้รับรางวัล (Reward) และหากทำการผิดจะได้รับการลงโทษหรือค่าปรับ (Punishment)

#### 2.1.4 การเรียนรู้เชิงลึก (Deep Learning)

คือหนึ่งในวิธีการของ ปัญญาประดิษฐ์ (AI) ที่ช่วยให้คอมพิวเตอร์สามารถประมวลผลข้อมูล ในลักษณะที่คล้ายคลึงกับการทำงานของสมองมนุษย์ โดยใช้แบบจำลองที่สามารถเรียนรู้และจดจำ รูปแบบที่ซับซ้อนในข้อมูลประเภทต่างๆ เช่น รูปภาพ ข้อความ เสียง และข้อมูลอื่นๆ การเรียนรู้ใน ลักษณะนี้ช่วยให้คอมพิวเตอร์สามารถสร้างข้อมูลเชิงลึกและทำการคาดคะเนที่แม่นยำ การใช้ Deep Learning สามารถทำให้คอมพิวเตอร์สามารถทำงานที่ต้องการความฉลาดของมนุษย์ได้โดยอัตโนมัติ เช่น การอธิบายสิ่งที่อยู่ในภาพ หรือการแปลงเสียงจากไฟล์เสียงเป็นข้อความ ซึ่งเป็นงานที่ต้องการ การวิเคราะห์และการเข้าใจบริบทที่ค่อนข้างซับซ้อน

โดยความแตกต่างระหว่าง Deep Learning และ Machine Learning คือ Machine Learning (ML) จะเป็นเทคนิคที่ใช้ในการทำให้คอมพิวเตอร์เรียนรู้จากข้อมูลเพื่อทำการนายหรือ การตัดสินใจโดยไม่ต้องมีการเขียนโปรแกรมใหม่ทุกครั้ง โดยใช้สถิติและคณิตศาสตร์ในการฝึกสอน และพัฒนาโมเดล ML เหมาะสำหรับงานที่มีข้อมูลที่มีโครงสร้างและมีป้ายกำกับ (Label) ชัดเจน เช่น การจำแนกรูปภาพหรือการทำนายผลลัพธ์จากข้อมูลที่มีคุณสมบัติเฉพาะ ส่วน Deep Learning (DL) เป็นสาขานึงของ Machine Learning ที่เน้นการใช้ Neural Networks (เครือข่ายประสาทเทียม) ที่ มีหลายชั้นในการเรียนรู้จากข้อมูลที่มีความซับซ้อน โดยสามารถเรียนรู้รูปแบบที่ซับซ้อนจากข้อมูลที่ ไม่มีโครงสร้าง เช่น รูปภาพ เสียง หรือข้อความได้โดยไม่ต้องแยกคุณสมบัติจากข้อมูลด้วยตนเอง เหมาะสำหรับงานที่ต้องการการประมวลผลข้อมูลที่ซับซ้อนและมีข้อมูลจำนวนมาก เช่น การจำแนก ภาพที่ซับซ้อนหรือการแปลงเสียงเป็นข้อความ

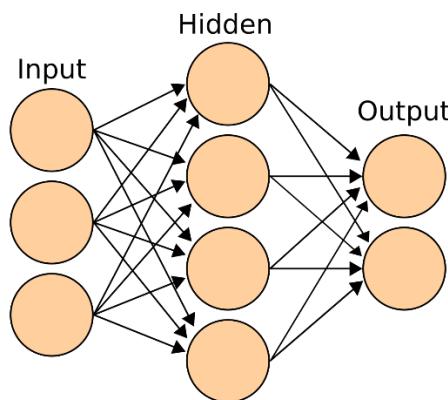
โดยทั่วไป Deep Learning เป็นวิธีการขั้นสูงใน Machine Learning และสามารถใช้เรียนรู้ จากข้อมูลที่ไม่ต้องการการกำหนดคุณสมบัติล่วงหน้า โดยสามารถปรับปรุงการทำงานด้วยการเรียนรู้ จากข้อมูลพลาด ขณะที่ Machine Learning มักจะต้องเลือกและแยกคุณสมบัติจากข้อมูลด้วยตัวเอง ซึ่ง Machine Learning จะใช้ทรัพยากรและข้อมูลน้อยกว่า Deep Learning ที่ต้องการข้อมูลจำนวน มากและทรัพยากรที่สูงกว่าในการฝึกสอนโมเดล

## แนวคิดของโครงข่ายประสาทเทียม (Artificial Neural Networks)

คือโมเดลทางคณิตศาสตร์ที่ได้รับแรงบันดาลใจจากการทำงานของสมองมนุษย์ โดยใช้โครงสร้างของ เซลล์ประสาท (Neurons) ที่เชื่อมโยงกันผ่านลิงค์หรือการเชื่อมต่อ (Connections) ในรูปแบบเครือข่ายที่ซับซ้อน จุดมุ่งหมายคือการจำลองกระบวนการที่สมองมนุษย์ใช้ในการเรียนรู้และตัดสินใจผ่านการประมวลผลข้อมูล

### โครงข่ายประสาทเทียมมีโครงสร้างที่ประกอบด้วย 3 ชั้น ดังนี้

1. **Input Layer** ชั้นแรกที่รับข้อมูลเข้าไปในระบบ ซึ่งอาจเป็นข้อมูลประเภทต่างๆ เช่น ข้อมูลภาพ ข้อความ หรือเสียง ข้อมูลในชั้นนี้จะถูกส่งผ่านไปยังชั้นถัดไป
2. **Hidden Layer** ชั้นที่ใช้ในการประมวลผลข้อมูลที่ได้รับจาก Input Layer โดยการคำนวณตามน้ำหนักที่กำหนดไว้ระหว่างเซลล์ประสาทและการใช้ฟังก์ชันการกระตุ้น (Activation Function) เพื่อช่วยในการตัดสินใจและเรียนรู้จากข้อมูล
3. **Output Layer** ชั้นที่ส่งผลลัพธ์สุดท้ายของการประมวลผล ซึ่งอาจเป็นการทำนายประเภท หรือค่าเฉพาะจากข้อมูลที่ได้รับมา เช่น การจำแนกประเภทของภาพ



ภาพที่ 2.8 โครงสร้างของโครงข่ายประสาทเทียม

## การประยุกต์ใช้ Deep Learning ใน การจัดการขยะรีไซเคิล

การใช้ Deep Learning ใน การจัดการขยะรีไซเคิลกำลังเป็นแนวทางที่ได้รับความสนใจอย่างมากในยุคปัจจุบัน เนื่องจากสามารถช่วยเพิ่มประสิทธิภาพในการคัดแยกขยะและการประมวลผลข้อมูลจากขยะที่มีความซับซ้อน เช่น ขยะที่มีรัศดุลากหลายประเภท การใช้ Deep Learning จะช่วยในการจำแนกประเภทของขยะอย่างอัตโนมัติผ่านการประมวลผลภาพหรือข้อมูลที่เกี่ยวข้องกับขยะชนิดต่างๆ เช่น ขวดพลาสติก กระดาษ แก้ว และโลหะ ด้วยการใช้ Convolutional Neural Networks (CNNs) ที่มีความสามารถในการแยกแยะและจัดลำกษณะของขยะจากราฟถ่ายได้อย่างมีประสิทธิภาพ

โดยหนึ่งในตัวอย่างของการนำ Deep learning ไปใช้เกี่ยวกับจัดการขยะในประเทศไทย คือโครงการของ จุฬาลงกรณ์มหาวิทยาลัย ที่นำเทคโนโลยี AI และ Deep Learning มาใช้ในการนับปริมาณขยะในช่วงเวลาหน้าฝน และติดตามเส้นทางขยะ การใช้ Deep Learning ในโครงการนี้ไม่เพียงแต่ช่วยในด้านการนับปริมาณขยะและการติดตามเส้นทางขยะในช่วงหน้าฝน แต่ยังช่วยในการวางแผนการจัดการขยะได้ดีขึ้นและลดผลกระทบต่อสิ่งแวดล้อมอีกด้วย

## 2.2 เทคโนโลยีที่เกี่ยวข้อง

### 2.2.1 TensorFlow

เป็นแพลตฟอร์มโอเพนซอร์สที่พัฒนาโดย Google Brain Team สำหรับการคำนวณเชิงตัวเลขและการพัฒนา Machine Learning และ Deep Learning โดยเฉพาะ โดยรองรับภาษาหลายภาษา เช่น Python, C++, และ R จุดเด่นของ TensorFlow คือมันสามารถทำงานได้ทั้งในสภาพแวดล้อม CPU, GPU TensorFlow ถูกออกแบบมาเพื่อสร้างและพัฒนามोเดล Deep Learning อย่างเช่น Deep Neural Networks (DNNs) และ Convolutional Neural Networks (CNNs) ซึ่งเหมาะสมกับการประมวลผลข้อมูลที่ซับซ้อน เช่น การจำแนกภาพ หรือการประมวลผลข้อมูลอื่นๆ ที่ต้องการเทคนิคการเรียนรู้ที่ลึกซึ้ง อีกทั้งยังสามารถปรับสเกลการประมวลผลให้เหมาะสมกับ CPUs และ GPUs ทำให้ TensorFlow เป็นเครื่องมือที่เหมาะสมสำหรับการพัฒนามोเดล Machine Learning และ Deep Learning ที่มีประสิทธิภาพ

### 2.2.2 Data Augmentation

การเสริมข้อมูล (Data Augmentation) คือกระบวนการสร้างข้อมูลใหม่จากข้อมูลต้นฉบับโดยการปรับเปลี่ยนข้อมูลเล็กน้อย เช่น การหมุนภาพ (rotation), การพลิกภาพ (flipping), การซูมเข้าออก (zoom), การปรับแสงหรือความคมชัด (brightness/contrast), หรือการเพิ่ม noise ซึ่งช่วยเพิ่มความหลากหลายให้กับชุดข้อมูลโดยไม่ต้องเก็บหรือรวมภาพเพิ่มเติมจากแหล่งใหม่

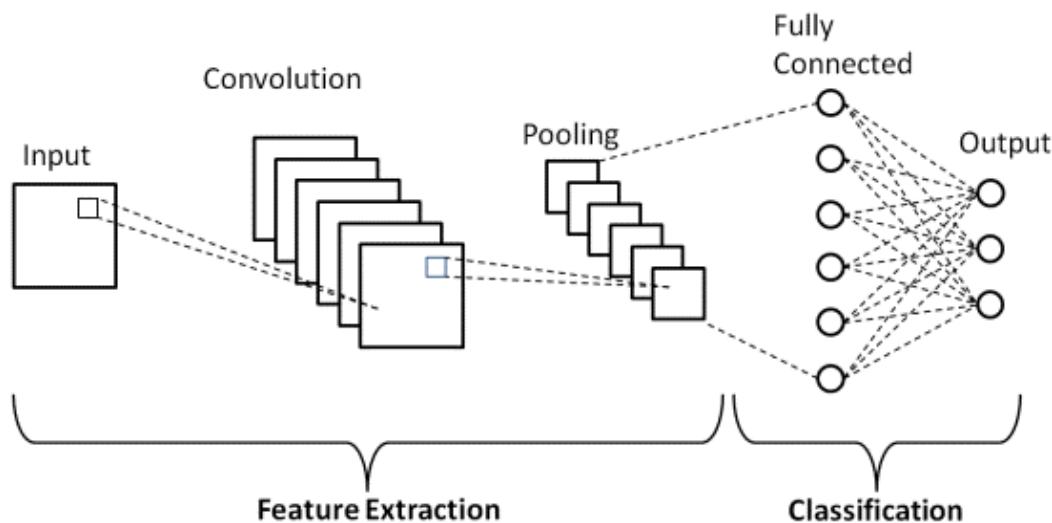
ในงานด้าน Machine Learning (ML) โดยเฉพาะ Deep Learning การมีชุดข้อมูลที่มีจำนวนมากและหลากหลายเป็นสิ่งสำคัญ เนื่องจากแบบจำลองต้องเรียนรู้จากตัวอย่างจำนวนมาก เพื่อให้สามารถทำนายหรือจำแนกข้อมูลใหม่ได้อย่างแม่นยำ อย่างไรก็ตาม การจัดหาข้อมูลปริมาณมากในโลกความจริงมักแพะชิญข้อจำกัดต่าง ๆ เช่น ข้อจำกัดด้านเวลา ทรัพยากร หรือกฎระเบียบด้านข้อมูล ด้วยเหตุนี้ การเสริมข้อมูลจึงเป็นเทคนิคที่ช่วยเพิ่มประสิทธิภาพของโมเดล โดยเฉพาะอย่างยิ่งในกรณีที่ข้อมูลมีจำนวนน้อย เทคนิคนี้จะทำให้โมเดลเรียนรู้จากตัวอย่างที่หลากหลายยิ่งขึ้น แม้ว่าจะอิงจากข้อมูลเดิมก็ตาม

ตัวอย่างเช่น หากมีภาพต้นฉบับอยู่ 1,000 ภาพ และตั้งค่าให้ augment 4 แบบ เช่น หมุน, พลิก, ซูม, และเปลี่ยนความสว่าง ระบบอาจสร้างภาพใหม่ได้เพิ่มอีก 3,000–4,000 ภาพ รวมเป็น 4,000–5,000 ภาพ ซึ่งจะถูกใช้ในการฝึกโมเดล โดยทั่วไปการเสริมข้อมูลจะไม่ทำให้จำนวนไฟล์ภาพในระบบเพิ่มขึ้นแบบชัดเจน แต่จะดำเนินการในขั้นตอนการฝึกโมเดล (training phase) เช่น ผ่านเครื่องมืออย่าง ImageDataGenerator ใน TensorFlow/Keras ซึ่งจะสร้างภาพที่ถูกแปลงแบบอัตโนมัติในแต่ละรอบของการฝึก ทำให้ประยัดพื้นที่จัดเก็บ และช่วยลดความซ้ำซ้อนในชุดข้อมูลเดิม ทั้งนี้ยังมีประโยชน์อย่างมากในงานประเภท Image Classification และ Object Detection ที่ต้องการข้อมูลที่หลากหลายสำหรับการเรียนรู้เชิงลึก

### 2.2.3 Convolutional Neural Networks (CNN)

CNN (Convolutional Neural Networks) เป็นรูปแบบหนึ่งของโครงข่ายประสาทเทียมที่ถูกออกแบบมาเพื่อประมวลผลข้อมูลในรูปแบบของภาพและข้อมูลเชิงพื้นที่อื่น ๆ โครงสร้างหลักที่ทำให้ CNN มีความโดดเด่นคือ Convolutional Layer ซึ่งเป็นเลเยอร์ชนิดพิเศษที่ทำหน้าที่สกัดคุณสมบัติสำคัญของภาพ เช่น เส้นขอบ ขอบเขต หรือลวดลาย โดยกระบวนการนี้จะทำให้ระบบสามารถเรียนรู้คุณลักษณะต่าง ๆ ในภาพได้อย่างละเอียดและแม่นยำ

ใน CNN จะใช้ Convolution Layer ร่วมกับเลเยอร์ชนิดอื่น เช่น Pooling Layer ซึ่งช่วยลดขนาดข้อมูลเพื่อป้องกัน Overfitting และเพิ่มประสิทธิภาพการประมวลผล เลเยอร์เหล่านี้จะถูกจัดกลุ่มและซ้อนกันหลายชั้น (Layer Stacking) โดยอาจปรับ Hyperparameters บางอย่าง เช่น ขนาดของฟิลเตอร์ใน Convolution Layer และจำนวนช่องสัญญาณ (Channel) เพื่อให้เหมาะสมกับลักษณะของข้อมูล กระบวนการจัดเรียงและออกแบบส่วนประกอบต่าง ๆ เหล่านี้เรียกว่า โครงสร้าง (Architecture) ของ CNN ซึ่งเป็นปัจจัยสำคัญต่อความสามารถของโมเดลในการแก้ปัญหาที่ซับซ้อน เช่น การจำแนกภาพและการตรวจจับวัตถุในภาพ

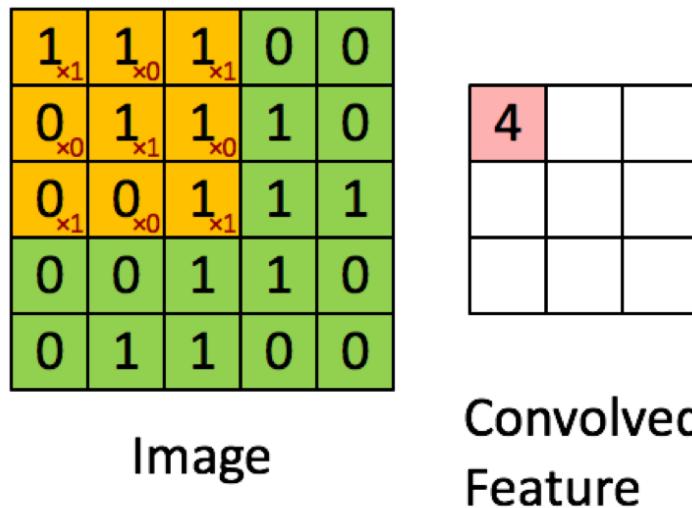


ภาพที่ 2.9 โครงสร้างของ CNN

### ● Convolutional Layer

ในการวิเคราะห์ภาพโดยใช้ CNN จะเริ่มจากการระบุคุณลักษณะสำคัญของภาพ เช่น รูปทรง เส้นขอบ หรือสี ผ่านกระบวนการที่เรียกว่า Convolution โดยใช้ Sliding Window หรือ Filter ซึ่งเป็นเมทริกซ์ขนาดเล็กที่เลื่อนไปตามภาพ ข้อมูลในภาพสีมักแบ่งออกเป็น 3 ช่องสัญญาณหลัก (Channels) ได้แก่ สีแดง (Red), สีน้ำเงิน (Blue) และสีเขียว (Green) โดยในแต่ละพิกเซลจะมีค่าตัวเลขตั้งแต่ 0-255 เพื่อระบุความเข้มของสี หากเป็นภาพขาวดำจะใช้เพียง 1 ช่องสัญญาณ โดยสีดำแทนด้วยค่า 1 และสีขาวแทนด้วยค่า 0

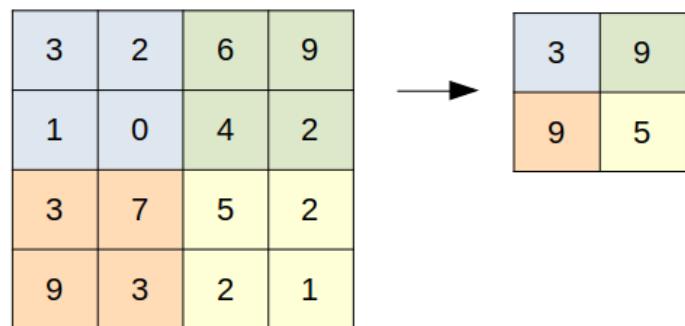
กระบวนการ Convolution เริ่มจากการนำ Filter ซึ่งเป็นเมทริกซ์ เช่น เมทริกซ์ขนาด  $3 \times 3$  ไปคูณจุดต่อจุดกับพิกเซลในภาพ และบันทึกผลลัพธ์ลงในเมทริกซ์ใหม่ที่เรียกว่า Convolved Feature หรือ Feature Map กระบวนการนี้ช่วยเน้นคุณลักษณะเด่น เช่น ขอบของวัตถุ หรือรูปทรงที่สำคัญในภาพ โดยข้อมูลใน Feature Map นี้จะถูกส่งต่อไปยังレイเยอร์ถัดไปของ CNN เพื่อการประมวลผลที่ซับซ้อนยิ่งขึ้น



ภาพที่ 2.10 Convolution layer

- Pooling Layer

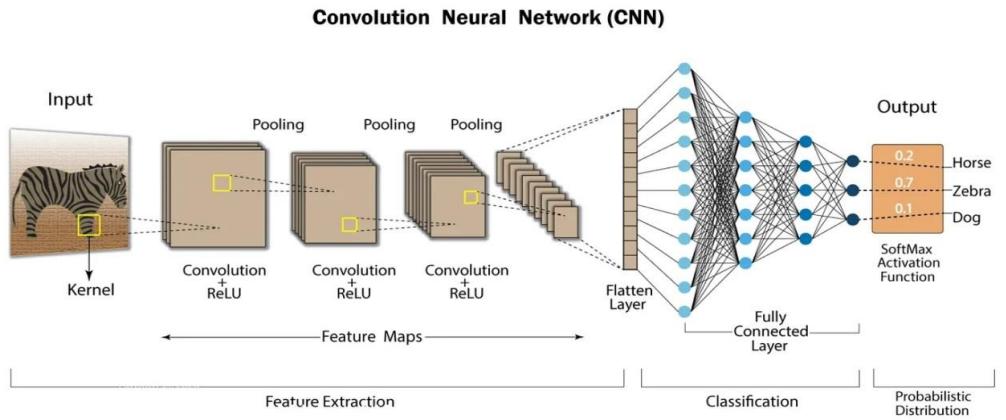
มีหน้าที่หลักในการลดขนาดข้อมูล (Feature Map) โดยยังคงรักษาสาระสำคัญของข้อมูลไว้ เพื่อเพิ่มประสิทธิภาพในการประมวลผลและลดความซับซ้อนของโมเดล กลไกของ Pooling Layer เรียบง่าย โดยจะเลือกค่าที่สำคัญจากกลุ่มพิกเซล เช่น การใช้ค่ามากที่สุด (Max Pooling) หรือค่าเฉลี่ย (Mean Pooling) ในแต่ละส่วนย่อยของภาพ ตัวอย่างเช่น เมื่อใช้ Max Pooling ขนาด  $2 \times 2$  พร้อมกับ Stride (ระยะการเลื่อน) เท่ากับ 2 ระบบจะเก็บเฉพาะค่าสูงสุดในแต่ละ Grid เพื่อนำไปสร้าง Output ที่เล็กลง กระบวนการนี้ช่วยลดขนาดของ Feature Map อย่างมีประสิทธิภาพ ทำให้การคำนวณรวดเร็วขึ้น และช่วยลดความเสี่ยงต่อปัญหา Overfitting โดย Pooling Layer ยังคงรักษารายละเอียดสำคัญที่ช่วยให้โมเดลเข้าใจข้อมูลภาพได้ดี แม้ว่าจะมีการลดมิติของข้อมูลลง



ภาพที่ 2.11 Pooling layer

- Fully Connected Layer

Fully Connected Layer (FC Layer) คือขั้นตอนสุดท้ายในโครงสร้างของ Convolutional Neural Networks ที่ทำหน้าที่รวมรวมและสังเคราะห์ข้อมูลที่ผ่านการประมวลผลจาก Convolution และ Pooling Layer โดยข้อมูลที่ได้รับจะถูกแปลงให้อยู่ในรูปแบบเวกเตอร์ จุดประสงค์ของขั้นตอนนี้เพื่อจะนำลักษณะเด่นๆ ที่สำคัญ ของรูปที่ได้จากกระบวนการก่อนหน้าป้อนเข้าสู่ Neural Network สำหรับการวิเคราะห์เพิ่มเติมและการทำนายผลลัพธ์ เช่น การจำแนกประเภทของรูปภาพ โดยที่ผลลัพธ์ที่ได้จะบ่งบอกถึงความมั่นใจในการทำนายประเภทของข้อมูลนั้นๆ (Confident)

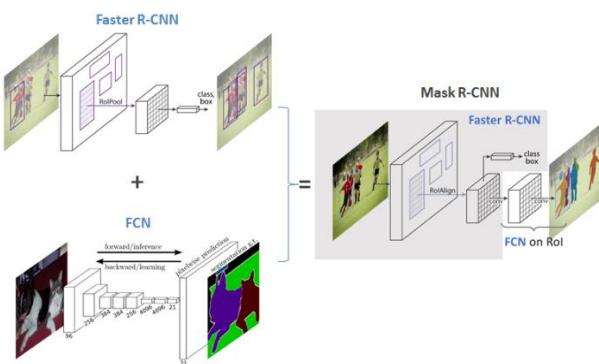


ภาพที่ 2.12 Fully Connected layer

Convolutional Neural Networks (CNN) เป็นเครื่องมือที่เหมาะสมในการจำแนกภาพ ขยะรีไซเคิลเนื่องจากมีความสามารถในการสกัดคุณสมบัติจากข้อมูลภาพได้โดยอัตโนมัติ เช่น ขอบของวัตถุ, รูปทรง และลักษณะเฉพาะที่สำคัญในการแยกแยะสิ่งของต่างๆ ที่เกี่ยวข้องกับการรีไซเคิล เช่น พลาสติก กระดาษ โลหะ หรือแก้ว โดยที่ไม่ต้องใช้คุณสมบัติที่ถูกออกแบบมาโดยมนุษย์ ทำให้ CNN สามารถเรียนรู้การจำแนกภาพในลักษณะที่มีประสิทธิภาพสูงและแม่นยำ

#### 2.2.4 Mask R-CNN (Mask Region-based Convolutional Neural Network)

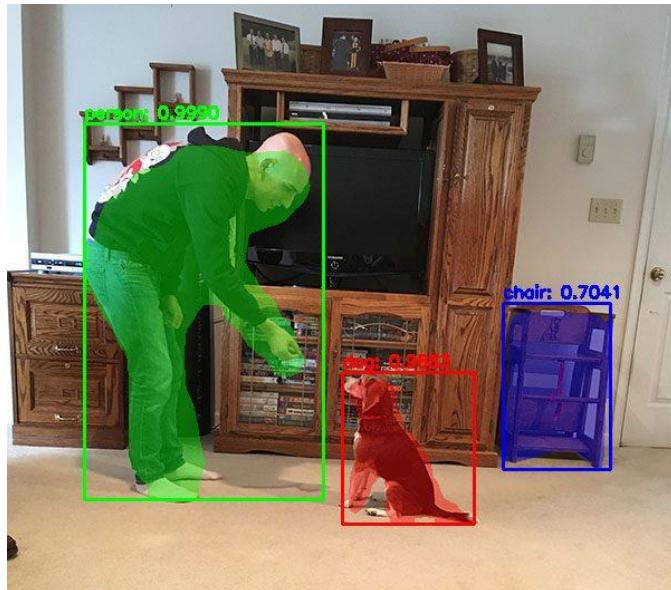
Mask R-CNN คือการรวมกันของ Faster R-CNN และ Fully Convolutional Networks (FCNs) โดยเพิ่มการทำงานที่สำคัญคือ แบ่งกลุ่มวัตถุภายในภาพ (Instance Segmentation) ที่ช่วยให้สามารถระบุตำแหน่งพิกเซลของวัตถุในภาพได้อย่างแม่นยำ การทำงานของ Mask R-CNN ไม่เพียงแต่สามารถตรวจสอบตำแหน่งและชนิดของวัตถุในภาพได้ แต่ยังสามารถแยกพิกเซลที่ตรงกับวัตถุแต่ละชิ้นออกจากกันได้อย่างชัดเจน เช่น การแยกแยะรถยนต์แต่ละคันหรือบุคคลแต่ละคนจากการเดียวกัน



ภาพที่ 2.13 โครงสร้างของ Mask R-CNN

ตัวอย่างการใช้ Mask R-CNN ในการทำ Instance Segmentation ได้แก่:

การตรวจจับและแยกแยะวัตถุในภาพ: Mask R-CNN สามารถใช้ในการแยกแยะวัตถุประเภทต่างๆ ภายในภาพได้อย่างละเอียดมากขึ้น เช่น การแยกแยะบุคคลในภาพ, รถยนต์ หรือสัตว์เลี้ยง โดยสามารถแยกวัตถุแต่ละชิ้นออกจากกันอย่างชัดเจนในระดับพิกเซล



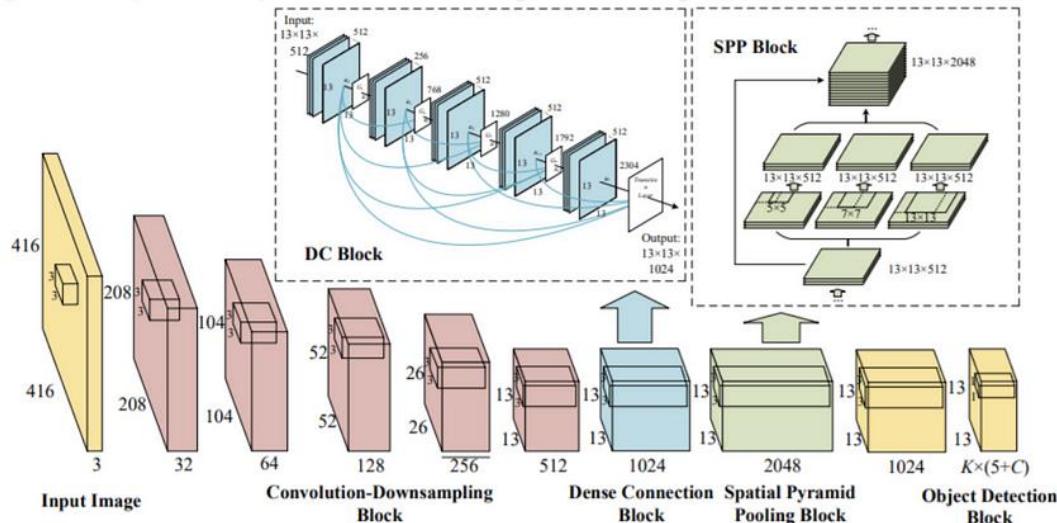
ภาพที่ 2.14 ตัวอย่างการใช้ Mask R-CNN

### 2.2.5 YOLO (You Only Look Once)

YOLO เป็นเทคนิคการตรวจจับวัตถุ (Object Detection) แบบเรียลไทม์ ที่ได้รับความนิยมอย่างกว้างขวางเนื่องจากมีจุดเด่นด้าน ความเร็วในการประมวลผล และ ความแม่นยำ ในการตรวจจับวัตถุในภาพเพียงครั้งเดียว หลักการทำงานของ YOLO คือ การนำภาพเข้ามาผ่านโครงข่ายประสาทเทียม (Neural Network) เพียงครั้งเดียว และระบบจะทำการระบุตำแหน่งของวัตถุโดยการหากรอบ (bounding box) ครอบวัตถุแต่ละชิ้นในภาพ โดยใช้การคำนวณจุดกึ่งกลางของวัตถุแล้ววัดกรอบครอบตามพิกัด

โมเดลพื้นฐานของ YOLO ถูกฝึกมาแล้วกับชุดข้อมูลขนาดใหญ่ เช่น COCO dataset ซึ่งครอบคลุมวัตถุกว่า 80 ประเภท ทำให้สามารถนำมาใช้งานตรวจจับวัตถุต่าง ๆ ได้ทันที พร้อมกับการแสดงผลค่า ความน่าจะเป็น ของการจำแนกแต่ละคลาส โดยข้อดีที่โดดเด่นของ YOLO คือ ความสามารถในการตรวจจับวัตถุที่ซ้อนกัน หรืออยู่ใกล้กันในภาพได้อย่างมีประสิทธิภาพ เนื่องจาก

YOLO ใช้โครงสร้างการแบ่งภาพออกเป็นกริด (grid) หลายระดับ ซึ่งในแต่ละเลเยอร์ (Layer) จะมีขนาดเล็กลง ทำให้สามารถจับรายละเอียดของวัตถุได้ดีขึ้น

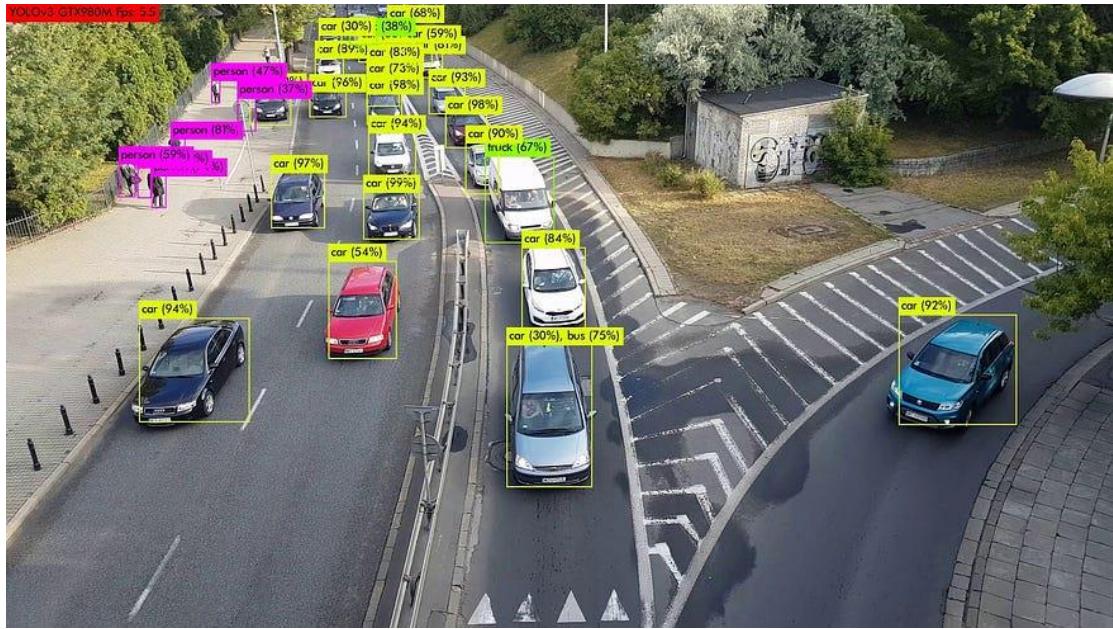


ภาพที่ 2.15 ตัวอย่างองค์ประกอบสถาปัตยกรรมโมเดลตรวจจับวัตถุ YOLO

### ตัวอย่างการใช้ YOLO ในการทำ Object Detection ได้แก่:

ภาคการจราจร (Traffic Sector) สามารถประยุกต์ใช้สำหรับการตรวจจับและวิเคราะห์ข้อมูลเกี่ยวกับยานพาหนะบนท้องถนน เช่น การตรวจจับจำนวนรถยนต์บนถนนในช่วงเวลาต่าง ๆ เพื่อประเมินความหนาแน่นของการจราจร นำไปสู่การจัดการระบบจราจรอย่างมีประสิทธิภาพมากขึ้น

นอกจากนี้ ระบบกล้องวงจรปิดที่ติดตั้งตามถนนหรือสี่แยกสามารถประยุกต์ใช้ YOLO เพื่อตรวจจับวัตถุและติดตามพฤติกรรมของรถยนต์ เช่น ตรวจสอบว่ารถขับเข้าใกล้เส้นทึบหรือเส้นท้ามแข้งหรือไม่ หรือใช้ตรวจจับความเร็วของรถร่วมกับระบบเซนเซอร์ เพื่อตรวจสอบการฝ่าฝืนกฎจราจร การใช้ YOLO ในลักษณะนี้สามารถเชื่อมต่อกับระบบควบคุมสัญญาณไฟจราจรหรือระบบแจ้งเตือนอัตโนมัติเพื่อเพิ่มความปลอดภัยและความมีประสิทธิภาพของการจัดการจราจรในเมืองใหญ่ได้อย่างมีประสิทธิผล



ภาพที่ 2.16 ตัวอย่างการใช้ YOLO ในภาคการจราจร

## 2.2.6 Roboflow

Roboflow คือแพลตฟอร์มออนไลน์ที่ช่วยในการจัดการและเตรียมข้อมูลภาพสำหรับงานด้าน Computer Vision ไม่ว่าจะเป็นการตรวจจับวัตถุ (Object Detection), การจำแนกภาพ (Image Classification) หรือการแบ่งแยกวัตถุจากพื้นหลัง (Image Segmentation) ผู้ใช้งานสามารถอัปโหลดภาพ ติดป้ายกำกับ (Annotation), ปรับแต่งข้อมูล (Data Augmentation) และแปลงชุดข้อมูลให้อยู่ในรูปแบบที่รองรับโดยเดลต่า ๆ เช่น YOLO, COCO, PASCAL VOC หรือ TensorFlow Object Detection API



ภาพที่ 2.17 Roboflow

## ข้อดีและประสิทธิภาพของ Roboflow

### 1. รองรับหลายประเภทของการติดป้ายกำกับ (Annotation)

Roboflow สามารถรองรับงานประเภทต่าง ๆ ได้แก่ การตรวจจับวัตถุ (Object Detection), การจำแนกภาพ (Image Classification) และการแบ่งแยกวัตถุจากภาพ (Segmentation) ทำให้เป็นเครื่องมือที่ครอบคลุมและยืดหยุ่นสำหรับการจัดเตรียมข้อมูลในงาน Computer Vision หลากหลายรูปแบบ

### 2. ใช้งานร่วมกับเครื่องมือพัฒนาโมเดล AI ได้หลากหลาย

Roboflow สามารถเชื่อมต่อและส่งออกข้อมูลในรูปแบบที่รองรับเฟรมเวิร์กยอดนิยม เช่น TensorFlow, PyTorch และ Keras ช่วยให้นักพัฒนาสามารถนำข้อมูลไปใช้งานต่อกับเครื่องมืออื่น ๆ ได้

### 3. รองรับการเตรียมข้อมูลที่มีคุณภาพสูง

ด้วยเครื่องมือและฟีเจอร์ต่าง ๆ ที่ครบครัน Roboflow ช่วยให้ผู้ใช้งานสามารถสร้างชุดข้อมูลที่มีคุณภาพ เหมาะสมสำหรับการฝึกโมเดล AI ที่มีประสิทธิภาพสูง

Roboflow เป็นเครื่องมือที่ช่วยในการติดป้ายกำกับข้อมูลภาพ เพื่อใช้ฝึกโมเดล

ปัญญาประดิษฐ์ โดยเฉพาะงานตรวจจับวัตถุ ซึ่งช่วยให้ผู้ใช้งานสามารถสร้างชุดข้อมูลคุณภาพสูงได้อย่างรวดเร็วและง่าย โดยมีฟีเจอร์สำหรับวัดกรอบล้อมรอบวัตถุ (Bounding Box) และกำหนดป้ายกำกับ (Label) ให้กับวัตถุในภาพ การใช้ Roboflow ช่วยประหยัดเวลาในการเตรียมข้อมูล และเพิ่มประสิทธิภาพในการพัฒนาโมเดล AI เมื่อชุดข้อมูลพร้อมแล้ว สามารถนำไปฝึกโมเดลเพื่อใช้ในแอปพลิเคชันต่าง ๆ เช่น การตรวจจับวัตถุ การจำแนกภาพ และการแปลงข้อความเป็นภาพ เป็นต้น

## 2.2.7 Google colab

Google Colab (Colaboratory) เป็นแพลตฟอร์มที่ช่วยให้ผู้ใช้สามารถเขียนและรันโปรแกรมภาษา Python ผ่านเว็บเบราว์เซอร์ได้โดยไม่ต้องติดตั้งซอฟต์แวร์เพิ่มเติมหรือกำหนดค่าระบบใด ๆ อีกทั้งยังรองรับการใช้งานหน่วยประมวลผลกราฟิก (GPU) ซึ่งช่วยเพิ่มประสิทธิภาพในการประมวลผลสำหรับงานที่ต้องการการคำนวนสูง เช่น งานด้าน Machine Learning หรือ Deep Learning โดย Colab notebook เป็นรูปแบบของ Jupyter notebook ที่ทำงานบนเว็บเบราว์เซอร์ และยังสามารถเชื่อมต่อและเข้าถึงบริการต่าง ๆ ของ Google ได้อย่างสะดวก



ภาพที่ 2.18 Google colab

#### ข้อดีและประสิทธิภาพของ Google colab

1. Google Colab ให้บริการฟรี ทำให้ผู้ใช้งานสามารถใช้งานได้โดยไม่ต้องเสียค่าใช้จ่าย
2. ไม่จำเป็นต้องตั้งค่าหรือกำหนดค่าใด ๆ ก่อนใช้งาน สามารถเริ่มเขียนโปรแกรมได้ทันที ผ่านเว็บเบราว์เซอร์
3. รองรับการเข้าถึงหน่วยประมวลผลกราฟิก (GPU) เพื่อเพิ่มประสิทธิภาพในการประมวลผลงานที่ต้องใช้พลังการคำนวนสูง
4. สามารถเชื่อมต่อและใช้งานบริการต่าง ๆ ของ Google ได้อย่างสะดวก เช่น Google Drive, BigQuery และอื่น ๆ

#### ข้อเสียของ Google colab

1. การใช้งาน Google Colab จำเป็นต้องเชื่อมต่ออินเทอร์เน็ตตลอดเวลา ทำให้ไม่สามารถทำงานแบบออฟไลน์ได้
2. มีข้อจำกัดด้านหน่วยความจำและระยะเวลาการรันโปรแกรม (runtime) ซึ่งอาจไม่เหมาะสมสำหรับงานที่ต้องการประมวลผลข้อมูลขนาดใหญ่หรืองานที่ต้องใช้เวลาประมวลผลนาน

### 2.2.8 Python

Python เป็นภาษาการเขียนโปรแกรมที่ได้รับความนิยมอย่างแพร่หลาย โดยเฉพาะในงานพัฒนาเว็บแอปพลิเคชัน วิทยาศาสตร์ข้อมูล และ Machine learning เนื่องจากมีความสะดวกในการเรียนรู้และใช้งาน ทำให้สามารถทำงานได้บนแพลตฟอร์มต่างๆ ได้ง่ายและรวดเร็ว นอกจากนี้ Python ยังเป็นซอฟต์แวร์โอเพนซอร์สที่สามารถดาวน์โหลดและใช้งานได้ฟรี โดยสามารถทำงานร่วมกับระบบปฏิบัติการหลายประเภทได้อย่างมีประสิทธิภาพ

ข้อดีของ Python คือมีไลบรารีจำนวนมากที่ช่วยให้นักพัฒนาสามารถใช้ฟังก์ชันที่เขียนไว้ล่วงหน้าในการจัดการงานที่ซับซ้อน เช่น การพัฒนาเว็บแอปพลิเคชัน โดยสามารถใช้เฟรมเวิร์กต่างๆ ที่มาพร้อมเครื่องมือช่วยในการสร้างแอปพลิเคชันได้อย่างรวดเร็ว โดยไม่ต้องเขียนโค้ดใหม่ทั้งหมด รวมถึงเครื่องมือสำหรับการทดสอบซอฟต์แวร์ที่ช่วยทำให้กระบวนการพัฒนาง่ายและเร็วขึ้น



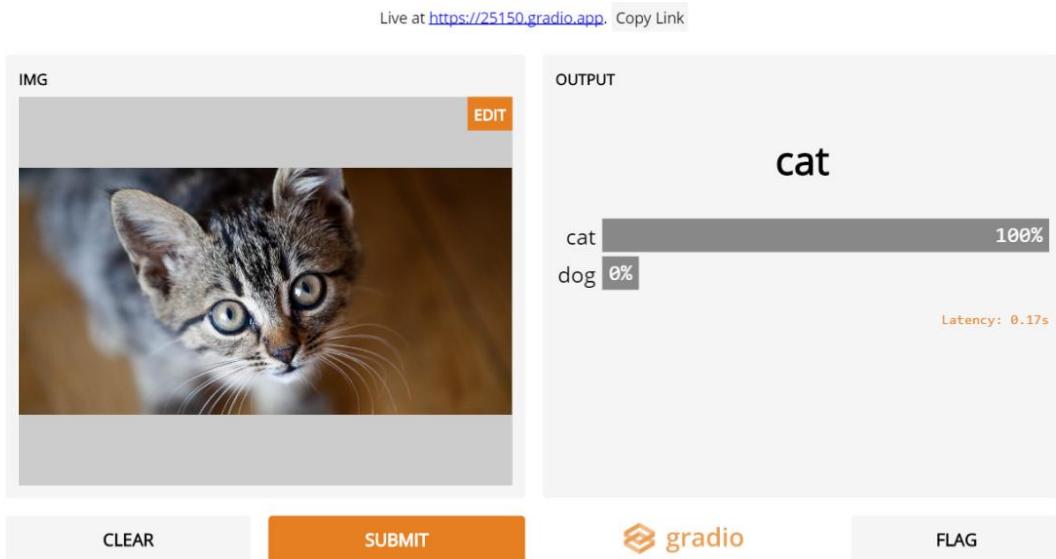
ภาพที่ 2.19 Python

### 2.2.9 Gradio

Gradio เป็นเครื่องมือโอเพนซอร์สที่ใช้สำหรับสร้างอินเทอร์เฟซบนเว็บแอปพลิเคชัน (Web Application Interface) เพื่อสนับสนุนการทำงานของโมเดลปัญญาประดิษฐ์ (AI) หรือโมเดลการเรียนรู้ของเครื่อง (Machine Learning) ในรูปแบบที่สามารถโต้ตอบกับผู้ใช้งานได้ (Interactive) โดยไม่จำเป็นต้องพัฒนาเว็บแอปพลิเคชันจากศูนย์หรือใช้เฟรมเวิร์กที่ซับซ้อน

โดยเครื่องมือนี้ถูกออกแบบมาเพื่ออำนวยความสะดวกให้กับนักพัฒนาและนักวิจัยในการทดสอบ ตรวจสอบ และนำเสนอผลลัพธ์ของโมเดลผ่านหน้าเว็บที่สามารถป้อนข้อมูลและแสดงผลได้อย่างเป็นระบบ ผู้ใช้งานสามารถอัปโหลดข้อมูล เช่น รูปภาพ ข้อความ หรือเสียง เพื่อให้โมเดลทำงานประมวลผล และแสดงผลลัพธ์กลับมาในรูปแบบที่เข้าใจง่าย

Gradio รองรับการทำงานร่วมกับไลบรารีโมเดลยอดนิยม เช่น TensorFlow, PyTorch, Scikit-learn และ Hugging Face ซึ่งช่วยให้การนำโมเดลเข้าสู่ระบบการทดสอบและสาธิตทำได้อย่างรวดเร็วและมีประสิทธิภาพ นอกจากนี้ Gradio ยังสามารถเผยแพร่เว็บแอปพลิเคชันไปยังสาธารณะผ่านทาง Public URL ได้ทันที จึงเหมาะสมสำหรับการใช้งานในงานนำเสนอ การสาธิตผลงาน หรือการทดสอบโมเดลกับผู้ใช้งานภายนอก



ภาพที่ 2.20 ตัวอย่างการใช้งาน Gradio

### 2.3 การวัดประสิทธิภาพและการประเมินผล

#### 1.) Accuracy (ความแม่นยำ)

จำนวนของผลลัพธ์ที่ถูกต้องทั้งหมดในชุดข้อมูล ซึ่งค่าความถูกต้องนี้เป็นค่าที่สามารถบ่งบอกได้ถึงประสิทธิภาพและความสามารถของเครื่องมือที่ใช้ในกระบวนการการวัดได้โดยคำนวณจากจำนวนการทำนายที่ถูกต้องหารด้วยจำนวนการทำนายทั้งหมด ดังสมการ

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

ค่าต่าง ๆ ที่ใช้ในสูตร:

**True Positives (TP):** จำนวนกรณีที่โมเดลทำนายประเภทของขยะถูกต้องและเป็นจริง

**True Negatives (TN):** จำนวนกรณีที่โมเดลทำนายประเภทของขยะไม่ถูกต้องและเป็นจริง

**False Positives (FP):** จำนวนกรณีที่โมเดลทำนายว่าขยะถูกต้อง แต่จริง ๆ แล้วไม่ถูกต้อง

**False Negatives (FN):** จำนวนกรณีที่โมเดลทำนายว่าขยะไม่ถูกต้อง แต่จริง ๆ แล้วถูกต้อง

#### 2.) Loss Function (ฟังก์ชันการสูญเสีย)

Loss function คือฟังก์ชันที่ใช้ในการวัดความแตกต่างระหว่างค่าที่ทำนายโดยโมเดล (Predicted Value) กับค่าจริง (Actual Value) โดยทั่วไปแล้วจะใช้ Cross-Entropy Loss สำหรับงานจำแนกประเภท (Classification) เช่น การคัดแยกขยะ

สูตรของ Cross-Entropy Loss สำหรับการจำแนกประเภทหลายคลาส (Multiclass Classification):

$$\text{Loss} = \sum_{i=1}^C y_i \log(\hat{y}_i)$$

ค่าต่าง ๆ ที่ใช้ในสูตร:

$C$  คือ จำนวนคลาส ทั้งหมด

$y_i$  คือ ค่าจริง (True label) สำหรับคลาสที่  $i$

- $y_i = 1$  หาก  $i$  คือคลาสที่ถูกต้อง (True class)
- $y_i = 0$  หาก  $i$  ไม่ใช่คลาสที่ถูกต้อง

$\hat{y}_i$  คือ ค่าความน่าจะเป็น ที่โมเดลทำนายว่าเป็นคลาส  $i$  ค่านี้จะอยู่ในช่วง 0 ถึง 1

และผลรวมของ  $\hat{y}_i$  ทั้งหมดจะต้องเท่ากับ 1 สำหรับแต่ละตัวอย่าง

$\log$  คือ พังก์ชันลอการิึม (โดยปกติจะใช้ลอการิึมฐาน  $e$ )

### 3.) Precision

ค่า Precision เป็นค่าที่ใช้ในการวัดความแม่นยำของผลลัพธ์ที่ถูกระบุว่าเป็นบวก (Positive) โดยแสดงให้เห็นว่าในจำนวนทั้งหมดที่ถูกทำนายว่าเป็นบวกนั้น มีจำนวนที่ทำนายถูกต้องจริงเท่าใด ซึ่งค่า Precision ที่สูงแสดงถึงความน่าเชื่อถือของผลลัพธ์ โดยสามารถคำนวณได้จากอัตราส่วนระหว่างจำนวนผลลัพธ์ที่ทำนายถูกต้องจริง (True Positive) ต่อจำนวนผลลัพธ์ทั้งหมดที่ถูกทำนายว่าเป็นบวก (True Positive + False Positive) ดังสมการ

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

ค่าต่าง ๆ ที่ใช้ในสูตร:

**True Positives (TP):** จำนวนกรณีที่โมเดลทำนายว่าเป็นบวก และผลลัพธ์นั้นถูกต้องจริง

**False Positives (FP):** จำนวนกรณีที่โมเดลทำนายว่าเป็นบวก แต่ผลลัพธ์นั้นไม่ถูกต้อง

#### 4.) Recall

การคำนวณค่า Recall ช่วยให้เราเข้าใจถึงประสิทธิภาพของระบบหรือวิธีการตรวจจับและระบุวัตถุในภาพ โดยวัดว่าในจำนวนวัตถุที่เราสนใจเท่านั้น ระบบสามารถตรวจจับหรือระบุได้ถูกต้องมากน้อยเพียงใด โดยแสดงอัตราส่วนของจำนวนข้อมูลที่เป็นบวกจริง ๆ ซึ่งถูกตรวจจับได้อย่างถูกต้อง ต่อจำนวนข้อมูลที่เป็นบวกทั้งหมดในชุดข้อมูล ดังสมการ

$$\text{Recall} = \frac{TP}{TP + FN}$$

ค่าต่าง ๆ ที่ใช้ในสูตร:

**True Positives (TP):** จำนวนกรณีที่โมเดลทำนายว่าเป็นบวก และผลลัพธ์นั้นถูกต้องจริง

**False Negatives (FN):** จำนวนกรณีที่โมเดลพลาดไม่สามารถทำนายว่าเป็นบวก ทั้งที่จริงแล้วเป็นบวก

#### 5.) F1-score

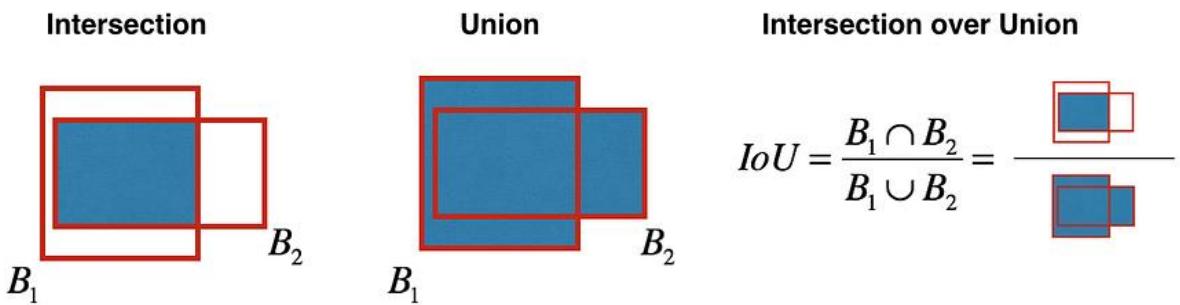
F1 Score เป็นค่าที่ใช้วัดความสมดุลระหว่างค่า Precision และ Recall ซึ่งจะช่วยประเมินผลได้อย่างมีประสิทธิภาพมากยิ่งขึ้น โดยคำนวณจากค่าเฉลี่ย harmonic ของ Precision และ Recall ดังสมการ

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

ค่า F1 Score จะอยู่ในช่วง 0 ถึง 1 โดยค่าที่ใกล้ 1 แสดงถึงโมเดลที่มีทั้ง Precision และ Recall สูง ซึ่งหมายถึงการทำนายที่แม่นยำและครอบคลุม

### 6.) Intersection over union(IoU)

IoU คือค่าสัดส่วนของพื้นที่ที่วัตถุจริงและพื้นที่ที่ไม่เดลทำนายทับซ้อนกัน เทียบกับพื้นที่ทั้งหมดของวัตถุจริงและพื้นที่ที่ไม่เดลทำนายรวมกัน โดยค่า IoU จะช่วยวัดความแม่นยำในการทำนายตำแหน่งและขนาดของวัตถุในภาพ ค่า IoU ที่สูงแสดงว่าไม่เดลทำนายตำแหน่งได้ใกล้เคียงกับวัตถุจริงมาก ค่าของ IoU อยู่ในช่วง 0 ถึง 1 โดยค่าใกล้ 1 แสดงถึงการทำนายตำแหน่งที่แม่นยำ



ภาพที่ 2.21 สูตรการคำนวณหาค่า Intersection over union

### 7.) mAP50 และ mAP50-95

mAP (mean Average Precision) เป็นค่าที่ใช้ประเมินผลการตรวจจับวัตถุ โดยคำนวณค่าความแม่นยำเฉลี่ย (Average Precision: AP) ของแต่ละคลาส และหาค่าเฉลี่ยรวม

- mAP@50 หมายถึงค่าความแม่นยำเฉลี่ยเมื่อกำหนดเกณฑ์ IoU  $\geq 0.50$
- mAP@50–95 หมายถึงค่าเฉลี่ยของ AP ที่คำนวณจากหลายเกณฑ์ของ IoU ตั้งแต่ 0.50 ถึง 0.95

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

โดยที่ N คือจำนวนคลาส และ  $AP_i$  คือค่า Average Precision ของแต่ละคลาส ค่า mAP ที่สูง แสดงถึงโมเดลที่มีประสิทธิภาพในการตรวจจับและระบุตำแหน่งของวัตถุในภาพได้ดี

## 2.4 งานวิจัยที่เกี่ยวข้อง

### 2.4.1 การจำแนกพระเครื่องด้วยการเรียนรู้เชิงลึก

อรรถพล เรืองสุข (2564) ได้นำเสนองานวิจัยเรื่องการจำแนกพระเครื่องด้วยการเรียนรู้เชิงลึก เพื่อแยกแยะลักษณะและประเภทของพระเครื่องจากภาพถ่าย โดยเป็นโมเดลแบบ 2 ขั้นตอน ขั้นตอนแรกเป็นการทำนายชื่อพระเครื่องโดยใช้โครงข่ายประสาทเทียมที่ประยุกต์จากโครงข่าย Xception เมื่อทราบชื่อของพระเครื่องนั้นแล้ว ขั้นตอนต่อไปจะทำนายรุ่นย่อยของพระเครื่อง โดยทำการแยกการสร้างโมเดลสำหรับแต่ละชื่อพระเครื่อง ซึ่งประยุกต์มาจากโครงข่ายประสาทเทียม CNN พบว่าผลการทดสอบให้ความแม่นยำสูงทั้งการทำนายชื่อพระเครื่อง และการทำนายชื่อพระเครื่องรุ่นย่อย ซึ่งให้ความแม่นยำสูงกว่าโมเดล CNN

#### 2.4.1.1 คุณสมบัติของการจำแนกพระเครื่องด้วยการเรียนรู้เชิงลึก

1. สามารถระบุชื่อพระเครื่องได้อย่างถูกต้อง
2. สามารถจำแนกพระเครื่องในรุ่นย่อยได้อย่างมีประสิทธิภาพ

#### 2.4.1.2 ข้อจำกัดของการจำแนกพระเครื่องด้วยการเรียนรู้เชิงลึก

1. ประสิทธิภาพของการจำแนกพระเครื่องยังมีความผิดพลาดในการทำนาย โดยเฉพาะในกรณีพระพึงซึ่งมีรอยแตกและไม่สมบูรณ์ ดูคล้ายคลึงกับพระเครื่องรุ่นอื่น จึงต้องเพิ่มตัวอย่างของข้อมูลภาพในลักษณะนี้เพิ่มขึ้นไปในอนาคต
2. ความท้าทายในการเก็บรวบรวมข้อมูลเนื่องจากรูปพระเครื่องโดยเฉพาะรุ่นเก่าๆ ขาดข้อมูลได้ยาก

### 2.4.2 การจำแนกภาพหมวดแบบเซ็ตเปิดด้วยโครงข่ายประสาทเทียมแบบคอนโวลูชัน

ศุภณัฐ จินตวัฒน์สกุล (2562) ได้นำเสนองานวิจัยเรื่องการจำแนกภาพหมวดแบบเซ็ตเปิดด้วยโครงข่ายประสาทเทียมแบบคอนโวลูชัน โดยขั้นตอนการทดสอบจะมีภาพเครื่องดื่มคลาสหลักอยู่ 3 class ที่ใช้ในการฝึกฝน และมีภาพประเภทที่อยู่นอกเหนือจากการฝึกฝน โดยเป้าหมายคือต้องการจำแนกภาพออกเป็นคลาสตามที่ได้ฝึกฝนเอาไว้ แต่ถ้าหากภาพนั้นไม่อยู่ในคลาสใดๆเลย จะต้องจำแนกออกมาระบบที่ไม่รู้จัก โดยสร้างแบบจำลองการจำแนกภาพออกเป็น 3 โครงสร้าง ดังนี้

1. N-Binary Classifier เป็นการฝึกแบบจำลองเท่ากับจำนวนประเภท (N Classifier) โดยแต่ละโมเดลจะใช้ข้อมูลภาพ 2 ประเภท (Binary Classification) คือภาพประเภทที่สนใจ 1 คลาส เท่านั้น และภาพประเภทอื่นๆ ซึ่งทำให้แต่ละแบบจำลองจะรับผิดชอบในการทำนายเพียงแค่ 1 คลาส ซึ่งคำตอบที่ได้จะเป็นใช่หรือไม่ใช่คลาสนั้นเท่านั้น

ด้วยวิธีการฝึกฝนแบบจำลองแบบนี้ ทำให้ได้แบบจำลองทั้งหมด N ชุดเท่ากับจำนวนคลาสที่มี (known Classes)

2. N+1 Class Classifier เป็นการฝึกแบบจำลองโดยใช้ข้อมูลภาพที่รู้จักทั้งหมด N คลาส (known Classes) และข้อมูลภาพคลาสเพิ่มเติม (Additional class) อีก 1 คลาส ซึ่งจะเป็นภาพที่อยู่นอกเหนือจาก N ประเภท (N Classes)
3. Two Layer Classification จะประกอบขึ้นด้วยแบบจำลองที่แบ่งออกเป็น 2 ชั้น ในชั้นแรกจะเป็นการตรวจสอบว่าภาพที่ใช้คำนายนี้เป็นภาพที่เป็นภาพที่อยู่ในคลาสใดๆ ในชั้นตอนการฝึกฝนหรือไม่ โดยใช้วิธีการตรวจหาสิ่งผิดปกติ (Anomaly detection) หากภาพใดเป็นภาพผิดปกติจะถูกจำแนกเป็นคลาสที่ไม่รู้จัก (unknown class) แต่หากภาพใดเป็นภาพปกติจะถูกส่งไปยังชั้นที่สองเพื่อทำการจำแนกภาพออกเป็นคลาส โดยผลการทดลองพบว่า N+1 Class Classifier มีความแม่นยำสูง 92% เนื่องจากวิธีการอื่นๆ

#### **2.4.2.1 คุณสมบัติของการจำแนกภาพขวดแบบเซ็ตเปิดด้วยโครงข่ายประสาทเทียมแบบคอนโวลูชัน**

1. สามารถจำแนกภาพขวดเครื่องดื่มออกเป็นประเภทที่กำหนด (Target class) และสามารถระบุว่าเป็นประเภทที่ไม่รู้จัก (Unknown) เมื่อภาพที่ใช้คำนายนี้ได้อยู่ในกลุ่มของคลาสใดในชั้นตอนการฝึกฝนแบบจำลอง
2. สามารถหัววิธีที่เหมาะสมในการจำแนกภาพที่ไม่รู้จัก (Unknown class) ซึ่งเป็นภาพที่ไม่อยู่ในคลาสใดในชั้นตอนการฝึกฝนแบบจำลอง

#### **2.4.2.2 ข้อจำกัดของการจำแนกภาพขวดแบบเซ็ตเปิดด้วยโครงข่ายประสาทเทียมแบบคอนโวลูชัน**

งานวิจัยนี้ได้สนใจเฉพาะวิธีการสำหรับโครงข่ายประสาทเทียมเท่านั้น จึงไม่มีการ pre-process ภาพ เช่น การปรับสี แสง ความคมชัด ตัดสัญญาณรบกวนเพื่อให้มีข้อมูลที่ดี ที่สุดในการสร้างแบบจำลอง

### 2.4.3 การพัฒนาและประยุกต์ใช้เทคนิค Convolutional Neural Network ในการจำแนกขยะชายฝั่งทะเล

รัชธรรม พุ่นลำพูน (2566) ได้นำเสนองานวิจัยเรื่องการพัฒนาและประยุกต์ใช้เทคนิค Convolutional Neural Network ในการจำแนกขยะชายฝั่งทะเลด้วยข้อมูลภาพถ่ายของขยะบริเวณชายฝั่งทะเล และการติดตามการแพร่กระจายของชนิดของขยะพลาสติกในทะเล บริเวณชายฝั่งทะเลอ่าวไทยตะวันออก โดยศึกษาชนิดของขยะพลาสติกที่ตรวจพบบริเวณชายฝั่ง จากนั้นทำการจำแนกภาพชายฝั่งทะเลด้วยเทคนิค Convolutional Neural Network (CNN) ซึ่งจากการทดสอบพบว่าเทคนิค CNN สามารถจำแนกขยะชายฝั่งทะเลทั้ง 4 ประเภทออกจากกันได้ดีโดยชุดข้อมูล Training set มีค่าความถูกต้องเฉลี่ยอยู่ที่ร้อยละ 85.08 ส่วนชุดข้อมูล Test set มีค่าความถูกต้องเฉลี่ยอยู่ที่ร้อยละ 53.83

#### 2.4.3.1 คุณสมบัติของการพัฒนาและประยุกต์ใช้เทคนิค Convolutional Neural Network ในการจำแนกขยะชายฝั่งทะเล

1. สามารถทำการจำแนกขยะชายฝั่งทะเล ด้วยข้อมูลภาพถ่าย เพื่อให้การจำแนกขยะชายฝั่งทะเลเป็นไปอย่างสะดวกและรวดเร็ว
2. สามารถติดตามการแพร่กระจายชนิดของขยะพลาสติกในทะเล

#### 2.4.3.2 ข้อจำกัดของการพัฒนาและประยุกต์ใช้เทคนิค Convolutional Neural Network ในการจำแนกขยะชายฝั่งทะเล

มีค่าความถูกต้องในการจำแนกที่ไม่สูงมาก เนื่องมาจากการตัวอย่างภาพถ่ายขยะบริเวณชายฝั่งทะเลในการศึกษา ครั้งนี้มีสิ่งรบกวนที่เกิดจากการปักคลุมวัตถุที่เป็นขยะของทรายหญ้า และกิ่งไม้ซึ่งอาจส่งผลให้ ประสิทธิภาพในการจำแนกขยะชายฝั่งทะเล ด้วยเทคนิค Convolutional Neural Network มีค่าความถูกต้องที่ลดลงโดยขึ้นอยู่กับปริมาณของสิ่งรบกวนที่ปรากฏบนตัวอย่างภาพถ่าย

#### 2.4.4 เครื่องคัดแยกขยะรีไซเคิลด้วยการประมวลผลภาพ

นโนโซชติ ภาชนะนัย, จักรกริช ปานเงิน, กรณิการ์ คนงาม, วรชัย ศรีสมุดคำ และวานา วงศ์ษา (2565) ได้นำเสนอผลงานวิจัยเรื่องเครื่องคัดแยกขยะรีไซเคิลด้วยการประมวลผลภาพให้สามารถควบคุมการคัดแยกขยะรีไซเคิล 4 ประเภทได้แก่ กระป่อง ขวดแก้ว กล่องกระดาษ และขวดพลาสติก โดยระบบประกอบด้วยกล้อง IMILAB Webcam รับภาพวิดีโอแบบเรียลไทม์เชื่อมต่อ กับ Raspberry Pi4 Model B เพื่อทำการประมวลผลจำแนกขยะด้วยโมเดลของ YOLOv5 ที่ผ่านการฝึกสอนด้วยชุดข้อมูลภาพ 1,321 ภาพ 250 รอบ มีค่า mAP50% 0.995, Recall 0.995, Precision 0.999 และค่า F-measure 0.997 โดยเชื่อมต่อ กับ Module L298N ในการควบคุมตำแหน่งการหมุน ของสเต็ปมอเตอร์ไปยังถังขยะເອົ້າຫຼຸດและควบคุมเซอร์โวมอเตอร์ในการเทขยะลงถังขยะแต่ละประเภท ผลการทดสอบประสิทธิภาพการทำงานพบว่าสามารถคัดแยกขยะได้ถูกต้องตามประเภทคิด เป็น 98.33%

##### 2.4.4.1 คุณสมบัติของเครื่องคัดแยกขยะรีไซเคิลด้วยการประมวลผลภาพ

1. สามารถคัดแยกขยะรีไซเคิลได้ 4 ประเภทตามโมเดลที่ฝึกสอน
2. สเต็ปมอเตอร์และเซอร์โวสามารถหมุนและเทขยะลงถังขยะแต่ละประเภทได้อย่างถูกต้อง

##### 2.4.4.2 ข้อจำกัดของเครื่องคัดแยกขยะรีไซเคิลด้วยการประมวลผลภาพ

1. ประสิทธิภาพการทดสอบพบว่ายังมีผลการคัดแยกขยะไม่ถูกต้อง อาจเนื่องมาจากการตรวจจับไม่ถูกต้อง เช่น ตรวจจับสีขาวซึ่งมีการสะท้อนของแสงทำให้ระบบแสดงผลการคัดแยกเป็นประเภทขวดแก้ว
2. การทดสอบทำในสภาวะแวดล้อมแบบเปิดซึ่งมีแสงสว่างที่มากหรืออาจน้อยเกินไป ส่งผลกระทบต่อการตรวจจับและคัดแยกวิดีโอด้วยในบางครั้ง
3. เมื่อมีวัตถุอื่นๆ นำมายังทดสอบระบบสามารถคัดแยกได้ในบางครั้งเท่านั้น หากวัตถุไม่ถูกตรวจพบว่าอยู่ใน 4 ประเภทของโมเดลนี้ระบบจะไม่สามารถแสดงกรอบเพื่อทำนายส่วนของวัตถุและค่า Confidence score บนภาพได้

#### 2.4.5 สรุปและเปรียบเทียบความแตกต่างของระบบ

	การจำแนก พระเครื่อง ด้วยการ เรียนรู้เชิงลึก	การจำแนก ภาพขาวด แบบเซ็ตเปิด ด้วยโครงข่าย ประสาท เทียมแบบ คอนโวจูชัน	การพัฒนาและ ประยุกต์ใช้เทคนิค Convolutional Neural Network ในการจำแนกขยะ ชายฝั่งทะเล	เครื่องคัด แยกขยะรี ไซเคิลด้วย การ ประมวลผล ภาพ	การคัดแยกขยะ รีไซเคิลจากภาพ ด้วยเทคนิคการ เรียนรู้เชิงลึก
เทคนิคการ เรียนรู้เชิงลึก	/	/	/	/	/
คัดแยกขยะรี ไซเคิล 4 ประเภท				/	/
ข้อมูลภาพ หลากหลาย รูปแบบ	/			/	/
สามารถ อัปโหลด รูปภาพที่ ต้องการ ทำนายได้					/
สามารถคัด แยกได้หลาย ชิ้นใน 1 ภาพ				/	
ระบบต้อง เชื่อมต่อ อินเทอร์เน็ต	/	/	/	/	/

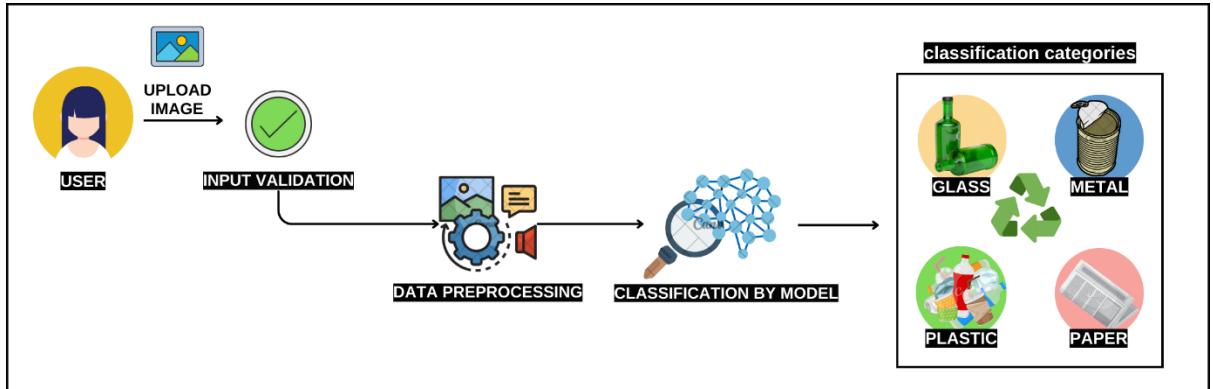
ตารางที่ 2.1 สรุปและเปรียบเทียบความแตกต่างของระบบ

เมื่อศึกษาวิจัยที่เกี่ยวข้องทั้งหมดสรุปได้ว่าแต่ละงานวิจัยมีข้อจำกัดที่แตกต่างกันออกไป โดยการคัดแยกขยายรีไซเคิลจากภาพด้วยเทคนิคการเรียนรู้เชิงลึกใช้เทคนิคการเรียนรู้เชิงลึก สามารถแยกขยายรีไซเคิลได้ทั้งหมด 4 ประเภทคือ แก้ว พลาสติก กระดาษ โลหะ อย่างไรก็ตาม ยังมีการปรับรูปภาพให้เหมาะสม เช่น การปรับสีและแสง เพื่อให้ระบบสามารถจำแนกประเภทขยะได้อย่างมีประสิทธิภาพ โดยการปรับแก้สีและแสงช่วยลดผลกระทบจากสภาพแวดล้อมที่ไม่เหมาะสม อีกทั้งยังสามารถอัปโหลดรูปภาพที่ต้องการทราบประเภทได้โดยการใช้งานต้องมีการเชื่อมต่ออินเทอร์เน็ต

## บทที่ 3

### วิธีการวิจัย

#### 3.1 ภาพรวมของโครงการ



ภาพที่ 3.1 สถาปัตยกรรมระบบ

#### คำอธิบายสถาปัตยกรรมของระบบ

##### 3.1.1 ผู้ใช้งาน (User)

ผู้ใช้งานคือบุคคลที่มีการโต้ตอบกับระบบ ในกรณีนี้ผู้ใช้งานทำการอัปโหลดภาพของวัสดุที่ต้องการจำแนกประเภท เช่น พลาสติก หรือโลหะ เพื่อช่วยในการคัดแยกขยะรีไซเคิลอย่างถูกต้อง

##### 3.1.2 การตรวจสอบข้อมูลนำเข้า (Input Validation)

เมื่อผู้ใช้งานอัปโหลดภาพ ระบบจะทำการตรวจสอบความถูกต้องของข้อมูลนำเข้า เช่น ตรวจสอบประเภทไฟล์ (เช่น JPEG, PNG) ขนาดของภาพ และตรวจสอบว่าไฟล์ไม่มีปัญหาหรือความเสียหาย เพื่อให้แน่ใจว่าเฉพาะข้อมูลที่ถูกต้องเท่านั้นที่จะถูกส่งไปยังขั้นตอนถัดไป

##### 3.1.3 การเตรียมข้อมูล (Data Preprocessing)

เป็นการเตรียมข้อมูลให้พร้อมก่อนที่จะนำไปจำแนกหมวดหมู่ โดยเริ่มจากการโหลดภาพที่ผู้ใช้อัปโหลดเข้าสู่ระบบและปรับขนาดของภาพให้มีมาตรฐานเดียวกัน เพื่อให้เหมาะสมกับกระบวนการประมวลผล จากนั้นทำการตรวจสอบรูปแบบและคุณภาพของภาพ นอกจากนี้ยังมีการปรับค่าของข้อมูลให้อยู่ในช่วงที่เหมาะสม เพื่อให้ข้อมูลพร้อมสำหรับการประมวลผลโดยไม่เดลออย่างมีประสิทธิภาพ

### 3.1.4 การจำแนกประเภทโดยโมเดล (Classification by Model)

ขั้นตอนนี้เป็นกระบวนการที่โมเดล Convolutional Neural Network (CNN) และ You Only Look Once (YOLO) ทำการวิเคราะห์ภาพที่ผ่านการเตรียมข้อมูล (Data Preprocessing) เพื่อจำแนกประเภทของขยะให้อยู่ในหมวดหมู่ต่าง ๆ เช่น แก้ว (Glass), โลหะ (Metal), พลาสติก (Plastic), กระดาษ (Paper)

### 3.1.5 หมวดหมู่การจำแนกประเภท (Classification Categories)

หมวดหมู่การจำแนกประเภทเป็นผลลัพธ์ที่ได้จากการประมวลผลโดยโมเดล ซึ่งจะจัดประเภทของขยะให้อยู่ในกลุ่มที่เหมาะสม โดยในระบบนี้หมวดหมู่หลักของการจำแนกประเภทมีดังนี้:

#### 1. แก้ว (Glass)

หมวดหมู่นี้ประกอบด้วยขยะที่ทำจากแก้ว เช่น ขวดแก้วหรือเศษแก้ว ซึ่งสามารถนำกลับมาใช้ค่าเพื่อผลิตแก้วใหม่ได้

#### 2. โลหะ (Metal)

รวมถึงขยะที่เป็นโลหะ เช่น กระป๋องอลูมิเนียม หรือเศษโลหะต่าง ๆ ซึ่งเป็นวัสดุที่มีน้ำหนักสูงและสามารถนำไปหยอดเพื่อใช้ซ้ำได้

#### 3. พลาสติก (Plastic)

ขยะในหมวดนี้ ได้แก่ ขวดพลาสติก ถุงพลาสติก หรือบรรจุภัณฑ์ที่ทำจากพลาสติก ซึ่งเป็นวัสดุที่มีผลกระทบต่อสิ่งแวดล้อมหากไม่ได้รับการจัดการอย่างเหมาะสม

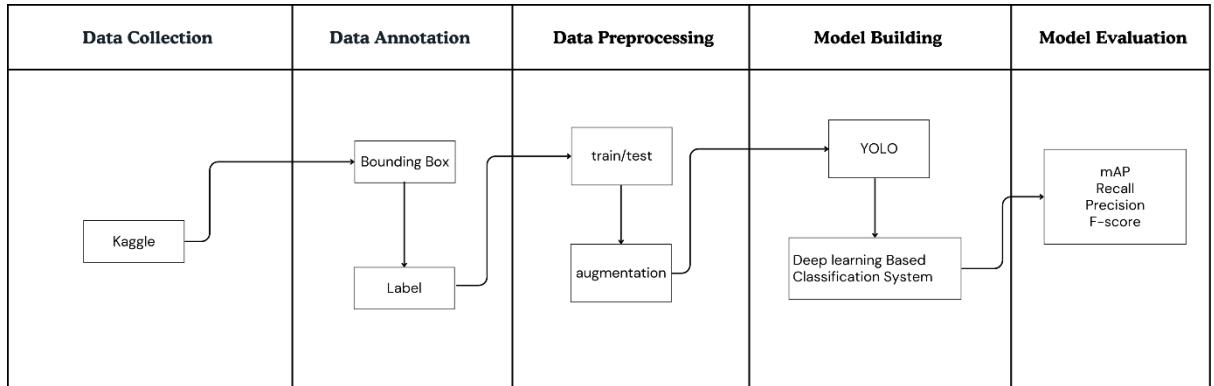
#### 4. กระดาษ (Paper)

ประกอบด้วยขยะประเภทกระดาษ เช่น หนังสือพิมพ์ กล่องกระดาษ หรือกระดาษเหลือใช้ ซึ่งสามารถนำไปรีไซเคิลเพื่อผลิตกระดาษใหม่ได้

## 3.2 การทดลองเบื้องต้นเพื่อเลือกแนวทางการพัฒนาโมเดล

เพื่อประเมินแนวทางที่เหมาะสมที่สุดสำหรับการพัฒนาโมเดลจำแนกประเภทขยะรีไซเคิล จึงได้ดำเนินการทดลองในเบื้องต้นโดยใช้วิธีการสองแนวทาง ได้แก่ YOLO ซึ่งเป็นอัลกอริズึมสำหรับงานตรวจจับวัตถุ และ CNN ซึ่งเป็นโมเดลสำหรับงานจำแนกประเภทภาพ โดยในการทดลองทั้งสองวิธีได้ใช้ชุดข้อมูลเดียวกัน คือ ภาพขยะรีไซเคิลจำนวน 4 ประเภท ได้แก่ แก้ว, โลหะ, กระดาษ และ พลาสติก ซึ่งแต่ละประเภทมีจำนวน 300 รูป รวมทั้งหมด 1,200 รูป ทั้งนี้เพื่อให้สามารถเปรียบเทียบประสิทธิภาพของแต่ละแนวทางได้อย่างเหมาะสมและเป็นธรรม

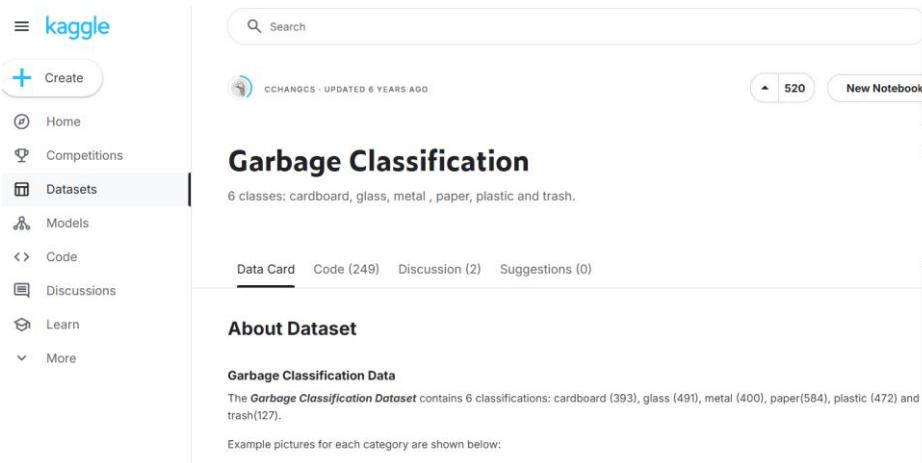
### 3.2.1 การทดลองด้วย YOLO (You Only Look Once)



ภาพที่ 3.2 ตัวอย่างการวิเคราะห์ขอบเขตและความต้องการของ YOLO

#### 3.2.1.1 การเก็บรวบรวมข้อมูล (Data Collection)

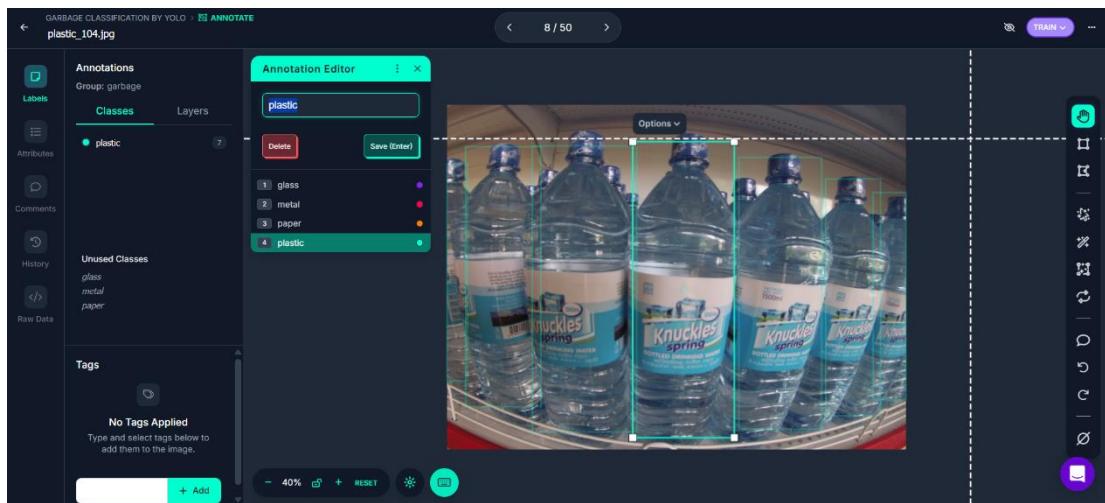
การรวบรวมข้อมูลเป็นขั้นตอนสำคัญในการพัฒนาโมเดล Deep Learning โดยในกระบวนการนี้ ข้อมูลภาพถูกรวบรวมจากแพลตฟอร์ม Kaggle โดยกำหนดเป้าหมายในการเก็บข้อมูลจำนวนประมาณ 300 รูปภาพต่อคลาส รวมทั้งหมด 1200 รูป เพื่อให้ข้อมูลมีความสมดุลและเพียงพอต่อการฝึกโมเดล YOLO ซึ่งเป็นโมเดลที่เน้นการตรวจจับวัตถุ (Object Detection) โดยข้อมูลภาพที่ได้มีความหลากหลาย ทั้งในเรื่องมุมมอง แสง และพื้นหลัง เพื่อช่วยให้โมเดลมีความสามารถในการจัดจำวัตถุในสถานการณ์จริงได้ดีขึ้น



ภาพที่ 3.3 การรวบรวมข้อมูลจาก Kaggle

### 3.2.1.2 การทำ Annotation (Data Annotation)

หลังจากการรวมข้อมูลครบถ้วนแล้ว ชุดข้อมูลภาพจะมีจำนวน 1200 ภาพ จะถูกอัปโหลดเข้าสู่แพลตฟอร์ม Roboflow ซึ่งเป็นเครื่องมือออนไลน์ที่ช่วยในการจัดการและเตรียมข้อมูลสำหรับงานด้าน Computer Vision จากนั้นได้ทำการ Annotate ข้อมูลด้วยการระบุตำแหน่งวัตถุ (Bounding Box) และติดป้ายกำกับ (Labeling) ของวัตถุในแต่ละภาพ สำหรับแต่ละประเภทของขยะ เพื่อให้เหมาะสมกับการฝึกโมเดล YOLO ดังภาพ



ภาพที่ 3.4 การทำ Annotation

### 3.2.1.3 การเตรียมข้อมูล (Data Preprocessing)

ในขั้นตอนการเตรียมข้อมูล ได้มีการแบ่งชุดข้อมูล โดยได้กำหนดสัดส่วนเป็น 70% สำหรับการฝึกสอนโมเดล (Training set), 20% สำหรับการตรวจสอบประสิทธิภาพระหว่างฝึก (Validation set) และ 10% สำหรับการทดสอบโมเดล (Test set) ดังนี้

Training set	840 ภาพ
Validation set	240 ภาพ
Test set	120 ภาพ

ตารางที่ 3.1 จำนวนข้อมูล

ต่อมา ได้มีการดำเนินการ เพิ่มจำนวนข้อมูลในชุด Training set ด้วยเทคนิค Data Augmentation โดยใช้เครื่องมือจากแพลตฟอร์ม Roboflow เพื่อเพิ่มความหลากหลายให้กับข้อมูล ซึ่งช่วยให้โมเดลสามารถเรียนรู้จากภาพในมุมมองหรือสภาพแวดล้อมที่แตกต่างกันได้ดีขึ้น เช่น การหมุนภาพ การปรับแสง การกลับด้าน หรือการซูมภาพ

จากการทำ Augmentation ตั้งกล่าว ส่งผลให้จำนวนภาพในชุด Training เพิ่มขึ้นจากเดิม 840 ภาพ เป็น 2,517 ภาพ ขณะที่ชุด Validation และ Test ยังคงเท่าเดิม นั่นคือ 240 และ 120 ภาพตามลำดับ การเพิ่มจำนวนข้อมูลด้วยวิธีนี้ มีวัตถุประสงค์หลักเพื่อ ลดปัญหา Overfitting ซึ่งอาจเกิดขึ้นเมื่อโมเดลเรียนรู้เฉพาะลักษณะของข้อมูลเดิมโดยไม่สามารถประยุกต์ใช้กับข้อมูลใหม่ได้ การใช้เทคนิค Augmentation จึงเป็นวิธีหนึ่งที่ช่วยให้โมเดลสามารถเรียนรู้ลักษณะของวัตถุได้ครอบคลุมยิ่งขึ้น ส่งผลให้โมเดลมีความแม่นยำและมีประสิทธิภาพมากขึ้นเมื่อนำไปใช้งานจริง

### 3.2.1.4 การสร้างโมเดล (Model Building)

ในการสร้างโมเดลสำหรับตรวจจับและจำแนกประเภทขยะที่สามารถรีไซเคิลได้ ได้นำไลบรารี Ultralytics YOLOv8 มาใช้ ซึ่งเป็นเวอร์ชันล่าสุดของโมเดล YOLO (You Only Look Once) ที่ได้รับการพัฒนาให้สามารถใช้งานได้สะดวก รวดเร็ว และแม่นยำ โดยใช้ภาษา Python เป็นหลัก

โดยได้ทำการฝึกโมเดลด้วยคำสั่ง ดังนี้

```
[ ] from ultralytics import YOLO

# โหลดโมเดล YOLOv8
model = YOLO('yolov8n.pt')

model.train(data='/content/garbage-classification-by-yolo-1/data.yaml',
            epochs=50,
            imgsz=640,
            batch=16,
            name='yolov8_recyclable_model',
            save=True)
```

ภาพที่ 3.5 โค้ดการฝึกโมเดล YOLO

ในการสร้างโมเดลสำหรับตรวจจับและจำแนกประเภทของตัวอักษรที่สามารถตัดสินใจได้โดยใช้เครื่องมือ Ultralytics YOLOv8 มาใช้ ซึ่งเป็นเวอร์ชันล่าสุดของโมเดล YOLO และมีรายละเอียดของพารามิเตอร์ต่าง ๆ ดังนี้

- **data:** ระบุไฟล์ .yaml ที่กำหนดเส้นทางของชุดข้อมูล พร้อมชื่อคลาสที่ต้องการให้โมเดลเรียนรู้ เช่น glass, paper, plastic, metal
- **epochs=50:** กำหนดจำนวนรอบในการฝึกโมเดลไว้ที่ 50 รอบ เนื่องจากเป็นจำนวนที่เพียงพอให้โมเดลสามารถเรียนรู้ลักษณะของภาพได้ครบถ้วนโดยไม่ใช้เวลานานเกินไป หากรอบน้อยเกินไปอาจทำให้โมเดลเรียนรู้ไม่เพียงพอ แต่หากมากเกินไปอาจเกิดปัญหา overfitting
- **imgsz=640:** ขนาดของภาพถูกตั้งไว้ที่ 640x640 พิกเซล ซึ่งเป็นค่าที่เหมาะสมตามค่าเริ่มต้นของ YOLOv8 โดยมีความสมดุลระหว่างคุณภาพของภาพที่นำเข้าและระยะเวลาในการประมวลผล ขนาดที่ใหญ่ขึ้นจะให้รายละเอียดภาพมากขึ้น แต่จะใช้หน่วยประมวลผลและเวลาในการฝึกมากขึ้นด้วย
- **batch=16:** กำหนด batch size ไว้ที่ 16 ภาพต่อรอบ เนื่องจากเป็นขนาดที่สมดุลระหว่างประสิทธิภาพการฝึกและการใช้ทรัพยากรคอมพิวเตอร์ โดยเฉพาะในสภาพแวดล้อมที่มีข้อจำกัด เช่น Google Colab ค่า batch ที่มากกว่านี้อาจทำให้หน่วยความจำไม่เพียงพอ
- **name='yolov8\_recyclable\_model':** กำหนดชื่อโมเดลและโฟลเดอร์ที่ใช้เก็บผลลัพธ์ การฝึก
- **save=True:** กำหนดให้บันทึกน้ำหนักของโมเดลที่ได้หลังจากฝึก เพื่อสามารถนำไปใช้งานตรวจจับจริงหรือประเมินผลต่อไปได้

### 3.2.1.5 การประเมินผลโมเดล YOLO (Model Evaluation)

คือขั้นตอนในการวัดประสิทธิภาพของโมเดลหลังจากที่มันได้รับการฝึกฝน (Training) โดยการใช้ข้อมูลที่ไม่เคยเห็นมาก่อน (Validation Data หรือ Test Data) ในที่นี้ เราจะใช้ Precision, Recall, F1-Score และ mAP เพื่อประเมินว่าโมเดลทำงานได้ดีเพียงใด

1. Precision : เป็นการวัดความแม่นยำในการทำนายค่าบวก

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

ค่าต่าง ๆ ที่ใช้ในสูตร:

True Positives (TP): จำนวนกรณีที่โมเดลทำนายว่าเป็นบวก และผลลัพธ์จริงนั้นถูกต้องจริง  
False Positives (FP): จำนวนกรณีที่โมเดลทำนายว่าเป็นบวก แต่ผลลัพธ์นั้นไม่ถูกต้อง

2. Recall : เป็นการวัดความสามารถของโมเดลในการทำนายค่าบวกทั้งหมดที่มีอยู่ในข้อมูลจริง

$$\text{Recall} = \frac{TP}{TP + FN}$$

ค่าต่าง ๆ ที่ใช้ในสูตร:

True Positives (TP): จำนวนกรณีที่โมเดลทำนายว่าเป็นบวก และผลลัพธ์จริงนั้นถูกต้องจริง  
False Negatives (FN): จำนวนกรณีที่โมเดลพลาดไม่สามารถทำนายว่าเป็นบวก ทั้งที่จริงแล้วเป็นบวก

3. F1-score : คำนวณจาก Precision และ Recall โดยการหาค่าเฉลี่ย

$$F1 - score = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

ค่า F1 Score จะอยู่ในช่วง 0 ถึง 1 โดยค่าที่ใกล้ 1 แสดงถึงโมเดลที่มีทั้ง Precision และ Recall สูง ซึ่งหมายถึงการทำนายที่แม่นยำและครอบคลุม

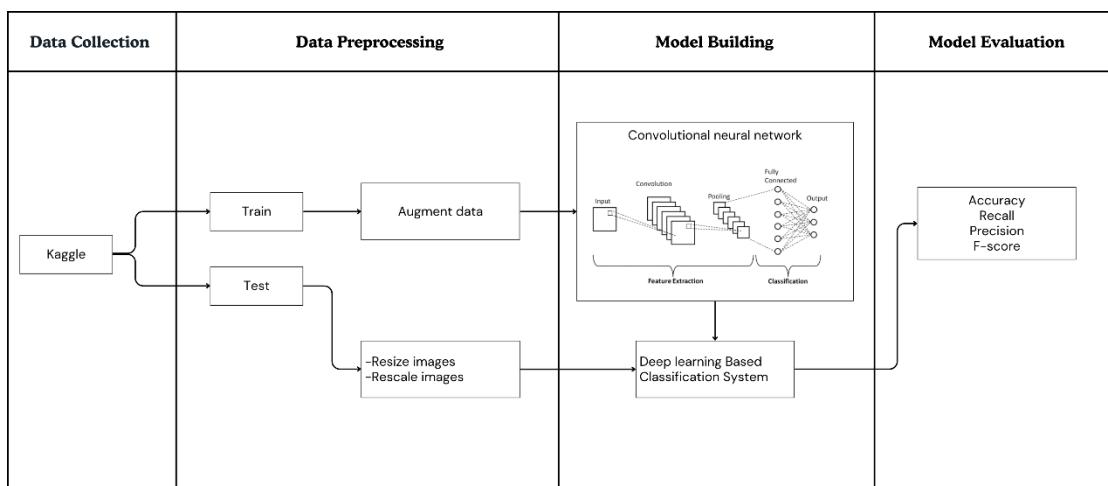
#### 4. mAP50 และ mAP50-95

- mAP@50 หมายถึงค่าความแม่นยำเฉลี่ยเมื่อกำหนดเกณฑ์ IoU  $\geq 0.50$
- mAP@50–95 หมายถึงค่าเฉลี่ยของ AP ที่คำนวณจากหลายเกณฑ์ของ IoU ตั้งแต่ 0.50 ถึง 0.95

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

โดยที่  $N$  คือจำนวนคลาส และ  $AP_i$  คือค่า Average Precision ของแต่ละคลาสค่า mAP ที่สูง แสดงถึงโมเดลที่มีประสิทธิภาพในการตรวจจับและระบุตำแหน่งของวัตถุในภาพได้ดี

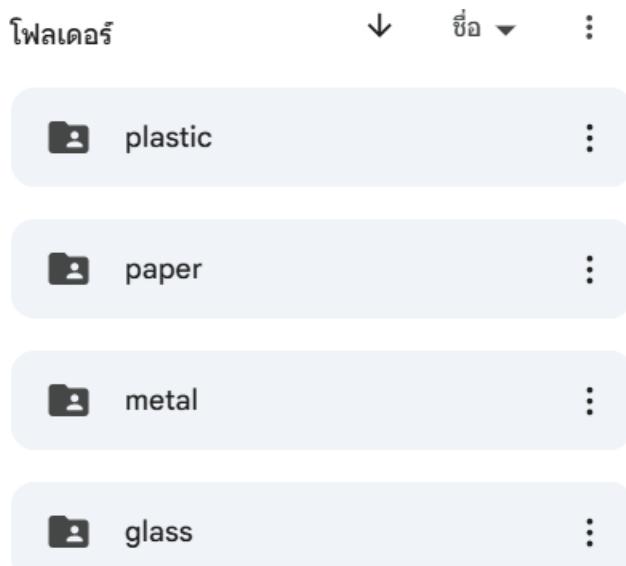
#### 3.2.2 การทดลองด้วย CNN (Convolutional Neural Networks)



ภาพที่ 3.6 ตัวอย่างการวิเคราะห์ขอบเขตและความต้องการของ CNN

### 3.2.2.1 การเก็บรวบรวมข้อมูล (Data Collection)

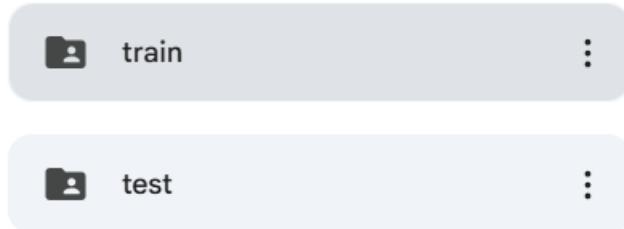
สำหรับการทดลองด้วยโมเดล Convolutional Neural Networks (CNN) ในการจำแนกประเภทของขยะที่สามารถรีไซเคิลได้ ได้ใช้ชุดข้อมูลเดียวกับที่ใช้ในการฝึกโมเดล YOLO เพื่อให้สามารถเปรียบเทียบประสิทธิภาพของทั้งสองแนวทางได้อย่างเหมาะสมและเป็นธรรม โดยจำนวนภาพที่รวบรวมมา 300 ภาพต่อคลาส รวมทั้งหมด 1,200 ภาพ และจัดเก็บภาพของขยะไว้ในโฟลเดอร์ตามประเภทของขยะ บน Google drive ดังนี้



ภาพที่ 3.7 การจัดเก็บภาพของขยะไว้บน Google Drive

### 3.2.2.2 การเตรียมข้อมูล (Data Preprocessing)

1. ข้อมูลฝึกอบรมและข้อมูลทดสอบ (Training and Test Data)  
ชุดข้อมูลที่รวบรวมมาถูกแบ่งออกเป็นโฟลเดอร์ 2 ส่วน ดังนี้



ภาพที่ 3.8 การแบ่งโฟลเดอร์ข้อมูลฝึกอบรมและข้อมูลทดสอบ

- **ข้อมูลฝึกอบรม (Training Dataset):** ใช้สำหรับการฝึกอบรมโมเดล โดยแยกเป็น 4 โฟลเดอร์ โดยในแต่ละโฟลเดอร์มีข้อมูลฝึกอบรมอยู่ร่วม 1080 ภาพ ดังนี้

Plastic	270 ภาพ
Paper	270 ภาพ
Metal	270 ภาพ
Glass	270 ภาพ

ตารางที่ 3.2 จำนวนข้อมูลฝึกอบรม

- **ข้อมูลทดสอบ (Test Dataset):** ใช้สำหรับการทดสอบโมเดล โดยแยกเป็น 4 โฟลเดอร์เช่นกัน โดยในแต่ละโฟลเดอร์มีข้อมูลฝึกทดสอบอยู่ รวม 120 ภาพ ดังนี้

Plastic	30 ภาพ
Paper	30 ภาพ
Metal	30 ภาพ
Glass	30 ภาพ

ตารางที่ 3.3 จำนวนข้อมูลทดสอบ

## 2. การเพิ่มปริมาณของข้อมูล (Data Augmentation)

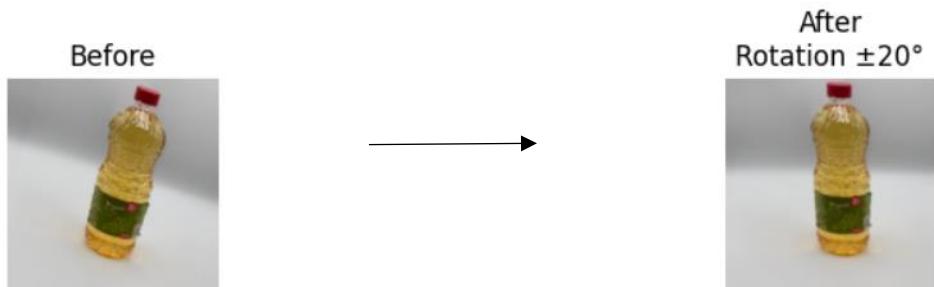
เพื่อให้การฝึกโมเดล Convolutional Neural Network (CNN) ให้สามารถจำแนกประเภทขยะได้อย่างแม่นยำและลดการ Overfitting จึงมีการใช้ `ImageDataGenerator` จาก Keras เพื่อทำ **Data Augmentation** หรือการแปลงภาพต้นฉบับให้หลากหลายขึ้นแบบสุ่ม

```
from tensorflow.keras.preprocessing.image import ImageDataGenerator
import os
train_datagen = ImageDataGenerator(
    rescale=1./255,
    validation_split=0.2,
    rotation_range=20,
    width_shift_range=0.1,
    height_shift_range=0.1,
    zoom_range=0.2,
    horizontal_flip=True
)
```

ภาพที่ 3.9 โค้ดการทำ Data Augmentation

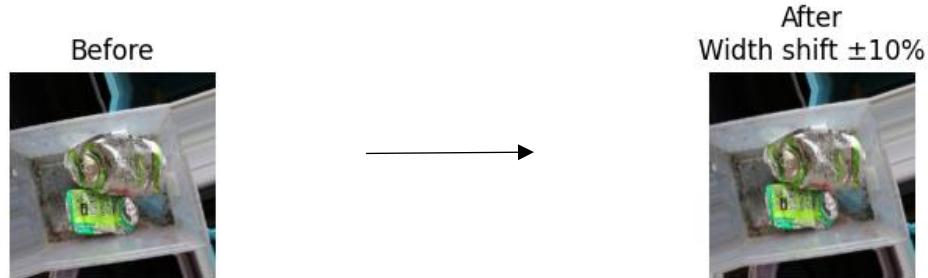
รายละเอียดของแต่ละพารามิเตอร์มีดังนี้

- rescale=1./255 เป็นการ ปรับค่าสีของพิกเซลให้อยู่ในช่วง 0 ถึง 1 โดยเดิมค่าพิกเซลของภาพจะอยู่ในช่วง 0–255 ซึ่งเป็นค่าความเข้มของสีแต่ละจุด
- validation\_split=0.2 เป็นการ กำหนดให้แบ่งข้อมูลฝึกอบรม ออกมาใช้เป็นข้อมูลตรวจสอบความถูกต้อง (Validation set) โดยอัตโนมัติ ในที่นี่ระบุค่า 0.2 หมายถึง 20% ของข้อมูล จะถูกใช้สำหรับ validation ใช้เพื่อวัดประสิทธิภาพของโมเดลระหว่างการฝึกโดยไม่ใช้ข้อมูลที่โมเดลได้เรียนรู้ไปแล้ว ช่วยให้สามารถตรวจสอบได้ว่าโมเดลมีปัญหา overfitting หรือไม่
- rotation\_range=20 คือการให้สูมหมุนภาพได้สูงสุด  $\pm 20$  องศา เพิ่มความหลากหลายของทิศทางวัดถูกในภาพ



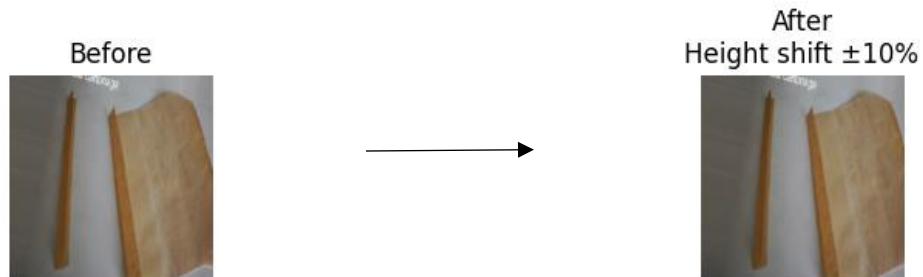
ภาพที่ 3.10 การทำ rotation\_range=20

- width\_shift\_range=0.1 คือการสุ่มเลื่อนภาพในแนวนอนประมาณ  $\pm 10\%$  ของความกว้าง



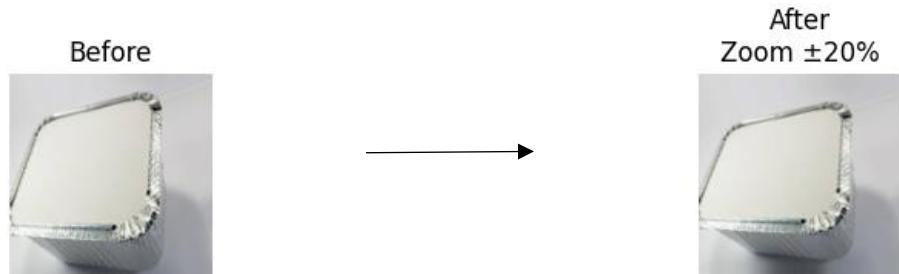
ภาพที่ 3.11 การทำ width\_shift\_range=0.1

- height\_shift\_range=0.1 คือการสุ่มเลื่อนภาพในแนวตั้งประมาณ  $\pm 10\%$  ของความสูง



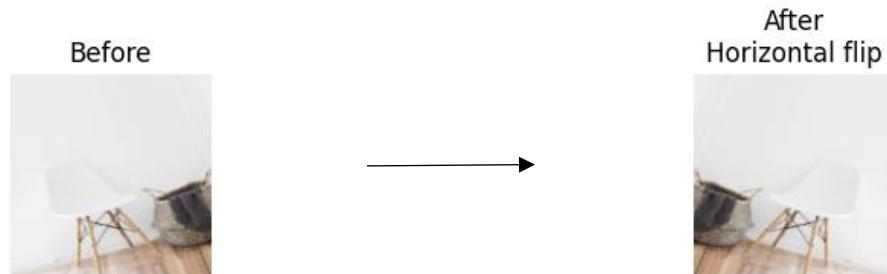
ภาพที่ 3.12 การทำ height\_shift\_range=0.1

- zoom\_range=0.2 คือการสุ่มซูมเข้า/ออกที่ช่วง 80%-120% ของภาพต้นฉบับ



ภาพที่ 3.13 การทำ zoom\_range=0.2

- horizontal\_flip=True คือการพลิกภาพในแนวอนแบบสูม (ซ้าย-ขวา)



ภาพที่ 3.14 การทำ horizontal\_flip=True

จากการกำหนดค่าดังกล่าว จะไม่ทำให้จำนวนภาพใน dataset เพิ่มขึ้นแบบถาวร แต่จะมีการสร้างภาพใหม่แบบสูมแบบ real-time ระหว่างการฝึกโมเดล ซึ่งหมายความว่าแม้ภาพต้นฉบับมีเพียง 1 ภาพ ระบบสามารถสร้างภาพใหม่ที่แตกต่างกันได้หลายรูปแบบในแต่ละ epoch โดยอาศัยการสุ่มหมุน เลื่อน ซูม หรือพลิกภาพตามค่าที่กำหนดไว้ ทำให้เกิดความหลากหลายมากขึ้นในการฝึกโมเดล ช่วยให้โมเดลเรียนรู้ได้ดีขึ้น มีความสามารถในการ generalize ต่อข้อมูลใหม่ และลดความเสี่ยงของปัญหา Overfitting โดยทำให้โมเดลไม่จดจำลักษณะเฉพาะของภาพต้นฉบับมากเกินไป

การทำ Data Augmentation จะถูกนำมาใช้เฉพาะกับข้อมูลฝึกอบรม (Training set) เท่านั้น โดยดำเนินการผ่าน ImageDataGenerator ที่ระบุพารามิเตอร์สำหรับการแปลงภาพแบบสุ่มระหว่างการฝึกโมเดล เพื่อช่วยให้โมเดลมีความสามารถในการเรียนรู้จากข้อมูลที่มีความหลากหลายมากขึ้น

สำหรับข้อมูลทดสอบ (Test set) จะไม่มีการทำ Data Augmentation ใด ๆ ทั้งสิ้น โดยจะดำเนินการเฉพาะ 2 ขั้นตอนหลัก ได้แก่:

- การปรับขนาดภาพ (Resize Images)  
เพื่อให้ภาพทุกภาพมีขนาดเท่ากันก่อนนำเข้าโมเดล ซึ่งเป็นข้อกำหนดของโครงข่ายประสาทเชิงมัตโน (Convolutional Neural Network - CNN) ที่ต้องการขนาดอินพุตที่คงที่และสม่ำเสมอ งานวิจัยนี้ได้กำหนดให้ภาพมีขนาด 224x224 พิกเซล ซึ่งเป็นค่ามาตรฐานที่สมดุลระหว่างความละเอียดของภาพและประสิทธิภาพในการประมวลผล ช่วยให้โครงข่ายสามารถเรียนรู้คุณลักษณะของภาพได้อย่างมีประสิทธิภาพ

- การปรับค่าสเกลภาพ (Rescale Images)

เป็นการแปลงค่าพิกเซลจากช่วง 0–255 ให้อยู่ในช่วง 0–1 โดยใช้  $\text{rescale}=1./255$  เพื่อให้เหมาะสมกับการประมวลผลของโมเดล การทำ rescaling มีบทบาทสำคัญในการเพิ่มเสถียรภาพของการเรียนรู้ ช่วยให้โมเดลสามารถปรับค่าน้ำหนักได้ดีขึ้นในแต่ละรอบของการฝึก ส่งผลให้การจำแนกภาพมีความแม่นยำและมีประสิทธิภาพยิ่งขึ้น

### 3.2.2.3 การสร้างโมเดล (Model Building)

ในการฝึกสอนโมเดลของระบบคัดแยกขยะเริ่มต้น เราใช้ชุดข้อมูลที่ฝ่ายการเตรียมข้อมูลและการเพิ่มข้อมูล (Data Augmentation) เป็นขั้นตอนเริ่มต้น จากนั้นเรานำข้อมูลเหล่านี้ไปฝึกสอนกับอัลกอริธึมการเรียนรู้ของเครื่อง คือ Convolutional Neural Networks (CNN) ซึ่งเป็นอัลกอริธึมที่มีประสิทธิภาพในการประมวลผลข้อมูลภาพ การใช้อัลกอริธึม CNN ช่วยให้โมเดลสามารถเรียนรู้ลักษณะเฉพาะของขยะในแต่ละประเภท เช่น แก้ว พลาสติก กระดาษ และโลหะ เพื่อนำไปใช้ในการจำแนกประเภทขยะได้อย่างแม่นยำ

โดยด้านล่างนี้เป็นตัวอย่างโค้ดการสร้างโมเดล CNN ที่ใช้งานวิจัยนี้ ซึ่งประกอบด้วยเลเยอร์ Convolution, MaxPooling, Flatten, และ Dense รวมถึงการตั้งค่า Optimizer และ Loss Function สำหรับการฝึกโมเดล

```
import tensorflow as tf
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Conv2D, MaxPooling2D, Flatten, Dense, Dropout

num_classes = train_generator.num_classes

model = Sequential([
    Conv2D(32, (3, 3), activation='relu', input_shape=(224, 224, 3)),
    MaxPooling2D(pool_size=(2, 2)),

    Conv2D(64, (3, 3), activation='relu'),
    MaxPooling2D(pool_size=(2, 2)),

    Conv2D(128, (3, 3), activation='relu'),
    MaxPooling2D(pool_size=(2, 2)),

    Flatten(),
    Dense(128, activation='relu'),
    Dropout(0.5),
    Dense(num_classes, activation='softmax')
])

model.compile(
    optimizer='adam',
    loss='categorical_crossentropy',
    metrics=['accuracy']
)
```

ภาพที่ 3.15 การสร้างโมเดล CNN

### คำอธิบายโดยคัด:

- Conv2D: เป็นเลเยอร์ convolution สำหรับดึงคุณลักษณะจากภาพ เช่น ขอบหรือรูปทรงต่าง ๆ โดยเริ่มจาก 32 ฟิลเตอร์และเพิ่มจำนวนขึ้นในแต่ละชั้น
- MaxPooling2D: ลดขนาดของภาพ เพื่อลดจำนวนพารามิเตอร์และการคำนวณ
- Flatten(): แปลงข้อมูลจากภาพ 2 มิติเป็นเวกเตอร์ 1 มิติ เพื่อเชื่อมต่อกับ Fully Connected Layer
- Dense: เป็นเลเยอร์ fully connected สำหรับทำการจำแนกประเภท (output layer ใช้ softmax สำหรับ multi-class classification)
- Dropout: ใช้สำหรับลด overfitting โดยสุ่มปิดบางนิวรอนระหว่างการฝึก
- compile: ตั้งค่าการฝึกด้วย optimizer แบบ Adam และ loss function เป็น categorical crossentropy เนื่องจากเป็นการจำแนกหลายคลาส

### การสร้างแบบจำลองการจำแนกประเภท (Classification Model)

ในการสร้างแบบจำลองการจำแนกประเภทเรียกเช่นเดิม เราใช้ Convolutional Neural Networks (CNN) ซึ่งเป็นโมเดลที่เหมาะสมที่สุดในการประมวลผลข้อมูลภาพ โดยในขั้นตอนนี้โมเดลจะได้รับข้อมูลภาพจากชุดข้อมูลฝึกและทำการเรียนรู้ลักษณะของขยะแต่ละประเภท เช่น ขยายที่เป็นแก้ว พลาสติก กระดาษ หรือโลหะ ผ่านกระบวนการฝึกสอน โดยใช้เทคนิคการเรียนรู้ลึก (Deep Learning) และการปรับน้ำหนักในแต่ละชั้นของโมเดล

```
from tensorflow.keras.callbacks import EarlyStopping

early_stopping = EarlyStopping(monitor='val_loss', patience=5, restore_best_weights=True)

history = model.fit(
    train_generator,
    validation_data=val_generator,
    epochs=30,
    callbacks=[early_stopping]
)
```

ภาพที่ 3.16 การฝึกสอนโมเดล CNN

### คำอธิบายโค้ด:

- EarlyStopping(monitor='val\_loss', patience=5, restore\_best\_weights=True)
 

เป็นคำสั่งที่ใช้สร้างอ็อบเจกต์ EarlyStopping ซึ่งทำหน้าที่เฝ้าติดตามค่าสูญเสีย (loss) ของชุดข้อมูล validation (val\_loss) หากค่า val\_loss ไม่ลดลงภายใน 5 epoch ติดต่อกัน (patience=5) ระบบจะหยุดการฝึกหัดที่ เพื่อป้องกันไม่ให้โมเดล overfitting และลดเวลาการฝึกที่ไม่จำเป็น

  - monitor='val\_loss' คือค่าที่ต้องการติดตาม
  - patience=5 คือจำนวนรอบที่อนุญาตให้ไม่มีการปรับปรุงก่อนหยุด
  - restore\_best\_weights=True คือการคืนค่าพารามิเตอร์ของโมเดลไปยังจุดที่ดีที่สุด (ค่าที่ได้จาก epoch ที่มี val\_loss ต่ำที่สุด)
- model.fit(...)
 

คือการสั่งให้โมเดลเริ่มฝึก (train) ด้วยข้อมูลภาพที่ถูกเตรียมไว้ใน train\_generator และตรวจสอบผลด้วย val\_generator ในแต่ละ epoch

  - train\_generator: ตัวแปรที่เก็บข้อมูลฝึก
  - validation\_data=val\_generator: ข้อมูลที่ใช้ตรวจสอบประสิทธิภาพของโมเดลในระหว่างการฝึก
  - epochs=30: ให้ฝึกโมเดลสูงสุด 30 รอบ
  - callbacks=[early\_stopping]: ใช้ EarlyStopping เป็นตัวช่วยหยุดการฝึกหากไม่พบการปรับปรุงของโมเดล

การใช้ EarlyStopping ร่วมกับการฝึกแบบหลาย epoch ทำให้ได้โมเดลที่มีประสิทธิภาพสูงและไม่ overfitting กับข้อมูลฝึกมากเกินไป ทั้งยังช่วยประหยัดเวลาในการฝึก โดยหยุดทันทีเมื่อประสิทธิภาพของโมเดลไม่ดีขึ้นแล้ว

### 3.2.2.4 การประเมินผลโมเดล CNN (Model Evaluation)

หลังจากที่โมเดลได้รับการฝึกฝนแล้ว ขั้นตอนการประเมินผลจะถูกนำมาใช้เพื่อตรวจสอบประสิทธิภาพของโมเดลในการจำแนกประเภทของรีไซเคิล โดยใช้ข้อมูลที่ไม่เคยผ่านการเรียนรู้มาก่อน (Test Set) เพื่อให้ผลการประเมินสะท้อนความสามารถที่แท้จริงของโมเดลในการทำงานกับข้อมูลใหม่

ในการประเมินประสิทธิภาพของโมเดล เราใช้ตัวชี้วัดหลักที่ได้กล่าวถึงแล้วในบทที่ 2 ได้แก่ Accuracy, Precision, Recall, F1-Score และ Loss โดยในบทนี้จะมุ่งเน้นที่การนำตัวชี้วัดเหล่านี้มาใช้เคราะห์ผลลัพธ์จากโมเดลที่ได้ฝึกฝนแล้ว โดยเฉพาะการใช้ข้อมูลชุดทดสอบที่ไม่มีการปรับแต่งเพิ่มเติมใด ๆ (ไม่มีการทำ Data Augmentation)

สำหรับวิธีการประเมินผล เราใช้คำสั่ง model.evaluate() เพื่อวัดค่า Loss และ Accuracy และใช้ไลบรารีจาก sklearn.metrics เพื่อคำนวณ Precision, Recall และ F1-Score จากผลการทำงานของโมเดล เปรียบเทียบกับป้ายชื่อจริง

1. Accuracy: เป็นตัววัดความถูกต้องของโมเดล ซึ่งคำนวณจากจำนวนครั้งที่โมเดลทำนายได้ถูกต้อง
2. Precision: เป็นการวัดความแม่นยำในการทำนายค่าบวก (True Positive) โดยเฉพาะในกรณีที่ข้อมูลไม่สมดุล
3. Recall: เป็นการวัดความสามารถของโมเดลในการทำนายค่าบวกทั้งหมดที่มีอยู่ในข้อมูลจริง
4. F1-Score: คำนวณจาก Precision และ Recall โดยการหาค่าเฉลี่ยเชิงเรขาคณิต ซึ่งจะช่วยให้เราเห็นภาพรวมของการทำนายได้ดีขึ้น
5. Loss: เป็นค่าที่บ่งบอกว่าโมเดลทำงานได้ดีแค่ไหน โดยคำนวณจากการเปรียบเทียบผลลัพธ์ที่โมเดลทำนายกับค่าจริง ซึ่งฟังก์ชันการสูญเสีย (Loss Function) ที่เราใช้ในที่นี้คือ Categorical Cross entropy สำหรับปัญหาหลายคลาส:
  - ค่า Loss ที่ต่ำหมายความว่าโมเดลสามารถทำนายได้ใกล้เคียงกับค่าจริง
  - ค่า Loss ที่สูงแสดงว่าโมเดลมีความผิดพลาดมากในการทำนาย

### 3.3 ประเด็นที่น่าสนใจและสิ่งที่ท้าทาย

#### 3.3.1 ความหลากหลายของลักษณะขยะ

ขยะประเภทเดียวกันอาจมีรูปร่างและลักษณะพื้นผิวที่แตกต่างกัน เช่น ขวดพลาสติกใสกับขวดพลาสติกสี หรือแก้วใสกับแก้วขุ่น ซึ่งอาจส่งผลให้โมเดลสับสนในการแยกประเภท

#### 3.3.2 ขยะที่มีลักษณะคล้ายคลึงกัน

1. แก้วและพลาสติก วัสดุทั้งสองประเภทอาจมีลักษณะโปรดักส์คล้ายกันในบางกรณี เช่น ขวดพลาสติกใสและเศษแก้วใส ทำให้โมเดลอาระบุผิดพลาด
2. กระดาษและพลาสติก บรรจุภัณฑ์พลาสติกบางชนิด เช่น ของอาหาร หรือพลาสติกเคลือบ อาจมีสีสันและลักษณะที่คล้ายกระดาษ ส่งผลให้โมเดลสับสนในการจำแนกประเภทหากลักษณะพื้นผิวไม่ชัดเจนพอ

### 3.3.3 ความท้าทายในการพัฒนาโมเดลโดยการใช้ CNN

#### 1. ความหลากหลายของภาพขยะ

ภาพของขยะในแต่ละประเภทมีลักษณะที่หลากหลาย เช่น สี รูปร่าง ขนาด มุมมอง และพื้นหลังที่แตกต่างกัน ทำให้โมเดลต้องสามารถเรียนรู้ลักษณะที่สำคัญได้อย่างแม่นยำ เพื่อหลีกเลี่ยงความสับสนระหว่างประเภทที่มีลักษณะใกล้เคียงกัน เช่น แก้วใสกับพลาสติก ใส

#### 2. ปริมาณข้อมูลไม่เพียงพอ

CNN ต้องการข้อมูลภาพจำนวนมากเพื่อให้สามารถเรียนรู้และจดจำลักษณะเฉพาะของแต่ละคลาสได้อย่างมีประสิทธิภาพ อย่างไรก็ตาม ข้อมูลภาพขยะที่รวบรวมได้มีจำนวนจำกัด โดยเฉพาะในบางประเภท ส่งผลให้ต้องเพิ่มพากการทำ Data Augmentation เพื่อเพิ่มความหลากหลายของข้อมูล

#### 3. เวลาในการฝึกและทรัพยากรคอมพิวเตอร์

การฝึกโมเดล CNN โดยเฉพาะเมื่อมีข้อมูลจำนวนมากหรือโมเดลมีความซับซ้อนสูง ต้องใช้เวลาในการประมวลผลนานและต้องการทรัพยากรคอมพิวเตอร์ที่มีประสิทธิภาพ เช่น GPU ซึ่งอาจเป็นข้อจำกัดในบางสภาพแวดล้อมการพัฒนา

### 3.3.4 ความท้าทายในการพัฒนาโมเดลโดยการใช้ YOLO

1. การใช้งาน YOLO (You Only Look Once) ต้องอาศัยชุดข้อมูลที่มีการระบุกรอบ (bounding box) อย่างแม่นยำรอบวัตถุที่ต้องการตรวจจับ เช่น ขาดพลาสติก หรือกระดาษในภาพ การ Annotate อย่างไม่ถูกต้องจะส่งผลให้โมเดลเรียนรู้ผิดพลาดและมีความแม่นยำลดลง

2. YOLO อาจตรวจจับวัตถุขนาดเล็ก หรือวัตถุที่มีการทับซ้อนกันในภาพได้ไม่แม่นยำนัก เมื่อเทียบกับเทคนิค segmentation เช่น Mask R-CNN

3. YOLO ต้องใช้ GPU ที่มีหน่วยความจำเพียงพอ เพื่อรับการประมวลผลอย่างมีประสิทธิภาพ

### 3.4 ผลลัพธ์ที่คาดหวัง

การคัดแยกขยะรีไซเคิลจากภาพด้วยเทคนิคการเรียนรู้ของเครื่องสามารถทำได้อย่างแม่นยำ และมีประสิทธิภาพ โดยระบบสามารถจำแนกขยะรีไซเคิลใน 4 ประเภทหลัก ได้แก่ แก้ว พลาสติก กระดาษ และโลหะ ได้อย่างถูกต้อง ช่วยลดความซับซ้อนและเพิ่มความรวดเร็วในกระบวนการจัดการขยะรีไซเคิล

### 3.5 ระบบต้นแบบและผลลัพธ์เบื้องต้น

#### 3.5.1 หน้าหลัก

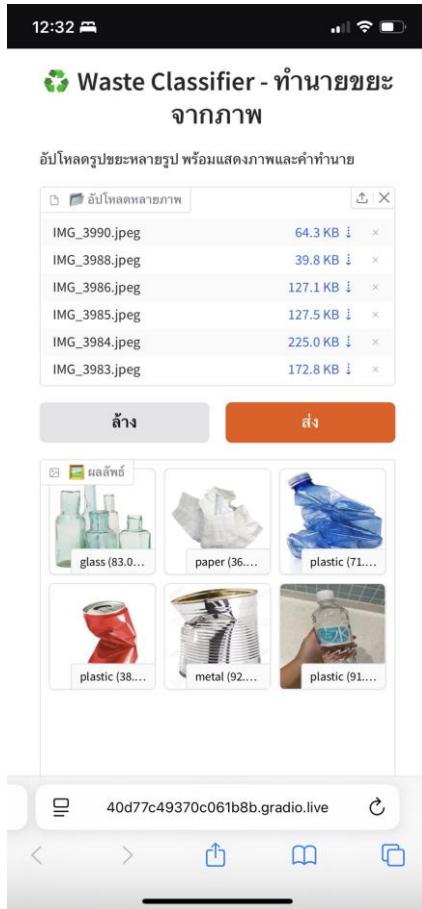
หน้าหลักของระบบการคัดแยกขยะรีไซเคิลจากภาพด้วยเทคนิคการเรียนรู้ของเครื่อง โดยสามารถอปป์โหลดหรือถ่ายรูปภาพของขยะที่ต้องการทราบประเภท โดยรองรับไฟล์ภาพในรูปแบบมาตรฐาน เช่น JPEG และ PNG



ภาพที่ 3.17 หน้าหลักของระบบ

### 3.5.2 หน้าการแสดงผลข้อมูล

เมื่อผู้ใช้งานอัปโหลดรูปภาพและกดส่ง ระบบจะแสดงหมวดหมู่ที่ได้ทำการจำแนก โดยผลลัพธ์ที่ได้จะแสดงประเภทของขยะ เช่น "Paper" พร้อมแสดงภาพต้นฉบับเพื่อยืนยันข้อมูล



ภาพที่ 3.18 หน้าการแสดงผลข้อมูล

ในการแสดงผลลัพธ์การทำนายภาพขยะผ่านอินเทอร์เฟซของ Gradio ระบบจะใช้ฟังก์ชัน predict\_multiple ในการประมวลผลภาพที่ผู้ใช้ส่งเข้ามา โดยเมื่อผู้ใช้อัปโหลดภาพหลายภาพพร้อมกัน ตัวระบบจะทำการเปิดภาพแต่ละไฟล์ แปลงขนาดให้เป็น 224x224 พิกเซล และทำการปรับค่าพิกเซลให้อยู่ในช่วง 0–1 ก่อนนำเข้าสู่โมเดล Convolutional Neural Network (CNN) เพื่อทำการทำนายประเภทของขยะ ภายหลังการทำนาย ระบบจะคำนวณค่าความมั่นใจ (confidence score) จากผลลัพธ์ของโมเดล ซึ่งเป็นค่าความน่าจะเป็นที่ได้จากการฟังก์ชัน softmax โดยเลือกค่าที่มีความน่าจะเป็นสูงที่สุด และแสดงผลในรูปแบบข้อความ เช่น "plastic (92.34%)" โดยหมายถึงว่าโมเดลมีความมั่นใจ 92.34% ว่าภาพนั้นเป็นขยะประเภทพลาสติก

ผลลัพธ์การทํานายทั้งหมดจะถูกแสดงในรูปแบบแกลเลอรี (gr.Gallery) ที่มีการจัดเรียงภาพในลักษณะจากซ้ายไปขวา จำนวน 3 ภาพต่อหนึ่งແຕວ ก่อนจะขึ้นແຕວใหม่เมื่อจำนวนภาพมากขึ้น การนำเสนอผลลัพธ์ในรูปแบบนี้ช่วยให้ผู้ใช้งานสามารถตรวจสอบภาพและคำทํานายของโมเดลได้อย่างชัดเจนและเข้าใจง่าย ทั้งยังสามารถประเมินความแม่นยำของโมเดลจากค่าความมั่นใจที่แสดงใต้ภาพแต่ละภาพ

## บทที่ 4

### ผลการดำเนินงาน

#### 4.1 การจัดเตรียมฮาร์ดแวร์และซอฟต์แวร์

##### 4.1.1 ฮาร์ดแวร์ที่ใช้ในการพัฒนา

(1) คอมพิวเตอร์ที่ใช้ระบบปฏิบัติการ Windows จำนวน 1 เครื่อง

ลักษณะ	คอมพิวเตอร์
ชื่อรุ่น (Name)	HP ENVY Laptop 13 -aq0xxx
ระบบประมวลผล (Processor)	Intel(R) Core(TM) i7-8565U CPU @ 1.80GHz 1.99 GHz
หน่วยความจำ (Memory)	8.00 GB
ส่วนประมวลผลภาพ (Graphic)	Intel Iris Xe Graphics
ส่วนเก็บข้อมูล (Storage)	SSD 1TB NVMe PCIe M.2
ระบบปฏิบัติการ (Operating System)	Windows 11

ตารางที่ 4.1 ข้อมูลฮาร์ดแวร์ที่ใช้ในการพัฒนา

#### 4.1.2 ซอฟต์แวร์ที่ใช้ในการพัฒนาโปรแกรม

#### 4.1.2.1 ภาษาที่ใช้ในการพัฒนา

Python ใช้เป็นภาษาหลักสำหรับการพัฒนาโปรแกรม และการสร้างโมเดล TensorFlow, การประมวลผลภาพ

#### 4.1.2.2 โปรแกรมที่ใช้ในการพัฒนา

Google colab pro แพลตฟอร์มออนไลน์ที่ช่วยในการเขียนและรันโค้ด Python พร้อมการรองรับ GPU สำหรับเร่งการประมวลผล

#### 4.1.2.3 เฟรมเวิร์กและไลบรารีที่ใช้

1. Tensorflow
  2. Keras
  3. Numpy และ Pandas
  4. Matplotlib และ Seaborn
  5. Scikit-learn
  6. Gradio

## 4.2 แผนการดำเนินงาน

ตารางที่ 4.2 การดำเนินงานที่ผ่านมาจนถึงปัจจุบัน



4.ทดสอบและรับปรุ่งโมเดล																	
5.จัดทำเอกสารโครงงาน																	
6.นำเสนอข้อเสนอโครงงานพิเศษ 2																	

### 4.3 ผลการทดลอง

เพื่อประเมินวิธีการที่เหมาะสมที่สุดสำหรับการพัฒนาโมเดลจำแนกประเภทขยะรีไซเคิล จึงได้ดำเนินการทดลองเบื้องต้นโดยเปรียบเทียบสองเทคนิค ได้แก่ YOLO ซึ่งเป็นอัลกอริธึมสำหรับตรวจจับวัตถุในภาพ และ CNN ซึ่งเน้นการจำแนกประเภทของภาพโดยตรง โดยทั้งสองแนวทางให้ผลลัพธ์ที่ชัดเจนและมีประสิทธิภาพมากกว่าเดิม สำหรับการทดลองนี้ได้ทำการทดลองทั้งหมด 300 รูป รวมจำนวน 1,200 รูป เพื่อให้การทดลองเป็นไปอย่างยุติธรรม

#### 4.3.1 ผลลัพธ์เบื้องต้นของโมเดล YOLO

##### 4.3.1.1 ค่าตัวชี้วัดการประเมินโมเดล YOLO (Evaluation Metrics)

เมื่อนำโมเดลมาทดสอบกับชุดทดสอบ มีรายละเอียดดังนี้

Class	จำนวนรูป	Instances	Precision	Recall	F1-Score	mAP50	mAP50-95
Glass	30	43	0.975	0.901	0.9365	0.960	0.567
Metal	30	45	0.872	0.756	0.8101	0.883	0.546
Paper	30	35	0.907	0.857	0.8811	0.860	0.472
Plastic	30	35	0.827	0.886	0.8552	0.933	0.614
รวมทั้งหมด	120	158	0.895	0.850	0.8721	0.909	0.550

ตารางที่ 4.3 ผลการประเมินประสิทธิภาพของโมเดล YOLO สำหรับการจำแนกประเภทขยะรีไซเคิลรายคลาส

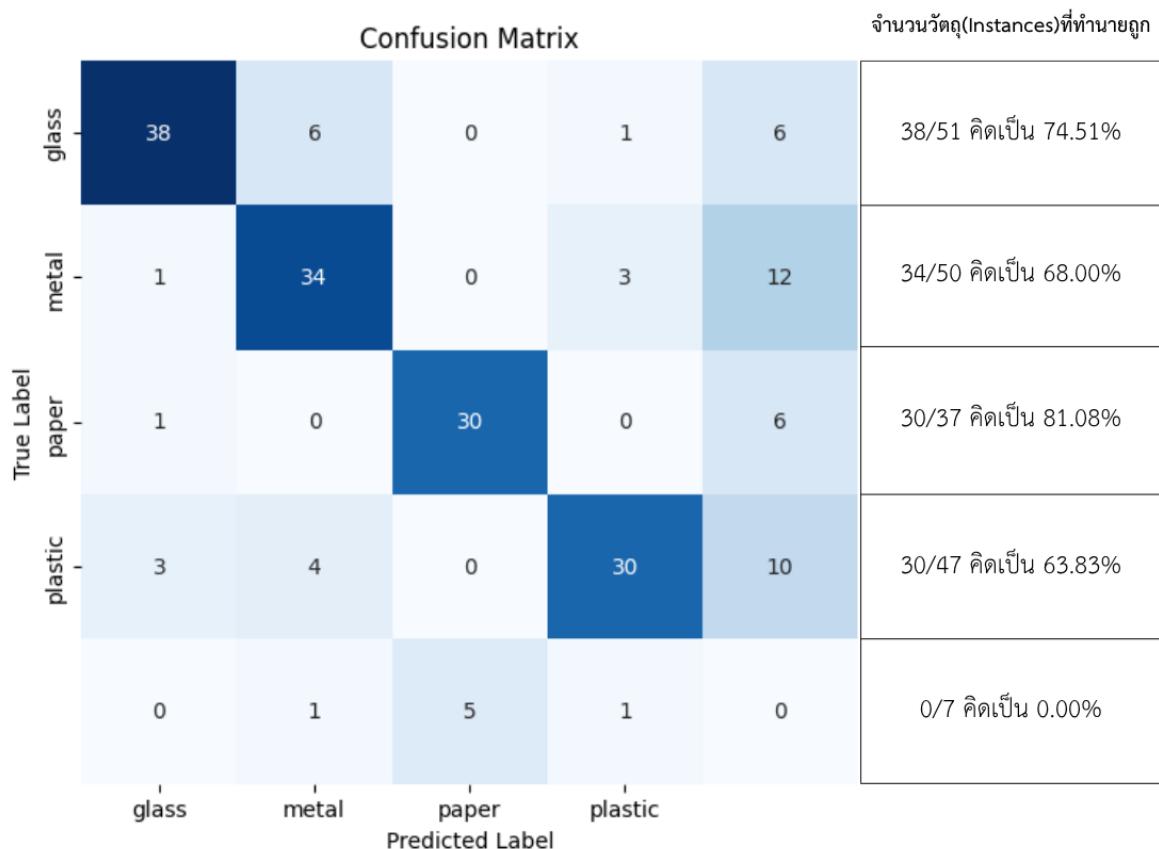
จากตารางผลการประเมินประสิทธิภาพของโมเดล YOLO บนชุดข้อมูลทดสอบพบว่าโมเดลสามารถจำแนกประเภทขยะได้ในระดับที่น่าพอใจ โดยมีค่าความแม่นยำเฉลี่ย

(mAP@0.5) เท่ากับ 0.909 และค่า ความแม่นยำเฉลี่ยแบบละเอียด (mAP@0.5:0.95) เท่ากับ 0.550 ซึ่งสะท้อนให้เห็นว่าโมเดลสามารถระบุตำแหน่งของวัตถุได้อย่างแม่นยำภายใต้เงื่อนไขที่แตกต่างกัน ของการซ่อนทับกันระหว่างกรอบทำนายและกรอบจริง

โดยเมื่อพิจารณาในแต่ละประเภท พบร่วม

- แก้ว มีผลการประเมินผลที่ดีที่สุดในทุกตัวชี้วัด โดยมีค่า Precision เท่ากับ 0.975, Recall เท่ากับ 0.901 และ mAP@0.5 เท่ากับ 0.960 แสดงให้เห็นว่าโมเดลสามารถตรวจจับและจำแนกภาพประเภทขวดแก้วได้อย่างแม่นยำและครอบคลุม
- พลาสติก เป็นประเภทเดียวที่มีค่า Recall สูงที่สุดที่ 0.886 และสูงกว่า Precision สะท้อนให้เห็นว่าโมเดลสามารถระบุวัตถุประเภทพลาสติกได้เกือบครบถ้วน แม้ Precision จะต่ำกว่าเล็กน้อย
- โลหะ มีค่า Recall ต่ำกว่าทุกประเภท โดยมีค่าอยู่ที่ 0.756 ซึ่งอาจเกิดจากลักษณะทางภาพที่คล้ายคลึงกับคลาสอื่น หรือมีความหลากหลายน้อยในภาพฝึก ทำให้โมเดลไม่สามารถเรียนรู้ลักษณะเฉพาะได้เพียงพอ
- กระดาษ มีค่าความแม่นยำเฉลี่ย ค่อนข้างต่ำที่สุดในกลุ่มที่ 0.860 และค่า mAP@0.5:0.95 ที่ 0.472 ซึ่งแสดงให้เห็นถึงข้อจำกัดของโมเดลในการแยกแยะภาพกระดาษ โดยอาจเนื่องจากลักษณะที่หลากหลายและคล้ายกับพลาสติกบางประเภท

#### 4.3.1.2 การประเมินโมเดล YOLO ด้วยค่า Confusion Matrix ของชุดทดสอบ



ภาพที่ 4.1 Confusion Matrix ของ YOLO

เมื่อดูภาพค่า Confusion Matrix จะเห็นได้ว่าโมเดลสามารถจำแนกภาพขยะประเภทต่าง ๆ ได้อย่างแม่นยำในระดับหนึ่ง โดยเฉพาะคลาส แก้ว และ โลหะ ที่มีจำนวนการทำนายถูกต้องสูง เพราะ มีจำนวน Instances ที่ค่อนข้างเยอะ แต่จะเห็นได้ว่าคลาสที่ทำนายผิดน้อยที่สุดคือคลาสกระดาษ ซึ่งผิดเพียงแค่ 7 ภาพ ทำให้เปอร์เซ็นต์ในการทำนายคลาสนี้สูงสุดถึง 81.09% ทั้งนี้ จำนวน instances ที่เพิ่มขึ้นจากเดิมในคลาสบางประเภท อาจเกิดจากลักษณะการทำงานของ YOLO ทำให้บางครั้งเกิด การจับช้าช้อนกันหรือจับวัตถุผิดตำแหน่ง ซึ่งส่งผลให้จำนวน instances ที่โมเดลประเมินในคลาสใดคลาสนั้นเพิ่มขึ้น และอาจทำให้การประเมินด้วย Confusion Matrix แสดงค่าที่มากกว่าจำนวนภาพจริงที่มีในชุดทดสอบ นอกจากนี้ยังพบว่ามีคลาสว่าง หรือ background โผล่ขึ้นมาใน Confusion Matrix แม้ว่าใน dataset จะไม่มีการสร้างคลาสนี้ไว้ก็ตาม ซึ่งเกิดจากการที่ YOLO พยายามตรวจจับภาพ และบางครั้งจะจัด Label ให้วัตถุที่จับไม่ได้ นั่นหมายความว่าคลาสนี้ เป็นคลาสที่โมเดลไม่สามารถจำแนกได้อย่างชัดเจน

ตัวอย่างการนำมายของโมเดลจากข้อมูลทดสอบ

### 1. ภาพที่มีการลากกรอบจากข้อมูลจริง (Ground Truth)

เป็นภาพที่แสดงตำแหน่งของขยะรีไซเคิลแต่ละชิ้นตามข้อมูลที่ได้จากผู้จัดทำขุดข้อมูล (label ที่ถูกต้อง) โดยจะระบุกรอบวัตถุและคลาส เช่น glass, plastic, paper หรือ metal เพื่อใช้เป็นมาตรฐานสำหรับประเมินผลการทำงานของโมเดล



ภาพที่ 4.2 ภาพที่มีการลากกรอบจากข้อมูลจริง ของ YOLO

## 2. ภาพที่ไม่เมเดลทำนาย (Prediction Results)

เป็นภาพที่แสดงผลการทำนายของ YOLO โดยระบบจะลากกรอบที่ตรวจจับได้ พร้อมแสดงชื่อคลาสที่ไม่เมเดลคาดการณ์ไว้ และค่า confidence score ซึ่งแสดงถึงความมั่นใจของโมเดลในการจำแนกวัตถุในแต่ละกรอบ



ภาพที่ 4.3 ภาพที่ไม่เมเดล YOLO ทำนาย

จากการเปรียบเทียบภาพห้องประภาก็คือ ภาพที่มีการลากกรอบจากข้อมูลจริง (Ground Truth) และภาพที่ไม่เมเดล YOLO ทำนาย (Prediction Results) พบว่าไม่เมเดลมีความสามารถในการตรวจจับวัตถุได้ดีในหลายกรณี อย่างไรก็ตาม จากการสังเกตภาพตัวอย่างจะเห็นว่าไม่เมเดล มีแนวโน้มที่จะ ทำนายผิดพลาดโดยเฉพาะในกรณีของวัตถุประเภท "แก้ว" ที่มักถูกทำนายเป็น "พลาสติก" อยู่บ่อยครั้ง

ข้อผิดพลาดนี้อาจเกิดจากลักษณะทางกายภาพของขยะห้องประภาก็คือความใกล้เคียงกันในบางสถานการณ์ เช่น แก้วใส่ที่มีรูปร่างคล้ายขวดพลาสติก หรือการสะท้อนแสงของพื้นผิววัตถุ ซึ่ง

อาจทำให้โมเดลแยกแยะได้ยาก โดยเฉพาะหากจำนวนตัวอย่างในชุดข้อมูลไม่เพียงพอหรือไม่มีความหลากหลายมากพอยในลักษณะของขยะประเภทแก้ว

#### 4.3.1.3 การทดสอบประสิทธิภาพโมเดล YOLO กับข้อมูลนอก Dataset

หลังจากที่โมเดล YOLO ได้รับการฝึกจากชุดข้อมูลที่ประกอบด้วยภาพขยะรีไซเคิล 4 ประเภท ได้แก่ แก้ว โลหะ กระดาษ และพลาสติก โดยผ่านกระบวนการ Annotation และ Data Augmentation แล้วนั้น ได้มีการนำโมเดลที่ได้ไปทดสอบกับชุดข้อมูลใหม่ซึ่งไม่ได้อยู่ในชุดฝึก (Training Set), ชุดตรวจสอบ (Validation Set) หรือชุดทดสอบ (Test Set) เพื่อทดสอบประสิทธิภาพของโมเดลในการตรวจจับและจำแนกประเภทขยะรีไซเคิล นี้ว่าตกลงมีความสามารถในการจำแนกประเภทของขยะที่มีหลายประเภทรวมอยู่ในภาพได้มากแค่ไหน ซึ่งมักจะเป็นภาพถ่ายของขยะที่มีหลายประเภทรวมอยู่ในภาพเดียวกัน โดยพิจารณาความแม่นยำในการตรวจจับวัตถุ ความสามารถในการจำแนกประเภท และความถูกต้องของตำแหน่งวัตถุที่ตรวจจับได้

ภาพที่	รูปภาพต้นฉบับ	รูปภาพที่ผ่านการทำนายโดย YOLO
1.ภาพหมู่ที่มีการกำหนดขอบเขตแต่ละวัตถุเท่ากัน		
2.ภาพหมู่ที่มีขอบเขตของวัตถุไม่ชัดเจน		
3.ภาพหมู่ที่หักซ้อนกันและไม่มีขอบเขต		

ตารางที่ 4.4 ผลการทดสอบโมเดล YOLO กับภาพถ่ายที่มีขยะหลายประเภทรวมอยู่ในภาพเดียว

ผู้วิจัยได้ทดสอบโดยใช้ภาพที่ประกอบด้วยวัตถุขยะหลายประเภทอยู่ร่วมกันในภาพเดียวดังตารางที่ 4.4 มีวัตถุขยะจำนวนหลายชิ้นวางอยู่ในตำแหน่งต่าง ๆ ภายในภาพเดียวกัน เพื่อประเมินความสามารถของโมเดล YOLO ในการตรวจจับและจำแนกวัตถุหลายรายการพร้อมกัน พบว่า ประสิทธิภาพของโมเดลลดลงตามความซับซ้อนของขอบเขตของขยะแต่ละชิ้นในภาพ

โมเดลสามารถตรวจจับและจำแนกวัตถุขยะได้บางส่วนเท่านั้น โดยมีกรณีที่โมเดล ไม่สามารถตรวจจับวัตถุได้ครบถ้วนในภาพที่ 2 และ 3 และบางชิ้นมีการตรวจจับช้าช้อน (ภาพที่ 3) หรือ ทำนายประเภทผิด ตัวอย่างเช่น วัตถุหนึ่งชิ้นถูกทำนายเป็น 2 คลาสพร้อมกัน หรือมีการทำนายประเภทผิดจากคลาสจริง ซึ่งสะท้อนให้เห็นถึงข้อจำกัดของโมเดลเมื่อต้องจัดการกับภาพที่มีความซับซ้อน หรือมีวัตถุจำนวนมากในบริบทเดียวกัน

ข้อสังเกตดังกล่าวแสดงให้เห็นว่า แม้โมเดล YOLO จะสามารถทำงานได้ดีในกรณีภาพที่มีวัตถุเดียวหรือพื้นหลังไม่ซับซ้อน แต่เมื่อใช้งานในสถานการณ์จริงที่มีความซับซ้อนสูงขึ้น โมเดลยังจำเป็นต้องได้รับการปรับปรุงเพิ่มเติม ทั้งในด้านคุณภาพของข้อมูลฝึก และเทคนิคการเพิ่มความสามารถของโมเดล เช่น การใช้ข้อมูลฝึกที่มีความหลากหลายมากขึ้น หรือการปรับแต่งพารามิเตอร์เพิ่มเติม

#### 4.3.2 ผลลัพธ์เบื้องต้นของโมเดล CNN (Convolutional Neural Networks)

##### 4.3.2.1 ค่าตัวชี้วัดการประเมินโมเดล CNN (Evaluation Metrics)

เมื่อนำโมเดลมาทดสอบกับชุดทดสอบ มีรายละเอียดดังนี้

Class	จำนวนรูป	Precision	Recall	F1-Score
Glass	30	<b>0.82</b>	0.47	0.60
Metal	30	0.57	<b>0.70</b>	<b>0.63</b>
Paper	30	0.67	0.60	<b>0.63</b>
Plastic	30	0.49	0.63	0.55
Accuracy		<b>0.60</b>		

ตารางที่ 4.5 ผลการประเมินประสิทธิภาพของโมเดล CNN สำหรับการจำแนกประเภทรีไซเคิลรายคลาส

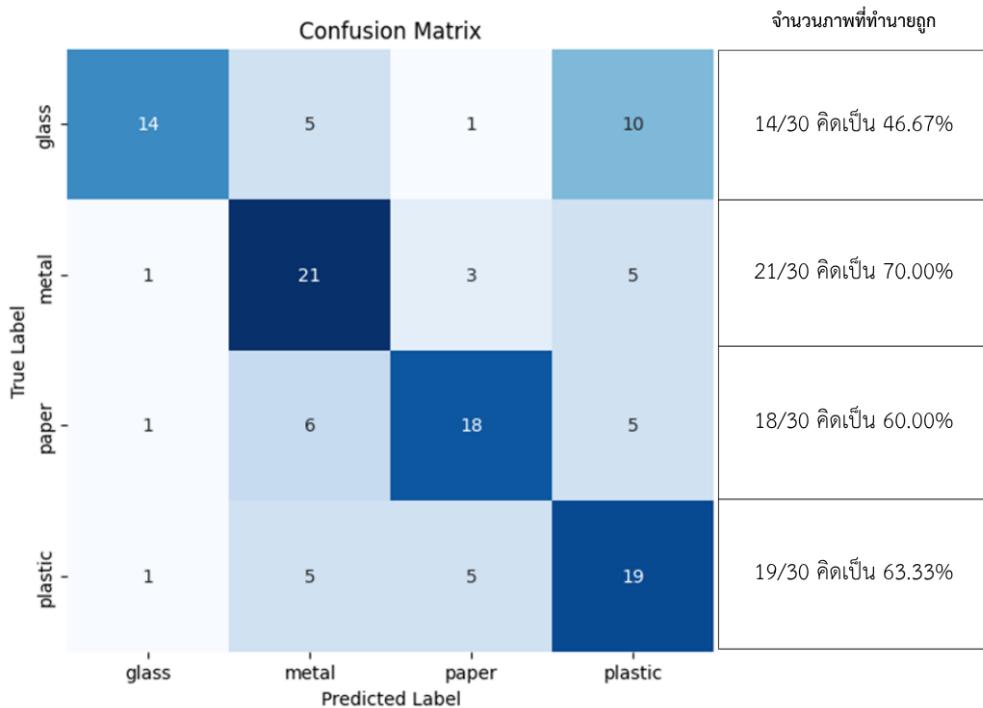
จากการประเมินประสิทธิภาพของโมเดล CNN บันชุดข้อมูลทดสอบ พบร่วมโมเดลเมื่อความสามารถในการจำแนกประเภทของรีไซเคิลในระดับปานกลาง โดยมีค่า Accuracy รวมอยู่ที่ประมาณ 60% ซึ่งสะท้อนถึงความแม่นยำโดยรวมของโมเดล

เมื่อพิจารณาค่า Precision, Recall และ F1-Score ในแต่ละคลาส พบร่วม

- คลาสแก้ว (Glass) มีค่า Precision เท่ากับ 0.82 และ Recall เท่ากับ 0.47 ทำให้ F1-Score เท่ากับ 0.60 แสดงว่าโมเดลมีความแม่นยำในการทำนายว่าภาพใดเป็นแก้ว (Precision สูง) แต่ยังไม่สามารถตรวจจับภาพแก้วทั้งหมดได้ครบถ้วน (Recall ต่ำ)
- คลาสโลหะ (Metal) มีค่า Precision เท่ากับ 0.57 และ Recall เท่ากับ 0.70 ทำให้ F1-Score เท่ากับ 0.63 ซึ่งแสดงว่าโมเดลสามารถระบุภาพโลหะได้ค่อนข้างครอบคลุม และมีความแม่นยำในระดับที่พอใช้ได้
- คลาสกระดาษ (Paper) มีค่า Precision เท่ากับ 0.67 และ Recall เท่ากับ 0.60 ทำให้ F1-Score อยู่ที่ 0.63 ซึ่งสะท้อนถึงประสิทธิภาพที่สมดุลพอสมควรของโมเดลในการจำแนกภาพประเภทกระดาษ
- คลาสพลาสติก (Plastic) มีค่า Precision เท่ากับ 0.49 และ Recall เท่ากับ 0.63 ทำให้ F1-Score เท่ากับ 0.55 แสดงว่าโมเดลสามารถตรวจจับพลาสติกได้ในระดับหนึ่ง แต่ยังมีความผิดพลาดในการทำนายภาพประเภทอื่นว่าเป็นพลาสติก (Precision ต่ำ)

โดยภาพรวม โมเดล CNN ยังมีข้อจำกัดในด้านความสามารถในการตรวจจับบางคลาสได้ครบถ้วนและแม่นยำในทุกรูปแบบ ซึ่งอาจเกิดจากจำนวนข้อมูลที่ยังไม่มากพอ หรือความซับซ้อนของลักษณะภาพในแต่ละคลาส ทำให้โมเดลยังมีโอกาสทำนายผิดพลาดในบางกรณี

#### 4.3.2.2 การประเมินโมเดล CNN ด้วยค่า Confusion Matrix ของชุดทดสอบ



ภาพที่ 4.4 Confusion Matrix ของ CNN

จาก Confusion Matrix พบร่วมกัน CNN สามารถจำแนกขยะประเภท โลหะ ได้ดีที่สุด โดยมีจำนวนการทำนายถูกต้องมากที่สุด จำนวน 21 รูป จาก 30 รูป ซึ่งคิดเป็น 70% ในขณะเดียวกันมีความสับสนสูงสุดในคลาสแก้ว ซึ่งทำนายว่าเป็นพลาสติก แสดงให้เห็นว่าโมเดลยังมีข้อจำกัดในการแยกแยะลักษณะที่คล้ายกันระหว่างบางประเภทของขยะ

- ตัวอย่างการทำนายที่ผิดพลาดของโมเดลจากข้อมูลทดสอบ



ภาพที่ 4.5 ตัวอย่างการทำนายที่ผิดพลาดของโมเดลจากข้อมูลทดสอบ CNN

จากการวิเคราะห์ผลภาพที่โมเดล CNN ทำนายผิด พบร้าภาพที่เป็นขวดแก้วจำนวนมากถูกจำแนกผิดให้เป็นพลาสติก ซึ่งอาจเกิดจากหลายปัจจัยร่วมกัน โดยเฉพาะลักษณะทางกายภาพของขวดแก้วและขวดพลาสติกที่มีความคล้ายคลึงกัน เช่น รูปร่างที่เป็นทรงกระบอก ความโปร่งใส และมีผ้าปิดเหมือนกัน ซึ่งอาจทำให้โมเดลสับสนได้ นอกจากนี้ หากลักษณะพื้นผิวหรือสีของขวดแก้วไม่เด่นชัด หรือมีแสงสะท้อนในภาพ อาจส่งผลให้โมเดลไม่สามารถแยกแยะลักษณะเฉพาะของแก้วออกจากพลาสติกได้อย่างชัดเจน อีกทั้งในกรณีที่จำนวนข้อมูลของคลาสแก้วในชุดฝึกมีน้อยกว่าหรือมีความหลากหลายน้อยกว่าคลาสพลาสติก ก็อาจทำให้โมเดลเกิดความเอนเอียงในการทำนาย และมีแนวโน้มจำแนกวัตถุที่ไม่ชัดเจนเป็นพลาสติกมากกว่าแก้ว ดังนั้น ข้อผิดพลาดในการจำแนกนี้สะท้อนถึงข้อจำกัดของโมเดลในการแยกแยะวัสดุที่มีลักษณะคล้ายกัน และเป็นประเด็นที่ควรพิจารณาเพิ่มเติมในการปรับปรุงโมเดลต่อไป

#### 4.3.2.3 การทดสอบประสิทธิภาพโมเดล CNN กับข้อมูลนอกชุด Dataset

ในการทดสอบประสิทธิภาพของโมเดล CNN กับข้อมูลนอกชุดฝึก (external test data) ได้เลือกใช้ภาพชุดเดียวกับที่ใช้ในการประเมินโมเดล YOLO เพื่อให้สามารถเปรียบเทียบผลการทำงานของทั้งสองโมเดลได้อย่างเป็นธรรม อย่างไรก็ตาม เนื่องจากข้อจำกัดของโมเดล CNN ซึ่งถูกออกแบบมาเพื่อการจำแนกประเภทเพียงชั้นเดียวต่อภาพ จึงไม่สามารถรองรับการทำนายหลายวัตถุใน

ภาพเดียวกันได้โดยตรง ดังนั้น จึงได้มีการตัดภาพออกเป็นส่วนย่อย ๆ โดยแยกแต่ละวัตถุออกจากกัน ก่อนนำเข้าสู่โมเดล CNN เพื่อให้สามารถประมวลผลและท่านายประเภทของขยะได้อย่างถูกต้อง วิธีนี้ อาจมีผลต่อความแม่นยำของการท่านาย เนื่องจากภาพที่ตัดออกมาอาจสูญเสียบริบทบางส่วนที่ จำเป็นต่อการจำแนกวัตถุ แต่ก็ถือเป็นแนวทางที่เหมาะสมภายใต้ข้อจำกัดของสถาปัตยกรรมโมเดล CNN

รูปภาพต้นฉบับ	รูปภาพที่ผ่านการทำนายโดย CNN
	<p>Predicted: glass (0.98)</p> 
	<p>Predicted: paper (0.49)</p> 

	Predicted: plastic (0.91)
	Predicted: metal (0.75)
	Predicted: plastic (0.43)

	Predicted: plastic (0.83)
	Predicted: metal (0.78)
	Predicted: plastic (0.43)

	Predicted: metal (0.68)
	Predicted: plastic (0.32)
	Predicted: metal (0.74)

	Predicted: metal (0.70)
	Predicted: plastic (0.93)
	Predicted: glass (0.91)



ตารางที่ 4.6 ผลการทดสอบโมเดล CNN กับข้อมูลนอก Dataset

เมื่อดำเนินการทดสอบกับข้อมูลนอกชุดฝึกด้วยวิธีดังกล่าว พบร่วมกัน CNN สามารถจำแนกประเภทขยะได้ในระดับที่น่าพึงพอใจ แม้จะมีกรณีที่โมเดลทำนายผิดบ้าง เช่น การสับสนระหว่างคลาสที่มีลักษณะใกล้เคียงกัน อย่างเช่น แก้วและพลาสติก อย่างไรก็ตาม โมเดลยังสามารถตรวจจับและจำแนกวัตถุได้ดีกว่า YOLO ในบางกรณี โดยเฉพาะในภาพที่ YOLO ไม่สามารถตรวจจับวัตถุใด ๆ ได้เลย ในขณะที่ CNN ยังสามารถประมวลผลจากภาพที่ถูกตัดย่ออยู่ก่อนมาและให้ผลการทำนายที่ตรงกับความเป็นจริงได้ในหลายภาพ สะท้อนให้เห็นถึงข้อได้เปรียบของโมเดล CNN

#### 4.3.3 การเปรียบเทียบและวิเคราะห์ความเหมาะสมของโมเดล CNN และ YOLO สำหรับการจำแนกประเภทขยะรีไซเคิล

เพื่อเลือกแนวทางที่เหมาะสมที่สุดสำหรับการพัฒนาโมเดลจำแนกประเภทขยะรีไซเคิล ได้มีการเปรียบเทียบประสิทธิภาพและข้อดีข้อเสียของโมเดล CNN และ YOLO ซึ่งเป็นโมเดลที่ใช้กันอย่างแพร่หลายในงานด้านการประมวลผลภาพ โดยทั้งสองโมเดลได้ผ่านการทดลองกับชุดข้อมูลเดียวกัน และมีการประเมินผลด้วยเกณฑ์ต่าง ๆ เพื่อวิเคราะห์ความเหมาะสมในเชิงการใช้งานจริง

##### 1. การเปรียบเทียบประสิทธิภาพของโมเดลด้วยค่าชี้วัด

ในการพัฒนาโมเดลเพื่อจำแนกประเภทขยะรีไซเคิลอย่างแม่นยำ โดยใช้ชุดข้อมูลเดียวกันในการประเมินเพื่อให้การเปรียบเทียบเป็นธรรม ทั้งสองโมเดลถูกประเมินด้วยค่าชี้วัด F1-Score ซึ่งสะท้อนถึงสมดุลระหว่างความแม่นยำ (Precision) และความครอบคลุม (Recall) ของการทำนายผลลัพธ์

โมเดล	F1-Score (Glass)	F1-Score (Metal)	F1-Score (Paper)	F1-Score (Plastic)	ค่าเฉลี่ย โดยรวม
YOLO	0.94	0.81	0.88	0.86	0.87
CNN	0.60	0.63	0.63	0.55	0.60

ตารางที่ 4.7 การเปรียบเทียบประสิทธิภาพของโมเดลด้วยค่าชี้วัด

จากผลการประเมินพบว่าโมเดล YOLO มีค่า F1-Score เฉลี่ยในแต่ละคลาสอยู่ในช่วงประมาณ 0.81 ถึง 0.94 ขณะที่โมเดล CNN มีค่า F1-Score อยู่ในช่วงประมาณ 0.55 ถึง 0.63 ซึ่งสะท้อนว่า YOLO มีความสามารถในการจำแนกภาพได้แม่นยำกว่าจากมุมมองของค่าชี้วัดทางเทคนิค โดยเฉพาะในคลาส “แก้ว” ที่ YOLO มีค่า F1-Score สูงถึง 0.94 ในขณะที่ CNN อยู่ที่เพียง 0.60 ซึ่งอาจเป็นผลมาจากการข้อได้เปรียบของ YOLO ในการตรวจจับวัตถุ ที่สามารถระบุขอบเขตของวัตถุ (bounding box) ได้อย่างแม่นยำและสามารถใช้บริบทของภาพโดยรอบในการช่วยตัดสินใจ

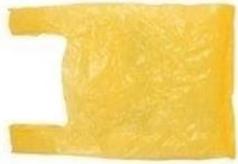
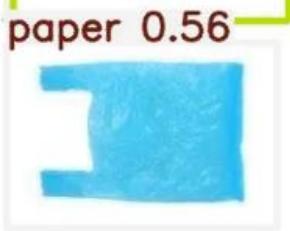
นอกจากนี้ YOLO ยังมีโครงสร้างที่สามารถจัดการกับการมีวัตถุหลายชิ้นในภาพเดียวganได้อย่างมีประสิทธิภาพ ขณะที่ CNN ที่ใช้ในการทดลองนี้เป็นสถาปัตยกรรมสำหรับการจำแนกภาพเพียงภาพหละหนึ่งคลาสเท่านั้น ซึ่งหมายความว่าในกรณีที่มีวัตถุหลายชิ้นหรือมีความซับซ้อนในภาพ CNN จะมีข้อจำกัดในการประมวลผลและอาจทำให้ผลการทำนายคลาดเคลื่อนได้มากขึ้น

## 2. การเปรียบเทียบประสิทธิภาพของโมเดลกับข้อมูลนอก Dataset

ทั้ง YOLO และ CNN มีการใช้ Dataset ชุดเดียวกันในการทดสอบ และเพื่อให้เปรียบเทียบกันได้อย่างชัดเจนจึงได้ทำการ crop รูปต้นฉบับ ให้เห็นภาพขยายเพียงชิ้นเดียวต่อหนึ่งภาพ และคุณลักษณะจากการทำนาย ดังตาราง

ภาพต้นฉบับ	ผลลัพธ์จาก YOLO	ผลลัพธ์จาก CNN	หมายเหตุ
			ผลลัพธ์ของ YOLO ตีกรอบผลลัพธ์ไม่ชัดเจนและทำนายไม่ถูกต้อง แต่ CNN ทำนายถูกต้อง คือ แก้ว

			YOLO ไม่ได้มีการอบรมกว่าสิ่งนี้คืออะไร แต่ CNN ทำนายถูกต้องคือกระดาษ
			YOLO ไม่ได้ตีกรอบบวกกว่าสิ่งนี้คืออะไร แต่ CNN ทำนายถูกต้องคือพลาสติก
			YOLO และ CNN ทำนายว่าเป็นโลหะซึ่งถูกต้องแต่ YOLO ตีกรอบไม่ชัดเจน
			ทั้ง 2 โมเดลทำนายผิดและ YOLO ตีกรอบไม่ชัดเจนว่าเป็นขยะขี้นที่

	<p><b>paper 0.56</b></p> 	 Predicted: plastic (0.83)	YOLO ทำนายคลาสพิเศษ แต่ CNN ได้ทำนายถูก คือพลาสติก
	<p><b>paper 0.57</b></p> 	 Predicted: metal (0.78)	YOLO ได้ทำนายถูกต้อง แต่ CNN ทำนายผิด
	<p><b>plastic 0.88</b></p> 	 Predicted: plastic (0.43)	ทั้ง 2 โมเดล ทำนายผิดทั้งคู่
	<p><b>plastic 0.48</b></p>  <b>----- 0.56 -----</b>	 Predicted: metal (0.68)	ทั้ง 2 โมเดล ทำนายผิดทั้งคู่ เพราะในภาพ คือกระดาษ
	<p><b>plastic 0.76</b></p>  <b>----- paper 0.57 -----</b>	 Predicted: plastic (0.32)	ทั้ง 2 โมเดล ทำนายผิดทั้งคู่ เพราะในภาพ คือกระดาษ

			YOLO ทำນາຍ ได้ถูกຕີກຕ້ອງແຕ່ນີ້ ການຕີກຮອບ ຂ້ອນ ແສດຈຳໄມ່ ຖຶງຄວາມໄມ່ ຫັດເຈນ ແລະ CNN ทำນາຍ ຜິດ
			YOLO ทำນາຍ ຜິດ ແຕ່ CNN ທຳນາຍໄດ້ຖືກ ເພຣະຄືວ ໂລຂະ
			YOLO ทำນາຍ ຖືກ ແຕ່ CNN ທຳນາຍຜິດ
			ທັງ 2 ໂມເດລ ທຳນາຍໄດ້ ຖືກຕ້ອງ
			ທັງ 2 ໂມເດລ ທຳນາຍໄດ້ ຖືກຕ້ອງ

ตารางที่ 4.8 การเปรียบเทียบประสิทธิภาพของโมเดลกับข้อมูลนอก Dataset

จากผลการทดสอบโมเดลกับภาพจริงนอก Dataset พบว่า แม้ว่า YOLO จะเป็นโมเดลแบบ object detection ที่มีความสามารถในการระบุตำแหน่งและจำแนกวัตถุหลายชิ้นในภาพเดียวกัน แต่จากผลการทดลองพบว่ามีหลายกรณีที่ YOLO ไม่สามารถตีกรอบวัตถุได้ชัดเจน หรือแม้แต่ไม่สามารถตรวจจับวัตถุในภาพได้เลย ในขณะที่โมเดล CNN แม้จะมีข้อจำกัดที่ต้องรับข้อมูลเพียงวัตถุเดียวต่อภาพ แต่กลับสามารถทำนายคลาสของขยะได้ถูกต้องในหลายกรณี

จากตัวอย่างที่ได้เปรียบเทียบ:

- CNN สามารถทำนายคลาสได้ถูกต้องในหลายภาพที่ YOLO ไม่สามารถตรวจจับได้เลย
- YOLO แม้จะทำนายคลาสถูกต้องบ้าง แต่บางครั้งตีกรอบไม่ชัดเจนหรือซ้อนกัน ทำให้ยากต่อการนำผลไปใช้งานจริง
- มีบางกรณีที่ทั้งสองโมเดลทำนายผิด หรือ YOLO ทำนายถูกแต่ CNN ทำนายผิด ซึ่งเป็นเรื่องที่พบได้จากข้อจำกัดของข้อมูลและโมเดลแต่ละแบบ

### 3. การเปรียบเทียบคุณสมบัติและผลลัพธ์ของโมเดล CNN และ YOLO

	CNN	YOLO
ประเภทงาน	classification (จำแนกประเภทภาพ)	Object Detection (ตรวจจับวัตถุ)
ความแม่นยำ F1-Score	60%	87%
การเตรียมข้อมูล	ไม่ต้อง Annotate	ต้อง Annotate กำหนดตำแหน่งวัตถุให้ชัดเจน
ความง่ายในการพัฒนา	ง่ายกว่า เนื่องจากต้องจำแนกภาพ	ซับซ้อนกว่า ต้องใช้เวลาในการเตรียมข้อมูล
ความสามารถในงานตรวจจับ	จำกัด ไม่สามารถระบุหลายวัตถุในภาพได้	สามารถตรวจจับหลายวัตถุในภาพเดียว
ความเร็วในการประมวลผล	ปานกลาง	เร็วมาก
ความต้องการทรัพยากร	ใช้ทรัพยากรในการประมวลผลน้อยกว่า	ต้องการทรัพยากรค่อนข้างสูง โดยเฉพาะในขั้นตอนการฝึกและการใช้งานจริง จำเป็นต้องใช้หน่วยประมวลผลกราฟิก

		(GPU) เพื่อให้ได้ประสิทธิภาพสูงสุด
--	--	------------------------------------

ตารางที่ 4.9 การเปรียบเทียบคุณสมบัติและผลลัพธ์ของโมเดล CNN และ YOLO ในงานจำแนกประเภทขยะรีไซเคิล

จากการทดลองและเปรียบเทียบโมเดล CNN และ YOLO โดยอาศัยจากแง่มุมภายนอกด้วยพบร่วมกัน โมเดล CNN มีข้อดีคือไม่จำเป็นต้องมีการระบุพิกัดวัตถุในภาพ (Annotation) จึงช่วยลดภาระในการเตรียมข้อมูลได้มาก เนื่องจากสามารถพัฒนาในระดับเริ่มต้นของกล้อง อีกทั้งยังสามารถพัฒนาและปรับแต่งได้รวดเร็ว และใช้ทรัพยากร่นอยกว่ามากเมื่อเทียบกับ YOLO ซึ่งต้องอาศัยหน่วยประมวลผลグラฟิก (GPU) อย่างหนักเพื่อให้ทำงานได้เต็มประสิทธิภาพ

จากการเปรียบเทียบโดยรวม จึงสรุปได้ว่า แม้ YOLO จะมีข้อได้เปรียบในด้านความแม่นยำและการตรวจจับหลายวัตถุในภาพเดียวกัน แต่โมเดล CNN กลับมีความเหมาะสมในเรื่องของการเรียนรู้และการพัฒนาและปรับแต่ง สำหรับการนำเข้าใช้งานจริงในบริบทของโครงงานนี้ จึงเลือกใช้ โมเดล CNN เป็นแนวทางหลักในการพัฒนาโมเดลการคัดแยกขยะรีไซเคิลจากภาพด้วยการเรียนรู้ของเครื่อง

#### 4.4 การพัฒนาและประเมินผลโมเดล CNN สำหรับการจำแนกประเภทขยะรีไซเคิล

หลังจากที่ได้ตัดสินใจเลือกโมเดล CNN เป็นแนวทางหลักในการพัฒนาโมเดลสำหรับโครงงานนี้ จึงได้ดำเนินการรวบรวมและเพิ่มจำนวนข้อมูลภาพมากขึ้นจากเดิม โดยเฉพาะภาพที่มีลักษณะหลากหลายทั้งในด้านแสง มุมมอง และพื้นหลัง เพื่อเสริมสร้างความสามารถในการเรียนรู้และการจำแนกรูปแบบที่ซับซ้อนได้ถูกยิ่งขึ้น ทำให้โมเดลสามารถรับมือกับความหลากหลายของข้อมูลในสถานการณ์จริงได้อย่างมีประสิทธิภาพ

โดยที่จากก่อนหน้านี้มีการพัฒนาด้วย ข้อมูลคลาสละ 300 รูป รวม 1200 รูป ซึ่งยังไม่ครอบคลุมลักษณะข้อมูลที่หลากหลายนัก แต่หลังจากเลือกใช้โมเดล CNN และต้องการเพิ่มประสิทธิภาพให้มากยิ่งขึ้น จึงได้รวบรวมและขยายขนาดชุดข้อมูลเพิ่มเติมจนมีจำนวนรวมทั้งสิ้น 8,031 ภาพ เพิ่มขึ้นจากเดิม 6,831 ภาพ หรือคิดเป็นการเพิ่มมากกว่า 5 เท่า เพื่อช่วยให้โมเดลสามารถเรียนรู้จากข้อมูลที่ครอบคลุมสถานการณ์ต่าง ๆ ได้มากขึ้น และลดความเอนเอียงของการเรียนรู้ในแต่ละคลาส

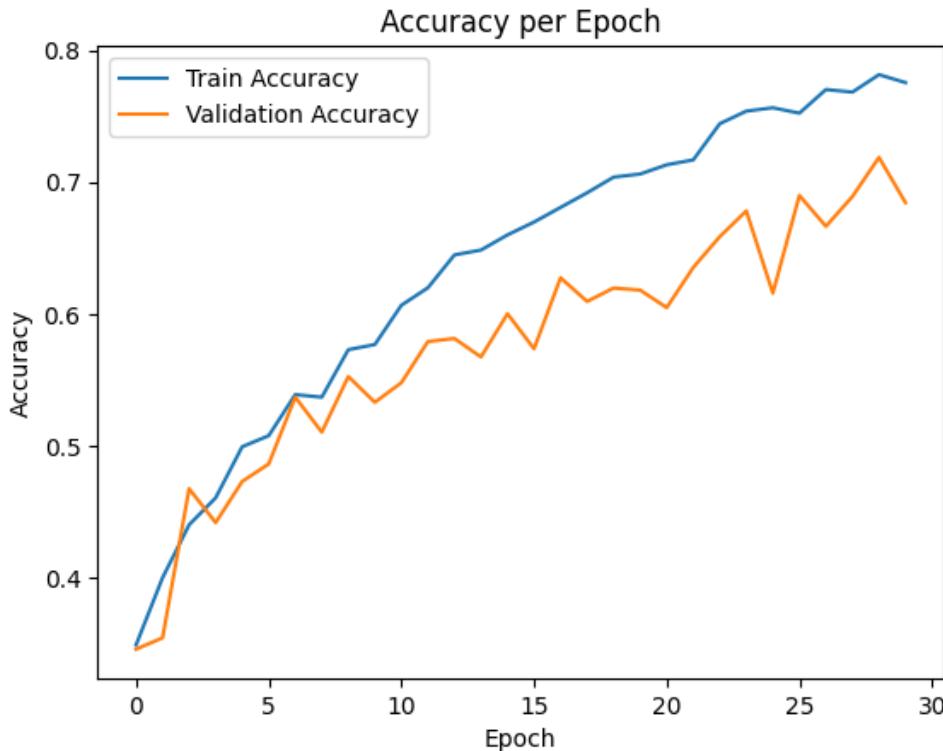
ประเภทขยะ	จำนวนภาพ (ก่อนเพิ่ม)	จำนวนภาพ (ปัจจุบัน)
Glass	300	1951
Metal	300	1824
Paper	300	2341
Plastic	300	1915
รวม	1200	8031

ตารางที่ 4.10 แสดงจำนวนภาพขยะก่อน-หลังการเพิ่ม

##### 4.4.1 ผลการฝึกโมเดล

ในการฝึกโมเดล CNN เพื่อจำแนกประเภทขยะรีไซเคิลในโครงงานนี้ เราใช้ Google Colab Pro ที่รองรับการใช้งาน GPU T4 และ High RAM เพื่อเพิ่มประสิทธิภาพในการประมวลผลและลดระยะเวลาในการฝึกโมเดล โดยโมเดลได้ทำการฝึกจำนวน 30 รอบ (epochs) ซึ่งใช้เวลาทั้งหมดประมาณ 35 นาที เพื่อฝึกข้อมูลทั้งหมด โดยถือว่าเป็นเวลาที่รวดเร็วเมื่อเทียบกับทรัพยากรที่จำกัดในสภาพแวดล้อมนี้

#### 4.4.1.1 ความแม่นยำ (Accuracy) ต่อ Epoch

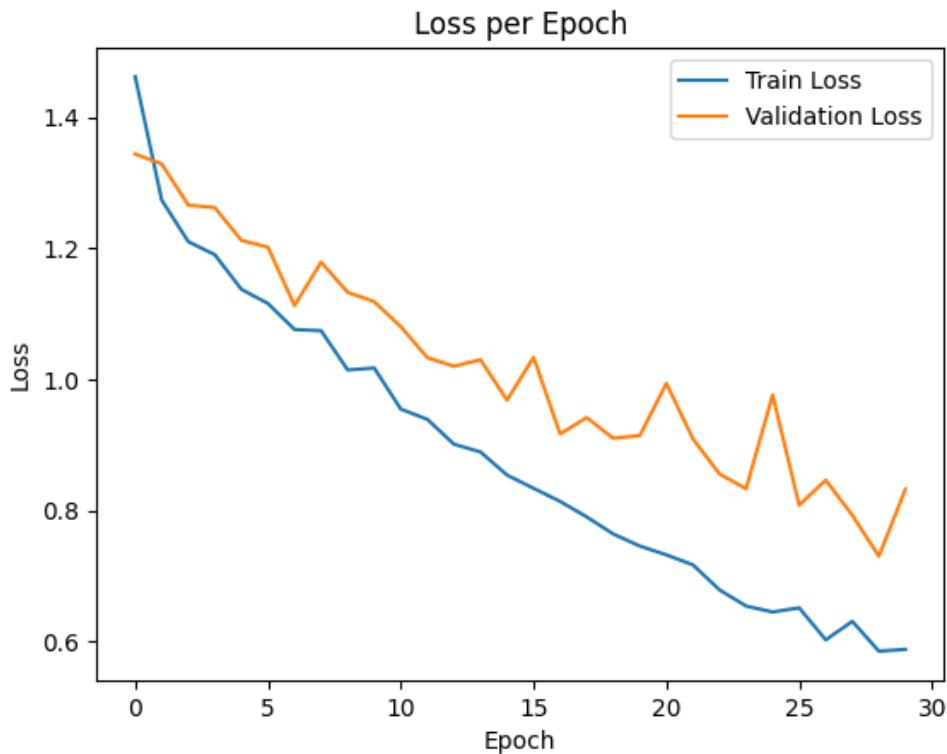


ภาพที่ 4.6 กราฟความแม่นยำ (Accuracy) ต่อ Epoch

การฝึกโมเดลทั้งหมดจำนวน 30 รอบ พบร่วมค่าความแม่นยำ ของชุดข้อมูลฝึกอบรม (train) และชุดข้อมูลตรวจสอบความถูกต้อง (validation) จากกราฟมีแนวโน้มเพิ่มขึ้นอย่างต่อเนื่อง โดยในช่วงเริ่มต้นโมเดลยังมีความแม่นยำที่ต่ำ โดยค่าความถูกต้องเริ่มต้นอยู่ที่ 31.40% สำหรับชุดฝึกอบรม และ 34.61% สำหรับชุดตรวจสอบความถูกต้อง ซึ่งแสดงให้เห็นว่าในช่วงเริ่มต้นโมเดลยังไม่สามารถเรียนรู้คุณลักษณะของภาพได้ดีนัก แต่เมื่อดำเนินฝึกต่อไปเรื่อยๆ ค่าความแม่นยำของทั้งสองชุดข้อมูลมีการเพิ่มขึ้นอย่างต่อเนื่อง โดยค่าความแม่นยำสูงสุดของชุดข้อมูลฝึกอบรมอยู่ที่ 78.18% และชุดข้อมูลตรวจสอบความถูกต้องอยู่ที่ 71.86 ในรอบการฝึก (epoch) ที่ 29 ซึ่งแสดงให้เห็นว่า โมเดลสามารถเรียนรู้ได้ดี และไม่พิตกับข้อมูลฝึกมากเกินไป แม้ว่าในรอบสุดท้ายของการฝึก ชุดข้อมูลตรวจสอบความถูกต้องจะลดลงเล็กน้อย แต่ยังถือว่าอยู่ในระดับที่ดี

จากการคำนวณ ค่าเฉลี่ยของ Train Accuracy ตลอดทั้ง 30 Epoch คือ 63.38% ซึ่งหมายความว่าโมเดลมีความแม่นยำเฉลี่ยในการจำแนกข้อมูลฝึกอบรมประมาณ 63.38% ในระหว่างการฝึก ทั้งนี้ ความแม่นยำนี้แสดงถึงประสิทธิภาพของโมเดลในการเรียนรู้จากชุดข้อมูลฝึกอบรมตลอดช่วงเวลาที่ฝึกมา โดยมีการปรับปรุงความแม่นยำขึ้นในแต่ละรอบ Epoch ตามลำดับ

#### 4.4.1.2 ค่าความสูญเสีย (Loss) ต่อ Epoch



ภาพที่ 4.7 กราฟค่าความสูญเสีย (Loss) ต่อ Epoch

ค่าความสูญเสีย (Loss) เป็นตัวชี้วัดว่าการพยากรณ์ของโมเดลใกล้เคียงกับค่าจริงมากน้อยเพียงใด ในช่วงต้นของการฝึก (Epoch 1) ค่าความสูญเสียของชุดข้อมูลฝึกสอนอยู่ที่ 1.9114 และชุดข้อมูลตรวจสอบอยู่ที่ 1.3438 ซึ่งบ่งชี้ว่าโมเดลยังมีความผิดพลาดสูงในการจำแนกภาพ และเมื่อดำเนินการฝึกไป จะเห็นได้ว่าค่าความสูญเสียลดลงอย่างต่อเนื่อง โดยตั้งแต่รอบการฝึกที่ 15 เป็นต้นไป ค่าความสูญเสียอยู่ในช่วงที่ต่ำกว่า 1.0 อย่างชัดเจน และค่าความสูญเสียต่ำที่สุดในช่วงรอบการฝึกที่ 29 การเปรียบเทียบค่าความสูญเสียของข้อมูลฝึกสอนและข้อมูลตรวจสอบพบว่าไม่มีช่องว่างห่างกันมากนัก แสดงให้เห็นว่าโมเดลไม่ฟิตกับข้อมูลฝึกสอนมากเกินไป และเรียนรู้ได้อย่างมีสติภาพ

จากการฝึกโมเดลจำนวน 30 รอบ พบร่วมกับความแม่นยำ (Accuracy) ของทั้งชุดข้อมูลฝึกอบรมและชุดตรวจสอบความถูกต้องเพิ่มขึ้นต่อเนื่อง ขณะที่ค่าความสูญเสีย (Loss) ลดลงอย่างมีนัยสำคัญ ทั้งสองค่านี้ความสอดคล้องกันและไม่พบช่องว่างที่ห่างกันมาก แสดงให้เห็นว่าโมเดลสามารถเรียนรู้ได้อย่างดี และยังไม่เกิดปัญหาการเรียนจำเพาะกับข้อมูลฝึกมากเกินไป (Overfitting)

#### 4.4.2 การประเมินประสิทธิภาพของโมเดลด้วยชุดข้อมูลทดสอบ (test set)

หลังจากที่โมเดลผ่านกระบวนการฝึกและตรวจสอบความถูกต้องเรียบร้อยแล้ว ได้มีการนำโมเดลไปประเมินผลกับชุดข้อมูลทดสอบ (Test set) ซึ่งเป็นชุดข้อมูลที่โมเดลไม่เคยเห็นมาก่อน เพื่อวัดความสามารถของโมเดลในการจำแนกประเภทของรีไซเคิลในสถานการณ์ที่ใกล้เคียงกับการใช้งานจริง โดยผลการประเมินจะพิจารณาจากค่า Accuracy โดยรวม ตลอดจนค่า Precision, Recall และ F1-score ในแต่ละประเภท รวมถึงการวิเคราะห์จาก Confusion Matrix เพื่อให้เข้าใจดูແร็ง จุดอ่อน และลักษณะการจำแนกของโมเดลในภาพรวม

##### 4.4.2.1 ค่าความแม่นยำ (Accuracy)

ผลการประเมินประสิทธิภาพของโมเดลด้วยชุดข้อมูลทดสอบ (Test Set) พบว่า โมเดลสามารถจำแนกประเภทของภาพขยะได้อย่างมีประสิทธิภาพ โดยมีค่า ความแม่นยำ (Test Accuracy) เท่ากับ 78% ซึ่งถือว่าเป็นผลลัพธ์ที่ดีเมื่อเทียบกับค่า ความแม่นยำของชุดข้อมูลฝึกสอน (Train Accuracy) ที่มีค่าเฉลี่ยอยู่ที่ประมาณ 63.38%

จากผลลัพธ์นี้แสดงให้เห็นว่าโมเดลสามารถเรียนรู้คุณลักษณะของข้อมูลได้ดี และยังสามารถ นำความรู้ที่ได้จากข้อมูลฝึกมาใช้กับข้อมูลใหม่ได้อย่างมีประสิทธิภาพ โดยไม่เกิดการเรียนรู้ที่มากเกินไปจากข้อมูลฝึกสอน ค่าความแม่นยำที่สูงของชุดข้อมูลทดสอบแสดงถึงความสามารถของ โมเดลในการทำงานได้ดีในสภาพแวดล้อมที่ไม่คุ้นเคยหรือกับข้อมูลที่ไม่เคยพบมาก่อน

##### 4.4.2.2 การวิเคราะห์ค่าความแม่นยำในแต่ละประเภท (Precision, Recall, F1-score)

จากการประเมินค่าความแม่นยำในแต่ละประเภท Precision, Recall และค่า F1-score ของโมเดลในชุดข้อมูลทดสอบ พบว่าโมเดลมีประสิทธิภาพในการจำแนกประเภทดังนี้

Class	precision	recall	f1-score	support
แก้ว	0.81	0.69	0.74	391
โลหะ	0.84	0.73	0.78	365
กระดาษ	0.77	0.91	0.83	469
พลาสติก	0.73	0.77	0.75	383

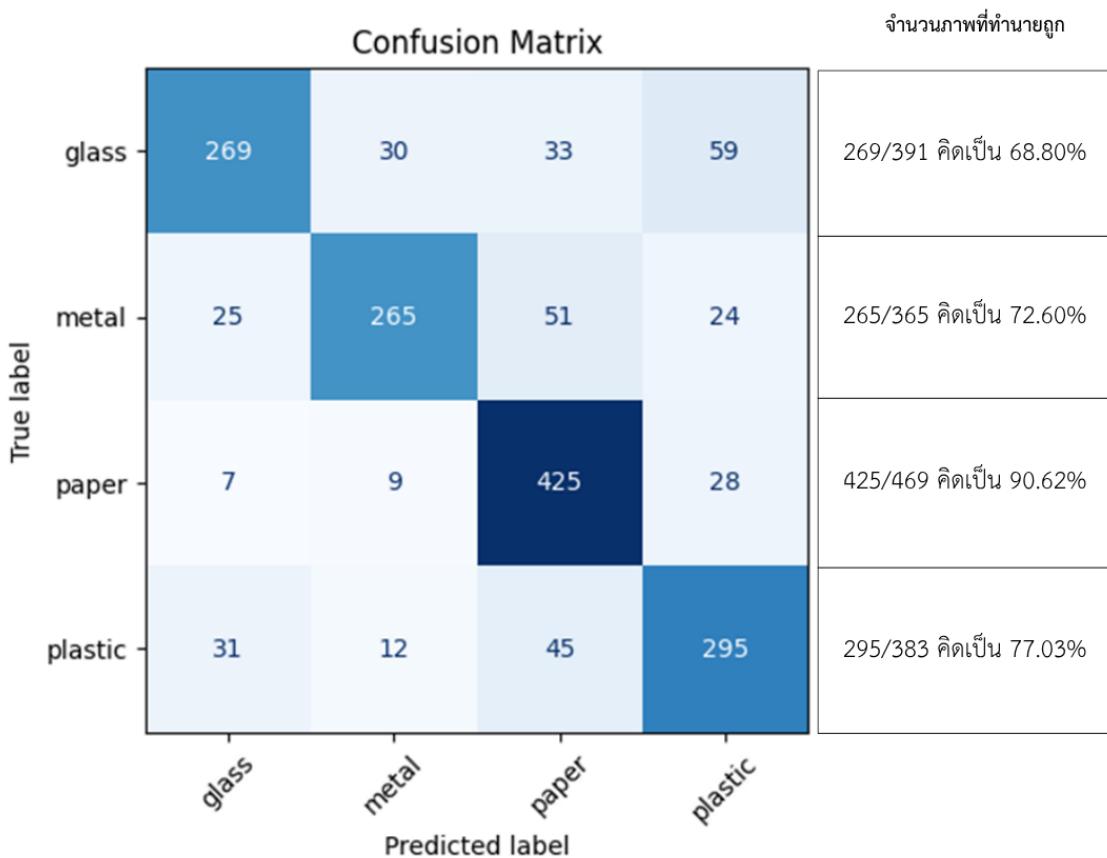
ตารางที่ 4.11 ค่าความแม่นยำในแต่ละประเภทหลังจากทดสอบโมเดล CNN ด้วยชุดข้อมูลใหญ่

จากตารางจะเห็นได้ว่า

- ประเกทกระดาษมีค่า Recall สูงที่สุดอยู่ที่ 0.91 และแสดงให้เห็นว่าโมเดลสามารถตรวจจับภาพขยะประเกทกระดาษได้ดี ในขณะที่ค่า F1-score อยู่ที่ 0.83 ซึ่งถือว่าเป็นค่าที่สูงที่สุดกว่าทุกประเกท และแสดงให้เห็นว่าโมเดลมีความสามารถที่ดีในการจำแนกประเกทกระดาษ
- ประเกทโลหะ มีค่า Precision สูงสุดในกลุ่มทั้งหมด 0.84 และแสดงว่าเมื่อโมเดลทำนายว่าเป็นโลหะ มีโอกาสสูงที่จะหายถูกจริง และมีค่า Recall อยู่ที่ 0.73 และแสดงว่ายังมีบางตัวอย่างที่โมเดลไม่สามารถตรวจจับได้
- ประเกทแก้ว มีค่า Recall ต่ำสุด 0.69 บ่งชี้ว่าโมเดลยังตรวจจับได้ไม่ครบถ้วน แต่ Precision อยู่ที่ 0.81 ซึ่งถือว่าสูง หมายความว่าเมื่อโมเดลบอกว่าเป็น แก้ว มักจะถูกต้อง อาจต้องเพิ่มข้อมูลฝึกสอนในประเกทนี้
- ประเกทพลาสติก ทั้ง Precision 0.73 และ Recall 0.77 อยู่ในระดับปานกลาง ส่งผลให้ค่า F1-score อยู่ที่ 0.75 สะท้อนว่าโมเดลมีความแม่นยำพอสมควรในการจำแนกพลาสติก แต่ยังมีโอกาสในการพัฒนาและปรับปรุงให้ดีขึ้นได้อีก

#### 4.4.2.3 การวิเคราะห์ตารางความสับสน (Confusion Matrix)

ตารางความสับสน (Confusion Matrix) เป็นเครื่องมือที่ช่วยแสดงให้เห็นถึงรายละเอียดของผลการจำแนกประเกทของโมเดลในแต่ละคลาส โดยแสดงจำนวนตัวอย่างที่โมเดลทำนายได้ถูกต้อง และจำนวนที่ทำนายผิดไปเป็นคลาสอื่น ๆ ซึ่งช่วยให้สามารถวิเคราะห์จุดแข็งและจุดที่ควรปรับปรุงของโมเดลได้ชัดเจนมากยิ่งขึ้น



ภาพที่ 4.8 ตาราง Confusion Matrix ของโมเดล CNN ด้วยชุดข้อมูลใหญ่

### 1. ประเภทที่โมเดลสามารถจำแนกได้อย่างถูกต้องมากที่สุด

จากตารางความสัมสุน พบร้า ขยายประเภทกระดาษ เป็นประเภทที่โมเดลสามารถจำแนกได้อย่างถูกต้องมากที่สุด โดยจำแนกได้ถูกต้องจำนวน 425 ภาพ จากทั้งหมด 469 ภาพ คิดเป็น ความแม่นยำที่ 90.62% ซึ่งถือเป็นผลลัพธ์ที่โดดเด่นเมื่อเปรียบเทียบกับคลาสอื่น

สาเหตุที่ทำให้โมเดลสามารถจำแนกประเภทกระดาษได้แม่นยามาก อาจมาจากการปัจจัยหลายประการ ดังนี้

- ลักษณะพื้นผิวของกระดาษ มีความเฉพาะ เช่น ผิวเรียบ รอยยับ หรือร่องรอย การใช้งานที่แตกต่างจากวัสดุอื่นอย่างชัดเจน ทำให้โมเดลสามารถเรียนรู้และจดจำได้ง่าย

- สีของกระดาษ ส่วนใหญ่นักเป็นสีขาว สีเทา หรือโทนอ่อน ทำให้ไม่เดลແຍກແຍະจากขยะที่มีโทนสีเข้ม เช่น โลหะหรือพลาสติก ได้ดี
- ลักษณะรูปทรง กระดาษมักมีรูปทรงแบบ ราบ หรือมีขอบที่ขัดเจน เช่น เอกสารใบปลิว กล่องกระดาษ ลักษณะนี้ช่วยให้ไม่เดลสามารถใช้ขับเขตของวัตถุใน การจำแนกได้อย่างมีประสิทธิภาพ

นอกจากนี้ ยังพบว่าในคลาสกระดาษมี จำนวนการทำนายผิดน้อยมาก เมื่อเทียบกับ คลาสอื่น เช่น ทำนายผิดเป็นพลาสติกเพียง 28 ครั้ง และผิดเป็นแก้วหรือโลหะเพียงเล็กน้อย สะท้อนว่าไม่เดลสามารถแยกลักษณะของกระดาษจากวัสดุใกล้เคียงได้ดีพอสมควร

## 2. คลาสที่มีความสับสนมากที่สุด

พบว่าเมื่อไม่เดลจะสามารถจำแนกภาพขยะได้อย่างมีประสิทธิภาพโดยรวม แต่ยังมี บางคลาสที่เกิดการทำนายผิดบ่อย ซึ่งสะท้อนถึงความสับสนของไม่เดลในบางลักษณะของ ข้อมูล โดยสามารถสรุปค่าสูงสุดของการทำนายผิดแต่ละประเภทได้ ดังนี้

- แก้วถูกทำนายเป็นพลาสติก จำนวน 59 ครั้ง

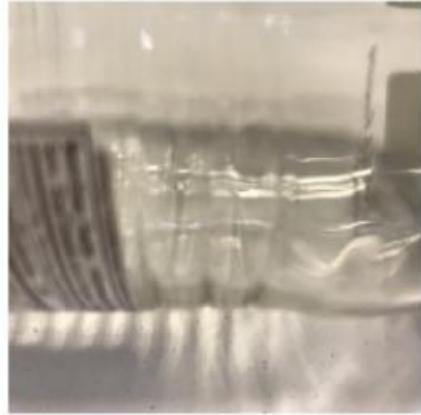
กรณีนี้เป็นการสับสนที่มากที่สุดในตาราง สะท้อนว่าไม่เดลมีความเข้าใจ คลาดเคลื่อนระหว่างแก้วกับพลาสติก ซึ่งอาจเกิดจากลักษณะพื้นผิวที่ ใส หรือ โปร่ง แสง คล้ายกัน เช่น ขวดน้ำพลาสติกใสและขวดแก้วบางประเภท อาจมีลักษณะทาง แสงที่ใกล้เคียงกันเมื่อถ่ายภาพในบางมุม

True: glass  
Pred: plastic



ภาพที่ 4.9 ตัวอย่างภาพที่แก้วถูกทำนายเป็นพลาสติก (1)

True: glass  
Pred: plastic



ภาพที่ 4.10 ตัวอย่างภาพที่แก้วุกทำนายเป็นพลาสติก (2)

- **โลหะถูกทำนายเป็นกระดาษ จำนวน 51 ครั้ง**

การทำนายผิดในลักษณะนี้อาจเกิดจากโลหะบางประเภท เช่น พอยล์อลูมิเนียม หรือกระปองที่ถูกบีบอัด มีลักษณะเรียบและบาง คล้ายกระดาษในเชิงรูปร่าง และในบางกรณีอาจมีสีเทาเงินที่ใกล้เคียงกับสีของกระดาษ

True: metal  
Pred: paper



ภาพที่ 4.11 ตัวอย่างภาพที่โลหะถูกทำนายเป็นกระดาษ (1)

True: metal  
Pred: paper



ภาพที่ 4.12 ตัวอย่างภาพที่โลหะถูกทำนายเป็นกระดาษ (2)

- พลาสติกถูกทำนายเป็นกระดาษ จำนวน 45 ครั้ง

การที่ไม่เดลทำนายผิดจากพลาสติกเป็นกระดาษในหลายกรณี อาจเกิดจากลักษณะภายนอกของวัสดุทั้งสองประเภทที่คล้ายคลึงกัน พลาสติกบางชิ้น เช่น ถุงพลาสติก แก้วพลาสติก หรือบรรจุภัณฑ์บางชนิด มีลักษณะเรียบ แบบ หรือป่อง แสง และมักเป็นสีขาวหรือโทนสีอ่อน ซึ่งทำให้คล้ายกับกระดาษทั่วไป

ในทางกลับกัน กระดาษบางประเภท เช่น ของจดหมาย กล่องกระดาษ หรือใบราชโองการ มีลักษณะที่คล้ายคลึงกับวัสดุพลาสติกที่มีการพิมพ์สีหรือมีฉลาก เช่น ถุงข้อปั๊วพลาสติกที่มีข้อความ ทำให้ไม่เดลสับสนได้

True: plastic  
Pred: paper



ภาพที่ 4.13 ตัวอย่างภาพที่พลาสติกถูกทำนายเป็นกระดาษ (1)

True: plastic  
Pred: paper



ภาพที่ 4.14 ตัวอย่างภาพที่พลาสติกถูกทำนายเป็นกระดาษ (2)

- กระดาษถูกทำนายเป็นพลาสติก จำนวน 28 ครั้ง

การที่กระดาษถูกทำนายเป็นพลาสติกจำนวน 28 ครั้ง แสดงให้เห็นว่าโมเดล CNN มีความสับสนระหว่างลักษณะของวัสดุทั้งสองประเภทในบางกรณี ซึ่งอาจเกิดจากลักษณะภายนอกของกระดาษบางประเภทที่คล้ายกับพลาสติก เช่น กระดาษที่มีการเคลือบผิวน้ำหรือกันน้ำ รวมถึงกล่องกระดาษหรือถุงกระดาษที่มีพิพิว言行 อาจสะท้อนแสงคล้ายกับพลาสติกได้ นอกจากนี้กระดาษบางชิ้นที่มีลวดลาย สีสัน หรือการพิมพ์ข้อความ อาจมีลักษณะใกล้เคียงกับถุงหรือบรรจุภัณฑ์พลาสติกที่มีลักษณะหรือภาพซึ่งกัน ในบางกรณีหากกระดาษอยู่ในสภาพยับยุ่ย หรือถูกบีบอัด อาจทำให้เสียรูปร่างจนคล้ายกับแผ่นพลาสติกหรือถุงขยะ จึงทำให้โมเดลเกิดความสับสนในการจำแนก ทั้งนี้ยังสะท้อนถึงความสำคัญของการเพิ่มความหลากหลายของชุดข้อมูลเพื่อช่วยให้โมเดลแยกแยะวัสดุที่มีลักษณะใกล้เคียงกันได้แม่นยำยิ่งขึ้น

True: paper  
Pred: plastic



ภาพที่ 4.15 ตัวอย่างภาพที่กระดาษถูกทำนายเป็นพลาสติก (1)

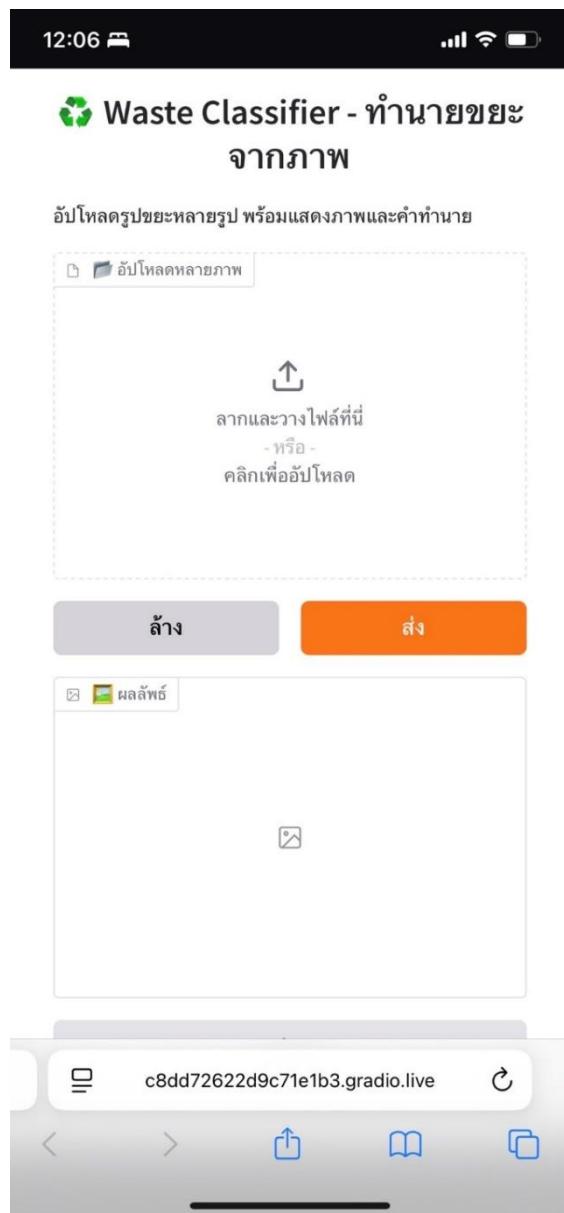
True: paper  
Pred: plastic



ภาพที่ 4.16 ตัวอย่างภาพที่กระดาษถูกทำนายเป็นพลาสติก (2)

#### 4.4.3 การทดสอบโมเดลกับข้อมูลภายนอกผ่าน Gradio Interface

การทดสอบโมเดลด้วยข้อมูลภายนอกเป็นขั้นตอนสำคัญในการประเมินความสามารถของโมเดลในสถานการณ์จริง โดยการใช้ Gradio Interface เป็นเครื่องมือที่ช่วยให้สามารถทดสอบและแสดงผลลัพธ์ได้อย่างสะดวกและรวดเร็ว



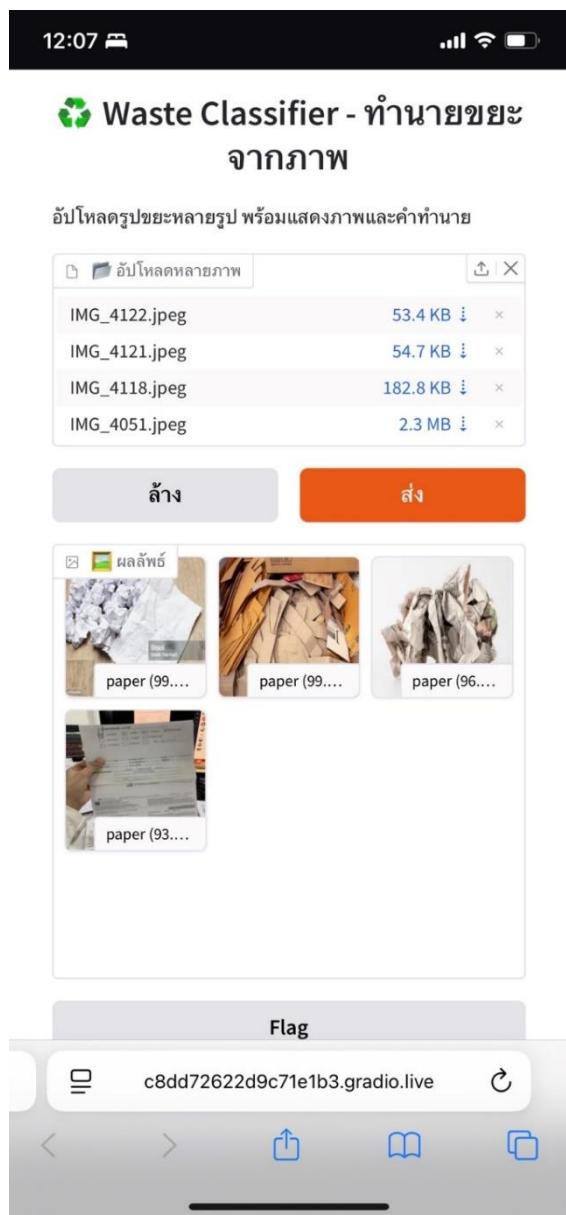
ภาพที่ 4.17 ตัวอย่าง Gradio Interface ที่ใช้ทดสอบโมเดลกับภาพขยะภายนอก

#### 4.4.3.1 ตัวอย่างภาพที่ทำนายถูกและผิด

จากการทดลอง พบร่วมโมเดลสามารถจำแนกประเภทของขยะได้ถูกต้องในหลายกรณี โดยเฉพาะประเภทที่มีลักษณะเฉพาะชัดเจน อย่างไรก็ตาม โมเดลยังมีความคลาดเคลื่อนในการทำนายในบางคลาส โดยเฉพาะเมื่อวัตถุมีลักษณะใกล้เคียงกัน เช่น สีรุ้งร่ารัง หรือพื้นผิว

- ตัวอย่างที่ทำนายถูกต้อง

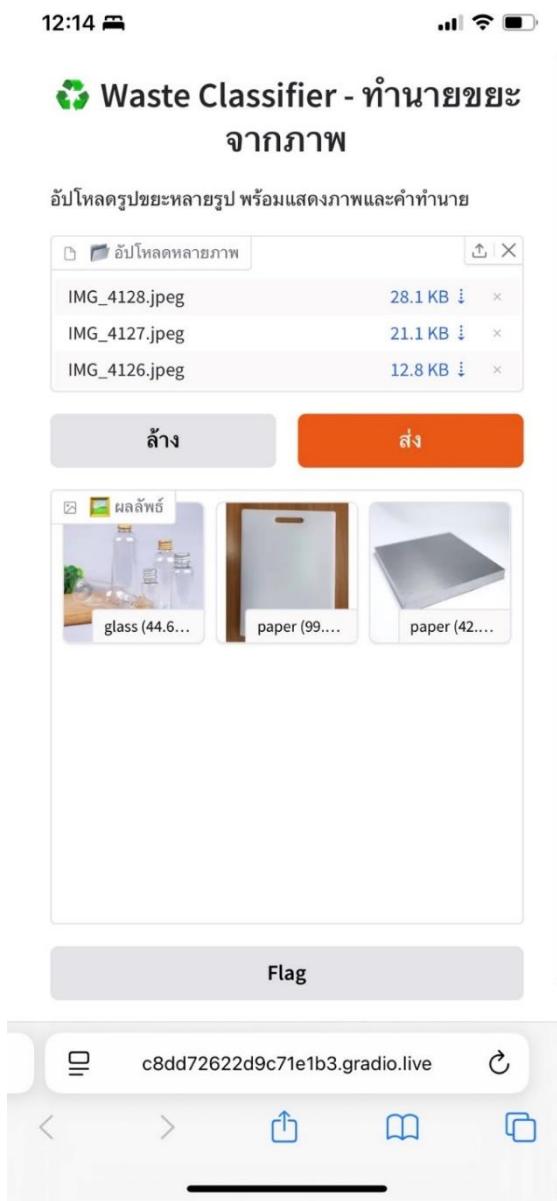
ไม่เดลสามารถจำแนกประเภทของขยะได้ถูกต้องในกรณีที่วัตถุมีลักษณะเด่นชัด เช่น ภาพของกล่องกระดาษที่มีรอยพับ และสีขาวอ่อน ซึ่งเป็นลักษณะเฉพาะของกระดาษ ส่งผลให้ไม่เดลสามารถวิเคราะห์และจำแนกได้อย่างแม่นยำ



ภาพที่ 4.18 ตัวอย่างภาพที่ไม่เดลสามารถทำนายได้ถูกต้อง

- ตัวอย่างที่ทำนายผิด

แม้โมเดลจะสามารถจำแนกภาพได้อย่างถูกต้องในหลายกรณี แต่ยังคงมีบางกรณีที่โมเดลทำนายผิด โดยเฉพาะในคลาสที่มีลักษณะคล้ายกัน เช่น พลาสติกใส่ที่ถูกทำนายว่าเป็นแก้ว หรือฟอยล์โลหะที่ถูกทำนายว่าเป็นกระดาษ



ภาพที่ 4.19 ตัวอย่างภาพที่ไม่เดลทำนายผิด

กรณีนี้จะท่อนให้เห็นถึงข้อจำกัดของโมเดลในการแยกแยะวัตถุที่มีพื้นผิวคล้ายกัน โดยเฉพาะในภาพที่มีแสงมากหรือพื้นหลังซับซ้อน ภาพของภานุภาพสติกที่มีพื้นผิวเรียบ และเป็นสีขาว ถูกโมเดลทำนายผิดว่าเป็นกระดาษ เนื่องจากลักษณะทางกายภาพที่คล้ายกัน เช่น ความบาง รูปร่างแบบ และโทนสีอ่อน ข้อผิดพลาดเหล่านี้เน้นย้ำถึงความจำเป็นในการเพิ่มความหลากหลายของชุดข้อมูลฝึกสอน โดยเฉพาะตัวอย่างที่มีความใกล้เคียงกันข้ามคลาส เพื่อเพิ่มความสามารถในการแยกแยะของโมเดล

#### 4.4.3.2 การวิเคราะห์ความสามารถของโมเดลในสถานการณ์จริง

จากการทดสอบการทำงานของโมเดลผ่าน Gradio Interface ซึ่งใช้ภาพจากผู้ใช้งานจริงหรือภาพภายนอกที่ไม่ได้อยู่ในชุดข้อมูลฝึกและทดสอบ พบร่วมกันความสามารถ จำแนกประเภทของขยะได้อย่างรวดเร็วและแม่นยำในหลายกรณี โดยเฉพาะในภาพที่มีลักษณะซัดเจนซึ่งแสดงถึงศักยภาพของโมเดลในการนำไปประยุกต์ใช้

อย่างไรก็ตาม ยังพบข้อจำกัดบางประการ โดยโมเดลมีแนวโน้มที่จะทำนายผิดในภาพที่มีองค์ประกอบอื่นแทรกซ้อนกับวัตถุหลัก เช่น การทำนายผิดระหว่างพลาสติกใสกับแก้ว หรือโลหะบางกับกระดาษ ซึ่งสะท้อนว่าโมเดลยังมีความอ่อนไหวต่อความซับซ้อนของภาพในโลกจริง

ดังนั้น เพื่อเพิ่มประสิทธิภาพของโมเดลในสถานการณ์จริง ควรมีการขยายและปรับปรุงชุดข้อมูลให้มีความหลากหลายมากยิ่งขึ้น เช่น การรวบรวมภาพจากสภาพแวดล้อมต่างๆ และการจัดองค์ประกอบของภาพที่หลากหลาย รวมถึงอาจพิจารณาการใช้เทคนิค preprocessing หรือ data augmentation เพิ่มเติม เพื่อช่วยให้โมเดลสามารถเรียนรู้ลักษณะเฉพาะที่แตกต่างกันได้ดียิ่งขึ้น และรองรับการใช้งานในสภาพแวดล้อมจริงได้อย่างมีเสถียรภาพมากขึ้น

## บทที่ 5

### สรุป

#### 5.1 สรุปผลการดำเนินงาน

โครงการนี้มีเป้าหมายเพื่อพัฒนาและปรับเปลี่ยนโมเดลจำแนกประเภทของรีไซเคิลจากภาพถ่ายโดยใช้เทคนิคการเรียนรู้ของเครื่อง (Machine Learning) และการเรียนรู้เชิงลึก (Deep Learning) เพื่อช่วยให้การคัดแยกขยะเป็นไปอย่างถูกต้องและมีประสิทธิภาพมากขึ้น โดยมุ่งเน้นที่ขยะรีไซเคิล 4 ประเภท ได้แก่ พลาสติก กระดาษ โลหะ และแก้ว

เพื่อปรับเปลี่ยนความสามารถของโมเดลต่าง ๆ ได้อย่างเป็นธรรม จึงได้เลือกโมเดลตัวแทนสองประเภท ได้แก่ YOLO (You Only Look Once) และ CNN (Convolutional Neural Network) ซึ่งเป็นโครงสร้างพื้นฐานของการจำแนกภาพในหลายงานวิจัย โดยใช้ชุดข้อมูล 1,200 ภาพ เท่ากันในการฝึกหัดสองโมเดล พบว่า YOLO ให้ค่า F1-Score เท่ากับ 87 ในขณะที่ CNN ให้ค่า F1-Score เท่ากับ 60 ซึ่งบ่งชี้ว่า YOLO มีความแม่นยำสูงกว่าในการทดลองเบื้องต้น

อย่างไรก็ตาม เมื่อพิจารณาด้านการนำไปใช้งานจริงในบริบทของโครงการ โดยเฉพาะในเรื่องของทรัพยากรที่ใช้ ความซับซ้อนในการเตรียมข้อมูล และความง่ายในการพัฒนา พบว่าโมเดล CNN มีความเหมาะสมกว่า เนื่องจากไม่ต้องเตรียม annotation หรือ bounding box เมื่อนอกจาก YOLO อีกทั้งยังใช้เวลาในการฝึกน้อยกว่า ใช้ทรัพยากรคอมพิวเตอร์น้อย และสามารถนำไปใช้พัฒนาแอปพลิเคชันต้นแบบได้จ่ายกว่า จึงได้มีการตัดสินใจเลือก CNN เป็นแนวทางหลักสำหรับการพัฒนาโมเดลในโครงการนี้

หลังจากเลือก CNN แล้ว ได้มีการปรับปรุงชุดข้อมูลใหม่โดยการรวมภาพเพิ่มเติมจนได้ภาพรวมทั้งหมดจำนวน 8,031 ภาพ โดยรักษาสัดส่วนของแต่ละคลาสให้สมดุล เพื่อให้โมเดลสามารถเรียนรู้ลักษณะเฉพาะของแต่ละประเภทขยะได้ดีขึ้น โดยเมื่อทำการฝึกโมเดล CNN ด้วยข้อมูลที่เพิ่มขึ้น พบว่าโมเดลสามารถเรียนรู้ได้อย่างมีประสิทธิภาพ โดยมีค่า Accuracy บนชุดทดสอบเท่ากับ 78% ซึ่งถือว่าพึงพอใจ และจากการวิเคราะห์ตารางสับสน (Confusion Matrix) พบว่าโมเดลมีความแม่นยำในภาพรวม แต่ยังมีความสับสนระหว่างบางคลาส เช่น แก้วมักถูกทำนายเป็นพลาสติก (59 ครั้ง) เนื่องจากความโปร่งแสงและพื้นผิวใสของวัสดุทั้งสองคล้ายกัน หรือโลหะที่ถูกทำนายผิดเป็นกระดาษ (51 ครั้ง) อาจมาจากรูปร่างที่เรียบและสีที่คล้ายกัน เช่น ฟอยล์อลูมิเนียม เป็นต้น

แม้ว่าโมเดล CNN จะยังมีข้อจำกัดในด้านความละเอียดในการจำแนกภาพที่มีลักษณะคล้ายกันสูง แต่จากการทดลองในภาพรวม แสดงให้เห็นว่าโมเดลมีความสามารถในการเรียนรู้ที่ดี เมื่อได้รับข้อมูลหลากหลายเพียงพอ อีกทั้งสามารถนำไปใช้งานจริงได้โดยไม่ต้องพึ่งพาทรัพยากรชั้นสูง

จึงเป็นทางเลือกที่ดีและมีความสมดุลระหว่างประสิทธิภาพ ความง่ายในการใช้งาน และความสามารถในการพัฒนาและปรับปรุงต่อยอดได้อย่างยั่งยืนในอนาคต

## 5.2 ข้อเสนอแนะ

จากการดำเนินโครงการในครั้งนี้ พบว่าโมเดล CNN มีความเหมาะสมในบริบทของการพัฒนาโมเดลจำแนกชั้นรีไซเคิลด้วยทรัพยากรที่จำกัด ทั้งในแง่ของการเตรียมข้อมูล ความเรียบง่ายในการฝึก และความสะดวกในการนำไปใช้งานจริง อย่างไรก็ตาม ยังมีข้อเสนอแนะเพื่อปรับปรุงและต่อยอดงานวิจัยในอนาคตดังนี้

### 1. เพิ่มปริมาณข้อมูลฝึกให้มากขึ้น

แม้การฝึกโมเดล CNN ด้วยภาพ 8,031 ภาพจะให้ผลลัพธ์ที่น่าพอใจ แต่การเพิ่มจำนวนภาพในอนาคต โดยเฉพาะภาพที่หลากหลายด้านมุมกล้อง แสง และพื้นหลัง จะช่วยเพิ่มความสามารถในการจำแนกของโมเดลในสถานการณ์จริงได้ดียิ่งขึ้น

### 2. เพิ่มเทคนิคการปรับแต่งข้อมูล (Data Augmentation)

เพื่อให้โมเดลสามารถเรียนรู้จากภาพที่มีความหลากหลายยิ่งขึ้น ควรพิจารณาใช้เทคนิคการปรับแต่งภาพ เช่น การหมุน การกลับด้าน การปรับแสงเงา หรือการซูมภาพ ซึ่งสามารถช่วยลดปัญหา overfitting และเพิ่มความยืดหยุ่นให้กับโมเดล

### 3. พิจารณาใช้งาน YOLO ในกรณีที่มีทรัพยากรเพียงพอ

แม้ในโครงการนี้จะเลือก CNN เป็นโมเดลหลัก แต่ YOLOv8 ยังมีจุดเด่นที่สามารถตรวจจับวัตถุหลายรายการในภาพเดียวได้พร้อมกัน ซึ่งเหมาะสมในกรณีที่หนึ่งภาพมีขยะหลายชิ้น หรือเมื่อนำมาใช้งานในสายพานอุตสาหกรรมจริง หากมีเวลาและทรัพยากรเพียงพอ ควรศึกษาและพัฒนาโมเดล YOLO ต่อเพิ่มเติม

### 4. ศึกษาโมเดลที่มีความละเอียดสูงกว่า เช่น Mask R-CNN

สำหรับโครงการในอนาคต หากต้องการแยกชิ้นขยะอย่างละเอียด หรือวิเคราะห์รูปร่างของขยะเป็นพิเศษ โมเดลแบบ Mask R-CNN ที่สามารถจำแนกพร้อมกับทำ segmentation ได้ อาจเป็นอีกทางเลือกที่เหมาะสม โดยเฉพาะในระบบที่ต้องการระบุตำแหน่งและขอบเขตของวัตถุอย่างชัดเจน

## รายการอ้างอิง

AWS. (ม.ป.ป.). *Data Augmentation* (การเสริมข้อมูล) คืออะไร. เข้าถึงได้จาก  
<https://aws.amazon.com/th/what-is/data-augmentation/>

aws. (ม.ป.ป.). อะไรคือความแตกต่างระหว่างแมชีนเลร์นนิ่งและดีปเลร์นนิ่ง. เข้าถึงได้จาก ม.ป.ป.:  
<https://aws.amazon.com/th/compare/the-difference-between-machine-learning-and-deep-learning/>

CYBER ELITE. (2563). *Machine Learning* เทคโนโลยีประยุกต์ครอบจักรวาล. เข้าถึงได้จาก  
<https://www.cyberelite.co.th/blog/machine-learning/>

Dailytech. (2563). *Image Processing* คือ ความฉลาดของเทคโนโลยีที่ทำให้ชีวิตเราง่ายขึ้น. เข้าถึง  
 ได้จาก <https://www.dailytech.in.th/image-processing/>

Mindphp Content Team 30. (2562). *OpenCV* คืออะไร ? เข้าถึงได้จาก  
<https://www.mindphp.com/%E0%B8%84%E0%B8%B9%E0%B9%88%E0%B8%A1%E0%B8%B7%E0%B8%AD/73-%E0%B8%84%E0%B8%B7%E0%B8%AD%E0%B8%AD%E0%B8%B0%E0%B9%84%E0%B8%A3/7061-what-is-opencv.html>

Nattakan Chuntra. (2561). *OpenCV* คืออะไร? เข้าถึงได้จาก  
<https://medium.com/@nnatchunn/opencv-%E0%B8%84%E0%B8%B7%E0%B8%AD%E0%B8%AD%E0%B8%B0%E0%B9%84%E0%B8%A3-8771e2a4c414>

ptg. (2559). การประมวลภาพ. เข้าถึงได้จาก  
[https://parntookgame.blogspot.com/2016/10/blog-post\\_62.html](https://parntookgame.blogspot.com/2016/10/blog-post_62.html)

กิจกรรมธนาคารฯ. (ม.ป.ป.). โครงการล่งเสริมพุทธิกรรมการคัดแยกขยะและรักษ์ชาลีน'แวดล้อม.  
 เข้าถึงได้จาก  
<https://www.maetha.ac.th/web/%E0%B9%80%E0%B8%A7%E0%B9%87%E0%B8%9A%20%E0%B9%82%E0%B8%84%E0%B8%A3%E0%B8%87%E0%B8%81%E0%B8%B2%E0%B8%A3%E0%B9%81%E0%B8%A2%E0%B8%81%E0%B8%82%E0%B8%A2%E0%B8%B0/page9.html>

ข้อมูลสถานการณ์ขยะมูลฝอยของประเทศไทย. (2566). เข้าถึงได้จาก กรมควบคุมมลพิษ:

<https://www.pcd.go.th/garbage/>

นางสาวศุภรัตน์ ฐิติรัตนอัศว์. (ม.ป.บ.). นักศึกษาฝึกงาน มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าราชนครเนื้อ. เข้าถึงได้จาก

<https://epo06.pcd.go.th/th/view/?file=pTMgMUqjGP5gATp3qQlcBKtlpQqgAKp1GQOgZzp5qQlcZKtgpQWgAUp4GQSgZJp3qQlcAKtjpQWgBKp1GQSgAzpjQAcZKtjpQWgZUplGP1gMJqfqTycMatipTlgoUqcGTMgY2q1qTycBUT1pQSgY3qxGTSgo2qfqUOcqKti&n=re&t=GTMgq2qxqS9cMUug>

ภัทรมน หุ่นลำพูน. (2566). การพัฒนาและประยุกต์ใช้เทคนิค Convolutional Neural Network ในการจำแนกขยะชายฝั่งทะเล. เข้าถึงได้จาก

[https://digital\\_collect.lib.buu.ac.th/dcms/files/62910162.pdf](https://digital_collect.lib.buu.ac.th/dcms/files/62910162.pdf)

ระบบสารสนเทศด้านการจัดการขยะมูลฝอยชุมชน. (ม.ป.บ.). ข้อมูลสถานการณ์ขยะมูลฝอยของประเทศไทย. เข้าถึงได้จาก กรมควบคุมมลพิษ:

[https://thaimsw.pcd.go.th/report\\_country.php](https://thaimsw.pcd.go.th/report_country.php)

ศุภณัฐ จินตวัฒน์สกุล. (2562). การจำแนกภาพหมวดแบบเบ็ดเตล็ดด้วยโครงข่ายประสาทเทียมแบบconvolutional. เข้าถึงได้จาก <http://irithesis.swu.ac.th/dspace/bitstream/123456789/488/1/gs601130181.pdf>

ศุภรัตน์ ฐิติรัตนอัศว์ และ สิรภัทร เทียนสันเทียะ. (ม.ป.บ.). ขยายรีไซเคิล. เข้าถึงได้จาก

<https://epo06.pcd.go.th/th/view/?file=pTMgMUqjGP5gATp3qQlcBKtlpQqgAKp1GQOgZzp5qQlcZKtgpQWgAUp4GQSgZJp3qQlcAKtjpQWgBKp1GQSgAzpjQAcZKtjpQWgZUplGP1gMJqfqTycMatipTlgoUqcGTMgY2q1qTycBUT1pQSgY3qxGTSgo2qfqUOcqKti&n=re&t=GTMgq2qxqS9cMUug>

Pattanan Tingpattana. (2019). คัมภีร์ YOLO LPR Object Detection (สอน Custom Model ,

ใช้ dataset ของตัวเองมาแทน ชิวๆ part1 เตรียม data). เข้าถึงได้จาก

<https://medium.com/@chamkung1412/%E0%B8%84%E0%B8%B1%E0%B8%A1%E0%B8%A0%E0%B8%B5%E0%B8%A3%E0%B9%8C-yolo-v2-v3-object-detection-%E0%B8%89%E0%B8%9A%E0%B8%B1%E0%B8%9A%E0%B8%A1%E0%B8%AD%E0%B9%83%E0%B8%AB%E0%B8%A1%E0%B9%88->

%E0%B8%AA%E0%B8%AD%E0%B8%99-custom-model-  
%E0%B9%83%E0%B8%8A%E0%B9%89-dataset-  
%E0%B8%82%E0%B8%AD%E0%B8%87%E0%B8%95%E0%B8%B1%E0%B8%A  
7%E0%B9%80%E0%B8%AD%E0%B8%87%E0%B8%A1%E0%B8%B2%E0%B9%  
80%E0%B8%97%E0%B8%A3%E0%B8%99-b69cb050cc4c.

Intouch Kunakorntum.(2017). Image detection โดยใช้ YOLOv5 จากต้นจนจบ (ตอน 1) .

เข้าถึงได้จาก https://medium.com/@intouchkunakorntum/image-detection-%E0%B9%82%E0%B8%94%E0%B8%A2%E0%B9%83%E0%B8%8A%E0%B9%89-yolov5-%E0%B8%88%E0%B8%B2%E0%B8%81%E0%B8%95%E0%B9%89%E0%B8%99-%E0%B8%88%E0%B8%99%E0%B8%88%E0%B8%9A-%E0%B8%95%E0%B8%AD%E0%B8%99-1-77faf04b0579.

Mindphp. (มป.). Roboflow ตัวช่วยสำคัญในการสร้างและปรับแต่ง Dataset. สืบค้น

จาก https://www.mindphp.com/บทเรียนออนไลน์/python-tensorflow/10210-roboflow-ตัวช่วยสำคัญในการสร้างและปรับแต่ง-dataset.html

ณัฐกิตติ์. (มป.). สร้าง AI ทำนายรูปภาพด้วย Roboflow. สืบค้นจาก

https://nattakit-nice2580.medium.com/สร้าง-ai-ทำนายรูปภาพด้วย-roboflow-6bb96a632f4d