

การจำแนกสายพันธุ์ข้าวไทยโดยใช้เทคนิคการประมวลผลภาพและการเรียนรู้
เชิงลึก

CLASSIFICATION OF THAI RICE VARIETIES USING IMAGE
PROCESSING AND DEEP LEARNING TECHNIQUES



ภัณฑ์ภัตสร คงมณี
PANPATSORN KONGMANEE

การค้นคว้าอิสระนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาวิทยาการข้อมูลและการวิเคราะห์
ศูนย์วิเคราะห์ข้อมูลดิจิทัลอัจฉริยะพระจอมเกล้าลาดกระบัง คณะวิทยาศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

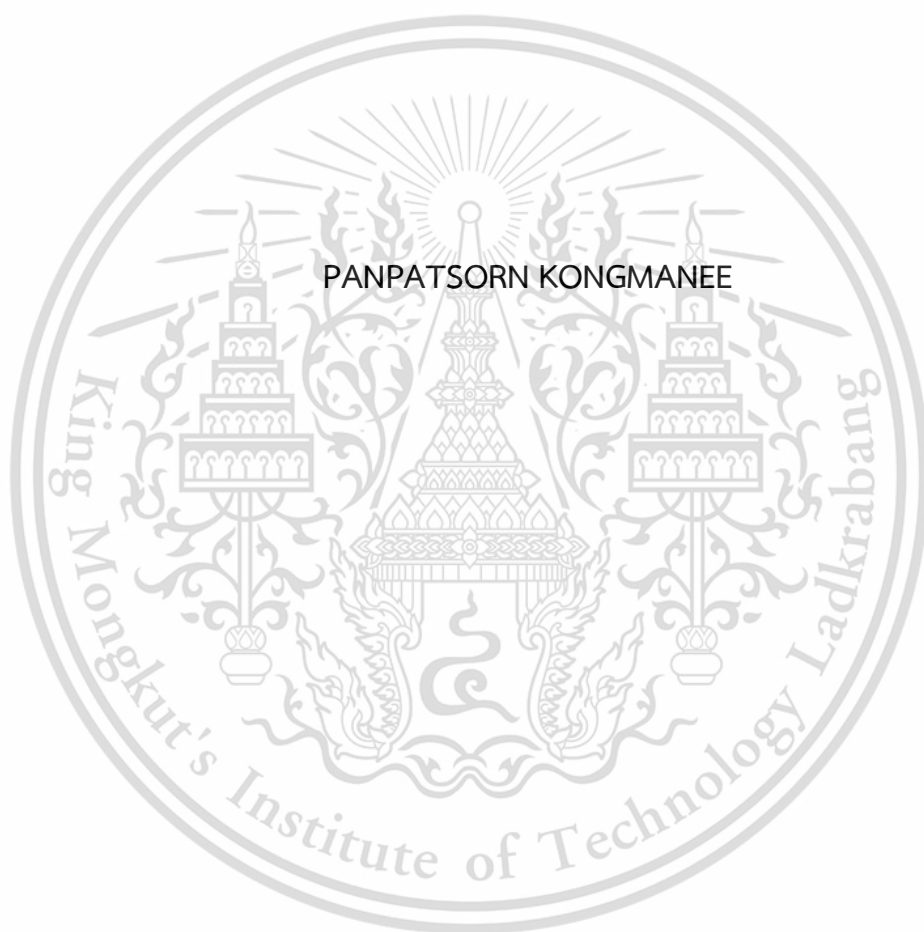
พ.ศ. 2566

KMITL-2023-SC-M-017-074

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

CLASSIFICATION OF THAI RICE VARIETIES USING IMAGE
PROCESSING AND DEEP LEARNING TECHNIQUES

PANPATSORN KONGMANEE



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

2023

KMITL-2023-SC-M-017-074

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



COPYRIGHT 2023

SCHOOL OF SCIENCE

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อการค้นคว้าอิสระ	การจำแนกสายพันธุ์ข้าวไทยโดยใช้เทคนิคการประมวลผลภาพและการเรียนรู้เชิงลึก
ชื่อนักศึกษา	ภณัทภัสสร คงมณี
รหัสประจำตัว	64605083
ปริญญา	วิทยาศาสตรมหาบัณฑิต (วิทยาการข้อมูลและการวิเคราะห์)
พ.ศ.	ศูนย์วิเคราะห์ข้อมูลดิจิทัลอัจฉริยะพระจอมเกล้าลาดกระบัง 2566
อาจารย์ที่ปรึกษาการค้นคว้าอิสระ	รองศาสตราจารย์ ดร.ละออ บุญเกษม

บทคัดย่อ

เนื่องจากปัจจุบันวิธีการตรวจพันธุ์ข้าวเปลือกไทยที่ใช้โดยทั่วไปยังมีความซับซ้อน ใช้เวลา และอาศัยความเชี่ยวชาญสูงเพื่อให้ได้ผลลัพธ์ที่แม่นยำ ผู้วิจัยจึงมีความสนใจศึกษาแบบจำลองที่สามารถบอกตำแหน่งและจำแนกสายพันธุ์ข้าวเปลือกได้โดยมีประสิทธิภาพทั้งความแม่นยำและความรวดเร็ว ซึ่งพันธุ์ข้าวเปลือกไทยดังกล่าวคือข้าวหอมมะลิและข้าวหอมปทุมธานี 1 ซึ่งเป็นพันธุ์ข้าวที่เป็นที่นิยมและมีราคาสูง แบบจำลองที่นำเสนอคือแบบจำลองแบ่งส่วนตัวอย่างของ YOLOv8 ซึ่งถูกเปรียบเทียบกับแบบจำลองแบ่งส่วนตัวอย่างอื่นที่ได้รับความนิยมในปัจจุบัน ประกอบด้วย YOLACT SOLOV2 Mask R-CNN และนำไปเปรียบเทียบกับแบบจำลอง YOLOv8 ที่ไม่ได้รับการฝึกสอนล่วงหน้า ทั้งยังหาค่าไฮเปอร์พารามิเตอร์ที่เหมาะสม โดยเปรียบเทียบทั้งด้านความแม่นยำ ความรวดเร็ว และความเสถียรของแบบจำลอง จากการทดลองพบว่าแบบจำลองแบ่งส่วนตัวอย่าง YOLOv8n-seg ซึ่งมีค่าพารามิเตอร์น้อยที่สุด ให้ค่าความแม่นยำที่ใกล้เคียงแบบจำลองแบ่งส่วนตัวอย่างอื่น ๆ ของ YOLOv8 ที่มีจำนวนพารามิเตอร์มากกว่า และใช้เวลาประมวลผลนานกว่า และเมื่อเปรียบเทียบกับแบบจำลองอื่นที่ได้รับความนิยมในปัจจุบัน พบว่า แบบจำลองที่นำเสนอมีความสามารถมากกว่าทั้งในด้านความแม่นยำ และด้านความเร็วในการประมวลผล

คำสำคัญ : การประมวลผลภาพ การเรียนรู้เชิงลึก แบบจำลองแบ่งส่วนตัวอย่าง พันธุ์ข้าวเปลือกไทย YOLOv8

Independent Study Title	Classification of Thai Rice Varieties using Image Processing and Deep Learning Techniques
Student Name	Panpatsorn Kongmanee
Student ID	64605083
Degree	Master of Science (Data Science and Analytics) KMITL Digital Analytics and Intelligence Center
Year	2023
Independent Study Advisor	Assoc.Prof. Dr. Laor Boongasame

Abstract

The methods for identifying Thai paddy varieties are currently complex, time-consuming, and require high expertise to achieve accurate results. This research explores models capable of efficiently classifying the strains of Thai paddy with high accuracy and speed. The focus paddy varieties are Thai Hommali and Hom Pathum Thani 1, both are popular and high-value varieties. The proposed model is based on an instance segmentation model of YOLOv8 which is compared against popular instance segmentation models such as YOLACT, SOLOv2, Mask R-CNN. Additionally, hyperparameter tuning is performed to ascertain the most optimal values. The evaluation includes model performance reported in the forms of mean average precision (mAP), inference time and model stability. Experimental results indicate that YOLOv8n-seg, with the fewest parameters, achieves accuracy comparable to other YOLOv8-based models with more parameters. When compared to other state-of-the-art models, the proposed model demonstrates superior performance in terms of both accuracy and processing speed.

Keywords : Image Processing, Deep Learning, Instance Segmentation, Thai Paddy Varieties, YOLOv8

กิตติกรรมประกาศ

การศึกษาค้นคว้าอิสระหัวข้อ การจำแนกสายพันธุ์ข้าวไทยโดยใช้เทคนิคการประมวลผลภาพ และการเรียนรู้เชิงลึก ผู้วิจัยขอขอบพระคุณ รองศาสตราจารย์ ดร.ละอ อบุญเกษม อาจารย์ที่ปรึกษาการค้นคว้าอิสระ ที่เสียสละเวลาให้คำปรึกษา แนะนำแนวทาง และการแก้ปัญหาต่าง ๆ ให้การค้นคว้าอิสระผ่านสำเร็จลุล่วงไปด้วยดี

ขอขอบพระคุณ ผศ.ดร. ปัทมา เจริญพร ที่กรุณาเป็นประธานกรรมการ และ ดร.จิรภัทร์ หยกรัตนศักดิ์ ที่กรุณาเป็นกรรมการในการสอบในการค้นคว้าอิสระนี้ และช่วยตรวจสอบแก้ไขให้การค้นคว้าอิสระเล่มนี้สมบูรณ์

ขอขอบพระคุณเพื่อนๆ และเจ้าหน้าที่ คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบังทุกท่าน ที่คอยช่วยอำนวยความสะดวกตลอดในหลายๆ เรื่อง

และขอขอบพระคุณบิดา มารดา และครอบครัวของข้าพเจ้า ถวายทอดประสพการณ์ความรู้ในเรื่องข้าวเปลือก ข้าวสาร ให้คำปรึกษาและคำแนะนำตลอดการศึกษา ทำให้ผู้วิจัยสามารถนำความรู้เหล่านั้นมาประยุกต์ใช้ในการศึกษาค้นคว้าอิสระหัวข้อนี้ ทั้งยังให้การสนับสนุน และเป็นกำลังใจที่สำคัญในการทำการค้นคว้าอิสระเล่มนี้ให้ผ่านไปได้โดยดี

นางสาวภัณฑิลาภัสสร คงมณี

สารบัญ

	หน้า
บทคัดย่อภาษาไทย	ก
บทคัดย่อภาษาอังกฤษ	ข
กิตติกรรมประกาศ	ค
สารบัญ	ง
สารบัญตาราง	ฉ
สารบัญรูป	ช
บทที่ 1 บทนำ	1
1.1 ความเป็นมาและความสำคัญของงานวิจัย/ปัญหา	1
1.2 วัตถุประสงค์ของงานวิจัย	2
1.3 ขอบเขตของงานวิจัย	2
1.4 ประโยชน์ที่คาดว่าจะได้รับ	2
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง	3
2.1 แนวคิดและทฤษฎีเกี่ยวกับข้าวสาร	3
2.1.1 ข้อมูลเบื้องต้นเกี่ยวกับข้าวหอมมะลิไทย และข้าวหอมปทุมธานี 1	3
2.1.2 วิธีการตรวจสอบความเป็นข้าวหอมมะลิไทย	4
2.2 แนวคิดและทฤษฎีเกี่ยวกับการประมวลผลภาพ	7
2.2.1 การดำเนินการตามจุด (Point Operation) หรือการแปลงความเข้ม (Intensity Transformation)	8
2.2.2 ตัวกรองเชิงพื้นที่ (Spatial Filters) หรือคอนโวลูชัน (Convolution)	8
2.2.3 การประมวลผลภาพกับรูปร่างและโครงสร้างของภาพ (Morphological Image Processing)	11
2.2.4 การหาเค้าโครงรูปร่าง (Contour detection)	12
2.2.5 การแบ่งส่วนรูปภาพ (Image segmentation)	13
2.3 แนวคิดและทฤษฎีเกี่ยวกับการเรียนรู้เชิงลึก	14
2.3.1 แนวคิดและทฤษฎีเกี่ยวกับโครงข่ายประสาทแบบคอนโวลูชัน	15
2.3.2 การตรวจจับวัตถุ และการแบ่งส่วนตัวอย่าง	17
2.3.3 การพัฒนาแบบจำลอง YOLO	19
2.4 การประเมินผล	26
2.5 โปรแกรมภาษาไพธอน (Python)	28
2.5.1 ไลบรารีจัดการข้อมูลรูปภาพ (Image Manipulation Library)	28

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

	หน้า
2.5.2 ไส้บรรณัฒนาแบบจำลอง (Model Implementation Library)	29
2.6 โปรแกรม Easyrice	29
2.7 งานวิจัยที่เกี่ยวข้อง	29
2.7.1 งานวิจัยที่เกี่ยวข้องกับคุณภาพข้าวสาร	29
2.7.2 งานวิจัยที่เกี่ยวข้องกับการคัดแยกเมล็ดข้าวเปลือก	31
2.7.3 งานวิจัยที่เกี่ยวข้องกับการคัดแยกเมล็ดข้าวสาร	32
บทที่ 3 วิธีการดำเนินงานวิจัย	35
3.1 ขั้นตอนการดำเนินงานภาพรวม	35
3.2 การเก็บรวบรวมข้อมูล	36
3.3 การจัดเตรียมข้อมูล (Data Preparation)	36
3.3.1 ความเปรียบต่าง (Contrast)	37
3.3.2 การปรับภาพให้เรียบ (Blur Image) หรือเครื่องกรองผ่านความถี่ต่ำ	38
3.3.3 การหาขอบภาพ (Edge detection)	39
3.3.4 การขยายขยาย (Dilation) และการกัดเซาะ (Erosion)	40
3.3.5 การหาเส้นเค้าโครง	40
3.4 การสร้างแบบจำลอง	41
3.5 การประเมินผล	43
บทที่ 4 ผลการวิจัยและการอภิปรายผล	44
4.1 แบบจำลองการแบ่งส่วนตัวอย่างอื่นที่ได้รับความนิยม	44
4.2 แบบจำลองการแบ่งส่วนตัวอย่าง YOLOv8 ที่ได้รับการฝึกสอนแล้ว	44
4.2.1 จำนวนรอบการฝึกสอน (Epoch)	45
4.2.2 จำนวนรายการข้อมูลที่คำนวณในหนึ่งครั้ง (Batch size)	45
4.2.3 อัตราการเรียนรู้ (Learning rate)	46
4.2.4 Optimizer	47
4.3 แบบจำลองการแบ่งส่วนตัวอย่าง YOLOv8 ที่ไม่ได้ผ่านการฝึกสอน	48
บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ	50
5.1 สรุปผลการวิจัย	50
5.2 ข้อเสนอแนะ	51
เอกสารอ้างอิง	52

สารบัญตาราง

ตารางที่	หน้า
2.1 ลักษณะประจำพันธุ์ข้าวหอมมะลิไทย	4
2.2 ระดับของการสลายของเมล็ดข้าวในตั้งแต่แต่ละเมล็ด	5
2.3 เมทริกซ์ความสับสน	26
2.4 สรุปงานวิจัยที่เกี่ยวข้อง	30
3.1 แบบจำลองการแบ่งส่วนตัวอย่าง YOLOv8 ที่ได้รับการฝึกสอน	41
3.2 กำหนดค่าในการทดลองการปรับค่าไฮเปอร์พารามิเตอร์	42
4.1 ผลการทดลองเปรียบเทียบแบบจำลองการแบ่งส่วนตัวอย่างแบบเรียลไทม์อื่น	44
4.2 ผลการทดลองแบบจำลองการแบ่งส่วนตัวอย่าง YOLOv8	45
4.3 ผลการทดลองการกำหนด Batch size ขนาดต่าง ๆ	46
4.4 ผลการทดลองที่อัตราการเรียนรู้ต่าง ๆ	46
4.5 ผลการทดลองใน Optimizer ที่อัตราการเรียนรู้ขนาดต่าง ๆ	47
4.6 ค่าไฮเปอร์พารามิเตอร์ที่เหมาะสมกับข้อมูลชุดนี้	48
4.7 ผลการทดลองระหว่างแบบจำลองที่ฝึกล่วงหน้าและไม่ได้ฝึกล่วงหน้า	49

สารบัญรูป

รูปที่	หน้า
2.1 การเคลื่อนที่ของเคอเนล (ก) หน้าต่างเคลื่อนที่ ขนาด 3×3 จุดภาพ เคลื่อนที่ไปบนภาพที่จะประมวลผล (ข) เลื่อนไปตามแถวจากจุดภาพหนึ่งไปยังอีกจุดภาพหนึ่งจนหมดแถว (ค) เลื่อนหน้าต่างเคลื่อนที่จากแถวหนึ่งไปยังอีกแถวหนึ่ง	9
2.2 แสดงความถี่เชิงพื้นที่แบบง่าย (ก) ความถี่เชิงพื้นที่เป็นศูนย์ เป็นภาพที่ทุกจุดภาพมีค่าเท่ากัน (ข) ความถี่เชิงพื้นที่ต่ำเป็นภาพที่ค่อย ๆ มีการเปลี่ยนแปลงของค่าหรือโทนสี (ค) ความถี่เชิงพื้นที่สูง เป็นภาพที่มีการตัดกันของโทนสีขาวดำอย่างชัดเจน หรือมีการเปลี่ยนโทนสีอย่างฉับพลัน	10
2.3 โครงสร้างของโครงข่ายประสาทแบบคอนโวลูชัน	16
2.4 การแบ่งประเภทของการตรวจจับวัตถุ	17
2.5 ความแตกต่างของอัลกอริทึมของการตรวจจับวัตถุทั้งสองประเภท	18
2.6 แสดงการพัฒนาของแบบจำลองตั้งแต่ YOLOv1 จนถึง YOLOv8	20
2.7 การเปรียบเทียบประสิทธิภาพระหว่างแบบจำลอง YOLOv8 และแบบจำลองอื่น	23
2.8 ส่วนโครงสร้างหัวการตรวจจับของ YOLOv5 ที่ถูกแสดงผ่าน netron.app	24
2.9 ส่วนโครงสร้างหัวการตรวจจับของ YOLOv8 ที่ถูกแสดงผ่าน netron.app	24
2.10 โครงสร้างแบบจำลองของแบบจำลองการตรวจจับวัตถุ YOLOv8	25
3.1 กระบวนการดำเนินงาน อธิบายถึงขั้นตอนโดยภาพรวม	35
3.2 ส่วนประกอบของเมล็ดข้าวเปลือก	38
3.3 ก) ส่วนของเมล็ดข้าวเปลือกที่มีดีกว่าส่วนอื่นและมีสีใกล้เคียงพื้นหลัง 3.3 ข) ส่วนของเมล็ดข้าวเปลือกหลังผ่านการปรับค่าความเปรียบต่าง	38
3.4 ก) ภาพข้าวเปลือกหลังจากใช้ตัวกรองภาพเลือนฟังก์ชันเกาส์ และทวิภาคี 3.4 ข) ภาพข้าวเปลือกหลังจากปรับภาพเป็นสีเทา	39
3.5 ภาพข้าวเปลือกหลังจากหาขอบภาพ	39
3.6 ก) ภาพข้าวเปลือกหลังจากใช้วิธีการขยายขนาด 3.6 ข) ภาพข้าวเปลือกหลังจากใช้วิธีการขยายกัดเซาะ	40
3.7 ก) คำอธิบายประกอบรูปภาพ 3.7 ข) รูปภาพนำเข้าพร้อมคำอธิบายประกอบรูปภาพ	40
3.8 ประสิทธิภาพแบบจำลองการแบ่งส่วนตัวอย่างแบบเรียลไทม์บนชุดข้อมูล MSCOCO	42
4.1 ค่า mAP50-95 ตลอดการฝึกสอน 100 รอบ ที่อัตราการเรียนรู้ขนาดต่าง ๆ	47
4.2 ค่า mAP50-95 ตลอดการฝึกสอนในอัตราการเรียนรู้ขนาดต่าง ๆ	48
4.3 กราฟค่า mAP50-95 ระหว่างแบบจำลองที่ฝึกล่วงหน้าและไม่ได้ฝึกล่วงหน้า	49

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ข้าวเป็นหนึ่งในพืชเศรษฐกิจสำคัญของประเทศไทย ทั้งในด้านการบริโภค (ข้าวเป็นอาหารหลักของคนไทย) และการเป็นสินค้าเกษตรส่งออกอันดับหนึ่งของประเทศไทย อีกทั้งยังมีความสำคัญต่อภาวะเศรษฐกิจภูมิภาค เนื่องจากเป็นพืชเกษตรที่ครอบคลุมพื้นที่เพาะปลูกมากที่สุด คิดเป็น 46% ของพื้นที่เพาะปลูกทั้งหมดในประเทศไทยในปี 2562 สำหรับช่วงปี 2565-2567 ปริมาณผลผลิตมีแนวโน้มเพิ่มขึ้นจากปี 2564 ขณะที่ราคามีแนวโน้มอ่อนตัวลง อย่างไรก็ตาม การแข่งขันที่ค่อนข้างสูงยังเป็นปัจจัยกดดันผลกำไรของผู้ประกอบการโรงสีข้าว ดังนั้นจึงมีความจำเป็นที่ธุรกิจโรงสีข้าวจะต้องบริหารจัดการให้ได้สินค้าที่มีคุณภาพและมีต้นทุนต่ำ ปัจจัยหลักซึ่งมีความสำคัญที่จะส่งผลต่อต้นทุนและการควบคุมคุณภาพข้าว คือรับซื้อข้าวเปลือกที่มีคุณภาพและตรงตามพันธุ์ข้าวที่ต้องการ

เนื่องจากข้าวเปลือกคือวัตถุดิบหลักของธุรกิจโรงสีข้าว ขั้นตอนการรับซื้อข้าวเปลือกเพื่อให้ได้ข้าวเปลือกที่มีคุณภาพและตรงตามพันธุ์ที่ต้องการ จึงเป็นปัจจัยที่กระทบโดยตรงต่อต้นทุนและการควบคุมคุณภาพข้าวสารโดยตรง ปัจจัยการพิจารณาเพื่อรับซื้อข้าวเปลือก ประกอบด้วย ความชื้น สิ่งเจือปน ลักษณะและสีของเมล็ดข้าวเปลือก และที่สำคัญคือพันธุ์ข้าวเปลือก เนื่องจากข้าวแต่ละพันธุ์มีราคาและลักษณะประจำพันธุ์แตกต่างกัน ยกตัวอย่างเช่น ข้าวสารหอมมะลิเมล็ดข้าวสุกจะอ่อนนุ่มมากกว่าข้าวเจ้าทั่วไป ราคา ณ ปัจจุบันอยู่ที่ 28.5 บาทต่อกิโลกรัม ข้าวสารหอมปทุมธานี 1 มีความนุ่มน้อยกว่าข้าวสารหอมมะลิ ราคาอยู่ที่ 22 บาทต่อกิโลกรัม ข้าวขาวหรือข้าวพื้นแข็ง ราคา 18.8 บาทต่อกิโลกรัม หากไม่สามารถควบคุมพันธุ์ข้าวเปลือกซื้อเข้าได้อย่างมีประสิทธิภาพ กล่าวคือมีข้าวพันธุ์อื่นปลอมปนมากจนเกินไป หลังจากกระบวนการสีข้าวจะได้ข้าวสารที่มีคุณภาพไม่แน่นอน ส่งผลต่อความน่าเชื่อถือของสินค้า และข้าวสารไม่ตรงตามมาตรฐานข้าวสารไทยที่ได้กำหนดอัตราส่วนข้าวไว้อย่างชัดเจน เช่น มาตรฐานข้าวหอมมะลิไทยต้องมีข้าวหอมมะลิไทยไม่น้อยกว่าร้อยละ 92 เป็นต้น ดังนั้นจึงมีความจำเป็นที่ต้องตรวจสอบพันธุ์ข้าวก่อนขั้นตอนการรับซื้อ

ปัจจุบันได้มีวิธีการตรวจสอบสายพันธุ์ข้าวเปลือก ซึ่งมีทั้งวิธีที่เชื่อถือได้คือการตรวจสอบสายพันธุ์กรรม (DNA) แต่ใช้เวลาและค่าใช้จ่ายสูง และวิธีที่ใช้จริงในทางปฏิบัติ ประกอบด้วย การหาค่าการสลายเมล็ดข้าวในด่าง การตรวจสอบด้วยการต้ม หรือการย้อมสีด้วยไอโอดีน หากแต่วิธีเหล่านี้ให้ผลลัพธ์ที่ใกล้เคียงกันมากระหว่างข้าวหอมมะลิ และข้าวหอมปทุมธานี 1 อาจเกิดความผิดพลาดในการวิเคราะห์ได้ง่าย ทั้งนี้ในทางปฏิบัติยังนิยมให้ผู้เชี่ยวชาญใช้การสังเกตด้วยสายตาและดมกลิ่น ซึ่งเป็นวิธีที่รวดเร็วและทำได้ง่าย หากแต่ก็มีโอกาสที่จะคลาดเคลื่อนได้เช่นกัน แต่อย่างไรก็ตาม การศึกษาในด้าน

วิทยาการข้อมูลที่เกี่ยวข้องกับคัดแยกเมล็ดพันธุ์ข้าวเปลือกนั้นมีจำนวนไม่มากนัก และในปัจจุบันมีเพียงเอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ผู้ให้บริการรายเดียวที่ออกแบบปัญญาประดิษฐ์ (Artificial Intelligence) เพื่อใช้วิเคราะห์อัตลักษณ์ของข้าวเปลือก และคัดแยกข้าวเปลือกสายพันธุ์ที่ต้องการออกจากสายพันธุ์อื่น หากแต่ใช้เวลาในการประมวลผลถึง 3 นาที ต่อหนึ่งภาพ

ดังนั้นในการศึกษาค้นคว้าครั้งนี้จึงมุ่งเน้นไปที่การบอกตำแหน่งและการจำแนกสายพันธุ์ข้าวเปลือกระหว่างข้าวหอมมะลิและข้าวหอมปทุมธานี 1 โดยไม่เพียงแต่วัดประสิทธิภาพแบบจำลองด้านความแม่นยำ แต่ยังวัดด้านความเร็วในการประมวลผลอีกเช่นกัน เนื่องจากยังไม่มีงานวิจัยที่เน้นด้านความสามารถในการบอกตำแหน่งและการจำแนกพันธุ์ข้าวเปลือกได้อย่างรวดเร็ว เพราะฉะนั้นจึงเลือกใช้แบบจำลองแบ่งส่วนตัวอย่าง YOLOv8 ที่ซึ่งถูกออกแบบให้มีทั้งความเร็วในการประมวลผลและมีความแม่นยำสูง

1.2 วัตถุประสงค์ของงานวิจัย

- 1) เพื่อศึกษาการจัดการกับข้อมูลประเภทภาพ และการทำคำอธิบายประกอบภาพ
- 2) เพื่อพัฒนาแบบจำลองที่สามารถใช้ในการบอกตำแหน่งและการจำแนกสายพันธุ์ข้าวเปลือกแบบเรียลไทม์
- 3) เพื่อเปรียบเทียบประสิทธิภาพแบบจำลองต่าง ๆ ที่ใช้ในการบอกตำแหน่งและการจำแนกสายพันธุ์ข้าวเปลือกแบบเรียลไทม์

1.3 ขอบเขตของงานวิจัย

- 1) การศึกษาค้นคว้ามุ่งศึกษาเฉพาะพันธุ์ข้าวเปลือกหอมมะลิเฉพาะที่ปลูกบริเวณพื้นที่จังหวัดนครราชสีมา และนครนายก ข้าวเปลือกพันธุ์ข้าวหอมปทุมธานี 1 เฉพาะที่ปลูกบริเวณพื้นที่จังหวัดชลบุรี และระยอง
- 2) การศึกษาค้นคว้านี้ทำคำอธิบายประกอบภาพโดยวิธีการหาขอบภาพโดยวิธีแคนนี่ (Canny edge detection) ร่วมกับวิธีการหาเส้นเค้าโครง (FindContour)
- 3) เครื่องมือสำหรับการจัดการข้อมูลและสร้างแบบจำลอง ผู้วิจัยใช้โปรแกรมภาษาไพธอน (Python)
- 4) เครื่องมือสำหรับการเก็บข้อมูลรูปภาพคือเครื่องสแกน Canon รุ่น CanoScan LiDE 300

1.4 ประโยชน์ที่คาดว่าจะได้รับ

- 1) สามารถจัดการกับข้อมูลประเภทภาพ และการทำคำอธิบายประกอบภาพ
- 2) งานวิจัยนี้ทำให้ทราบถึงแบบจำลองที่สามารถใช้ในการบอกตำแหน่งและการจำแนกสายพันธุ์ข้าวเปลือกแบบเรียลไทม์

บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

การศึกษาค้นคว้าอิสระในครั้งนี้ ผู้ค้นคว้าได้ศึกษาเอกสาร แนวคิด ทฤษฎี และงานวิจัยต่าง ๆ ที่เกี่ยวข้องกับเรื่องจำแนกสายพันธุ์ข้าวด้วยการประมวลผลภาพ โดยใช้ประเด็นศึกษา ดังต่อไปนี้

2.1 แนวคิดและทฤษฎีเกี่ยวกับข้าวสาร

2.1.1 ข้อมูลเบื้องต้นเกี่ยวกับข้าวหอมมะลิไทย และข้าวหอมปทุมธานี 1

ข้าวหอมมะลิเป็นสายพันธุ์ข้าวที่มีถิ่นกำเนิดในประเทศไทย เป็นพันธุ์ข้าวที่มีที่มาจากสีของข้าวที่ขาวเหมือนดอกมะลิ มีลักษณะกลิ่นหอมคล้ายใบเตย เป็นพันธุ์ข้าวที่ทำให้ข้าวไทยเป็นสินค้าส่งออกที่รู้จักไปทั่วโลก ในปี พ.ศ.2502 คณะกรรมการพิจารณาพันธุ์ข้าวได้อนุมัติให้เป็นพันธุ์ส่งเสริมแก่เกษตรกร โดยเกษตรกรทั่วไปเรียกว่า ข้าวขาวดอกมะลิ 105 ต่อมาได้มีการปรับปรุงพันธุ์ข้าวขาวดอกมะลิ 105 จนได้ข้าวพันธุ์ กข 15 ซึ่งกระทรวงพาณิชย์ประกาศให้ข้าวทั้ง 2 พันธุ์เป็นข้าวหอมมะลิไทย มีระยะพักตัวในเวลาประมาณ 8 สัปดาห์ โดยมีลักษณะประจำพันธุ์ ดังตารางที่ 2.1

ลักษณะที่สำคัญของข้าวหอมมะลิ คือ พันธุ์ข้าวหอมที่ไวต่อช่วงแสง มีกลิ่นหอมตามธรรมชาติของข้าวใหม่หรือข้าวเก่า เมื่อหุงหรือนึ่งสุกแล้วเมล็ดข้าวสุกจะอ่อนนุ่มมากกว่าข้าวเจ้าทั่วไป แต่ร่วนน้อยกว่า และเป็นข้าวที่ผลิตในประเทศไทยในฤดูนาปี (ปลูกได้ปีละครั้ง) ราคา ณ ปัจจุบัน อ้างอิงตามรายงานราคาข้าวขาวส่งตลาดกรุงเทพฯ สมาคมโรงสีข้าวไทย อยู่ที่ 2,920-3,000 บาท/100กิโลกรัมมาตรฐานข้าวหอมมะลิไทยอ้างอิงตามสำนักงานมาตรฐานสินค้าเกษตรและอาหารแห่งชาติกระทรวงเกษตรและสหกรณ์ กล่าวว่า ข้าวเปลือกของข้าวหอมมะลิไทย ต้องมีคุณลักษณะเฉพาะ ดังต่อไปนี้

ก) เปลือกมีสีฟาง

ข) เมื่อกะเทาะเป็นข้าวกล้องแล้ว ความยาวเฉลี่ยของข้าวเต็มเมล็ดที่ไม่มีส่วนใดหัก ต้องไม่ต่ำกว่า 7.0 มิลลิเมตร และอัตราส่วนความยาวเฉลี่ยต่อความกว้างเฉลี่ยของข้าวเต็มเมล็ดที่ไม่มีส่วนใดหัก ต้องไม่ต่ำกว่า 3.2 : 1

ค) เมื่อขัดสีเป็นข้าวขาวแล้ว มีกลิ่นหอมโดยธรรมชาติของข้าวใหม่หรือข้าวเก่า

ง) มีความชื้นไม่เกิน 15% กรณีข้าวเปลือกที่จะนำไปเก็บรักษาต้องมีความชื้นไม่เกิน 14%

จ) มีข้าวเปลือกหอมมะลิไทยไม่น้อยกว่า 95% โดยปริมาณ

ฉ) กรณีข้าวเปลือกแห้ง ให้มีคุณภาพการขัดสีได้ข้าวเต็มเมล็ดและต้นข้าวขาวตั้งแต่ 36% ขึ้นไปโดยน้ำหนัก

ตารางที่ 2.1 ลักษณะประจำพันธุ์ข้าวหอมมะลิไทย

ลักษณะประจำพันธุ์	พันธุ์ข้าวหอมมะลิไทย	
	ชาวดอกมะลิ 105	กข 15
ปริมาณแอมิโลส (เปอร์เซ็นต์โดยน้ำหนัก)	13.0 ถึง 18.0	13.0 ถึง 18.0
ความไวต่อช่วงแสง	ไวต่อช่วงแสง	ไวต่อช่วงแสง
สีของข้าวเปลือก	ฟาง	ฟาง
หางของข้าวเปลือก	ไม่มี	ไม่มี
ความยาวเมล็ดข้าวเปลือก (มิลลิเมตร)	9.7 ถึง 11.5	9.8 ถึง 11.6
ความยาวเมล็ดข้าวกล้อง (มิลลิเมตร)	7.0 ถึง 8.1	7.0 ถึง 8.1
อัตราส่วนความยาวต่อความกว้างของเมล็ดข้าว กล้อง	$\geq 3:2:1$	$\geq 3:2:1$
น้ำหนักของข้าวเปลือก 100 เมล็ด (กรัม)	2.50 ถึง 2.97	2.50 ถึง 2.97
ระดับค่าการสลายเมล็ดในต่าง	6.0 ถึง 7.0	6.0 ถึง 7.0

(ที่มา: https://www.acfs.go.th/standard/download/Thai-Hom-Mali_Rice-60.pdf)

ข้าวหอมปทุมธานี 1 หรือ ข้าวหอมปทุม มีกลิ่นหอมใกล้เคียงกับข้าวหอมมะลิ โดยเฉพาะข้าวใหม่ช่วงต้นฤดูจะมีกลิ่นหอมมากเป็นพิเศษ [1] แต่ไม่ใช่ข้าวหอมมะลิ โดยจะความนุ่มน้อยกว่า เมื่อวางข้าวที่หุงแล้วทิ้งไว้สักพัก ข้าวจะไม่นุ่มเท่าตอนหุงใหม่ ๆ ต่างจากข้าวหอมมะลิที่ยังคงความนุ่มไว้ได้ อีกข้อหนึ่งที่แตกต่างกันคือข้าวหอมปทุมธานี 1 นี้สามารถปลูกได้ตลอดทั้งปี และแต่ละรอบก็ให้ผลผลิตปริมาณสูง หากมีชลประทาน (น้ำ) อุดมสมบูรณ์อาจได้ผลผลิตถึง 3 รอบต่อปี สามารถปลูกได้ในท้องถิ่นทั่วไป เป็นทางเลือกใหม่ของผู้ที่ชอบรับประทานข้าวที่มีความนุ่มและหอมในแบบคุ้มค่าคุ้มราคา ส่วนข้าวหอมมะลิปลูกได้ปีละครั้งและให้ผลผลิตที่น้อยกว่า ราคาข้าวหอมปทุมธานี 1 ณ ปัจจุบัน อ้างอิงตามรายงานราคาข้าวขาวส่งตลาดกรุงเทพฯ สมาคมโรงสีข้าวไทย อยู่ที่ 1,800-1,850 บาท/100 กิโลกรัม

2.1.2 วิธีการตรวจสอบความเป็นข้าวหอมมะลิไทย

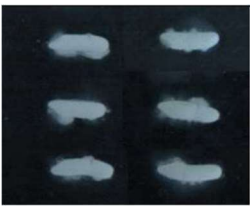
การตรวจสอบที่สามารถยืนยันว่าเป็นข้าวหอมมะลิแท้หรือไม่ ปัจจุบันมีเพียงวิธีเดียวคือการตรวจสอบสายพันธุ์กรรม (DNA) [2] ซึ่งมีสถาบันที่สามารถตรวจสอบได้อยู่บ่อย มีค่าใช้จ่ายในการตรวจสอบค่อนข้างสูง ทั้งยังต้องใช้เวลาในการตรวจสอบพอสมควร สำหรับวิธีการอื่นที่นิยมใช้เพื่อทดแทน ซึ่งเป็นวิธีการทดสอบที่เกี่ยวข้องกับปริมาณแอมิโลส (amylose) ในข้าวสาร โดยแ

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์สำหรับการใช้งานเพื่อการศึกษาเท่านั้น เมื่อได้เนื้อหาไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

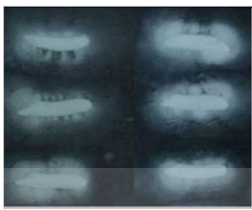
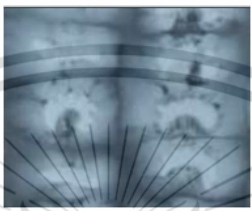
มีโลสคือแป้งหรือคาร์โบไฮเดรตชนิดหนึ่ง โดยทั่วไปมักนิยมแบ่งประเภทข้าวโดยกล่าวถึงแอมิโลสเป็นหลักสำคัญ มีผลทำให้เมื่อหุงเป็นข้าวสวยจะมีความอ่อนนุ่มหรือกระด้างแตกต่างกันไป ข้าวที่มีแอมิโลสสูง ในระหว่างการหุงต้มจะดูดน้ำได้มากกว่าข้าวที่มีแอมิโลสต่ำ ทำให้ข้าวสุกมีความเหนียวลดลงหรือร่วนมากขึ้น ทั้งนี้ปริมาณแอมิโลสที่สูงขึ้นเมื่อหุงต้มสุกจะทำให้ข้าวมีความร่วนและกระด้างมากขึ้น [3] โดยสามารถทดสอบได้โดยวิธีดังต่อไปนี้

- 1) การหาค่าการสลายเมล็ดข้าวในด่าง (alkali spreading value) หมายถึง อัตราการสลายของแป้งในเมล็ดข้าว เมื่อแช่ข้าวที่ขัดสีเยื่อหุ้มเมล็ดออกแล้วในสารละลายโพแทสเซียมไฮดรอกไซด์เข้มข้น 1.7% โดยให้เมล็ดข้าวทุกเมล็ดจมอยู่ในสารละลาย และให้แต่ละเมล็ดอยู่ห่างกันพอสมควร ปิดฝาทิ้งไว้ที่อุณหภูมิห้อง (30 + 5 องศาเซลเซียส) โดยไม่เขย่าเขยื้อนเป็นเวลา 23 ชั่วโมง เมล็ดข้าวที่มีระดับการสลายในด่าง ตั้งแต่ระดับ 1 ถึงระดับ 5 เป็นเมล็ดข้าวที่ไม่ใช่ข้าวหอมมะลิไทย เนื่องจากความอ่อนนุ่มของเมล็ดข้าวนั้น นอกจะจะมีความสัมพันธ์กับปริมาณแอมิโลสแล้ว ยังมีความสัมพันธ์โดยตรงกับอุณหภูมิแป้งสุกที่สามารถตรวจสอบได้โดยวิธีการสลายตัวในด่าง [4] ข้าวที่มีค่าการสลายตัวในด่างสูง (อุณหภูมิแป้งสุกต่ำ) เป็นข้าวที่มีความอ่อนนุ่มมากกว่าข้าวที่มีค่าการสลายตัวในด่างต่ำ (อุณหภูมิแป้งสุกสูง) ดังตารางที่ 2.2 โดยข้าวหอมมะลิไทย และข้าวหอมปทุมธานี 1 จะมีระดับการสลายในด่างที่ใกล้เคียงกัน

ตารางที่ 2.2 ระดับของการสลายของเมล็ดข้าวในด่างแต่ละเมล็ด

ระดับการสลายของเมล็ดข้าว	ภาพลักษณะการสลายของเมล็ดข้าว	ลักษณะของเมล็ดข้าวที่สลายในด่าง
1		ลักษณะของเมล็ดข้าวไม่เปลี่ยนแปลงเลย
2		เมล็ดข้าวพองตัว
3		เมล็ดข้าวพองตัวและมีแป้งกระจายออกมาจากบางส่วนของเมล็ดข้าว

ตารางที่ 2.2 ระดับของการสลายของเมล็ดข้าวในต่างแต่ละเมล็ด (ต่อ)

ระดับการสลาย ของเมล็ดข้าว	ภาพลักษณะการสลาย ของเมล็ดข้าว	ลักษณะของเมล็ดข้าว ที่สลายในต่าง
4		เมล็ดข้าวพองตัวและมีแป้งกระจาย ออกมารอบเมล็ดข้าวเป็นบริเวณ กว้าง
5		ผิวของเมล็ดข้าวปริทางขวางหรือ ทางยาว และมีแป้งกระจายออกมา รอบเมล็ดเป็นบริเวณกว้าง
6		เมล็ดข้าวสลายตัวตลอดทั้งเมล็ด มีลักษณะเป็นเมือกขุนขาว
7		เมล็ดข้าวสลายตัวทั้งเมล็ดและมี ลักษณะเป็นแป้งเมือกใส

(ที่มา: https://www.acfs.go.th/standard/download/Thai-Hom-Mali_Rice-60.pdf)

2) การตรวจสอบด้วยการต้ม เพื่อตรวจสอบค่าไตของเมล็ดข้าว เนื่องด้วยความอ่อนนุ่มนั้น มีความสัมพันธ์กับอุณหภูมิแป้งสุกดังที่ได้กล่าวไป วิธีนี้จึงเป็นวิธีการตรวจสอบข้าวหอมมะลิต่างง่าย ซึ่งมักถูกใช้ในกระบวนการตรวจข้าวของผู้ส่งออก โดยการนำเมล็ดข้าวที่ได้จากการต้มในน้ำเดือดเป็นเวลา 17 นาที ไปจุ่มลงในน้ำเย็นทันที แล้วจึงกดลงบนกระจก เพื่อตรวจดูภายในของเมล็ดข้าวทั้ง 100 เมล็ด ถ้าปรากฏว่าข้าวเมล็ดใดยังเป็นไต โดยมีลักษณะเป็นจุดขุนขาวของแป้งดิบปรากฏภายในเมล็ด ให้ถือว่าเป็นข้าวที่ยังไม่สุกสมบูรณ์ ซึ่งไม่ใช่ข้าวหอมมะลิไทย

3) การย้อมสีด้วยไอโอดีน เป็นอีกหนึ่งวิธีที่นิยมใช้เพื่อตรวจหาเมล็ดพันธุ์ข้าวอื่นที่จัดว่าเป็นข้าวที่มีแอมิโลสสูงโดยเฉพาะข้าวที่ไม่ไวต่อช่วงแสง (ข้าวแดงหรือข้าวนาปรัง) ที่ปะปนมาในเมล็ดพันธุ์ข้าวหอมมะลิ (พันธุ์ขาวดอกมะลิ 105 และ กข15) ซึ่งมีแอมิโลสต่ำ อันจะ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เป็นข้อมูลช่วยสนับสนุนการตัดสินใจในการวิเคราะห์ความบริสุทธิ์เมล็ดพันธุ์ [5] โดยวิธีคือ ใช้สารละลายไอโอดีน 40 มิลลิลิตร และแกว่งเป็นเวลา 45 วินาที แล้วจึงเทสารทิ้ง ซึ่งวิเคราะห์ผลโดยการแบ่งการติดสีของเมล็ดเป็น 2 กลุ่ม ดังนี้

กลุ่มที่ 1 เมล็ดติดสีชมพู ถึง ไม่ติดสี เป็นข้าวพื้นนุ่ม หรือ แอมิโลสต่ำ อยู่ในกลุ่ม ข้าวเจ้านุ่ม เช่น ข้าวพันธุ์ขาวดอกมะลิ 105 กข15 ข้าวหอมปทุมธานี1

กลุ่มที่ 2 เมล็ดติดสีน้ำเงินหรือม่วงเข้ม เป็นข้าวพื้นแข็ง หรือ แอมิโลสปานกลาง หรือแอมิโลสสูงในกลุ่มข้าวเจ้าร้อน หรือข้าวเจ้าแข็ง เช่น ข้าวพันธุ์พิษณุโลก2 ชัยนาท1

หากพบเมล็ดที่ติดสีน้ำเงินหรือม่วงเข้ม มากกว่า 1 เมล็ด แสดงว่า มีปริมาณพันธุ์ ข้าวเจ้าปะปนในตัวอย่างข้าวหอมมะลิเกินกว่ามาตรฐาน โดยที่การย้อมสีด้วย ไอโอดีน วิธีนี้สามารถทำได้ง่าย แต่มีข้อเสียคือสามารถแยกได้เพียงข้าวพื้นนุ่ม และ ข้าวพื้นแข็ง แต่ไม่สามารถแยกระหว่างข้าวหอมมะลิ และข้าวหอมปทุมธานี 1 ได้ เนื่องจากมีปริมาณแอมิโลสที่ใกล้เคียงกัน

จะเห็นได้ว่าวิธีดังกล่าวข้างต้นนั้นใช้เวลาพอสมควรในการตรวจสอบสายพันธุ์ข้าว อีกทั้ง ในหลาย ๆ วิธียังเป็นเพียงการแยกระหว่างข้าวพื้นนุ่มและข้าวพื้นแข็งเท่านั้น กล่าวคือจากการ ตรวจสอบโดยใช้วิธีทดแทนการตรวจสอบสายพันธุ์กรรม ผลลัพธ์ที่ได้ระหว่างข้าวหอมมะลิไทย และ ข้าวหอมปทุมธานี 1 นั้นใกล้เคียงกัน สำหรับในทางปฏิบัติอาจใช้ผู้เชี่ยวชาญในการสังเกตด้วย สายตาและดมกลิ่น เนื่องจากเป็นวิธีที่รวดเร็วและทำได้ง่าย หากแต่ก็มีโอกาสที่จะคลาดเคลื่อนได้ เนื่องมาจากประสบการณ์และความผิดพลาดของมนุษย์

2.2 แนวคิดและทฤษฎีเกี่ยวกับการประมวลผลภาพ

การประมวลผลภาพ (Image Processing) คือการนำภาพมาประมวลผลหรือคำนวณด้วยเครื่อง คอมพิวเตอร์เพื่อให้ได้ข้อมูลที่ต้องการทั้งในเชิงปริมาณและเชิงคุณภาพ โดยมีขั้นตอนต่าง ๆ ที่สำคัญคือ การทำให้ภาพมีความคมชัดมากขึ้น การกำจัดสัญญาณรบกวน การแบ่งส่วนของวัตถุที่สนใจ เพื่อนำภาพ วัตถุที่ได้ไปวิเคราะห์หาข้อมูลเชิงปริมาณ เช่น ขนาด รูปร่าง และทิศทางการเคลื่อนของวัตถุในภาพ จากนั้นสามารถนำข้อมูลเชิงปริมาณที่ได้เหล่านี้ไปวิเคราะห์เพื่อใช้ประโยชน์ในงานด้านอื่น เช่น การรู้จำ ลายนิ้วมือ หรือการตรวจจบบั้ววัตถุในภาพ เป็นต้น

การปรับปรุงคุณภาพของภาพ (Image Enhancement) เป็นหัวข้อย่อยในเรื่องการประมวลผล ภาพ ซึ่งมีความหมายถึงการเน้นรายละเอียดภายในภาพ ทั้งยังลดหรือลบองค์ประกอบที่ไม่จำเป็น เช่น การขจัดสัญญาณรบกวนหรือจุดต่างในภาพ การเปิดเผยในส่วนที่คลุมเครือ การเน้นขอบหรือการหา ขอบภาพ เป็นต้น โดยทั่วไปการปรับปรุงภาพแบ่งได้เป็น 2 ประเภทหลัก ประกอบด้วย

1. โดเมนเชิงพื้นที่ (Spatial Domain) เป็นการปรับปรุงภาพโดยการดำเนินการแต่ละพิกเซลตาม

พิกัดเชิงพื้นที่ที่ความละเอียดเฉพาะ [6]

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. โดเมนความถี่ (Frequency Domain) เป็นการปรับปรุงภาพโดยใช้การแปลงฟูริเยร์ (Fourier Transform) กับโดเมนเชิงพื้นที่ การจัดการพิกเซลในกลุ่มและทางอ้อม [6]

สำหรับงานวิจัยนี้เกี่ยวข้องกับเทคนิคที่นำไปใช้ในโดเมนเชิงพื้นที่ สามารถแบ่งย่อยได้ 2 ประเภท คือ

2.2.1 การดำเนินการตามจุด (Point Operation) หรือการแปลงความเข้ม (Intensity Transformation)

คือการใช้การเปลี่ยนแปลงกับแต่ละพิกเซลหรือจุดในภาพระดับสีเทา ซึ่งเป็นการเปลี่ยนแปลงชนิดเดียวกันทั้งหมด โดยผลลัพธ์ที่ได้จะขึ้นอยู่กับค่าของพิกเซลดั้งเดิม ไม่ขึ้นอยู่กับตำแหน่งหรือพิกเซลข้างเคียง มักใช้เพื่อเปลี่ยนช่วงระดับและการกระจายของสีเทา ดังนั้นทุกพิกเซลบนรูปภาพใหม่จะถูกคำนวณจากฟังก์ชันการแปลงที่กำหนดไว้ล่วงหน้า ตัวอย่างของตัวดำเนินการดังกล่าว ได้แก่ การปรับความสว่าง (brightness adjustment) การปรับค่าความเปรียบต่าง (contrast adjustment) เป็นต้น รายละเอียดมีดังนี้

$$g(x) = \alpha f(x) + \beta \quad (2.1)$$

1) การปรับความสว่าง คือการบวกด้วยค่าคงที่เข้าไปที่แต่ละพิกเซล ดังค่า β ในสมการที่ 2.1 ความหมายคือจากค่าความเข้มแสงในแต่ละพิกเซล ซึ่งมีค่าระหว่าง 0–255 นั้น ถูกบวกเข้าด้วยค่าคงที่ จะทำให้แต่ละพิกเซลสว่างเพิ่มมากขึ้น

2) การปรับค่าความเปรียบต่าง (Contrast) ซึ่งค่าความเปรียบต่าง คือ ความชัดหรือความต่างของสีแต่ละจุดบนภาพนั้น เกิดจากการมองเห็นจุดที่อยู่ข้างกันแตกต่างกัน หากสองจุดมีความแตกต่างกันมาก จะหมายถึงมีค่าความเปรียบต่างมาก โดยการปรับค่าความเข้มแสงในแต่ละพิกเซลให้สามารถมองเห็นความต่างได้มากขึ้น จะต้องคูณจำนวนเต็มหนึ่ง ๆ ให้กับแต่ละพิกเซล ดังค่า α ในสมการที่ 2.1 แต่หากต้องการให้ความต่างน้อยลง ต้องคูณด้วยจำนวนทศนิยมที่น้อยกว่า 1 (ไม่ใช่ตัวเลขติดลบ)

การปรับแก้ความสว่างและความเปรียบต่างนี้ สามารถใช้ฟังก์ชันในไลบรารี opencv เพื่อช่วยเรื่องดำเนินการดังกล่าว คือใช้ฟังก์ชัน `cv.convertScaleAbs()` โดยมีพารามิเตอร์ดังต่อไปนี้ 1. รูปภาพนำเข้า 2. ค่า α ที่เกี่ยวข้องกับการปรับค่าความเปรียบต่าง และ 3. ค่า β ที่เกี่ยวข้องกับการปรับค่าความสว่าง

2.2.2 ตัวกรองเชิงพื้นที่ (Spatial Filters) หรือคอนโวลูชัน (Convolution)

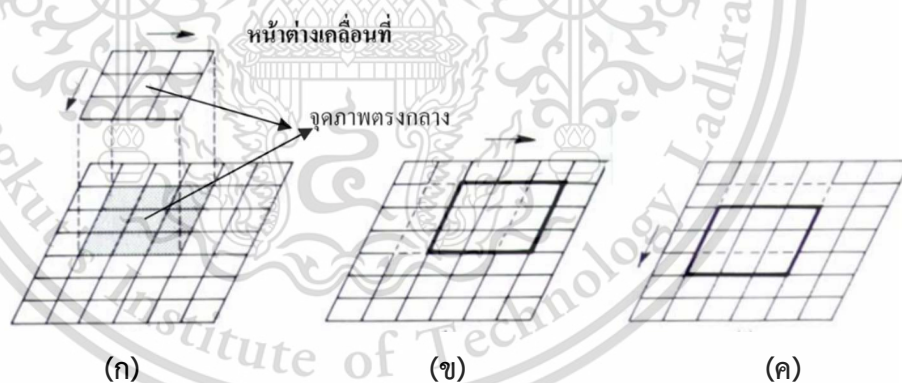
ผลลัพธ์ของการดำเนินการเหล่านี้ขึ้นอยู่กับค่าของฟังก์ชัน $f(x, y)$ และพื้นที่ใกล้เคียงที่มีความหมายว่าความเข้มที่ถูกแปลงนั้นถูกกำหนดโดยค่าสีเทาของจุดภายในบริเวณใกล้เคียง ซึ่งการดำเนินการดังกล่าวเรียกว่า คอนโวลูชัน หรือ สังวัตนาการ มีความหมายถึงการม้วนทับ [7]

เป็นวิธีการทางคณิตศาสตร์ ทำหน้าที่เฉลี่ยค่าของจุดภาพภายในเซตเล็ก ๆ และทำไปกับทั่วทั้งเอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษานาน เมื่ออนุญาตให้ใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

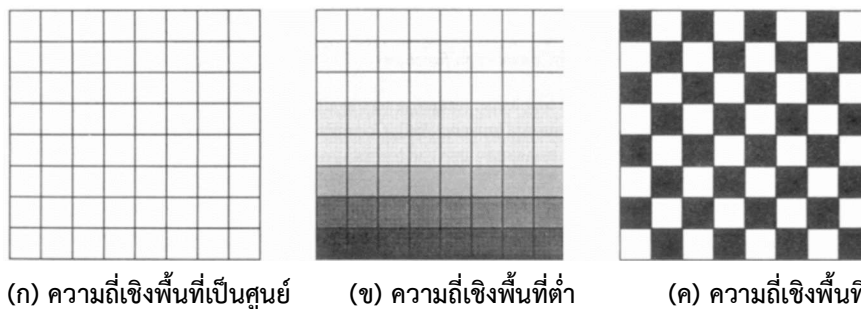
ภาพ โดยวิธีการคือ นำเมตริกขนาดเล็ก ที่เรียกว่า เคอร์เนล (kernel) หรือเรียกอีกอย่างว่า ตัวกรอง (filter) คือตารางสี่เหลี่ยมของจุดภาพมีขนาดเป็นจำนวนเลขคี่ เช่น 3×3 หรือ 5×5 หรือ 7×7 จุดภาพ ซึ่งได้จากค่าสัมประสิทธิ์ (coefficients) หรือปัจจัยน้ำหนัก (weighting factors) ที่ใช้เป็นตัวถ่วงน้ำหนัก ทำหน้าที่คล้ายตะแกรงที่มีช่องขนาดต่าง ๆ [8] แล้วเคลื่อนที่ไปลักษณะดังรูป 2.1 โดยจะเคลื่อนไปที่ละพิกเซลตลอดทั่วทั้งรูปภาพ แล้วค่าตัวเลขของพิกเซลตรงกลางของเคอร์เนลจะถูกคำนวณใหม่ โดยมีวิธีคำนวณคือ การคูณสัมประสิทธิ์แต่ละตัวในเคอร์เนล ด้วยค่าตัวเลขของภาพดั้งเดิมที่เคอร์เนลเคลื่อนที่ไป จากนั้นรวมผลคูณทั้งหมด ค่าที่ได้จะถูกนำไปแทนค่าของพิกเซลตรงกลางเคอร์เนลนั้น ๆ ทำให้ภาพมีลักษณะไม่เหมือนเดิม สำหรับในเรื่องของคอนโวลูชัน เคอร์เนลและตัวกรอง อาจมองว่ามีความหมายเดียวกันได้ ใช้คำทดแทนกันได้ [7]

การนำเรื่องคอนโวลูชันไปใช้นั้นมีวัตถุประสงค์หลายอย่าง เช่น การกรองภาพให้ภาพชัด ภาพนุ่มนวล เพื่อการตรวจสอบขอบ หรือเน้นขอบของสิ่งของในภาพ การลบจุดขาว-จุดดำในภาพ [8] โดยทั่วไปสามารถแบ่งประเภทได้ ดังนี้

- 1) เครื่องกรองผ่านความถี่ต่ำ (low-pass filter) เป็นเทคนิคการกรองภาพที่บริเวณที่มีความถี่เชิงพื้นที่ต่ำสามารถผ่านไปได้ ดังรูปที่ 2.2 (ข) ซึ่งความถี่เชิงพื้นที่ หมายถึง ความแตกต่างระหว่างค่าสูงสุดและต่ำสุดของกลุ่มพิกเซลที่ติดกัน หรือใกล้เคียงกัน หรือหมายถึง จำนวนการเปลี่ยนแปลงค่าความสว่างต่อหน่วยระยะทางในส่วนหนึ่งของภาพ



รูปที่ 2.1 การเคลื่อนที่ของเคอร์เนล (ก) หน้าต่างเคลื่อนที่ ขนาด 3×3 จุดภาพเคลื่อนที่ไปบนภาพที่จะประมวลผล (ข) เลื่อนไปตามแถวจากจุดภาพหนึ่งไปยังอีกจุดภาพหนึ่งจนหมดแถว (ค) เลื่อนหน้าต่างเคลื่อนที่จากแถวหนึ่งไปยังอีกแถวหนึ่ง
ดัดแปลงจาก : Lillesand and Kiefer, 1994



รูปที่ 2.2 แสดงความถี่เชิงพื้นที่แบบง่าย (ก) ความถี่เชิงพื้นที่เป็นศูนย์ เป็นภาพที่ทุกจุดภาพมีค่าเท่ากัน (ข) ความถี่เชิงพื้นที่ต่ำเป็นภาพที่ค่อย ๆ มีการเปลี่ยนแปลงของค่าหรือโทนสี (ค) ความถี่เชิงพื้นที่สูง เป็นภาพที่มีการตัดกันของโทนสีขาวดำอย่างชัดเจน หรือมีการเปลี่ยนโทนสีอย่างฉับพลัน

สามารถแบ่งแบบง่ายได้เป็น 3 แบบ 1. ความถี่เชิงพื้นที่เป็นศูนย์ 2. ความถี่เชิงพื้นที่ต่ำ 3. ความถี่เชิงพื้นที่สูง ดังรูปที่ 2.2 สำหรับเครื่องกรองผ่านความถี่ต่ำนั้น บริเวณใดมีความถี่สูงกว่ากำหนดจะถูกเคอร์เนล กรองให้ลดความถี่ลง ส่งผลให้ภาพที่มีความถี่สูงนั้นมีความนุ่มนวลมากขึ้น ภาพเบลอกว่าภาพเดิม มักถูกใช้เพื่อการลบหรือลดสัญญาณรบกวน หรือจุดต่าง (noise) ซึ่งหากเคอร์เนลมีขนาดใหญ่ขึ้น ภาพจะยิ่งเบลอมากขึ้น เนื่องจากตามหลักการแล้ว จุดภาพตรงกลางนั้นถูกคำนวณจากจุดภาพจำนวนมากขึ้น ตัวอย่างการกรองเช่นนี้มีหลากหลาย ตัวอย่างจากฟังก์ชันของ OpenCV เช่น

1. ฟังก์ชัน `cv2.filter2D()` คือทำคอนโวลูชันโดยใช้ค่าในตัวกรองทุกช่องเป็นค่าเท่ากัน รวมกันเป็น 1 ซึ่งคล้ายกับฟังก์ชัน `cv2.blur()` ต่างที่ ฟังก์ชัน `cv2.blur()` เพียงกำหนดขนาดของตัวกรอง ก็จะถูกสร้างให้มีค่าเท่ากันและรวมกันเป็น 1 เช่น กำหนดเคอร์เนลหรือตัวกรองที่ใช้คือ เมตริก 5×5 โดยทุกค่าในเมตริกนี้เป็น $1/25$ ซึ่งมีค่ารวมในเคอร์เนลนี้เท่ากับหนึ่ง ($0.04 \times 25 = 1$) [9]
2. ฟังก์ชัน `cv2.GaussianBlur()` ใช้ตัวกรองเป็นฟังก์ชันเกาส์ โดยมีค่าขนาดใหญ่ที่สุดตรงกลาง และเล็กลงเรื่อย ๆ โดยใช้สัญลักษณ์ σ (ซิกมา) เพื่อกำหนดค่าว่าจะลดลงเท่าใด ยิ่งค่า σ น้อยก็จะยิ่งลดลงมาก หาก σ มีค่าสูงมากผลที่ได้ก็จะใกล้เคียงกับการใช้ `cv2.blur()` ซึ่งมีความหมายถึงการเฉลี่ยให้ค่าทุกตัวบนตัวกรองเท่ากันทั้งหมด
3. ฟังก์ชัน `cv2.bilateralFilter()` หรือเรียกว่าตัวกรองทวิภาคี (bilateral filter) หลักการคำนวณโดยพื้นฐานคือใช้ฟังก์ชันเกาส์ เพียงแต่พิจารณาเฉพาะบริเวณที่ใกล้เคียงกันเท่านั้น ใช้สำหรับกรณีที่ต้องการให้ภาพเบลอในขณะที่แยกส่วนขอบออกจากกัน (ดังเช่นงานวิจัยนี้) เนื่องจากการใช้ตัวกรองเกาส์ ผลลัพธ์ที่ได้จะพบว่าบริเวณ

ขอบแบ่ง ตำแหน่งที่สีมีการตัดเปลี่ยนไปค่อนข้างกะทันหัน จะเบลอมากกว่า เพราะเกิดการผสมของสีบริเวณที่ต่างกันมาก

4. ฟังก์ชัน `cv2.medianBlur()` หรือตัวกรองมัธยฐาน (median filter) มักใช้เพื่อกำจัด ส่วนปนเปื้อนในภาพ ทำให้ภาพดูเรียบขึ้น โดยการหาค่ามัธยฐานจากจุดทั้งหมด ภายในบริเวณตัวกรอง (ค่ามัธยฐานคือค่าของตัวที่มีค่าอยู่ในลำดับกึ่งกลางเมื่อนำ ข้อมูลทั้งหมดที่พิจารณาเรียงต่อกัน)

2) เครื่องกรองผ่านความถี่สูง (high-pass filter) คือการกรองที่สกัดเอาองค์ประกอบของ ภาพที่มีความแปรปรวนเชิงพื้นที่สูง ซึ่งหมายถึงค่าของกลุ่มพิกเซลที่อยู่ใกล้เคียงกันที่มีความแปรปรวนอย่างฉับพลัน ดังรูปที่ 2.2 (ค) การประยุกต์ใช้เทคนิคการกรองภาพแบบ เครื่องกรองผ่านความถี่สูง เช่น การหาขอบภาพ (Edge detection) ซึ่งเป็นเทคนิคหนึ่งในการประมวลผลภาพ ใช้เพื่อหาขอบเขตในภาพที่ได้จากการทำให้ขอบของภาพมีความเด่นชัดขึ้น โดยเกิดจากความแตกต่างของความเข้มของแสงจากจุดภาพหนึ่งไปอีกจุด หนึ่งที่มีความต่อเนื่องกัน หากความต่างนี้มีค่ามาก จะทำให้เห็นขอบภาพได้อย่างชัดเจน ซึ่งวิธีการหาขอบภาพแบ่งได้เป็น 2 ประเภทหลัก ประเภทแรก คือ เกรเดียนต์เมทอด เป็นการหาขอบโดยการหาจุดต่ำสุดและจุดสูงสุดในรูปแบบของอนุพันธ์อันดับหนึ่งของภาพ โดยจุดที่เป็นขอบจะอยู่ในส่วนที่เหนือค่าเกณฑ์ (Threshold) จึงอาจทำให้เส้นขอบที่ได้มีลักษณะ หนา ประกอบด้วยหลายวิธี เช่น โซเบล (Sobel) แคนนี่ (Canny) โรเบิร์ต (Robert) เป็นต้น โดยงานศึกษานี้ได้เลือกใช้วิธีแคนนี่ และการหาขอบภาพประเภทที่สอง คือลาปลาเซียน เมทอด (Laplacian Methods) เป็นการหาขอบเขตภาพโดยการใช้อนุพันธ์อันดับ 2 [10] คือการให้ความแตกต่างระหว่างจุดภาพตรงกลางกับจุดภาพรอบข้างมาก ใช้ได้ดีในกรณีที่ ข้อมูลที่มีขอบแบบความถี่เชิงพื้นที่ที่มีความลาดชัน สำหรับในทางปฏิบัตินั้น เรื่องการหา ขอบภาพ มักพบปัญหา จากภาพที่สักรบกวนที่ทำให้ภาพหาขอบได้ยาก เช่น สักรบกวน หรือจุดต่างของภาพ จึงต้องใช้เครื่องกรองผ่านความถี่ต่ำก่อน เพื่อลดสักรบกวนดังกล่าว แล้วจึงจะหาเส้นขอบ โดยใน Opencv มีฟังก์ชันที่หาเส้นขอบหลายฟังก์ชันตามวิธีที่ ต้องการใช้ เช่น `Canny()` `Sobel()` `Laplacian()` และ `Scharr()`

2.2.3 การประมวลผลภาพกับรูปร่างและโครงสร้างของภาพ (Morphological Image Processing)

คือการประมวลผลภาพโดยการเปลี่ยนแปลงลักษณะรูปร่างหรือโครงสร้างของภาพ เป็นการดำเนินการพื้นฐานในการประมวลผลภาพ คำว่าสัณฐานวิทยาหมายถึงการดำเนินการที่ วิเคราะห์รูปร่าง โดยมีจุดประสงค์ต่าง ๆ เช่น เติมลงในรูเล็ก ๆ ขจัดสัญญาณรบกวน หรือสกัดหา เส้นเค้าโครง เป็นต้น การดำเนินการทางสัณฐานวิทยาแตกต่างจากกระบวนการคอนโวลูชัน คือ การดำเนินการนี้จะเป็นเชิงตรรกะ ซึ่งคือ และ (AND) หรือ (OR) และไม่ (NOT) มากกว่าที่จะเป็น การคำนวณเช่นเดียวกับกระบวนการคอนโวลูชัน การดำเนินการทางตรรกะจะใช้กับ 0 หรือ 1 พิก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เซลเท่านั้น ดังนั้นเทคนิคทางสัณฐานวิทยามักจะประมวลผลภาพไบนารีเป็นส่วนใหญ่ ภาพไบนารีคือภาพที่มีความเข้มแค่ 2 ค่า คือ 0 และ 1 ดังนั้นแต่ละพิกเซลจึงมีค่าได้เพียง 2 ค่านี้ โดยที่หากพิกเซลใดที่มีค่าเป็น 0 พิกเซลนั้นจะแสดงสีดำ และหากพิกเซลใดที่มีค่าเป็น 1 ก็จะแสดงสีขาว อย่างไรก็ตามวิธีการบางอย่างของการดำเนินการทางสัณฐานวิทยาสามารถขยายไปสู่ภาพระดับสีเทาได้เช่นกัน การดำเนินการทางสัณฐานวิทยาดั้งเดิม ประกอบด้วย การขยายขนาด (Dilation) การกัดเซาะ (Erosion)

1) การขยายขนาด (Dilation) เป็นการทำงานพื้นฐานในการทำสัณฐานวิทยา (Morphology) กับภาพไบนารี (Binary Image Dilation) เป็นการขยายให้กับภาพพื้นหน้า (Foreground) หรือคือการเพิ่มส่วนสีขาวหรือขนาดของวัตถุเบื้องหน้า [11] จึงมีประโยชน์มากในการหาขอบของภาพเนื่องจากเป็นวิธีที่ทำให้ขอบของภาพชัดเจนยิ่งขึ้น

2) การกัดเซาะ (Erosion) เป็นลักษณะของการลบข้อมูลภาพบริเวณขอบของภาพ หรือลดความเข้มแสงบริเวณขอบของภาพตามขนาดของเคอร์เนล (Kernel) ถ้าหากใช้กับภาพไบนารีจะเป็นการเปลี่ยนค่าจาก 1 เป็น 0 ในตำแหน่งของขอบเขตของวัตถุที่ไม่ถูกปิดล้อมโดยสมบูรณ์ ซึ่งคือการลดสีขาว ขยายสีดำ ดังนั้นความหนาหรือขนาดของวัตถุจะลดลง (หรือพื้นที่ที่มีสีขาวก็จะลดลง) [11] ซึ่งเป็นประโยชน์ในการลดสัญญาณรบกวนขนาดเล็กหรือจุดต่างในภาพ และใช้กับการแยกวัตถุที่เชื่อมต่อกัน

2.2.4 การหาเค้าโครงรูปร่าง (Contour detection)

เป็นอีกเทคนิคสำคัญในเรื่องการมองเห็นด้วยคอมพิวเตอร์ (Computer Vision) ซึ่งสามารถระบุและแยกขอบเขตของวัตถุในภาพได้ เค้าโครงรูปร่างคือเส้นโค้งที่เชื่อมจุดที่ต่อเนื่องกันทั้งหมดตามแนวขอบเขตซึ่งมีสีหรือความเข้มเท่ากัน [12] โดยปกติแล้ววัตถุจะมีพิกเซลที่มีความเข้มเท่ากันหรือใกล้เคียงกัน ส่งผลให้สามารถแยกแยะวัตถุจากพื้นหลังได้ [13]

ความแตกต่างที่สำคัญระหว่างการหาขอบภาพและการหาเค้าโครงรูปร่าง คือ การเปลี่ยนแปลงพิกเซลอย่างกะทันหันทุกครั้งสามารถเป็นขอบของภาพได้ ในขณะที่ขอบด้านนอกสุดเท่านั้นจึงจะเป็นเค้าโครงรูปร่าง และข้อจำกัดเกี่ยวกับการหาขอบภาพคือ ไม่มีเอกลักษณ์ (identity) ไม่มีความต่อเนื่อง (continuity) ไม่สามารถนับได้และไม่สามารถดำเนินการต่อได้ง่าย ไม่สามารถแยกออกจากกันและกันได้ และไม่มีรูปร่างที่ชัดเจนโดยตัวเอง ซึ่งแตกต่างจากเค้าโครงรูปร่าง ที่มีคุณสมบัติคือเป็นโครงร่างที่แสดงหรือล้อมรอบรูปร่างของบางสิ่ง รูปร่างได้รับการออกแบบโดยใช้ขอบภาพ และมีคุณสมบัติคือเป็นขอบที่มีเอกลักษณ์ เป็นพารามิเตอร์ทางเรขาคณิต และมีความต่อเนื่อง ซึ่งส่งผลให้มีประโยชน์สำหรับการวิเคราะห์รูปร่างและการจดจำวัตถุ [14]

สำหรับการนำเรื่องการหาเค้าโครงรูปร่างไปใช้ สามารถใช้ฟังก์ชัน findContours() ของ

ไลบรารี opencv มาช่วยในการหาโครงร่างได้ โดยมีอาร์กิวเมนต์ที่สำคัญ 3 ตัว ประกอบด้วย 1. เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปภาพใบนารี 2. วิธีการดึงข้อมูล ซึ่งประกอบด้วย 4 วิธี คือ CV_RETR_EXTERNAL ดึงเฉพาะรูปทรงที่อยู่ด้านนอก CV_RETR_LIST ดึงข้อมูลรูปทรงทั้งหมดโดยไม่ต้องสร้างความสัมพันธ์แบบลำดับชั้น CV_RETR_CCOMP ดึงรูปทรงทั้งหมดและมีการจัดให้เป็นลำดับชั้น (hierarchy) สองระดับ และ CV_RETR_TREE ดึงข้อมูลรูปทรงทั้งหมดและสร้างโครงสร้างลำดับชั้นที่สมบูรณ์ของรูปทรงที่ซ้อนกัน 3. วิธีการประมาณรูปร่าง ประกอบด้วย CHAIN_APPROX_NONE หมายถึงให้เก็บทุกจุดคอนทัวร์ และ CHAIN_APPROX_SIMPLE คือการบีบอัดส่วนแนวนอน แนวตั้ง และแนวทแยง และเหลือเพียงจุดสิ้นสุดเท่านั้น

2.2.5 การแบ่งส่วนรูปภาพ (Image segmentation)

คือหนึ่งในงานย่อยของการประมวลผลภาพ ซึ่งหมายถึงการแบ่งภาพออกเป็นส่วนๆ เช่น การแบ่งแยกวัตถุที่สนใจออกจากพื้นหลัง เป็นต้น โดยสามารถแบ่งประเภท ออกเป็น 2 ประเภทใหญ่ ดังนี้

1) Partial Segmentation เป็นการแบ่งภาพออกเป็นส่วนๆ โดยไม่ได้แบ่งแยกออกเป็นวัตถุ ก. หรือ ข. หรือพื้นหลังอย่างชัดเจน แต่จะแบ่งตามคุณสมบัติที่เรากำหนดขึ้น เช่น สี หรือความสว่าง ตัวอย่างเช่น การแบ่งภาพตามเข้มของแสง คือการแบ่งภาพออกเป็น 2 ส่วน คือ บริเวณที่เป็นวัตถุ และบริเวณที่เป็นพื้นหลัง เรียกว่า Thresholding หรือการแบ่งภาพตามสี คือ การแบ่งแยกวัตถุที่เราสนใจออกจากพื้นหลังโดยเลือกเฉพาะบริเวณที่มีสีเดียวเดียวกับวัตถุ

2) การแบ่งส่วนรูปภาพแบบสมบูรณ์ (Complete Segmentation) เป็นการแบ่งภาพออกเป็นส่วน ๆ ตามวัตถุบนโลกจริง หรือตามสิ่งที่เราเห็น โดยไม่มีส่วนที่ซ้อนทับกัน ซึ่งต้องใช้ข้อมูลที่เฉพาะเจาะจงในการแบ่งแยกวัตถุ ตัวอย่างเช่น การแบ่งภาพออกเป็น เรือ สะพาน ภูเขา เป็นต้น ตัวอย่าง การทำแบบ Complete Segmentation เช่น

1. การตรวจจับวัตถุ (Object Detection)
2. การให้ความหมายแต่ละจุดภาพ (Semantic Segmentation)
3. การแบ่งส่วนตัวอย่าง (Instance Segmentation)

ซึ่งหากพิจารณาเปรียบเทียบข้อแตกต่างของทั้ง 3 รูปแบบของการแบ่งส่วนรูปภาพแบบสมบูรณ์ มีดังนี้

การตรวจจับวัตถุ (Object detection) สามารถวาดกล่องรอบ ๆ วัตถุและติดป้ายว่าเป็นวัตถุชนิดใด กล่าวคือนอกจากการจัดประเภทหรือการจำแนกรูปภาพแล้ว ยังสามารถแบ่งแยกวัตถุที่สนใจออกจากพื้นหลัง ระบุตำแหน่งของวัตถุเพื่อทราบพื้นที่ที่วัตถุครอบคลุม ซึ่งแตกต่างจากการจำแนกข้อมูลรูปภาพ (Image Classification) ที่ทำได้เพียงบอกว่าวัตถุนั้นเป็นวัตถุอะไร โดยที่ไม่สามารถวาดกล่องรอบ ๆ วัตถุนั้น ๆ ได้ หรือในอีกความหมายคือไม่ทราบได้ว่าวัตถุนั้นอยู่ส่วนใดของภาพ

การให้ความหมายแต่ละจุดภาพ (Semantic Segmentation) นั้นสามารถรู้ละเอียดถึงระดับจุดภาพ ว่าแต่ละจุดภาพในภาพนั้น เป็นจุดภาพที่อยู่ในวัตถุหมวดใดบ้าง ซึ่งเป็นการแบ่งกลุ่มระดับถัดไปของการตรวจจับวัตถุ (Object detection) กล่าวคือเป็นการแทนที่กล่องที่ถูกร่างขึ้นรอบ ๆ วัตถุ ด้วยการวาดขอบเขตรอบวัตถุทุกชิ้น โดยที่สามารถกำหนดสีให้กับทุกหมวดได้ ผลที่ได้คือภาพที่มีการแบ่งเป็นพื้นที่สี และแต่ละสีหมายความว่าถึงหมวดต่าง ๆ ในรูปภาพ ซึ่งการแบ่งส่วนประเภทนี้มีประโยชน์มากในการขับเคลื่อนอัตโนมัติ หรือมักถูกใช้สำหรับการฝึกอบรมแบบจำลองการรับรู้วัตถุสภาพแวดล้อมทางธรรมชาติ

การแบ่งส่วนตัวอย่าง หรือการแบ่งส่วนจำเพาะวัตถุ (Instance Segmentation) มีการเพิ่มเติมจากการให้ความหมายแต่ละจุดภาพ (Semantic Segmentation) ที่เป็นเพียงการแสดงผลภาพวัตถุหลายชิ้นในหมวดหมู่เดียวกันเป็นเอนทิตีเดียวหรือของชิ้นเดียว แต่สำหรับการแบ่งส่วนตัวอย่าง (Instance Segmentation) เป็นการจำแนกวัตถุออกจากกันโดยจำเพาะว่าเป็นวัตถุที่ 1, 2 หรือ 3 กล่าวคือ มีความสามารถในการบอกแยกรายชิ้น (Instance) ของวัตถุในแต่ละหมวดได้ [15] ตัวอย่างเช่น หากต้องการแบ่งแยกต้นไม้หลายต้นในรูปภาพ วิธีการให้ความหมายแต่ละจุดภาพนั้น ต้นไม้ทั้งหมดจะถูกถมเป็นสีเดียวกัน และรวมกันว่าเป็นต้นไม้ แต่วิธีการแบ่งส่วนตัวอย่าง จะสามารถแบ่งแยกต้นไม้แต่ละต้นได้เป็นรายชิ้น โดยต้นไม้ทั้งหมดจะมีสีหรือป้ายกำกับที่ต่างกัน ซึ่งในงานวิจัยนี้จะมุ่งเน้นไปที่การแบ่งส่วนตัวอย่าง ที่อยู่ในส่วนของการแบ่งส่วนรูปภาพแบบสมบูรณ์ (Complete Segmentation)

สำหรับก่อนการมีการเรียนรู้เชิงลึกนั้น การแบ่งส่วนรูปภาพต้องสกัดเอาคุณลักษณะเด่นแบ่งภาพตามเฉดสี ตามการตัดกันของสี หรือตามขอบ ซึ่งในปัจจุบันสามารถใช้วิธีการเรียนรู้เชิงลึกในการฝึกสอนตั้งแต่ต้นจนจบกระบวนการ [15]

2.3 แนวคิดและทฤษฎีเกี่ยวกับการเรียนรู้เชิงลึก

เรียนรู้เชิงลึกคือเทคนิคในการสร้างปัญญาประดิษฐ์โดยใช้โครงข่ายประสาทเทียม (Artificial Neural Networks) หรือข่ายงานประสาทเทียมหลาย ๆ ชั้น (Layer) เป็นหนึ่งในเทคนิคของการทำเหมืองข้อมูล (Data Mining) ซึ่งคือโมเดลทางคณิตศาสตร์ สำหรับประมวลผลสารสนเทศด้วยการคำนวณแบบคอนเนกชันนิสต์ (Connectionist) [16] ด้วยวัตถุประสงค์ที่จะสร้างเครื่องมือซึ่งมีความสามารถในการเรียนรู้การจดจำรูปแบบ (Pattern Recognition) และการสร้างความรู้ใหม่ (Knowledge Extraction) เป็นการเลียนแบบการทำงานของโครงข่ายประสาทของมนุษย์ (Neurons) เพื่อให้มีความสามารถเช่นเดียวกับที่มีในสมองมนุษย์

ในการทำงานของการเรียนรู้เชิงลึก (Deep Learning) นี้จะเหมาะกับการทำสิ่งที่ซับซ้อนมาก มีข้อมูลจำนวนมาก ตัวอย่างเช่น การแยกแยะวัตถุ (Object recognition) ระบบจดจำใบหน้า โดยระบบโครงข่ายประสาทนี้จะถูกแบ่งเป็น 3 ชั้น ได้แก่ ชั้นข้อมูลเข้า (Input layer) ชั้นซ่อน (Hidden layer) เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

และชั้นข้อมูลออก (Output layer) การทำงานจะเป็นรับข้อมูล ประมวลผล และได้ผลลัพธ์ตามลำดับ ในแต่ละชั้น จะมีสิ่งที่เรียกว่าสถานีเชื่อมโยง (Node) ซึ่งในชั้นซ่อน (Hidden layer) จะเป็นชั้นที่สำคัญที่สุด เนื่องจากการประมวลผลจากการใช้ค่าน้ำหนัก (Weight) และ ฟังก์ชันกระตุ้น (Activation functions) วิธีคิดของการเรียนรู้เชิงลึกจะแตกต่างไปจากคอมพิวเตอร์ทั่ว ๆ ไป ในส่วนของการหาวิธีการแก้ปัญหาด้วยการให้ข้อมูลและสอนว่าผลลัพธ์นั้นจะเป็นอย่างไร แต่คอมพิวเตอร์ทั่ว ๆ ไปจะเป็นการให้ข้อมูลตามด้วยวิธีการที่มนุษย์ต้องคิดขึ้นมาทุกครั้งเพื่อให้ได้ผลลัพธ์

การเรียนรู้เชิงลึกสามารถแบ่งคร่าวๆ เป็นสองประเภท คือ โครงข่ายประสาทแบบคอนโวลูชัน (Convolutional Neural Network: CNN) ซึ่งเป็นโครงข่ายประสาทเทียมที่มองชุดข้อมูลที่ได้รับผ่านชั้นข้อมูลป้อนเข้าเหมือนภาพภาพหนึ่ง เปรียบเทียบได้กับจอประสาทตาของมนุษย์ ที่มีการรับแสงที่ตกกระทบมาจากวัตถุต่าง ๆ อีกประเภทคือโครงข่ายประสาทแบบวนกลับ (Recurrent Neural Network: RNN) คือ โครงข่ายประสาทเทียมที่รับชุดข้อมูลผ่านข้อมูลป้อนเข้าแบบอนุกรมเวลา (Time Series) หรือข้อมูลที่มีลักษณะเป็นลำดับ (Sequence) เป็นวิธีการที่ถูกนำมาใช้ในการวิจัยเกี่ยวกับการรู้จำเสียง (Speech recognition) และการประมวลผลภาษาธรรมชาติ (Natural Language Processing) ซึ่งในที่นี้จะขอกล่าวถึงเพียงโครงข่ายประสาทแบบคอนโวลูชันที่มีลักษณะเหมาะกับข้อมูลประเภทรูปภาพ

2.3.1 แนวคิดและทฤษฎีเกี่ยวกับโครงข่ายประสาทแบบคอนโวลูชัน

งานการแยกประเภทรูปภาพ (Image Classification) ด้วยการเรียนรู้เชิงลึก ตัวแบบที่นิยมใช้คือโครงข่ายประสาทแบบคอนโวลูชัน ซึ่งคือ โครงข่ายประสาทหลายชั้นที่มีโครงสร้างเฉพาะตัว เป็นรูปแบบสถาปัตยกรรมหนึ่งของโครงข่ายประสาทแบบป้อนไปข้างหน้า (Feed-forward neural networks) ที่ซึ่งข้อมูลสามารถผ่านหน่วยประมวลผลได้เพียงทางเดียว ไม่ได้นำข้อมูลผลลัพธ์มาใช้ซ้ำ จัดเป็นการเรียนรู้เชิงลึกเช่นกัน โดยร่วมมือระหว่างชั้นคอนโวลูชัน (Convolutional Layer) และ ชั้นการพูลลิงแบบค่าสูงสุด (Max-pooling Layer) เพื่อบีบและค้นจนได้ข้อมูลขนาดเล็กพอที่เป็นลักษณะเด่นที่สำคัญ (Feature) เพื่อจะเป็นตัวแทนของภาพ

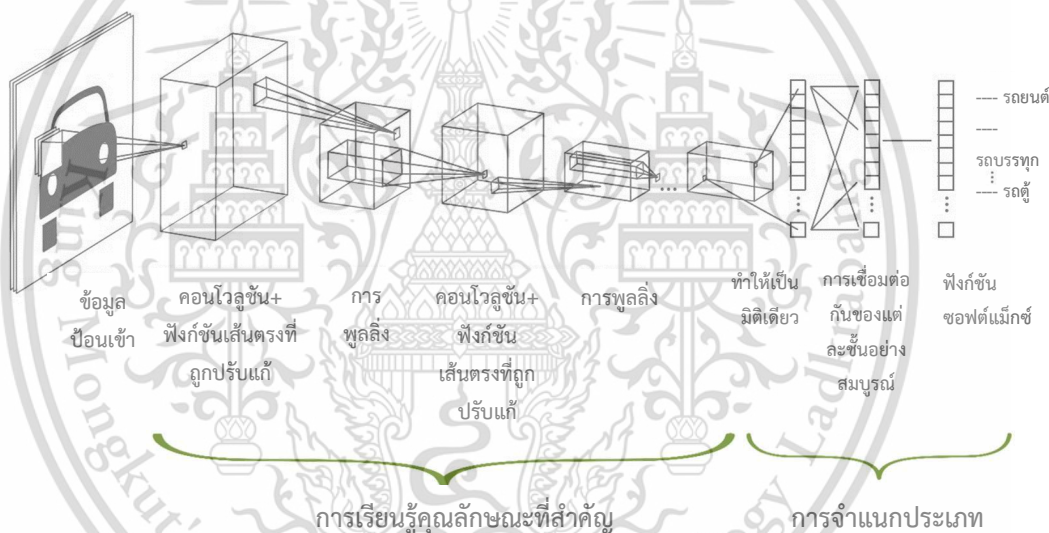
โครงข่ายประสาทแบบคอนโวลูชันนั้นตอบโจทย์ปัญหาประเภทการรับรู้ (Perceptual tasks) อย่างมาก [17] มักจะถูกใช้เพื่อการสกัดลักษณะเด่นที่สำคัญจากข้อมูลประเภทที่ไม่ค่อยเป็นระเบียบหรือไม่ได้มีโครงสร้างเป็นรูปแบบเฉพาะตัว (Unstructured data) [17] อย่างเช่นรูปภาพ เป็นต้น โดยจะจำลองการมองเห็นของมนุษย์ในพื้นที่ย่อย ๆ จากการแยกแยะคุณลักษณะเชิงภาพ เช่น สี ลายเส้น และอื่น ๆ จากนั้นจึงค่อยนำสิ่งที่สกัดได้ที่เป็นเวกเตอร์ (Feature vector) มาผ่านกระบวนการสุดท้ายในการจำแนกประเภทว่าเป็นภาพวัตถุในหมวดใด จากนั้นนำมาผสานกันเพื่อทำนายว่าเป็นภาพอะไร

โครงสร้างของโครงข่ายประสาทแบบคอนโวลูชันโดยปกติประกอบด้วย 2 ส่วน คือ ชั้นฐานแบบคอนโวลูชัน (Convolutional Base Layer) ซึ่งจะเป็นส่วนที่มีหน้าที่ในการสกัดลักษณะเด่น

ที่สำคัญของภาพ (Feature Extraction) และชั้นจำแนกประเภท (Classifier Layer) มีหน้าที่ในการแยกประเภทภาพ ดังแผนภาพที่ 2.3

สำหรับในการทำงานของโครงข่ายประสาทแบบคอนโวลูชันจะมี 4 กระบวนการ [18] คือ

1. คอนโวลูชัน (Convolution) จุดประสงค์ของคือเพื่อต้องการระบุคุณลักษณะที่สำคัญและเกี่ยวข้องกับภาพ โดยขั้นตอนนี้จะทำการสร้าง ตัวกรอง (Filter หรือ Sliding window) มาสแกนรูปภาพเพื่อแยกองค์ประกอบต่างๆ เช่น รูปทรงของเส้นขอบ สี โดยปกติภาพจะมี 3 สีหลัก คือ สีแดง สีน้ำเงิน และสีเขียว แบ่งเป็น 3 ช่องสัญญาณ (Channel) ซึ่งแต่ละจุดภาพสามารถแทนค่าด้วยตัวเลขเพื่อบอกความเข้มของสี ตั้งแต่ 0-255 ในการทำภาพขาวดำจะใช้เพียง 1 ช่องสัญญาณ สีดำแทนด้วยเลข 1 และสีขาวแทนด้วยเลข 0 จากนั้นจะใช้ตัวกรอง ทำการคอนโวลูชันกับภาพขาวดำ เพื่อเก็บค่าไว้ในเมทริกซ์ชุดใหม่ที่เรียกว่าคอนโวลูฟเจอร์ (Convolved Feature) หรือ ฟังก์ชันคุณลักษณะ (Feature map) [18]



รูปที่ 2.3 โครงสร้างของโครงข่ายประสาทแบบคอนโวลูชัน

(แปลงจาก: <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>)

2. ฟังก์ชันเส้นตรงที่ถูกปรับแก้ (Rectified Linear Unit หรือ ReLU) เป็นฟังก์ชันกระตุ้น (Activation functions) จุดประสงค์ของคือเพื่อแปลงเมทริกซ์พีเอเจอร์แมพ จากข้อ 1 ให้อยู่ในรูปที่ไม่มีลักษณะเป็นเชิงเส้น (nonlinear) และช่วยแก้ปัญหาการไล่ระดับสีที่หายไป (gradient vanishing) โดยจะทำการแทนที่ค่าในจุดภาพที่เป็นลบด้วย 0 หลังจากทำ ReLU แล้วค่าที่ได้จะเป็นค่าบวกเท่านั้นส่วนใหญ่นิยมใช้ฟังก์ชัน ซอฟต์แมกซ์ (Softmax) เพื่อความง่ายในการคำนวณและประสิทธิภาพของผลลัพธ์ [18]

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3. การพูลลิ่ง (Pooling) เป็นการลดมิติของฟังก์ชันลักษณะให้มีขนาดเล็กลงแต่ยังคงรักษารายละเอียดข้อมูลสำคัญของภาพไว้ [18] การพูลลิ่งมีประโยชน์ในเรื่องความไวในการคำนวณและแก้ปัญหาเกินพอดี (overfitting) การพูลลิ่งสามารถจำแนกเป็นประเภทต่างๆ ได้เช่น พูลลิ่งด้วยค่าสูงสุด (Max Pooling) ค่าเฉลี่ย (Mean Pooling)

4. ชั้นเชื่อมต่อกันของแต่ละชั้นอย่างสมบูรณ์ (Fully Connected Layer) จากกระบวนการทั้ง 3 ขั้นตอนที่กล่าวไว้ข้างต้นคือ

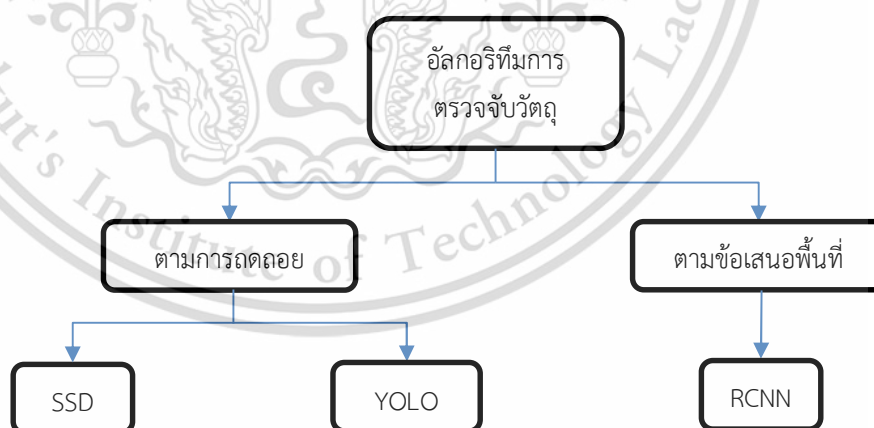
1) คอนโวลูชัน (Convolution) ระบุคุณลักษณะที่สำคัญ ๆ และเกี่ยวข้องกับภาพ

2) Rectified Linear Unit (ReLU) แปลงเมทริกซ์ชุดใหม่หรือฟังก์ชันลักษณะ

3) การพูลลิ่ง (Pooling) ลดมิติของฟังก์ชันลักษณะให้มีขนาดเล็กลงแต่ยังคงรักษารายละเอียดข้อมูลสำคัญของภาพไว้ ในขั้นตอนนี้จะเป็นการทำซ้ำตั้งแต่ขั้นตอนคอนโวลูชันถึงการพูลลิ่งจนกว่าจะเกิดการเชื่อมต่อกันของแต่ละชั้นอย่างสมบูรณ์ จุดประสงค์ของขั้นตอนนี้เพื่อจะนำลักษณะเด่นที่สำคัญของรูปที่ได้ จากกระบวนการก่อนหน้านี้มาสร้างเป็นโครงข่ายประสาทสำหรับการเรียนรู้และทำนายประเภทของรูปภาพโดยจะคัดกรองรูปที่รับเข้ามาและจัดให้อยู่ในรูปของประเภท (Classes) และผลลัพธ์ที่ได้จะแสดงค่าความเชื่อมั่น (Confident) ในทำนายประเภทของรูปภาพ [18]

2.3.2 การตรวจจำวัตถุ และการแบ่งส่วนตัวอย่าง

การเรียนรู้เชิงลึกถูกนำมาใช้อย่างกว้างขวางสำหรับงานตรวจจำวัตถุ โดยหลักสามารถแบ่งได้เป็นสองประเภท ดังรูปที่ 2.4 มีรายละเอียดดังนี้

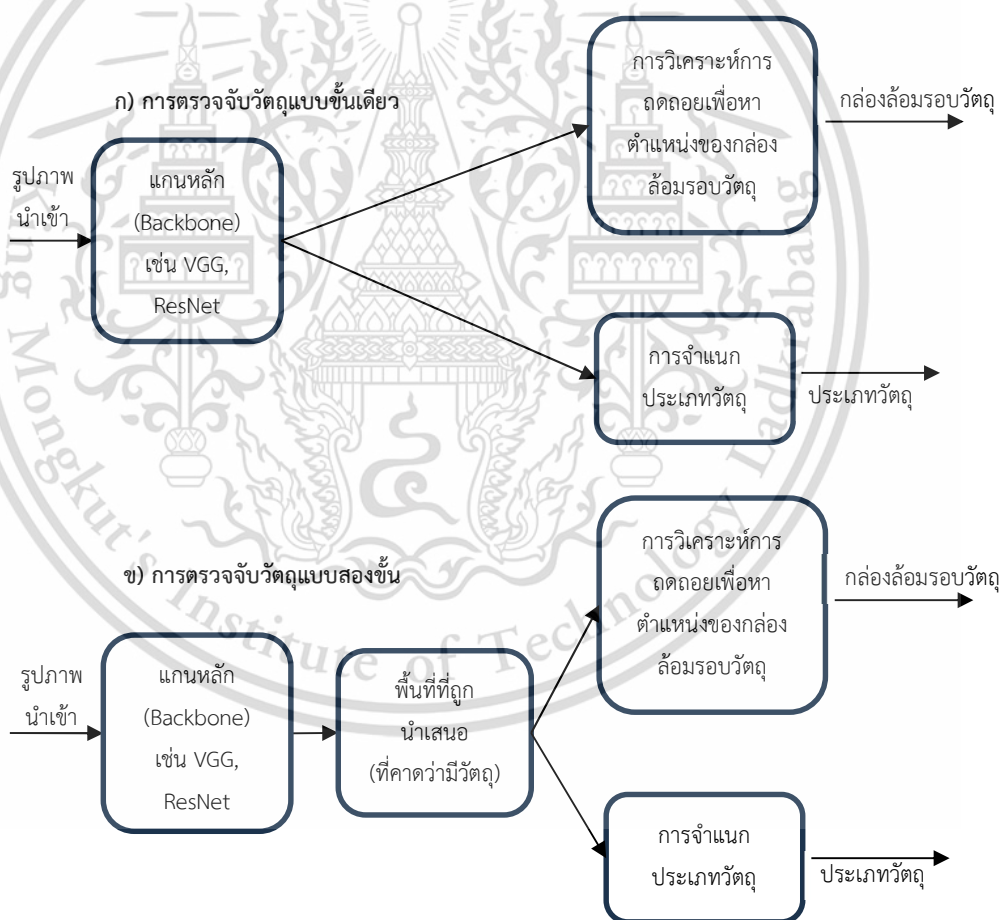


รูปที่ 2.4 การแบ่งประเภทของการตรวจจำวัตถุ

แปลงจาก : <https://www.mdpi.com/1424-8220/23/13/5843>

1. การตรวจจำวัตถุแบบขั้นเดียว (One-stage object detection หรือ Single-stage object detection) ซึ่งจะมองว่างานการตรวจจำวัตถุเป็นปัญหาการถดถอย เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(Regression problem) [19] การตรวจจับวัตถุประเภทนี้คือการหาตำแหน่งและจัดประเภทของวัตถุที่ปรากฏในภาพภายใต้การส่งผ่านข้อมูลภาพเข้าสู่กระบวนการ (pipeline) เพียงครั้งเดียว [20] กล่าวคือสามารถทำนายตำแหน่งและขนาดกรอบล้อมรอบวัตถุ รวมถึงความน่าจะเป็นของประเภทวัตถุภายในรูปภาพออกมาพร้อมกัน ดังรูปที่ 2.5 ประโยชน์ของการตรวจจับวิธีนี้คือเพื่อแก้ไขด้านความเร็วของการฝึกอบรมและการอนุมาน (inference) [21] ตัวอย่างตัวแบบเช่น Single Shot MultiBox Detector (SSD) และ You Only Look Once (YOLO) สกัดคุณลักษณะ ในชั้นคุณลักษณะแบบคอนโวลูชัน (convolutional feature layers) จึงจะทำนายผล และเนื่องจาก SSD ไม่ได้มีขั้นตอนการเลือกพื้นที่ที่สนใจ เช่นเดียวกับ YOLO จึงทำให้ความเร็วในการตรวจจับเพิ่มขึ้นเมื่อเทียบกับตัวตรวจจับแบบสองขั้นตอน [21] สำหรับรายละเอียดของแบบจำลอง YOLO มีการกล่าวถึงในส่วนถัดไป



รูปที่ 2.5 ความแตกต่างของอัลกอริทึมของการตรวจจับวัตถุทั้งสองประเภท
แปลงจาก : Semantic Image Cropping (2021) by Oriol Corcoll

2. การตรวจจับวัตถุแบบสองชั้น (Two-stage object detection) คือแบบจำลองที่ตรวจจับ ระบุ และบ่งบอกถึงตำแหน่งของวัตถุโดยผ่านเครือข่ายประสาทเทียมสองรอบ [28] โดยการทำงานคือการเริ่มจากการระบุตำแหน่งหรือพื้นที่ที่สนใจเป็นอันดับแรก แล้วจะถูกจำแนกออกเป็นประเภทวัตถุในภายหลัง เช่น แบบจำลอง R-CNN (Regions with CNN features) รายละเอียดคือการนำเอาโครงข่ายประสาทแบบคอนโวลูชันมาช่วยงานด้านการตรวจจับวัตถุ แบบจำลองนี้ถูกพัฒนาขึ้นในปี พ.ศ. 2557 โดยทีมวิจัยบริษัท ไมโครซอฟท์ (Microsoft Research) ซึ่งตัวแบบนี้เป็นการผสมกันระหว่าง 1. อัลกอริทึมที่นำเสนอพื้นที่ที่คาดว่าจะมีวัตถุ (region proposal algorithms) ซึ่งช่วยให้ตัวแบบแบ่งประเภทรูปภาพในขั้นตอนที่สองได้แม่นยำยิ่งขึ้น และ 2. โครงข่ายประสาทแบบคอนโวลูชัน (CNNs) เพื่อตรวจจับและระบุตำแหน่งของวัตถุในรูป ดังรูปที่ 2.5 การตรวจจับวัตถุแบบสองชั้นนี้สามารถทำงานได้ดีในงานตรวจจับวัตถุต่างๆ แต่การฝึกต้องใช้เวลามาก และความเร็วในการตรวจจับนั้นช้ากว่าการตรวจจับวัตถุแบบชั้นเดียว

วิธีการตรวจจับวัตถุที่ใช้การเรียนรู้เชิงลึกที่ประสบความสำเร็จในช่วงแรก คือแบบจำลอง R-CNN (Regions with CNN features) ซึ่งเป็นตัวตรวจจับแบบสองชั้นตอน จากนั้นจึงมีการพัฒนาจึงเกิดเป็นการตรวจจับวัตถุแบบชั้นเดียว ข้อแตกต่างหลักระหว่างตัวตรวจจับแบบชั้นเดียว และตัวตรวจจับแบบสองชั้นตอน คือ เน็ตเวิร์คที่นำเสนอพื้นที่ที่คาดว่าจะมีวัตถุ (Region Proposal Network : RPN) เนื่องจากว่าทั้งสองการตรวจจับนั้นนำรูปภาพนำเข้ามามีผลผ่านเน็ตเวิร์คแกนหลักเพื่อสกัดเอาคุณสมบัติที่สำคัญเหมือนกัน จากนั้นหากเป็นการตรวจจับแบบสองชั้น ขั้นตอนถัดมาคือการดึงเอาพื้นที่ที่สนใจที่คาดว่าจะมีวัตถุ แต่สำหรับการตรวจจับแบบชั้นเดียว การตรวจจับจะทำผ่านเน็ตเวิร์คเดียวโดยที่ไม่ต้องดึงเอาพื้นที่ที่สนใจที่คาดว่าจะมีวัตถุออกมาก่อน สำหรับการตรวจจับวัตถุ มักจะมีการแลกเปลี่ยนระหว่างความเร็วและความแม่นยำ [22] เครื่องตรวจจับแบบชั้นตอนเดียวมีความเร็วในการตรวจจับสูง ซึ่งเหมาะสำหรับการเฝ้าระวัง การขับขี่แบบอัตโนมัติ การตรวจจับป้ายทะเบียน ฯลฯ ในทางตรงกันข้าม เครื่องตรวจจับแบบสองชั้นตอนสามารถให้ความถูกต้องหรือความแม่นยำที่ดีกว่า เช่น การจดจำใบหน้า ลายนิ้วมือ และการจดจำรูปแบบ ฯลฯ [23] ดังนั้นการเลือกใช้วิธีการตรวจจับวัตถุแต่ละประเภทจะขึ้นกับความเหมาะสมของงาน

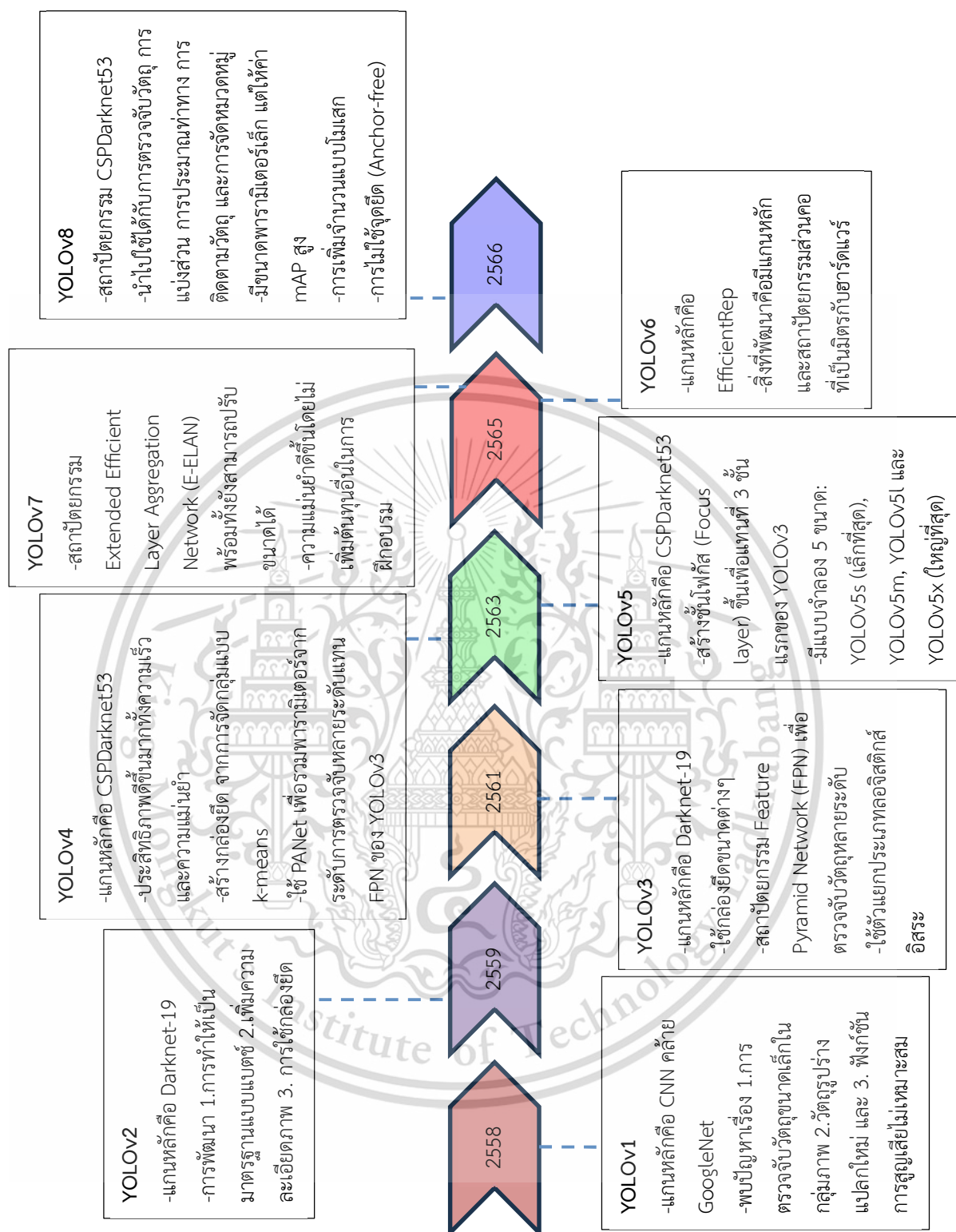
2.3.3 การพัฒนาแบบจำลอง YOLO

YOLO หรือ You Only Look Once เสนอการใช้โครงข่ายประสาทเทียมจากต้นจนจบ ในการทำงานรอบขอบเขตวัตถุและความน่าจะเป็นของประเภทวัตถุทั้งหมดในครั้งเดียวด้วยความช่วยเหลือของเลเยอร์ที่เชื่อมต่ออย่างสมบูรณ์เพียงชั้นเดียว [24] ในขณะที่แบบจำลองที่ใช้เน็ตเวิร์คที่นำเสนอพื้นที่ที่คาดว่าจะมีวัตถุจะต้องดำเนินการวนซ้ำหลายครั้งสำหรับรูปภาพเดิม [24]

โดยการพัฒนาของแบบจำลองตั้งแต่ YOLOv1 จนถึง YOLOv8 ดังรูปที่ 2.6

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.6 แสดงการพัฒนาของแบบจำลองตั้งแต่ YOLOv1 จนถึง YOLOv8

แบบจำลอง YOLO (2558) ใช้โครงข่ายประสาทเทียมเชิงลึกอย่างง่ายเพื่อตรวจจับวัตถุในภาพ โดยสถาปัตยกรรมของแบบจำลองโครงข่ายประสาทแบบคอนโวลูชันที่เป็นแกนหลักนั้นเหมือนกับสถาปัตยกรรม GoogleNet [25] มีปัญหาเรื่องการตรวจจับวัตถุขนาดเล็กภายในกลุ่มเอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพ เช่น กลุ่มบุคคล เนื่องจากสถาปัตยกรรมถูกออกแบบสำหรับการตรวจจับวัตถุแบบขึ้นเดียว ทั้งยังไม่สามารถตรวจจับวัตถุที่ไม่รูปทรงที่แตกต่างออกไป และมีฟังก์ชันการสูญเสียที่ไม่เหมาะสมเนื่องจากให้ค่าความผิดพลาดระหว่างกล่องล้อมรอบวัตถุขนาดเล็กและขนาดใหญ่เท่ากัน

YOLOv2 หรือ YOLO9000 (2559) สถาปัตยกรรมโครงข่ายประสาทแบบคอนโวลูชันที่เป็นแกนหลัก ถูกเปลี่ยนเป็น Darknet-19 ซึ่งเป็นรูปแบบหนึ่งของสถาปัตยกรรม VGGNet และการปรับปรุงหลักอย่างหนึ่ง คือการใช้กล่องยึด (Anchor box) ซึ่งคือชุดกรอบที่กำหนดไว้ล่วงหน้า ด้วยขนาด อัตราส่วนกว้างยาว และตำแหน่งเฉพาะ กล่องยึดของ YOLOv2 มีขนาดเท่ากันทั้งหมด ซึ่งช่วยให้อัลกอริธึมสามารถจัดการกับขนาดวัตถุและอัตราส่วนภาพได้กว้างขึ้น นอกจากนี้ สิ่งที่เปลี่ยนแปลงเพิ่มคือ YOLOv2 ใช้การทำให้เป็นมาตรฐานแบบแบตช์ (Batch Normalization) เป็นวิธีในการทำการปรับขนาดข้อมูล (Data Scaling) เพื่อปรับค่าข้อมูลให้อยู่ในขอบเขตที่กำหนด ซึ่งช่วยปรับปรุงความแม่นยำและความเสถียรของแบบจำลอง ทั้งนี้ยังใช้ฟังก์ชันการสูญเสีย (Loss function) ใหม่ที่เหมาะสมกับงานการตรวจจับวัตถุมากขึ้น ซึ่งคือผลรวมของข้อผิดพลาดกำลังสอง (Sum of the squared errors) ระหว่างกล่องขอบเขตรอบวัตถุที่แบบจำลองทำนาย และความจริง และความน่าจะเป็นของประเภทวัตถุนั้น ๆ

YOLOv3 (2561) ใช้สถาปัตยกรรมโครงข่ายประสาทแบบคอนโวลูชันใหม่ที่เรียกว่า Darknet-53 ซึ่งเป็นอีกรูปแบบหนึ่งของสถาปัตยกรรม ResNet และได้รับการออกแบบมาโดยเฉพาะสำหรับงานตรวจจับวัตถุ Darknet-53 นั้นมีใหญ่กว่า เร็วกว่า และแม่นยำกว่ามากเมื่อเทียบกับ Darknet-19 ซึ่งเป็นสถาปัตยกรรมแกนหลักของ YOLOv2 การปรับปรุงอีกประการหนึ่งคือใช้กล่องยึดที่มีขนาดและอัตราส่วนภาพที่แตกต่างกัน โดยถูกปรับขนาดเพื่อให้เข้ากับขนาดและรูปร่างของวัตถุที่ตรวจพบได้ดียิ่งขึ้น และได้ใช้สถาปัตยกรรม Feature Pyramid Network (FPN) เป็นสถาปัตยกรรมของโครงข่ายประสาทแบบคอนโวลูชันที่ใช้ในการตรวจจับวัตถุหลายระดับ โดยสร้างพีระมิดจากแผนที่คุณลักษณะ โดยแต่ละระดับของพีระมิดจะถูกใช้ในการตรวจจับวัตถุในระดับที่แตกต่างกัน ซึ่งจะช่วยปรับปรุงประสิทธิภาพการตรวจจับวัตถุขนาดเล็กได้ดียิ่งขึ้น นอกจากนี้สำหรับฟังก์ชันที่ใช้ทำนายประเภทวัตถุเปลี่ยนจากการใช้ฟังก์ชัน Softmax ของ YOLOv2 เป็นตัวแยกประเภทลอจิสติกส์อิสระ (independent logistic classifiers) ซึ่งจะช่วยให้การทำนายกรณีวัตถุซ้อนทับกัน เนื่องจากฟังก์ชัน Softmax มีข้อจำกัดคือแต่ละกล่องจะมีเพียงวัตถุประเภทเดียวซึ่งไม่จริงเสมอไป

YOLOv4 (2563) พัฒนาโดย Bochkovskiy และคณะ โดยใช้สถาปัตยกรรมโครงข่ายประสาทแบบคอนโวลูชันใหม่ที่เรียกว่า CSPDarknet53 (Cross Stage Partial Network) เป็นอีกรูปแบบหนึ่งของสถาปัตยกรรม ResNet สิ่งที่ปรับเปลี่ยนเพิ่มเติมคือการสร้างกล่องยึดโดยใช้การจัดกลุ่มแบบ k-means ซึ่งคือใช้จุดศูนย์กลางของกลุ่ม (centroid) เป็นกล่องยึด ซึ่งช่วยให้กล่องยึดนั้นใกล้เคียงกับขนาดและรูปร่างของวัตถุที่ตรวจพบได้มากขึ้น นอกจากนี้ยังเปลี่ยนจากเอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การใช้ FPN ของ YOLOv3 เป็น PANet เพื่อใช้รวมพารามิเตอร์จากการตรวจจับหลายระดับ ทั้งนี้ YOLOv4 ได้ใช้ฟังก์ชันการสูญเสีย GHM เป็นตัวแปรหนึ่งของฟังก์ชันการสูญเสียแบบโฟกัส (Focal loss)

YOLOv5 (2563) พัฒนาโดย Ultralytics ใช้สถาปัตยกรรม CSPDarknet53 เช่นเดียวกับ YOLOv4 หัวข้อหลักที่พัฒนาขึ้นคือการสร้างชั้นโฟกัส (Focus layer) ขึ้นเพื่อแทนที่ 3 ชั้นแรกของ YOLOv3 ซึ่งการบูรณาการนี้มีวัตถุประสงค์เพื่อลดจำนวนชั้น ลดพารามิเตอร์ ลดดำเนินการจุดลอยต่อวินาที (FLOPS) ลดหน่วยความจำ เพิ่มความเร็วในขณะที่ยังคงผลกระทบต่อค่า mAP น้อยที่สุด และ YOLOv5 ยังได้ปรับปรุงเรื่องการรวมกลุ่มพีระมิดเชิงพื้นที่ (Spatial Pyramid Pooling : SPP) จาก YOLOv4 ให้มีประสิทธิภาพมากขึ้น ซึ่งการรวมกลุ่มพีระมิดเชิงพื้นที่นี้เป็นขั้นตอนการรวมกลุ่มประเภทหนึ่งที่ใช้เพื่อลดความละเอียดเชิงพื้นที่ของแผนที่คุณลักษณะ โดยถูกใช้เพื่อปรับปรุงประสิทธิภาพการตรวจจับบนวัตถุขนาดเล็ก เนื่องจากช่วยให้แบบจำลองมองเห็นวัตถุได้หลายระดับ ทั้งยังได้นำเสนอฟังก์ชันการสูญเสีย CIOU ซึ่งเป็นรูปแบบหนึ่งของฟังก์ชันการสูญเสีย พื้นที่ทับซ้อนเหนือพื้นที่ทั้งหมด (IoU) ที่ออกแบบมาเพื่อปรับปรุงประสิทธิภาพของโมเดลบนชุดข้อมูลที่ไม่สมดุล (Imbalanced Datasets)

YOLOv6 (2565) ไม่ได้เป็นส่วนหนึ่งของกลุ่ม YOLO อย่างเป็นทางการ แต่สถาปัตยกรรมแกนหลักได้อิทธิพลมาจากอัลกอริทึม YOLO แบบขั้นเดียว สิ่งที่พัฒนาจาก YOLOv5 คือมีแกนหลักและสถาปัตยกรรมส่วนคอที่เป็นมิตรกับฮาร์ดแวร์ สถาปัตยกรรมส่วนหัวที่มีหน้าที่คำนวณผลลัพธ์นั้นใช้แบบแยกส่วนที่มีประสิทธิภาพ และกลยุทธ์การฝึกอบรมที่มีประสิทธิภาพมากขึ้น โดยแกนหลักที่มีหน้าที่สกัดคุณสมบัติที่สำคัญ แล้วทำเป็นแผนที่คุณลักษณะ (Feature Map) นั้นใช้ EfficientRep ซึ่งช่วยให้หน่วยประมวลผลทำได้มีประสิทธิภาพ สถาปัตยกรรมส่วนคอที่มีหน้าที่ดึงลักษณะพิเศษบางอย่างออกมานั้นใช้ Rep-PAN

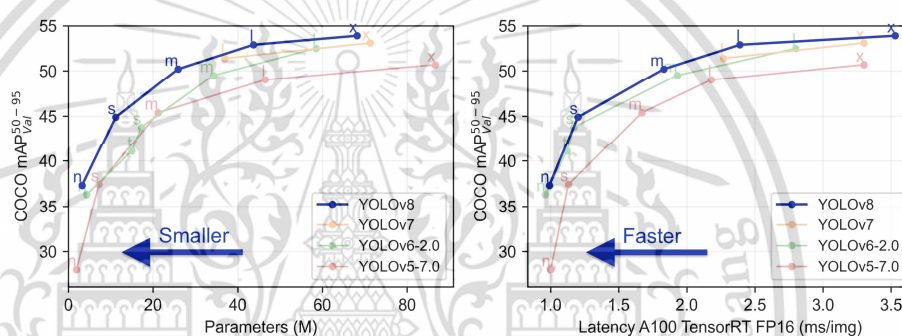
YOLOv7 มีการเปลี่ยนแปลงหลัก คือ สถาปัตยกรรม ซึ่งใช้เป็น Extended Efficient Layer Aggregation Network (E-ELAN) ซึ่งจะช่วยให้โมเดลเรียนรู้คุณสมบัติที่หลากหลายมากขึ้น ส่งผลต่อการเรียนรู้ที่ดีขึ้น ทั้งยังสามารถปรับขนาดสถาปัตยกรรมได้ ซึ่งช่วยให้แบบจำลองสามารถตอบสนองความต้องการในเรื่องความเร็วในการอนุมานที่แตกต่างกันได้ การพัฒนาหลักอีกอย่างคือการปรับปรุงความแม่นยำของแบบจำลองโดยไม่ต้องเพิ่มต้นทุนอื่นในการฝึกอบรม และความเร็วในการตรวจจับไม่ลดลง ซึ่งถูกใช้คำว่าTrainable bag-of-freebies จึงเป็นเหตุผลที่ YOLOv7 มีความเร็วในการอนุมาน และความแม่นยำเพิ่มขึ้น สำหรับเรื่องอื่น ๆ ของ YOLOv7 คือ ใช้กล่องยึด 9 กล่อง ซึ่งช่วยให้ตรวจจับรูปร่างและขนาดของวัตถุได้หลากหลายขึ้น จึงช่วยลดจำนวนผลบวกหลง (False Positive) ลง ทั้งยังมีการปรับปรุงฟังก์ชันการสูญเสียใหม่ซึ่งคือฟังก์ชันการสูญเสียแบบโฟกัส (focal loss) ช่วยแก้ปัญหาการตรวจจับวัตถุขนาดเล็กให้มีประสิทธิภาพมากยิ่งขึ้น โดยลดน้ำหนัก (weight) การสูญเสียสำหรับตัวอย่างที่ได้รับการจำแนกประเภทอย่างดี และมุ่งเน้นไปที่ตัวอย่างที่ยากซึ่งเป็นวัตถุที่ตรวจพบได้ยาก นอกจากนี้ YOLOv7 ยังมีภาพที่มี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ความละเอียดสูงขึ้น โดยจะประมวลผลภาพที่ความละเอียด 608×608 พิกเซล ซึ่งยิ่งเพิ่มให้แบบจำลองสามารถตรวจจับวัตถุขนาดเล็กและมีความแม่นยำโดยรวมสูงขึ้น ข้อดีหลักประการหนึ่งของ YOLOv7 คือความเร็ว สามารถประมวลผลภาพในอัตรา 155 เฟรมต่อวินาที ซึ่งเร็วกว่าอัลกอริธึมการตรวจจับวัตถุที่ล้ำสมัยอื่นๆ มาก

YOLOv8 ถูกพัฒนาโดยทีม Ultralytics รองรับการคอมพิวเตอร์วิทัศน์อย่างเต็มรูปแบบ ซึ่งรวมถึงการตรวจจับวัตถุ การแบ่งส่วน การประมาณท่าทาง การติดตามวัตถุ และการจัดหมวดหมู่ จากการแสดงประสิทธิภาพของแบบจำลองเปรียบเทียบกับวิธีการตรวจจับวัตถุแบบเรียลไทม์อื่น ๆ บนชุดข้อมูล COCO ซึ่งเป็นชุดข้อมูลมาตรฐาน พบว่าแบบจำลองตรวจจับวัตถุ YOLOv8 มีขนาดพารามิเตอร์เล็กกว่า ซึ่งส่งผลต่อความเร็วในการประมวลผล (หน่วย: มิลลิวินาทีต่อรูป) ที่รวดเร็วกว่า แต่ให้ค่า mAP ที่สูงกว่าแบบจำลองอื่น ดังรูปที่ 2.7



รูปที่ 2.7 การเปรียบเทียบประสิทธิภาพระหว่างแบบจำลอง YOLOv8 และแบบจำลองอื่น

ที่มา : <https://github.com/ultralytics/ultralytics>

สำหรับการเปลี่ยนแปลงหลักของแบบจำลอง YOLOv8 ประกอบด้วย

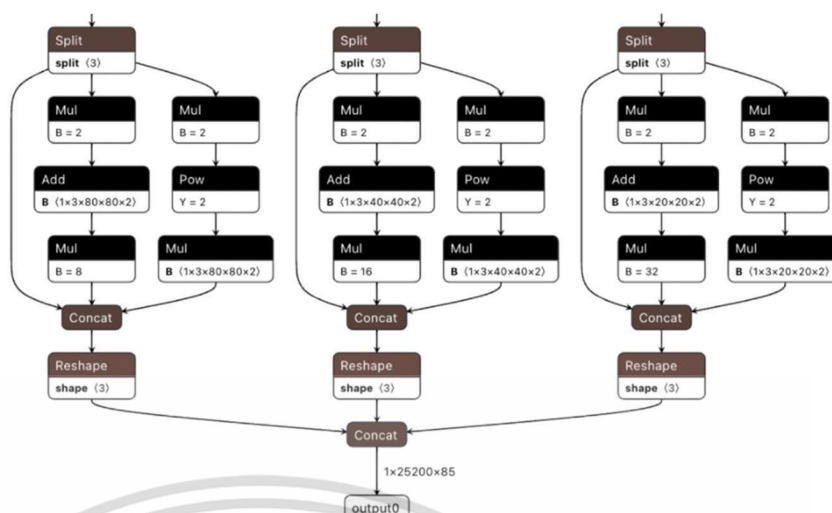
1. การเพิ่มรูปภาพ (Augmentation) ระหว่างการฝึกสอน เป็นส่วนสำคัญของความสำเร็จ

โดยในแต่ละรอบของการฝึกสอน แบบจำลองจะมองเห็นรูปภาพที่มีรูปแบบแตกต่างกันเล็กน้อย ซึ่งวิธีการนี้เรียกว่าการเพิ่มจำนวนแบบโมเสก (Mosaic augmentation) ซึ่งเป็นการเพิ่มจำนวนโดยการต่อภาพสี่ภาพเข้าด้วยกัน เพื่อบังคับให้แบบจำลองเรียนรู้วัตถุในตำแหน่งใหม่ในการบดบังบางส่วน โดยจากการทดลองเพื่อให้ได้ประโยชน์สูงสุด จะต้องไม่ใช้การเพิ่มจำนวนแบบโมเสกดังกล่าวใน 10 รอบสุดท้ายของการฝึกอบรม

2. การไม่ใช้จุดยึด (Anchor-free) กล่าวคือระบบจะคาดการณ์หรือทำนายจุดศูนย์กลางของวัตถุโดยตรง แทนการเริ่มต้นจากจุดยึดดังกล่าวที่กำหนดค่าไว้ เนื่องจากข้อเสียจุดยึดคือเป็นส่วนที่ซับซ้อนและยุ่งยากอย่างมากสำหรับแบบจำลอง YOLO อื่น ๆ

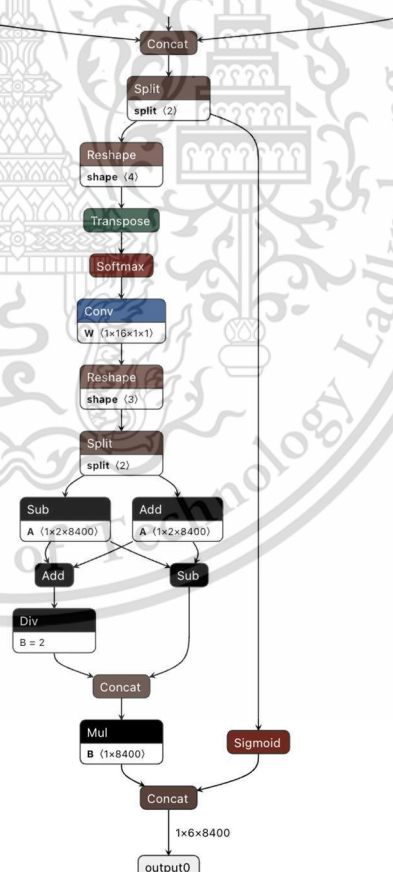
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.8 ส่วนโครงสร้างหัวการตรวจจับ YOLOv5 ที่ถูกแสดงผ่าน netron.app

ที่มา : <https://blog.roboflow.com/whats-new-in-yolov8/>

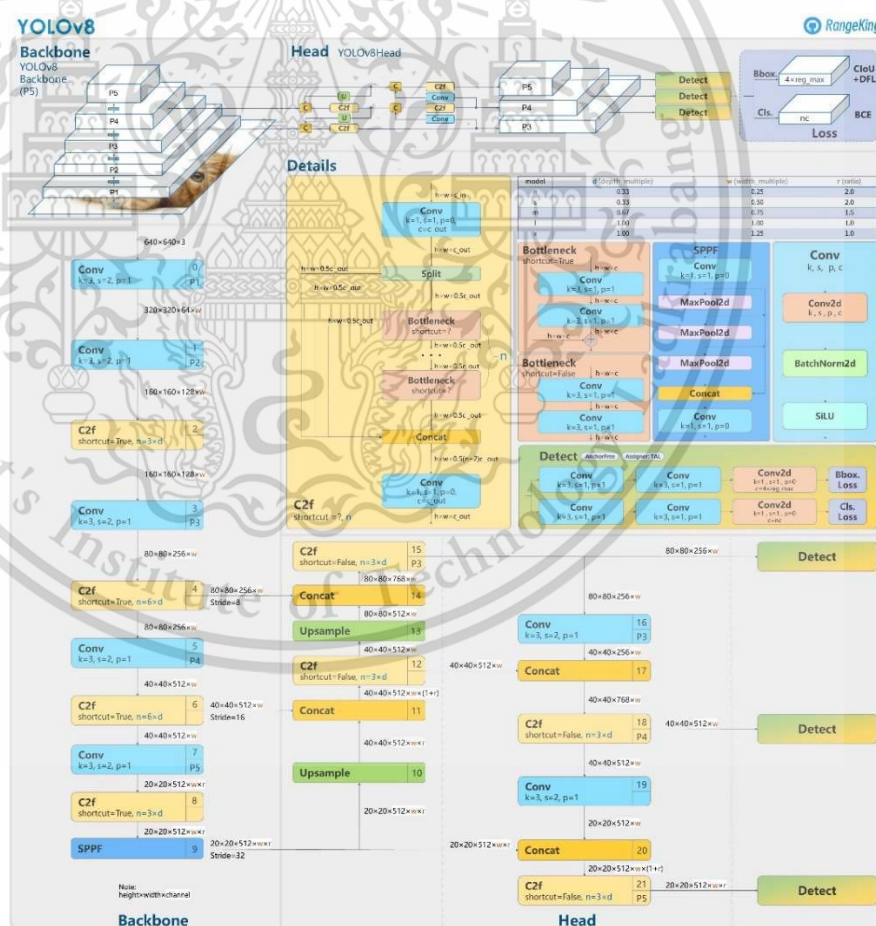


รูปที่ 2.9 ส่วนโครงสร้างหัวการตรวจจับ YOLOv8 ที่ถูกแสดงผ่าน netron.app

ที่มา : <https://blog.roboflow.com/whats-new-in-yolov8/>

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

และเป็นสิ่งที่ถูกกำหนดค่ามาจากการกระจายตัวของกล่องขอบเขตรอบวัตถุของชุดข้อมูลเป้าหมายที่เป็นเกณฑ์มาตรฐานในการวัดประสิทธิภาพแบบจำลอง จากรูปที่ 2.8 คือสถาปัตยกรรมในส่วนหัวการตรวจจับ (Detection head) ของ YOLOv5 พบว่าเมื่อเปรียบเทียบกับในส่วนหัวการตรวจจับของ YOLOv8 ดังรูปที่ 2.9 พบว่าแบบจำลอง YOLOv8 จะช่วยลดจำนวนกล่องล้อมรอบวัตถุที่ถูกคาดการณ์ ซึ่งจะช่วยเร่งความเร็วในขั้นตอนการพิจารณาขอบที่เป็นไปได้ (Non-Maximum Suppression : NMS) ซึ่งเป็นขั้นตอนที่ซับซ้อนหลังจากการประมวลผล โดยทั่วไปในการตรวจจับวัตถุนั้น กล่องขอบเขตล้อมรอบรูปภาพหลายกล่องจะถูกสร้างขึ้นสำหรับวัตถุเพียงชิ้นเดียวในรูปภาพ กล่องเหล่านี้อาจทับซ้อนกันหรืออยู่ในตำแหน่งที่แตกต่างกัน แต่ทั้งหมดแสดงถึงวัตถุเดียวกัน ดังนั้น NMS ใช้เพื่อระบุและลบกล่องขอบเขตที่ซ้ำซ้อนหรือไม่ถูกต้อง โดยพิจารณาพร้อมกับค่าอัตราส่วนพื้นที่ทับซ้อนเหนือพื้นที่รวม และส่งออกเพียงกล่องขอบเขตเดียวซึ่งเป็นผลลัพธ์ที่ดีที่สุดสำหรับแต่ละวัตถุในรูปภาพ



เล่มงานวิจัยของแบบจำลอง YOLOv8 นี้ยังไม่ได้ถูกตีพิมพ์อย่างเป็นทางการ แต่อย่างไรก็ตาม ผู้ใช้ชื่อ RangeKing จาก GitHub ได้ทำสรุปโครงสร้างสถาปัตยกรรมของแบบจำลองการตรวจจับวัตถุ YOLOv8 ดังรูปที่ 2.10

แบบจำลองการแบ่งส่วนตัวอย่างของ YOLOv8 ที่ใช้ในงานวิจัยนี้ เป็นส่วนขยายจากแบบดั้งเดิม ซึ่งคือแบบจำลองการตรวจจับวัตถุที่กล่าวมาข้างต้น โดยโครงสร้างหลักของส่วนขยายนี้ใช้ CSPDarknet53 เป็นตัวสกัดคุณสมบัติ แล้วตามด้วยโมดูล C2f เพื่อแทนสถาปัตยกรรมในส่วนคอของ YOLOv8 แบบดั้งเดิม จากนั้นจะเป็นสถาปัตยกรรมในส่วนหัวถูกแบ่งสองส่วนเพื่อใช้ในการเรียนรู้ที่จะทำนายมาส์กรูปภาพนำเข้า แบบจำลองการแบ่งส่วนตัวอย่างนี้มีส่วนหัวที่ใช้ตรวจจับวัตถุเหมือนแบบดั้งเดิม คือประกอบด้วยโมดูลการตรวจจับหัวโมดูล ตามด้วยชั้นการทำนาย (prediction layer) [26] จุดเด่นของ YOLOv8 คือมีค่าความแม่นยำ (accuracy) สูง ง่ายต่อใช้งานและนำไปการพัฒนา เนื่องจากมีแฟ้มเคจโปรแกรมภาษาไพทอน มีส่วนประสานงานกับผู้ใช้ประเภทคอมมานด์ไลน์ (Command-line user interface) ที่ใช้งานง่าย และแบบจำลอง YOLOv8 นั้นถูกใช้งานอย่างกว้างขวาง ทั้งนี้ยังมีแบบจำลองที่ได้รับการฝึกสอนแล้ว (Pre-trained Model) เพื่อให้ไปใช้งาน สำหรับการแบ่งส่วนตัวอย่าง แบบจำลองดังกล่าวได้รับการฝึกสอนล่วงหน้ากับชุดข้อมูล COCO ซึ่งเป็นชุดข้อมูลมาตรฐานขนาดใหญ่สำหรับการตรวจจับวัตถุ มีวัตถุทั้งหมด 80 ประเภท โดยมีทั้งหมด 5 แบบจำลองที่ได้รับการฝึกสอนแล้ว หากเรียงตามขนาดของพารามิเตอร์ของแบบจำลองจากน้อยไปหามาก ประกอบด้วย YOLOv8n-seg YOLOv8s-seg YOLOv8m-seg YOLOv8l-seg YOLOv8x-seg

2.4 การประเมินผล

สำหรับการประเมินผลแบบจำลองการตรวจจับวัตถุ นิยมใช้ความแม่นยำเฉลี่ย (Average Precision : AP) และค่าเฉลี่ยความแม่นยำเฉลี่ย (mean Average Precision : mAP) เป็นตัวชี้วัดประสิทธิภาพแบบจำลอง ซึ่งส่วนประกอบย่อยเพื่อใช้หาค่าเฉลี่ยของความแม่นยำเฉลี่ย มีรายละเอียดดังนี้

เมทริกซ์ความสับสน (Confusion Matrix) คือตารางเฉพาะเจาะจงที่นำผลลัพธ์จากการทำนายมาใส่ในรูปตารางเมทริกซ์ ดังตารางที่ 2.3 ซึ่งช่วยให้ง่ายต่อการมองเห็นค่าทำนายของอัลกอริทึม โดยในการสร้างเมทริกซ์ความสับสน จำเป็นต้องมีค่าของคุณลักษณะ 4 ประการ ดังนี้

ตารางที่ 2.3 เมทริกซ์ความสับสน

		การทำนาย	
		ผลบวก (1)	ผลลบ (0)
ผลเฉลยจริง	ผลบวก (1)	ผลบวกจริง (TP)	ผลลบจริง (FN)
	ผลลบ (0)	ผลบวกจริง (FP)	ผลลบจริง (TN)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้ทำซ้ำโดยไม่ได้รับอนุญาต

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1. ผลบวกจริง (True Positive : TP) คือจำนวนข้อมูลที่ถูกจำแนกได้อย่างถูกต้องในกลุ่มของประเภทที่ต้องการ (Positive class) หรือกลุ่มที่สนใจ
2. ผลลบจริง (True Negatives : TN) คือจำนวนข้อมูลที่ถูกจำแนกถูกต้องในกลุ่มประเภทที่ไม่ใช่เป้าหมาย (Negative class) หรือกลุ่มที่ไม่สนใจ
3. ผลบวกหลง (False Positive : FP) คือจำนวนข้อมูลที่แบบจำลองทำนายเป็นประเภทที่สนใจ แต่ผลเฉลยจริงไม่ใช่ประเภทที่สนใจ (ข้อผิดพลาดชนิดที่ 1)
4. ผลลบหลง (False Negative : FN) คือจำนวนข้อมูลที่ถูกจำแนกเป็นกลุ่มของประเภทที่ไม่สนใจ แต่ผลเฉลยจริงคือประเภทที่สนใจ (ข้อผิดพลาดชนิดที่ 2)

โดยค่าดังกล่าวใช้ในการคำนวณตัวชี้วัดประสิทธิภาพต่าง ๆ ที่เกี่ยวข้องกับค่าความแม่นยำเฉลี่ยดังต่อไปนี้

1. ค่าความเที่ยงตรง (Precision) คือ ความแม่นยำผลทำนาย โดยสนใจผลทำนาย จำนวนจากสัดส่วนผลการทำนายผลบวก (Positive) และ ผลลบ (Negative) โดยแสดงผลการทำนายจากร้อยละที่ทำนายถูกต้อง แสดงได้ดังสมการที่ (1)

$$Precision = \frac{TP}{(TP+FP)} = \frac{TP}{All\ Predict} \quad (1)$$

2. ค่าความครบถ้วน (Recall) คือการวัดความแม่นยำ ที่สนใจจากจำนวนข้อมูลที่ทำนายว่าเป็นผลบวกต่อจำนวนผลบวกจริง แสดงได้ดังสมการที่ (2)

$$Recall = \frac{TP}{(TP+FN)} = \frac{TP}{All\ Ground\ Truth} \quad (2)$$

3. อัตราความเที่ยงตรงเฉลี่ย (mean Average Precision: mAP) คำนวณได้จากการค้นหาความเที่ยงตรงเฉลี่ย (AP) สำหรับแต่ละหมวดหมู่ แล้วนำไปหาค่าเฉลี่ยจากทุกหมวดหมู่อีกครั้ง ดังสมการที่ 3 ซึ่งค่าความเที่ยงตรงเฉลี่ย (Average Precision) คือการนำค่าความเที่ยงตรง (Precision) และค่าความครบถ้วน (Recall) ทำเป็นกราฟ สำหรับพื้นที่ใต้กราฟคือค่าความเที่ยงตรงเฉลี่ย (Average Precision)

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \quad (3)$$

การใช้ mAP จะเกี่ยวข้องกับเรื่องอัตราส่วนพื้นที่ที่ทับซ้อนเหนือพื้นที่ทั้งหมด (Intersection over union : IoU) ซึ่งคือการคำนวณจากอัตราส่วนระหว่างพื้นที่ทับซ้อน

กันของทั้ง 2 กล้องล้อมรอบวัตถุ ประกอบด้วย กล้องล้อมรอบวัตถุผลเฉลย (Ground

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Truth) และกล่องล้อมรอบวัตถุที่ถูกทำนาย (Predict) หาดด้วยพื้นที่รวมของกล่องล้อมรอบวัตถุทั้งสอง ดังรูปที่ 3.8 โดยค่า mAP ที่ 0.5:0.95 คือ mAP ที่อัตราส่วนพื้นที่ทับซ้อนเหนือพื้นที่ทั้งหมด (IoU) มีค่าเท่ากับ 0.5 – 0.95 ซึ่งหมายถึงการเปรียบเทียบค่าของอัตราส่วนระหว่างพื้นที่ทับซ้อนกันต่อพื้นที่รวมทั้งหมดระดับฟิสิกเซลของกล่องล้อมรอบวัตถุที่คาดการณ์ เทียบกับกล่องล้อมรอบวัตถุจริง และค่าเกณฑ์ ใช้เพื่อพิจารณาว่าการทำนายกล่องล้อมรอบวัตถุนั้นถูกต้องหรือไม่

2.5 โปรแกรมภาษาไพธอน (Python)

ภาษาโปรแกรมไพธอน (Python) คือภาษาโปรแกรมคอมพิวเตอร์ระดับสูง ถูกออกแบบมาให้เป็นภาษาสคริปต์ที่อ่านง่าย โดยตัดความซับซ้อนของโครงสร้างและไวยากรณ์ของภาษาออก ทั้งยังถูกพัฒนาขึ้นมาโดยไม่ยึดติดกับแพลตฟอร์ม กล่าวคือสามารถรันภาษาไพธอนได้ทั้งบนระบบ Unix, Linux, Windows NT, Windows 2000, Windows XP หรือแม้แต่ระบบ FreeBSD การทำงานของภาษาไพธอนเป็นการแปลชุดคำสั่งที่ละบรรทัด (Interpreter) คือเพื่อป้อนเข้าสู่หน่วยประมวลผลให้คอมพิวเตอร์ นอกจากนั้นภาษาโปรแกรมไพธอน ยังสามารถนำไปใช้ในการเขียนโปรแกรมได้หลากหลายประเภท โดยไม่ได้จำกัดอยู่ที่งานเฉพาะทางใดทางหนึ่ง (General-purpose language) สำหรับงานวิจัยนี้ใช้ Google Colaboratory ซึ่งคือสภาพแวดล้อมสำหรับการพัฒนาแบบเบ็ดเสร็จ (Integrated Development Environment : IDE) ออนไลน์ที่ใช้สำหรับเขียนโปรแกรมภาษาไพธอนหรือภาษาอื่น ๆ บนคลาวด์ (Cloud) โดยไม่ต้องติดตั้งโปรแกรมใด ๆ เพิ่มเติม

ไพธอนไลบรารี (Python library) เปรียบเสมือนโปรแกรมสำเร็จรูปในที่เก็บฟังก์ชันการทำงานที่เฉพาะทาง กล่าวคือเป็นการนำคำสั่งต่างๆของไพธอนมาสร้างเป็นชุดคำสั่ง หรือเป็นแพ็คเกจ (Package) เพื่อให้สามารถทำงานตามวัตถุประสงค์แต่ละด้านได้อย่างมีประสิทธิภาพมากขึ้น ซึ่งประโยชน์ของไลบรารีสำเร็จรูปคือผู้ใช้งานสามารถนำไลบรารีที่ได้มีการพัฒนาไว้อยู่แล้วมาใช้งานได้เลย ไม่จำเป็นต้องสร้างขึ้นใหม่ทั้งหมด ไลบรารีส่วนใหญ่ของไพธอนทั้งที่ถูกพัฒนาโดยผู้พัฒนาหลักและผู้ใช้งานทั่วไปจะถูกสร้างและอัปโหลดขึ้นเก็บในเซิร์ฟเวอร์ของดัชนีแพ็คเกจไพธอน (Python Package Index : PyPI) หรือเซิร์ฟเวอร์ของคลาวด์ของอนาคอนด้า (Anaconda Cloud) โดยไลบรารีที่สำคัญที่ใช้สำหรับการพัฒนาแบบจำลอง ประกอบด้วย

2.5.1 ไลบรารีจัดการข้อมูลรูปภาพ (Image Manipulation Library)

1. Pillow ใช้ในการเปิด จัดการรูปภาพ เช่น การจัดการต่อพิกเซล การปกปิดและการจัดการความโปร่งใส การกรองภาพ เช่น การเบลอ การปรับเส้นขอบ การปรับให้เรียบ หรือการค้นหาขอบ การปรับปรุงภาพ และสามารถบันทึกรูปแบบไฟล์ต่าง ๆ

2. OpenCV ใช้สำหรับงานคอมพิวเตอร์วิทัศน์ (Computer Vision) และฟังก์ชันการประมวลผลภาพและวิดีโอเพื่อระบุรูปร่าง วัตถุ ข้อความ เป็นต้น ในงานวิจัยนี้นำมาใช้ในการอ่าน แสดง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

และจัดการรูปภาพ เช่น การปรับขนาดและการครอบตัด และการดำเนินการทางสัญญาณวิทยา การกรองและการปรับปรุงภาพ เช่น การเบลอบเบกาส์เซียน การตรวจจับขอบ และการทำให้คมชัดหาเส้นเคาโครง เป็นต้น

2.5.2 ไบรารีพัฒนาแบบจำลอง (Model Implementation Library)

1. Ultralytics ปรับปรุงงานต่างๆ เช่น การผลเฉลยชุดข้อมูล (คำอธิบายประกอบรูปภาพ) การฝึกอบรม การแสดงภาพ และการจัดการแบบจำลอง เช่นตรวจจับวัตถุ การติดตามหลายวัตถุ การแบ่งส่วนตัวอย่าง การประมาณท่าทาง และการจัดหมวดหมู่รูปภาพ สำหรับในงานวิจัยนี้ใช้คำสั่ง YOLO ในไลบรารีดังกล่าวเพื่อพัฒนาเป็นแบบจำลอง

2. MMDetection ใช้เป็นเครื่องมือในการสร้างแบบจำลองตรวจจับวัตถุ การแบ่งส่วนตัวอย่าง และวิธีการแบ่งส่วนแบบพิกเซล ทดลองจนส่วนประกอบและโมดูลที่เกี่ยวข้อง ซึ่ง MMDetection ประกอบด้วย 7 ส่วนหลัก ได้แก่ ส่วนต่อประสานโปรแกรมประยุกต์(API) โครงสร้าง (structure) ชุดข้อมูล(dataset) แบบจำลอง(model) กลไก(engine) การประเมินผล(evaluation) และการแสดงภาพ(visualization)

2.6 โปรแกรม Easyrice

โปรแกรม EasyRice คือโปรแกรมที่ให้บริการด้านการวิเคราะห์อัตลักษณ์และคัดแยกข้าวเปลือกสายพันธุ์ที่ต้องการออกจากสายพันธุ์อื่นผ่านรูปภาพสแกน โดยมีฐานข้อมูลจากศูนย์เมล็ดพันธุ์และศูนย์วิจัยข้าวทั่วประเทศซึ่งเป็นแหล่งข้อมูลที่น่าเชื่อถือ แบบจำลองมีความแม่นยำที่ 95% และระยะเวลาประมวลผลประมาณ 3-5 นาที [27] สำหรับงานวิจัยนี้ โปรแกรม Easyrice ดังกล่าว ได้ถูกใช้ในการระบุประเภทข้อมูล (Data Labeling) ในขั้นตอนการเก็บรวบรวมข้อมูล เพื่อจำแนกว่าข้าวเปลือกแต่ละเมล็ดเป็นข้าวสายพันธุ์ที่ต้องการหรือไม่ ทั้งนี้ยังถูกใช้เพื่อเปรียบเทียบประสิทธิภาพของแบบจำลองในด้านความเร็วอีกเช่นกัน

2.7 งานวิจัยที่เกี่ยวข้อง

จากทฤษฎีทั้งหมดที่กล่าวมานี้ จะนำมาประยุกต์ใช้กับการศึกษาการจำแนกสายพันธุ์ข้าวเปลือกไทยระหว่างข้าวหอมมะลิไทยและข้าวหอมปทุมธานี 1 โดยใช้การประมวลผลภาพ ดังนั้นงานวิจัยนี้จึงได้คัดเลือกงานวิจัยที่เกี่ยวข้องและมีความใกล้เคียงกับงานวิจัยที่กำลังศึกษา ดังตัวอย่างต่อไปนี้

2.7.1 งานวิจัยที่เกี่ยวข้องกับคุณภาพข้าวสาร

งานวิจัยประเภทนี้มีจุดประสงค์เพื่อจำแนกข้าวเสียหายหรือไม่ได้มาตรฐาน ตัวอย่างเช่น ข้าวเมล็ดหัก ข้าวเมล็ดเหลือง ข้าวท้องไข หรือข้าวเมล็ดลีบ เป็นต้น ออกจากข้าวเมล็ดดีประโยชน์เพื่อตรวจคุณภาพการสีข้าว หรือคุณภาพข้าวที่รับซื้อจากผู้ขาย

ตารางที่ 2.4 สรุปงานวิจัยที่เกี่ยวข้อง

ประเภท	ชื่องานวิจัย	เนื้อหาโดยย่อ	ตัวแบบที่ใช้	ผลที่ได้
คุณภาพข้าวสาร	Deep CNN-based damage classification of milled rice grains using a high-magnification image dataset	แยกประเภทเมล็ดข้าวสารที่เสียหายหลังผ่านกระบวนการสีข้าว โดยใช้ภาพที่มีกำลังขยายสูง (4.5 เท่า)	ตัวแบบ CNN ทั้งหมด 5 ตัวแบบ คือ efficientNet-50, ResNet-50, InceptionV3, MobileNetV2, MobileNetV3	EfficientNet-B0 มีค่าความแม่นยำ 98.37%
คัดแยกเมล็ดข้าวเปลือก	แบ่งข้อมูลภาพเมล็ดข้าวด้วยการเรียนรู้เชิงลึก	แบ่งแยกเมล็ดข้าวเปลือกออกจากพื้นหลัง ในรูปที่มีสภาพแวดล้อมต่างกัน และถ่ายโดยใช้กล้องโทรศัพท์มือถือ	ตัวแบบการเรียนรู้เชิงลึก 1. U-net 2. Mask R-CNN	U-net ทำงานได้ดีเมื่อสภาพแวดล้อมของรูปทดสอบเหมือนกับรูปฝึกสอน Mask R-CNN ทำงานได้ในบริบทที่หลากหลายกว่า
	Localization and Classification of Rice-grain Images Using Region Proposals-based Convolutional Neural Network	จำแนกข้าวเปลือก 5 พันธุ์ของไทยที่ข้อมูลมีความใกล้เคียงกันมาก โดยทดลองหลายสถานการณ์	1. Mask R-CNN ผ่านการฝึกสอนด้วยชุดข้อมูล COCO 2. วิธีซัพพอร์ตเวกเตอร์แมชชีน	ตัวแบบมีประสิทธิภาพที่สุดคือการแยกแยะหว่างข้าวเปลือกข้าวเหนียวและข้าวสาร ส่วนสถานการณ์แยก 5 สายพันธุ์ ทำได้ 75%
คัดแยกเมล็ดข้าวสาร	Identification of milled rice varieties using machine vision	ศึกษาลักษณะเด่น 3 กลุ่มใหญ่คือ สันฐานวิทยาพื้นผิว และสี เพื่อดูสัดส่วนใดที่ดีที่สุด	วิธีกำลังสองน้อยสุดบางส่วนหาความสัมพันธ์ระหว่างลักษณะเด่นและพันธุ์ข้าว	การใช้สันฐานวิทยาและพื้นผิวร่วมกันได้คะแนนมากที่สุด
	Contrastive analysis of rice grain classification techniques: multi-class support vector machine vs artificial neural network	จำแนกเมล็ดข้าวสาร 3 พันธุ์ ประกอบด้วย 1. Ponni 2. Basmati และ 3. Brown rice	เปรียบเทียบตัวแบบที่ตรงข้ามกัน 1. วิธีซัพพอร์ตเวกเตอร์แมชชีนแบบหลายคำตอบ 2. วิธีโครงข่ายประสาทเทียม	ทั้ง 2 ตัวแบบได้ผลดีวิธีโครงข่ายประสาทเทียมให้ค่าความแม่นยำที่สูงกว่า คือ 93.34% และลักษณะทางสันฐานวิทยาของรูปร่างและสีเป็นข้อมูลป้อนเข้าที่ดี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 2.4 สรุปงานวิจัยที่เกี่ยวข้อง (ต่อ)

ประเภท	ชื่องานวิจัย	เนื้อหาโดยย่อ	ตัวแบบที่ใช้	ผลที่ได้
คัดแยกเมล็ดข้าวสาร (ต่อ)	Deep Convolutional Neural Network Architecture for Boosting Image Discrimination Accuracy of Rice Species	หาาระบบจำแนกที่อัตโนมัติและแม่นยำเพื่อจำแนกเมล็ดข้าวสาร 3 สายพันธุ์ที่มีลักษณะเมล็ดแตกต่างกันสูง	ตัวแบบ 1. วิธีโครงข่ายประสาทแบบคอนโวลูชัน 2. วิธีดั้งเดิม คือวิธีเพื่อบ้านใกล้สุด K ตัว และวิธีซัพพอร์ตเวกเตอร์แมชชีน	วิธีโครงข่ายประสาทแบบคอนโวลูชันโดดเด่นที่สุด ค่าความแม่นยำคือ 95.5%
	Convolutional Neural Network with transfer learning for rice type classification	แยกประเภทเมล็ดข้าวสารที่มีลักษณะเหมือนกันสูง 5 พันธุ์ และแยกระหว่างข้าวหักและข้าวเต็มเมล็ด	เปรียบเทียบตัวแบบวิธีโครงข่ายประสาทแบบคอนโวลูชัน (CNN) ระหว่าง 1. พัฒนาขึ้นเอง 2. ผ่านการถ่ายทอดการเรียนรู้ ซึ่งคือ VGG16	ตัวแบบที่ผ่านการถ่ายทอดการเรียนรู้ให้ค่าความแม่นยำเพิ่มขึ้นมาก จาก 86.8% เป็น 94.2% ส่วนการแยกระหว่างข้าวหักและข้าวเต็มเมล็ด VGG16 ให้ค่า 99.31%

1. “Deep CNN-based damage classification of milled rice grains using a high-magnification image dataset” ได้ค้นคว้าเรื่องการแยกประเภทเมล็ดข้าวสารที่เสียหายหลังจากผ่านกระบวนการสีข้าว โดยใช้ภาพที่มีกำลังขยายสูง (4.5 เท่า) จำนวน 8,048 ภาพ ซึ่งแบ่งภาพข้าวเสียหายออกเป็น 7 ประเภท โดยใช้ตัวแบบโครงข่ายประสาทแบบคอนโวลูชัน (CNN) เปรียบเทียบทั้งหมด 5 ตัวแบบ ซึ่งทั้งหมดล้วนเป็นการถ่ายทอดการเรียนรู้ (Transfer Learning) โดยผ่านการปรับปรุงแก้ไข ประกอบด้วย EfficientNet-B0, ResNet-50, InceptionV3, MobileNetV2, MobileNetV3 โดยตัวแบบที่ให้ประสิทธิภาพดีที่สุดซึ่งมีค่าความแม่นยำในภาพรวมอยู่ที่ 98.37% คือ EfficientNet-B0 ใช้เวลาในการประมวลผลน้อย ทั้งยังสามารถแยกประเภทย่อยของข้าวเมล็ดทองไข่ ประกอบด้วย ข้าวทองไข่เต็มเมล็ด ครึ่งเมล็ด และ ทองไข่ผิวดลี

2.7.2 งานวิจัยที่เกี่ยวข้องกับการคัดแยกเมล็ดข้าวเปลือก

จุดประสงค์งานวิจัยประเภทนี้ เพื่อศึกษาวิธีที่เหมาะสมในการจำแนกสายพันธุ์ข้าวเปลือกภายใต้สภาพแวดล้อมที่ต่างกันออกไป เพื่อให้ทราบถึงวิธีที่ให้ค่าความแม่นยำสูงที่สุด ตัวอย่างการใช้ประโยชน์คือการประเมินราคาที่ยุติธรรมในการรับซื้อข้าวเปลือกจากชาวนา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1. งานวิจัย “แบ่งข้อมูลภาพเมล็ดข้าวด้วยการเรียนรู้เชิงลึก” มุ่งเน้นการแยกเมล็ดข้าวเปลือก (Rice Image Segmentation) ในรูปที่มีสภาพแวดล้อมต่างกันที่ถ่ายโดยกล้องโทรศัพท์มือถือ เพื่อนำผลลัพธ์ที่ได้ไปใช้ในการจำแนกสายพันธุ์ต่อไป โดยตัวแบบที่ใช้คือการเรียนรู้เชิงลึก ด้วยวิธี U-net และ Mask R-CNN ผลที่ได้คือ สำหรับวิธี U-net หากพื้นหลังต่างไปจากที่ได้รับการฝึกสอน ความสามารถในการแบ่งแยกเมล็ดข้าวสารจากพื้นหลังจะทำได้ไม่ค่อยดีนัก จึงมีการขยายข้อมูล (Data Augmentation) โดยการแปลงแสง เปลี่ยนพื้นหลัง ซึ่งผลที่ได้ดีขึ้น แต่ก็ยังไม่มีความสามารถทั่วไปมากพอที่จะนำไปใช้งานจริง และตัวแบบ Mask R-CNN พบปัญหาการตัดสินใจจากขนาด และวัตถุอื่นที่ขนาดใกล้เคียงเมล็ดข้าวจะถูกตรวจจับเป็นข้าว จึงใช้วิธีการขยายข้อมูล (Data Augmentation) โดยลดขนาด และเพิ่มภาพพื้นหลังให้หลากหลาย พบว่าตัวแบบสามารถทนต่อแสง พื้นผิว พื้นหลัง และระยะการถ่ายภาพที่แตกต่างกันได้เป็นอย่างดี เพียงพอต่อการใช้งานจริงในบริบทที่หลากหลาย

2. “Localization and Classification of Rice-grain Images Using Region Proposals-based Convolutional Neural Network” กล่าวว่า งานที่เกี่ยวข้องที่ผ่านมาไม่ได้จำแนกพันธุ์ข้าวที่ลักษณะทางกายภาพมีความใกล้เคียงกันมากเช่นงานวิจัยนี้ โดยจำแนกข้าวเปลือก 5 สายพันธุ์ของไทย (ข้าวเหนียว 1 สายพันธุ์ และข้าวเจ้า 4 สายพันธุ์) ด้วยวิธีการเรียนรู้เชิงลึก คือ Mask R-CNN ที่ผ่านการถ่ายทอดการเรียนรู้ (Transfer Learning) และได้ค่าน้ำหนัก (weights) มาจากชุดข้อมูล COCO เพื่อทั้งระบุตำแหน่งและจำแนกชนิดเมล็ดข้าว เปรียบเทียบกับวิธีซัพพอร์ตเวกเตอร์แมชชีน (SVM) โดยได้ทดลองหลายสถานการณ์ คือ 1. แยกระหว่างข้าวเปลือก 2. ประเภทย่อยระหว่างข้าวเจ้าและข้าวเหนียว 2. แยกประเภทย่อยของข้าวเจ้า 4 ชนิดที่มีลักษณะทางกายภาพที่เหมือนกันค่อนข้างมาก 3. แยกประเภททั้ง 5 ชนิด และสถานการณ์สุดท้าย 4. แยกระหว่างข้าวเหนียว ข้าวเปลือกคุณภาพดี และข้าวเปลือกคุณภาพไม่ดี โดยมาจากการอ้างอิงสถานการณ์ที่เกิดขึ้นจริง ทั้งยังได้เปรียบเทียบกับการใช้ผู้เชี่ยวชาญทำนาย 4 คน สำหรับการวัดผลประสิทธิภาพโดยรวม ใช้ค่า Precision เฉลี่ย ซึ่งสถานการณ์ที่ 1 เนื่องจากข้าวเปลือกมีความแตกต่างกัน ตัวแบบจึงสามารถทำได้ดีที่สุด ส่วนสถานการณ์ที่ 3 ทำได้เพียง 75% และสถานการณ์สุดท้ายได้ ค่า Precision เฉลี่ย 80% ซึ่งทำได้ดีกว่าผู้เชี่ยวชาญ

2.7.3 งานวิจัยที่เกี่ยวข้องกับการคัดแยกเมล็ดข้าวสาร

งานวิจัยประเภทนี้เพื่อศึกษาวิธีการคัดแยกสายพันธุ์ของข้าวสาร ประกอบด้วยทั้งวิธีแบบดั้งเดิม หรือวิธีที่เกี่ยวข้องกับการเรียนรู้เชิงลึก โดยมีข้อดี ข้อเสีย และความเหมาะสมกับงานที่ใช้แตกต่างกันออกไป

1. งานวิจัย “Identification of milled rice varieties using machine vision” ศึกษาลักษณะเด่นที่สำคัญ (feature) 3 กลุ่มใหญ่ ประกอบด้วย สัณฐานวิทยา (morphology) พื้นผิว (texture) และ สี (color) ซึ่งแต่ละลักษณะเด่นนั้นก็ยังมีกลุ่มย่อย ๆ แยกเพิ่มเติม และหาสัดส่วนที่เหมาะสมระหว่างลักษณะเด่นนั้น ๆ ที่จะให้ประสิทธิภาพดีที่สุด โดยใช้วิธีการกำลังสองน้อยสุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บางส่วน (Partial least square) มาช่วยในการหาความสัมพันธ์ระหว่างลักษณะเด่น (feature) และ พันธุ์ข้าว โดยพบว่าการใช้ลักษณะเด่นที่สำคัญ (feature) คือ สัณฐานวิทยา (morphology) พื้นผิว (texture) ร่วมกัน ได้คะแนนมากที่สุด และหากใช้ลักษณะเด่นที่สำคัญ (feature) มากจนเกินไป ก็อาจทำให้ได้ผลที่ไม่ดี เช่น การใช้ลักษณะเด่นที่สำคัญ (feature) ทั้งหมดก็ไม่สามารถได้คะแนนที่มากที่สุด เนื่องจากเกิดการซ้ำซ้อน

2. “Contrastive analysis of rice grain classification techniques: multi-class support vector machine vs artificial neural network” จำแนกเมล็ดข้าวสาร 3 สายพันธุ์ ประกอบด้วย Ponni, Basmati และ Brown rice นำมาผ่านกระบวนการแปลงภาพ ซึ่งใช้เปรียบเทียบ 3 วิธีคือ Level Sweep, Gradient Magnitude และ Hysteresis Threshold และตัวแบบที่ใช้จำแนกประเภทพันธุ์ข้าว ได้เลือกใช้วิธีที่ตรงกันข้ามคือวิธีซัพพอร์ตเวกเตอร์แมชชีนแบบหลายคำตอบ (Multi-class SVM) และวิธีโครงข่ายประสาทเทียม (ANN) พบว่าทั้ง 2 ตัวแบบสามารถให้ค่าความแม่นยำที่สูงทั้งคู่ และลักษณะทางสัณฐานวิทยาของรูปร่างและสีเป็นข้อมูลป้อนเข้าที่ดีของทั้ง 2 ตัวแบบ และหากพิจารณาความสามารถในการจำแนกภาพรวม วิธีโครงข่ายประสาทเทียมจะให้ค่าความแม่นยำที่สูงกว่า คือ 93.34%

3. “Deep Convolutional Neural Network Architecture for Boosting Image Discrimination Accuracy of Rice Species” จุดมุ่งหมายของงานวิจัยนี้คือหาระบบจำแนกที่อัตโนมัติและแม่นยำ เพื่อลดแรงงานในการปรับพารามิเตอร์ที่ซับซ้อน โดยใช้วิธีโครงข่ายประสาทแบบคอนโวลูชัน (CNN) โดยใช้ Stochastic Gradient Descent เป็นกลไกในการปรับค่าน้ำหนัก (optimizer) เพื่อจำแนกเมล็ดข้าวสาร 3 สายพันธุ์ที่มีลักษณะเมล็ดแตกต่างกันสูง ประกอบด้วยสายพันธุ์ Japonica เมล็ดขนาดกลาง พันธุ์ Glutinous เมล็ดกลม และพันธุ์ Indica เมล็ดยาว โดยผ่านเครื่องสแกนภาพ และเปรียบเทียบตัวแบบกับวิธีดั้งเดิม คือวิธีเพื่อนบ้านใกล้สุด K ตัว (KNN) และวิธีซัพพอร์ตเวกเตอร์แมชชีน (SVM) ซึ่งวิธีโครงข่ายประสาทแบบคอนโวลูชันให้ประสิทธิภาพที่โดดเด่นที่สุด ค่าความแม่นยำโดยรวมคือ 95.5%

4. “Convolutional Neural Network with transfer learning for rice type classification” เป็นงานวิจัยที่แยกประเภทเมล็ดข้าวสารซึ่งชุดข้อมูลมีความยากขึ้นมาจากความเหมือนกันของเมล็ดข้าว เนื่องจากงานวิจัยที่ผ่านมา แยกประเภทพันธุ์เมล็ดข้าวสารที่มีลักษณะทางกายภาพที่ต่างกันชัดเจน หรือเป็นเพียงการจัดกลุ่มตามความเหมือน (Clustering) หรือแบ่งประเภทตามคุณภาพข้าวเท่านั้น (Grading) ดังนั้นจึงไม่สามารถเปรียบเทียบกับงานวิจัยนี้ได้ โดยงานวิจัยนี้ทำ 2 ชิ้นงาน คือการแยกประเภทข้าวสารทั้งหมด 5 สายพันธุ์ โดยใช้ข้อมูลรูปภาพ 4,000 รูป งานที่สองคือการแยกระหว่างข้าวหักและข้าวเต็มเมล็ด ใช้ข้อมูล 1,600 รูป และเก็บข้อมูลรูปภาพผ่านเครื่องสแกน ตัวแบบที่ใช้จำแนกคือ วิธีโครงข่ายประสาทแบบคอนโวลูชัน (CNN) โดยเปรียบเทียบระหว่างตัวแบบที่พัฒนาขึ้นเอง และตัวแบบที่ผ่านการถ่ายทอดการเรียนรู้ (Transfer Learning) ซึ่งคือ VGG16 พบว่าในงานแยกพันธุ์ข้าวสาร มีเมล็ดข้าวประเภทที่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1 กับ 2 และ 1 กับ 4 ที่มีความใกล้เคียงสูง การใช้ตัวแบบที่ผ่านการถ่ายทอดการเรียนรู้ (transfer learning) ซึ่งคือ VGG16 ช่วยให้ค่าความแม่นยำโดยรวมเพิ่มขึ้นเป็นอย่างมาก จาก 86.8% เป็น 94.2% ส่วนการแยกระหว่างข้าวหักและข้าวเต็มเมล็ด VGG16 ให้ค่าความแม่นยำที่ 99.31%

จากการศึกษางานวิจัยที่เกี่ยวข้องเพื่อให้ทราบถึงการศึกษาในปัจจุบันที่เกี่ยวข้องกับเรื่องการจัดจำแนกสายพันธุ์ข้าวเปลือกไทย พบว่า งานวิจัยส่วนใหญ่จำแนกสายพันธุ์โดยมีข้อมูลที่มีลักษณะรูปทรงที่แตกต่างกันสูง ซึ่งแตกต่างจากงานที่กำลังศึกษา มีเพียงสองงานวิจัยที่จำแนกข้อมูลที่มีลักษณะใกล้เคียงกันสูง ประกอบด้วย งานวิจัยข้าวเปลือกไทย 5 พันธุ์ ชื่อว่า “Localization and Classification of Rice-grain Images Using Region Proposals-based Convolutional Neural Network” ใช้แบบจำลอง Mask R-CNN ที่ผ่านการถ่ายทอดการเรียนรู้มาจากชุดข้อมูล COCO และงานวิจัยข้าวสาร 5 พันธุ์ “Convolutional Neural Network with transfer learning for rice type classification” ที่ใช้วิธีโครงข่ายประสาทแบบคอนโวลูชันเปรียบเทียบกับตัวแบบที่ผ่านการถ่ายทอดการเรียนรู้ คือ VGG16 จึงใช้งานวิจัยดังกล่าวเป็นตัวอย่างในการศึกษา และเปรียบเทียบประสิทธิภาพแบบจำลองทั้งด้านความแม่นยำและความรวดเร็ว นอกจากนี้ยังรวมถึงวิธีการเก็บข้อมูลเมล็ดข้าวเปลือก กล่าวคือได้เรียนรู้วิธีการและใช้ขนาดรูปภาพตามงานวิจัยดังกล่าว

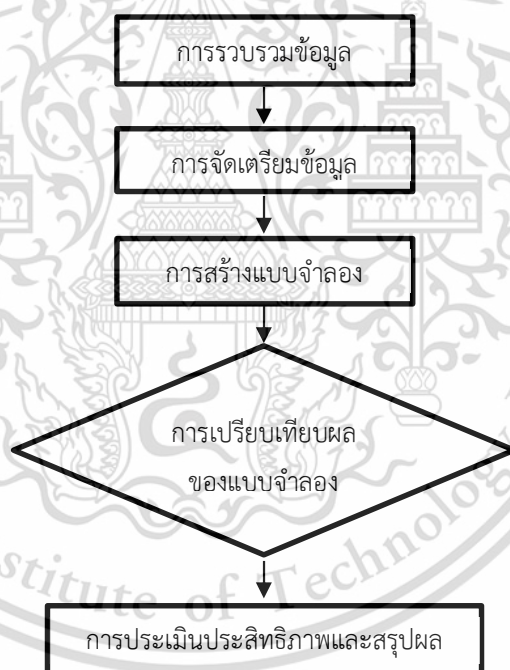
บทที่ 3

วิธีการดำเนินงานวิจัย

การศึกษาค้นคว้าอิสระหัวข้อเรื่อง การจำแนกสายพันธุ์ข้าวไทยโดยใช้เทคนิคการประมวลผลภาพและการเรียนรู้เชิงลึก ในบทนี้อธิบายถึงขั้นตอนการดำเนินงานภาพรวม การเก็บรวบรวมข้อมูล การจัดเตรียมข้อมูล การสร้างแบบจำลอง การประเมินผลแบบจำลอง ภาษาและซอฟต์แวร์ที่ใช้ในการพัฒนาแบบจำลอง

3.1 ขั้นตอนการดำเนินงานภาพรวม

ขั้นตอนการดำเนินงานภาพรวมถูกแสดงในรูปแบบผังงาน ดังรูปภาพที่ 3.1 ซึ่งมีทั้งหมด 4 ขั้นตอน ดังต่อไปนี้



รูปที่ 3.1 กระบวนการดำเนินงาน อธิบายถึงขั้นตอนโดยภาพรวม

1. เก็บรวบรวมข้อมูล โดยเก็บภาพสแกนเมล็ดข้าวเปลือกแต่ละสายพันธุ์ ใช้เครื่องสแกน Canon รุ่น CanoScan LiDE 300 ได้ภาพสแกนขนาด 4960 x 7015 พิกเซล
2. จัดเตรียมข้อมูลรูปภาพ ตัดแบ่งเป็นรูปขนาด 1024 x 1024 พิกเซล และระบุตำแหน่งเมล็ดข้าวในแต่ละภาพ โดยใช้วิธีการหาขอบภาพ ร่วมกับวิธีการหาเส้นเค้าโครง และเขียนในรูปแบบ JSON เพื่อใช้เป็นคำอธิบายประกอบรูปภาพ และนำรูปภาพพร้อมคำอธิบาย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประกอบรูปภาพอัฟโฟลด์ลง RoboFlow โดยแบ่งชุดข้อมูลรูปภาพเป็น 2 ชุดข้อมูล ประกอบด้วย ชุดข้อมูลเรียนรู้ จำนวน 80% และ ชุดข้อมูลทดสอบ จำนวน 20%

3. สร้างแบบจำลอง ตั้งค่าพารามิเตอร์ และฝึกอบรวมแบบจำลอง เพื่อฝึกสอนชุดข้อมูลให้เกิดการเรียนรู้
4. การเปรียบเทียบผลของแบบจำลอง คำนวณ และแสดงค่าเฉลี่ยการทดสอบประสิทธิภาพของชุดข้อมูลการทดสอบในแต่ละรอบ

3.2 การเก็บรวบรวมข้อมูล

ข้อมูลนำเข้าได้จากการเก็บภาพสแกนเมล็ดข้าวเปลือกแต่ละสายพันธุ์ผ่านเครื่องสแกน Canon รุ่น CanoScan LiDE 300 ได้ภาพสแกนขนาด 4960 x 7015 พิกเซล (กว้างคูณยาว) ซึ่งภาพสแกนเมล็ดข้าวเปลือกดังกล่าวจะถูกตรวจสอบสายพันธุ์ผ่านโปรแกรม EasyRice เพื่อตรวจสอบว่าข้าวเปลือกแต่ละเมล็ดว่าเป็นสายพันธุ์ที่ต้องการหรือไม่ เนื่องด้วยอาจมีการปลอมปน หรือปะปนสายพันธุ์อื่น สำหรับงานวิจัยนี้ สายพันธุ์ข้าวเปลือกที่เก็บรวบรวมข้อมูลประกอบด้วย 2 สายพันธุ์ ได้แก่

1. ข้าวเปลือกพันธุ์หอมมะลิไทย (ข้าวขาวดอกมะลิ 105 และข้าวพันธุ์ กข 15) ที่ปลูกบริเวณพื้นที่จังหวัดนครราชสีมา และนครนายก
2. ข้าวเปลือกพันธุ์ข้าวหอมปทุมธานี 1 ที่ปลูกบริเวณพื้นที่จังหวัดชลบุรี และระยอง

3.3 การจัดเตรียมข้อมูล (Data Preparation)

ข้อมูลที่ถูกเก็บรวบรวมจะถูกนำมาวิเคราะห์โดยเริ่มจากขั้นตอนการจัดเตรียมรูปภาพ หรือกระบวนการก่อนการประมวลผลภาพ (Pre-processing) เป็นขั้นตอนสำคัญในการทำความสะอาดข้อมูลภาพก่อนที่จะพร้อมใช้งานในคอมพิวเตอร์ต่อไป โดยในขั้นตอนนี้ นำภาพที่ได้จากการสแกนมาตัดเป็นรูปย่อยขนาด 1024 x 1024 พิกเซล เพื่อให้ภาพนำเข้าในขั้นตอนการเรียนรู้ผ่านแบบจำลองไม่ถูกย่อขนาดซึ่งจะส่งผลให้สูญเสียคุณลักษณะที่สำคัญลง สำหรับเครื่องมือที่ช่วยในการตัดรูปภาพคือไลบรารี opencv โปรแกรมภาษาไพธอน ผ่าน Google Colaboratory เพื่อความรวดเร็ว แม่นยำ และลดความผิดพลาดของมนุษย์

การแบ่งข้อมูลเพื่อใช้เป็นข้อมูลชุดฝึกสอน (Training Data) และข้อมูลชุดทดสอบ (Testing Data) ซึ่งจะใช้หลักการสุ่มกลุ่มตัวอย่างอย่างง่าย (Simple Random Sampling) โดยใช้คอมพิวเตอร์ในการสุ่มตัวอย่าง ในสัดส่วนข้อมูลชุดฝึกสอนต่อข้อมูลชุดทดสอบ ร้อยละ 80:20 ซึ่งรายละเอียดมีดังนี้

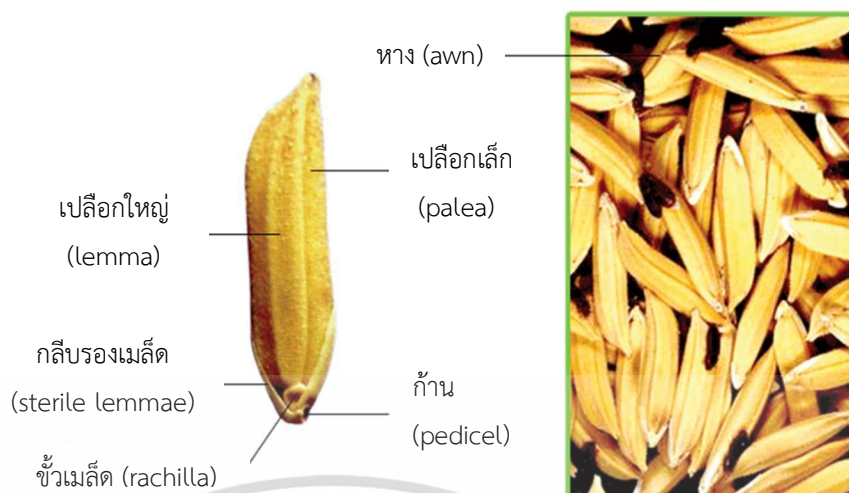
1. ข้าวเปลือกพันธุ์หอมมะลิไทย (ข้าวขาวดอกมะลิ 105 และข้าวพันธุ์ กข 15) ที่ปลูกบริเวณพื้นที่จังหวัดนครราชสีมา และนครนายก จำนวน ภาพ 756 มีจำนวนเมล็ด 4565 เมล็ด
2. ข้าวเปลือกพันธุ์ข้าวหอมปทุมธานี 1 ที่ปลูกบริเวณพื้นที่จังหวัดชลบุรี และระยอง ทั้งหมดจำนวน 648 ภาพ มีจำนวนเมล็ด 4381 เมล็ด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากนั้นทำคำอธิบายประกอบรูปภาพ (Image annotation) ซึ่งเป็นขั้นตอนการติดป้ายกำกับรูปภาพในชุดข้อมูลนั้น ๆ เพื่อใช้เป็นผลเฉลยในการฝึกแบบจำลอง สำหรับการแบ่งส่วนตัวอย่าง (Instance segmentation) ต้องตีกรอบล้อมรอบตามรูปทรงวัตถุซึ่งในที่นี้คือเมล็ดข้าวเปลือก เพื่อบ่งบอกถึงตำแหน่งที่ตั้งของเมล็ดข้าวเปลือกในชุดข้อมูลฝึกสอนว่าแต่ละเมล็ดอยู่ส่วนใดของภาพ เนื่องด้วยวัตถุที่ใช้ในการศึกษาครั้งนี้คือเมล็ดข้าวเปลือกเพียงอย่างเดียว และสภาพแวดล้อมหรือภาพพื้นหลังไม่ซับซ้อน จึงได้นำวิธีการหาขอบภาพโดยวิธีแคนนี่ (Canny edge detection) ร่วมกับวิธีการหาเส้นเค้าโครง (FindContour) ของไลบรารี opencv มาช่วยในการทำคำอธิบายประกอบรูปภาพเพื่อช่วยให้การทำคำอธิบายประกอบรูปภาพมีความแม่นยำ สม่ำเสมอ และรวดเร็วมากยิ่งขึ้น โดยก่อนการหาขอบภาพจะต้องปรับปรุงคุณภาพรูปภาพเพื่อให้สามารถหาขอบได้ครบถ้วน ประกอบด้วย การปรับค่าความเปรียบต่าง (Contrast) เพื่อปรับให้ทุกส่วนของเมล็ดข้าวเปลือกสามารถเห็นได้ชัด จากนั้นปรับภาพให้เรียบ (Blur Image) เพื่อกำจัดสัญญาณรบกวน แล้วจึงหาขอบภาพโดยวิธีแคนนี่ (Canny edge detection) เพื่อนำขอบภาพที่ได้ไปทำให้ขอบภาพชัดเจนยิ่งขึ้นในขั้นตอนการขยายขนาด (Dilation) เนื่องจากมีบางส่วนของขอบภาพเมล็ดข้าวเปลือกไม่สามารถตรวจจับได้ แล้วจึงใช้วิธีการกัดเซาะ (Erosion) เพื่อลดสัญญาณรบกวนที่ยังคงเหลืออยู่อีกครั้ง จากนั้นจึงหาเส้นเค้าโครงเป็นขั้นตอนสุดท้าย โดยแต่ละขั้นตอนมีรายละเอียดดังนี้

3.3.1 ความเปรียบต่าง (Contrast)

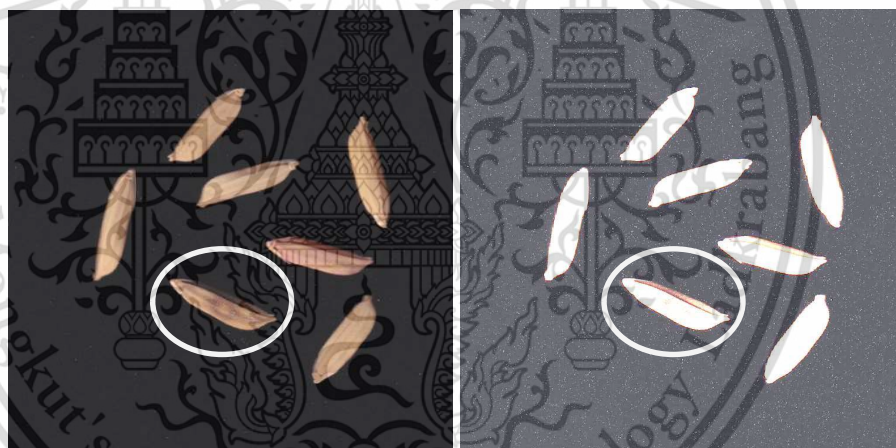
เปลือกของเมล็ดข้าวเปลือกโดยทั่วไปจะมีส่วนกลางหรือบริเวณเปลือกใหญ่ที่มีลักษณะเป็นสันขึ้นมาเล็กน้อย ดังรูปที่ 3.2 แสดงส่วนประกอบของเมล็ดข้าวเปลือก ภาพที่ได้จากการสแกนในบางครั้ง ส่วนสันนี้จะบดบังแสงที่จะตกกระทบไปยังส่วนข้างเคียง ส่งผลให้ส่วนข้างเคียงเหล่านั้นมีสีที่มืดกว่าส่วนอื่นของเมล็ด ซึ่งมีสีใกล้เคียงพื้นหลัง ดังรูปที่ 3.3 ก) ซึ่งจะกระทบต่อการหาขอบเพื่อทำคำอธิบายประกอบรูปภาพ ดังนั้นในลำดับแรกจึงต้องปรับค่าความเปรียบต่างของรูปภาพเพื่อเพิ่มความสว่างให้ส่วนที่มืดนั้น ผลลัพธ์จะเห็นได้ว่าบริเวณดังกล่าวเกิดความแตกต่างขึ้นมาเมื่อเทียบกับพื้นหลัง ส่งผลให้ตรวจหาขอบได้สมบูรณ์มากขึ้น ดังรูปที่ 3.3 ข)



รูปที่ 3.2 ส่วนประกอบของเมล็ดข้าวเปลือก

ที่มา : องค์ความรู้เรื่องข้าว กองวิจัยและพัฒนาข้าว กรมการข้าว

<https://ricecloud.ricethailand.go.th/RKB/important.php?id=2>



รูปที่ 3.3 ก) ส่วนของเมล็ดข้าวเปลือกที่มีดกว่าส่วนอื่นและมีสีใกล้เคียงพื้นหลัง

รูปที่ 3.3 ข) ส่วนของเมล็ดข้าวเปลือกหลังผ่านการปรับค่าความเปรียบต่าง

3.3.2 การปรับภาพให้เรียบ (Blur Image) หรือเครื่องกรองผ่านความถี่ต่ำ

หลังจากการปรับความเปรียบต่าง ภาพนำเข้ามีสัญญาณรบกวนหรือจุดต่างจำนวนมาก ดังนั้นจึงใช้วิธีการปรับภาพให้เรียบ (Blur Image) หรือฟิลเตอร์ภาพเลื่อน เพื่อลดจุดต่างในภาพ สำหรับความเหมาะสมในชุดข้อมูลรูปภาพที่ใช้ในงานศึกษานี้ ได้ใช้ฟิลเตอร์ภาพเลื่อนทั้งหมด 2 วิธี ประกอบด้วย การใช้ตัวกรองเป็นฟังก์ชันเกาส์ (Gaussian Blur) และตัวกรองทวิภาคี (Bilateral filter) ตามลำดับ ดังรูปที่ 3.4 ก) โดยการเลือกใช้ฟิลเตอร์ภาพเลื่อนวิธีใด

ๆ ได้มาจากการทดลองจนกว่าจะได้รูปภาพที่เหมาะสมสำหรับการทำคำอธิบายประกอบเอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

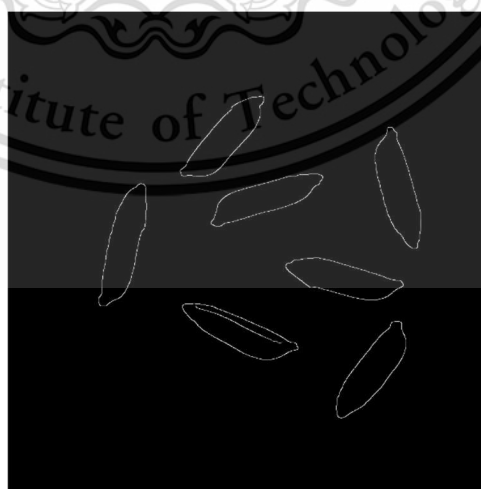
รูปภาพ พิจารณาจากความสามารถในการลบจุดต่าง โดยที่ไม่สูญเสียรูปร่างเมล็ดข้าวเปลือกเดิม หลังจากนั้นจึงปรับภาพที่ได้เป็นสีเทา ดังรูปที่ 3.4 ข)



รูปที่ 3.4 ก) ภาพข้าวเปลือกหลังจากใช้ตัวกรองภาพเลือนฟังก์ชันเกาส์ และทวิภาคี
รูปที่ 3.4 ข) ภาพข้าวเปลือกหลังจากปรับภาพเป็นสีเทา

3.3.3 การหาขอบภาพ (Edge detection)

ขั้นตอนถัดมาคือการหาขอบภาพ โดยในงานศึกษานี้ได้ใช้วิธีแคนนี่ (Canny edge detection) ดังรูปที่ 3.5 และจากการทดลองการหาขอบภาพนั้นยังไม่สมบูรณ์ทุกเมล็ด มีขอบภาพของข้าวเปลือกบางเมล็ดขาดหาย ซึ่งจะส่งผลต่อการทำคำอธิบายประกอบรูปภาพ ดังนั้นจึงใช้วิธีการขยายขนาด และวิธีการกัดเซาะเพิ่มเติม



รูปที่ 3.5 ภาพข้าวเปลือกหลังจากหาขอบภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.3.4 การขยายขยาย (Dilation) และการกัดเซาะ (Erosion)

วิธีการขยายขนาดเพื่อให้ขอบของภาพชัดเจนยิ่งขึ้น จึงมีประโยชน์มากในการหาขอบของภาพเนื่องจากเป็นวิธีที่ทำให้ขอบของภาพชัดเจนยิ่งขึ้น แสดงดังรูปที่ 3.6 ก) จากนั้นใช้วิธีการกัดเซาะ ซึ่งเป็นวิธีที่ทำให้ความหนาหรือขนาดของวัตถุจะลดลง นำมาใช้ประโยชน์เรื่องการลดสัญญาณรบกวนที่ยังคงเหลือในภาพ แสดงดังรูปที่ 3.6 ข)



รูปที่ 3.6 ก) ภาพข้าวเปลือกหลังจากใช้วิธีการขยายขนาด
รูปที่ 3.6 ข) ภาพข้าวเปลือกหลังจากใช้วิธีการขยายกัดเซาะ

3.3.5 การหาเส้นเค้ไครง



รูปที่ 3.7 ก) คำอธิบายประกอบรูปภาพ

รูปที่ 3.7 ข) รูปภาพนำเข้าพร้อมคำอธิบายประกอบรูปภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นตอนสุดท้ายในการทำคำอธิบายประกอบรูปภาพ คือการใช้ฟังก์ชันการหาเส้นเค้าโครง (FindContour) และฟังก์ชันการวาดเส้นเค้าโครง (drawContours) ของไลบรารี opencv เพื่อวาดขอบบนรูปภาพ ดังรูปที่ 3.7 ก) ซึ่งข้อมูลเส้นเค้าโครงรูปนี้จะถูกนำไปเขียนเป็นไฟล์นามสกุล JSON ซึ่งเป็นที่นิยมในการใช้สำหรับการทำคำอธิบายประกอบรูปภาพ ซึ่งในไฟล์นี้จะประกอบด้วยข้อมูลชื่อรูปภาพ ลำดับเมล็ดข้าวเปลือกและพิกัดล้อมรอบแต่ละเมล็ด พร้อมทั้งบ่งบอกสายพันธุ์ข้าวเปลือก (ผลเฉลี่ย) เมล็ดดังกล่าว สำหรับรูปที่ 3.7 ข) คือการวาดเส้นเค้าโครงที่ได้ลงบนรูปภาพนำเข้านั้น ๆ

3.4 การสร้างแบบจำลอง

ผู้วิจัยสร้างแบบจำลองการแบ่งส่วนตัวอย่างของ YOLOv8 โดยรวบรวมข้อมูลข้าวเปลือกพันธุ์หอมมะลิไทย (ข้าวขาวดอกมะลิ 105 และข้าวพันธุ์ กข 15) จำนวน 756 ภาพ มีจำนวนเมล็ด 4565 เมล็ด และข้าวเปลือกพันธุ์ข้าวหอมปทุมธานี 1 จำนวน 648 ภาพ มีจำนวนเมล็ด 4381 เมล็ด แบ่งข้อมูลชุดสอน สำหรับข้าวเปลือกพันธุ์หอมมะลิไทย จำนวน 605 ภาพ ข้าวเปลือกพันธุ์ข้าวหอมปทุมธานี 1 จำนวน 518 ภาพ โดยคิดเป็นร้อยละประมาณ 80 และข้อมูลชุดทดสอบ ข้าวเปลือกพันธุ์หอมมะลิไทย จำนวน 151 ภาพ ข้าวเปลือกพันธุ์ข้าวหอมปทุมธานี 1 จำนวน 130 ภาพ คิดเป็นร้อยละ 20 จากนั้นได้สร้างแบบจำลองการแบ่งส่วนตัวอย่างที่ได้รับการฝึกล่วงหน้าบนชุดข้อมูล COCO ทั้งหมด 5 แบบย่อย ประกอบด้วย YOLOv8n-seg YOLOv8s-seg YOLOv8m-seg YOLOv8l-seg YOLOv8x-seg ซึ่งแต่ละแบบจำลองนั้นมีจำนวนพารามิเตอร์ ความเร็วของการประมวลผล รวมถึงประสิทธิภาพที่ต่างกัน หากเรียงตามขนาดของพารามิเตอร์ของแบบจำลองจากน้อยไปหามาก จะสามารถแสดงได้ตามตารางที่ 3.1 จากนั้นนำแบบจำลองที่ให้ค่าที่ดีที่สุดไปเปรียบเทียบประสิทธิภาพกับแบบจำลองที่สร้างขึ้นใหม่โดยไม่ได้รับการฝึกล่วงหน้าบนชุดข้อมูล COCO

ตารางที่ 3.1 แบบจำลองการแบ่งส่วนตัวอย่าง YOLOv8 ที่ได้รับการฝึกสอน

Model	Image size	mAPbox 50-95	mAPmask 50-95	Speed CPU ONNX (ms)	Speed A100 TensorRT (ms)	params (M)	FLOPs (B)
YOLOv8n-seg	640	36.7	30.5	96.1	1.21	3.4	12.6
YOLOv8s-seg	640	44.6	36.8	155.7	1.47	11.8	42.6
YOLOv8m-seg	640	49.9	40.8	317.0	2.18	27.3	110.2
YOLOv8l-seg	640	52.3	42.6	572.4	2.79	46.0	220.5
YOLOv8x-seg	640	53.4	43.4	712.1	4.02	71.8	344.1

ที่มา : <https://github.com/ultralytics/ultralytics>

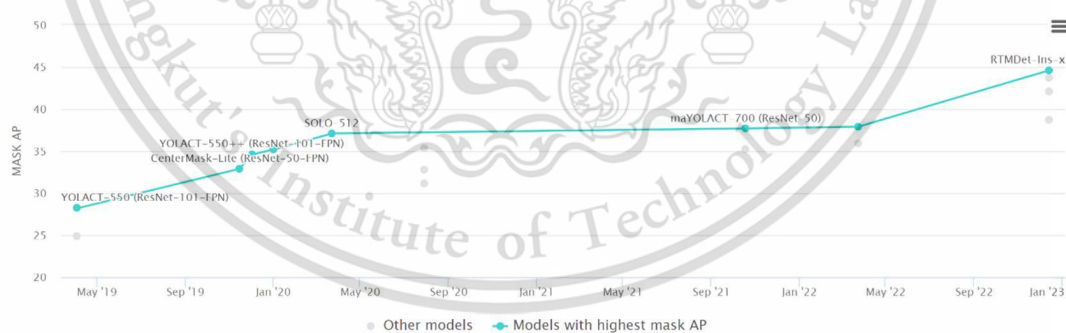
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ระหว่างกระบวนการฝึกสอน สามารถกำหนดค่าไฮเปอร์พารามิเตอร์ที่ใช้ในการฝึกสอน การตั้งค่าเหล่านี้ส่งผลต่อประสิทธิภาพ ความเร็ว และความแม่นยำของแบบจำลอง ไฮเปอร์พารามิเตอร์ที่สำคัญดังกล่าวที่ควรทดลองหาค่าที่เหมาะสม มีรายละเอียดดังตารางที่ 3.2

ตารางที่ 3.2 กำหนดค่าในการทดลองการปรับค่าไฮเปอร์พารามิเตอร์

ไฮเปอร์พารามิเตอร์	กำหนดค่าที่ทดลอง
จำนวนรอบการฝึกสอน (epoch)	100, 500
จำนวนข้อมูลที่เรียนรู้ในหนึ่งรอบ (batch size)	4, 8, 16, 32, 64
อัตราการเรียนรู้ (learning rate)	0.1, 0.01, 0.001, 0.0001, 0.01-0.0001(ค่อย ๆ ลดจาก 0.01 จนถึง 0.0001)
เครื่องมือเพิ่มประสิทธิภาพ (optimizer)	SGD, Adam

งานการแบ่งส่วนตัวอย่างแบบเรียลไทม์ มีแบบจำลองอื่นที่มีความทันสมัย เป็นที่ยอมรับ ถูกใช้ในงานวิจัยหลากหลาย และมีประสิทธิภาพที่ดีที่สุดในปัจจุบัน (State-of-the-Art: SOTA) อ้างอิงจากตารางการเปรียบเทียบแบบจำลองการแบ่งส่วนตัวอย่างแบบเรียลไทม์บนชุดข้อมูลมาตรฐาน MSCOCO ของเว็บไซต์ Papers With Code ประกอบด้วย แบบจำลอง RTMDet YOLACT SOLOV2 และ Mask R-CNN ดังรูปที่ 3.8 จึงได้นำมาเปรียบเทียบประสิทธิภาพกับแบบจำลอง YOLOv8 ในชุดข้อมูลนี้



รูปที่ 3.8 ประสิทธิภาพแบบจำลองการแบ่งส่วนตัวอย่างแบบเรียลไทม์บนชุดข้อมูล MSCOCO
ที่มา : <https://paperswithcode.com/sota/real-time-instance-segmentation-on-mscoco>

3.5 การประเมินผล

สำหรับการประเมินผลแบบจำลองการตรวจจับวัตถุด้านความแม่นยำของแบบจำลอง นิยมใช้ค่าเฉลี่ยความแม่นยำเฉลี่ย (mean Average Precision : mAP) เป็นตัวชี้วัดประสิทธิภาพแบบจำลอง ซึ่งเป็นค่าที่ได้จากค่าความเที่ยงตรงเฉลี่ย (AP) สำหรับแต่ละหมวดหมู่ แล้วนำไปหาค่าเฉลี่ยจากทุกหมวดหมู่อีกครั้ง และโดยทั่วไปจะดูจากค่า mAP ที่ 0.5:0.95 คือ mAP ที่อัตราส่วนพื้นที่ทับซ้อนเหนือพื้นที่ทั้งหมด (IoU) มีค่าเท่ากับ 0.5 – 0.95 นอกจากนี้ยังประเมินประสิทธิภาพแบบจำลองด้านความเร็วในการอนุมาน (Inference time) ว่าแบบจำลองสามารถประมวลผลโดยใช้เวลากี่มิลลิวินาทีต่อหนึ่งภาพ และสามารถประมวลผลได้จำนวนกี่ภาพในหนึ่งวินาที



บทที่ 4

ผลการวิจัยและการอภิปรายผล

งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาการจำแนกสายพันธุ์ข้าวไทย ประกอบด้วยข้าวเปลือกพันธุ์หอมมะลิไทย (ข้าวขาวดอกมะลิ 105 และข้าวพันธุ์ กข 15) และข้าวเปลือกพันธุ์หอมปทุมธานี 1 โดยสร้างแบบจำลองการแบ่งส่วนตัวอย่าง YOLOv8 สำหรับการอภิปรายผล ในบทนี้จะกล่าวถึงผลลัพธ์ของการทดลองประสิทธิภาพแบบจำลอง

4.1 แบบจำลองการแบ่งส่วนตัวอย่างอื่นที่ได้รับความนิยม

ผลการทดลองโดยใช้แบบจำลองอื่นที่ได้รับความนิยมและเป็นที่ยอมรับ เช่น YOLACT SOLOV2 และ Mask R-CNN การทดลองแสดงผลดังตารางที่ 4.1

ตารางที่ 4.1 ผลการทดลองเปรียบเทียบแบบจำลองการแบ่งส่วนตัวอย่างแบบเรียลไทม์อื่น

Model	Image size	Backbone	mAP all (0.50:0.95)	mAP50	mAP75	Speed (ms)	Speed (FPS)
YOLACT	550x550	ResNet50	0.9190	0.9890	0.9890	72.1	13.87
SOLOV2	1333x800	ResNet50	0.9830	0.9900	0.9900	197	5.08
Mask R-CNN	1333x800	ResNet50	0.9880	0.9890	0.9890	119.5	8.37

จากการทดลอง พบว่าแบบจำลอง Mask R-CNN มีประสิทธิภาพดีที่สุด เนื่องจากให้ค่าความแม่นยำเฉลี่ยสูงสุด คือ 0.9880 วัดผลโดยค่า mAP ที่ 0.50-0.95 ซึ่งเป็นค่าหลักที่ใช้พิจารณาประสิทธิภาพของแบบจำลอง [8] มีความหมายถึงค่าความแม่นยำเฉลี่ย (AP) ของอัตราส่วนพื้นที่ทับซ้อนเหนือพื้นที่ทั้งหมด (IoU) เริ่มจาก IoU ที่ 0.5 จนถึง 0.95 โดยเพิ่มทีละ 0.05 จากนั้นจึงคำนวณเป็นค่าเฉลี่ยอีกครั้ง ค่า IoU ได้จากการคำนวณอัตราส่วนระหว่างพื้นที่ทับซ้อนที่เป็นผลเฉลยกับที่แบบจำลองทำนาย หากด้วยพื้นที่รวมของกล่องล้อมรอบวัตถุทั้งสอง หากพิจารณาถึงความรวดเร็วในการทำนาย แบบจำลอง YOLACT สามารถทำนายได้รวดเร็วที่สุด ใช้เวลาต่อหนึ่งเฟรมอยู่ที่ 72.1 มิลลิวินาที หรือสามารถประมวลผลได้ 13.87 เฟรมต่อวินาที แต่ให้ค่า mAP(50-95) ต่ำที่สุด

4.2 แบบจำลองการแบ่งส่วนตัวอย่าง YOLOv8 ที่ได้รับการฝึกสอนแล้ว

ผลการทดลองแบบจำลองการแบ่งส่วนตัวอย่าง YOLOv8 ทั้ง 5 แบบจำลอง ดังตารางที่ 4.2

ตารางที่ 4.2 ผลการทดลองแบบจำลองการแบ่งส่วนตัวอย่าง YOLOv8

Model	Image size	mAP50-95 (Box)	mAP50-95 (Mask)	Speed Tesla T4 GPU		Params (M)	FLOPs
				ms	FPS		
YOLOv8n-seg	640	0.9950	0.9903	4.8	208.33	3.26	12.0
YOLOv8s-seg	640	0.9950	0.9903	6.6	151.52	11.78	42.4
YOLOv8m-seg	640	0.9950	0.9910	13.7	72.99	27.22	110.0
YOLOv8l-seg	640	0.9950	0.9910	21.4	46.73	45.91	220.1
YOLOv8x-seg	640	0.9950	0.9906	30.7	32.57	71.72	343.7

จากตารางที่ 4.2 พบว่าทุกแบบจำลองให้ค่า mAP50-95 ที่ใกล้เคียงกัน คือ mAP สำหรับกล่องล้อมรอบวัตถุของทุกแบบจำลองมีค่าเท่ากับ 0.9950 และ ค่า mAP ของกรอบล้อมรอบตามรูปทรงเมล็ดข้าวของทุกแบบจำลองมีค่าที่ใกล้เคียงกันมากในช่วง 0.9903-0.9910 ดังนั้นจึงพิจารณาถึงขนาดของแบบจำลอง ซึ่งส่งผลต่อความเร็วในการประมวลผล และจำนวนภาพที่สามารถประมวลผลได้ในหนึ่งวินาที พบว่า YOLOv8n-seg ให้ค่าที่ดีที่สุด แบบจำลองมีขนาดเล็กที่สุด และความเร็วในการประมวลผลมากที่สุด ดังนั้นจึงสามารถสรุปได้ว่า แบบจำลอง YOLOv8n-seg ให้ค่าที่ดีที่สุดเมื่อพิจารณาถึงการแลกเปลี่ยนระหว่างความเร็วในการประมวลผลและความแม่นยำ (Speed and Accuracy trade-off)

การปรับค่าไฮเปอร์พารามิเตอร์ที่สำคัญของแบบจำลอง

4.2.1 จำนวนรอบการฝึกสอน (Epoch)

ผลทดลองเปรียบเทียบระหว่างการฝึกสอนจำนวน 100 รอบ และ 500 รอบ พบว่าสำหรับค่าเฉลี่ยความแม่นยำเฉลี่ยที่ใช้วัดความถูกต้องของแบบจำลองนั้นไม่ได้แตกต่างกัน เนื่องจากแบบจำลองเข้าสู่ค่าที่ดีที่สุดก่อนการฝึกสอนรอบที่ 100 และค่าสูญเสียของการฝึกสอนจำนวน 500 รอบนั้นลดลงเพียงเล็กน้อย ไม่ได้มีนัยสำคัญ ดังนั้นงานวิจัยนี้จึงใช้รอบการฝึกสอน 100 รอบ สำหรับแบบจำลองเปรียบเทียบต่าง ๆ

4.2.2 จำนวนรายการข้อมูลที่คำนวณในหนึ่งครั้ง (Batch size)

จากการทดลองปรับค่า ให้มีขนาด 4, 8, 16, 32 และ 64 พบว่า ไม่ได้มีผลต่อค่าเฉลี่ยของความแม่นยำเฉลี่ยของแบบจำลองอย่างมีนัยสำคัญ แต่มีผลเรื่องระยะเวลาของการฝึกสอนแบบจำลอง ความเสถียรของแบบจำลอง และจำนวนหน่วยความจำของหน่วยประมวลผลภาพกราฟิก (GPU memory) ดังตารางที่ 4.3 สรุปได้ว่าแบบจำลองที่กำหนดค่า Batch size ให้เท่ากับ 16 ใช้ระยะเวลาการฝึกสอนสั้นที่สุด และสำหรับความ

เสถียรของแบบจำลอง ขนาดรายการข้อมูลที่คำนวณในหนึ่งครั้งที่น้อยกว่าจะให้ค่าเฉลี่ยของความแม่นยำเฉลี่ยที่เสถียรกว่า

ตารางที่ 4.3 ผลการทดลองการกำหนด Batch size ขนาดต่าง ๆ

ขนาด แบตช์	จำนวนชั่วโมง การฝึกสอน	จำนวนหน่วย ความจำ (GB)	mAP50-95 (Box)	mAP50-95 (Mask)	Speed (ms)
4	2.995	1.29	0.9950	0.9893	6.0
8	2.939	2.62	0.9950	0.9903	4.8
16	2.735	4.52	0.9950	0.9904	3.7
32	2.835	9.52	0.9950	0.9903	4.1
64	2.959	14.5	0.9950	0.9909	3.6

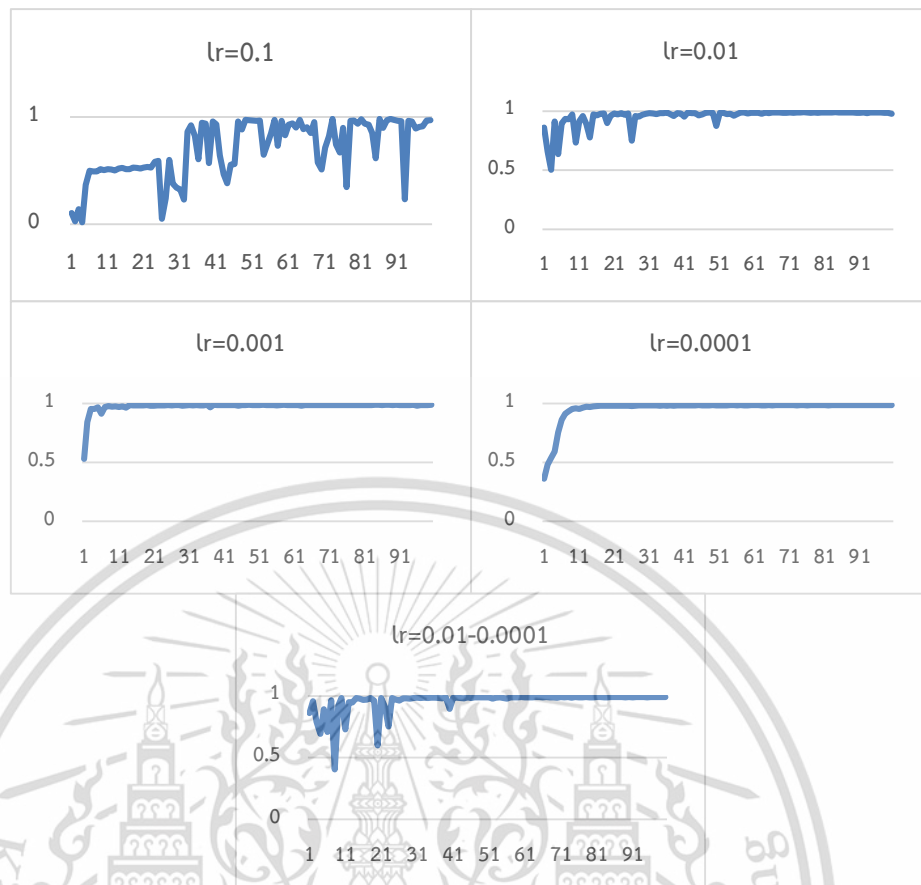
4.2.3 อัตราการเรียนรู้ (Learning rate)

จากการเปรียบเทียบอัตราการเรียนรู้ทั้งหมด 4 ค่า ประกอบด้วย 0.01, 0.001, 0.0001 และ 0.01-0.0001 พบว่าอัตราการเรียนรู้ที่ 0.01-0.0001 ที่ซึ่งค่อย ๆ ลดอัตราการเรียนรู้ในแต่ละรอบการฝึกสอนในอัตราที่เหมาะสม แบบจำลองมีอัตราการเรียนรู้สูงสุด คือ มี mAP50-95 ที่ 0.9904 ตารางที่ 4.4

ตารางที่ 4.4 ผลการทดลองที่อัตราการเรียนรู้ต่าง ๆ

อัตรา การเรียนรู้	lr=0.1	lr=0.01	lr=0.001	lr=0.0001	lr=0.01- 0.0001
mAP50-95 (Mask)	0.9831	0.9886	0.9879	0.9860	0.9904

หากพิจารณาจากค่า mAP50-95 ที่อัตราการเรียนรู้ในค่าต่าง ๆ ตลอดการฝึกสอนทั้ง 100 รอบ พบว่า อัตราการเรียนรู้ที่ 0.1 (lr=0.1) กราฟจะมีการแกว่งขึ้นลงอย่างมาก ถึงแม้ค่าเฉลี่ยความแม่นยำเฉลี่ยที่ดีที่สุดนั้นมีค่าสูง แต่แบบจำลองไม่เสถียร ขณะที่อัตราการเรียนรู้ที่ 0.01, 0.001, 0.0001 และ 0.01-0.0001 แบบจำลองประสบความสำเร็จในการเรียนรู้ค่อนข้างมาก และมีค่า mAP ที่ใกล้เคียงกัน แต่อัตราการเรียนรู้ที่ 0.001 และ 0.0001 นั้นจะลู่เข้าหาค่าที่ดีที่สุดตั้งแต่รอบการฝึกสอนต้น ๆ ซึ่งค่อนข้างเร็วกว่าอัตราการเรียนรู้อื่นอย่างเห็นได้ชัด ดังรูปที่ 4.1



รูปที่ 4.1 ค่า mAP50-95 ตลอดการฝึกสอน 100 รอบ ที่อัตราการเรียนรู้ขนาดต่าง ๆ

4.2.4 Optimizer

ทดลองเปรียบเทียบ Optimizer 2 ตัว เพื่อหาสิ่งที่เหมาะสมกับข้อมูลชุดนี้มากที่สุด ประกอบด้วย Stochastic gradient descent (SGD) และ Adam ซึ่งทั้ง 2 วิธีได้รับความนิยมในปัจจุบัน ซึ่งทดลองที่อัตราการเรียนรู้ที่ 0.001, 0.0001 และ 0.01-0.0001 กับ Optimizer ทั้ง 2 ตัว เนื่องจากการทดลองก่อนหน้านี้ พบว่าที่อัตราการเรียนรู้นี้ดังกล่าว แบบจำลองให้ผลลัพธ์ที่ดีที่สุด กล่าวคือแบบจำลองมีความเสถียร ผลการทดลองเปรียบเทียบ Optimizer ดังตารางที่ 4.5

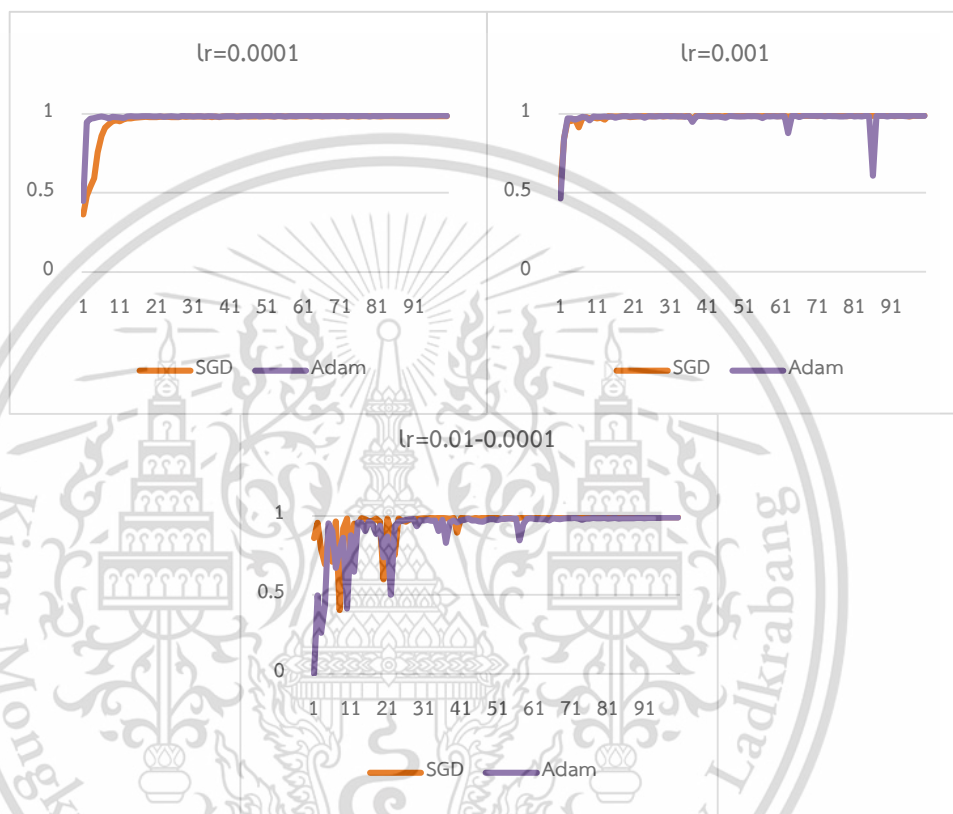
ตารางที่ 4.5 ผลการทดลองใน Optimizer ที่อัตราการเรียนรู้ขนาดต่าง ๆ

Learning rate	mAP50-95 (Mask)	
	Optimizer = SGD	Optimizer = Adam
lr=0.001	0.9879	0.9894
lr=0.0001	0.9860	0.9896
lr=0.01-0.0001	0.9904	0.9880

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากตารางที่ 4.5 พบว่า อัตราการเรียนรู้ที่ 0.001 และ 0.0001 Optimizer Adam ให้ค่าที่ดีกว่า แต่ที่ 0.01-0.0001 นั้น Optimizer SGD ให้ค่าที่ดีกว่า และดีที่สุดเมื่อเปรียบเทียบกับอัตราการเรียนรู้อื่น สำหรับการพิจารณาถึงความเสถียรในการเรียนรู้ของแบบจำลองที่อัตราการเรียนรู้ขนาดต่าง ๆ พบว่าทั้งสอง Optimizer มีความเสถียรใกล้เคียงกัน ดังรูปที่ 4.2



รูปที่ 4.2 ค่า mAP50-95 ตลอดการฝึกสอนในอัตราการเรียนรู้ขนาดต่าง ๆ

4.3 แบบจำลองการแบ่งส่วนตัวอย่าง YOLOv8 ที่ไม่ได้ผ่านการฝึกสอน

ตารางที่ 4.6 ค่าไฮเปอร์พารามิเตอร์ที่เหมาะสมกับข้อมูลชุดนี้

ไฮเปอร์พารามิเตอร์	รายละเอียด
จำนวนรอบการฝึกสอน (epoch)	100
จำนวนข้อมูลที่เรียนรู้ในหนึ่งรอบ (batch size)	16
อัตราการเรียนรู้ (learning rate)	0.01-0.0001
เครื่องมือเพิ่มประสิทธิภาพ (optimizer)	SGD

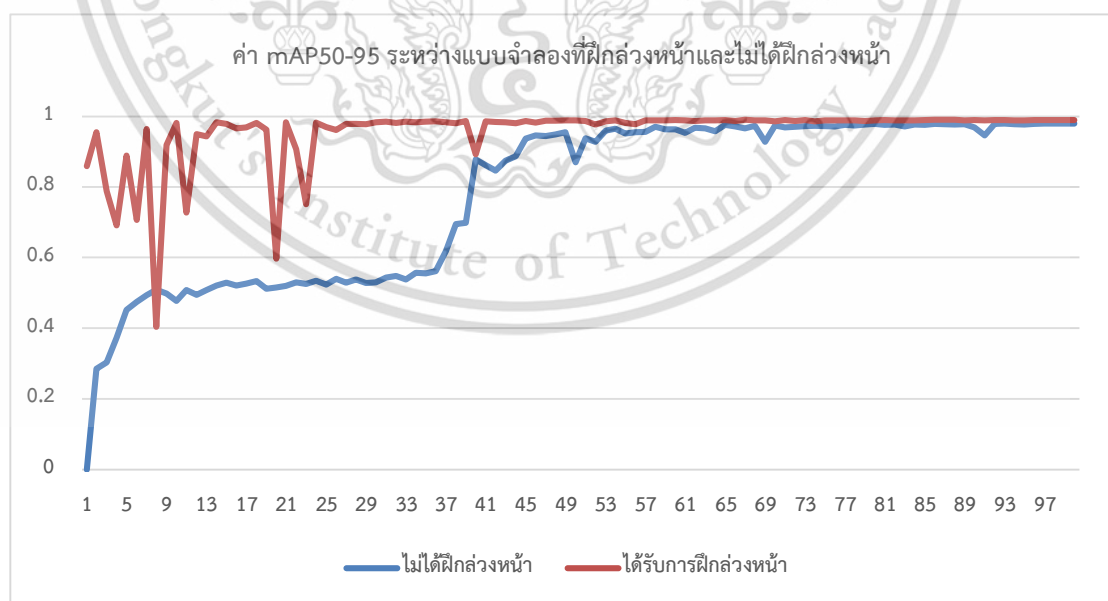
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากการทดลองข้างต้น จะได้แบบจำลองที่ดีและเหมาะสมกับข้อมูลชุดนี้ ซึ่งแบบจำลองดังกล่าวคือคือ YOLOv8n-seg ที่ได้รับการฝึกสอนล่วงหน้าบนชุดข้อมูล COCO และกำหนดค่าไฮเปอร์พารามิเตอร์ โดยมีรายละเอียดดังตารางที่ 4.6 นำมาเปรียบเทียบกับแบบจำลองที่ไม่ได้รับการฝึกสอนล่วงหน้าที่มีขนาดพารามิเตอร์เท่ากัน กำหนดค่าไฮเปอร์พารามิเตอร์เหมือนกัน ผลลัพธ์จากการทดลองเปรียบเทียบแบบจำลอง ถูกแสดงในตารางที่ 4.7

ตารางที่ 4.7 ผลการทดลองระหว่างแบบจำลองที่ฝึกล่วงหน้าและไม่ได้ฝึกล่วงหน้า

Model	Image size	mAP50-95 (Box)	mAP50-95 (Mask)
YOLOv8n-seg ฝึกสอนบนชุดข้อมูล COCO	640	0.9950	0.9904
YOLOv8n-seg ไม่ได้รับการฝึกสอน	640	0.9943	0.9803

พบว่าจากการเปรียบเทียบทั้งค่า mAP ของกล่องสี่เหลี่ยมล้อมรอบวัตถุ (mAP_{box}50-95) และ ค่า mAP ของมาสก์หรือกรอบล้อมตามรูปทรงวัตถุ (mAP_{mask}50-95) แบบจำลองที่ได้รับการฝึกล่วงหน้า นั้นสูงกว่าแบบที่ไม่ได้รับการฝึกเล็กน้อย ผลลัพธ์ดังตารางที่ 4.7 แต่ค่าของแบบจำลองที่ได้รับการฝึกสอนล่วงหน้าจะเข้าสู่ค่าที่ดีที่สุดรวดเร็วกว่าอย่างมีนัยสำคัญ ดังรูปที่ 4.3 โดยแบบที่ได้รับการฝึกสอนล่วงหน้าให้ค่าที่ดีที่สุดและเสถียรตั้งแต่ประมาณรอบที่ 25 ในขณะที่แบบที่ไม่ได้รับการฝึกสอนล่วงหน้า แบบจำลองเริ่มเสถียรช่วงประมาณรอบที่ 70



รูปที่ 4.3 กราฟค่า mAP50-95 ระหว่างแบบจำลองที่ฝึกล่วงหน้าและไม่ได้ฝึกล่วงหน้า

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

5.1 สรุปผลการวิจัย

งานวิจัยนี้เพื่อศึกษาการจำแนกสายพันธุ์ข้าวเปลือกไทย ประกอบด้วย ข้าวเปลือกพันธุ์หอมมะลิไทย (ข้าวขาวดอกมะลิ 105 และข้าวพันธุ์ กข 15) และข้าวเปลือกพันธุ์หอมปทุมธานี 1 โดยใช้เทคนิคการประมวลผลภาพและการเรียนรู้เชิงลึก ผู้วิจัยได้จัดเตรียมรูปภาพก่อนการประมวลผล สร้างแบบจำลอง และทดสอบกับเมล็ดข้าวเปลือก ได้ผลลัพธ์ดังนี้

5.1.1 การทำคำอธิบายประกอบรูปภาพ (Image annotation) ซึ่งคือการลากเส้นล้อมรอบตามรูปทรงวัตถุ เพื่อบ่งบอกถึงตำแหน่งและประเภทวัตถุ มีความสำคัญมากต่อการเรียนรู้หรือจดจำรูปภาพของแบบจำลอง ในงานศึกษานี้ใช้วิธีการหาขอบภาพโดยวิธีแคนนี่ (Canny edge detection) ร่วมกับวิธีการหาเส้นเค้าโครง (FindContour) เนื่องจากเป็นวิธีที่ใช้งานง่าย สามารถอธิบายได้ มีผลลัพธ์ชัดเจน และสามารถหาขอบได้แม่นยำซึ่งส่งผลต่อประสิทธิภาพแบบจำลอง

5.1.2 ผลการทดลองใช้แบบจำลองการแบ่งส่วนตัวอย่างอื่นที่ได้รับความนิยม พบว่าแบบจำลอง Mask R-CNN มีประสิทธิภาพดีที่สุด ให้ค่า mAP50-95 ที่ 0.9880 โดยใช้ความเร็วในการประมวลผลคือ 119.5 มิลลิวินาที สำหรับแบบจำลอง YOLACT นั้นรวดเร็วที่สุดคือใช้เวลา 72.1 วินาที แต่ให้ค่า mAP50-95 น้อยกว่าแบบอื่น ๆ อย่างมีนัยสำคัญ

5.1.3 แบบจำลองแบ่งส่วนตัวอย่าง YOLOv8 ทั้งหมด 5 แบบจำลอง ผลลัพธ์คือทุกแบบจำลองให้ค่า mAP50-95 ของกล่องล้อมรอบวัตถุเท่ากันที่ 0.9950 และ ค่า mAP ของกรอบล้อมรอบตามรูปทรงเมล็ดข้าวใกล้เคียงกันคือช่วง 0.9903-0.9910 แต่แบบจำลอง YOLOv8n-seg มีขนาดพารามิเตอร์เล็กที่สุด ใช้เวลาในการประมวลผลน้อยที่สุด จึงเป็นแบบจำลองที่ดีที่สุด

5.1.4 ผลการทดลองการปรับค่าไฮเปอร์พารามิเตอร์เพื่อให้ได้ค่าที่เหมาะสมที่สุด มีดังนี้ จำนวนรอบการฝึกแบบจำลอง 100 รอบ จำนวนรายการข้อมูลที่คำนวณในหนึ่งครั้ง เท่ากับ 16 อัตราการเรียนรู้ค่อย ๆ ลดในแต่ละรอบการฝึกสอนเริ่มจาก 0.01 จนถึง 0.0001 และ Optimizer คือ SGD จะส่งผลให้ได้แบบจำลองที่ดีที่สุด วัดโดยให้ค่า mAP50-95 สูงที่สุด ใช้เวลาฝึกสอนไม่มาก แบบจำลองมีความเสถียร

5.1.5 จากการเปรียบเทียบระหว่างแบบจำลอง YOLOv8n-seg ที่ถูกฝึกสอนล่วงหน้าบนชุดข้อมูล COCO และที่ไม่ถูกฝึกสอนล่วงหน้า โดยกำหนดค่าไฮเปอร์พารามิเตอร์ดังข้อ

5.1.3 ผลคือ แบบจำลองที่ได้รับการฝึกล่วงหน้าให้ค่า mAP สูงกว่า ลู่เข้าสู่ค่าที่ดีที่สุดเร็วกว่า และมีความเสถียรมากกว่าเมื่อฝึกสอน 100 รอบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากการศึกษาครั้งนี้พบว่าแบบจำลองแบ่งส่วนตัวอย่างที่เหมาะสมกับข้อมูลชุดนี้มากที่สุด พิจารณาถึงค่าการแลกเปลี่ยนระหว่างความเร็วในการประมวลผลและความแม่นยำ (Speed and Accuracy trade-off) คือ YOLOv8n-seg ที่ถูกกำหนดค่าไฮเปอร์พารามิเตอร์ คือรอบการฝึกแบบจำลอง 100 รอบ จำนวนรายการข้อมูลที่คำนวณในหนึ่งครั้ง เท่ากับ 16 อัตราการเรียนรู้ค่อย ๆ ลดในแต่ละรอบการฝึกสอนเริ่มจาก 0.01 จนถึง 0.0001 และ Optimizer คือ SGD ให้ค่า mAP50-95 ของกล่องล้อมรอบวัตถุที่ 0.9950 และ ค่า mAP ของกรอบล้อมรอบตามรูปทรงเมล็ดข้าวคือ 0.9904 ความเร็วในการประมวลผลอยู่ที่ 3.7 มิลลิวินาที

5.2 ข้อเสนอแนะ

- 5.2.1 สำหรับงานวิจัยนี้เก็บเมล็ดข้าวเปลือกโดยใช้โปรแกรม Easyrice ในการให้ผลเฉลย โดยโปรแกรมดังกล่าวถูกใช้เพื่อวิเคราะห์อัตลักษณ์และคัดแยกข้าวเปลือกสายพันธุ์ที่ต้องการออกจากสายพันธุ์อื่นผ่านรูปภาพสแกน ซึ่งกล่าวว่าแบบจำลองมีความแม่นยำที่ 95% ดังนั้นชุดข้อมูลรูปภาพของการศึกษาในครั้งนี้จึงขึ้นอยู่กับประสิทธิภาพและความแม่นยำของโปรแกรม Easyrice แต่เนื่องจากข้อจำกัดจึงไม่สามารถหาชุดข้อมูลที่มีความแม่นยำที่มากขึ้นได้
- 5.2.2 พันธุ์ข้าวหอมมะลิ (ข้าวขาวดอกมะลิ 105 และข้าวพันธุ์ กข 15) เก็บตัวอย่างเฉพาะที่ปลูกบริเวณพื้นที่จังหวัดนครราชสีมา และนครนายก ข้าวเปลือกพันธุ์ข้าวหอมปทุมธานี 1 เฉพาะที่ปลูกบริเวณพื้นที่จังหวัดชลบุรี และระยอง ดังนั้น หากนำแบบจำลองไปใช้งานจริง จึงควรเก็บตัวอย่างในจังหวัดที่หลากหลายยิ่งขึ้น
- 5.2.3 เนื่องจากแบบจำลองถูกฝึกสอนเฉพาะข้าวเปลือก 2 สายพันธุ์ คือ ข้าวหอมมะลิและข้าวหอมปทุมธานี 1 หากพิจารณาในทางปฏิบัติ เช่น การรับซื้อข้าวเปลือก และมีสายพันธุ์อื่นนอกจากที่ได้เรียนรู้ป้อนมา แบบจำลองจะไม่สามารถทำนายสายพันธุ์อื่นได้ อาจส่งผลต่อประสิทธิภาพของแบบจำลอง
- 5.2.4 การทำคำอธิบายประกอบรูปภาพ (Image annotation) อาจมีวิธีอื่นที่สามารถทำแล้วให้ผลได้แม่นยำเช่นเดียวกัน

เอกสารอ้างอิง

- [1] มหาวิทยาลัยราชภัฏสวนสุนันทา วิทยาเขตสมุทรสงคราม. 2565. ข้าวหอมปทุมธานี 1. สืบค้นเมื่อ 17 พฤศจิกายน 2565. [ระบบออนไลน์]. แหล่งที่มา : <https://skm.ssru.ac.th/th/news/view/wut049>.
- [2] ธนรัตน์ ศรีรุ่งเรือง และวีระศักดิ์ เลิศสิริโยธิน. 2555. “การวิเคราะห์องค์ประกอบทางเคมีที่แตกต่างกันของข้าวหอมมะลิและข้าวปทุมธานี 1 โดยใช้เทคนิคสเปกโตรสโคป.” *การประชุมวิชาการสมาคมวิศวกรรมเกษตรแห่งประเทศไทย*. 13(1) : 648-652.
- [3] กองวิจัยและพัฒนาข้าว กรมการข้าว. (ม.ป.ป.). *องค์ความรู้เรื่องข้าว*. สืบค้นเมื่อ 17 พฤศจิกายน 2565. [ระบบออนไลน์]. แหล่งที่มา : <https://newwebs2.ricethailand.go.th/webmain/rkb3/title-index.php-file=content.php&id=3-2.htm>.
- [4] ปวีณา เตจาคำ. 2562. “ผลผลิต คุณภาพเมล็ด และความหอมของข้าวหอมพื้นเมืองที่สูงพันธุ์ป้อนอมู.” *แกนเกษตร*. 47(2) : 317-326.
- [5] ศูนย์เมล็ดพันธุ์ข้าวสุรินทร์ กรมการข้าว. (ม.ป.ป.). *การทดสอบไอโอดีน*. สืบค้นเมื่อ 17 พฤศจิกายน 2565. [ระบบออนไลน์]. แหล่งที่มา : <https://srn-rsc.ricethailand.go.th>.
- [6] Dynamsoft. 2019. *Image Processing 101 Chapter 2.1: Image Enhancement*. Retrieved September 24, 2023. [Online]. Available : <https://www.dynamsoft.com/blog/insights/image-processing/image-processing-101-image-enhancement>.
- [7] Phyblas. 2563. *opencv-python เบื้องต้น บทที่ ๘: ตัวกรองคอนโวลูชันและการทำภาพเบลอ*. สืบค้นเมื่อ 24 กันยายน 2566. [ระบบออนไลน์]. แหล่งที่มา : <https://phyblas.hinaboshi.com/oshi08>.
- [8] คณะทรัพยากรธรรมชาติ มหาวิทยาลัยสงขลานครินทร์. (ม.ป.ป.). *การสำรวจข้อมูลระยะไกลและการแปลภาพถ่ายทางอากาศ*. สืบค้นเมื่อ 24 กันยายน 2566. [ระบบออนไลน์]. แหล่งที่มา : <https://natres.psu.ac.th/Department/EarthScience/remote1/chapter8.pdf>.
- [9] นายธีระพล ลิ้มศรัทธา. (ม.ป.ป.). *การประมวลผลมิติของภาพ เอกสารประกอบการสอน วิทยาลัยเซาธ์อีสท์บางกอก บางนา กรุงเทพฯ*. สืบค้นเมื่อ 24 กันยายน 2566. [ระบบออนไลน์]. แหล่งที่มา : http://www.theerapone.com/sbc/courses/openCV/doc/03.Edges_1.2.pdf.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- [10] ยุภาวดี อินถาเครือ และรัฐสิทธิ์ สุขะหุต. 2554. “การสกัดคุณลักษณะของภาพโดยใช้เทคนิคโซเบลสำหรับการแทนรูปภาพอักษรเบรลล์.” *วารสารเทคโนโลยีสารสนเทศ*. 7(13) : 1-6
- [11] Advance Innovation Centre (AIC) Open Knowledge ภาควิชาวิศวกรรมไฟฟ้า – มหาวิทยาลัยบูรพา. (ม.ป.ป.). **Morphology Transformations**. สืบค้นเมื่อ 24 กันยายน 2566. [ระบบออนไลน์]. แหล่งที่มา : <https://docs.aic-eeec.com/artificial-intelligence-ai/computer-vision-for-python/lab-2-basic-cv/morphology-transformations>.
- [12] Educative. (n.d.). **Edge detection vs. contour detection in computer vision**. Retrieved September 24, 2023. [Online]. Available : <https://www.educative.io/answers/edge-detection-vs-contour-detection-in-computer-vision>.
- [13] Baeldung. (n.d.). **What Are Contours in Computer Vision?**. Retrieved September 24, 2023. [Online]. Available : <https://www.baeldung.com/cs/computer-vision-contours>.
- [14] Akula Hemanth Kumar. 2020. **Contours in Images**. Retrieved September 24, 2023. [Online]. Available : <https://medium.com/swlh/contours-in-images-a58b4c12c0ff>.
- [15] Surapong Kanoktipsatharporn. 2562. **Image Segmentation คืออะไร**. สืบค้นเมื่อ 24 กันยายน 2566. [ระบบออนไลน์]. แหล่งที่มา : <https://www.bualabs.com/archives/835/what-is-image-segmentation-semantic-segmentation-camvid-machine-learning-unet-deep-learning-image-segmentation-ep-1/>.
- [16] นคร ละลอกน้ำ. 2564. “การใช้เทคนิคการทำเหมืองข้อมูลเพื่อคัดกรองเด็กที่มีความบกพร่องทางการเรียนรู้.” *วิทยานิพนธ์วิศวกรรมศาสตรมหาบัณฑิต สาขาวิชาวิศวกรรมคอมพิวเตอร์และโทรคมนาคม วิทยาลัยนวัตกรรมการด้านเทคโนโลยีและวิศวกรรมศาสตร์ มหาวิทยาลัยธุรกิจบัณฑิต*.
- [17] สมาคมโปรแกรมเมอร์ไทย. 2561. **Deep learning คืออะไร ?**. สืบค้นเมื่อ 24 กันยายน 2566. [ระบบออนไลน์]. แหล่งที่มา : <https://www.thaiprogrammer.org/2018/12/deep-learning-%E0%B8%84%E0%B8%B7%E0%B8%AD%E0%B8%AD%E0%B8%B0%E0%B9%84%E0%B8%A3/>.

- [18] พิมพ์ ชีวาประกอบกิจ. 2562. “การปรับปรุงประสิทธิภาพในการจำแนกภาพด้วยโครงข่ายประสาทแบบคอนโวลูชันโดยใช้เทคนิคการเพิ่มภาพ.” *TNI Journal of Engineering and Technology*. 7(1) : 60-61.
- [19] พุฒิพงศ์ จันทร์แจ่ม. 2563. “การปรับปรุงประสิทธิภาพของกระบวนการบริการลูกค้าโดยใช้การวิเคราะห์จากกล้องวงจรปิด” วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต การวิเคราะห์ธุรกิจและวิทยาการข้อมูล คณะสถิติประยุกต์ สถาบันบัณฑิตพัฒนบริหารศาสตร์.
- [20] Teddy Sangtongdee. 2564. **สร้างโมเดล Object Detection สำหรับ Aqua Pests แบบมี Pipeline**. สืบค้นเมื่อ 3 กันยายน 2566. [ระบบออนไลน์]. แหล่งที่มา : <https://medium.com/botnoi-classroom/สร้างโมเดล-object-detection-สำหรับ-aqua-pets-แบบมี-pipeline-41ca7411c2e0>.
- [21] Nabin Sharma. 2023. “Parking Time Violation Tracking Using YOLOv8 and Tracking Algorithms.” *Sensors*. 23(13) : 5843.
- [22] Parab, C. U., Mwitta, C., Hayes, M., Schmidt, J. M., Riley, D., Fue, K., ... & Rains, G. C. 2022. “Comparison of Single-Shot and TwoShotDeep Neural Network Models for Whitefly Detection in IoT Web Application.” *AgriEngineering*. 4(2) : 507-522
- [23] Sanjog Tamang. 2022. “Enhancing COVID-19 Safety: Exploring YOLOv8 Object Detection for Accurate Face Mask Classification.” *International Journal of INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING*. 11(2) : 892-897.
- [24] V7Labs. 2023. **YOLO: Algorithm for Object Detection Explained**. Retrieved October 15, 2023. [Online]. Available : <https://www.v7labs.com/blog/yolo-object-detection>.
- [25] datacamp. 2022. **YOLO Object Detection Explained**. Retrieved October 15, 2023. [Online]. Available : <https://www.datacamp.com/blog/yolo-object-detection-explained>.
- [26] ultralytics. 2023. **Brief summary of YOLOv8 model structure**. Retrieved October 15, 2023. [Online]. Available : <https://github.com/ultralytics/ultralytics/issues/189>.
- [27] EASYRICE. 2023. **เกี่ยวกับ EASYRICE**. สืบค้นเมื่อ 21 พฤศจิกายน 2566. [ระบบออนไลน์]. แหล่งที่มา : <https://easyrice.ai/th/about-us>.