

การวิเคราะห์เปรียบเทียบโมเดลการเรียนรู้เชิงลึกสำหรับการสกัดอาคารจาก
ภาพถ่ายดาวเทียมความละเอียดสูง

COMPARATIVE ANALYSIS OF DEEP LEARNING MODELS FOR
BUILDING EXTRACTION FROM HIGH-RESOLUTION SATELLITE
IMAGERY



วิทยานิพนธ์นี้สำหรับการศึกษิตตามหลักสูตร
ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์
ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

พ.ศ. 2567

KMITL-2024-SC-M-002-058

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

COMPARATIVE ANALYSIS OF DEEP LEARNING MODELS FOR
BUILDING EXTRACTION FROM HIGH-RESOLUTION SATELLITE
IMAGERY



A THESIS SUBMITTED IN FULFILLMENT OF THE REQUIREMENT FOR THE
DEGREE OF MASTER OF SCIENCE IN COMPUTER SCIENCE
DEPARTMENT OF COMPUTER SCIENCE SCHOOL OF SCIENCE
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG
2024

KMITL-2024-SC-M-002-058

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



COPYRIGHT 2024

SCHOOL OF SCIENCE

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อวิทยานิพนธ์	การวิเคราะห์เปรียบเทียบโมเดลการเรียนรู้เชิงลึกสำหรับการสกัดอาคารจากภาพถ่ายดาวเทียมความละเอียดสูง
ชื่อนักศึกษา	นายเดชสิทธิ์ ชื่อประเสริฐ
รหัสนักศึกษา	61605035
ปริญญา	วิทยาศาสตรมหาบัณฑิต (วิทยาการคอมพิวเตอร์)
ภาควิชา	วิทยาการคอมพิวเตอร์
พ.ศ.	2567
อาจารย์ที่ปรึกษาวิทยานิพนธ์	ผู้ช่วยศาสตราจารย์ ดร.ศรัณย์ อินทโกสุม

บทคัดย่อ

งานวิจัยนี้ นำเสนอแนวทางการสกัดอาคารจากภาพถ่ายดาวเทียมความละเอียดสูงของบริษัท Google โดยการศึกษามุ่งเน้นการเปรียบเทียบโมเดลการเรียนรู้เชิงลึก ได้แก่ ยูเน็ต (U-Net) อาร์โอยูเน็ต (RIU-Net) ยูเน็ตพลัสพลัส (U-Net++) เรสยูเน็ต (Res-U-Net) และดีพแล็บวีทีพีพลัส (DeepLabV3+) ซึ่งทดลองเรียนรู้ด้วยชุดข้อมูลที่ผ่านมากระบวนการประมวลผลล่วงหน้า ทั้งนี้ ระหว่างการเรียนรู้ได้ใช้ดัชนีชี้วัดความคล้ายคลึง Intersection-over-Union (IoU) และ Dice Similarity Coefficient (DSC) จากผลการทดลอง โมเดลที่มีประสิทธิภาพดีที่สุดในกระบวนการสกัดอาคารจากภาพถ่ายดาวเทียม คือ Res-U-Net และ DeepLabV3+ โดย Res-U-Net เป็นโครงสร้างปรับปรุงของโมเดลรุ่นดั้งเดิมอย่าง U-Net โดยมีการประยุกต์ใช้แบ็กโบน (Backbone) เสริมเข้าช่วยเพื่อให้การตรวจจับคุณสมบัติสำคัญของรูปภาพดียิ่งขึ้น และได้รับคะแนนจากดัชนีชี้วัด F1 ที่ 85.43% เมื่อใช้ชุดข้อมูลภาพสีในการเรียนรู้ ทำนองเดียวกัน DeepLabV3+ ทำประสิทธิภาพสูงสุดด้วยชุดข้อมูลภาพสีแบบปรับปรุง ได้คะแนน F1 ที่ 85.18% ด้วยเหตุนี้ การวิจัยได้ค้นพบหนึ่งในปัจจัยสำคัญของการตรวจจับคุณสมบัติวัตถุในรูปภาพคือค่าสี อันอยู่ภายใต้กระบวนการแบ่งส่วนความหมายจากภาพถ่ายดาวเทียมเพื่อให้ได้ผลลัพธ์ที่แม่นยำ ท้ายที่สุด การวิจัยมีส่วนช่วยในการปรับปรุงวิธีการในการระบุสิ่งปลูกสร้างจากภาพถ่ายดาวเทียม เอื้อประโยชน์ต่อการวางผังเมืองและคาดการณ์ประมาณประชากร และการประยุกต์ใช้การจัดการภัยพิบัติ

คำสำคัญ : การแบ่งส่วนความหมาย การประมวลผลรูปภาพ การระบุสิ่งปลูกสร้าง การเรียนรู้เชิงลึก ภาพถ่ายดาวเทียม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Thesis Title	Comparative Analysis of Deep Learning Models for Building Extraction from High-Resolution Satellite Imagery
Student Name	Mr.Tachasit Chueprasert
Student ID	61605035
Degree	Master of Science (Computer Science)
Department	Computer Science
Year	2024
Thesis Advisor	Assistant Professor Dr.Sarun Intagosum

Abstract

This research proposes an approach to extract buildings from Google's satellite imagery. The study compares various deep learning models (U-Net, RIU-Net, U-Net++, Res-U-Net, DeepLabV3+) experimenting on pre-processed datasets. The models were trained using similarity metrics of Intersection over Union (IoU) and Dice Similarity Coefficient (DSC). The best-performing models among the compared segmentation techniques were Res-U-Net and DeepLabV3+. Res-U-Net, an enhanced version of the traditional U-Net model that incorporates residual connections for improved feature propagation, achieved an F1 score of 85.43% when using the RGB dataset. Similarly, DeepLabV3+ also achieved high performance with the Enhanced RGB dataset, obtaining an F1 score of 85.18% after applying pre-processing techniques. This research highlights the significance of color as a dominant feature for accurate building extraction from satellite images. The findings contribute to improved methodologies for building identification, benefiting urban planning, and disaster management applications.

Keywords : semantic segmentation, image processing, building extraction, deep learning, satellite imagery

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

วิทยานิพนธ์ การวิเคราะห์เปรียบเทียบรูปแบบการเรียนรู้เชิงลึกสำหรับระบุสิ่งปลูกสร้างจากภาพถ่ายดาวเทียมความละเอียดสูง เกิดขึ้นและดำเนินการจนสำเร็จลงได้ด้วยวิสัยทัศน์ ความเพียร และความร่วมมือประกอบการชี้แนะโดยบุคลากรสายวิชาการ บุคลากรสายสนับสนุนวิชาการจากสถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง บุคลากรสายวิชาการจากมหาวิทยาลัยมหิดล ตลอดจนหน่วยงานรัฐ และบริษัทเอกชนที่มีอากล่าวานามและกิตติคุณไว้ได้หมด ณ ที่นี้

ผู้วิจัยขอขอบพระคุณ คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง โดยเฉพาะท่านคณบดี รศ. ดร.สุธี ชูดีไพจิตร ที่ได้อนุมัติและรับรองการวิจัยสู่การเป็นวิทยานิพนธ์ฉบับนี้ ขอขอบพระคุณเหล่าบุคลากรสายวิชาการ สาขาวิชาวิทยาการคอมพิวเตอร์ ผศ. ดร.อนันตพร หาราชคุณาฉัย อาจารย์บัณฑิตประจำหลักสูตร ผศ. ดร.ศรีณย์ อินทโกสุม อาจารย์ที่ปรึกษาวิทยานิพนธ์ ดร.อัคเดช อุดมชัยพร และผศ. ดร.วรางคณา กัมปาน อาจารย์ประจำภาควิชา ผู้ให้คำปรึกษาและสนับสนุนการดำเนินงานทุกขั้นตอนในการวิจัยจนสามารถดำเนินการสำเร็จลงได้เป็นอย่างดี นอกจากนี้ ยังช่วยต่อยอดองค์ความรู้และเสริมสร้างพัฒนาทักษะคู่ความสามารถให้กับผู้วิจัย ขอขอบคุณบุคลากรสายสนับสนุนวิชาการ ผู้คอยทำหน้าที่ติดต่อ ประสานงาน และให้คำชี้แนะ ตลอดจนวิทยานิพนธ์บรรลุลผลตามขั้นตอนทุกประการ

ขอขอบพระคุณ ผศ. ดร.ธนสนี เพียรตระกูล ประธานกรรมการ ผู้ทรงคุณวุฒิจากภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยมหิดล ที่ได้ให้เกียรติมาร่วมให้คำปรึกษาชี้แนะแนวทางในการทำวิจัย รวมทั้งตรวจสอบขั้นตอนการทำวิจัยให้อยู่ในขอบเขตคู่ทางที่เป็นอันควร

ขอขอบพระคุณหน่วยงานภาครัฐ สำนักงานสถิติแห่งชาติ สังกัดกระทรวงดิจิทัลเพื่อเศรษฐกิจและสังคม ที่ได้เอื้อข้อมูลสถิติอันมีประโยชน์ต่อการทำวิจัย ขอขอบคุณบริษัทเอกชน Google ที่ได้เผยแพร่ข้อมูลภาพถ่ายดาวเทียมให้ผู้วิจัยได้นำมาเป็นข้อมูลสำคัญในการดำเนินเงิน รวมถึงทรัพยากรซึ่งใช้ในการประมวลผลผ่านเครือข่ายอินเทอร์เน็ต

ท้ายที่สุดนี้ ผู้วิจัยคาดหวังให้วิทยานิพนธ์ฉบับนี้เป็นส่วนสำคัญในการเสริมสร้างองค์ความรู้ และได้เผยแพร่ส่งต่อการค้นพบเพื่อเป็นประโยชน์ต่องานวิจัยอื่นที่เกี่ยวข้อง หรือแม้แต่นำกระบวนการไปประยุกต์ใช้เพื่อต่อยอดอย่างสร้างสรรค์สืบไป

นายเตชสิทธิ์ ชื้อประเสริฐ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

	หน้า
บทคัดย่อภาษาไทย	ก
บทคัดย่อภาษาอังกฤษ	ข
กิตติกรรมประกาศ	ค
สารบัญ.....	ง
สารบัญตาราง.....	ก
สารบัญรูป.....	ก
บทที่ 1 บทนำ.....	1
1.1 ความเป็นมาและความสำคัญของปัญหา.....	1
1.2 วัตถุประสงค์ของงานวิจัย.....	2
1.3 ขอบเขตงานวิจัย.....	3
1.4 ประโยชน์ที่คาดว่าจะได้รับ	3
1.5 ระเบียบวิธีวิจัย.....	3
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง.....	5
2.1 การสำมะโนประชากรและเคหะในประเทศไทย	5
2.2 ทฤษฎีและเทคโนโลยีการสำรวจระยะไกล.....	7
2.2.1 การสำรวจระยะไกล.....	7
2.2.2 หลักการและกระบวนการของดาวเทียม	8
2.2.3 คุณสมบัติของดาวเทียมสำรวจทรัพยากรโลก.....	9
2.2.4 ผู้ให้บริการข้อมูลภาพถ่ายดาวเทียมและคุณสมบัติของดาวเทียม.....	10
2.2.5 คุณลักษณะภาพถ่ายดาวเทียม.....	15
2.3 ทฤษฎีเกี่ยวกับการประมวลผลภาพ.....	18
2.3.1 การแปลงภาพนาล็อกเป็นภาพดิจิทัล.....	18
2.3.2 ภาพดิจิทัลและประเภทของภาพดิจิทัล	21
2.3.3 ประเภทของภาพดิจิทัล.....	23
2.3.4 คุณลักษณะของภาพ	25
2.3.5 การประมวลผลภาพดิจิทัล	27
2.4 ทฤษฎีแบบจำลองปัญญาประดิษฐ์และโมเดลการเรียนรู้เชิงลึก	29
2.4.1 การเรียนรู้ของเครื่อง (Machine Learning).....	29
2.4.2 การเรียนรู้เชิงลึก (Deep Learning).....	30
2.4.3 การถ่ายโอนความรู้ (Transfer Learning).....	33

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.4.4	โมเดลการเรียนรู้เชิงลึกยูเน็ต (U-Net)	34
2.4.5	โมเดลการเรียนรู้เชิงลึกอาร์ไอยูเน็ต (RIU-Net)	35
2.4.6	โมเดลการเรียนรู้เชิงลึกยูเน็ตพลัสพลัส (U-Net++)	37
2.4.7	โมเดลการเรียนรู้เชิงลึกเรสยูเน็ต (Res-U-Net)	37
2.4.8	โมเดลการเรียนรู้เชิงลึกดีพีแล็บวีทีพีพลัส (DeepLabV3+)	39
2.5	ดัชนีชี้วัดความคล้ายคลึงระหว่างรูปภาพ	40
2.5.1	เมทริกซ์ชี้วัดความสับสน (Confusion Matrix)	41
2.5.2	Jaccard Index (Intersection Over Union; IoU)	42
2.5.3	Dice Similarity Coefficient (DSC)	43
2.6	งานวิจัยที่เกี่ยวข้อง	43
2.6.1	การสกัดอาคารจากภาพถ่ายดาวเทียมด้วยการใช้ฐานฐานวิทยาทางคณิตศาสตร์แบบ อัตโนมัติ	43
2.6.2	การสกัดอาคารจากภาพถ่ายดาวเทียมด้วยโมเดลยูเน็ต	45
2.6.3	การสกัดอาคารจากภาพถ่ายดาวเทียมด้วยโมเดลอาร์ไอยูเน็ต	46
2.6.4	การสกัดอาคารจากภาพถ่ายดาวเทียมด้วยโมเดลเรสยูเน็ต	47
2.6.5	การสกัดอาคารจากภาพถ่ายดาวเทียมด้วยโมเดลดีพีแล็บวีทีพีพลัส	47
บทที่ 3	วิธีการดำเนินงานวิจัย	49
3.1	การกำหนดพื้นที่ศึกษาและการรวบรวมข้อมูล	49
3.1.1	ลักษณะทางภูมิศาสตร์และการปกครองจังหวัดเลย	50
3.1.2	การรวบรวมข้อมูลภาพถ่ายดาวเทียมจาก Google Earth ด้วย SAS.Planet	52
3.2	การจัดทำชุดคำตอบ	52
3.3	การจัดเตรียมชุดข้อมูล (Data Preparation)	53
3.3.1	การประมวลผลข้อมูลล่วงหน้า (Data Preprocessing)	53
3.3.2	การเสริมชุดข้อมูล (Data Augmentation)	55
3.4	การออกแบบขั้นตอนการสกัดอาคารจากภาพถ่ายดาวเทียม	55
3.4.1	วิธีประมวลผลภาพแบบสัณฐานวิทยา	56
3.4.2	วิธีประยุกต์ใช้โมเดลการเรียนรู้เชิงลึก	57
บทที่ 4	ผลการวิจัยและการอภิปรายผล	58
4.1	การแบ่งส่วนอาคารด้วยเทคนิคสัณฐานวิทยา	58
4.1.1	การลบพื้นที่ป่าด้วยกระบวนการแปลงสีภาพ	58
4.1.2	การจัดเตรียมข้อมูลภาพ	60
4.1.3	การระบุส่วนที่สว่างและมีดของพื้นที่ในภาพด้วยเทคนิคสัณฐานวิทยา	61
4.1.4	การลบพื้นที่ป่าด้วยกระบวนการแปลงสีภาพ	63

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.1.5 การประยุกต์ใช้ Threshold เพื่อแบ่งส่วนอาคาร.....	64
4.2 การสกัดอาคารด้วยโมเดลการเรียนรู้เชิงลึก.....	67
4.2.1 การวัดผลความคล้ายคลึง.....	70
4.2.2 วิเคราะห์ผลและเปรียบเทียบ.....	72
4.2.3 ระยะเวลาที่ใช้ในการฝึกฝน.....	77
บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ.....	80
5.1 สรุปผลการวิจัย.....	80
5.2 ข้อเสนอแนะ.....	82
ภาคผนวก.....	83
เอกสารอ้างอิง.....	84
ประวัติผู้เขียน.....	92



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญญภาพ

รูปที่ 2.1 แบบสัมภาษณ์สำมะโนประชากรและเคหะประเทศไทย ปี พ.ศ. 2553	6
รูปที่ 2.2 เปรียบเทียบกำลังขยายต่ำ (ซ้าย) และกำลังขยายสูง (ขวา) ของซอฟต์แวร์ Google Earth 10	
รูปที่ 2.3 การแปลงคลื่นสัญญาณที่ได้และจัดเก็บในรูปแบบภาพดิจิทัล 2 มิติ.....	16
รูปที่ 2.4 สเปกตรัมแม่เหล็กไฟฟ้า (Electromagnetic Spectrum, Louis E. Keiner, Coastal Carolina University, ไม่ทราบปีที่เผยแพร่).....	17
รูปที่ 2.5 เปรียบเทียบรายละเอียดภาพจากรายละเอียดสูง (ซ้าย) ไปยังรายละเอียดต่ำ (ขวา)	17
รูปที่ 2.6 ภาพถ่ายดาวเทียมจาก Google เปรียบเทียบสภาพเมือง Kamaishi ประเทศญี่ปุ่นก่อน (ซ้าย) และหลัง (ขวา) ประสบภัยพิบัติคลื่นสึนามิ.....	18
รูปที่ 2.7 สเปกตรัมของแสงที่มนุษย์สามารถสัมผัสด้วยเซลล์รูปกรวยในเรตินา	19
รูปที่ 2.8 ภาพกระบวนการ Image Sampling และ Quantization.....	19
รูปที่ 2.9 แสดงตัวอย่างการทำกระบวนการสุ่มตัวอย่าง	20
รูปที่ 2.10 ภาพตัวอย่างกรณี Sampling rate ตามทฤษฎี	20
รูปที่ 2.11 แสดงตัวอย่างการจัดระดับสัญญาณ (Quantization)	21
รูปที่ 2.12 ตัวอย่างภาพบิตแมป (Bitmap Image).....	23
รูปที่ 2.13 ตัวอย่างภาพเวกเตอร์ (Vector Image)	23
รูปที่ 2.14 ตัวอย่างภาพขาวดำ (Binary Image).....	24
รูปที่ 2.15 ตัวอย่างภาพระดับสีเทา (Gray Scale Image).....	24
รูปที่ 2.16 ตัวอย่างภาพสี (Color Image)	25
รูปที่ 2.17 ตัวอย่างฮิสโตแกรมของภาพ (Image Histogram).....	26
รูปที่ 2.18 ตัวอย่างโครงสร้างเพอร์เซปตรอน (A Perceptron Structure)	30
รูปที่ 2.19 เปรียบเทียบค่าของผลลัพธ์ที่ได้จากสมการฟังก์ชันซิกมอยด์ (Sigmoid) และเพอร์เซปตรอนแบบดั้งเดิม (Perceptron, Unit step: threshold) ด้วยรูปแบบกราฟเส้น.....	31
รูปที่ 2.20 ตัวอย่างโครงสร้างสถาปัตยกรรมโมเดล U-Net.....	34
รูปที่ 2.21 โครงสร้างโครงข่ายในส่วนของการเข้ารหัส (Encoder) ของ RIU-Net: (a) Module A, (b) Module B, (c) Module C, (d) Reduction A, and (e) Reduction B	35
รูปที่ 2.22 โครงสร้างโครงข่ายในส่วนของการถอดรหัส (Decoder) ของ RIU-Net: (a) Module D, and (b) Upsampling module.....	36

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 2.23 ตัวอย่างโครงสร้างสถาปัตยกรรมโมเดล RIU-Net (Encoder/Decoder อิงตามรูปที่ 2.21 และ 2.22 ตามลำดับ).....36

รูปที่ 2.24 ตัวอย่างโครงสร้างสถาปัตยกรรม U-Net++37

รูปที่ 2.25 ตัวอย่างโครงสร้างสถาปัตยกรรม Res-U-Net38

รูปที่ 2.26 ตัวอย่างโครงสร้างสถาปัตยกรรม DeepLabV3+39

รูปที่ 2.27 เปรียบเทียบตัวอย่างโครงสร้างสถาปัตยกรรมในมุมมองตัวเข้ารหัสและถอดรหัส โดยภาพ a) คือโครงสร้างสถาปัตยกรรมบนแนวคิด Spatial Pyramid Pooling ของ DeepLabV3 b) คือโครงสร้างสถาปัตยกรรมบนแนวคิด Encoder-Decoder ของ U-Net และ c) คือโครงสร้างสถาปัตยกรรมบนแนวคิด a) และ b) ของ DeepLabV3+40

รูปที่ 2.28 (a) ลบส่วนที่มีความผิดเพี้ยนตามอัตราส่วนแกน (b) ลบส่วนที่ผิดเพี้ยนตามความโค้งมนหรือกลม (c) ลบส่วนที่ผิดเพี้ยนตามขนาด.....44

รูปที่ 2.29 เปรียบเทียบชุดคำตอบกับผลการทำนายการแบ่งส่วนอาคารด้วยโมเดลที่ผ่านการเรียนรู้เชิงลึกด้วยโครงสร้าง U-Net45

รูปที่ 2.30 เปรียบเทียบผลลัพธ์การทำนายของแต่ละโมเดลในงานวิจัยของ Saritürk และผู้ร่วมวิจัย.46

รูปที่ 3.1 แผนภาพแสดงขั้นตอนการดำเนินงานวิจัย.....49

รูปที่ 3.2 แผนที่ท้องจังหวัดเลยแสดงเส้นทาง ขอบเขต และแม่น้ำในจังหวัด50

รูปที่ 3.3 จำนวนประชากรและบ้านในแต่ละตำบลของอำเภอเมืองเลย.....51

รูปที่ 3.4 แผนที่ (ซ้าย) และภาพถ่ายดาวเทียม (ขวา) ของตำบลศรีสองรักจาก Google Map และ Google Earth.....51

รูปที่ 3.5 ตัวอย่างภาพชุดคำตอบ (Ground-truth) โดยชุดภาพ (A) แสดงถึงคุณลักษณะภาพถ่ายที่ปรากฏอาคารหนาแน่น ชุดภาพ (B) ปรากฏอาคารเล็กน้อย และชุดภาพ (C) ไม่ปรากฏอาคาร52

รูปที่ 3.6 ตัวอย่างภาพชุดข้อมูล ก) RGB ข) Enh-RGB ค) Gray และ ง) Enh-Gray ตามลำดับ54

รูปที่ 3.7 ตัวอย่างการเสริมข้อมูลด้วยการพลิกภาพแบบแนวตั้ง แนวนอน การหมุน 90 องศาทวนเข็มนาฬิกา และตามเข็มนาฬิกา55

รูปที่ 3.8 ขั้นตอนกระบวนการสกัดอาคารด้วยวิธีพื้นฐานวิทยา.....56

รูปที่ 3.9 ขั้นตอนกระบวนการสกัดอาคารด้วยวิธีประยุกต์ใช้โมเดลการเรียนรู้เชิงลึก57

รูปที่ 4.1 ตัวอย่าง Gaussian blur (filter) ซึ่งใช้ Kernel ขนาด 5x559

รูปที่ 4.2 ผลลัพธ์จากการดำเนินการใช้ Gaussian blur (ซ้าย) การแปลงค่าสีเป็น HSV (กลาง) และการลบพื้นที่สีเขียว (ขวา).....59

รูปที่ 4.3 ภาพถ่ายดาวเทียมและหลังดำเนินการใช้เทคนิคลักษณะพื้นฐานวิทยาแบบ Dilation60

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 4.4 ภาพถ่ายดาวเทียมกระจายตัวของค่าสีระหว่างก่อนและหลังประยุกต์ใช้เทคนิค CLAHE....61

รูปที่ 4.5 ผลลัพธ์จากการใช้ตัวดำเนินการ Open (ซ้าย) และ close (ขวา).....62

รูปที่ 4.6 ผลลัพธ์ก่อน (ซ้าย) และหลังการปรับปรุงคุณภาพโดยใช้วิธีตามสมการที่ 6 (กลาง) และการประยุกต์ลักษณะพื้นฐานวิทยาแบบ open (ขวา).....62

รูปที่ 4.7 กลุ่มสีซึ่งถูกจัดเรียงด้วย K-means (ซ้าย) และค่ากลางของแต่ละกลุ่มสีทั้ง 5 กลุ่ม(ขวา)....63

รูปที่ 4.8 ผลลัพธ์ก่อนและหลัง (ขวา) ประยุกต์ใช้ K-means63

รูปที่ 4.9 ผลลัพธ์ก่อนและหลัง (ขวา) การประยุกต์ใช้ Binary Threshold64

รูปที่ 4.10 ภาพในกระบวนการหาผลลัพธ์จากการแบ่งส่วนอาคารที่มีแสงสว่างน้อย65

รูปที่ 4.11 การเปรียบเทียบภาพระหว่างต้นฉบับ ชุดคำตอบ และหลังการแบ่งส่วนด้วยลักษณะพื้นฐานทางวิทยา.....65

รูปที่ 4.12 การเปรียบเทียบภาพระหว่างต้นฉบับและคำตอบ (บน) กับผลลัพธ์การแบ่งส่วนอาคารที่สว่างและมีด (ล่าง).....66

รูปที่ 4.13 การแสดงผลข้อมูลวิจัยผ่านทางแดชบอร์ด.....69

รูปที่ 4.14 การประเมินเปรียบเทียบ Res-U-Net (ResNet-50) ที่ฝึกฝนและทดสอบด้วยชุดข้อมูลภาพสี RGB แสดงค่าความแม่นยำและข้อผิดพลาดในการทำนายด้วยคะแนน F1, IoU และ DSC71

รูปที่ 4.15 ผลคะแนน F1 และคะแนนอื่น ๆ ของโมเดล 9 อันดับแรก72

รูปที่ 4.16 การเปรียบเทียบการสกัดอาคารโดยโมเดล Res-U-Net (ResNet-50; RGB/DSC) และ โมเดล DeepLabV3+ (ResNet-50; Enh-RGB/DSC)73

รูปที่ 4.17 เปรียบเทียบผลคะแนนโมเดล Res-U-Net (ResNet-50/RGB) ที่ใช้ดัชนีต่างกัน.....73

รูปที่ 4.18 ผลคะแนน F1 ของโมเดล 9 อันดับแรกและแผนภาพเปรียบเทียบคะแนนเฉลี่ย F1 โดยอิงจากชุดข้อมูล ดัชนีชี้วัด และการประยุกต์ใช้และไม่ใช้แบ็กโบน.....74

รูปที่ 4.19 ตัวอย่างผลลัพธ์ที่ได้รับจากแต่ละโมเดลโดยใช้ดัชนี DSC ระหว่างการฝึกฝน ประกอบกับตัวเลขแสดงคะแนน F1 ของแต่ละภาพ75

รูปที่ 4.20 ตัวอย่างผลลัพธ์ที่ได้รับจากแต่ละโมเดลโดยใช้ดัชนี IoU ระหว่างการฝึกฝน ประกอบกับตัวเลขแสดงคะแนน F1 ของแต่ละภาพ.....76

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตารางที่ 2.1	อุปกรณ์บันทึกข้อมูลของดาวเทียม LANDSAT – 8.....	11
ตารางที่ 2.2	อุปกรณ์บันทึกข้อมูลของดาวเทียม SENTINEL-2A.....	12
ตารางที่ 2.3	อุปกรณ์บันทึกข้อมูลของดาวเทียม SPOT 6 และ 7	13
ตารางที่ 2.4	อุปกรณ์บันทึกข้อมูลของดาวเทียม Pleiades 1A และ 1B.....	13
ตารางที่ 2.5	อุปกรณ์บันทึกข้อมูลของดาวเทียม QUICKBIRD	15
ตารางที่ 4.1	เปรียบเทียบการวัดประสิทธิภาพของแต่ละโมเดลโดยอิงคะแนน F1 จากมากไปน้อย...68	
ตารางที่ 4.2	เปรียบเทียบเวลาที่ใช้ประมวลผลในการฝึกฝนของแต่ละโมเดล.....	77
ตารางที่ 4.3	เปรียบเทียบผลลัพธ์และจัดลำดับใหม่ด้วยค่า F1 จากสมการที่ (4.1) พร้อมระยะเวลาที่ใช้ในการฝึกฝน	78



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

การสกัดอาคารจากภาพถ่ายดาวเทียม หรือการระบุสิ่งปลูกสร้างที่มีหลังคาปรากฏอย่างเด่นชัดในภาพถ่ายดาวเทียม มีการนำไปใช้งานซึ่งพบได้หลากหลายรูปแบบทั้งภาครัฐและเอกชน โดยตัวอย่างวัตถุประสงค์ อาทิ การวางแผนจัดการภัยพิบัติ การวางแผนผังเมือง การติดตามการขยายตัวของพื้นที่อยู่อาศัย การประมาณจำนวนประชากรตามขนาดและจำนวนของสิ่งปลูกสร้าง สำหรับประเทศไทย การทำสำมะโนประชากรจะเกิดขึ้นทุก 10 ปี โดยมีผู้รับผิดชอบคือ สำนักงานสถิติแห่งชาติ และสำนักงานฯ ได้ทำการสำมะโนประชากรครั้งแรกในปี พ.ศ. 2503 และตั้งแต่ปี พ.ศ. 2513 เป็นต้นมา ได้ทำการสำมะโนเคหะพร้อมกับการทำสำมะโนประชากรร่วมด้วย ทว่าอย่างไรก็ตาม นับตั้งแต่เดือนธันวาคม พ.ศ. 2562 ได้เกิดการแพร่ระบาดของไวรัสโคโรนาสายพันธุ์ใหม่ไปทั่วโลก ชื่อว่า ไวรัสโคโรนา 2019 (COVID-19; โควิด-19) ส่งผลให้การดำเนินการสำมะโนประชากรและเคหะได้รับผลกระทบในเวลาถัดมา เนื่องจากมาตรการเรื่องความปลอดภัยด้านสุขภาพ ทำให้การดำเนินงานที่จำเป็นต้องลงพื้นที่ปฏิบัติไม่สามารถดำเนินการได้โดยทั่วกัน และสะท้อนถึงจุดอ่อนของวิธีการแบบดั้งเดิม

จากตัวอย่างสถานการณ์แพร่ระบาดดังกล่าว ชี้ให้เห็นถึงความท้าทายในการรวบรวมข้อมูลประชากรและเคหะสถาน ด้วยเหตุนี้ การประยุกต์ใช้เทคโนโลยีภาพถ่ายดาวเทียมจึงถูกนำมาพิจารณา เหตุเพราะมีความสามารถในการเก็บรวบรวมข้อมูลทางภูมิศาสตร์ ประกอบกับความก้าวหน้าสมัยใหม่ของเทคโนโลยี ที่ช่วยให้การบันทึกและการจัดเก็บภาพพื้นผิวโลกที่มีความละเอียดสูง [1] ทั้งยังครอบคลุมหลายช่วงเวลาและหลายภูมิภาคพื้นที่ที่ได้รับอนุญาตให้บันทึก แน่นนอนว่า การเข้าถึงภาพถ่ายดาวเทียมความละเอียดสูง [2] สามารถเข้าถึงได้โดยผู้ให้บริการหลายราย แต่ถึงกระนั้น ในการศึกษาได้นำภาพถ่ายดาวเทียมของผู้ให้บริการจากบริษัท Google มาใช้เป็นชุดข้อมูลเพื่อการศึกษา ซึ่งไม่ขัดกับวัตถุประสงค์ทางวิชาการที่ทางผู้ให้บริการระบุไว้ [3]

อย่างไรก็ตาม องค์ประกอบที่อยู่ในการบันทึกด้วยภาพถ่ายดาวเทียมมีความแตกต่างกันไปขึ้นอยู่กับหลายปัจจัย ไม่ว่าจะเป็นตำแหน่งที่ตั้งและลักษณะทางภูมิศาสตร์ ตัวอย่างเช่น อาคาร ป่าไม้ พื้นที่แห้งแล้ง พื้นที่เกษตรกรรม ถนน และอื่น ๆ นอกจากนี้ หากพิจารณาถึงความท้าทายที่เกี่ยวข้องกับคุณภาพของภาพถ่ายดาวเทียม เช่น ค่าสีในภาพไม่สอดคล้องกันตามช่วงเวลาที่บ้านเกิดเนื่องจากการโคจรของดาวเทียม การมีเมฆปกคลุมหรือบดบังความชัดเจนของวัตถุในภาพ ด้วยเหตุนี้ การสกัดอาคารจากภาพถ่ายดาวเทียมโดยอิงลักษณะเฉพาะ ตำแหน่ง และโครงสร้างของอาคารด้วยแรงงานมนุษย์อาจใช้เวลานาน มีค่าใช้จ่ายสูง รวมถึงมีแนวโน้มการเกิดข้อผิดพลาด

ดังนั้น งานวิจัยนี้จึงได้นำเสนอวิธีการอัตโนมัติในการสกัดพื้นที่อาคารจากภาพถ่ายดาวเทียม เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ด้วยการประยุกต์ใช้เทคนิคการประมวลผลภาพ (Image Processing) และการเรียนรู้ของเครื่อง (Machine Learning) โดยมุ่งเน้นไปที่วิธีการการเรียนรู้เชิงลึก (Deep Learning) ที่สามารถแบ่งส่วนภาพ (Image Segmentation) ได้อย่างแม่นยำและมีประสิทธิภาพในบริบทพื้นที่ประเทศไทย การศึกษานี้ให้ความสนใจกับภูมิภาคที่ห่างไกลจากเมืองหลวง โดยเฉพาะกับภูมิภาคแบบชนบท ซึ่งระหว่างการตรวจสอบพื้นที่ต่าง ๆ พบว่าอาคารมีการออกแบบตามสถาปัตยกรรมที่ต่างกัน เช่น สี หลังคา ขนาด และรูปทรง นอกจากนี้ การปลูกต้นไม้สูงใกล้ชายคาบ้านส่งผลให้พื้นที่หลังคาถูกบดบัง และมีรูปทรงไม่เด่นชัด แต่ทว่า ปัจจัยด้านสีและความเข้มของแสงโดยรอบช่วยให้เห็นความแตกต่างที่ชัดเจนระหว่างอาคารและสภาพแวดล้อมทางธรรมชาติ

จากวิธีการที่นำเสนอ วัตถุประสงค์หลักของการวิจัยคือการใช้ประโยชน์จากภาพถ่ายดาวเทียม โดยนำมาสร้างชุดข้อมูลซึ่งใช้สำหรับการเรียนรู้และทดสอบด้วยโมเดลการเรียนรู้เชิงลึก นอกจากนี้ ระเบียบการวิจัยยังให้ความสำคัญต่อการวิเคราะห์เชิงเปรียบเทียบ ทั้งในส่วนการแบ่งหมวดหมู่ชุดข้อมูลตามกระบวนการการประมวลผลรูปภาพล่วงหน้า (Image preprocessing) และการเปรียบเทียบความสามารถของสถาปัตยกรรมแต่ละโมเดลการเรียนรู้เชิงลึก ได้แก่ U-Net, RIU-Net, U-Net++, Res-U-Net และ DeepLabV3+ ทั้งนี้ สำหรับโมเดลที่มีความสามารถในการปรับแต่งโครงสร้างแกนหลักแบ็กโบน (Backbone) จะถูกดำเนินการโดยใช้ ResNet-50, ResNet-101 และ ResNet-152

เนื่องด้วยวิธีการการสกัดอาคารจากภาพถ่ายดาวเทียม การศึกษานี้มีส่วนช่วยในการพัฒนาและปรับปรุงเทคนิคก่อนการประมวลผลเพื่อเพิ่มความแม่นยำในการแบ่งส่วน การเปรียบเทียบและวัดผลเพื่อระบุโมเดลการเรียนรู้เชิงลึกที่ให้ประสิทธิภาพสูงสุดในการตรวจจับอาคาร อย่างไรก็ตาม แม้วางานวิจัยไม่ได้นำเสนอโมเดลรูปแบบใหม่ แต่ได้มุ่งเน้นไปที่การปรับปรุงระเบียบวิธีเพื่อให้ได้ประสิทธิภาพมากยิ่งขึ้น นอกจากนี้ ในด้านการนำไปประยุกต์ใช้งานจริง ยังสามารถนำไปเป็นต้นแบบเพื่อดำเนินการติดตามการขยายตัวของเมือง เติมเต็มความสมบูรณ์ให้กับแผนที่ หรือแม้แต่ นำไปต่อยอดเพื่อประมาณค่าความหนาแน่นของประชากรแบบอัตโนมัติ ในพื้นที่ห่างไกลและพื้นที่ที่มีความเสี่ยง ทั้งนี้ วิธีการที่นำเสนอสามารถดำเนินการได้ในพื้นที่กว้าง รวดเร็ว และลดต้นทุนการลงพื้นที่จริง

1.2 วัตถุประสงค์ของงานวิจัย

- 1) พัฒนารูปแบบการสกัดอาคารจากภาพถ่ายดาวเทียมความละเอียดสูงด้วยโมเดล หรือแบบจำลองปัญญาประดิษฐ์การเรียนรู้เชิงลึก ร่วมกับเทคนิคการประมวลผลรูปภาพ
- 2) สร้างและปรับแต่งชุดข้อมูลด้วยเทคนิคการประมวลผลรูปภาพ เพื่อเพิ่มประสิทธิภาพในการสกัดอาคารจากภาพถ่ายดาวเทียม
- 3) วิเคราะห์และเปรียบเทียบชุดข้อมูลที่ถูกนำไปดำเนินการร่วมกับโมเดลการเรียนรู้เชิงลึก เพื่อค้นหาโมเดลที่มีความแม่นยำสูงสุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.3 ขอบเขตงานวิจัย

- 1) เป้าหมายพื้นที่ที่นำมาศึกษาคือ ตำบลกกทอง ขนาดพื้นที่ประมาณ 64 ตารางกิโลเมตร และตำบลศรีสองรัก ขนาดพื้นที่ประมาณ 129 ตารางกิโลเมตร ซึ่งทั้งสองตำบลเป็นส่วนหนึ่งในอำเภอเมืองเลย จังหวัดเลย ประเทศไทย ทั้งนี้ตามเว็บไซต์ระบบสถิติทางการทะเบียน รายงานสถิติจำนวนประชากรและบ้านประจำปี พ.ศ. 2566 [4] ตำบลกกทองมีจำนวนประชากร 2,740 คน จำนวนบ้าน 732 หลัง และตำบลศรีสองรัก มีจำนวนประชากร 7,476 คน จำนวนบ้าน 3,373 หลัง
- 2) เก็บรวบรวมภาพถ่ายดาวเทียมความละเอียดสูงจากผู้ให้บริการ Google ซึ่งให้สิทธิ์การเข้าถึงเพื่อวัตถุประสงค์ทางวิชาการและการวิจัย ตัวเลือกรับประกันการปฏิบัติตามมาตรฐานทางกฎหมายและจริยธรรม ขณะเดียวกัน ก็ให้รายละเอียดที่เพียงพอสำหรับการวิเคราะห์ที่มีประสิทธิภาพ
- 3) การจัดทำหมวดหมู่ชุดข้อมูลรูปภาพออกเป็นสองกลุ่มที่แตกต่างกัน ได้แก่ กลุ่มรูปภาพต้นฉบับที่ผ่านการประมวลผลเพียงการปรับขนาด จำนวน 1 ชุด และกลุ่มรูปภาพที่ผ่านการประมวลผลเพิ่มเติม จำนวน 3 ชุด
- 4) พัฒนา เปรียบเทียบ และทดสอบการสกัดอาคารจากภาพถ่ายดาวเทียมของโมเดลการเรียนรู้เชิงลึก ได้แก่ U-Net, RIU-Net, U-Net++, Res-U-Net และ DeepLabV3+ พร้อมด้วยการเปรียบเทียบแกนหลักแบ็กโบน ResNet-50, ResNet-101 และ ResNet-152 (สำหรับ Res-U-Net และ DeepLabV3+ เท่านั้น)

1.4 ประโยชน์ที่คาดว่าจะได้รับ

- 1) ค้นพบวิธีการการสกัดอาคารจากภาพถ่ายดาวเทียมความละเอียดสูงแบบอัตโนมัติ โดยทราบถึงโมเดลที่ให้ประสิทธิภาพความแม่นยำในการแบ่งส่วนที่ดีที่สุดตามดัชนีชี้วัด
- 2) ค้นพบขั้นตอนการปรับปรุงรูปภาพก่อนการประมวลผลรูปภาพเพื่อให้ผลลัพธ์ที่แม่นยำ
- 3) สามารถนำไปตีพิมพ์ผลงานทางวิชาการเพื่อแบ่งปันสิ่งที่ค้นพบ ความรู้ และระเบียบวิธีวิจัยที่ใช้ในการทดลอง ซึ่งมีส่วนช่วยให้เกิดองค์ความรู้ที่กว้างขวางแก่กลุ่มผู้ศึกษาการสำรวจระยะไกล (Remote sensing) การเรียนรู้ของเครื่อง และการแบ่งส่วนภาพ
- 4) สามารถนำไปต่อยอดพัฒนาเป็นเครื่องมือซอฟต์แวร์ หรือระบบที่สามารถนำไปใช้โดยนักวางผังเมือง นักสถิติ หน่วยงานภาครัฐ และผู้มีส่วนได้ส่วนเสียอื่น ๆ ที่เกี่ยวข้อง

1.5 ระเบียบวิธีวิจัย

- 1) ศึกษางานวิจัยและทฤษฎีที่เกี่ยวข้อง รวมทั้งออกแบบการทดลองเพื่อแก้ไขปัญหา
 - 2) ดำเนินการเก็บรวบรวมข้อมูลภาพถ่ายดาวเทียมความละเอียดสูงตามขอบเขตพื้นที่ศึกษา
- เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- 3) จัดทำชุดคำตอบให้กับข้อมูลที่ดำเนินการจัดเก็บ ซึ่งจำเป็นต้องระบุตำแหน่งอาคารหรือสิ่งปลูกสร้างในภาพถ่ายดาวเทียม ทั้งนี้ ดำเนินการโดยผู้วิจัย
- 4) จัดเตรียมชุดข้อมูลโดยการแบ่งหมวดหมู่และแบ่งชุดข้อมูล ดำเนินการด้วยกระบวนการประมวลผลภาพ รวมทั้งประยุกต์ใช้การเสริมชุดข้อมูล (Data Augmentation)
- 5) ทดลองสกัดอาคารจากภาพถ่ายดาวเทียม ด้วยวิธีการประมวลรูปภาพแบบสัณฐานวิทยา และวิธีการประยุกต์ใช้ปัญญาประดิษฐ์ โมเดลการเรียนรู้เชิงลึก ซึ่งครอบคลุมกระบวนการแบ่งชุดข้อมูลเพื่อการเรียนรู้และการทดสอบ ได้แก่ ชุดข้อมูลเรียนรู้ (Training Set) ร่วมกับชุดข้อมูลตรวจสอบ (Validation Set) และชุดข้อมูลทดสอบ (Test set)
- 6) ปรับแต่งค่าไฮเปอร์พารามิเตอร์ (Hyperparameter Tuning) เพื่อให้สอดคล้องกับการเรียนรู้ต่อชุดข้อมูลที่จัดเตรียม และเพื่อให้ได้ผลลัพธ์ความแม่นยำที่ดียิ่งขึ้น
- 7) วิเคราะห์ เปรียบเทียบ และประเมินผลความแม่นยำจากการสกัดอาคาร โดยอ้างอิงจากสมการที่ใช้สามารถตรวจสอบความคล้ำยคลึงของรูปภาพ
- 8) อภิปรายผลลัพธ์ สรุปผลการวิจัย และบันทึกสิ่งจำเป็นที่ต้องปรับปรุงในงานวิจัย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

เนื่องด้วยวัตถุประสงค์และขอบเขตงานวิจัยตามที่ระบุในบทที่ 1 เนื้อหาบทนี้จะกล่าวถึงทฤษฎีและการศึกษางานวิจัยที่เกี่ยวข้องกับการแยก ระบุ สกัด อาคารจากภาพถ่ายดาวเทียมความละเอียดสูง เพื่อให้เป็นไปตามกรอบการดำเนินงานและระเบียบวิจัย

ดังนั้น การทำความเข้าใจวิธีการดั้งเดิมอย่างตัวอย่างปัญหาการสำมะโนประชากรในประเทศไทยจึงจำต้องพิจารณา รวมถึง การวิวัฒนาการและความเป็นมาของเทคโนโลยีการสำรวจทางไกลซึ่งได้ถูกนำมาใช้ในงานวิจัยนี้ การศึกษาทฤษฎีที่เกี่ยวกับการประมวลผลภาพดิจิทัล และการประยุกต์ร่วมกับโมเดลการเรียนรู้เชิงลึก อันเป็นส่วนสำคัญราวปฏิวัติความก้าวหน้าในสาขาการประมวลผลและการแบ่งส่วนภาพ โดยตัวอย่างสถาปัตยกรรมแบบจำลองปัญญาประดิษฐ์ที่ได้กล่าวถึง ไม่ว่าจะเป็น U-Net, RIU-Net, U-Net++, Res-U-Net และ DeepLabV3+ จะได้รับการกล่าวถึง นอกจากนี้ เพื่อให้เป็นที่ยอมรับในการประเมินและวัดผลความคล้ายคลึง การศึกษาและนำเสนอสมการ Intersection over Union (IoU) และ Dice Similarity Coefficient (DSC) ได้ถูกหยิบมาใช้เป็นเกณฑ์การประเมินภาพผลลัพธ์จากการทำนายของโมเดลการเรียนรู้เชิงลึก

อนึ่ง เนื้อหาในบทที่ยังครอบคลุมถึงงานวิจัยหรือการศึกษาก่อนหน้านี้ที่ได้รับการตีพิมพ์วารสารวิชาการ เพื่อเป็นการศึกษาตัวอย่างระเบียบวิจัย ปัญหาที่พบ การนำเทคโนโลยีและองค์ความรู้มาประยุกต์ใช้อย่างถูกต้องตามหลักการวิชาการ รวมทั้งเป็นการตอกย้ำให้เห็นว่า ปัญหาของงานวิจัยนี้เป็นที่ได้รับความสนใจในแวดวงการ ซึ่งสนับสนุนให้เห็นถึงแนวทางที่เลือก และการมีส่วนร่วมในการพัฒนาองค์ความรู้ด้านการสำรวจระยะไกลควบคู่กับปัญญาประดิษฐ์สืบไป

2.1 การสำมะโนประชากรและเคหะในประเทศไทย

การสำมะโนประชากรและเคหะ คือ การเก็บรวบรวมข้อมูลเกี่ยวกับประชากรทุกคนในประเทศตามที่อยู่จริง รวมทั้งข้อมูลที่อยู่อาศัยของประชากร เพื่อนำเสนอผลในภาพรวมของประเทศ ในระดับหมู่บ้าน เทศบาล ตำบล อำเภอ จังหวัด และประเทศ เพื่อประเทศจะได้มีข้อมูลประชากรตามที่อยู่จริงในแต่ละท้องถิ่น เพื่อนำมากำหนดนโยบายและวางแผนงานในการระดับท้องถิ่น เพื่อพัฒนาประเทศ

การสำมะโนประชากรและเคหะในประเทศไทยจะดำเนินการทุก 10 ปี ตามข้อเสนอแนะขององค์การสหประชาชาติ ซึ่งการสำมะโนประชากรและเคหะอย่างเป็นทางการครั้งล่าสุดคือปี พ.ศ. 2553 [2.2] โดยมีผู้รับผิดชอบคือสำนักงานสถิติแห่งชาติ (เอกสารเล่มนี้จัดพิมพ์เมื่อ เดือนพฤษภาคม ปีพ.ศ. 2567) และเป็นที่ทราบกันดีว่า การสำมะโนประชากรและเคหะจำเป็นต้องอาศัยความร่วมมือจากหน่วยงานทั้งภาครัฐและเอกชน รวมถึงประชาชนผู้มีสัญชาติไทยทุกคนในประเทศจำเป็นต้องทำแบบ

สัมภาษณ์ ไม่ว่าจะเป็นลายลักษณ์อักษร หรือการกรอกข้อมูลผ่านอินเทอร์เน็ต เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ทั้งนี้ รายละเอียดที่ต้องกรอกในแบบสัมภาษณ์ จะต้องเป็นไปตามที่สำนักงานสถิติแห่งชาติระบุไว้ในแต่ละรอบที่ทำสำมะโนประชากรและเคหะ สำหรับการติดตามข่าวสารและข้อมูลสามารถค้นหาที่ เว็บไซต์สำนักงานสถิติแห่งชาติ (สสช.)

- 4 -

เขตจังหวัด หรือประเทศ	หมู่ที่ของเขต หรือบ้านเลขที่	บ้านเลขที่ของครัวเรือน หรือบ้านเลขที่ของครัวเรือน	ชื่อของครัวเรือน	ชื่อของครัวเรือน	ชื่อของครัวเรือน
(19)	(20)	(21)	(22)	(23)	(24)

๓) *ครัวเรือนใดบ้างที่ไม่ได้ส่ง และใช้เกณฑ์ใด

๓) *ครัวเรือนใดบ้างที่ไม่ได้ส่ง และใช้เกณฑ์ใด

๓) *ครัวเรือนใดบ้างที่ไม่ได้ส่ง และใช้เกณฑ์ใด

สำมะโนประชากรและเคหะ พ.ศ. 2553

แบบ ๒/1 (แบบแจ้งพื้นที่)

เรียน ท่านเจ้าบ้าน

ตามที่กระทรวงมหาดไทยได้แจ้งให้ดำเนินการสำมะโนประชากรและเคหะ พ.ศ. 2553 นั้น เพื่อให้ทราบถึงจำนวนประชากรและเคหะของประเทศไทย

๑) สำมะโนประชากรและเคหะ พ.ศ. 2553

๒) สำมะโนประชากรและเคหะ พ.ศ. 2553

๓) สำมะโนประชากรและเคหะ พ.ศ. 2553

๔) สำมะโนประชากรและเคหะ พ.ศ. 2553

๕) สำมะโนประชากรและเคหะ พ.ศ. 2553

๖) สำมะโนประชากรและเคหะ พ.ศ. 2553

๗) สำมะโนประชากรและเคหะ พ.ศ. 2553

๘) สำมะโนประชากรและเคหะ พ.ศ. 2553

๙) สำมะโนประชากรและเคหะ พ.ศ. 2553

๑๐) สำมะโนประชากรและเคหะ พ.ศ. 2553

- 2 -

โปรดกรอกรายละเอียดของครัวเรือนในกรณีนี้ คนละหนึ่งบรรทัด (โดยไม่ทำตัวว่างหรือเขียนชื่อย่อ)

ชื่อของครัวเรือน	ความเกี่ยวข้อง	เขต	หมู่ (บ้าน)	ตำบล	อำเภอ	จังหวัด	ชื่อของครัวเรือน	ชื่อของครัวเรือน	ชื่อของครัวเรือน
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)

- 3 -

การอยู่อาศัย	รายละเอียดของครัวเรือน	สถานะของครัวเรือน	อาชีพหลักในปัจจุบัน	ประเภท	สาขาอาชีพ	ทางไปรษณีย์
(11)	(12)	(13)	(14)	(15)	(16)	(17)

รูปที่ 2.1 แบบสัมภาษณ์สำมะโนประชากรและเคหะประเทศไทย ปี พ.ศ. 2553

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 2.1 จะพบว่าภายในแบบสัมภาษณ์มีการเก็บข้อมูล ทั้งข้อมูลส่วนบุคคลและข้อมูลของสถานที่อยู่อาศัย (ครัวเรือน) รวมอยู่ด้วย ซึ่งรายละเอียดดังกล่าวไม่สามารถดำเนินการจัดเก็บและรวบรวมผ่านการถ่ายภาพดาวเทียม แต่อย่างไรก็ดี หากพินิจในแง่มุมการต่อยอดเพื่อนำไปทำการประมาณประชากรจากการระบุจำนวนสิ่งปลูกสร้าง วิธีการที่นำเสนอยังคงมีความเป็นไปได้

ถึงแม้ว่า การทำสำมะโนประชากรและเคหะจะมีข้อมูลที่ต้องเก็บรวบรวมหลากหลายด้านมากกว่าแค่จำนวนประชากรในพื้นที่ แต่จำเป็นต้องใช้ทรัพยากรทั้งบุคคล เวลา และงบประมาณอย่างมากในการดำเนินการ อีกทั้งยังมีความเสี่ยงในการเกิดข้อผิดพลาดจากมนุษย์ (Human Error) และมีความถี่ที่ล่าช้าในการปรับปรุงรอบข้อมูล หรือกล่าวคือ ข้อมูลจะเกิดขึ้นในทุก 10 ปี ทั้งนี้ ยังไม่รวมระยะเวลาการเก็บรวบรวม วิเคราะห์ และจัดทำบันทึกเอกสารอย่างเป็นทางการก่อนการเผยแพร่

ด้วยวัตถุประสงค์ของงานวิจัยและความพยายามในการเสนอวิธีการแก้ไขปัญหา ผู้วิจัยจึงเริ่มค้นคว้าทฤษฎีเกี่ยวกับการสำรวจระยะไกล และผู้ให้บริการข้อมูลเพื่อที่จะสามารถนำกลุ่มตัวอย่างของข้อมูลมาทดลอง จากนั้น จึงเริ่มศึกษางานวิจัยที่เกี่ยวข้องในการใช้กระบวนการประมวลผลรูปภาพเพื่อแบ่งส่วนที่สนใจออกจากภาพพื้นหลัง หรือกล่าวคือ การแบ่งส่วนเฉพาะสิ่งที่เป็นสิ่งปลูกสร้างอย่างอาคารบ้านเรือนจากภาพถ่ายดาวเทียม

2.2 ทฤษฎีและเทคโนโลยีการสำรวจระยะไกล

การสำรวจระยะไกลเป็นวิทยาการหรือเทคโนโลยีที่นิยมใช้กันอย่างแพร่หลายเนื่องจากข้อมูลที่จัดเก็บมาได้สามารถนำไปวิเคราะห์เพื่อแก้ไขปัญหาได้ในหลากหลายด้าน อีกทั้ง ข้อมูลที่ได้มายังมีคุณสมบัติที่แตกต่างกับข้อมูลประเภทอื่น ๆ สำหรับนิยามกระบวนการเกี่ยวกับการสำรวจระยะไกลด้วยดาวเทียม และการประยุกต์ใช้ข้อมูลที่ได้จากการจัดเก็บนับว่าเป็นพื้นฐานสำคัญในการศึกษา

2.2.1 การสำรวจระยะไกล

การสำรวจระยะไกล หรือ รีโมทเซนซิง (Remote Sensing) [5, 6] เป็นวิทยาศาสตร์และเทคโนโลยีแขนงหนึ่งที่ใช้ในการบ่งบอก จำแนก หรือวิเคราะห์คุณลักษณะของวัตถุผ่านการบันทึกคุณลักษณะของวัตถุ (Object) หรือปรากฏการณ์ (Phenomena) จากการสะท้อนแสง หรือการแผ่รังสีพลังงานแม่เหล็กไฟฟ้า (Electromagnetic Radiation) ด้วยเครื่องมือที่ใช้วัดค่าพลังงานแม่เหล็กไฟฟ้าที่สะท้อนหรือแผ่ออกจากวัตถุ เรียกว่า เครื่องวัดจากระยะไกล (Remote sensor) หรือ เครื่องวัด (Sensor) ซึ่งติดอยู่กับยานพาหนะสำรวจ โดยปราศจากการสัมผัสโดยตรง

คำว่ารีโมทเซนซิง เป็นศัพท์เทคนิคที่ใช้ครั้งแรกในประเทศสหรัฐอเมริกาปี พ.ศ. 2503 ซึ่งมีความหมายรวมถึงการทำแผนที่ การแปลภาพถ่าย ธรณีวิทยาเชิงภาพถ่าย แต่แล้วหลังจากนั้นก็แพร่หลายมากขึ้นในปี พ.ศ. 2515 เนื่องจากการส่งดาวเทียมสำรวจทรัพยากร

ดวงแรกนามว่า LANDSAT-1 ขึ้นไปโคจรรอบโลก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การสำรวจระยะไกลทำให้ได้รับข้อมูลวัตถุด้วยการที่ตัวรับสัญญาณจะรับสัญญาณค่า การสะท้อนแสงหรือการแผ่รังสีจากแต่ละวัตถุหรือสภาพแวดล้อมที่มีค่าเฉพาะตัวและแตกต่างกัน โดยวิเคราะห์วัตถุนั้นจาก 3 คุณลักษณะ คือ การสะท้อนช่วงคลื่นของวัตถุ (Spectral) รูปทรงสัญญาณของวัตถุนั้นบนพื้นโลก (Spatial) และการเปลี่ยนแปลงของวัตถุตามช่วงเวลา (Temporal) ทั้งนี้การใช้อุปกรณ์ในการจัดเก็บค่าหรือเครื่องวัดย่อมให้คุณลักษณะของข้อมูลที่แตกต่างกัน และเหตุนี้เองภาพถ่ายดาวเทียม (Satellite Image) ก็เป็นข้อมูลประเภทหนึ่งจากเทคโนโลยีที่ใช้ในการสำรวจระยะไกลด้วยดาวเทียม

2.2.2 หลักการและกระบวนการของดาวเทียม

ดาวเทียม [7] คือ สิ่งประดิษฐ์ที่สามารถโคจรรอบโลกโดยอาศัยแรงดึงดูดในลักษณะเดียวกันกับดวงจันทร์โคจรรอบโลก และโลกโคจรรอบดวงอาทิตย์อาจกล่าวได้ว่าเป็นการเลียนแบบดาวบริวารของดาวเคราะห์ แต่ทว่าดาวเทียมแต่ละดวงแต่ละประเภทก็จะมีการติดตั้งอุปกรณ์หรือเครื่องวัดต่าง ๆ เพื่อตามความเหมาะสม สำหรับวัตถุประสงค์ของดาวเทียมประเภทดาวเทียมสำรวจทรัพยากรก็เพื่อการสำรวจ ติดตามทรัพยากรและสิ่งแวดล้อม ศึกษา ลักษณะทางภูมิศาสตร์ของโลก รวมไปถึงการนำมาใช้เพื่อการนำทางหรือเป็นแผนที่ เนื่องจากดาวเทียมประเภทนี้ได้รับการออกแบบให้มีความสามารถในการถ่ายภาพความละเอียดสูง และสามารถจำแนกประเภทของวัตถุได้อย่างชัดเจน

ดาวเทียมเคลื่อนที่เป็นวงรอบโลก หรือที่เรียกว่า วงโคจรของดาวเทียม [8, 9] สามารถแบ่งเป็น 2 ประเภทขึ้นกับความสูงและลักษณะการโคจร ในขณะเดียวกันก็บ่งบอกถึงคุณลักษณะของภาพถ่ายดาวเทียมด้วย

2.2.2.1 วงโคจรแบบสัมพันธ์กับดวงอาทิตย์ (Sun-synchronous Orbit)

การโคจรของดาวเทียมโดยอิงตามทิศแนวเหนือ-รุ่งใต้ หรือกล่าวคือการเคลื่อนผ่านแนวละติจูดหนึ่ง ๆ ที่เวลาท้องถิ่นเดียวกันซึ่งส่วนใหญ่เป็นวงโคจรสำหรับดาวเทียมสำรวจทรัพยากร โดยแบ่งได้อีกเป็น 2 ประเภท คือ

1) วงโคจรผ่านขั้วโลก (Polar orbit) เป็นวงโคจรที่มีรูปลักษณะเป็นวงกลมโดยมีเส้นผ่านศูนย์กลางในแนวขั้วโลก โคจรที่ระดับความสูง 500-1,000 กิโลเมตร จากพื้นโลก เป็นวงโคจรระดับต่ำ (Low Earth Orbit “LEO”) ภาพที่ถ่ายจะมีรายละเอียดสูง แต่เพราะอยู่ใกล้พื้นผิวโลกจึงครอบคลุมพื้นที่ที่เป็นบริเวณแคบและไม่สามารถครอบคลุมบริเวณใดบริเวณหนึ่งได้นาน เนื่องจากความเร็วในการเคลื่อนที่ซึ่งส่งผลต่อระยะเวลาในการโคจร กระนั้น วงโคจรลักษณะดังกล่าวสามารถให้พื้นที่การบริการครอบคลุมได้ทั่วโลก

2) วงโคจรเอียง (Inclined orbit) เป็นวงโคจรที่มีรูปลักษณะเป็นทั้งวงกลมและวงรี เป็นวงโคจรที่มีอยู่จำนวนมากแตกต่างกันไปตามความเอียง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาค้นคว้าเท่านั้น เมื่อนำไปเผยแพร่โดยไม่ได้รับอนุญาตให้ถือว่าผิดกฎหมายและไม่ว่ากรณิใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(Incline) หรือมุมที่ทำกับระนาบศูนย์สูตร และความรีของวงโคจรว่าเล็กน้อยเพียงใด โคจรที่ระดับความสูง 5,000-35,000 กิโลเมตร จากพื้นโลกสามารถให้พื้นที่บริการบริเวณละติจูดสูงหรือต่ำมาก ๆ ได้ หรืออาจครอบคลุมพื้นที่ขั้วโลกได้ด้วย อีกกล่าวว่าเป็นวงโคจรระยะปานกลาง (Medium Earth Orbit “MEO”) การถ่ายภาพและส่งสัญญาณวิทยุสามารถครอบคลุมพื้นที่ได้กว้างกว่าวงโคจรแบบต่ำ แต่หากต้องการให้ครอบคลุมทั้งโลกต้องใช้ดาวเทียมหลายดวงทำงานร่วมกัน

2.2.2.2 วงโคจรระนาบศูนย์สูตร (Equatorial Orbit)

การโคจรในแนวระนาบหรือหมุนเหวี่ยงตั้งฉากกับโลก มีลักษณะการโคจรเป็นรูปวงกลม โคจรในแนวระนาบกับเส้นศูนย์สูตร ในกรณีที่โคจรที่ระดับความสูงประมาณ 36,000 กิโลเมตรจากพื้นโลก ดาวเทียมจะเคลื่อนที่ด้วยความเร็วเท่ากับความเร็วที่โลกหมุนรอบตัวเอง คือ ใช้เวลาในการโคจร 24 ชั่วโมง/รอบ ซึ่งเสมือนว่าตำแหน่งของดาวเทียมคงที่ตลอดเวลา เรียกว่า “วงโคจรค้างฟ้า (Geo-stationary orbit)” และเรียกดาวเทียมที่โคจรในวงโคจรนี้ว่า “ดาวเทียมค้างฟ้า (Geo-stationary satellite)”

2.2.3 คุณสมบัติของดาวเทียมสำรวจทรัพยากรโลก

ดาวเทียมสำรวจทรัพยากรเข้ามามีบทบาทอย่างมาก ในการศึกษาแบบรีโมทเซ็นซิง เนื่องจากคุณสมบัติที่เหมาะสมดังนี้

- 1) การบันทึกข้อมูลเป็นบริเวณกว้าง (Synoptic view) ภาพจากดาวเทียมสำรวจทรัพยากรครอบคลุมพื้นที่กว้าง ทำให้ได้ข้อมูลในลักษณะต่อเนื่องในระยะเวลาบันทึกภาพสั้น กล่าวคือภาพที่ได้ในสภาพแวดล้อมหนึ่งเป็นสภาพแวดล้อมในระยะเวลาเดียวกันหรือใกล้เคียงกัน
- 2) การบันทึกภาพได้หลายช่วงคลื่น (Multispectral Resolution) ดาวเทียมสำรวจทรัพยากรมีระบบบันทึกภาพได้หลายช่วงคลื่นจากเครื่องวัดหรือเซ็นเซอร์ที่ถูกติดตั้งไว้ โดยช่วงคลื่นจะครอบคลุมตั้งแต่ช่วงคลื่นที่เห็นได้ด้วยตาเปล่าจนถึงคลื่นที่ไม่สามารถมองเห็นด้วยตามนุษย์ สาเหตุก็เพื่อจำแนกหรือแยกแยะวัตถุต่าง ๆ บนพื้นผิวโลกได้อย่างชัดเจน
- 3) การบันทึกภาพบริเวณเดิม (Repetitive coverage, Temporal resolution) ดาวเทียมสำรวจทรัพยากรมีวงโคจรจากเหนือลงใต้ และกลับมายังจุดเดิมในเวลาท้องถิ่นอย่างสม่ำเสมอในช่วงเวลาที่แน่นอน
- 4) การให้รายละเอียดเชิงพื้นที่หลายระดับ (Spatial resolution) รายละเอียดที่กล่าวถึงจะมีหน่วยเป็นเมตร ซึ่งค่ายิ่งน้อยนั้นหมายถึงรายละเอียดสูง
- 5) ภาพจากดาวเทียมสามารถให้ภาพสีผสม (False color composite) การ

นำค่าสีมาผสมกันขึ้นกับวัตถุประสงค์ที่ต้องการขยายรายละเอียดเฉพาะให้เด่นชัดขึ้น มักเอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น เมื่อนำไปเผยแพร่บนเว็บไซต์สาธารณะโดยไม่ผ่านการอนุญาตเห็นาเบเซบระยะชันดานการคําไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

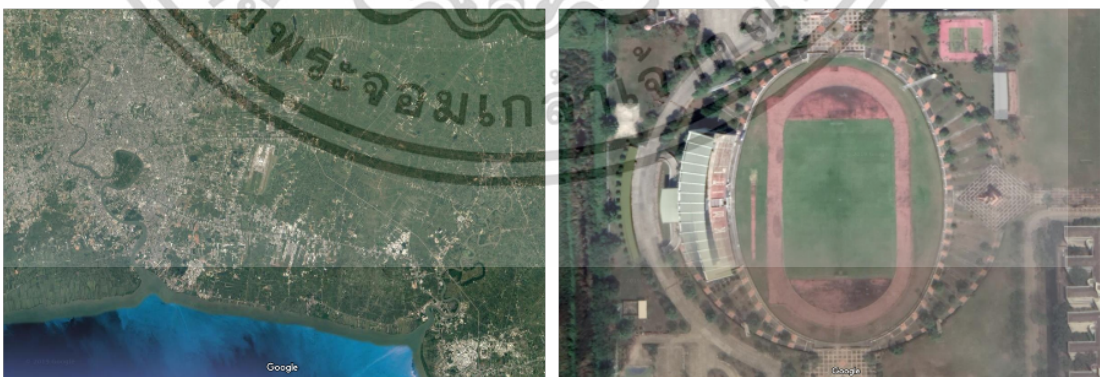
นำมาใช้การจำแนกหรือแยกสีที่แตกต่างจากสภาพแวดล้อม

6) การเน้นคุณภาพของภาพ (Image enhancement) ภาพจากดาวเทียมต้นฉบับสามารถนำมาปรับปรุงคุณภาพให้มีรายละเอียดเพิ่มขึ้น โดยหนึ่งในเทคนิคที่นิยมใช้กันเรียกว่า Pan-Sharpned เป็นการผสมภาพสี (Multispectral) เข้ากับภาพขาวดำรายละเอียดสูง (Panchromatic)

2.2.4 ผู้ให้บริการข้อมูลภาพถ่ายดาวเทียมและคุณสมบัติของดาวเทียม

จากคุณสมบัติของดาวเทียมสำรวจทรัพยากรทำให้ผู้วิจัยค้นคว้าถึงผู้ให้บริการข้อมูลดังกล่าวที่สามารถเข้าถึงได้ง่ายและไม่มีค่าใช้จ่ายในการบริการ ซึ่งก็คือ Google โดย Google ได้มีบริการการเข้าถึงข้อมูลภาพถ่ายดาวเทียมในชื่อว่า Google Earth

ภาพถ่ายดาวเทียม Google ที่นำมาให้บริการเกิดจากความร่วมมือกับบริษัทด้านอวกาศและภูมิศาสตร์รวมถึงการให้บริการดาวเทียม นามว่า DigitalGlobe [10] หรือกล่าวได้ว่าบริษัท DigitalGlobe เป็นผู้ให้บริการภาพถ่ายดาวเทียมร่วมกับ Google โดยที่ Google เป็นผู้เปิดช่องทางการเข้าถึงข้อมูลต่ออีกที ในส่วนของเทคโนโลยีการแสดงผลภาพถ่ายดาวเทียมเกิดจากการซ้อนกันหลายชั้น (Layer) จากข้อมูลภาพถ่ายดาวเทียมหลายดาวเทียมซึ่งให้คุณลักษณะของภาพถ่ายดาวเทียมที่ต่างกันออกไปด้วยเช่นกัน ยกตัวอย่างเช่น เมื่อผู้ใช้ใช้งาน Google Earth จะมีฟังก์ชันการทำงานหนึ่งชื่อว่า กำลังขยาย (Zoom) สามารถปรับได้ตามความต้องการของผู้ใช้เพื่อเป็นการเพิ่มรายละเอียดในส่วนพื้นที่ที่สนใจตามรูปที่ 2.2 และด้วยเหตุนี้เองกำลังขยายที่ต่ำที่สุดกับกำลังขยายที่สูงที่สุดย่อมมีการแสดงรายละเอียดที่แตกต่างกันไปรวมไปถึงการจัดเก็บข้อมูลที่ต่างกัน ฉะนั้นแล้ว การซ้อนภาพที่มีรายละเอียดต่างกันจึงเป็นวิธีการหนึ่งที่ได้ถูกนำมาใช้ในจุดนี้



รูปที่ 2.2 เปรียบเทียบกำลังขยายต่ำ (ซ้าย) และกำลังขยายสูง (ขวา) ของซอฟต์แวร์ Google Earth

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำหรับดาวเทียมที่ถูกนำมาใช้ร่วมกับบริการของ Google Earth ได้แก่

2.2.4.1 ดาวเทียม LANDSAT-8

ดาวเทียม LANDSAT-8 [11] ใช้สำหรับสำรวจทรัพยากรของประเทศสหรัฐอเมริกา ได้รับการพัฒนาโดยความร่วมมือระหว่างองค์การ NASA และ USGS เริ่มปฏิบัติการวันที่ 30 พฤษภาคม พ.ศ. 2556 โคจรสูงเหนือพื้นโลก 705 กิโลเมตร

ตารางที่ 2.1 อุปกรณ์บันทึกข้อมูลของดาวเทียม LANDSAT - 8 [12]

LANDSAT-8: Operational Land Imager และ Thermal Infrared Sensor		
แบนด์	ความยาวคลื่น (ไมโครเมตร)	รายละเอียดภาพ (เมตร)
1	0.43 - 0.45 (Coastal Aerosol)	30
2	0.45 - 0.51 (Blue)	30
3	0.53 - 0.59 (Green)	30
4	0.64 - 0.67 (Red)	30
5	0.85 - 0.88 (Near Infrared NIR)	30
6	1.57 - 1.65 (SWIR 1)	30
7	2.11 - 2.29 (SWIR 2)	30
8	0.50 - 0.68 (Panchromatic)	15
9	1.36 - 1.38 (Cirrus)	30
10	10.60 - 11.19 (Thermal Infrared - TIRS 1)	100
11	11.50 - 12.51 (Thermal Infrared - TIRS 2)	100

เครื่องวัด (Sensor) ที่ถูกติดตั้งมากับ LANDSAT-8 [13] มีทั้งสิ้น 2 เครื่องวัด ดังนี้

- 1) Operational Land Imager (OLI) จะทำได้เก็บภาพในแบนด์ที่ 1 ถึง 9 โดยครอบคลุมพื้นที่ผิวโลก 185×185 ตารางกิโลเมตร และ
- 2) Thermal Infrared Sensor ถูกติดตั้งมาเพื่อการเฝ้าติดตามการใช้ทรัพยากรน้ำและพื้นดิน อาทิ การเฝ้าติดตามปริมาณน้ำที่ถูกใช้ในไร่นาสวนข้าวโพด

2.2.4.2 ดาวเทียม SENTINEL-2A

ดาวเทียม SENTINEL-2A [14] ใช้สำหรับสำรวจทรัพยากรที่ถูกพัฒนาขึ้น โดยผู้นำเทคโนโลยี และอุตสาหกรรมอากาศยาน Airbus Defence and Space ในเครือของ AIRBUS Group ซึ่งเป็นบริษัทผลิตอากาศยานและระบบป้องกันประเทศของยุโรป โดยการส่งดาวเทียม SENTINEL 2A ขึ้นสู่วงโคจรและเริ่มปฏิบัติการวันที่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

16 กรกฎาคม พ.ศ. 2558 ที่วงโคจรความสูงเหนือพื้นโลก 786 กิโลเมตร ซึ่งอยู่ในความรับผิดชอบขององค์การอวกาศยุโรป (European Space Agency – ESA)

ตารางที่ 2.2 อุปกรณ์บันทึกข้อมูลของดาวเทียม SENTINEL-2A

SENTINEL-2A: Multispectral Imager (MSI)		
แบนด์	ค่ากลางความยาวคลื่น (ไมโครเมตร)	รายละเอียดภาพ (เมตร)
1	0.443 (Coastal Aerosol)	60
2	0.490 (Blue)	10
3	0.560 (Green)	10
4	0.665 (Red)	10
5	0.705 (Vegetation Red Edge)	20
6	0.740 (Vegetation Red Edge)	20
7	0.783 (Vegetation Red Edge)	20
8	0.842 (Near Infrared NIR)	10
8A	0.865 (Vegetation Red Edge)	20
9	0.945 (Water vapour)	60
10	1.375 (SWIR - Cirrus)	60
11	1.610 (SWIR)	20
12	2.190 (SWIR)	20

ด้วยคุณสมบัติของอุปกรณ์บันทึกข้อมูล (Sensor) MSI [15] ซึ่งสามารถตรวจจับค่าความยาวช่วงคลื่นได้มากถึง 13 แบนด์ จากเดิมเพียงช่วงที่สามารถมองเห็นได้ด้วยตาเปล่าและช่วงที่มีค่าใกล้อินฟราเรด VNIR (Visible and Near Infrared) มาเป็นช่วงความยาวคลื่นสั้นอินฟราเรด SWIR (Short-Wave Infrared) และครอบคลุมพื้นที่ผิวโลก 290×290 ตารางกิโลเมตร สำหรับวัตถุประสงค์ของดาวเทียม SENTINEL-2A คือการเฝ้าติดตามแก้ไขปัญหาสิ่งแวดล้อม ประกอบกับการตัดสินใจดำเนินนโยบายของสหภาพยุโรปในด้านต่าง ๆ เช่นเกษตรกรรม สาธารณสุข ตลอดจนการบริหารจัดการภัยธรรมชาติ

2.2.4.3 ดาวเทียม SPOT-6 และ SPOT-7

ดาวเทียม SPOT-6 และ SPOT-7 [16] ด้านการสำรวจทรัพยากรธรรมชาติ อยู่ในความรับผิดชอบขององค์การอวกาศฝรั่งเศสร่วมกับประเทศในกลุ่มยุโรป โดย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

SPOT-6 เริ่มปฏิบัติการวันที่ 9 กันยายน พ.ศ. 2555 และ SPOT-7 เริ่มปฏิบัติการวันที่ 30 มิถุนายน พ.ศ. 2557

ตารางที่ 2.3 อุปกรณ์บันทึกข้อมูลของดาวเทียม SPOT 6 และ 7 [17]

SPOT – 6 และ 7: 2 x New AstroSat Optical Modular Instrument (NAOMI)		
แบนด์	ความยาวคลื่น (ไมโครเมตร)	รายละเอียดภาพ (เมตร)
1	0.455 – 0.525 (Blue)	6
2	0.530 – 0.590 (Green)	6
3	0.625 – 0.695 (Red)	6
4	0.760 – 0.890 (Near Infrared NIR)	6
5	0.45 - 0.745 (Panchromatic)	1.5

ดาวเทียม SPOT-6 โคจรอยู่แนวเส้นศูนย์สูตรฝั่งตะวันตก ในขณะที่ดาวเทียม SPOT-7 โคจรอยู่แนวเส้นศูนย์สูตรฝั่งตะวันออก ทั้งนี้สำหรับรายละเอียดการจัดเก็บข้อมูลของภาพถ่ายจะครอบคลุมพื้นที่ผิวโลก 60×60 ตารางกิโลเมตร ตามคุณสมบัติของเครื่องบันทึก NAOMI [18]

2.2.4.4 ดาวเทียม Pleiades 1A และ 1B

ดาวเทียมสำรวจทรัพยากรอวกาศยุโรปเช่นเดียวกับ SENTINEL-2A ที่ถูกพัฒนาโดยบริษัท AIRBUS Defence & Space ดาวเทียมทั้งสองนี้ถูกออกแบบให้ทำงานร่วมกันแบบ Constellation ซึ่งเป็นรูปแบบหนึ่งของการทำงานซึ่งสอดคล้องกับการเคลื่อนที่ที่ห่างกัน 180 องศา (ในกรณี Pleiades 1A และ 1B) ด้วยเหตุนี้ดาวเทียมทั้งสองจึงมีคุณสมบัติที่เหมือนกัน

ตารางที่ 2.4 อุปกรณ์บันทึกข้อมูลของดาวเทียม Pleiades 1A และ 1B [19]

Pleiades 1A และ 1B: MS (Primary) และ Ortho		
แบนด์	ความยาวคลื่น (ไมโครเมตร)	รายละเอียดภาพ (เมตร)
1	0.470 – 0.830 (Panchromatic)	0.5
2	0.430 – 0.550 (Blue)	2
3	0.500 – 0.620 (Green)	2
4	0.590 – 0.710 (Red)	2
5	0.740 – 0.940 (Near Infrared NIR)	2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ดาวเทียม Pleiades 1A เริ่มปฏิบัติการวันที่ 17 ธันวาคม พ.ศ. 2554 และในปีถัดมา Pleiades 2B ได้เริ่มปฏิบัติการวันที่ 2 ธันวาคม พ.ศ. 2555 ดาวเทียมทั้งสองโคจรสูงจากพื้นโลก 694 กิโลเมตร ตามวงโคจรแบบสัมพันธ์กับดวงอาทิตย์ สำหรับวัตถุประสงค์การใช้งานหลัก คืองานที่ต้องการความถูกต้องและแม่นยำสูง เช่น งานด้านพลเรือน การทหาร งานด้านภัยพิบัติ รวมถึงการเฝ้าติดตามการเปลี่ยนแปลงด้านทรัพยากรและพื้นที่ ทั้งนี้หลักการโคจรที่ทั้งสองทำงานร่วมกัน ทำให้สามารถถ่ายภาพซ้ำบริเวณพื้นที่เดียวกันได้ทุกวัน กล่าวคือสามารถส่งถ่ายภาพได้วันละ 3 รอบ ทุก 8 ชั่วโมง เพื่อให้มั่นใจว่าจะได้รับภาพที่ทันต่อเหตุการณ์

2.2.4.5 ดาวเทียม WorldView

ข้อมูลภาพถ่ายจากดาวเทียมในรุ่น WorldView ถูกใช้จาก WorldView-3 และ WorldView-4 ซึ่งทั้งสองเป็นดาวเทียมรายละเอียดสูงในระบบ Multi-Payload ภายใต้การพัฒนาของบริษัท DigitalGlobe สำหรับ WorldView-3 และ WorldView-4 ถูกส่งขึ้นวงโคจรในวันที่ 13 สิงหาคม พ.ศ. 2557 และ วันที่ 11 พฤศจิกายน พ.ศ. 2559 ตามลำดับ ที่ระดับความสูงเดียวกันที่ 617 กิโลเมตร

1) ดาวเทียม WorldView-3 [20, 21] สามารถบันทึกรายละเอียดภาพชนิดขาวดำ (Panchromatic) ที่ 0.3 เมตร และภาพสี (Multispectral) ที่ 1.24 เมตรนอกจากนี้ยังมีแบนด์ความยาวคลื่นสั้นอินฟราเรด (SWIR) และแบนด์สำหรับตรวจจับผ่านชั้นบรรยากาศ (CAVIS) ด้วยความอุปกรณ์บันทึกข้อมูลทั้งสองแบบนี้ทำให้ดาวเทียม WorldView-3 มีคุณสมบัติที่จะเก็บข้อมูลในพื้นที่ที่มีหมอก ควัน และฝุ่นได้ รวมทั้งยังมีรายละเอียดภาพที่สูง เหมาะกับการนำไปประยุกต์ใช้ในการทำแผนที่ การเฝ้าติดตามสภาพสิ่งแวดล้อมและทรัพยากร แม้แต่การติดตามวัตถุหรือการแยกแยะสิ่งปลูกสร้างที่มนุษย์สร้างขึ้น

2) ดาวเทียม WorldView-4 [22] มีคุณสมบัติพื้นฐานและการบันทึกข้อมูลภาพถ่ายดาวเทียมเช่นเดียวกับ WorldView-3 แต่เมื่อเดือนมกราคม พ.ศ. 2562 มีรายงานว่าดาวเทียมดวงนี้ได้สูญเสียการควบคุมแกนการหมุน (Control moment gyroscopes)

2.2.4.6 ดาวเทียม QUICKBIRD

ดาวเทียมเชิงพาณิชย์ของบริษัท Digital-Globe ถูกส่งขึ้นวงโคจรวันที่ 18 ตุลาคม พ.ศ. 2544 ที่ระดับความสูง 450 กิโลเมตร

ตารางที่ 2.5 อุปกรณ์บันทึกข้อมูลของดาวเทียม QUICKBIRD [23]

QUICKBIRD: Multispectral และ Panchromatic		
แบนด์	ความยาวคลื่น (ไมโครเมตร)	รายละเอียดภาพ (เมตร)
1	0.450 – 0.900 (Panchromatic)	0.61
2	0.450 – 0.520 (Blue)	2.44
3	0.520 – 0.600 (Green)	2.44
4	0.630 – 0.690 (Red)	2.44
5	0.760 – 0.900 (Near Infrared NIR)	2.44

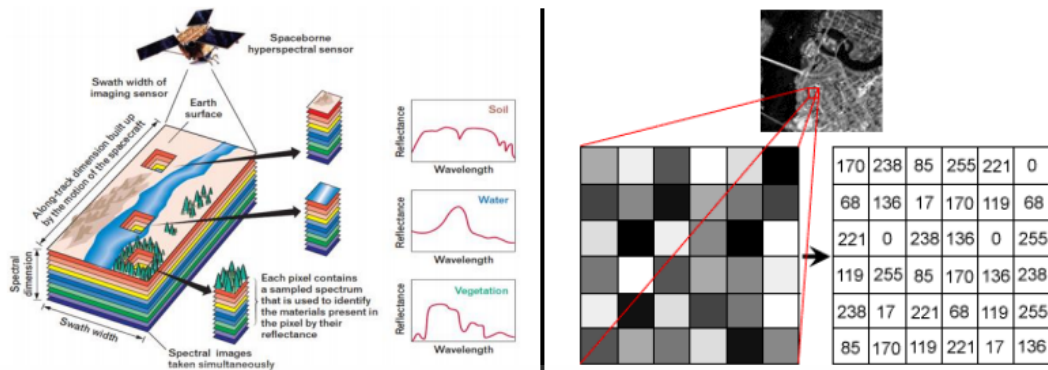
อุปกรณ์ที่ถูกติดตั้งในการบันทึกข้อมูลของ QuickBird นับเป็นดาวเทียมเชิงพาณิชย์ที่สามารถนำมาใช้ประโยชน์ในด้านการสำรวจทรัพยากรได้เช่นกัน [24] แม้ว่าจะมีความกว้างในการบันทึกภาพพื้นที่ผิวโลกที่ 16.5×16.5 ตารางกิโลเมตรก็ตาม แต่เมื่อเทียบกับดาวเทียมเชิงพาณิชย์ดวงอื่น ๆ แล้วยังคงมีประสิทธิภาพที่ดีกว่า รวมถึงรายละเอียดที่สามารถบันทึกได้ยังสามารถแยกแยะประเภทของทรัพยากรได้ดีกว่าข้อมูลจากดาวเทียม LANDSAT และ SPOT เช่น การแยกแยะสวนผลไม้แต่ละชนิดออกจากกัน

คุณสมบัติของดาวเทียมแต่ละรุ่นจะถูกติดตั้งอุปกรณ์บันทึกข้อมูลซึ่งมีคุณสมบัติและนำมาใช้เพื่อจุดประสงค์ที่แตกต่าง [24] ทว่าล้วนแล้วมีพื้นฐานที่สำคัญ คือ กระบวนการบันทึกและจัดเก็บข้อมูล ซึ่งทำให้ภาพถ่ายดาวเทียมมีคุณลักษณะที่แตกต่างกันออกไป

2.2.5 คุณลักษณะภาพถ่ายดาวเทียม

อุปกรณ์บันทึกหรือตัวรับสัญญาณ (Sensor) โดยส่วนมากใช้หลักการที่รับการสะท้อนพลังงานจากพื้นผิวโลก และแปลงค่าสัญญาณให้เป็นข้อมูลตัวเลขที่เรียกว่าตัวเลขดิจิทัล แล้วจึงแสดงผลออกมาเป็นภาพ ซึ่งประกอบไปด้วยจุดสีที่เชื่อมแต่ละจุดที่เรียกว่า พิกเซล (Pixel) วางเรียงต่อกันในรูปแบบคณิตศาสตร์ที่มีจุดตัดแกน x และ y หรือที่เรียกกันว่าภาพ 2 มิติ โดย Pixel จะทำหน้าที่เก็บค่าตัวเลขดิจิทัลที่บ่งบอกถึงค่าสีที่ต่างกันออกไป ทั้งนี้ขึ้นกับรูปแบบการจัดเก็บด้วยเช่นกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.3 การแปลงคลื่นสัญญาณที่ได้และจัดเก็บในรูปแบบภาพดิจิทัล 2 มิติ

ที่มา: พยานหลักฐานดิจิทัลในคดีสิ่งแวดล้อม: ศึกษากรณีเทคโนโลยีการสำรวจระยะไกล [6]

สำหรับรายละเอียดคุณสมบัติที่ทำให้บ่งบอกความเป็นวัตถุ จะอ้างอิงจากคุณสมบัติทั้ง 3 ข้อดังนี้

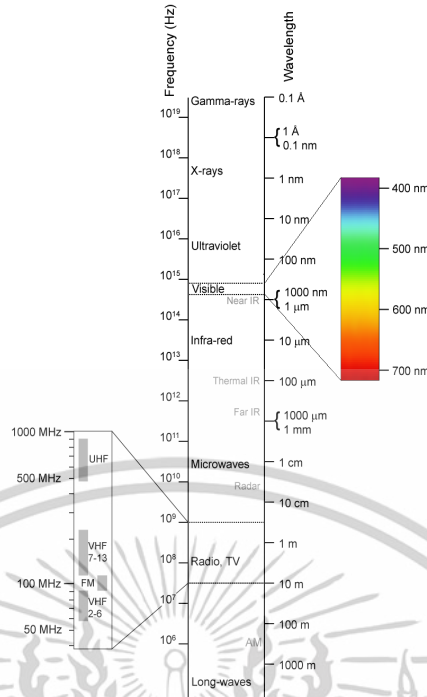
2.2.5.1 การสะท้อนช่วงคลื่นแม่เหล็กไฟฟ้า (Spectral Resolution)

คุณลักษณะใช้บ่งบอก “ช่วงคลื่นและการให้สี (Band and Color)” โดยมีความสัมพันธ์กับความยาวช่วงคลื่นแสงในแต่ละแบนด์ เนื่องจากวัตถุจะสะท้อนแสงในแต่ละช่วงคลื่นไม่เท่ากัน ทำให้สีของวัตถุในภาพแต่ละแบนด์แตกต่างกันในระดับสีเทา-ดำ (Gray Scale) และทำให้สีแตกต่างในภาพสีผสมด้วย โดยช่วงคลื่นที่ใช้ประกอบในการรับรู้จากระยะไกล แบ่งได้เป็น 2 กลุ่ม คือ

- 1 ช่วงคลื่นเชิงแสง (Optical Wavelength) ความยาวคลื่น 0.4 - 14 ไมโครเมตรสามารถถ่ายภาพและบันทึกภาพด้วยฟิล์มถ่ายรูปและเครื่องรับรู้ เป็นช่วงคลื่นที่ตามองเห็น (Visible Light) อยู่ระหว่าง 0.4 - 0.7 ไมโครเมตร แบ่งเป็น 3 ช่วง คือ น้ำเงิน เขียวและแดง ถัดมาเป็นช่วงคลื่นสั้นอินฟราเรด (Near Infrared) หรืออินฟราเรดสะท้อนซึ่งอยู่ระหว่าง 0.7 - 3 ไมโครเมตร และอินฟราเรดความร้อน (Thermal Infrared) ระหว่าง 3 - 15 ไมโครเมตร

- 2 ช่วงคลื่นไมโครเวฟ (Microwave) ความยาวคลื่น 1 มิลลิเมตรถึง 1 เมตร มักนิยมเรียกหน่วยนับเป็นหน่วยความถี่ ต่างจากกลุ่มช่วงคลื่นเชิงแสงที่มีหน่วยเป็นความยาวคลื่น เช่น ระบบเรดาร์ จะบันทึกข้อมูลในช่วงคลื่นความถี่ระหว่าง 3 - 12.5 GHz (ความยาวคลื่นระหว่าง 2.4 - 100 เซนติเมตร) นิยมใช้ตัวอักษรบอกช่วงคลื่น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.4 สเปกตรัมแม่เหล็กไฟฟ้า (Electromagnetic Spectrum, Louis E. Keiner, Coastal Carolina University, ไม่ทราบปีที่เผยแพร่)

2.2.5.2 ขนาดพื้นที่และรูปร่าง (Spatial Resolution)

คุณลักษณะใช้บ่งบอก“ขนาดพื้นที่ที่ให้รายละเอียดได้ (Area and Detailed)” หมายถึง พื้นที่ขนาดเล็กที่สุดของวัตถุที่จะถูกตรวจจับได้โดยอุปกรณ์บันทึกข้อมูล หรือถูกแสดงออกมาในรูปของจุดภาพ ความละเอียดเชิงพื้นที่ที่มีความแตกต่างกันตามมาตราส่วนและรายละเอียดภาพจากดาวเทียม เช่น ความละเอียดที่ 70 เมตร จะหยابกว่าความละเอียดที่ 10 เมตร ซึ่งจะสัมพันธ์กับมาตราส่วนภาพ ทั้งนี้ อัตราส่วนของระยะทางบนแผนที่จะสัมพันธ์กับระยะทางจริงบนพื้นที่โลก ตัวอย่างภาพดังรูปที่ 2.5



รูปที่ 2.5 เปรียบเทียบรายละเอียดภาพจากรายละเอียดสูง (ซ้าย) ไปยังรายละเอียดต่ำ (ขวา)

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์: What is Remote Sensing? The Definitive Guide (GISGeography, 2024) ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2.5.3 ช่วงเวลาการถ่ายภาพ (Temporal Resolution)

คุณลักษณะใช้ปรงรอบและระยะเวลา หรือความถี่ในการกลับมาถ่ายภาพซ้ำ ตรงตำแหน่งเดิม เป็นการติดตามการเปลี่ยนแปลงของวัตถุ เช่น การเปลี่ยนแปลงตามช่วงฤดูกาล การเปลี่ยนแปลงรายปี หรือรายคาบ โดยลักษณะการเปลี่ยนแปลงดังกล่าวทำให้มีความแตกต่างของระดับสีในภาพทำให้เราสามารถใช้อัตราดาวเทียมที่ถ่ายซ้ำที่เดิมในช่วงเวลาต่าง ๆ มาติดตามความเปลี่ยนแปลงได้



รูปที่ 2.6 ภาพถ่ายดาวเทียมจาก Google เปรียบเทียบสภาพเมือง Kamaishi ประเทศญี่ปุ่นก่อน (ซ้าย) และหลัง (ขวา) ประสบภัยพิบัติคลื่นสึนามิ

ที่มา: Empowerment and Education Saves Lives: Tendenko in Kamaishi (MIRIANNA, 2014)

คุณลักษณะทั้ง 3 ประการที่กล่าวมามีความเกี่ยวข้องกับกระบวนการประมวลผลภาพ (Image Processing) ทั้งการแปลงค่าที่ได้รับจากอุปกรณ์แบบอนาล็อกเพื่อนำไปบันทึกในรูปแบบดิจิทัล ตลอดจนนำไปประยุกต์ใช้ด้วยวิธีการต่าง ๆ

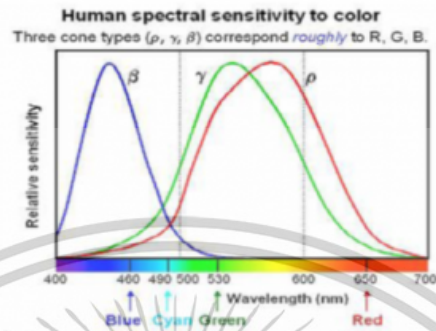
2.3 ทฤษฎีเกี่ยวกับการประมวลผลภาพ

2.3.1 การแปลงภาพอนาล็อกเป็นภาพดิจิทัล

การแปลงภาพอนาล็อกเป็นภาพดิจิทัล เป็นการได้มาของภาพที่ได้จากการบันทึกภาพโดยอุปกรณ์ที่เลียนแบบการรับภาพของมนุษย์ให้เป็นภาพสองมิติ และจะเรียกภาพที่บันทึกได้ว่า ภาพเชิงกายภาพ (Physical Image) ซึ่งเป็นภาพที่เกิดจากกระบวนการแสงที่ตกกระทบกับวัตถุ โดยที่แสงสะท้อนจากวัตถุออกมาเข้าสู่ฉากรับแสงมีความยาวคลื่นแบบต่อเนื่องที่แตกต่างกันตามสีของวัตถุ ดังรูปที่ 2.4 แสดงแถบสเปกตรัมของแสงในความถี่ที่มนุษย์สามารถรับรู้ได้ซึ่งมีสีที่แตกต่างกันตามความยาวคลื่น เนื่องจากเซลล์รูปกรวยภายในเรตินา (Cones) สามารถซึมซับแสงสีแดง เขียวและน้ำเงิน (Red Green Blue : RGB) ได้เพียงผิวเผิน โดยสีที่รับรู้ได้มีลักษณะของการรวมกันแบบไม่คงที่ของสีเริ่มต้น (Primary Colors) คือ แดง เขียว และน้ำเงิน ในรูปที่ 2.7 แสดงเส้นโค้งที่อธิบายการตอบสนองการซึมซับแสงของ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

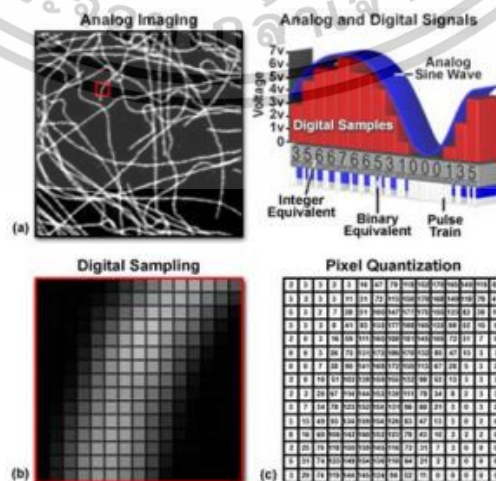
เซลล์รูปกรวยภายในเรตินา ดังนั้น ภาพเชิงกายภาพเป็นการเลียนแบบการทำงานของเซลล์รับรูสี (Cones Cell) และเซลล์รับรูมืดของแสง (Rods Cell) ของเรตินาโดยแทนที่เรตินาด้วยฉากรับแสงในการบันทึกภาพ



รูปที่ 2.7 สเปกตรัมของแสงที่มนุษย์สามารถซึมซับสีด้วยเซลล์รูปกรวยในเรตินา

ภาพที่ถูกบันทึกมาแล้ว ระบบจะนำรูปที่รับเข้ามาไปคำนวณ โดยเข้ากระบวนการสุ่มตัวอย่าง (Sampling) และการจัดระดับสัญญาณ (Quantization) และส่งข้อมูลออกมาในรูปแบบดิจิทัล คอมพิวเตอร์จะเก็บข้อมูลภาพลงหน่วยความจำ โดยจะอยู่ในรูปของแผ่นตารางโดยแต่ละช่องจะเป็นส่วนหนึ่งของภาพหรืออักษร เรียกแต่ละจุดหรือช่องนั้นว่า พิกเซล (pixel) แต่ละพิกเซลจะถูกกำหนดให้มีระดับของความเข้ม (สีดำ สีขาว สีเทา หรือสีอื่น ๆ) ซึ่งแสดงให้อยู่ในรูปของรหัสไบนารี (Binary ได้แก่ 0 และ 1) แต่ละพิกเซลจะแทนด้วยไบนารีดิจิทัล (Binary digital หรือ bits) จะถูกเก็บเป็นลำดับในคอมพิวเตอร์และโดยทั่วไปจะถูกลดขนาดลงด้วยวิธีการทางคณิตศาสตร์ หรือกล่าวคือการบีบอัดให้เล็กลง

การดำเนินการ Sampling คือ การแปลงค่าพิกัดให้เป็นเลขจำนวนเต็ม แล้วจึงนำไปการแปลงค่าของฟังก์ชันให้เป็นเลขจำนวนเต็มด้วยกระบวนการ Quantization ดังรูปที่ 2.8



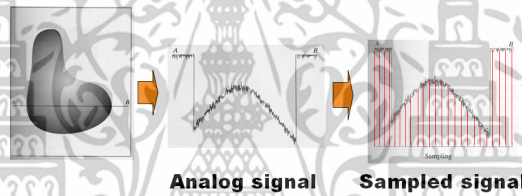
รูปที่ 2.8 ภาพกระบวนการ Image Sampling และ Quantization

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น เมื่ออนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ลำดับต่อไป จะกล่าวถึงรายละเอียดและวิธีการดำเนินการของกระบวนการสุ่มตัวอย่าง (Sampling) และกระบวนการการจัดระดับสัญญาณ (Quantization)

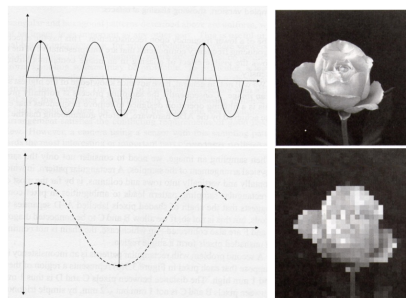
2.3.1.1 กระบวนการสุ่มตัวอย่าง (Sampling)

ขั้นตอนการสร้างภาพดิจิทัล ซึ่งเป็นตัวแปลงข้อมูลต่อเนื่อง (Continuous Interval) ให้อยู่ในช่วงไม่ต่อเนื่องที่กำหนด (Discrete interval) ซึ่งจะสัมพันธ์กับจำนวนพิกเซล คือ จำนวนจุด หรือ สีเหลี่ยมจัตุรัสเล็ก ๆ ที่เก็บค่าระดับความเข้มแสงในภาพ ซึ่งความถี่ในการสุ่มตัวอย่างมีผลต่อจำนวนพิกเซล และจำนวนพิกเซลมีผลต่อความละเอียดของภาพ ดังนั้นหากมี Sampling Rate สูง ภาพจะมีความละเอียดมากและขนาดของไฟล์ภาพจะมีขนาดใหญ่ นอกจากนี้ในกระบวนการบันทึกหรือรับภาพจะทำได้ซ้ำ (Single Sensor / Row Sensor) และยังต้องเพิ่มตัวประมวลผลเพื่อจัดเก็บค่า CCD (Array Sensor) โดยตัวอย่างนำสัญญาณอนาล็อกผ่านกระบวนการสุ่มตัวอย่างเป็นดังรูปที่ 2.9



รูปที่ 2.9 แสดงตัวอย่างการทำกระบวนการสุ่มตัวอย่าง

ตามทฤษฎีของ Nyquist จะกำหนดความถี่ในการสุ่มตัวอย่างให้มากกว่า 2 เท่าของความถี่สูงสุดของสัญญาณ หากกำหนด Sampling Rate ไม่เป็นไปตามทฤษฎีของ Nyquist จะทำให้เกิดผลกระทบหรือปรากฏการณ์ที่เรียกว่า Aliasing คือ การบิดเบือนของสัญญาณทำให้การแปลงค่าผิดเพี้ยนไป เมื่อนำมาแสดงผลจะทำให้ลักษณะของภาพที่ได้ผิดเพี้ยนไป อธิบายได้ดังภาพที่ 2.10 ซึ่งแสดงสัญญาณกรณี Sampling rate เป็นไปตามทฤษฎีและผลลัพธ์ ก็กรณี Sampling rate ไม่เป็นไปตามทฤษฎีและผลลัพธ์



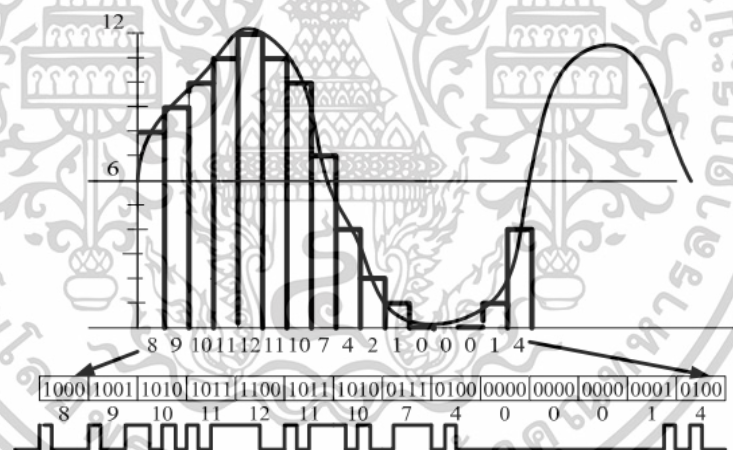
รูปที่ 2.10 ภาพตัวอย่างกรณี Sampling rate ตามทฤษฎี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการแข่งขันเพื่อการศึกษาเท่านั้น เมื่อผู้ผู้ใดเห็นประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.3.1.2 กระบวนการจัดระดับสัญญาณ (Quantization)

การนำผลของการสุ่มค่ามาเทียบค่าของจุดในการสุ่มระดับสัญญาณต่อ 1 ลูกคลื่น แล้วจะทำการวัดค่าของสัญญาณที่สุ่มตัวอย่างมา ซึ่งค่าที่ได้จะเป็นจำนวนจริง จึงจำเป็นต้องแปลงค่าเป็นจำนวนเต็มที่สามารถเก็บเป็นเลขฐานสองได้ โดยจำนวนเต็มที่ใช้ คือ ค่าลำดับขั้นของการจัดระดับสัญญาณ (quantize) หากจำนวนขั้นมาก ๆ จะทำให้สามารถแทนระดับของสัญญาณภาพได้ถูกต้องมากขึ้น เนื่องจากคอมพิวเตอร์เก็บค่าเป็นเลขฐานสอง ลำดับขั้นที่เก็บระดับสัญญาณจะเป็น $n = 2^b$ ช่วง โดยมีค่าตั้งแต่ 0 ถึง $2^b - 1$ (b เป็นจำนวน bit ในการทำ quantize)

การแสดงผลจะเริ่มตั้งแต่สีขาวไล่ความเข้มจนถึงสีดำ เรียกจำนวนขั้นของ quantize ว่า Gray Level และเรียกการเปลี่ยนระดับค่าของสีขาวเป็นดำ ว่า Grayscale ทั้งนี้ภาพถ่ายดิจิทัลส่วนใหญ่จะใช้ 8 bit มาดำเนินการ quantize ทำให้ได้ Gray Level 256 ระดับ (ค่า Grayscale 0 – 255) แต่สำหรับภาพถ่ายทางการแพทย์หรือภาพถ่ายดาวเทียม ซึ่งจำเป็นต้องใช้รายละเอียดมากก็จะใช้ค่า bit ที่สูงกว่าภาพทั่วไปมาดำเนินการ สำหรับภาพตัวอย่างทางทฤษฎีเป็นดังรูปที่ 2.11



รูปที่ 2.11 แสดงตัวอย่างการจัดระดับสัญญาณ (Quantization)

ภาพดิจิทัลเกิดจากการแสดงผลของค่าสีที่เก็บตามการแปลงสัญญาณ ทว่าทฤษฎีของภาพดิจิทัลและรูปแบบการจัดเก็บของภาพ จะส่งผลถึงแตกต่างของภาพ หรือกล่าวได้ว่าการเก็บข้อมูลของภาพดิจิทัลที่ต่างกันทำให้มีประเภทของภาพดิจิทัลเกิดขึ้นมา ซึ่งจะกล่าวถึงรายละเอียดในหัวข้อ 2.3.2

2.3.2 ภาพดิจิทัลและประเภทของภาพดิจิทัล

ภาพดิจิทัล คือ function 2 มิติ ($f(x, y)$) โดยที่ x และ y คือ พิกัดในระนาบ (spatial/plane coordinate) และค่าของ f สำหรับ (x, y) ใด ๆ คือ ความเข้มแสงหรือค่า

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ระดับเทา (intensity/gray level) เมื่อ x, y และ f มีค่าแน่นอน และเป็นจำนวนเต็ม ซึ่งจะแสดงผลภาพในลักษณะสองมิติในหน่วยที่เรียกว่าพิกเซล

ภาพดิจิทัลสามารถนิยามเป็นฟังก์ชันสองมิติ $f(x, y)$ โดยที่ x และ y เป็นพิกัดของภาพ และแอมพลิจูดของ f ที่พิกัด (x, y) ใด ๆ ภายในภาพ คือ ค่าความเข้มแสงของภาพ (Intensity) ที่ตำแหน่งนั้น ๆ และเมื่อ x, y และแอมพลิจูดของ f เป็นค่าจำกัด (Finite value) จึงเรียกรูปภาพนี้ว่าเป็นภาพดิจิทัล (Digital Image) และถ้ากำหนดให้ภาพ $f(x, y)$ มีขนาด M แถวและ N คอลัมน์ และพิกัดของจุดกำเนิดของภาพคือที่ตำแหน่ง $(x, y) = (0, 0)$ แล้ว จะสามารถเขียนสมการให้อยู่ในรูปเมทริกซ์ได้ดังนี้

$$f(x, y) = \begin{bmatrix} f(0,0) & f(0,1) & \dots & f(0,N-1) \\ f(1,0) & f(1,1) & \dots & f(1,N-1) \\ \vdots & \vdots & \ddots & \vdots \\ f(M-1,0) & f(M-1,1) & \dots & f(M-1,N-1) \end{bmatrix}$$

ค่าแต่ละค่าที่อยู่ในเมทริกซ์จะเรียกว่า พิกเซล โดยตำแหน่ง $(0,0)$ จะอยู่ทางด้านซ้ายมือสุดด้านบนของภาพ การจัดลำดับตำแหน่งของจุดภาพจะเรียงจากซ้ายไปขวาในแต่ละเส้นจุด และจัดลำดับของเส้นจุดจะเรียงจากบนลงล่างการเก็บค่าของความเข้มแสงของภาพดิจิทัลลงหน่วยความจำในลักษณะเส้นจุด (raster) นี้จะเรียกภาพบิตแมป (bit-map image) หรือภาพราสเตอร์ (Raster image) แต่ภาพที่จัดเก็บในลักษณะนี้จะมีขนาดใหญ่จึงมีการบีบอัดภาพ (image compression) เพื่อให้ข้อมูลภาพมีขนาดเล็กลง โดยรายละเอียดของการจัดภาพและแสดงผลเป็นดังนี้

1) จุดภาพ (Pixel) เป็นหน่วยพื้นฐานของภาพ คือจุดภาพบนจอแสดงผล หรือจุดภาพในรูปภาพที่รวมกันเป็นภาพขึ้น โดยภาพหนึ่ง ๆ จะประกอบไปด้วยจุดภาพหรือพิกเซลจำนวนมาก และแต่ละภาพที่สร้างขึ้นจะมีความหนาแน่นของจุดภาพหรือบางครั้งแทนว่าความละเอียด (ความคมชัด) ที่แตกต่างกันไป จึงใช้ในการบอกคุณสมบัติของภาพ จอภาพ หรือ อุปกรณ์แสดงผลภาพได้

2) ภาพบิตแมป (Bitmap Image) หรือเรียกอีกอย่างหนึ่งว่าภาพแบบราสเตอร์ (Raster Image) เป็นภาพที่ประกอบขึ้นจากจุดขนาดเล็ก ๆ ที่เรียกว่า พิกเซลจำนวนมากที่เรียงต่อกันเป็นภาพภาพหนึ่ง เพื่อให้เห็นภาพลักษณะนี้ชัดเจนยิ่งขึ้น

ภาพบิตแมปนี้ จะมีจำนวนจุดขนาดเล็ก ๆ จำนวนมาก ดังนั้น ดวงตาของมนุษย์ไม่สามารถที่จะมองเห็นและแยกแยะรายละเอียดส่วนย่อยเล็ก ๆ นั้นได้ แต่เมื่อลองขยายภาพดูจะเห็นเป็นรูปตาราง ยิ่งขยายมากเท่าไร ตารางสี่เหลี่ยมจะยิ่งมีขนาดใหญ่ขึ้นจนทำให้มองเห็นจุดของภาพหรือพิกเซล จึงมีผลทำให้รายละเอียดของ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพมีความไม่ชัดเจนมากขึ้น โดยทั่วไปแล้วภาพบิตแมป (Bitmap Image) เป็นประเภทภาพที่นิยมใช้กันมาในภาพถ่าย ตัวอย่างดังรูปที่ 2.12



รูปที่ 2.12 ตัวอย่างภาพบิตแมป (Bitmap Image)

3) ภาพเวกเตอร์ (Vector Image) เป็นภาพที่ประกอบด้วยเส้นตรง เส้นโค้ง และรูปทรงต่าง ๆ ภาพที่ได้จะสร้างขึ้นจากคำสั่งที่บอกถึงลักษณะของภาพในรูปทรงเรขาคณิตด้วยสมการทางคณิตศาสตร์ ซึ่งมีจุดเด่น คือ เมื่อขยายภาพ คอมพิวเตอร์จะคำนวณค่าต่าง ๆ ให้ใหม่ทุกครั้งที่ทำกรขยาย ทำให้ภาพที่เกิดขึ้นมามีความคมชัด ภาพเวกเตอร์จึงเหมาะสำหรับงานที่มีความแม่นยำและต้องการความละเอียดสูง เช่น การสร้างภาพสามมิติ การสร้างแบบร่างทางวิศวกรรม ตัวอย่างดังรูปที่ 2.13



รูปที่ 2.13 ตัวอย่างภาพเวกเตอร์ (Vector Image)

ภาพดิจิทัลและประเภทภาพดิจิทัลที่กล่าวถึง ได้เชื่อมโยงไปสู่ประเภทของภาพอีกหลายประเภท โดยมีความนิยมนำไปใช้งานที่ต่างกันออกไปตามวัตถุประสงค์การใช้งาน ทั้งนี้รวมไปถึงการประยุกต์ใช้งานทางด้านประมวลผลภาพ หรือกล่าวคือ ภาพแต่ละประเภทอาจแสดงคุณลักษณะที่แตกต่างกันได้ ซึ่งจะกล่าวถึงในส่วนถัดไป

2.3.3 ประเภทของภาพดิจิทัล

การจัดเก็บข้อมูลของภาพดิจิทัลทำให้คุณลักษณะของภาพแตกต่างกันออกไป เช่น การจัดเก็บค่าแบบ 1 บิต (Bit) คือ การแทนค่าแต่ละพิกเซลเป็นค่า 0 หรือ 1 ซึ่ง 0 แสดงจะ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

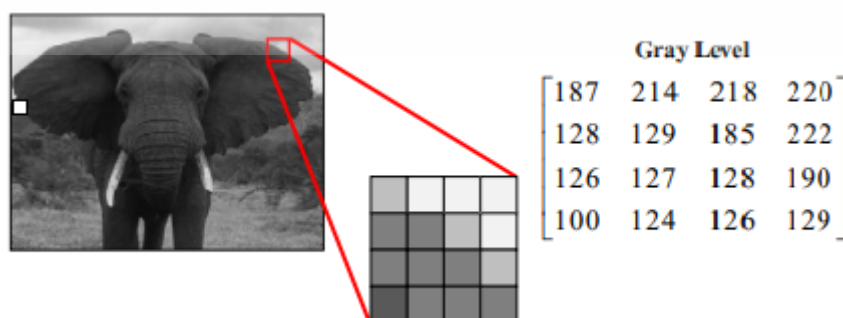
ให้ค่าสีดำ และ 1 จะให้ค่าสีขาว ทั้งนี้ประเภทของภาพมักถูกนำไปใช้งานที่แตกต่างกัน สำหรับในงานวิจัยนี้ได้ยกตัวอย่างประเภทของภาพดิจิทัลที่ถูกนำมาใช้ในการทดลอง ดังนี้

- 1) ภาพขาวดำ (Binary Image) คือภาพที่แต่ละพิกเซล (Pixel) จะแสดงด้วยค่าแบบไบนารี (Binary) คือ มี 1 บิต ซึ่งจะประกอบไปด้วยค่า 0 และ 1 โดยที่ 1 หมายถึง จุดภาพสีขาว และ 0 หมายถึง จุดภาพสีดำ ภาพประเภทนี้ถูกนำมาใช้งานวิจัยสำหรับภาพคำตอบจาก หรือผลลัพธ์จากการทำนายด้วยการเรียนรู้เชิงลึกในกระบวนการแบ่งส่วน ตัวอย่างดังรูปที่ 2.14



รูปที่ 2.14 ตัวอย่างภาพขาวดำ (Binary Image)

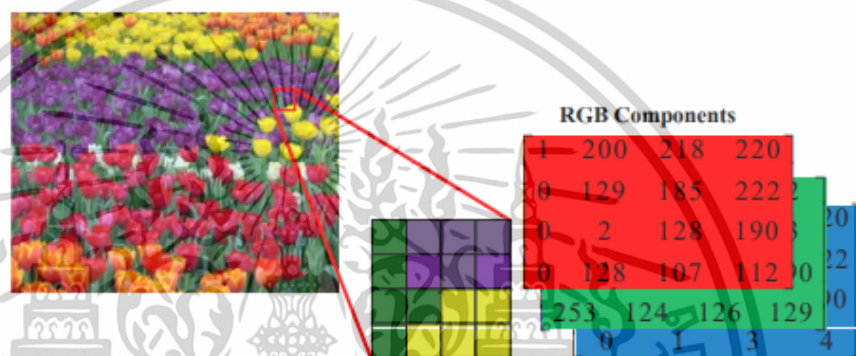
- 2) ภาพระดับสีเทา (Grayscale Image) เป็นภาพที่แต่ละพิกเซล จะมีค่าความเข้มของแสงในแต่ละระดับที่แตกต่างกันไป ตั้งแต่สีขาวไปยังสีดำ โดยสามารถกำหนดระดับความเข้มของแสงนั้นได้ โดยใช้ค่าระดับเทา โดยปกติแล้วภาพแบบระดับสีเทามีความละเอียด (Resolution) เท่ากับ 8 บิต ซึ่งภาพจะมีค่าระดับความเข้มแสงของสีดำเท่ากับ 0 ส่วนค่าระดับความเข้มแสงของสีขาวจะมีค่าเท่ากับ 255 ภาพประเภทนี้ได้ถูกนำมาใช้จากกรณีการแปลงค่าสีของภาพเพื่อเป็นการลดมิติของข้อมูลภาพ และใช้เพื่อแสดงความเข้มแสงของสีในแต่ละวัตถุ ตัวอย่างดังรูปที่ 2.15



รูปที่ 2.15 ตัวอย่างภาพระดับสีเทา (Gray Scale Image)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3) ภาพสี (Color Image – RGB) เป็นภาพที่แต่ละพิกเซลจะเก็บค่าระดับความเข้มแสงแต่ละแถบแสงของแม่สี 3 สีที่ซ้อนกัน คือ สีแดง เขียว และน้ำเงิน (Red Green Blue : RGB) ซึ่งในแต่ละพิกเซล ก็จะแสดงผลของค่าสีแต่ละพิกเซลตามระดับความเข้มในแต่ละแถบแสงนั้น ภาพประเภทนี้มีคุณลักษณะที่แสดงความแตกต่างของวัตถุภายในภาพได้ชัดเจนมากที่สุด เนื่องจากวัตถุแต่ละชนิดมีการสะท้อนค่าแสงที่ต่างกัน เมื่อนำไปบันทึกเป็นภาพจะแสดงความเป็นวัตถุแต่ละชนิดได้อย่างชัดเจน เช่น พื้นที่ป่าภายในภาพถ่ายดาวเทียมจะแสดงออกเป็นสีเขียว เป็นต้น โดยตัวอย่างภาพดังรูปที่ 2.16



รูปที่ 2.16 ตัวอย่างภาพสี (Color Image)

ภาพแต่ละประเภทเกิดจากการถูกจัดเก็บด้วยค่าและมิติของข้อมูลที่ต่างกัน ซึ่งรวมไปถึงคุณลักษณะของภาพด้วยเช่นกัน ดังนั้นการใช้งานภาพแต่ละประเภทควรพิจารณาจากคุณลักษณะของภาพตามความเหมาะสม โดยในลำดับต่อไปจะอธิบายถึงลักษณะของภาพตามหัวข้อ 2.3.4

2.3.4 คุณลักษณะของภาพ

ภาพดิจิทัลซึ่งถูกจัดเก็บด้วยค่าตัวเลขในแต่ละตำแหน่งที่เรียกกันว่าพิกเซล เมื่อนำไปประยุกต์ใช้งานมักมีคุณภาพที่ต่างกันไป เช่น ภาพถ่ายดาวเทียมแต่ละดวงมีความละเอียดภาพที่ไม่เหมือนกัน เมื่อนำไปใช้วิเคราะห์หรือใช้งาน ก็จะมีคุณสมบัติที่ต่างกันออกไป นอกจากนี้ ยังรวมไปถึงเทคนิคในการวิเคราะห์และปรับปรุงภาพเพื่อให้ภาพมีคุณลักษณะที่เหมาะสมกับการนำไปใช้งานต่อ ผู้วิจัยจึงได้นำตัวอย่างคุณลักษณะของภาพที่เกี่ยวข้องในงานวิจัยนี้มานำเสนอ ดังนี้

- 1) ค่าความละเอียดภาพ (Resolution) คือ ค่าความละเอียดของภาพ ๆ หนึ่ง ซึ่งกำหนดเป็นจำนวนพิกเซลต่อหนึ่งหน่วยความยาวของภาพ ตัวอย่างเช่น หากภาพนั้นมีค่า Resolution เท่ากับ 150 pixels/inches แสดงว่าในพื้นที่ 1 ตารางนิ้วของ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพนั้นประกอบไปด้วยพิกเซลจำนวน 150 พิกเซล หรือหากกล่าวภาพที่นำมาใช้ มีขนาด 256×256 pixels แสดงว่าภาพดังกล่าวมีความกว้างคูณความยาว คือ 256×256 เท่ากับ 65,536 pixels

2) ความสว่างของภาพ (Image Brightness) คือ ค่าเฉลี่ยของค่าระดับเทาในแต่ละพิกเซลของภาพทั้งหมด ซึ่งถ้าเรากำหนดให้ภาพมีขนาด $N \times M$ และ $I(x, y)$ คือ ค่าระดับความเข้มเทาของแต่ละพิกเซลที่จะพิจารณา เราจะสามารถคำนวณค่าความสว่างของภาพได้จากสมการต่อไปนี้

3) ค่าความแตกต่างสี (Image Contrast) คือ ความแตกต่างของความสว่างและความมืดบนภาพ ซึ่งช่วยให้แยกรายละเอียดของภาพออกจากกันได้ ค่าความแตกต่างสีมี 2 ประเภท คือ 3.1) สเกลคอนทราสต์สั้น (short scale contrast) หมายถึง ค่าความแตกต่างสีของภาพที่แสดงให้เห็นระดับความแตกต่างของสีดำและสีขาว 3.2) สเกลคอนทราสต์ยาว (long scale contrast) หมายถึง ค่าความแตกต่างสีของภาพที่แสดงให้เห็นระดับความแตกต่างระหว่างสีดำกับสีขาว กล่าวคือ ส่วนที่เป็นสีเทาใช้ความเข้มของสีเทาจะต่างกันออกไป

4) ฮิสโตแกรมของภาพ (Image Histogram) คือ กราฟที่แสดงถึงจำนวนของพิกเซลที่ความสว่างต่างๆของภาพทั้งในระบบค่าแสงสีแดง เขียว และน้ำเงิน (Red Green Blue : RGB) และค่าระดับสีเทา (Gray scale) โดยแบ่งระดับความสว่างออกเป็น 256 ระดับตั้งแต่ 0 ถึง 255 โดยในแกนนอนไล่ระดับความสว่างซ้ายมือที่มีค่าความสว่างน้อย (สีดำ) ไปยังด้านขวามือที่มีค่าความสว่างมาก (สีขาว) และแกนตั้งแสดงถึงจำนวนพิกเซล ในแต่ละระดับความสว่าง ตัวอย่างดังรูปที่ 2.17



รูปที่ 2.17 ตัวอย่างฮิสโตแกรมของภาพ (Image Histogram)

จากที่กล่าวไปข้างต้น คุณลักษณะของภาพมีผลต่อการนำไปใช้งาน ดังนั้นแล้ว

กระบวนการประมวลผลภาพ คือ การจัดการกับคุณลักษณะของภาพไม่ว่าจะความเข้ม ความเอกซารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สว่างของสี ลักษณะของวัตถุที่ปรากฏภายในภาพโดยอาศัยความต่างระหว่างสีของขอบวัตถุ กับพื้นหลังของภาพ แน่แน่นอนว่าเทคนิคการประมวลผลภาพมีอยู่หลายวิธี ฉะนั้นในหัวข้อถัดไป จะนำเสนอเทคนิคการประมวลผลภาพที่ถูกนำมาประยุกต์ในการทดลองของงานวิจัยชิ้นนี้

2.3.5 การประมวลผลภาพดิจิทัล

การประมวลผลภาพดิจิทัล เป็นสาขาที่กล่าวถึงเทคนิคและอัลกอริทึมต่าง ๆ ที่ใช้การประมวลผลภาพที่อยู่ในรูปแบบดิจิทัล เพื่อปรับปรุงภาพให้สามารถนำมาใช้งานได้ง่ายมากขึ้น จัดการข้อมูลภาพสำหรับการเก็บและส่งภาพ ช่วยวิเคราะห์รูปภาพอัตโนมัติ มักถูกเรียกรย่อ ๆ ว่า การประมวลผลภาพ (Image Processing) หรือก็คือ การเรียกใช้ขั้นตอนหรือกรรมวิธีใด ๆ มากกระทำกับภาพ โดยมีวัตถุประสงค์เพื่อการปรับปรุงคุณภาพของภาพ ให้ได้ภาพใหม่ที่มีคุณสมบัติตามต้องการ เช่น ความคมชัด หรือการประหยัดพื้นที่ในการเก็บข้อมูล หรือใช้สำหรับการประมวลผลระดับสูง เช่น การจดจำรูปร่างลักษณะให้ได้อย่างแม่นยำ โดยทั่วไปแล้วระดับของการประมวลผลภาพแบ่งได้เป็น 3 ระดับคือ

- 1) ระดับต่ำ (Low-Level) เป็นการจัดการเกี่ยวกับค่าสี และจุดภาพ เช่น การแปลงปริภูมิสี (Color Space) ภาพไบนารี การวิเคราะห์ตัดขี้นี้สี การแยกสี
- 2) ระดับกลาง (Mid-Level) เป็นนำเสนอความรู้ (Knowledge) จากภาพ เช่น การสกัดลักษณะเด่น (Feature Extraction) การสกัดรูปทรงและวิเคราะห์รูปร่าง (Geometry Extraction)
- 3) ระดับสูง (High-Level) เป็นการวิเคราะห์และนำเสนอความรู้และการรู้จำลักษณะ หรือลวดลายของภาพ (Recognition)

ฉะนั้นจึงขึ้นกับว่าต้องการนำภาพมาดำเนินการในระดับใด แต่โดยพื้นฐานของการประยุกต์ใช้เทคนิคการประมวลผลภาพแล้วจะมีขั้นตอน ดังนี้

- 1) การได้มาของรูป (Image Acquisition) คือ การนำภาพเข้าระบบโดยอาจจะมี การประมวลผลภาพก่อน (Image Processing) เช่น การปรับขนาดของภาพ การลดสัญญาณรบกวน
- 2) การปรับปรุงภาพ (Image Enhancement) กระบวนการปรับปรุงภาพให้ดีขึ้น เพื่อประโยชน์การแปลภาพด้วยการมองด้วยตา (Visual Interpretation) โดยที่ไม่มีการเปลี่ยนแปลงเนื้อหาของภาพ ซึ่งทำให้ภาพมีความชัดเจนขึ้น หรือทำให้คุณลักษณะที่สำคัญเด่นขึ้น
- 3) การคืนสภาพของภาพ (Image Restoration) การทำให้ภาพคืนสู่สภาพเดิม หรือการปรับปรุงภาพให้เหมาะสมกับการมองเห็นซึ่งจะเกี่ยวข้องกับความสัมพันธ์ (Degradation) ของรูปภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- 4) การประมวลผลภาพสี (Color Image Processing) คือ การประมวลผลภาพโดยใช้รูปแบบของสีที่สำคัญภายในภาพที่สนใจมาช่วยในการแยกคุณลักษณะ
- 5) การบีบอัดรูปภาพ (Image Compression) คือ การบีบอัดข้อมูลภาพซึ่งจะทำให้ภาพมีขนาดเล็กลง มีอยู่ด้วยกัน 2 แบบ คือ การบีบอัดแบบไม่มีการสูญเสียรายละเอียดข้อมูล (Lossless compression) ซึ่งค่าความสว่างของแต่ละจุดภาพจะยังคงอยู่เหมือนเดิมทุกประการ หรือไม่มีการเปลี่ยนแปลงค่าของแต่ละจุดภาพ ซึ่งการบีบอัดวิธีนี้จะอาศัยเทคนิคการจัดเก็บข้อมูลเชิงตัวเลขในการลดขนาดของข้อมูล และการบีบอัดแบบสูญเสียรายละเอียดข้อมูล (Lossy compression) วิธีการนี้จะมีการเปลี่ยนแปลงค่าความสว่างของจุดภาพนั้น หมายความว่า วิธีการนี้ไม่เหมาะสมสำหรับข้อมูลภาพที่ต้องมีการจำแนกข้อมูล (Classification)
- 6) การเปลี่ยนแปลงลักษณะรูปร่างหรือลักษณะสัณฐานวิทยา (Morphological processing) คือ การประมวลผลของภาพทางด้านโครงสร้าง โดยเกี่ยวกับการแยกส่วนประกอบของภาพออกเพื่อใช้ในการแสดงรูปร่างด้วยตัวดำเนินการเมทริกซ์ (Matrix) โดยจะใช้กับภาพไบนารี (Binary Image)
- 7) การแบ่งข้อมูลภาพ (Image Segmentation) คือ การแบ่งรูปออกเป็นหลายพื้นที่ (regions) หรือแบ่งออกตามวัตถุที่สนใจ ในกระบวนการนี้จำเป็นต้องวิเคราะห์ถึงคุณลักษณะของวัตถุภายในภาพที่สนใจ และประยุกต์ใช้เทคนิคการประมวลผลภาพตามความเหมาะสม แต่ในปัจจุบันได้มีการนำเทคนิคการเรียนรู้ด้วยเครื่องจักร (Machine Learning) มาประยุกต์ร่วมด้วยเช่นกัน

ขั้นตอนการประมวลผลภาพที่ได้นำเสนอมาทั้ง 7 ขั้นตอน ควรได้รับการพิจารณานำไปประยุกต์ตามความเหมาะสมในการใช้งาน โดยจำเป็นต้องวิเคราะห์กับคุณลักษณะของข้อมูลว่าควรดำเนินการด้วยกระบวนการใด トラบเท่าที่ได้ศึกษางานวิจัยที่ผ่านมาเกี่ยวกับการประมวลผลภาพเพื่อแก้ปัญหาต่าง ๆ พบว่า ข้อมูลชนิดเดียวกัน ลักษณะคล้ายคลึงกัน แต่ถ้าหากใช้เทคนิคและลำดับขั้นตอนการประยุกต์ที่ต่างกัน จะส่งผลให้ได้ผลลัพธ์ที่ต่างกันออกไป

ดังนั้น การศึกษาครั้งนี้จึงให้ความสำคัญในการแบ่งหมวดหมู่ชุดข้อมูลรูปภาพ เพื่อเปรียบเทียบคุณลักษณะพื้นฐานของภาพ ก่อนดำเนินการประยุกต์ใช้การเรียนรู้เครื่องจักร ทว่าในรายละเอียดการปรับปรุงและประมวลผลรูปภาพ จะกล่าวถึงในบทที่ 3

การทำความเข้าใจเกี่ยวกับทฤษฎีพื้นฐาน ทั้งกระบวนการทำงานของดาวเทียม และการประมวลผลภาพเบื้องต้น ทำให้ทราบถึงหลักการที่สามารถนำไปประยุกต์ใช้งานในหลากหลายรูปแบบ โดยเฉพาะกับงานวิจัยที่เกี่ยวข้องและได้นำมาศึกษา ทั้งเพื่อประยุกต์ใช้ทฤษฎีและวิเคราะห์ถึงผลลัพธ์ที่ได้จากกระบวนการวิจัย รวมถึงข้อดีและข้อเสียจากกระบวนการดำเนินการและการกำหนดตัวแปร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำหรับเนื้อหาลำดับถัดไป จะกล่าวถึงทฤษฎีและรายละเอียดเกี่ยวกับแบบจำลองปัญญาประดิษฐ์ที่ได้หยิบยกมาร่วมศึกษาในครั้งนี้

2.4 ทฤษฎีแบบจำลองปัญญาประดิษฐ์และโมเดลการเรียนรู้เชิงลึก

การเรียนรู้ของเครื่อง (Machine Learning) เป็นศาสตร์แขนงหนึ่งของปัญญาประดิษฐ์ หรือที่นิยมเรียกโดยย่อว่า “เอไอ (AI)” ซึ่งย่อมาจากศัพท์ภาษาอังกฤษคำว่า “Artificial Intelligence: AI” ทั้งนี้ การเรียนรู้เชิงลึก (Deep Learning) ก็เป็นศาสตร์แขนงหนึ่งที่อยู่ในขอบเขตของการเรียนรู้ของเครื่องเช่นกัน ฉะนั้น ในหัวข้อนี้จะกล่าวถึงทฤษฎีพื้นฐานที่เกี่ยวกับโมเดลการเรียนรู้เชิงลึก ซึ่งเกี่ยวข้องกับโมเดลที่ได้นำมาศึกษา วิจัย และวัดผลเปรียบเทียบประสิทธิภาพตามวัตถุประสงค์การวิจัย

อนึ่ง ความก้าวหน้าของโครงสร้างโมเดลการเรียนรู้เชิงลึกได้ถูกตีพิมพ์ และเผยแพร่ผลงานวิชาการเป็นจำนวนมาก เกี่ยวกับการนำมาประยุกต์ใช้เพื่อการแบ่งส่วนภาพ (Image Segmentation with Deep Learning) รวมถึงทฤษฎีที่เกี่ยวข้องกับการเรียนรู้ จดจำ และเข้าใจภาพของเครื่อง (Computer Vision)

2.4.1 การเรียนรู้ของเครื่อง (Machine Learning)

การเรียนรู้ของเครื่อง [25] มุ่งเน้นไปที่การสร้างองค์ความคิด การตัดสินใจ โดยเรียนรู้จากข้อมูลซึ่งถูกป้อนเข้าระบบ ซึ่งต่างจากการเขียนโปรแกรมแบบดั้งเดิม ที่ใช้การตัดสินใจตามกฎหรือเงื่อนไขที่กำหนดไว้ล่วงหน้า ดังนั้น การเรียนรู้ของเครื่องช่วยให้คอมพิวเตอร์สามารถตัดสินใจตามประสบการณ์ หรือข้อมูลที่ผ่านการประมวลผลก่อนหน้า โดยไม่ต้องตั้งโปรแกรมอย่างชัดเจนหรือเฉพาะเจาะจง อย่างไรก็ตาม ความสามารถผลลัพธ์ของการตัดสินใจและประสิทธิภาพการทำงานขึ้นอยู่กับข้อมูลที่ป้อนให้เครื่องเรียนรู้

ประเภทของการเรียนรู้ของเครื่องแบ่งออกเป็น 3 ประเภท โดยแต่ละประเภทขึ้นอยู่กับวิธีการเข้าถึงข้อมูล คุณลักษณะข้อมูล และการเรียนรู้ แนนอนว่าส่งผลถึงวัตถุประสงค์การนำไปใช้งาน

- 1) การเรียนรู้แบบมีผู้สอน (Supervised Learning) คือ การเรียนรู้ชุดข้อมูลแบบมีป้ายกำกับ (Label) ของคำตอบหรือผลลัพธ์ เพื่อฝึกฝนอัลกอริทึม (Algorithm) ในการจำแนกข้อมูลหรือทำนายผลลัพธ์
- 2) การเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning) คือ การเรียนรู้เพื่อวิเคราะห์และจัดกลุ่มชุดข้อมูลโดยไม่มีป้ายกำกับคำตอบ ซึ่งอัลกอริทึมสามารถเรียนรู้ได้และระบุความเหมือนหรือแตกต่างของข้อมูลได้
- 3) การเรียนรู้แบบลองผิดลองถูก (Reinforcement Learning) เป็นการเรียนรู้ที่มีลักษณะคล้ายกับการเรียนรู้แบบมีผู้สอน ทว่าอัลกอริทึมจะไม่ได้รับการฝึกฝนโดยใช้ข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

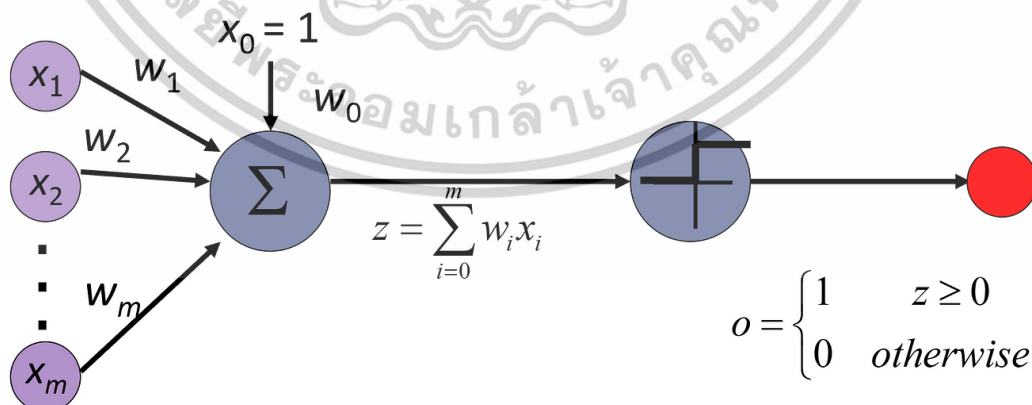
ตัวอย่าง หรืออาจกล่าวได้ว่า โมเดลเรียนรู้แบบลองผิดลองถูกโดยอิงจากการกระทำที่ตัดสินใจ ซึ่งจุดสำคัญในกระบวนการเรียนรู้คือหลักการ “การให้รางวัล (Reward)” ฉะนั้น เพื่อให้ประสิทธิภาพของการเรียนรู้ออกมาดีและตรงตามวัตถุประสงค์ที่ตั้งเป้า โมเดลจะพยายามตัดสินใจการกระทำ (Action) แบบลองผิดลองถูกเพื่อให้ได้ Reward มากที่สุดใน

2.4.2 การเรียนรู้เชิงลึก (Deep Learning)

การเรียนรู้เชิงลึก [26] เป็นส่วนหนึ่งของการเรียนรู้ของเครื่องที่ใช้โครงข่ายประสาทเทียม [27] (Neural Network) ซึ่งได้รับแนวคิดการสร้างอัลกอริทึมจากการจำลองการตัดสินใจที่ซับซ้อนของสมองมนุษย์ เพื่อระบุและจำแนกปัญหาแบบเฉพาะเจาะจง โดยรูปแบบโครงสร้างโครงข่ายประสาทเทียมจะมีผลต่อการทำนายและการตัดสินใจของโมเดล

ทั้งนี้ โครงข่ายประสาทเทียมนับเป็นหัวใจสำคัญของการเรียนรู้เชิงลึก อันเนื่องมาจากการทำงานร่วมกันของโหนด (Node) ในแต่ละเลเยอร์ (Layer) ที่เปรียบเสมือนเซลล์ประสาทเทียม อาทิ อินพุตเลเยอร์ (Input Layer) ฮิดเดนเลเยอร์ (Hidden Layer) เอาต์พุตเลเยอร์ (Output Layer) ซึ่งกระบวนการคำนวณยังประยุกต์เกณฑ์การชั่งน้ำหนักตัวเลือก (Weight Options) เพื่อให้ได้ข้อสรุปการทำนายจากการฝึกฝนหรือการเรียนรู้ที่ดีที่สุด โดยสมการตั้งต้นที่จำต้องกล่าวถึง คือ สมการถดถอยเชิงเส้น (Linear Regression) อันประยุกต์ใช้ฟังก์ชัน $f(x) = 1$ if $\sum w_1x_1 + b \geq 0$; 0 if $\sum w_1x_1 + b < 0$

อย่างไรก็ตาม เมื่อนำสมการมาประยุกต์ใช้ประกอบกับโครงสร้างแบบโครงข่ายประสาทเทียม ซึ่งรวมทั้งอินพุตเลเยอร์ตลอดจนเอาต์พุตเลเยอร์จึงได้รูปแบบตั้งต้นของโครงข่ายที่เรียกว่า เพอร์เซปตรอน (Perceptrons)



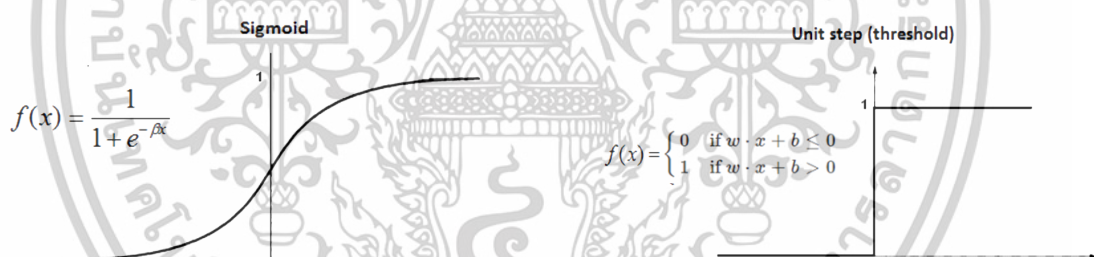
รูปที่ 2.18 ตัวอย่างโครงสร้างเพอร์เซปตรอน (A Perceptron Structure)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ฉะนั้น หากมีการกำหนด ลำดับ ปรับแต่งโครงสร้างเลเยอร์ในรูปแบบใด ๆ ย่อมส่งผลต่อประสิทธิภาพด้านความแม่นยำในการทำนายของโมเดลเพื่อแก้ปัญหา รวมถึงระยะเวลาที่ใช้ในการฝึกฝน

2.4.2.1 โครงข่ายประสาทเทียมแบบคอนโวลูชันนอล (Convolutional Neural Networks - CNNs)

ตามที่กล่าวเรื่องโครงสร้างเลเยอร์ของโครงข่ายประสาทเทียมซึ่งสามารถถูกต่อยอด และแบ่งประเภทได้หลายประเภทตามการวัตถุประสงค์เพื่อการใช้งาน ทว่าอย่างไรก็ตาม สำหรับการศึกษาระดับปริญญาโทและปริญญาเอกได้มุ่งเน้นไปที่โครงข่ายประสาทเทียมแบบคอนโวลูชันนอล [28] ซึ่งเหมาะสำหรับการจดจำรูปภาพ (Image Recognition) การจำจากรูปแบบ (Pattern Recognition) การมองเห็นและเข้าใจภาพของเครื่อง (Computer Vision) ทั้งนี้ แทนที่จะใช้เพอร์เซปตรอน (Perceptron) แบบดั้งเดิมในโครงข่ายประสาทเทียม โครงข่ายประสาทเทียมแบบคอนโวลูชันนอลได้ประยุกต์ใช้เซลล์ประสาทซิกมอยด์ (Sigmoid Neurons) เหตุอันเนื่องมาจากปัญหาในโลกแห่งความจริง บางผลลัพธ์ไม่อาจเป็นได้แค่ 0 หรือ 1 โดยเฉพาะอย่างยิ่งกับวัตถุประสงค์การใช้งานโครงข่ายประสาทเทียมแบบคอนโวลูชันนอล



รูปที่ 2.19 เปรียบเทียบค่าของผลลัพธ์ที่ได้จากสมการฟังก์ชันซิกมอยด์ (Sigmoid) และเพอร์เซปตรอนแบบดั้งเดิม (Perceptron, Unit step: threshold) ด้วยรูปแบบกราฟเส้น

2.4.2.2 ประเภทของเลเยอร์ (Types of Layers)

การวิจัยนี้มุ่งเน้นไปที่การใช้ประโยชน์จากโครงข่ายประสาทเทียมแบบคอนโวลูชันนอลเพื่อสกัดอาคารจากภาพถ่ายดาวเทียม หรือกล่าวคือ ต้องการให้โมเดลสามารถเรียนรู้และจดจำคุณลักษณะของอาคารจากภาพถ่ายดาวเทียมได้ ดังนั้นแล้ว ประเภทของเลเยอร์ อันเป็นทฤษฎีพื้นฐานเกี่ยวกับการตรวจจำวัตถุในรูปภาพของโมเดล U-Net และโมเดลอื่น ๆ ที่เกี่ยวข้องเป็นตามรายละเอียดเบื้องต้นดังต่อไปนี้

- 1) อินพุตเลเยอร์ (Input Layer) จุดเริ่มต้นของโครงข่ายประสาทเทียม สำหรับรูปภาพสี เลเยอร์นี้ยอมรับอินพุตเป็นรูปภาพที่มีสามช่อง

สัญญาณที่สอดคล้องกับสีหลัก: แดง เขียว และน้ำเงิน (RGB) การตั้งค่านี้นช่วย

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ให้โครงข่ายสามารถจัดการข้อมูลสีของรูปภาพ ที่จะถูกประมวลผลผ่านเลเยอร์ถัดไป ทว่าสำหรับชุดข้อมูลภาพที่มีช่องสัญญาณสีเพียงหนึ่ง หรือช่วงสีขาวดำ (Gray Scale) เลเยอร์อินพุตจะถูกกำหนดเป็น 1 แทน 3

2) คอนโวลูชันนอลเลเยอร์ (Convolutional Layer - Conv Layer) ถูกนำไปใช้เพื่อแยกคุณสมบัติออกจากอินพุต หรือกล่าวคือการดึงคุณลักษณะสำคัญออกจากภาพ (Feature Extraction) โดยกระบวนการตั้งค่าน้ำหนักไฮเพอร์พารามิเตอร์ที่สำคัญในเลเยอร์นี้ คือตัวกรองที่จะเลื่อนผ่านรูปภาพอินพุต เพื่อสร้างแผนที่คุณลักษณะ (Stride) หมายถึง จำนวนพิกเซลที่ฟิลเตอร์เคลื่อนผ่านอินพุตในแต่ละขั้นตอน และการแพดดิ้ง (Padding) เกี่ยวข้องกับการเพิ่มพิกเซลพิเศษรอบ ๆ รูปภาพอินพุตเพื่อให้ฟิลเตอร์ครอบคลุมบริเวณเส้นขอบได้อย่างสมบูรณ์

3) แอกทีเวชันฟังก์ชันเลเยอร์ (Activation Layer (ReLU - Rectified Linear Unit)) ฟังก์ชันการปรับค่าน้ำหนักที่นิยมใช้กันอย่างแพร่หลายในโครงข่ายประสาทเทียม เพื่อปรับปรุงค่าความน่าจะเป็นจากการคาดเดาที่ไม่เป็นเชิงเส้น ทำให้แบบจำลองสามารถเรียนรู้และทำนายรูปแบบข้อมูลที่ซับซ้อนมากขึ้น นอกจากนี้ ReLU ยังทำหน้าที่ช่วยแปลงอินพุตเชิงลบให้เป็นศูนย์ ในขณะที่ยังคงอินพุตเชิงบวกหรือมากกว่าศูนย์ไว้ไม่เปลี่ยนแปลงอย่างใดก็ได้ ค่าที่ได้จากฟังก์ชันจะอยู่ในช่วง 0 ถึง 1

4) พูลลิงเลเยอร์ (Pooling Layer (Max Pooling)) เลเยอร์การรวมกลุ่มเพื่อลดขนาดเชิงพื้นที่ (เช่น ความกว้างและความสูง) ของคุณลักษณะอินพุต นิยมใช้หลังจาก Conv Layer แล้ว ทั้งนี้ ประโยชน์เพื่อลดภาระในการคำนวณ และช่วยป้องกันไม่ให้เกิดปัญหาการเรียนรู้แบบโอเวอร์ฟิตติง (Overfitting) โดยในการวิจัยนี้เลือกใช้ Pooling Layer แบบ Max Pooling ซึ่งจะดำเนินการลดขนาดโดยการเลือกค่าสูงสุดจากขนาดของอินพุตที่กำหนดตามขนาด Kernel ที่เคลื่อนผ่าน

5) อัปแซมพลิงเลเยอร์ (Up-sampling Layer) การทำงานซึ่งตรงกันข้ามกับ Pooling Layer หรือกล่าวเป็นกระบวนการที่ใช้ขยายคุณลักษณะเชิงพื้นที่แทนที่จะลดลง โดยในการศึกษานี้ใช้ Transposed Convolution เพื่อเพิ่มขนาดคุณลักษณะของอาคารจากภาพถ่ายดาวเทียม หรืออาจกล่าวได้ว่าเป็นการขยายคุณลักษณะให้เด่นชัดเพื่อเพิ่มความแม่นยำอย่างเจาะจง

6) คอนแคตเทนชันเลเยอร์ (Concatenation Layer) คือเลเยอร์ที่ทำหน้าที่รวมคุณลักษณะจากทั้งสองเลเยอร์เข้าด้วยกัน เปรียบเสมือนการต่อชุดข้อมูลเพื่อรวมคุณสมบัติจากการประมวลผล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตามรายละเอียดพอสังเขปเกี่ยวกับการเรียนรู้เชิงลึกและโครงข่ายประสาทเทียม ซึ่งให้เห็นว่าการพัฒนาโครงสร้างของโครงข่ายหรือที่นิยมเรียกว่า สถาปัตยกรรมโมเดล โดยเริ่มจากศูนย์ เพื่อให้บรรลุตามวัตถุประสงค์ตามเป้าหมายงานวิจัยเป็นไปได้ยาก ดังนั้น การศึกษาจากสถาปัตยกรรมหรือโมเดลที่ได้รับการตีพิมพ์ผลงานวิชาการอยู่แล้วจึงเป็นทางเลือกที่เหมาะสม โดยเนื้อหาในส่วนถัดไปจะกล่าวถึงการถ่ายโอนองค์ความรู้ (Transfer Learning) ซึ่งอาศัยสถาปัตยกรรมที่มีอยู่แล้วประกอบกับการเรียนรู้ล่วงหน้า (Pretrained) ด้วยชุดข้อมูล ImageNet จากนั้น จะกล่าวถึงโมเดลที่เลือกตามวัตถุประสงค์งานวิจัย

2.4.3 การถ่ายโอนความรู้ (Transfer Learning)

การถ่ายโอนความรู้หรือการเรียนรู้การถ่ายโอน [29] เป็นเทคนิคของการเรียนรู้ของเครื่องรูปแบบหนึ่ง โดยหลักการแล้วคือการนำโมเดลที่ได้รับการฝึกฝนล่วงหน้าด้วยวัตถุประสงค์งานใดงานหนึ่ง มาปรับแต่งรายละเอียดเพื่อการใช้งานในวัตถุประสงค์ใหม่ที่ใกล้เคียงกัน ตัวอย่างเช่น มีโมเดลที่สามารถระบุภาพสุนัขได้ นำมาฝึกฝนเพิ่มเพื่อให้ระบุแมวได้ ทั้งนี้ใช้ชุดข้อมูลภาพแมวที่มีจำนวนน้อยกว่า และมีความแตกต่างกันของคุณสมบัติระหว่างสุนัขกับแมว ดังนั้น เทคนิคดังกล่าวจึงเป็นกระบวนการที่ช่วยลดการใช้ข้อมูลปริมาณมากในการฝึกฝน ลดระยะเวลาและทรัพยากรการคำนวณในการประมวลผล

นอกจากนี้ หากพิจารณาตามขอบเขตการวิจัยเรื่องปริมาณข้อมูลในประเทศไทย การประยุกต์ใช้เทคนิคการถ่ายโอนความรู้ นับเป็นประโยชน์อย่างมากต่อการพัฒนาโมเดลใหม่ที่มีต้นทุนจำกัดด้านปริมาณข้อมูล ทั้งนี้ เมื่อกล่าวถึงเทคนิคการถ่ายโอนความรู้จึงเป็นเรื่องที่เกี่ยวข้องไม่ได้ที่ต้องทำความเข้าใจเกี่ยวกับ การนำโครงข่ายประสาทเทียมที่มีอยู่แล้วมาประยุกต์ใช้ร่วมกับโครงสร้างในโมเดลใหม่ หรือที่เรียกกันว่าการใช้ แแบ็กโบน (Backbone) รวมถึงแบ็กโบนบางตัวมีคุณสมบัติผ่านการฝึกฝนจากชุดข้อมูลใดชุดข้อมูลหนึ่งมาก่อนแล้ว จะเรียกสถานะดังกล่าวว่า 프리เทรน (Pretrained)

ดังนั้นแล้ว การวิจัยนี้ได้นำเทคนิคการถ่ายโอนความรู้รูปแบบ Inductive Transfer ของโมเดล Residual Network (ResNet) ประเภท 50 101 และ 152 เลเยอร์ มาประยุกต์ใช้เป็นแบ็กโบนในโมเดล Res-U-Net และ DeepLabV3+ โดย ResNet ทั้ง 3 ประเภทดังกล่าวผ่านการฝึกฝนในโครงสร้างเลเยอร์ที่ต่างกันด้วยชุดข้อมูล ImageNet เหมือนกัน (Pretrained with ImageNet) ทว่าวัตถุประสงค์เดิมคือใช้เพื่อจำแนกรูปภาพ (Image Classification) แต่การวิจัยนี้นำมาประยุกต์ด้วยวัตถุประสงค์การแบ่งส่วนภาพ (Image Segmentation) [30]

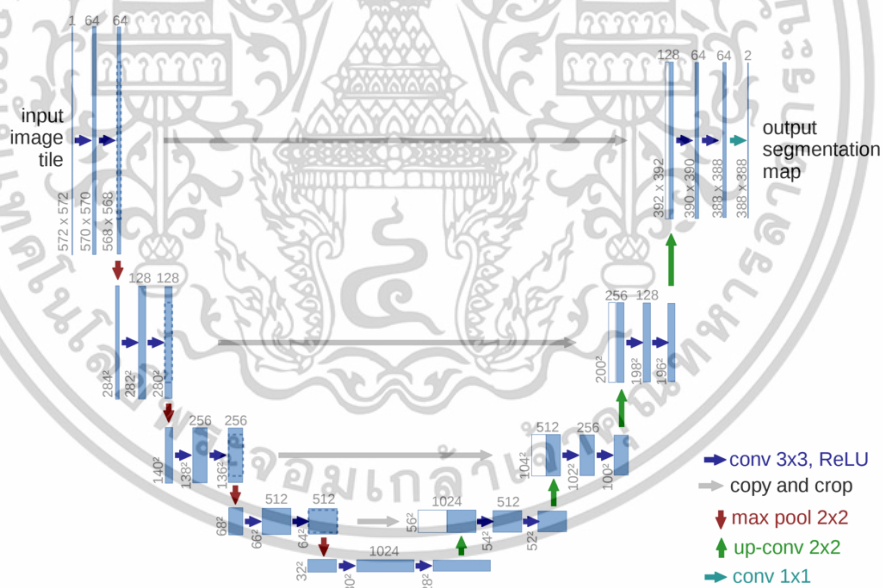
ฉะนั้น การวิจัยนี้จึงออกแบบการทดลองเพื่อเปรียบเทียบศักยภาพของโครงสร้างแบบดั้งเดิมของ U-Net ตลอดจนโครงสร้างปรับปรุงแบบ RIU-Net และ U-Net++ ในการสกัดอาคารจากภาพถ่ายดาวเทียม กระนั้น เพื่อวิเคราะห์ความแตกต่างเรื่องประสิทธิภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

นอกจากแค่การเปลี่ยนแปลงโครงสร้างโครงข่ายประสาทเทียม จึงได้ประยุกต์ใช้เทคนิคการถ่ายโอนความรู้ ซึ่งทำให้เกิดการประยุกต์ใช้แบ็กโบนที่ผ่านการฝึกฝน ResNet มาเป็นส่วนหนึ่งในโครงสร้าง Res-U-Net ทั้งนี้ สำหรับโมเดล DeepLabV3+ ก็มีส่วนของโครงสร้างที่คล้ายกับโครงสร้าง U-Net ในทางทฤษฎี ซึ่งจะกล่าวถึงในลำดับถัดไป

2.4.4 โมเดลการเรียนรู้เชิงลึกยูเน็ต (U-Net)

U-Net เป็นสถาปัตยกรรมโครงข่ายประสาทเทียมแบบคอนโวลูชันนอล (CNNs) ที่นำมาใช้สำหรับการแบ่งส่วนภาพทางการแพทย์ (Biomedical Image Segmentation) โดยเฉพาะการแบ่งส่วนโครงสร้างเซลล์ประสาทในเนื้อเยื่อสมอง [31] ซึ่งโครงสร้างหลักประกอบด้วยเส้นทางตัวเข้ารหัส (Encoder; Downsampling) สำหรับการรวบรวมข้อมูลคุณลักษณะ และเส้นทางตัวถอดรหัสสำหรับการแปลที่แม่นยำ (Decoder; Upsampling) นอกจากนี้ ยังมีการทำงานประสานกันระหว่างฝั่งเข้ารหัสและถอดรหัส ด้วยการเชื่อมต่อกันคุณลักษณะ (Concatenation) ซึ่งมีส่วนช่วยให้สามารถผสมผสานคุณลักษณะระดับต่ำและระดับสูง โครงสร้างดังรูปที่ 2.20



รูปที่ 2.20 ตัวอย่างโครงสร้างสถาปัตยกรรมโมเดล U-Net

จากการออกแบบ สถาปัตยกรรมตัวเข้ารหัสและถอดรหัสของ U-Net ได้ประยุกต์ใช้ Conv Layer แบบสลับกับ Pooling Layer ในฝั่งเข้ารหัสเพื่อสกัดเอาคุณลักษณะแบบต่ำ (การสกัดเบื้องต้น) จากนั้น ดำเนินการประยุกต์ใช้ Up-Conv Layer เพื่อขยายคุณลักษณะให้มีความละเอียดชัดเจนยิ่งขึ้นสลับกับการทำ Conv Layer นอกจากนี้ ในส่วนของการ

ดำเนินการ Concatenation (copy and crop) คือการผสมผสานคุณลักษณะตามที่กล่าว

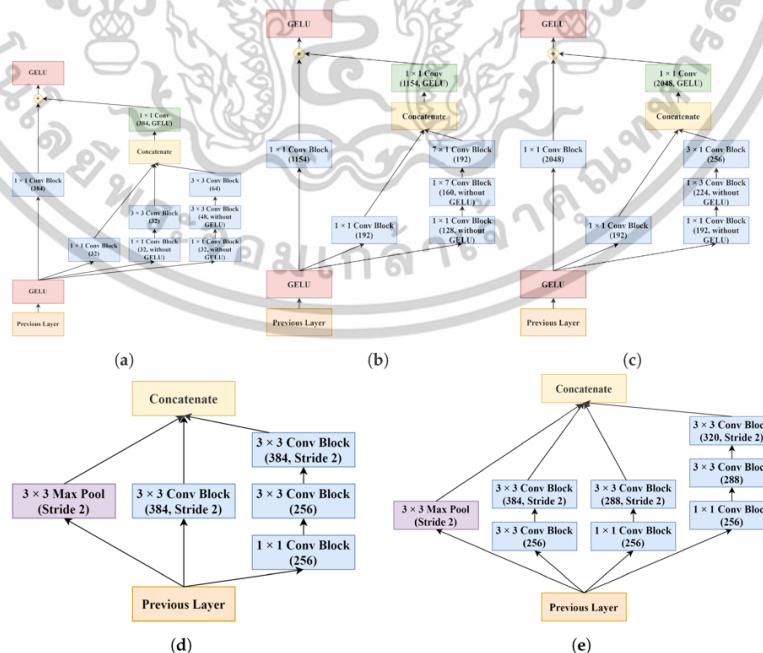
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น เมื่ออนุญาตให้เผยแพร่เป็นประโยชน์ทางการศึกษา

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

อย่างไรก็ดี เพื่อจัดการกับผลลัพธ์การปรับค่าน้ำหนักและความเอนเอียงในโครงข่าย การเลือกใช้ Activation Function ReLu จึงช่วยขจัดปัญหาผลลัพธ์ที่ไม่ใช่เชิงเส้น 0 หรือ 1 นอกจากนี้ แนวทางของ U-Net ยังเอื้อประโยชน์ต่อการรองรับข้อมูลที่มีขนาดไม่มากนักเพื่อการฝึกฝน อย่างไรก็ตาม การเสริมชุดข้อมูลจะกล่าวถึงภายหลังในส่วนของการดำเนินงานวิจัยแน่นอนว่าขนาดของชุดข้อมูลย่อมเป็นความท้าทายต่อการฝึกฝนโมเดลให้เกิดความแม่นยำ แต่กระนั้น ด้วยการออกแบบโครงสร้างให้รองรับกับภาพถ่ายทางการแพทย์ซึ่งอาจไม่มีชุดข้อมูลภาพที่มากนักทำให้เป็นผลพลอยได้ในการนำมาทดลองวิจัยในภาพถ่ายดาวเทียมที่จำต้องเก็บชุดข้อมูลด้วยตนเองให้เกิดความก้าวหน้าในการใช้งานทั้งทางการแพทย์และไม่ใช้ทางการแพทย์

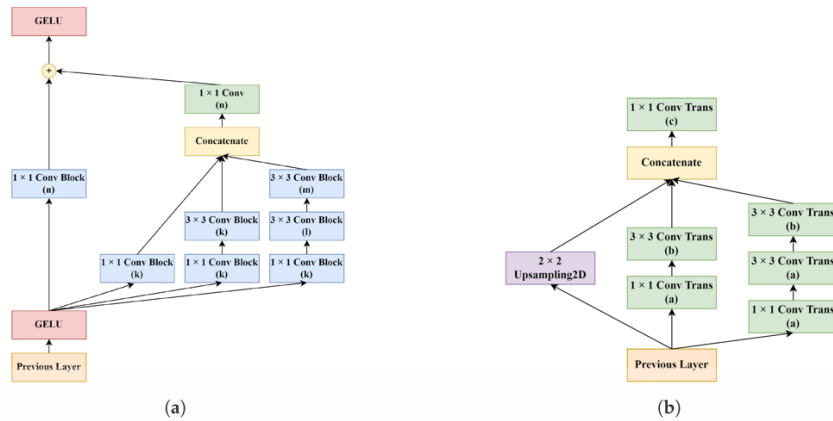
2.4.5 โมเดลการเรียนรู้เชิงลึกอาร์ไอยูเน็ต (RIU-Net)

RIU-Net หรือ Residual Inception U-Net เป็นการปรับขั้นสูงของสถาปัตยกรรม U-Net โดยการประยุกต์ใช้โมดูล Residual และ Inception อันทำหน้าที่เพิ่มประสิทธิภาพการแยกคุณลักษณะ [32, 33] โดยโมดูลที่กล่าวมีส่วนช่วยในการจัดการ Gradient Propagation ระหว่างการฝึกฝน เพื่อช่วยให้การส่งถ่ายข้อมูลคุณลักษณะมีประสิทธิภาพมากยิ่งขึ้น [34] ขณะเดียวกันที่โมดูล Inception ได้รับการออกแบบให้มีความยืดหยุ่นในการตรวจจับคุณลักษณะที่หลากหลายขนาดในภาพ ทั้งนี้ การประยุกต์ใช้โมดูล Residual และ Inception ได้ถูกประยุกต์เข้ากับฝั่งตัวเข้ารหัสและตัวถอดรหัสของ U-Net แบบดั้งเดิม ดังรูปที่ 2.21

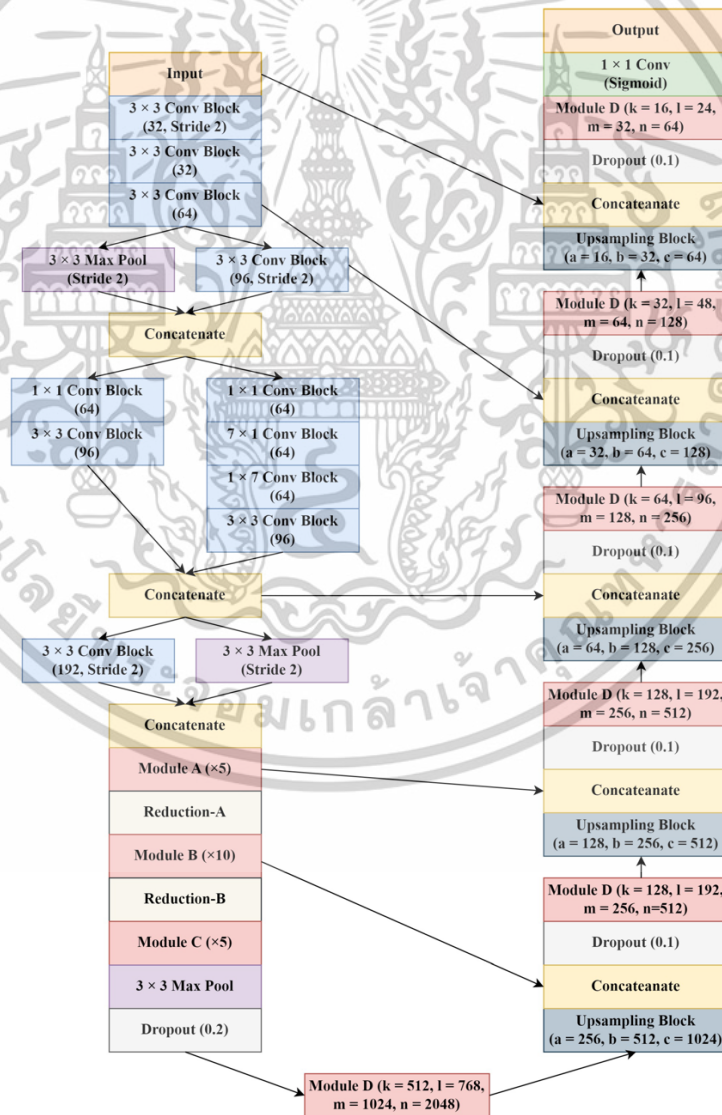


รูปที่ 2.21 โครงสร้างโครงข่ายในส่วนของการเข้ารหัส (Encoder) ของ RIU-Net: (a) Module A, (b) Module B, (c) Module C, (d) Reduction A, and (e) Reduction B

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.22 โครงสร้างโครงข่ายในส่วนของการถอดรหัส (Decoder) ของ RIU-Net: (a) Module D, and (b) Upsampling module.

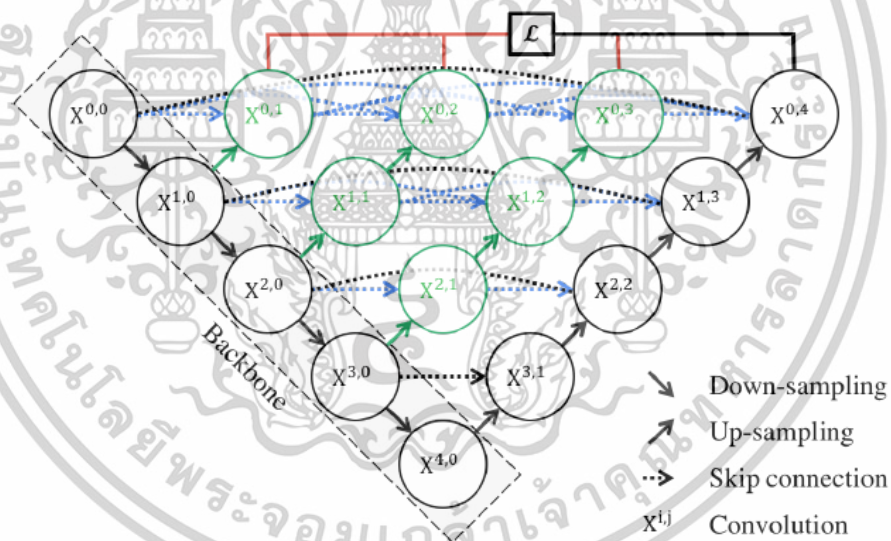


รูปที่ 2.23 ตัวอย่างโครงสร้างสถาปัตยกรรมโมเดล RIU-Net (Encoder/Decoder อิงตามรูปที่ 2.21 และ 2.22 ตามลำดับ)
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.4.6 โมเดลการเรียนรู้เชิงลึกยูเน็ตพลัสพลัส (U-Net++)

U-Net++ ขยายสถาปัตยกรรม U-Net ด้วยการเชื่อมต่อแบบ Nested and Dense Skip Connections มาประยุกต์ใช้ดังรูปที่ 2.24 เพื่อตรวจจับภาพคุณสมบัติที่หลากหลาย และปรับปรุงการถ่ายโอนข้อมูลภายในโครงข่าย [35] ผลจากการเชื่อมต่อเหล่านี้ช่วยให้ U-Net++ สามารถใช้ประโยชน์จากคุณลักษณะแบบลำดับขั้นได้ (Hierarchical Features) ซึ่งถือเป็นแนวคิดอย่างหนึ่งที่ต่อยอดการเชื่อมต่อข้อมูลคุณลักษณะ และอ้างอิงจากผลการวิจัย โครงสร้างลักษณะดังกล่าวมีส่วนช่วยในการตรวจจับคุณลักษณะรูปภาพในหลายบริบท

การแยกส่วนการออกแบบของ U-Net++ เพิ่มเติมเผยให้เห็นสถาปัตยกรรมที่เรียกว่า Skip Pathway ซึ่งแตกต่างจาก U-Net แบบดั้งเดิมที่เพียงเชื่อมต่อคุณลักษณะในระดับเลเยอร์เดียวกัน ทว่าด้วยการออกแบบของ Skip Pathway ทำให้ U-Net++ สามารถเชื่อมต่อคุณลักษณะแบบข้ามเลเยอร์ถึงกันได้ และเอื้อให้เกิดการผสมผสานคุณลักษณะระดับต่ำเข้ากับคุณลักษณะระดับสูง เหตุอันนี้ด้วยความคาดหวังว่ากระบวนการดังกล่าวจะปรับปรุงความแม่นยำในการแบ่งส่วนภาพ



รูปที่ 2.24 ตัวอย่างโครงสร้างสถาปัตยกรรม U-Net++

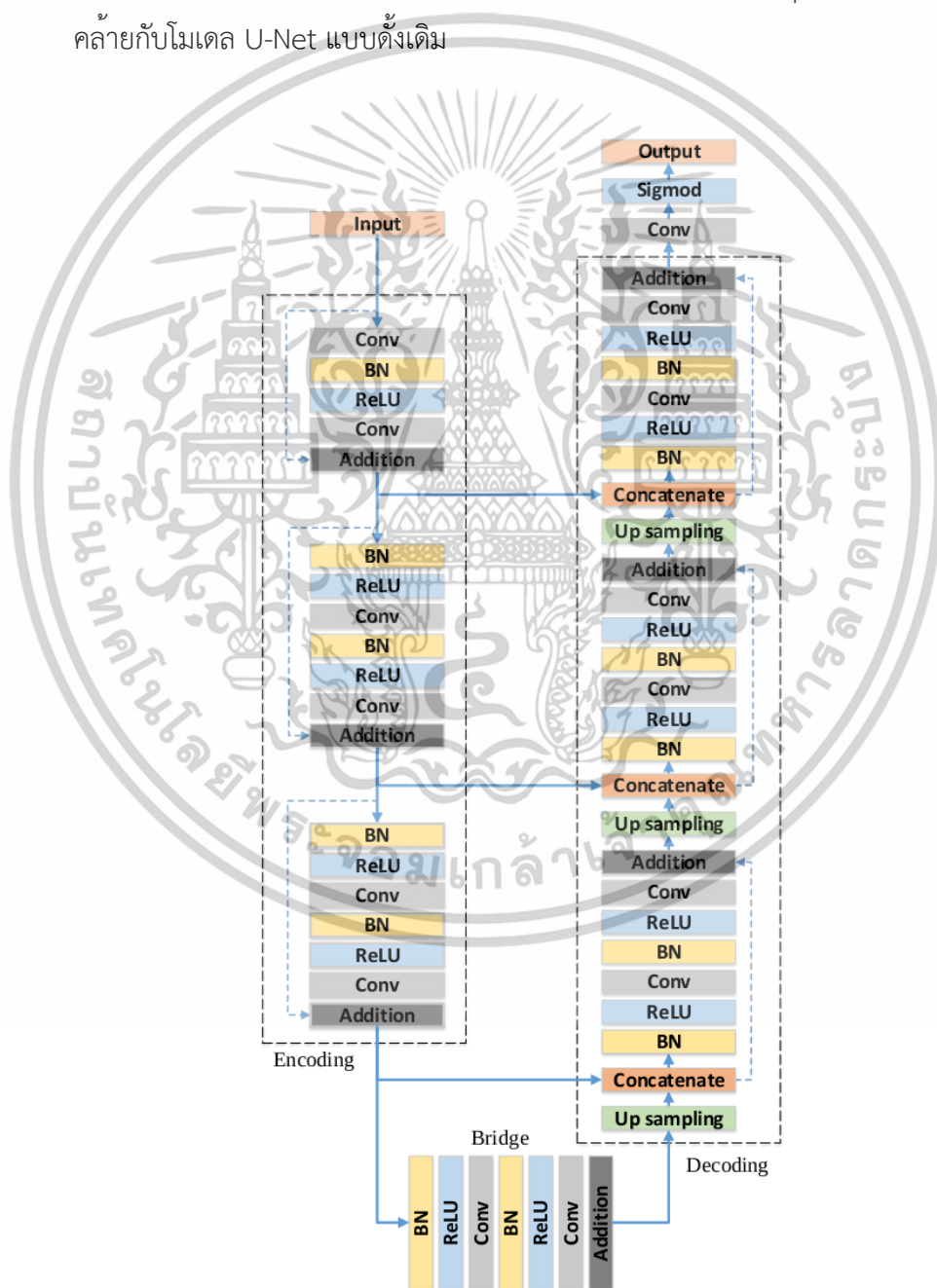
2.4.7 โมเดลการเรียนรู้เชิงลึกเรซูเน็ต (Res-U-Net)

Res-U-Net เป็นการผสมผสานของสถาปัตยกรรมพื้นฐานของ U-Net กับโครงข่ายแกนหลักแบ็กโบน (Backbone) ที่ได้รับการฝึกฝนล่วงหน้า (Pretrained) อย่าง ResNet (ใช้ในกรณีศึกษาวิจัยนี้) ในฝั่งตัวเข้ารหัส (Encoder; Addition Layer) [36] การประยุกต์ใช้เทคนิคการถ่ายโอนความรู้นี้ เพื่อปรับปรุงประสิทธิภาพของการสกัดคุณลักษณะในการตรวจจับทั้งระดับต่ำและระดับสูง อีกทั้งยังมีความยืดหยุ่นในการปรับแต่ง Pretrained weight (การวิจัยนี้ใช้ pretrained weight จาก ImageNet) และการปรับแต่ง Backbone

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้เพื่อการวิจัยเท่านั้น เมื่อผู้เผยแพร่เห็นใบเซอร์ไอเซนด้าเนกการศา
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีกรนำไปใช้

ตัวต่าง ๆ อาทิ ResNet-50 ResNet-101 และ ResNet-152 ซึ่งมีส่วนใช้ชีวิตความแตกต่างถึงประสิทธิภาพในการสกัดคุณลักษณะ

ในด้านสถาปัตยกรรมของ Res-U-Net การนำ ResNet มาเป็น Backbone ของตัวเข้ารหัสถือเป็นหัวใจสำคัญที่สร้างความแตกต่างจากโครงสร้าง U-Net RIU-Net และ U-Net++ หรือกล่าวคือแนวคิดที่ประยุกต์ใช้ความสามารถจากโครงข่ายซึ่งได้รับการฝึกฝนมาแล้วร่วมกับการฝึกฝนด้วยชุดข้อมูลใหม่ ทำให้เป็นเรื่องที่น่าสนใจในการเปรียบเทียบและวัดผลประสิทธิภาพ กระนั้นตามทฤษฎีแล้ว การปรับแต่งโครงสร้างดังกล่าวควรช่วยเพิ่มความแม่นยำและประสิทธิภาพ อย่างไรก็ตาม สำหรับกระบวนการเชื่อมต่อคุณลักษณะยังมีแนวทางที่คล้ายกับโมเดล U-Net แบบดั้งเดิม



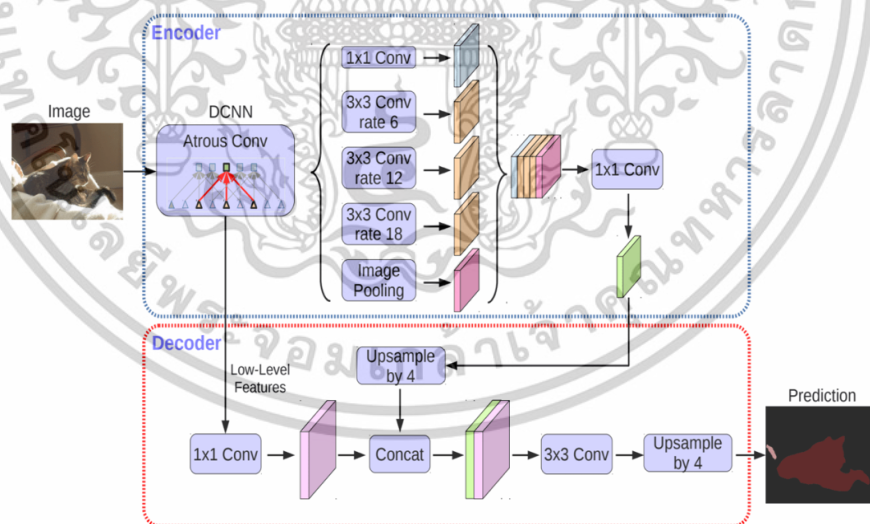
รูปที่ 2.25 ตัวอย่างโครงสร้างสถาปัตยกรรม Res-U-Net

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.4.8 โมเดลการเรียนรู้เชิงลึกดีฟล์แอนด์บิวิทีฟลัส (DeepLabV3+)

DeepLabV3+ คือโมเดลการเรียนรู้เชิงลึกสำหรับงานการแบ่งส่วนภาพ โดยอาศัยหลักสกัดคุณลักษณะที่เรียกว่าการบิดแบบขยาย (Atrous Convolution) เพื่อรวบรวมข้อมูลบริบทคุณลักษณะในหลายระดับของรายละเอียดภาพ [37, 38] นอกจากนี้ DeepLabV3+ ยังใช้โมดูลกลุ่มพีระมิดเชิงพื้นที่ (Spatial Pyramid Pooling Module) เพื่อรวบรวมคุณสมบัติหลายขนาดได้อย่างมีประสิทธิภาพ

การศึกษาศาปตยกรรม DeepLabV3+ เผยให้เห็นว่าการประยุกต์ใช้ Atrous Convolution และ Spatial Pyramid Pooling ดังรูปที่ 2.26 เป็นสิ่งสำคัญในการรวบรวมข้อมูลคุณลักษณะเชิงพื้นที่ของรูปภาพอย่างมีประสิทธิภาพ โดย Atrous Convolution มีส่วนช่วยให้โมเดลสามารถสกัดคุณลักษณะที่สำคัญได้อย่างถูกต้อง และแม่นยำในรายละเอียดเล็กน้อยบนภาพ การดำเนินการเปรียบคล้ายกับการใช้แว่นขยายส่องหารายละเอียด ส่วน Spatial Pyramid Pooling เป็นการประยุกต์ใช้ Pooling Layer ในหลากหลายความละเอียดเพื่อเป็นการรวมและดึงเอาคุณลักษณะออกจากภาพ จากนั้นนำมารวมด้วย Concatenation Layer ทั้งนี้ วิธีการดังกล่าวมีลักษณะคล้ายกับปิรามิดหลายขนาดที่ซ้อนทับกัน ลำดับแต่ละชั้นสะท้อนถึงรายละเอียดของการดึงเอาคุณลักษณะของภาพ ดังนั้น แนวคิดดังกล่าวจึงสนับสนุนให้เห็นถึงการสกัดเอาข้อมูลคุณลักษณะของภาพในหลากหลายรายละเอียด

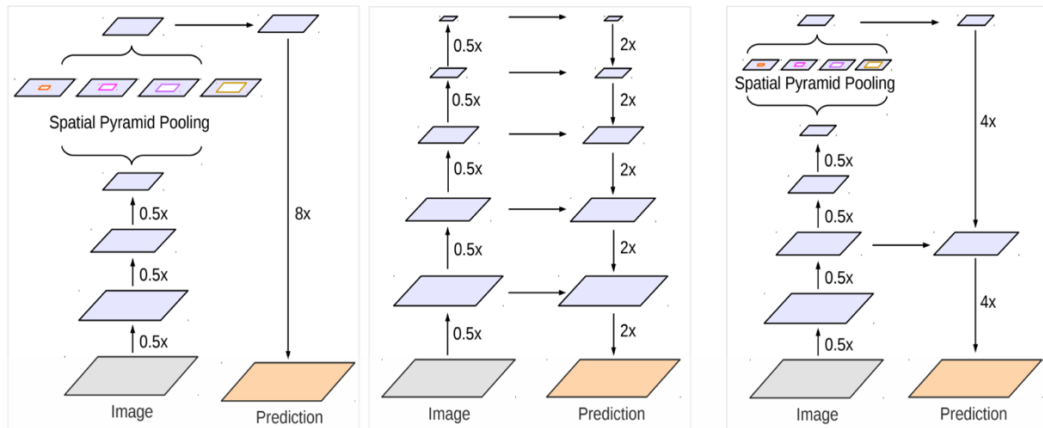


รูปที่ 2.26 ตัวอย่างโครงสร้างสถาปัตยกรรม DeepLabV3+

ด้วยโครงสร้างสถาปัตยกรรมตัวเข้ารหัสและถอดรหัส ที่ประยุกต์ใช้กระบวนการสกัดคุณลักษณะดังกล่าว ทำให้โมเดล DeepLabV3+ เป็นตัวเลือกที่น่าสนใจในการนำมาเปรียบเทียบและวัดผลโครงสร้างที่เกี่ยวกับการสกัดอาคารจากภาพถ่ายดาวเทียม อย่างไรก็ตาม ด้วยแนวคิดตัวเข้ารหัสและถอดรหัสที่มีความคล้ายกับโมเดล U-Net ด้วยเช่นกัน ดังรูปที่ 2.27

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้ในเพื่อการศึกษาค้นคว้าเท่านั้น เมื่ออนุญาตให้เผยแพร่ไปยังเว็บไซต์อื่น การค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(a) Spatial Pyramid Pooling

(b) Encoder-Decoder

(c) Encoder-Decoder with Atrous Conv

รูปที่ 2.27 เปรียบเทียบตัวอย่างโครงสร้างสถาปัตยกรรมในมุมมองตัวเข้ารหัสและถอดรหัส โดยภาพ a) คือโครงสร้างสถาปัตยกรรมบนแนวคิด Spatial Pyramid Pooling ของ DeepLabV3 b) คือโครงสร้างสถาปัตยกรรมบนแนวคิด Encoder-Decoder ของ U-Net และ c) คือโครงสร้างสถาปัตยกรรมบนแนวคิด a) และ b) ของ DeepLabV3+

จากการศึกษาทฤษฎีที่เกี่ยวข้องกับโมเดลการเรียนรู้เชิงลึกตามที่กล่าวมา ซึ่งให้เห็นว่าโครงสร้างสถาปัตยกรรมของโมเดล U-Net, RIU-Net, U-Net++, Res-U-Net และ DeepLabV3+ มีประสิทธิภาพในการแบ่งส่วนภาพ ซึ่งเป็นประโยชน์และคุ้มค่าแก่การนำมาวิเคราะห์เพื่อเปรียบเทียบประสิทธิผล อย่างไรก็ตาม เพื่อเป็นการชี้วัดความสามารถของแต่ละโมเดล ดังนั้น เนื้อหาในส่วนถัดไปจะกล่าวถึงดัชนีชี้วัดความคล้ายคลึงของภาพ โดยวัตถุประสงค์คือการวัดความคล้ายคลึงระหว่างภาพคำตอบที่ได้ทำการระบุอาคารไว้แล้ว กับภาพผลลัพธ์ที่ได้จากการทำนายจากแต่ละโมเดล

2.5 ดัชนีชี้วัดความคล้ายคลึงระหว่างรูปภาพ

การประเมินประสิทธิภาพการแบ่งส่วนภาพของโมเดลจำเป็นต้องใช้เกณฑ์ชี้วัด ที่ชี้ให้เห็นถึงความแตกต่างระหว่างคำตอบและผลทำนาย หรือกล่าวได้ว่าเป็นการเปรียบเทียบความคล้ายคลึงของรูปภาพสองรูปว่ามีส่วนที่เหมือนกันมากน้อยเพียงใดผ่านดัชนีชี้วัด แน่แน่นอนว่าหากดัชนีชี้วัดให้คะแนนความคล้ายคลึงสูงย่อมสะท้อนถึงประสิทธิภาพของโมเดลที่ทำนายผลได้อย่างแม่นยำ สำหรับการวิจัยนี้ได้หยิบยกดัชนีชี้วัดพื้นฐาน 3 ดัชนีมาดำเนินการชี้วัดความคล้ายคลึง ได้แก่ เมทริกซ์ชี้วัดความสับสน (Confusion Matrix) ดัชนีความคล้ายคลึง Jaccard Index (Intersection over Union; IoU) และดัชนีค่าสัมประสิทธิ์ Dice Similarity Coefficient (DSC) โดยดัชนีชี้วัดทั้งสามจะถูกนำมาใช้ประเมินผลลัพธ์การทำนายของโมเดล นอกจากนี้ ดัชนีชี้วัด IoU และ DSC ได้ถูกนำมาใช้วัดผลระหว่างการฝึกฝนด้วยชุดข้อมูลฝึกฝนของแต่ละโมเดลควบคู่กัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.5.1 เมทริกซ์ชี้วัดความสับสน (Confusion Matrix)

หนึ่งในเมทริกซ์พื้นฐานที่นิยมใช้วัดประสิทธิภาพการทำนายของโมเดล โดยหลักการคำนวณของเมทริกซ์ชี้วัดความสับสน [39] ได้กำหนดเงื่อนไขการวิเคราะห์ไว้ดังนี้ (ประยุกต์ใช้และอ้างอิงถึงความหมายที่เกี่ยวข้องกับบริบทงานวิจัยนี้)

- True Positive (TP) จำนวนครั้งของผลทำนายพื้นที่ในภาพเป็นอาคาร และสอดคล้องกับคำตอบในความเป็นจริงว่าพื้นที่ดังกล่าวเป็นอาคาร (Building)
- True Negative (TN) จำนวนครั้งของผลทำนายพื้นที่ในภาพไม่ใช่อาคาร และสอดคล้องกับคำตอบในความเป็นจริงว่าพื้นที่ดังกล่าวไม่ใช่อาคาร (non-Building)
- False Positive (FP) จำนวนครั้งของผลทำนายพื้นที่ในภาพเป็นอาคาร และไม่สอดคล้องกับคำตอบในความเป็นจริงที่ว่าพื้นที่ดังกล่าวไม่ใช่อาคาร (ทำนายว่าเป็นอาคารแต่ความเป็นจริงไม่ใช่)
- False Negative (FN) จำนวนครั้งของผลทำนายพื้นที่ในภาพไม่ใช่อาคาร และไม่สอดคล้องกับคำตอบในความเป็นจริงที่ว่าพื้นที่ดังกล่าวเป็นอาคาร (ทำนายว่าไม่ใช่อาคารแต่ความเป็นจริงใช่)

จากเงื่อนไขข้างต้นทั้ง 4 เงื่อนไข นำไปสู่สมการชี้วัดได้ดังต่อไปนี้

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2.1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2.2)$$

$$Recall = \frac{TP}{TP + FN} \quad (2.3)$$

$$F1\ Score = \frac{2 * TP}{2 * TP + FP + FN} \quad (2.4)$$

สำหรับความหมายของสมการ (2.1) ถึง (2.4) มีรายละเอียดดังต่อไปนี้

2.5.1.1 ค่าความแม่นยำ (Accuracy)

การประเมินค่าความแม่นยำ โดยสนใจในบริบทที่ว่าโมเดลทำนายถูกมากน้อยเพียงใด โดยจะพิจารณาส่วนที่ทำนายถูก ทั้งใช่อาคารและไม่ใช่อาคาร ซึ่งค่า TP และ TN สะท้อนถึงการทำนายที่ตรงกับความเป็นจริง อย่างไรก็ตาม ในกรณีที่โมเดลสามารถทำนายส่วนใช่อาคารและส่วนที่ไม่ใช่อาคารได้มาก ย่อมส่งผลให้ค่าความแม่นยำสูง หรือกล่าวได้ว่า โมเดลมีความแม่นยำในการระบุทั้งส่วนที่ใช่และไม่ใช่อาคาร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.5.1.2 ค่าความแม่นยำในการทำนายเชิงบวก (Precision)

การประเมินค่าความแม่นยำเฉพาะการทำนายที่ให้ผลเป็น Positive หรือกล่าวได้ว่าโมเดลมีความสามารถในการทำนายอาคารถูกต้องมากน้อยเพียงใด โดยพิจารณาจากค่า TP และ FP หากโมเดลไม่มีการทำนายอาคารในส่วนที่ไม่ใช่อาคาร ย่อมหมายความว่าโมเดลมีความแม่นยำในการทำนายอาคารสูง

2.5.1.3 ค่าความแม่นยำในการตรวจจับ (Recall)

การประเมินค่าความแม่นยำในการทำนายตามความเป็นจริง หรือมักเรียกว่าค่าความแม่นยำในการตรวจจับ กล่าวคือ สนใจความเป็นจริงในบริบท Positive เทียบกับการทำนายของโมเดลว่าทำนายตรงกับความเป็นจริงมากน้อยเพียงใด ดังนั้น ในกรณีนี้สามารถใช้เป็นตัวชี้วัดประสิทธิภาพการตรวจจับอาคารของโมเดลได้ โดยพิจารณาจากค่า TP และ FN หากโมเดลทำนายว่าอาคารเป็นอาคาร (TP) มากกว่าไม่ใช่อาคารทั้งที่เป็นอาคาร (FN) ย่อมหมายความว่าโมเดลมีความสามารถในการตรวจจับอาคารจากภาพได้ดี

2.5.1.4 คะแนน F1 (F1 Score)

ค่าประเมินความสมดุลระหว่าง Precision และ Recall หรืออีกชื่อหนึ่งคือ ค่าเฉลี่ยแบบ Harmonic Mean ใช้วัดผลความสามารถของโมเดลอย่างเป็นกลาง ในเรื่องความแม่นยำการทำนายและความสามารถในการตรวจจับ

2.5.2 Jaccard Index (Intersection Over Union; IoU)

Jaccard Index หรือที่เรียก IoU เป็นตัวชี้วัดการประเมินที่ใช้กันทั่วไปในงานการแบ่งส่วนภาพและงานตรวจจับวัตถุภายในภาพ ถูกเสนอขึ้นครั้งแรกโดย P. Jaccard ในปี 1901 [40] และตั้งแต่นั้นเป็นต้นมา ดัชนีชี้วัดนี้ได้กลายเป็นมาตรฐานสำหรับการประเมินความแม่นยำของอัลกอริทึมการแบ่งภาพ รวมไปถึงโมเดลการเรียนรู้เชิงลึกที่ถูกนำมาใช้ด้วยวัตถุประสงค์เดียวกัน IoU พิจารณาจากการทับซ้อนกันระหว่างภาพการทำนาย (Predict Image) และภาพคำตอบจริงที่ทำป้ายกำกับไว้ (Ground truth) โดยการคำนวณอัตราส่วนของจุดซ้อนทับเข้าด้วยกันแบบทั้งหมด ดังนั้น IoU จึงสามารถวัดผลเป็นคะแนนเชิงปริมาณได้อย่างแม่นยำได้ สูตรสำหรับดัชนี IoU เป็นดังสมการที่ (2.5)

$$IoU = \frac{TP}{TP + FP + FN} \quad (2.5)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.5.3 Dice Similarity Coefficient (DSC)

ค่าสัมประสิทธิ์ Dice Similarity Coefficient (DSC) เป็นตัวชี้วัดความคล้ายคลึงกันที่ใช้กันอย่างแพร่หลายสำหรับการประเมินแบบโมเดลแบ่งส่วนภาพ เช่นเดียวกับดัชนีชี้วัด IoU ทั้งนี้ DSC ถูกเสนอขึ้นโดย Sørensen ในปี 1948 [41, 42] และได้รับการเลือกเป็นมาตรชี้วัดมาตรฐานสำหรับการวิเคราะห์ภาพทางการแพทย์และการใช้งานอื่นๆ โดยค่าสัมประสิทธิ์ DSC จะวัดพื้นที่การทับซ้อนระหว่างภาพทำนายและภาพคำตอบ แบบให้น้ำหนักของการทำนายถูก (TP) ซึ่งเปรียบเสมือนดัชนีชี้วัดความสามารถในการตรวจจับกรณีทำนายถูก (Sensitivity — The ability to capture positives) สูตรสำหรับดัชนีค่าสัมประสิทธิ์ DSC เป็นตามสมการที่ (2.6)

$$DSC = \frac{2 * TP + smooth}{2 * TP + FP + FN + smooth} \quad (2.6)$$

อย่างไรก็ดี ในการฝึกฝนโมเดลด้วยการใช้ดัชนี DSC เป็นฟังก์ชันความผันเปลี่ยนค่าสูญเสีย (Loss Function) จำเป็นต้องได้รับการเพิ่มค่าคงที่ไว้ในสมการ (smooth) เพื่อคงการคำนวณไม่ให้เกิดกรณีที่ตัวหารเป็น 0 และทำให้การฝึกฝนของโมเดลไม่สามารถประเมินค่า Loss เพื่อปรับ weight ให้เหมาะสมตามการคำนวณที่ควรจะเป็น

จากทฤษฎีทุกองค์ประกอบที่ได้กล่าวมาข้างต้น นำมาสู่การศึกษางานวิจัยที่เกี่ยวข้อง และเป็น การตอบย้าถึงความสอดคล้องของทฤษฎีที่ควรศึกษาประกอบการนำไปประยุกต์ใช้งานในเชิงการวิจัย

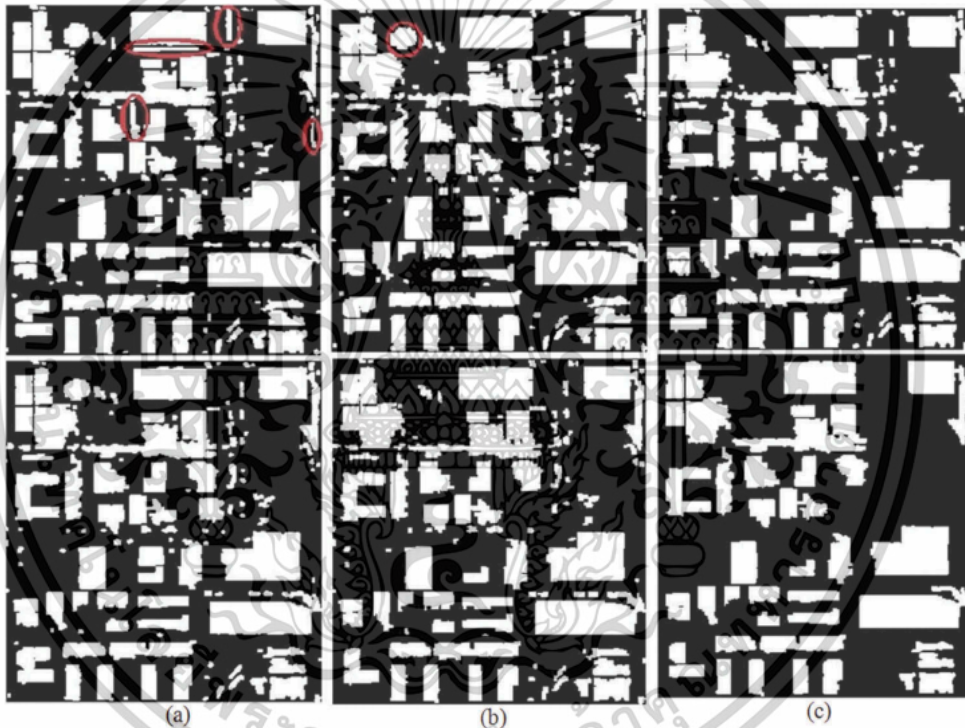
2.6 งานวิจัยที่เกี่ยวข้อง

การประยุกต์ใช้เทคนิคต่าง ๆ ในการสกัดพื้นที่อาคารจากภาพถ่ายดาวเทียมได้รับการศึกษา และสำรวจอย่างกว้างขวาง เทคนิคเหล่านี้ครอบคลุมแนวทางที่หลากหลาย รวมถึงวิธีการประมวลผลภาพ การใช้อัลกอริทึมการเรียนรู้เชิงลึก และการประยุกต์ใช้การประมวลผลภาพ ดังนั้น ในส่วนต่อไปนี้จะนำเสนอการศึกษางานวิจัยที่เกี่ยวข้อง โดยการค้นคว้าและศึกษางานวิจัยที่เกี่ยวข้องมีหลักการอ้างอิงจากวัตถุประสงค์ ทฤษฎี และวิธีดำเนินการตามระเบียบวิจัย

2.6.1 การสกัดอาคารจากภาพถ่ายดาวเทียมด้วยการใช้สัญญาณวิทยาทางคณิตศาสตร์แบบอัตโนมัติ

การสกัดอาคารจากภาพถ่ายดาวเทียมด้วยด้วยกระบวนการใช้การปฏิบัติเชิงสัญญาณวิทยาทางคณิตศาสตร์ (Mathematical morphology operation) [43] ร่วมกับวิธีการแบ่งกลุ่มสีด้วย K-means กับข้อมูลภาพถ่ายดาวเทียมแพนโครเมติกจากดาวเทียม IKONOS เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

และผลลัพธ์ที่ระบุในงานวิจัยนี้พบว่า วิธีการที่ใช้สามารถสกัดบ้านจากภาพถ่ายดาวเทียมมีความสมบูรณ์ (Completeness) ร้อยละ 88 ความถูกต้อง (Correctness) ร้อยละ 90 และคุณภาพ (Quality) ที่ร้อยละ 83 จากเทคนิคการประเมินด้วยจำนวนวัตถุ (Object-based Evaluation) และความสมบูรณ์ร้อยละ 82 ความถูกต้องร้อยละ 88 และคุณภาพที่ร้อยละ 76 ด้วยการประเมินด้วย Pixel (Pixel-based Evaluation) ทั้งนี้ เบื้องหลังได้มีดำเนินขั้นตอนที่สำคัญต่อผลลัพธ์ คือ ขั้นตอนการลบข้อผิดพลาดจากการตรวจพบ (Removal of false detected building) เป็นขั้นตอนที่เสนอหลักเกณฑ์คุณสมบัติในการคัดเลือกอาคาร โดยดูจากอัตราส่วนตามแกน ความโค้งมนรูปทรงวงรี และพื้นที่ ซึ่งในขั้นตอนนี้ได้มีส่วนช่วยลดข้อผิดพลาดในการใช้กระบวนการแบ่งส่วนหรือการสกัดที่กล่าวในข้างต้นได้ ดังรูปที่ 2.28



รูปที่ 2.28 (a) ลบส่วนที่มีความผิดพลาดตามอัตราส่วนแกน (b) ลบส่วนที่ผิดพลาดตามความโค้งมนหรือกลม (c) ลบส่วนที่ผิดพลาดตามขนาด

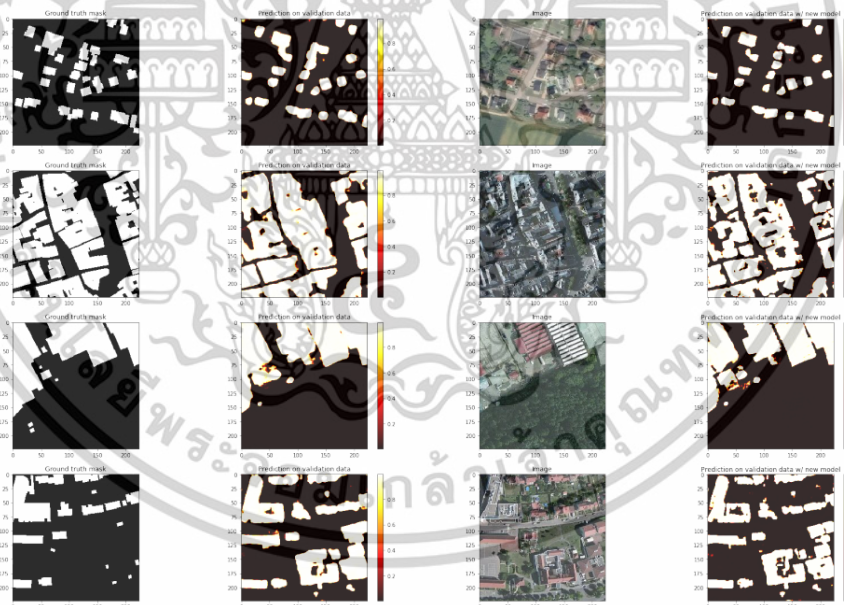
แม้ว่าการลดข้อผิดพลาดในขั้นตอน Removal of false detected building จะได้ผล แต่ในขณะเดียวกันนั้นก็พบว่ายังมีอาคารที่ความสว่างใกล้เคียงกับวัตถุอื่น ๆ อาทิ ถนน ลานจอดรถ สนามในสวนสาธารณะ ซึ่งส่งผลให้กระบวนการที่ใช้ในการสกัดยังไม่สามารถสกัดอาคารออกมาได้ ทว่าเรื่องประสิทธิภาพ ความเร็ว การดำเนินการที่เรียบง่าย และไม่ต้องใช้ข้อมูลเสริมใด ๆ ทำให้เป็นข้อได้เปรียบที่ชัดเจนของงานวิจัยชิ้นนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.6.2 การสกัดอาคารจากภาพถ่ายดาวเทียมด้วยโมเดลยูเน็ต

การศึกษาก่อนหน้านี้ได้ระบุข้อจำกัดในด้านต่าง ๆ โดยเฉพาะอย่างยิ่งเมื่อต้องเผชิญกับภาพถ่ายดาวเทียมที่มีส่วนประกอบทางธรรมชาติผันผวน ไม่ว่าจะเป็นต้นไม้ เงา พื้นที่เกษตรกรรม ตลอดจนความเป็นปัจเจกคุณลักษณะของหลังคาอาคาร ความท้าทายเหล่านี้เป็นอุปสรรคต่อประสิทธิภาพของวิธีการประมวลผลภาพสำหรับการสกัดอาคาร ตัวอย่างงานวิจัยต่อไปได้นำเสนอโมเดลการเรียนรู้เชิงลึก U-Net เข้ามาดำเนินการ

เริ่มจากงานวิจัยในปี พ.ศ. 2560 Chhor และผู้ร่วมวิจัย [44] นำเสนอการประยุกต์ใช้โครงข่ายประสาทเทียมแบบ Convolutional Neural Network (CNN) บนสถาปัตยกรรม U-Net ซึ่งเดิมออกแบบมาสำหรับการแบ่งส่วนภาพทางการแพทย์ ทว่างานวิจัยดังกล่าวนำมาใช้เพื่อสกัดอาคารออกจากภาพถ่ายดาวเทียม ประเด็นสำคัญที่น่าสนใจของงานวิจัยนี้คือ การเสริมข้อมูล (Data Augmentation) และการประยุกต์ใช้ตัวชี้วัดการประเมินผล โดยเฉพาะดัชนี Jaccard Index (IoU) และค่าสัมประสิทธิ์ DSC ซึ่งให้คะแนน 59% และ 74% ตามลำดับ ในส่วนของผลลัพธ์เป็นดังรูปที่ 2.29 อย่างไรก็ตาม งานวิจัยดังกล่าวไม่ได้ใช้เทคนิคหลังการประมวลผล (Post-processing)



รูปที่ 2.29 เปรียบเทียบชุดคำตอบกับผลการทำนายการแบ่งส่วนอาคารด้วยโมเดลที่ผ่านการเรียนรู้เชิงลึกด้วยโครงสร้าง U-Net

กระนั้น อีกหนึ่งการวิจัยที่ได้ประยุกต์ใช้ U-Net สำหรับการสกัดอาคารจากภาพถ่ายดาวเทียม Ivanovsky และผู้ร่วมวิจัย ปีพ.ศ. 2562 [45] ได้นำแนวทางใหม่ในการจัดหมวดหมู่ชุดข้อมูลเป็น "อาคาร" และ "ไม่ใช่อาคาร" เพื่อฝึกฝนโมเดล ผลลัพธ์ถูกนำมาเปรียบเทียบกับ

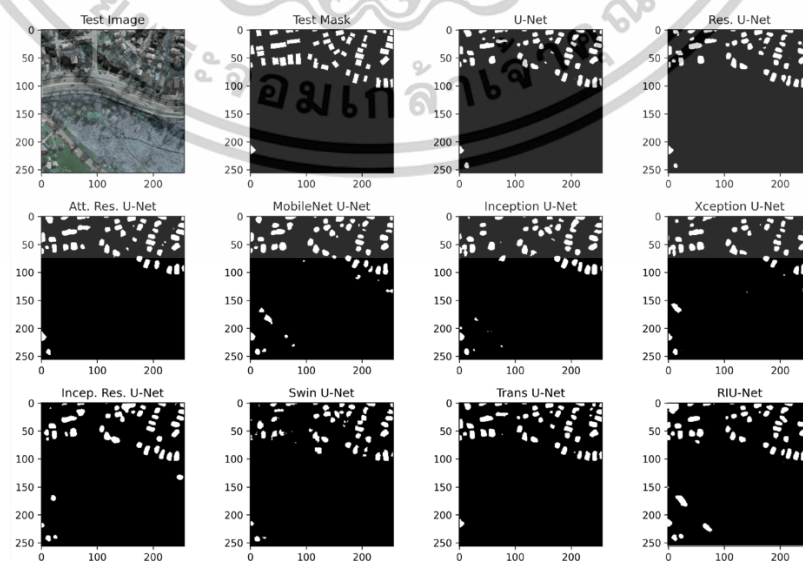
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โมเดล LinkNet ซึ่งเผยให้เห็นประสิทธิภาพที่เหนือกว่าของ U-Net ในแง่ของคุณภาพการสกัดอาคาร โดยมีการใช้ค่าสัมประสิทธิ์ DSC ในการประเมินประสิทธิภาพของการทดลอง และผลลัพธ์ค่าสัมประสิทธิ์ได้เท่ากับ DSC 77%

ท้ายที่สุด งานวิจัยของ Daranagama ในปี พ.ศ. 2564 [46] การศึกษานี้นำเสนอวิธีการในการแยกพื้นที่อาคารจากภาพถ่ายดาวเทียมและภาพทางอากาศของพาหนะ UAV ที่มีความละเอียดสูง แนวทางดังกล่าวได้ประยุกต์ใช้เทคนิคการประมวลผลข้อมูลล่วงหน้า (Pre-processing) เพื่อเพิ่มความแม่นยำก่อนการนำไปฝึกฝนด้วยสถาปัตยกรรม U-Net นอกจากนี้ สถาปัตยกรรมยังได้รับการแก้ไขเพื่อให้สอดคล้องกับเทคนิคการประมวลผลล่วงหน้า ทั้งนี้ อัลกอริทึมการปรับปรุงคุณภาพภาพซึ่งใช้ในขั้นตอนก่อนการประมวลผลคือ Contrast Limited Adaptive Histogram Equalization (CLAHE) ที่ได้รับการพิสูจน์แล้วว่ามีความมีประสิทธิภาพในการปรับปรุง

2.6.3 การสกัดอาคารจากภาพถ่ายดาวเทียมด้วยโมเดลอาร์ไอยูเน็ต

นอกเหนือจากการใช้งานโมเดล U-Net แบบดั้งเดิมสำหรับการสกัดอาคารจากภาพถ่ายดาวเทียมแบบอัตโนมัติ ยังมีการสำรวจแนวคิดใหม่เกี่ยวกับการพัฒนาสถาปัตยกรรมของโมเดล U-Net เพื่อเพิ่มประสิทธิภาพในการแยกคุณสมบัติ (Feature Extraction) ในส่วนตัวเข้ารหัส (Encoder) ของโมเดลดั้งเดิมอีกด้วย โดย Sariturk และผู้ร่วมวิจัยได้นำเสนอการใช้ Residual-Inception [47] ในปีพ.ศ. 2565 ผลการวิจัยระบุว่า RIU-Net ให้ผลลัพธ์ที่ดีที่สุดในช่วงข้อมูล Inria เมื่อเปรียบเทียบกับโมเดลอื่น ๆ อีกหลายรุ่น เช่น U-Net, Residual U-Net และ Trans U-Net เป็นต้น ตัวอย่างดังรูปที่ 2.30 ดังนั้น โมเดล RIU-Net จึงถูกพิจารณานำมาทดลองเปรียบเทียบในการวิจัยนี้



รูปที่ 2.30 เปรียบเทียบผลลัพธ์การทำนายของแต่ละโมเดลในงานวิจัยของ Sariturk และผู้ร่วมวิจัย

ไม่ว่าการณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.6.4 การสกัดอาคารจากภาพถ่ายดาวเทียมด้วยโมเดลเรสยูเน็ต

แนวคิดในเรื่องการปรับปรุงตัวเข้ารหัส ในการสกัดคุณลักษณะอาคารจากภาพถ่ายดาวเทียมได้รับการวิจัยอย่างหลากหลายมิติในงานวิจัยต่าง ๆ ทว่าแนวทางหนึ่งที่น่าสนใจคือ การประยุกต์ใช้สถาปัตยกรรม Residual Network (ResNet) เข้ากับโครงสร้าง U-Net ชื่อ Res-U-Net การปรับปรุงโครงสร้างนี้ช่วยให้การส่งถ่ายคุณลักษณะแต่ละเลเยอร์มีประสิทธิภาพมากขึ้น โดยเฉพาะอย่างยิ่งกับกระบวนการคำนวณ Gradient Propagation ระหว่างการฝึกฝนของโมเดล นอกจากนี้ ด้วยเทคนิคการถ่ายโอนความรู้ที่ใช้ Pretrained weight ImageNet ของโครงข่าย ResNet ยังมีส่วนช่วยให้โมเดลมีความรู้พื้นฐานเกี่ยวกับรูปภาพมากกว่าการประยุกต์ใช้ U-Net ในการศึกษาก่อนหน้านี้

งานวิจัยที่นำเสนอโดย Xu และผู้ร่วมวิจัยในปีพ.ศ. 2561 [48] Res-U-Net ผสมผสาน ResNet และกระบวนการหลังการประมวลผล ทำให้ผลลัพธ์มีความแม่นยำสูง โดยค่าความแม่นยำโดยรวม (OA) อยู่ที่ 97.71% และ 96.91% ในชุดข้อมูล Vaihingen และ Potsdam ตามลำดับ ทำนองเดียวกัน ในการศึกษา ที่เสนอโดย Alsabhan และผู้ร่วมวิจัยในปีพ.ศ. 2565 [49] U-Net พร้อมด้วย ResNet ซึ่งเป็นสถาปัตยกรรมแกนหลักถูกรวมเข้าด้วยกันเพื่อสกัดอาคารจากภาพถ่ายดาวเทียม ได้ผลลัพธ์ที่บ่งชี้ประสิทธิภาพที่ดีขึ้นเมื่อเทียบกับ U-Net แบบดั้งเดิม ซึ่งประเมินด้วยดัชนี IoU และค่าสัมประสิทธิ์ DSC ท้ายที่สุด การศึกษางานวิจัยที่นำเสนอโดย Alsabhan และผู้ร่วมวิจัยในปีพ.ศ. 2565 [50] แสดงให้เห็นว่าการใช้ ResNet-50 และ ResNet-152 เป็นสถาปัตยกรรมแกนหลักทำให้ได้ผลลัพธ์ที่ดีกว่าการใช้ CNN ดั้งเดิม

2.6.5 การสกัดอาคารจากภาพถ่ายดาวเทียมด้วยโมเดลดีพแล็บวีทีพีพลัส

ตัวอย่างการนำ CNNs ไปใช้งานในด้านการสกัดอาคารจากภาพถ่ายดาวเทียมยังถูกค้นพบในอีกรูปแบบที่น่าสนใจ คือ DeepLabv3+ ทั้งนี้ การศึกษางานวิจัยต่อไปนี้เกี่ยวข้องกับผลลัพธ์ที่เทียบกับโมเดล U-Net เริ่มจากงานวิจัยของ Aslantaş และผู้ร่วมวิจัย [51] ที่ได้นำเสนอเทคนิคในการสกัดอาคารโดยใช้ DeepLabv3+ ในปีพ.ศ. 2564 ซึ่งใช้สถาปัตยกรรม DeepLabV3+ ที่คล้ายคลึงกับโมเดลดั้งเดิม และการทดลองมุ่งเน้นไปที่การปรับไฮเปอร์พารามิเตอร์เพื่อให้ได้ผลลัพธ์ที่ดีที่สุดสำหรับชุดข้อมูลอาคารทางอากาศของมหาวิทยาลัยหูอู่ฮั่น (Wuhan University — WHU) ผลการวิจัยระบุว่ามีความแม่นยำสูงในการทำนายด้วยดัชนีชี้วัด IoU อยู่ที่ 98.23% นอกจากนี้ การศึกษาที่นำเสนอโดย Han และผู้ร่วมวิจัย ในปีพ.ศ. 2565 [52] มีวัตถุประสงค์เพื่อปรับปรุงประสิทธิภาพของการสกัดอาคารโดยแก้ไขปัญหาต่าง ๆ เช่น ความเร็วในการสกัดที่ช้า และการแบ่งส่วนขอบอาคารที่ไม่สมบูรณ์ ซึ่งสามารถทำได้โดยการปรับเปลี่ยน Backbone และ Pyramid Pooling Module ในการทดลอง วิธีการที่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

นำเสนอได้นำมาเปรียบเทียบกับแบบจำลองการทำนายอื่น ๆ ไม่ว่าจะเป็น U-Net, SegNet, PSPNet และ DeepLabV3+ โดยใช้ชุดข้อมูล 2 ชุด ได้แก่ WHU และ Massachusetts ผลการวิจัยพบว่า วิธีที่เสนอรูปแบบใหม่ในการปรับเปลี่ยนโครงสร้างมีประสิทธิภาพเหนือกว่าโมเดลตัวอื่น ๆ ที่ได้กล่าวถึง

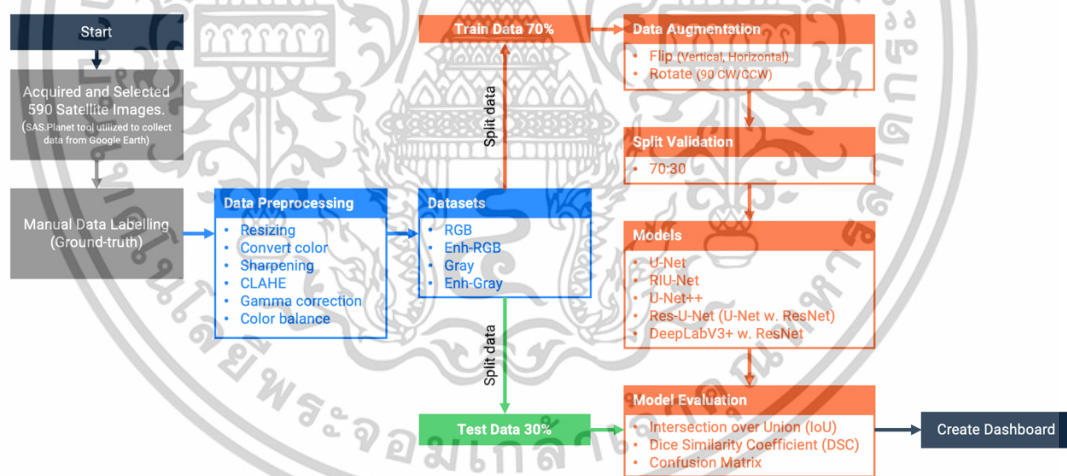
อย่างไรก็ตาม เมื่อได้ตรวจสอบและศึกษาการวิจัยเกี่ยวกับโมเดลการเรียนรู้เชิงลึกสำหรับการแบ่งส่วนภาพในการแยกอาคาร ผู้วิจัยได้พบการกล่าวถึงในแง่เปรียบเทียบ U-Net และ DeepLabV3+ ตัวอย่างเช่น ในงานวิจัยที่เสนอโดย Bakirman และผู้ร่วมวิจัยในปีพ.ศ. 2565 [53] นอกจากนี้ ยังมีการนำเสนอสถาปัตยกรรม U-Net++ ที่ได้รับการปรับปรุงโดยใช้ตัวเข้ารหัส SE-ResNeXt101 ที่ได้รับการฝึกอบรมล่วงหน้าด้วย ImageNet การทดลองของพวกเขาครอบคลุมการดำเนินการเพื่อเปรียบเทียบตัวเข้ารหัส (Encoder) ที่ใช้สำหรับสกัดคุณลักษณะของรูปภาพ และผลลัพธ์เผยให้เห็นความแม่นยำด้วยดัชนี IoU ที่ 75.39% และ 92.53% บนชุดข้อมูล Inria และ Massachusetts ตามลำดับ

จากการศึกษางานวิจัยที่เกี่ยวข้อง เมื่อพิจารณาถึงข้อจำกัดของการประมวลผลภาพแบบดั้งเดิม ในการสกัดอาคารจากภาพถ่ายดาวเทียมเพื่อให้ได้ผลลัพธ์ที่แม่นยำ โดยเฉพาะอย่างยิ่งในภูมิภาคชนบทและชานเมือง การวิจัยครั้งนี้จึงมุ่งเน้นไปที่โมเดลการเรียนรู้เชิงลึก ซึ่งเป็นที่ทราบกันดีว่ามีความสามารถในการแบ่งส่วนภาพ ทั้งนี้ งานวิจัยที่เกี่ยวข้องได้นำเสนอผลลัพธ์ที่น่าพึงพอใจในการสกัดอาคารจากภาพถ่ายดาวเทียม ไม่ว่าจะเป็นโมเดล U-Net, RIU-Net, U-Net++ และ Res-U-Net ล้วนชี้ให้เห็นถึงความแม่นยำในการทำนาย อย่างไรก็ตาม โมเดล DeepLabV3+ มีโครงสร้างสถาปัตยกรรมที่แตกต่างจากโมเดลตระกูล U-Net นอกจากนี้ บางงานวิจัยได้มีการประยุกต์ใช้เทคนิคการประมวลผลภาพร่วมในการทดลองเพื่อชี้ถึงความสำเร็จในการเพิ่มประสิทธิภาพ ดังนั้น รายละเอียดการวิจัยจะถูกกล่าวถึงในบทถัดไป

บทที่ 3

วิธีการดำเนินงานวิจัย

สาระสำคัญในบทนี้ กล่าวถึงขั้นตอนและวิธีการดำเนินงานวิจัยอันเป็นผลพวงจากการศึกษา งานวิจัยที่เกี่ยวข้อง เพื่อให้สอดคล้องตามวัตถุประสงค์งานวิจัย ระเบียบดำเนินงานต่อไปนี้ได้รับการ ออกแบบที่ครอบคลุมตั้งแต่ กระบวนการรวบรวมข้อมูลภาพถ่ายดาวเทียมความละเอียดสูงในจังหวัด เลย ประเทศไทย กระบวนการจัดทำชุดคำตอบหรือป้ายกำกับคำตอบ เพื่อใช้เปรียบเทียบกับดัชนีชี้ วัดความคล้ายคลึง รวมถึงใช้สำหรับฝึกฝนโมเดลการเรียนรู้เชิงลึกแบบมีผู้สอน การประยุกต์ใช้เทคนิค การเสริมชุดข้อมูลเพื่อเพิ่มปริมาณข้อมูลฝึกฝนให้แก่โมเดลการเรียนรู้เชิงลึก การแบ่งหมวดหมู่ชุด ข้อมูลและการแบ่งสัดส่วนที่ใช้ภายใต้กระบวนการฝึกฝนและทดสอบ ตลอดจนการจัดทำแดชบอร์ด (Dashboard) เพื่อแสดงผลและวิเคราะห์ข้อมูล โดยกระบวนการสามารถอธิบายได้ดังภาพที่ 3.1 นอกจากนี้ เนื้อหาในบทนี้ยังครอบคลุมถึงแผนผังแสดงลำดับขั้นตอนการทดลองของการสกัดอาคาร จากภาพถ่ายดาวเทียม ด้วยวิธีประมวลผลภาพแบบสัณฐานวิทยา และวิธีการประยุกต์ใช้โมเดลการ เรียนรู้เชิงลึก



รูปที่ 3.1 แผนภาพแสดงขั้นตอนการดำเนินงานวิจัย

3.1 การกำหนดพื้นที่ศึกษาและการรวบรวมข้อมูล

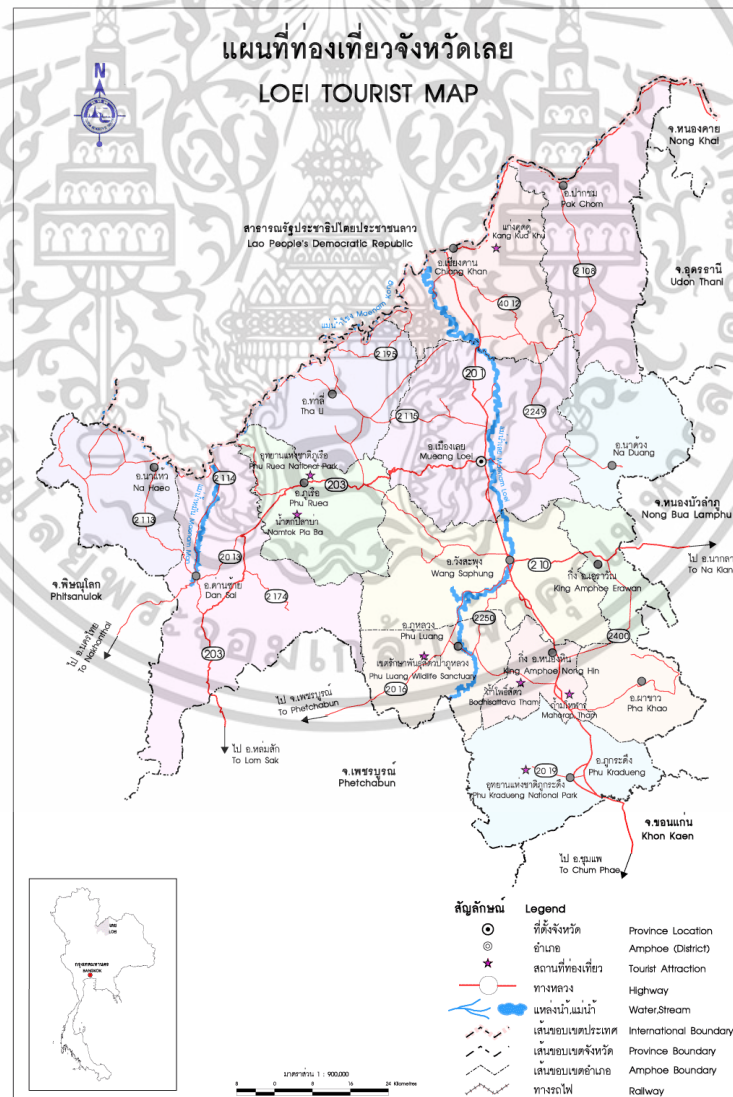
งานวิจัยนี้มุ่งเน้นไปที่การรวบรวมข้อมูลภาพถ่ายดาวเทียมในพื้นที่ภูเขาที่ห่างไกล และมีความ ทำทายทางภูมิศาสตร์ในบริบทชนบท ชานเมือง พื้นที่ห่างไกลเมืองหลวง โดยลักษณะภูมิประเทศ ปรากฏเป็นป่าทึบ พื้นที่ทางการเกษตร ตลอดจนชุมชนอาศัย ซึ่งตัวอย่างที่กล่าวมาได้นำไปสู่ความ ทำทายที่น่าสนใจในการวิจัย กล่าวคือ ลักษณะอันโดดเด่นของพื้นที่ที่เลือกมาศึกษา คือการทับซ้อน

ระหว่างสิ่งแวดล้อมทางธรรมชาติและสิ่งปลูกสร้างอย่างอาคารบ้านเรือนของมนุษย์ เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ดังนั้น ข้อมูลภาพถ่ายดาวเทียมที่นำมาประกอบการวิจัยจำเป็นต้องได้รับกล่าวถึง เหตุเพราะ ข้อมูลภาพถ่ายดาวเทียมความละเอียดสูง ซึ่งมีคุณสมบัติพร้อมในด้านคุณภาพ ความสะดวกในการเข้าถึง อนุมัติลิขสิทธิ์การนำไปใช้วิจัย มีรอบความถี่สม่ำเสมอในการปรับปรุงข้อมูลเป็นปัจจุบัน และที่สำคัญคือไม่มีค่าใช้จ่ายเป็นข้อจำกัดการวิจัย ดังนั้น หนึ่งในผู้ให้บริการข้อมูลตามคุณสมบัติที่กล่าวมา คือบริษัท Google ซึ่งทางผู้วิจัยและคณะต้องขอแสดงความขอบคุณเป็นอย่างสูง

3.1.1 ลักษณะทางภูมิศาสตร์และการปกครองจังหวัดเลย

จังหวัดเลยมีพื้นที่ประมาณ 11,424 ตารางกิโลเมตร ทางทิศเหนือของจังหวัดติดต่อกับประเทศสาธารณรัฐประชาธิปไตยประชาชนลาว สภาพแวดล้อมทางภูมิศาสตร์เป็นพื้นที่ราบลุ่มระหว่างหุบเขาและมีแนวเทือกเขา อีกทั้งยังมีแม่น้ำโขงและแม่น้ำเลยตัดผ่านจังหวัด สภาพภูมิอากาศแบบร้อนชื้นแต่หากเจอมรสุมพาดผ่านจะทำให้มีฝนตกติดต่อกันหลายวัน



รูปที่ 3.2 แผนที่ท่องเที่ยวจังหวัดเลยแสดงเส้นทาง ขอบเขต และแม่น้ำในจังหวัด
เอกสารนี้เป็นเอกสารที่สละส่วนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 3.2 แสดงให้เห็นว่าอำเภอเมืองเลยตั้งอยู่ตอนกลางค่อนข้างไปทางตอนบน การปกครองภายในจังหวัดเลยแบ่งออกเป็น 14 อำเภอ 89 ตำบล โดยผู้วิจัยได้เลือกอำเภอเมืองเลยซึ่งแบ่งส่วนการปกครองย่อยออกเป็น 9 ตำบล มีประชากรรวม 59,155 คน จำนวนบ้าน 24,811 หลังคาเรือน อ้างอิงจำนวนประชากรและจำนวนบ้านในปีพ.ศ. 2566 เดือน พฤษภาคม จากเว็บไซต์สำนักงานสถิติแห่งชาติ ดังรูปที่ 3.3 ทั้งนี้ การเลือกพื้นที่ศึกษาต้องคำนึงถึงกรณีที่ต้องต้องลงพื้นที่สำรวจที่จริง และจากรูปที่ 3.3 ตำบลที่มีจำนวนประชากรและบ้านสูงสุด คือ ตำบลเมือง

สถิติจำนวนประชากรและบ้าน		พื้นที่ อำเภอเมืองเลย			
ปี	ชาย (คน)	หญิง (คน)	รวม (คน)	บ้าน (หลัง)	
รวมทั้งหมด	29,791	29,364	59,155	24,811	
ตำบลเมือง	6,414	6,829	13,243	8,206	
ตำบลกกสุ้ย	3,716	3,471	7,187	2,462	
ตำบลน้ำมน	2,011	1,992	4,003	1,606	
ตำบลศรี	1,877	1,848	3,725	1,249	
ตำบลน้ำสวย	3,504	3,334	6,838	2,099	
ตำบลเชียงคาน	3,481	3,641	7,042	2,756	
ตำบลนาพระ	3,409	3,492	6,901	2,328	
ตำบลศรีสองรัก	4,112	3,354	7,476	3,373	
ตำบลเชียง	1,347	1,393	2,740	732	

รูปที่ 3.3 จำนวนประชากรและบ้านในแต่ละตำบลของอำเภอเมืองเลย

ทว่าความเป็นตำบลเมืองจะไม่ตรงกับคุณสมบัติที่ต้องการจะศึกษา ผู้วิจัยจึงเลือกตำบลศรีสองรักซึ่งครอบคลุมพื้นที่ระหว่างเส้นละติจูดที่ 17.59° ถึง 17.69° N และเส้นลองจิจูดที่ 101.67° ถึง 101.81° E หรือปกคลุมพื้นที่ประมาณ 128.50 ตารางกิโลเมตร มีจำนวนประชากรและบ้านเป็นอันดับรองลงมา อีกทั้งยังมีคุณสมบัติตรงกับกรณีที่ต้องการจะศึกษาเพราะมีทั้งแม่น้ำเลยและเทือกเขาอยู่ในพื้นที่ ดังรูปที่ 3.4



รูปที่ 3.4 แผนที่ (ซ้าย) และภาพถ่ายดาวเทียม (ขวา) ของตำบลศรีสองรักจาก Google Map และ

Google Earth เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีกรนำมาใช้

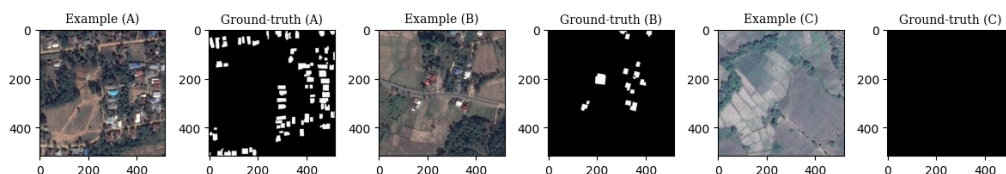
3.1.2 การรวบรวมข้อมูลภาพถ่ายดาวเทียมจาก Google Earth ด้วย SAS.Planet

ชุดข้อมูลภาพถ่ายดาวเทียมความละเอียดสูงที่ได้รับจาก Google Earth [54] ถูกนำมาที่ใช้ในการวิจัยนี้ อย่างไรก็ดี เป็นที่ทราบกันว่าการศึกษาและงานวิจัยอื่น ๆ ที่ได้รับการตีพิมพ์ผลงานวิชาการได้ใช้ภาพถ่ายดาวเทียมเพื่อจุดประสงค์ที่คล้ายกัน [55, 56, 57] สำหรับการรวบรวมข้อมูลได้ใช้เครื่องมือ SAS.Planet [58] ซึ่งอำนวยความสะดวกในการดึงภาพจาก Google Earth ตามพื้นที่ที่สนใจตามที่ระบุก่อนหน้า และดำเนินการแบ่งจำนวนรูปภาพด้วยฟังก์ชัน Selection Manager (Tab Stitch) ดังนั้น จึงได้ชุดข้อมูลภาพสี (3-channels; RGB) จำนวน 1,672 รูป โดยแต่ละภาพมีขนาด 517 x 517 พิกเซล และความละเอียดเชิงพื้นที่ประมาณ 0.57 - 0.6 เมตร หรือเทียบเท่ากับกำลังขยาย 19 เท่าของข้อมูลต้นฉบับ โดยไฟล์รูปภาพที่ได้จากเครื่องมือ SAS.Planet เป็นไฟล์รูปภาพสกุล Bitmap (BMP)

หลังดำเนินการตรวจสอบภาพอย่างละเอียดโดยผู้วิจัย พบว่า ภาพส่วนมากปรากฏซึ่งป่าไม้และพื้นที่รกร้าง ไม่มีอาคารบ้านเรือน ดังนั้น จากภาพที่ได้มาทั้งหมดจำต้องนำมาดำเนินการแบ่งกลุ่มออกเป็น 2 กลุ่ม คือ กลุ่มภาพที่มีอาคาร และกลุ่มภาพที่ไม่มีอาคาร ตามด้วยการคัดเลือกภาพแบบสุ่มจากแต่ละกลุ่ม ท้ายที่สุด จำกัดภาพเหลือเพียง 590 รูป แบ่งเป็นภาพที่มีอาคาร (Building) จำนวน 294 รูป และภาพที่ไม่มีอาคาร (non-Building) จำนวน 296 รูป ทั้งนี้ เพื่อเป็นการถ่วงดุลปริมาณข้อมูลประกอบการฝึกฝน และหลีกเลี่ยงไม่ให้โมเดลเรียนการรู้เชิงลึก ฝึกฝนคุณลักษณะข้อมูลประเภทใดประเภทหนึ่งมากเกินไป

3.2 การจัดทำชุดคำตอบ

ชุดคำตอบถูกจัดทำขึ้นทั้งสิ้น 590 รูปโดยผู้วิจัย คุณลักษณะของรูปในชุดคำตอบเป็นไฟล์ภาพแบบไบนารีมาสก์ (Binary mask) ขนาด 517 x 517 พิกเซล โดยรายละเอียดภายในภาพถูกแทนสีด้วยสีสองสี ได้แก่ สีขาว แทนพื้นที่อาคารที่ปรากฏในภาพถ่ายดาวเทียม และสีดำ แทนพื้นที่ที่ไม่ใช่อาคาร ดังรูปที่ 3.5 อย่างไรก็ตาม สิ่งสำคัญคือการระบุตำแหน่งและขนาดของอาคารให้ใกล้เคียงต่อความเป็นจริง ดังนั้น ผู้วิจัยจึงได้เลือกใช้เครื่องมือหรือซอฟต์แวร์คอมพิวเตอร์ชื่อว่า Adobe Photoshop ในการระบุอาคารจากภาพถ่าย ซึ่งเครื่องมือดังกล่าวสามารถสร้างเลเยอร์ภาพซ้อนทับและปรับความโปร่งแสงของการตีป้ายกำกับอาคารได้อย่างแม่นยำ



รูปที่ 3.5 ตัวอย่างภาพชุดคำตอบ (Ground-truth) โดยชุดภาพ (A) แสดงถึงคุณลักษณะภาพถ่ายที่ปรากฏอาคารหนาแน่น ชุดภาพ (B) ปรากฏอาคารเล็กน้อย และชุดภาพ (C) ไม่ปรากฏอาคาร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

วิธีการรวบรวมข้อมูลและจำทำชุดคำตอบในการศึกษาค้นคว้าครั้งนี้แตกต่างจากการวิจัยอื่น ๆ ที่ได้หยิบยกมาศึกษา เนื่องจากผู้วิจัยเป็นผู้รวบรวมและจัดทำชุดคำตอบด้วยตนเอง ทั้งนี้ เพื่อดำเนินงานวิจัยตามขอบเขตวัตถุประสงค์ในบริบทประเทศไทยที่ซึ่งไม่มีชุดข้อมูลพร้อมชุดคำตอบ

อย่างไรก็ตาม ด้วยกระบวนการดังกล่าวส่งผลให้เกิดข้อจำกัดในกระบวนการฝึกฝนโมเดลการเรียนรู้เชิงลึก เนื่องจากปริมาณชุดข้อมูลที่มีขนาดค่อนข้างเล็ก ดังนั้น ในลำดับถัดไปจะกล่าวถึงกระบวนการจัดเตรียมข้อมูล

3.3 การจัดเตรียมชุดข้อมูล (Data Preparation)

ชุดข้อมูลและชุดคำตอบถูกแบ่งออกเป็นสองกลุ่ม คือ กลุ่มภาพที่ปรากฏอาคาร และกลุ่มภาพไม่ปรากฏอาคาร และเพื่อหลีกเลี่ยงปัญหาการเรียนรู้แบบ Overfitting ซึ่งมีส่วนทำให้โมเดลมีความแม่นยำสูงเกินกว่าที่ควรกับชุดข้อมูลเรียน แต่ด้อยประสิทธิภาพกับชุดข้อมูลใหม่ที่ไม่รู้จัก ดังนั้น จึงได้มีการประยุกต์ใช้เทคนิคการแบ่งสัดส่วนข้อมูลแบบสุ่ม Train-Test-Split โดยจัดสรรข้อมูลในอัตราส่วนร้อยละ 70 สำหรับใช้ฝึกฝน และร้อยละ 30 สำหรับใช้ทดสอบ ทั้งนี้ ดำเนินการคัดเลือก 70:30 จากกลุ่มภาพทั้งสองกลุ่ม ท้ายที่สุด ได้ภาพสำหรับฝึกฝน 413 รูป และภาพสำหรับทดสอบ 177 รูป ฉะนั้น ภาพสำหรับฝึกฝนและทดสอบจะถูกคละรวมไว้ซึ่งทั้งภาพที่มีอาคารและไม่มีอาคาร

3.3.1 การประมวลผลข้อมูลล่วงหน้า (Data Preprocessing)

กระบวนการประมวลผลข้อมูลล่วงหน้าได้ถูกพิจารณาประยุกต์ใช้ เพื่อเปรียบเทียบและค้นหากรรมวิธีที่สามารถเพิ่มความแม่นยำให้กับโมเดลต่อการทำนายอาคารจากภาพถ่ายดาวเทียม โดยชุดคำสั่งถูกเขียนด้วยภาษา Python ร่วมกับไลบรารีที่จำเป็นต่อการประมวลผลภาพ อาทิ โอเพนซีวี (OpenCV) ไพอิมเมจ (Pylmage)

ขั้นต้น การประมวลผลภาพที่นำมาประยุกต์ใช้คือการปรับขนาดภาพ (Resize) เพื่อลดความต้องการทรัพยากรและเวลาในการประมวลผล โดยปรับขนาดรูปภาพจากกว้างคูณยาว 517 พิกเซล เหลือ 256 พิกเซล ทั้งนี้ ดำเนินการกับทั้งชุดข้อมูลฝึกฝนและชุดข้อมูลทดสอบ กระนั้น กระบวนการลดขนาดภาพยังคงไว้ซึ่งคุณลักษณะสำคัญของภาพ หรือกล่าวคือไม่ส่งผลเสียอย่างมีนัยยะในการสกัดคุณลักษณะและเป็นอุปสรรคต่อการเรียนรู้ นอกจากนี้ การศึกษางานวิจัยก่อนหน้าพบว่า Lin และผู้ร่วมวิจัย [59] ได้นำเสนอประสิทธิภาพที่ชี้ว่าเทคนิคการประมวลผลล่วงหน้ามีส่วนช่วยเพิ่มคุณภาพการทำนายได้ดียิ่งขึ้น

ดังนั้น งานวิจัยนี้ได้นำเสนอรูปแบบการประยุกต์ใช้เทคนิคการประมวลผลล่วงหน้าที่แตกต่างกัน 3 รูปแบบ เข้ากับชุดข้อมูลตั้งต้น โดยมีคุณลักษณะตามข้อกำหนดต่อไปนี้

ก) ชุดข้อมูลภาพสี (RGB) ชุดข้อมูลนี้ประกอบด้วยภาพต้นฉบับที่ผ่านการปรับขนาด โดยมีมิติข้อมูลกว้างคูณยาวคูณแชนแนลสี หรือเขียนแทนด้วยสัญลักษณ์ (256x256x3)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข) ชุดข้อมูลภาพสีปรับปรุง (Enh-RGB) เป็นการนำชุดข้อมูลภาพสี ก) RGB ไปปรับแก้ค่าแกรมมา (Gamma) โดยกำหนดเป็นค่าคงที่ที่ 0.75 ซึ่งส่งผลให้เฉดสีของภาพมีความเข้มมากยิ่งขึ้น ทั้งนี้ การปรับแก้ค่าแกรมมามีวัตถุประสงค์เพื่อให้ค่าการสะท้อนแสงของอาคารเด่นชัดกว่าสภาพแวดล้อมทางธรรมชาติที่มีค่าสะท้อนแสงต่ำกว่า เนื่องจากพื้นที่ป่าไม่มีสีเขียวเป็นสีโทนเข้ม จากนั้น ดำเนินการปรับปรุงสมมูลค่าสีแซนแนล CIELAB [60] ซึ่งการปรับสมมูลค่าสีนี้ได้รับแรงบันดาลใจจากแนวคิดที่ว่า การรับรู้ค่าสีของดวงตามนุษย์ไวต่อแซนแนลสีดังกล่าว [61] นอกจากนี้ ค่าสี CIELAB ยังครอบคลุมขอบเขตสีมากกว่าช่วงค่าสี RGB ดังนั้น ด้วยสมมุติฐานที่ว่า หากมนุษย์สามารถระบุสิ่งปลูกสร้างในภาพถ่ายได้ง่ายขึ้น ฉะนั้น โมเดลการเรียนรู้เชิงลึกก็อาจใช้คุณลักษณะสีเป็นคุณลักษณะสำคัญในการระบุอาคารได้เช่นกัน สุดท้าย นำเข้ากระบวนการเบลอแบบเกาส์เซียน (Gaussian Blur) [62] เพื่อช่วยลดค่าสีรบกวน (Noise) และจากนั้น นำเข้ากระบวนการปรับน้ำหนักขอบเส้นวัตถุ (Weighted Image) [63] เพื่อให้ขอบของวัตถุในภาพเด่นชัดขึ้นโดยทั่วกัน

ค) ชุดข้อมูลภาพช่วงสีเทา (Gray) ชุดข้อมูลนี้นำชุดข้อมูลภาพสี ก) RGB มาดำเนินการปรับค่าขอบเขตสีเป็นช่วงสีขาวถึงสีดำ หรือที่เรียกกันว่าช่วงสีเทา (Gray Scale) โดยมีมิติขนาดข้อมูลจะถูกเปลี่ยนเป็น (256x256x1) และอยู่ในค่าตัวเลขแทนสี 0 – 255 โดยค่า 0 แสดงเป็นสีดำ ในขณะที่ค่า 255 คือสีขาว

ง) ชุดข้อมูลภาพช่วงสีเทาแบบปรับปรุง (Enh-Gray) เป็นกระบวนการนำชุดข้อมูลภาพสี ค) Gray มาปรับปรุงด้วยเทคนิคเพิ่มความคมชัดของภาพด้วยการปรับน้ำหนักขอบเส้นวัตถุ ตามต่อด้วยเทคนิคการปรับค่าสีคอนทราสต์ Contrast Limited Adaptive Histogram Equalization (CLAHE) [64] เพื่อปรับปรุงการกระจายตัวค่าความเข้มแสงและสีด้วยการถ่วงดุลค่าคอนทราสต์ของภาพ

จากกระบวนการแบ่งชุดข้อมูลออกเป็น 4 ชุด ก) ถึง ง) เผยให้เห็นความแตกต่างของชุดข้อมูลดังรูปที่ 3.6



รูปที่ 3.6 ตัวอย่างภาพชุดข้อมูล ก) RGB ข) Enh-RGB ค) Gray และ ง) Enh-Gray ตามลำดับ

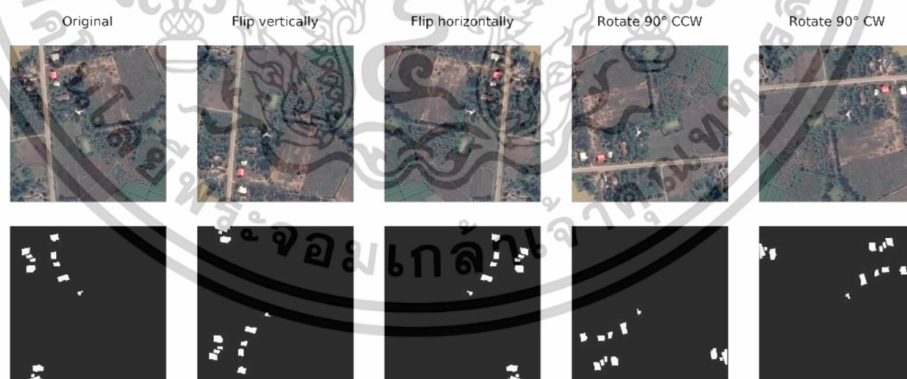
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.3.2 การเสริมชุดข้อมูล (Data Augmentation)

เนื่องจากชุดข้อมูลฝึกฝนที่ใช้สำหรับการเรียนรู้ของโมเดลมีขนาดจำกัด ผู้วิจัยจึงประยุกต์ใช้เทคนิคการเสริมข้อมูลเพื่อเพิ่มปริมาณ โดยการศึกษางานวิจัยที่นำเสนอโดย Shorten และผู้ร่วมวิจัย [65] ได้ชี้ให้เห็นว่าการเรียนรู้ของโมเดลเชิงลึกสามารถเพิ่มข้อมูลโดยใช้วิธีการจัดการค่าสี และวิธีการอื่น ๆ

แนวทางการเสริมข้อมูลที่เลือกประกอบไปด้วยวิธีการเปลี่ยนแปลงสองรูปแบบ ได้แก่ การหมุนและการพลิก (Rotation and Flipping) โดยการหมุนจะหมุนสองทิศทางคือ การหมุนแบบ 90 องศาตามเข็มนาฬิกาและทวนเข็มนาฬิกา เหตุเพราะการหมุนไม่แปรผันต่อคุณลักษณะการวางของอาคารภายในภาพถ่ายดาวเทียม อีกทั้งอาจมีส่วนช่วยให้โมเดลสามารถตรวจจับอาคารได้ดียิ่งขึ้น ขณะเดียวกัน การพลิกจะดำเนินการตามแกนแนวดิ่งและแนวนอน (Horizontal and Vertical) โดยการพลิกแนวดิ่งจะกลับภาพจากบนลงล่าง ในขณะที่การพลิกแนวนอนจะสะท้อนภาพตามแนวกึ่งกลางแนวดิ่ง โดยการพลิกกลับเหล่านี้มีส่วนช่วยให้โมเดลมีความยืดหยุ่นต่อการเปลี่ยนแปลงมุมมองและการจัดตำแหน่งอาคาร

แม้ว่ากระบวนการประมวลผลภาพเพื่อเสริมข้อมูลเหล่านี้ก่อให้เกิดการเปลี่ยนแปลงรูปลักษณ์ทางกายภาพของวัตถุภายในภาพ แต่ทว่า คุณลักษณะสำคัญของอาคารและความสมบูรณ์ยังคงอยู่ ฉะนั้น เมื่อดำเนินการเสริมข้อมูลเข้ากับแต่ละชุดข้อมูลการฝึกฝน ทำให้ปริมาณภาพเพิ่มขึ้นเป็น 2,065 รูปในแต่ละชุด ทั้งนี้ ตัวอย่างรูปแสดงให้เห็นถึงการประยุกต์ใช้วิธีการเสริมข้อมูลเป็นดังรูปที่ 3.7



รูปที่ 3.7 ตัวอย่างการเสริมข้อมูลด้วยการพลิกภาพแบบแนวดิ่ง แนวนอน การหมุน 90 องศาทวนเข็มนาฬิกา และตามเข็มนาฬิกา

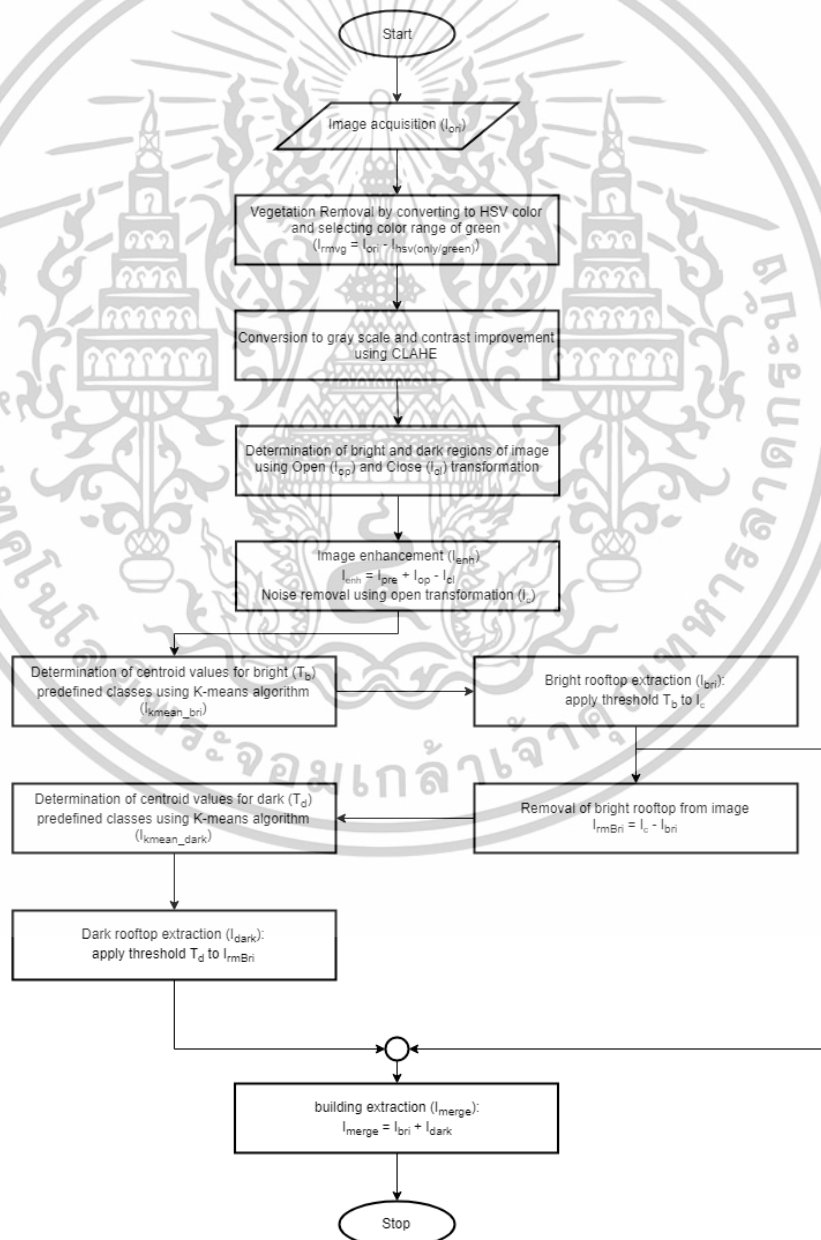
3.4 การออกแบบขั้นตอนการสกัดอาคารจากภาพถ่ายดาวเทียม

การสกัดอาคารจากภาพถ่ายดาวเทียมในงานวิจัยนี้แบ่งการทดลองออกเป็นสองวิธี ได้แก่ วิธีประมวลผลภาพแบบสัญญาณวิทยา โดยวิธีการนี้จะไม่พึ่งพาโมเดลการเรียนรู้เชิงลึก และอีกหนึ่งวิธีคือ เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น เมื่ออนุญาตให้ไปใช้ประโยชน์ด้านการศึกษาไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

วิธีประยุกต์ใช้โมเดลการเรียนรู้เชิงลึก ซึ่งได้จัดเตรียมชุดข้อมูลไว้แล้วเสร็จ รวมถึงจำแนกหมวดหมู่ชุดข้อมูลเพื่อเปรียบเทียบ อย่างไรก็ตาม ทั้งสองวิธีที่กล่าวมาได้รับการออกแบบลำดับขั้นตอนดำเนินการดังต่อไปนี้

3.4.1 วิธีประมวลผลภาพแบบสัณฐานวิทยา

การแปลงลักษณะสัณฐานวิทยา (Morphological Transformation) เป็นเทคนิคที่นิยมนำมาดำเนินการกับรูปร่างของวัตถุภายในภาพขาวดำ (Binary Image) แต่ทั้งนี้ การประยุกต์ใช้เทคนิคการแปลงลักษณะสัณฐานวิทยา และการปรับพารามิเตอร์เป็นเพียงส่วนหนึ่งของกระบวนการทั้งหมด ซึ่งมีขั้นตอนของกระบวนการทั้งหมดดังรูปที่ 3.8

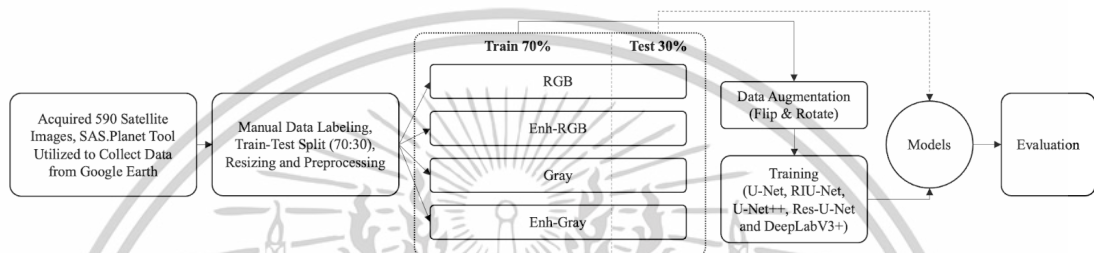


รูปที่ 3.8 ขั้นตอนกระบวนการสกัดอาคารด้วยวิธีสัณฐานวิทยา

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์สำหรับการใช้ในพิธีการวิชาการเท่านั้น เมื่อผู้เผยแพร่เอกสารนี้ไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.4.2 วิธีประยุกต์ใช้โมเดลการเรียนรู้เชิงลึก

การประยุกต์ใช้โมเดลการเรียนรู้เชิงลึกแต่ละโมเดล ผักฝนกับชุดข้อมูลที่กำลังกับ คำตอบและจำแนกหมวดหมู่ไว้ รวมถึงการเสริมและวัดประสิทธิภาพความแม่นยำในการ ทำนาย สามารถอธิบายลำดับขั้นตอนการดำเนินงานได้ตามรูปที่ 3.9 ทั้งนี้ โครงข่ายประสาท เทียมของแต่ละโมเดลได้ถูกนำมาพัฒนา และปรับแต่งให้เข้ากับคุณลักษณะของข้อมูลที่ได้ จัดเตรียม อาทิ แชนแนลสีของชุดข้อมูล ขนาดกว้างคูณยาวของภาพ และการกำหนดไฮเปอร์ พารามิเตอร์พื้นฐาน



รูปที่ 3.9 ขั้นตอนกระบวนการสกัดอาคารด้วยวิธีประยุกต์ใช้โมเดลการเรียนรู้เชิงลึก

ขั้นตอนดำเนินงานทั้งสองวิธีมีความแตกต่างกันอย่างสิ้นเชิง โดยกระบวนการประมวลผลภาพ แบบสัญญาณวิทยาและเทคนิคอื่น ๆ สะท้อนถึงความพยายามสกัดคุณลักษณะอาคารจากภาพถ่าย ดาวเทียม ซึ่งดำเนินการด้วยชุดคำสั่งที่ไม่ซับซ้อนและทรัพยากรประมวลผลน้อยกว่า แน่นนอนว่าหาก ประสบผลสำเร็จ ขั้นตอนวิธีดังกล่าวจะช่วยลดต้นทุนทั้งในด้านเวลาและค่าใช้จ่ายสำหรับการ ประมวลผล ทว่ากลับกัน วิธีการประยุกต์ใช้โมเดลการเรียนรู้เชิงลึก สะท้อนให้เห็นถึงความสามารถ ของอัลกอริทึมที่อาจหาคุณลักษณะ และปัจจัยได้หลายมิติมากกว่าแค่เพียงชุดคำสั่งการประมวลผล ภาพ กระนั้น เป็นที่ยืนยันได้ว่าใช้ต้นทุนสูงกว่าในทุกด้านเมื่อเทียบกับวิธีการแรก อย่างไรก็ตาม สำหรับผลการวิจัยและการอภิปรายผลการดำเนินงานจะถูกกล่าวถึงในบทถัดไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4

ผลการวิจัยและการอภิปรายผล

การนำเสนอผลลัพธ์จากการทดลอง การวิเคราะห์เปรียบเทียบ และการอภิปรายเพื่อประเมินประสิทธิภาพของแนวทางที่นำเสนอ สำหรับการสกัดอาคารจากภาพถ่ายดาวเทียมความละเอียดสูงจะถูกกล่าวถึงในบทนี้ โดยเบื้องต้น ได้ทำการทดลองสกัดอาคารด้วยเทคนิคสัณฐานวิทยา ซึ่งอาศัยเทคนิคการประมวลผลรูปภาพแบบไม่ใช้ปัญญาประดิษฐ์ อันเป็นจุดเริ่มต้นก่อนก้าวเข้าสู่การประยุกต์ใช้โมเดลการเรียนรู้เชิงลึก ทั้งนี้ การวิเคราะห์และเปรียบเทียบอิงตามเกณฑ์การประเมิน ได้แก่ ดัชนีชี้วัดความคล้ายคลึง Intersection over Union (IoU) ดัชนีชี้วัดความคล้ายคลึง Dice Similarity Coefficient (DSC) และดัชนีชี้วัดความสับสน (Confusion Matrix) นอกจากนี้ ยังระบุปัจจัยด้านเวลาที่ใช้ในการเรียนรู้ของแต่ละโมเดล รวมทั้งอภิปรายจุดแข็งและจุดอ่อนตามการทดลอง

4.1 การแบ่งส่วนอาคารด้วยเทคนิคสัณฐานวิทยา

กระบวนการสกัดอาคารจากภาพถ่ายดาวเทียมด้วยวิธีสัณฐานวิทยา มีขั้นตอนตั้งแต่นำรูปเข้าสู่การประมวลผล (ขั้นตอนที่ 1: I_{on}) ตามต่อด้วยกระบวนการและผลลัพธ์ในแต่ละขั้นตอนดังต่อไปนี้

4.1.1 การลบพื้นที่ป่าด้วยกระบวนการแปลงสีภาพ

ภาพถ่ายดาวเทียมที่นำเข้ามาใช้ในการทดลองมีคุณลักษณะ คือ เป็นภาพสี (RGB) ฉะนั้น ส่วนที่เป็นพื้นที่ป่าไม้โดยส่วนมากปรากฏเป็นสีเขียว ผู้วิจัยจึงได้ใช้เทคนิคการแปลงสีภาพจากระบบสี RGB ไปเป็น HSV (Hue, Saturation, Value) [66] ตามสมการ (4.1) และกำหนดค่าคุณสมบัติของช่วงสีที่ครอบคลุมสีเขียว เพื่อนำมาแบ่งส่วนพื้นที่ป่าไม้ออกจากภาพ นอกจากนี้ เพื่อให้มีประสิทธิภาพในการจัดการกับค่าสี จึงได้มีการประยุกต์ใช้เทคนิคการเบลอภาพแบบเกาส์เซียน [67, 68] ที่ Kernel ขนาด 15×15 (4.2) ซึ่งมีส่วนช่วยให้สีที่ใกล้เคียงกันรวมเป็นกลุ่มสีเดียวกัน

$$\begin{aligned} V &\leftarrow \max(R, G, B) \\ S &\leftarrow \begin{cases} V - \min(R, G, B), & \text{if } V \neq 0 \\ 0, & \text{otherwise} \end{cases} \\ H &\leftarrow \begin{cases} \frac{60(G - B)}{V - \min(R, G, B)}, & \text{if } V = R \\ 120 + \frac{60(B - R)}{V - \min(R, G, B)}, & \text{if } V = G \\ 240 + \frac{60(R - G)}{V - \min(R, G, B)}, & \text{if } V = b \end{cases} \end{aligned} \quad (4.1)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ถ้าค่า H น้อยกว่า 0 แล้วค่า $H \leftarrow H+360$ โดยที่ค่าผลลัพธ์จะอยู่ในช่วง $0 \leq V \leq 1$, $0 \leq S \leq 1$ และ $0 \leq H \leq 360$

$$G(x, y; \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} = A e^{-\frac{(x-\mu_x)^2}{2\sigma_x^2} - \frac{(y-\mu_y)^2}{2\sigma_y^2}} \quad (4.2)$$

จากสมการ (4.2) μ คือค่ากลาง (mean) และ σ^2 คือความแปรปรวนตามแต่ละตัวแปร x และ y ; Kernel ดำเนินการกับภาพแบบ 2 มิติ สามารถคำนวณได้จากสมการ (4.3)

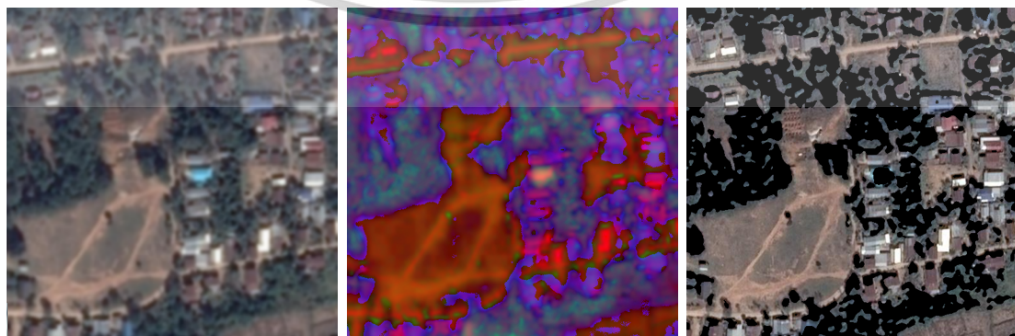
$$K(n \times n) = \frac{1}{2^n} [x_0, x_1, \dots, x_n] \times \frac{1}{2^n} \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{bmatrix} \quad (4.3)$$

จากสมการ (4.3) การกำหนดขนาดของ Kernel (K) ที่ n ใด ๆ ด้วยวิธีการเบลอภาพแบบเกาส์เซียน ซึ่งสามารถนำเสนอได้ดังรูปที่ 4.1 ด้วยขนาด $K(5 \times 5)$; $n = 5$

$$K \equiv \frac{1}{256} \begin{bmatrix} 1 & 4 & 6 & 4 & 1 \\ 4 & 16 & 24 & 16 & 4 \\ 6 & 24 & 36 & 24 & 6 \\ 4 & 16 & 24 & 16 & 4 \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix} = \frac{1}{16} [1 \ 4 \ 6 \ 4 \ 1] * \frac{1}{16} \begin{bmatrix} 1 \\ 4 \\ 6 \\ 4 \\ 1 \end{bmatrix}$$

รูปที่ 4.1 ตัวอย่าง Gaussian blur (filter) ซึ่งใช้ Kernel ขนาด 5×5

ผลลัพธ์จากการประยุกต์ใช้วิธีการเบลอภาพแบบเกาส์เซียน การแปลงค่าสี และการลบค่าสีเขียวออกจากช่วงที่กำหนดทำให้ได้ผลลัพธ์ดังรูปที่ 4.2



รูปที่ 4.2 ผลลัพธ์จากการดำเนินการใช้ Gaussian blur (ซ้าย) การแปลงค่าสีเป็น HSV (กลาง) และ

การลบพื้นที่สีเขียว (ขวา)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในเทคนิคการลบพื้นทีสีเขียวจากภาพทำให้เกิดผลข้างเคียง 2 ประการ คือ 1) อาคารที่มีหลังคาสีเขียวถูกลบออกไปบางส่วน 2) พื้นทีสีเขียวที่ปกคลุมหลังคาของอาคารถูกลบออก ส่งผลให้หลังคาส่วนที่ถูกปกคลุมถูกลบออกไปด้วย ดังนั้น ผู้วิจัยจึงได้นำเทคนิคลักษณะสัมพันธ์วิทยาด้วยตัวดำเนินการแบบ Dilation ซึ่งกำหนดขนาดตัวดำเนินการแบบสี่เหลี่ยมขนาด 5x5 (Rectangular Structuring Element) ดำเนินการกับภาพผลลัพธ์ล่าสุด (รูปที่ 4.2 (ขวา)) และได้ผลลัพธ์ (ขั้นตอนที่ 2 : I_{rmvg}) ดังรูปที่ 4.3 (ซ้าย)



รูปที่ 4.3 ภาพถ่ายดาวเทียมและหลังดำเนินการใช้เทคนิคลักษณะสัมพันธ์วิทยาแบบ Dilation

การใช้เทคนิคลักษณะสัมพันธ์วิทยาส่งผลให้ผลข้างเคียงประการแรกถูกกู้คืนกลับมาได้เพียงบางส่วน และในประการที่สอง พื้นที่ซึ่งถูกสีเขียวปกคลุมได้ถูกกู้คืนกลับมาอย่างสมบูรณ์ ทว่าการใช้เทคนิคดังกล่าวส่งผลให้สูญเสียรายละเอียดของภาพ แต่ในขณะเดียวกันอาคารที่มีหลังคาสว่างจะปรากฏขึ้นอย่างชัดเจนทั้งขนาดและรูปทรง ดังนั้นขั้นตอนต่อไปจะเป็นการจัดเตรียมข้อมูลภาพเพื่อแบ่งส่วนอาคารที่สว่างออกจากภาพ

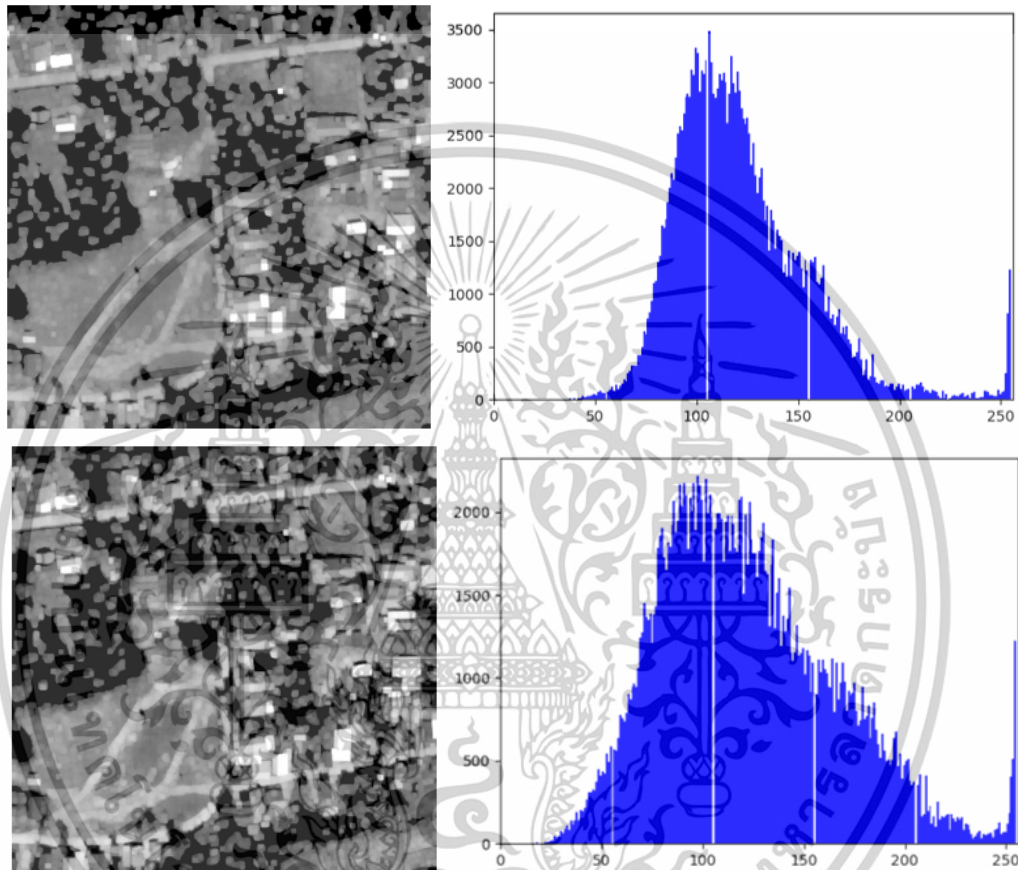
4.1.2 การจัดเตรียมข้อมูลภาพ

ภาพผลลัพธ์ (I_{rmvg}) มีคุณลักษณะเด่นที่ตำแหน่งของอาคารซึ่งมีความสว่าง ฉะนั้นการจัดเตรียมข้อมูลภาพจะดำเนินการแปลงจากภาพสี (RGB) ไปเป็นภาพระดับสีเทา (Grayscale) (สมการ 4.4) เพื่อลดมิติของข้อมูลและแสดงถึงความสว่างของสีในแต่ละพื้นที่ในภาพ จากนั้นจึงประยุกต์ใช้การปรับค่าสมดุการกระจายตัวของช่วงสีเทา (Contrast)

$$RGB \text{ to Gray: } Y \leftarrow 0.299 \cdot R + 0.587 \cdot G + 0.114 \cdot B \quad (4.4)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพผลลัพธ์จากกระบวนการแปลงค่าสีเป็น Grayscale ถูกนำเข้าสู่กระบวนการ Contrast Limited Adaptive Histogram Equalization (CLAHE) ได้ผลลัพธ์ดังรูปที่ 4.4 โดยจะเปรียบเทียบการกระจายของค่าสีด้วย Histogram ระหว่างก่อนและหลังการดำเนินการ ทั้งนี้การแสดงผลจะไม่นำจำนวน Pixel ที่มีค่า 0 ถึง 1 มาแสดงผล เนื่องจากมีปริมาณมากกว่าช่วงค่าสีอื่น ๆ ส่งผลให้ไม่เห็นถึงการกระจายตัวที่ชัดเจน



รูปที่ 4.4 ภาพถ่ายดาวเทียมกระจายตัวของค่าสีระหว่างก่อนและหลังประยุกต์ใช้เทคนิค CLAHE

การประยุกต์ใช้เทคนิค CLAHE สามารถช่วยปรับสมดุลการกระจายตัวของค่าสีได้อย่างชัดเจน ส่งผลให้ภาพได้รายละเอียดค่าสีที่ชัดเจนมากขึ้นโดยเฉพาะขอบและเงาของวัตถุ อีกทั้งยังช่วยจัดกลุ่มสีของวัตถุขึ้นเดียวกันให้มีความแตกต่างระหว่างค่าสีลดน้อยลง (ขั้นตอนที่ 3 : I_{pre})

4.1.3 การระบุส่วนที่สว่างและมีดของพื้นที่ในภาพด้วยเทคนิคสัณฐานวิทยา

ภาพผลลัพธ์ (I_{pre}) จากขั้นตอนที่ 4.1.2 มีคุณลักษณะความต่างระหว่างค่าสีที่สามารถนำมาระบุส่วนที่สว่างและมีดของอาคารได้ ดังนั้นแล้วจึงใช้เทคนิคลักษณะสัณฐานวิทยาด้วยตัวดำเนินการแบบ Open และ Closing เข้ากับภาพ (I_{pre}) และได้ผลลัพธ์ดังภาพที่ 4.5

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.5 ผลลัพธ์จากการใช้ตัวดำเนินการ Open (ซ้าย) และ close (ขวา)

งานวิจัยการสกัดอาคารจากภาพถ่ายดาวเทียมคุณภาพสูง ด้วยการใช้สัณฐานวิทยาทางคณิตศาสตร์แบบอัตโนมัติ ได้นำเสนอวิธีการปรับปรุงคุณภาพข้อมูลด้วยวิธีการนำภาพที่ผ่านเทคนิคปรับสมดุลการกระจายตัวของค่าสี (I_{pre}) ผสมกับภาพส่วนที่สว่าง (I_{bri}) และห้ก้ล้าส่วนที่มีมืด (I_{dark}) ตามสมการดำเนินการที่ดังนี้

$$I_{enh} = I_{pre} + I_{bri} - I_{dark} \quad (4.5)$$

การประยุกต์ใช้เทคนิคนี้ จะส่งผลให้การแบ่งแยกขอบของวัตถุที่มีความต่างระหว่างสีชัดเจนมากขึ้น และจำเป็นต้องใช้เทคนิคลักษณะสัณฐานวิทยาด้วยตัวดำเนินการ Open เพื่อกลบส่วนสีขาวซึ่งเกิดจากการหักล้างกันในกระบวนการก่อนหน้า หรือกล่าวได้ว่า การใช้เทคนิคลักษณะสัณฐานวิทยาแบบ Open มีคุณสมบัติการกลบหลุม (Fill-hole) ในภาพ โดยที่หลุม คือ กลุ่ม Pixel สีขาวลักษณะเป็นจุดหรือวงกลมขนาดเล็กซึ่งถูกล้อมรอบไปด้วยกลุ่ม Pixel สีดำ สำหรับผลลัพธ์ในการปรับปรุงคุณภาพ (ขั้นตอนที่ 4 : 1) เป็นดังรูปที่ 4.6



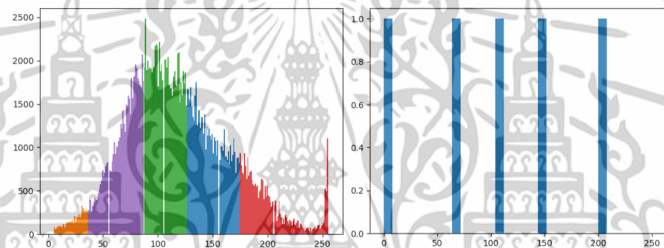
รูปที่ 4.6 ผลลัพธ์ก่อน (ซ้าย) และหลังการปรับปรุงคุณภาพโดยใช้วิธีตามสมการที่ 6 (กลาง) และการประยุกต์ลักษณะสัณฐานวิทยาแบบ open (ขวา)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.1.4 การลบพื้นที่ป่าด้วยกระบวนการแปลงสีภาพ

K-means คือ วิธีการแบ่งกลุ่มข้อมูล (Clustering) แบบการเรียนรู้โดยไม่อาศัยชุดคำตอบ (Unsupervised Learning) โดยทั่วไปนิยมใช้กับกลุ่มข้อมูลทางสถิติ แต่ทว่าข้อมูลของรูปภาพซึ่งถูกจัดเก็บในเป็นค่าตัวเลขของค่าสีในแต่ละ Pixel จึงสามารถนำเทคนิค K-means มาประยุกต์ใช้ร่วมกันได้กับทั้งภาพสี RGB และภาพระดับสีเทา (Grayscale)

กระบวนการพื้นฐานของ K-means และการจัดกลุ่มค่าสี เริ่มจากการกำหนดกลุ่มที่ต้องการจะแบ่งแทนด้วย K กลุ่ม กำหนดรูปแบบจุดกึ่งกลางเริ่มต้น และฟังก์ชันคำนวณระยะห่าง (Distance function) โดยในการทดลองนี้ได้กำหนดพารามิเตอร์ทั้ง 3 ได้แก่ กำหนด K = 5 กำหนดรูปแบบจุดกึ่งกลางเริ่มต้นแบบสุ่ม (Random centroids) และกำหนดฟังก์ชันคำนวณระยะห่างด้วยวิธีการคำนวณแบบ Euclidean จากนั้นจึงนำไปประยุกต์ใช้กับภาพตามขั้นตอนที่ 5 และ 8 (ประยุกต์ใช้เทคนิค K-means กับภาพ (I_c))



รูปที่ 4.7 กลุ่มสีซึ่งถูกจัดเรียงด้วย K-means (ซ้าย) และค่ากลางของแต่ละกลุ่มสีทั้ง 5 กลุ่ม(ขวา)

ผลลัพธ์จากการใช้เทคนิค K-means เป็นดังรูปที่ 4.7 โดยแกน X แสดงถึงค่าสีช่วงสีตั้งแต่ 0 (ดำ) ถึง 256 (ขาว) ซึ่งเป็นรูปแบบการจัดเก็บค่าชนิดภาพสี 8 บิต และแกน Y แสดงถึงจำนวน Pixel นอกจากนี้ค่าสีที่ปรากฏในภาพทางซ้ายแสดงถึงกลุ่มของสีซึ่งถูกแบ่งกลุ่ม เช่น สีแดงคือกลุ่มของค่าสีตั้งแต่สิ้นสุดช่วงกลุ่มสีฟ้าจนถึงค่า 255 และภาพทางขวาแสดงค่ากลาง (Centroids) ของแต่ละกลุ่มสีโดยมีลำดับดังนี้ 4, 68, 105, 147 และ 201 สำหรับภาพผลลัพธ์ (I_{kmean_br}) เป็นดังรูปที่ 4.8



รูปที่ 4.8 ผลลัพธ์ก่อนและหลัง (ขวา) ประยุกต์ใช้ K-means

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 4.8 (ขวา) จะสังเกตเห็นถึงความสว่างของสีภายในภาพที่ใกล้เคียงกันจะถูก รวมกันแล้วจึงแบ่งออกเป็น 5 กลุ่ม ตามที่กล่าวมาแล้วข้างต้น ในขั้นตอนต่อไป คือ การ ดำเนินการประยุกต์ Thresholding แบบ Binary เข้ากับภาพจากกระบวนการก่อนหน้า (I_{kmean_bri}) โดยพิจารณาจากค่ากลางของกลุ่มสีทั้ง 5 กลุ่ม

4.1.5 การประยุกต์ใช้ Threshold เพื่อแบ่งส่วนอาคาร

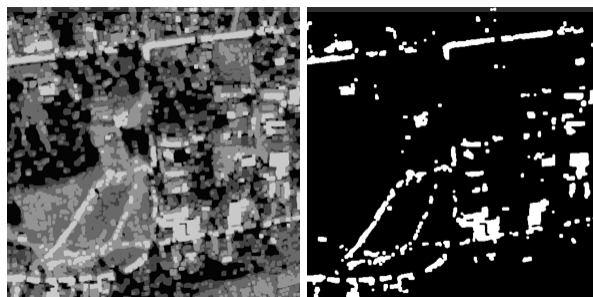
Image Thresholding คือ หนึ่งในเทคนิคการแบ่งส่วนที่สนใจจากรูปภาพ (Image Segmentation) เพื่อนำไปวิเคราะห์ โดยผู้วิจัยได้นำหนึ่งในประเภทของการทำ Thresholding มาใช้ คือ Threshold Binary [69] (4.6)

$$dst(x,y) = f(x) = \begin{cases} maxVal, & \text{if } src(x,y) > thresh \\ 0, & \text{otherwise} \end{cases} \quad (4.6)$$

หลักการทำงานของ Threshold binary คือ เมื่อนำภาพระดับสีเทามาดำเนินการ จะมีการกำหนดค่า Threshold จากนั้นจึงใช้การเปรียบเทียบการค่าสีแต่ละ Pixel ของภาพที่ นำเข้ามา หากว่าค่าสีของ Pixel เกินกว่าค่า Threshold ที่กำหนด ค่าของ Pixel ที่นำมา เปรียบเทียบจะมีค่าเท่ากับ 1 ในขณะที่ถ้าหากว่าค่าสีของ Pixel ต่ำกว่าค่า Threshold ที่ กำหนดก็จะมีค่าเท่ากับ 0

การทดลองนี้ผู้วิจัยได้นำค่าที่ 4 ของการประยุกต์ใช้ K-means มากำหนดเป็นค่า Threshold คือ 147 (T_b) ดังนั้นหากนำรูปภาพอื่น ๆ เข้ามาค่ากลางที่ได้จากเทคนิค K-means จะแตกต่างกันออกไป กระบวนการรวมของเทคนิคประยุกต์นี้เรียกว่า Dynamic - Thresholding กล่าวคือ ค่า Threshold จะถูกกำหนดให้ขึ้นตามความเหมาะสมแต่ละภาพ โดยใช้ K-means เป็นตัวช่วยในการหาค่า Threshold

ภาพผลลัพธ์หลังการใช้เทคนิค Threshold Binary (ขั้นตอนที่ 6: I_{bri}) เป็นดังภาพที่ 4.9 และในขั้นตอนต่อไปจะเป็นการลบส่วนที่ถูกแบ่งออกจากขั้นตอนที่ 7 แล้วจึงนำไปเข้า ขั้นตอนที่ 8 และ 9 เพื่อแบ่งส่วนอาคารที่มีความสว่างน้อยกว่า (ส่วนที่มีด) ในรูปที่ 4.10



รูปที่ 4.9 ผลลัพธ์ก่อนและหลัง (ขวา) การประยุกต์ใช้ Binary Threshold

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น เมื่อนำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



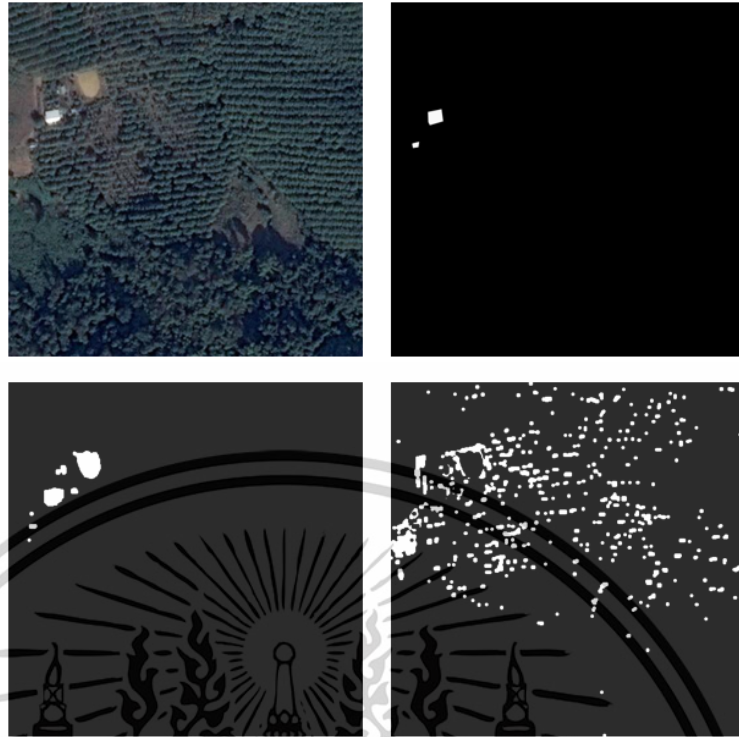
รูปที่ 4.10 ภาพในกระบวนการหาผลลัพธ์จากการแบ่งส่วนอาคารที่มีแสงสว่างน้อย

ผลลัพธ์จากการแบ่งส่วนอาคารที่สว่าง (I_{br}) รูปที่ 4.9 (ขวา) และการแบ่งส่วนอาคารที่สว่างน้อย (I_{dark}) รูปที่ 4.10 (ขวา) พบว่ามีหลายส่วนที่ไม่มีคุณลักษณะของอาคารเลยแม้แต่น้อยเมื่อพิจารณาด้วยตาเปล่า ดังนั้นผู้วิจัยจึงทดลองนำภาพผลลัพธ์ทั้งสองมาผสมผสานรวมกัน จึงได้ผลลัพธ์ดังรูปที่ 4.11 ซึ่งเปรียบเทียบกับภาพต้นฉบับที่นำเข้าสู่กระบวนการ (I_{br})



รูปที่ 4.11 การเปรียบเทียบภาพระหว่างต้นฉบับ ชุดคำตอบ และหลังการแบ่งส่วนด้วยลักษณะสัญญาณทางวิทยา

ภาพผลลัพธ์สุดท้ายที่ได้จากการแบ่งส่วนด้วยลักษณะสัญญาณทางวิทยา แสดงให้เห็นว่าส่วนที่ทำการแบ่งส่วนออกมาจากภาพมีทั้งวัตถุซึ่งเป็นอาคาร ถนน พื้นที่ดิน และจุดเล็ก ๆ เปรียบเสมือนกับส่วนผิดพลาดขนาดเล็ก (Noise) และถ้าหากพิจารณาและประเมินคุณภาพด้วยตาเปล่าโดยไม่ใช้ตัวแปรชี้วัดก็สามารถสรุปได้ว่าวิธีการที่ใช้ยังไม่มีประสิทธิภาพในการจะระบุทั้งตำแหน่ง ขนาด และรูปร่างของอาคาร แต่เมื่อได้ดำเนินการกับชุดข้อมูลอื่นที่นำมาทดลอง พบว่าวิธีการนี้สามารถแบ่งส่วนอาคารที่สว่างและตั้งอยู่กลางป่าซึ่งมีสีเขียวปกคลุมทั่วบริเวณภาพได้ค่อนข้างมีประสิทธิภาพ ดังรูปภาพที่ 4.12



รูปที่ 4.12 การเปรียบเทียบภาพระหว่างต้นฉบับและคำตอบ (บน) กับผลลัพธ์การแบ่งส่วนอาคารที่สว่างและมีด (ล่าง)

ผลลัพธ์ที่ได้จากตัวอย่างชุดข้อมูลหลังผ่านกระบวนการแบ่งส่วน พบว่าตำแหน่งของอาคารได้รับการแบ่งส่วนที่ชัดเจนในขั้นตอนการแบ่งส่วนอาคารที่สว่าง จากภาพประกอบตัวอย่างผลการทดลองของข้อมูลภาพทั้งสองชุดนี้ สามารถสรุปได้ว่าวิธีการแบ่งส่วนที่ใช้ ไม่สามารถดำเนินการกับภาพที่มีความซับซ้อนของวัตถุจำนวนมาก และไม่สามารถดำเนินการกับภาพซึ่งมีสภาพแวดล้อมที่แตกต่างกันอย่างมีประสิทธิภาพ นอกจากนี้กระบวนการอีกส่วนที่สำคัญ คือ ต้องใช้เทคนิคในการลบส่วนที่ถูกระบุอย่างผิดพลาดหรือการแบ่งส่วนที่ผิดพลาด เช่น จุดสีขาวขนาดเล็กที่มีลักษณะเป็นวงกลม ถนนที่มีความยาวซึ่งมีอัตราส่วนตามแนวแกนที่ไม่ตรงกับคุณลักษณะของบ้าน นอกจากนี้ยังต้องดำเนินการแยกกันระหว่างภาพส่วนที่เป็นอาคารสว่างและส่วนที่เป็นอาคารมืด จากนั้นจึงนำผลลัพธ์ทั้งสองมาผสมผสานกัน

แนวทางการพัฒนาและการปรับปรุงประสิทธิภาพที่ได้กล่าวถึงในข้อสรุปที่ผ่านมา มีส่วนของเทคนิคซึ่งจำเป็นต้องใช้กระบวนการ Postprocessing กระนั้น ด้วยความไม่แน่นอนรวมถึงประกันความเสี่ยงในการวิจัย เนื่องด้วยผลลัพธ์ไม่เป็นไปอย่างที่คาดหวัง ผู้วิจัยจึงได้ตัดสินใจใช้วิธีการสกัดอาคารด้วยโครงข่ายประสาทเทียมเสียแทน ด้วยโมเดลยูเน็ตและโมเดลการเรียนรู้เชิงลึกอื่น ๆ เพื่อเปรียบเทียบผลลัพธ์ และตัดสินใจว่าควรนำกระบวนการสกัดอาคารแบบใดที่ให้ประสิทธิภาพสูงสุดในกรณีศึกษา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2 การสกัดอาคารด้วยโมเดลการเรียนรู้เชิงลึก

กระบวนการสกัดอาคารด้วยโมเดลการเรียนรู้เชิงลึกทุกขั้นตอน ตั้งแต่ การเตรียมข้อมูล การพัฒนาโครงข่ายประสาทเทียม ตลอดจนประเมินผลความคล้ยคลึงด้วยดัชนีชี้วัด ถูกพัฒนาขึ้นด้วยภาษา Python เวอร์ชัน 3 (Python 3.10) ทั้งนี้ หนึ่งในไลบรารีสำคัญหรือชุดคำสั่งสำเร็จรูปที่พร้อมด้วยฟังก์ชันนำไปพัฒนาโครงข่ายประสาทเทียมคือ Tensonflow [70] และใช้ทรัพยากรการประมวลผลของ Google Colab [71] นอกจากนี้ การฝึกฝนโมเดลการเรียนรู้เชิงลึกได้ใช้วีดีโอกราฟิกในการประมวลผล (GPU) คือรุ่น NVIDIA Tesla V100 SXM2 16 GB GPU ซึ่งช่วยลดระยะเวลาในการฝึกฝนของโมเดลได้ดีกว่าเมื่อเทียบกับการใช้ CPU ประมวลผล อย่างไรก็ตาม สิ่งสำคัญคือชุดข้อมูลชุดคำสั่ง และโมเดลที่ได้จากการฝึกฝน ดังนั้น ชุดข้อมูลและไฟล์จึงถูกจัดเก็บไว้บน Google Drive [72] เนื่องจากสะดวกต่อการนำไปใช้ร่วมกันระหว่างเครื่องมือที่ได้กล่าวมา

โมเดลการเรียนรู้เชิงลึก ได้รับการฝึกฝนด้วยการกำหนดไฮเปอร์พารามิเตอร์ของแต่ละเลเยอร์ภายใต้แต่ละโมเดล ให้สอดคล้องกับคุณลักษณะของชุดข้อมูล โดยการปรับแต่งเบื้องต้นเป็นมีดังนี้ การแบ่งชุดข้อมูลฝึกฝนและทดสอบด้วยอัตราส่วนร้อยละ 70 ต่อ 30 ตามลำดับ โมเดลจะดำเนินเรียนรู้พร้อมทั้งใช้ชุดข้อมูลทดสอบ (Validation) จำนวน 250 รอบ (250 epochs) ด้วยขนาดปริมาณข้อมูลภาพแบบแบ่งย่อยทีละ 8 (Batch = 8) กล่าวคือ ในการฝึกฝน 1 รอบ โมเดลจะนำภาพจำนวน 8 รูป เข้าสู่เลเยอร์เพื่อสกัดคุณลักษณะ และจะทำจนกว่าจะครบทุกรูปที่มีเป็นการครบ 1 รอบ epoch นอกจากนี้ เพื่อให้หลีกเลี่ยงการฝึกฝนโมเดลดำเนินการปรับค่าน้ำหนัก (Weight) และ (Bias) อย่างมีประสิทธิภาพ และหลีกเลี่ยงไม่ให้ใช้เวลานานเกินไปจนกว่าค่า Loss จะต่ำสุดที่เป็นไปได้ ดังนั้น ตัวปรับปรุงประสิทธิภาพ (Optimizer) ชื่อว่า Root Mean Square Propagation (RMSProp) [73] จึงถูกพิจารณานำมาใช้ โดยมีลักษณะเด่นคือการใช้สมการคำนวณ Exponentially weighted moving average (EMA) ของอนุพันธ์กำลังสองมาปรับพารามิเตอร์ (Parameter) ทั้งค่า Weight และ Bias ร่วมกับการกำหนดอัตราการเรียนรู้ (Learning Rate) ที่ 0.001 ซึ่งใช้ถ่วงน้ำหนักสมการปรับปรุงอนุพันธ์ของ Optimizer ก่อนลบออกจากค่า Parameter และเพื่อวัดประสิทธิภาพของโมเดล ดัชนีชี้วัดความคล้ายคลึง IoU และ DSC จึงถูกนำมากำหนดไว้ในเมตริกชี้วัดผลระหว่างการฝึกฝน อย่างไรก็ตาม การออกแบบวิธีดำเนินงานวิจัยก่อให้เกิดการเปรียบเทียบโมเดลทั้งสิ้น 48 แบบ โดยอิงตามชุดข้อมูลและดัชนีชี้วัดเพื่อเปรียบเทียบ ดังตารางที่ 4.1

เนื่องด้วยโมเดล Res-U-Net และ DeepLabV3+ ใช้มีการประยุกต์ใช้ Backbone ในโครงสร้าง และมีการใช้ประโยชน์จากการฝึกฝนล่วงหน้า (Pretrained weight) ด้วยชุดข้อมูล ImageNet ซึ่งออกแบบมาสำหรับภาพสี RGB 3 แชนแนล ส่งผลให้การทดลองด้วยชุดข้อมูลภาพสีเทา (Gray) และสีเทาปรับปรุง (Enh-Gray) ที่มีแชนแนลสีเพียง 1 แชนแนล ไม่สามารถนำมาทดลองกับโมเดล 2 โมเดลดังกล่าวได้ เหตุเพราะแต่ละพิกเซลในภาพเหล่านี้ครอบคลุมช่วงความเข้มของระดับสีเทาที่ 0-255 ซึ่งต่างจากภาพสี RGB ที่พิกเซลครอบคลุมค่าสีมากกว่า (0-255, 0-255, 0-255) เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ด้วยเหตุนี้ ชุดข้อมูลภาพสีเทา (Gray) และสีเทาปรับปรุง (Enh-Gray) แล้วจึงไม่สามารถใช้สำหรับการฝึกโมเดลเฉพาะเหล่านี้ได้ ดังนั้น โมเดล Res-U-Net และ DeepLabV3+ จึงได้รับการฝึกฝนโดยชุดข้อมูลภาพสี (RGB) และภาพสีปรับปรุง (Enh-RGB) เท่านั้น

ตารางที่ 4.1 เปรียบเทียบการวัดประสิทธิภาพของแต่ละโมเดลโดยอิงคะแนน F1 จากมากไปน้อย

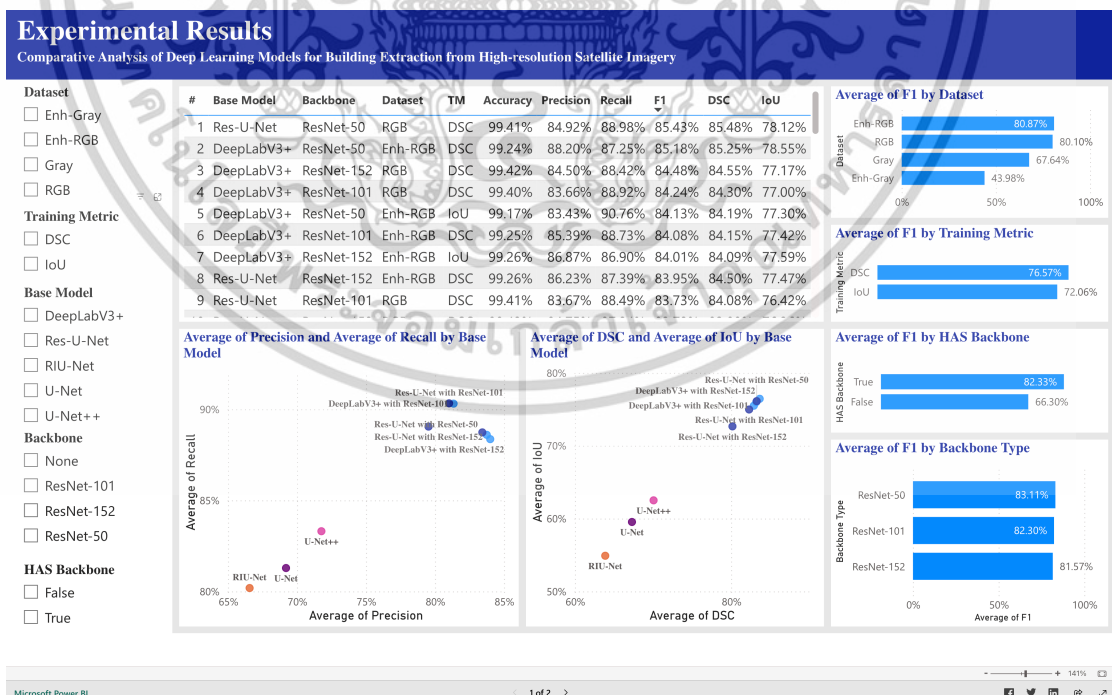
Model	Dataset	Training Metric	Accuracy	Precision	Recall	F1	DSC	IoU
RUN w. RN-50	RGB	DSC	0.9941	0.8492	0.8898	0.8543	0.8548	0.7812
DLv3+ w. RN-50	Enh-RGB	DSC	0.9924	0.8820	0.8725	0.8518	0.8525	0.7855
DLv3+ w. RN-152	RGB	DSC	0.9942	0.8450	0.8842	0.8448	0.8455	0.7717
DLv3+ w. RN-101	RGB	DSC	0.9940	0.8366	0.8892	0.8424	0.8430	0.7700
DLv3+ w. RN-50	Enh-RGB	IoU	0.9917	0.8343	0.9076	0.8413	0.8419	0.7730
DLv3+ w. RN-101	Enh-RGB	DSC	0.9925	0.8539	0.8873	0.8408	0.8415	0.7742
DLv3+ w. RN-152	Enh-RGB	IoU	0.9926	0.8687	0.8690	0.8401	0.8409	0.7759
RUN w. RN-152	Enh-RGB	DSC	0.9926	0.8623	0.8739	0.8395	0.8450	0.7747
RUN w. RN-101	RGB	DSC	0.9941	0.8367	0.8849	0.8373	0.8408	0.7642
RUN w. RN-152	RGB	DSC	0.9942	0.8475	0.8724	0.8372	0.8389	0.7636
DLv3+ w. RN-152	Enh-RGB	DSC	0.9924	0.8600	0.8693	0.8355	0.8366	0.7705
RUN w. RN-152	Enh-RGB	IoU	0.9920	0.8429	0.8942	0.8354	0.8374	0.7691
RUN w. RN-101	Enh-RGB	IoU	0.9923	0.8310	0.9046	0.8336	0.8344	0.7680
RUN w. RN-101	Enh-RGB	DSC	0.9923	0.8310	0.9046	0.8336	0.8344	0.7680
RUN w. RN-50	Enh-RGB	DSC	0.9926	0.8583	0.8738	0.8323	0.8330	0.7687
U-Net++	RGB	DSC	0.9934	0.8208	0.8828	0.8277	0.8286	0.7515
DLv3+ w. RN-50	RGB	DSC	0.9941	0.8426	0.8628	0.8230	0.8248	0.7500
RUN w. RN-50	Enh-RGB	IoU	0.9927	0.8517	0.8705	0.8222	0.8267	0.7588
DLv3+ w. RN-101	Enh-RGB	IoU	0.9918	0.8026	0.9167	0.8201	0.8219	0.7516
U-Net++	Enh-RGB	IoU	0.9920	0.8380	0.8820	0.8183	0.8196	0.7520
DLv3+ w. RN-152	RGB	IoU	0.9935	0.7863	0.9121	0.8164	0.8229	0.7400
RUN w. RN-50	RGB	IoU	0.9925	0.7779	0.9153	0.8153	0.8160	0.7364
U-Net	RGB	IoU	0.9929	0.7979	0.8865	0.8143	0.8152	0.7361
U-Net	RGB	DSC	0.9933	0.8259	0.8595	0.8136	0.8147	0.7382
DLv3+ w. RN-50	RGB	IoU	0.9935	0.7910	0.9012	0.8085	0.8095	0.7321
U-Net++	Enh-RGB	DSC	0.9920	0.8320	0.8733	0.8050	0.8069	0.7375
DLv3+ w. RN-101	RGB	IoU	0.9932	0.7607	0.9194	0.8009	0.8096	0.7243
RIU-Net	RGB	DSC	0.9931	0.8374	0.8113	0.7956	0.8014	0.7170
U-Net	Enh-RGB	DSC	0.9914	0.7892	0.8766	0.7787	0.7820	0.7093
U-Net	Enh-RGB	IoU	0.9914	0.7892	0.8766	0.7787	0.7820	0.7093
U-Net++	RGB	IoU	0.9925	0.7386	0.9089	0.7779	0.7797	0.6983
RUN w. RN-101	RGB	IoU	0.9927	0.7420	0.9191	0.7756	0.7822	0.6990
U-Net++	Gray	DSC	0.9922	0.7730	0.7990	0.7373	0.7405	0.6628
RIU-Net	Enh-RGB	DSC	0.9907	0.7136	0.8590	0.7048	0.7206	0.6322
U-Net++	Gray	IoU	0.9883	0.6729	0.8876	0.7011	0.7032	0.6240
RIU-Net	Gray	DSC	0.9923	0.7805	0.7668	0.7010	0.7102	0.6300
RUN w. RN-152	RGB	IoU	0.9918	0.6284	0.9217	0.6770	0.6851	0.5992
U-Net	Gray	DSC	0.9911	0.7163	0.7875	0.6732	0.6799	0.5960
U-Net	Gray	IoU	0.9908	0.6659	0.7978	0.6569	0.6634	0.5788
RIU-Net	RGB	IoU	0.9908	0.6135	0.9009	0.6556	0.6695	0.5732
RIU-Net	Enh-RGB	IoU	0.9908	0.6592	0.8562	0.6450	0.6796	0.5732
RIU-Net	Gray	IoU	0.9904	0.5885	0.8278	0.5889	0.6060	0.5121
RIU-Net	Enh-Gray	DSC	0.9874	0.7270	0.6493	0.5193	0.5393	0.4573
U-Net++	Enh-Gray	DSC	0.9870	0.6017	0.6827	0.4907	0.4960	0.4230
U-Net	Enh-Gray	DSC	0.9864	0.5535	0.6876	0.4580	0.4650	0.3895
U-Net++	Enh-Gray	IoU	0.9852	0.4641	0.7484	0.4255	0.4308	0.3517
U-Net	Enh-Gray	IoU	0.9842	0.3980	0.7309	0.3767	0.3817	0.3060
RIU-Net	Enh-Gray	IoU	0.9856	0.4051	0.7438	0.3685	0.3837	0.2974

หมายเหตุ* ในตารางที่ 4.1 ใช้ตัวย่อ ตัวอย่างเช่น RUN w. RN-50 หมายถึง โมเดล Res-U-Net ที่ประยุกต์ใช้ Backbone ResNet-50 และ DLv3+ w. RN-50 หมายถึง DeepLabV3+ ที่ประยุกต์ใช้ Backbone ResNet-50

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การวิเคราะห์ผลตามตารางที่ 4.1 ได้ถูกนำมาสร้างเป็นแผนภาพการแสดงผลแดชบอร์ดด้วยโปรแกรมพาวเวอร์บีไอเดสก์ทอป (Power BI Desktop) ตามรูปที่ 4.13 โดยรายละเอียดการแสดงผลประกอบไปด้วยหัวข้อดังต่อไปนี้

1. หน้าต่างกรองข้อมูล (ด้านซ้าย) ใช้สำหรับกรองข้อมูลในทุกแผนภาพ โดยมีหัวข้อชุดข้อมูลที่ใช้ในการเรียนรู้ของแต่ละโมเดล ดัชนีชี้วัด โมเดลการเรียนรู้เชิงลึก ประเภทของแบ็กโบน และโมเดลที่มีหรือไม่มีการประยุกต์แบ็กโบน
2. แผนภาพตารางแจกแจงคะแนนดัชนีชี้วัดซึ่งเรียงจากค่าคะแนน F1 มากไปน้อย
3. แผนภาพแสดงการกระจายตัวแบบจุดของโมเดล โดยอิงจากค่าเฉลี่ยของคะแนน Precision และ Recall ซึ่งใช้เปรียบเทียบผลความสมดุลของคะแนนทั้งสองค่า
4. แผนภาพแสดงการกระจายตัวแบบจุดของโมเดล โดยอิงจากค่าเฉลี่ยของคะแนนดัชนีชี้วัดความคล้ายคลึง IoU และ DSC
5. แผนภาพแสดงคะแนนเฉลี่ย F1 แบบแท่ง โดยอิงจากชุดข้อมูล (Dataset)
6. แผนภาพแสดงคะแนนเฉลี่ย F1 แบบแท่ง โดยอิงจากดัชนีชี้วัด (Training Metric)
7. แผนภาพแสดงคะแนนเฉลี่ย F1 แบบแท่ง โดยอิงจากการประยุกต์ใช้และไม่ใช้แบ็กโบน (Has backbone)
8. แผนภาพแสดงคะแนนเฉลี่ย F1 แบบแท่ง โดยอิงจากประเภทของแบ็กโบน (Backbone Type)



รูปที่ 4.13 การแสดงผลข้อมูลวิจัยผ่านทางแดชบอร์ด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2.1 การวัดผลความคล้ายคลึง

การวัดประเมินผลความคล้ายคลึง หรือการวัดประสิทธิภาพของโมเดลในการสกัดอาคารจากภาพถ่ายดาวเทียม ดำเนินการโดยนำโมเดลที่ได้จากการฝึกฝนจนแล้วเสร็จมาทดสอบกับชุดข้อมูลทดสอบที่ได้แบ่งกลุ่มไว้ ทั้งนี้ แต่ละโมเดลจะได้รับการทดสอบด้วยชุดทดสอบที่สอดคล้องกับชุดข้อมูลการฝึกฝน เช่น โมเดล Res-U-Net ฝึกฝนด้วยชุดข้อมูลภาพสี RGB จะได้รับการทดสอบกับชุดข้อมูลสี RGB เท่านั้น ไม่นำไปทดสอบกับภาพสีแบบปรับปรุง Enh-RGB และภาพสีเทา นอกจากนี้ การวัดผลใช้ดัชนีชี้วัดสามตัว ได้แก่ ดัชนี IoU ดัชนี DSC และเมทริกซ์ชี้วัดความสับสน Confusion Matrix สำหรับเมทริกซ์ชี้วัดความสับสน คะแนน F1 จะถูกนำมาพิจารณาเป็นหลักเมื่อต้องลำดับคะแนนจากมากไปน้อย อย่างไรก็ตาม ค่าความแม่นยำในการทำนาย (Precision) และค่าความแม่นยำในการตรวจจับ (Recall) ย่อมถูกนำมาอภิปรายด้วยเช่นกัน กระนั้น ค่าความแม่นยำรวม (Accuracy) ได้ถูกนำมาอภิปรายในแง่ความสามารถของโมเดลที่สมดุลในการทำนายอาคารและไม่ใช่อาคาร เนื่องจากเป็นสมการที่นำค่าการทำนายตามเงื่อนไข True Negative (TN) มาวิเคราะห์ ตามหัวข้อ 2.5.1

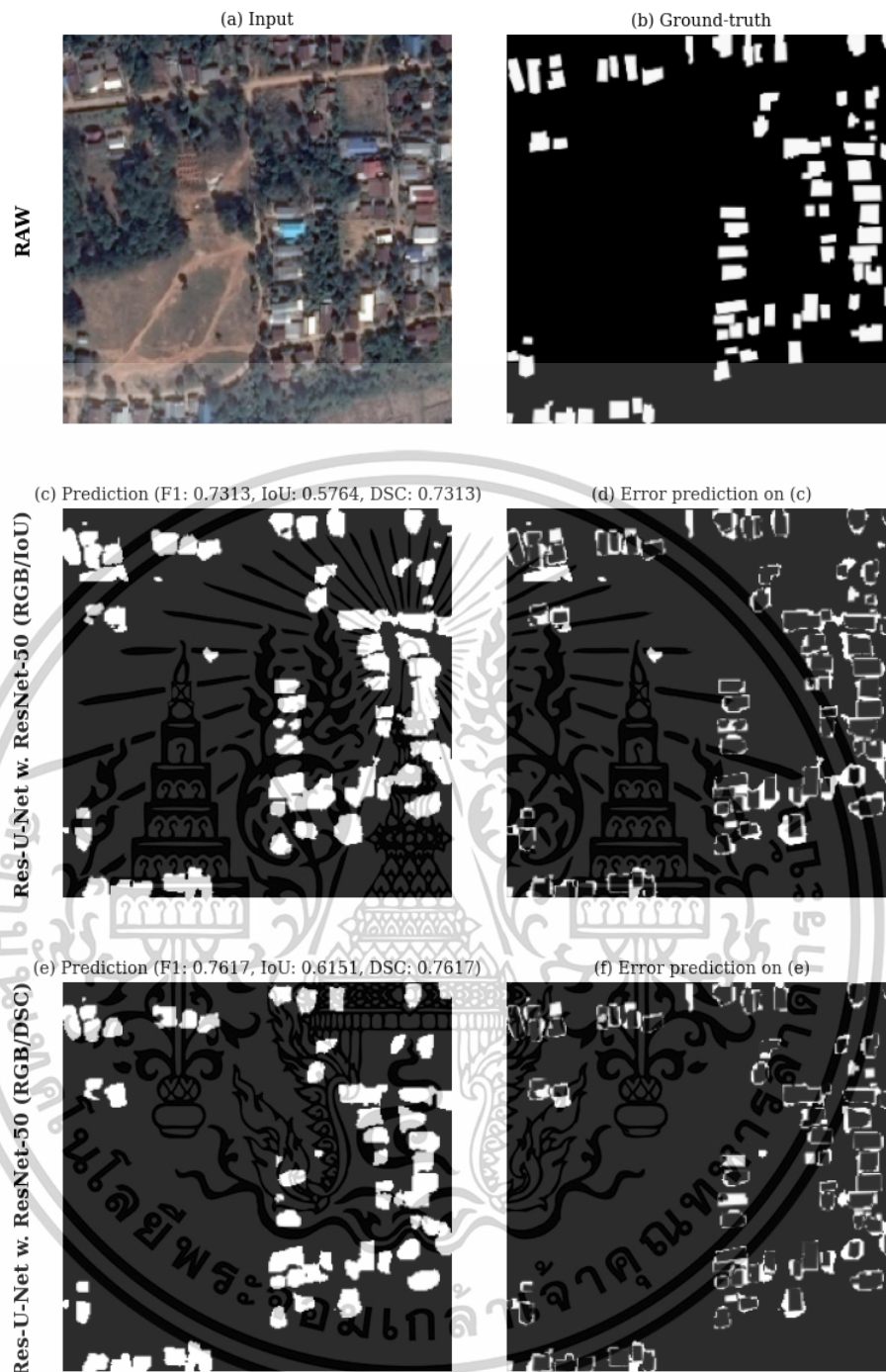
ตัวอย่างการทดสอบโมเดล Res-U-Net (ResNet-50) ที่ได้รับการฝึกบนชุดข้อมูล RGB เทียบกับชุดข้อมูลทดสอบ ดังที่ปรากฏในรูปที่ 4.14 โดยมีรายละเอียดดังนี้

รูปที่ (a) แสดงภาพนำเข้าทดสอบเพื่อทำนายโครงสร้างอาคาร

รูปที่ (b) แสดงถึงคำตอบหรือป้ายกำกับ ที่ระบุตำแหน่งของโครงสร้างอาคารทั้งหมดในภาพ ซึ่งนำมาใช้เป็นตัวชี้วัด 3 ดัชนีชี้วัดที่กล่าวถึง

รูปที่ (c) ภาพการแยกอาคารที่ได้จากการทำนายด้วยโมเดล (พื้นที่สีขาว) โดยใช้ดัชนี IoU ระหว่างการฝึกฝน ได้คะแนน F1 เท่ากับ 73.13% คะแนน IoU และ DSC เท่ากับ 57.64% และ 73.13% ตามลำดับ

รูปที่ (d) แสดงข้อผิดพลาดในการทำนายที่คลาดเคลื่อนกับคำตอบ รูปที่ (b) เช่นเดียวกัน รูปที่ (e) และรูปที่ (f) เป็นผลลัพธ์การทำนายและข้อผิดพลาดที่สอดคล้องกัน โดยได้รับคะแนนการประเมิน F1 เท่ากับ 76.17% คะแนน IoU และ DSC เท่ากับ 61.51% และ 76.17% ตามลำดับ



รูปที่ 4.14 การประเมินเปรียบเทียบ Res-U-Net (ResNet-50) ที่ฝึกฝนและทดสอบด้วยชุดข้อมูลภาพสี RGB แสดงค่าความแม่นยำและข้อผิดพลาดในการทำนายด้วยคะแนน F1, IoU และ DSC

ดังนั้น เมื่อนำโมเดลทั้ง 48 โมเดลเข้าสู่ขั้นตอนการทดสอบโดยใช้ชุดการทดสอบที่กำหนด และต่อมาใช้ดัชนีชี้วัดประเมินผลกับผลลัพธ์ที่ได้ทำนาย คะแนนผลลัพธ์การทดสอบจึงเป็นไปตามรายละเอียดในตารางที่ 4.1 และเนื่องด้วยการออกแบบการวิจัยนี้ก่อให้เกิดรูปแบบและผลลัพธ์ที่หลากหลาย รายละเอียดการอภิปรายจะกล่าวถึงสิ่งที่ค้นพบตามลำดับ

ความสำคัญ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2.2 วิเคราะห์ผลและเปรียบเทียบ

วิเคราะห์ผลและเปรียบเทียบจากภาพรวม โมเดลที่ให้ผลลัพธ์การทำนายที่ดีที่สุดเมื่อประเมินด้วยคะแนน F1 คือ Res-U-Net (ResNet-50) ซึ่งใช้ชุดข้อมูล RGB และได้รับการประเมินระหว่างการฝึกฝนด้วยดัชนี DSC โดยได้คะแนน F1 อยู่ที่ 85.43% และเมื่อประเมินด้วยหน่วยวัด DSC และ IoU คะแนนที่ได้เท่ากับ 85.48% และ 78.12% ตามลำดับ

#	Base Model	Backbone	Dataset	TM	Accuracy	Precision	Recall	F1	DSC	IoU
1	Res-U-Net	ResNet-50	RGB	DSC	99.41%	84.92%	88.98%	85.43%	85.48%	78.12%
2	DeepLabV3+	ResNet-50	Enh-RGB	DSC	99.24%	88.20%	87.25%	85.18%	85.25%	78.55%
3	DeepLabV3+	ResNet-152	RGB	DSC	99.42%	84.50%	88.42%	84.48%	84.55%	77.17%
4	DeepLabV3+	ResNet-101	RGB	DSC	99.40%	83.66%	88.92%	84.24%	84.30%	77.00%
5	DeepLabV3+	ResNet-50	Enh-RGB	IoU	99.17%	83.43%	90.76%	84.13%	84.19%	77.30%
6	DeepLabV3+	ResNet-101	Enh-RGB	DSC	99.25%	85.39%	88.73%	84.08%	84.15%	77.42%
7	DeepLabV3+	ResNet-152	Enh-RGB	IoU	99.26%	86.87%	86.90%	84.01%	84.09%	77.59%
8	Res-U-Net	ResNet-152	Enh-RGB	DSC	99.26%	86.23%	87.39%	83.95%	84.50%	77.47%
9	Res-U-Net	ResNet-101	RGB	DSC	99.41%	83.67%	88.49%	83.73%	84.08%	76.42%

รูปที่ 4.15 ผลคะแนน F1 และคะแนนอื่น ๆ ของโมเดล 9 อันดับแรก

อย่างไรก็ตาม เมื่อพิจารณาความสัมพันธ์ระหว่าง Precision และ Recall จะพบว่า DeepLabV3+ (ResNet-50) ซึ่งฝึกฝนด้วยชุดข้อมูลภาพสีปรับปรุง Enh-RGB ประกอบการประเมินด้วยดัชนี DSC ระหว่างการฝึกฝน ซึ่งให้เห็นว่าทำคะแนนได้ดีกว่า ทว่าผลลัพธ์คะแนน F1 กลับอยู่ที่ 84.18% อย่างไรก็ตาม การคำนวณ F1 ในงานวิจัยนี้ ไม่ได้มาจากการดำเนินการรวมกันของคะแนน Precision และ Recall เนื่องจากอาจส่งผลให้เกิดค่าคลาดเคลื่อนสูงเมื่อปิดเศษทศนิยม

ดังนั้น ในการวิเคราะห์ผลลัพธ์ที่ได้จากการประเมินจึงจำเป็นต้องทำการเปรียบเทียบผลการทำนาย ระหว่างโมเดลทั้งสองที่กล่าวมาข้างต้น โดยรูปที่ 4.16 ได้ชี้ถึงข้อสังเกตที่สำคัญของอาคารที่ได้จากการทำนายโดย Res-U-Net (ResNet-50; RGB/DSC) สามารถระบุและแยกแยะรอยต่อระหว่างโครงสร้างอาคารได้ชัดเจนกว่าการทำนายของโมเดล DeepLabV3+ (ResNet-50; Enh-RGB/DSC) ซึ่งมีจุดหรือส่วนของการทำนายที่ผสมผสานโครงสร้างอาคารเข้าเป็นพื้นที่เดียว นอกจากนี้ พื้นที่ทำนายโดยรอบของวัตถุยังมีขนาดใหญ่กว่าคำตอบ ฉะนั้น จึงสอดคล้องกับค่าความแม่นยำรวม (Accuracy) ที่พิจารณา TN ในการคำนวณ อย่างไรก็ตาม เมื่อเปรียบเทียบกับชุดคำตอบแล้ว โมเดล DeepLabV3+ (ResNet-50; Enh-RGB) ทำนายขอบเขตของอาคารได้ครอบคลุมกว่า

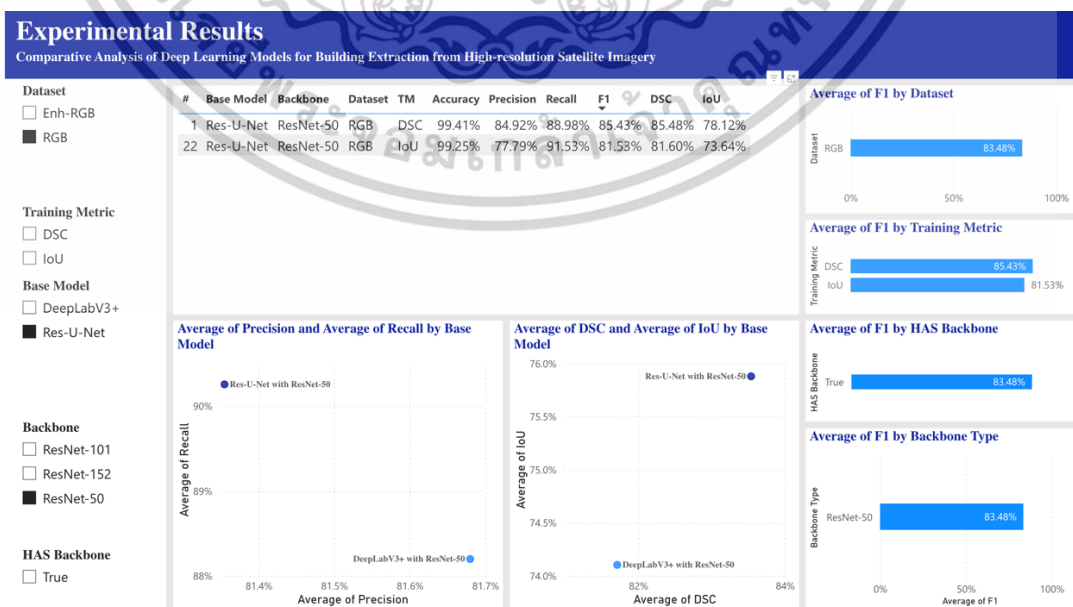
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.16 การเปรียบเทียบการสกัดอาคารโดยโมเดล Res-U-Net (ResNet-50; RGB/DSC) และโมเดล DeepLabV3+ (ResNet-50; Enh-RGB/DSC)

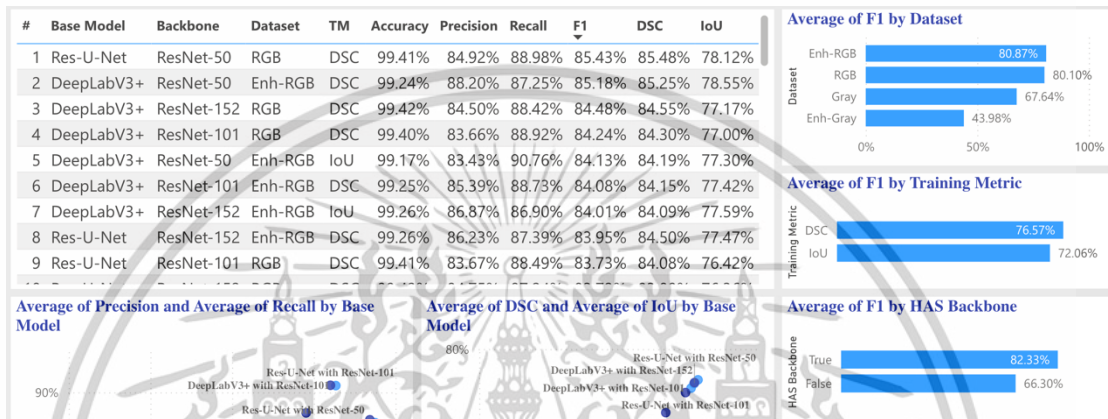
นอกจากนี้ หากเปรียบเทียบโมเดลประเภทเดียวกันกับชุดข้อมูลฝึกฝนเดียวกัน ทว่าใช้ดัชนีชี้วัด IoU ระหว่างการฝึกฝน พบว่า DeepLabV3+ (ResNet-50; Enh-RGB/IoU) ได้คะแนน F1 เท่ากับ 84.13% (ลดลงจาก DSC 1.05%) เทียบกับ Res-U-Net (ResNet-50; RGB/IoU) ได้คะแนน F1 เท่ากับ 81.53% (ลดลงจาก DSC 3.9%) ดังนั้น จากการตรวจสอบจึงสรุปได้ว่าความห่างระหว่างคะแนนของโมเดล DeepLabV3+ ด้วยคุณลักษณะดังที่กล่าวไม่มีนัยสำคัญเมื่อเทียบกับ Res-U-Net

นอกจากนี้ เมื่อเปรียบเทียบโมเดลประเภทและฟีเจอร์ชุดข้อมูลที่เหมือนกัน แต่ใช้ตัววัด IoU ตลอดระยะเวลาการฝึกอบรม ผลลัพธ์ที่ได้คะแนน F1 อยู่ที่ 84.13% สำหรับ DeepLabV3+ พร้อม ResNet-50 (Enh-RGB, IoU) และ 81.53% สำหรับ Res-U-Net พร้อม ResNet-50 (RGB, IoU) จากการตรวจสอบ จะเห็นได้ชัดว่าช่องว่างคะแนนระหว่าง DeepLabV3+ ทั้งสองเวอร์ชันนั้นไม่มีนัยสำคัญมากเมื่อเทียบกับ Res-U-Net (อันดับ 1 กับ 22 ทั้งที่เป็นโมเดลและชุดข้อมูลฝึกฝนเดียวกัน แต่ชี้วัดระหว่างการเรียนรู้ด้วยดัชนีต่างกัน)



รูปที่ 4.17 เปรียบเทียบผลคะแนนโมเดล Res-U-Net (ResNet-50/RGB) ที่ใช้ดัชนีต่างกัน เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการศึกษาเท่านั้น ไม่อนุญาตให้นำไปเผยแพร่หรือใช้งานโดยไม่ได้รับอนุญาตจากเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประการถัดมาจากภาพรวม อีกหนึ่งปัจจัยที่ชี้ให้เห็นถึงความแตกต่างอย่างชัดเจนของการวัดประสิทธิภาพของโมเดล คือ โมเดลที่มีการประยุกต์ใช้สถาปัตยกรรม Backbone พร้อมด้วย Pretrained weight ให้คะแนนการประเมินที่สูงกว่าโมเดลที่ไม่มีการประยุกต์ใช้ อย่างเห็นได้ชัด นอกจากนี้ โมเดลที่ได้ฝึกฝนด้วยชุดข้อมูลภาพสี RGB และภาพสีปรับปรุง Enh-RGB ให้คะแนนกว่าประเมินโดยเฉลี่ยสูงกว่าชุดภาพสีเทาทั้งสองประเภท เช่นเดียวกับกับการใช้ดัชนีชี้วัด DSC ระหว่างการฝึกฝน (รูปที่ 4.18)

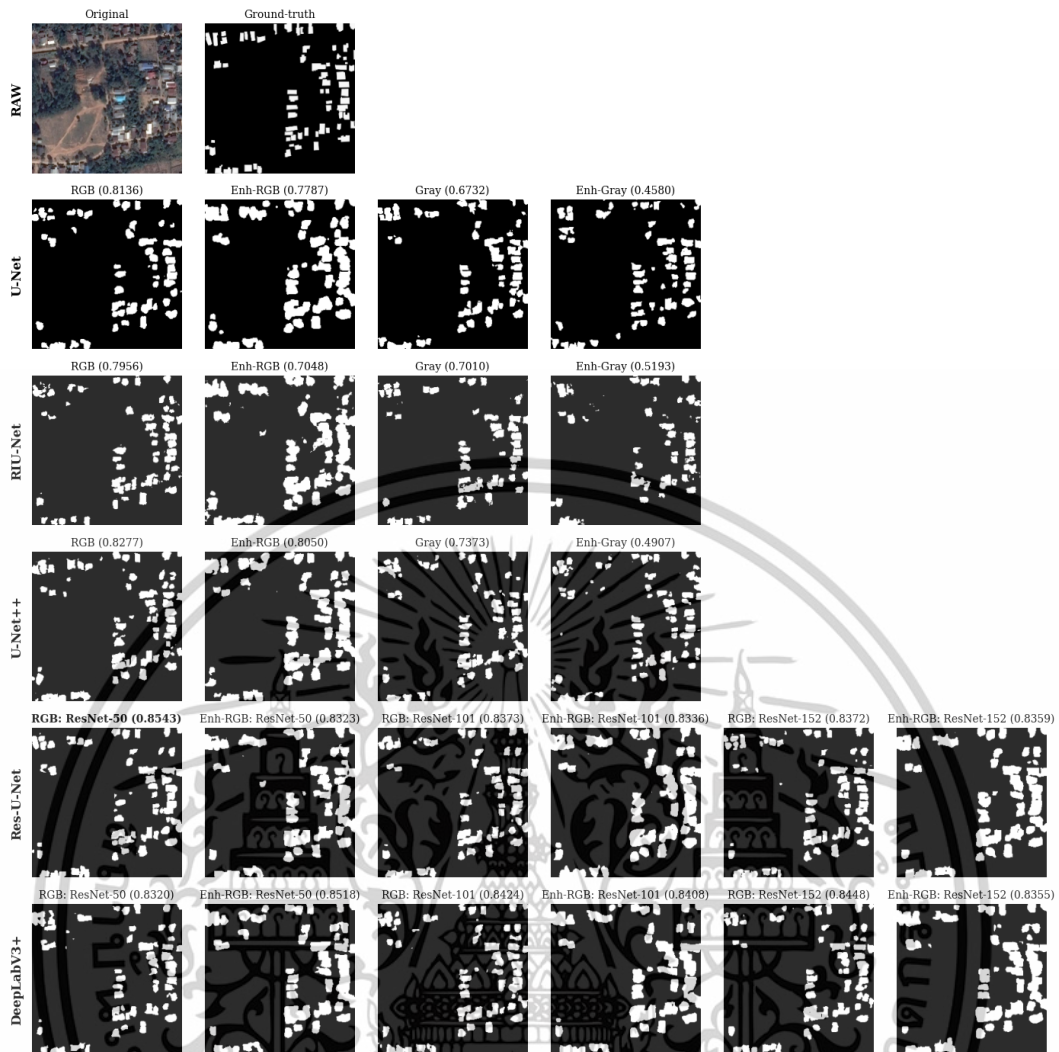


รูปที่ 4.18 ผลคะแนน F1 ของโมเดล 9 อันดับแรกและแผนภาพเปรียบเทียบคะแนนเฉลี่ย F1 โดยอิงจากชุดข้อมูล ดัชนีชี้วัด และการประยุกต์ใช้และไม่ใช้แบ็กโบน

ดังนั้น จึงสรุปได้ว่าโมเดลที่ได้รับการฝึกฝนด้วยภาพสี (RGB และ Enh-RGB) และประเมินผลระหว่างฝึกฝนด้วยดัชนีชี้วัด DSC ให้ประสิทธิภาพในการสกัดอาคารจากภาพถ่ายดาวเทียมได้ดี โดยเฉพาะโมเดลอันดับแรก ๆ อย่างโมเดล Res-U-Net (ResNet-50; RGB) และ DeepLabV3+ (ResNet-50; Enh-RGB) ที่ทำคะแนน DSC ได้เท่ากับ 85.48% และ 85.25% ตามลำดับ อย่างไรก็ตาม สิ่งสำคัญอีกประการหนึ่งคือโมเดล DeepLabV3+ สกัดอาคารและได้ผลลัพธ์ที่ดีเป็นอันดับสองเมื่อชี้วัดด้วย F1 แต่กระนั้น โมเดลดังกล่าวใช้ชุดข้อมูลภาพสีปรับปรุงในการฝึกฝน

อย่างไรก็ตาม เมื่อตรวจสอบภาพการทำนายตามตัวอย่าง รูปที่ 4.19 ซึ่งเป็นผลลัพธ์การทำนายของแต่ละโมเดลที่ประเมินผลระหว่างฝึกฝนด้วยดัชนี DSC พบว่า ประสิทธิภาพการทำนายของโมเดลเป็นไปตามข้อสรุปก่อนหน้า ทั้งการประยุกต์ใช้ Backbone ซึ่ง U-Net RIU-Net และ U-Net++ ไม่มีการประยุกต์ใช้ ทำให้ผลลัพธ์การทำนายคะแนนต่ำกว่าการฝึกฝนด้วยชุดข้อมูลภาพสี ตลอดจนตัวประเมินผลระหว่างฝึกฝน

สำหรับความคลาดเคลื่อนอันเกิดจากการทำนายด้วยโมเดลการเรียนรู้เชิงลึก แม้ว่าโมเดล Res-U-Net จะทำนายพื้นที่อาคารได้ชัดเจนและมีคะแนน F1 ที่สูงกว่า DeepLabV3+ แต่กระนั้น ก็ปรากฏจุดสีขาวขนาดเล็กซึ่งมีลักษณะคล้ายกับ Noise ในภาพที่สูงกว่า



รูปที่ 4.19 ตัวอย่างผลลัพธ์ที่ได้รับจากแต่ละโมเดลโดยใช้ดัชนี DSC ระหว่างการฝึกฝน ประกอบกับตัวเลขแสดงคะแนน F1 ของแต่ละภาพ

โมเดลสองอันดับแรกที่ได้รับการประเมินโดยใช้ดัชนี IoU ระหว่างการฝึกฝน และทดสอบกับชุดทดสอบตามที่กำหนด พบว่าโมเดล DeepLabV3+ ที่ประยุกต์ใช้ Backbone ResNet-152 และ ResNet-50 และรับการฝึกฝนบนชุดข้อมูลภาพสีแบบปรับปรุง Enh-RGB ได้คะแนน IoU ที่ 77.59% และ 77.30% ตามลำดับ ในทำนองเดียวกัน คะแนน F1 อยู่ที่สอดคล้องกันคือ 84.01% และ 84.13% ตามลำดับ

นอกจากนี้ เมื่อตรวจสอบผลลัพธ์การทำนายดังรูปที่ 4.20 ซึ่งเป็นผลการทำนายแบบเปรียบเทียบของแต่ละโมเดลที่ได้รับการประเมินระหว่างฝึกฝนด้วยดัชนี IoU พบว่า โมเดลที่ได้รับการฝึกฝนด้วยดัชนี IoU มีประสิทธิภาพต่ำกว่าเมื่อเทียบกับดัชนี DSC แต่กระนั้น เป็นเพราะว่า IoU มีแนวคิดทางทฤษฎีที่ตรงไปตรงมามากกว่าเกี่ยวกับการชี้วัดความคล้ายคลึง และพื้นที่ที่ทับซ้อนกันของภาพ ต่างกับดัชนี DSC ที่ให้สมดุลในด้านความแม่นยำและการ

ตรวจจับ ด้วยเหตุนี้ จึงสรุปได้ว่าดัชนีชี้วัด DSC เหมาะสมกับงานจำพวกการระบุ แยก สกัด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปเผยแพร่ในวงกว้างโดยไม่ผ่านการอนุมัติจากเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หรือการแบ่งส่วนภาพ (Image Segmentation) ได้ดีกว่า ทั้งนี้ ยังเห็นได้ชัดว่าโมเดล DeepLabV3+ ที่ใช้ Backbone ResNet-152 และ ResNet-50 มีประสิทธิภาพในการระบุโครงสร้างอาคารได้อย่างแม่นยำ โดยปราศจาก Noise ซึ่งเป็นพิกเซลสีขาวขนาดเล็กที่น้อยกว่าอย่างมีนัยสำคัญ แต่กระนั้น แม้ว่าโครงสร้างอาคารจะเชื่อมติดกัน แต่การทำนายของโมเดลยังคงแม่นยำและได้คะแนนสูง โดยรวมแล้ว ชุดข้อมูลภาพสี RGB และภาพสีปรับปรุง RGB มีส่วนช่วยโมเดลมีประสิทธิภาพที่เหนือกว่า ชุดข้อมูลภาพระดับสีเทา Gray และระดับสีเทาปรับปรุง Enh-Gray



รูปที่ 4.20 ตัวอย่างผลลัพธ์ที่ได้รับจากแต่ละโมเดลโดยใช้ดัชนี IoU ระหว่างการฝึกฝน ประกอบกับตัวเลขแสดงคะแนน F1 ของแต่ละภาพ

สิ่งที่ค้นพบจากการทดลองนี้แสดงให้เห็นว่า โมเดลที่ได้รับการฝึกบนชุดข้อมูลที่แตกต่างกันและประเมินโดยใช้ดัชนีชี้วัด DSC จะให้ประสิทธิภาพการทำนายของโมเดลที่ดีกว่าเมื่อเปรียบเทียบกับโมเดลที่ประเมินโดยใช้ดัชนีชี้วัด IoU และเมื่อพิจารณาจากคะแนนเอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น เมื่อนำมาใช้ในเชิงพาณิชย์ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

F1 โมเดล Res-U-Net (ResNet-50; RGB/DSC) ได้คะแนน F1 สูงสุดที่ 85.43% ซึ่งมีประสิทธิภาพเหนือกว่าโมเดล DeepLabV3+ (ResNet-50; Enh-RGB/DSC) ซึ่งได้คะแนน 85.18% นอกจากนี้ โมเดลที่ใช้สถาปัตยกรรมแกนหลัก Backbone ที่ได้รับการฝึกฝนล่วงหน้า (Pretrained weight) ให้ผลลัพธ์ที่เหนือกว่าเช่นกัน และคุณสมบัติของภาพเรื่องสี นับว่ามีบทบาทสำคัญอย่างมากในการสกัดอาคารจากภาพถ่ายดาวเทียม

4.2.3 ระยะเวลาที่ใช้ในการฝึกฝน

ข้อสังเกตเกี่ยวกับระยะเวลาที่ใช้ประมวลผลในการฝึกฝนโมเดล มีจุดที่น่าสนใจดังต่อไปนี้ ประการแรก ดัชนีชี้วัดที่ใช้ประเมินระหว่างการพัฒนาฝึกฝนไม่ส่งผลต่อระยะเวลาที่ใช้ในการฝึกฝน ทั้งดัชนี IoU และ DSC แม้ว่าจะให้ความแตกต่างกันในการทำนาย ประการที่สอง ความซับซ้อนของสถาปัตยกรรมแต่ละโมเดลส่งผลอย่างเห็นได้ชัดต่อระยะเวลาที่ใช้ โดยโมเดลที่มีความซับซ้อนของโครงข่ายน้อยที่สุด มีจำนวนพารามิเตอร์ในแต่ละเลเยอร์น้อยที่สุดคือ U-Net และผลการจับเวลาชี้ให้เห็นว่าโมเดลดังกล่าวใช้เวลาน้อยที่สุดในการฝึกฝน ประการที่สาม โมเดลที่มีการประยุกต์ใช้ Backbone ในโครงสร้าง หากมีการปรับเปลี่ยน Backbone ไปใช้ตัวที่มีโครงสร้างซับซ้อนก็จะส่งผลให้ใช้เวลาในการฝึกฝนมากขึ้นเช่นกัน ดังนั้น ประการที่สามจึงสนับสนุนข้อสรุปประการที่สอง โดยข้อมูลเปรียบเทียบเวลาที่ใช้ประมวลผลในการฝึกฝนเป็นไปตามตารางที่ 4.2

ตารางที่ 4.2 เปรียบเทียบเวลาที่ใช้ประมวลผลในการฝึกฝนของแต่ละโมเดล

Model	Dataset	ระยะเวลาที่ใช้ประมวลผล (ชั่วโมง)	ระยะเวลาที่ใช้ประมวลผล (นาที)
U-Net	RGB	1.264	75.84
	Enh-RGB	1.344	80.64
	Gray	1.344	80.64
	Enh-Gray	1.346	80.76
RIU-Net	RGB	2.082	124.92
	Enh-RGB	2.181	130.86
	Gray	1.977	118.62
	Enh-Gray	2.121	127.26
U-Net++	RGB	1.561	93.66
	Enh-RGB	1.584	95.04
	Gray	1.583	94.98
	Enh-Gray	1.579	94.74
Res-U-Net with ResNet-50	RGB	2.416	144.96
	Enh-RGB	2.135	128.1
Res-U-Net with ResNet-101	RGB	1.976	118.56
	Enh-RGB	2.211	132.66
Res-U-Net with ResNet-152	RGB	2.446	146.76
	Enh-RGB	2.352	141.12
DeepLabV3+ with ResNet-50	RGB	2.346	140.76
	Enh-RGB	2.297	137.82
DeepLabV3+ with ResNet-101	RGB	2.374	142.44
	Enh-RGB	2.352	141.12
DeepLabV3+ with ResNet-152	RGB	2.435	146.1
	Enh-RGB	2.451	147.06

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

นอกจากนี้ หากการพิจารณาความสมดุลระหว่างประสิทธิภาพของโมเดลกับระยะเวลาที่ใช้ในการฝึกฝน ด้วยดัชนีชี้วัดและสมการตามทฤษฎีอ้างอิง จึงกล่าวสรุปได้ว่าโมเดล Res-U-Net (ResNet-50) เป็นตัวเลือกที่เหมาะสมที่สุดสำหรับการสกัดอาคารจากภาพถ่ายดาวเทียม ทว่าในขณะเดียวกัน หากพิจารณาคุณภาพการทำนายจาก DeepLabV3+ (ResNet-50) และระยะเวลาที่ใช้ในการฝึกฝน โมเดลดังกล่าวกลับเป็นตัวเลือกที่น่าสนใจมากกว่าหากได้นำไปต่อยอด เหตุเพราะค่าเฉลี่ยระหว่าง Precision (88.20%) และ Recall (87.25%) คำนวณร่วมกันให้ผลลัพธ์เท่ากับ 87.73% ซึ่งสูงกว่า Res-U-Net (ResNet-50) ที่ประเมินแล้วได้ค่า Precision เท่ากับ 84.92 และ Recall เท่ากับ 88.98 ดังนั้น ค่าเฉลี่ยจึงได้เท่ากับ 86.95% อีกประการหนึ่ง หากคำนวณ F1 Score ด้วยสมการ (4.7) ซึ่งสมการพื้นฐานตามทฤษฎีที่ใช้กันแพร่หลาย กระนั้น โลเบรารีคำนวณ F1 Score ของ scikit-learn เวอร์ชัน 1.4.2 [74] ที่นำมาใช้เป็นไปตามสมการที่ (2.4)

$$F1 \text{ Score (Original)} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4.7)$$

จากสมการที่ (4.7) หากนำค่า Precision และ Recall ของโมเดลสองอันดับแรกตามเดิมมาพิจารณา พบว่าโมเดล DeepLabV3+ (ResNet-50; Enh-RGB/DSC) และโมเดล Res-U-Net (ResNet-50; RGB/DSC) ให้ผลลัพธ์ F1 Score (4.7) เท่ากับ 87.72% และ 86.90% ตามลำดับ ทั้งนี้ หากนำสมการที่ (4.1) ใช้แทนสมการที่ (2.4) และจัดลำดับใหม่ จะได้ผลลัพธ์ตามตารางที่ 4.3

ตารางที่ 4.3 เปรียบเทียบผลลัพธ์และจัดลำดับใหม่ด้วยค่า F1 จากสมการที่ (4.1) พร้อมระยะเวลาที่ใช้ในการฝึกฝน

Model	Dataset	Training Metric	Accuracy	Precision	Recall	F1 (4.7)	ระยะเวลาที่ใช้ในการฝึกฝน (นาที)
DLv3+ w. RN-50	Enh-RGB	DSC	0.9924	0.8820	0.8725	0.8772	137.82
DLv3+ w. RN-101	Enh-RGB	DSC	0.9925	0.8539	0.8873	0.8703	141.12
DLv3+ w. RN-50	Enh-RGB	IoU	0.9917	0.8343	0.9076	0.8694	140.76
RUN w. RN-50	RGB	DSC	0.9941	0.8492	0.8898	0.8690	144.96
DLv3+ w. RN-152	Enh-RGB	IoU	0.9926	0.8687	0.8690	0.8688	147.06
RUN w. RN-152	Enh-RGB	DSC	0.9926	0.8623	0.8739	0.8680	141.12
RUN w. RN-152	Enh-RGB	IoU	0.9920	0.8429	0.8942	0.8678	141.12
RUN w. RN-101	Enh-RGB	IoU	0.9923	0.8310	0.9046	0.8662	132.66
RUN w. RN-101	Enh-RGB	DSC	0.9923	0.8310	0.9046	0.8662	132.66
RUN w. RN-50	Enh-RGB	DSC	0.9926	0.8583	0.8738	0.8660	128.1

ด้วยเหตุนี้ จึงสรุปได้ว่า DeepLabV3+ (ResNet-50) ซึ่งฝึกฝนด้วยชุดข้อมูลภาพสี่แบบปรับปรุง Enh-RGB มีแนวโน้มการนำไปต่อยอดได้ดีที่สุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากการศึกษาครั้งนี้ พบว่าโมเดลการเรียนรู้เชิงลึกที่ได้เลือกมาทดสอบมีศักยภาพในการแบ่งส่วนภาพ หรือนำมาสกัดอาคารได้ดีกว่ากระบวนการประมวลผลรูปภาพแบบดั้งเดิม อย่างไรก็ตามแต่ละโมเดลยังคงมีข้อจำกัดที่ควรวิเคราะห์และพิจารณา

เริ่มต้นด้วยโมเดล U-Net แม้ว่าจะใช้เวลาฝึกฝนน้อยที่สุดและให้ผลลัพธ์ที่น่าพอใจ แต่กระนั้นข้อจำกัดของสถาปัตยกรรมโครงข่ายก็มีข้อจำกัดในการตรวจจับคุณสมบัติจากรูปภาพ เนื่องจากโครงข่ายดั้งเดิมถูกออกแบบมาให้สอดคล้องกับบริบทภาพถ่ายทางการแพทย์ ดังนั้น ด้วยกระบวนการที่ศึกษาในครั้งนี้นี้ไม่ได้ปรับแต่งโครงสร้างเพียงปรับค่าไฮเปอร์พารามิเตอร์ให้สอดคล้อง จึงอาจกล่าวได้ว่าการเรียนรู้ของโมเดลยังมีข้อจำกัดเรื่องความลึกและขอบเขตการสกัดคุณลักษณะที่สำคัญของวัตถุขนาดเล็กในภาพขนาดใหญ่ [75] นอกจากนี้ยังต้องใช้ข้อมูลจำนวนมากเพื่อปรับปรุงประสิทธิภาพให้ดียิ่งขึ้น ลำดับถัดมา โมเดล RIU-Net ซึ่งรวมโมดูล Residual และ Inception เพื่อปรับปรุงการสกัดคุณสมบัติ ส่งผลให้ความซับซ้อนของสถาปัตยกรรมโมเดลเพิ่มมากขึ้น โดยต้องใช้เวลาและทรัพยากรการคำนวณที่มากตามอย่างไม่อาจเลี่ยง ทั้งนี้ ผลการทดลองระบุว่าเป็นโมเดลที่ไม่มีการประยุกต์ Backbone แต่กลับต้องใช้เวลากว้างขวางเท่ากับโมเดลที่ใช้ และด้วยความซับซ้อนที่เพิ่มเข้ามา ทำให้การปรับแต่งไฮเปอร์พารามิเตอร์ยุ่งยากขึ้น เช่นเดียวกับ โมเดล U-Net++ ที่มีสถาปัตยกรรมแบบ Nested and Dense Skip Connections การเพิ่มจำนวนพารามิเตอร์ทำให้ต้องเพิ่มกำลังประมวลผล แน่นนอนว่าการฝึกฝนโมเดลจำเป็นต้องใช้ระยะเวลามากกว่าโมเดล U-Net แบบดั้งเดิม แต่ทั้งนี้ ด้วยโครงสร้างสถาปัตยกรรมมีข้อจำกัดเรื่องปริมาณข้อมูลที่ใช้ ซึ่งอาจสรุปได้ว่าจำนวนข้อมูลในกรณีศึกษานี้มีไม่มากพอจะทำให้โมเดลมีประสิทธิภาพที่โดดเด่น

สำหรับ Res-U-Net ซึ่งให้คะแนน F1 ตามสมการของไลบรารี Scikit-Learn สูงสุด แต่กลับต้องกังวลเรื่องความซับซ้อนและความจำเป็นที่ต้องพึ่งพา Backbone กระนั้น หากการเลือก Backbone ที่มีโครงสร้างสถาปัตยกรรมไม่เหมาะสมกับงานที่นำไปใช้ก็อาจทำให้ประสิทธิภาพลดลงได้เช่นกัน หรือกล่าวได้ว่า Backbone อาจเป็นการจำกัดความสามารถไปโดยปริยาย ซึ่งกรณีนี้นับว่าเป็นส่วนที่ควรค่าแก่การวิจัยเปรียบเทียบต่อไปในอนาคต นอกจากนี้ การใช้ Backbone ยังนำมาซึ่งข้อจำกัดในการถ่ายโอนการเรียนรู้ เพราะขึ้นกับน้ำหนักหรือองค์ความรู้ของพารามิเตอร์ที่ได้รับการฝึกฝนไว้ล่วงหน้า

สุดท้าย โมเดล DeepLabV3+ แม้ว่าจะคะแนน F1 ตามไลบรารีที่กล่าวถึงข้างต้น จะให้ผลประสิทธิภาพที่ต่ำกว่าเล็กน้อยเมื่อเทียบกับ Res-U-Net ทว่าหากใช้สมการ (4.7) คำนวณ โมเดลได้แสดงให้เห็นถึงความโดดเด่นในแง่ความสมดุลระหว่าง Precision และ Recall อย่างไรก็ตาม การใช้โครงสร้างสถาปัตยกรรมที่มี Atrous Convolution ร่วมในการประมวลผล อาจเพิ่มความซับซ้อนในการคำนวณและส่งผลกระทบต่อระยะเวลาการฝึกอบรม โดยเฉพาะอย่างยิ่งกับภาพที่มีความละเอียดสูง เหตุเนื่องจากแนวคิดการทำ Spatial Pyramid Pooling ซึ่งออกแบบมาเพื่อรวบรวมข้อมูลหลายขนาด นอกจากนี้ การพึ่งพา Backbone ก็มีข้อจำกัดคล้ายกรณีของ Res-U-Net

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

เนื้อหาภายในบทนี้กล่าวถึงการสรุปผลการวิจัยจากการศึกษา เพื่อสะท้อนให้เห็นถึงสิ่งที่บรรลุตามวัตถุประสงค์ และความท้าทายในการดำเนินงานวิจัยที่เกี่ยวข้องกับการเปรียบเทียบวิธีการแยกระบุ สกัดพื้นที่อาคารจากภาพถ่ายดาวเทียมความละเอียดสูง ซึ่งบทสรุปจะเชื่อมโยงกับผลของการค้นพบตามระเบียบ ขอบเขต และวัตถุประสงค์ตามที่กำหนดไว้ในบทนำ นอกจากนี้ ยังนำเสนอข้อเสนอแนะสำหรับงานวิจัยในอนาคตที่จะสามารถนำไปพัฒนา ต่อยอดองค์ความรู้ ตลอดจนแนวทางปรับปรุง โดยเป็นผลประเมินจากผู้เชี่ยวชาญและผู้มีส่วนเกี่ยวข้อง เพื่อพัฒนาคุณภาพงานวิจัยสืบไป

5.1 สรุปผลการวิจัย

การวิจัยนี้นำเสนอวิธีการที่ครอบคลุมเกี่ยวกับ การสกัดอาคารจากภาพถ่ายดาวเทียมความละเอียดสูงในพื้นที่จังหวัดเลย ประเทศไทย ทั้งนี้ นับตั้งแต่กระบวนการจัดเก็บข้อมูล การประยุกต์ใช้เทคนิคการประมวลผลภาพ ตลอดจนการใช้งานแบบจำลองปัญญาประดิษฐ์หรือโมเดลการเรียนรู้เชิงลึกที่หลากหลายเพื่อเปรียบเทียบผลการทดลอง ซึ่งจากการทดลองแสดงให้เห็นว่าโมเดล Res-U-Net ที่ใช้แกนสถาปัตยกรรมแบ็กโบน ResNet-50 ประกอบการเรียนรู้ด้วยชุดข้อมูลภาพสี RGB และดัชนีชี้วัดความคล้ายคลึง DSC ให้ความแม่นยำสูงสุดในการแบ่งส่วนภาพอาคาร ขณะเดียวกัน ผลลัพธ์จากการแบ่งส่วนภาพอาคารด้วยโมเดล DeepLabV3+ ประกอบการเรียนรู้ด้วยชุดข้อมูลภาพสี Enh-RGB และดัชนีชี้วัดความคล้ายคลึง DSC ก็ให้ความแม่นยำได้ดีเป็นลำดับรองลงมา อย่างไรก็ตาม ประเด็นที่น่าสนใจคือผลคะแนน Precision ของโมเดล DeepLabV3+ (88.20%) สูงกว่า Res-U-Net (84.92%) ทว่าเมื่อพิจารณาผลการแบ่งส่วนภาพจะพบว่า ภาพจากโมเดล DeepLabV3+ ให้ชิ้นส่วนอาคารมีขนาดใหญ่กว่าและมีลักษณะเชื่อมติดกันมากกว่าเพียงน้อยเมื่อเทียบชุดคำตอบ

ดังนั้น หากพิจารณาในแง่ผลลัพธ์และคะแนนจากดัชนีชี้วัดความคล้ายคลึง ปัจจัยหลักที่มีนัยยะสำคัญต่อการเพิ่มประสิทธิภาพความแม่นยำในการทำงานของโมเดล คือการประยุกต์ใช้สถาปัตยกรรมแบ็กโบน (Backbone) ร่วมกับโมเดลอย่าง Res-U-Net และ DeepLabV3+ นอกจากนี้ ผลการประเมินยังชี้ให้เห็นว่าแบ็กโบน ResNet-50 สามารถทำคะแนน F1 ได้สูงที่สุดกับทั้งโมเดล Res-U-Net และ DeepLabV3+ ด้วยดัชนีชี้วัดความคล้ายคลึงเดียวกัน

อนึ่ง ปัจจัยสำคัญคือค่าสี (Color channels) อันเป็นคุณลักษณะที่ชี้ให้เห็นถึงความแตกต่างจากการทดลองแบ่งหมวดหมู่ชุดข้อมูลด้วยกระบวนการก่อนการประมวลผล ทั้งนี้ โมเดลที่ได้รับการเรียนรู้ด้วยชุดข้อมูลภาพสี RGB และ Enh-RGB (3-channels) ให้ประสิทธิภาพความแม่นยำเหนือกว่าโมเดลที่ได้รับเรียนรู้ด้วยชุดข้อมูลภาพระดับสีเทา Gray และ Enh-Gray (1-channels) ผลที่ได้จากการประเมินด้วยดัชนีชี้วัดสะท้อนถึงความสำคัญของค่าสีซึ่งมีส่วนสำคัญในกระบวนการแบ่งส่วนภาพอย่างเอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ชัดเจน หรืออาจกล่าวได้ว่า ค่าสีเป็นคุณลักษณะสำคัญอย่างหนึ่งในกระบวนการแบ่งส่วนภาพอาคาร จากภาพถ่ายดาวเทียมความละเอียดสูง ไม่พอแค่นั้น อีกหนึ่งปัจจัยที่ส่งผลต่อคะแนนการประเมินความแม่นยำในการแบ่งส่วนภาพยังรวมถึง ค่าสัมประสิทธิ์ดัชนีชี้วัดความคล้ายคลึง โดยสมการวัดค่าความคล้ายคลึง Dice Similarity Coefficient (DSC) ให้ประสิทธิภาพที่เหนือกว่า Intersection-over-Union (IoU) อย่างไรก็ตาม หากพิจารณาจากการวิจัยและพัฒนา การใช้สัมประสิทธิ์ชี้วัดความคล้ายคลึงทั้งสองสามารถใช้ร่วมกันได้ เนื่องจากไม่ส่งผลกระทบต่อขีดจำกัดด้านทรัพยากรการประมวลผล อีกทั้ง ดัชนีชี้วัดทั้งสองสามารถใช้เพื่อพิสูจน์แนวโน้มประสิทธิภาพของโมเดลได้ในเวลาเดียวกัน แต่หากทำที่สุด มีข้อจำกัดด้านทรัพยากรการจัดเก็บข้อมูล การเลือกใช้ดัชนีชี้วัด DSC จึงเป็นตัวเลือกลำดับแรกที่ควรตัดสินใจ

ด้วยเหตุและปัจจัยที่กล่าวมา หากพิจารณาค่าคะแนน F1 และคะแนนจากดัชนีชี้วัด DSC ที่ดีที่สุด จึงสรุปได้ว่า Res-U-Net/ResNet-50 ให้ผลลัพธ์การทำงานที่ดีที่สุด แต่กระนั้น หากพิจารณาจากความสมดุลระหว่างค่าเฉลี่ยของ Precision และ Recall ผลลัพธ์ที่ดีที่สุดคือ DeepLabV3+/ResNet-50 นอกจากนี้ โมเดลดังกล่าวได้ใช้ชุดข้อมูล Enh-RGB ในการเรียนรู้ และจุดนี้ชี้ให้เห็นว่ากระบวนการการประมวลผลล่วงหน้าสามารถช่วยเพิ่มประสิทธิภาพการทำงาน เพื่อแบ่งส่วนภาพได้ดียิ่งขึ้นอย่างชัดเจน แต่ทว่า ควรมีการพิจารณาความเสี่ยงบางประการร่วมด้วย เนื่องจากการนำภาพในพื้นที่ต่าง ๆ ซึ่งนอกเหนือจากชุดข้อมูลที่นำมาทดสอบไปเข้ากระบวนการตามที่น่าเสนอ อาจส่งผลให้เกิดความแปรปรวนต่างกันไป เหตุจากปัจจัยเรื่อง เวลา สภาพอากาศ และทิศทางทางรัศมีของแสงซึ่งเกิดจากการถ่ายภาพของดาวเทียม ประกอบกับ พื้นที่ที่นำมาศึกษาเป็นเขตชนบทหรือชานเมือง ทำให้มีพื้นที่ป่าไม้ปะปนเป็นส่วนมาก และในขณะเดียวกัน หากโมเดลสามารถเรียนรู้และทำนายส่วนที่ไม่ใช่อาคารได้อย่างแม่นยำ ผลลัพธ์จะมีแนวโน้มส่งผลให้ความแม่นยำสูงขึ้นด้วย

นอกเหนือจากคุณลักษณะชุดข้อมูล อีกสิ่งหนึ่งที่เป็นข้อสรุปการวิจัย คือโครงสร้างสถาปัตยกรรมโมเดลและระยะเวลาที่ใช้ในการเรียนรู้ แน่แน่นอนว่าโครงสร้างที่ซับซ้อนย่อมใช้เวลาในการเรียนรู้มากกว่า ถึงแม้โมเดล U-Net จะมีโครงสร้างที่ซับซ้อนน้อยที่สุดและใช้ระยะเวลาที่น้อยที่สุด แต่ผลลัพธ์ไม่ได้ด้อยไปกว่าโมเดลประเภทอื่นถึงขั้นเป็นลำดับสุดท้าย ดังนั้น หากต้องการใช้เพื่อพิสูจน์แนวทางการแบ่งส่วนภาพโดยใช้โมเดลการเรียนรู้เชิงลึกร่วมกับชุดข้อมูลที่ผ่านการประมวลผล รวมถึงชุดข้อมูลที่แปลกใหม่เพื่อทดสอบข้อสมมุติฐาน โมเดล U-Net จึงเป็นตัวเลือกที่ดี และสำหรับโมเดลที่ประยุกต์ใช้โครงสร้างร่วมกับแบ็กโบน แม้จะใช้ระยะเวลามากกว่าแต่ผลลัพธ์ที่ได้มีประสิทธิภาพที่เหนือกว่า

สำหรับแนวทางการพัฒนาในอนาคต ประเด็นสำคัญหลายประการควรได้รับการพิจารณาโดยแบ่งได้ตามขั้นตอนต่าง ๆ เช่น การปรับปรุงและประยุกต์ใช้เทคนิคการประมวลผลล่วงหน้าเพื่อเพิ่มความสมบูรณ์ของภาพ หรือแม้แต่ทำให้คุณลักษณะบางคุณลักษณะเด่นชัดมากยิ่งขึ้น ทว่าสิ่งสำคัญอย่างการปรับปรุงคุณภาพข้อมูลก็มีความเสี่ยงที่ควรคำนึง [76, 77] นอกจากนี้ ยังสามารถปรับการทดลองโดยเปรียบเทียบร่วมกับโมเดลการเรียนรู้เชิงลึกตัวอื่น ๆ ซึ่งอาจเป็นตัวเลือกที่ดีกว่า รวมไปถึง

การปรับใช้แบ็กโบนตัวอื่น ๆ ด้วยเช่นกัน ทั้งนี้ จากการตรวจสอบงานวิจัยและการศึกษาก่อนหน้า [57] เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

พบว่าโมเดลชื่อ SRBuildingSeg ได้รับการประเมินว่ามีประสิทธิภาพในการสกัดอาคารจากภาพถ่ายดาวเทียมได้แม่นยำเหนือกว่า เมื่อเปรียบเทียบกับโมเดลตระกูล U-Net และ DeepLabV3+

5.2 ข้อเสนอแนะ

1. ดำเนินการทดลองกับชุดข้อมูลที่มีความเป็นสากลมากยิ่งขึ้น เพื่อพิสูจน์ข้อสมมุติฐานและข้อสรุปการวิจัย ทั้งนี้ เนื่องจากชุดข้อมูลที่ใช้ในการวิจัยครั้งนี้เป็นชุดข้อมูลที่ดำเนินการจัดเก็บโดยผู้วิจัย อาจมีบางส่วนที่ไม่น่าเชื่อถือหรือมีความผิดพลาดในการทำชุดคำตอบ
2. ดำเนินการทดลองกับโมเดลการเรียนรู้ตัวอื่น ๆ รวมทั้งศึกษากระบวนการ Fine-tuning เพื่อปรับแต่งและเพิ่มประสิทธิภาพการทำนาย
3. ดำเนินการทดลอง Postprocessing เพื่อปรับแต่งผลลัพธ์โครงสร้างจากการสกัดอาคารให้มีความเป็นรูปทรงเลขาคณิตมากยิ่งขึ้น ซึ่งมีแนวโน้มส่งผลให้คะแนนความแม่นยำสูงขึ้น
4. ดำเนินการปรับค่าพารามิเตอร์ภายใต้กระบวนการ Preprocessing ให้มีความสอดคล้องกับปัจจัยของภาพมากขึ้น หรือกล่าวคือการทำไดนามิกพารามิเตอร์ให้ตรงกับคุณสมบัติของภาพนั้น ๆ ตัวอย่างเช่น ไดนามิกแกมมา (Dynamic gamma correction) โดยคำนวณจากค่าการกระจายตัวของค่าแสงในภาพ
5. ดำเนินการทดลองการจัดสรรชุดข้อมูล โดยไม่คำนึงถึงจำนวนรูปภาพของแต่ละกลุ่มประเภทระหว่างภาพที่มีบ้านและภาพที่ไม่มีบ้าน เนื่องจากเป้าหมายคือความพยายามทำการแบ่งส่วนภาพ ดังนั้น มีความเป็นไปได้ที่อาจไม่ต้องสมตุลภาพดังกล่าวในชุดข้อมูล



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เอกสารอ้างอิง

- [1] McCarthy, M. J., & Halls, J. N., 2014. "Habitat mapping and change assessment of coastal environments: an examination of WorldView-2, QuickBird, and IKONOS satellite imagery and airborne LiDAR for mapping barrier island habitats". *ISPRS International Journal of Geo-Information*, 3(1), pp. 297-325. <https://doi.org/10.3390/ijgi3010297>.
- [2] Boyle, S. A., Kennedy, C. M., Torres Monges, J. C., Colman, K., Pérez Estigarribia, P. E., & de la Sancha, N.U., 2014. "High-resolution satellite imagery is an important yet underutilized resource in conservation biology". *PLoS One*, 9(1), e86908. <https://doi.org/10.1371/journal.pone.0086908>.
- [3] Google. 2567. **Google Earth**. [Online]. Available : <https://about.google/brand-resource-center/products-and-services/geo-guidelines/#google-earth>.
- [4] ส่วนบริหารและพัฒนาเทคโนโลยีการทะเบียน, สำนักบริหารการทะเบียน, กรมการปกครอง. 2567. **ระบบสถิติทางการทะเบียน**. [Online]. Available : <https://stat.bora.dopa.go.th/stat/statnew/statyear/#/TableTemplate4/Area/statpop?yymm=66&ccDescจังหวัดเลย&topic=statpop&ccNo=42&rcodeNo=4201&rcodeDesc=อำเภอเมืองเลย>.
- [5] สำนักงานคณะกรรมการวิจัยแห่งชาติ. 2540. **คำบรรยายเรื่องการสำรวจจากระยะไกล : Remote Sensing Note**. พิมพ์ครั้งที่ 1. ISBN : 974-8027-59-7.
- [6] จริญญา กิติไพศาลนนท์ และผู้แต่งคนอื่นๆ. 2560. **พยานหลักฐานดิจิทัลในคดีสิ่งแวดล้อม: ศึกษากรณีเทคโนโลยีการสำรวจระยะไกล**. ม.ป.ท.:มหาวิทยาลัยธรรมศาสตร์. https://doi.nrct.go.th/ListDoi/listDetail?Resolve_DOI=10.14457/TU.the.2017.1296.
- [7] สำนักงานพัฒนาเทคโนโลยีอวกาศและภูมิสารสนเทศ (องค์การมหาชน). 2558. **ดาวเทียม (Satellites)**. [Online]. Available : https://www.gistda.or.th/news_view.php?n_id=2406&lang=TH.
- [8] สำนักงานพัฒนาเทคโนโลยีอวกาศและภูมิสารสนเทศ (องค์การมหาชน). 2567. **วงโคจรของดาวเทียม (Satellite orbit)**. [Online]. Available : <http://learn.gistda.or.th/article/บทความนักวิชาการ/remote-sensing/วงโคจรของดาวเทียม-satellite-orbit/>.
- [9] Ship Expert Technology. 2567. **วงโคจรดาวเทียมมีกี่ประเภท? แล้วอินเทอร์เน็ตที่ใช้กันในทะเลนั้นมันใช้วงโคจรไหน?**. [Online]. Available : <https://blog.shipexpert.net/blog/satellite-orbit>.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- [10] Prasanna Sivakumar. 2017. **Which satellite takes the image of Google Maps?** [Online]. Available : <https://www.quora.com/Which-satellite-takes-the-image-of-Google-Maps>.
- [11] สำนักงานพัฒนาเทคโนโลยีอวกาศและภูมิสารสนเทศ (องค์การมหาชน). 2567. **Landsat 8**. [Online]. Available : https://www.gistda.or.th/news_view.php?n_id=2446.
- [12] Imagery & Remote Sensing. 2013. **Band Combinations for Landsat 8**. [Online]. Available : <https://www.esri.com/arcgis-blog/products/product/imagery/band-combinations-for-landsat-8/>.
- [13] NASA Official. 2024. **LANDSAT 8 MISSION DETAILS**. [Online]. Available : <https://landsat.gsfc.nasa.gov/satellites/landsat-8/landsat-8-mission-details/>.
- [14] Satellite Imaging Corporation. 2024. **Sentinel-2A Satellite Sensor (10m)**. [Online]. Available : <https://www.satimagingcorp.com/satellite-sensors/other-satellite-sensors/landsat-8/>.
- [15] eoPortal. 2012. **Copernicus: Sentinel-2**. [Online]. Available : <https://www.eoportal.org/satellite-missions/copernicus-sentinel-2>.
- [16] Satellite Imaging Corporation. 2024. **SPOT-7 Satellite Sensor (1.5m)**. [Online]. Available : <https://www.satimagingcorp.com/satellite-sensors/spot-7/>.
- [17] Satellite Imaging Corporation. 2024. **SPOT-6 Satellite Sensor (1.5m)**. [Online]. Available : <https://www.satimagingcorp.com/satellite-sensors/spot-6/>.
- [18] eoPortal. 2023. **SPOT-6 and SPOT-7 (Azersky)**. [Online]. Available : <https://www.eoportal.org/satellite-missions/spot-6-7>.
- [19] สำนักงานพัฒนาเทคโนโลยีอวกาศและภูมิสารสนเทศ (องค์การมหาชน). 2567. **กลุ่มข้อมูลดาวเทียมรายละเอียดสูงจากฝรั่งเศส**. [Online]. Available : <http://learn.gistda.or.th/2016/02/18/กลุ่มข้อมูลดาวเทียมราย/>.
- [20] Satellite Imaging Corporation. 2024. **WorldView-3 Satellite Sensor (30cm)**. [Online]. Available : <https://www.satimagingcorp.com/satellite-sensors/worldview-3/>.
- [21] สำนักงานพัฒนาเทคโนโลยีอวกาศและภูมิสารสนเทศ (องค์การมหาชน). 2558. **ดาวเทียม WorldView-3 ประสบความสำเร็จในการขึ้นสู่อวกาศเมื่อวันที่ 13 ส.ค. 2557**. [Online]. Available : https://www.gistda.or.th/news_view.php?n_id=2385&lang=TH.
- [22] Satellite Imaging Corporation. 2024. **WorldView-4 Satellite Sensor (30cm)**. [Online]. Available: <https://www.satimagingcorp.com/satellite-sensors/geoeye-2/>.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษเท่านั้น เมื่ออนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- [23] สำนักงานพัฒนาเทคโนโลยีอวกาศและภูมิสารสนเทศ (องค์การมหาชน). 2567. QuickBird. [Online]. Available : https://gistda.or.th/ewtadmin//ewt/gistda_web/article_attachment/articlefile_2021081813502299167.pdf.
- [24] สำนักงานพัฒนาเทคโนโลยีอวกาศและภูมิสารสนเทศ (องค์การมหาชน). 2567. ตำราเทคโนโลยีอวกาศและภูมิสารสนเทศศาสตร์. [Online]. Available : <http://learn.gistda.or.th/wp-content/uploads/book/Space%20techology%20and%20GEO-informatics.pdf>.
- [25] IBM. 2024. What is machine learning (ML)?. [Online]. Available : <https://www.ibm.com/topics/machine-learning>.
- [26] IBM. 2024. What is deep learning?. [Online]. Available : <https://www.ibm.com/topics/deep-learning>.
- [27] IBM. 2024. What is a neural network?. [Online]. Available : <https://www.ibm.com/topics/neural-networks>.
- [28] IBM. 2024. What are convolutional neural networks?. [Online]. Available : <https://www.ibm.com/topics/convolutional-neural-networks>.
- [29] Amazon (AWS). 2024. Transfer Learning (การถ่ายโอนความรู้) คืออะไร?. [Online]. Available : <https://aws.amazon.com/th/what-is/transfer-learning/>.
- [30] IBM. 2024. What is image segmentation?. [Online]. Available : <https://www.ibm.com/topics/image-segmentation>.
- [31] Ronneberger, O., Fischer, P., & Brox, T., 2015. "U-net: Convolutional networks for biomedical image segmentation". *Proceeding of Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference*, Munich, Germany, October 5-9, 2015, pp. 234-241. https://doi.org/10.1007/978-3-319-24574-4_28.
- [32] He, K., Zhang, X., Ren, S., & Sun, J., 2016. "Deep residual learning for image recognition". *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*, Las Vegas, NV, USA, June 27-30, 2016, pp. 770-778.
- [33] Sariturk, B., & Seker, D. Z., 2022. "A Residual-Inception U-Net (RIU-Net) Approach and Comparisons with U-Shaped CNN and Transformer Models for Building Segmentation from High-Resolution Satellite Images". *Sensors*, 22(19), 7624. <https://doi.org/10.3390/s22197624>.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- [34] Alom, M. Z., Yakopcic, C., Hasan, M., Taha, T. M., & Asari, V. K., 2019. Recurrent residual U-Net for medical image segmentation. *Journal of Medical Imaging*, 6(1). 014006. <https://doi.org/10.1117/1.JMI.6.1.014006>.
- [35] Zhou, Z., Rahman Siddiquee, M. M., Tajbakhsh, N., & Liang, J., 2018. "Unet++: A nested u-net architecture for medical image segmentation". In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018*, pp. 3-11.
- [36] Diakogiannis, F. I., Waldner, F., Caccetta, P., & Wu, C., 2020. "ResUNet-a: A deep learning framework for semantic segmentation of remotely sensed data". *ISPRS Journal of Photogrammetry and Remote Sensing*, 162, pp. 94-114. <https://doi.org/10.1016/j.isprsjprs.2020.01.013>
- [37] Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., & Adam, H., 2018. "Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation". *Proceedings of the European Conference on Computer Vision (ECCV2018)*, Munich, Germany, September 8-14, 2018, pp. 801-818. <https://doi.org/10.48550/arXiv.1802.02611>.
- [38] Liu, M., Fu, B., Xie, S., He, H., Lan, F., Li, Y., ... & Fan, D., 2021. "Comparison of multi-source satellite images for classifying marsh vegetation using DeepLabV3 Plus deep learning algorithm". *Ecological Indicators*, 125, 107562. <https://doi.org/10.1016/j.ecolind.2021.107562>.
- [39] Kohavi, R., Provost, F., 1998. "Glossary of terms. Machine Learning—Special Issue on Applications of Machine Learning and the Knowledge Discovery Process". *Machine Learning*, 30, pp. 271-274. <https://doi.org/10.1023/A:1017181826899>.
- [40] Jaccard, P., 1912, "THE DISTRIBUTION OF THE FLORA IN THE ALPINE ZONE.1". *New Phytologist*, 11(2), pp. 37-50. <https://doi.org/10.1111/j.1469-8137.1912.tb05611.x>.
- [41] Dice, L. R., 1945. "Measures of the amount of ecologic association between species". *Ecology*, 26(3), pp. 297-302. <https://doi.org/10.2307/1932409>
- [42] Sorensen, T., 1948. "A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analyses of the vegetation on Danish commons". *Biologiske skrifter*, 5, pp. 1-34.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น เมื่ออนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- [43] Gavankar, N. L., & Ghosh, S. K., 2018. "Automatic building footprint extraction from high-resolution satellite image using mathematical morphology". *European Journal of Remote Sensing*, 51(1), pp. 182-193. <https://doi.org/10.1080/22797254.2017.1416676>
- [44] Chhor, G., Aramburu, C. B., & Bougdal-Lambert, I.. 2017. **Satellite image segmentation for building detection using U-Net**. [Online]. Available: <http://cs229.stanford.edu/proj2017/final-reports/5243715.pdf>.
- [45] Ivanovsky, L., Khryashchev, V., Pavlov, V., & Ostrovskaya, A., 2019. "Building detection on aerial images using U-NET neural networks". *Proceedings of the 24th Conference of Open Innovations Association (FRUCT)*, Moscow, Russia, April 8-12, 2019, pp. 116-122. <https://doi.org/10.23919/FRUCT.2019.8711930>.
- [46] Daranagama, S., & Witayangkum, A., 2021. "Automatic building detection with polygonizing and attribute extraction from high-resolution images". *ISPRS International Journal of Geo-Information*, 10(9), 606. <https://doi.org/10.3390/ijgi10090606>
- [47] Sariturk, B., & Seker, D. Z., 2022. "A Residual-Inception U-Net (RIU-Net) Approach and Comparisons with U-Shaped CNN and Transformer Models for Building Segmentation from High-Resolution Satellite Images". *Sensors*, 22(19), 7624. <https://doi.org/10.3390/s22197624>.
- [48] Xu, Y., Wu, L., Xie, Z., & Chen, Z., 2018. "Building extraction in very high resolution remote sensing imagery using deep learning and guided filters". *Remote Sensing*, 10(1), 144. <https://doi.org/10.3390/rs10010144>.
- [49] Alsabhan, W., Alotaiby, T., & Dudin, B., 2022. "Detecting Buildings and Nonbuildings from Satellite Images Using U-Net". *Computational Intelligence and Neuroscience*, 4831223, <https://doi.org/10.1155/2022/4831223>.
- [50] Alsabhan, W., & Alotaiby, T., 2022. "Automatic building extraction on satellite images using Unet and ResNet50". *Computational Intelligence and Neuroscience*, 5008854, <https://doi.org/10.1155/2022/5008854>.
- [51] Aslantaş, N., Bayram, B., Bakırman, T., 2021. "Building Segmentation from VHR Aerial Imagery using DeepLabv3+ Architecture". *Proceedings of the 42nd Asian Conference on Remote Sensing (ACRS2021)*, Can Tho city, Vietnam, November 22-24, 2021, pp. 134-142.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- [52] Han, J., Wang, Z., Wang, Y., & Hou, W., 2022. "Building extraction algorithm from remote sensing images based on improved DeepLabv3+ network". *Journal of Physics: Conference Series*, 2303(1), 012010. <https://doi.org/10.1088/1742-6596/2303/1/012010>
- [53] Bakirman, T., Komurcu, I., & Sertel, E., 2022. "Comparative analysis of deep learning based building extraction methods with the new VHR Istanbul dataset". *Expert Systems with Applications*, 202, 117346. <https://doi.org/10.1016/j.eswa.2022.117346>.
- [54] Google. 2023. **Google Earth**. [Online]. Available : <https://earth.google.com/web/>.
- [55] Chen, S., Ogawa, Y., Zhao, C., & Sekimoto, Y., 2023. "Large-scale individual building extraction from open-source satellite imagery via super-resolution-based instance segmentation approach". *ISPRS Journal of Photogrammetry and Remote Sensing*. 195, pp. 129-152. <https://doi.org/10.1016/j.isprsjprs.2022.11.006>.
- [56] Wen, Q., Jiang, K., Wang, W., Liu, Q., Guo, Q., Li, L., & Wang, P., 2019. "Automatic building extraction from Google Earth images under complex backgrounds based on deep instance segmentation network". *Sensors*, 19(2), 333. <https://doi.org/10.3390/s19020333>.
- [57] Zhang, L., Dong, R., Yuan, S., Li, W., Zheng, J., & Fu, H., 2021. "Making low-resolution satellite images reborn: a deep learning approach for super-resolution building extraction". *Remote Sensing*, 13(15), 2872. <https://doi.org/10.3390/rs13152872>.
- [58] GIS English. 2023. **GIS**. [Online]. Available: <https://gisenglish.geojamal.com/>.
- [59] Lin, B. S., Michael, K., Kalra, S., & Tizhoosh, H. R., 2017. "Skin lesion segmentation: U-nets versus clustering". *Proceeding of the 2017 IEEE Symposium Series on Computational Intelligence (SSCI)*, Honolulu, HI, USA, November 27 - December 1, 2017, pp. 1-7. <https://doi.org/10.1109/SSCI.2017.8280804>.
- [60] International Commission on Illumination. 2004. **CIE 15: Technical Report: Colorimetry, 3rd edition**. [Online]. Available : <https://archive.org/details/gov.law.cie.15.2004/page/n1/mode/2up>.
- [61] Kurt A. Rosentrater, A.D. Evers, 2018. **Kent's Technology of Cereals, Chapter 7 - Flour treatments, applications, quality, storage and transport pp. 515-564**.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- Fifth Edition. UK, Woodhead Publishing. <https://doi.org/10.1016/B978-0-08-100529-3.00007-4>.
- [62] Gedraite, E. S., & Hadad, M., 2011. "Investigation on the effect of a Gaussian Blur in image filtering and segmentation". *Proceedings of International Symposium on Electronics in Marine (ELMAR2011)*, Zadar, Croatia, September 14-16, 2011, pp. 393-396.
- [63] Li, Z., Zheng, J., Zhu, Z., Yao, W., & Wu, S., 2014. "Weighted guided image filtering". *IEEE Transactions on Image processing*, 24(1), pp. 120-129. <https://doi.org/10.1109/TIP.2014.2371234>.
- [64] Vidhya, G. R., & Ramesh, H., 2017. "Effectiveness of contrast limited adaptive histogram equalization technique on multispectral satellite imagery". *Proceedings of the International Conference on Video and Image Processing (ICVIP)*, Singapore, Singapore, December 27-29, 2017, pp. 234-239. <https://doi.org/10.1145/3177404.3177409>.
- [65] Shorten, C., & Khoshgoftaar, T. M., 2019. "A survey on image data augmentation for deep learning". *Journal of Big Data*, 6(1), pp. 1-48. <https://doi.org/10.1186/s40537-019-0197-0>.
- [66] OpenCV (Open Source Computer Vision), doxygen. 2024. **Color conversions**. [Online]. Available : https://docs.opencv.org/3.4/de/d25/imgproc_color_conversions.html.
- [67] Richard Szeliski. 2022. *Computer Vision: Algorithms and Applications*. พิมพ์ครั้งที่ 2. ISBN : 978-3-030-34372-9.
- [68] OpenCV (Open Source Computer Vision), doxygen. 2024. **Smoothing Images**. [Online]. Available : https://docs.opencv.org/3.4/dc/dd3/tutorial_gaussian_median_blur_bilateral_filter.html.
- [69] OpenCV (Open Source Computer Vision), doxygen. 2024. **Basic Thresholding Operations**. [Online]. Available : <https://docs.opencv.org/2.4/doc/tutorials/imgproc/threshold/threshold.html>.
- [70] Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D.G., Steiner, B., Tucker, P.A., Vasudevan, V., Warden, P., Wicke, M., Yu, Y., & Zhang, X.,

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2016. "TensorFlow: A system for large-scale machine learning". *Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation*, Savannah, GA, USA, November 2-4, 2016, pp. 265-283. <https://doi.org/10.48550/arXiv.1605.08695>.
- [71] Google. 2023. **Google Colab**. [Online]. Available : <https://colab.google/>.
- [72] Google. 2023. **Personal cloud storage & file sharing platform – Google**. [Online]. Available : <https://www.google.com/drive/>.
- [73] Keras. 2024. **RMSprop**. [Online]. Available : <https://keras.io/api/optimizers/rmsprop/>.
- [74] scikit-learn developers. 2024. **sklearn.metrics.f1_score**. [Online]. Available : https://scikit-learn.org/stable/modules/generated/sklearn.metrics.f1_score.html.
- [75] Wu, Y., Wang, G., Wang, Z., Wang, H., & Li, Y., 2022. "DI-Unet: Dimensional interaction self-attention for medical image segmentation". *Biomedical Signal Processing and Control*, 78, 103896. <https://doi.org/10.1016/j.bspc.2022.103896>.
- [76] Namdeo, A., & Bhadoriya, S. S., 2016. "A review on image enhancement techniques with its advantages and disadvantages". *International Journal for Science and Advance Research in Technology (IJSART)*, 2(5), pp. 171-182.
- [77] Zhang, Y., Chen, W., Chen, Y., & Tang, X., 2018. "A post-processing method to improve the white matter hyperintensity segmentation accuracy for randomly-initialized U-net". *Proceeding of the 2018 IEEE 23rd International Conference on Digital Signal Processing (DSP)*, Shanghai, China, November 19-21, 2018, pp. 1-5. <https://doi.org/10.1109/ICDSP.2018.8631858>.

ประวัติผู้เขียน

ชื่อ	นายเดชสิทธิ์ ชื้อประเสริฐ
วัน เดือน ปีเกิด	วันที่ 16 เมษายน พ.ศ. 2539
ที่อยู่ปัจจุบัน	บ้านเลขที่ 33/183 หมู่ที่ 4 ตำบลทับมา อำเภอเมืองระยอง จังหวัดระยอง รหัสไปรษณีย์ 21000
ประวัติการศึกษา	(2560) วิทยาศาสตร์บัณฑิต สาขาวิทยาการคอมพิวเตอร์ เกรดเฉลี่ย 3.61 (สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง) (2567) วิทยาศาสตรมหาบัณฑิต สาขาวิทยาการคอมพิวเตอร์ (สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง)
ผลงานทางวิชาการ	1. Yasothon, P., Chueprasert, T., Towongpaichayont, W., 2011. "Death Escape: A Case Study of Merging Ubiquitous Activities into a Hardcore Computer Game". <i>Proceeding of the 2018 ICCCV International Conference on Control and Computer Vision</i> , Singapore, June 15-18, 2018, pp. 121-126. https://doi.org/10.1145/3232651.3232655 . 2. Chueprasert, T., Udomchaipom, A., Intakosum, S., 2024. "Comparative Analysis of Deep Learning Models for Building Extraction from High-resolution Satellite Imagery". <i>Current Applied Science and Technology</i> , 25(1). https://doi.org/10.55003/cast.2024.260846 .

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้