



การสกัดข้อมูลรอยพิมพ์ฐานอาคารจากภาพออร์โธจริงด้วยแบบจำลองการแยกส่วนทุกสิ่งเชิงภูมิปริภูมิ
BUILDING FOOTPRINT EXTRACTION FROM TRUE ORTHOPHOTO BY GEOSPATIAL
SEGMENT ANYTHING MODEL

เทพชัย ศรีน้อย^{1*}, ธีรวัฒน์ บรรณกุลพิพัฒน์² และไพศาล สันติธรรมนนท์³

¹นิสิตมหาบัณฑิต สาขาวิศวกรรมสำรวจ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

²นักวิจัย, สาขาวิศวกรรมสำรวจ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

³รองศาสตราจารย์, สาขาวิศวกรรมสำรวจ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

*Corresponding author, E-Mail: thepchaisrinoi@gmail.com

บทคัดย่อ

ข้อมูลอาคารเป็นข้อมูลพื้นฐานในการแสดงถึงข้อมูลประชากร สภาพแวดล้อมในพื้นที่หนึ่ง วิธีการสร้างชั้นข้อมูลด้วยวิธีดั้งเดิมเป็นการคัดลอกเส้นจากภาพถ่ายเทียมหรือภาพออร์โธซึ่งใช้ทรัพยากรมนุษย์ เวลา เป็นอย่างมาก ปัจจุบันมีข้อมูลรอยพิมพ์ฐานอาคารที่เปิดให้ใช้ได้โดยไม่เสียค่าใช้จ่าย แต่ความสมบูรณ์ของรอยพิมพ์ฐานอาคารยังไม่มากนักโดยเฉพาะในพื้นที่เมือง ล่าสุดมีการเผยแพร่แบบจำลองการแยกแยะทุกสิ่งอย่างเชิงภูมิปริภูมิให้ทดลองใช้ได้ ประกอบกับมีโครงการทำแผนที่ภายในเมืองด้วยอากาศยานไร้คนขับซึ่งได้ภาพออร์โธจริงที่ระยะการสุ่มบนพื้นดินเป็น 5 เซนติเมตร จึงเกิดเป็นการศึกษาสกัดข้อมูลรอยพิมพ์ฐานอาคารด้วยแบบจำลองภายในพื้นที่คณะวิทยาศาสตร์และวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ตรวจสอบความสมบูรณ์ของรอยพิมพ์ฐานอาคารด้วยค่าไอโอยูพบว่าได้มากกว่า 0.9 บ่งว่ารอยพิมพ์ฐานอาคารที่สกัดได้ใกล้เคียงกับการคัดลอกเส้นด้วยตนเอง เป็นทางเลือกหนึ่งในการสร้างรอยพิมพ์ฐานอาคารที่มีความถูกต้องเชิงตำแหน่งและความสมบูรณ์ของรอยพิมพ์ฐานอาคาร ได้ดี

คำสำคัญ: การเรียนรู้เชิงลึก; การตรวจจับวัตถุ; การสกัดรอยพิมพ์ฐานอาคาร; ภาพออร์โธจริง; แบบจำลองการแยกส่วนทุกสิ่งเชิงภูมิปริภูมิ

ABSTRACT

Building data is essential for visualizing the population or environment in some areas. The classic method of creating data layers is delineation from satellite imagery or orthophoto, which takes more human and time resources. Presently, there are free and open building footprints, but the completeness of buildings is not great, especially in urban areas. Recently, there is an available geospatial segment anything model that we can experiment with, and there is an urban mapping project from an unmanned aerial vehicle that processed true orthophoto with a ground sampling distance of 5 centimeters. Accordingly, there was a study of

Thepchai Srinoi^{1*}, Thirawat Bannakulpiphat² and Pisan Santitamnont³

¹Graduate Student, Department of Survey Engineering, Faculty of Engineering, Chulalongkorn University, Thailand.

²Researcher, Department of Survey Engineering, Faculty of Engineering, Chulalongkorn University, Thailand.

³Associate Professor, Department of Survey Engineering, Faculty of Engineering, Chulalongkorn University, Thailand.

building footprint extraction with this model in the faculty of science and engineering at Chulalongkorn University. We do a completeness assessment of the building footprint with Intersection over Union (IOU). The research showed that IOUs are higher than 0.9 in both areas. Those said that extracted footprint data is nearly the same as self-delineation, which is the alternative way for building footprint generation with high positioning accuracy and completeness.

KEYWORDS: Deep Learning; Object Detection; Building Footprint Extraction; True Orthophoto; Geospatial Segment Anything Model (GeoSAM)

1. บทนำ

อาคารเป็นสถานที่ที่รวมมนุษย์อยู่ด้วยกัน มีพื้นที่จำกัดเพื่อดำรงชีวิต ทำกิจกรรมส่วนตัว ประกอบอาชีพของตน ปกป้องตนเองจากสภาพแวดล้อมหรือสภาพอากาศภายนอก มนุษย์มีการสร้างอาคารทั้งในชนบทหรือในเมือง แต่ละหลังมีการใช้ประโยชน์ที่หลากหลาย เช่น แหล่งที่อยู่อาศัย อุตสาหกรรม พาณิชยกรรม สถานที่ราชการ ศาสนสถาน เป็นต้น สำหรับการกระจายตัวของอาคารก็มีลักษณะการสร้างที่หลากหลายรูปแบบ เช่น การสร้างเรียงเป็นแนวเส้นตรงตามเส้นทางการคมนาคมหรือแหล่งน้ำ การสร้างแบบกระจัดในพื้นที่เมือง การสร้างกระจายไปตามพื้นที่การเกษตรในชนบท โดยข้อมูลตำแหน่งอาคารสามารถบอกการกระจายตัว หรือความหนาแน่นของประชากรได้ บอกสภาพแวดล้อมของเมืองหรือชนบท การใช้พลังงานในครัวเรือน [1] ข้อมูลที่ได้นำไปประยุกต์ใช้ในการนำทางสามมิติ และการแจ้งเตือนภัยฉุกเฉิน [2] นำไปใช้วางแผนหรือการบูรณาการสิ่งต่าง ๆ ภายในเมือง [3]

การจัดทำฐานข้อมูลรอยพิมพ์ฐานอาคารในพื้นที่หนึ่ง โดยเฉพาะในพื้นที่เมืองจำเป็นต้องใช้ทรัพยากรจำนวนมาก ทั้งแรงงานมนุษย์ เทคโนโลยี เวลา โดยเฉพาะวิธีการดั้งเดิมคือ การสร้างข้อมูลเวกเตอร์รูปปิด (Polygon) ของอาคารจากภาพถ่ายทางอากาศ หรือการสำรวจจริงวัดรายละเอียดภาคสนาม ภายหลังมีการวิจัยการสกัดข้อมูลรอยพิมพ์ฐานอาคารโดยหาแบบจำลองทางปัญญาประดิษฐ์ (Artificial Intelligence) ที่เหมาะสมเป็นจำนวนมาก แต่การทดลองใช้นั้นมีข้อจำกัดที่ทรัพยากรคอมพิวเตอร์ต้องมีประสิทธิภาพสูง และผู้ใช้งานต้องมีความรู้เพื่อปรับค่าพารามิเตอร์ให้ได้ผลลัพธ์ที่ต้องการ



รูปที่ 1 ตัวอย่างรอยพิมพ์ฐานอาคารจาก Microsoft (ซ้าย) และจาก Google (ขวา) บริเวณอนุสาวรีย์ชัยสมรภูมิ

ปัจจุบันมีการใช้เทคนิคการเรียนรู้เชิงลึก (Deep Learning) ร่วมกับข้อมูลดาวเทียมเชิงทัศนความละเอียดสูงและข้อมูลตัวอย่างรอยพิมพ์ฐานอาคารในการสร้างแบบจำลองรอยพิมพ์ฐานอาคาร (Building Footprint) จำนวนมากเกือบทั่วทั้งโลกจากผลิตภัณฑ์ของ

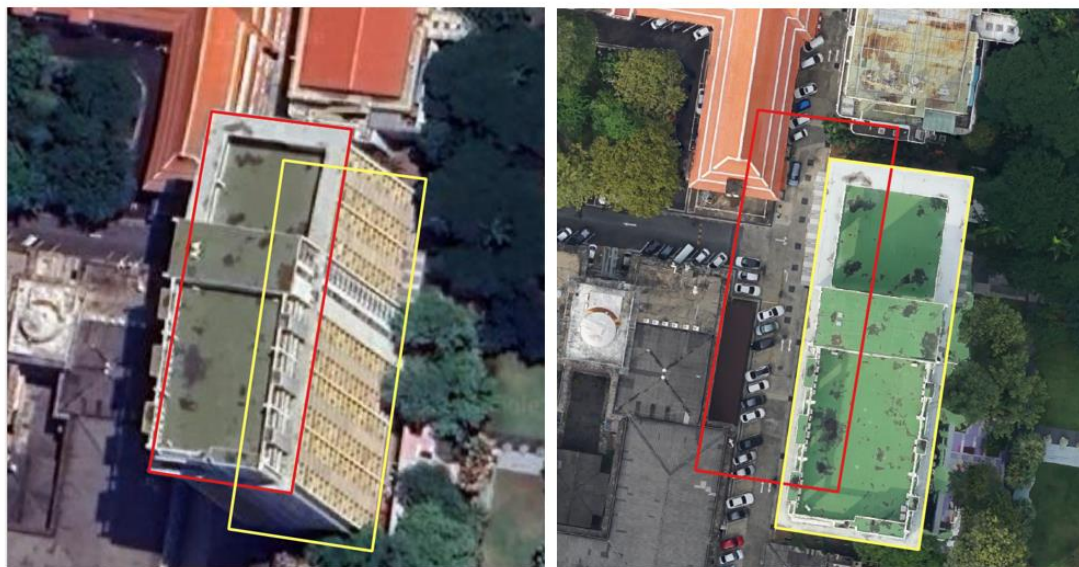
Microsoft Global Building Footprint [4] หรือ Google Open Building [5] ทั้งสองแหล่งนี้ครอบคลุมประเทศไทย ผู้ใช้ในประเทศสามารถดาวน์โหลดข้อมูลออกมาเป็นข้อมูลเวกเตอร์ของรอยพิมพ์ฐานอาคารได้ ตัวอย่างแสดงดังรูปที่ 1

แต่ทว่าข้อมูลดาวเทียมดังกล่าวสามารถพบปรากฏการณ์การเลื่อนของจุดภาพอันเนื่องมาจากความสูงต่างของสิ่งปลูกสร้าง (Relief Displacement) พบได้ตามอาคารสูงในพื้นที่เมือง สมมติจุดสองจุดบนอาคาร จุดแรกอยู่ที่ฐานอาคาร จุดที่สองอยู่ที่เพดานอาคาร สองจุดนี้กำหนดให้มีพิกัดทางราบเท่ากัน แต่มีพิกัดทางดิ่งต่างกันเพราะความสูงต่างกัน เมื่อบันทึกภาพแล้วสองจุดนี้จะปรากฏที่พิกัดภาพสองมิติที่แตกต่างกัน เกิดเป็นภาพอาคารเอียง การนำภาพที่บันทึกไปใช้ทำแผนที่ที่มีความถูกต้องจำเป็นต้องทำการประมวลผลภาพออร์โธจริง (True Orthophoto) ซึ่งใช้ข้อมูลภาพถ่ายทางอากาศ มีการคำนวณองค์ประกอบภายนอกของแต่ละภาพ (Exterior Orientation) ซึ่งบอกตำแหน่งถ่ายภาพและการวางตัวของภาพเทียบกับกรอบพิกัดภาคพื้นดิน และมีการทำแบบจำลองความสูงเชิงพื้นผิว (Digital Terrain Model) สร้างข้อมูลราสเตอร์กริดในพื้นที่งาน นำค่าพิกัดทางราบในแต่ละช่องกริดทำการคำนวณค่าสีจากภาพผ่านชุดสมการสภาวะร่วมเส้น เพื่อเติมลงในช่องกริดสร้างเป็นภาพออร์โธจริงขึ้น ในลำดับการคำนวณนั้นหากพิกัดทางราบในกริดตรงกับตำแหน่งของอาคารสูงเดิม อัลกอริทึมจะไปค้นค่าทางดิ่งจากแบบจำลองความสูงเชิงพื้นผิวซึ่งได้ความสูงของเพดานอาคาร ไม่ใช่ความสูงของฐานอาคาร ทำให้ได้ภาพอาคารที่ไม่เอียง แสดงดังรูปที่ 2



รูปที่ 2 ภาพจากดาวเทียม Google Earth (ซ้าย) และภาพออร์โธจริงจากอากาศยานไร้คนขับ (ขวา)

การมีภาพออร์โธจริงทำให้สามารถใช้เป็นแผนที่ฐาน (Base map) ของการคัดลอกลายเส้น (Delineation) [6] สร้างรอยพิมพ์ฐานอาคารได้โดยการคัดลอกบริเวณขอบหลังคาอาคาร (Roof line) แล้วใช้เป็นรูปรอยพิมพ์ฐานอาคารได้ทันที เพราะสองรอยนี้ตรงกัน แตกต่างจากการใช้ข้อมูลดาวเทียมที่ไม่ผ่านการทำออร์โธจริงหรือผ่านการทำออร์โธพื้นดิน (Ground Orthophoto) ซึ่งใช้แบบจำลองความสูงพื้นภูมิประเทศ (Digital Terrain Model) แทนการใช้ความสูงพื้นผิว ทำให้ตำแหน่งขอบหลังคาอาคารไม่ตรงกับรอยพิมพ์ฐาน ต้องนำเส้นขอบหลังคาเลื่อนลงมายังบริเวณที่คาดว่าเป็นฐานอาคารแล้วทุกครั้งเพื่อให้ได้รอยพิมพ์ฐานอาคาร ตัวอย่างการคัดลอกลายอาคารจากออร์โธจริงและพื้นดินแสดงดังรูปที่ 3

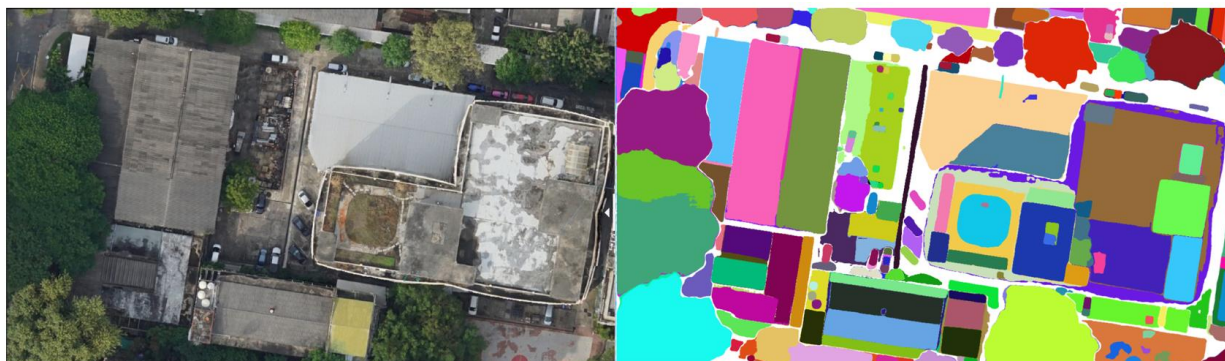


รูปที่ 3 การคัดลอกลายขอบหลังคาสีแดงจากภาพถ่ายดาวเทียม Google Earth (ซ้าย) และรอยพิมพ์ฐานอาคารจากภาพออร์โธจริงจากอากาศยานไร้คนขับ (ขวา) พบว่าขอบหลังคาต้องเลื่อนมาที่ฐานอาคารในกรณีภาพถ่ายดาวเทียม

ล่าสุดมีการพัฒนาแบบจำลองการแยกทุกอย่างเชิงภูมิปริภูมิ (Geospatial Segment Anything Model : GeoSAM) [7] มาจากแบบจำลองการแยกทุกอย่างของทีมงานพัฒนา Meta Facebook [8] ซึ่งผู้ใช้งานรูปภาพเข้าแบบจำลองแล้วระบบจะทำการแยกวัตถุที่พบได้ในภาพออกมา ภายหลังจากมีการปรับใช้ทำงานได้กับข้อมูลแผนที่ มีรูปแบบให้ใช้งานได้ทั้ง การนำภาพถ่ายดาวเทียมเข้ามาแยกวัตถุจากภาพ (Satellite Image Segmentation) แสดงในรูปที่ 4 การแยกวัตถุที่สนใจโดยเฉพาะทั้งแบบที่ใช้ป้อนพิกัดของวัตถุ (Point Prompt) และแบบที่ป้อนคำศัพท์ (Text Prompt) นอกจากนี้ยังมีการนำไปใช้กับข้อมูลพอยต์คลาวด์ (Point Cloud) จากผลผลิตของอากาศยานไร้คนขับ (Unmanned Aerial Vehicle) หรือเลเซอร์สแกน (Laser Scanning) เพื่อแยกกลุ่มพอยต์คลาวด์ [9] แบบจำลองการแยกส่วนทุกอย่างเชิงภูมิปริภูมิดังกล่าวสามารถนำมาใช้ร่วมกับภาพออร์โธจริงที่ความละเอียดแบบ Ground Sampling Distance (GSD) 5 เซนติเมตร ที่ได้จากโครงการแผนที่สามมิติของเมืองด้วยอากาศยานไร้คนขับของการไฟฟ้านครหลวง ในพื้นที่เขตปทุมวัน ส่วนหนึ่งของเขตวัฒนาและคลองเตย กรุงเทพมหานคร ศึกษาแนวทางการใช้แบบจำลองนี้ในการทำงานกับภาพออร์โธจริง เพื่อสร้างข้อมูลรอยพิมพ์ฐานอาคารในพื้นที่ตัวอย่างคณะวิทยาศาสตร์และวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย และตรวจสอบกับข้อมูลจากการคัดลอกลายเส้นด้วยตนเอง

2. การศึกษาแบบจำลองการแยกทุกอย่างเชิงภูมิปริภูมิ

แบบจำลองการแยกทุกอย่างเชิงภูมิปริภูมิมี 3 องค์ประกอบ ได้แก่ การแปลงภาพที่นำเข้าแบบจำลอง (Image Encoder) การแปลงตัวช่วยที่นำเข้าแบบจำลอง (Flexible prompt encoder) และการแปลงหน้ากากของวัตถุที่ได้จากแบบจำลอง (Fast Mask Decoder) หน้ากากในที่นี้แสดงด้วยข้อมูลราสเตอร์ที่บอกริเวณวัตถุที่แยกออกมาได้เป็นผลลัพธ์ปลายทาง

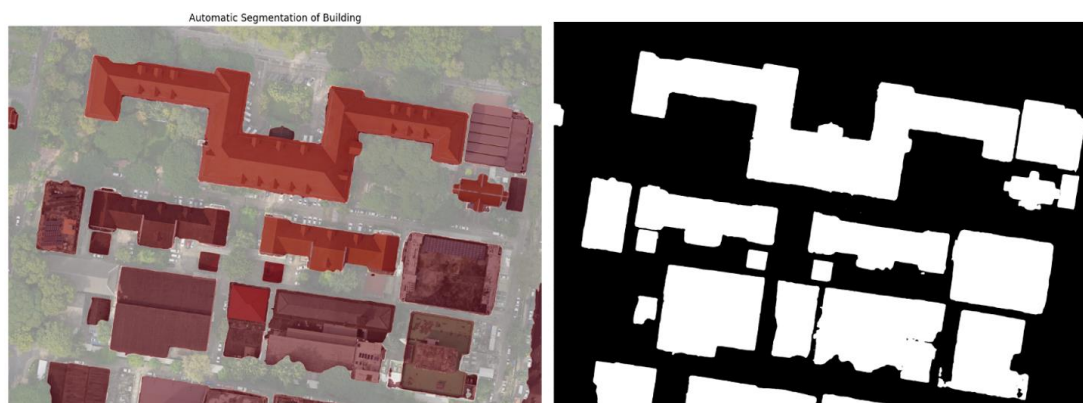


รูปที่ 4 ตัวอย่างการแยกแยะทุกสิ่งอย่างจากภาพออร์โธจริง บริเวณหน้าอาคารเรียนศัลยวิชานิเทศ จุฬาลงกรณ์มหาวิทยาลัย

แนวทางสกัดเฉพาะวัตถุที่สนใจ เช่น รอยพิมพ์ฐานอาคารจากภาพออร์โธจริงด้วยแบบจำลองการแยกส่วนทุกสิ่งเชิงภูมิปริภูมิ สามารถทำได้ 2 วิธี คือ การสกัดด้วยข้อความ (text prompt) และ การสกัดด้วยจุด (point prompt)

2.1 การสกัดด้วยข้อความ (text prompt)

การสกัดด้วยข้อความ ใช้แบบจำลอง LangSAM กับภาพออร์โธจริงในพื้นที่ศึกษา สกัดรอยพิมพ์ฐานอาคารจากโดยการใส่ `text_prompt = "building"` ลงในฟังก์ชัน `predict` พร้อมทั้งกรอกพารามิเตอร์ของแบบจำลองสองอย่างคือ `box_threshold` และ `text_threshold` วิธีนี้เป็นการใช้แบบจำลองที่ผ่านการเรียนรู้แล้ว ผู้ใช้เพียงกรอกข้อความของวัตถุที่สนใจเพื่อให้ระบบทำการสกัดบริเวณที่สอดคล้องกับข้อความดังกล่าว งานวิจัยนี้สนใจการปรับ พารามิเตอร์ `box_threshold` ดยพบว่าค่าที่ทดลองได้อยู่ที่ 0.1 ถึง 0.5 ยิ่งใช้ค่าน้อยลงจะตรวจพบวัตถุเล็กลงได้ แต่ใช้ทรัพยากรเครื่องจำนวนมาก การใช้ค่ามากมีโอกาสนี้ไม่พบวัตถุที่สนใจได้ ผลลัพธ์เริ่มต้นเป็นข้อมูลราสเตอร์ของบริเวณอาคาร แล้วแปลงเป็นข้อมูลเวกเตอร์รูปปิดของอาคาร แสดงดังรูปที่ 5



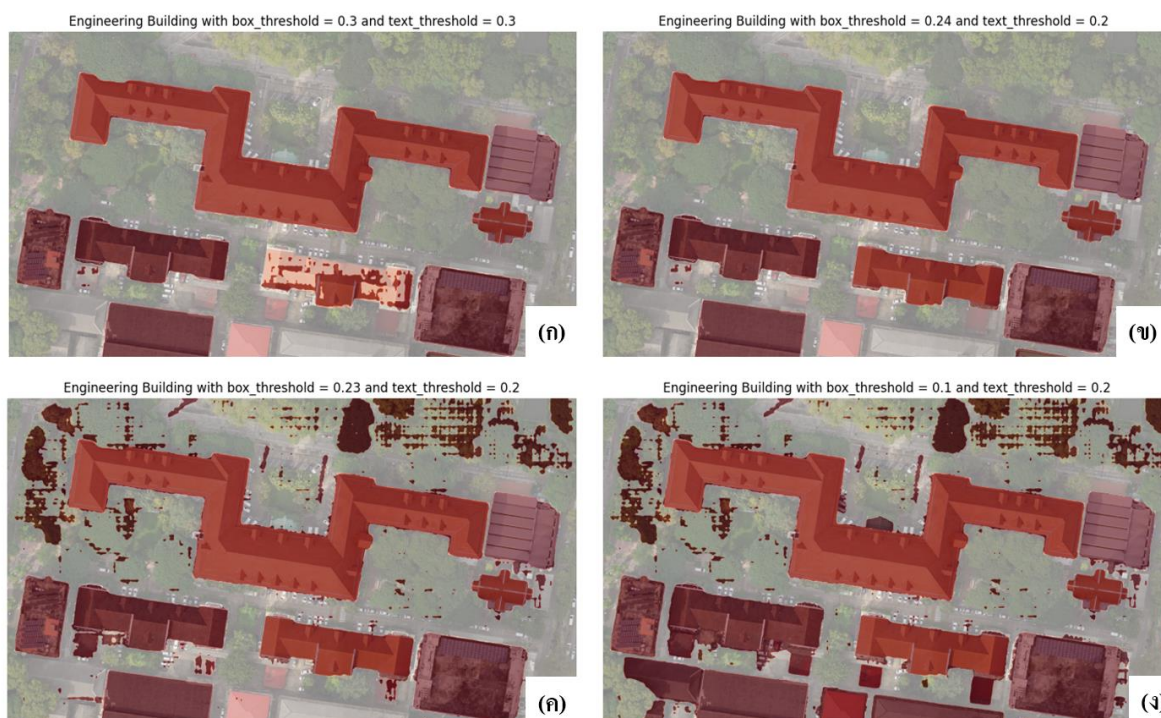
รูปที่ 5 ผลการสกัดรอยพิมพ์ฐานอาคารแสดงผลกับภาพออร์โธจริง (ซ้าย) แสดงแบบขาวดำก่อนแปลงเป็นเวกเตอร์ (ขวา)

ข้อจำกัดที่พบได้คือรอยพิมพ์ฐานอาคารไม่สมบูรณ์ถ้ามีต้นไม้บัง ขอบอาคารจะหยักตามที่มีการบัง แตกต่างจากการคัดลอกด้วยมือซึ่งมนุษย์เข้าใจได้ว่ารูปร่างขีดได้ตรงไม่ต้องสนเรื่องการบังจากรูปที่ 6 พบว่า `box_threshold` ไม่น้อยกว่า 0.24 สามารถสกัดอาคารที่มีขนาดใหญ่ออกมาได้จำนวนหนึ่ง หากลดค่าดังกล่าวลงมานอกจากจะตรวจพบวัตถุที่ขนาดเล็กลงแล้วยังได้วัตถุที่ไม่

พึงประสงค์ได้เช่น ส่วนหนึ่งของต้นไม้ ลานจอดรถ ส่งผลให้เมื่อแปลงเป็นเวกเตอร์แล้วต้องจัดการลบส่วนที่ไม่ต้องการออก ภายหลัง ทั้งนี้การปรับ text_threshold ไม่ทำให้ผลเปลี่ยนแปลง จึงกำหนดให้ใช้ค่า 0.2 ตลอดการวิจัย

2.2 การสกัดด้วยจุด (point prompt)

การสกัดด้วยจุด ใช้แบบจำลอง SAM กับภาพออร์โธจิริงในพื้นที่ศึกษา ผู้ใช้ต้องเตรียมคำพิกัดภูมิศาสตร์ตัวแทนของวัตถุที่ต้องการ ป้อนเข้าโปรแกรมเพื่อสกัดรอยพิมพ์ฐานอาคารตรงตำแหน่งที่กำหนด จากรูปที่ 7 พบว่ากรณีป้อนหนึ่งจุด ระบบสามารถสกัดอาคารได้อย่างสมบูรณ์ ถ้าเลือกตำแหน่งตัวแทนได้เหมาะสม การสกัดทำได้ดีเมื่ออาคารมีเพดานเรียบ หรือมีหลังคาเดียว หากเป็นอาคารที่มีหลายหลังคา ระบบจะติกรอบเพียงหลังคาเดียวได้ หรือมีชั้นสูงแทรกกลาง ระบบจะติกรอบอาคารเพียงตรงกลางได้ แต่เมื่อเพิ่มจุดที่สองเป็นต้นไป การสกัดอาคารจะเริ่มไม่สมบูรณ์ บางหลังขึ้นผลมาไม่เต็มทัวอาคาร ถ้าเพิ่มตัวแทนจุดอาคารทุกจุดลงในแบบจำลอง รอยพิมพ์ฐานอาคารยังคงไม่สมบูรณ์ จำเป็นต้องคัดกรองปรับรอยพิมพ์ฐานอาคารในภายหลัง

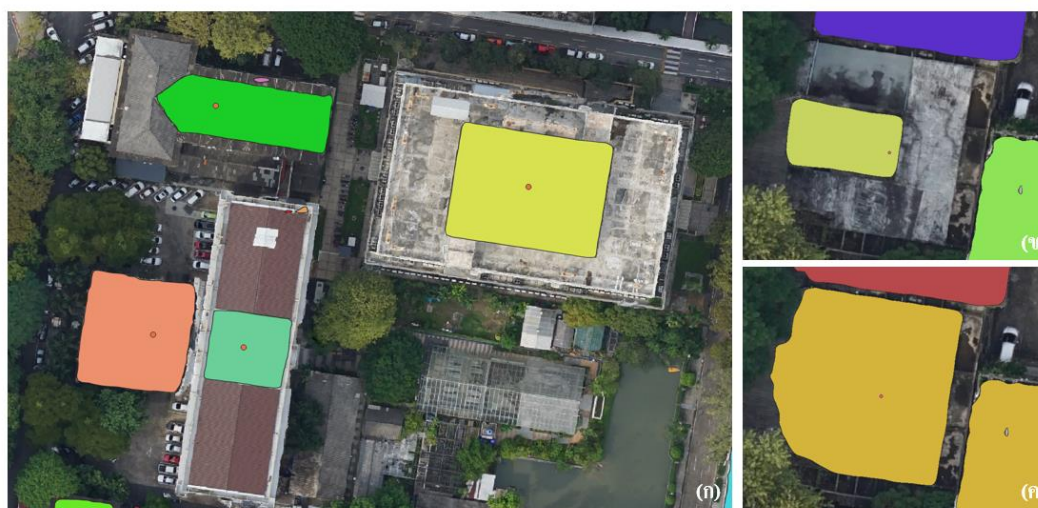


รูปที่ 6 ตัวอย่างการสกัดรอยพิมพ์ฐานอาคารเมื่อปรับ box threshold เป็น (ก) 0.3 (ข) 0.24 (ค) 0.23 และ (ง) 0.1

ข้อสรุปคือเริ่มจากสกัดด้วยข้อความเพื่อสกัดอาคารขนาดใหญ่ก่อน จากนั้นจึงใช้การสกัดด้วยจุดเพื่อเติมเต็มข้อมูลอาคารขนาดเล็กในภายหลัง ข้อเสนอแนะคือควรป้อนทีละจุดและให้สกัดรายอาคารจากนั้นแปลงเป็นข้อมูลเวกเตอร์แล้วนำมารวมกันเป็นข้อมูลชุดเดียว โดยตัวแทนจุดควรอยู่บนพื้นหลังที่เป็นส่วนใหญ่ของอาคาร เพื่อให้ได้ทั้งหลังคา ดังรูปที่ 8



รูปที่ 7 ตัวอย่างการสกัดรอยพิมพ์ฐานอาคารแบบจุด กรณีสกัดจุดเดียว (ซ้าย) กรณีสกัดหลายจุด (ขวา)



รูปที่ 8 ตัวอย่างผลการสกัดอาคาร กรณีเลือกขอดกลางอาคารบางครั้งทำให้สกัดได้เนื้อที่อาคารไม่สมบูรณ์ (ก และ ข) และผลการสกัดอาคารเมื่อเลือกตัวแทนพื้นที่อาคารส่วนใหญ่ (ค)

ข้อจำกัดของวิธีนี้คือมีโอกาสที่จะได้รอยพิมพ์ฐานอาคารซ้ำซ้อนกัน เกิดได้ในกรณีอาคารติดกัน มีพิกัดตัวแทนใกล้เคียงกัน หากนำพิกัดเข้าแบบจำลองแล้วสกัดออกมาเป็นรูปเดียวกัน จะเกิดการบันทึกซ้ำได้ จำเป็นต้องตรวจสอบผลด้วยตนเองอีกครั้ง ดังนั้นสามารถสรุปความแตกต่างของวิธีสกัดแบบข้อความและจุดได้ ดังตารางที่ 1 และรูปแบบแผนการสกัดรอยพิมพ์ฐานอาคารบริเวณพื้นที่คณะวิทยาศาสตร์และวิศวกรรมศาสตร์ในการศึกษานี้ ดังรูปที่ 9

ตารางที่ 1 การเปรียบเทียบการสกัดวัตถุจากแบบจำลองแยกแยะทุกสิ่งอย่าง

ประเด็นเปรียบเทียบ	การสกัดด้วยข้อความ (Text Prompt)	การสกัดด้วยจุด (Point Prompt)
ข้อมูลนำเข้า	- ข้อมูลภาพที่ตรงคำพิกัดโลก - ข้อความเพื่อสกัดวัตถุ	- ข้อมูลภาพที่ตรงคำพิกัดโลก - คำพิกัดภูมิศาสตร์ของจุดตัวแทนอาคาร
พารามิเตอร์การสกัดวัตถุ	box_threshold และ text_threshold	ไม่มี
ลักษณะผลลัพธ์	ข้อมูลราสเตอร์บอกขอบเขตวัตถุ สามารถแปลงเป็นข้อมูลเวกเตอร์รายพิมพ์ฐานอาคารได้	ข้อมูลราสเตอร์บอกขอบเขตวัตถุสามารถแปลงเป็นข้อมูลเวกเตอร์รายพิมพ์ฐานอาคารได้
ข้อได้เปรียบ	การสกัดวัตถุที่มีขนาดใหญ่ จำนวนมาก เมื่อใช้ box_threshold ที่เหมาะสม	การสกัดวัตถุที่เฉพาะตำแหน่งที่สนใจ ให้ผลที่สมบูรณ์ กรณีป้อนจุดเดียว โดยเฉพาะวัตถุที่มีขนาดเล็ก
ข้อจำกัด	การปรับ box_threshold น้อยเกินไปทำให้สกัดวัตถุที่สนใจในขนาดเล็กพร้อมกับวัตถุไม่พึงประสงค์และใช้ทรัพยากรเครื่องจำนวนมาก	ผู้ใช้งานกรอกพิกัดด้วยตนเอง ถ้ากรอกพิกัดจำนวนมาก สามารถสกัดวัตถุจำนวนมากได้แต่ความสมบูรณ์ไม่ดีเท่าผลที่ได้จากการกรอกพิกัดจุดเดียว



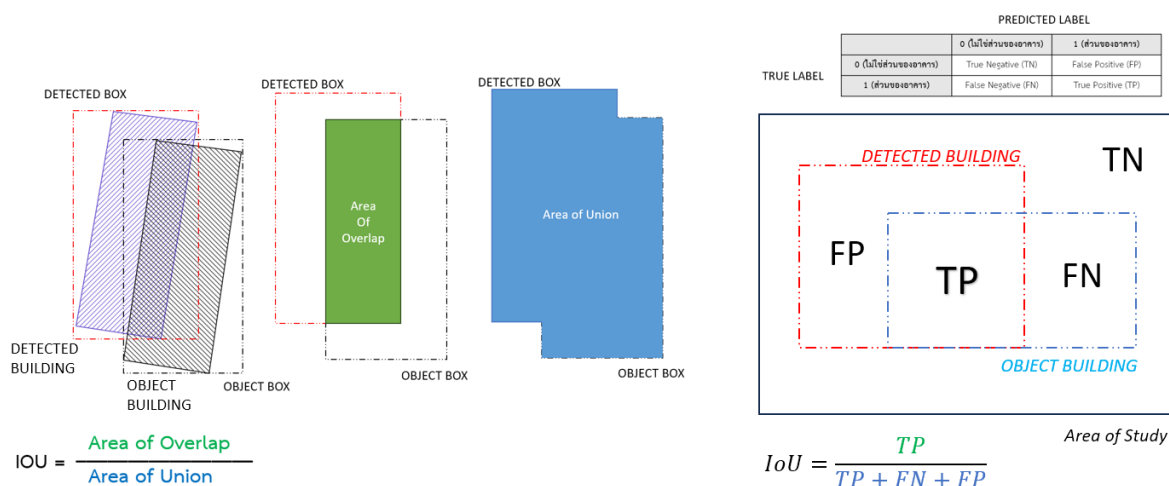
รูปที่ 9 แผนการสกัดรายพิมพ์ฐานอาคารในการศึกษา

3. การตรวจสอบความสมบูรณ์ของรายพิมพ์ฐานอาคาร

ดำเนินการสกัดรายพิมพ์ฐานอาคารตามข้อสรุปดังกล่าวในพื้นที่คณะวิศวกรรมศาสตร์และวิทยาศาสตร์ ตรวจสอบความสมบูรณ์กับข้อมูลตัวอย่างรายพิมพ์ฐานอาคารจากการลอกด้วยตนเองกับผลจาก Microsoft, Google และ GeoSAM ซึ่งสกัดจากภาพออร์โธจริง ซึ่งงานวิจัยนี้ใช้ตัวชี้วัด Intersection Over Union (IoU) ซึ่งบอกสัดส่วนของพื้นที่ซ้อนทับกัน (Area of Overlap) กับพื้นที่รวม (Area of Union) เมื่อใช้ข้อมูลเวกเตอร์ คัดจากพื้นที่สี่เหลี่ยมรอบวัตถุที่สกัดได้จากแบบจำลองเทียบกับสี่เหลี่ยมรอบวัตถุที่เป็นผลคำตอบ นอกจากการคำนวณจากข้อมูลเวกเตอร์ สามารถแปลงเป็นข้อมูลราสเตอร์แล้วเทียบราสเตอร์ของรายพิมพ์ฐานอาคารจากแบบจำลองเทียบกับแผนที่ละจุดภาพ (pixel) ได้ ค่าในจุดภาพเป็น 1 ถ้าตรงกับรายพิมพ์ฐานอาคาร และเป็น 0 ถ้าไม่ตรง

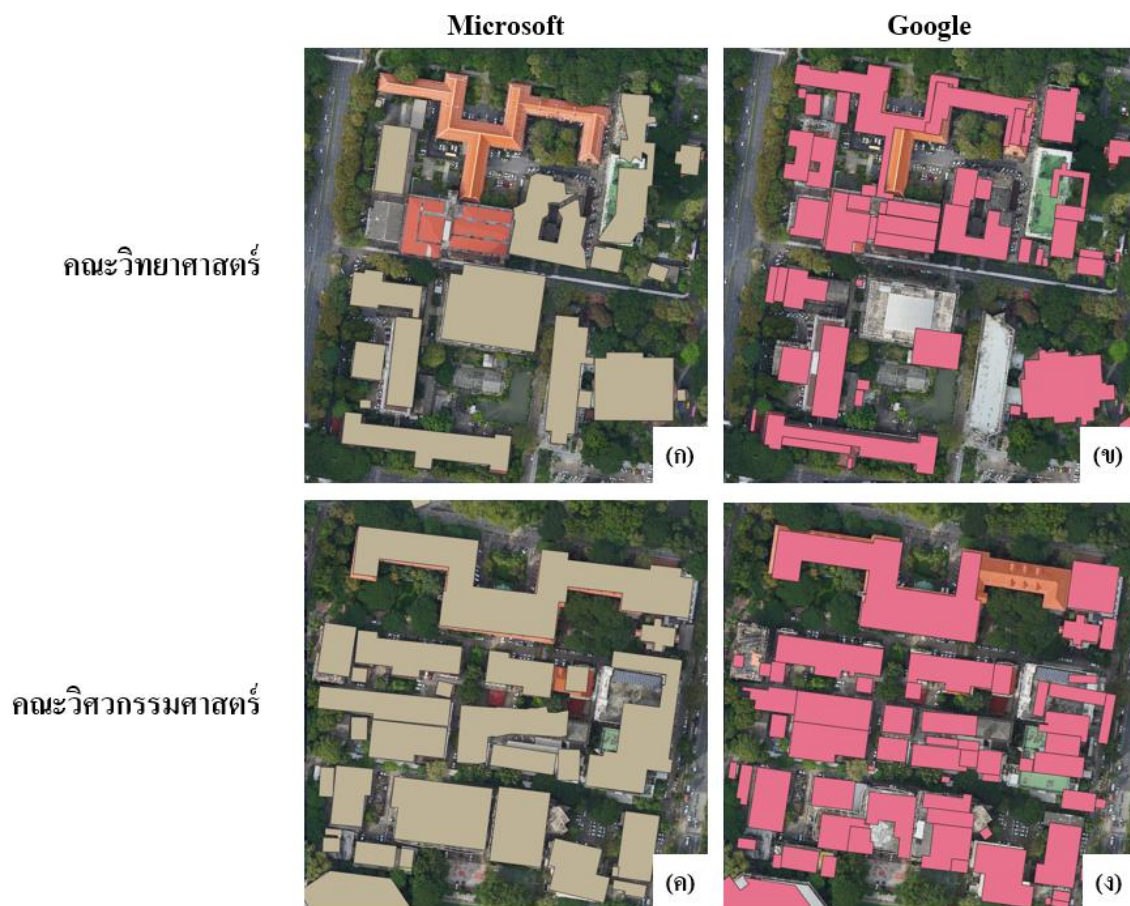
กับรอยพิมพ์ฐานอาคาร การวิเคราะห์เชิงเรขาคณิตนี้ ทำให้ทราบส่วนที่ซ้อนกัน (ค่าจุดภาพจากแบบจำลองและเฉลยเท่ากัน เท่ากับ 1) และส่วนที่รวมกันได้ (ค่าจุดภาพจากแบบจำลองและเฉลย ทั้งสองต้องไม่เป็น 0) สร้างตาราง Confusion Matrix แล้วทำการคำนวณตัวชี้วัด IoU ได้ ลักษณะการคำนวณ IoU ทั้งแบบข้อมูลเวกเตอร์และราสเตอร์แสดงดังรูปที่ 10

จากการศึกษาข้อมูลรอยพิมพ์ฐานอาคารจาก Microsoft และ Google ดังรูปที่ 11 พบว่าทั้งสองผลผลิตยังได้รอยพิมพ์ฐานอาคารไม่ครบถ้วน มีอาคารหลายหลังที่ยังไม่สมบูรณ์ ดังนั้นการใช้ภาพออร์โธจริงสักรอยพิมพ์ฐานอาคารมีความจำเป็นในการทำในพื้นที่เมือง เพื่อให้ได้รอยพิมพ์ฐานอาคารที่มีความถูกต้องทางเรขาคณิต ทั้งนี้ความคลาดเคลื่อนของตำแหน่งรอยพิมพ์ฐานอาคารอาจมาจากพื้นหลักฐานของรอยพิมพ์ฐานอาคารที่แตกต่างกับภาพออร์โธจริง ลักษณะของการสักรอยพิมพ์ฐานอาคารจากลักษณะแผ่นหลังคา (ตัวอย่างบริเวณอาคารตรงกลางภาพพื้นที่คณะวิทยาศาสตร์) พบได้จากรอยพิมพ์ฐานอาคารของ Google ซึ่งพบได้ในการสักรั้วด้วยจุดกับภาพออร์โธด้วย



รูปที่ 10 การคำนวณ Intersection Over Union (IoU) ทั้งจากข้อมูลเวกเตอร์และราสเตอร์

ผลการสักรอยพิมพ์ฐานอาคารด้วยข้อความดังรูปที่ 12 (คอลัมน์ซ้าย) พบว่าสามารถให้รอยพิมพ์ฐานอาคารของคณะวิทยาศาสตร์ได้เกือบครบถ้วน สามารถสักรั้วอาคารซึ่งมีหลังคามากมายในหนึ่งอาคารได้ ส่วนคณะวิศวกรรมศาสตร์สามารถดึงรอยพิมพ์ฐานอาคารได้จำนวนหนึ่ง มีอาคารตรงกลางที่ได้รับอิทธิพลจากต้นไม้บัง ยังมีอาคารขนาดเล็กที่ไม่สามารถขึ้นรูปได้ ต้องใช้การสักรั้วด้วยจุดเสริมเข้าไป ผลการเติมเต็มรอยพิมพ์ฐานอาคารด้วยการสักรั้วด้วยจุดจากแบบจำลอง แสดงดังรูปที่ 12 (คอลัมน์กลาง) เปรียบเทียบกับรอยพิมพ์ฐานอาคารซึ่งสร้างด้วยการคัดลอกจากภาพออร์โธจริง แสดงดังรูปที่ 12 (คอลัมน์ขวา) ทำการแปลงข้อมูลเวกเตอร์อาคารเป็นราสเตอร์แสดงรอยพิมพ์ฐานอาคารเพื่อคำนวณ IoU เพื่อตรวจสอบความสมบูรณ์ของรอยพิมพ์ฐานอาคาร ผลการประเมิน IoU ทั้งจาก Microsoft, Google และ GeoSAM จากภาพออร์โธจริงแสดงดังตารางที่ 2



รูปที่ 11 รอยพิมพ์ฐานอาคารจาก Microsoft (ซ้าย) เทียบกับ Google (ขวา) ในคณะวิทยาศาสตร์ (แถวบน ก และ ข) และคณะวิศวกรรมศาสตร์ (แถวล่าง ค และ ง)

ตารางที่ 2 ผลการตรวจสอบความสมบูรณ์ของรอยพิมพ์ฐานอาคาร

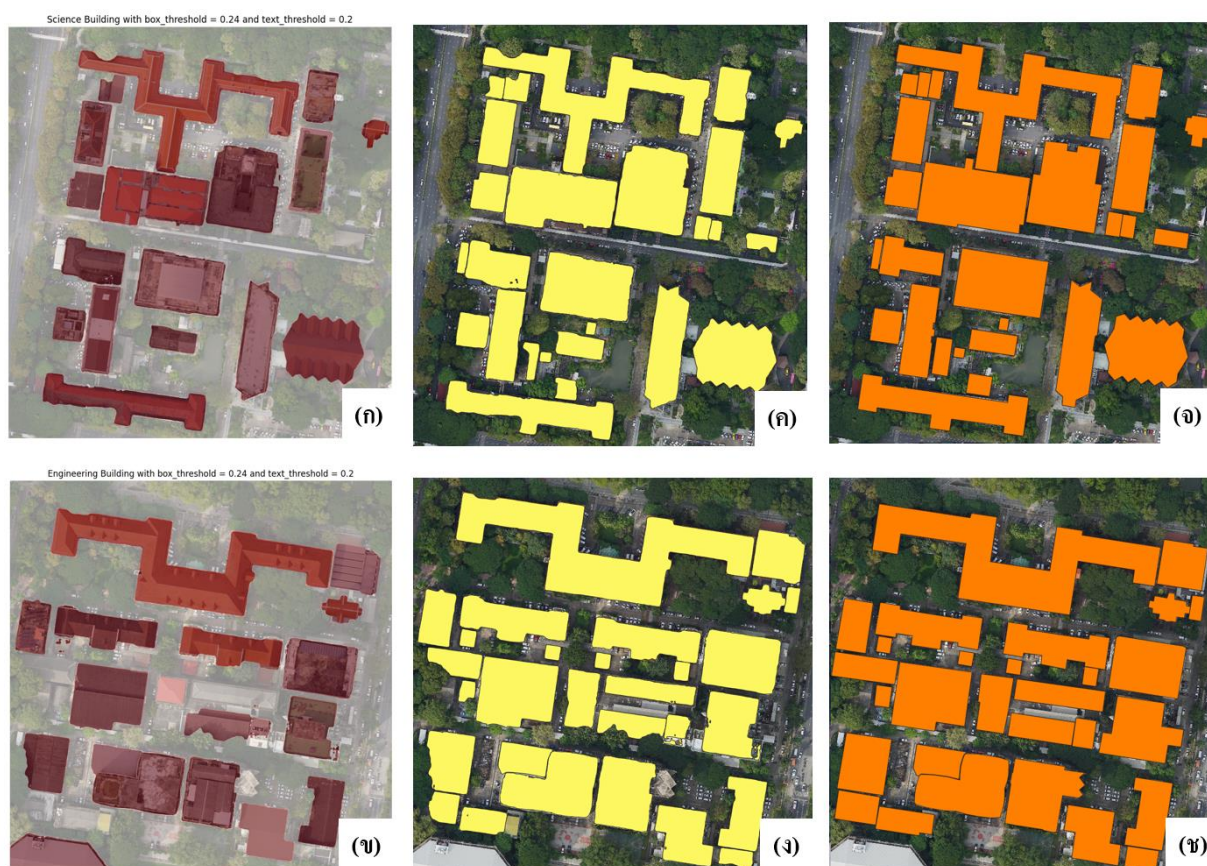
พื้นที่ศึกษา	คณะวิทยาศาสตร์	คณะวิศวกรรมศาสตร์
Microsoft Global Building Footprint	0.5136	0.7277
Google Open Building	0.5168	0.6369
GeoSAM from True Orthophoto	0.9196	0.9113

ค่า IoU จากผลผลิตของ Microsoft, Google มีค่าประมาณ 0.51 สำหรับคณะวิทยาศาสตร์ และ 0.72 กับ 0.63 ตามลำดับ สำหรับคณะวิศวกรรมศาสตร์ พิจารณาจากรูปที่ 11 ในพื้นที่คณะวิทยาศาสตร์ จะพบว่ารอยพิมพ์ฐานอาคาร 2 หลังใหญ่ที่ด้านล่างของถนนตรงกลาง Microsoft สก๊ตได้ แต่ Google ไม่สามารถสก๊ตได้ และมีรอยพิมพ์ฐานอาคารส่วนบนของถนนตรงกลางที่ Microsoft สก๊ตไม่ได้ แต่ Google สก๊ตได้ การมีส่วนที่ขาดและส่วนที่เกินเมื่อเทียบกันสองแหล่ง ในสัดส่วนเท่ากันทำให้ค่า IoU มีค่าใกล้เคียงกัน ส่วนของคณะวิศวกรรมศาสตร์ ผลการสก๊ตจาก Microsoft ให้รอยพิมพ์ฐานอาคารที่มีเนื้อที่กว้างได้มากกว่า Google จึงทำให้ IoU สูงกว่า ทั้งนี้ผลลัพธ์จากวิธีล่าสุดที่ได้เป็นที่น่าสนใจ IoU จำนวนได้มากกว่า 0.9 แสดงถึงศักยภาพของแบบจำลอง GeoSAM ใน

การใช้สักรอยพิมพ์ฐานอาคารแทนที่การคัดลอกด้วยตนเองได้ดี แต่ทว่าต้องมีการตรวจสอบผิดพลาดด้วยตนเองภายหลัง กรณีการบดบังด้วยต้นไม้ หรือกรณีอาคารที่ติดกันแล้วระบบมองรวมเป็นหลังคาเดียวกัน การเปรียบเทียบตัวอย่างรอยพิมพ์ฐานอาคารจาก Microsoft, Google, GeoSAM และจากการลอกถ่าย แสดงดังตารางที่ 3

4. สรุปผลและข้อเสนอแนะ

การสักรอยพิมพ์ฐานอาคารด้วยแบบจำลอง GeoSAM กับภาพออร์โธจริง เป็นตัวเลือกหนึ่งในการทำงานเพื่อสร้างชั้นข้อมูลอาคารที่มีความสมบูรณ์ใกล้เคียงกับวิธีดั้งเดิมด้วยการคัดลอกด้วยตนเอง แทนที่การใช้ข้อมูลรูปรอยพิมพ์ฐานอาคารที่เปิดเผยให้ใช้ได้โดยไม่เสียค่าใช้จ่าย ผลการวิเคราะห์ความสมบูรณ์ด้วยตัวชี้วัด Intersection over Union (IoU) พบว่าได้ค่า 0.9196 สำหรับคณะวิทยาศาสตร์ และ 0.9113 สำหรับคณะวิศวกรรมศาสตร์ ดีกว่าผลผลิตรอยพิมพ์ฐานอาคารที่มีมาก่อนหน้า ปัจจัยส่งผลให้ค่าดังกล่าวน้อยเพราะการบดบังของต้นไม้และความหลากหลายของสิ่งปกคลุมหลังคาอาคาร



รูปที่ 12 รอยพิมพ์ฐานอาคารจากการสกัดด้วยข้อความ (ก และ ข) ต่อมาเติมเต็มด้วยการสกัดด้วยจุด (ค และ ง) เปรียบเทียบกับข้อมูลจากการคัดลอกด้วยตนเอง (จ และ ฉ) ในพื้นที่คณะวิทยาศาสตร์ (แถวบน) และวิศวกรรมศาสตร์ (แถวล่าง)

ตารางที่ 3 การเปรียบเทียบตัวอย่างรอยพิมพ์ฐานอาคาร ณ ตำแหน่งเดียวกันในบริเวณพื้นที่การศึกษา

True Orthophoto	Microsoft	Google	GeoSAM	Delineation
				
				
				
				
				
				
				
				
				

ข้อจำกัดของวิธีนี้คือต้องหาภาพออร์โธจริงความละเอียดสูง การประมวลผลภาพต้องใช้คอมพิวเตอร์ที่มีประสิทธิภาพสูง การทดลองในพื้นที่เล็กนี้ทำผ่านระบบ Cloud Computing คือ Google Colab ซึ่งมี GPU ที่มีศักยภาพในการประมวลผลในพื้นที่เล็กได้ ภาพออร์โธที่ใช้ต้องทำการตัดเฉพาะพื้นที่ที่ต้องการนำเข้าแบบจำลอง (ขนาดภาพในแต่ละพื้นที่อยู่ที่ 100 – 200 MB ภาพพื้นที่เล็กแต่ขนาดใหญ่มากจากการใช้ภาพความละเอียดสูงระดับ 5 เซนติเมตร) การนำเข้ารูปที่ขนาดใหญจะทำให้ระบบไม่สามารถประมวลผลได้ หากต้องการสกัดอาคารในพื้นที่ใหญ่ควรใช้คอมพิวเตอร์ส่วนตัวที่มีประสิทธิภาพพร้อม ปัจจุบันวิธีดังกล่าวไม่สามารถสกัดได้แบบอัตโนมัติในครั้งเดียว ผู้ใช้ควรทดลองปรับพารามิเตอร์ในการสกัดด้วยข้อความให้มั่นใจ สักครอยพิมพ์ฐานอาคารก่อน แล้วค่อยเก็บอาคารเล็กด้วยการสกัดด้วยจุดพิกัด

งานวิจัยที่ใช้ภาพความละเอียดสูงนอกจากการสกัดครอยพิมพ์ฐานอาคารคือการสกัดลักษณะของหลังคา [10] ซึ่งสามารถนำไปใช้วิเคราะห์เมือง วางตำแหน่งที่รับพลังงานแสงอาทิตย์ได้ โดยใช้เทคนิคการเรียนรู้เชิงลึกเช่นกัน กับมีความพยายามสร้างแบบจำลองในการสกัดข้อมูลอาคารที่ใช้เวลาเร็วและมีความถูกต้อง [11]

ความก้าวหน้าทางเทคโนโลยีทำให้งานซึ่งใช้ทรัพยากรมนุษย์อย่างหนักนั้นทดแทนด้วยปัญญาประดิษฐ์ แต่ทว่าบทบาทของมนุษย์ยังคงมีอยู่เพื่อตรวจสอบความถูกต้องของผลลัพธ์ที่ได้ต่อไป แม้จะมีผลงานใหม่ออกมาอย่างไม่หยุดยั้ง

ผลประโยชน์ทับซ้อน

ผู้เขียนขอประกาศว่าบทความนี้ไม่มีผลประโยชน์ทับซ้อน

กิตติกรรมประกาศ

ผู้เขียนขอขอบพระคุณ การไฟฟ้านครหลวง สำหรับผลผลิตโครงการทำแผนที่สามมิติในกรุงเทพมหานครอย่างภาพออร์โธจริงในพื้นที่เมืองซึ่งหาได้ยาก เป็นประโยชน์อย่างยิ่งต่อการศึกษาด้านการแผนที่ในปัจจุบัน และ ขอขอบพระคุณอาจารย์เอไอโนใจ ผู้เขียน Hiroyuki Miyazaki, Ph.D. สำหรับองค์ความรู้เกี่ยวกับปัญญาประดิษฐ์สำหรับข้อมูลภาพดาวเทียม

เอกสารอ้างอิง

- [1] Güneralp, B., Zhou, Y., Ürge-Vorsatz, D., Gupta, M., Yu, S., Patel, P.L., Fragkias, M., Li, X., Seto, K.C. Global Scenarios of Urban Density and Its Impacts on Building Energy Use Through 2050, *Proceedings of the National Academy of Sciences*, 2017, 114, pp.8945–8950.
- [2] Zheng Y, Weng Q, Zheng Y. A Hybrid Approach for Three-Dimensional Building Reconstruction in Indianapolis from Lidar Data. *Remote Sensing*, 2017, 9 (4), pp.310.
- [3] Pieterse, E. Building with Ruins and Dreams: Some Thoughts on Realising Integrated Urban Development In South Africa Through Crisis. *Urban Studies*, 2006, 43, pp.285–304.
- [4] Microsoft Bing Maps. Worldwide Building Footprints Derived from Satellite Imagery: Github, 2022. Retrieved from: <https://github.com/microsoft/GlobalMLBuildingFootprints#will-there-be-more-data-coming-for-other-geographies>.
- [5] W. Sirko, S. Kashubin, M. Ritter, A. Annkah, Y.S.E. Bouchareb, Y. Dauphin, D. Keyzers, M. Neumann, M. Cisse, J.A. Quinn. Continental-scale Building Detection from High Resolution Satellite Imagery. arXiv:2107.12283, 2021.
- [6] ไพศาล สันติธรรมนนท์. (2564). *การรังวัดด้วยภาพดิจิทัล Digital Photogrammetry*. กรุงเทพมหานคร: สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย
- [7] Qiusheng Wu, & Lucas Osco. SAMGEO: A Python Package for Segmenting Geospatial Data with The Segment Anything Model (SAM). 2023. Zenodo. <https://doi.org/10.5281/zenodo.7966658>

- [8] Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A., Lo, W., Dollár, P. & Girshick, R. SEGMENT ANYTHING. 2023. {ArXiv:2304.02643}.
- [9] Yarroudh, A. Lidar Automatic Unsupervised Segmentation Using Segment-anything Model (Sam) From Meta AI: Github, 2023. Retrieved from <https://github.com/Yarroudh/segment-lidar>.
- [10] Wufan Zhao, Claudio Persello, Alfred Stein End-to-end Roofline Extraction From Very High-resolution Remote Sensing Images. 2021 *IEEE International Geoscience and Remote Sensing Symposium IGARSS*
- [11] Miao, Y.; Jiang, S.; Xu, Y.; Wang, D. Feature Residual Analysis Network for Building Extraction from Remote Sensing Images. *Applied Sciences*, 2022, 12, pp.5095. <https://doi.org/10.3390/app12105095>