



## การศึกษาเปรียบเทียบแบบจำลองโครงข่ายปฏิกิริยเชิงกำเนิดในการสร้างภาพความละเอียดสูง เพื่อการตรวจจับวัตถุขนาดเล็ก

### A Comparative Study of Super-Resolution GAN Models for Small Object Detection

ภัทรชนน สกุลคุ<sup>1\*</sup> และ อัครินทร์ ไพบูลย์พานิช<sup>2</sup>

Patchanon Sakunkoo<sup>1\*</sup> and Akarin Phaibulpanich<sup>2</sup>

<sup>1</sup> นักศึกษาระดับปริญญาโท, หลักสูตรวิทยาศาสตรมหาบัณฑิต, ภาควิชาสถิติ, คณะพาณิชยศาสตร์และการบัญชี,  
จุฬาลงกรณ์มหาวิทยาลัย

<sup>1</sup> Graduate student, Master of Science, Department of Statistics, Faculty of Commerce and  
Accountancy, Chulalongkorn University.

<sup>2</sup> ผู้ช่วยศาสตราจารย์, ภาควิชาสถิติ, คณะพาณิชยศาสตร์และการบัญชี, จุฬาลงกรณ์มหาวิทยาลัย

<sup>2</sup> Assistant Professor, Department of Statistics, Faculty of Commerce and Accountancy, Chulalongkorn  
University.

\*Corresponding author, E-mail: 6680280626@cbs.chula.ac.th

#### บทคัดย่อ

การตรวจจับวัตถุขนาดเล็ก (Small Object Detection) เป็นหนึ่งในความท้าทายด้านคอมพิวเตอร์วิทัศน์ เนื่องจากภาพที่มีความละเอียดต่ำและมีข้อจำกัดด้านรายละเอียดของวัตถุ งานวิจัยนี้มุ่งเน้นการเปรียบเทียบประสิทธิภาพของแบบจำลอง Super-Resolution GAN คือ SRGAN ESRGAN Real-ESRGAN และ BSRGAN กับวิธีการสร้างภาพความละเอียดสูงแบบดั้งเดิม คือ Bilinear และ Bicubic เพื่อศึกษาว่าการสร้างภาพความละเอียดสูงสามารถช่วยให้การตรวจจับวัตถุขนาดเล็กมีความแม่นยำสูงขึ้นหรือไม่ โดยการทดลองดำเนินการกับภาพต้นฉบับขนาด  $160 \times 160$  พิกเซล และสร้างภาพความละเอียดสูงขนาด  $640 \times 640$  พิกเซลก่อนนำไปประเมินผลการตรวจจับ โดยใช้ค่าเฉลี่ยความแม่นยำเป็นตัวชี้วัด ในสามชุดข้อมูล ได้แก่ ภาพสายเคเบิล เซลล์ Microglia และวัชพืช Ridderzuring ผลการทดลองพบว่า SRGAN สามารถเพิ่มความแม่นยำได้สูงสุดในทุกชุดข้อมูล โดยสามารถเพิ่มความแม่นยำได้ 2.2 – 6.7 เปอร์เซ็นต์ แม้ว่า SRGAN จะมีความแม่นยำสูงสุด แต่ใช้เวลาในการสร้างภาพมากกว่าวิธีแบบดั้งเดิมประมาณ 4-7 เท่า ทำให้ต้องพิจารณาความสมดุลระหว่างความเร็วและความแม่นยำ ทั้งนี้ ภาพต้นฉบับความละเอียดสูงยังคงมีค่าเฉลี่ยสูงสุดในทุกชุดข้อมูล ซึ่งสะท้อนว่าการใช้เทคนิค Super-Resolution GAN สามารถลดช่องว่างระหว่างภาพความละเอียดต่ำและภาพความละเอียดสูงได้ และสามารถนำไปประยุกต์ใช้กับงานเกี่ยวกับการตรวจจับวัตถุจากภาพความละเอียดต่ำได้จริง

**คำสำคัญ:** การตรวจจับวัตถุขนาดเล็ก, โครงข่ายปฏิกิริยเชิงกำเนิด, การสร้างภาพความละเอียดสูง



## Abstract

Small object detection is challenging in computer vision due to low image resolution and limited object details. This research compares Super-Resolution GAN models—SRGAN, ESRGAN, Real-ESRGAN, and BSRGAN—with traditional methods (Bilinear and Bicubic interpolation) to determine if enhancing resolution improves detection accuracy. Images were upscaled from  $160 \times 160$  pixels to  $640 \times 640$  pixels, and object detection accuracy was evaluated using mean Average Precision (mAP) across three datasets: cables, Microglia cells, and Ridderzuring weeds. Results showed SRGAN achieved the highest accuracy improvement, increasing mAP by 2.2–6.7% compared to traditional methods. However, SRGAN required 4–7 times longer processing, highlighting a trade-off between accuracy and computational speed. Original high-resolution images consistently yielded the highest mAP, indicating that Super-Resolution GAN techniques effectively narrow the accuracy gap between low and high resolution images, making them practical for enhancing small object detection in low-resolution imagery.

**Keywords:** Small Object Detection, GAN, Super-Resolution

## บทนำ

การตรวจจับวัตถุขนาดเล็ก (Small Object Detection) เป็นงานที่ท้าทายในด้านคอมพิวเตอร์วิทัศน์ เนื่องจากวัตถุขนาดเล็กมักมีรายละเอียดน้อย ประกอบกับรูปภาพมีคุณภาพและความละเอียดต่ำ ซึ่งส่งผลให้การตรวจจับทำได้ยากและขาดความแม่นยำปัญหานี้ จึงมีความสำคัญมากขึ้นเมื่อระบบตรวจจับวัตถุถูกนำมาใช้ในภาคอุตสาหกรรมหรือในงานวิจัยที่ต้องการความแม่นยำสูง เช่น การจัดการขยะในสิ่งแวดล้อม การเฝ้าระวังความปลอดภัยในที่สาธารณะ และการควบคุมการจราจรทางอากาศ การพัฒนาระบบตรวจจับวัตถุที่มีความสามารถในการตรวจจับวัตถุขนาดเล็กได้อย่างแม่นยำจึงเป็นสิ่งสำคัญ

เทคนิคที่นิยมนำมาใช้เพื่อเพิ่มความสามารถในการตรวจจับวัตถุขนาดเล็ก คือ เทคนิคการสร้างภาพความละเอียดสูง (Super-Resolution) เพื่อเพิ่มความละเอียดของภาพก่อนการตรวจจับวัตถุ โดยหนึ่งในสถาปัตยกรรมที่มีความโดดเด่นเรื่องความสามารถในการเพิ่มความละเอียดของภาพที่มีความละเอียดต่ำเพื่อให้รายละเอียดของวัตถุขนาดเล็กชัดเจนขึ้น คือ Super-Resolution Generative Adversarial Networks (SRGAN) ที่ได้รับการพัฒนาและปรับปรุงอย่างต่อเนื่อง ซึ่งจะช่วยปรับปรุงผลลัพธ์ในการตรวจจับวัตถุขนาดเล็กได้อย่างมีประสิทธิภาพ

## วัตถุประสงค์ของการวิจัย

1. เพื่อเปรียบเทียบประสิทธิภาพในการสร้างภาพความละเอียดสูงของแบบจำลองโครงข่ายประสาทเชิงกำเนิด จำนวน 4 แบบ ได้แก่ SRGAN, ESRGAN, Real-ESRGAN และ BSRGAN



2. เพื่อประเมินความสามารถของแบบจำลองโครงข่ายปฏิบัติเชิงกำเนิด และเทคนิคดั้งเดิมในการปรับปรุงภาพสำหรับการตรวจจับวัตถุขนาดเล็ก

### แนวคิด ทฤษฎี กรอบแนวคิด

#### 1. โครงข่ายปฏิบัติเชิงกำเนิด (Generative Adversarial Networks: GAN)

โครงข่ายปฏิบัติเชิงกำเนิด (GAN) เป็นนวัตกรรมสำคัญในด้านการเรียนรู้ของเครื่องและปัญญาประดิษฐ์ ซึ่งได้รับความสนใจอย่างกว้างขวางนับตั้งแต่ Goodfellow, Pouget-Abadie, Mirza, Xu, Warde-Farley, Ozair, Courville และ Bengio (2014) ด้วยความสามารถในการสร้างข้อมูลจำลองคุณภาพสูงที่ใกล้เคียงกับข้อมูลจริง GAN มีโครงสร้างโดดเด่นจากการมีเครือข่ายย่อยสองส่วนที่ฝึกฝนร่วมกันในลักษณะของเกมทฤษฎี ได้แก่ ตัวสร้าง (Generator) และตัวแยกแยะ (Discriminator) โดยตัวสร้างรับสัญญาณสุ่ม เช่น การแจกแจงแบบ Uniform หรือ Gaussian เพื่อผลิตข้อมูลจำลอง ส่วนตัวแยกแยะมีหน้าที่เปรียบเทียบข้อมูลจำลองกับข้อมูลจริง และพยายามแยกแยะระหว่างข้อมูลทั้งสอง เป้าหมายของตัวสร้าง คือ การผลิตข้อมูลที่หลอกตัวแยกแยะได้สำเร็จ ขณะที่ตัวแยกแยะพยายามพัฒนาความแม่นยำในการตรวจจับข้อมูลจำลอง การทำงานร่วมกันของทั้งสองเครือข่ายนี้ส่งผลให้เกิดการพัฒนาประสิทธิภาพอย่างต่อเนื่อง

#### 2. การสร้างภาพความละเอียดสูง (Super-Resolution)

การสร้างภาพความละเอียดสูง หมายถึงกระบวนการปรับปรุงหรือฟื้นฟูภาพที่มีความละเอียดต่ำให้กลายเป็นภาพที่มีความละเอียดสูง โดยมีวิธีการที่ได้รับความนิยมอย่างแพร่หลาย ดังนี้

##### 2.1) เทคนิคการสร้างภาพความละเอียดสูงแบบดั้งเดิม

การสร้างภาพความละเอียดสูงโดยวิธี Interpolation แบบ Nearest Neighbor Bilinear และ Bicubic เป็นวิธีการเริ่มต้นที่ใช้ในการเพิ่มความละเอียดภาพ โดยประมาณค่าพิกเซลจากข้อมูลรอบข้าง วิธี Bicubic เป็นที่นิยมมากเนื่องจากความเรียบง่าย แต่ไม่สามารถกู้คืนรายละเอียดความถี่สูงที่สูญหายไป

##### 2.2) เทคนิคการสร้างภาพความละเอียดสูงโดยใช้โครงข่ายปฏิบัติเชิงกำเนิด

2.2.1) Super-Resolution Generative Adversarial Network (SRGAN) เป็นสถาปัตยกรรมที่นำเสนอโดย Ledig, Theis, Huszar, Caballero, Cunningham, Acosta, Aitken, Tejani, Totz and Wang (2017) ซึ่งนับเป็นความก้าวหน้าสำคัญในการเพิ่มความละเอียดของภาพด้วยการเรียนรู้เชิงลึก โดยสามารถสร้างภาพความละเอียดสูงจากภาพความละเอียดต่ำได้อย่างสมจริง ทั้งในด้านรายละเอียดและพื้นผิวของภาพ โครงสร้างของ SRGAN ประกอบด้วยสองส่วนหลัก คือ เครือข่ายสร้างภาพและเครือข่ายแยกแยะ ตามแนวคิดของโครงข่ายปฏิบัติเชิงกำเนิดที่เสนอโดย Goodfellow et.al., (2014) โดย SRGAN ได้นำแนวคิดนี้มาประยุกต์ใช้เฉพาะสำหรับการเพิ่มความละเอียดของภาพ

2.2.2) Enhanced Super-Resolution Generative Adversarial Networks (ESRGAN) ถูกเสนอโดย Wang, Yu, Wu, Gu, Liu, Dong, Qiao and Change Loy (2018) ได้แก้ไขข้อบกพร่องของ SRGAN



โดยการพัฒนาคุณภาพของการสร้างภาพ รวมทั้งเพิ่มความสามารถในการจับรายละเอียดของภาพทั้งที่มีความถี่สูงและโครงสร้างของภาพด้วย

2.2.3) Real-Enhanced Super-Resolution Generative Adversarial Networks (Real-ESRGAN) ซึ่งถูกพัฒนาโดย Wang, Xie, Dong and Shan (2021) มุ่งเน้นที่การแก้ปัญหาของการสร้างภาพในโลกความเป็นจริง โดยปรับปรุงแบบจำลอง ESRGAN ให้สามารถรับมือกับการบิดเบือนของภาพได้ดียิ่งขึ้น

2.2.4) Blind Super-Resolution Generative Adversarial Networks (BSRGAN) ได้รับการพัฒนาโดย Zhang, Liang, Van Gool และ Timofte (2021) ซึ่งจะเน้นที่การแก้ปัญหาเรื่องการเสื่อมสภาพของภาพ (Image Degradation) กระบวนการนี้ช่วยให้แบบจำลองสามารถทำงานได้ดีขึ้นในสถานการณ์ที่ภาพมีการเสื่อมสภาพหลายรูปแบบ

### 3. การตรวจจับวัตถุด้วย YOLO

YOLO (You Only Look Once) เป็นอัลกอริทึมสำหรับการตรวจจับวัตถุที่ถูกนำเสนอครั้งแรกโดย Redmon & Farhadi (2017) ซึ่งทำให้เกิดการเปลี่ยนแปลงวิธีการตรวจจับวัตถุ โดยมองปัญหาเป็นงานเดียวที่ใช้การถดถอย (Regression) เพื่อทำนายทั้งขอบเขตของวัตถุและความน่าจะเป็นของประเภทในครั้งเดียว การใช้วิธีนี้ช่วยให้ YOLO สามารถทำการตรวจจับวัตถุได้รวดเร็วมาก

### 4. mAP (Mean Average Precision)

Padilla, Netto & Silva (2020) ได้ทำการศึกษารวบรวมและเสนอแนวทางเกี่ยวกับตัวชี้วัดสำคัญที่ใช้ในงานตรวจจับวัตถุ โดยหนึ่งในตัวชี้วัดที่ได้รับความนิยมอย่างมากคือ Mean Average Precision (mAP) ซึ่งใช้ประเมินประสิทธิภาพของแบบจำลองผ่านการวัดค่าความถูกต้องของผลลัพธ์ที่ได้ โดยพิจารณาจาก Precision (P) และ Recall (R) ทั้งนี้ Mean Average Precision (mAP) เป็นตัวชี้วัดที่ได้รับการยอมรับอย่างแพร่หลายในการประเมินผลแบบจำลองตรวจการตรวจจับวัตถุ โดยคำนวณจากพื้นที่ใต้กราฟ Precision-Recall (PR Curve) และคำนวณค่าเฉลี่ยของความแม่นยำสำหรับทุกคลาสในชุดข้อมูลที่กำหนด

## ที่มาและความสำคัญของปัญหา

การตรวจจับวัตถุขนาดเล็กเป็นงานที่ท้าทายในด้านคอมพิวเตอร์วิทัศน์ เนื่องจากวัตถุขนาดเล็กมักมีรายละเอียดน้อย ประกอบกับรูปภาพมีคุณภาพและความละเอียดต่ำ ซึ่งส่งผลให้การตรวจจับทำได้ยากและขาดความแม่นยำปัญหานี้ จึงมีความสำคัญมากขึ้นเมื่อระบบตรวจจับวัตถุถูกนำมาใช้ในภาคอุตสาหกรรมหรือในงานวิจัยที่ต้องการความแม่นยำสูง การพัฒนาระบบตรวจจับวัตถุที่มีความสามารถในการตรวจจับวัตถุขนาดเล็กได้อย่างแม่นยำจึงเป็นสิ่งสำคัญ

เพื่อศึกษาผลกระทบของการใช้เทคนิค Super-Resolution ต่อประสิทธิภาพการตรวจจับวัตถุ งานวิจัยนี้จึงเลือกใช้แบบจำลอง Super-Resolution ที่ได้รับความนิยมและมีประสิทธิภาพสูงจำนวน 4 แบบ ได้แก่ SRGAN, ESRGAN, Real-ESRGAN และ BSRGAN ทั้งนี้เนื่องจากแต่ละแบบจำลองมีลักษณะเฉพาะที่ต่างกัน และสามารถเป็นตัวแทนของวิวัฒนาการด้านสถาปัตยกรรม Super-Resolution



ได้อย่างครบถ้วน SRGAN เป็นแบบจำลองพื้นฐานที่ริเริ่มนำแนวคิดการใช้ Generative Adversarial Networks (GAN) เข้ามาสร้างภาพความละเอียดสูง ในขณะที่ ESRGAN ได้มีการปรับปรุงโครงสร้างของ Generator และการประเมินคุณภาพภาพ เพื่อให้ได้ภาพที่มีรายละเอียดคมชัดยิ่งขึ้น Real-ESRGAN พัฒนาต่อเนื่องจาก ESRGAN โดยเน้นการประมวลผลภาพที่หลากหลายในโลกความจริง ส่วน BSRGAN ออกแบบมาเพื่อรองรับการเพิ่มความละเอียดของภาพที่มีความเสื่อมสภาพจากสภาพแวดล้อมจริง โดยใช้กลยุทธ์การเสื่อมสภาพภาพแบบสุ่ม

การเลือกแบบจำลองทั้ง 4 แบบนี้ จึงมีความเหมาะสมเพื่อให้นักศึกษาเปรียบเทียบครอบคลุมทั้งด้านคุณภาพของภาพ ความสามารถในการรองรับข้อมูลที่หลากหลาย และความสอดคล้องกับสภาพปัญหาที่พบในงานตรวจจบบัตถุจริงในปัจจุบัน

### งานวิจัยที่เกี่ยวข้อง

Telçeken, Akgun, Kacar & Bingol (2024) ได้นำเสนอวิธีการใหม่ในการปรับปรุงการตรวจจบบัตถุในภาพถ่ายดาวเทียมและภาพถ่ายทางอากาศ โดยใช้เทคนิคการสร้างภาพความละเอียดสูงร่วมกับอัลกอริทึมการแบ่งภาพและการแบ่งส่วนขั้นสูงเป็นระบบ Super-Resolution Object Detection (SROD) ระบบตรวจจบบัตถุนี้เป็นการรวมเทคนิค SRGAN เข้ากับ YOLO โดยที่ SRGAN จะช่วยเพิ่มความละเอียดของภาพเพื่อปรับปรุงความแม่นยำในการตรวจจบบัตถุ พบว่าระบบ SROD ที่ผสมผสาน SRGAN แสดงผลลัพธ์ที่ดีกว่าการใช้ YOLO รุ่นปกติ และชี้ให้เห็นถึงการปรับปรุงอย่างมีนัยสำคัญในค่า mAP (mean Average Precision) และเมตริกอื่น ๆ โดยเฉพาะการตรวจจบบัตถุขนาดเล็ก

Yigit (2023) ได้นำเสนองานวิจัยเรื่อง “การปรับปรุงการตรวจจบบัตถุจากภาพถ่ายโดรนด้วย Real-ESRGAN” โดยมุ่งเน้นการปรับปรุงการตรวจจบบัตถุจากภาพถ่ายที่ได้จากโดรนซึ่งมักมีความละเอียดต่ำ ซึ่งส่งผลให้ความแม่นยำในการตรวจจบบัตถุลดลงเพื่อลดข้อจำกัดนี้ นักวิจัยได้ใช้เทคนิค Super-Resolution ผ่านแบบจำลอง Real-ESRGAN ร่วมกับ YOLOv7 ในการเพิ่มคุณภาพของภาพก่อนนำไปใช้ตรวจจบบัตถุ ผลการทดลองพบว่าสามารถตรวจจบบัตถุได้มากขึ้น และมีความแม่นยำสูงขึ้น

จากงานวิจัยที่เกี่ยวข้องแสดงให้เห็นว่า ภาพที่ผ่านกระบวนการ Super-Resolution สามารถช่วยให้ตรวจจบบัตถุได้มากขึ้นและมีความแม่นยำสูงขึ้น โดยเฉพาะกับวัตถุขนาดเล็กนอกจากนี้ ยังพบว่าค่าความแม่นยำเฉลี่ย (mAP) เพิ่มขึ้นอย่างมีนัยสำคัญเมื่อใช้ภาพที่ได้รับการปรับปรุงความละเอียด

### วิธีดำเนินการวิจัย

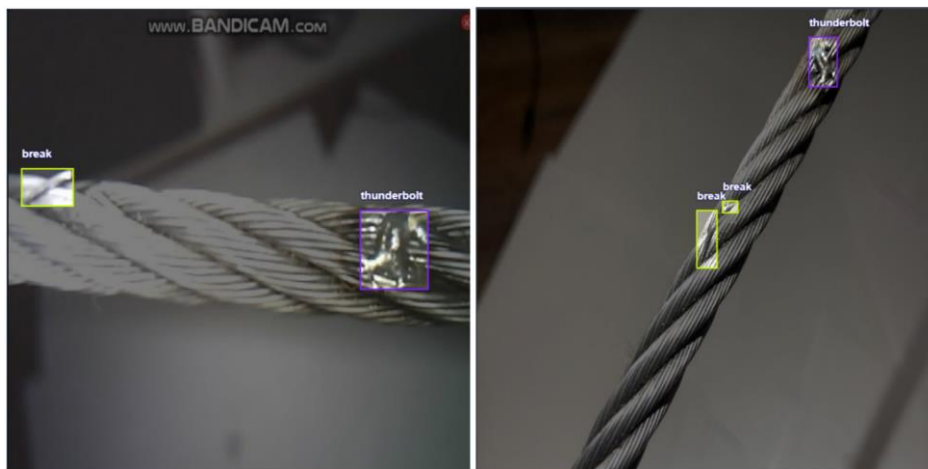
ในการวิจัยด้าน Super-Resolution และการตรวจจบบัตถุขนาดเล็ก มีความจำเป็นอย่างยิ่งที่จะต้องควบคุมตัวแปรที่อาจส่งผลกระทบต่อผลลัพธ์ของแบบจำลอง เพื่อให้สามารถประเมินประสิทธิภาพของเทคนิค Super-Resolution ได้อย่างแม่นยำ งานวิจัยนี้จึงเลือกใช้วิธีการสร้างชุดข้อมูลความละเอียดต่ำ (Low-Resolution) โดยการลดขนาด (Downsampling) จากภาพต้นฉบับที่มีความละเอียดสูง (High-

Resolution Images) แทนที่จะนำภาพความละเอียดต่ำที่มีอยู่แล้วมาใช้งานโดยตรง เพื่อควบคุมลักษณะของการเสื่อมสภาพของข้อมูลได้อย่างเป็นระบบ ทำให้มั่นใจได้ว่าภาพที่นำมาใช้ในกระบวนการฝึกสอนและทดสอบนั้นมีคุณสมบัติที่เป็นมาตรฐาน และไม่มีตัวแปรแฝงอื่น ๆ โดยมีขั้นตอนในการวิจัย ดังนี้

#### 1. การเตรียมชุดข้อมูล

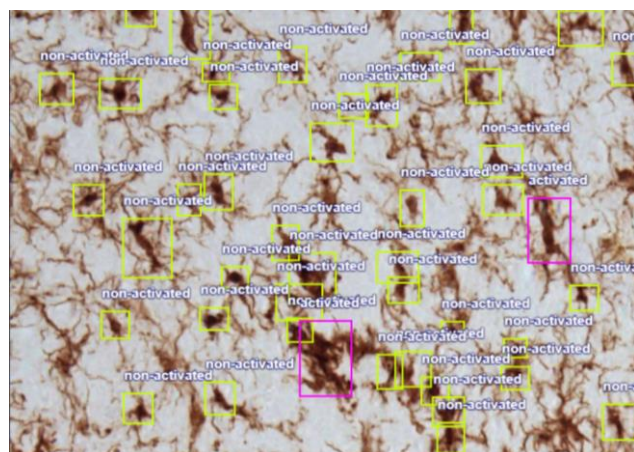
ชุดข้อมูลที่ใช้ในการวิจัยนี้มี 3 ชุดข้อมูล ซึ่งได้มาจากแหล่งข้อมูลออนไลน์ Roboflow Universe ที่ถูกรวบรวมโดย Ciaglia, Zuppichini, Guerrie, McQuade และ Solawetz (2022) ซึ่งเป็นข้อมูลที่มีภาพที่มีขนาด 640 x 640 พิกเซล และถูกแบ่งออกเป็น 3 ส่วน ได้แก่ ชุดข้อมูลสอน (train set) ชุดข้อมูลตรวจสอบ (validation set) และชุดข้อมูลทดสอบ (test set) พร้อมทั้งไฟล์ Annotation ซึ่งเป็นไฟล์ข้อความที่ระบุตำแหน่งวัตถุบนภาพ

1) ชุดข้อมูลภาพความเสียหายของสายเคเบิล มีข้อมูลภาพทั้งหมด 1,318 ภาพ พร้อมทั้งข้อมูลการระบุตำแหน่งวัตถุ ทั้งหมด 1,512 ตำแหน่ง ซึ่งแบ่งออกเป็น 2 ประเภท คือ ความเสียหายที่เกิดจากการขาด (Cable Breakage) และความเสียหายที่เกิดจากความร้อน (Heat Damage, Thunderbolt) วัตถุในชุดข้อมูลนี้มีขนาดเล็กและบางครั้งกลมกลืนกับพื้นหลังของสายเคเบิล ทำให้เป็นกรณีศึกษาที่เหมาะสมสำหรับการประเมินประสิทธิภาพของ Super-Resolution ในการเพิ่มความเด่นชัดของตำแหน่งความเสียหาย ตัวอย่างของวัตถุที่ระบุไว้แสดงในภาพที่ 1



ภาพที่ 1 ตัวอย่างของความเสียหายที่เกิดจากการขาด (Cable Breakage) และความเสียหายที่เกิดจากความร้อน (Heat Damage, Thunderbolt)

2) ชุดข้อมูล Microglia ประกอบด้วยข้อมูลภาพทั้งหมด 1,422 ภาพ พร้อมทั้งข้อมูลการระบุตำแหน่งวัตถุ ทั้งหมด 32,265 ตำแหน่ง ซึ่งแบ่งออกเป็น 2 ประเภท คือ ชุดข้อมูลที่ถูกกระตุ้น (activated) และชุดข้อมูลที่ไม่ได้ถูกกระตุ้น (non-activated) ลักษณะของ Microglia มีความละเอียดซับซ้อนสูงและมีขนาดเล็กมากในภาพ จึงเป็นชุดข้อมูลที่มีความท้าทายสูง และเหมาะสำหรับการทดสอบว่าการเพิ่มความละเอียดสามารถช่วยปรับปรุงการจำแนก และการตรวจจับวัตถุขนาดเล็กในระดับเซลล์ตัวอย่างของวัตถุที่ระบุไว้แสดงในภาพที่ 2



ภาพที่ 2 ตัวอย่างของเซลล์ไมโครเกลียที่ถูกกระตุ้น (แสดงด้วยกรอบสีชมพู) และเซลล์ไมโครเกลียที่ไม่ได้ถูกกระตุ้น (แสดงด้วยกรอบสีเหลือง)

3) ชุดข้อมูลภาพวัชพืช Ridderzuring มีข้อมูลภาพทั้งหมด 2,486 ภาพ พร้อมทั้งข้อมูลการระบุตำแหน่งของวัชพืช Ridderzuring ทั้งหมด 6,814 ตำแหน่ง วัชพืชเหล่านี้มักมีขนาดเล็กกระจายอยู่ในพื้นที่กว้าง และพื้นหลังมีความหลากหลาย เช่น พุ่มหญ้า ดิน หรือแปลงเกษตร การตรวจจับวัชพืชในสภาพแวดล้อมจริงที่ซับซ้อนจึงต้องอาศัยคุณภาพของภาพสูง การนำเทคนิค Super-Resolution เข้ามาช่วยจึงเป็นแนวทางที่เหมาะสมในการเพิ่มความแม่นยำ ตัวอย่างของวัตถุที่ระบุไว้แสดงในภาพที่ 3



ภาพที่ 3 ตัวอย่างของวัชพืช Ridderzuring

## 2. การเตรียมสภาพแวดล้อมที่ใช้ในการทดลอง

การวิจัยในครั้งนี้ มีการใช้ Google Colaboratory ซึ่งเป็นโฮสต์โปรแกรม Jupyter Notebook ที่ทำงานบนระบบคลาวด์ เพื่อใช้ในการเขียนโปรแกรมสำหรับการสอนแบบจำลองตรวจการจับวัตถุรวมถึงการทดสอบผลลัพธ์ที่ได้จากการสร้างภาพความละเอียดสูง โดยเน้นการใช้งานโครงข่ายปฏิบัติเชิงกำเนิดทั้ง 4 แบบ ได้แก่ SRGAN ESRGAN Real-ESRGAN และ BSRGAN

ภาษาหลักที่ใช้ในการเขียนโปรแกรมคือ Python โดยทำงานร่วมกับเฟรมเวิร์ก ได้แก่ TensorFlow

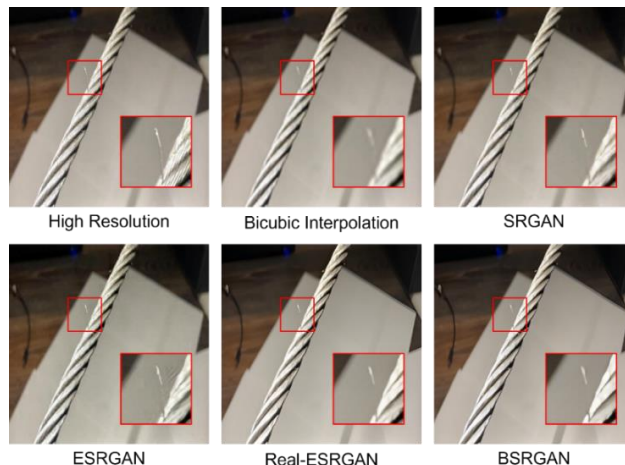
Keras และ PyTorch ซึ่งช่วยให้การสร้างและสอนแบบจำลองเป็นไปอย่างมีประสิทธิภาพและยืดหยุ่น นอกจากนี้ ยังมีการใช้ค่าน้ำหนักของแบบจำลองที่ผ่านการสอนมาแล้ว ทำให้ลดเวลาในการสอนแบบจำลอง ตั้งแต่เริ่มต้น และช่วยให้สามารถสร้างภาพความละเอียดสูง พร้อมทั้งประเมินประสิทธิภาพของแบบจำลอง ได้อย่างแม่นยำและรวดเร็ว

### 3. การสอนแบบจำลองการตรวจจับวัตถุ

กระบวนการสอนแบบจำลองการตรวจจับวัตถุ ผู้วิจัยเลือกใช้แบบจำลอง YOLOv10m ที่ได้รับการพัฒนาโดย Wang, Chen, Liu, Chen, Lin, Han และ Ding (2024) ซึ่งเป็นแบบจำลองที่มีประสิทธิภาพ และเหมาะสมสำหรับการตรวจจับวัตถุในภาพความละเอียดสูง โดยเริ่มต้นจากการโหลดแบบจำลอง YOLOv10m ที่ยังไม่มี การสอน และใช้ชุดข้อมูลมีการจัดเตรียมการระบุตำแหน่งของวัตถุบนภาพ และ คลาสของวัตถุอย่างครบถ้วน การสอนนี้ถูกตั้งค่าให้ใช้จำนวน epoch ทั้งหมด 100 รอบ พร้อมกำหนด ขนาดภาพสำหรับการสอนเป็น 640x640 พิกเซล ซึ่งเป็นขนาดเดียวกันกับขนาดภาพของชุดข้อมูลต้นฉบับ

### 4. การสร้างแบบจำลอง Super-Resolution GAN และสร้างภาพความละเอียดสูง

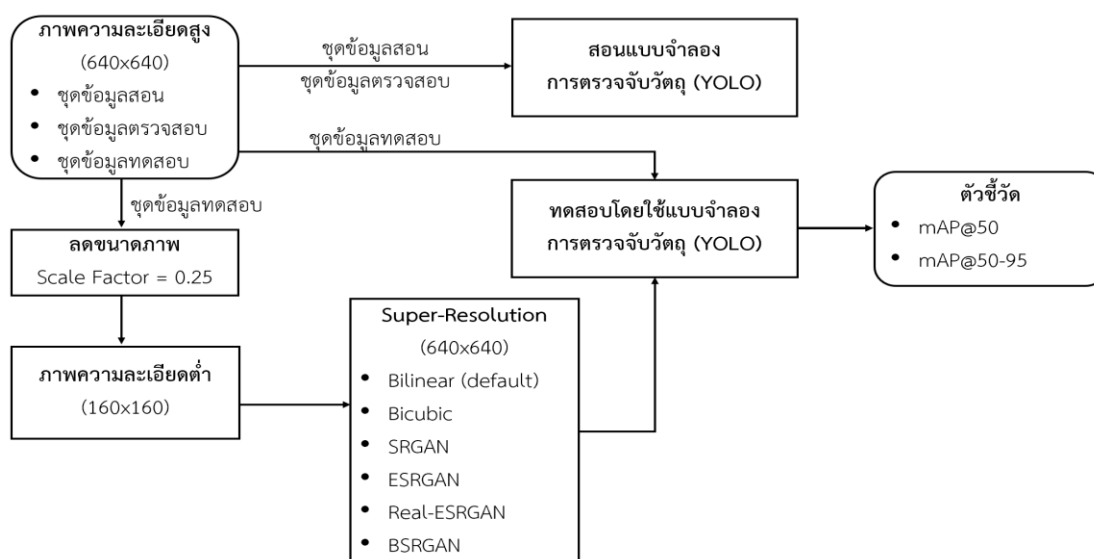
การสร้างแบบจำลอง Super-Resolution GAN ในงานวิจัยนี้เป็นการรวบรวมจากแหล่งข้อมูล ที่เผยแพร่ผ่าน GitHub และ TensorFlow Hub โดยแบบจำลอง SRGAN นำมาจากโปรเจกต์ Tensor Layer แบบจำลอง ESRGAN นำมาจากโปรเจกต์ TensorFlow Hub แบบจำลอง Real-ESRGAN นำมาจาก โปรเจกต์ Real-ESRGAN และแบบจำลอง BSRGAN นำมาจากโปรเจกต์ BSRGAN



ภาพที่ 4 ตัวอย่างผลลัพธ์การสร้างภาพความละเอียดสูงจากแบบจำลองต่าง ๆ

### 5. การวัดประสิทธิภาพของการสร้างภาพความละเอียดสูงผ่านการตรวจจับวัตถุ

การประเมินแบบจำลอง Super-Resolution GAN คือ การนำแบบจำลองแต่ละแบบมาคำนวณ ตัวชี้วัดเพื่อเปรียบเทียบประสิทธิภาพในการเพิ่มความละเอียดของภาพ เพื่อการตรวจจับวัตถุขนาดเล็ก โดยตัววัดประสิทธิภาพที่ใช้ในงานวิจัยนี้ ได้แก่ mAP@50 (Mean Average Precision at IoU=0.50) mAP@50-95 (Mean Average Precision at IoU=0.50-0.95) และเวลาที่ใช้ในการสร้างภาพความละเอียด สูงเฉลี่ย



ภาพที่ 5 กระบวนการดำเนินงานวิจัยสำหรับ 1 ชุดข้อมูล

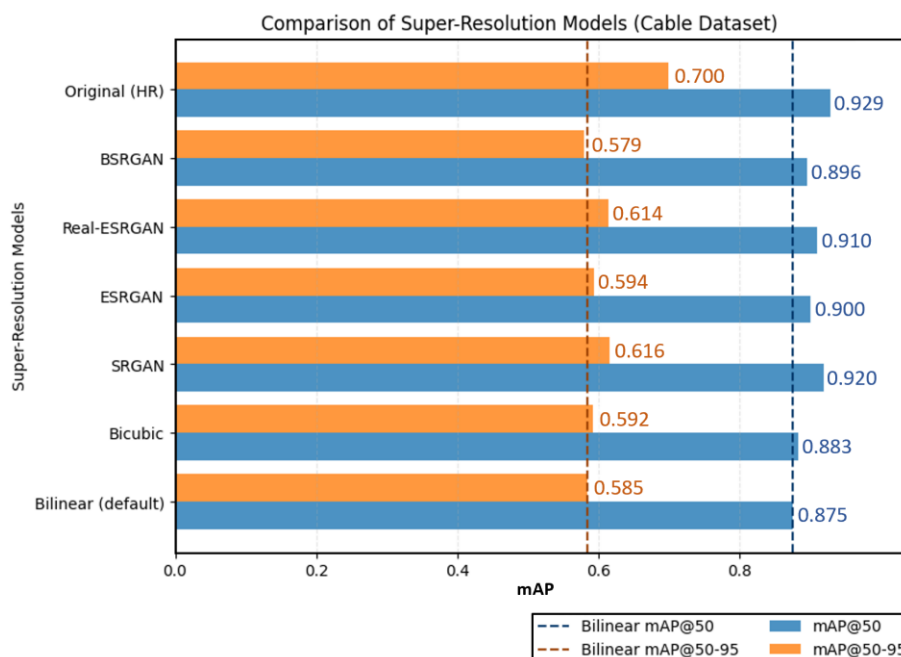
## ผลการวิจัย

จากผลการทดลองที่ได้ สามารถเปรียบเทียบประสิทธิภาพในการสร้างภาพความละเอียดสูงของแบบจำลอง Super-Resolution GAN ได้แก่ SRGAN ESRGAN Real-ESRGAN และ BSRGAN กับประสิทธิภาพในการสร้างภาพความละเอียดสูงแบบดั้งเดิม ได้แก่ Bilinear และ Bicubic Interpolation ในบริบทของการตรวจจบบัณฑิตขนาดเล็กได้ โดยพิจารณาจาก Mean Average Precision (mAP) และเวลาในการประมวลผลต่อภาพ (Inference Time per Image) ดังตารางที่ 1-3 และภาพที่ 6-8

**ตารางที่ 1** เวลาในการสร้างภาพ (Inference Time) บนชุดข้อมูลภาพสายเคเบิล

| Model                            | Bilinear<br>(default) | Bicubic | SRGAN | ESRGAN | Real-<br>ESRGAN | BSRGAN |
|----------------------------------|-----------------------|---------|-------|--------|-----------------|--------|
| Inference Time<br>(second/image) | -                     | 0.034   | 0.129 | 0.154  | 0.282           | 0.293  |

ผลการวัดประสิทธิภาพแบบจำลอง Super-Resolution GAN ของชุดข้อมูลภาพสายเคเบิล โดยแยกตามตัวชี้วัด ดังแผนภูมิต่อไปนี้



ภาพที่ 6 ผลการวัดประสิทธิภาพแบบจำลอง Super-Resolution GAN ของชุดข้อมูลภาพสายเคเบิล

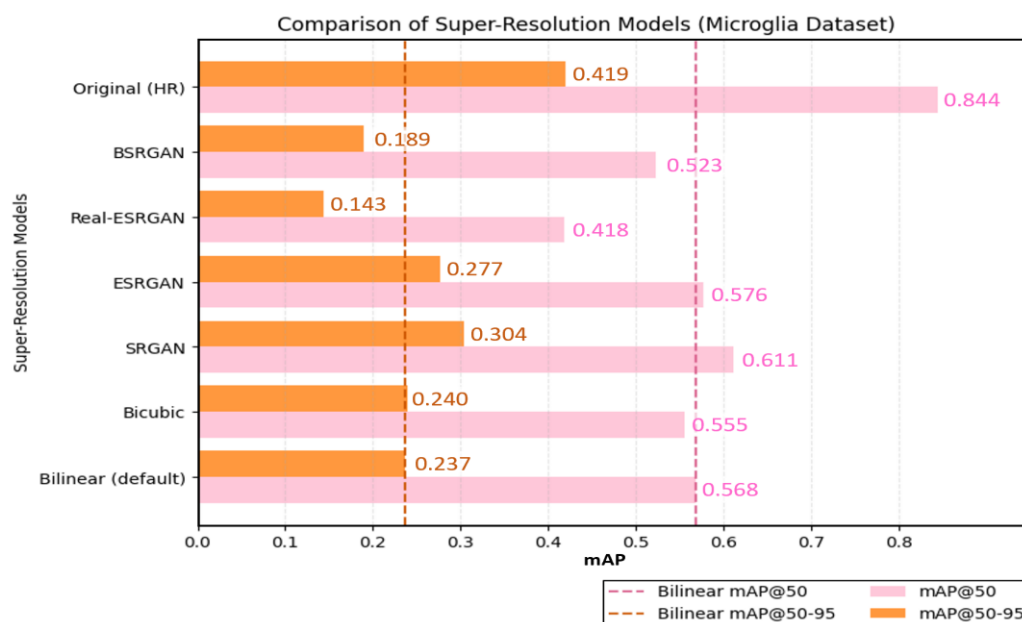
เมื่อพิจารณาถึงค่า mAP ซึ่งเป็นตัวชี้วัดสำคัญของการประเมินผลการตรวจจับวัตถุ พบว่า SRGAN ให้ค่าที่ดีที่สุด ใน mAP@50 (0.920) และค่าที่ดีที่สุด ใน mAP@50-95 (0.616) ซึ่งแสดงให้เห็นว่าแบบจำลองนี้ช่วยให้การตรวจจับวัตถุมีประสิทธิภาพสูงขึ้นเมื่อเทียบกับวิธีการแบบดั้งเดิม อย่างไรก็ตาม Real-ESRGAN ก็ให้ค่า mAP@50 (0.910) และ mAP@50-95 (0.614) ที่ใกล้เคียงกัน แสดงให้เห็นถึงความสามารถของแบบจำลองนี้ ในการช่วยปรับปรุงการตรวจจับวัตถุได้ดีเช่นกัน

ในแง่ของระยะเวลาการประมวลผล พบว่า Bicubic Interpolation ใช้เวลาน้อยที่สุดเพียง 0.034 วินาทีต่อภาพ ในขณะที่ Super-Resolution GAN ใช้เวลานานขึ้นอย่างเห็นได้ชัด โดย SRGAN ใช้เวลา 0.129 วินาที ESRGAN ใช้เวลา 0.154 วินาที Real-ESRGAN ใช้เวลา 0.282 วินาที และ BSRGAN ใช้เวลานานที่สุดที่ 0.293 วินาทีต่อภาพ แสดงให้เห็นว่าการใช้แบบจำลองที่ซับซ้อนขึ้นมีต้นทุนในด้านเวลา อย่างไรก็ตาม การปรับปรุงคุณภาพของการตรวจจับวัตถุอาจทำให้การประมวลผลที่นานขึ้นนั้นคุ้มค่าหากความแม่นยำเป็นสิ่งสำคัญ

ตารางที่ 2 เวลาในการสร้างภาพ (Inference Time) บนชุดข้อมูลภาพ Microglia

| Model                         | Bilinear (default) | Bicubic | SRGAN | ESRGAN | Real-ESRGAN | BSRGAN |
|-------------------------------|--------------------|---------|-------|--------|-------------|--------|
| Inference Time (second/image) | -                  | 0.018   | 0.135 | 0.156  | 0.325       | 0.293  |

ผลการวัดประสิทธิภาพแบบจำลอง Super-Resolution GAN ของชุดข้อมูลภาพ Microglia โดยแยกตามตัวชี้วัด แสดงในแผนภูมิแท่งต่อไปนี้



ภาพที่ 7 ผลการวัดประสิทธิภาพแบบจำลอง Super-Resolution GAN ของชุดข้อมูลภาพ Microglia

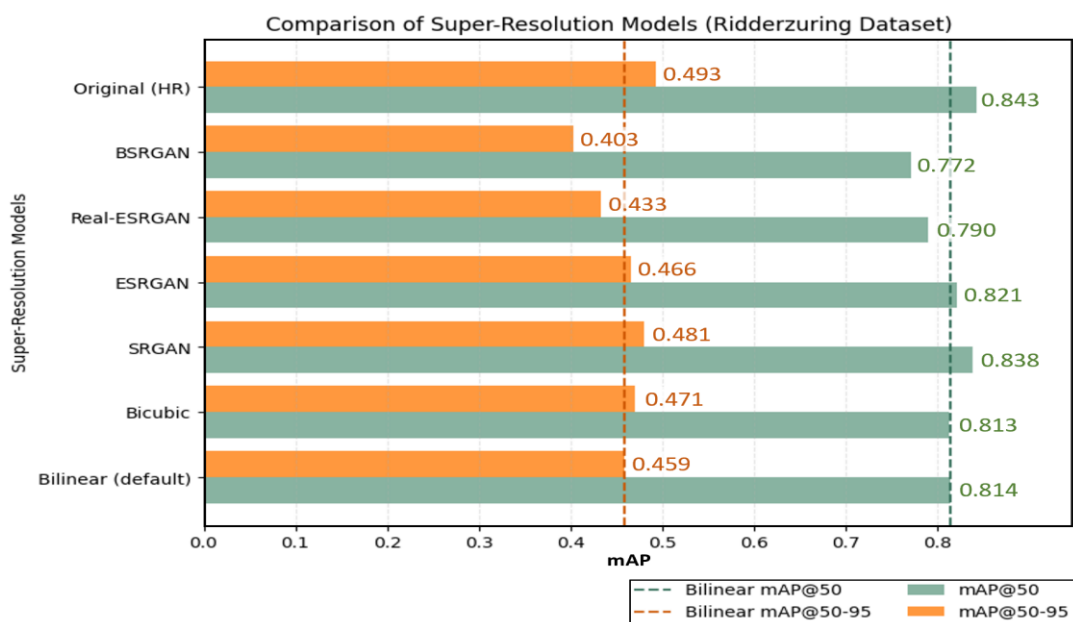
เมื่อพิจารณาถึงค่า mAP ซึ่งเป็นตัวชี้วัดสำคัญของการประเมินผลการตรวจจับวัตถุ พบว่า SRGAN ให้ค่า mAP@50 (0.611) และ mAP@50-95 (0.304) สูงที่สุด ซึ่งแสดงให้เห็นว่าแบบจำลองนี้สามารถช่วยให้การตรวจจับวัตถุมีประสิทธิภาพสูงขึ้นเมื่อเทียบกับวิธีการสร้างภาพความละเอียดสูงแบบดั้งเดิม อย่างไรก็ตาม ESRGAN ก็ให้ค่า mAP@50 (0.576) และ mAP@50-95 (0.277) ที่ใกล้เคียงกัน แสดงให้เห็นถึงความสามารถของแบบจำลองนี้ ในการช่วยปรับปรุงการตรวจจับได้บางส่วน แม้ว่าประสิทธิภาพโดยรวมจะยังเป็นรอง SRGAN

ในแง่ของระยะเวลาการประมวลผล พบว่า Bicubic ใช้เวลาน้อยที่สุดเพียง 0.018 วินาทีต่อภาพ ในขณะที่แบบจำลอง Super-Resolution GAN ใช้เวลานานขึ้นอย่างเห็นได้ชัด โดย SRGAN ใช้เวลา 0.135 วินาที ESRGAN ใช้เวลา 0.156 วินาที BSRGAN ใช้เวลา 0.305 วินาที และ Real-ESRGAN ใช้เวลานานที่สุดที่ 0.325 วินาทีต่อภาพ

ตารางที่ 3 เวลาในการสร้างภาพ (Inference Time) บนชุดข้อมูลภาพวชิพีซ Ridderzuring

| Model                         | Bilinear (default) | Bicubic | SRGAN | ESRGAN | Real-ESRGAN | BSRGAN |
|-------------------------------|--------------------|---------|-------|--------|-------------|--------|
| Inference Time (second/image) | -                  | 0.022   | 0.149 | 0.163  | 0.348       | 0.349  |

ผลการวัดประสิทธิภาพแบบจำลอง Super-Resolution GAN ของชุดข้อมูลภาพวชิพีซ Ridderzuring โดยแยกตามตัวชี้วัด แสดงในแผนภูมิแท่งต่อไปนี้



ภาพที่ 8 แสดงผลการวัดประสิทธิภาพแบบจำลอง Super-Resolution GAN ของชุดข้อมูลภาพเวชภาพ Ridderzuring

จากผลการทดลองพบว่า SRGAN ให้ค่า mAP@50 สูงสุดที่ 0.838 และ mAP@50-95 ที่ 0.481 ซึ่งสูงกว่าการใช้วิธีการสร้างภาพความละเอียดสูงแบบ Bicubic และ Bilinear เล็กน้อย ซึ่งได้ค่า mAP@50 เท่ากับ 0.813 และ 0.814 และ mAP@50-95 เท่ากับ 0.471 และ 0.459 ตามลำดับ แสดงให้เห็นว่าแบบจำลองนี้ สามารถช่วยเพิ่มประสิทธิภาพของการตรวจจับวัตถุได้ดีที่สุดเมื่อเทียบกับวิธีอื่น ๆ ในขณะเดียวกัน ESRGAN ต่ำกว่าเล็กน้อย โดยได้ค่า mAP@50 เท่ากับ 0.821 และ mAP@50-95 เท่ากับ 0.466 อย่างไรก็ตาม แบบจำลองที่มีความแม่นยำต่ำสุด คือ Real-ESRGAN และ BSRGAN โดยมีค่า mAP@50 เท่ากับ 0.790 และ 0.772 และ mAP@50-95 เท่ากับ 0.433 และ 0.403 ตามลำดับ

ในแง่ของระยะเวลาการประมวลผล พบว่า Bicubic ใช้เวลาน้อยที่สุดเพียง 0.022 วินาทีต่อภาพ ซึ่งเร็วกว่าวิธีการอื่น ๆ ในขณะที่ SRGAN และ ESRGAN ใช้เวลาประมวลผลที่ 0.149 และ 0.163 วินาทีต่อภาพ ตามลำดับ ส่วน Real-ESRGAN และ BSRGAN ใช้เวลานานที่สุดที่ 0.348 และ 0.349 วินาทีต่อภาพ

### สรุปและอภิปรายผล

การศึกษานี้เปรียบเทียบประสิทธิภาพของแบบจำลอง Super-Resolution GAN คือ SRGAN, ESRGAN, Real-ESRGAN และ BSRGAN กับวิธีสร้างภาพความละเอียดสูงแบบดั้งเดิม คือ Bilinear และ Bicubic โดยวัดผลลัพธ์ผ่านตัวชี้วัด mAP@50 และ mAP@50-95 สำหรับการตรวจจับวัตถุขนาดเล็กในภาพถ่ายสามชุดข้อมูล ได้แก่ ภาพสายเคเบิล Microglia และ วัชพืช Ridderzuring ซึ่งเป็นข้อมูลที่มีรายละเอียดซับซ้อนและต้องการความคมชัดสูง เพื่อให้สามารถตรวจจับวัตถุได้อย่างแม่นยำ



จากผลการทดลอง พบว่า SRGAN มีประสิทธิภาพสูงที่สุดในทุกชุดข้อมูล โดยให้ค่า mAP@50 เท่ากับ 0.920 สำหรับชุดข้อมูลภาพสายเคเบิล 0.611 สำหรับชุดข้อมูล Microglia และ 0.722 สำหรับชุดข้อมูลวัชพืช Ridderzuring ซึ่งสูงกว่าการสร้างภาพความละเอียดสูงด้วย Bicubic และ Bilinear นอกจากนี้ยังพบว่า Real-ESRGAN และ BSRGAN มีค่า mAP@50 และ mAP@50-95 ต่ำกว่าการสร้างภาพความละเอียดสูงแบบอื่น ๆ ในบางกรณี เช่น ชุดข้อมูล Microglia ที่ Real-ESRGAN และ BSRGAN ให้ค่า mAP@50 ที่ 0.418 และ 0.523 ตามลำดับ ซึ่งต่ำกว่าวิธีอื่น ๆ อย่างชัดเจน

งานวิจัยนี้ ได้รับการออกแบบการทดลองภายใต้เงื่อนไขที่ควบคุมตัวแปรให้เหมือนกันทุกประการ เช่น ขนาดของภาพ แบบจำลองที่ใช้ในการตรวจจับ ใช้ชุดข้อมูลเดียวกัน และจำนวนรอบการทดลองที่เท่ากัน ความแตกต่างที่ปรากฏในค่า mAP@50 และ mAP@50-95 จึงสามารถสะท้อนถึงผลของเทคนิค Super-Resolution โดยตรงอย่างมีนัยยะในเชิงวิศวกรรม

ในงานวิจัยด้านการเรียนรู้เชิงลึก โดยเฉพาะในบริบทของการตรวจจับวัตถุและการประมวลผลภาพ มักใช้การประเมินผลเชิงเปรียบเทียบด้วยชุดข้อมูลมาตรฐาน (benchmark datasets) ร่วมกับตัวชี้วัดเชิงประสิทธิภาพ เช่น mAP โดยตรง มากกว่าการใช้การทดสอบนัยสำคัญทางสถิติแบบดั้งเดิม ความแตกต่างของค่าตัวชี้วัดที่เกิดขึ้น จึงสามารถนำมาใช้ในการวิเคราะห์ประสิทธิภาพของแบบจำลองแต่ละเทคนิคได้อย่างน่าเชื่อถือและถือว่าเพียงพอ ในการสรุปแนวโน้มของประสิทธิภาพที่เกิดขึ้นจริง

เมื่อเปรียบเทียบกับงานวิจัยที่ผ่านมา พบว่า ผลลัพธ์ในงานวิจัยนี้มีความสอดคล้องกับข้อค้นพบของ Telçeken et.al., (2024) ซึ่งได้นำเสนอระบบ Super-Resolution Object Detection (SROD) โดยผสมผสานเทคนิคการเพิ่มความละเอียดด้วย SRGAN เข้ากับการตรวจจับวัตถุด้วย YOLO งานวิจัยดังกล่าว แสดงให้เห็นว่าการเพิ่มความละเอียดก่อนการตรวจจับวัตถุในภาพถ่ายดาวเทียม และภาพโดรน ช่วยเพิ่มค่า mAP ได้อย่างมีนัยสำคัญ โดยเฉพาะในกรณีของวัตถุขนาดเล็ก ซึ่งสอดคล้องกับงานวิจัยนี้ที่ SRGAN ช่วยเพิ่มความแม่นยำในการตรวจจับวัตถุขนาดเล็ก ได้ดีกว่าวิธีการสร้างภาพความละเอียดสูงแบบดั้งเดิม

นอกจากนี้ ผลการทดลองในงานวิจัยนี้ยังสอดคล้องกับงานของ Yigit (2023) ซึ่งศึกษาการใช้ Real-ESRGAN เพื่อเพิ่มคุณภาพของภาพถ่ายทางอากาศจากโดรนก่อนนำไปตรวจจับวัตถุด้วย YOLOv7 โดยพบว่าการใช้ Real-ESRGAN สามารถช่วยเพิ่มค่า mAP ได้อย่างมีนัยสำคัญ อย่างไรก็ตามเมื่อเปรียบเทียบกับผลการทดลองในงานวิจัยนี้ พบว่า Real-ESRGAN แม้จะมีศักยภาพในการปรับปรุงคุณภาพของภาพ แต่ในบางกรณี เช่น ชุดข้อมูล Microglia กลับมีค่า mAP@50 ต่ำกว่าวิธีการอื่น ซึ่งอาจสะท้อนถึงข้อจำกัดของ Real-ESRGAN ในการรักษารายละเอียดโครงสร้างเฉพาะที่สำคัญสำหรับการตรวจจับวัตถุขนาดเล็กในบริบทที่มีความละเอียดสูงมากหรือมีลักษณะวัตถุซับซ้อน

แม้ว่าแบบจำลอง SRGAN จะสามารถเพิ่มความแม่นยำของการตรวจจับวัตถุขนาดเล็กได้ แต่ก็มาพร้อมกับต้นทุนด้านเวลาในการประมวลผลที่สูงขึ้น โดยใช้เวลาในการสร้างภาพความละเอียดสูงประมาณ 0.129 - 0.149 วินาทีต่อภาพ ซึ่งมากกว่าวิธี Bicubic ที่ใช้เวลาเพียง 0.018 - 0.034 วินาทีต่อภาพ ทำให้ต้องพิจารณาความสมดุลระหว่างความเร็ว และความแม่นยำในการใช้งานจริง



นอกจากนี้ ยังพบว่าข้อมูลต้นฉบับความละเอียดสูงให้ค่าคะแนนสูงสุดในทุกชุดข้อมูล ซึ่งเป็นสิ่งที่คาดการณ์ได้ เนื่องจากไม่มีการสูญเสียรายละเอียดจากการขยายภาพเทคนิค Super-Resolution GAN จึงมีบทบาทสำคัญในการลดช่องว่างระหว่างภาพความละเอียดต่ำและภาพต้นฉบับ ทำให้สามารถนำไปปรับใช้กับงานตรวจจับวัตถุขนาดเล็กในสภาพแวดล้อมที่ความละเอียดของภาพเป็นข้อจำกัด

งานวิจัยนี้ มีการนำเสนอองค์ความรู้ใหม่ที่สำคัญหลายประการงานวิจัย ได้นำเสนอการประเมินเปรียบเทียบประสิทธิภาพของเทคนิค Super-Resolution ที่แตกต่างกัน ได้แก่ SRGAN, ESRGAN, Real-ESRGAN และ BSRGAN ร่วมกับการตรวจจับวัตถุขนาดเล็กบนชุดข้อมูลที่มีความหลากหลาย ทั้งในบริบทของสายเคเบิลทางวิศวกรรม เซลล์ชีวภาพ Microglia และวัชพืชในสิ่งแวดล้อมธรรมชาติ การวิเคราะห์นี้ช่วยให้เห็นภาพรวมที่ชัดเจนของประสิทธิภาพ แต่ละเทคนิคเมื่อใช้งานกับข้อมูลที่มีลักษณะต่างกัน

นอกจากนี้ งานวิจัยยังชี้ให้เห็นถึงความสำคัญของการเลือกเทคนิค Super-Resolution ที่เหมาะสมกับลักษณะข้อมูล โดยผลการทดลองพบว่า แม้แบบจำลองบางประเภท เช่น Real-ESRGAN และ BSRGAN จะสามารถสร้างภาพที่มีคุณภาพการมองเห็นสูง แต่กลับไม่สามารถเพิ่มความแม่นยำในการตรวจจับวัตถุได้ดีในบางกรณี ในทางตรงกันข้าม SRGAN ซึ่งเน้นการรักษาโครงสร้างของภาพมากกว่าความสมจริงในการมองเห็น กลับแสดงผลลัพธ์ที่ดีกว่าอย่างชัดเจนในทุกชุดข้อมูลที่ศึกษา ข้อค้นพบนี้เน้นย้ำว่าการเลือกเทคนิคการเพิ่มความละเอียดต้องพิจารณาตามลักษณะของข้อมูลและเป้าหมายการใช้งาน

สุดท้าย งานวิจัยนี้ยังช่วยตอกย้ำศักยภาพของการประยุกต์ใช้เทคนิค Super-Resolution ในการสนับสนุนการพัฒนาเทคโนโลยี ตรวจจับวัตถุขนาดเล็กจากข้อมูลภาพความละเอียดต่ำ ซึ่งมีความสำคัญในหลายอุตสาหกรรม เช่น การสำรวจและติดตามสภาพแวดล้อมด้วยโดรน การเฝ้าระวังด้านความมั่นคง และการวินิจฉัยทางการแพทย์จากภาพจุลทรรศน์ ผลลัพธ์ที่ได้สามารถนำไปต่อยอดเพื่อพัฒนาระบบตรวจจับวัตถุที่มีความแม่นยำสูง และใช้งานได้จริงในสภาวะแวดล้อมที่ซับซ้อนในอนาคต

### ข้อเสนอแนะ

1. การเพิ่มชุดข้อมูลที่มีความหลากหลายทั้งในด้านขนาด ปริมาณ หรือประเภทของวัตถุที่ต้องการตรวจจับ จะช่วยให้สามารถเปรียบเทียบแบบจำลอง Super-Resolution GAN ได้ในหลายแง่มุมมากขึ้น
2. การทดสอบกับแบบจำลองการตรวจจับวัตถุอื่น ๆ เช่น Faster R-CNN หรือ DETR เพื่อเปรียบเทียบประสิทธิภาพและเพิ่มความน่าเชื่อถือของผลลัพธ์

### เอกสารอ้างอิง

Ciaglia, F., Zuppichini, F. S., Guerrie, P., McQuade, M., & Solawetz, J. (2022). Roboflow 100: A rich, multi-domain object detection benchmark. arXiv preprint arXiv:2211.13523.



- Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. C., & Bengio, Y. (2014). Generative Adversarial Nets. *Neural Information Processing Systems*.
- Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., & Wang, Z. (2017). Photo-realistic single image super-resolution using a generative adversarial network. *Proceedings of the IEEE conference on computer vision and pattern recognition*.
- Padilla, R., Netto, S. L., & Silva, E. A. B. d. (2020). A Survey on Performance Metrics for Object-Detection Algorithms. *2020 International Conference on Systems, Signals and Image Processing (IWSSIP)*.
- Redmon, J., & Farhadi, A. (2017). YOLO9000: better, faster, stronger. *Proceedings of the IEEE conference on computer vision and pattern recognition*.
- Telçeken, M., Akgun, D., Kacar, S., & Bingol, B. (2024). A New Approach for Super Resolution Object Detection Using an Image Slicing Algorithm and the Segment Anything Model. *Sensors*, 24(14).
- Wang, A., Chen, H., Liu, L., Chen, K., Lin, Z., Han, J., & Ding, G. (2024). Yolov10: Real-time end-to-end object detection. *arXiv preprint arXiv:2405.14458*.
- Wang, X., Xie, L., Dong, C., & Shan, Y. (2021). Real-esrgan: Training real-world blind super-resolution with pure synthetic data. *Proceedings of the IEEE/CVF international conference on computer vision*.
- Wang, X., Yu, K., Wu, S., Gu, J., Liu, Y., Dong, C., Qiao, Y., & Change Loy, C. (2018). Esrgan: Enhanced super-resolution generative adversarial networks. *Proceedings of the European conference on computer vision (ECCV) workshops*.
- Yigit, S. (2023). Improving object detection of UAV images with Real-ESRGAN. *Recent Advances in Science and Engineering*.
- Zhang, K., Liang, J., Van Gool, L., & Timofte, R. (2021). Designing a practical degradation model for deep blind image super-resolution. *Proceedings of the IEEE/CVF International Conference on Computer Vision*.