

Road defect detection using Mask R-CNN

PhoPhrai PhoJundee^{1,*}, Warinthon Nuprakhon¹, Keng Channual¹, Nitinan Mata¹
and Saengdao Noppitak¹

ABSTRACT

Road damage can impede traffic flow and potentially cause accidents, leading to loss of life and property. With the advancement of deep learning, it is now possible to automatically detect road surface defects. This study proposes a road defect detection model based on Mask R-CNN, a state-of-the-art framework for object detection and image segmentation. The model was trained on 240 images (80% of the dataset) and validated on 30 images, with different epoch settings and learning rates compared to determine the optimal configuration. Performance evaluation was then conducted using a separate test set of 30 images. Experimental results indicate that the proposed model achieved an average precision (AP) of 23.35 at a 70% confidence threshold, obtained with a learning rate of 0.001 and 100 training epochs.

Keywords: Road defect, Object Detection, Convolutional Neural Network, Mask R-CNN

Published Online: 24 September 2025

ISSN: 3057-0778 (Online)

PhoPhrai PhoJundee^{1,*}

¹Faculty of Science Buriram Rajabhat University, Buriram, Thailand
(catwater1233@gmail.com)

Warinthon Nuprakhon¹

¹Faculty of Science Buriram Rajabhat University, Buriram, Thailand
(630112418019@bru.ac.th)

Keng Channual¹

¹Faculty of Science Buriram Rajabhat University, Buriram, Thailand
(keng.ch@bru.ac.th)

Nitinan Mata¹

¹Faculty of Science Buriram Rajabhat University, Buriram, Thailand
(nitinan.mt@bru.ac.th)

Saengdao Noppitak¹

¹Faculty of Science Buriram Rajabhat University, Buriram, Thailand
(sangdaow.np@bru.ac.th)

*Corresponding Author

Received date: 16 March 2023

Revised date: 15 May 2023

Accepted date: 28 August 2025

การตรวจจับพื้นผิวถนนชำรุดโดยใช้ Mask R-CNN

โพธิ์ไพร โพธิ์จันดี^{1,*} วรินทร์ หนูประโคน¹ เก่ง จันทรนวล¹ นิธินันท์ มาตา¹ และแสงดาว นพพิทักษ์¹

บทคัดย่อ

ถนนที่มีความเสียหายหรือชำรุดทำให้การจราจรช้าลงหรือ ก่อให้เกิดอุบัติเหตุ และการสูญเสียชีวิตหรือทรัพย์สิน ปัจจุบันสามารถนำเทคโนโลยีการเรียนรู้เชิงลึกสำหรับตรวจจับพื้นผิวถนนชำรุดได้งานวิจัยนี้มุ่งเน้นเพื่อพัฒนาแบบจำลองตรวจจับพื้นผิวถนนชำรุดโดยใช้ Mask R-CNN ซึ่งเป็นเทคนิคที่มีประสิทธิภาพสำหรับงานด้านการตรวจจับวัตถุและแบ่งส่วนภาพการทดลองเริ่มจากการฝึกฝนโมเดลกับชุดข้อมูลสำหรับการเรียนรู้จำนวน 240 ภาพ (80%) และชุดข้อมูลสำหรับตรวจสอบ 30 ภาพ โดยเปรียบเทียบจำนวนรอบในการเรียนรู้และอัตราการเรียนรู้เพื่อหาโมเดลที่มีประสิทธิภาพดีที่สุด จากนั้นนำไปทดสอบกับชุดข้อมูลสำหรับทดสอบจำนวน 30 ภาพ ผลการทดลอง พบว่า ประสิทธิภาพของโมเดลที่มีความแม่นยำเฉลี่ยในการตรวจจับตำแหน่งถนนชำรุด มีความแม่นยำเฉลี่ยที่ระดับความมั่นใจ 70% ที่ดีที่สุดคือ 23.35 ได้จากการทดลองที่อัตราการเรียนรู้ 0.001 และจำนวนรอบที่ใช้ในการรัน 100 รอบ

คำสำคัญ: พื้นผิวถนนชำรุด การตรวจจับวัตถุ โครงข่ายประสาทแบบคอนโวลูชัน มาร์ค อา-ซีเอ็นเอ็น

Published Online: 24 กันยายน 2568

ISSN: 2730-3829

โพธิ์ไพร โพธิ์จันดี^{1*}

¹คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏบุรีรัมย์ จังหวัดบุรีรัมย์
(catwater1233@gmail.com)

วรินทร์ หนูประโคน¹

¹คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏบุรีรัมย์ จังหวัดบุรีรัมย์
(630112418019@bru.ac.th)

เก่ง จันทรนวล¹

¹คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏบุรีรัมย์ จังหวัดบุรีรัมย์
(keng.ch@bru.ac.th)

นิธินันท์ มาตา¹

¹คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏบุรีรัมย์ จังหวัดบุรีรัมย์
(nitinan.mt@bru.ac.th)

แสงดาว นพพิทักษ์¹

¹คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏบุรีรัมย์ จังหวัดบุรีรัมย์
(sangdaow.np@bru.ac.th)

*Corresponding Author

Received date: 16 มีนาคม 2566

Revised date: 15 พฤษภาคม 2566

Accepted date: 28 สิงหาคม 2568

1. บทนำ

ถนนมีความสำคัญในการเดินทาง การขนส่งสินค้า ในปัจจุบันมีการใช้รถใช้ถนนในการเดินทาง มีการใช้รถยนต์หลากหลายประเภทบนท้องถนน การขับขี่ยานพาหนะที่ผิดประเภท และการบรรทุกเกินขนาดเกินกว่ามาตรฐานของถนนที่จะรับน้ำหนักได้ ทำให้อุบัติเหตุหรือการจราจรที่ช้าลงก่อให้เกิดอุบัติเหตุ และการสูญเสียชีวิตหรือทรัพย์สินเสียหาย (พิชยุท บัญธน และยุทธนา สีขวา, 2564) จากสำรวจการแจ้งซ่อมถนนชำรุด เช่น กรณีขององค์การบริหารส่วนตำบลบ้านกลาง อ.หล่มสัก จ.เพชรบูรณ์ (นายสมพงษ์ มาลา, 2565) ได้เปิดโอกาสให้ชาวบ้านร้องเรียนเรื่องถนนชำรุดภายในหมู่บ้านผ่านตัวเว็บไซต์ โดยการถ่ายภาพพร้อมพิกัดสถานที่ กรณีของกรมทางหลวงชนบท (กรมทางหลวงชนบท, 2565) รับแจ้งเหตุให้ความช่วยเหลือประชาชน โดยวิธีโทรเข้าที่สายด่วน 1146 ตลอด 24 ชั่วโมง และสำนักการช่าง อบจ.ขอนแก่น (สำนักการช่าง อบจ.ขอนแก่น, 2565) เปิดให้บริการร้องทุกข์ ร้องเรียนทางอินเทอร์เน็ต เป็นช่องทางให้บุคคลทั่วไปได้สามารถร้องทุกข์ ร้องเรียนมายังหน่วยงานได้โดยตรง

ปัจจุบันเทคโนโลยีด้านการเรียนรู้ของเครื่อง (Machine Learning) ถูกพัฒนาไปจนทำให้คอมพิวเตอร์สามารถจำแนกภาพถ่ายได้ด้วยการเรียนรู้จากภาพเรียกว่าการเรียนรู้เชิงลึก (Deep Learning) ซึ่งทำให้คอมพิวเตอร์สามารถเรียนรู้ และพัฒนาตัวเองได้คล้ายกับสมองมนุษย์ โดยเรียนรู้จากข้อมูลเป็นจำนวนมาก จึงสามารถคัดแยกข้อมูลได้อย่างรวดเร็ว และแม่นยำ เช่น ภาพคน สัตว์ สิ่งของ เป็นต้น การตรวจจับวัตถุด้วยการเรียนรู้เชิงลึก (Object Detection) การตรวจจับวัตถุเป็นเทคโนโลยีในทางคอมพิวเตอร์ที่ใช้ Deep Learning และ Image Processing โดยการพยายามทำให้คอมพิวเตอร์สามารถตรวจจับวัตถุได้ จะใช้วิธีแบ่งภาพออกเป็นหลาย ๆ ส่วน แล้วนำภาพทั้งหมดตรวจสอบด้วย Convolutional Neural Network (CNN) โดยการตรวจจับวัตถุด้วยการเรียนรู้เชิงลึกมีหลายวิธีได้แก่ R-CNN, Fast R-CNN, Faster R-CNN, และ Mask R-CNN โครงการวิจัยนี้สนใจเทคนิคสนใจ Mask R-CNN เนื่องจากเป็นแบบจำลองที่พัฒนามาจาก R-CNN, Fast R-CNN และ Faster R-CNN ความสามารถของ Mask R-CNN นอกจากจะตรวจจับวัตถุในภาพได้แล้ว ยังสามารถแยกหรือระบุตำแหน่งของวัตถุแต่ละอย่างในภาพได้ในระดับพิกเซลได้อีกด้วย เช่น งานวิจัยของ พิชิตชัย พิมพ์โคตร (2564) เลือกใช้โมเดล Mask R-CNN ประยุกต์ใช้กับการตรวจจับ และการคัดแยกโลโก้สินค้าซึ่งสามารถนำไปประยุกต์ใช้ทั้งในภาครัฐ และภาคธุรกิจ หรือใช้ในการตรวจจับรูปภาพเครื่องมือแอลกอฮอล์ที่ผิดกฎหมายโฆษณาโดยใช้เทคนิค (รัตนโชติ พันธุ์วิไล, 2562)

วัตถุประสงค์

1. เพื่อพัฒนาแบบจำลองตรวจจับพื้นผิวถนนชำรุดโดยใช้ Mask R-CNN
2. เพื่อหาประสิทธิภาพแบบจำลองตรวจจับพื้นผิวถนนชำรุดโดยใช้ Mask R-CNN

2. วิธีการวิจัย

คณะผู้วิจัยได้ทำการศึกษาทฤษฎีที่เกี่ยวข้องกับงานวิจัย ต่อมาได้ทำการสร้างชุดข้อมูลที่ถูกเก็บรวบรวมรูปภาพถนนชำรุดมาจาก Google map street view โดยชุดข้อมูลนี้มีจำนวนทั้งหมด 300 ภาพ ภาพทั้งหมดจะถูกนำมาทำผลเฉลย (Ground truth) หลังจากการทำผลเฉลยเรียบร้อยแล้วทำการแบ่งชุดข้อมูลสำหรับการเรียนรู้ และทดสอบ โดยการสุ่ม แบ่งเป็นชุดข้อมูลสำหรับการเรียนรู้ 80% จะมี 240 ภาพ ชุดข้อมูลสำหรับตรวจสอบ 10% มี 30 ภาพ และชุดข้อมูลสำหรับการทดสอบ 10% มี 30 ภาพ หลังจากการแบ่งชุดข้อมูลสำหรับการเรียนรู้ และทดสอบ จะทำการฝึกฝนโมเดลโดยการเปรียบเทียบจำนวนรอบที่ใช้ในการรัน และอัตราการเรียนรู้โดยตารางจะแสดงจำนวนรอบที่ใช้ในการรัน (100, 300, 500, 1000) Epoch และอัตราการเรียนรู้ (0.001 และ 0.0001) โดยมีการศึกษาทฤษฎีที่เกี่ยวข้องกับงานวิจัยดังนี้

ทฤษฎีที่เกี่ยวข้องกับงานวิจัยจากการสืบค้นเอกสารพบว่า Detection มีหลากหลายวิธี และวิธีที่เป็นนิยม 1 คือ mask R-CNN เป็นแนวคิดในการตรวจจับ และแบ่งส่วนวัตถุ มีการพัฒนาต่อยอดมาจาก R-CNN, Fast R-CNN, Faster R-CNN เป็นการนำ Object Detection ซึ่งมีข้อจำกัดหลายอย่าง ต่อมามีการพัฒนาเพื่อแก้ไขข้อจำกัด

Mask R-CNN จะเป็น Instance Segmentation ที่เป็นแบบ Multiple Object เฟรมเวิร์คของ Mask R-CNN ดัง Figure 1 (พิชิตชัย พิมพ์โคตร, 2564)

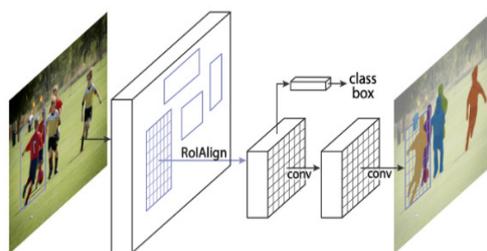


Figure1 Framework of Mask R-CNN

ค่าไอโอยู (Intersection over Union: IoU) ค่าไอโอยูเป็นค่าวัดความแม่นยำของการตรวจจับวัตถุ ซึ่งเป็น การวัดอัตราส่วนพื้นที่ที่ทับซ้อนกัน (Area of Overlap) ระหว่างพื้นที่ที่ทำนาย (Prediction Area) กับพื้นที่จริงของ วัตถุ (Ground Truth Area)หารด้วยพื้นที่รวมทั้งหมด (Area of Union) ระหว่างพื้นที่ ที่ทำนายกับพื้นที่ที่จริงของ วัตถุ ดัง Figure 2 (Rezatofighi et al., 2019)

$$\text{Intersection over Union (IoU)} = \frac{\text{Area of Overlap}}{\text{Area of Union}}$$

Figure 2 Equation for calculating the Intersection over Union (IoU)

ความแม่นยำเฉลี่ย (Average precision) (พิชยุทธ บุญตน และ ยุทธนา สีขาว, 2564) ใช้วัดความแม่นยำ ของโมเดล ตรวจจับวัตถุโดยใช้ค่า Precision และ Recall เพื่อเรียงความมั่นใจที่ทำนายสูงสุด ตามลำดับ เมื่อกำหนดให้ IOU มากกว่าหรือ เท่ากับ 0.5 จากนั้นปรับค่า Precision ด้วย ฟังก์ชัน Interpolated Precision เพื่อหาพื้นที่ใต้กราฟ (AUC) แล้วจะแบ่งกราฟออกเป็น ส่วน ๆ เพื่อคำนวณความแม่นยำเฉลี่ย ในขั้นตอนสุดท้ายดัง สมการ 1 และ 2

$$P_{in}(r) = \max_{\tilde{r}: \tilde{r} \geq r} p(\tilde{r}) \quad (1)$$

โดยที่ p คือ Precision

\tilde{r} คือ Recall

การแบ่งกราฟออกเป็น ส่วน ๆ เพื่อหาพื้นที่ใต้กราฟทำได้ดังนี้

$$AP = \sum (r_{n+1} - r_n) p_{interp}(r_{n+1}) \quad (2)$$

โดยที่ผลลัพธ์ของ AP คือ ความแม่นยำเฉลี่ยของการทำนาย

งานวิจัยนี้ได้ใช้ชุดข้อมูลที่ถูกรวบรวมรูปภาพถนนชำรุดมาจาก Google map street view โดยชุดข้อมูล นี้มีจำนวนทั้งหมด 300 ภาพ และภาพทั้งหมดจะถูกนำมาทำผลเฉลย (Ground truth) การแบ่งชุดข้อมูลสำหรับการ เรียนรู้ และทดสอบโดยใช้การสุ่ม โดยแบ่งเป็นชุดข้อมูลสำหรับการเรียนรู้ 80% (240 ภาพ) ชุดข้อมูลสำหรับตรวจสอบ 10% (30 ภาพ) ชุดข้อมูลสำหรับทดสอบ 10% (30 ภาพ) ดัง Figure 3-5



Figure 3 Example images of damaged road surfaces: (a) original image, and (b) result



Figure 4 Example images of damaged road surfaces: (a) original image, and (b) result



Figure 5 Example images of damaged road surfaces: (a) original image, and (b) result

3. ผลการวิจัย

งานวิจัยนี้ ได้นำเสนอวิธีการตรวจจับพื้นผิวถนนชำรุดด้วย Mask R-CNN การทดลองทั้งหมดรันบน Google Colab Pro โดยเริ่มจากการฝึกฝนโมเดลจำนวนภาพ 240 ภาพ เปรียบเทียบ Epoch (100, 300, 500, 1000) และอัตราการเรียนรู้ (0.001 และ 0.0001) ดัง Figure 6 จากนั้นนำไปทดสอบกับชุดข้อมูลทดสอบจำนวน 30 ภาพ โดยมีผลการทดลองดัง Table 1



Figure 6 Model training error

จาก Figure 6 แสดงค่าความผิดพลาดของการฝึกฝนโมเดลจำนวนรอบการรัน 1000 Epoch ในการฝึกฝนโมเดลค่าความผิดพลาดลดลงอย่างต่อเนื่อง

Table 1 Average precision for detecting road

| No. | Epoch | Learning Rate | AP70 | AP50 |
|-----|-------|---------------|---------------|---------------|
| 1. | 100 | 0.0001 | 1.3270 | 5.611 |
| 2. | 300 | 0.0001 | 12.651 | 27.348 |
| 3. | 500 | 0.0001 | 19.689 | 40.312 |
| 4. | 1000 | 0.0001 | 21.533 | 39.536 |
| 5. | 100 | 0.0010 | 23.348 | 43.243 |
| 6. | 300 | 0.0010 | 21.051 | 47.726 |
| 7. | 500 | 0.0010 | 17.740 | 34.368 |
| 8. | 1000 | 0.0010 | 16.243 | 30.469 |

จาก Table 1 ผลการทดลองเพื่อหาความแม่นยำเฉลี่ยของการตรวจจับตำแหน่งถนนชำรุด พบว่าความแม่นยำเฉลี่ยที่ระดับความมั่นใจ 70% (AP70) ที่ดีที่สุดคือ 23.35 ได้จากการทดลองที่อัตราการเรียนรู้ 0.001 จำนวนรอบที่ใช้ในการรัน 100 Epoch และความแม่นยำเฉลี่ยที่ระดับความมั่นใจ 50% (AP50) ที่ดีที่สุดคือ 47.73 ได้จากการทดลองที่อัตราการเรียนรู้ 0.001 จำนวนรอบที่ใช้ในการรัน 300 Epoch ดัง Figure 7-9

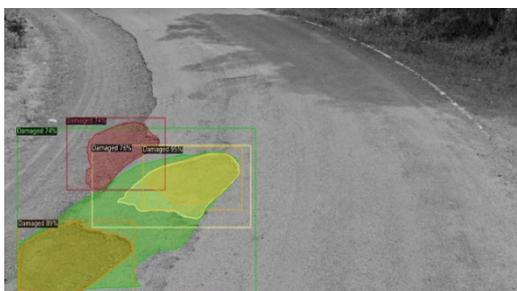


Figure 7 Road damage detection results

จาก Figure 7 ค่าแม่นยำที่ดีที่สุดอยู่ที่ 95% เนื่องจากสภาพของถนนชำรุดมีขอบหลุมชัดเจน ทำให้ได้ค่าความแม่นยำสูง ส่วนค่าความแม่นยำต่ำสุดอยู่ที่ 74% เนื่องจากชุดข้อมูลมีผลเฉลยที่ใช้ในการเรียนรู้มีลักษณะไม่ครอบคลุมลักษณะของพื้นผิวถนนชำรุดจึงทำให้ไม่สามารถเรียนรู้ได้



Figure 8 Road damage detection results

จาก Figure 8 ค่าแม่นยำที่ดีที่สุดอยู่ที่ 93% เนื่องจากสภาพของการชำรุดที่มีขอบหลุมชัดเจนทำให้ได้ค่าความแม่นยำสูง ส่วนค่าความแม่นยำต่ำสุดอยู่ที่ 72% เนื่องจากชุดข้อมูลที่นำมาทำผลเฉลยที่ใช้ในการเรียนรู้มีชุดข้อมูลถนนชำรุดที่น้อยเกินไปทำให้ไม่สามารถเรียนรู้ได้จึงต้องทำผลเฉลยของชุดข้อมูลถนนให้มากขึ้น และครอบคลุมเพื่อเพิ่มความแม่นยำ



Figure 9 Road damage detection results

จาก Figure 9 ค่าแม่นยำที่ดีที่สุดอยู่ที่ 94% เนื่องจากสภาพของถนนชำรุดที่มีขอบหลุมชัดเจน ทำให้ได้ค่าความแม่นยำสูง ส่วนค่าความแม่นยำต่ำสุดอยู่ที่ 74% เนื่องจากชุดข้อมูลมีผลเฉลยที่ใช้ในการเรียนรู้มีลักษณะไม่ครอบคลุมลักษณะของพื้นผิวถนนชำรุดจึงทำให้ไม่สามารถเรียนรู้ได้

ข้อเสนอแนะ

1. ในงานวิจัยครั้งต่อไป คณะผู้วิจัยจะแก้ปัญหาโดยการเพิ่มจำนวนชุดข้อมูลที่มากขึ้น
2. ปรับปรุงประสิทธิภาพการตรวจจับพื้นผิวถนนชำรุดด้วยวิธีการเพิ่มภาพ (Data Augmentation) หรือ Generative Adversarial Networks (GANs) เป็นต้น เพื่อที่จะนำไปพัฒนาระบบแจ้งถนนชำรุดผ่านเว็บแอปพลิเคชันอย่างมีประสิทธิภาพ

4. สรุปผลการวิจัย

งานวิจัยนี้ ได้กล่าวถึงวิธีการตรวจจับพื้นผิวถนนชำรุดด้วยโดยใช้การเรียนรู้เชิงลึกด้วยวิธี Mask R-CNN โดยชุดข้อมูลนี้มีจำนวนทั้งหมด 300 ภาพ ข้อมูลทั้งหมดจะถูกนำมาทำผลเฉลย (Ground truth) จำนวน 538 จุด จากนั้นแบ่งเป็นชุดข้อมูลสำหรับการเรียนรู้ 80% (240 ภาพ) ชุดข้อมูลสำหรับตรวจสอบ 10% (30 ภาพ) ชุดข้อมูลสำหรับทดสอบ 10% (30 ภาพ) ได้ทำการทดลองโดยเริ่มจากการฝึกฝนโมเดล จำนวนภาพ 240 ภาพ เปรียบเทียบ Epoch (100, 300, 500, 1000) และอัตราการเรียนรู้ (0.001 และ 0.0001) พบว่า ค่าความผิดพลาดในการฝึกฝนโมเดลลดลงอย่างต่อเนื่อง และใช้ข้อมูลสำหรับทดสอบจำนวน 30 ภาพ เพื่อหาความแม่นยำเฉลี่ยของการตรวจจับตำแหน่งถนนชำรุด พบว่า ความแม่นยำเฉลี่ยที่ระดับความมั่นใจ 70% ที่ดีที่สุดคือ 23.35 ได้จากการทดลองที่อัตราการเรียนรู้ 0.001 จำนวนรอบที่ใช้ในการรัน 100 Epoch และความแม่นยำเฉลี่ยที่ระดับความมั่นใจ 50% ที่ดี

ที่สุดคือ 47.73 ได้จากการทดลองที่อัตราการเรียนรู้ 0.001 จำนวนรอบที่ใช้ในการรัน 300 Epoch จากความแม่นยำเฉลี่ย ที่ระดับความมั่นใจ 50% จากผลการทดลองประสิทธิภาพการตรวจจับพื้นผิวถนนชำรุดยังอยู่ในระดับต่ำ เนื่องจากมีจำนวนภาพถนนชำรุดน้อยเกินไป

REFERENCES

- Boonton, P., & Sikwa, Y. (2021). *Road damage classification* (Undergraduate thesis). Mahasarakham University.
- Department of Rural Roads. (2022). Department of Rural Roads addresses areas at risk of road subsidence. Retrieved 15 December 2022. Form <https://drr.go.th/?p=13824>.
- Khon Kaen Provincial Administrative Organization, Department of Public Works. (2022). Public Services: Complaints and Grievances. Retrieved 15 December 2022. From http://www.kkpao.go.th/dep/publicworks/menu7-Complaint_cate7-menuid27
- Phanwilai, R. (2019). Automatic detection of diseased trees using aerial images from drones and deep learning techniques (Master's thesis). Dhurakij Pundit University.
- Pimkhot, P. (2021). *Alcohol product logo detection using deep learning models* (Master's thesis). Dhurakij Pundit University.
- Rezatofghi, H., Tsoi, N., Gwak, J., Sadeghian, A., Reid, I., & Savarese, S. (2019). Generalized intersection over union :A metric and a loss for bounding box regression .Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 658–666.
- Somphong Mala. (2022). Ban Klang Subdistrict Administrative Organization. Retrieved 15 December 2022. From [https://www.banklang.go.th/webboard_detail.php?doIP=1&checkIP=%20\\$chkIP&hd=1&id=1035&checkAdd=\\$chkAd&dum=379156_ypk](https://www.banklang.go.th/webboard_detail.php?doIP=1&checkIP=%20$chkIP&hd=1&id=1035&checkAdd=$chkAd&dum=379156_ypk).