การตรวจจับสภาพพื้นผิวถนนที่ผิดปกติโดยใช้โครงข่ายประสาทเทียม แบบคอนโวลูชันเชิงลึกด้วยเทคนิคการถ่ายโอนการเรียนรู้
The Road Surface Anomalies Detection using Deep Convolutional Neural Networks with Transfer Learning Technique

ปิติภูมิ โพสาวัง (Pitiphum Posawang)* สถิตย์โชค โพธิ์สอาด (Satidchoke Phosaard)*
ศุภกฤษฎ์ นิวัฒนากูล (Suphakit Niwattanakul)* และ วสันต์ ภัทรอธิคม (Wasan Pattara-atikom)**

Received : August 16, 2019
Revised : October 22, 2019
Accepted : October 25, 2019

บทคัดย่อ

การตรวจจับสภาพพื้นผิวถนนที่ผิดปกติมีความสำคัญต่องานบำรุงรักษาถนนและการขนส่งทางถนนเพื่อประกันความปลอดภัยในการขับขี่ของผู้ใช้ถนน งานวิจัยนี้ศึกษาวิธีการตรวจจับสภาพพื้นผิวถนนที่ผิดปกติโดยใช้โครงข่ายประสาทเทียมแบบคอนโวลูชันเชิงลึกด้วยเทคนิคการถ่ายโอนการเรียนรู้ โดยมีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของโครงข่ายประสาทเทียมแบบคอนโวลูชันเชิงลึกที่ผ่านการฝึกมาแล้วล่วงหน้าจำนวน 5 โครงสร้าง ได้แก่ 1) Faster R-CNN_NAS 2) Faster R-CNN Inception_ResNetV2 Atrous 3) SSD ResNet50 FPN 4) SSD MobileNetV1 FPN และ 5) Faster R-CNN ResNet101 ชุดข้อมูลรูปภาพของสภาพพื้นผิวถนนที่ผิดปกติตามสภาพแวดล้อมทางธรรมชาติที่นำมาใช้ในการเรียนรู้และทดสอบมีจำนวนทั้งสิ้น 6 ประเภท ได้แก่ 1) หลุมบ่อ 2) รอยแตกตามแนวยาว 3) รอยแตกตามแนวขวาง 4) รอยแตกหนังจระเข้ 5) รอยปะซ่อม และ 6) เนินชะลอความเร็ว ผลการทดลอง พบว่า ตัวแบบ SSD ResNet50 FPN สามารถตรวจจับสภาพพื้นผิวถนนที่ผิดปกติได้ค่าเฉลี่ยของค่าเฉลี่ยความแม่นยำสูงที่สุด คือ 87.38% และใช้เวลาในการทดสอบน้อยที่สุด วิธีการและตัวแบบที่เสนอสามารถประยุกต์ใช้ในการตรวจจับสภาพพื้นผิวถนนที่ผิดปกติในพื้นที่ต่าง ๆ เพื่อใช้เป็นข้อมูลสนับสนุน ปรับปรุงคุณภาพประสิทธิภาพของถนนและเพิ่มความปลอดภัยในการขับขี่

คำสำคัญ: การตรวจจับสภาพพื้นผิวถนน ตัวแบบที่ผ่านการฝึกมาแล้วล่วงหน้า โครงข่ายประสาทเทียมแบบคอนโวลูชันเชิงลึก การถ่ายโอนการเรียนรู้

Abstract

The road surface anomaly detection is an important task in road maintenance and road transportation to safety assurance for all road users. This research study the road surface anomalies detection using deep convolutional neural networks with transfer learning technique. The primary objective was to compare the performance of 5 deep convolutional neural networks architectures that are the pre-trained models, such as 1) Faster R-CNN_NAS, 2) Faster R-CNN Inception_ ResNetV2 Atrous, 3) SSD ResNet50 FPN, 4) SSD Mobile-NetV1 FPN, and 5) Faster R-CNN ResNet101. The images dataset of road surface anomalies in the natural environment for training and testing have 6 categories, such as 1) pothole 2) longitudinal cracks, 3) transverse cracks, 4) alligator cracks, 5) patch, and 6) speed bumps. The results show that the SSD ResNet50 FPN model provides the highest mean average precision of 87.38% and least test times. The study provides practical methodology and models that could be applied to automatically detect road surface anomalies in other regions to support road informatics, improve quality, efficiency and driving safety.

Keyword: Road surface detection, Pre-trained model, Deep convolutional neural networks, Transfer learning.

* สาขาวิชาเทคโนโลยีสารสนเทศ สำนักวิชาเทคโนโลยีสังคม มหาวิทยาลัยเทคโนโลยีสุรนารี
* School of Information Technology, Institute of Social Technology, Suranaree University of Technology.
**ห้องปฏิบัติการวิจัยระบบขนส่งและจราจรอัจฉริยะ ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ
** Intelligent Transportation System Laboratory, National Electronics and Computer Technology Center.

1. บทนำ

การตรวจจับสภาพพื้นผิวถนนที่ผิดปกติมีความสำคัญต่องานบำรุงรักษาถนนเพื่อประกันความปลอดภัยในการขับขี่ เพิ่มประสิทธิภาพการขนส่งทางถนน ก่อให้เกิดประโยชน์ทางเศรษฐกิจและสังคมของประเทศ [1-2] ปัจจุบันประเทศไทยมีการก่อสร้างโครงข่ายถนนที่เชื่อมโยงไปทั่วทุกภูมิภาค รวมระยะทางกว่า 116,624 กิโลเมตร [3-4] อย่างไรก็ดี เมื่อถนนเปิดใช้งานไปในระยะเวลาหนึ่ง สภาพพื้นผิวถนนมักเกิดความชำรุดเสียหาย พื้นผิวถนนเสื่อมและเป็นหลุมเป็นบ่อ ส่งผลให้หน่วยงานที่รับผิดชอบต้องใช้งบประมาณในการบำรุงรักษาถนนต่อปีเป็นจำนวนมากกว่า 40,786 ล้านบาท [3-4] เพื่อใช้สำรวจ ประเมินและวางแผนบำรุงรักษาตามสภาพความเสียหายของพื้นผิวถนน

สภาพความเสียหายของพื้นผิวถนนถือว่าเป็นข้อมูลหลักในการพิจารณาวางแผน ตัดสินใจ หาวิธีซ่อมบำรุงรักษาให้ถูกต้องตามกำหนดเวลาบำรุงรักษา [5-6] ดังนั้น เพื่อป้องกันมิให้ความเสียหายลุกลามแผ่ขยายออกไปและบำรุงรักษาถนนให้อยู่ในสภาพพร้อมใช้งานตลอดเวลาจึงจำเป็นต้องสำรวจติดตามและตรวจสอบ เพื่อเร่งดำเนินการซ่อมแซมและแจ้งเตือนผู้ใช้รถใช้ถนนทันทีเมื่อตรวจพบสภาพพื้นผิวถนนที่ผิดปกติ

การตรวจจับสภาพพื้นผิวถนนที่ผิดปกติบนโทรศัพท์สมาร์ทโฟนเป็นอีกวิธีหนึ่งที่มีต้นทุนต่ำในการพัฒนา สามารถเก็บรวบรวมข้อมูลได้เป็นจำนวนมากและในบริเวณกว้าง โดยไม่จำเป็นต้องติดตั้งอุปกรณ์พิเศษเพิ่มเติมที่มีราคาแพง ดังเช่น เครื่องสแกนวัตถุสามมิติด้วยแสงเลเซอร์ (3D laser scanner) ซึ่งเป็นวิธีที่ต้องใช้การคำนวณและประมวลผลสูง ไม่เหมาะกับการนำมาใช้ในสภาพแวดล้อมแบบเวลาจริง [7]

ที่ผ่านมาการตรวจจับสภาพพื้นผิวถนนที่ผิดปกติบนโทรศัพท์สมาร์ทโฟนใช้เครื่องรับรู้มาตรความเร่ง (Accelerometer) เพื่อวัดแรงสั่นสะเทือนและใช้กล้องบันทึกภาพวิดีโอเพื่อรวบรวมข้อมูลรูปภาพ แต่อย่างไรก็ตาม การใช้เครื่องรับรู้มาตรความเร่งอาจให้ผลลัพธ์ผิดพลาด เช่น รอยต่อบนถนนที่ถูกตรวจจับเป็นหลุมบ่อ และหลุมบ่อกลางช่องจราจรไม่สามารถตรวจจับได้หากใช้เครื่องรับรู้มาตรความเร่งเพียงอย่างเดียว [7-9] การใช้กล้องบันทึกภาพวิดีโอสามารถตรวจจับสภาพพื้นผิวถนนได้ในบริเวณกว้างและมีความถูกต้องสูง [7-8] ดังนั้น การใช้กล้องบันทึกภาพวิดีโอจึงเป็นวิธีที่ดีกว่าการใช้เครื่องรับรู้มาตรความเร่งสำหรับการตรวจจับสภาพพื้นผิวถนนที่ผิดปกติบนโทรศัพท์สมาร์ทโฟน

อย่างไรก็ดี ข้อมูลรูปภาพจากกล้องบันทึกภาพวิดีโอบนโทรศัพท์สมาร์ทโฟนถูกนำมาใช้เพื่อตรวจจับสภาพพื้นผิวถนนที่ผิดปกติโดยวิธีการสกัดคุณลักษณะด้วยมือ (Hand-crafted feature extraction) ด้วยขั้นตอนวิธีการประมวลผลภาพ (Image processing) [10-12] ความถูกต้องในการตรวจจับสภาพพื้นผิวถนนที่ผิดปกติขึ้นอยู่กับคุณภาพของการสกัดคุณลักษณะพิเศษของข้อมูล (Features extraction)

ปัจจุบันการใช้โครงข่ายประสาทเทียมแบบคอนโวลูชันเชิงลึก (Deep convolutional neural networks) เป็นขั้นตอนวิธีที่สามารถเรียนรู้และคำนวณหาคุณลักษณะพิเศษของรูปภาพและจำแนกประเภทข้อมูลแบบอัตโนมัติโดยไม่จำเป็นต้องใช้วิธีการสกัดคุณลักษณะพิเศษด้วยมือ [13-19] แม้ว่าการใช้โครงข่ายประสาทเทียมแบบคอนโวลูชันเชิงลึกสามารถเรียนรู้และคำนวณหาคุณลักษณะพิเศษของรูปภาพที่มีประสิทธิภาพมากขึ้น แต่ในทางกลับกัน ส่งผลให้กระบวนการเรียนรู้ใช้เวลานานและต้องการชุดข้อมูลเป็นจำนวนมาก [14-17] ด้วยเหตุนี้ การถ่ายโอนการเรียนรู้ (Transfer learning) จึงถูกนำมาประยุกต์ใช้ผ่านตัวแบบที่ถูกฝึกมาล่วงหน้า (Pre-trained model) ซึ่งการถ่ายโอนการเรียนรู้นี้เป็นการปรับปรุงการเรียนรู้ในงานใหม่ผ่านการถ่ายทอดความรู้จากตัวแบบที่ผ่านการฝึกมาล่วงหน้า ส่งผลให้ประหยัดเวลาในการเรียนรู้และไม่จำเป็นต้องใช้ชุดข้อมูลฝึกฝนเป็นจำนวนมากยิ่งไปกว่านั้นตัวแบบที่ถูกฝึกมาล่วงหน้านี้ยังมีประสิทธิภาพที่ดีกว่าตัวแบบทั่วไปที่ถูกสร้างขึ้นมาเองตั้งแต่เริ่มต้น [20]

งานวิจัยนี้จึงศึกษาวิธีการตรวจจับสภาพพื้นผิวถนนที่ผิดปกติโดยใช้โครงข่ายประสาทเทียมแบบคอนโวลูชันเชิงลึกด้วยเทคนิคการถ่ายโอนการเรียนรู้ เพื่อเปรียบเทียบประสิทธิภาพของตัวแบบที่ผ่านการฝึกมาแล้วล่วงหน้าซึ่งมีโครงสร้างที่แตกต่างกัน โดยใช้ตัวแบบมีค่าเฉลี่ยของค่าเฉลี่ยความแม่นยำ (Mean average precision: mAP) บนชุดข้อมูลรูปภาพของโคโค (COCO dataset) สูงสุดจำนวน 5 ลำดับแรก ได้แก่ 1) Faster R-CNN_NAS 2) Faster R-CNN Inception_ResNetV2 Atrous 3) SSD ResNet50 FPN 4) SSD MobileNetV1 FPN และ 5) Faster R-CNN ResNet101 [21] นอกจากนี้ได้เสนอชุดข้อมูลรูปภาพสภาพพื้นผิวถนนลาดยางแอสฟัลท์ที่ผิดปกติ (Road surface anomalies dataset: RSAD) จำนวน 6 ประเภท ตามสภาพแวดล้อมทางธรรมชาติ ได้แก่

1) หลุมบ่อ 2) รอยแตกตามแนวยาว 3) รอยแตกตามแนวขวาง 4) รอยแตกหนังจระเข้ 5) รอยปะซ่อม และ 6) เนินชะลอความเร็วซึ่งลักษณะของสภาพพื้นผิวถนนที่ผิดปกติทั้ง 6 ประเภทนี้มาจากคู่มือตรวจสอบและประเมินสภาพความเสียหายของผิวทางในประเทศไทยโดยสำนักวิเคราะห์และตรวจสอบ สำนักบริหารบำรุงทาง กรมทางหลวง [5] จากนั้นจึงนำตัวแบบที่เหมาะสมมาตรวจจับสภาพพื้นผิวถนนที่ผิดปกติต่อไป

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

การตรวจจับวัตถุโดยทั่วไปใช้ตัวตรวจจับวัตถุ (Object detector) ด้วยวิธีการเปลี่ยนแปลงขนาดและตำแหน่ง เพื่อเสนอขอบเขตพื้นที่ (Region proposals) ที่น่าจะมีวัตถุอยู่ภายในภาพ จากนั้นจึงจำแนกประเภท (Classifier) ของวัตถุนั้น ๆ [15] หลายปีที่ผ่านมาประสิทธิภาพของตัวแบบที่ผ่านการเรียนรู้มาแล้วล่วงหน้าบนชุดข้อมูลมาตรฐานขนาดใหญ่ เช่น ชุดข้อมูลของโคโค (COCO dataset) พบว่า มีความเร็วและความแม่นยำสูงเหมาะสำหรับนำมาใช้ในงานตรวจจับวัตถุอื่น ๆ [21-22] โครงสร้างของตัวแบบดังกล่าว ประกอบด้วย 2 ส่วน ได้แก่ ตัวตรวจจับวัตถุ (Object detector) และตัวสกัดคุณลักษณะ (Feature extractor) สำหรับงานวิจัยนี้ใช้โครงสร้างของตัวตรวจจับวัตถุ 2 รูปแบบและโครงสร้างของตัวสกัดคุณลักษณะ 5 รูปแบบ ดังนี้

2.1 โครงสร้างของตัวตรวจจับวัตถุ

2.1.1 โครงสร้างแบบ Faster R-CNN

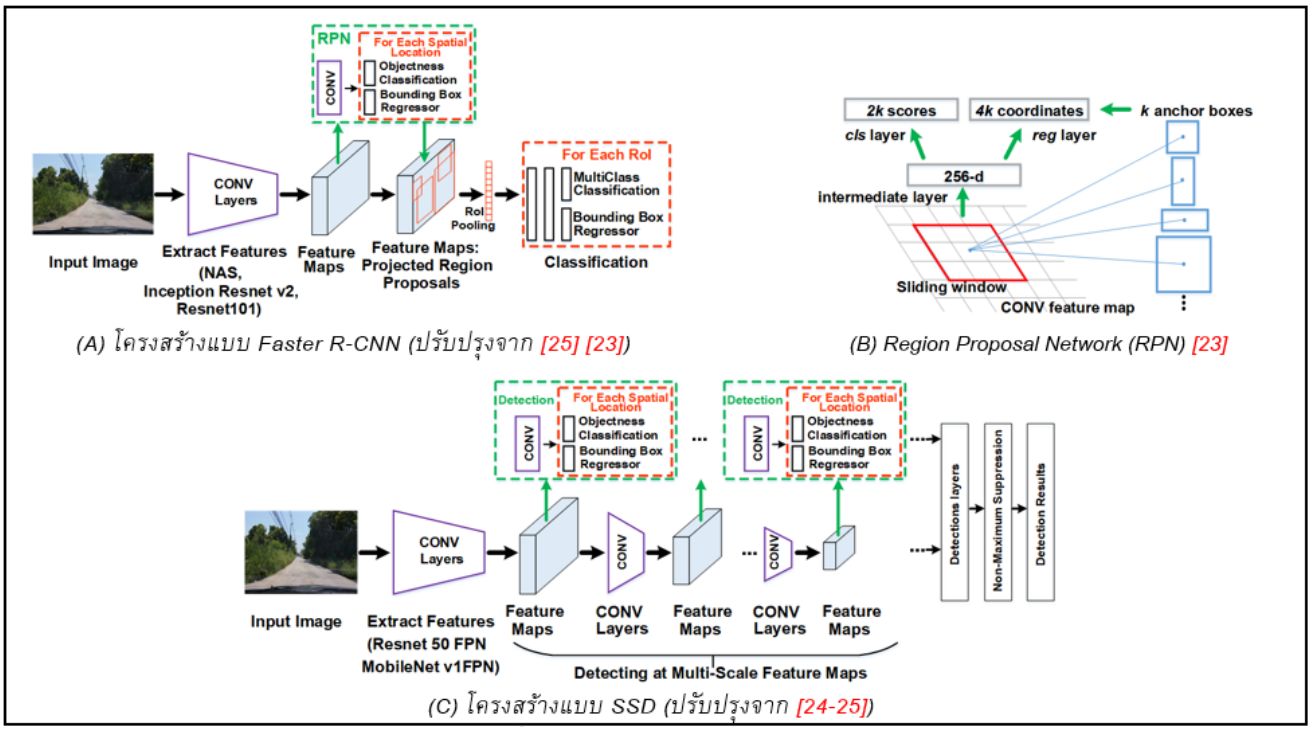
Ren และคณะ [23] เสนอโครงสร้างแบบ Faster R-CNN ในปี ค.ศ. 2017 ประกอบด้วย 2 ขั้นตอน ดังนี้

1) ขั้นตอนแรก เป็นการเสนอขอบเขตพื้นที่ที่น่าจะมีวัตถุภายในภาพโดยใช้โครงข่ายเสนอพื้นที่เรียกว่า Region Proposal Network (RPN) จากภาพที่ 1 (A) เมื่อรูปภาพผ่านชั้นคอนโวลูชัน (CONV layers) ในชั้นนี้จะทำการสกัดคุณลักษณะที่สำคัญในรูปภาพ โดยสามารถเลือกใช้โครงสร้างของตัวสกัดคุณลักษณะรูปแบบต่าง ๆ ที่มีโครงสร้างของการถ่ายโอนการเรียนรู้อยู่ภายในได้ เช่น NAS, Inception Resnet v2, และ Resnet101 ผลลัพธ์ที่ผ่านการทำคอนโวลูชันในตำแหน่งต่าง ๆ ของพื้นที่ส่วนย่อยในรูปภาพ เรียกว่า แผนภาพคุณลักษณะ (Feature map) จากนั้นจึงส่งต่อไปยัง RPN เพื่อหาขอบเขตพื้นที่ที่น่าจะมีวัตถุในภาพ ซึ่งต้องผ่านชั้นคอนโวลูชันเพิ่มเติม เพื่อคำนวณโอกาสค้นพบวัตถุในแต่ละพื้นที่ย่อยเล็ก ๆ รวมถึงเสนอกรอบล้อมวัตถุมาให้ด้วย (Bounding box) สำหรับขอบเขตพื้นที่ของ RPN ที่นำมาพิจารณา คือ กรอบสี่เหลี่ยมหลากหลายขนาด เรียกว่า แองเคิล (Anchor box) ซึ่งครอบคลุม Feature map ในตำแหน่งต่าง ๆ ดังภาพที่ 1 (B) จากจำนวน Anchor box ทั้งหมดของ RPN จะทำการหาขอบเขตที่น่าจะเป็นวัตถุมากที่สุดมาจำนวนหนึ่ง เพื่อเสนอขอบเขตพื้นที่ที่น่าสนใจสุดท้ายหรืออาร์โอไอ (Region of interest: RoI) จากนั้นจึงส่งไปขั้นตอนที่สอง

2) ขั้นตอนที่สอง ประกอบด้วยขั้นตอนอาร์โอไอพูลลิง (RoI pooling) และการจำแนก (Classification) ดังภาพที่ 1 (A) ขั้นตอนการทำ RoI pooling เป็นการนำขอบเขตพื้นที่ที่น่าสนใจ (RoI) จาก Feature map ที่มีขนาดต่าง ๆ มาลดขนาดลง (Down sample) ให้ได้ค่าคุณลักษณะที่มีขนาดคงที่ (Fixed size feature vector) โดยใช้ฟังก์ชันหาค่าสูงสุดในการคำนวณที่เรียกว่า Max pooling แล้วส่งต่อไปยังโครงข่ายชั้นเชื่อมโยงแบบทั่วถึงถัดไป (Fully connected layer) ส่วนสุดท้ายเป็นส่วนการจำแนกประเภท (Multiclass classification) เพื่อทำการจำแนกว่าวัตถุนั้นจัดอยู่ในประเภทอะไร และส่วนการทำนายขอบเขตของวัตถุจากพื้นที่ที่ถูกเสนอมา (Bounding box regressor) เพื่อคำนวณหาตำแหน่งของวัตถุให้แม่นยำมากยิ่งขึ้น

2.1.2 โครงสร้างแบบ SSD

Liu และคณะ [24] เสนอโครงสร้างแบบ SSD ในปี ค.ศ. 2016 เพื่อตรวจจับวัตถุให้เร็วขึ้นโดยไม่สูญเสียความแม่นยำในการตรวจจับวัตถุมากเกินไป จากภาพที่ 1 (C) เมื่อรูปภาพผ่านชั้นคอนโวลูชันซึ่งเป็นการสกัดคุณลักษณะที่สำคัญในรูปภาพ โดยสามารถเลือกใช้โครงสร้างของตัวสกัดคุณลักษณะรูปแบบต่าง ๆ ได้ เช่น Resnet 50 FPN และ MobileNet v1 FPN เป็นต้น จากนั้นจึงส่งผลลัพธ์ไปยังชั้นคอนโวลูชันที่มีขนาดต่าง ๆ (Multiple CONV layers) ในแต่ละชั้นคอนโวลูชันจะทำการสกัดคุณลักษณะที่สำคัญในรูปภาพพร้อมทั้งจำแนกประเภทของวัตถุนั้นและทำนายขอบเขตของวัตถุมาให้ด้วย ซึ่งผลลัพธ์ที่ได้จากชั้นคอนโวลูชันนี้ คือ Feature map ที่มีขนาดต่าง ๆ (Multi scale feature maps) วิธีนี้ช่วยเพิ่มความแม่นยำในการตรวจจับวัตถุอย่างมีนัยสำคัญและเหมาะกับภาพที่มีวัตถุ



ภาพที่ 1 *โครงสร้างตัวตรวจจับวัตถุ*

ขนาด (Scale) ที่แตกต่างกันมากได้

2.2 โครงสร้างของตัวสกัดคุณลักษณะ

2.2.1 โครงสร้างแบบ NASNet

Zoph และคณะ [26] ทีมพัฒนาของ Google เสนอโครงสร้างแบบ NASNet (Neural Architecture Search Network) ที่เรียกว่า NASNet-A เป็นโครงสร้างที่ใช้ค้นหาโครงข่ายประสาทเทียมแบบคอนโวลูชันเชิงลึกที่ดีที่สุดเพื่อสกัดคุณลักษณะสำคัญของรูปภาพ การนำตัวแบบที่ผ่านการฝึกมาล่วงหน้าที่มีโครงสร้างแบบ Faster R-CNN NAS มาใช้ในงานตรวจจับวัตถุบนชุดข้อมูลของโคโค พบว่าได้ค่า mAP สูงถึง 43% [21]

2.2.2 โครงสร้างแบบ Inception Resnet v2

Szegedy และคณะ [27] ทีมพัฒนาของ Microsoft Research เสนอโครงสร้างแบบ Inception Resnet v2 ที่รวมโครงสร้างการเชื่อมต่อแบบ Residual network (Resnet) ของ He และคณะ [28] และส่วนของ Inception V4 units [27] เข้าไว้ด้วยกัน เพื่อเพิ่มประสิทธิภาพและลดเวลาการคำนวณ การนำตัวแบบที่ผ่านการฝึกมาล่วงหน้าที่มีโครงสร้างแบบ Faster R-CNN Inception ResNetV2 Atrous มาใช้ตรวจจับวัตถุบนชุดข้อมูลของโคโค พบว่า ได้ค่า mAP สูงถึง 37% [21]

2.2.3 โครงสร้างแบบ Resnet 50 FPN

Lin และคณะ [29] ทีมพัฒนาของ Facebook AI Research เสนอโครงสร้างแบบ Resnet 50 FPN ที่ประกอบด้วยโครงสร้างแบบ Residual Network (Resnet) จำนวน 50 ชั้น ที่ถูกพัฒนาโดย He และคณะ [28] และเพิ่มชั้น Feature Pyramid Network (FPN) [30] ซึ่งเป็นการลดขนาดและเพิ่มขนาดของรูปภาพสำหรับสกัดคุณลักษณะสำคัญของรูปภาพ เพื่อเพิ่มความแม่นยำในการรู้จำและการจำแนกข้อมูล การนำตัวแบบที่ผ่านการฝึกมาล่วงหน้าที่มีโครงสร้างแบบ SSD Resnet 50 FPN มาใช้ตรวจจับวัตถุบนชุดข้อมูลของโคโค พบว่า ได้ค่า mAP สูงถึง 35% [21]

2.2.4 โครงสร้างแบบ MobileNet v1 FPN

Howard และคณะ [31] ทีมพัฒนาของ Google เสนอโครงสร้างแบบ MobileNet v1 ที่ถูกออกแบบมาเพื่อลดการคำนวณและลดขนาดของตัวแบบให้เหมาะสมกับการใช้งานบนโทรศัพท์สมาร์ทโฟน และเพิ่มชั้น FPN [30] เพื่อลดขนาดและเพิ่มขนาดของรูปภาพสำหรับสกัดคุณลักษณะสำคัญของรูปภาพ การนำตัวแบบที่ผ่านการฝึกมาล่วงหน้าที่มีโครงสร้างแบบ MobileNet v1 FPN มาใช้ตรวจจับวัตถุบนชุดข้อมูลของโคโค พบว่า ได้ค่า mAP สูงถึง 32% [21]

2.2.5 โครงสร้างแบบ Resnet101

He และคณะ [28] ทีมพัฒนาของ Microsoft Research เสนอโครงสร้างแบบ Resnet101 ซึ่งมีจำนวน 101 ชั้น เพื่อสกัดคุณลักษณะสำคัญของรูปภาพ การนำตัวแบบที่ผ่านการฝึกมาล่วงหน้าที่มีโครงสร้างแบบ Resnet101 มาใช้ตรวจจับวัตถุบนชุดข้อมูลของโคโค พบว่า ได้ค่า mAP สูงถึง 32% [21]

2.3 งานวิจัยที่เกี่ยวข้อง

งานวิจัยที่ใช้โครงข่ายประสาทเทียมแบบคอนโวลูชันเชิงลึกที่ถูกฝึกมาล่วงหน้าด้วยเทคนิคการถ่ายโอนการเรียนรู้สำหรับตรวจจับสภาพพื้นผิวถนนที่ผิดปกติ มีดังนี้

Maeda และคณะ [32] พัฒนาชุดข้อมูลสำหรับตรวจจับความเสียหายบนถนนในญี่ปุ่นจำนวน 8 ประเภท ได้แก่ รอยแตกตามแนวยาวบริเวณล้อรถ รอยแตกตามแนวยาวบริเวณการก่อสร้าง รอยแตกตามแนวยาว รอยแตกหนังจระเข้ หลุมบ่อ เส้นช่องจราจรชำรุดและเส้นข้ามทางม้าลายชำรุด โดยใช้ตัวแบบ SSD Inception_V2 และ SSD MobileNet (ค่าไอโอยู > 0.5) ผลการทดลองพบว่า ตัวแบบทั้งสองสามารถตรวจจับความเสียหายบนถนนได้ค่าความแม่นยำและค่าความระลึกมากกว่า 75% ยิ่งไปกว่านั้น Singh และคณะ [33] ใช้ชุดข้อมูลของ Maeda และคณะ [32] เพื่อตรวจจับความเสียหายบนถนนด้วยตัวแบบ Mask region-based convolutional network (Mask R-CNN) (ค่าไอโอยู > 0.5) ผลการทดลองพบว่า ตัวแบบ Mask R-CNN ตรวจจับความเสียหายบนถนนได้ค่าตัววัดค่าเอฟ (F1-score) เท่ากับ 52.8% นอกจากนี้ Koh และคณะ [34] ตรวจจับสภาพพื้นผิวถนนจากชุดข้อมูลหลุมบ่อในแอฟริกาใต้โดยใช้ตัวแบบ Region-based Fully Convolutional Networks (R-FCN) ผลการทดลองพบว่า ตัวแบบ R-FCN สามารถตรวจจับหลุมบ่อได้ค่า mAP เท่ากับ 89%

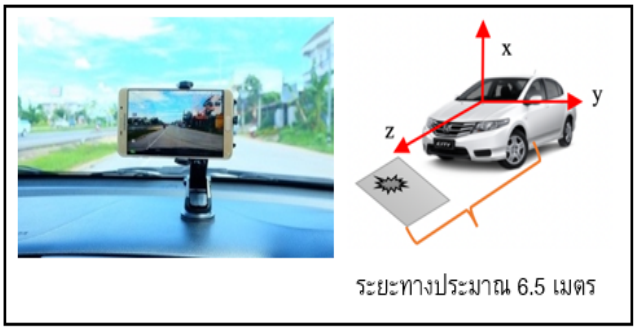
จากงานวิจัยที่ผ่านมาข้างต้น พบว่า การตรวจจับสภาพพื้นผิวถนนที่ผิดปกติใช้วิธีการถ่ายโอนการเรียนรู้จากตัวแบบ SSD Inception_V2, SSD MobileNet, Mask R-CNN และ R-FCN ด้วยไลบรารีสำหรับตรวจจับวัตถุของเทนเซอร์โฟล (TensorFlow object detection API) สำหรับงานวิจัยนี้ใช้วิธีการถ่ายโอนการเรียนรู้จากตัวแบบที่ให้ค่า mAP สูงสุด 5 ลำดับแรกบนชุดข้อมูลรูปภาพทดสอบของโคโค [21] เพื่อเปรียบเทียบประสิทธิภาพแต่ละตัวแบบและนำตัวแบบที่เหมาะสมมาใช้ในการตรวจจับสภาพพื้นผิวถนนที่ผิดปกติตามสภาพแวดล้อมทางธรรมชาติในประเทศไทยต่อไป

3. วิธีดำเนินการวิจัย

3.1 การเก็บรวบรวมข้อมูล

ข้อมูลรูปภาพของสภาพพื้นผิวถนนถนนลาดยางแอสฟัลท์ถูกเก็บรวบรวมในจังหวัดนครราชสีมา จากกล้องบันทึกวิดีโอในโทรศัพท์รุ่น Samsung Galaxy A9 Pro ปี 2016 ที่ถูกติดตั้งบนคอนโซลหน้ารถยนต์รุ่น Honda City 2010 ระยะทางจากจุดตั้งโทรศัพท์ไปจนถึงจุดที่สามารถเห็นภาพได้ชัดเจน คือ 6.5 เมตร ดังภาพที่ 2

โปรแกรมที่ใช้บันทึกภาพวิดีโอถูกพัฒนาขึ้นบนระบบปฏิบัติการแอนดรอยด์ ความละเอียดภาพขนาด 1920×1080 พิกเซล และมีอัตราจำนวนเฟรม 30 เฟรม/วินาที ขับรถเก็บรวบรวมข้อมูลด้วยความเร็ว 40 - 60 กิโลเมตร/ชั่วโมง



***ภาพที่ 2** การติดตั้งโทรศัพท์สมาร์ทโฟนและระยะทางที่ใช้ในการตรวจจับสภาพพื้นผิวถนน*

รวมระยะทางประมาณ 350 กิโลเมตร

ข้อมูลที่ถูกจัดเก็บ ได้แก่ วัน เวลา ละติจูด ลองติจูด ภาพวิดีโอสภาพพื้นผิวถนนทั้ง 6 ประเภท ประกอบด้วย 1) หลุมบ่อ (Pothole: PO) 2) รอยแตกตามแนวยาว (Longitudinal cracks: LC), 3) รอยแตกตามแนวขวาง (Transverse cracks: TC), 4) รอยแตกหนังจระเข้ (Alligator cracks: AC), 5) รอยปะซ่อม (Patch: PA) และ 6) เนินชะลอความเร็ว (Speed bumps: SB) ตามสภาพแวดล้อมทางธรรมชาติที่แตกต่างกัน เช่น ขนาด รูปทรง แสงเงาของต้นไม้ รอยคราบสกปรก คราบน้ำมันและเครื่อง

ภาพที่ 3 *ตัวอย่างชุดข้อมูล RSAD ที่ถูกสร้างกรอบล้อมวัตถุและระบุประเภทสภาพพื้นผิวถนนที่ผิดปกติ*

หมายจราจรบนผิวถนน ดังภาพที่ 3

3.2 การเตรียมชุดข้อมูล

ข้อมูลรูปภาพถูกดึงมาจากไฟล์วิดีโอที่บันทึกสภาพพื้นผิวถนนและลดขนาดรูปภาพลงมาที่ 960×540 พิกเซล จากนั้นจึงใช้โปรแกรม LabelImg [35] เพื่อสร้างกรอบล้อมวัตถุ (Bounding boxes) พร้อมทั้งระบุประเภทสภาพพื้นผิวถนนโดยผู้ใช้ถนนและผู้เชี่ยวชาญทางถนน

ชุดข้อมูล RSAD มีทั้งหมด 6,584 ภาพประกอบด้วยกรอบล้อมวัตถุพร้อมทั้งระบุประเภทสภาพพื้นผิวถนนซึ่งมีจำนวนทั้งสิ้น 10,240 วัตถุ จากนั้นสุ่มแบ่งข้อมูลรูปภาพจำนวน 70% (7,168) ใช้เป็นชุดข้อมูลเรียนรู้ (Training dataset) และ 30% (3,072) ใช้เป็นชุดข้อมูลทดสอบ (Testing dataset) ดังตารางที่ 1

ตารางที่ 1 *ชุดข้อมูล RSAD ในแต่ละประเภท*

สภาพพื้นผิวถนน	จำนวนของวัตถุ (Bounding boxes)		
	ฝึกฝน	ทดสอบ	รวม
หลุมบ่อ	1,294	565	1,859
รอยแตกตามแนวยาว	1,399	595	1,994
รอยแตกตามแนวขวาง	1,100	443	1,543
รอยแตกหนังจระเข้	1,244	565	1,809
รอยปะซ่อม	1,292	539	1,831
เนินชะลอความเร็ว	839	365	1,204
รวม	**7,168**	**3,072**	**10,240**

3.3 ตัวแบบที่ใช้ตรวจจับสภาพพื้นผิวถนน

ตัวแบบที่ใช้ตรวจจับสภาพพื้นผิวถนนที่ผิดปกติ คือ ตัวแบบที่ผ่านการฝึกมาแล้วล่วงหน้า (Pre-trained model) ที่ถูกพัฒนามาจากเทนเซอร์โฟล (TensorFlow) ซึ่งเป็นไลบรารีโอเพ่นซอร์สสำหรับงานด้านการเรียนรู้ด้วยเครื่อง ในงานวิจัยนี้ใช้ตัวแบบที่ให้ค่าเฉลี่ยของค่าเฉลี่ยความแม่นยำ (Mean average precision: mAP) บนชุดข้อมูลรูปภาพของโคโค (COCO dataset) สูงสุดจำนวน 5 ลำดับแรก [21] ดังตารางที่ 2 จากนั้นจึงใช้เทคนิคการถ่ายโอนการเรียนรู้เพื่อปรับปรุงการเรียนรู้ใหม่สำหรับงานตรวจจับสภาพพื้นผิวถนนที่ผิดปกติ

การกำหนดค่าพารามิเตอร์สำหรับปรับแต่งการถ่ายโอนการเรียนรู้ใช้ตามรูปแบบข้อกำหนดใน [21] เช่น จำนวนประเภทสภาพพื้นผิวถนนเท่ากับ 6 ประเภท

ตารางที่ 2 *ตัวแบบที่ผ่านการฝึกมาล่วงหน้าที่มีค่า mAP สูงที่สุด 5 ลำดับแรกบนชุดข้อมูลทดสอบ*

ตัวแบบ	ชื่อตัวแบบ		ความเร็ว	mAP
	ตัวตรวจจับวัตถุ	ตัวสกัดคุณลักษณะ		
FRN	Faster R-CNN	NAS	1.833	43
FRIR	Faster R-CNN	Inception ResNetV2 Atrous	0.620	37
SRF	SSD	ResNet50 FPN	0.076	35
SMF	SSD	MobileNetV1 FPN	0.056	32
FRR	Faster R-CNN	ResNet101	0.106	32

ของโคโค [21]

(num_classes = 6) กำหนดจำนวนรอบในการฝึกฝนเท่ากับ 30,000 รอบ (num_step = 30,000) ซึ่งในระหว่างการฝึกฝนตัวแบบ ผู้วิจัยทำการสังเกตค่าสูญเสียการสอน (Training loss) และค่าสูญเสียการตรวจสอบ (Validation loss) ซึ่งจะค่อย ๆ ลดลงตามการฝึกฝนในแต่ละรอบ หากการสูญเสียการตรวจสอบนั้นสูงขึ้นหรือเริ่มคงที่ ผู้วิจัยจะทำการหยุดการสอนที่รอบการสอนนั้น กำหนดอัตราการเรียนรู้ (Learning rate) เท่ากับ 0.0001 และกำหนดพาธเก็บไฟล์ของตัวแบบที่ผ่านการฝึกมาล่วงหน้า เป็นต้น สำหรับการกำหนดค่าพารามิเตอร์อื่น ๆ ใช้ค่าตามไฟล์ต้นฉบับของแต่ละตัวแบบที่อ้างอิงใน [21] ซึ่งเป็นการกำหนดค่าเบื้องต้นที่เหมาะสมตามโครงสร้างของตัวตรวจจับวัตถุและโครงสร้างของตัวสกัดคุณลักษณะของแต่ละตัวแบบที่ผ่านการฝึกมาล่วงหน้าแล้ว

3.4 การประเมินประสิทธิภาพของตัวแบบ

งานวิจัยนี้ประเมินประสิทธิภาพของตัวแบบตรวจจับสภาพพื้นผิวถนนที่ผิดปกติด้วยค่าไอโอยู ค่าเฉลี่ยความแม่นยำและค่าเฉลี่ยของค่าเฉลี่ยความแม่นยำ [25] ดังนี้

3.4.1 ค่าไอโอยู (Intersection Over Union: IoU) เป็นค่าที่นิยมใช้วัดความแม่นยำในการตรวจจับวัตถุ ซึ่งวัดได้จากอัตราส่วนพื้นที่ทับซ้อนกัน (Area of overlap) ระหว่างพื้นที่บริเวณทำนาย (Prediction area) กับพื้นที่จริง (Ground truth area) หารด้วยพื้นที่รวมทั้งหมด (Area of union)

$$\text{IoU} = \frac{\text{Area of overlap}}{\text{Area of union}} \qquad (1)$$

ระหว่างพื้นที่บริเวณทำนายและพื้นที่จริง ดังสมการที่ 1

เมื่อวัตถุที่ถูกตรวจจับอยู่ในคลาสนั้นถูกต้องและมีค่าไอโอยูมากกว่าหรือเท่ากับ 0.5 ถือว่าตัวแบบสามารถตรวจจับวัตถุของสภาพพื้นผิวถนนที่ผิดปกติได้ถูกต้อง

3.4.2 ค่าเฉลี่ยความแม่นยำ (Average precision: AP) เป็นค่าที่นิยมใช้วัดประสิทธิภาพของตัวตรวจจับวัตถุซึ่งได้มาจากการคำนวณพื้นที่ใต้กราฟของค่าความแม่นยำและ

$$AP = \sum_{i=1}^{n-1} (r_{i+1} - r_i)\, p_{interp}(r_{i+1}) \qquad (2)$$

ค่าความระลึกในแต่ละระดับของแต่ละคลาส ดังสมการที่ 2

เมื่อ r คือ ค่าความระลึกในแต่ละระดับที่ค่าความแม่นยำมีการเปลี่ยนแปลงค่าครั้งแรก (First interpolate)

n คือ จำนวนรูปภาพทั้งหมด

pinterp คือ การกำหนดค่าใหม่ให้ค่าความแม่นยำซึ่งค่าสูงสุดของค่าความแม่นยำจะเท่ากับค่าสูงสุดของค่าความระลึกในแต่ละระดับ

โดยค่าความแม่นยำ (Precision) หมายถึง ประสิทธิภาพของตัวตรวจจับวัตถุโดยดูจากอัตราส่วนของจำนวนวัตถุที่ตรวจจับได้ถูกต้อง (TP) จากจำนวนวัตถุที่ถูกตรวจจับได้ทั้งหมด (TP+FP) ดังสมการที่ 3

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \qquad (3)$$

ค่าความระลึก (Recall) หมายถึง ประสิทธิภาพของตัวตรวจจับวัตถุโดยดูจากอัตราส่วนของจำนวนวัตถุที่ตรวจจับได้ถูกต้อง (TP) จากวัตถุที่ถูกต้องทั้งหมด (TP+FN)

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \qquad (4)$$

ดังสมการที่ 4

เมื่อ TP = วัตถุที่ตัวแบบทำนายว่าเป็นสภาพพื้นผิวถนนที่ผิดปกติในคลาส Ci และวัตถุนั้นเป็นสภาพพื้นผิวถนนที่ผิดปกติในคลาส Ci และมีค่าไอโอยู > 0.5

FP = วัตถุที่ตัวแบบทำนายว่าเป็นสภาพพื้นผิวถนนที่ผิดปกติในคลาส Ci แต่วัตถุนั้นไม่ใช่สภาพพื้นผิวถนนที่ผิดปกติในคลาส Ci หรือวัตถุนั้นมีค่าไอโอยู < 0.5 หรือวัตถุนั้นไม่มีพื้นที่ทับซ้อนกัน (No overlap) ระหว่างพื้นที่บริเวณทำนายกับพื้นที่จริง

FN = ตัวแบบไม่สามารถตรวจจับสภาพพื้นผิวถนนที่ผิดปกติในคลาส Ci นั้นได้

Ci = คลาสของสภาพพื้นผิวถนนที่ผิดปกติจำนวน 6 ประเภท เมื่อ $1 < i < 6$

3.4.3 ค่าเฉลี่ยของค่าเฉลี่ยความแม่นยำ (Mean average precision: mAP) เป็นการนำค่า AP ของทุกคลาส

$$mAP = \frac{\sum_{i=1}^{K} AP_i}{K} \qquad (5)$$

มาทำการหาค่าเฉลี่ย ดังสมการที่ 5

เมื่อ K = จำนวนคลาสทั้งหมด

APi = ค่าเฉลี่ยความแม่นยำของคลาสที่ i

4. ผลการทดลองและการอภิปราย

ตัวแบบสำหรับตรวจจับสภาพพื้นผิวถนนที่ผิดปกติถูกฝึกฝนและทดสอบบนกูเกิลคราวด์แพลตฟอร์ม (Google Cloud Platform) ด้วย GPU รุ่น NVIDIA Tesla T4 แรม 16GB ผลการประเมินประสิทธิภาพของตัวแบบ มีดังนี้

ตารางที่ 3 *ผลการประเมินความแม่นยำของตัวแบบสำหรับตรวจจับสภาพพื้นผิวถนนที่ผิดปกติ*

ตัวแบบ	mAP	ค่า AP เมื่อ IoU ≥ 0.5					
		PO	LC	TC	AC	PA	SB
FRN	81.83	81.01	76.60	82.95	80.01	70.77	**99.67**
FRIR	84.34	80.90	**79.35**	82.02	87.47	77.33	98.97
SRF	**87.38**	**84.39**	79.02	**92.77**	**87.55**	**80.94**	99.63
SMF	85.21	82.90	76.25	89.56	86.70	76.41	99.47
FRR	83.53	77.26	78.85	88.43	86.33	70.77	99.33

ตารางที่ 4 *ผลการประเมินประสิทธิภาพของตัวแบบสำหรับตรวจจับสภาพพื้นผิวถนนที่ผิดปกติ*

ตัวแบบ	ระยะเวลาฝึกฝน (ชั่วโมง)	ระยะเวลาทดสอบ (ชั่วโมง)	ระยะเวลาทดสอบต่อ 1 ภาพ (วินาที)
FRN	~ 41	~ 1.45	~ 2.0507
FRIR	~ 15	~ 0.27	~ 0.5273
SRF	~ 28	~ 0.03	~ 0.0585
SMF	~ 26	~ 0.02	~ 0.0390
FRR	~ 6	~ 0.07	~ 0.1367

จากตารางที่ 3 พบว่า ตัวแบบ SRF มีค่า mAP บนชุดข้อมูลทดสอบ RSAD (กำหนดค่าไอโอยู > 0.5) คือ 87.38% รองลงมา คือ ตัวแบบ SMF, FRIR, FRR และ FRN ซึ่งได้ค่า mAP เท่ากับ 85.21%, 84.34%, 83.53% และ 81.83% ตามลำดับ เมื่อเปรียบเทียบค่าเฉลี่ยความแม่นยำ (AP) ของแต่ละคลาส (กำหนดค่าไอโอยู > 0.5) พบว่า ตัวแบบ SRF ให้ค่าเฉลี่ยความแม่นยำสูงที่สุด จำนวน 4 คลาส ได้แก่ หลุมบ่อ (PO) รอยแตกตามแนวขวาง (TC) รอยแตกหนังจระเข้ (AC) และรอยปะซ่อม (PA)

จากตารางที่ 4 พบว่า ตัวแบบที่ใช้เวลาเรียนรู้มากที่สุด คือ FRN ซึ่งใช้เวลาประมาณ 41 ชั่วโมง รองลงมา คือ ตัวแบบ SRF, SMF, FRIR และ FRR ซึ่งใช้เวลาเรียนรู้ประมาณ 28, 26, 15 และ 6 ชั่วโมง ตามลำดับ ส่วนตัวแบบ SMF ใช้เวลาทดสอบน้อยที่สุด ซึ่งใช้เวลาประมาณ 2 นาที (1 ภาพใช้เวลาประมาณ 0.039 วินาที) รองลงมา คือ ตัวแบบ SRF, FRR, FRIR และ FRN ซึ่งใช้เวลาทดสอบประมาณ 3, 7, 27 นาที และ 1 ชั่วโมง 45 นาที ตามลำดับ

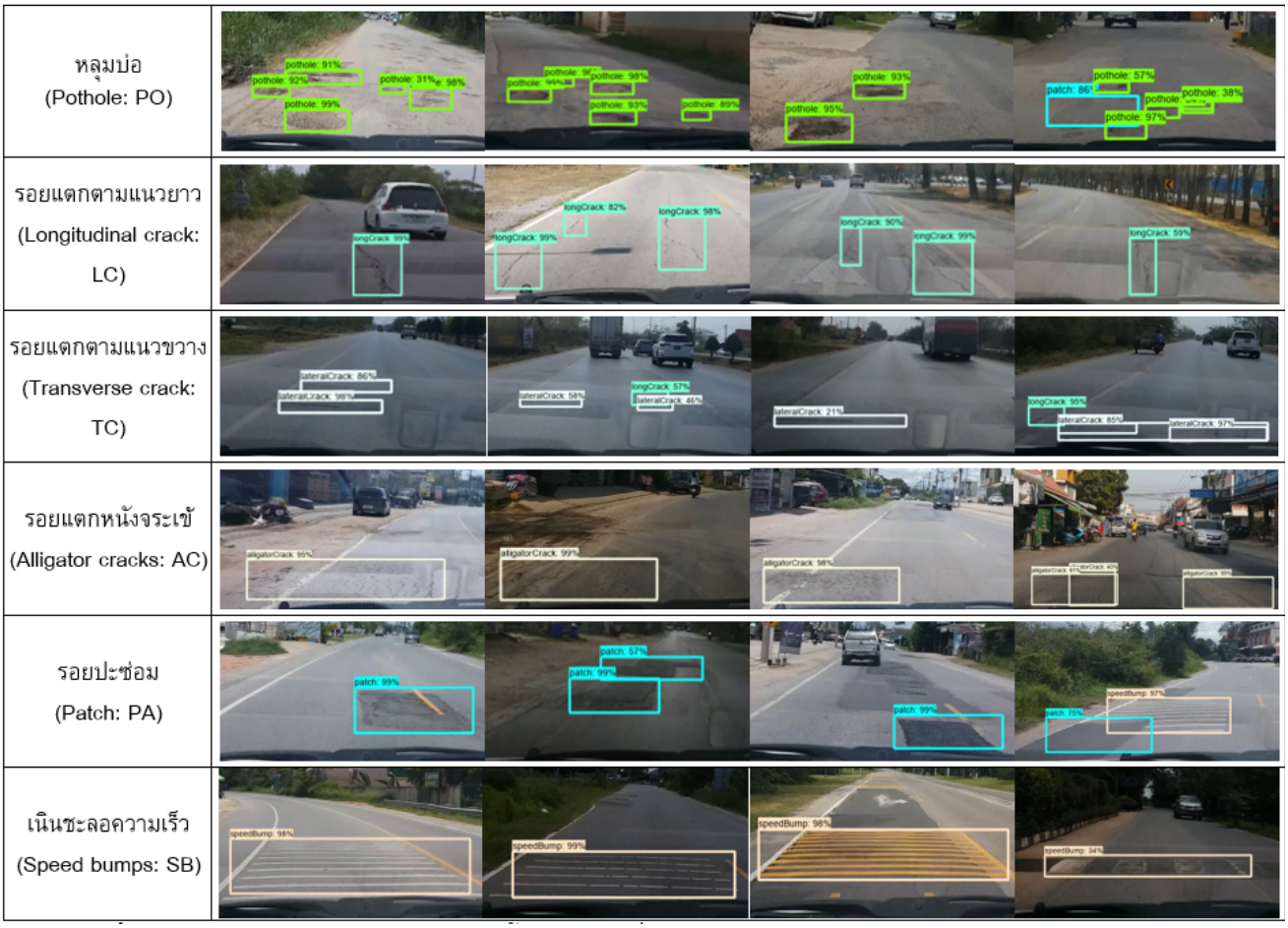
แม้ว่าตัวแบบ FRN และ FRIR มีค่า mAP มากกว่าตัวแบบ SRF บนชุดข้อมูลทดสอบของโคโค แต่อย่างไรก็ตาม เมื่อนำตัวแบบ FRN และ FRIR มาถ่ายโอนการเรียนรู้เพื่อปรับปรุงการเรียนรู้ใหม่สำหรับงานตรวจจับสภาพพื้นผิวถนนที่ผิดปกติ พบว่า ทั้งสองตัวแบบให้ mAP น้อยกว่าตัวแบบ SRF บนชุดข้อมูลทดสอบ RSAD ดังนั้น การนำโครงข่ายประสาทเทียมแบบคอนโวลูชันเชิงลึกที่ผ่านการฝึกมาแล้วล่วงหน้าที่มีโครงสร้างของตัวตรวจจับวัตถุและโครงสร้างของตัวสกัดคุณลักษณะที่แตกต่างกัน มาใช้เทคนิคการถ่ายโอนการเรียนรู้เพื่อปรับปรุงการเรียนรู้ใหม่สำหรับงานตรวจจับวัตถุอื่น ๆ ควรพิจารณาเปรียบเทียบผลการประเมินประสิทธิภาพของแต่ละตัวแบบ เพื่อเลือกใช้ตัวแบบที่เหมาะสมกับชุดข้อมูลสำหรับงานตรวจจับวัตถุนั้น

ตัวแบบที่เหมาะสมสำหรับการตรวจจับสภาพพื้นผิวถนนที่ผิดปกติ คือ ตัวแบบ SRF ซึ่งได้ mAP สูงที่สุด โดยมีโครงสร้างของตัวตรวจจับวัตถุแบบ SSD และมีโครงสร้างของการสกัดคุณลักษณะแบบ Resnet 50 FPN จุดเด่นของโครงสร้างแบบ SSD คือ ผลลัพธ์ที่ได้จากการทำคอนโวลูชันในแต่ละชั้นที่มีขนาดแตกต่างกัน ช่วยเพิ่มความแม่นยำในการตรวจจับวัตถุได้อย่างมีนัยสำคัญและเหมาะกับภาพที่มีวัตถุขนาดต่างกันมากได้ [24] สอดคล้องกับผลการทดลองนี้ คือ ตัวแบบ SRF และ SMF ที่มีโครงสร้างแบบ SSD ให้ค่า mAP สูงกว่าตัวแบบ FRIR, FRR และ FRN ที่มีโครงสร้างแบบ Faster R-CNN นอกจากนี้ผลการประเมินประสิทธิภาพด้านเวลาของตัวแบบสำหรับตรวจจับสภาพพื้นผิวถนนที่ผิดปกติบนชุดข้อมูลทดสอบ RSAD พบว่า ตัวแบบ SRF และ SMF ที่มีโครงสร้างแบบ SSD ใช้เวลาในการตรวจจับวัตถุต่อ 1 ภาพ น้อยกว่าตัวแบบ FRIR, FRR และ FRN ที่มีโครงสร้างแบบ Faster R-CNN ตัวอย่างผลการตรวจจับสภาพพื้นผิวถนนที่ผิดปกติ

จากรูปภาพทดสอบทั้ง 6 ประเภทตามสภาพแวดล้อมทางธรรมชาติที่แตกต่างกัน โดยใช้ตัวแบบ SRF (ค่าไอโอยู > 0.5) แสดงดังภาพที่ 4

5. สรุป

งานวิจัยนี้เสนอวิธีการประยุกต์ใช้โครงข่ายประสาทเทียมแบบคอนโวลูชันเชิงลึกที่ผ่านการฝึกมาแล้วล่วงหน้าด้วยเทคนิคการถ่ายโอนการเรียนรู้ในการทดลองได้เปรียบเทียบประสิทธิภาพของตัวแบบที่มีโครงสร้างที่แตกต่างกันและมีค่าเฉลี่ยของค่าเฉลี่ยความแม่นยำบนชุดข้อมูลรูปภาพของโคโคสูงสุดจำนวน 5 ลำดับแรก ผลการทดลอง พบว่าตัวแบบ SSD ResNet50_FPN (SRF) สามารถตรวจจับสภาพพื้นผิวถนนที่ผิดปกติบนชุดข้อมูลทดสอบได้ค่าเฉลี่ยของค่าเฉลี่ยความแม่นยำสูงที่สุด คือ 87.38% และใช้เวลาในการทดสอบน้อยที่สุด นอกจากนี้โครงสร้างของตัวตรวจจับวัตถุแบบ SSD ที่มีจุดเด่นในชั้นคอนโวลูชันที่มีขนาดแตกต่างกันยังช่วยเพิ่มความแม่นยำในการตรวจจับวัตถุและเหมาะกับภาพที่มีวัตถุขนาดต่างกันมากได้ ดังเช่นชุดข้อมูลสภาพพื้นผิวถนนที่ผิดปกติที่ลักษณะทางกายภาพของหลุมบ่อ รอยแตกตามแนวยาว รอยแตกตามแนวขวาง รอยแตกหนังจระเข้ รอยปะซ่อมและเนินชะลอความเร็ว ที่มีขนาดและรูปทรงที่แตกต่างกันหลายหลายรูปแบบ ยิ่งไปกว่านั้นงานวิจัยนี้ยังได้เสนอชุดข้อมูลสภาพพื้นผิวถนนที่ผิดปกติ (RSAD) จำนวน 6 ประเภทตามสภาพแวดล้อมทางธรรมชาติ ซึ่งสามารถใช้เป็นชุดข้อมูลพื้นฐานสำหรับเรียนรู้และพัฒนาตัวแบบให้มีประสิทธิภาพมากยิ่งขึ้นได้ งานวิจัยนี้สามารถนำไปประยุกต์ใช้ตรวจจับวัตถุอื่น ๆ โดยใช้โครงข่ายประสาทเทียมแบบคอนโวลูชันเชิงลึกที่ผ่านการฝึกมาแล้วล่วงหน้าด้วยเทคนิคการถ่ายโอนการเรียนรู้ที่เหมาะสมได้



ภาพที่ 4 *ตัวอย่างผลการตรวจจับสภาพพื้นผิวถนนที่ผิดปกติของตัวแบบ SRF บนชุดข้อมูลทดสอบ RSAD*

6. เอกสารอ้างอิง

[1] World Road Association, *The importance of road maintenance*. The World Road Association (PIARC). Rep. 2014R02EN, 2014.

[2] A. Hartmann, and F. Y. Y. Ling, "Value creation of road infrastructure networks: A structural equation approach." *Journal of traffic and transportation engineering.*, Vol.3, No.1, pp.28–36, 2016.

[3] Department of highways, *Annual report*, 2016.

[4] Department of rural roads, *Annual report*, 2016.

[5] Department of highways, Pavement *distress identi fication manual*. Bangkok: Bureau of Highway Maintenance Management, 2007.

[6] Department of rural roads, *Road Routine Manual*. Department of rural roads and Transportation Institute Chulalongkorn University, 2014.

[7] T. Kim, and S. K. Ryu, "Review and analysis of pothole detection methods," *Journal of Emerging Trends in Computing and Information Sciences.*, Vol.5, No.8, pp.603-608, 2014.

[8] T. B. Coenen, and A. Golroo, "A review on automated pavement distress detection methods," *Cogent Engineering.*, Vol.4, No.1, p.1374822, 2017.

[9] G. Chugh, D. Bansal, S. Sofat, "Road condition detection using smartphone sensors: A Survey." *International Journal of Electronic and Electrical Engineering*, Vol.7, No.6, pp.595-601, 2014.

[10] V. F. Enigo, T. V. Kumar, S. Vijay, and K. G. Prabu, "Crowd Sourcing based online petitioning system for pothole detection using android platform," *Procedia Computer Science,* Vol.87, pp.316-321, 2016.

[11] Y. Jo, and S. Ryu, "Pothole detection using android smartphone with a video camera," *Journal of emerging trends in computing and information sciences,* Vol.8, No.1, pp.25-29, 2017.

[12] A. Tedeschi, and F. Benedetto, "A real-time automatic pavement crack and pothole recognition system for mobile Android-based devices," *Advanced Engineering Informatics,* vol.32, pp.11-25, 2017.

[13] H. Maeda, Y. Sekimoto, and T. Seto, "Lightweight road manager: smartphone-based automatic determination of road damage status by deep neural

[14] K. Gopalakrishnan, S. K. Khaitan, A. Choudhary, and A. Agrawal, "Deep convolutional neural networks with transfer learning for computer vision-based data-driven pavement distress detection," *Constrution and Building Materials.*, vol.157, pp.322-330, 2017.

[15] H. Maeda, Y. Sekimoto, T. Seto, T. Kashiyama, and H. Omata, "Road damage detection using deep neural networks with images captured through a smartphone," arXiv:*1801.09454*, 2018.

[16] P. Posawang, S. Phosaard, S. Niwattanakul, and W. Pattara-atikom, "Classification of road surface anomalies using ensemble methods with deep convolutional neural networks," *16th ITS Asia-Pacific Forum FUKUOKA*, pp.824-836, 2018.

[17] J. Singh, and S. Shekhar, "Road damage detection and classification in smartphone captured images using Mask R-CNN," *arXiv:1811.04535*, 2018.

[18] V. Pereira, S. Tamura, S. Hayamizu, and H. Fukai, "A Deep Learning-Based Approach for Road Pothole Detection in Timor Leste," *2018 IEEE International Conference Service Operations and Logistics, and Informatics (SOLI)*, pp.279-284, 2018.

[19] Y. Bhatia, R. Rai, V. Gupta, N. Aggarwal, and A. Akula, "Convolutional neural networks-based potholes detection using thermal imaging," *Journal of King Saud University-Computer and Information Sciences*., 2019.

[20] L. Torrey, and J. Shavlik, "Transfer learning," *Handbook of research on machine learning applications and trends: algorithms, methods, and techniques*., pp.242-264, 2010.

[21] Pkulzc, *TensorFlow detection model zoo*, Available Online at https://github.com/tensorflow/models/blob/master/research/object_detection/g3doc/detection_model_zoo.md, accessed on 18 March 2019.

[22] J. Huang, et al., "Speed/accuracy trade-offs for modern convolutional object detectors," *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp.7310-7311, 2017.

[23] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE transactions on pattern analysis and machine intelligence*., Vol.39, No.6, pp.1137-1149, 2017.

[24] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C. Y. Fu, and A. C. Berg, "Ssd: Single shot multibox detector," *European conference on computer vision.*, pp.21-37, 2016.

[25] L. Liu, W. Ouyang, X. Wang, P. Fieguth, J. Chen, X. Liu, and M. Pietikäinen, "Deep learning for generic object detection: A survey," *arXiv:1809.02165,* 2018.

[26] B. Zoph, V. Vasudevan, J. Shlens, and Q. V. Le. "Learning transferable architectures for scalable image recognition," *Proceedings of the IEEE conference on computer vision and pattern recognition,* pp.8697-8710, 2018.

[27] C. Szegedy, S. Ioffe, and V. Vanhoucke, "Inception-v4, inception-resnet and the impact of residual connections on learning," *arXiv:1602.07261,* 2016.

[28] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *Proceedings of the IEEE conference on computer vision and pattern recognition,* pp.770-778, 2016.

[29] T. Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," *In Proceedings of the IEEE international conference on computer vision,* pp.2980-2988, 2017.

[30] T. Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection," *Proceedings of the IEEE conference on computer vision and pattern recognition,* pp. 2117-2125, 2017.

[31] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, ... and H. Adam, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," arXiv:1704.04861, 2017.

[32] H. Maeda, Y. Sekimoto, T. Seto, T. Kashiyama, and H. Omata, "Road damage detection using deep neural networks with images captured through a smartphone," arXiv:1801.09454, 2018.

[33] J. Singh, and S. Shekhar, "Road damage detection and classification in smartphone captured images using Mask R-CNN," arXiv:1811.04535, Nov. 2018.

[34] J. J. Koh, T. T. V. Yap, H. Ng, V. T. Goh, H. L. Tong, C. C. Ho, and T. Y. Kuek, "Autonomous Road Potholes Detection on Video," Computational Science and Technology, pp.137-143, 2019.

[35] Tzutalin, Labelimg, Available Online at https://github.com/tzutalin/labelImg, accessed on: 7 Apr 2019.