

การพัฒนาระบบเฝ้าระวังช้างป่าบุกรุกในพื้นที่เกษตรกรรมเพื่อลดความขัดแย้งระหว่างคนและ ช้างด้วยเทคโนโลยีโครงข่ายประสาทเทียมแบบคอนโวลูชัน

วสุพล เผือกนำผล^{1,*}, วิชาญ ทุมทอง², ปฎิคม ทองจริง³

^{1,2}สาขาวิชาเทคโนโลยีเครือข่ายคอมพิวเตอร์

³สาขาวิทยาการคอมพิวเตอร์

คณะวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ

มหาวิทยาลัยราชภัฏรำไพพรรณี

Received: 30 June 2022

Revised: 20 September 2023

Accepted: 24 November 2023

บทคัดย่อ

ความขัดแย้งระหว่างคนกับช้างเกิดขึ้นเนื่องจากการอพยพของช้างจากถิ่นที่อยู่อาศัย ไปยังพื้นที่เกษตรกรรมของมนุษย์เพื่อหาอาหารและแหล่งน้ำ งานวิจัยนี้จึงได้นำเสนอเทคโนโลยีโครงข่ายประสาทเทียมแบบคอนโวลูชันเพื่อลดความขัดแย้งระหว่างคนกับช้าง ด้วยกระบวนการทำงานที่เรียกว่า โยโลรุ่นที่ 5 เพื่อดำเนินการตรวจจับวัตถุแบบเรียลไทม์จากวิดีโอ โดยใช้อุปกรณ์สมองกลแบบฝังตัวเพื่อเฝ้าระวังช้างป่าบุกรุกในพื้นที่เกษตรกรรม ขั้นตอนแรกนักวิจัยได้รวบรวมชุดข้อมูลภาพช้างจากพื้นที่เฝ้าระวัง โดยใช้ประโยชน์จากเฟรมเวิร์คของปัญญาประดิษฐ์วิวัฒนาการเชิงลึก เพื่อพัฒนาแบบจำลองสถาปัตยกรรมเครือข่ายโวล์ และสร้างสถาปัตยกรรมที่เหมาะสม กับอุปกรณ์สมองกลฝังตัวที่มีทั้งความเร็วและความแม่นยำที่เหมาะสม งานนี้นักวิจัยได้ปรับค่าไฮเปอร์พารามิเตอร์สำหรับแบบจำลองโวล์ ได้แก่ แบบจำลองโวล์เอ็น โวล์เอส และโวล์เอ็ม ซึ่งเป็นแบบจำลองการคำนวณที่ลดความซับซ้อนลง และแบบจำลองที่ได้แนะนำมานี้มีความเหมาะสมกับอุปกรณ์สมองกลฝังตัว หลังจากทดสอบพบว่า แบบจำลองโวล์เอสมีค่าเฉลี่ยความแม่นยำอยู่ที่ 95.68% mAP@.5 โดยมีความเร็วกว่าแบบจำลองโวล์เอ็มถึง 50% ที่ความแม่นยำสูงสุด 95.46% mAP@.5 จากนั้น ก่อนนำแบบจำลองไปติดตั้งที่อุปกรณ์สมองกลฝังตัว นักวิจัยได้เพิ่มชุดข้อมูลรูปภาพช้างจากอินเทอร์เน็ตให้กับโวล์เอส เพื่อเพิ่มความแม่นยำจนได้ค่าใหม่คือ 98.02% mAP@.5 หลังจากติดตั้งระบบพบว่าการเรียนรู้เชิงลึกสามารถตรวจจับการบุกรุกของช้างป่าในพื้นที่เกษตรกรรมได้อย่างแม่นยำ แม้บางครั้งอาจช้ากว่าที่มนุษย์สังเกตเห็นอยู่เล็กน้อย

คำสำคัญ: ช้างป่าบุกรุก พื้นที่เกษตร ลดความขัดแย้ง โครงข่ายประสาทเทียมคอนโวลูชัน

* ผู้ประสานงานหลัก; อีเมล: vasupon.p@rbu.ac.th

The Development of Monitoring System for Elephant Intrusion Detection in Agricultural Areas to Reduce Human-elephant Conflict with Convolutional Neural network technology

Vasupon Phueaknamphol^{1,*}, Wichan Thumthong², Patikom Thongjing³

^{1,2}Department of Computer Network Technology

³Department of Computer Science,

Faculty of Computer Science and Information Technology,

Rambhai Barni Rajabhat University

Received: 30 June 2022

Revised: 20 September 2023

Accepted: 24 November 2023

Abstract

Human-elephant conflict occurs due to the migration of elephants from their habitat to human agricultural areas in search of food and water resource. This research has therefore introduced Convolutional Neural Network Technology to reduce Human-elephant conflicts, with a framework called YOLOv5 is utilized for real-time object detection from video footage using embedded neural brain devices to monitor and prevent wild elephant intrusions into agricultural areas. First, researchers collected elephant image datasets from monitoring areas, leveraging the advancements in deep learning frameworks to develop Yolo-based architecture models suitable for embedded devices, ensuring both speed and accuracy. In this work, the researchers adjusted the hyperparameters for the YOLO model variants: YOLOv5N, YOLOv5S and YOLOv5M. The computations were conducted with reduced complexity and the proposed models are well-suited for embedded devices. After testing, it was observed that the YOLOv5S model achieved an average accuracy of 95.68% mAP@.5 with a speed increase of up to 50% compared to the YOLOv5M model, which had a maximum accuracy of 95.46% mAP@.5. Then, before deploying the models on embedded devices, researchers augmented the elephant image dataset from the internet for the YOLOv5S model. This augmentation improved the accuracy to a new value of 98.02% mAP@.5. After deployment, it was found that deep learning could accurately detect instances of wild elephant intrusions into agricultural areas, even though it may sometimes be slightly slower than human observation.

Keywords: Elephant Intrusion, Agricultural Area, Reducing Conflict, Convolutional Neural Network

*Corresponding Author; E-mail: vasupon.p@rbru.ac.th

ความเป็นมาและความสำคัญของปัญหา

สถานการณ์ความขัดแย้งระหว่างคนกับช้างป่าในประเทศไทยพบมากในพื้นที่อนุรักษ์ที่มีช้างป่าอยู่ 41 แห่ง คิดเป็น 22% ของพื้นที่อนุรักษ์ทั้งหมด ปัญหาความขัดแย้งระหว่างคนกับช้างรุนแรงที่สุดมีทั้งคนและช้างบาดเจ็บ เสียชีวิต คือ กลุ่มป่าตะวันออก บริเวณป่ารอยต่อ 3 จังหวัด เมื่อพื้นที่ป่าลดลง ทรัพยากร อาหารในป่าก็เปลี่ยน การหาอาหารของช้างก็ลดลง หากินในป่าต่อไม่ได้ต้องหากินไกลออกไปถึงชายป่าหรือเขตชุมชน เมื่อพบพืชผลเกษตรทั้งพืชสวนพืชไร่และแหล่งน้ำ ช้างป่าจึงสามารถหากินหลบซ่อนในบริเวณดังกล่าวมากขึ้น เมื่อที่อยู่อาศัยของมนุษย์เพิ่มขึ้น พื้นที่เกษตรกรรมเพิ่มขึ้น เส้นทางคมนาคมมากขึ้น ส่งผลให้เกิดการเผชิญหน้ากันระหว่างคนและช้าง ในช่วงแรกต่างคนต่างเลียงกันไป พอนานเข้ามักจะแก้ปัญหาด้วยความรุนแรง (Noonto et al., 2018) จากข้อมูลกรมอุทยานแห่งชาติฯ ช่วงปี 2559 - 2560 ยืนยันรายงานช้างป่าออกนอกพื้นที่อนุรักษ์ 6,213 ครั้ง ส่งผลกระทบกับทรัพย์สินชาวบ้านเสียหาย 228 ราย พืชผลเสียหาย 2,355 ราย จากปัญหาความขัดแย้งระหว่างคนกับช้างป่าในช่วงปี พ.ศ. 2555-2561 มีคนและช้างบาดเจ็บและเสียชีวิตรวมทั้งสิ้น 107 เหตุการณ์ พบว่าคนบาดเจ็บ 30 คน เสียชีวิต 45 คน ช้างบาดเจ็บ 7 เชือกและตาย 25 เชือก ซึ่งมีแนวโน้มความรุนแรงที่เพิ่มมากขึ้นอย่างมีนัยสำคัญ สาเหตุการตายหลักของช้างป่าเกิดจากรั่วไฟฟ้าถึงร้อยละ 72 และได้รับบาดเจ็บจากรถยนต์ขณะข้ามถนนร้อยละ 57 ขณะที่สาเหตุของช้างบาดเจ็บสูงที่สุดคือ ถูกรถยนต์ชน และถูกทำร้ายในขณะที่คนกำลังผลักดันช้างออกจากพื้นที่เกษตรกรรม สาเหตุของคนบาดเจ็บและเสียชีวิตมาจากถูกช้างทำร้ายขณะผลักดันหรือเผชิญหน้ากัน การพัฒนาระบบเฝ้าระวังช้างป่าบุกรุกในพื้นที่เกษตรกรรมเพื่อลดความขัดแย้งระหว่างคนและช้าง มีวัตถุประสงค์เพื่อนำเทคโนโลยีการมองเห็นด้วยคอมพิวเตอร์มาเฝ้าระวังช้างป่าบุกรุกพื้นที่เกษตรกรรม แทนอาสาสมัครในพื้นที่ตำบลทุ่งควายกิน อำเภอกงหรา จังหวัดระยอง บริเวณรอยต่อ หมู่ 7 บ้านขุนนมสูง และหมู่ 8 หมู่บ้านเขาจุก เพื่อลดปัญหาการเผชิญหน้าระหว่างคนและช้าง รวมถึงการวางแผนการผลักดันช้างออกจากพื้นที่โดยให้เกิดความสูญเสียน้อยที่สุด

วัตถุประสงค์ของการวิจัย

1. เพื่อพัฒนาแบบจำลองตรวจจับภาพของช้างด้วยโครงข่ายประสาทเทียมแบบคอนโวลูชัน
2. เพื่อทดสอบประสิทธิภาพการแจ้งเตือนช้างป่าบุกรุกในพื้นที่เฝ้าระวัง เพื่อลดการเผชิญหน้ากันระหว่างคนกับช้าง

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

1. ชุดข้อมูลสำหรับการเรียนรู้เชิงลึก

ชุดข้อมูล (Dataset) ใช้ฝึกฝนการเรียนรู้เชิงลึก (Deep Learning) (Kumar et al., 2020) ที่อาศัยรูปแบบ (Pattern) ของข้อมูลที่มากเพื่อสกัดเอาคุณสมบัติ หรือความสัมพันธ์ที่เกิดขึ้น การได้รับข้อมูลที่เพียงพอ อาจจะต้องอาศัยชุดข้อมูลที่เปิดเผย (Open Source Dataset) เพื่อเพิ่มประสิทธิภาพ เช่น ชุดข้อมูลรูปภาพของอิมเมจเน็ต (Imagenet) มีรูปภาพมากกว่า 14 ล้านภาพ ครอบคลุมมากกว่า 20,000 หมวดหมู่ มีมากกว่า 1,000,000 ภาพที่มีคำอธิบายและตำแหน่งของวัตถุเป้าหมาย โคโค (COCO Common Objects in Context) เป็นชุดข้อมูลของการรู้จำมีการแบ่งส่วนและคำบรรยาย ประกอบด้วยภาพขนาดเล็ก 200,000 ภาพ ที่มีคุณสมบัติสำหรับแต่ละภาพ เช่น การแบ่งส่วนวัตถุมากกว่า

การล้อมกล่องวัตถุ และคำบรรยายหัวข้อความต้อภาพ เช่น “เครื่องบินแอร์บัส A380 ขึ้นสู่เมฆ” และ “เครื่องบินที่บินผ่านท้องฟ้าสีคราม” เป็นต้น (Zhou et al., 2017)

2. โครงข่ายงานประสาทเทียม

โครงข่ายงานประสาท (Neural Networks) คือการเรียนรู้ด้วยเครื่องที่ได้แรงบันดาลใจจากโครงข่ายประสาทชีวภาพที่เลียนแบบการทำงานของสมองสิ่งมีชีวิต ประกอบด้วยนิวรอน (Neurons) ที่ใช้เชื่อมโยงการแก้ปัญหา ในหนึ่งโครงข่ายประสาทประกอบไปด้วยข้อมูลขาเข้า (Vector x) หรือเซตของหลายเซลล์ประสาท ในแต่ละเซลล์ต่อเข้ากับชั้นซ่อนตัว (Hidden Layer) ของหลายเซลล์ ชั้นซ่อนมีผลต่อประสิทธิภาพการเรียนรู้ด้วยแบบจำลอง ผ่านเซตของน้ำหนัก (Weight) ที่ใช้ระบุความแข็งแรงในแต่ละเซลล์ประสาท (w) จำนวนชั้นและจำนวนนิวรอนส่งผลต่อความสามารถในการสกัดข้อมูล สำหรับข้อมูลขาออกของเซลล์ซ่อนตัว (Hidden Neuron Outputs) ณ j^{th} เมื่อสมการ (1) ฟังก์ชันกระตุ้น (Activation Function) รับผลรวมทั้งหมดจากทุกข้อมูลขาเข้า (Dendrite) แล้วพิจารณาว่าจะส่งต่อผลลัพธ์ (Output Neuron Outputs) ไปยังสมการ (2) ถ้าต้องการความน่าจะเป็นเหล่านั้น สามารถแปลงชั้นข้อมูลขาออก (Output Layer) ด้วยสมการ (3) ฟังก์ชันซอฟต์แม็กซ์ (SoftMax) โดยสามารถสรุปผลการทำงานเป็นดังภาพที่ 1 (Girshick et al., 2014)

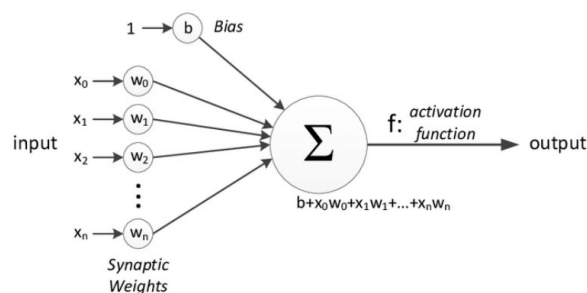
$$h_j = \phi(\sum_i w_{ij}x_i) \quad (1)$$

$$y_i = \phi(\sum_j v_{ij}h_j) \quad (2)$$

$$\text{softmax}(x_n) = \frac{e^{x_n}}{\sum_m e^{x_m}} \quad (3)$$

โดยที่ ϕ = ฟังก์ชันกระตุ้น, w = ค่าน้ำหนัก และ x = ข้อมูลขาเข้า

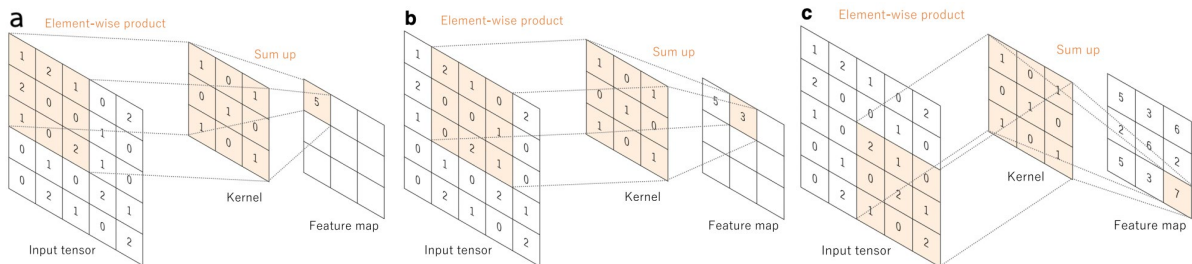
ฟังก์ชันทั่วไปของฟังก์ชันกระตุ้น (Common Activation Function) สำหรับ ϕ คือฟังก์ชันซิกมอยด์ (Sigmoid Function, $\sigma(x)$) เป็นการแปลงเลขให้อยู่ในช่วง 0 ถึง 1 หรือ hyperbolic tangent, $\tanh(x)$ สำหรับการแปลงให้อยู่ในช่วง -1 ถึง 1 หรือ rectified linear unit, $\text{ReLU}(x) = \max(0, x)$



ภาพที่ 1 การทำงานของประสาทเทียมที่รับข้อมูลขาเข้ามาคูณกับตัวน้ำหนักที่ได้จากการสุ่มหาผลรวมของแต่ละขารวมเข้ากับ Bias ป้อนเข้าสู่ฟังก์ชันกระตุ้นสรุปผลออกไปยังประสาทชั้นต่อไป (Sanguansat, 2019)

3. คอนโวลูชันนิวรอนเน็ตเวิร์ค

คอนโวลูชันนิวรอนเน็ตเวิร์ค (Convolutional Neural Network) เป็นการคำนวณทางคณิตศาสตร์เชิงเส้นที่ใช้หาความสัมพันธ์ของพิกเซล (Pixel) ที่อยู่บริเวณใกล้เคียงและจัดเก็บในรูปแบบเมทริกซ์ (Matrix) คอนโวลูชันจะอยู่ในขั้นตอนของแบบจำลองเพื่อต่อยอดความรู้ประกอบกันเป็นจุด, เส้นแนวตั้ง, เส้นแนวนอน, แนวทแยง, กากบาท, มุม, เส้นโค้ง, วงกลม, ลวดลาย, ดวงตา, หู, จมูก, ปาก, ใบหน้า เป็นต้น ด้วยฟิวเจอร์ดีเทกเตอร์ที่แตกต่างของวิธีการคำนวณแบบ dot product กับเคอร์เนล ดังภาพ 2 เช่น 3×3 , 5×5 หรือ 7×7 เป็นต้น ผลลัพธ์ที่ได้จากการคำนวณเรียกว่า ฟิวเจอร์แมปไปจนถึงขั้นลึก ๆ ที่เป็นนามธรรมมากขึ้น แบบจำลองส่วนใหญ่จะจบด้วยชั้นฟูลคอนเน็คเตด (Fully Connected Layer) (Unzueta, 2021) เพื่อปรับผลลัพธ์ (Output) ของแบบจำลองให้ตรงกับงานที่ต้องการ เช่น การจำแนกประเภทหรือคลาส จากความน่าจะเป็น (Probability) ด้วยสมการซอร์ฟแม็กซ์ ที่ให้ผลลัพธ์อยู่ในช่วง 0 ถึง 1 และเลือกใช้ฟังก์ชันกระตุ้นด้วยซิกมอยด์ หรือฟังก์ชันเรลู (ReLU) หรือฟังก์ชันไฮเพอร์โบลิคแทนเจนต์ (Tanh) ได้

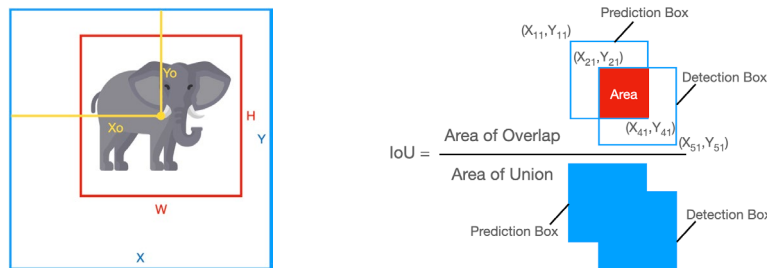


ภาพที่ 2 ขั้นตอนคอนโวลูชันที่มีขนาดเคอร์เนล 3×3 ไม่มีช่องว่างภายใน และขยับที่ 1 step. ตามลำดับภาพ a, b และ c ด้วยการคำนวณคอนโวลูชัน เพื่อสกัดหาความสัมพันธ์ของพิกเซล (Yamashita et al., 2018)

4. ขั้นตอนวิธีของโยโลเวอร์ชันที่ห้าและโรโบโฟว์

YOLO (You Only Look Once) (Gur-Arie, 2022) เป็นงานด้านคอมพิวเตอร์วิทัศน์ (Computer Vision) มีขั้นตอนวิธีการตรวจจับวัตถุที่แบ่งรูปเป็นระบบกริด แต่ละเซลล์มีหน้าที่ตรวจจับวัตถุภายในตัวเอง โดดเด่นด้านการตรวจจับ (Detect) วัตถุที่ซ้อนกัน มีโครงสร้างเป็นชั้น (Layers) ประกอบด้วยชั้นคอนโวลูชันที่ซ้อนกันหลายขนาด ในช่วงแรกของนิวรอน มีความเร็วและความแม่นยำ (Ultralytics, 2021) สามารถทำความเร็ว 140 FPS บนหน่วยประมวลผล Nvidia Tesla P100 เมื่อเทียบกับวิธีโวลัวตัวเก่า (YoloV4) ที่ความเร็ว 50 FPS โยโลวีไฟฟ์ (YoloV5) มีขนาด 27MB ใช้สถาปัตยกรรม Darknet ได้รับการสืบทอดคุณสมบัติที่ดีจากโยโลวีโฟว์และปรับโครงสร้างประสาทให้ทำงานเร็วขึ้น ใช้หน่วยความจำน้อยลงด้วยวิธี เช่น Mosaic training, Self-Adversarial Training (SAT) และ Multi-Channel Feature แทนที่ FPN fusion ด้วย PANet (Liu et al., 2018) แบบจำลองโครงข่ายประสาทของโยโลมีให้เลือกหลากหลายตามความเหมาะสมของงาน เช่น YOLOV5s เป็นแบบจำลองขนาดเล็กสำหรับภาพสี (RGB) ขนาด 640×640 พิกเซล

โรโบฟโฟว์ (Roboflow Platform) เป็นซอฟต์แวร์สำหรับเตรียมชุดข้อมูลให้พร้อมก่อนเข้าเน็ตเวิร์คในรูปแบบการติดฉลาก (Labeling Format) เพื่อกำหนดขอบเขตเป้าหมายในภาพ เช่น จุดศูนย์กลางวัตถุไปยังแกน x (X Center) และแกน y (Y Center) ขนาดความกว้างของกรอบ (Box Width) และความสูงของกรอบ (Box Height) เป็นกรอบสี่เหลี่ยมล้อมรอบ (Bounding-Box Box) ตำแหน่งหรือคำอธิบายจะถูกปรับให้อยู่ในมาตรฐาน 0 ถึง 1 เพื่อใช้หาค่าไอโอยู (Solawetz, 2020) ดังภาพที่ 3 การฝึกแบบจำลองจะแบ่งชุดข้อมูลออกเป็น 3 คือ (1) ส่วนฝึกอบรม (Training) (2) ส่วนตรวจสอบและเรียนรู้เพิ่มเติมจากผลการตรวจ (Validation) และ (3) ส่วนทดสอบ (Test Set) แบบจำลองที่เลือกใช้ต้องเหมาะสมกับขนาดภาพ เช่น ขนาด 640*640px ต้องใช้ YOLOv5n (นาโน), YOLOv5s (เล็ก), YOLOv5m (กลาง), YOLOv5l (ใหญ่), YOLOv5x (ใหญ่พิเศษ) ขนาด 1280*1280px ใช้ YOLOv5n6, YOLOv5s6, YOLOv5m6, YOLOv5l6, YOLOv5x6 เป็นต้น เหมาะกับงานที่ต้องการความละเอียดสูง (Ultralytics, 2021)



ภาพที่ 3 การกำหนดขอบเขตเป้าหมายของวัตถุในภาพ เพื่อใช้ทำนายผลความแม่นยำด้วยค่าไอโอยู (IoU) ระหว่างผลเฉลยของตำแหน่งวัตถุในภาพ กับผลการทำนายที่แบบจำลองทำนายได้มีหน่วยเป็น mAP

5. การประเมินประสิทธิภาพแบบจำลอง

Intersection over Union (IoU) เป็นการวัดประสิทธิภาพแบบจำลองสำหรับงานตรวจจับวัตถุ (Object Detection) หาได้จากอัตราส่วนระหว่างพื้นที่ (Area) ที่ทับซ้อน (Intersection) ของพื้นที่ทั้ง 2 (Bounding Box) หารด้วยพื้นที่รวมทั้งหมด (Area) ดังภาพที่ 4 หรือเรียกว่า ดัชนี (Jaccard) เป็นวิธีในการหาเปอร์เซ็นต์ทับซ้อนระหว่างผลเฉลย (Ground Truth) และผลทำนาย (Predict) ถ้าได้ความแม่นยำที่ทับซ้อนมากกว่า 0.5 ถือว่ายอมรับได้ ค่า AP และ mean Average Precision (mAP) เป็นการหาค่าเฉลี่ยความแม่นยำ ซึ่งคำนวณจากสมการ (4) ค่าความแม่นยำ (5) ค่าความรำลึก (6) ค่าความถูกต้อง และ RMSE (Root Mean Square Error) รากที่สองของความคลาดเคลื่อนเฉลี่ย เพื่อคำนวณความผิดพลาดการพยากรณ์เข้าใกล้ศูนย์แสดงว่าใกล้เคียงกับความจริง มากกว่า 0.5 เท่ากับจำนวน Class ที่มีค่าความแม่นยำที่ทับซ้อนมากกว่า 0.5 TP+FP มีค่าเท่ากับจำนวน Class ที่ทำนายได้ และ TP+FN เท่ากับผลเฉลย (Kumar et al., 2020)

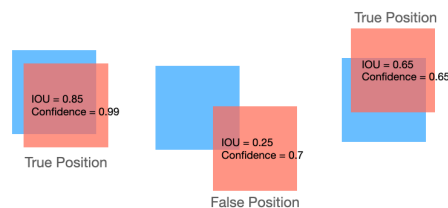
$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (4)$$

$$Recall = \frac{TP}{TP+FN} = \frac{count(TP)}{\#ground truths} \quad (5)$$

$$Precision = \frac{TP}{TP+FP} = \frac{count(TP)}{\#predictions} \quad (6)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i' - Y_i)^2} \quad (7)$$

โดยที่ True Positive (TP) คือข้อมูลเป็นจริง และโปรแกรมทำนายว่าจริง
True Negative (TN) คือข้อมูลไม่เป็นจริง และโปรแกรมทำนายว่าไม่จริง
False Positive (FP) คือข้อมูลเป็นจริง แต่โปรแกรมทำนายว่าไม่จริง
False Negative (FN) คือข้อมูลไม่เป็นจริง แต่โปรแกรมทำนายว่าจริง
 Y_i คือ ค่าข้อมูลที่เกิดขึ้นจริง, Y_i' คือ ค่าที่ได้จากการพยากรณ์
 n คือ จำนวนข้อมูลเข้าทั้งหมด



ภาพที่ 4 การหาความแม่นยำที่ทับซ้อนจากพื้นที่วัตถุในภาพ (สีฟ้า) และพื้นที่คำทำนาย (สีแดง)

ซึ่งมีค่า Precision = count(true positive) / count(all red boxes) = 2/3 = 0.67

และค่าของ Recall = count(true positive) / count(all blue boxes) = 2/3 = 0.67

6. งานวิจัยที่เกี่ยวข้อง

Madheswaran et al. (2019) ได้สร้างเสียงผึ้งและเสียงเสือคำรามที่ช่างไม่ชอบ โดยเปรียบเทียบผลการใช้แบบจำลองตรวจจับวัตถุ 4 แบบคือ SSD mobilenet v2 model, SSDlite mobilenet v2 model, SSD inception v2 model และ Fast R-CNN inception v2 พบว่า mobilenet v2 มีความแม่นยำมากที่สุด

Jagannathan et al. (2019) และคณะได้ทดสอบเปรียบเทียบโครงข่ายประสาทเทียมแบบคอนโวลูชันแบบดั้งเดิมกับการใช้เทคนิคการสับทอดเพื่อระบุตัวช่างจากฟีดวิดีโอที่ได้รับ เสียงผึ้ง เสียงเสือคำราม เพื่อให้ช่างลดการเดินทางไปในพื้นที่อยู่อาศัย พบว่าแบบจำลองที่มีการสับทอดความรู้ VGG16 CNN มีความแม่นยำกว่าที่ค่าความแม่นยำ 94% สำหรับการใช้งานแบบเรียลไทม์ได้ใช้แบบจำลอง SqueezeNet CNN เนื่องจากเวลาในการคำนวณต่ำกว่า (ต่างกันที่ 0.02 – 0.05 วินาที) และมีความแม่นยำระดับกลางที่ 92.67%

Duporge et al. (2020) ได้ใช้แบบจำลองโครงข่ายประสาทเทียมแบบคอนโวลูชันตรวจจับและนับจำนวนช่างแอฟริกาด้วยภาพถ่ายดาวเทียมเพื่อเทียบความแม่นยำของ CNN กับมนุษย์ นอกจากนี้ได้ใช้แบบจำลองนี้ทดสอบภาพถ่ายดาวเทียมจากเคนยา โดยไม่ฝึกอบรมเพิ่ม เพื่อทดสอบข้อมูลประชากรช่างที่อยู่นอกพื้นที่ได้หรือไม่ ผลที่ได้คือ CNN มีความแม่นยำสูงเทียบเท่ากับความสามารถในการตรวจจับของมนุษย์ เช่น การตรวจจับด้วย CNN คือ 0.78 ในพื้นที่ต่างกัน และ 0.73 ในพื้นที่เดียวกัน เมื่อเทียบกับมนุษย์ คือ 0.77 ในพื้นที่ต่างกัน และ 0.80 ในพื้นที่เดียวกัน ซึ่งแสดงให้เห็นว่า CNN สามารถนำไปใช้สำรวจระยะไกลได้

Gur-Arie (2022) ได้ใช้การระบุวัตถุในภาพด้วยเครื่องคอมพิวเตอร์ ด้วยอัลกอริทึมตรวจจับวัตถุแบบเรียลไทม์ YOLO (you only look once) พัฒนาโดย Ultralytics ที่ใช้เทคนิคการสืบทอดความรู้ (Transfer Learning) เพื่อจำแนกภาพนกเพนกวิน 250 ภาพจากวิดีโอ ด้วยโรโบฟโฟว์ (Roboflow Platform) เทคนิคการถ่ายโอนความรู้ส่งผลให้ได้ผลลัพธ์ที่ดีกว่าการฝึกตั้งแต่เริ่มต้น ผลการวิจัยพบว่าแบบจำลองทำนายได้ถูกต้อง 98% ถ้าต้องการความแม่นยำเพิ่มสามารถทำได้โดย เพิ่มเวลาทดสอบ (Test-Time Augmentations TTA) การพลิกภาพ กลับภาพ และความเอียงภาพ (Horizontal, Rotation, Flip and Different Resolutions)

Zhou et al. (2017) นำเสนอการตรวจจับวัตถุคือ ผู้เล่นฟุตบอล ผลการแข่งขัน มุมธง และลูกฟุตบอล จากชุดข้อมูลที่สร้างขึ้นใหม่ โดยใช้เครือข่ายประสาทเทียมแบบ R-CNN (Girshick et al., 2014) ที่เร็วกว่าเครือข่ายประสาทเทียมแบบ CNN ผลการเรียนรู้แสดงให้เห็นว่าแต่ละคลาสสามารถทำได้ดีเมื่อภาพอยู่ในมุมที่ใกล้ โดยมีความแม่นยำเฉลี่ย (MAP) คือ player = 79.02, soccer goal = 83.77, corner flag = 35.08 และ football = 47.52

Hongboonmee and Sitthichokchaisiri (2021) ได้จำแนกภาพด้วยเทคนิคโครงข่ายประสาทเทียมการเรียนรู้เชิงลึกแบบคอนโวลูชันที่ใช้แบบจำลองโมบายเน็ตผ่านไลบรารีเทนเซอร์โฟลว์ จำแนกภาพโรคตาแดง ตาгүйยัง ต้อลม ต้อเนื้อ ต้อกระจก ต้อหินและตาปกติ ประเภทละ 100 ภาพ ฝึก 500 รอบมีความแม่นยำ 96.40%

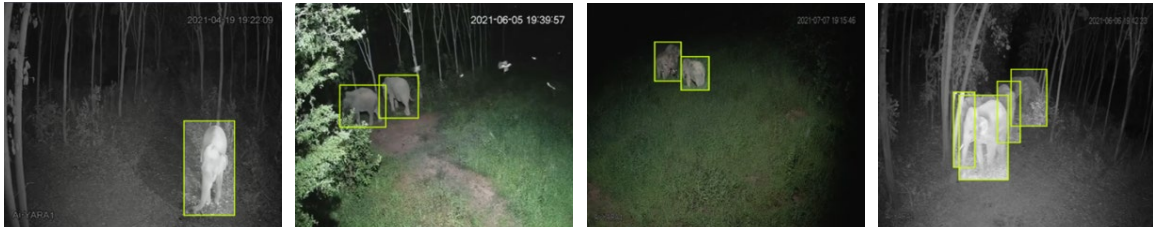
Kaewtathanawatthana and Brorewongtrakhul (2020) เสนอผลวิเคราะห์ความขัดแย้งระหว่างคนกับช้างป่าและศึกษาแนวทางการอยู่ร่วมกันอย่างยั่งยืน ตำบลพวา อำเภอแก่งหางแมว จังหวัดจันทบุรี โดยใช้วิธีการวิเคราะห์เชิงคุณภาพ (Qualitative Research) สัมภาษณ์เจาะลึก (In-Depth Interview) มีคำถามเป็นเครื่องมือเก็บข้อมูลจากผู้ใช่ 2 กลุ่ม คือผู้นำชุมชน 5 หมู่บ้าน และกลุ่มเกษตรกรที่ได้รับผลกระทบ 20ครัวเรือน ผลที่ได้แบ่งเป็นสามแนวทางคือ ระยะเร่งด่วน ตั้งชุดปฏิบัติการเฝ้าระวังช้างป่า ชุดปฏิบัติการเคลื่อนที่เร็วลาดตระเวนและให้คำแนะนำตลอดจนการช่วยเหลือ ระยะกลาง จัดทำแนวคูกันช้างป่าในจุดที่เป็นอันตรายมีประชาชนสัญจร และปลูกไฟให้หนาแน่นตลอดแนวคูกัน ทำรั้วไฟฟ้าพื้นที่สวนผลไม้ ระยะยาวประกอบไปด้วย 4 แนวทาง (1) สร้างทัศนคติที่ดีเพื่อการอยู่ร่วมกัน (2) ปรับรูปแบบการทำเกษตรกรรมรอบพื้นที่ (3) ปรับปรุงระเบียบราชการต้องมีแนวปฏิบัติที่ชัดเจน รวดเร็วในการให้ความช่วยเหลือ (4) ประสานกับหน่วยงานกระทรวงทรัพยากรธรรมชาติและสิ่งแวดล้อมเพื่อเคลื่อนย้ายช้างป่าออกไปเมื่อมีปริมาณมากขึ้น

วิธีดำเนินการวิจัย

การรวบรวมและเตรียมข้อมูล

งานวิจัยนี้ได้เก็บข้อมูลภาพช้างจากกล้องวงจรปิดในพื้นที่เฝ้าระวังจำนวน 2,278 ภาพ มีทั้งภาพ ด้านหน้า ด้านข้าง ด้านหลัง ของช้าง ช้างหลายเชือก ช้างยืนซ้อนกัน ช้างเชือกเดียว ช้างหลบหลังต้นไม้หรือพุ่มไม้ ช้างมีงา ช้างไม่มีงา ภาพส่วนใหญ่เป็นภาพถ่ายตอนกลางวัน จากนั้นแบ่งข้อมูลเป็น 3 ส่วนด้วยวิธีสุ่ม พร้อมทำป้ายกำกับช้างในภาพ (Data and Ground Truth Labels) ดังภาพที่ 5 สำหรับส่วนฝึกอบรมมีจำนวน 1,600 ภาพ คิดเป็นร้อยละ 70 ส่วนตรวจสอบผลการฝึกอบรมและเรียนรู้เพิ่มเติมจากการฝึก จำนวน 452 ภาพ คิดเป็นร้อยละ 20 และส่วนสำหรับทดสอบ ที่เป็นข้อมูลที่ไม่เคยพบมาก่อน 226 ภาพ คิดเป็นร้อยละ 10 สำหรับการทำนายผลโดยไม่มีการปรับจูนพารามิเตอร์ใด เพื่อทดสอบความสามารถว่าทำงานได้ดีเพียงใดเมื่อต้องเจอกับสถานการณ์จริง การเลือกใช้ข้อมูลในพื้นที่เพื่อให้การฝึกอบรมคล้ายกับข้อมูลที่ต้องพบมากที่สุดและลดปัญหาความสับสน (Overfitting) แบบจำลองที่ถูก

วิวัฒนาการมาแล้วสามารถช่วยเพิ่มความแม่นยำได้ โดยเพิ่มชุดข้อมูลภาพช้างจากอินเทอร์เน็ตรวมเข้าไปในชุดข้อมูลเดิมเป็นจำนวน 3,788 ภาพ และแบ่งเป็น 3 ชุด ด้วยวิธีสุ่ม คือชุดฝึกอบรม จำนวน 2,640 ภาพ คิดเป็นร้อยละ 70 ข้อมูลสำหรับตรวจสอบผลการอบรมและเรียนรู้เพิ่มเติม จำนวน 756 ภาพ คิดเป็นร้อยละ 20 และชุดข้อมูลสำหรับทดสอบ กับข้อมูลที่ไม่เคยพบมาก่อน 392 ภาพ คิดเป็นร้อยละ 10 โดยชุดข้อมูลทั้ง 2 มีขนาด 640px*640px เป็นภาพสี



ภาพที่ 5 สร้างป้ายกำกับและข้อมูลตำแหน่งวัตถุภายในภาพ ด้วยซอฟต์แวร์ของโรโบโฟร์

การฝึกอบรมและสร้างแบบจำลองที่เหมาะสมกับหน่วยประมวลผลสมองกลฝังตัว

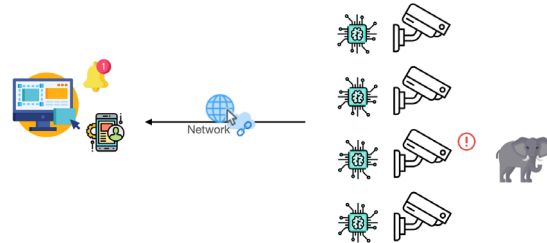
ใช้หน่วยประมวลผลกราฟิกส์ (GPU: Graphics Processing Unit) Nvidia tesla p100 หน่วยความจำ 16 GB มีขั้นตอนคือ (1) ใช้แบบจำลอง yolov5n, yolov5s และ yolov5m ที่พัฒนาโดย Ultralytics ซึ่งมีขนาดการคำนวณเล็กไปใหญ่ตามลำดับ ขนาดภาพไม่เกิน 640px ด้วยค่าไฮเปอร์พารามิเตอร์เริ่มต้น ฝึกอบรม (Epochs) 300 รอบ ป้อนภาพเข้าสู่แบบจำลอง (Batch Size) จำนวน 32 ภาพ (2) ทดลองใช้ไฮเปอร์พารามิเตอร์ที่วิวัฒนาการมาแล้ว (Hyperparameter Evolution) จากอัลกอริทึม Genetic Algorithm (GA) ไปให้กับแบบจำลอง yolov5s, yolov5n และ yolov5m จากการศึกษางานวิจัยเทคนิคการสืบทอดความรู้สำหรับการสุ่มค่าเบื้องต้นที่ดีจะให้ผลลัพธ์ในขั้นสุดท้ายที่ดีกว่า โยโลวีไฟฟ์ มีไฮเปอร์พารามิเตอร์ที่เรียกว่า VOC ที่เคยผ่านการฝึกฝนมาแล้วด้วยชุดข้อมูล COCO128 ที่ปรับแต่งมาอย่างละเอียด ด้วยแบบจำลอง yolov5s (3) ทดสอบการป้อนภาพเข้าสู่แบบจำลองที่เหมาะสมกับชุดข้อมูล คือ 16, 32, 64 และ 128 ตามลำดับ ด้วยค่าไฮเปอร์พารามิเตอร์ VOC กับแบบจำลอง yolov5n yolov5s และ yolov5m (4) ทดสอบเพิ่มประสิทธิภาพแบบจำลองจากแบบจำลองที่เหมาะสมที่สุดด้วยการเพิ่มชุดข้อมูลบนอินเทอร์เน็ต (5) ทดสอบเพิ่มความเร็วด้วยการลดขนาดภาพลงจาก 640 เหลือ 416 และ 320 เพื่อเปรียบเทียบความแม่นยำกับความเร็วที่เพิ่มขึ้น จากนั้นนำไปใช้กับหน่วยประมวลผลฝังตัว (Embedded Device) ที่จุดเฝ้าระวัง

การประเมินประสิทธิภาพแบบจำลอง

ผลการฝึกอบรมจะใช้การหาค่าเฉลี่ย (mAP) บนค่าความแม่นยำที่ทับซ้อนที่แตกต่างกันระหว่าง 0.5 ถึง 0.95 ด้วยระยะ 0.05 (0.5, 0.55, 0.6, 0.65, 0.7, 0.75, 0.8, 0.85, 0.9, 0.95) $mAP@IoU=0.5:0.95$ ($mAP@[.95]$) และหน่วย $mAP@IoU = 0.5$ ($mAP@.5$) ซึ่งเป็นหน่วยมาตรฐานของ MS COCO ตามด้วยค่าความผิดพลาดของการฝึกอบรม train/obj_loss ค่าความผิดพลาดของการตรวจสอบผลทำนาย val/obj_loss และความน่าจะเป็นที่แบบจำลองสามารถตรวจจับ (Recall) จากนั้นประเมินการใช้หน่วยความจำ (GPU Memory) และระยะเวลาที่ใช้ประโยชน์จากหน่วยประมวลผล (GPU Utilization)

ติดตั้งแบบจำลองที่จุดเฝ้าระวังข้างป่า

นำแบบจำลองที่ได้ไปติดตั้งกับหน่วยประมวลผลแบบฝังตัวของ Nvidia รุ่น Jetson Nano สำหรับการรับภาพจากกล้องวงจรปิด (IP Camera) ขนาด 1024p 30 เฟรมต่อวินาที (30 fps) ด้วย FTP โพรโทคอล (File Transfer Protocol) พร้อมประมวลผลสำหรับเฟรมเรทที่มีการเปลี่ยนแปลงมากกว่าร้อยละ 20 เพื่อประหยัดการประมวลผล โดยมีผังการติดตั้งเป็นดังภาพที่ 6



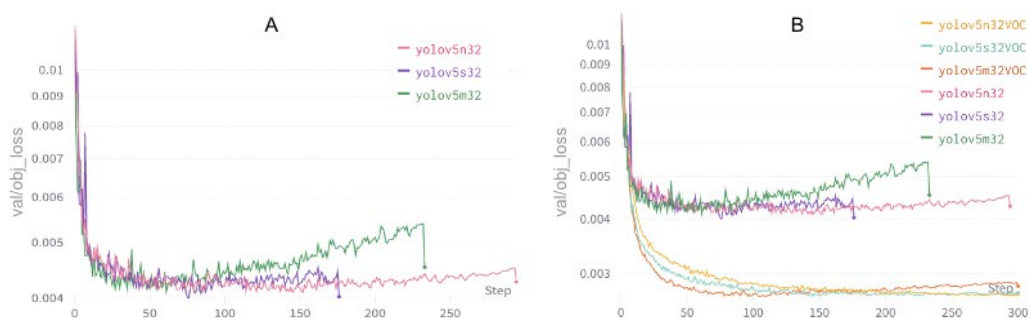
ภาพที่ 6 แผนผังการติดตั้งระบบกล้องวงจรปิดเพื่อรับภาพไปประมวลผลต่อที่หน่วยประมวลผลแบบฝังตัวเพื่อการแจ้งเตือนไปยังแอปพลิเคชันไลน์ของผู้ที่เกี่ยวข้อง

ประเมินผลการแจ้งเตือน

หลังจากติดตั้งระบบ ผู้วิจัยร่วมกับผู้นำหมู่บ้าน กลุ่มอาสาสมัครเฝ้าระวังและผลักดันช้างป่า ทำการทดสอบประสิทธิภาพการแจ้งเตือนผ่านกลุ่มไลน์ ด้วยการเฝ้าดูช้างที่เดินเข้ามาในจุดเฝ้าระวังผ่านกล้องวงจรปิดในเวลาที่ยังชอบออกหากินคือ ช่วงพระอาทิตย์ตกดิน และกลับเข้าป่าก่อนพระอาทิตย์ขึ้น ตามคำบอกเล่าของชาวบ้านในพื้นที่เป็นระยะเวลา 1 สัปดาห์ เพื่อประเมินความน่าเชื่อถือของระบบ

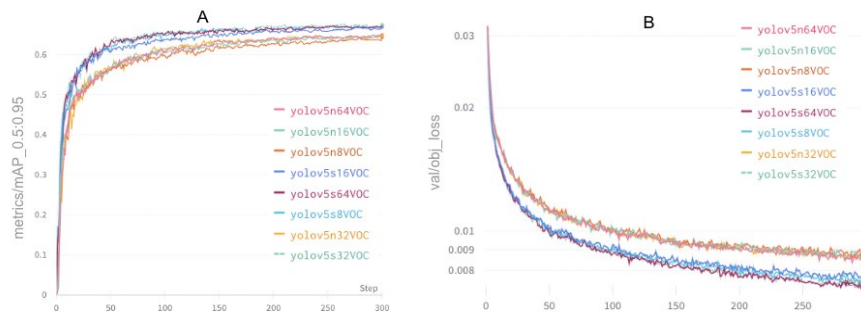
ผลการวิจัย

การฝึกด้วยไฮเปอร์พารามิเตอร์เริ่มต้นทำให้เกิดการโอเวอร์ฟิต เมื่อใช้ชุดข้อมูลสำหรับทดสอบผลฝึกอบรม ดังภาพที่ 7A กราฟแสดงความถูกต้องกับชุดข้อมูลการตรวจสอบ พบว่าเมื่อฝึกอบรมไปได้ 75 รอบ แบบจำลองทั้งสาม เริ่มทำได้ไม่ดีกับข้อมูลที่ไม่เคยพบ เห็นได้ชัดว่าแบบจำลองที่มีชั้นคำนวณที่ซับซ้อน yolov5m (สีเขียว) มีปัญหาโอเวอร์ฟิตมากที่สุด จากการศึกษาทางวิจัยพบว่าเทคนิคการสับทอดไฮเปอร์พารามิเตอร์สามารถช่วยให้ได้ผลลัพธ์ที่ดีสำหรับการใช้ค่า VOC กับแบบจำลองทั้งสาม (สีเหลือง สีเขียวอ่อน สีส้ม) เห็นได้ว่าแบบจำลองทั้งสาม สามารถแก้ปัญหาได้ดีกับข้อมูลที่ไม่เคยพบดังภาพที่ 7B



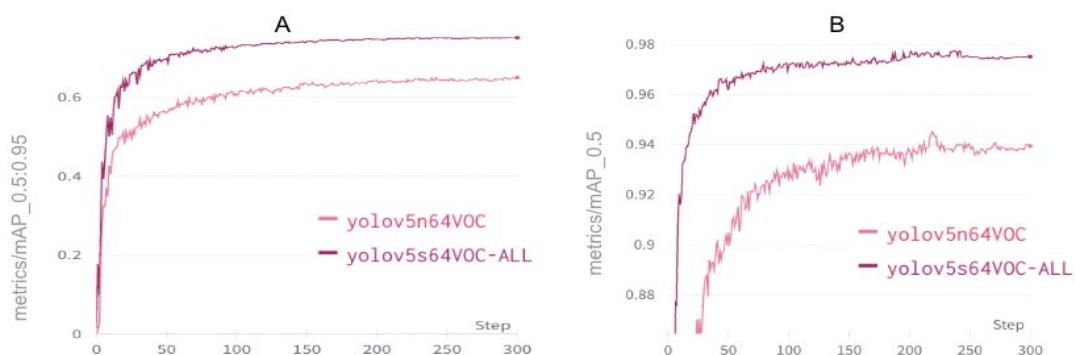
ภาพที่ 7 กราฟ (A) ผลการผิ่กอบรมแบบจำลองทั้ง 3 ด้วยไฮเปอร์พารามิเตอร์เริ่มต้นที่เกิดโอเวอร์ฟิตตั้ง
และกราฟ (B) ผลเปรียบเทียบการผิ่กระหว่างกราฟ (A) และการสืบทอดไฮเปอร์พารามิเตอร์ VOC

ทดลองแบ่งภาพเข้าสู่แบบจำลองทั้งสามด้วยขนาด 16, 32, 64 และ 128 พบว่าแบบจำลอง yolov5m ไม่สามารถใช้ขนาด 64 ได้เนื่องจากหน่วยความจำไม่พอ ที่ขนาด 32 ใช้เวลาประมวลผล 408.767 นาที นานกว่า yolov5s ที่ใช้เวลา 172.25 นาที คิดเป็นร้อยละ 57.86 และ yolov5n ที่ใช้เวลา 99.96 นาที คิดเป็นร้อยละ 75.54 การใช้หน่วยความจำ yolov5m ใช้หน่วยความจำที่ 82.84% มากกว่า yolov5s ที่ใช้หน่วยความจำ 50.03% และ yolov5n ใช้หน่วยความจำ 32.25% ดังนั้นการทดลองนี้จึงใช้เฉพาะแบบจำลอง yolov5n และ yolov5s จากภาพที่ 8A เห็นได้ชัดว่าค่าเฉลี่ยผลการฝึกอบรมระหว่าง 0.5 – 0.95 ของแบบจำลอง yolov5s ทำได้ดีกว่า yolov5n ในทุกขนาดแบบ (น้ำเงิน เลือดหมู ฟ้า เส้นประเขียว) ซึ่งสอดคล้องกับ ภาพที่ 8B แบบจำลอง yolov5s มีค่าความผิดพลาดของการตรวจสอบที่น้อยกว่าแบบจำลอง yolov5n ทุกขนาดแบบ



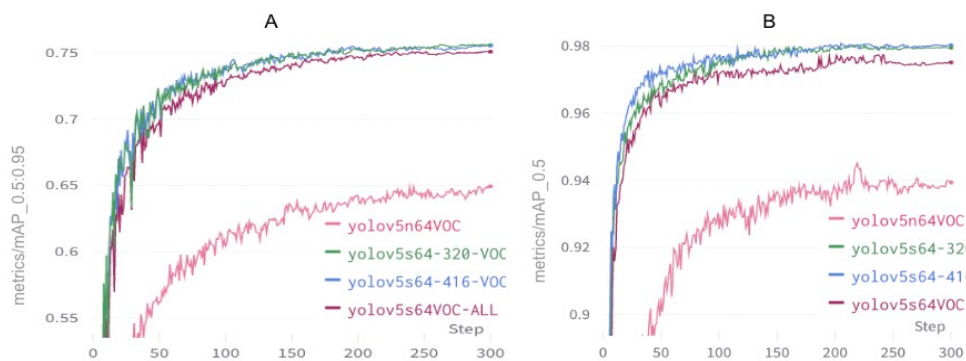
ภาพที่ 8 ผลทดสอบการแบ่งชุดข้อมูลภาพเข้าสู่แบบจำลอง yolov5s และ yolov5n

เมื่อวิเคราะห์จากภาพที่ 8 พบว่าค่า Batch 64 จาก yolov5s มีค่าความผิดพลาดเฉลี่ยของการตรวจสอบข้อมูลที่ไม่เคยพบมาก่อนต่ำที่สุดตามด้วย Batch 32 จาก yolov5s ที่ประมวลเร็วกว่า แต่มีค่าความแม่นยำต่ำกว่า ดังนั้นการทดสอบในขั้นตอนถัดไปจึงได้เลือกใช้แบบจำลอง yolov5s ที่ขนาดแบบเท่ากับ 64 สำหรับการทดสอบเพิ่มภาพจากอินเทอร์เน็ตพบว่า ความแม่นยำที่ทับซ้อนเฉลี่ยระหว่าง 50% ถึง 95% มีค่าสูงขึ้นจากเดิม 64.93% mAP@.5-0.95 เป็น 75.09% mAP@.5-0.95 ดังภาพ 9A ส่วนความแม่นยำที่ทับซ้อนมากกว่า 50% คือจากเดิม 93.94% mAP@.5 เป็น 97.51% mAP@.5 ดังภาพ 9B

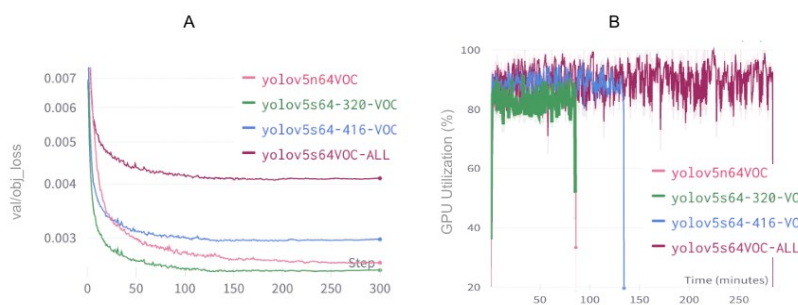


ภาพที่ 9 ผลการเพิ่มภาพจากอินเทอร์เน็ตพบว่าความแม่นยำเฉลี่ย mAP สูงขึ้น ทั้งค่าที่ได้จาก IoU ของพื้นที่ทับซ้อนระหว่าง 50% ถึง 95% และความแม่นยำที่ทับซ้อนมากกว่า 50% ด้วย yolov5s ที่ batch 64

จากนั้นลดขนาดรูปภาพจาก 640 เป็น 416 และ 320 เพื่อทดสอบผลลัพธ์ที่เกิดขึ้นพบว่าขนาด 416 ดังภาพที่ 10 กราฟสีฟ้า มีคะแนนที่สูงกว่าเดิมทั้ง mAP ที่ 50-95% มีความแม่นยำเฉลี่ยอยู่ที่ 75.55 mAP@.5-0.95 และ mAP ที่ 50% มีความแม่นยำอยู่ที่ 98.02mAP@.5 สำหรับภาพขนาด 320 ดังภาพที่ 10 กราฟสีเขียว มีความแม่นยำเฉลี่ยอยู่ที่ 75.59 mAP@.5-0.95 แทบไม่ต่างกับภาพขนาด 416 และ mAP ที่ 50% มีความแม่นยำอยู่ที่ 97.95mAP@.5 น้อยกว่าภาพ 416 อยู่เล็กน้อย เมื่อไปดูที่ค่าความผิดพลาดของการตรวจสอบพบว่าภาพขนาด 320 สามารถรับมือกับข้อมูลที่ไม่เคยพบเจอมาก่อนได้ดีที่สุดดังภาพที่ 11A กราฟสีเขียวมีคะแนน 0.002531 ตามด้วยกราฟสีชมพูที่ใช้ภาพขนาด 640 มีคะแนน 0.002634 ซึ่งเป็นแบบจำลองที่ยังไม่ได้เพิ่มภาพจากอินเทอร์เน็ตเข้าไปจึงอนุมานได้ว่าความสับสนจากชุดข้อมูลมีน้อยกว่าชุดข้อมูลภาพข้างที่มาจากอินเทอร์เน็ตที่หลากหลายกว่า ตามด้วยกราฟสีน้ำเงินที่ใช้ภาพขนาด 416 มีคะแนน 0.002981 สุดท้ายกราฟสีเลือดหมูมีคะแนน 0.004121 เมื่อดูที่การใช้งานหน่วยประมวลผล ภาพที่ 11B พบว่าแบบจำลองที่ใช้ภาพ 320 (สีเขียว) มีการใช้งานหน่วยประมวลผลน้อยมาก ซึ่งใกล้เคียงกับแบบจำลองที่ยังไม่ได้เพิ่มภาพจากอินเทอร์เน็ต (สีชมพู) ตามด้วยแบบจำลองที่ใช้ภาพ 416 และ 640 ตามลำดับ



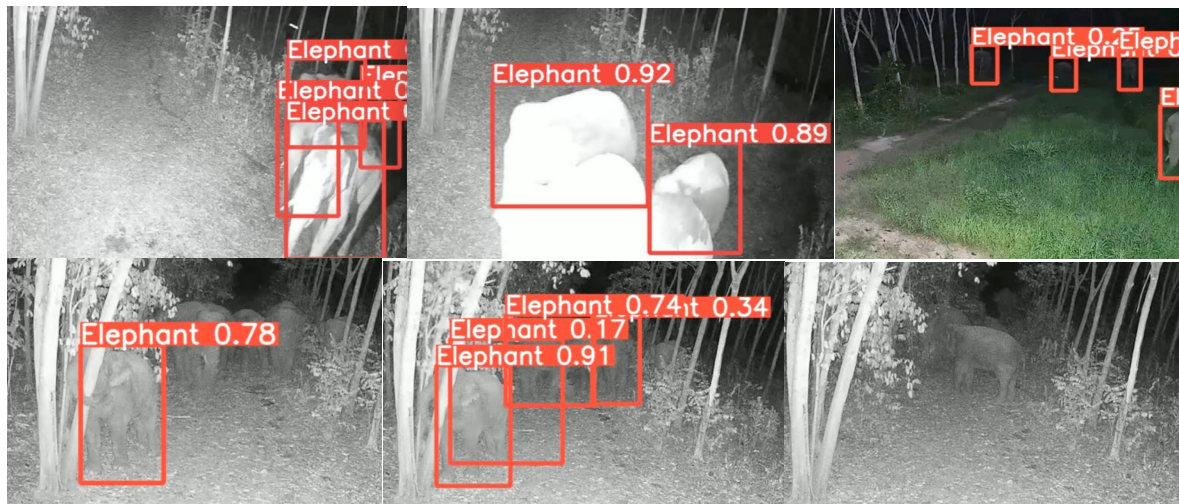
ภาพที่ 10 กราฟเปรียบเทียบการลดขนาดรูปภาพที่เพิ่มรูปจากอินเทอร์เน็ตเข้าไปกับ แบบจำลองที่ยังไม่ได้เพิ่ม



ภาพที่ 11 กราฟแสดงการแก้ปัญหาของแบบจำลองเมื่อเจอข้อมูลที่ไม่เคยพบกับการใช้งานหน่วยประมวลผล

ดังนั้นงานวิจัยนี้จึงเลือกใช้แบบจำลอง Yolov5s ที่ใช้การสับทอไฮเปอร์พารามิเตอร์ VOC และขนาดแบทเท่ากับ 32 ประมวลผลภาพที่ 320px หลังจากติดตั้งระบบ 1 สัปดาห์ พบว่ากล้องวงจรปิดทั้ง 4 ตัวในจุดติดตั้งสามารถตรวจจับข้างได้ทันทีที่ข้างเดินผ่านกล้อง ซึ่งใกล้เคียงกับการมองเห็นของมนุษย์ที่สามารถระบุได้ว่าเป็นข้างจากภาพที่ 12 (ซ้ายบน) ระบบสามารถตรวจจับข้างได้ 4 เชือก ตรงตามที่มนุษย์เห็น เมื่อเดินเข้ามากลางภาพระบบตรวจจับได้ 2 เชือก (กลางบน) เนื่องจากแสงตกกระทบบนตัวข้างทำให้แยกไม่ออก ระบบสามารถตรวจจับข้างที่อยู่ไกลออกไปในมุมมืดได้ (ขวาบน) ข้างหลบหลังต้นไม้ ยืนซ้อนกันก็สามารถแจ้งเตือนได้ (ซ้ายและกลางล่าง) แต่เมื่อข้าง

อยู่หนึ่ง ๆ ระบบจะไม่ดักจับ (ขว้าง) ข้อสังเกตระบบสามารถแจ้งเตือนได้ดีถึงแม้จะเป็นภาพขนาดเล็ก ซึ่งต่างกับการมองเห็นของมนุษย์ที่รับรู้ว่าเป็นช้างได้เมื่อมองที่จอภาพขนาดใหญ่



ภาพที่ 12 ช้าง 4 เชือกเดินเข้ามาหากล้องจากมุมมืด และภาพช้างระยะไกลในมุมมืด ช้างยืนซ้อนและอยู่หนึ่ง

อภิปรายผล

โครงข่ายประสาทเทียมแบบคอนโวลูชันสามารถใช้ในงานเฝ้าระวัง เพื่อลดการเผชิญหน้ากันระหว่างช้างกับคนได้ ชุดข้อมูลที่ดีทำให้การประมวลผลด้วยโครงข่ายประสาทเทียมแบบคอนโวลูชัน สามารถฝึกให้มีความแม่นยำสูงกับหน่วยประมวลผลขนาดเล็กได้ ด้วยขนาดช้างที่ใหญ่กับพื้นหลังที่แตกต่าง และการเคลื่อนไหวที่ แม้สภาพแสง และพฤติกรรมของช้างที่ชอบออกหากินเวลากลางคืน แบบจำลองขนาดเล็กสามารถทำงานได้ดี แม้ในสภาพแสงน้อยได้ โดยมีความแม่นยำเฉลี่ย $mAP_{0.5:0.95}$ อยู่ที่ 0.7559 $mAP_{0.5}$ อยู่ที่ 0.9795 และค่า Recall อยู่ที่ 0.957 มีค่าความผิดพลาดการทำนายผลอยู่ที่ 0.002531 และการสูญเสียการถดถอยขอบเขต (box_loss) อยู่ที่ 0.0071 เมื่อเทียบกับการมองด้วยสายตามนุษย์พบว่า การแจ้งเตือนช้างถูกต้อง แต่บางครั้งอาจจะช้ากว่าที่คนเห็นเล็กน้อย โดยเฉพาะช้างในมุมมืดที่อยู่ห่างไกลออกไป ดังนั้นงานวิจัยนี้จึงสามารถช่วยลดการทำงานของมนุษย์ลงได้ด้วยการแจ้งเตือนผ่านแอปพลิเคชันไลน์ ทำให้ผู้ที่เกี่ยวข้องสามารถรับทราบข้อมูลเพื่อวางแผนผลักดันช้างได้รวดเร็ว จากงานวิจัยของ Kaewtathanawatthana and Brorewongtrakhul (2020) ชี้ให้เห็นว่างานวิจัยนี้สามารถแก้ไขปัญหาในระยะต้นที่ช่วยให้มนุษย์สามารถอาศัยอยู่ร่วมกับช้างโดยลดการเผชิญหน้าลงได้

ข้อเสนอแนะ

1. แบบจำลองที่ใช้มีขนาดเล็กเหมาะกับหน่วยประมวลผลขนาดเล็ก ซึ่งมีความแม่นยำไม่มาก ถ้ามีชุดข้อมูลการเรียนรู้ที่ดี ตรงตามเป้าหมายที่ระบบต้องเจอ จะทำให้สามารถแก้ปัญหาดังกล่าวได้ดี ถ้าต้องเพิ่มคลาสอื่น เช่น คน หรือยานพาหนะ อาจจะต้องเพิ่มข้อมูลให้ครอบคลุม ถ้าความแม่นยำไม่เป็นที่พอใจอาจต้องเปลี่ยนแบบจำลองให้ใหญ่ขึ้นเพื่อสกัดคุณลักษณะเฉพาะของภาพนั้นออกมา

2. แบบจำลองนี้ถูกฝึกฝนด้วยข้อมูลภาพในพื้นที่เฝ้าระวัง เมื่อนำไปใช้ในพื้นที่อื่น พบว่ามีความผิดพลาดเกิดขึ้น เช่น เห็นเสาไฟฟ้าหรือใบไม้ยับ ระบบคิดว่าเป็นช้าง ดังนั้นควรที่จะรวบรวม ชุดข้อมูลรูปภาพช้างในจุดติดตั้งอื่นใหม่เพื่อฝึกฝนเพิ่มเติมเสริมความแข็งแกร่งให้กับแบบจำลอง

3. ระบบดังกล่าวต้องอาศัยไฟฟ้าและอินเทอร์เน็ต ดังนั้นถ้าในพื้นที่ไม่มีไฟฟ้าอาจต้องอาศัยพลังงานแสงอาทิตย์ รวมถึงสัญญาณอินเทอร์เน็ตมือถือ

กิตติกรรมประกาศ

ขอขอบคุณผู้ใหญ่บ้านหมู่ 10 และหมู่ 12 ตำบลทุ่งควายกิน อำเภอแกลง จังหวัดระยอง อาสาสมัครเฝ้าระวัง และผลักดันช้างป่า ที่อนุเคราะห์ข้อมูล และอำนวยความสะดวกในการลงพื้นที่ศึกษาเก็บข้อมูลการวิจัยและบริการวิชาการในครั้งนี้ รวมถึงบริษัท 3bb บริษัท Things Analytic และมหาวิทยาลัยราชภัฏรำไพพรรณีที่ได้สนับสนุนทั้งสัญญาณอินเทอร์เน็ต อุปกรณ์ และงบประมาณ

เอกสารอ้างอิง

- Duporge, L., Isupova, O., Reece, S., Macdonald, D., and Wang, T. (2020). Using very-high-resolution satellite imagery and deep learning to detect and count African elephants in heterogeneous landscapes. *Remote Sensing in Ecology and Conservation*, 7, 369-381.
- Girshick, R., Donahue, J., Darrell, T., and Malik, J. (2014). Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. *The Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 23-28 June 2014, at USA, 580-587.
- Gur-Arie, L. (2022). *The practical guide for Object Detection with YOLOv5 algorithm* [Online]. Retrieved December 11, 2021, from: <https://towardsdatascience.com/the-practical-guide-for-object-detection-with-yolov5-algorithm-74c04aac4843>.
- Hongboonmee, N., and Sitthichokchaisiri, T. (2021). The Development of Application for Analysis of Eye Health with Image Recognition Using Deep Learning Technique. *Sripatum Review of Science and Technology*, 13(1), 7-21. (in Thai)
- Jagannathan, S., Kumar, V., and Meganathan, D. (2019). Design and implementation of in-situ human-elephant conflict management system. *Journal of Intelligent & Fuzzy Systems*, 3, 2005-2013.
- Kaewtathanawatthana, P., and Broewongtrakul, S. (2020). Guidelines for Solving Conflicts between human and Wild Elephants A case study: Phawa Sub-district, Kaeng Hang Maeo District, Chanthaburi. *Journal of Humanities and Social Sciences*, 12(2), 113-127. (in Thai)
- Kumar, V., Recupero, D., Riboni, D., and Helaoui, R. (2020). Ensembling Classical Machine Learning and learning approaches for Morbidity Identification from Clinical Notes. *Institute of Electrical and Electronics Engineers (IEEE 2021)*, 9, 7107-7126.

- Liu, S., Qi, L., Qin, H., Shi, J., and Jia, J. (2018). Path Aggregation Network for Instance Segmentation. *The Proceedings of the 2018 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 18-23 June 2018 at Salt Lake City, Utah, USA, 8759-8768.
- Madheswaran, K.M.S., Veerappan, K., and Kumar, V. (2019). Region Based Convolutional Neural Network for Human-Elephant Conflict Management System. *The Proceedings of the International Conference on Computational Intelligence in Data Science (ICCIDS 2019)*, 21-23 February 2019 at Chennai in India. (IEEE 2019), 1, 1-5.
- Noonto, B., Savini, C., Srikrachang, M., and Maneesrikum, C. (2018). *The Voice of the People, the Voice of the Wild Elephants: The Trend of Community Responses to the Management of Human-Elephant Conflict in Thailand*. Bangkok: Community-Based Research Division, The Thailand Research Fund (TRF). (in Thai)
- Sanguansat, P. (2019). *Artificial Intelligence with Machine Learning, AI can be created using machine learning*. Nonthaburi: ICD-premier. (in Thai)
- Solawetz, J. (2020). *yolov5 improvements and evaluation*. [Online]. Retrieved January 22, 2022, from: <https://blog.roboflow.com/yolov5-improvements-and-evaluation>.
- Ultralytics. (2021). *yolov5*. [Online]. Retrieved January 22, 2022, from: <https://github.com/ultralytics/yolov5>
- Unzueta, D. (2021). *Convolutional Layers vs Fully Connected Layers*. [Online]. Retrieved January 12, 2022, from: www.towardsdatascience.com/convolutional-layers-vs-fully-connected-layers
- Yamashita, R., Nishio, M., Kinoshita, R., and Togashi, K. (2018). Convolutional neural networks: an overview and application in radiology. *Insights into Imaging*, 9(4), 611–629.
- Zhou, X., Gong, W., Fu, W., and Du, F. (2017). Application of deep learning in object detection. *The Proceedings of the IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS 2017)*, 10-13 December 2017 at COEX Convention Centre (Seoul) in South Korea, 631-634.