การระบุสายพันธุ์แคคตัสด้วยโมเดลการเรียนรู้เชิงลึกร่วมกับเทคนิคตรวจจับวัตถุ
Automatic Cactus Species Identification using Deep Learning Model and Object Detection Techniques

จักรินทร์ สันติรัตนภักดี[1]* และนิชาภัทร ตุลาธาร[1]
Chakkarin Santirattanaphakdi[1]* and Nichapat Tulathan[1]
[1]สาขาวิชาเทคโนโลยีธุรกิจดิจิทัล คณะบริหารธุรกิจ มหาวิทยาลัยวงษ์ชวลิตกุล
[1]Department of Digital Business Technology, Faculty of Business Administration, Vongchavalitkul University
*Corresponding author: chakkarin_san@vu.ac.th
Received: December 18, 2024
Revised: February 17, 2025
Accepted: February 25, 2025

บทคัดย่อ

การระบุสายพันธุ์แคคตัสสำหรับผู้เริ่มต้นนับเป็นเรื่องที่ท้าทาย อันเนื่องมาจากความหลากหลายของสายพันธุ์และความคล้ายคลึงของคุณลักษณะเฉพาะที่ใกล้เคียงกัน ซึ่งอาจส่งผลต่อความถูกต้องในการจำแนก การวิจัยนี้มีวัตถุประสงค์ คือ (1) เพื่อพัฒนาชุดข้อมูลรูปภาพและคำอธิบายรูปภาพสายพันธุ์แคคตัส (2) เพื่อพัฒนาตัวแบบการระบุสายพันธุ์แคคตัสด้วยโมเดลการเรียนรู้เชิงลึกและเทคนิคตรวจจับวัตถุ และ (3) เพื่อประเมินความถูกต้องการระบุสายพันธุ์แคคตัส การดำเนินงานเริ่มจากการรวบรวมรูปภาพแคคตัส 10 สายพันธุ์ที่ได้รับความนิยมจากกลุ่มผู้เพาะเลี้ยงแคคตัสเชิงพาณิชย์ในจังหวัดนครราชสีมา พร้อมสร้างคำอธิบายรูปภาพในรูปแบบ XML รวมทั้งสิ้น 1,508 รูปภาพ โดยใช้โมเดล MobileNetV2 ที่ได้รับการถ่ายโอนการเรียนรู้มาฝึกฝนบนชุดข้อมูลฝึกฝน และใช้ออปติไมเซอร์เป็น RMSprop พร้อมกำหนดฟังก์ชันการสูญเสียที่เหมาะสม เพื่อปรับปรุงความแม่นยำของโมเดลตามกระบวนการแพร่กระจายย้อนกลับในการปรับค่าถ่วงน้ำหนักของโมเดลให้เหมาะสมกับคุณลักษณะสำคัญของแคคตัสในแต่ละสายพันธุ์ ร่วมกับเทคนิคการตรวจจับวัตถุแบบขั้นตอนเดียวที่ใช้การตรวจจับ และการจำแนกประเภทของวัตถุในภาพที่รวดเร็วและมีประสิทธิภาพสูง ผลลัพธ์จากการฝึกฝนโมเดลพบว่า ค่าความถูกต้องเฉลี่ยบนชุดข้อมูลฝึกฝนอยู่ที่ 92.64% และบนชุดข้อมูลตรวจสอบอยู่ที่ 91.05% ซึ่งบ่งชี้ว่า โมเดลสามารถเรียนรู้รูปแบบของข้อมูลได้อย่างมีประสิทธิภาพ และไม่มีสัญญาณของการเกิด Overfitting หรือ Underfitting อย่างมีนัยสำคัญ เช่นเดียวกับค่าความแม่นยำเฉลี่ยและค่าการเรียกคืนเฉลี่ยในการตรวจจับวัตถุขนาดใหญ่ที่สูงถึง 0.991 และ 0.993 ตามลำดับ อย่างไรก็ดี ยังคงพบข้อจำกัดในการตรวจจับวัตถุขนาดกลางและขนาดเล็ก ซึ่งยังต้องได้รับการปรับปรุงในอนาคต และเมื่อพิจารณาผลการประเมินในสถานการณ์จริงกับชุดข้อมูลที่ไม่เคยเห็นมาก่อน พบว่า ความถูกต้องเฉลี่ยอยู่ที่ 83.33% โดยทำงานได้ดีในกรณีที่มีวัตถุหนึ่งชิ้นที่สามารถจำแนกได้เป็นหลายคลาส และในกรณีที่มีหลายวัตถุในคลาสเดียวกัน แต่ความถูกต้องลดลงเมื่อรูปภาพมีวัตถุหลายชนิดจากหลายคลาส

คำสำคัญ: การจำแนกรูปภาพ สายพันธุ์แคคตัส การเรียนรู้เชิงลึก การตรวจจับวัตถุ

Abstract

Identifying cactus species can be challenging for beginners due to the diversity of species and the similarities in distinctive features, which may affect the accuracy of classification. This research aims to 1) develop a dataset of images and image descriptions for cactus species, 2) develop a cactus species identification model using deep learning and object detection techniques, and 3) evaluate the accuracy of cactus species identification. The work involves collecting images of 10 popular cactus species from commercial cactus growers in Nakhon Ratchasima province, along with creating image descriptions in XML format, totaling 1,508 images. The MobileNetV2 model, which has been pre-trained with transfer learning, was used to train on the training dataset. The optimizer used was RMSprop, with an appropriate loss function to improve model accuracy during backpropagation for adjusting the model weights to suit the key features of each cactus species. Additionally, a single-shot object detection technique was employed for fast and efficient object detection and classification. The results from training the model showed an average accuracy of 92.64% on the training dataset and 91.05% on the validation dataset, indicating that the model was able to effectively learn data patterns without significant signs of overfitting or underfitting. The average precision and average recall for large object detection were high, at 0.991 and 0.993, respectively. However, limitations in detecting medium and small objects were observed and require future improvement. When evaluating the model in real-world scenarios with unseen datasets, the average accuracy was 83.33%. The model performed well in cases where a single object could be classified into multiple classes or when multiple objects were in the same class. However, accuracy decreased when images contained multiple objects from different classes.

Keywords: image classification, cactus species, deep learning, object detection

บทนำ

แคคตัส (Cactus) เป็นหนึ่งในไม้ประดับที่ได้รับความนิยมจากทั่วโลกด้วยเอกลักษณ์เฉพาะตัวที่หลากหลายในแต่ละสายพันธุ์ ตลอดจนคุณสมบัติที่ทนทานต่อสภาพอากาศแห้งแล้งและง่ายต่อการดูแลรักษา จึงเป็นที่แพร่หลายในหลายประเทศ เช่น เม็กซิโก สหรัฐอเมริกา และประเทศในแถบเมดิเตอร์เรเนียน (Anderson, 2001) ในปี 2020 ตลาดแคคตัสมีมูลค่าประมาณ 2.06 พันล้านดอลลาร์สหรัฐ และคาดว่าจะสูงถึงประมาณ 3.21 พันล้านดอลลาร์สหรัฐภายในปี 2027 โดยมีอัตราการเติบโตเฉลี่ยประมาณร้อยละ 6.7 ต่อปี (Verified Market Research, 2021) โดยเฉพาะในช่วงที่เกิดการแพร่ระบาดของโควิด-19 ซึ่งได้กระตุ้นความนิยมในการจัดสวนและการปลูกพืชในบ้าน สอดคล้องกับตลาดแคคตัสทั่วโลกมีมูลค่า 5.12 พันล้านดอลลาร์สหรัฐในปี 2023 และคาดว่าจะสูงถึง 21.03 พันล้านดอลลาร์สหรัฐภายในปี 2032 โดยมีอัตราการเติบโตเฉลี่ยต่อปี (CAGR) ที่ 17% ในช่วงเวลาที่คาดการณ์ไว้ และมีแนวโน้มเพิ่มขึ้นเรื่อย ๆ ในอนาคต การเติบโตในตลาดนี้เป็นผลมาจากการเปลี่ยนแปลงวิถีชีวิตและการเข้าถึงสินค้าออนไลน์ที่เพิ่มมากขึ้น ทำให้ผู้เพาะพันธุ์แคคตัสและผู้ค้าปลีกสามารถขยายตลาดได้อย่างมีประสิทธิภาพ (Business Research Insights, 2023)

อย่างไรก็ดี ความหลากหลายของสายพันธุ์แคคตัสที่มีมากกว่า 1,500 สายพันธุ์ทั่วโลก (Nobel, 2002) ทำให้การระบุสายพันธุ์สำหรับผู้เริ่มต้นเป็นเรื่องที่ท้าทาย อันเนื่องมาจากคุณลักษณะเฉพาะของแต่ละสายพันธุ์ เช่น รูปทรง หนาม สีของลำต้น และขนาดที่แตกต่างกัน ตลอดจนการเติบโตในสภาพแวดล้อมที่แตกต่างกันยังทำให้คุณลักษณะของสายพันธุ์เดียวกันเปลี่ยนแปลงไป จะเห็นได้ว่าการจำแนกสายพันธุ์แคคตัสที่ไม่ถูกต้องอาจส่งผลเสียต่อการเพาะปลูก เนื่องจากแคคตัสแต่ละสายพันธุ์ต้องการการดูแลที่แตกต่างกัน เช่น ปริมาณน้ำ แสงแดด และสภาพดินที่เหมาะสม ซึ่งส่งผลต่อการเจริญเติบโตของแคคตัส หรืออาจกระทบต่อคุณภาพของผลผลิตได้

จากที่กล่าวมา จึงเป็นที่มาของการนำเทคโนโลยีเข้ามาช่วยในการระบุสายพันธุ์แคคตัส โดยความสำคัญการจำแนกสายพันธุ์ไม่เพียงแต่ช่วยให้สามารถเลือกซื้อแคคตัสสายพันธุ์ที่ตรงกับความสนใจเท่านั้น แต่ยังช่วยแบ่งปันความรู้เกี่ยวกับคุณลักษณะเฉพาะของแต่ละสายพันธุ์ ทำให้ผู้ที่สนใจสามารถเข้าถึงข้อมูล และสร้างความเข้าใจที่ถูกต้องเกี่ยวกับพืชชนิดนี้ ดังนั้นการระบุสายพันธุ์แคคตัสด้วยโมเดลการเรียนรู้เชิงลึกร่วมกับเทคนิคตรวจจับวัตถุนั้นจะสามารถลดความผิดพลาดที่อาจเกิดจากการจำแนกโดยมนุษย์ และประหยัดเวลาในการค้นหาข้อมูล นอกจากนี้ยังเป็นการตอบสนองต่อความต้องการของกลุ่มผู้ชื่นชอบแคคตัสที่กำลังได้รับความนิยมเพิ่มขึ้นอย่างมากในทุกเพศทุกวัย ตลอดจนเป็นการประยุกต์ใช้เทคโนโลยีทางการเกษตรเชิงพาณิชย์ อันจะเป็นแนวทางสำหรับการวิจัยในอนาคต

วัตถุประสงค์การวิจัย

1. เพื่อพัฒนาชุดข้อมูลรูปภาพ และคำอธิบายรูปภาพสายพันธุ์แคคตัส

2. เพื่อพัฒนาตัวแบบการระบุสายพันธุ์แคคตัสด้วยโมเดลการเรียนรู้เชิงลึกและเทคนิคตรวจจับวัตถุ

3. เพื่อประเมินความถูกต้องการระบุสายพันธุ์แคคตัส

แนวคิดทฤษฎีที่เกี่ยวข้อง

1) การเรียนรู้เชิงลึก (deep learning) ถูกพัฒนามาจากการเรียนรู้ของเครื่อง (machine learning) ที่มุ่งเน้นการสร้างและฝึกฝนโครงข่ายประสาทเทียม (neural networks) เพื่อให้สามารถเรียนรู้ข้อมูลผ่านโครงข่ายแต่ละชั้นที่มีความซับซ้อนแบบส่งข้อมูลไปข้างหน้า (forward propagation) โดยข้อมูลจะถูกป้อนเข้าไปในเครือข่ายประสาทเทียมผ่านหลายเลเยอร์ โดยจะคำนวณผลลัพธ์ขั้นสุดท้าย ร่วมกับการแพร่กระจายย้อนกลับ (backpropagation) เพื่อคำนวณค่าการสูญเสีย (loss) และปรับน้ำหนักของโครงข่ายผ่านกระบวนการ Gradient Descent เพื่อทำให้ค่าการสูญเสียลดลงและทำให้โมเดลสามารถทำนายผลลัพธ์ที่ใกล้เคียงกับค่าจริงมากที่สุดในครั้งถัดไป (Goodfellow, Bengio & Courville, 2016) อย่างไรก็ดี การปรับน้ำหนักในโครงข่ายจะทำการเรียนรู้ทีละขั้น โดยแต่ละรอบของการฝึกจะทำให้เครือข่ายประสาทเทียมสามารถเข้าใจรูปแบบและคุณลักษณะของข้อมูลได้ดีขึ้น ซึ่งกระบวนการนี้จะถูกทำซ้ำหลายรอบจนกว่าโมเดลจะมีความแม่นยำสูงสุดในการพยากรณ์ผลลัพธ์

2) โครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional Neural Network: CNN) เป็นเทคนิคสำคัญในสาขาการเรียนรู้เชิงลึกที่ได้รับการพัฒนาเพื่อประมวลผลข้อมูลที่มีโครงสร้างแบบตาราง จึงถูกนำมาในงานวิเคราะห์ภาพอย่างแพร่หลาย (Goodfellow, Bengio & Courville, 2016) จากความสามารถในการเรียนรู้คุณลักษณะสำคัญของข้อมูลโดยอัตโนมัติผ่านการสกัดคุณลักษณะ (feature extraction) ซึ่งช่วยลดความยุ่งยากในการพัฒนาโมเดล และยังช่วยเพิ่มความแม่นยำของการประมวลผลภาพ โครงสร้างประกอบด้วยชั้นต่าง ๆ ได้แก่ ชั้นคอนโวลูชัน (Convolution Layer) ซึ่งทำหน้าที่ค้นหาและสกัดคุณลักษณะ เช่น เส้นขอบหรือรูปร่างของวัตถุ จากนั้นชั้นพูลลิง (Pooling Layer) จะทำหน้าที่ลดขนาดของผังคุณลักษณะ (feature map) และเพิ่มความทนทานต่อการแปรผันของข้อมูล เช่น การหมุนหรือเปลี่ยนตำแหน่งของวัตถุ ในขั้นสุดท้ายข้อมูลจะถูกแปลงเป็นเวกเตอร์ในชั้นแบนราบ (flattening layer) และนำเข้าสู่ชั้นเชื่อม

ต่อสมบูรณ์ (fully connected layer) เพื่อตัดสินใจที่มีผลลัพธ์เป็นคำตอบตามที่คาดหวัง

อย่างไรก็ตาม การพัฒนาโมเดลที่มีประสิทธิภาพสูงต้องอาศัยข้อมูลและทรัพยากรจำนวนมาก ด้วยเหตุนี้จึงมีการพัฒนาโมเดลที่ได้รับการฝึกฝนล่วงหน้า (pre-trained model) เพื่อนำมาใช้พัฒนาโมเดลสำหรับงานเฉพาะทางด้วยความสามารถในการถ่ายโอนความรู้ (transfer learning) จากการฝึกฝนบนชุดข้อมูลขนาดใหญ่สามารถนำไปปรับใช้กับงานใหม่ได้ง่าย โดยไม่ต้องสร้างโมเดลใหม่ทั้งหมด ร่วมกับการปรับแต่งโมเดลเพิ่มเติม (fine-tuning) ให้เหมาะสมกับชุดข้อมูลด้วยการปรับค่าน้ำหนักบางส่วนของโมเดลหรือฝึกฝนเฉพาะบางชั้นของโครงข่าย ทำให้สามารถใช้งานได้อย่างมีประสิทธิภาพแม้ในสภาพแวดล้อมที่แตกต่างจากชุดข้อมูลเดิม ปัจจุบันมีหลายโมเดลที่ได้รับการฝึกฝนล่วงหน้าที่นำสถาปัตยกรรมของโครงข่ายประสาทเทียมแบบคอนโวลูชันมาพัฒนาเพื่อจำแนกรูปภาพ ในเบื้องต้นคณะผู้วิจัยได้คัดเลือกโมเดลเพื่อทดสอบกับชุดข้อมูลรูปภาพต่าง ๆ พบว่า ทั้ง 5 โมเดล ประกอบด้วย ResNet34 VGG16 MobileNetV2 DenseNet121 และ SqueezeNet นั้นมีค่าความถูกต้องในการจำแนกข้อมูลที่สูงในระดับใกล้เคียงกัน อย่างไรก็ดี แต่ละโมเดลนั้นมีความแตกต่างกันอย่างชัดเจน ทั้งในด้านความเร็วและขนาดของโมเดล ดังนั้นในงานวิจัยนี้จึงเลือกใช้โมเดล MobileNetV2 เป็นโครงสร้างพื้นฐาน (baseline model) ร่วมกับการตรวจจับวัตถุแบบขั้นตอนเดียว เนื่องจากมีประสิทธิภาพสูงในการประมวลผลภาพ โดยมีจำนวนพารามิเตอร์ที่น้อยและสามารถทำงานได้อย่างรวดเร็วบนอุปกรณ์ที่มีทรัพยากรจำกัด เช่น อุปกรณ์พกพาหรือระบบที่ต้องการการตอบสนองแบบเรียลไทม์ เป็นต้น

3) MobileNetV2 เป็นหนึ่งในโมเดลที่ได้รับการฝึกฝนล่วงหน้าสำหรับการจำแนกรูปภาพที่ได้รับความนิยมในปัจจุบัน (Sandler, Howard, Zhu, Zhmoginov & Chen, 2018) โดยเฉพาะการประมวลผลบนอุปกรณ์ที่มีข้อจำกัดด้านทรัพยากรด้วยเทคนิค Depthwise separable convolutions ซึ่งช่วยลดจำนวนพารามิเตอร์และทรัพยากรในการคำนวณ แบ่งเป็นการประมวลผลคุณลักษณะด้วยฟิลเตอร์ที่แยกกัน และการรวมผลลัพธ์ด้วยฟิลเตอร์ทั่วไป อีกทั้งโมเดลนี้ยังถูกเรียนรู้ล่วงหน้าบนชุดข้อมูลขนาดใหญ่ เช่น ImageNet ก่อนที่จะนำไปปรับแต่งเพื่อใช้กับข้อมูลเฉพาะด้าน และการเปลี่ยนฟังก์ชันกระตุ้น (activated function) จากเดิมคือ ReLU เป็น ReLU6 ที่ควบคุมค่าเอาต์พุตได้ดีกว่า นอกจากนี้ยังใช้แนวคิด Linear bottlenecks เพื่อปรับปรุงการไหลของข้อมูลภายในให้สามารถจัดการกระจายฟีเจอร์อย่างเหมาะสม จึงสามารถเรียนรู้จากข้อมูลที่มีความซับซ้อนได้ดีขึ้นโดยไม่สูญเสียคุณลักษณะที่สำคัญ

4) การตรวจจับวัตถุ (object detection) เป็นหนึ่งในศาสตร์ด้านวิทยาการคอมพิวเตอร์ที่เกี่ยวข้องกับการประมวลผลภาพและคอมพิวเตอร์วิทัศน์ ซึ่งผสมผสานระหว่างการจำแนกรูปภาพ (image classification) และการระบุตำแหน่งของวัตถุ (localization) ในรูปภาพ (Zhang, Yang, Han, Wang & Gao, 2014) แนวคิดพื้นฐานมาจากการใช้โครงข่ายประสาทเทียมเชิงลึกที่เลียนแบบการทำงานของสมองมนุษย์ในการประมวลผลภาพผ่านชั้นต่าง ๆ ของโครงข่าย เพื่อดึงคุณลักษณะสำคัญตั้งแต่รายละเอียดย่อยจนถึงคุณลักษณะเด่นของวัตถุ เช่น รูปทรง สี และพื้นผิวที่จำเป็นสำหรับการระบุวัตถุ โดยจะให้ผลลัพธ์ในรูปแบบของ Bounding box หรือกรอบสี่เหลี่ยมที่ล้อมรอบวัตถุที่ตรวจพบ พร้อมกับระบุประเภทของวัตถุ (object classes) และคลาสความน่าจะเป็นจากการจำแนก (class probability) ของวัตถุที่ตรวจจับ

การตรวจจับวัตถุแบ่งออกเป็นสองประเภท ได้แก่ โมเดลตรวจจับวัตถุแบบขั้นตอนเดียว (single-stage detectors) ได้แก่ YOLO: You Only Look Once (Redmon, Divvala, Girshick & Farhadi, 2016) และการตรวจจับวัตถุในครั้งเดียว (Single Shot Detector: SSD) (Liu et al., 2016) ที่จะตรวจจับและจำแนกวัตถุในภาพพร้อมกันภายในขั้นตอนเดียว ทำให้มีความเร็วสูงและเหมาะสำหรับการใช้งานที่ต้องการการประมวลผลแบบเรียลไทม์ และโมเดลตรวจจับวัตถุแบบสองขั้นตอน (Two-stage detectors) เช่น R-CNN และ Faster-RCNN (Girshick, 2015) ที่จะตรวจหาพื้นที่ที่น่าสนใจก่อน (region proposal) จากนั้นจึงทำการตรวจจับและจำแนกวัตถุในพื้นที่นั้น อย่างไรก็ตาม

แม้การตรวจจับวัตถุแบบขั้นตอนเดียวอาจมีข้อจำกัดด้านความแม่นยำเมื่อเทียบกับการตรวจจับวัตถุแบบสองขั้นตอนที่มีประสิทธิภาพสูงกว่าในการแยกแยะรายละเอียดย่อยของวัตถุ แต่ต้องแลกกับการใช้เวลาและทรัพยากรการประมวลผลมากขึ้น ดังนั้นโมเดลตรวจจับวัตถุแบบขั้นตอนเดียวจึงมักถูกนำไปประยุกต์ใช้ในหลายด้าน

5) ตัวตรวจจับวัตถุแบบหลายกล่องในครั้งเดียว (single shot multi box detector) (Liu et al., 2016) เป็นหนึ่งในโมเดลที่ได้รับการฝึกฝนล่วงหน้าสำหรับการตรวจจับวัตถุแบบขั้นตอนเดียวสำหรับการตรวจจับวัตถุที่มีความสามารถในการระบุตำแหน่งและจำแนกประเภทของวัตถุภายในภาพที่ได้รับความนิยมอย่างสูง โดยใช้โครงข่ายประสาทเทียมแบบคอนโวลูชันเป็นโครงสร้างพื้นฐานและเป็นตัวดำเนินการตรวจจับวัตถุในภาพแบบขั้นตอนเดียว

การตรวจจับวัตถุในครั้งเดียวใช้เทคนิคตรวจจับแบบหลายกล่อง (multi box) ซึ่งเป็นกลไกที่สามารถคาดการณ์พิกัดของกรอบล้อมรอบวัตถุ (bounding boxes) และคะแนนความน่าจะเป็นของคลาสเป้าหมายได้โดยตรง จากการใช้ Default boxes หรือกล่องสมมุติในหลายอัตราส่วน (aspect ratios) ซึ่งช่วยลดความซับซ้อนของกระบวนการคัดเลือกกล่องเป้าหมาย (anchor boxes) และสามารถปรับปรุงความแม่นยำผ่านการใช้ฟีเจอร์หลายระดับ (multi-scale feature maps) ที่ได้จากชั้นคอนโวลูชันของโครงข่าย ส่งผลให้การตรวจจับวัตถุนั้นมีความเร็วสูงในการประมวลผล อีกทั้งสามารถใช้งานได้อย่างมีประสิทธิภาพในบริบทที่ต้องการการตอบสนองแบบเรียลไทม์

6) การประยุกต์ใช้เทคโนโลยีที่ใช้ในการตรวจจับและจำแนกสายพันธุ์พืช ปัจจุบันมีการนำเสนอผ่านงานวิจัยเป็นจำนวนมาก โดยผสมหลากหลายเทคนิคและวิธีการต่าง ๆ ที่แตกต่างกัน โดยสามารถแบ่งออกเป็น 3 กลุ่มหลัก ได้แก่

6.1) การประมวลผลภาพ (image processing) ซึ่งช่วยให้สามารถวิเคราะห์คุณลักษณะวัตถุจากการแบ่งส่วนภาพ (image segmentation) เช่น การทำเธรสโฮลด์ (Thresholding) และอัลกอริทึม Watershed ที่เป็นเทคนิคที่ใช้ในการแยกส่วนของภาพออกเป็นส่วน ๆ โดยใช้การคำนวณระยะทางแบบยุคลิด (Euclidian distance) และการจัดกลุ่มแบบเคมีน (K-Means clustering) เพื่อแยกวัตถุออกจากพื้นหลัง ร่วมกับการสกัดคุณลักษณะ (feature extraction) เช่น สี รูปร่าง และพื้นผิวในการจำแนกความแตกต่างของพืชแต่ละสายพันธุ์ ตลอดจนการจับคู่คุณลักษณะ (feature matching) โดยใช้อัลกอริทึม เช่น SIFT: Scale-Invariant Feature Transform (Thomkaew & Intakosum, 2023) หรือ SURF: Speeded-Up Robust Features (Wu, Fang, Yu & Yang, 2024) ที่ช่วยระบุชนิดของพืชโดยเปรียบเทียบกับชุดข้อมูลที่กำหนดไว้ล่วงหน้า จึงเหมาะสำหรับงานที่ต้องการจำแนกพันธุ์พืชที่มีคุณลักษณะใกล้เคียงกัน

6.2) การเรียนรู้ของเครื่อง (machine learning) โดยใช้แบบจำลองทางสถิติเพื่อเรียนรู้คุณลักษณะของพืชโดยอาศัยชุดข้อมูลตัวอย่าง (Rani, Nikhila, Bhargavi & Sushma, 2024) อัลกอริทึมที่นิยมใช้ ได้แก่ ป่าแบบสุ่ม (Random Forest: RF) จากการรวมผลลัพธ์จากต้นไม้การตัดสินใจ (decision trees) เพื่อเพิ่มความแม่นยำ เพื่อนบ้านที่ใกล้ที่สุด K อันดับ (K-Nearest Neighbors: KNN) ในการคำนวณความคล้ายคลึงของพืชแต่ละภาพกับข้อมูลในฐานข้อมูล และซัพพอร์ตเวคเตอร์แมชชีน (Support Vector Machine: SVM) (Kaur & Kaur, 2019) ใช้สำหรับจำแนกภาพโดยอาศัยการหาขอบเขตของกลุ่มข้อมูล

6.3) การเรียนรู้เชิงลึก (deep learning) ได้เข้ามามีบทบาทสำคัญในการจำแนกพันธุ์พืช เนื่องจากสามารถเรียนรู้คุณลักษณะจากข้อมูลได้โดยอัตโนมัติ โดยโมเดลที่ได้รับความนิยม ได้แก่ โครงข่ายประสาทเทียมแบบคอนโวลูชัน (Arun, Sharma, Singh & Hazra, 2025) เช่น AlexNet VGG16 ResNet และ EfficientNet ที่สามารถเรียนรู้คุณลักษณะของภาพได้โดยตรง จึงสามารถวิเคราะห์และจดจำคุณลักษณะได้อย่างแม่นยำ ตั้งแต่โครงสร้างที่สามารถเรียนรู้ตั้งแต่ลักษณะทั่วไปจนถึงรายละเอียดเฉพาะของแต่ละสายพันธุ์ นอกจากนี้ การใช้โมเดล Vision in Transformers: ViT ที่สามารถเรียนรู้คุณลักษณะเชิงพื้นที่ของภาพพืชได้อย่างแม่นยำโดยไม่ต้องกำหนดคุณลักษณะล่วงหน้า ตลอดจนโมเดลที่พัฒนาเพิ่มเติม เช่น YOLO (Ariwa, Markus,

Teneke, Adamu & George, 2024) Faster R-CNN (Gong & Zhang, 2023) และตัวตรวจจับวัตถุแบบหลายกล่องในครั้งเดียว (Ali, Keller & Huang, 2023) มักถูกนำมาใช้สำหรับการตรวจจับและจำแนกวัตถุจำนวนมากในภาพ เนื่องจากสามารถทำงานได้อย่างรวดเร็วและมีประสิทธิภาพสูง

วิธีดำเนินการวิจัย

การพัฒนาตัวแบบการระบุสายพันธุ์แคคตัสด้วยการเรียนรู้เชิงลึกและเทคนิคตรวจจับวัตถุพัฒนาขึ้นด้วยภาษาไพธอน (Python) บน Google Colaboratory งานวิจัยนี้เลือกใช้โมเดล MobileNetV2 ร่วมกับการตรวจจับวัตถุแบบขั้นตอนเดียวในการระบุสายพันธุ์แคคตัสในภาพ มีรายละเอียด ดังนี้

1) การสร้างชุดข้อมูล (create a dataset) ที่มีการกำหนดป้ายกำกับประเภทของวัตถุในแต่ละภาพ รวมถึงขอบเขตของวัตถุที่ต้องการตรวจจับ เนื่องจากคุณภาพของชุดข้อมูลที่ใช้ในการฝึกฝนส่งผลโดยตรงต่อความถูกต้องของโมเดลการเรียนรู้เชิงลึก ขั้นตอนการดำเนินงานประกอบด้วย

1.1) การเก็บรวบรวมข้อมูล (data collecting) รูปภาพแคคตัส 10 สายพันธุ์ที่ได้รับความนิยมจากกลุ่มผู้เพาะเลี้ยงแคคตัสเชิงพาณิชย์ในจังหวัดนครราชสีมาที่คณะผู้วิจัยรวบรวมเองโดยใช้กล้องหลักของโทรศัพท์มือถือความละเอียด 50 ล้านพิกเซล (Megapixels) ที่สามารถเก็บรายละเอียดของวัตถุได้อย่างคมชัด โดยเฉพาะเมื่อต้องการซูมหรือครอบตัดภาพโดยไม่สูญเสียคุณภาพมากนัก ในการถ่ายภาพแคคตัสที่มีระยะห่างจากกล้องประมาณ 10-20 เซนติเมตร ในหลากหลายสภาพแวดล้อม เพื่อให้ได้ข้อมูลที่หลากหลายสำหรับการวิเคราะห์ผล ประกอบด้วยสายพันธุ์กระบองทอง คามิเน่ แมมดูไว แมมนกฮูก นิ้วทอง แมมขนแมว แมมแม่เฒ่า ยิมโน ลูกกอล์ฟยักษ์ และแอสโตร รวมทั้งสิ้น 1,508 รูปภาพที่แตกต่างกันใน 3 รูปแบบ คือ ขนาดของวัตถุในภาพ (scale variant) ทิศทางของวัตถุในภาพ (view-point variant) และแสงสว่างในภาพ (illumination) ดังภาพ 1 ตัวอย่างรูปภาพของแคคตัส 10 สายพันธุ์ที่มีความแตกต่างกัน

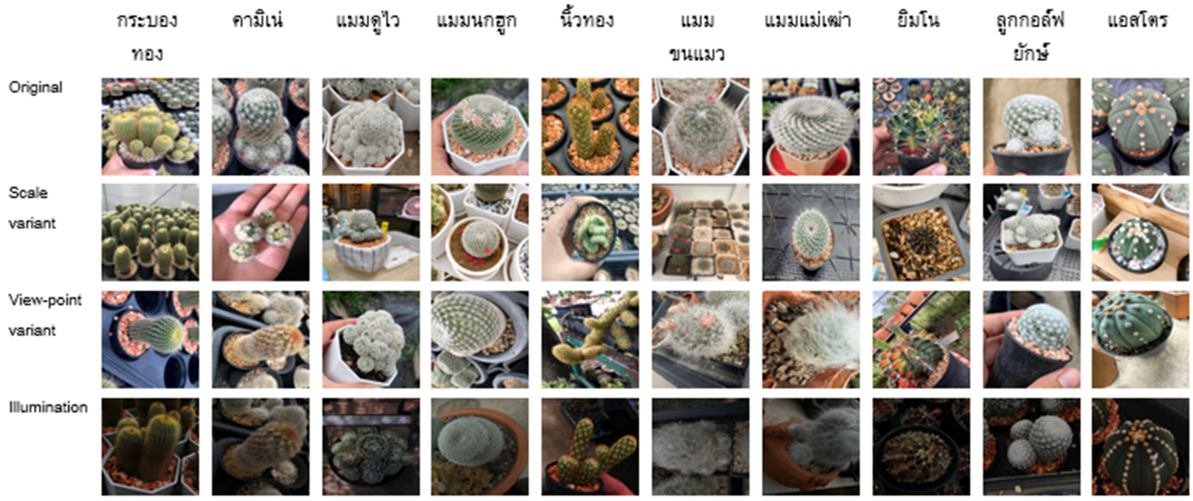
จากภาพ 1 แคคตัสเป็นพืชที่มีความหลากหลายทางสายพันธุ์ โดยแต่ละสายพันธุ์มีลักษณะเฉพาะที่แตกต่างกันทั้งในด้านรูปร่าง ขนาด สีสัน และลักษณะของหนาม ซึ่งเป็นปัจจัยสำคัญที่ช่วยในการจำแนกประเภทความแตกต่างเหล่านี้ส่งผลต่อการเพาะปลูก วิธีการดูแล และการนำไปใช้ประโยชน์ในด้านต่าง ๆ หากแบ่งตามลักษณะการเติบโตหรือชนิดของต้นที่พบบ่อยในกลุ่มแคคตัส (Barthlott & Hunt, 1993) สามารถแบ่งออกเป็น 3 กลุ่มตามลักษณะดังนี้ (1) กลุ่มที่มีลักษณะทรงพุ่มหรือกลม (globular cactus) ได้แก่ แอสโตร (Astrophytum) เป็นแคคตัสที่มีลักษณะพิเศษ คือ ไม่มีหนาม แต่มีตุ่มสีขาวกระจายทั่วลำต้น ลักษณะลำต้นเป็นทรงกลมแบน และออกดอกสีเหลืองขนาดใหญ่ที่ยอดของลำต้น เช่นเดียวกับยิมโน (Gymnocalycium) ซึ่งมีลำต้นทรงกลมขนาดใหญ่ สีเขียวเข้ม และมีหนามขนาดใหญ่เรียงตัวรอบตุ่มหนาม การเจริญเติบโตของสายพันธุ์นี้ค่อนข้างช้า แต่สามารถเติบโตจนมีขนาดใหญ่หลายสิบเซนติเมตร และมักออกดอกสีขาวหรือชมพูบริเวณยอด นอกจากนี้แมมนกฮูก (owl eye) มีลำต้นทรงกลมรี และมีลักษณะการเรียงตัวของหนามที่คล้ายกับดวงตาของนกฮูก สายพันธุ์นี้มักแตกหน่อเป็นกลุ่มและเติบโตเป็นพุ่มขนาดใหญ่ พร้อมออกดอกสีชมพูอ่อนหรือสีขาวบริเวณยอด ส่วนแมมขนแมว (mammillaria bocasana) มีจุดเด่นที่ขนสีขาวละเอียดปกคลุมทั่วลำต้น และมีหนามสีแดงบางเส้นแซมอยู่ การออกดอกของสายพันธุ์นี้มีลักษณะเป็นสีขาวนวลหรือสีชมพู (2) กลุ่มแคคตัสที่มีลักษณะยาวหรือแหลม (cylindrical cactus) ได้แก่ กระบองทอง (yellow tower) มีลักษณะเด่นคือลำต้นทรงกระบอกตั้งตรง สีเขียวอมเหลืองและปกคลุมด้วยหนามสีเหลืองละเอียด ทำให้ดูนุ่มคล้ายขน เมื่อลำต้นเติบโตเต็มที่อาจสูงถึง 1 เมตร และสามารถออกดอกสีเหลืองบริเวณยอด ส่วนคามิเน่ (Carmenae) มีลักษณะลำต้นขนาดเล็กถึงปานกลาง ปกคลุมด้วยหนามสีขาวหรือสีเหลืองทองหนาแน่น และมักออกดอกสีชมพูหรือสีขาวบริเวณยอด ลักษณะการแตกหน่อเป็นกลุ่มช่วยให้สามารถขยายพันธุ์ได้ง่าย และ (3) กลุ่มที่มีลักษณะพิเศษหรือชนิดผสม (hybrid or uncommon cactus) ได้แก่ แมมดูไว (mammillaria Duweii) เป็นสายพันธุ์ที่มีขนาดเล็ก ลำต้นทรงกลมสีเขียวเข้ม มีหนามสีขาวละเอียดและกระจุกตัวแน่นหนา จุดเด่นของสายพันธุ์นี้ คือ การออกดอกสีเหลืองอ่อนบริเวณยอด และการเจริญเติบโตค่อนข้างช้า ทำให้เหมาะสำหรับการปลูกในพื้นที่จำกัด นอกจากนั้นนิ้วทอง (lady finger) มีลำต้นเป็นทรงกระบอกยาวคล้ายเรียวนิ้ว

มือ ปกคลุมด้วยหนามสีเหลืองทองหรือสีขาว และมีลักษณะการแตกหน่อที่ง่าย ทำให้เติบโตเป็นกอขนาดใหญ่ได้ ตลอดจนแมมแม่เฒ่า (mammillaria oly lady) เป็นสายพันธุ์ที่ได้รับความนิยมอย่างมากเนื่องจากมีขนสีขาวหนาแน่นและมีหนามสีแดงแซมบางจุด การออกดอกของสายพันธุ์นี้มักเป็นสีชมพูหรือสีแดงบริเวณยอด และสุดท้ายลูกกอล์ฟยักษ์ (mammillaria humboldtii elegans) เป็นแคคตัสที่มีลำต้นทรงกลมขนาดใหญ่ สีเขียวเข้ม และมีหนามขนาดใหญ่เรียงตัวรอบตุ่มหนาม การเจริญเติบโตของสายพันธุ์นี้ค่อนข้างช้า แต่สามารถเติบโตเป็นขนาดใหญ่ได้หลายสิบเซนติเมตร และมักออกดอกสีขาวหรือชมพูบริเวณยอด จะเห็นได้ว่าแคคตัสแต่ละสายพันธุ์มีลักษณะเฉพาะที่แตกต่างกัน ทั้งในด้านรูปร่าง สีสัน ขนาด และลักษณะของหนาม ซึ่งเป็นปัจจัยสำคัญในการจำแนกสายพันธุ์และการดูแล การทำความเข้าใจความแตกต่างเหล่านี้ช่วยให้สามารถเลือกปลูกแคคตัสให้เหมาะสมกับสภาพแวดล้อมและความต้องการของผู้เลี้ยง รวมถึงสามารถใช้เป็นแนวทางในการศึกษาวิจัยเกี่ยวกับการปรับปรุงพันธุ์และการพัฒนาเทคนิคการเพาะเลี้ยงแคคตัสในอนาคต

1.2) การประมวลผลข้อมูลล่วงหน้า (data preprocessing) สำหรับการตรวจจับวัตถุจะต้องมีการปรับข้อมูลให้เหมาะสมและพร้อมใช้งานกับโมเดล ซึ่งจะรวมถึงการปรับขนาดภาพ (image resizing) ให้มีขนาด 224x224 พิกเซล เพื่อให้สามารถป้อนเข้าสู่โมเดลที่มีขนาดอินพุตคงที่ได้ และรวมถึงการปรับรูปแบบข้อมูล (format conversion) ให้อยู่ในรูปแบบที่เหมาะสม เช่น การแปลงภาพเป็นอาร์เรย์ (Array) หรือเทนเซอร์ (Tensor) เพื่อให้สามารถนำเข้าสู่โมเดลได้ก่อนจะทำการปรับค่าพิกเซลของภาพ (normalization) ให้อยู่ในช่วง 0 และ 1 ซึ่งจะช่วยให้โมเดลทำงานได้ดีขึ้น โดยการแบ่งค่าพิกเซลเท่ากับ 255 เพื่อให้ทุกพิกเซลอยู่ในช่วงที่เหมาะสม และการจัดการกับข้อมูลการระบุตำแหน่ง (bounding box processing) ที่บ่งชี้ตำแหน่งของวัตถุในภาพที่จะต้องได้รับการปรับให้สอดคล้องกับขนาดใหม่ของภาพ

1.3) การแบ่งข้อมูล (data splitting) สำหรับการตรวจจับวัตถุจะต้องพิจารณาทั้งภาพและข้อมูลการระบุตำแหน่งของวัตถุ ซึ่งแบ่งได้เป็นชุดข้อมูลฝึกฝน (training dataset) โดยจะมีข้อมูลภาพพร้อมทั้งข้อมูลการระบุตำแหน่งของวัตถุ และป้ายกำกับข้อมูล และชุดข้อมูลตรวจสอบ (validation dataset) เพื่อประเมินประสิทธิภาพของโมเดลในการตรวจจับวัตถุบนชุดข้อมูลที่ไม่เคยเห็นมาก่อน งานวิจัยนี้กำหนดอัตราส่วนชุดข้อมูลฝึกฝน และชุดข้อมูลตรวจสอบเป็น 80:20 คิดเป็น 1,206 และ 302 รูปภาพตามลำดับ ดังตาราง 1

จากตาราง 1 ตัวอย่างรูปภาพของแคคตัส 10 สายพันธุ์ที่มีความแตกต่างกันใน 3 รูปแบบ คือ ขนาดของวัตถุ ทิศทางของวัตถุ และแสงสว่างในภาพ เมื่อแบ่งออกเป็นชุดข้อมูลฝึกฝน และชุดข้อมูลตรวจสอบ ดังภาพ 2

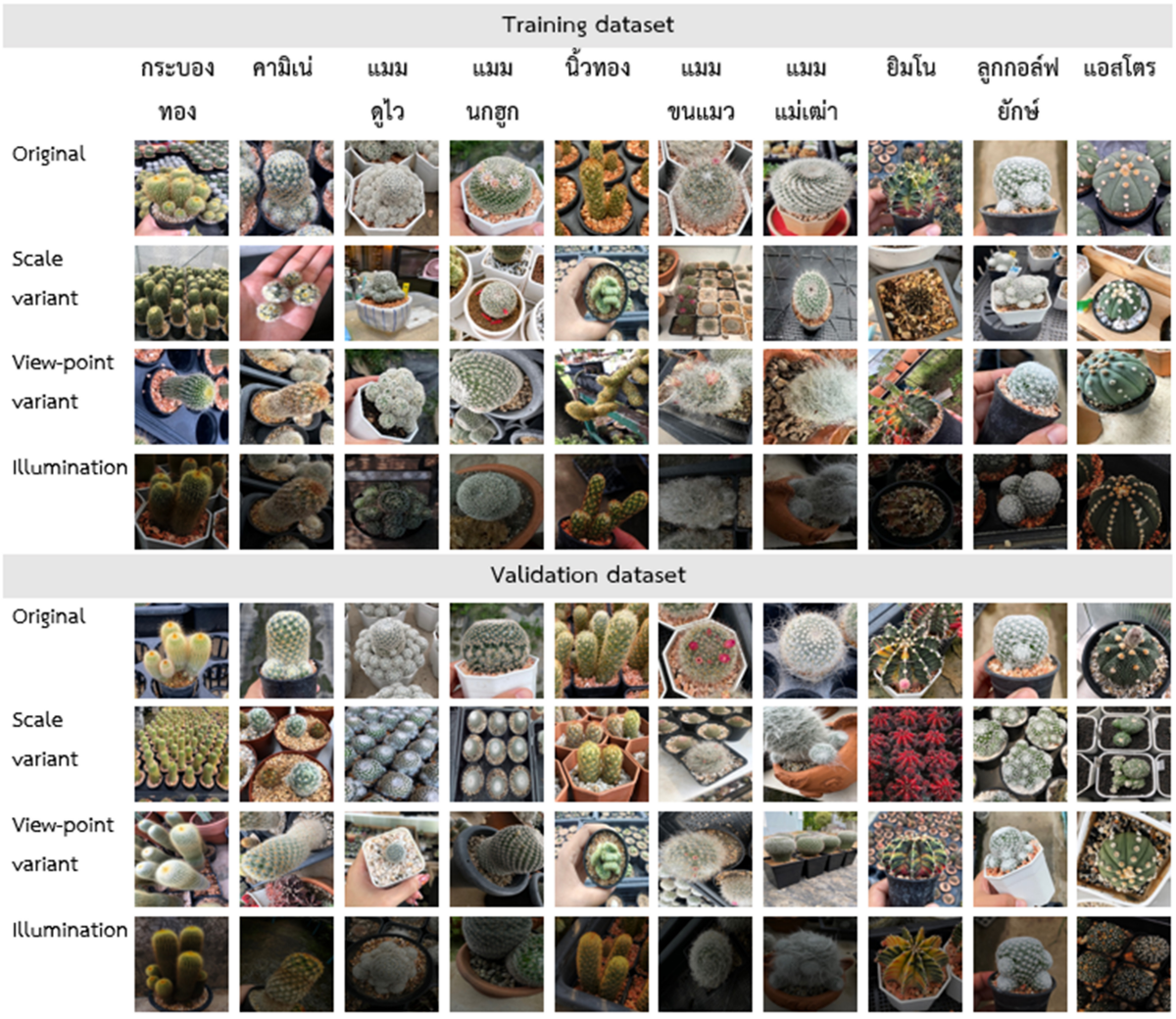


ภาพ 1 ตัวอย่างรูปภาพของแคคตัส 10 สายพันธุ์ที่มีความแตกต่างกัน

ตาราง 1

จำนวนรูปภาพแคคตัสแต่ละสายพันธุ์ เมื่อแบ่งเป็นชุดข้อมูลฝึกฝนและชุดข้อมูลตรวจสอบ

สายพันธุ์	จำนวนรูปภาพ (ภาพ)	ชุดข้อมูลฝึกฝน (ภาพ)	ชุดข้อมูลตรวจสอบ (ภาพ)
กระบองทอง	130	104	26
คามิเน่	156	125	31
แมมดูไว	169	135	34
แมมนกฮูก	154	123	31
นิ้วทอง	130	104	26
แมมขนแมว	142	114	28
แมมแม่เฒ่า	140	112	28
ยิมโน	133	106	27
ลูกกอล์ฟยักษ์	174	139	35
แอสโตร	180	144	36
รวม	**1,508**	**1,206**	**302**



ภาพ 2 ตัวอย่างรูปภาพของแคคตัส 10 สายพันธุ์ที่แบ่งออกเป็นชุดข้อมูลฝึกฝนและชุดข้อมูลตรวจสอบ

2) การกำหนดค่าไฮเปอร์พารามิเตอร์ (model hyperparameter configuration) งานวิจัยนี้กำหนดค่าพารามิเตอร์ที่สำคัญดังตารางที่ 2 สำหรับการฝึกฝนโมเดล MobileNetV2 ทำหน้าที่เป็นตัวสกัดคุณลักษณะ (feature extractor) ซึ่งใช้โครงสร้างของโครงข่ายประสาทเทียมแบบคอนโวลูชันที่ออกแบบมาเพื่อให้มีประสิทธิภาพสูงในด้านการคำนวณ และลดความซับซ้อนของโมเดล โดยกระบวนการสกัดคุณลักษณะนี้จะช่วยแปลงภาพอินพุตให้อยู่ในรูปแบบของผังคุณลักษณะ (feature maps) ซึ่งเป็นตัวแทนเชิงพื้นที่ของข้อมูลที่สามารถนำไปใช้ในการตรวจจับวัตถุได้

ค่าไฮเปอร์พารามิเตอร์ที่กำหนดประกอบด้วยการกำหนดขนาดภาพ (image size) ที่ 160x160 พิกเซล ซึ่งเป็นขนาดที่เหมาะสมระหว่างการรักษาความละเอียดของข้อมูลและการใช้ทรัพยากรการประมวลผลอย่างมีประสิทธิภาพ ขนาดของชุดข้อมูล (batch size) ถูกตั้งไว้ที่ 32 เพื่อปรับสมดุลระหว่างความเร็วและความเสถียรในการเรียนรู้ นอกจากนี้ยังมีการตั้งค่าการสับเปลี่ยนลำดับข้อมูล (shuffle buffer size) ที่ 1,000 เพื่อช่วยให้การฝึกฝนโมเดลไม่เกิดการจดจำลำดับข้อมูล การกำหนดอัตราการเรียนรู้ (base learning rate) ที่ 0.0001 ช่วยให้การปรับค่าพารามิเตอร์มีความละเอียด ลดโอกาสเกิดปัญหา Overshoot ที่จุดต่ำสุดของฟังก์ชันการสูญเสีย สำหรับจำนวนรอบการฝึกฝนเบื้องต้น (initial epochs) กำหนดเป็น 20 รอบ เพื่อให้โมเดลได้เรียนรู้คุณสมบัติพื้นฐานก่อนเข้าสู่ขั้นตอนการปรับแต่ง (fine tune at) ซึ่งเริ่มต้นที่เลเยอร์ที่ 100 ของโมเดล MobileNetV2 โดยเลือกปรับเฉพาะเลเยอร์ชั้นบนสุดที่เกี่ยวข้องกับงานที่ต้องการ ขณะที่เลเยอร์ก่อนหน้านั้นยังคงถูกแช่แข็ง และในขั้นตอนนี้กำหนดรอบการฝึกฝนเพิ่มเติม (fine tune epochs) อีก 10 รอบ เพื่อเพิ่มประสิทธิภาพการจำแนก นอกจากนี้ยังกำหนดขั้นตอนในการประเมินผล (validation steps) ที่ 20 เพื่อช่วยลดเวลาในการตรวจสอบความถูกต้องโดยไม่กระทบคุณภาพของผลลัพธ์โดยรวม

3) การฝึกฝนโมเดล (model training) เป็นกระบวนการที่สำคัญในการปรับค่าพารามิเตอร์ของโมเดลตามแนวคิดการแพร่กระจายย้อนกลับ ประกอบด้วยค่าน้ำหนัก (weights) และค่าไบแอส (biases) สำหรับปรับปรุงประสิทธิภาพของโมเดล กระบวนการเริ่มต้นจากการส่งข้อมูลเข้าสู่โมเดลเป็นกลุ่มย่อย (batch) ซึ่งช่วยเพิ่มประสิทธิภาพในการฝึกโมเดลด้วยการคำนวณในช่วงของข้อมูลที่มีขนาดเล็กและการประเมินผลระหว่างการฝึก (evaluation during training) โดยในระหว่างการฝึกฝน โมเดลจะถูกประเมินผลด้วยค่าความถูกต้อง (accuracy) และค่าการสูญเสีย (loss) บนชุดข้อมูลตรวจสอบ เพื่อให้แน่ใจว่าโมเดลสามารถแสดงประสิทธิภาพที่ดีทั้งในชุดข้อมูลฝึกฝนและข้อมูลตรวจสอบ และช่วยป้องกันปัญหา Overfitting ซึ่งอาจเกิดขึ้นได้เมื่อโมเดลมีประสิทธิภาพสูงมากบนชุดข้อมูลฝึกฝน แต่ประสิทธิภาพจะลดลงเมื่อต้องทำงานกับข้อมูลใหม่ที่ไม่เคยเห็นมาก่อน เมื่อดำเนินการเสร็จสิ้นแล้วจึงส่งออกโมเดล (model exportation) ที่ผ่านการฝึกฝนแล้วให้อยู่ในรูปแบบไฟล์นามสกุล .tflite เพื่อให้สะดวกต่อการนำไปใช้งาน

ตาราง 2

การกำหนดค่าไฮเปอร์พารามิเตอร์ในการฝึกฝนโมเดล MobileNetV2

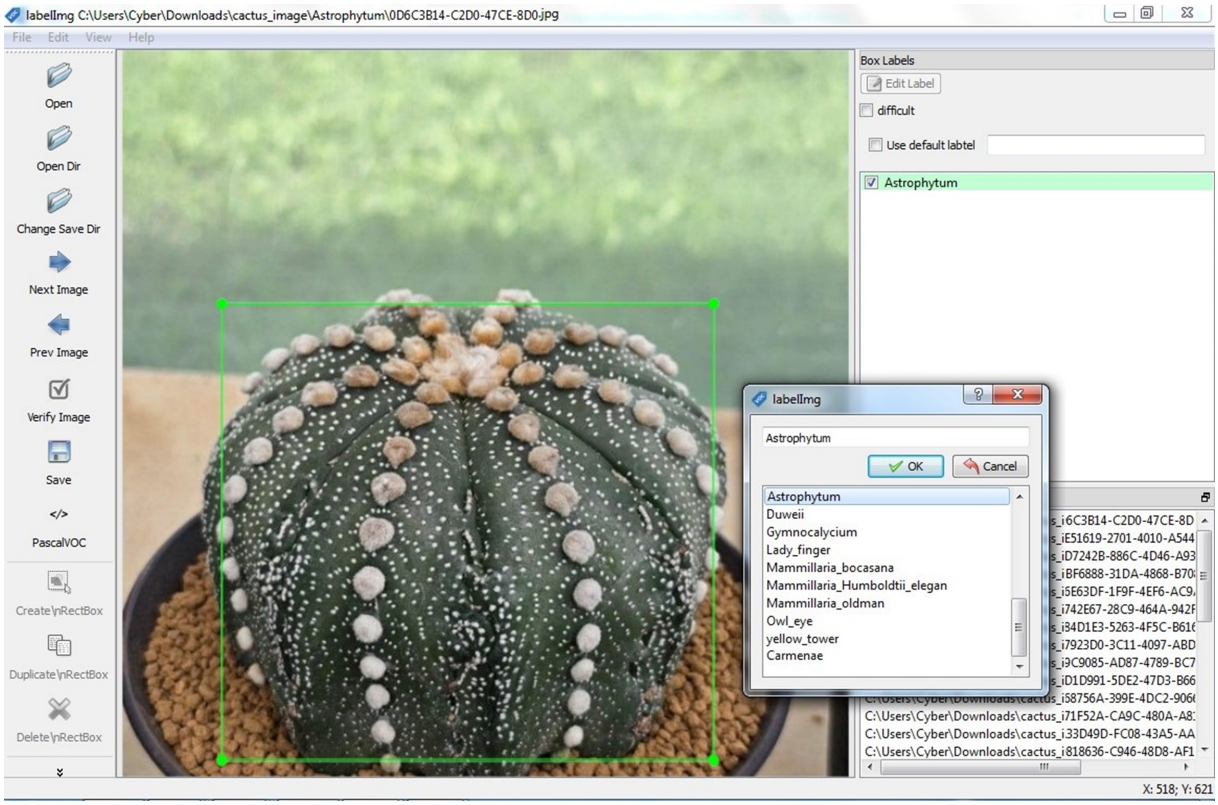
ไฮเปอร์พารามิเตอร์	ค่า
Image size	160 x 160 pixels
Batch size	32
Shuffle buffer size	1000
Base learning rate	0.0001
Initial epochs	20
Fine tune at	Layer 100 of MobileNetV2
Fine tune epochs	10
Validation step	20

ผลการวิจัย

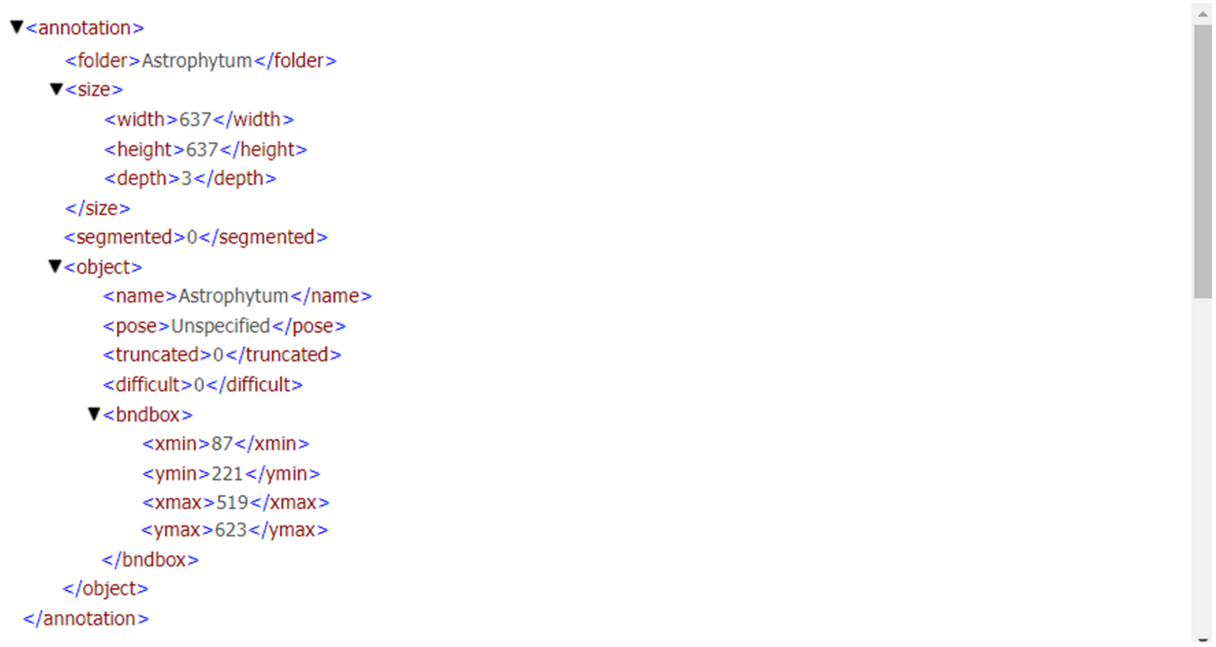
งานวิจัยนี้แบ่งผลการวิจัยออกเป็น 3 ส่วนตามวัตถุประสงค์ ประกอบด้วยผลการพัฒนาชุดข้อมูลรูปภาพและคำอธิบายรูปภาพสายพันธุ์แคคตัส ผลการพัฒนาตัวแบบการระบุสายพันธุ์แคคตัสด้วยโมเดลการเรียนรู้เชิงลึกและเทคนิคตรวจจับวัตถุ และผลการประเมินความถูกต้องการระบุสายพันธุ์แคคตัส มีรายละเอียด ดังนี้

1) ผลการพัฒนาชุดข้อมูลรูปภาพและคำอธิบายรูปภาพสายพันธุ์แคคตัส จำนวน 10 สายพันธุ์รวมทั้งสิ้น 1,508 ภาพ ก่อนจะนำมาสร้างคำอธิบายรูปภาพเพื่อกำกับข้อมูลแต่ละภาพด้วยกรอบ Bounding box ด้วยโปรแกรม LabelImg (Konyushkova, Pishchulin & Matuszek, 2018) ประกอบด้วยการกำหนดตำแหน่งของวัตถุที่ต้องการตรวจจับในภาพในรูปแบบของพิกัด ซึ่งบ่งชี้ถึงตำแหน่งของสี่เหลี่ยมที่ครอบคลุมวัตถุ ประกอบด้วยความกว้าง ความสูง และตำแหน่งของ Bounding box รวมถึงชื่อประเภทของวัตถุที่ถูกกำกับและป้ายกำกับที่บ่งบอกถึงสายพันธุ์ของแคคตัสในภาพ ดังภาพ 3

จากภาพ 3 และ 4 งานวิจัยนี้คำนึงถึงความหลากหลายของขนาดวัตถุ มุมมอง และสภาพแสง จึงปรับ Bounding box ให้ตรงกับวัตถุในภาพที่มีแสงน้อยหรือมุมกล้องที่แตกต่างไปจากภาพอื่น ๆ เพื่อให้โมเดลสามารถเรียนรู้ได้ดีจากข้อมูลที่ถูกกำกับอย่างถูกต้องในทุกสภาวะ ครอบคลุมพื้นที่ของวัตถุอย่างถูกต้อง ไม่ครอบคลุมพื้นที่เกินความจำเป็นหรือเล็กเกินไปจนทำให้โมเดลเข้าใจข้อมูลผิด จากนั้นจะแปลงข้อมูลให้อยู่ในรูปแบบ COCO format แล้วบันทึกข้อมูลคำอธิบายรูปภาพในรูปแบบ XML



ภาพ 3 การกำกับข้อมูลด้วยกรอบ Bounding Box ด้วยโปรแกรม LabelImg



ภาพ 4 ไฟล์ XML ที่เก็บข้อมูลของวัตถุที่ถูกกำกับ

เมื่อบันทึกข้อมูลเสร็จสิ้น ต้องทำการตรวจสอบคุณภาพข้อมูลก่อนนำไปใช้งาน โดยสุ่มตรวจสอบรูปภาพจากชุดข้อมูลในสัดส่วนร้อยละ 10 ของจำนวนรูปภาพในแต่ละสายพันธุ์ เพื่อตรวจสอบว่ากรอบ Bounding box กำกับแคคตัสในรูปภาพได้อย่างถูกต้องหรือไม่ หากพบว่า ไม่ถูกต้องให้ดำเนินการแก้ไขใหม่ พร้อมตั้งชื่อรูปภาพและคำอธิบายรูปภาพแต่ละไฟล์ให้สอดคล้องกัน และเรียกใช้งานผ่าน Google drive เพื่อให้สะดวกต่อการใช้งาน

2) ผลการพัฒนาตัวแบบการระบุสายพันธุ์แคคตัสด้วยโมเดลการเรียนรู้เชิงลึกและเทคนิคตรวจจับวัตถุ ประกอบด้วยขั้นตอนต่าง ๆ ประกอบด้วยการนำเข้าชุดข้อมูล การฝึกฝนและการปรับแต่งโมเดล การประเมินผลการตรวจจับวัตถุ และการส่งออกโมเดลที่ผ่านการฝึกฝนแล้ว มีรายละเอียด ดังนี้

2.1) การนำเข้าชุดข้อมูล (dataset importing) ประกอบด้วยชุดข้อมูลฝึกฝน และชุดข้อมูลตรวจสอบที่ผ่านการปรับขนาดเป็น 160x160 พิกเซล เพื่อลดภาระในการประมวลผล โดยข้อมูลของภาพและคำอธิบายรูปภาพแต่ละคู่จะได้รับการประมวลผลให้พร้อมสำหรับการนำเข้าสู่โมเดล หลังจากนั้นข้อมูลจะถูกสับเปลี่ยนและแบ่งออกเป็นชุดขนาดเล็กจำนวน 32 ภาพต่อชุด เพื่อเพิ่มประสิทธิภาพการฝึกฝนและลดความเสี่ยงจากการเกิดปัญหา Overfitting ก่อนจะผ่านตัวสกัดคุณลักษณะ (feature extractor) สำหรับแปลงภาพจากขนาดเดิม คือ 160x160x3 โดย 3 คือ ช่องสีรูปแบบ RGB ประกอบด้วยสีแดง สีเขียว และสีน้ำเงิน กระบวนการนี้ช่วยลดขนาดของข้อมูลในแต่ละภาพลงกลายเป็นบล็อกคุณลักษณะ (feature block) เป็นขนาด 5x5x1280 โดยการลดมิติของภาพผ่านกระบวนการต่าง ๆ ในโครงข่ายประสาทเทียม ซึ่งมีบทบาทในการสกัดคุณลักษณะสำคัญจากภาพ โดยยังคงรักษาความสำคัญของข้อมูลไว้ในรูปแบบฟีเจอร์ที่มี 1,280 สำหรับนำไปใช้ในขั้นตอนถัดไปของการเรียนรู้เชิงลึก ซึ่งคุณลักษณะเหล่านี้จะมีความละเอียดต่ำลง แต่ยังคงแสดงถึงคุณลักษณะสำคัญที่สามารถใช้ในการวิเคราะห์และประมวลผลรูปภาพได้ครบถ้วน

2.2) การฝึกฝนและการปรับแต่งโมเดล (model training and fine-tuning) งานวิจัยนี้ใช้โมเดล MobileNetV2 ที่ได้รับการฝึกฝนล่วงหน้าบนชุดข้อมูล ImageNet เป็นโมเดลฐาน ซึ่งเรียกใช้โดยไม่รวมเลเยอร์การจำแนกประเภทชั้นบนสุด เพื่อให้สามารถใช้เลเยอร์ฐานในการสกัดคุณลักษณะของข้อมูลได้ เมื่อได้เลเยอร์ฐานแล้วจะมีการตรึงเลเยอร์เหล่านี้เพื่อป้องกันไม่ให้พารามิเตอร์ถูกปรับระหว่างการฝึกฝน จากนั้นจะเพิ่มส่วนหัวสำหรับการจำแนกประเภทโดยใช้ GlobalAveragePooling2D เพื่อแปลงคุณลักษณะจากเลเยอร์ฐานเป็นเวกเตอร์ขนาด 1,280 องค์ประกอบ และใช้

Dense layer เพื่อพยากรณ์ประเภทของภาพ โดยกำหนดออปติไมเซอร์ (Optimizer) เป็น RMSprop: Root Mean Square Propagation ที่เป็นอัลกอริทึมปรับอัตราการเรียนรู้บนชุดข้อมูลที่มีการกระจายตัวไม่แน่นอน โดยแบ่งค่าอัตราการเรียนรู้ด้วยค่าเฉลี่ยของค่าความชันยกกำลังสอง (mean squared gradients) จากค่าความชันในแต่ละรอบ (Iteration) และถ่วงน้ำหนักตามเวลาล่าสุด ดังนั้น RMSprop จะปรับปรุงอัตราการเรียนรู้แบบไดนามิก โดยให้ค่าอัตราการเรียนรู้น้อยลงเรื่อย ๆ สำหรับพารามิเตอร์ที่มีค่าความชันขนาดใหญ่ ตรงกันข้าม หากค่าความชันมีขนาดเล็กจะให้อัตราการเรียนรู้ที่มากขึ้นตามไปด้วย ซึ่งช่วยป้องกันไม่ให้เกิดการเปลี่ยนแปลงน้ำหนักที่รุนแรงเกินไป และทำให้การเรียนรู้ของโมเดลเสถียรขึ้นกระบวนการทำงานเริ่มจากคำนวณหาค่าความผิดพลาดที่เกิดขึ้นในผลลัพธ์ของโมเดล จากนั้นจึงค่อย ๆ คำนวณย้อนกลับไปยังทุกชั้นของโมเดลเพื่อหาทิศทางของการเปลี่ยนแปลงน้ำหนักที่เหมาะสม ฟังก์ชันการสูญเสีย (loss function) กำหนดเป็น BinaryCrossentropy สอดคล้องกับปัญหาการจำแนกประเภทแบบไบนารี และมาตรวัด (metrics) กำหนดเป็นค่าความถูกต้อง (accuracy)

การฝึกฝนโมเดลกำหนด 20 รอบบนชุดข้อมูลฝึกฝนเพื่อให้โมเดลได้เรียนรู้คุณลักษณะพื้นฐานของข้อมูลก่อนที่จะเข้าสู่กระบวนการปรับแต่งเพิ่มเติม (fine-tuning) การเลือกจำนวนรอบการฝึกที่เหมาะสมเป็นปัจจัยสำคัญที่ช่วยให้โมเดลสามารถพัฒนาได้โดยไม่เกิดปัญหา Overfitting หรือ Underfitting โดยในช่วงเริ่มต้นของการฝึกเลเยอร์ทั้งหมดของโมเดล MobileNetV2 จะถูกแช่แข็ง (frozen) และมีการปรับค่าพารามิเตอร์เฉพาะส่วนที่เพิ่มเข้าไปใหม่ซึ่งเป็นชั้นสุดท้ายของโมเดลที่ทำหน้าที่จำแนกประเภท การฝึกฝนในระยะนี้ช่วยให้โมเดลสามารถเรียนรู้คุณลักษณะที่จำเป็นจากชุดข้อมูลฝึกฝน โดยอาศัยการถ่ายโอนความรู้ (transfer learning) จากการฝึกบนชุดข้อมูล ImageNet ดังนั้นการกำหนดจำนวนรอบการฝึกที่ 20 ครั้งช่วยให้โมเดลสามารถปรับค่าเวกเตอร์น้ำหนักในชั้นใหม่ได้อย่างเหมาะสม และหลีกเลี่ยงการเรียนรู้ที่เร็วเกินไปซึ่งอาจนำไปสู่ความไม่เสถียรของโมเดล นอกจากนี้ ยังช่วยให้สามารถสังเกตพฤติกรรมของค่าความแม่นยำและค่าการสูญเสีย ทั้งบนชุดข้อมูลฝึกฝนและชุดข้อมูลทดสอบ หากโมเดลมีค่าความแม่นยำเพิ่มขึ้นอย่างต่อเนื่อง และแนวโน้มที่ค่าการสูญเสียลดลงเรื่อย ๆ แสดงว่า ผลการฝึกฝนกำลังดำเนินไปอย่างเหมาะสม ในทางตรงกันข้ามหากพบว่า ค่าการสูญเสียของชุดข้อมูลฝึกฝนลดลงมาก แต่ค่าการสูญเสียของชุดข้อมูลทดสอบกลับเพิ่มขึ้น อาจเป็นสัญญาณของการเกิด Overfitting ซึ่งอาจนำไปสู่การพิจารณาใช้เทคนิคหยุดการฝึกก่อนกำหนด (early stopping) หรือการลดอัตราการเรียนรู้ (learning rate scheduling) เพื่อป้องกันการสูญเสียความสามารถในการจำแนกประเภท และไม่ใช้ทรัพยากรในการประมวลผลมากเกินไป หลังจากการฝึกฝนจะมีการประเมินผลเพื่อวัดประสิทธิภาพของโมเดล ประกอบด้วย ค่าความถูกต้อง ค่าการสูญเสียบนชุดข้อมูลฝึกฝน ค่าความถูกต้อง และค่าการสูญเสียบนชุดข้อมูลตรวจสอบตามลำดับ

โมเดล MobileNetV2 ที่ใช้ในการวิจัยนี้ได้รับการถ่ายโอนการเรียนรู้ (transfer Learning) จากชุดข้อมูล ImageNet โดยนำโครงสร้างของโมเดลที่ได้รับการฝึกฝนล่วงหน้ามาประยุกต์ใช้กับชุดข้อมูลแคคตัส โมเดลถูกฝึกฝนโดยใช้ชุดข้อมูลฝึกอบรมและประเมินผลด้วยการทดสอบแบบไขว้ทบ (K-fold cross validation) แบบ 10 ส่วน เป็นหนึ่งในเทคนิคที่ได้รับความนิยมโดยเฉพาะอย่างยิ่งเมื่อใช้กับข้อมูลที่มีจำนวนจำกัด เนื่องจากช่วยลดอคติที่เกิดจากการแบ่งชุดข้อมูลแบบสุ่มและช่วยให้การประเมินผลมีความน่าเชื่อถือมากขึ้น เริ่มจากการแบ่งชุดข้อมูลทั้งหมดออกเป็น 10 ส่วนที่มีขนาดใกล้เคียงกัน โดยแต่ละรอบของการประเมินผล โมเดลจะถูกฝึกฝนด้วย 9 ส่วนของข้อมูล และใช้ส่วนที่เหลืออีก 1 ส่วนเป็นชุดข้อมูลตรวจสอบ กระบวนการนี้จะทำซ้ำทั้งหมด 10 ครั้ง โดยแต่ละครั้งจะมีการเปลี่ยนชุดข้อมูลตรวจสอบไปเรื่อย ๆ จนครบทุกส่วน หลังจากดำเนินการครบ 10 รอบแล้ว ค่าความแม่นยำที่ได้จากแต่ละรอบจะถูกนำมาคำนวณค่าเฉลี่ย เพื่อใช้เป็นตัวบ่งชี้ประสิทธิภาพของโมเดลให้การประเมินผลมีความเป็นกลางมากขึ้น เนื่องจากทุกตัวอย่างในชุดข้อมูลจะถูกใช้ทั้งในการฝึกฝนและการตรวจสอบ ส่งผลให้สามารถลดความลำเอียงที่อาจเกิดจากการแบ่งข้อมูลเพียงครั้งเดียว นอกจากนี้ยังช่วยลดปัญหาการเกิด Overfitting เพราะโมเดลไม่ได้เรียนรู้จากชุดข้อมูลใดชุดหนึ่งมากเกินไป

ระหว่างกระบวนการฝึกฝนโมเดล มีการใช้เทคนิคการปรับแต่งค่าพารามิเตอร์ (fine-tuning) เพื่อปรับค่า

ถ่วงน้ำหนักของโมเดลให้เหมาะสมกับลักษณะเฉพาะของชุดข้อมูล โดยใช้ออปติไมเซอร์เป็น RMSprop ควบคู่กับฟังก์ชันค่าการสูญเสีย เพื่อเพิ่มประสิทธิภาพและความแม่นยำของโมเดลผ่านกระบวนการแพร่กระจายย้อนกลับ (backpropagation) ซึ่งเป็นแนวคิดพื้นฐานของการเรียนรู้เชิงลึก กระบวนการดังกล่าวช่วยให้โมเดลสามารถเรียนรู้คุณลักษณะที่สำคัญของแต่ละสายพันธุ์แคคตัสได้อย่างมีประสิทธิภาพ

ผลลัพธ์จากการฝึกฝนโมเดล MobileNetV2 สำหรับการจำแนกประเภทแสดงให้เห็นถึงพัฒนาการที่ต่อเนื่องในแต่ละรอบ ทั้งค่าความถูกต้อง และค่าการสูญเสียบนชุดข้อมูลฝึกฝน ซึ่งสะท้อนถึงประสิทธิภาพของกระบวนการเรียนรู้และความสามารถในการจำแนกประเภทที่เพิ่มขึ้นอย่างมีนัยสำคัญตลอดช่วงการฝึกฝน โดยค่าความถูกต้องของชุดข้อมูลฝึกฝนเริ่มต้นที่ 70.51% และเพิ่มขึ้นอย่างต่อเนื่องจนถึง 92.86% ในรอบที่ 20 ขณะที่ค่าการสูญเสียลดลงจาก 0.5643 ในรอบแรกไปที่ 0.1796 ในรอบที่ 20 ซึ่งบ่งชี้ว่า โมเดลสามารถเรียนรู้คุณลักษณะของข้อมูลและลดข้อผิดพลาดจากการจำแนกได้อย่างมีประสิทธิภาพ เช่นเดียวกับบนชุดข้อมูลตรวจสอบที่พบว่า ค่าความถูกต้องเพิ่มขึ้นจาก 74.73% ในรอบแรกไปถึง 90.66% ในรอบที่ 20 ในขณะที่ค่าการสูญเสียลดลงจาก 0.5222 ไปที่ 0.1933 ซึ่งสะท้อนถึงความสามารถของโมเดลในการจำแนกประเภทภาพในชุดข้อมูลที่ไม่เคยเห็นมาก่อน ดังภาพ 5

อย่างไรก็ดี งานวิจัยนี้ให้ความสำคัญค่าความคลาดเคลื่อนเชิงสถิติ (statistical error) ที่ยากจะควบคุมหรือกำจัดให้หมดไป เนื่องจากอยู่นอกเหนือการควบคุม จึงต้องทำการวัดซ้ำหลาย ๆ ครั้ง ในงานวิจัยนี้เมื่อฝึกฝนโมเดลครบใน 1 รอบจะทำการประเมินอีกจนครบ 3 รอบ แล้วนำผลการประเมินมาหาค่าเฉลี่ยเพื่อให้ความคลาดเคลื่อนลดน้อยลงจนถึงในระดับที่ยอมรับได้ ดังตาราง 3 เมื่อพิจารณาผลการฝึกฝนโมเดลทั้ง 3 รอบ พบว่า ค่าความแม่นยำเฉลี่ยของชุดข้อมูลฝึกฝนอยู่ที่ 92.64% และชุดข้อมูลตรวจสอบอยู่ที่ 91.05% ซึ่งบ่งชี้ว่า โมเดลสามารถเรียนรู้รูปแบบของข้อมูลได้อย่างมีประสิทธิภาพ และไม่มีสัญญาณของการเกิด Overfitting หรือ Underfitting อย่างมีนัยสำคัญ เช่นเดียวกับค่าการสูญเสียเฉลี่ยของชุดข้อมูลฝึกฝนอยู่ที่ 0.1853 และชุดข้อมูลตรวจสอบอยู่ที่ 0.1955 ซึ่งแสดงให้เห็นว่าโมเดลสามารถลดค่าการสูญเสียได้อย่างต่อเนื่องโดยไม่มีความแตกต่างมากนักระหว่างสองชุดข้อมูล

จากตาราง 3 เมื่อพิจารณาในรายละเอียด พบว่าความสัมพันธ์ระหว่างค่าความถูกต้องและค่าการสูญเสียที่สอดคล้องกันในทั้งสองชุดข้อมูลชี้ให้เห็นว่า โมเดลไม่ได้ประสบปัญหาการเกิดปัญหาการจดจำคุณลักษณะเฉพาะจนเกินไป และลดปัญหาที่โมเดลมีประสิทธิภาพสูงเมื่อทำงานกับชุดข้อมูลฝึกฝน แต่ประสิทธิภาพจะลดลงเมื่อนำไปใช้งานกับข้อมูลจริง (poor real-world performance)

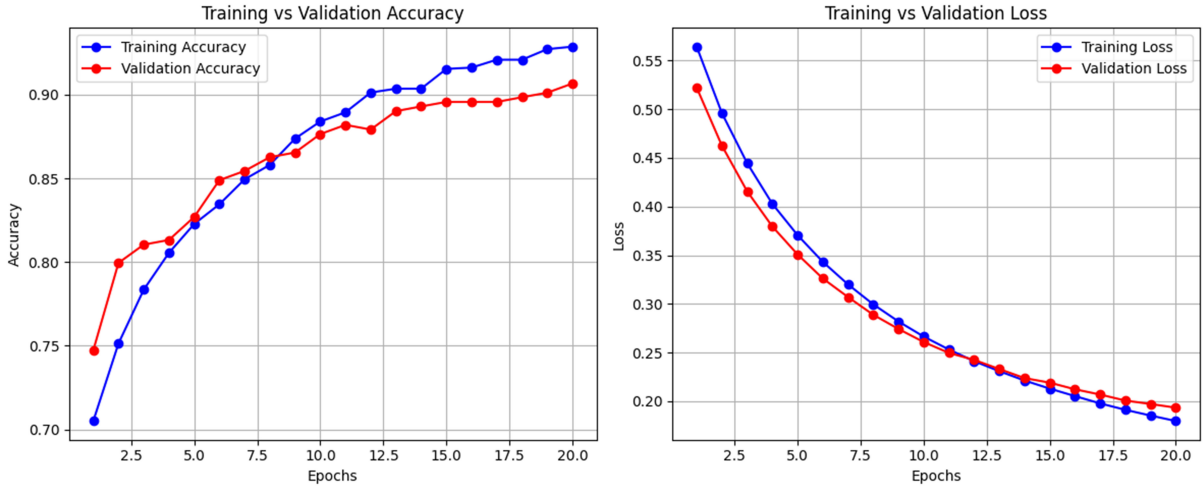
2.3) การประเมินผลการตรวจจับวัตถุ (object detection evaluation) บนชุดข้อมูลตรวจสอบด้วยเมทริกซ์ที่ออกแบบและคำนวณในลักษณะเดียวกับที่ใช้ในชุดข้อมูล COCO: Common Objects in Context ซึ่งเป็นมาตรฐานที่ใช้กันอย่างแพร่หลายในการประเมินโมเดลสำหรับงานตรวจจับวัตถุ โดยใช้ค่าความแม่นยำเฉลี่ย และค่าการเรียกคืนเฉลี่ย ซึ่งถูกคำนวณในช่วงค่าดัชนี IoU: Intersection over Union ต่าง ๆ เพื่อตรวจสอบความถูกต้องและความครบถ้วนของการตรวจจับวัตถุ มีรายละเอียด ดังนี้

2.3.1) ค่าความแม่นยำเฉลี่ย (Average Precision: AP) เป็นค่าที่ใช้ประเมินประสิทธิภาพของการตรวจจับวัตถุ โดยคำนวณจากค่าความแม่นยำ (precision) และค่าการเรียกคืน (recall) ที่ระดับต่าง ๆ ของอัตราส่วนพื้นที่ซ้อนทับต่อพื้นที่รวม (Intersection over Union: IoU) โดยค่าความแม่นยำคือค่าที่บอกจำนวนที่โมเดลทำนายเป็นวัตถุว่ามีจำนวนเท่าไหร่ที่ถูกต้อง และค่าการเรียกคืนคือค่าที่บอกความสามารถในการทำนายวัตถุทั้งหมดที่มีในภาพได้ครบถ้วน โดยอัตราส่วนพื้นที่ซ้อนทับต่อพื้นที่รวมคือค่าที่วัดว่าขอบเขตที่โมเดลทำนายมาทับซ้อนกับขอบเขตของวัตถุจริงมากน้อยแค่ไหน จะอยู่ระหว่าง 0 ถึง 1 โดยค่า 1 หมายถึงขอบเขตทำนายทับซ้อนกับวัตถุจริง 100% ดังนั้นค่าความแม่นยำเฉลี่ย จะถูกคำนวณจากการวัดความแม่นยำและการเรียกคืนที่ระดับต่าง ๆ ของอัตราส่วนพื้นที่ซ้อนทับต่อพื้นที่รวม

2.3.2) ค่าการเรียกคืนเฉลี่ย (Average Recall: AR) เป็นค่าที่วัดความสามารถในการทำนายวัตถุทุกชิ้นในภาพ ยิ่งค่า AR สูงหมายถึง โมเดลสามารถตรวจจับวัตถุได้ครบถ้วน แบ่งเป็นค่า ARmax1 ARmax10 และ ARmax100 คือค่าการเรียกคืนสูงสุดสำหรับการตรวจจับวัตถุได้สูงสุดจำนวน 1 10 และ 100 ชิ้นตามลำดับ ดังนั้นหากค่า AR สูงจะหมายถึง โมเดลสามารถตรวจจับวัตถุได้

ครบถ้วนหรือเกือบครบถ้วนในภาพ ในทางกลับกันหากค่าค่าการเรียกคืนเฉลี่ยต่ำ หมายถึง โมเดลพลาดการตรวจจับวัตถุจำนวนมากในภาพ และไม่ได้ทำการตรวจจับวัตถุครบถ้วนนั่นเอง

จากตาราง 4 ค่าความแม่นยำเฉลี่ยในการตรวจจับวัตถุบนชุดข้อมูลตรวจสอบอยู่ในระดับที่สูงมาก โดยเฉพาะในช่วงค่าอัตราส่วนพื้นที่ซ้อนทับต่อพื้นที่รวม (IoU Range) ที่เท่ากับ 0.50 และ 0.75 เท่ากับ 0.991 ซึ่งบ่งชี้ว่าการตรวจจับวัตถุในภาพมีความแม่นยำสูง เมื่อพิจารณาจากทั้งพื้นที่ (area) ของภาพที่มีวัตถุขนาดเล็กไปจนถึงขนาดใหญ่ อย่างไรก็ตาม ค่าความแม่นยำเฉลี่ยสำหรับวัตถุขนาดเล็กและขนาดกลาง เท่ากับ -1.000 แสดงถึงข้อจำกัดของโมเดลในการตรวจจับวัตถุที่มีขนาดเล็กและขนาดกลาง อาจเป็นผลมาจากข้อมูลที่ไม่เพียงพอสำหรับฝึกฝนเพื่อการตรวจจับวัตถุหลากหลายขนาด ตลอดจนขนาดของวัตถุอาจเล็กเกินไปเมื่อเทียบกับขนาดภาพ ในขณะที่ค่าความแม่นยำเฉลี่ยสำหรับวัตถุขนาดใหญ่เท่ากับ 0.988 ซึ่งอยู่ในระดับดีมาก



ภาพ 5 ผลลัพธ์จากการฝึกฝนรอบที่ 1 ของโมเดล MobileNetV2 สำหรับการจำแนกประเภทในด้านค่าความถูกต้องและค่าการสูญเสียบนชุดข้อมูลฝึกฝน เปรียบเทียบกับค่าความถูกต้องและค่าการสูญเสียบนชุดข้อมูลตรวจสอบ

ตาราง 3

ค่าความถูกต้อง ค่าการสูญเสียบนชุดข้อมูลฝึกฝน และค่าความถูกต้อง ค่าการสูญเสียบนชุดข้อมูลตรวจสอบ จากการฝึกฝนโมเดลจำนวน 3 รอบ

การประเมินรอบที่	Training dataset		Validation dataset	
	Accuracy	Loss	Accuracy	Loss
1	0.9286	0.1796	0.9138	0.1910
2	0.9271	0.1852	0.9105	0.1955
3	0.9234	0.1911	0.9073	0.2001
ค่าเฉลี่ย (Average)	**0.9264**	**0.1853**	**0.9105**	**0.1955**

ตาราง 4

ค่าความแม่นยำเฉลี่ย และค่าการเรียกคืนเฉลี่ยในการตรวจจับวัตถุบนชุดข้อมูลตรวจสอบ

เมทริกซ์	ช่วงค่าอัตราส่วนพื้นที่ซ้อนทับต่อพื้นที่รวม	พื้นที่	จำนวนกรอบตรวจจับสูงสุด	ค่า
ค่าความแม่นยำเฉลี่ย Average Precision: AP	0.50:0.95	วัตถุทั้งหมด (all)	100	0.988
	0.5	วัตถุทั้งหมด (all)	100	0.991
	0.75	วัตถุทั้งหมด (all)	100	0.991
	0.50:0.95	วัตถุขนาดเล็ก (small)	100	-1.000
	0.50:0.95	วัตถุขนาดกลาง (medium)	100	-1.000
	0.50:0.95	วัตถุขนาดใหญ่ (large)	100	0.988
ค่าการเรียกคืนเฉลี่ย Average Recall: AR	0.50:0.95	วัตถุทั้งหมด (all)	1	0.993
	0.50:0.95	วัตถุทั้งหมด (all)	10	0.993
	0.50:0.95	วัตถุทั้งหมด (all)	100	0.993
	0.50:0.95	วัตถุขนาดเล็ก (small)	100	-1.000
	0.50:0.95	วัตถุขนาดกลาง (medium)	100	-1.000
	0.50:0.95	วัตถุขนาดใหญ่ (large)	100	0.993

เช่นเดียวกับค่าการเรียกคืนเฉลี่ยที่สูงถึง 0.993 ในหลายหมวด โดยค่าจำนวนกรอบตรวจจับสูงสุด (MaxDets) เท่ากับ 1 10 และ 100 ตามลำดับ แสดงให้เห็นถึงความสามารถในการตรวจจับวัตถุที่มีการระบุวัตถุถูกต้องในระดับสูง เช่นเดียวกับค่าการเรียกคืนเฉลี่ยสำหรับการตรวจจับวัตถุขนาดเล็กและขนาดกลางที่เท่ากับ -1.000 ก็แสดงถึงปัญหาเดียวกับค่าความแม่นยำเฉลี่ยในการตรวจจับวัตถุขนาดเล็กและขนาดกลาง ซึ่งยังต้องได้รับการปรับปรุงในอนาคต

2.4) การส่งออกโมเดลที่ผ่านการฝึกฝนแล้ว (exporting the trained model) ในรูปแบบที่สามารถนำไปใช้งานได้จริง รวมถึงการแปลงโมเดลให้เหมาะสมกับการทำงานในสภาพแวดล้อมที่มีข้อจำกัดด้านพื้นที่จัดเก็บและด้านการใช้ทรัพยากร เช่น สมาร์ทโฟน หรืออุปกรณ์ IoT: Internet of Things โดยจะมีไฟล์ที่ประกอบด้วยเมทาดาตา (Metadata) ที่จำเป็นเพื่อให้สามารถใช้งานโมเดลในภายหลังได้อย่างถูกต้อง โดยงานวิจัยนี้บันทึกโมเดลที่ผ่านการฝึกฝนเป็นไฟล์ชื่อ cactus.tflite เพื่อให้สะดวกในการเรียกใช้งานในขั้นตอนต่อไป

3) ผลการประเมินความถูกต้องการระบุสายพันธุ์แคคตัสด้วยโมเดลการเรียนรู้เชิงลึกและเทคนิคตรวจจับวัตถุเพื่อเป็นการตรวจสอบความสมบูรณ์ของโมเดล (validation of exported model) ให้มั่นใจว่าโมเดลสามารถใช้งานได้จริงในสภาพแวดล้อมที่แตกต่างกัน งานวิจัยนี้จึงนำไปทดสอบผ่าน MediaPipe Studio ซึ่งเป็นเครื่องมือที่พัฒนาโดยกูเกิล เพื่อช่วยทดสอบการตรวจจับและวิเคราะห์วัตถุในภาพ ดังภาพ 6 การกำหนด Model Selection เพื่อเรียกใช้โมเดล cactus.tflite ที่ผ่านการฝึกฝนสำหรับทดสอบกับรูปภาพแคคตัสในสถานการณ์จริงผ่าน MediaPipe Studio ก่อนจะเลือกภาพที่ต้องการทดสอบ จากนั้นจะปรากฏผลการตรวจจับวัตถุ พร้อมแสดงสายพันธุ์ที่มีค่าความน่าจะเป็นสูงสุด และค่าความน่าจะเป็นของสายพันธุ์ที่ตรวจพบในภาพ

การประเมินความถูกต้องของการระบุสายพันธุ์โดยผู้เชี่ยวชาญกับภาพแคคตัสที่ตัวแบบไม่เคยเห็นมาก่อน (unseen dataset) โดยนำโมเดลที่ผ่านการฝึกฝนมาทดสอบกับรูปภาพจำนวน 300 รูปภาพ โดยให้ผู้เชี่ยวชาญทั้ง 3 ท่าน

ทำการประเมินผลท่านละ 100 รูปภาพผ่าน MediaPipe Studio การประเมินความถูกต้องโดยผู้เชี่ยวชาญหลายท่านจะช่วยลดความคลาดเคลื่อนของการประเมินที่อาจเกิดขึ้นจากการพิจารณาโดยผู้เชี่ยวชาญเพียงท่านเดียว และสะท้อนถึงความพร้อมของโมเดลต่อการใช้งานในบริบทที่ผู้ใช้งานแต่ละคนมีแหล่งข้อมูลและเงื่อนไขที่แตกต่างกัน เช่น รูปภาพที่มีความแตกต่างด้านระยะถ่าย มุมมอง แสงสว่างตลอดจนพื้นหลัง จะช่วยทดสอบความยืดหยุ่นของโมเดลในการจัดการกับความแปรปรวนของข้อมูล เช่น ภาพถ่ายในระยะที่แตกต่างกัน อาจทำให้รายละเอียดของลำต้นหรือหนามเปลี่ยนไป หรือมุมมองจากด้านบนหรือด้านข้างอาจส่งผลต่อคุณลักษณะเด่นของสายพันธุ์ นอกจากนี้สภาพแสงที่แตกต่างกัน เช่น แสงธรรมชาติ แสงจากหลอดไฟ หรือการมีเงาในภาพ ยังเป็นปัจจัยที่เพิ่มความซับซ้อนในการจำแนกสายพันธุ์ ดังนั้นหากโมเดลสามารถรักษาความถูกต้องในระดับสูงได้กับข้อมูลที่ไม่เคยเห็นมาก่อน จะช่วยเพิ่มศักยภาพในการนำไปใช้งานจริง การทดสอบนี้จึงถือเป็นขั้นตอนสำคัญที่ช่วยให้มั่นใจได้ว่าโมเดลสามารถตอบโจทย์ในสถานการณ์จริงได้อย่างมีประสิทธิภาพและสอดคล้องกับความต้องการของผู้ใช้งาน

กำหนดให้ หากผลการระบุสายพันธุ์แคคตัสในภาพถูกต้องทั้งหมด เท่ากับ 1 คะแนน ในทางตรงกันข้าม หากผลการระบุสายพันธุ์แคคตัสในภาพมีบางส่วนไม่ถูกต้องหรือไม่ถูกต้องทั้งหมด เท่ากับ 0 คะแนน เมื่อทำการประเมินแล้วจะนำคะแนนทั้งหมดที่ได้จากผู้เชี่ยวชาญทั้ง 3 ท่านมาคำนวณเป็นค่าความถูกต้องเฉลี่ย เพื่อประเมินประสิทธิภาพของตัวแบบดังสมการที่ 1

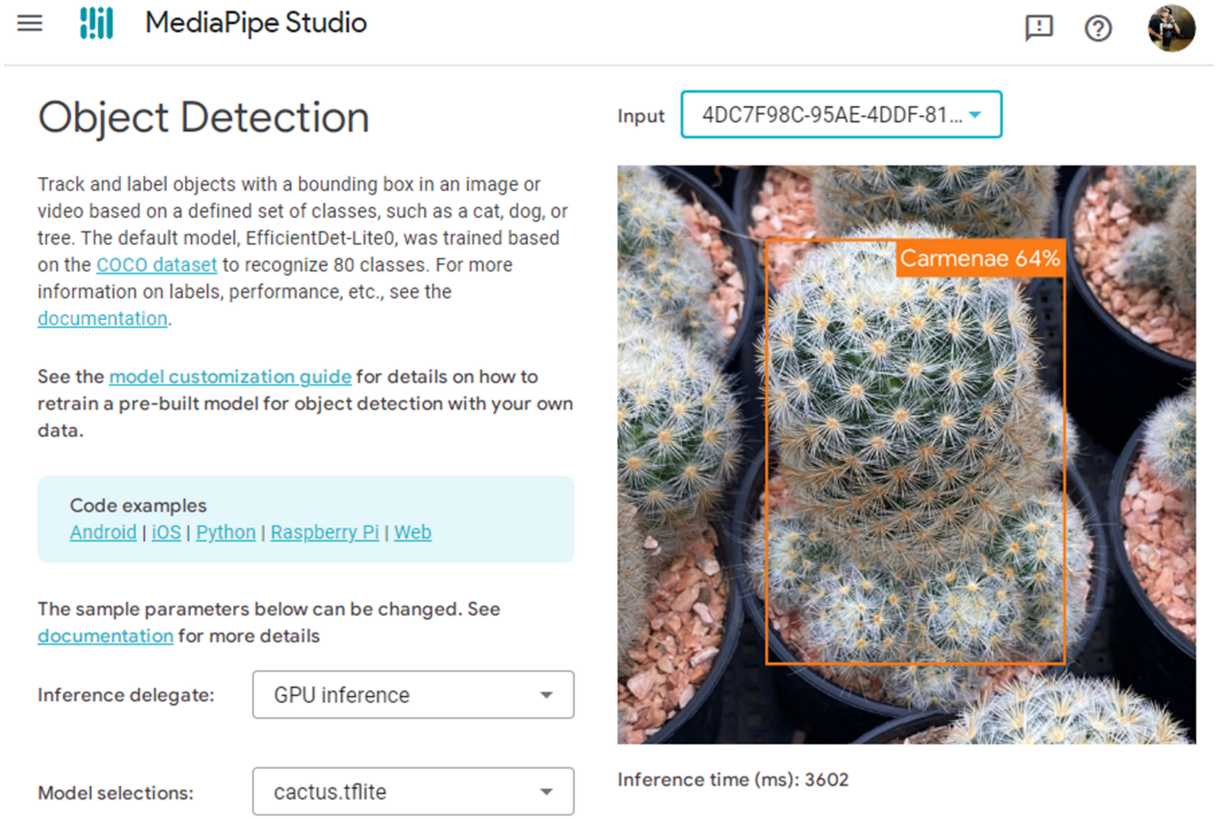
จากภาพ 7 เมื่อผู้เชี่ยวชาญนำเข้ารูปภาพแคคตัสจำนวน 10 สายพันธุ์ที่ตัวแบบไม่เคยเห็นมาก่อนเข้าสู่ MediaPipe Studio ที่เรียกใช้โมเดล cactus.tflite ที่ผ่านการฝึกฝนมาแล้วนั้น โดยผลลัพธ์จะเป็นกรอบสี่เหลี่ยมที่ล้อมรอบวัตถุที่ตรวจพบ (bounding box) พร้อมกับระบุประเภทของวัตถุและคลาสความน่าจะเป็นจากการจำแนกของวัตถุที่ตรวจจับ ก่อนจะนำไปประเมินความถูกต้องของการระบุสายพันธุ์แคคตัสทีละรูปภาพจนครบ 300 รูปภาพ โดยให้ผู้เชี่ยวชาญทั้ง 3 ท่านทำการประเมินผลท่านละ 100 รูปภาพแบบเป็นอิสระต่อกัน ดังภาพ 8 ตัวอย่างผลลัพธ์จาก MediaPipe Studio ประกอบด้วยกรอบสี่เหลี่ยมที่ล้อมรอบวัตถุที่ตรวจพบ พร้อมกับระบุประเภทของวัตถุและคลาสความน่าจะเป็นจากการจำแนกของวัตถุที่ตรวจจับกับรูปภาพใน 3 ประเภท ได้แก่ (1) Single-object, multiple classes (2) Multiple-objects, one Class และ (3) Multiple-objects, multiple classes

$$Accuracy = \frac{Number\ of\ correct\ predictions}{Total\ number\ of\ predictions} \quad (1)$$

โดยที่ $Number\ of\ correct\ predictions$ คือ จำนวนตัวอย่างที่โมเดลจำแนกได้ถูกต้อง (ภาพ)

$Total\ number\ of\ predictions$ คือ จำนวนตัวอย่างทั้งหมดที่ใช้ในการทดสอบ (ภาพ)

งานวิจัยนี้กำหนดรูปแบบของรูปภาพที่ใช้ในการทดสอบใน 3 ประเภท เพื่อให้ครอบคลุมทุกบริบทในการตรวจจับวัตถุ ประกอบด้วย (1) Single-object, multiple classes หมายถึง ภาพที่มีวัตถุเพียงหนึ่งชิ้นแต่สามารถจำแนกได้เป็นหลายคลาส (2) Multiple-objects, one class หมายถึง ภาพที่มีวัตถุหลายชิ้นที่กระจายอยู่ในภาพแต่ทั้งหมดเป็นคลาสเดียวกัน และ (3) Multiple-objects, multiple classes หมายถึง ภาพที่มีวัตถุหลายชนิดและทั้งหมดสามารถจำแนกได้เป็นหลายคลาส ดังภาพ 7



ภาพ 6 โมเดล MobileNetV2 ที่ผ่านการฝึกฝนและบักทึกไฟล์ชื่อ cactus.tflite เมื่อนำมาระบุสายพันธุ์แคคตัสโดยใช้เทคนิคตรวจจับวัตถุผ่าน MediaPipe Studio



ภาพ 7 ตัวอย่างรูปภาพแคคตัสประเภท Single-object, multiple classes, multiple-objects, one class และ multiple-objects, multiple Classes



ภาพ 8 ตัวอย่างผลการตรวจจับและระบุสายพันธุ์แคคตัสจาก MediaPipe Studio กับรูปภาพประเภท Single-object, multiple classes Multiple-objects, one class และ Multiple-objects, multiple classes

ตาราง 5

ผลการตรวจจับและระบุสายพันธุ์แคคตัสกับรูปภาพประเภท Single-object, multiple classes Multiple-objects, one class และ Multiple-objects, multiple classes

ประเภทรูปภาพ	จำนวน	ถูกต้อง	ไม่ถูกต้อง	ความถูกต้อง (เฉลี่ย)
Single-object, multiple classes	100	91	9	91.00%
Multiple-objects, one class	100	83	17	83.00%
Multiple-objects, multiple classes	100	76	24	76.00%
รวม	**300**	**250**	**50**	**83.33%**

การอภิปรายผล

งานวิจัยนี้มุ่งเน้นการตรวจจับและจำแนกสายพันธุ์ของแคคตัส โดยใช้เทคนิคการตรวจจับวัตถุแบบขั้นตอนเดียวที่จะช่วยในการระบุตำแหน่งของแคคตัสในภาพได้อย่างแม่นยำและรวดเร็ว ขณะที่โมเดล MobileNet จะช่วยลดการใช้ทรัพยากรในการคำนวณ เนื่องจากมีโครงสร้างมีขนาดเล็กและมีประสิทธิภาพสูง ซึ่งเหมาะสำหรับการประมวลผลในสภาพแวดล้อมที่มีข้อจำกัดด้านทรัพยากร ดังนั้นจากการผสมผสานระหว่างการเรียนรู้เชิงลึกและการตรวจจับวัตถุจะช่วยให้โมเดลสามารถระบุสายพันธุ์แคคตัสจากภาพถ่ายที่มีความแตกต่างหลากหลายรูปแบบ เช่น ขนาดของวัตถุ ทิศทางของวัตถุ และแสงสว่างได้อย่างมีประสิทธิภาพ

การพัฒนาชุดข้อมูลรูปภาพแคคตัส 10 สายพันธุ์ รวม 1,508 ภาพ พร้อมกำกับข้อมูลด้วยกรอบ Bounding box ผ่านโปรแกรม LabelImg และบันทึกคำอธิบายรูปภาพในรูปแบบ XML ผลการฝึกฝนโมเดล พบว่า ค่าความแม่นยำเฉลี่ยบนชุดข้อมูลฝึกฝน และชุดข้อมูลตรวจสอบอยู่ที่ 92.64% และ 91.05% ตามลำดับ อยู่ในระดับดีมาก เช่นเดียวกับค่าการสูญเสียเฉลี่ยบนชุดข้อมูลฝึกฝน และชุดข้อมูลตรวจสอบอยู่ที่ 0.1853 และ 0.1955 ตามลำดับ ซึ่งบ่งชี้ว่าโมเดลสามารถเรียนรู้รูปแบบของข้อมูลได้อย่างมีประสิทธิภาพ และไม่มีสัญญาณของ Overfitting หรือ Underfitting อย่างมีนัยสำคัญ อีกทั้งยังสามารถนำไปใช้งานกับข้อมูลใหม่ได้อย่างมีประสิทธิภาพ เมื่อพิจารณาผลการตรวจจับวัตถุบนชุดข้อมูลตรวจสอบ พบว่า ตัวแบบที่พัฒนาขึ้นมีค่าความแม่นยำเฉลี่ย และค่าการเรียกคืนเฉลี่ยได้สูงถึง 0.991 และ 0.993 ตามลำดับ ในการจำแนกและตรวจจับวัตถุขนาดใหญ่ในภาพได้อย่างถูกต้องในสภาพแสงและมุมมองที่แตกต่างกัน อย่างไรก็ตาม ผลการทดสอบยังพบข้อจำกัดในการตรวจจับวัตถุขนาดกลางและขนาดเล็ก ซึ่งเป็นปัญหาที่พบได้ทั่วไปในโมเดลตรวจจับวัตถุที่ต้องทำงานกับข้อมูลที่มีความหลากหลายของขนาดวัตถุ ดังนั้นยังคงต้องได้รับการปรับปรุงและพัฒนาเพิ่มเติมในอนาคต

ผลการประเมินความถูกต้องจากผู้เชี่ยวชาญด้วยชุดข้อมูลที่โมเดลไม่เคยเห็นมาก่อน เพื่อทดสอบประสิทธิภาพของตัวแบบในสถานการณ์จริง พบว่า มีค่าความถูกต้องเฉลี่ยอยู่ที่ 83.33% นอกจากนี้ พบว่า ตัวแบบสามารถทำงานได้ดีในการตรวจจับวัตถุในภาพที่มีเพียงหนึ่งชิ้นแต่สามารถจำแนกได้เป็นหลายคลาส เช่นเดียวกับกรณีที่มีหลายวัตถุแต่อยู่ในคลาสเดียวกัน แต่ค่าความถูกต้องจะลดลงเมื่อในภาพประกอบด้วยวัตถุหลายชนิดที่สามารถจำแนกได้เป็นหลายคลาส แต่ผลลัพธ์ที่ได้ยังถือว่ามีประสิทธิภาพสูงและสามารถนำไปเป็นต้นแบบสำหรับแอปพลิเคชันทางการเกษตรเชิงพาณิชย์ได้

อย่างไรก็ดี คุณลักษณะของแคคตัสแต่ละสายพันธุ์ที่มีความคล้ายคลึงกันนั้นส่งผลโดยตรงต่อความคลาดเคลื่อนในการระบุสายพันธุ์ โดยเฉพาะในกรณีที่ขนาดและรูปร่างของลำต้น ลักษณะหนาม หรือสีสันที่ใกล้เคียงกัน อาจส่งผลให้โมเดลไม่สามารถเรียนรู้ความแตกต่างระหว่างสายพันธุ์ได้อย่างมีประสิทธิภาพ ตลอดจนหากภาพมีความละเอียดต่ำหรือมีแสงสว่างที่ไม่เหมาะสม จะส่งผลให้ประสิทธิภาพการจำแนกสายพันธุ์ลดลงอย่างมาก และหากในชุดข้อมูลมีภาพบางสายพันธุ์มากเกินไปในขณะที่สายพันธุ์อื่นมีน้อย อาจทำให้โมเดลมีแนวโน้มจะจดจำคุณลักษณะของสายพันธุ์ที่ปรากฏบ่อยกว่า ดังนั้นการแก้ปัญหาความคลาดเคลื่อนนี้จำเป็นต้องปรับปรุงคุณภาพของชุดข้อมูล ตลอดจนการปรับแต่งโมเดลเพื่อให้สามารถเรียนรู้คุณลักษณะเชิงลึกได้ดีขึ้น อันจะส่งผลต่อประสิทธิภาพในการจำแนกข้อมูลและการตรวจจับวัตถุต่อไป

ข้อเสนอแนะ

งานวิจัยครั้งต่อไปควรพิจารณานำเทคนิคและวิธีการอื่น ๆ เข้ามาช่วยเพิ่มประสิทธิภาพ เช่น เทคนิค Feature Pyramid Network: FPN ที่ช่วยให้โมเดลสามารถเรียนรู้และจำแนกคุณลักษณะที่ซับซ้อนในหลายระดับของภาพได้ดียิ่งขึ้น โดยเฉพาะในกรณีของวัตถุที่มีขนาดแตกต่างกัน ร่วมกับการนำ Attention Mechanism มาใช้เพื่อกำหนดความสำคัญกับพื้นที่ในภาพที่มีข้อมูลสำคัญ เช่น บริเวณหนามหรือสีสันเฉพาะของแต่ละสายพันธุ์ จะทำให้โมเดลสามารถโฟกัสที่คุณลักษณะสำคัญที่จำเป็นในการจำแนกสายพันธุ์ได้ดีขึ้น นอกจากนั้นควรขยายขอบเขตการวิจัยให้ครอบคลุมแคคตัสสายพันธุ์อื่น ๆ เนื่องจากการศึกษาที่จำกัดเฉพาะบางสายพันธุ์อาจทำให้โมเดลขาดความสามารถในการจำแนกแคคตัสที่มีลักษณะใกล้เคียงกันหรือสายพันธุ์ที่ไม่ได้รวมอยู่ในการฝึกฝน อีกทั้งการเพิ่มจำนวนสายพันธุ์ยังช่วยให้สามารถพัฒนาโมเดลที่มีความสามารถในการจำแนกแคคตัสในระดับสกุล (genus) หรือในระดับชนิด (species) ได้อย่างแม่นยำมากขึ้น โดยเฉพาะในสายพันธุ์ที่มีความหลากหลายทางสัณฐานวิทยา (morphological diversity) หรือสายพันธุ์ที่มีลักษณะคล้ายคลึงกัน เช่น แคคตัสในสกุล Mammillaria ซึ่งมีความหลากหลายของรูปร่าง ขนาด และลักษณะของหนาม การศึกษาที่ครอบคลุมจะช่วยให้โมเดลสามารถเรียนรู้และแยกแยะความแตกต่างของสายพันธุ์ที่คล้ายกันได้ดีขึ้น และเป็นประโยชน์ในการนำไปพัฒนาระบบการจำแนกสายพันธุ์แคคตัสในสถานการณ์จริงต่อไป

References

Ali, M., Keller, C., & Huang, M. (2023). Fruits detection using single shot multibox detector. *In BSCI '23: Proceedings of the 5th ACM International Symposium on Blockchain and Secure Critical Infrastructure* (pp. 140–144). New York: Association for Computing Machinery

Anderson, E. F. (2001). *The cactus family*. Portland, OR: Timber Press.

Ariwa, R. N., Markus, C., Teneke, N. G., Adamu, S., & George, F. K. (2024). Plant disease detection using YOLO machine learning approach. *British Journal of Computer, Networking and Information Technology, 7*(2), 115–129. https://doi.org/10.52589/BJCNIT-EJWGFW6D

Arun, A., Sharma, S., Singh, B., & Hazra, T. (2025). Identification of plant species using convolutional neural network with transfer learning. *Journal of Phytopathology, 173*(1). e70032. https://doi.org/10.1111/jph.70032

Barthlott, W., & Hunt, D. R. (1993). *Cacti: An illustrated guide to the plants and their cultivation*. Portland, OR: Timber Press.

Business Research Insights. (2023). *Succulent plant market report* (Market Report No. 100171). Retrieved from https://www.businessresearchinsights.com/market-reports/succulent-plant-market-100171

Girshick, R. (2015). Fast R-CNN. *In Proceedings of the IEEE International Conference on Computer Vision (ICCV)* (pp. 1440–1448). Santiago: IEEE

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. Cambridge, MA: MIT Press.

Gong, X., & Zhang, S. (2023). An analysis of plant diseases identification based on deep learning methods. *The Plant Pathology Journal, 39*(4), 320–324. https://doi.org/10.5423/PPJ.OA.02.2023.0034

Kaur, S., & Kaur, P. (2019). Plant species identification based on plant leaf using computer vision and machine learning techniques. *Journal of Multimedia Information System, 6*(2), 49–59. https://doi.org/10.33851/JMIS.2019.6.2.49

Konyushkova, K., Pishchulin, L., & Matuszek, C. (2018). Learning intelligent dialogs. *In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 1–10). UT: IEEE

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., & Berg, A. C. (2016). SSD: Single shot multibox detector. In B. Leibe et al. (Eds.), *Proceedings of the European Conference on Computer Vision (ECCV 2016)* (pp. 21–37). Amsterdam: Springer International Publishing.

Nobel, P. S. (2002). *Cacti: Biology and uses*. CA: University of California Press.

Rani, A., Nikhila, M., Bhargavi, M., & Sushma, P. (2024). Supervised learning based plant species classification for precise e-agriculture. *Journal for Educators, Teachers and Trainers, 15*(5), 405–415. Retrieved from https://hdl.handle.net/10481/103749

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 779–788). NV: IEEE

Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L.-C. (2018). *MobileNetV2: Inverted residuals and linear bottlenecks. arXiv, 1801*(04381), 4510-4520. https://doi.org/10.48550/arXiv.1801.04381

Thomkaew, J., & Intakosum, S. (2023). Plant species classification using leaf edge feature combination with morphological transformations and SIFT key point. *Journal of Image and Graphics, 11*(1), 91–97. https://doi.org/10.18178/joig.11.1.91-97

Verified Market Research. (2021). *Cactus extract market – Global industry trends and forecast to 2027.* Retrieved from https://www.verifiedmarketresearch.com/product/cactus-extract-market/

Wu, H., Fang, L., Yu, Q., & Yang, C. (2024). Composite descriptor based on contour and appearance for plant species identification. *Engineering Applications of Artificial Intelligence, 133*(3), 108291. https://doi.org/10.1016/j.engappai.2024.108291

Zhang, X., Yang, Y.-H., Han, Z., Wang, H., & Gao, C. (2014). Object class detection: A survey. *ACM Computing Surveys, 46*(1), Article 10, 1–53. https://doi.org/10.1145/2522968.2522978