

การจำแนกสายพันธุ์กระบองเพชรด้วยโมเดลการเรียนรู้เชิงลึก

จักรินทร์ สันติรัตนภักดี^{1,*}, นิชาภัทร ตูลาธาร²

^{1,2} สาขาวิชาเทคโนโลยีธุรกิจดิจิทัล คณะบริหารธุรกิจ มหาวิทยาลัยวงษ์ชวลิตกุล

Received: 22 September 2024

Revised: 28 November 2024

Accepted: 4 December 2024

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาโมเดลการจำแนกสายพันธุ์กระบองเพชรด้วยการเรียนรู้เชิงลึก และเพื่อประเมินความถูกต้องของโมเดลการจำแนกสายพันธุ์กระบองเพชร โดยใช้ 3 โมเดลการเรียนรู้เชิงลึกที่ได้รับความนิยม ได้แก่ ResNet34, VGG16 และ MobileNetV3 ที่ฝึกฝนบนชุดข้อมูลรูปภาพกระบองเพชร 10 สายพันธุ์ที่ได้รับความนิยมจากผู้บริโภคจากข้อมูลของร้านขายกระบองเพชรในพื้นที่อำเภอเมืองนครราชสีมา จังหวัดนครราชสีมา จำนวน 4,512 ภาพ ที่ผ่านการดัดแปลงรูปภาพจากต้นฉบับ เพื่อเพิ่มขนาดของชุดข้อมูลโดยการแปลงรูปภาพเดิมให้หลากหลายมากขึ้น ช่วยให้โมเดลเรียนรู้ได้ดีขึ้น และไม่จำเพาะเจาะจงกับรูปภาพแบบใดแบบหนึ่งมากเกินไป กระบวนการฝึกฝนโมเดล โดยการใช้ค่าไฮเปอร์พารามิเตอร์ ร่วมกับการใช้เทคนิคการเรียนรู้แบบรอบเดียว (One cycle policy) ในการปรับอัตราการเรียนรู้และโมเมนตัม เพื่อปรับค่าน้ำหนักของโครงข่ายประสาทเทียมด้วยกระบวนการถ่ายทอดย้อนกลับ ผลการประเมินความถูกต้องบนชุดข้อมูลตรวจสอบ พบว่า โมเดล MobileNetV3 มีค่าความถูกต้องสูงสุดที่ 91.36% เช่นเดียวกับ ค่าความแม่นยำ ค่าการเรียกคืน และค่าประสิทธิภาพโดยรวม เท่ากับ 91.46%, 91.36% และ 91.20% ตามลำดับ เมื่อทดสอบกับชุดข้อมูลทดสอบที่โมเดลไม่เคยเห็นมาก่อนจำนวน 400 รูปภาพ เพื่อทดสอบประสิทธิภาพของโมเดล ในสถานการณ์ใช้งานจริง โดยผ่านการประเมินความถูกต้องโดยผู้เชี่ยวชาญ พบว่า สามารถจำแนกสายพันธุ์กระบองเพชรได้ถูกต้องถึง 81.50%

คำสำคัญ: การจำแนกรูปภาพ สายพันธุ์กระบองเพชร การเรียนรู้เชิงลึก

* ผู้ประสานงานหลัก; อีเมล: chakkarin_san@vu.ac.th

Classification of Cactus Species Using Deep Learning Model

Chakkarin Santirattanaphakdi^{1,*}, Nichapat Tulathan²

^{1,2} Department of Digital Business Technology, Faculty of Business Administration,
Vongchavalitkul University

Received: 22 September 2024

Revised: 28 November 2024

Accepted: 4 December 2024

Abstract

This research aims to develop a deep learning model for cactus species classification and evaluate its accuracy. Three popular deep learning models—ResNet34, VGG16 and MobileNetV3—are trained on a dataset of 4,512 images representing 10 cactus species that are popular among consumers, based on data from cactus shops in Mueang Nakhon Ratchasima district, Nakhon Ratchasima province. The dataset is augmented through various image transformations to increase its size and diversity, allowing the model to learn better and avoid overfitting. The training process includes hyperparameter tuning and the application of the one-cycle policy to adjust the learning rate and momentum for optimal weight adjustment using backpropagation. The evaluation results on the validation dataset show that MobileNetV3 achieves the highest accuracy of 91.36%, with precision, recall and f-measure values of 91.46%, 91.36%, and 91.20%, respectively. When tested on an unseen dataset of 400 images to assess the model's performance in real-world scenarios, the expert-verified cactus species classification accuracy reaches 81.50%.

Keywords: Image Classification, Cactus Species, Deep Learning Model

* Corresponding Author; E-mail: chakkarin_san@vu.ac.th

ความเป็นมาและความสำคัญของปัญหา

กระบองเพชรนับว่าเป็นหนึ่งในไม้ประดับที่ได้รับความนิยมจากทั่วโลก อันเนื่องมาจากเอกลักษณ์ อันโดดเด่น และความหลากหลายทางสายพันธุ์ ประกอบกับเป็นพืชที่สามารถทนทานต่อสภาพอากาศแห้งแล้ง ง่ายต่อการดูแลรักษา จึงเป็นที่นิยมทั้งในกลุ่มนักสะสมและเกษตรกรในหลายประเทศ (Anderson, 2001) เช่น เม็กซิโก สหรัฐอเมริกา และประเทศในแถบเมดิเตอร์เรเนียนที่มีการซื้อขายอย่างแพร่หลาย เช่นเดียวกับ ตลาดกระบองเพชรในแถบเอเชียตะวันออกเฉียงใต้ โดยเฉพาะอย่างยิ่งในประเทศไทยก็เติบโตอย่างรวดเร็วเช่นกัน ด้วยความนิยมในการปลูกกระบองเพชรเป็นงานอดิเรกและในเชิงพาณิชย์ (Shmida & Burgess, 1988) นอกจากนี้ การประกวดกระบองเพชรและไม้อวบน้ำในประเทศไทยยังมีบทบาทสำคัญในการส่งเสริมการเพาะปลูกกระบองเพชรสายพันธุ์ต่าง ๆ ให้แพร่หลายมากยิ่งขึ้น ส่งผลให้ตลาดการซื้อขายมีมูลค่าเติบโตอย่างต่อเนื่องในช่วงหลายทศวรรษที่ผ่านมา และมีแนวโน้มที่จะเพิ่มขึ้นอีกในอนาคต

อย่างไรก็ดี ความหลากหลายของสายพันธุ์กระบองเพชรที่มีมากกว่า 1,500 สายพันธุ์ทั่วโลก โดยมีลักษณะต่าง ๆ ที่เป็นเอกลักษณ์เฉพาะตัว เช่น รูปทรงหนาม สีลำต้น และขนาดลำต้นที่แตกต่างกัน (Hunt et al., 2006) ส่งผลให้การระบุสายพันธุ์กระบองเพชรสำหรับผู้เริ่มต้นเป็นสิ่งที่ท้าทายเนื่องจากความคล้ายคลึงกันทางสัณฐานวิทยา (Morphology) ที่อาจทำให้เกิดความสับสน เช่น สายพันธุ์ Echinopsis ที่มีลักษณะเด่นคือหนามและลักษณะลำต้นที่คล้ายคลึงกันกับสายพันธุ์ Mammillaria (Barthlott & Hunt, 1993) จะเห็นได้ว่าการจำแนกสายพันธุ์กระบองเพชรที่ไม่ถูกต้องอาจส่งผลเสียต่อการเพาะปลูก เนื่องจากกระบองเพชรแต่ละสายพันธุ์ต้องการการดูแลที่แตกต่างกัน เช่น ปริมาณน้ำ แสงแดด และสภาพดินที่เหมาะสม ซึ่งส่งผลต่อการเจริญเติบโตของกระบองเพชร หรืออาจกระทบต่อคุณภาพของผลผลิต ด้วยความซับซ้อนในการจำแนกสายพันธุ์ จึงเป็นที่มาของการนำเทคโนโลยีเข้ามาช่วยในการระบุสายพันธุ์กระบองเพชรอย่างถูกต้อง

การเรียนรู้เชิงลึก (Deep learning) นับเป็นหนึ่งในแนวทางที่มีประสิทธิภาพสูงในการจำแนกรูปภาพ (Image classification) ด้วยความสามารถในการเรียนรู้คุณลักษณะจากข้อมูลจำนวนมากโดยไม่ต้องพึ่งพาการออกแบบคุณลักษณะ (Feature engineering) อย่างละเอียด ดังนั้นจึงมีการนำมาใช้กับการจำแนกรูปภาพในหลากหลายด้าน รวมถึงด้านการเกษตรและพืชพันธุ์ (Kamilaris & Prenafeta-Boldú, 2018) ด้วยจุดเด่นจากการถ่ายโอนความรู้ (Transfer learning) และการปรับแต่งเฉพาะด้าน (Fine-tuning) เพื่อใช้กับการจำแนกคุณลักษณะที่ซับซ้อน (LeCun et al., 2015) จะเห็นได้ว่าการเรียนรู้เชิงลึกเป็นเทคนิคที่มีประสิทธิภาพสูงจึงถูกนำมาใช้งานอย่างแพร่หลายในการจำแนกภาพได้อย่างรวดเร็วและมีประสิทธิภาพ

อย่างไรก็ดี เนื่องจากสายพันธุ์ของกระบองเพชรมีเป็นจำนวนมาก ดังนั้นการตรวจสอบสายพันธุ์ของกระบองเพชรจึงมีความสำคัญ ในปัจจุบันการจำแนกสายพันธุ์กระบองเพชรสามารถทำได้จากลักษณะเฉพาะของแต่ละสายพันธุ์ เช่น รูปทรงหนาม สีของลำต้น และขนาดที่แตกต่างกัน อีกทั้งการเติบโตในสภาพแวดล้อมที่แตกต่างกันยังอาจทำให้ลักษณะเฉพาะของสายพันธุ์เปลี่ยนแปลงไป จึงต้องอาศัยผู้เชี่ยวชาญที่มีประสบการณ์เข้ามาช่วยจำแนก แต่ผลกระทบจากวิธีการดังกล่าวคือสิ้นเปลืองเวลา และมีความยุ่งยากซับซ้อน โดยเฉพาะอย่างยิ่งกับผู้เริ่มต้นเพาะเลี้ยงกระบองเพชร จากเหตุผลดังกล่าว คณะผู้วิจัยจึงได้นำเสนอการพัฒนาโมเดลการจำแนกสายพันธุ์กระบองเพชรด้วยการเรียนรู้เชิงลึก เพื่อใช้ในการวิเคราะห์และจำแนกสายพันธุ์กระบองเพชรจากรูปภาพ ผลลัพธ์จากงานวิจัยนี้จะเป็นประโยชน์ต่อผู้ที่ต้องการจำแนกสายพันธุ์กระบองเพชร เนื่องจากสามารถจำแนกสายพันธุ์ได้ตรงตามลักษณะ พร้อม

แสดงผลการจำแนกได้ทันที อีกทั้งยังมีการเชื่อมโยงไปยังข้อมูลเบื้องต้นของกระบอกเพชรสายพันธุ์นั้น ๆ เพื่อเสริมความรู้แก่ผู้ใช้งานเพิ่มเติมอีกด้วย

วัตถุประสงค์ของการวิจัย

1. เพื่อพัฒนาโมเดลการจำแนกสายพันธุ์กระบอกเพชรด้วยการเรียนรู้เชิงลึก
2. เพื่อประเมินความถูกต้องของโมเดลการจำแนกสายพันธุ์กระบอกเพชร

เอกสารและงานวิจัยที่เกี่ยวข้อง

การจำแนกรูปภาพสามารถทำได้ด้วยหลากหลายวิธีการ ซึ่งหลายเทคนิคถูกใช้ก่อนที่การเรียนรู้เชิงลึกจะกลายเป็นแนวทางหลักในการจำแนกรูปภาพ เทคนิคเหล่านี้มักมุ่งเน้นการวิเคราะห์พีเจอร์พื้นฐานในภาพ และใช้วิธีการทางสถิติหรืออัลกอริทึมการเรียนรู้ของเครื่องแบบดั้งเดิมเพื่อช่วยในการจำแนกรูปภาพ ตัวอย่างของเทคนิคที่ได้รับความนิยม ได้แก่

1) การสร้างพีเจอร์ด้วยถุงคำภาพ (Bag of Visual Words: BoVW) แนวคิดหลักคือการเปลี่ยนข้อมูลภาพเป็นคำ และนำไปสร้างพีเจอร์เวกเตอร์ ซึ่งอาศัยแนวคิดแบบเดียวกับการใช้ถุงคำ (Bag of Words: BoW) ในการประมวลผลภาษาธรรมชาติ โดย Sivic & Zisserman (2003) นำเสนอการใช้ถุงคำภาพ เพื่อจับคู่วัตถุในวิดีโอโดยวิธีการแยกพีเจอร์จากวัตถุ แล้วนำมาสร้างเป็นคลังคำภาพ (Dictionary of visual words) เพื่อใช้ระบุสิ่งที่ปรากฏในแต่ละวิดีโอ ก่อนจะถูกนำไปประยุกต์ใช้กับการจำแนกภาพนิ่งในหลากหลายโดเมน

2) การแปลงพีเจอร์ที่คงทนต่อการเปลี่ยนสเกล (Scale-Invariant Feature Transform: SIFT) และ พีเจอร์ที่เสริมความเร็วและคงทนต่อการเปลี่ยนแปลง (Speeded-Up Robust Features: SURF) ซึ่งเป็นเทคนิคการสกัดพีเจอร์ที่คงทนต่อการเปลี่ยนแปลงของภาพ (Image transformation) เช่น การย่อขยายภาพ และการหมุนภาพ โดย Lowe (1999) นำเสนอการใช้การแปลงพีเจอร์ที่คงทนต่อการเปลี่ยนสเกลในการตรวจจับพีเจอร์เฉพาะจุดในภาพ ซึ่งสามารถใช้จำแนกวัตถุได้แม้ภาพจะเปลี่ยนแปลงไปในบางแง่มุม ต่อมา Bay et al. (2006) ได้พัฒนาพีเจอร์ที่เสริมความเร็วและคงทนต่อการเปลี่ยนแปลง ซึ่งเพิ่มประสิทธิภาพ ที่ทำให้การตรวจจับพีเจอร์เร็วขึ้น และเป็นที่ยอมรับในด้านการจำแนกรูปภาพในเวลาต่อมา

3) การสกัดพีเจอร์ด้วยฮิสโตแกรมของทิศทางการไล่ระดับ (Histograms of Oriented Gradients: HOG) เป็นเทคนิคการสกัดพีเจอร์ที่อาศัยฮิสโตแกรมของทิศทางการเปลี่ยนแปลงของพิกเซลในภาพเพื่อระบุรูปร่างและลักษณะของวัตถุ โดย Dalal & Triggs, (2005) ใช้เทคนิคฮิสโตแกรมของทิศทางการไล่ระดับในการตรวจจับมนุษย์จากรูปภาพ ทำให้ได้พีเจอร์ที่เด่นชัดและไม่ไวต่อแสงเงาของภาพมากนัก จึงถูกใช้อย่างกว้างขวางในการจำแนกและตรวจจับวัตถุต่าง ๆ โดยเฉพาะในภาพนิ่งและวิดีโอ

4) ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machines: SVM) เป็นอัลกอริทึมที่มักใช้ร่วมกับพีเจอร์ที่สกัดจากเทคนิคพื้นฐานอื่น ๆ สำหรับการจำแนกรูปภาพ โดย Boser et al. (1992) นำเสนอการใช้ซัพพอร์ตเวกเตอร์แมชชีนเป็นส่วนหนึ่งในการจำแนกข้อมูลด้วยการเพิ่มระยะห่างระหว่างข้อมูลของแต่ละคลาส ทำให้การจำแนกมีความถูกต้องสูงขึ้น ส่งผลให้ถูกนำมาใช้ร่วมกับการสกัดพีเจอร์ต่าง ๆ สำหรับการจำแนกรูปภาพที่มีความซับซ้อนในภายหลัง

5) ป่าการตัดสินใจแบบสุ่มและต้นไม้การตัดสินใจ (Random forests and decision trees) ซึ่งถูกนำมาใช้ในการจำแนกรูปภาพโดยอาศัยการสร้างต้นไม้ตัดสินใจจากพีเจอร์ทึ่มพื้นฐาน เช่น สี รูปร่าง และพื้นผิว โดย Breiman (2001) นำเสนอป่าการตัดสินใจแบบสุ่ม ซึ่งเป็นการรวมเอาการตัดสินใจแบบหลายเงื่อนไขจาก ต้นไม้การตัดสินใจมารวมกัน เพื่อให้ได้ผลลัพธ์ที่ถูกต้องและมีความทนทานต่อการเปลี่ยนแปลงของภาพ

แม้เทคนิคเหล่านี้จะยังมีการใช้งานอยู่บ้างในบางกรณี แต่พัฒนาการของการเรียนรู้เชิงลึกได้ทำให้วิธีการเหล่านี้ค่อย ๆ ถูกลบทบหาในการจำแนกรูปภาพ เนื่องจากการเรียนรู้เชิงลึกสามารถสร้างพีเจอร์ทึ่มที่ซับซ้อนและเรียนรู้จากข้อมูลได้โดยอัตโนมัติ ซึ่งได้รับการยอมรับอย่างกว้างขวางว่ามีประสิทธิภาพมากขึ้นเมื่อเปรียบเทียบกับวิธีการจำแนกรูปภาพแบบเดิม จากความสามารถในการเรียนรู้พีเจอร์ทึ่มที่ซับซ้อนได้โดยอัตโนมัตินอกจากนี้ โมเดลที่ได้รับการฝึกฝนบนชุดข้อมูลขนาดใหญ่ยังสามารถถ่ายโอนมาใช้ในการกับจำแนกรูปภาพในชุดข้อมูลที่เล็กลงได้อย่างมีประสิทธิภาพ จึงช่วยลดระยะเวลาและทรัพยากรในการฝึกฝนโมเดลใหม่ตั้งแต่ต้น (Yosinski et al., 2014) อีกทั้งการปรับแต่งโมเดลเพื่อใช้งานเฉพาะทางจากการปรับน้ำหนักของโมเดลบางส่วนหรือทั้งหมดให้เหมาะสมกับลักษณะเฉพาะของชุดข้อมูลเฉพาะด้าน (Yosinski et al., 2014) จากที่กล่าวมาแสดงถึงประสิทธิภาพของการเรียนรู้เชิงลึกในการจำแนกรูปภาพ งานวิจัยนี้มุ่งศึกษาการจำแนกรูปภาพด้วย 3 โมเดล ประกอบด้วย

1) โมเดล ResNet34 (Residual Network) (He et al., 2016) เป็นหนึ่งในสถาปัตยกรรมที่สำคัญจากพัฒนาการของรูปแบบของโครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional Neural Networks: CNN) โดยเฉพาะสำหรับงานที่ต้องใช้เครือข่ายที่มีความลึกมากด้วยคุณลักษณะสำคัญคือการเชื่อมต่อส่วนต่าง (Residual connections) ระหว่างผลลัพธ์ที่คาดหวังกับค่าที่ได้ ซึ่งใช้ในบริบทของการเพิ่มการเชื่อมต่อโดยตรงเพื่อลดปัญหาการสูญเสียมูลค่า ซึ่งมีส่วนสำคัญในการช่วยให้โครงข่ายสามารถเรียนรู้ได้อย่างมีประสิทธิภาพมากขึ้น แม้ในสภาพแวดล้อมที่ซับซ้อน ร่วมกับเทคนิคการเชื่อมต่อแบบข้ามชั้น (Skip connections) ของบางส่วนในเครือข่าย โดยข้อมูลจะถูกส่งผ่านไปยังชั้นที่ลึกกว่าทันที จึงสามารถเชื่อมต่อข้อมูลระหว่างชั้นที่อยู่ห่างกันหลายชั้นส่งผ่านข้อมูลแบบตรง (Direct mapping) ซึ่งช่วยลดปัญหาการสูญหายของกราเดียนต์ (Vanishing gradients) ที่มักเกิดขึ้นเมื่อเครือข่ายมีความลึกมากโดยโมเดล ResNet34 ประกอบไปด้วย 34 ชั้น โดยใช้โครงสร้างของบล็อกส่วนต่าง (Residual blocks) ที่ช่วยเพิ่มความสามารถในการเรียนรู้พีเจอร์ทึ่มที่มีความหลากหลาย และซับซ้อนด้วยการถ่ายทอดย้อนกลับทำให้สามารถปรับน้ำหนักในแต่ละชั้นให้เหมาะสมยิ่งขึ้น ส่งผลให้โมเดลมีประสิทธิภาพสูงในการจำแนกรูปภาพ

2) โมเดล VGG16 (Simonyan & Zisserman, 2014) ถือเป็นหนึ่งในสถาปัตยกรรมโครงข่ายประสาทเทียมแบบคอนโวลูชันที่มีการออกแบบโครงสร้างอย่างเป็นระบบ เพื่อให้สามารถวิเคราะห์และดึงพีเจอร์ทึ่มที่สำคัญจากภาพได้อย่างมีประสิทธิภาพ คุณลักษณะสำคัญคือการใช้ชั้นคอนโวลูชัน (Convolutional layers) ขนาด 3*3 และชั้นพูลลิง (Pooling layers) ขนาด 2*2 เพื่อให้สามารถดึงพีเจอร์ทึ่มจากภาพได้อย่างมีประสิทธิภาพและลดขนาดของภาพลงเป็นชั้น ๆ ดังนั้นหลังจากการประมวลผลในชั้นคอนโวลูชัน ข้อมูลจะถูกส่งผ่านไปยังชั้นเชื่อมต่อแบบสมบูรณ์ (Fully connected layers) ซึ่งจะทำหน้าที่จำแนกภาพตามคลาสของความน่าจะเป็น ซึ่งวิธีการนี้ยังใช้กระบวนการถ่ายทอดย้อนกลับเพื่อคำนวณค่าความผิดพลาดจากชั้นผลลัพธ์ (Output layer) และปรับพารามิเตอร์ของแต่ละชั้นในเครือข่ายอย่างต่อเนื่อง ซึ่งช่วยให้การเรียนรู้และปรับปรุงโมเดลมีประสิทธิภาพยิ่งขึ้น โดยโมเดล VGG16 ประกอบด้วย 16 ชั้น แบ่งเป็นชั้นคอนโวลูชันจำนวน 13 ชั้น และชั้นเชื่อมต่อแบบสมบูรณ์อีก 3 ชั้นจากโครงสร้างที่เป็นระบบ และมีมาตรฐานนี้ทำให้โมเดล VGG16 ได้รับความนิยมใช้ในงานจำแนกรูปภาพที่มีความซับซ้อนสูง

3) โมเดล MobileNetV3 เป็นอีกหนึ่งในสถาปัตยกรรมโครงข่ายประสาทเทียมแบบคอนโวลูชันที่ออกแบบมาให้มีความกะทัดรัด โดยโมเดล MobileNetV3 ถูกปรับปรุงและพัฒนาสถาปัตยกรรมจากโมเดล MobileNet รุ่นก่อนหน้า โดยใช้กระบวนการค้นหาสถาปัตยกรรมโครงข่ายประสาทเทียม (Neural Architecture Search: NAS) ที่เหมาะสมที่สุดสำหรับการใช้งานบนอุปกรณ์พกพา (Howard et al., 2019) คุณสมบัติที่สำคัญคือการใช้เทคนิคคอนโวลูชันแยกส่วนตามความลึก (Depthwise separable convolutions) ซึ่งแยกการคำนวณคอนโวลูชันออกเป็นสองส่วน ได้แก่ คอนโวลูชันความลึก (Depthwise convolution) สำหรับแต่ละช่องสัญญาณของภาพ และการคอนโวลูชันแบบจุด (Pointwise convolution) ที่ใช้การคูณแบบจุดที่ใช้ในการรวมข้อมูลจากหลายช่องสัญญาณ เทคนิคนี้จะช่วยลดจำนวนพารามิเตอร์และความซับซ้อนของการคำนวณ ซึ่งทำให้สามารถประมวลผลได้อย่างรวดเร็วและมีประสิทธิภาพ นอกจากนี้ยังมีการปรับปรุงในด้านโครงสร้างให้เหมาะสมยิ่งขึ้นด้วยการใช้เทคนิคเพิ่มเติม เช่น ฟังก์ชันการกระตุ้นแบบฮาร์ดสวิช (Hard-swish activation) และบล็อกปรับน้ำหนักตามช่องสัญญาณ (Squeeze-and-excitation blocks) ที่ช่วยเพิ่มประสิทธิภาพในการทำงานโดยไม่เพิ่มความซับซ้อนในโครงสร้าง ซึ่งเน้นการประหยัดพลังงานและการใช้ทรัพยากรต่ำ จึงเหมาะสำหรับการใช้งานในอุปกรณ์ที่มีทรัพยากรจำกัด เช่น สมาร์ทโฟน หรืออุปกรณ์พกพาอื่น ๆ

ปัจจุบันมีงานวิจัยที่นำการเรียนรู้เชิงลึกมาประยุกต์ใช้ในหลากหลายบริบท ซึ่งแต่ละงานวิจัยก็มีเทคนิคและวิธีการที่แตกต่างกัน โดย Sae-Lim et al. (2019) นำเสนอแนวทางการจำแนกประเภทแผลผิวหนัง (Skin lesion) ให้มีประสิทธิภาพมากขึ้น โดยใช้โมเดลการเรียนรู้เชิงลึกบนชุดข้อมูล Human Against Machine จำนวน 10,000 รูปภาพ (HAM 1000) หลังจากนั้นได้มีการใช้เทคนิคการเพิ่มข้อมูล เพื่อสร้างความหลากหลายให้กับข้อมูลฝึกฝน ทำให้โมเดลเรียนรู้คุณลักษณะเฉพาะได้ดียิ่งขึ้น ร่วมกับการใช้เทคนิคการถ่ายโอนการเรียนรู้จากโมเดลที่ถูกฝึกฝนไว้ล่วงหน้า ซึ่งช่วยให้กระบวนการเรียนรู้รวดเร็วขึ้นและมีประสิทธิภาพมากขึ้น จากนั้น Zhu et al. (2023) พัฒนาโมเดลการจำแนกคุณภาพของเห็ดโดยใช้การเรียนรู้เชิงลึกบนชุดข้อมูลรูปภาพเห็ดที่มีพื้นหลังซับซ้อนจากแหล่งต่าง ๆ หลังจากนั้นมีการใช้เทคนิคการเพิ่มข้อมูล เพื่อสร้างความหลากหลายให้กับข้อมูลในการฝึกฝน ทำให้โมเดลสามารถเรียนรู้ลักษณะเฉพาะของเห็ดได้ดียิ่งขึ้น ในขั้นตอนการฝึกฝนโมเดลโดยใช้วิธีการถ่ายโอนการเรียนรู้โดยปรับน้ำหนักของโมเดลให้สอดคล้องกับลักษณะเฉพาะของภาพเห็ด เพื่อสร้างเครื่องมือที่สามารถช่วยในการประเมินคุณภาพของเห็ดได้อย่างมีประสิทธิภาพในวงการเกษตรและอุตสาหกรรมอาหาร เช่นเดียวกับ Hongboonmee and Thamditae (2023) ที่นำเสนอการจำแนกชนิดอัญมณีเบื้องต้นผ่านโมบายแอปพลิเคชันด้วยกระบวนการเรียนรู้เชิงลึกบนชุดข้อมูลรูปภาพของ อัญมณีแต่ละชนิดจากการใช้เทคนิคการเพิ่มข้อมูลเพื่อสร้างความหลากหลายให้กับข้อมูลฝึกฝน หลังจากนั้นจะใช้เทคนิคถ่ายโอนการเรียนรู้ และฝึกฝนโมเดลบนชุดข้อมูลที่เตรียมไว้ เพื่อลดเวลาในการฝึกฝนและเพิ่มประสิทธิภาพการจำแนกชนิดอัญมณีได้อย่างถูกต้องและรวดเร็ว

วิธีดำเนินการวิจัย

งานวิจัยนี้เป็นงานวิจัยเชิงประยุกต์เพื่อโมเดลการจำแนกสายพันธุ์กระบองเพชรด้วยการเรียนรู้เชิงลึก ก่อนจะพัฒนาเป็นแอปพลิเคชันที่สมบูรณ์ต่อไป แบ่งออกเป็น 3 ขั้นตอนหลัก ได้แก่ 1) การเตรียมรูปภาพก่อนการประมวลผล 2) การฝึกฝนโมเดลและการถ่ายทอดย้อนกลับ และ 3) การประเมินประสิทธิภาพบนชุดข้อมูลตรวจสอบ มีรายละเอียดดังนี้

1) การเตรียมรูปภาพก่อนการประมวลผล (Data preparation) โดยเริ่มจากการรวบรวมรูปภาพของกระบองเพชร 10 สายพันธุ์ที่ได้รับความนิยมจากผู้บริโภคจากข้อมูลของร้านขายกระบองเพชรในพื้นที่อำเภอเมืองนครราชสีมา จังหวัดนครราชสีมาจำนวน 4 ร้าน ประกอบด้วย Astrophytum, Lady finger, Carmenae, Yellow tower, Mammillaria bocasana, Owl eye, Mammillaria oldman, Gymnocalycium, Duweii และ Mammillaria humboldtii elegans ที่รวบรวมโดยคณะผู้วิจัยโดยใช้กล้องโทรศัพท์มือถือในการถ่ายภาพกระบองเพชรที่มีระยะห่างจากกล้องประมาณ 10-20 เซนติเมตรในทิศทาง และแสงสว่างแตกต่างกัน ขนาดของรูปภาพทั้งหมดได้รับการปรับให้มีความละเอียดเป็น 224*224 พิกเซล รวมทั้งสิ้น 4,512 ภาพ ตัวอย่างรูปภาพของกระบองเพชร 10 สายพันธุ์ที่ใช้ในงานวิจัยนี้ ดังภาพที่ 1



ภาพที่ 1 ตัวอย่างรูปภาพของกระบองเพชร 10 สายพันธุ์ที่ใช้ในงานวิจัยนี้

จากภาพที่ 1 การเตรียมรูปภาพก่อนการประมวลผล จะดำเนินการแบ่งข้อมูลด้วยวิธีการสุ่ม (Random sampling) ออกเป็นชุดข้อมูลฝึกฝน (Training set) ร้อยละ 80 สำหรับการฝึกฝนแบบจำลอง และชุดข้อมูลตรวจสอบ (Validation set) ร้อยละ 20 สำหรับตรวจสอบประสิทธิภาพของแต่ละโมเดลในการจำแนกสายพันธุ์กระบองเพชร ดังตารางที่ 1

ตารางที่ 1 จำนวนรูปภาพกระบองเพชรแต่ละสายพันธุ์

สายพันธุ์	ข้อมูลทั้งหมด (ภาพ)	ข้อมูลฝึกฝน (ภาพ)	ข้อมูลตรวจสอบ (ภาพ)
1) Astrophytum	484	387	97
2) Lady finger	474	379	95
3) Carmenae	430	344	86
4) Yellow tower	448	358	90
5) Mammillaria bocasana	433	346	87
6) Owl eye	461	369	92
7) Mammillaria oldman	440	352	88
8) Gymnocalycium	456	365	91
9) Duweii	444	355	89
10) Mammillaria humboldtii elegans	442	354	88
รวม	4,512	3,610	902

จากตารางที่ 1 การเตรียมรูปภาพก่อนการประมวลผลจะนำรูปภาพผ่านการดัดแปลงรูปภาพจากต้นฉบับ (Image augmentation) เป็นเทคนิคที่ช่วยเพิ่มขนาดของชุดข้อมูลโดยการแปลงรูปภาพเดิมให้หลากหลายมากขึ้น (Perez & Wang, 2017) ช่วยให้แบบจำลองเรียนรู้คุณลักษณะสำคัญของรูปภาพได้ดีขึ้นและไม่จำเพาะเจาะจงกับรูปภาพแบบใดแบบหนึ่งมากเกินไป ก่อนจะตรวจสอบคุณภาพและจัดการข้อมูล (Data verification and management) โดยสุ่มตรวจสอบรูปภาพจากชุดข้อมูลในสัดส่วนร้อยละ 10 ของจำนวนรูปภาพในแต่ละสายพันธุ์ เพื่อตรวจสอบป้ายกำกับว่าจำแนกสายพันธุ์กระบองเพชรได้ถูกต้องหรือไม่ หากพบว่าไม่ถูกต้องให้ดำเนินการแก้ไขใหม่

2) การฝึกฝนโมเดลและการถ่ายทอดย้อนกลับ (Model training and backpropagation) งานวิจัยนี้ตั้งค่าไฮเปอร์พารามิเตอร์ (Hyperparameter) เป็นค่าที่กำหนดไว้ล่วงหน้า (Manually set) ก่อนการฝึกฝนโมเดลและไม่สามารถเรียนรู้ได้จากข้อมูล ในการฝึกฝนโครงข่ายประสาทเทียมแบบคอนโวลูชัน ดังตารางที่ 2 เพื่อเป็นเงื่อนไขในการทดสอบให้มีความสอดคล้องกันในทุกด้าน และสามารถเปรียบเทียบผลลัพธ์ได้อย่างน่าเชื่อถือ

ตารางที่ 2 การตั้งค่าไฮเปอร์พารามิเตอร์หลักในการการฝึกฝนโครงข่ายประสาทเทียมแบบคอนโวลูชัน

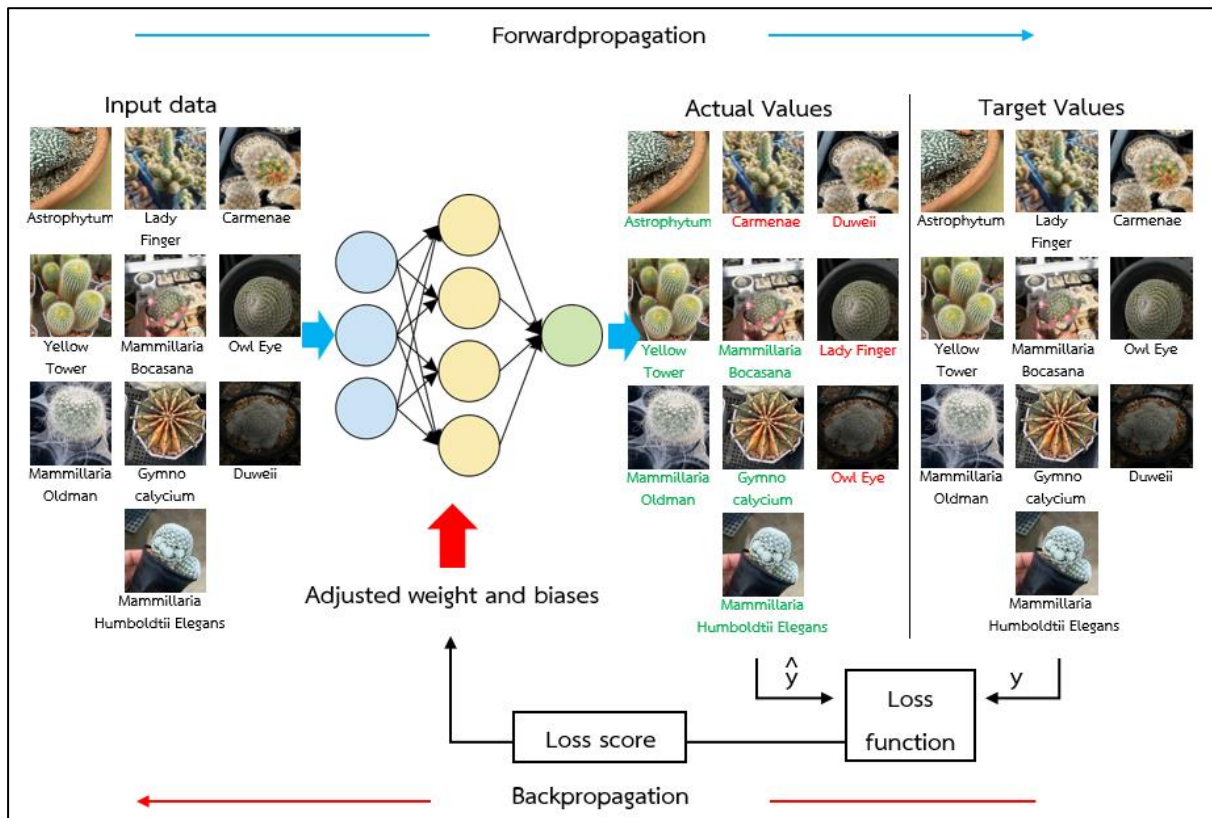
ไฮเปอร์พารามิเตอร์	ค่าที่กำหนด
No. of convolutional layers	2
No. of pooling layers	2
Activation function for convolutional layers	Relu
Stride	1
Padding	Same
Hidden layers	3

ตารางที่ 2 (ต่อ)

ไฮเปอร์พารามิเตอร์	ค่าที่กำหนด
Activation function for FC layers	Relu
Activation function for the output layer	Softmax
Loss function	Categorical crossentropy
Metric	Accuracy
Optimizer	Adam
epochs	10
Batch size	128

จากตารางที่ 2 การตั้งค่าไฮเปอร์พารามิเตอร์ ประกอบด้วย จำนวนของชั้นคอนโวลูชัน (No. of convolutional layers) ซึ่งเป็นชั้นหลักที่ช่วยดึงคุณลักษณะ (Features extraction) สำคัญจากภาพ จำนวนของชั้นพูลลิง (No. of pooling layers) เพื่อช่วยลดขนาดของข้อมูล (Down-sampling) และเพิ่มความคงทนต่อความแปรปรวนในข้อมูลภาพ ฟังก์ชันกระตุ้นสำหรับชั้นคอนโวลูชัน (Activation function for convolutional layers) ซึ่งช่วยเพิ่มความไม่เป็นเชิงเส้น (Non-linear) ในโครงข่ายและช่วยให้การเรียนรู้มีประสิทธิภาพ สไตรด์ (Stride) คือการเคลื่อนที่ของฟิลเตอร์ในแต่ละครั้งหน่วยเป็นพิกเซล แพดดิ้ง (Padding) โดยเติม 0 เข้าไป เพื่อรักษารูปร่างของภาพในแต่ละส่วนให้อยู่ในขนาดเดิม (Same padding) จำนวนชั้นซ่อนเร้น (Hidden layers) ใช้สำหรับการประมวลผลข้อมูลในลำดับที่ซับซ้อนยิ่งขึ้น ฟังก์ชันกระตุ้นสำหรับชั้นเชื่อมต่อสมบูรณ์ (Activation function for FC layers) ใช้ฟังก์ชัน ReLU เพื่อเพิ่มความไม่เป็นเชิงเส้นในการประมวลผล ฟังก์ชันกระตุ้นสำหรับชั้นผลลัพธ์ (Activation function for the output layer) ด้วยฟังก์ชันซอฟต์แวร์แมกซ์ ซึ่งเหมาะสำหรับการจำแนกประเภทแบบหลายคลาส (Multi-class classification) ฟังก์ชันความสูญเสีย (Loss function) ด้วย Categorical crossentropy ซึ่งเหมาะสำหรับวัดความผิดพลาดในการจำแนกประเภท มาตรวัด (Metric) ด้วยค่าความถูกต้องเป็นมาตรวัดหลัก ออปติไมเซอร์ (Optimizer) โดยใช้ Adam ซึ่งเป็นอัลกอริทึมการปรับค่าพารามิเตอร์ที่มีประสิทธิภาพสูง จำนวนรอบการฝึกฝน (epochs) ที่ 10 รอบ เพื่อให้โครงข่ายสามารถเรียนรู้จากข้อมูลได้อย่างเหมาะสม และขนาดของแบตช์ (Batch size) เท่ากับ 128 ซึ่งช่วยให้การประมวลผลข้อมูลมีความสมดุลระหว่างความเร็วและประสิทธิภาพ

งานวิจัยนี้กำหนดโมเดลการเรียนรู้เชิงลึกในการจำแนกรูปภาพสายพันธุ์กระบองเพชร ประกอบด้วยโมเดล ResNet34, VGG16 และ MobileNetV3 สำหรับเรียนรู้บนชุดข้อมูลฝึกฝน ประกอบด้วยสายพันธุ์กระบองเพชร 10 สายพันธุ์ โดยผ่านการคำนวณหลายชั้นของโครงข่ายประสาทเทียม ดังภาพที่ 2



ภาพที่ 2 การฝึกฝนโมเดล และการถ่ายทอดย้อนกลับของโมเดลการเรียนรู้เชิงลึก
ในการจำแนกรูปภาพสายพันธุ์กระบองเพชร

จากภาพที่ 2 การฝึกฝนโมเดลและการถ่ายทอดย้อนกลับของโมเดลการเรียนรู้เชิงลึกประกอบด้วยสองกระบวนการที่สำคัญคือการส่งข้อมูลไปข้างหน้า (Forwardpropagation) จะเห็นได้ว่ากระบวนการนี้เป็นการส่งข้อมูลไปในทิศทางจากซ้ายไปขวา และการถ่ายทอดย้อนกลับ (Backpropagation) โดยทั้งสองกระบวนการทำงานร่วมกันเพื่อให้โครงข่ายสามารถเรียนรู้และปรับค่าพารามิเตอร์ต่าง ๆ ให้เหมาะสมต่อการจำแนกผลลัพธ์มากที่สุด

กระบวนการเริ่มต้นด้วยการรับข้อมูลนำเข้าที่ป้อนเข้าสู่โหนดในชั้นแรก จากนั้นข้อมูลจะถูกประมวลผลผ่านโหนดต่าง ๆ ในแต่ละชั้น โดยน้ำหนักของการเชื่อมต่อระหว่างโหนดและฟังก์ชันกระตุ้นในแต่ละชั้นมีบทบาทสำคัญในการคำนวณผลลัพธ์ที่ส่งต่อไปยังชั้นถัดไป เมื่อข้อมูลถูกส่งต่อจนถึงชั้นสุดท้ายตัวแบบจะสร้างค่าผลลัพธ์ที่คาดการณ์ไว้ (Actual values) จากนั้นค่าที่ได้จะถูกเปรียบเทียบกับค่าความจริง (Target values) โดยใช้ฟังก์ชันความสูญเสีย (Loss function) เพื่อประเมินว่าตัวแบบมีความถูกต้องเพียงใด แล้วจะทำการส่งค่าความสูญเสียไปใช้ในกระบวนการปรับปรุงค่าพารามิเตอร์ เช่น ค่าน้ำหนัก (Weights) และค่าไบอัส (Biases) ซึ่งเป็นส่วนสำคัญในการการถ่ายทอดย้อนกลับ ดังนั้นการฝึกฝนโมเดลและการถ่ายทอดย้อนกลับจะวนซ้ำกันไปเรื่อย ๆ เพื่อปรับตัวแบบให้สามารถลดค่าความสูญเสียลงในแต่ละรอบจากการฝึกฝนจนได้ค่าที่เหมาะสมที่สุด (Rumelhart et al., 1986)

การค้นหาค่าอัตราการเรียนรู้ (Learning rate) (Kingma & Ba, 2014) ที่เหมาะสมที่สุดสำหรับการฝึกโมเดลเนื่องจากหากอัตราการเรียนรู้สูงเกินไป โมเดลอาจเรียนรู้ได้ไม่มั่นคงและไม่บรรลุค่าที่เหมาะสมที่สุด (Optimum) ขณะที่อัตราการเรียนรู้ที่ต่ำเกินไปจะทำให้กระบวนการเรียนรู้ช้า (Ruder, 2016) โดยทั่วไปจะเริ่มต้นด้วยค่าอัตราการ

เรียนรู้ต่ำแล้วค่อย ๆ เพิ่มขึ้น เพื่อดูว่าค่าความสูญเสียจะลดลงอย่างไร ซึ่งบ่งชี้ถึงอัตราการเรียนรู้ที่เหมาะสมและช่วยในการเลือกค่าที่ดีที่สุดสำหรับการฝึกฝนโมเดล งานวิจัยนี้ใช้เทคนิคการเรียนรู้แบบรอบเดียว (One cycle policy) (Smith, 2018) ที่เป็นแนวทางการตั้งค่าอัตราการเรียนรู้ เพื่อเพิ่มประสิทธิภาพในการฝึกฝนโครงข่ายประสาทเทียม โดยช่วงเริ่มต้นของการฝึกฝนอัตราการเรียนรู้จะถูกตั้งไว้ในระดับต่ำเพื่อให้โมเดลได้เรียนรู้และสร้างโครงสร้างเบื้องต้นของข้อมูล (Data structure) เพื่อลดความเสี่ยงที่จะทำให้โมเดลเรียนรู้เร็วเกินไปและเสี่ยงต่อการปรับค่าพารามิเตอร์ผิดพลาดในช่วงแรก ต่อมาช่วงกลางของการฝึกฝน อัตราการเรียนรู้จะเพิ่มขึ้นเรื่อย ๆ จนถึงระดับสูงสุด (Maximum learning rate) เพื่อช่วยให้โมเดลสามารถเรียนรู้ข้อมูลที่มีความซับซ้อนได้อย่างมีประสิทธิภาพ และลดปัญหาที่โมเดลจะติดอยู่ในค่าขั้นต่ำเฉพาะที่ (Local minima) ซึ่งเป็นปัญหาที่พบบ่อยในการฝึกฝนโมเดลที่มีความลึกสูง (Deep models) และช่วงท้ายของการฝึกฝนนั้นอัตราการเรียนรู้จะค่อย ๆ ลดลงสู่ระดับต่ำกว่าอัตราเริ่มต้นในช่วงแรก เพื่อให้โมเดลเกิดความเสถียรจากการหาค่าที่ดีที่สุดของฟังก์ชันความสูญเสีย (Loss function) และส่งผลให้โมเดลสามารถทำงานได้ดีขึ้นเมื่อนำไปทดสอบกับข้อมูลที่ไม่เคยเห็นมาก่อน (Generalization) เมื่อครบกระบวนการจะเรียกว่าหนึ่งรอบการฝึกฝน (One cycle) ส่งผลให้ใช้ระยะเวลาฝึกฝนที่น้อยลงและสามารถปรับปรุงประสิทธิภาพให้ดียิ่งขึ้นตามจำนวน 10 รอบการฝึกฝน ดังนั้นอัตราการเรียนรู้จะถูกปรับเปลี่ยนแบบไดนามิกให้เปลี่ยนแปลงในช่วงเวลาที่เหมาะสม โมเดลจะสามารถเรียนรู้ข้อมูลได้ดีในเวลาที่เหมาะสม อีกทั้งมีความเสถียร และลดโอกาสการเกิดปัญหาการจดจำลักษณะเฉพาะจนเกินไป (Overfitting) ได้อย่างมีประสิทธิภาพ

3) การประเมินประสิทธิภาพบนชุดข้อมูลตรวจสอบ (Performance evaluation on validation set) หลังจากโมเดลผ่านการฝึกฝนแล้วจะถูกทดสอบประสิทธิภาพบนชุดข้อมูลตรวจสอบ ด้วยตารางเมทริกซ์ความสับสน (Confusion matrix) คือเครื่องมือที่ใช้ในการประเมินประสิทธิภาพของโมเดลการเรียนรู้เชิงลึกสำหรับปัญหาการจำแนกประเภท (Classification) โดยแสดงข้อมูลในรูปแบบตารางที่เปรียบเทียบผลลัพธ์ที่โมเดลคาดการณ์กับค่าความจริงสำหรับการจำแนกประเภทแบบไบนารี (Binary classification) เมทริกซ์จะมีขนาด 2×2 ประกอบด้วย

TP (True Positive)	คือ ผลการจำแนกสายพันธุ์ “ถูกต้อง”	เทียบกับป้ายกำกับ “ถูกต้อง”
TN (True Negative)	คือ ผลการจำแนกสายพันธุ์ “ไม่ถูกต้อง”	เทียบกับป้ายกำกับ “ไม่ถูกต้อง”
FN (False Negative)	คือ ผลการจำแนกสายพันธุ์ “ไม่ถูกต้อง”	เทียบกับป้ายกำกับ “ถูกต้อง”
FP (False Positive)	คือ ผลการจำแนกสายพันธุ์ “ถูกต้อง”	เทียบกับป้ายกำกับ “ไม่ถูกต้อง”

งานวิจัยนี้วัดประสิทธิภาพของการจำแนกรูปภาพด้วยค่าความถูกต้อง (Accuracy) จากสัดส่วนของการทำนายที่ถูกต้องทั้งหมดต่อจำนวนการทดสอบทั้งหมด (Sokolova & Lapalme, 2009) ดังสมการที่ 1

$$\text{ค่าความถูกต้อง} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

อย่างไรก็ดี ข้อจำกัดจากการรวบรวมข้อมูลรูปภาพส่งผลกระทบต่อประสิทธิภาพของแต่ละสายพันธุ์มีจำนวนรูปภาพไม่เท่ากัน เรียกว่าความไม่สมดุลของคลาส (Class imbalance) ดังนั้นการใช้เพียงค่าความถูกต้องอาจส่งผลให้โมเดลมีประสิทธิภาพสูงในคลาสที่มีรูปภาพจำนวนมาก ตรงกันข้ามประสิทธิภาพจะลดลงเมื่อรูปภาพในคลาสมีจำนวนไม่มากนัก งานวิจัยนี้จึงนำค่าความแม่นยำ (Precision) ค่าการเรียกคืน (Recall) และค่าประสิทธิภาพโดยรวม (F-measure) เข้ามาร่วมพิจารณาประสิทธิภาพของโมเดล (Powers, 2011) ดังสมการที่ 2, 3 และ 4

$$\text{ค่าความแม่นยำ} = \frac{TP}{TP + FP} \quad (2)$$

$$\text{ค่าการเรียกคืน} = \frac{TP}{TP + FN} \quad (3)$$

$$\text{ค่าประสิทธิภาพโดยรวม} = 2 \times \frac{\text{ค่าความแม่นยำ} \times \text{ค่าการเรียกคืน}}{\text{ค่าความแม่นยำ} + \text{ค่าการเรียกคืน}} \quad (4)$$

ผลการวิจัย

งานวิจัยนี้มุ่งเน้นการพัฒนาโมเดลการจำแนกสายพันธุ์กระบองเพชรด้วยการเรียนรู้เชิงลึกที่สามารถเรียนรู้จากคุณลักษณะของรูปภาพที่หลากหลาย แบ่งผลการวิจัยออกเป็น 2 ส่วนตามวัตถุประสงค์ ประกอบด้วย

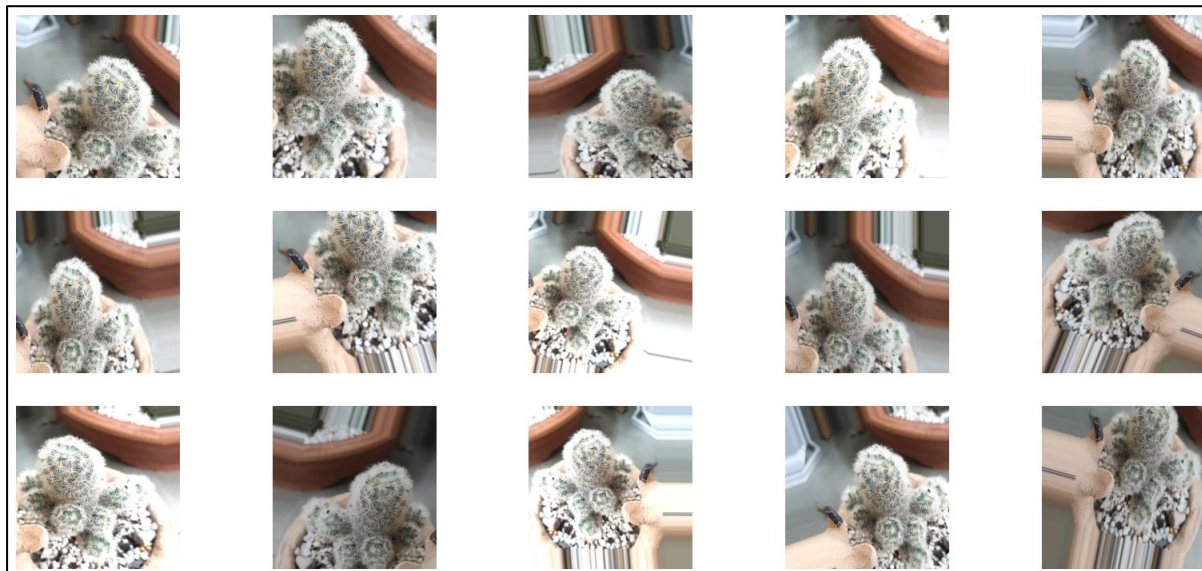
1) ผลการพัฒนาโมเดลการจำแนกสายพันธุ์กระบองเพชรด้วยการเรียนรู้เชิงลึก การดำเนินงานแบ่งออกเป็น 3 มอดูลหลัก ได้แก่ 1) การเตรียมรูปภาพก่อนการประมวลผล 2) การฝึกฝนโมเดลและการถ่ายทอดย้อนกลับ และ 3) การประเมินประสิทธิภาพบนชุดข้อมูลตรวจสอบ มีรายละเอียดดังนี้

1.1) มอดูลการเตรียมข้อมูลก่อนการประมวลผล เริ่มต้นด้วยการดัดแปลงรูปภาพจากต้นฉบับ เพื่อขยายความหลากหลายของชุดข้อมูลด้วยเทคนิคต่าง ๆ ดังภาพที่ 3 ประกอบด้วย การเลื่อนตำแหน่งภาพ (Image shift) ไปในแนวนอนหรือแนวตั้ง ซึ่งช่วยให้โมเดลสามารถเรียนรู้ตำแหน่งที่แตกต่างกันของวัตถุได้ เช่น เลื่อนไปทางซ้าย ขวาชั้นหรือลง การกลับด้านภาพ (Image flips) ในแนวนอนหรือแนวตั้ง เพื่อให้โมเดลรู้จักกับวัตถุในทิศทางที่หลากหลาย เช่น กลับด้านซ้าย-ขวาในกรณีภาพวัตถุที่สามารถอยู่ทั้งสองทิศทางได้ การหมุนภาพ (Image rotation) ในมุมต่าง ๆ เช่น 15 องศาหรือ 30 องศา เพื่อให้โมเดลสามารถแยกแยะภาพที่มีมุมมองแตกต่างกันได้ การซูมเข้า-ออกภาพ (Image zoom) เพื่อเปลี่ยนขนาดของวัตถุในภาพ ช่วยให้โมเดลเรียนรู้การรู้จำวัตถุในขนาดที่แตกต่างกัน การบิดเฉียงภาพ (Image shear) จะทำให้ภาพมีลักษณะเอียงไปในทิศทางหนึ่ง คล้ายกับการมองภาพจากมุมที่แตกต่างไป เพื่อช่วยให้โมเดลทนทานต่อการเอียงหรือการเบี่ยงของวัตถุ และการปรับความสว่างภาพ (Image brightness) เพื่อจำลองแสงที่แตกต่างกัน ช่วยให้แบบจำลองสามารถรู้จำวัตถุในสภาพแสงที่แตกต่างกันได้ เช่น ในที่ที่มีแสงสว่างมากหรือน้อยกว่าปกติ



ภาพที่ 3 ตัวอย่างผลลัพธ์จากการดัดแปลงรูปภาพจากต้นฉบับ

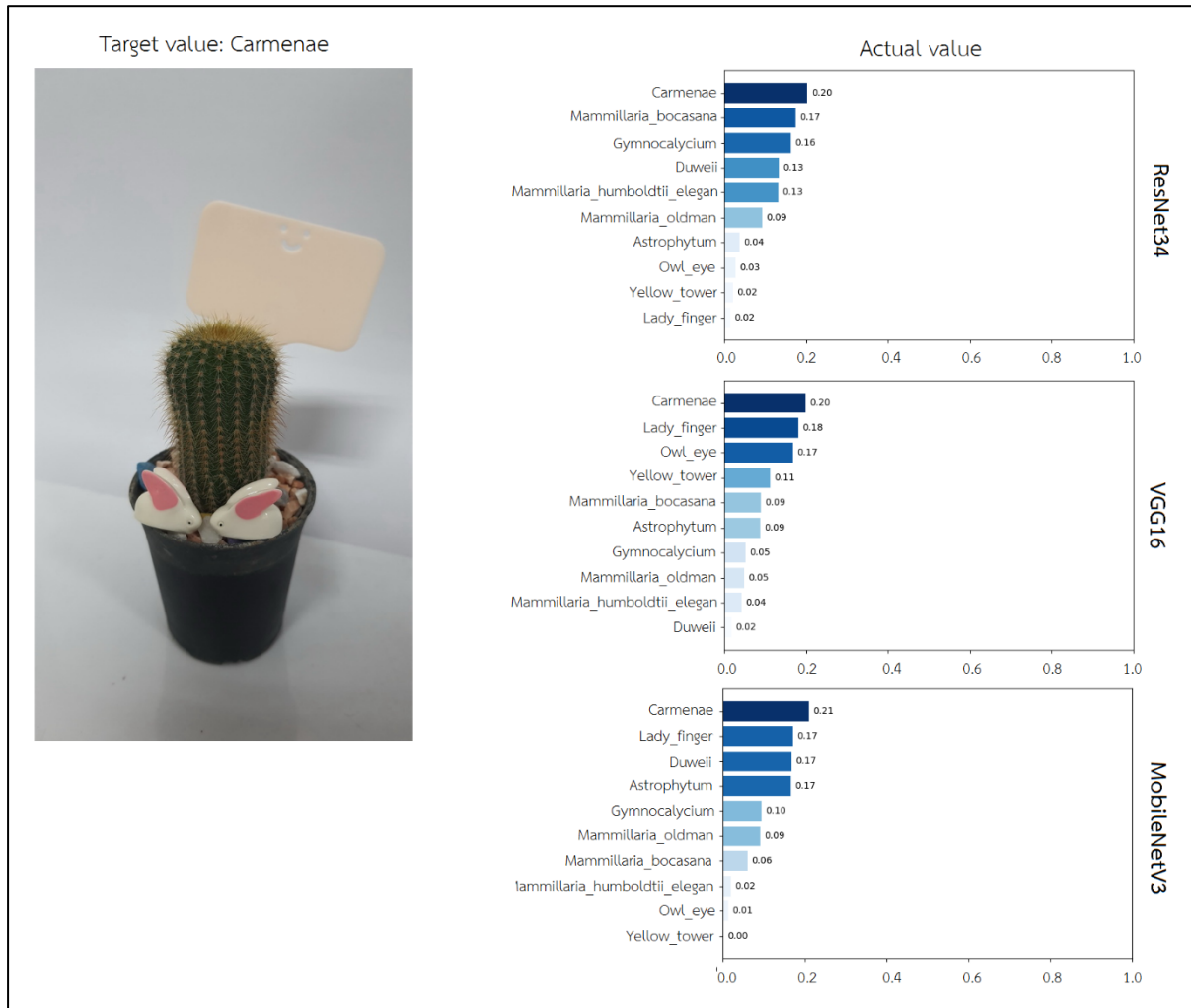
จากภาพที่ 3 เมื่อนำผลลัพธ์จากการดัดแปลงรูปภาพจากต้นฉบับมาผสมผสานเข้าด้วยกัน เกิดเป็นภาพใหม่ที่มีลักษณะแตกต่างจากต้นฉบับ ดังภาพที่ 4 เพื่อสร้างความหลากหลายของภาพที่มีรูปแบบแตกต่างกัน และช่วยให้โมเดลได้ฝึกฝนกับข้อมูลที่คล้ายกับสถานการณ์จริงในสภาพแวดล้อมที่หลากหลายขึ้น เช่น ภาพที่มาจากมุมมองหรือสภาพแสงที่ต่างกัน ซึ่งอาจไม่ได้มีในชุดข้อมูลดั้งเดิม ตลอดจนป้องกันการเกิด การจำลักษณะเฉพาะจนเกินไปที่โมเดลจดจำลักษณะเฉพาะของภาพที่ซ้ำกันในชุดข้อมูล (Shorten & Khoshgoftaar, 2019) ซึ่งจะลดการจดจำรูปแบบที่ไม่ทั่วไป และทำให้โมเดลเรียนรู้คุณสมบัติที่สำคัญจริง ๆ ของวัตถุได้ดีขึ้น ดังนั้นการผสมผสานการแปลงภาพหลายรูปแบบเป็นหนึ่งในวิธีการเพิ่มคุณภาพของข้อมูลภาพสำหรับการเรียนรู้เชิงลึก และลดปัญหาที่โมเดลมีประสิทธิภาพสูงเมื่อทำงานกับชุดข้อมูลฝึกฝน แต่ประสิทธิภาพจะลดลงเมื่อนำไปใช้งานกับข้อมูลจริง (Poor real-world performance)



ภาพที่ 4 การผสมผสานการแปลงภาพหลายรูปแบบเข้าด้วยกัน

1.2) โมดูลการฝึกฝนโมเดลและการถ่ายทอดย้อนกลับ โดยโมเดลการเรียนรู้เชิงลึกประกอบด้วย ResNet34, VGG16 และ MobileNetV3 จะถูกฝึกฝนบนชุดข้อมูลรูปภาพกระบองเพชร 10 สายพันธุ์ที่ผ่านกระบวนการดัดแปลงรูปภาพจากต้นฉบับ ซึ่งเป็นหนึ่งในการควบคุมสภาพแวดล้อมในการทดสอบโมเดล เพื่อความสอดคล้องและน่าเชื่อถือ ร่วมกับการตั้งค่าไฮเปอร์พารามิเตอร์ที่สำคัญ เช่น อัตราการเรียนรู้ ขนาดของแบตช์ จำนวนรอบการฝึกฝน และออปติไมเซอร์ เพื่อให้การเรียนรู้ของโมเดลเป็นไปในทิศทางเดียวกัน เช่นเดียวกับการควบคุมฮาร์ดแวร์และซอฟต์แวร์ที่มีผลต่อการคำนวณ งานวิจัยนี้จึงกำหนดการทดสอบบน Google colabatory แบบมีค่าใช้จ่าย โดยใช้ GPU เป็นหน่วยประมวลผล และใช้หน่วยความจำขนาด 25 กิกะไบต์ในโหมด High-RAM ตลอดจนเรียกใช้เวอร์ชันของเฟรมเวิร์กและไลบรารีให้เหมือนกัน

กระบวนการฝึกฝนโมเดลด้วยการเรียนรู้แบบรอบเดียวในการปรับอัตราการเรียนรู้และโมเมนตัม ซึ่งช่วยให้โมเดลเรียนรู้ได้อย่างมีประสิทธิภาพในระยะเวลา 10 รอบด้วยขั้นตอนการฝึกฝน การตรวจสอบความถูกต้อง และการปรับแต่งค่าพารามิเตอร์ (Parameter tuning) ในแต่ละรอบซ้ำไปเรื่อย ๆ โดยมีเป้าหมายเพื่อให้ค่าความถูกต้องสูงสุด และลดค่าสูญเสียจากการฝึกฝนให้น้อยที่สุดเท่าที่จะเป็นไปได้ อย่างไรก็ตาม ระดับที่เหมาะสมจะขึ้นอยู่กับปัจจัยหลายอย่าง เช่น ชุดข้อมูลที่ฝึกฝน โมเดลที่เลือก และการปรับพารามิเตอร์ หากโมเดลมีการเรียนรู้ที่ดี ค่าความถูกต้องจะมีเสถียรภาพและค่าสูญเสียจากการฝึกฝนจะลดลงจนถึงจุดหนึ่งที่โมเดลไม่สามารถปรับปรุงเพิ่มเติมได้ (Convergence) จากนั้นจะคำนวณความแตกต่างระหว่างค่าที่โมเดลพยากรณ์กับป้ายกำกับจริง และปรับน้ำหนักและไบแอสผ่านการถ่ายทอดย้อนกลับ เพื่อลดความผิดพลาดในแต่ละรอบของการฝึกฝน จนกระทั่งได้โมเดลที่มีประสิทธิภาพสูงสุดกับชุดข้อมูลตรวจสอบ ดังภาพที่ 5 ผลการจำแนกสายพันธุ์กระบองเพชรของแต่ละโมเดลที่สามารถจำแนกสายพันธุ์กระบองเพชรได้ถูกต้อง แต่มีค่าความเชื่อมั่น (Confidence) ที่แตกต่างกัน



ภาพที่ 5 ผลการจำแนกสายพันธุ์กระบองเพชรของแต่ละโมเดล

1.3) มอดูลการประเมินประสิทธิภาพบนชุดข้อมูลตรวจสอบ เพื่อประเมินประสิทธิภาพการจำแนกสายพันธุ์กระบองเพชรผ่านค่าความถูกต้อง งานวิจัยนี้ให้ความสำคัญค่าความคลาดเคลื่อน (error) คือ ผลต่างระหว่างค่าที่วัดได้กับค่าที่แท้จริง ถ้าค่าที่วัดได้ใกล้เคียงกับค่าจริงมากแสดงว่าการวัดนั้นมีความถูกต้องสูง (เรวัต แสงสุริยงค์, 2565) โดยการวัดทุกครั้งมักมีความคลาดเคลื่อนเกิดขึ้นเสมอ แบ่งออกเป็น 3 ชนิด ได้แก่ 1) ความคลาดเคลื่อนที่เกิดจากผู้วัด (Human error) อันเกิดจากความประมาท เลินเล่อในการวัด 2) ความคลาดเคลื่อนเชิงระบบ (Systematic error) มักเกิดจากเครื่องมือวัดไม่ได้ประสิทธิภาพ และ 3) ความคลาดเคลื่อนเชิงสถิติ (Statistical error) ซึ่งความคลาดเคลื่อนประเภทที่ 1 และ 2 สามารถจัดการได้ด้วย ความรอบคอบ ระวัง และ การพัฒนาเครื่องมือวัดให้มีประสิทธิภาพ แต่สำหรับความคลาดเคลื่อนประเภทที่ 3 นั้น ไม่มีทางกำจัดให้หมดไปได้ เนื่องจากอยู่นอกเหนือการควบคุม จึงมักใช้การวัดซ้ำหลาย ๆ ครั้งนั้นเมื่อครบรอบจะทำการประเมินอีกครั้งจนครบ 3 รอบ แบบผลลัพธ์เป็นอิสระต่อกัน แล้วนำผลการประเมินมาหาค่าเฉลี่ยเพื่อให้ความคลาดเคลื่อนลดน้อยลงจนถึงในระดับที่ยอมรับได้ ดังตารางที่ 3

ตารางที่ 3 ผลประเมินประสิทธิภาพการจำแนกบนชุดข้อมูลตรวจสอบด้วยค่าความถูกต้องของโมเดล ResNet34, VGG16 และ MobileNetV3

รอบ	ResNet34	VGG16	MobileNetV3
1	89.01%	90.01%	91.01%
2	90.03%	91.03%	91.36%
3	91.05%	92.05%	91.71%
ค่าเฉลี่ย	90.03%	91.03%	91.36%

จากตารางที่ 3 พบว่าโมเดลทั้ง 3 ตัว ประกอบด้วย ResNet34, VGG16 และ MobileNetV3 มีค่าความถูกต้องอยู่ในระดับสูงมาก เท่ากับ 90.03%, 91.03% และ 91.36% ตามลำดับ เมื่อพิจารณาในรายละเอียด พบว่าการประเมินทั้ง 3 รอบแต่ละโมเดลมีค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.83, 0.83 และ 0.29 ตามลำดับ

อย่างไรก็ดี ผลการทดสอบประสิทธิภาพบนชุดข้อมูลตรวจสอบของโมเดล ResNet34 พบว่าค่าความถูกต้องเฉลี่ยอยู่ที่ 90.03% ซึ่งต่ำที่สุดในกลุ่มโมเดลที่พิจารณา โดยความถูกต้องในแต่ละรอบเพิ่มขึ้นอย่างต่อเนื่องจาก 89.01% เป็น 91.05% โดยมีความแปรปรวนต่ำ แสดงถึงความเสถียรในการเรียนรู้ของโมเดล โครงสร้างบล็อกส่วนต่างของโมเดล ResNet34 ช่วยแก้ปัญหาการสูญหายของกราฟิก และทำให้การฝึกฝนโมเดลลึกมีประสิทธิภาพ อย่างไรก็ตาม โมเดลนี้อาจไม่เหมาะกับชุดข้อมูลที่ใช้ในการทดสอบ เนื่องจากไม่สามารถเพิ่มค่าความถูกต้องได้เทียบเท่ากับโมเดลอื่น ๆ สำหรับโมเดล VGG16 มีค่าความถูกต้องเฉลี่ยสูงขึ้นที่ 91.03% ซึ่งดีกว่าโมเดล ResNet34 ในทุก ๆ รอบ โดยความถูกต้องในแต่ละรอบเพิ่มขึ้นสม่ำเสมอ เท่ากับ 90.01% ถึง 92.05% โดยโครงสร้างที่ลึกของโมเดล VGG16 ซึ่งประกอบด้วยชั้นคอนโวลูชันแบบเรียงซ้อนกัน ช่วยให้โมเดลสามารถจับลักษณะเฉพาะของข้อมูลได้ดี ส่งผลให้ผลลัพธ์มีความถูกต้องสูง อย่างไรก็ตาม ข้อเสียสำคัญของโมเดล VGG16 คือจำนวนพารามิเตอร์ที่มาก ทำให้ต้องใช้ทรัพยากรในการประมวลผลสูง จึงอาจไม่เหมาะสมกับระบบที่มีข้อจำกัดด้านฮาร์ดแวร์ ในทางกลับกันโมเดล MobileNetV3 ซึ่งเป็นโมเดลที่แสดงประสิทธิภาพดีที่สุดในกลุ่ม มีค่าความถูกต้องเฉลี่ยสูงถึง 91.36% และค่าความถูกต้องในแต่ละรอบเพิ่มขึ้นอย่างต่อเนื่อง เท่ากับ 91.01% ถึง 91.71% โดยโครงสร้างของโมเดล MobileNetV3 ถูกออกแบบมาให้เหมาะกับการใช้งานที่ต้องการประสิทธิภาพสูงและการใช้ทรัพยากรต่ำโดยใช้คอนโวลูชันแยกส่วนตามความลึก เพื่อลดความซับซ้อนของการคำนวณ จึงเหมาะอย่างยิ่งสำหรับการใช้งานในระบบที่มีข้อจำกัดด้านทรัพยากร

เมื่อพิจารณาประสิทธิภาพของโมเดลโดยนำมาตรวัดอื่น ๆ เข้ามาร่วมพิจารณาประสิทธิภาพของโมเดล ประกอบด้วยค่าความแม่นยำ ค่าการเรียกคืน และค่าประสิทธิภาพโดยรวม ดังตารางที่ 4 จากการประเมินทั้ง 3 รอบ พบว่า โมเดล VGG16 เหมาะสมที่สุดในแง่ของความแม่นยำและความสมดุลในการประมวลผลข้อมูลใหม่ รองลงมาคือโมเดล MobileNetV3 ที่มีประสิทธิภาพใกล้เคียงกัน แต่เหมาะกับการใช้งานที่ต้องการทรัพยากรน้อย ในขณะที่โมเดล ResNet34 อาจต้องการการปรับปรุงเพิ่มเติมเพื่อให้สามารถแข่งขันกับโมเดลอื่นในบริบทการจำแนกสายพันธุ์

ตารางที่ 4 ผลประเมินประสิทธิภาพการจำแนกบนชุดข้อมูลตรวจสอบด้วยค่าความแม่นยำ ค่าการเรียกคืน และค่าประสิทธิภาพโดยรวมของโมเดล ResNet34, VGG16 และ MobileNetV3

โมเดล	ค่าความแม่นยำ	ค่าการเรียกคืน	ค่าประสิทธิภาพโดยรวม
ResNet34	91.12%	90.03%	89.94%
VGG16	91.83%	91.03%	91.10%
MobileNetV3	91.46%	91.36%	91.20%

จากตารางที่ 4 พบว่า โมเดล VGG16 มีค่าประสิทธิภาพโดยรวมสูงสุดในทุกเกณฑ์ เมื่อพิจารณาในรายละเอียด พบว่า ค่าความแม่นยำ ค่าการเรียกคืน และค่าประสิทธิภาพโดยรวม อยู่ที่ 91.83%, 91.03% และ 91.10% ตามลำดับ ซึ่งบ่งบอกถึงความสามารถของโมเดลในการจำแนกสายพันธุ์กระบองเพชรได้อย่างแม่นยำ และสมดุลระหว่างการลดอัตราการทำนายผิดพลาด และการจำแนกสายพันธุ์ทั้งหมดในชุดข้อมูลตรวจสอบได้อย่างครบถ้วน เช่นเดียวกับโมเดล MobileNetV3 แสดงค่าประสิทธิภาพรองลงมา โดยค่าความแม่นยำอยู่ที่ 91.46% และค่าการเรียกคืนอยู่ที่ 91.36% ทำให้ค่าประสิทธิภาพโดยรวมอยู่ที่ 91.20% ซึ่งใกล้เคียงกับโมเดล VGG16 อย่างมาก ซึ่งโมเดล MobileNetV3 มีข้อได้เปรียบในด้านการประมวลผลที่รวดเร็ว และมีประสิทธิภาพสูงในสภาพแวดล้อมที่มีข้อจำกัดด้านทรัพยากรแม้ว่าค่าประสิทธิภาพโดยรวมจะต่ำกว่าโมเดล VGG16 เล็กน้อย แต่โมเดลนี้ยังคงเหมาะสมสำหรับการใช้งานในสถานการณ์จริงที่ต้องการความเร็วและใช้ทรัพยากรอย่างคุ้มค่า ในส่วนของโมเดล ResNet34 มีค่าประสิทธิภาพต่ำที่สุดในกลุ่ม โดยค่าความแม่นยำ ค่าการเรียกคืน และค่าประสิทธิภาพโดยรวมอยู่ที่ 91.12%, 90.03% และ 89.94% ตามลำดับ แสดงถึงข้อจำกัดของโมเดลในการรักษาความสมดุลระหว่างการลดข้อผิดพลาด และการจำแนกข้อมูลที่หลากหลาย อย่างไรก็ตาม โมเดลนี้ยังคงแสดงความเสถียรในระดับที่ยอมรับได้ และต้องการการปรับแต่งเพิ่มเติมเพื่อเพิ่มประสิทธิภาพในอนาคต

ดังนั้นจากการเปรียบเทียบทั้ง 3 โมเดล พบว่า โมเดล MobileNetV3 แสดงประสิทธิภาพสูงสุดทั้งในด้านความถูกต้องและการประหยัดทรัพยากร รองลงมาคือโมเดล VGG16 ซึ่งให้ความถูกต้องใกล้เคียงกัน แต่มีข้อเสียในด้านการใช้ทรัพยากรสูง ขณะที่โมเดล ResNet34 มีความถูกต้องต่ำสุดในกลุ่ม แต่ยังคงมีความเสถียรและมีความสามารถในการเรียนรู้ที่เหมาะสมในภาพรวม ทั้งนี้การเลือกใช้โมเดลที่เหมาะสมควรพิจารณาเป้าหมายของงาน และข้อจำกัดด้านทรัพยากรที่มีอยู่ร่วมด้วย

2) ผลการประเมินความถูกต้องของโมเดลการจำแนกรูปภาพสายพันธุ์กระบองเพชร งานวิจัยนี้เลือกโมเดลที่มีค่าความถูกต้องสูงสุดคือโมเดล MobileNetV3 มาทดสอบกับชุดข้อมูลทดสอบ (Test set) ที่ประกอบด้วยรูปภาพกระบองเพชรแต่ละสายพันธุ์ที่โมเดลไม่เคยเห็นมาก่อน (Unseen data) เนื่องจากการที่โมเดลสามารถจำแนกสายพันธุ์กระบองเพชรจากข้อมูลที่ไม่เคยถูกใช้ในการฝึกฝน จะเป็นตัววัดความสามารถของโมเดลในการทำงานกับข้อมูลใหม่ที่มีความหลากหลาย และสะท้อนถึงประสิทธิภาพของโมเดลในสถานการณ์จริงที่มีผู้ใช้งานด้วยรูปภาพในบริบทที่แตกต่างกัน เช่น รูปภาพที่มีความแตกต่างด้านระยะถ่าย มุมมอง แสงสว่าง ตลอดจนพื้นหลัง จะช่วยทดสอบความยืดหยุ่นของโมเดลในการจัดการกับความแปรปรวนของข้อมูล เช่น ภาพถ่ายในระยะที่แตกต่างกัน อาจทำให้รายละเอียดของลำต้นหรือหนามเปลี่ยนไป หรือมุมมองจากด้านบนหรือด้านข้างอาจส่งผลต่อลักษณะเด่นของสายพันธุ์ นอกจากนี้สภาพแสงที่แตกต่างกัน เช่น แสงธรรมชาติ แสงจากหลอดไฟ หรือการมีเงาในภาพ ยังเป็นปัจจัยที่เพิ่มความซับซ้อนในการจำแนกสายพันธุ์

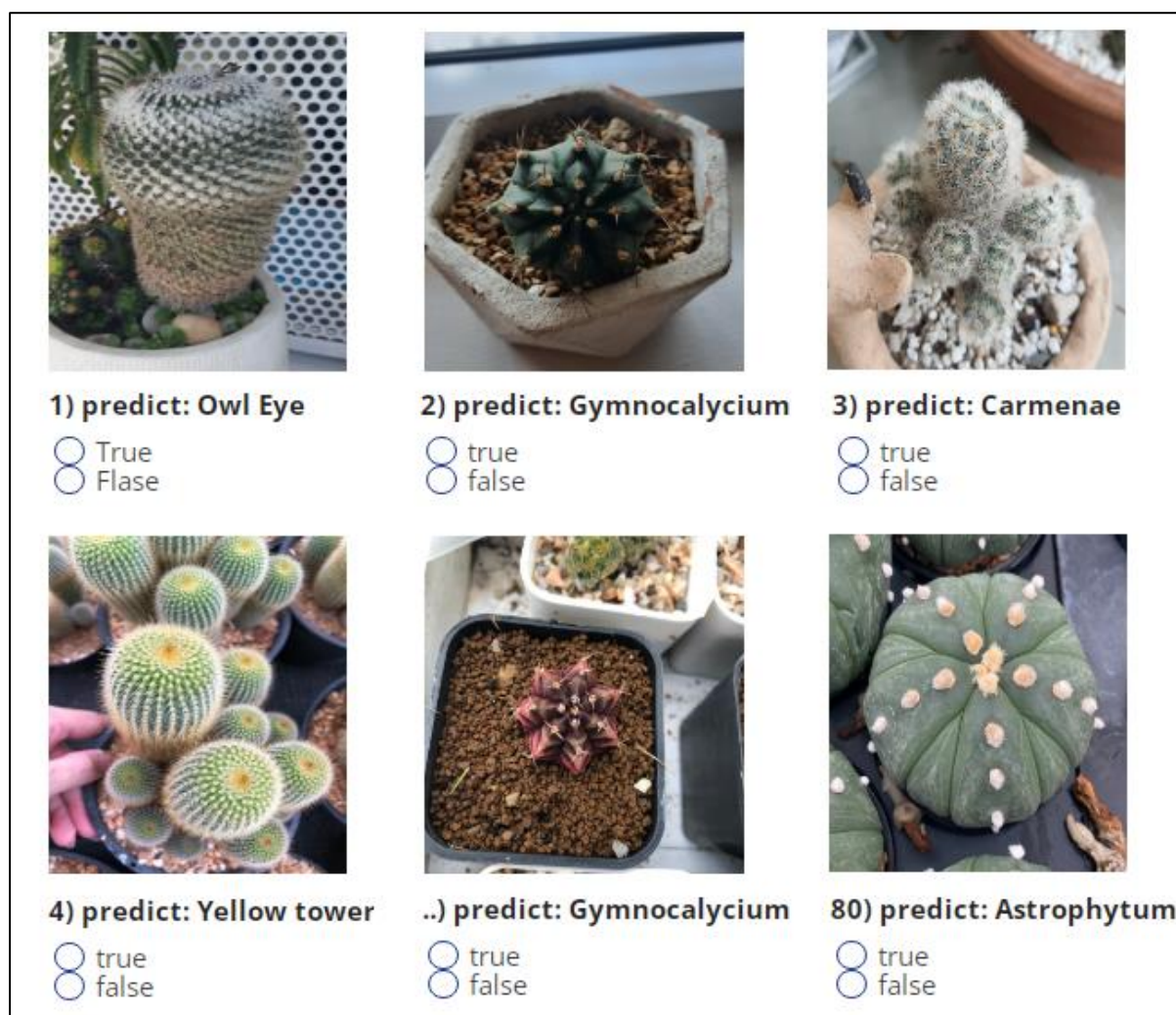
ดังนั้นเพื่อเป็นการสะท้อนถึงความพร้อมของโมเดลต่อการใช้งานในวงกว้าง โดยเฉพาะในบริบทที่ผู้ใช้งานแต่ละคนมีแหล่งข้อมูลและเงื่อนไขที่แตกต่างกัน เนื่องจากหากโมเดลสามารถรักษาความถูกต้องในระดับสูงได้กับข้อมูลที่ไม่เคยเห็นมาก่อน จะช่วยเพิ่มศักยภาพในการนำไปใช้งานจริง การทดสอบนี้จึงถือเป็นขั้นตอนสำคัญที่ช่วยให้มั่นใจได้ว่าโมเดลสามารถตอบโจทย์ในสถานการณ์จริงได้อย่างมีประสิทธิภาพและสอดคล้องกับความต้องการของผู้ใช้งาน จึงเป็นที่มาของการให้ผู้เชี่ยวชาญเข้ามามีบทบาทในการประเมินความถูกต้อง เพื่อเพิ่มความน่าเชื่อถือและลดข้อผิดพลาดที่อาจเกิดขึ้นจากกระบวนการอัตโนมัติหรือข้อจำกัดของแบบจำลองในการจำแนกข้อมูลเฉพาะทาง

การประเมินจากผู้เชี่ยวชาญหลายคนยังช่วยลดความคลาดเคลื่อนของการประเมินที่อาจเกิดขึ้นจากการพิจารณาโดยผู้เชี่ยวชาญเพียงท่านเดียว ทำให้ผลการประเมินมีความน่าเชื่อถือมากขึ้น งานวิจัยนี้กำหนดจำนวนผู้เชี่ยวชาญ 5 ท่านนำเข้ารูปภาพอย่างอิสระ ภายใต้ขอบเขตของกระเบื้องเพชร 10 สายพันธุ์ตามที่กำหนดไว้ ท่านละ 80 รูปภาพรวมทั้งสิ้น 400 ภาพ ก่อนจะบันทึกผลประเมินผ่านแบบประเมินผลออนไลน์ ดังภาพที่ 6 กำหนดให้ หากผลการจำแนกสายพันธุ์กระเบื้องเพชรถูกต้องเท่ากับ 1 คะแนน ในทางตรงกันข้าม หากผลการจำแนกสายพันธุ์กระเบื้องเพชรไม่ถูกต้องเท่ากับ 0 คะแนน เมื่อทำการประเมินแล้ว จะนำคะแนนทั้งหมดที่ได้จากผู้เชี่ยวชาญมาคำนวณเป็นค่าความถูกต้องเพื่อประเมินประสิทธิภาพของโมเดลจากสมการที่ 5

$$\text{ค่าความถูกต้อง} = \frac{\text{จำนวนรูปภาพที่จำแนกถูกต้อง}}{\text{จำนวนรูปภาพทั้งหมด}} \quad (5)$$

การแปลความหมายของระดับคะแนน (บุญชม ศรีสะอาด, 2556) จากการประเมินความถูกต้องจากผู้เชี่ยวชาญต่อผลลัพธ์การจำแนกสายพันธุ์กระเบื้องเพชรของโมเดลที่พัฒนาขึ้น กำหนดให้ 80%–100% อยู่ในระดับดีมาก 70%–79% อยู่ในระดับดี 60%–69% อยู่ในระดับพอใช้ 50%–59% อยู่ในระดับผ่าน และ 0%–49% อยู่ในระดับต้องปรับปรุง

เมื่อตรวจสอบพบว่า จำนวนรูปภาพที่จำแนกสายพันธุ์กระเบื้องเพชรถูกต้องเท่ากับ 326 รูปภาพจากจำนวนรูปภาพทั้งหมด 400 รูปภาพ คิดเป็นค่าความถูกต้องเท่ากับ 81.50% อยู่ในระดับดีมาก ถือว่าเป็นระดับที่น่าพอใจก่อนจะพัฒนาให้ใช้งานบนแอปพลิเคชันต่อไป



ภาพที่ 6 ตัวอย่างแบบประเมินผลออนไลน์การจำแนกรูปภาพสายพันธุ์กระบองเพชร

อภิปรายผล

การพัฒนาโมเดลการจำแนกรูปภาพสายพันธุ์กระบองเพชรด้วยการเรียนรู้เชิงลึกเพื่อจำแนกคุณลักษณะที่ซับซ้อนของกระบองเพชรสายพันธุ์ต่าง ๆ โดยงานวิจัยนี้ได้ใช้ 3 โมเดลหลัก ได้แก่ โมเดล ResNet34, VGG16 และ MobileNetV3 ฝึกฝนบนชุดข้อมูลรูปภาพของกระบองเพชร 10 สายพันธุ์ จำนวน 4,512 ภาพที่ผ่านการดัดแปลงรูปภาพจากต้นฉบับ เพื่อเพิ่มขนาดของชุดข้อมูลเดิมให้หลากหลายมากขึ้น กระบวนการฝึกฝนโมเดลโดยการตั้งค่าไฮเปอร์พารามิเตอร์ ร่วมกับการใช้เทคนิคการเรียนรู้แบบรอบเดียว ในการปรับอัตราการเรียนรู้ และโมเมนตัม ก่อนปรับค่าน้ำหนักของเครือข่ายประสาทเทียมด้วยกระบวนการถ่ายทอดย้อนกลับ ซึ่งช่วยให้โมเดลสามารถจำแนกสายพันธุ์กระบองเพชรได้อย่างมีประสิทธิภาพ

ผลการประเมินประสิทธิภาพของโมเดล ResNet34, VGG16 และ MobileNetV3 บนชุดข้อมูลตรวจสอบพบว่า โมเดล MobileNetV3 มีประสิทธิภาพดีที่สุด ด้วยค่าความถูกต้องเฉลี่ย 91.36% และความแม่นยำสูงโดยเทคนิคคอนโวลูชันแยกส่วนตามความลึกช่วยลดความซับซ้อนและเหมาะกับการใช้งานในสภาพแวดล้อมที่มีข้อจำกัดด้านทรัพยากร เช่นเดียวกับ โมเดล VGG16 มีค่าความถูกต้องเฉลี่ย 91.03% ใกล้เคียงกับโมเดล MobileNetV3 โดยใช้ชั้น

คอนโวลูชันแบบเรียงซ้อนช่วยเพิ่มประสิทธิภาพ แต่ต้องแลกมาด้วยการใช้ทรัพยากรที่สูงกว่า และโมเดล ResNet34 มีค่าความถูกต้องเฉลี่ยต่ำสุดที่ 90.03% แม้จะแสดงความเสถียรและมีประสิทธิภาพในการเรียนรู้จากโครงสร้างบล็อก ส่วนต่าง แต่ยังมีข้อจำกัดในการดึงคุณสมบัติของข้อมูลได้อย่างเต็มที่ งานวิจัยนี้เลือกโมเดลที่มีค่าความถูกต้องสูงสุด คือโมเดล MobileNetV3 มาทดสอบกับรูปภาพที่โมเดลไม่เคยเห็นมาก่อน จำนวน 400 ภาพ โดยผลการประเมิน ความถูกต้องจากผู้เชี่ยวชาญจำนวน พบว่า สามารถจำแนกรูปภาพได้ถูกต้องถึง 81.50% ของรูปภาพทั้งหมด ซึ่งถือว่า อยู่ในระดับดีมาก

อย่างไรก็ดี ความคล้ายคลึงของลักษณะทางสัณฐานวิทยาของกระบองเพชรนั้นส่งผลต่อความคลาดเคลื่อนในการจำแนกรูปภาพอย่างมีนัยสำคัญ โดยเฉพาะในกรณีสายพันธุ์ที่มีลักษณะคล้ายคลึงกัน เช่น ขนาดและรูปร่างของลำต้น ลักษณะหนาม หรือสีสันที่ใกล้เคียงกัน อาจส่งผลให้โมเดลไม่สามารถเรียนรู้ความแตกต่างระหว่างสายพันธุ์ได้อย่างมีประสิทธิภาพ ตลอดจนคุณภาพของภาพ เนื่องจากหากภาพมีความละเอียดต่ำหรือมีแสงสว่างที่ไม่เหมาะสม จะส่งผลให้ประสิทธิภาพของการจำแนกสายพันธุ์ลดลงอย่างมาก อีกทั้งหากในชุดข้อมูลมีภาพกระบองเพชรบางสายพันธุ์ มากเกินไปในขณะที่สายพันธุ์อื่นมีน้อย อาจทำให้โมเดลมีแนวโน้มที่จะจดจำคุณลักษณะของภาพในสายพันธุ์ที่ปรากฏ บ่อยกว่า ดังนั้นการแก้ปัญหาคความคลาดเคลื่อนนี้จำเป็นต้องปรับปรุงคุณภาพของชุดข้อมูล เช่น การเพิ่มจำนวนภาพ ของแต่ละสายพันธุ์ให้ใกล้เคียงกัน และการเพิ่มความหลากหลายของรูปภาพในชุดข้อมูล เป็นต้น ตลอดจนการปรับแต่ง โมเดลเพื่อให้สามารถเรียนรู้คุณลักษณะเชิงลึกได้ดีขึ้น เช่น การใช้ Attention mechanism เพื่อโฟกัสไปที่ส่วนสำคัญของภาพ หรือการเลือกใช้โมเดลที่มีความซับซ้อนมากขึ้นในการจำแนกสายพันธุ์ เช่น โมเดล EfficientNet หรือโมเดล ViT (Vision in Transformer) ดังนั้นการจัดการกับความคล้ายคลึงของลักษณะทางสัณฐานวิทยาอย่างเหมาะสมจะมีความสำคัญอย่างยิ่งในการเพิ่มความถูกต้องในการจำแนกสายพันธุ์กระบองเพชร

เอกสารอ้างอิง

- Anderson, E. F. (2001). *The cactus family*. Portland, OR: Timber Press.
- Barthlott, W., & Hunt, D. R. (1993). *Cacti: An illustrated guide to the plants and their cultivation*. Portland, OR: Timber Press.
- Bay, H., Tuytelaars, T., & Van Gool, L. (2006). SURF: Speeded up robust features. *The Proceedings of the European Conference on Computer Vision (ECCV)*, 7-13 May 2006, Graz, Austria, 404-417. https://doi.org/10.1007/11744023_32
- Boser, B. E., Guyon, I. M., & Vapnik, V. N. (1992). A training algorithm for optimal margin classifiers. *The Proceedings of the Fifth Annual Workshop on Computational Learning Theory*, 27-29 July 1992, Pittsburgh, PA, USA, 144-152. <https://doi.org/10.1145/130385.130401>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5-32. <https://doi.org/10.1023/A:1010933404324>
- Dalal, N., & Triggs, B. (2005). Histograms of oriented gradients for human detection. *The Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 20-25 June 2005, San Diego, CA, USA, 886-893. <https://doi.org/10.1109/CVPR.2005.177>

- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *The Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 27-30 June 2016, Las Vegas, NV, USA, 770-778. <https://doi.org/10.1109/CVPR.2016.90>
- Hongboonmee, N., & Thamditae, A. (2023). Preliminary classification of gemstones via mobile application using deep learning processes. *Sripatum Review of Science and Technology*, 15, 71-85. (in Thai)
- Howard, A., Sandler, M., Chu, G., Chen, L. C., Chen, B., & Tan, M. (2019). Searching for MobileNetV3. *The Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 27 Oct-2 Nov 2019, Seoul, South Korea, 1314-1324. <https://doi.org/10.1109/ICCV.2019.00139>
- Hunt, D., Taylor, N., & Charles, G. (2006). *The new cactus lexicon*. Milborne Port, UK: DH Books.
- Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture*, 147, 70-90. <https://doi.org/10.1016/j.compag.2018.02.016>
- Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. [Online]. Retrieved from: <https://arXiv preprint arXiv: 1412.6980>.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444. <https://doi.org/10.1038/nature14539>
- Lowe, D. G. (1999). Object recognition from local scale-invariant features. *The Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 20-27 September 1999, Corfu, Greece, 1150-1157. <https://doi.org/10.1109/ICCV.1999.790410>
- Perez, L. & Wang, J. (2017). The effectiveness of data augmentation in image classification using deep learning. [Online]. Retrieved from: <https://arXiv preprint arXiv:1712.04621>.
- Powers, D. M. (2011). Evaluation: From precision recall and F-measure to ROC informedness markedness and correlation. *Journal of Machine Learning Technologies*, 2(1), 37-63.
- Ruder, S. (2016). An overview of gradient descent optimization algorithms. [Online]. Retrieved from: <https://arXiv preprint arXiv: 1609.04747>.
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323(6088), 533-536. <https://doi.org/10.1038/323533a0>
- Sae-Lim, W., Wettayaprasit, W., & Aiyarak, P. (2019). Convolutional Neural Networks Using MobileNet for Skin Lesion Classification. *The Proceedings of 16th International Joint Conference on Computer Science and Software Engineering (JCSSE)*, 10-12 July 2019 at Chonburi, Thailand, 242-247.
- Sangsuriyong, R. (2022). The risks of measurement error in quantitative research in sociology. *Journal of Humanities and Social Sciences, Burapha University*, 30(1), 158-185. (in Thai)

- Sivic, J., & Zisserman, A. (2003). Video Google: A text retrieval approach to object matching in videos. *The Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 13-16 October 2003, Nice, France, 1470-1477. <https://doi.org/10.1109/ICCV.2003.1238663>
- Shmida, A. & Burgess, T. L. (1988). Plant growth form strategies and vegetation processes in the deserts of Israel and the Sinai. *Plant Species Biology*, 3(1), 59-69.
- Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data*, 6(1), 1-48. <https://doi.org/10.1186/s40537-019-0197-0>
- Simonyan, K. & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. [Online]. Retrieved from: <https://arXiv preprint arXiv:1409.1556>.
- Smith, L. N. (2018). A disciplined approach to neural network hyper-parameters: Part 1 - Learning rate, batch size, momentum, and weight decay. [Online]. Retrieved from: <https://arXiv preprint arXiv:1803.09820>.
- Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing & Management*, 45(4), 427-437.
- Srisaat, B. (2013). *Introduction to research*. 9th ed. Bangkok: Suweiryasarn.
- Yosinski, J., Kirsch, A., Ramsey, J., Schneider, J., & Soh, M. (2014). How transferable are features in deep neural networks? [Online]. Retrieved from: <https://arxiv.org/abs/1411.1792>
- Zhu, F., Sun, Y., Zhang, Y., Zhang, W., & Qi, J. (2023). An Improved MobileNetV3 Mushroom Quality Classification Model Using Images with Complex Backgrounds. *Agronomy*, 13(12), 2924. <https://doi.org/10.3390/agronomy13122924>.