

การจำแนกภาพเมล็ดข้าววัชพืชรบกวนจากภาพโทรศัพท์เคลื่อนที่ โดยใช้โครงข่ายประสาทเทียม ด้วยวิธีการถ่ายโอนความรู้

นฤมิตร กิติเวช^{1,*}, ธนาสัย สุคนธ์พันธุ์², อานนท์ ศักดิ์วรวิชญ์³

^{1,2,3}สาขาวิศวกรรมธุรกิจและวิทยาการข้อมูล คณะสถิติประยุกต์ สถาบันบัณฑิตพัฒนบริหารศาสตร์

Received: 22 May 2021

Revised: 20 July 2022

Accepted: 21 July 2022

บทคัดย่อ

งานวิจัยนี้ได้เสนอการใช้โครงข่ายประสาทเทียม (Neural network) ในการคัดแยกข้าววัชพืชรบกวนออกจากข้าวปลูกสำหรับรูปภาพที่ถ่ายด้วยกล้องโทรศัพท์เคลื่อนที่โดยใช้วิธีการถ่ายโอนการเรียนรู้จากตัวแบบที่เทรนด้วยภาพจากกล้องสะท้อนภาพเลนส์เดี่ยวด้วยระบบดิจิทัล (DSLR) ที่มีรายละเอียดของภาพสูงและถ่ายในสภาพแวดล้อมแบบปิดเพื่อนำภาพไปใช้เป็นตัวตั้งต้นสำหรับสร้างตัวแบบเพื่อทำนายภาพโทรศัพท์เคลื่อนที่ โดยงานวิจัยได้ศึกษาแนวทางด้านการสกัดคุณลักษณะ (Feature extraction) ซึ่งมีข้อจำกัดที่ต้องใช้งานในสภาพแวดล้อมแบบปิดจึงไม่เหมาะสมกับภาพโทรศัพท์เคลื่อนที่ที่มีสภาพแวดล้อมของภาพหลากหลาย ขั้นตอนการเตรียมข้อมูล ทางผู้วิจัยได้แยกภาพเมล็ดข้าวออกจากพื้นหลังและฝุ่นผงและทำการปรับสเกลของภาพเมล็ดข้าวที่ถ่ายด้วยกล้องโทรศัพท์เคลื่อนที่ (Scaled mobile) ให้ใกล้เคียงขนาดจริงโดยอ้างอิงขนาดของเหรียญบาท เมื่อทดสอบตัวแบบที่เทรนด้วยภาพ DSLR พบว่ามีค่าความแม่นยำในการทำนายด้วยภาพโทรศัพท์เคลื่อนที่ (Original mobile) เพิ่มขึ้นจากร้อยละ 49 เป็นร้อยละ 67 ผู้วิจัยเลือกใช้ตัวแบบ VGG16 และ Resnet50 ทดลองหาไฮเปอร์พารามิเตอร์ที่เหมาะสม และได้แบ่งวิธีการป้อนข้อมูลเป็น 3 วิธีดังนี้ (1) Mix method: สร้างตัวแบบจากข้อมูลภาพ DSLR และภาพ Scaled mobile (2) Retrain1 method: นำตัวแบบที่เทรนด้วยภาพ DSLR นำมาเทรนซ้ำด้วยภาพ Scaled mobile ให้ชุด validate set มีภาพจากทั้งสองชุด และ (3) Retrain2 method: ดัดแปลงมาจากวิธีที่ (2) โดยเพิ่มชุดภาพ DSLR อัตราสามในสิบส่วนจากตัวแบบต้นแบบ จากการทดลองพบว่าวิธีที่ให้ความแม่นยำที่สุดคือวิธีที่ (1) Mix method มีความแม่นยำร้อยละ 95 และร้อยละ 92 ในชุดทดสอบ DSLR และ Scaled mobile ตามลำดับ สำหรับวิธีที่ (3) Retrain2 method ให้ผลดีรองลงมา แต่มีประโยชน์ในด้านการประหยัดเวลาในการเทรนตัวแบบและใช้ขนาดข้อมูลที่น้อย

คำสำคัญ: การถ่ายโอนความรู้ การจำแนกพันธุ์เมล็ดข้าว การเรียนรู้เชิงลึก

* ผู้ประสานงานหลัก; อีเมล: narumitkitiweth@gmail.com

Classification of Weedy Rice-grain from Mobile Images Using Transfer learning Neural network

Narumit Kitiweth^{1,*}, Tanasai Sucontphunt², Arnond Sakworawich³

^{1,2,3}Business Analytics and Data Science Program, Graduate school of Applied Statistics,
National Institute of Development Administration

Received: 22 May 2021

Revised: 20 July 2022

Accepted: 21 July 2022

Abstract

This research presents a rice weed classification framework using neural network with transfer learning for mobile images. The transfer learning technique is employed so that high-res images by Digital single-lens reflex (DSLR) in the closed environment setting can be reused in the mobile environment setting. Other existing related works mainly rely on feature extraction in closed environment. In pre-processing steps, image segmentation technique is used to separate background and dust from a rice image and re-scale the rice grain image using known-size of a one-baht coin. This step improves accuracy from 49% to 67% using DSLR model. This work constructs VGG16 and Resnet50 then find-tunes with empirical hyperparameters. Also, an evaluation is conducted on three input and model combinations which are 1) DSLR and mobile images, 2) mobile images retrained with DSLR images, and 3) mobile images and 30% of DSLR images retrained with DSLR images. The first method gives highest accuracy of 95%. However, the third method also gives high accuracy with less training data.

Keywords: Transfer Learning, Rice Grain Classification, Deep Learning

* Corresponding Author; E-mail: narumitkitiweth@gmail.com

ความเป็นมาและความสำคัญของปัญหา

ข้าววัชพืช เป็นพันธุ์ข้าวป่าที่มีการเจริญเติบโตดีและสร้างความเสียหายกับนาข้าวอย่างมาก เนื่องจากเมื่อข้าววัชพืชโตแล้ว จะสูงบังและแย่งสารอาหารจากข้าวปลูกและยังกำจัดได้ยากเนื่องจากเจริญเติบโตได้ไวและเติบโตไม่พร้อมกัน ทำให้กำจัดได้ยากและฝังตัวอยู่ในดินได้นาน เมื่อชาวนานำข้าวที่ปลูกไปขายให้แก่โรงสีหากพบเจอข้าววัชพืชจะถูกกดราคาลงอย่างมาก สร้างความเสียหายและทำให้ชาวนามีต้นทุนในการกำจัดข้าววัชพืชสูงขึ้น ปัจจุบันข้าววัชพืชได้ระบาดไปทุกภาคส่วนและการตรวจสอบคุณภาพของข้าวเป็นไปได้อย่างยากลำบาก เนื่องจากบางสายพันธุ์มีลักษณะทางกายภาพใกล้เคียงระหว่างข้าววัชพืช เช่น ขน รูปทรง และ สีโดยเฉพาะสีจะมีความคล้ายคลึงกับข้าวปลูก ทางผู้วิจัยจึงสนใจที่จะพัฒนาตัวแบบเพื่อหวังว่าจะช่วยทุ่นแรงและลดระยะเวลาการตรวจสอบเมล็ดพันธุ์ข้าวได้

ในงานวิจัยก่อนหน้านี้ ส่วนมากจะเป็นการวิเคราะห์ภาพโดยตรงจากการสกัดคุณลักษณะ โดยมีงานวิจัยอยู่หลายงานที่เกี่ยวข้อง ได้แก่ งานวิจัยของ Lursthut and Pornpanomchai (2017) และ Silva et al. (2013) ได้ทำการสกัดข้อมูลจากภาพเมล็ดข้าวได้แก่ สี สัณฐานวิทยา (Morphological) และพื้นผิว (Texture) เพื่อนำไปสร้างตัวแปรทำนาย แต่วิธีนี้มีข้อจำกัดที่ต้องถ่ายในสภาพแวดล้อมแบบปิดเท่านั้น กล่าวคือจะต้องควบคุมปัจจัยต่างๆ เช่น แสงที่ตกกระทบเพื่อไม่ให้เกิดเงา และคงระยะห่างระหว่างเลนส์กับภาพ เพื่อไม่ให้มีผลต่อตัวแปรที่นำไปจำแนกเมล็ดพันธุ์ข้าว จากตัวอย่างข้าววัชพืชในภาพที่ 7 สีของข้าววัชพืชมีความใกล้เคียงกับข้าวปลูกมากทำให้วิธีสกัดข้อมูลอาจจะไม่เหมาะกับงานวิจัยนี้ สำหรับงานวิจัยที่ใกล้เคียงจะได้แก่ งานของ Aukkapinyo et al. (2020) ได้ใช้การเรียนรู้เชิงลึก (Deep learning) ในการแยกสายพันธุ์ข้าวออกจากกัน โดยลักษณะภายนอกของข้าวทั้ง 5 สายพันธุ์ ซึ่งมีความคล้ายคลึงกันทางด้านกายภาพ และสร้างความยากลำบากในการจำแนกของผู้เชี่ยวชาญข้าว ดังนั้นการคัดแยกพันธุ์ข้าวด้วยวิธีการสกัดฟีเจอร์ระดับสูงหรือการทำ Conventional classification techniques ทำให้วิธีดังกล่าวไม่เพียงพอต่อการแก้ปัญหานี้ได้ ซึ่งงานวิจัยดังกล่าวมีความแตกต่างจากงานของผู้วิจัยตรงที่งานวิจัยของผู้วิจัยจะมุ่งเน้นไปเพื่อสร้างตัวแบบให้สามารถคัดแยกข้าววัชพืชจากภาพโทรศัพท์เคลื่อนที่ได้ ซึ่งตัวภาพทั้งสองชุดมีความแตกต่างกันทั้งจากสภาพแวดล้อมและจากคุณภาพของกล้องเอง ทางผู้วิจัยจึงต้องการจะสร้างตัวแบบที่สามารถจำแนกภาพเมล็ดข้าววัชพืชออกจากข้าวปลูกเพื่อที่จะสามารถช่วยทุ่นแรงและลดระยะเวลาในการตรวจสอบคุณภาพเมล็ดข้าว และสามารถให้ชาวนาสามารถใช้ตรวจสอบได้ด้วยตนเองอีกด้วย

วัตถุประสงค์ของการวิจัย

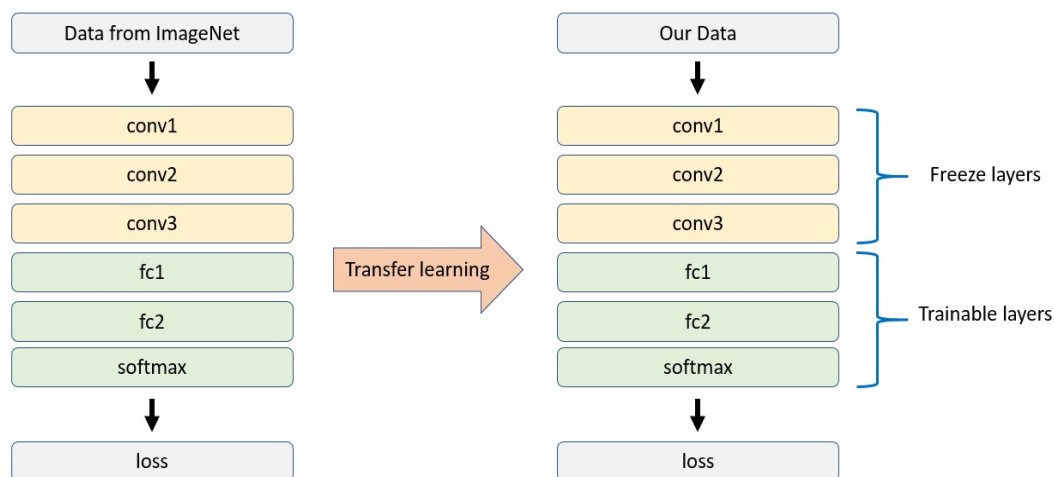
เพื่อพัฒนาตัวแบบโดยการทำการถ่ายโอนความรู้เพื่อให้ตัวแบบสามารถจำแนกเมล็ดข้าวปลูกและเมล็ดข้าววัชพืชจากรูปภาพโทรศัพท์เคลื่อนที่ได้

เอกสารและงานวิจัยที่เกี่ยวข้อง

จากงานวิจัยของ Aukkapinyo et al. (2020) ได้ใช้ Mask-RCNN ในการวิเคราะห์ภาพถ่ายโดยไม่ได้มุ่งเน้นไปที่การสกัดคุณลักษณะเช่นเดียวกับงานวิจัยอื่นที่เคยทำมาก่อนหน้า และได้ใช้ตัวแบบการถ่ายโอนความรู้จากชุดข้อมูล COCO-dataset เพื่อจำแนกภาพเมล็ดข้าว 5 สายพันธุ์ประกอบไปด้วยข้าวปลูกและข้าวเหนียว ในขั้นตอนการเก็บรวบรวมรูปภาพจะใช้กล้องดิจิทัลโดยมีการควบคุมปัจจัย แสง ระยะเลนส์ และ สมดุลแสงสีขาว โดยถ่ายบน

กระดาษสีดำ จากนั้นนำมาตัดและหมุนภาพให้ได้เฉพาะเมล็ดข้าวก่อนจะนำเข้าตัวแบบ ในการเตรียมชุดข้อมูลภาพ จะต้องระบุพื้นที่ของเมล็ดข้าวที่เราสนใจให้ตัวแบบ Mask-RCNN โดยตัวแบบจะใช้วิธี masker-based watershed algorithm ซึ่งเป็นอัลกอริทึมที่ทำการแยกภาพพื้นหลังออกจากเมล็ดข้าวจากการแยกเซตสีที่ต่างกันภายในภาพ จากนั้นจะส่งลักษณะเด่น (Feature) เข้าไปในชั้น convolutional เพื่อทำนายว่าจะป็นรูปของข้าวสายพันธุ์ใด โดยผลลัพธ์ที่ได้จากตัวแบบนี้มีความแม่นยำอยู่ที่ร้อยละ 81 เมื่อเทียบกับผู้เชี่ยวชาญข้าวมีความแม่นยำแค่เพียง ร้อยละ 76 สำหรับงานวิจัยอื่นที่เกี่ยวข้อง ทางด้าน Silva et al. (2013) ใช้วิเคราะห์ภาพโดยตรงโดยใช้วิธีการสกัด คุณลักษณะ โดยใช้กล่องดิจิตอลโดยติดตั้งขาตั้งกล้องเก็บรวบรวมรูปภาพและมีการใช้แสงไฟไปที่ภาพถ่าย เพื่อทำให้ไม่มีเงาเกิดขึ้นที่เมล็ดข้าว ข้าวที่ทำการทดลองเป็นสายพันธุ์ข้าวศรีลังกา 9 สายพันธุ์ ทำการแยกภาพพื้นหลังออกจากเมล็ดข้าว และนำฟิเจอร์ เช่น สี, สัณฐานวิทยา และ พื้นผิว ส่งเข้าไปใน MLP ที่มีการปรับแต่งเลเยอร์ต่างกัน ทำการเทรนตัวแบบและเปรียบเทียบความแม่นยำ ผลลัพธ์ที่ได้มีความแม่นยำอยู่ที่ร้อยละ 91

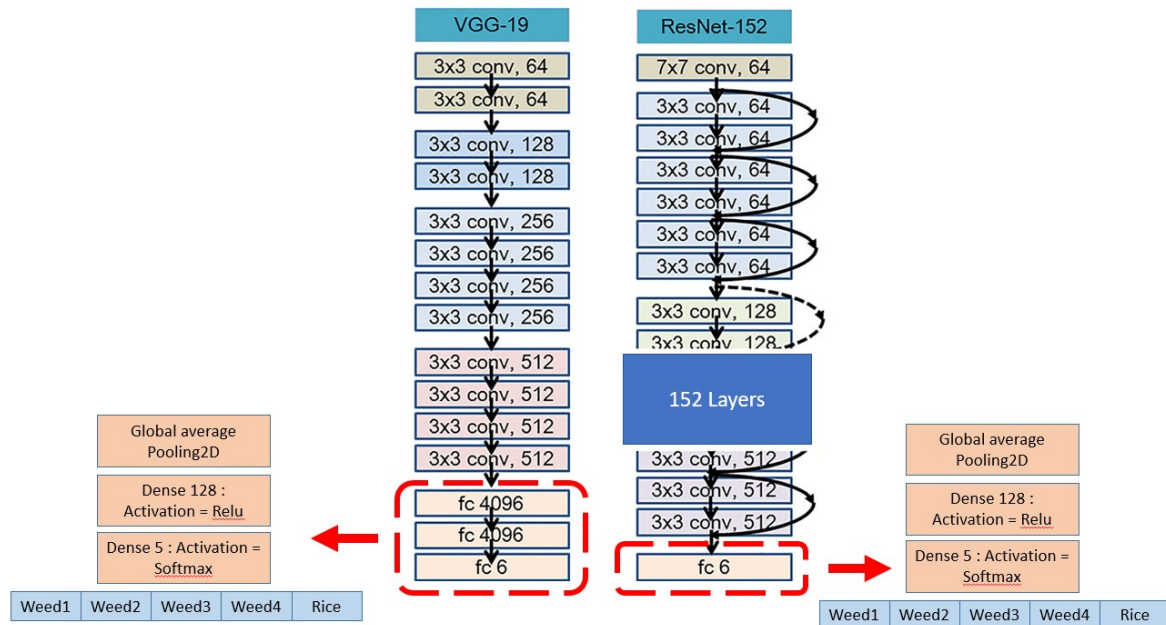
ในงานวิจัยของ Patel and Vaibhav (2014) การถ่ายโอนความรู้จะเป็นการนำค่าถ่วงน้ำหนักที่ถูกเทรนจาก ตัวแบบ ที่นักพัฒนาท่านอื่นได้เทรนตัวแบบมาแล้วและนำค่าถ่วงน้ำหนักนั้นมาปรับใช้กับชุดข้อมูลของเรา ผลที่ได้จะ ทำให้ประหยัดเวลาในการเทรนข้อมูลและสามารถนำมาประยุกต์ใช้ได้ทันที ค่าถ่วงน้ำหนักที่นักวิจัยท่านนี้ได้นำมาจาก ตัวแบบที่ถูกเทรนด้วยชุดข้อมูลที่ชื่อว่า image-net dataset ด้วยซึ่งเทรนด้วยรูปภาพจำนวน 1,000 คลาส รวมกว่า 1.2 ล้านภาพ ในการทำ Transfer Learning ใน CNN (Convolutional neural network) เลเยอร์ชุดหน้าของ CNN จะทำการจดจำและเรียนรู้จากคุณลักษณะเด่น เช่น เส้นหรือรูปทรงในแต่ละภาพ ในการถ่ายโอนความรู้ค่าถ่วงน้ำหนัก จากเลเยอร์ข้างต้นนี้ที่ถูกเทรนมาก่อนแล้วจะถูกนำมาเป็นค่าเริ่มต้นของโครงข่ายประสาทเทียมของ ดังตัวอย่างภาพที่ 1



ภาพที่ 1 ภาพตัวอย่างแสดงการทำการถ่ายโอนความรู้ด้วยตัวแบบ CNN

การทบทวนวรรณกรรมด้านตัวแบบ VGG และ Resnet ทางผู้วิจัยต้องการนำตัวแบบทั้งสองมาศึกษาทำการทดลองโดยการถ่ายโอนความรู้และใช้ค่าถ่วงน้ำหนักเริ่มต้นจาก Image-net ซึ่งทั้งสองตัวแบบเป็นตัวแบบที่ประกอบไปด้วยชั้น CNN จำนวนมากเช่นเดียวกัน โดยที่ VGG16 เป็นตัวแบบที่ชนะเลิศรางวัล ImageNet ILSVRC ใน ปี2014 ส่วน Resnet ชนะรางวัล ILSVRC ใน ปี2015 (Shallu and Mehra, 2018) ความแตกต่างสำหรับทั้งสองตัวแบบนี้ ในส่วนของตัวแบบ VGG ผู้พัฒนาตัวแบบ ต้องการจะสร้างตัวแบบที่มีขนาดเล็กแต่เปี่ยมไปด้วยประสิทธิภาพ

ภายในตัวแบบจะประกอบไปด้วยชั้น CNN ขนาดเล็กหลายชั้นประกอบเข้าด้วยกัน สำหรับตัวแบบ Resnet ทางผู้พัฒนาเชื่อว่าตัวแบบที่ประกอบไปด้วยเลเยอร์หลายชั้นและยิ่งลึกก็จะมีประสิทธิภาพ แต่มีข้อจำกัดที่ว่าตัวแบบที่มีเลเยอร์มากจะเริ่มมีปัญหาที่เลเยอร์ด้านท้าย ค่าที่ส่งผ่านมาตั้งแต่ต้นจะเกิดการเลือนหาย (Vanishing of the gradient) ทำให้ต้องมี Skip layers ซึ่งคอยเรียนรู้ว่าเลเยอร์ใดควรจะข้ามไป ตัวแบบ Resnet จึงมีความยาวกว่า VGG ถึง 8 เท่า แต่ตัวแบบยังคงมีความซับซ้อนต่ำกว่า (Lower complexity) และยังเร็วกว่าอีกด้วย (Simonyan et al., 2014)



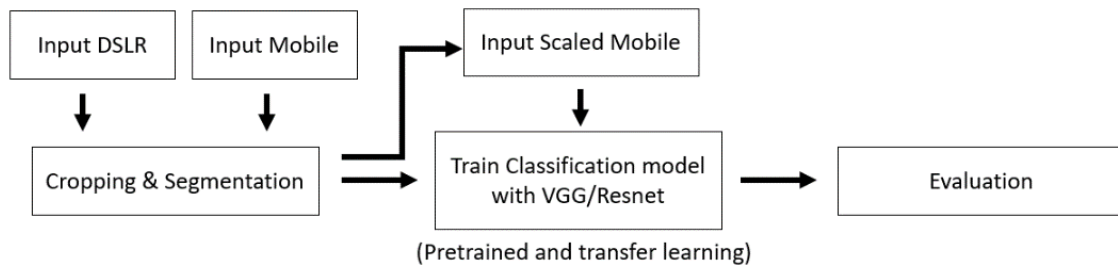
ภาพที่ 2 ภาพแสดงเลเยอร์ที่ถูกตัดเพื่อทำการถ่ายโอนความรู้ของตัวแบบ VGG19 และ Resnet152 (Han, 2018)

ผู้วิจัยนำตัวแบบทั้งสองมาทำการถ่ายโอนความรู้ โดยเลือกที่จะตัดเลเยอร์ fully connect หลังชั้น convolutional ออก แทนที่ด้วยเลเยอร์ดังต่อไปนี้ Global average pooling2D, Dense128 และ Dense5 ซึ่งเลเยอร์ชุดนี้จะทำการแปลง (Transform) ข้อมูลก่อนที่จะนำไปทำนาย (Classification) และในชั้นเลเยอร์สุดท้ายทางผู้วิจัยได้แทนที่ชั้น Fully Connected Layer ที่แต่เดิมมีจำนวนมิติของข้อมูลเท่ากับ 1000 คลาส จากชุดข้อมูล image-net นำมาแทนที่ด้วย Dense เท่ากับ 5 ซึ่งหมายความว่ามิติของข้อมูลจะลดลงเท่ากับจำนวนคลาสที่เราต้องการทำนาย (5คลาส) จากนั้นนำตัวเลขเหล่านั้นไปทำนายด้วยชั้น Softmax เพื่อทำนายว่ารูปภาพข่าวที่เราป้อนเข้าไปนั้นเป็นคลาสใด

วิธีดำเนินการวิจัย

กระบวนการทำวิจัย

การทดลองนี้แบ่งกระบวนการทำวิจัยเป็นดังนี้ (1) การเก็บรวบรวมชุดข้อมูลภาพ (2) การแยกภาพเมล็ดข้าวออกจากภาพพื้นหลังและเตรียมภาพเพื่อเทรนตัวแบบ (3) การเทรนตัวแบบและปรับไฮเปอร์พารามิเตอร์ และ (4) ทดลองแบ่งชุดข้อมูลและทดสอบตัวแบบ



ภาพที่ 3 แผนภาพวิธีวิจัยนี้แสดงถึงกระบวนการทำวิจัย

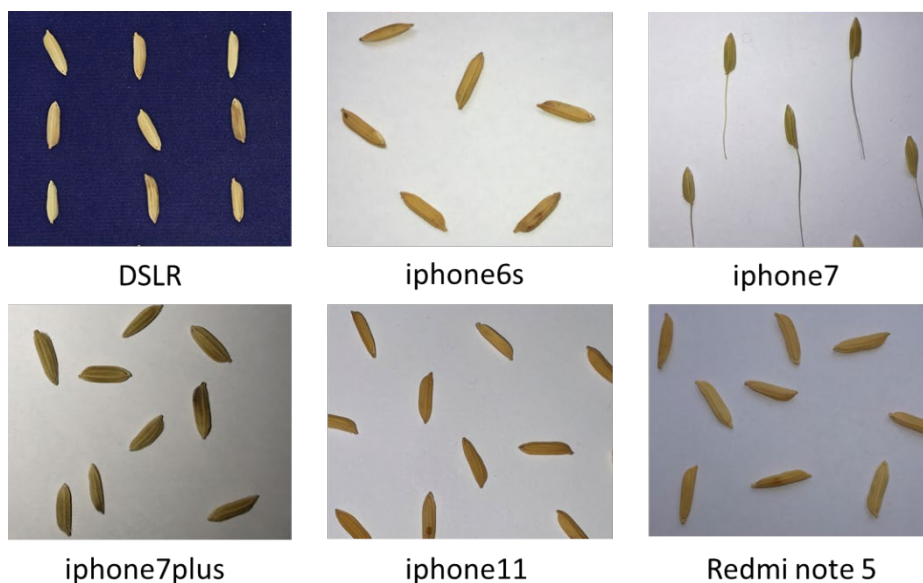
จากภาพที่ 3 ในการพัฒนาตัวแบบสำหรับจำแนกข้าววัชพืชออกจากข้าวปลูก จะนำภาพที่รวบรวมมาจากทั้ง DSLR และภาพจากกล้องโทรศัพท์เคลื่อนที่ นำมาแยกพื้นหลังและดึงเฉพาะภาพเมล็ดข้าวออกมา จากนั้นทำการสเกลภาพเมล็ดข้าวที่ถ่ายจากกล้องโทรศัพท์เคลื่อนที่ก่อนที่จะนำไปเทรนตัวแบบด้วยภาพจาก DSLR และภาพจากกล้องโทรศัพท์เคลื่อนที่ที่ผ่านการสเกลภาพแล้ว จากนั้นทำการวัดผลและสรุปผล

ประชากรและตัวอย่าง

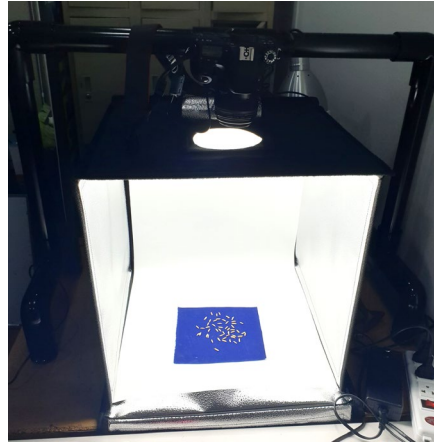
ประชากรที่ศึกษาคือ เมล็ดพันธุ์ข้าวที่นำมาจากศูนย์วิจัยข้าวปทุมธานี เป็นข้าวปลูก 6 สายพันธุ์ และข้าววัชพืช 4 สายพันธุ์ โดยแต่ละสายพันธุ์ไม่ปะปนกับสายพันธุ์อื่นๆ

เครื่องมือที่ใช้ในการวิจัย

การเก็บรวบรวมรูปภาพทางผู้วิจัยได้ใช้กล้อง Cannon 77D เลนส์ fix 35 mm. ถ่ายในระบบปิดควบคุมสภาพแวดล้อมทั้งแสงและมุมในการถ่าย ถ่ายบนพื้นหลังเป็นผ้าสีกหาลาดสีน้ำเงินโดยใช้ขาตั้งกล้องและกล้องสำหรับถ่ายภาพตลอดการเก็บรวบรวมข้อมูล และใช้กล้องโทรศัพท์เคลื่อนที่รุ่น iPhone 6s, iPhone 7, iPhone 7 Plus, iPhone 11 และ Redmi Note 5 ถ่ายบนพื้นหลังกระดาษสีขาวและถ่ายแบบระบบเปิดโดยเปิดแฟลชโทรศัพท์เคลื่อนที่ตลอดการเก็บรวบรวมข้อมูล



ภาพที่ 4 ตัวอย่างภาพถ่ายเมล็ดข้าวด้วยอุปกรณ์กล้องแต่ละแบบ ได้แก่ DSLR, iPhone 6s, iPhone 7, iPhone 7 Plus, iPhone 11 และ Redmi Note 5



ภาพที่ 5 ภาพตัวอย่างการถ่ายภาพเมล็ดข้าวด้วยกล้อง DSLR โดยมีการยิงแสงไฟจากด้านล่างไปที่เมล็ดข้าวเพื่อไม่ให้เกิดเงา และมีการควบคุมระยะเลนส์โดยใช้ขาตั้งกล้องตลอดการถ่ายภาพ

การแยกภาพเมล็ดข้าวออกจากพื้นหลัง (Image Segmentation) สำหรับภาพ DSLR ทางผู้วิจัยได้สร้างโปรแกรมในการแยกภาพเมล็ดข้าวออกจากพื้นหลัง โดยเทียบความต่างของแสงและสีโดยซ้อนทับด้วยเลเยอร์ mask หลายชั้น จากนั้นจึงเฉพาะเมล็ดข้าวออกมา สำหรับภาพจากกล้องโทรศัพท์เคลื่อนที่ ทางผู้วิจัยได้ทำการ mask พื้นที่ที่เป็นเมล็ดข้าวโดยใช้เครื่องมือ Magic tools ในโปรแกรม Photoshop จากนั้นทำการดึงเฉพาะภาพเมล็ดข้าวออกมาและปรับเป็นขนาดภาพเป็น 128*128 pixel ก่อนจะนำไปเข้าโปรแกรม

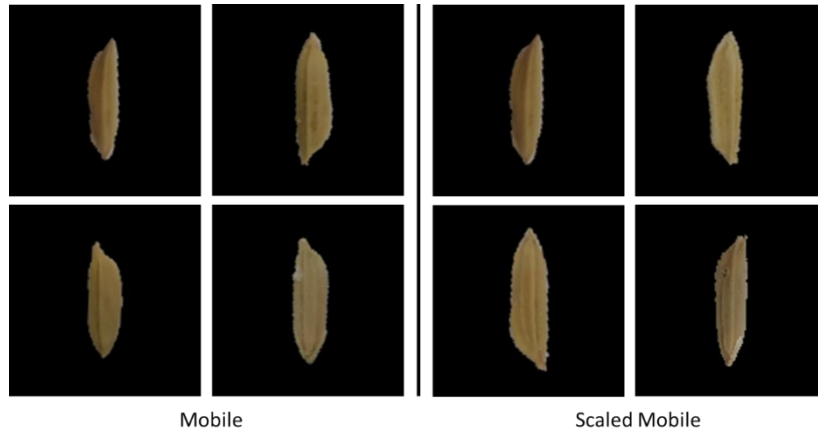


ภาพที่ 6 ภาพตัวอย่างข้าวปลูก 6 สายพันธุ์ที่ถ่ายจากกล้อง DSLR หลังจากทำ Segmentation



ภาพที่ 7 ภาพตัวอย่างข้าววัชพืช 4 สายพันธุ์ที่ถ่ายจากกล้อง DSLR หลังจากทำ Segmentation

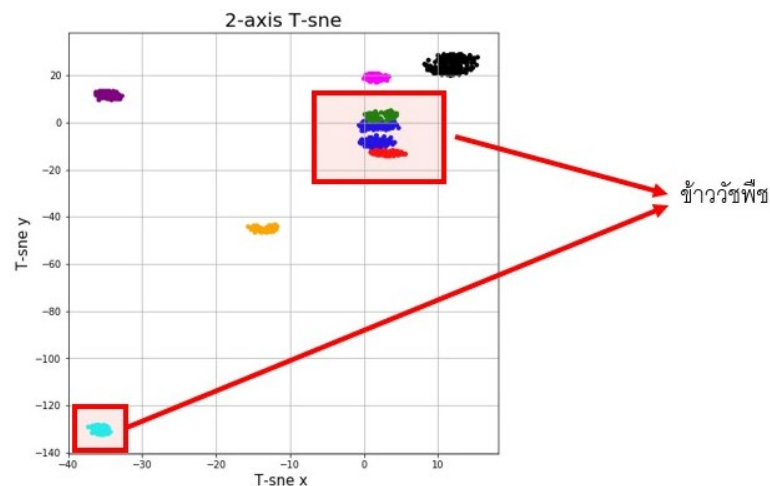
ทางผู้วิจัยทำการสเกลภาพโทรศัพท์เคลื่อนที่เนื่องจากขนาดของภาพเมล็ดข้าวที่ถ่ายระหว่างโทรศัพท์เคลื่อนที่แต่ละรุ่นและกล้อง DSLR มีความต่างกัน โดยทางผู้วิจัยได้อ้างอิงขนาดจากขนาดของเหรียญบาท สาเหตุที่ผู้วิจัยเลือกใช้เหรียญบาทเป็นตัวอ้างอิงเนื่องจากเหรียญบาทสามารถแยกออกจากเมล็ดข้าวและพื้นหลังได้โดยง่าย นำมาประยุกต์ใช้ได้ง่ายและมีขนาดที่ค่อนข้างคงที่



ภาพที่ 8 ซ้าย:ตัวอย่างภาพจากกล้องโทรศัพท์เคลื่อนที่ ขวา:ตัวอย่างภาพที่นำไปปรับสเกลโดยใช้เหรียญบาทเป็นตัวอ้างอิง



ภาพที่ 9 ตัวอย่างภาพที่ถ่ายด้วยกล้องโทรศัพท์เคลื่อนที่และใช้เหรียญบาทในการอ้างอิงสเกล



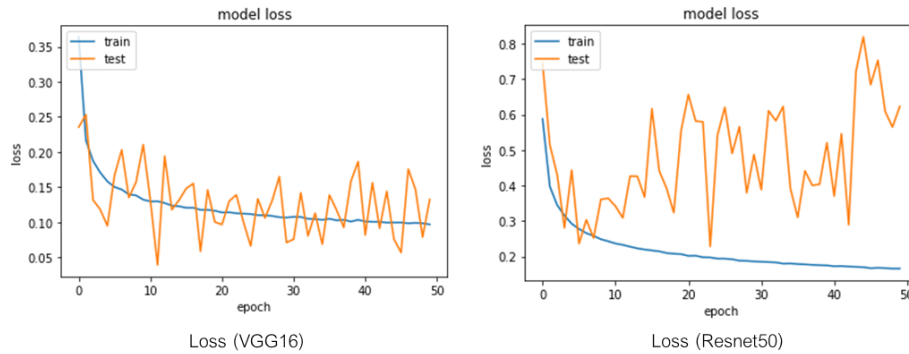
ภาพที่ 10 ภาพการจับกลุ่มของข้าวเมื่อนำฟีเจอร์ทได้จาก VGG16 เลย์เยอร์ก่อนสุดท้ายมาทำ Clustering ด้วย T-sne พบว่าแยกข้าววิเศษออกมาได้ค่อนข้างชัด

การนำตัวแบบที่ถูกเทรนแล้วและการถ่ายโอนความรู้มาใช้ (Pretrained model and Transfer learning) เริ่มต้นผู้วิจัยใช้ค่าถ่วงน้ำหนักจากชุดข้อมูล image-net นำมาเทรนเฉพาะชุดเลเยอร์ส่วนท้าย (สำหรับVGG16 จะเป็น 3 layers สุดท้าย) เพื่อให้ตัวแบบเกิดการเรียนรู้และปรับค่าถ่วงน้ำหนักในแต่ละโหนดภายในตัวแบบกับชุดข้อมูลภาพเมล็ดข้าวที่เตรียมไว้ โดยจะใช้ภาพจาก DSLR และภาพจากโทรศัพท์เคลื่อนที่ ซึ่งในไลบรารีของ keras มีตัวแบบที่ถูกเทรนไว้แล้วให้ใช้ได้แก่ Xception, VGG, ResNet, Inception, MobileNet, DenseNet, NASNet และ EfficientNet ทางผู้วิจัยเลือกสร้างตัวแบบด้วย VGG และ Resnet เนื่องจากเป็นตัวแบบที่ถูกเทรนด้วยรูปภาพจาก Image-net ที่มีรูปถึง 1.2 ล้านภาพและชนะรางวัลด้านการทำตัวแบบถ่ายโอนความรู้ทั้งคู่ ซึ่งคาดว่าจะได้ Perceptual features แบบเดียวที่มนุษย์มองหรือสังเกตมา ประกอบกับมีจำนวนพารามิเตอร์จำนวนมากทำให้มีความ Generalize ได้ดี ทางผู้วิจัยได้นำคุณลักษณะที่ตัวแบบใช้ในการทำนาย มาแสดงผลการจัดกลุ่มอย่างคร่าว ด้วย T-sne (Rezende et al., 2018) ดังภาพที่ 10 จะเห็นได้ว่าภาพของข้าววัชพืชนั้น มีคุณลักษณะคล้ายคลึงกันระดับหนึ่งที่จะคัดแยกออกจากข้าวปลูกได้ ในการเลือกวิธีการป้อนข้อมูลเพื่อเทรนตัวแบบ จากการทดลองเบื้องต้นพบว่าเมื่อนำตัวแบบมาทำการถ่ายโอนความรู้ ตัวแบบจะปรับตัวไปกับชุดข้อมูลล่าสุดและลืมชุดข้อมูลก่อนหน้า ทางผู้วิจัยจึงทดลองการนำข้อมูลไปเทรนตัวแบบในรูปแบบที่ต่างกัน เพื่อหารูปแบบที่เหมาะสมกับการเทรนตัวแบบด้วยการถ่ายโอนความรู้ต่อไป

การสร้างตัวแบบทางผู้วิจัยได้ทำการทดลองตัวแบบ VGG16 และ Resnet50 มาทำการถ่ายโอนความรู้เพื่อทำนายข้อมูล 2 คลาส (คลาส Rice และ Weed) แบ่งขนาดชุดข้อมูลเป็นอัตราส่วน 70:15:15 โดยขนาดข้อมูลที่ใช้เทรน (Train set) มีขนาด 70,000 รูป (ร้อยละ 70) ขนาดข้อมูลที่ใช้ตรวจสอบและปรับค่าถ่วงน้ำหนัก (Validate set) มีขนาด 15,000 รูป (ร้อยละ 15) และขนาดข้อมูลที่ใช้ในการทดสอบสุดท้าย (Test set) 15,000 รูป (ร้อยละ 15) พบว่า VGG16 มีความแม่นยำโดยรวมดีกว่าที่ร้อยละ 94.60 ในขณะที่ตัวแบบ Resnet50 ค่าความสามารถในการจดจำ (Recall) คลาส Rice อยู่ในระดับที่ต่ำ ทางผู้วิจัยจึงตัดสินใจเลือกใช้ตัวแบบ VGG16 ในการทดลองในขั้นต่อไป

ตารางที่ 1 แสดง Confusion matrix ในการจำแนกข้าววัชพืช โดยใช้ชุดทดสอบภาพ DSLR

Class	VGG16		Resnet50	
	Rice	Weed	Rice	Weed
Precision	89%	96%	89%	79%
Recall	85%	97%	2%	100%
F1-score	87%	93%	5%	94%
Accuracy	94.60%		79.3%	



ภาพที่ 11 แสดงถึงกราฟ Loss ในขณะเทรนตัวแบบ VGG16 และ Resnet50 พบว่า Resnet50 เรียนรู้ข้อมูลได้ดี แต่ไม่สามารถจดจำข้อมูลใน test set ได้ หมายความว่าตัวแบบจะมีความ Overfitting

จากนั้นทางผู้วิจัยได้ทำการ Fine-tuning model เพื่อหาไฮเปอร์พารามิเตอร์ที่เหมาะสม โดยแบ่งชุดข้อมูลออกเป็นอัตราส่วนตามข้างต้น กล่าวคือแบ่ง Train: Validated: Test เท่ากับ 70:15:15 โดยการสุ่มเลือกชุดข้อมูลแบบ Uniform ในแต่ละคลาส และกำหนดคลาสของข้าวเป็น 5 คลาส ได้แก่ ข้าววัชพืช 4 คลาส (ได้แก่ ดิดตันสูง ดิดเตี้ย ข้าวลาย และ ข้าวหาง) และข้าวปลูก 6 สายพันธุ์ ผสมรวมเป็น 1 คลาส (ได้แก่ ชัยนาท2 ดอกมะลิ105 เหลืองปะทิว123 ปทุมธานี1 พิษณุโลก2 และ กข.43) เนื่องจากในแต่ละคลาสมีจำนวนตัวอย่างภาพไม่เท่ากันจึงใช้การคำนวณค่าถ่วงน้ำหนักในแต่ละคลาส ทำการทดลองเลือก Optimizer โดยเปรียบเทียบทั้ง 3 รูปแบบ ได้แก่ SGD, adam และ RMSprop นำมาเทรนและทดสอบด้วยรูปภาพ DSLR พบว่าตัวแบบที่ใช้ SGD เป็น Optimizer มีค่าเฉลี่ย F1-score และ Accuracy ดีที่สุดที่ร้อยละ 89 และร้อยละ 94 ตามลำดับ

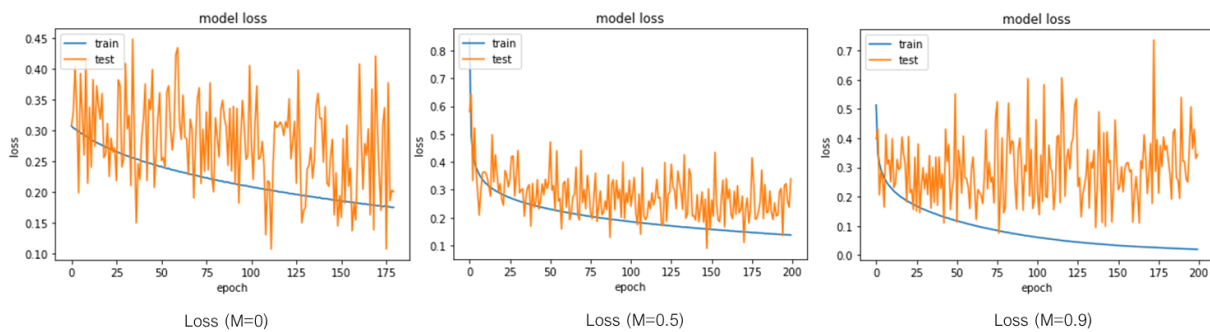
ตารางที่ 2 เปรียบเทียบค่าเฉลี่ย F1-score และ Accuracy ของตัวแบบโดยใช้ Optimizer ด้วย SGD, adam และ RMSprop

Optimizer Selection	TEST	
	F1-score	Accuracy
TRAIN		
DSL (SGD)	89%	94%
DSL (adam)	76%	89%
DSL (RMSprop)	76%	89%

ทำการเปรียบเทียบค่า Momentum ที่ 0, 0.5 และ 0.9 ตามลำดับ นำมาเทรนตัวแบบโดยแบ่งชุดข้อมูลออกเป็นอัตราส่วนตามข้างต้น กล่าวคือแบ่ง Train: Validated: Test เท่ากับ 70:15:15 จากชุดข้อมูลทั้งหมด 10,000 ภาพ พบว่า Momentum ที่ค่า 0.9 มีค่า F1-score และ Accuracy ดีที่สุด ที่ร้อยละ 99.90 และร้อยละ 99.90 ตามลำดับ

ตารางที่ 3 เปรียบเทียบค่าเฉลี่ย F1-score และ Accuracy ของตัวแบบโดยใช้ Momentum ที่ 0, 0.5 และ 0.9 ตามลำดับ

Momentum	TEST	
TRAIN	F1-score	Accuracy
DSLR (m=0)	89%	94%
DSLR (m=0.5)	90%	95%
DSLR (m=0.9)	99.90%	99.90%



ภาพที่ 12 แสดงถึงกราฟของ Loss เมื่อเทรนด้วย SGD เป็น Optimizer และใช้ Momentum ที่ 0, 0.5 และ 0.9 ตามลำดับ ในส่วนของตัวแบบที่ใช้ Momentum เท่ากับ 0.9 ตัวแบบมีความแม่นยำสูง แต่ระหว่าง train set และ test set มีความแตกต่างกันมาก ซึ่งหมายความว่าตัวแบบมีความ low bias และ high variance ส่งผลให้เกิด Overfitting ได้

นำตัวแบบที่ใช้ Backbone เป็น VGG16 ใช้ SGD เป็น Optimizer ของตัวแบบ ทำการสร้างตัวแบบด้วยการถ่ายโอนความรู้จากชุดข้อมูล แบ่งเป็น 6 แบบ ดังนี้

1. เทรนตัวแบบด้วยภาพจากกล้อง DSLR เท่านั้น (DSLR)
2. เทรนตัวแบบด้วยภาพกล้องโทรศัพท์เคลื่อนที่ต้นฉบับ (Original Mobile)
3. เทรนตัวแบบด้วยภาพกล้องโทรศัพท์เคลื่อนที่ที่ปรับขนาดด้วยตัวอ้างอิงสเกล (Scaled mobile)
4. เทรนตัวแบบด้วยการนำภาพจาก DSLR และ Scaled mobile ในครั้งเดียว (MIX)
5. นำตัวแบบที่เทรนจากภาพถ่าย DSLR มาเทรนซ้ำด้วยภาพ Scaled mobile และให้ Validate set มีรูปภาพจากทั้งสองแบบ (Retrain1)

6. นำตัวแบบที่เทรนจากภาพถ่าย DSLR มาเทรนซ้ำด้วยภาพ Scaled mobile และภาพจาก DSLR โดยภาพนำมาเพียงสามในสิบส่วนของตัวแบบที่ใช้เทรนในข้อที่ 1 (Retrain2)

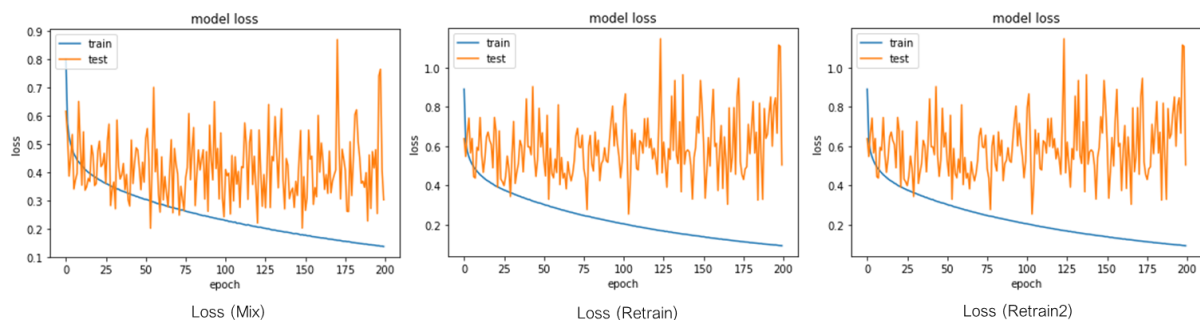
จากนั้นทำการวัดผลโดยนำตัวแบบมาทดสอบด้วยชุดข้อมูลทั้ง 3 ชุด ได้แก่ ชุดภาพทดสอบ DSLR 1,500 รูป ชุดภาพทดสอบ Mobile 1,500 รูป และ ชุดภาพทดสอบ Scaled mobile 1,500 รูป ตามที่แบ่งโดยข้างต้น

ผลการวิจัย

ทำการทดลองด้วยการนำตัวแบบ VGG16 ที่ถูกเทรนทั้ง 6 แบบในข้างต้น มาทดสอบด้วยชุดข้อมูลภาพ DSLR, Original Mobile และ Scaled mobile โดยการแบ่งภาพเมล็ดข้าวเป็น 5 คลาส (ข้าววัชพืช 4 คลาส และ ข้าวปลูก 1 คลาส) พบว่าผลการทดสอบตัวแบบที่ถูกเทรนด้วยภาพ DSLR เมื่อนำมาทำการทดสอบด้วยภาพ Scaled mobile พบว่ามีความแม่นยำมากกว่าตัวแบบที่เทรนด้วยชุดข้อมูลภาพโทรศัพท์เคลื่อนที่ปกติที่ร้อยละ 66 และ ร้อยละ 71 ตามลำดับ จึงคาดว่าในการทำ Scaled mobile เพื่อให้ขนาดของรูปใกล้เคียงกันจะทำให้ได้ผลลัพธ์ที่ดีขึ้น จากนั้นทางผู้วิจัยได้สร้างตัวแบบจากชุดข้อมูลภาพ DSLR และ Scaled mobile แบ่งวิธีป้อนข้อมูลเป็นดังนี้ (1) Mix method โดยนำภาพจาก DSLR และ Scaled mobile นำไปสร้างตัวแบบ (2) Retrain1 method นำตัวแบบที่ถูกเทรนเสร็จด้วยภาพจาก DSLR นำไปเทรนซ้ำอีกครั้งโดยใช้ภาพจาก Scaled mobile แต่ปรับแต่ง validate set ให้มีทั้งภาพจาก DSLR และ Scaled mobile (3) Retrain2 method เช่นเดียวกับวิธีก่อนหน้านี้ แต่ในการเทรนครั้งที่สอง จะป้อนภาพ DSLR ในอัตราส่วนสามในสิบส่วนในการและคง validate set เช่นเดียวกับวิธีที่ (2) ผลลัพธ์ที่ได้พบว่า ตัวแบบจะปรับค่าถ่วงน้ำหนักให้ใกล้เคียงกับชุดข้อมูลล่าสุด โดยค่าความแม่นยำในการทดสอบด้วย Scaled mobile ถึงแม้ตัวแบบจะมีความแม่นยำมากถึง 100% แต่ตัวแบบจะลืมชุดข้อมูลที่เทรนด้วยชุดภาพ DSLR ก่อนหน้าไป ดังนั้น วิธีที่ให้ผลลัพธ์ที่ดีจึงเป็นวิธี Mix และ Retrain2 ที่มีข้อมูลของทั้งสองรูปแบบอยู่ในการเทรนตัวแบบ

ตารางที่ 4 แสดงค่าเฉลี่ยของผลลัพธ์ Transfer learning ในการวัดผลด้วยภาพจาก DSLR, Mobile (ต้นฉบับ) และ Scaled mobile

Experimental	TEST SET					
	DSLR		Original Mobile		Scaled Mobile	
TRAIN SET	F1-score	Accuracy	F1-score	Accuracy	F1-score	Accuracy
DSLR	89%	94%	15%	49%	38%	67%
Original Mobile	48%	66%	76%	87%	64%	78%
Scaled Mobile	33%	71%	30%	62%	74%	87%
MIX	90%	95%	37%	65%	85%	92%
Retrain1	43%	71%	29%	60%	99%	100%
Retrain2	77%	89%	38%	63%	84%	92%



ภาพที่ 13 แสดงถึงกราฟของ Accuracy และ Loss เมื่อเทรนด้วยข้อมูลแบบ Mix, Retrain และ Retrain2 จะเห็นได้ว่าค่าเฉลี่ย loss ของวิธี mix มีค่าน้อยที่สุดในการเทรนตัวแบบ

จากข้างต้นจึงสรุปได้ว่า ในการนำตัวแบบไปใช้ทำการถ่ายโอนความรู้หรือนำตัวแบบมาเทรนซ้ำด้วยข้อมูลชุดใหม่ ควรนำข้อมูลที่ถูกลเทรนไปก่อนหน้าบางส่วนนำกลับมาเทรนพร้อมกัน เช่น ภาพจากกล้องโทรศัพท์รุ่นใหม่ หรือภาพจากอุปกรณ์อื่น เพื่อให้ตัวแบบยังคงจดจำและเรียนรู้จากข้อมูลทุกแบบได้

อภิปรายผล

งานวิจัยนี้ได้มุ่งเน้นไปที่การปรับไฮเปอร์พารามิเตอร์และวิธีการแบ่งชุดข้อมูลในการเทรนตัวแบบ เพื่อให้ตัวแบบสามารถจำแนกภาพเมล็ดข้าววัชพืชและข้าวปลูกได้อย่างแม่นยำ จากการทดลองพบว่าไฮเปอร์พารามิเตอร์ที่ทางผู้วิจัยเลือกใช้ได้แก่ตัวแบบ VGG16 และเลือกใช้ Optimizer เป็น SGD ปรับพารามิเตอร์ Momentum เท่ากับ 0.9 และ learning rate เท่ากับ 0.0001 ทั้งนี้ การสร้างตัวแบบโดยวิธีถ่ายโอนความรู้ในแต่ละครั้ง ตัวแบบจะยังคงทำนายแม่นยำกับชุดข้อมูลที่ถูกลเทรนเข้าไปในครั้งล่าสุดเสมอ ทางผู้วิจัยจึงเสนอว่าควรจะมีตัวแบบบางส่วนจากการเทรนตัวแบบครั้งก่อนหน้าไปด้วยเสมอเพื่อให้ตัวแบบจะยังทำนายและสามารถจดจำชุดข้อมูลที่ถูกลเทรนไปแล้วในรอบก่อน แต่ชุดข้อมูลจะมีขนาดเล็กลง จากตารางที่ 4 พบว่าหากเราผสมชุดข้อมูลที่มีความหลากหลายมากและใกล้เคียงกับการนำไปใช้งานเช่น วิธีที่ (1) Mix method ที่รวมภาพ DSLR และ Scaled mobile มีผลการทดสอบให้ความแม่นยำดีกว่าตัวแบบที่ถูกลเทรนด้วยข้อมูลภาพชุดเดียว

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลวิจัยไปใช้จริง หากชุดข้อมูลภาพมีความแตกต่างจากชุดข้อมูลที่ถูกลเทรนให้ตัวแบบ ยกตัวอย่างเช่น เป็นภาพโทรศัพท์เคลื่อนที่รุ่นใหม่ อาจส่งผลให้ตัวแบบมีความแม่นยำลดลงได้ ในงานวิจัยในอนาคตอาจเพิ่มชุดข้อมูลหรือทำ Data augmentation เพื่อเพิ่มความหลากหลายของชุดข้อมูล หรือใช้โครงข่ายประสาทเทียม เช่น Mask-RCNN เพื่อแยกภาพเมล็ดข้าวจากพื้นหลังได้หลากหลายมากขึ้น
2. ข้อเสนอแนะในการวิจัยครั้งต่อไป งานวิจัยนี้ได้มุ่งเน้นไปที่การปรับไฮเปอร์พารามิเตอร์และการแบ่งชุดข้อมูลเพื่อนำไปเทรนตัวแบบ สำหรับงานวิจัยในอนาคตอาจทดลองเจาะลึกไปที่การถ่ายโอนความรู้ตัวแบบ และเปรียบเทียบตัวแบบโครงข่ายประสาทเทียม เช่น VGG, ResNet, MobileNet, DenseNet, EfficientNet และ Yolo โดยปรับแต่งเลเยอร์ส่วนท้ายของโครงข่ายประสาทเทียมเพื่อหาว่าชั้นของโครงข่ายประสาทเทียมหรือตัวแบบแบบใดเหมาะสมกับรูปภาพหรืองานในลักษณะนี้

กิตติกรรมประกาศ

ผลงานนี้ได้รับการอุดหนุนการวิจัยจาก โครงการสร้าง Mobile Application สำหรับชาวนาเพื่อการตรวจสอบการจำแนกเมล็ดพันธุ์ข้าวออกจากข้าววัชพืชด้วยการประมวลผลภาพถ่ายโทรศัพท์เคลื่อนที่ สำหรับชาวนา โดยสำนักงานการวิจัยแห่งชาติ

เอกสารอ้างอิง

- Aukkapinyo, K., Sawangwong, S., P., P., and Kusakunniran, W. (2020). Localization and classification of rice-grain images using region proposals-based convolutional neural network. *International Journal of Automation and Computing*, 17, 233-246.

- Han, S. S., Park, G. H., Lim, W., Kim, M. S., Na, J. I., Park, I., and Chang, S. E. (2018). Deep neural networks show an equivalent and often superior performance to dermatologists in onychomycosis diagnosis: Automatic construction of onychomycosis datasets by region-based convolutional deep neural network. *PLoS ONE*, 13.
- He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. *The Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 27-30 June 2016, Las Vegas, NV, USA, 770-778.
- Joshi, M., and Patel, V. (2018). Convolutional neural network with transfer learning for rice type classification. *The proceedings of Tenth international Conference on Machine Vision (ICMV 2017)*, 13 April 2018, Vol. 10696, pp. 282-289.
- Lurstwut, B., and Pornpanomchai, C. (2017). Image analysis based on color, shape and texture for rice seed (*Oryza sativa* L.) germination evaluation. *Agriculture and Natural Resources*, 51(5), 383-389.
- Mehra, R. (2018). Breast cancer histology images classification: Training from scratch or transfer learning?. *ICT Express*, 4(4), 247-254.
- Rezende, E., Ruppert, G., Carvalho, T., Theophilo, A., Ramos, F., and de Geus, P. (2018). Malicious software classification using VGG16 deep neural network's bottleneck features. In *Information Technology-New Generation*. Springer, Cham., 51-59.
- Silva, C. S., and Sonnadara, D. (2013). Classification of rice grains using neural networks. *The Proceedings of the Technical Sessions (IPSL)*. Sri Lanka., 29, 9-14.
- Simonyan, K., and Zisserman, A. (2014). *Very deep convolutional networks for large-scale image recognition*. [Online]. Retrieved May 15, 2021 from: <https://arxiv.org/abs/1409.1556>.