การจำแนกหมวดหมู่สินค้าบนระบบอีคอมเมิร์ซโดยใช้เทคนิคการเรียนรู้เชิงลึกวิเคราะห์ภาพสินค้า

ภาณุพันธ์ เมฆสุวรรณ์[1] และ ฐิติรัตน์ ศิริบวรรัตนกุล[2*]
สถาบันบัณฑิตพัฒนบริหารศาสตร์ ถ.เสรีไทย แขวงคลองจั่น เขตบางกะปิ กรุงเทพฯ 10240

[*] Corresponding Author: thitirat@as.nida.ac.th
[1] นักศึกษาระดับบัณฑิตศึกษา คณะสถิติประยุกต์
[2] อาจารย์ คณะสถิติประยุกต์

ข้อมูลบทความ

ประวัติบทความ :
รับเพื่อพิจารณา : 18 เมษายน 2565
แก้ไข : 15 พฤษภาคม 2566
ตอบรับ : 19 พฤษภาคม 2566
DOI : 10.14456/kmuttrd.2023.10

คำสำคัญ : การเรียนรู้เชิงลึก / การจำแนกประเภทของรูปภาพ / การจำแนกหมวดหมู่ของสินค้า / อีคอมเมิร์ซ

บทคัดย่อ

งานวิจัยนี้มีจุดประสงค์เพื่อสร้างตัวแบบการเรียนรู้เชิงลึกสำหรับใช้จำแนกหมวดหมู่ของสินค้าบนระบบอีคอมเมิร์ซจากภาพถ่ายของสินค้าด้วยเทคนิคต่าง ๆ อีกทั้งเพื่อเปรียบเทียบประสิทธิภาพและหาตัวแบบที่ดีที่สุด โดยทดลองวิจัยจากข้อมูลรูปภาพของระบบอีคอมเมิร์ซ Shopee จำนวน 38 หมวดหมู่ รวม 106,309 รูป ซึ่งเป็นรูปสินค้าที่ใช้งานจริงและมีความซับซ้อนสูงในเชิงของการวิเคราะห์ข้อมูลจากรูปภาพ ทดลองใช้ตัวแบบทั้งหมด 6 ตัวแบบ และใช้ Loss Function ที่แตกต่างกัน 2 แบบ พบว่า ตัวแบบ EfficientNetB5 เป็นตัวแบบที่ดีที่สุดเมื่อพิจารณาจากค่า Accuracy จากขนาดของตัวแบบ และจากเวลาที่ใช้ในการจำแนกภาพ โดยมีค่า Accuracy เมื่อทดสอบกับชุดข้อมูลทดสอบ 84% ค่า Accuracy เมื่อทดสอบกับชุดข้อมูลทดสอบเพิ่มเติม 92% ใช้เวลาในการจำแนก 0.068 วินาทีต่อรูป และมีขนาดของตัวแบบเป็น 141.9 MB เมื่อนำตัวแบบดังกล่าวไปทดสอบประสิทธิภาพในการจำแนกภาพเปรียบเทียบกับแอปพลิเคชัน Shopee พบว่าตัวแบบ EfficientNetB5 สามารถจำแนกได้แม่นยำมากกว่าถึง 31.5% และเมื่อเปรียบเทียบกับการจำแนกภาพโดยผู้เข้าร่วมการทดลองจำนวน 4 รายก็พบว่าแม้ตัวแบบของผู้วิจัยจะมีความแม่นยำน้อยกว่าผู้เข้าร่วมการทดลอง 7.16% แต่สามารถจำแนกประเภทของภาพได้เร็วกว่าผู้เข้าร่วมการทดลอง ถึงประมาณ 25.5 เท่า

Classifying e-Commerce Product Categories Using Image-based Deep Learning

Panupun Maksuwan[1] and Thitirat Siriborvornratanakul[2*]
National Institute of Development Administration, Bangkapi, Bangkok 10240

[*] Corresponding Author: thitirat@as.nida.ac.th
[1] Student, Graduate School of Applied Statistics.
[2] Lecturer, Graduate School of Applied Statistics.

Article Info

Article History:
Received: April 18, 2022
Revised: May 15, 2023
Accepted: May 19, 2023
DOI : 10.14456/kmuttrd.2023.10

Keywords : Deep Learning / Image Classification / Categorizing products / E-commerce

Abstract

This research aimed to create a deep learning model to classify categories of products that are being sold on an e-commerce platform by analyzing images of the products. Different techniques were comparatively assessed and their efficiencies were compared to arrive at the best model. The public image dataset of the Shopee e-commerce system, consisting of 38 product categories, with a total of 106,309 images, was used. As these images belong to actual product images in the Shopee e-commerce system, they are highly complex in terms of computer vision and image analytics. Six models were tested along with 2 different loss functions. The results revealed that EfficientNetB5 was the best modelin terms of accuracy, considering the size of the model and the time used in classifying the images; the accuracy on the test set was noted to be 84% while the value was 92% on the additional test set. The inference time used was 0.068 seconds per image with the model size of 141.9 MB. When the model was tested for the task of image classification in comparison with the actual Shopee application, the EfficientNetB5 model was 31.5% more accurate. When compared to four human participants, although our model was 7.16% less accurate, the model was 25.5 times faster in classifying each image.

1. บทนำ

อีคอมเมิร์ซเป็นการผสมผสานกันของเทคโนโลยีอินเทอร์เน็ตและการแลกเปลี่ยนซื้อขายสินค้าที่ผู้ใช้งานสามารถซื้อขายสินค้าได้ทุกที่ทุกเวลาผ่านแอปพลิเคชันหรือเว็บไซต์ ตลาดอีคอมเมิร์ซมีสินค้าให้เลือกหลากหลาย อีกทั้งผู้ซื้อสามารถเปรียบเทียบราคาจากแหล่งต่างๆ ได้ง่าย ทำให้ปัจจุบันการซื้อขายสินค้าผ่านระบบอีคอมเมิร์ซได้รับความนิยมมากขึ้น จากรายงานผลการสำรวจมูลค่าพาณิชย์อิเล็กทรอนิกส์ประเทศไทย ปี 2562 [1] เฉพาะในประเทศไทย ใน พ.ศ. 2561 ธุรกิจอีคอมเมิร์ซมีมูลค่าสูงถึง 3,767,045.45 ล้านบาท และใน พ.ศ.2562 มีมูลค่าเพิ่มขึ้นเป็น 4,027,277.83 ล้านบาท หรือคิดเป็นอัตราการเติบโต 6.91%

สถานการณ์การแพร่ระบาดของโรคโควิด19 เป็นปัจจัยสำคัญปัจจัยหนึ่งที่ทำให้การซื้อขายผ่านระบบอีคอมเมิร์ซได้รับความนิยมมากขึ้นทั่วโลก จากรายงานของ The Future Shopper report [2] พบว่าเมื่อเทียบช่วงเวลาปกติกับช่วงที่มีการแพร่ระบาดของโรคโควิด 19 ในประเทศไทย การซื้อขายสินค้าผ่านระบบอีคอมเมิร์ซและระบบออนไลน์ในช่วงที่มีการแพร่ระบาดของโรคโควิด19 มีอัตราการเติบโตสูงขึ้นเฉลี่ย 11% และทั่วโลกมีอัตราการเติบโตสูงขึ้นเฉลี่ย 6% นอกจากนี้การสำรวจยังพบว่า 92% ของผู้บริโภคชาวไทยที่ตอบแบบสอบถามจะยังคงซื้อสินค้าบนระบบอีคอมเมิร์ซและระบบออนไลน์ต่อไปหลังสิ้นสุดการระบาดของโรคโควิด 19

ด้วยอัตราการเติบโตที่สูงอย่างต่อเนื่องของธุรกิจอีคอมเมิร์ซส่งผลให้สินค้าที่วางขายอยู่ในระบบอีคอมเมิร์ซมีปริมาณมากและมีประเภทอีกทั้งความหลากหลายของสินค้าสูง โดยเฉพาะในระบบอีคอมเมิร์ซแบบ C2C (Consumer-to-Consumer) [3] ที่เป็นรูปแบบธุรกิจที่ผู้ใช้งานสามารถแลกเปลี่ยนซื้อขายสินค้ากันเองได้ เช่น Lazada, Shopee, AliExpress ในการนี้การจัดหมวดหมู่หรือแบ่งแยกประเภทของสินค้าให้ถูกต้องถือเป็นกุญแจสำคัญสำหรับธุรกิจอีคอมเมิร์ซที่จะช่วยสร้างประสบการณ์ที่ดีให้กับผู้ใช้งาน ช่วยให้ผู้ใช้งานสามารถค้นหาและเข้าถึงสินค้าที่ต้องการจากสินค้าจำนวนมากบนระบบอีคอมเมิร์ซได้อย่างถูกต้องและรวดเร็ว [4]

ปัจจุบัน การตรวจสอบและแก้ไขหมวดหมู่ของสินค้าจำนวนมากที่ถูกวางขายในแต่ละวันให้ถูกต้องนั้นจะทำโดยพนักงานของบริษัทอีคอมเมิร์ซ ยิ่งมีจำนวนสินค้ามากขึ้นก็ยิ่งต้องใช้จำนวนพนักงานที่มากขึ้นในการตรวจสอบและจัดหมวดหมู่สินค้าให้ถูกต้องซึ่งส่งผลให้บริษัทอีคอมเมิร์ซมีต้นทุนมากขึ้น ทำให้ปัจจุบันบริษัทอีคอมเมิร์ซและนักวิจัยได้เริ่มศึกษาในการใช้เทคนิคการเรียนรู้เชิงลึก (Deep learning) มาช่วยจำแนกหมวดหมู่ของสินค้าแบบอัตโนมัติเพื่อเพิ่มประสิทธิภาพในการจำแนกหมวดหมู่และช่วยลดต้นทุน เช่น Shopee, Taobao และ Cdiscount อย่างไรก็ตามประสิทธิภาพความแม่นยำของตัวแบบ Deep learning ที่นำมาใช้ยังคงเป็นเรื่องท้าทายของบริษัทอีคอมเมิร์ซหลาย ๆ แห่งโดยเฉพาะการจำแนกหมวดหมู่สินค้าจากรูปภาพ (Image Classification) เนื่องจากในความเป็นจริงข้อมูลรูปภาพสินค้าบนระบบอีคอมเมิร์ซค่อนข้างมีปัญหาในหลายๆด้าน เช่น มีจำนวนหมวดหมู่ของสินค้าที่ค่อนข้างมาก ภาพสินค้าในแต่ละหมวดหมู่ค่อนข้างมีความหลากลาย มีพื้นหลังที่ซับซ้อน จำนวนรูปภาพในแต่ละหมวดหมู่มีไม่เท่ากัน คุณภาพของภาพสินค้าบางภาพมีคุณภาพต่ำ เป็นต้น

ปัจจุบันเทคนิค Deep learning ได้รับการยอมรับและถูกนำมาใช้อย่างกว้างขวางมากขึ้นโดยเฉพาะในงานที่เกี่ยวข้องกับคอมพิวเตอร์วิทัศน์ (Computer vision) และการวิเคราะห์ข้อมูลจากภาพ (Image Analytics) เพื่อใช้ในงานด้านการจำแนกประเภทของรูปภาพ (Image classification) ซึ่งได้ถูกนำมาประยุกต์ใช้กับงานและธุรกิจหลากหลายประเภท โดยในงานวิจัยนี้จะพูดถึงการทดลองและการสร้างตัวแบบ Deep learning แบบต่างๆ ในการจำแนกหมวดหมู่ของสินค้าแบบอัตโนมัติจากรูปภาพสินค้าของจริงที่ถูกใช้งานบนระบบอีคอมเมิร์ซ เพื่อเปรียบเทียบประสิทธิภาพและหาตัวแบบที่มีประสิทธิภาพและประสิทธิผลดีที่สุด

2. การทบทวนงานวิจัยที่เกี่ยวข้อง

โดยทั่วไปการจำแนกหมวดหมู่ของสินค้าสามารถใช้ข้อมูลในการจำแนกได้หลายวิธี เช่น การจำแนกด้วยคำบรรยายของสินค้าหรืองานของ Partalas และ Balikas [5] ที่เสนอการจำแนกด้วยวิธี Text classification แต่ในงานวิจัยนี้ผู้วิจัยจะพูดถึงเฉพาะเทคนิคการจัดหมวดหมู่สินค้าจากภาพเท่านั้น Oyewole และ Olugbara [6] ได้เสนอวิธีจำแนกหมวดหมู่สินค้าจากภาพ (Image classification) ด้วยเทคนิค Traditional machine learning โดยใช้ Eigen Color ของภาพในการทำ Feature extraction ร่วมกับตัวแบบ Support vector machine (SVM) และ Multilayer

Perceptron (MLP) - Artificial neural network (ANN) ซึ่งให้ประสิทธิภาพสูงสุดที่ Accuracy 87.20% ข้อสังเกตของงานวิจัยนี้คือใช้ชุดข้อมูลรูปสินค้าส่วนใหญ่ที่มีภาพพื้นหลังเป็นสีขาวหรือสีเรียบๆ กล่าวคือหากทำ Feature extraction โดยใช้ Eigen Color กับรูปภาพสินค้าที่มีพื้นหลังที่ค่อนข้างหลากหลายอาจส่งผลให้ตัวแบบมีประสิทธิภาพลดลงได้ ปัจจุบันงานวิจัยส่วนใหญ่ที่เกี่ยวข้องกับการจำแนกหมวดหมู่ของภาพ (Image classification) มักจะใช้เทคนิค Transfer Learning บนพื้นฐานของ Convolutional neural network (CNN) ที่ผ่านการทำ Pre-training บนฐานข้อมูล ImageNet ซึ่งเป็นส่วนหนึ่งของการแข่งขัน ImageNet Large Scale Visual Recognition Challenge [7] เช่น VGG19 [8] ResNet [9] EfficientNet [10] เนื่องจากการใช้ตัวแบบ Transfer Learning ช่วยลดเวลาและทรัพยากรในการ Training ลงได้ โดยที่ผ่านมา Kumar และคณะ [7] ได้ทดลองจำแนกหมวดหมู่ของสินค้าจากรูปภาพบนระบบอีคอมเมิร์ซโดยใช้ตัวแบบ Transfer learning 3 ตัวแบบเพื่อเปรียบเทียบประสิทธิภาพ ซึ่งประกอบไปด้วย VGG19, Inception V3 และ ResNet50 พบว่า VGG19 ให้ผลที่แม่นยำที่สุด Accuracy 88% ซึ่งคล้ายกับงานวิจัยของ Li และคณะ [11] ที่ใช้ตัวแบบ Transfer learning VGG19 และ Inception V3 ในการทดลองจำแนกหมวดหมู่ของสินค้าจากรูปภาพบนระบบอีคอมเมิร์ซ ซึ่งผลประสิทธิภาพของทั้งสองตัวแบบออกมาใกล้เคียงกันที่ Accuracy 80% นอกจากนี้ยังมีงานวิจัยของ Herdian และคณะ [4] ที่ได้ทดลองใช้ตัวแบบ Transfer learning อีก 5 ตัวแบบ ได้แก่ MobileNetV1, MoblieNetV2, NASNetMobile, NASNetLarge และ DensNet121 ซึ่งจากผลการทดลองพบว่า NASNetLarge มีประสิทธิภาพในการจำแนกหมวดหมู่ของภาพสินค้าได้ดีที่สุดที่ Accuracy 84%

งานวิจัยทั้งสามงานข้างต้นเน้นไปทางด้านการทดลองและเปรียบเทียบประสิทธิภาพในการจำแนกสินค้าจากรูปภาพบนระบบอีคอมเมิร์ซของตัวแบบ Transfer learning แบบต่าง ๆ แต่จุดสังเกตของทั้งสามงานคือการใช้ข้อมูลรูปภาพสินค้าที่ไม่ซับซ้อนโดยเฉพาะภาพพื้นหลังส่วนมากจะเป็นสีขาว คุณภาพของรูปสินค้าค่อนข้างดี และทำการทดลองเฉพาะสินค้าในหมวดหมู่เครื่องแต่งกายและเครื่องประดับเท่านั้น ซึ่งแตกต่างจากงานวิจัยนี้ที่ใช้ข้อมูลรูปภาพสินค้าจาก Shopee ที่มีความซับซ้อนและความหลากหลายของภาพ มีพื้นหลังที่หลากหลาย ซึ่งอาจจะส่งผลต่อประสิทธิภาพของตัวแบบในการจำแนกหมวดหมู่โดยเฉพาะรูปที่มีสัดส่วนของพื้นหลังค่อนข้างเยอะ [12] คุณภาพของรูปภาพสินค้าบางภาพมีคุณภาพต่ำซึ่งอาจส่งผลให้กับประสิทธิภาพของตัวแบบด้วยเช่นกัน [13] อีกทั้งมีหมวดหมู่ของสินค้าที่หลากหลายไม่ใช่แค่หมวดหมู่เครื่องแต่งกายและเครื่องประดับ เช่น เจลแอลกอฮอล์ หน้ากากอนามัย หม้อหุงข้าว ขวดนมเด็ก ขนม เป็นต้น

3. วิธีดำเนินการวิจัย

3.1 การรวบรวมและเตรียมข้อมูล

ในการวิจัยนี้จะแบ่งข้อมูลออกเป็น 2 ชุด ได้แก่ 1) ข้อมูลที่ใช้สำหรับการสร้างตัวแบบเพื่อใช้สำหรับการจำแนกหมวดหมู่ของสินค้า ซึ่งมีความหลากหลายของรูปภาพในแต่ละหมวดหมู่ ความละเอียดของรูปภาพ และ 2) ข้อมูลที่ใช้สำหรับการทดสอบร่วมกับแอปพลิเคชัน Shopee เพื่อเปรียบเทียบประสิทธิภาพระหว่างตัวแบบที่ได้จากงานวิจัยนี้และแอปพลิเคชัน Shopee

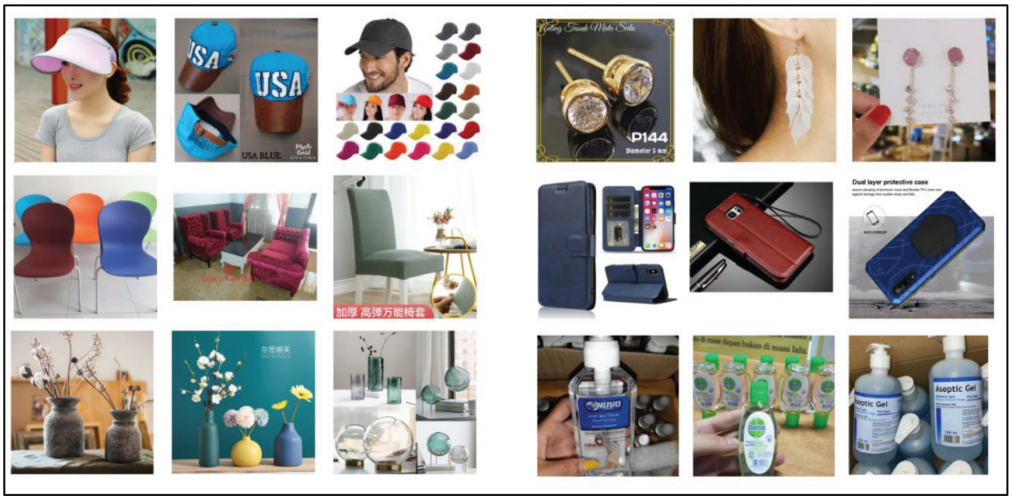
3.1.1 ข้อมูลที่ใช้สำหรับการสร้างโมเดล

ใช้รูปภาพสินค้าที่เป็นภาพนิ่งจากข้อมูลสาธารณะของบริษัท Shopee ซึ่งเป็นหนึ่งในหัวข้อของการแข่งขัน Shopee Code League 2020 Data Science (www.kaggle.com/c/shopee-product-detection-open/data) กำหนดหมวดหมู่ของสินค้าในการจำแนกทั้งหมด 35 หมวดหมู่ โดยอ้างอิงจำนวนหมวดหมู่จากแหล่งที่มาของข้อมูล Shopee Code League 2020 ซึ่งประกอบไปด้วยจำนวนหมวดหมู่ตั้งต้นทั้งหมด 42 หมวดหมู่ แต่จากการสำรวจข้อมูลเบื้องต้นผู้วิจัยพบว่ามีบางหมวดหมู่ที่ข้อมูลมีปัญหาไม่เหมาะในการนำมาใช้ เช่น บางหมวดหมู่มีรูปที่ผิดจำนวนมาก บางหมวดหมู่มีความคล้ายกันมากโดยเฉพาะหมวดหมู่ที่เกี่ยวข้องกับเสื้อผ้าผู้หญิง ทำให้ต้องมีการตัดและยุบรวมบางหมวดหมู่ ทำให้ได้ออกมาเป็น 35 หมวดหมู่สำหรับใช้ในงานวิจัยนี้

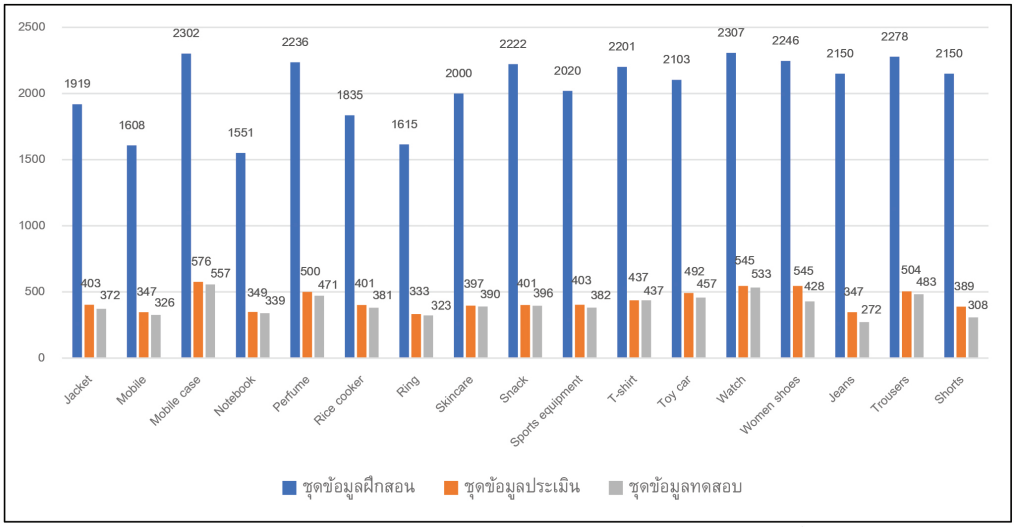
นอกจากนี้ ผู้วิจัยยังทำการเพิ่มเติมข้อมูลภาพสินค้าในหมวดหมู่กางเกงจากแอปพลิเคชัน Shopee อีก 3 หมวดหมู่ รวมเป็น 38 หมวดหมู่ เนื่องจากชุดข้อมูลเดิมจาก Shopee Code League 2020 แบ่งข้อมูลรูปภาพหมวดหมู่กางเกงออกเป็น 2 หมวดหมู่ ได้แก่ กางเกงผู้ชายและกางเกงผู้หญิง ซึ่งไม่สอดคล้องกับการแบ่งหมวดหมู่บน แอปพลิเคชัน Shopee บนโทรศัพท์มือถือที่มีการใช้งานจริงซึ่งแบ่งหมวดหมู่ของกางเกงตามประเภท เช่น กางเกงยีนส์ กางเกงขายาว และกางเกงขาสั้น

ทำให้การทดสอบเปรียบเทียบประสิทธิภาพระหว่างตัวแบบในงานวิจัยนี้และแอปพลิเคชัน Shopee อาจทำได้ยาก ด้วยเหตุที่กล่าวไปข้างต้นผู้วิจัยจึงเลือกแบ่งหมวดหมู่ของกางเกงออกเป็น 3 หมวดหมู่ ได้แก่ กางเกงยีนส์ กางเกงขายาวและกางเกงขาสั้นแทนที่จะเป็นหมวดหมู่กางเกงผู้ชายและกางเกงผู้หญิง

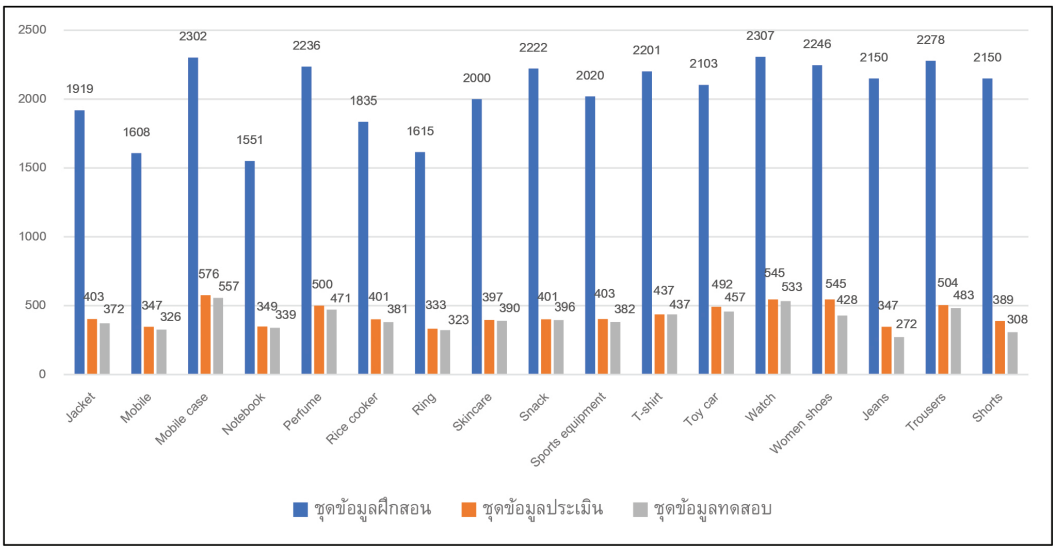
แต่เมื่อแบ่งหมวดหมู่กางเกงออกมาเป็น 3 หมวดหมู่แล้วก็เกิดปัญหาว่าจำนวนข้อมูลในแต่ละหมวดหมู่เหลือน้อยลงมากและไม่เพียงพอต่อการนำไปใช้งาน ผู้วิจัยจึงจำเป็นต้องหาชุดข้อมูลรูปภาพของสินค้าเพิ่มเติมจากแอปพลิเคชัน Shopee ทำให้รวมทั้งสิ้นมีจำนวนรูปภาพสินค้าทั้งหมด 106,309 รูป โดยผู้วิจัยแบ่งชุดข้อมูลภาพถ่ายออกเป็น 3 ชุดด้วยวิธีการสุ่ม ทำให้ได้ชุดข้อมูลสำหรับฝึกสอนตัวแบบ (Training set) จำนวน 75,307 รูป หรือประมาณ 70% ของข้อมูลชุดรูปภาพทั้งหมด ชุดข้อมูลสำหรับประเมิน (Validation set) จำนวน 16,140 รูป หรือประมาณ 15% ของข้อมูลชุดรูปภาพทั้งหมด และชุดข้อมูลสำหรับทดสอบ (Test set) จำนวน 14,892 รูป หรือประมาณ 15% ของข้อมูลชุดรูปภาพทั้งหมด รูปที่ 1 ถึง 3 แสดงตัวอย่างรูปภาพสินค้าในแต่ละหมวดหมู่และจำนวนของรูปภาพในแต่ละชุดข้อมูล



รูปที่ 1 ภาพตัวอย่างของรูปภาพสินค้าจากข้อมูลสาธารณะของ Shopee



รูปที่ 2 กราฟแสดงจำนวนรูปภาพสินค้าในแต่ละหมวดหมู่ (กราฟที่ 1)



รูปที่ 3 กราฟแสดงจำนวนรูปภาพสินค้าในแต่ละหมวดหมู่ (กราฟที่ 2)

จากรูปที่ 2 และ 3 จะเห็นว่ามีข้อมูลบางหมวดหมู่ที่จำนวนข้อมูลมีปริมาณค่อนข้างน้อยกว่าหมวดหมู่อื่น ๆ (Imbalanced data) ซึ่งอาจส่งผลต่อความแม่นยำในการสร้างตัวแบบ เพื่อแก้ปัญหานี้ผู้วิจัยจึงเพิ่มการทำ Data Augmentation ด้วยคำสั่ง Preprocessing ImageDataGenerator ในไลบรารี Keras (กำหนดค่าพารามิเตอร์ zoom range = 0.5-1.3, rotation range = 45, width shift = 0.5, height shift = 0.5, brightness range = 0.7-1.5, horizontal flip = true, vertical flip = true, shear range = 0.5) สำหรับชุดข้อมูลสำหรับเรียนรู้ (Training set) เพื่อเพิ่มจำนวนรูปภาพให้แต่ละหมวดหมู่ให้มีจำนวนพอๆกัน อีกทั้งการทำ Data Augmentation ยังทำให้รูปภาพมีความหลากหลายมากขึ้นช่วยให้ตัวแบบสามารถนำไปใช้งานจริงได้ดี (Generalization) และลดปัญหาการเกิด Overfitting โดยเมื่อเสร็จสิ้นกระบวนการ Data Augmentation แล้วจะได้จำนวนรูปภาพทั้งหมด 38 หมวดหมู่ รวมทั้งสิ้นจำนวน 128,736 รูป แบ่งชุดข้อมูลภาพถ่ายออกเป็น 3 ชุดด้วยวิธีการสุ่มซึ่งประกอบไปด้วย ชุดข้อมูลสำหรับฝึกสอนตัวแบบจำนวน 92,682 รูป ชุดข้อมูลสำหรับประเมินจำนวน 21,246 รูป และชุดข้อมูลสำหรับทดสอบจำนวน 14,892 รูป

3.1.2 ข้อมูลที่ใช้สำหรับการสร้างโมเดล

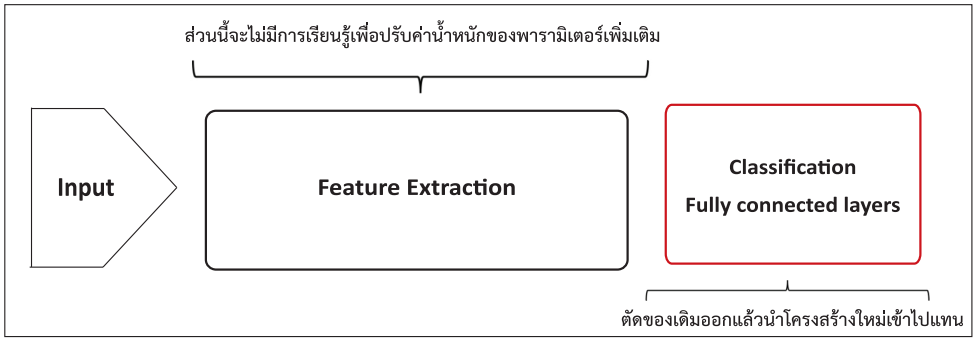
ชุดข้อมูลทดสอบเพิ่มเติม (Additional test set) หมายถึง ชุดข้อมูลรูปภาพจำนวน 1,900 รูป โดยแบ่งเป็น 50 รูป ต่อ 1 หมวดหมู่ ชุดข้อมูลนี้ผู้วิจัยรวบรวมมาจากอินเทอร์เน็ต โดยเลือกมาเฉพาะรูปภาพสินค้าที่ไม่มีปรากฏอยู่ในแอปพลิเคชัน Shopee จุดประสงค์ของชุดข้อมูลทดสอบเพิ่มเติมนี้คือเพื่อใช้ทำการทดสอบประสิทธิภาพเชิงเปรียบเทียบ ระหว่างแบบจำลองที่ผู้วิจัยพัฒนาขึ้นกับฟีเจอร์ค้นหาด้วยรูปภาพของแอปพลิเคชัน Shopee เอง

3.2 การฝึกสอนและการสร้างตัวแบบ

งานวิจัยนี้ใช้ตัวแบบการเรียนรู้เชิงลึก ได้แก่ VGG19, ResNet50, ResNet152, EfficientNetB0, EfficientNetB5 และ EfficientNetB7 โดยทั้งหมดคือแบบจำลองการเรียนรู้เชิงลึกประเภท Convolutional Neural Network ที่ถูกฝึกสอน (pre-train) ด้วยชุดข้อมูลภาพ ImageNet มาแล้วเหตุผลในการเลือกตัวแบบ VGG และ ResNet เนื่องจากตัวแบบทั้ง 2 ได้รับความนิยมสูงและให้ประสิทธิภาพที่ดีในงานวิจัยต่างๆ โดยเฉพาะงานด้าน Image classification ตัวอย่างเช่น ในงานการจำแนกหมวดหมู่สินค้าจากรูปของKumar และคณะ [7] รวมถึง Li และคณะ [7] ที่มีการใช้ตัวแบบ VGG และ ResNet ในงานวิจัย ในส่วนของเหตุผลในการเลือกตัวแบบ EfficientNet นั้น เนื่องจากเป็นตัวแบบขนาดเล็กที่ให้ประสิทธิภาพสูง และใช้จำนวนพารามิเตอร์น้อย โดยอ้างอิงจากงานวิจัยของ Tan และ Le [10] และจากการทบทวนวรรณกรรมที่เกี่ยวข้อง ผู้วิจัยยังไม่พบว่ามีงานวิจัยด้านการจำแนกหมวดหมู่สินค้าจาก

รูปที่ใช้ตัวแบบ EfficientNet

ในงานวิจัยนี้ผู้วิจัยจะใช้เทคนิคโอนถ่ายความรู้ (Transfer learning) ซึ่งเป็นการนำโครงข่ายของตัวแบบที่ผ่านการทำ Pre-training มาแล้ว มาตัดส่วนที่เป็นการทำ Classification (fully connected layer) ของเดิมออกแล้วนำโครงสร้างใหม่เข้าไปเพิ่มแทน โดยในการเรียนรู้ฝึกฝนตัวแบบจะไม่มีการเรียนรู้เพื่อปรับค่าน้ำหนักของพารามิเตอร์ในส่วนของการทำ Feature Extraction เดิม แต่จะฝึกฝนตัวแบบเพื่อปรับค่าของพารามิเตอร์เฉพาะส่วนที่เป็น Classification (fully connected layer) ที่ถูกเพิ่มเข้าไปในโครงข่ายของตัวแบบเท่านั้นดังแสดงในรูปที่ 4 ทั้งนี้ส่วนของกรอบสีแดงซึ่งเป็นส่วนต่อเติมเองของผู้วิจัยนั้นประกอบไปด้วยลำดับชั้น GlobalAveragePooling2D, Fully connected layers, Dropout และ Output layers ตามลำดับ



รูปที่ 4 การใช้เทคนิค Transfer learning

ในงานวิจัยนี้ผู้วิจัยทดลองใช้ Loss function 2 แบบ ได้แก่ Categorical cross entropy และ Focal loss [14] เพื่อเปรียบเทียบประสิทธิภาพระหว่างกัน โดยข้อดีของ Focal loss ในสมการที่ 1 คือสามารถช่วยเพิ่มประสิทธิภาพให้กับตัวแบบในกรณีที่จำนวนของข้อมูลในแต่ละหมวดหมู่ไม่สมดุลกัน (Imbalanced dataset) นอกจากนี้งานวิจัยของ Zhao และคณะ [15] ยังกล่าวไว้ว่า Focal loss สามารถช่วยเพิ่มประสิทธิภาพให้ตัวแบบในกรณีที่ตัวอย่างชุดข้อมูลมีความซับซ้อนและจำแนกได้ยาก โดย Focal loss จะช่วยให้ตัวแบบพุ่งความสนใจไปที่ข้อมูลที่จำแนกได้ยากมากขึ้นและลดความสำคัญของข้อมูลที่ตัวแบบจำแนกได้ถูกต้องด้วยความมั่นใจสูงอยู่แล้วลง

$$FL(p_t) = -\alpha_t(1-p_t)^{\gamma}\log(p_t) \qquad (1)$$

เมื่อ p_t คือ ค่าความน่าจะเป็น (Probability) ของ ground truth class หรือเป็นค่าความมั่นใจของการทำนายผลลัพธ์ และ α_t คือ ค่าน้ำหนักที่ผู้วิจัยต้องกำหนดสำหรับควบคุมเรื่อง Imbalanced dataset โดยทั่วไปจะถูกกำหนดค่าอยู่ที่ 0.25 และ γ คือ ค่าที่ผู้วิจัยต้องกำหนดว่าจะให้ค่า loss มากน้อยเพียงใดเมื่อเทียบกับค่าความมั่นใจของการทำนายผลลัพธ์ (Probability ของ ground truth class) โดยทั่วไปจะถูกกำหนดค่าเท่ากับ 2

โดยตัวแบบแต่ละตัวที่ทดลองจะมีจำนวนพารามิเตอร์อ้างอิงตามตารางที่ 1 และมีการกำหนดขนาด Batch size = 32 ขนาดของ Input Image = (300,300) ใช้ Optimizer คือ Adam กำหนด Learning rate สำหรับตัวแบบ VGG19, ResNet50 และ ResNet152 ที่ค่า 0.001 และ สำหรับตัวแบบ EfficientNet ที่ค่า 0.0001 โดยทุกตัวแบบจะถูกฝึกสอนผ่านเครื่องมือ Python เวอร์ชัน 3.7 บน Google colab pro+ ใช้ GPU Tesla V100-SXM2-16GB และใช้ Library TensorFlow Keras เวอร์ชัน 2.7.0

ตารางที่ 1 จำนวนพารามิเตอร์ของแต่ละตัวแบบ

ตัวแบบ	จำนวนพารามิเตอร์ทั้งหมด	จำนวนพารามิเตอร์ที่เรียนรู้ (Trainable parameters)	จำนวนพารามิเตอร์ที่ไม่ได้เรียนรู้ (Non-trainable parameters)
VGG19	20,428,134	403,750	20,024,384
ResNet50	24,777,894	1,190,182	23,587,712
ResNet152	59,561,126	1,190,182	58,370,944
EfficientNetB0	6,027,209	1,977,638	4,049,571
EfficientNetB5	31,277,597	2,764,070	28,513,527
EfficientNetB7	67,386,045	3,288,358	28,513,527

จากตารางที่ 1 จำนวนพารามิเตอร์ของแต่ละตัวแบบเป็นจำนวนจากตัวแบบที่ใช้งานแบบ Transfer learning ซึ่งจะเป็นการใช้เทคนิค Fixed Feature extractor (Head) ทำให้พารามิเตอร์ในส่วนที่เป็น Feature extractor ของในแต่ละตัวแบบไม่ถูกใช้งานและไม่รวมอยู่ในตารางที่ 1

3.3 การฝึกสอนและการสร้างตัวแบบ

3.3.1 Accuracy, Precision, Recall และ F1-Score

ในงานวิจัยนี้จะประเมินผลของตัวแบบด้วยตัวชี้วัดที่สามารถคำนวณได้ 4 ค่า ได้แก่ Accuracy, Precision, Recall และ F1-Score ซึ่งมีวิธีการคำนวณแสดงในสมการที่ 2 ถึง 5 อย่างไรก็ตามในงานวิจัยนี้จะใช้การวัดค่า Accuracy เป็นหลักเพื่อวัดประสิทธิภาพโดยรวมของแต่ละตัวแบบ และเน้นการพิจารณาค่า Recall ของแต่ละหมวดหมู่ เพื่อดูความแม่นยำและประสิทธิภาพในการจำแนกเป็นรายหมวดหมู่

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

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

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

$$F1\ score = \frac{2*Precision*Recall}{Precision+Recall} \quad (5)$$

เมื่อ TP (True Positive) คือ จำนวนของข้อมูลที่เป็นหมวดหมู่ที่กำลังพิจารณา แล้วตัวแบบจำแนกถูกว่าเป็นหมวดหมู่ที่กำลังพิจารณา

FP (False Positive) คือ จำนวนของข้อมูลที่ไม่ใช่หมวดหมู่ที่กำลังพิจารณา แต่ตัวแบบจำแนกว่าเป็นหมวดหมู่ที่กำลังพิจารณา

TN (True Negative) คือ จำนวนของข้อมูลที่ไม่ใช่หมวดหมู่ที่กำลังพิจารณา แล้วตัวแบบจำแนกถูกว่าไม่ใช่หมวดหมู่ที่กำลังพิจารณา

FN (False Negative) คือ จำนวนของข้อมูลที่เป็นหมวดหมู่ที่กำลังพิจารณา แต่ตัวแบบจำแนกว่าไม่ใช่หมวดหมู่ที่กำลังพิจารณา

Accuracy คือ การวัดความถูกต้องแม่นยำของตัวแบบโดยพิจารณารวมทุกหมวดหมู่

Precision คือ การวัดความถูกต้องแม่นยำของตัวแบบในแต่ละหมวดหมู่ โดยวัดจำนวนข้อมูลที่ถูกจำแนกว่าเป็นหมวดหมู่ที่กำลังพิจารณาและถูกต้องเทียบกับจำนวนข้อมูลทั้งหมดที่ถูกจำแนกว่าเป็นหมวดหมู่ที่กำลังพิจารณาทั้งถูกและผิด

Recall คือ การวัดความถูกต้องแม่นยำของตัวแบบในแต่ละหมวดหมู่ โดยวัดจำนวนข้อมูลที่ถูกจำแนกว่าเป็นหมวดหมู่ที่กำลังพิจารณาและถูกต้องเทียบกับจำนวนข้อมูลทั้งหมดของหมวดหมู่กำลังพิจารณา

F1-Score คือ การวัดความถูกต้องแม่นยำของตัวแบบในแต่ละหมวดหมู่ โดยคำนวณจากค่า Precision และ Recall ประกอบกัน

3.3.2 การประเมินผลแอปพลิเคชัน Shopee เพื่อเปรียบเทียบผลลัพธ์กับตัวแบบ

ในงานวิจัยนี้นอกจากจะมีการประเมินตัวแบบหลายตัวเปรียบเทียบกันเองแล้ว ผู้วิจัยยังทำการทดสอบความถูกต้องแม่นยำของตัวแบบเปรียบเทียบกับแอปพลิเคชัน Shopee อีก

ด้วย โดยการประเมินผลจะใช้คุณสมบัติการค้นหาสินค้าด้วยรูปบนแอปพลิเคชัน Shopee เพื่อตรวจสอบว่าแอปพลิเคชัน Shopee สามารถจำแนกหมวดหมู่ของสินค้าจากรูปภาพได้ดีเพียงใด โดยอ้างอิงหลักการในการประเมินผลตามงานวิจัยเรื่องการประเมินประสิทธิภาพการค้นหาด้วยรูปภาพบน Google ของ Bitirim และคณะ [16] ที่นำรูปภาพสินค้า 25 หมวดหมู่ จำนวน 2,500 รูป แบ่งเป็นหมวดหมู่ละ 100 รูป มาทดสอบประสิทธิภาพการค้นหาด้วยรูปภาพบน Google โดยในการทดสอบ Bitirim และคณะได้แบ่งข้อมูลรูปภาพในแต่ละหมวดหมู่ออกเป็น 5 กลุ่ม หรือเรียกว่าการแบ่งข้อมูลออกเป็น Cut-off point โดยจำนวนรูปภาพในแต่ละ Cut-off point จะมีจำนวน 20, 40, 60, 80 และ 100 รูป ตามลำดับ เพื่อเปรียบเทียบประสิทธิภาพระหว่าง Cut-off point ว่าจำนวนข้อมูลรูปภาพมีผลต่อประสิทธิภาพในการค้นหาด้วยรูปภาพบน Google อย่างไร

ในงานวิจัยนี้ ผู้วิจัยทำการทดลองด้วยแอปพลิเคชัน Shopee บนโทรศัพท์มือถือระบบปฏิบัติการ Android แอปพลิเคชันเวอร์ชัน 2.81.31 ดำเนินการทดสอบระหว่างวันที่ 10 ถึง 21 มกราคม พ.ศ. 2565 โดยใช้ชุดข้อมูลทดสอบเพิ่มเติม (รายละเอียดในหัวข้อ 3.1.2) จำนวน 1,900 รูป ซึ่งแบ่งเป็น 50 รูป ต่อ 1 หมวดหมู่ ในการประเมินผลจะแบ่งข้อมูลออกเป็น 4 Cut-off point โดยในแต่ละ Cut-off point จะมีการประเมินค่า Recall ซึ่งการแบ่งข้อมูลเป็น Cut-off point สามารถทำได้ 2 วิธี ได้แก่

1) แบ่งข้อมูลออกเป็น Cut-off point ที่ 12 ภาพ, 24 ภาพ, 37 ภาพ และ 50 ภาพ ของในแต่ละหมวดหมู่ ซึ่งจำนวนรูปในแต่ละ Cut-off point มีการกำหนดขึ้นมาจากการแบ่งแต่ละ Cut-off point เป็นเปอร์เซ็นต์ โดย Cut-off point ที่ 12 ภาพ จะเท่ากับประมาณ 25% ของรูปภาพทั้งหมด Cut-off point ที่ 24 ภาพ จะเท่ากับประมาณ 50% ของรูปภาพทั้งหมด Cut-off point ที่ 37 ภาพ จะเท่ากับประมาณ 75% ของรูปภาพทั้งหมด และ Cut-off point ที่ 50 ภาพ จะเท่ากับ 100% ของรูปภาพทั้งหมด โดยที่ Cut-off point ที่ 24 ภาพ จะมีรูปภาพเดิม 12 ภาพจาก cut-off point ที่ 12 ภาพ และรูปภาพใหม่อีก 12 ภาพ ที่ cut-off point ที่ 37 ภาพ จะมีรูปภาพเดิม 24 ภาพจาก cut-off point ที่ 24 ภาพ และรูปภาพใหม่อีก 13 ภาพ

2) แบ่งข้อมูลออกเป็น Cut-off point ที่ 25% ของรูปภาพทั้งหมด, 50% ของรูปภาพทั้งหมด, 75% ของรูปภาพทั้งหมด, 100% ของรูปภาพทั้งหมด ของในแต่ละหมวดหมู่

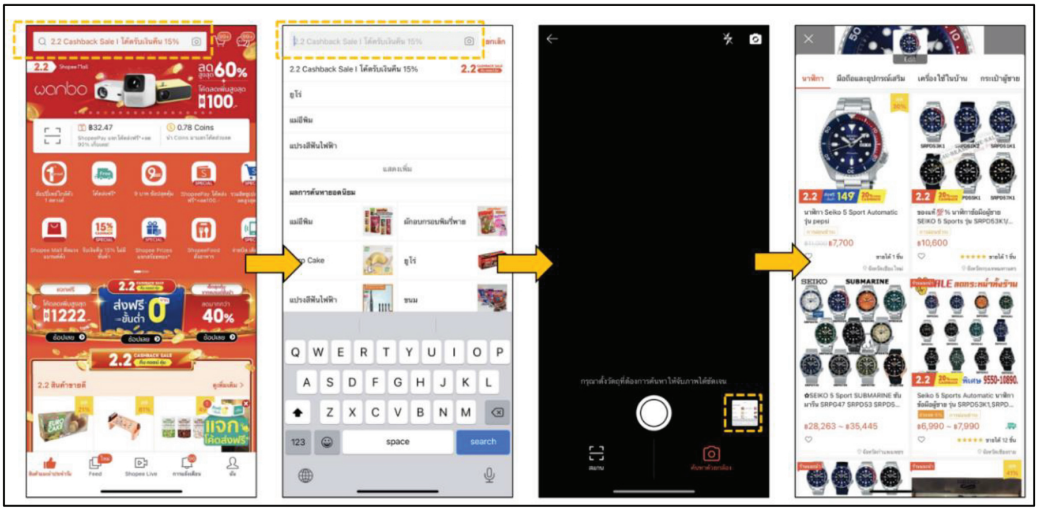
เกณฑ์ที่ผู้วิจัยใช้ในการประเมินผลการค้นหาที่ได้จากฟีเจอร์ค้นหาสินค้าด้วยรูปของแอปพลิเคชัน Shopee มีดังต่อไปนี้

1) ตรวจสอบผลที่ได้จากการค้นหาว่าหมวดหมู่ที่แสดงออกมาเป็นหมวดหมู่สินค้าที่ถูกต้องหรือไม่

2) ถ้าหมวดหมู่ของสินค้าที่แสดงออกมามีลักษณะที่เจาะจงคุณลักษณะของสินค้าก็สามารถยอมรับได้ เช่น ถ้ารูปภาพสินค้าเป็นกางเกงยีนส์ซึ่งหมวดหมู่ที่เราคาดหวังว่าจะได้จากการค้นหาคือหมวดหมู่ “กางเกงยีนส์” แต่เมื่อนำไปค้นหาแล้วกลับได้ผลออกมาเป็นหมวดหมู่ “กางเกงยีนส์ทรงกระบอก” ซึ่งเป็นหมวดหมู่สินค้าที่มีการเจาะจงคุณลักษณะ ในกรณีนี้สามารถยอมรับว่าแอปพลิเคชันจำแนกได้ถูกต้อง

3) ในกรณีที่แอปพลิเคชันแสดงหมวดหมู่ที่ได้จากการค้นหาไม่ถูกต้องหรือไม่มีหมวดหมู่ที่กำลังพิจารณา จะใช้วิธีการดูสินค้าที่แอปพลิเคชันแนะนำว่ามีความสอดคล้องกับภาพของสินค้าที่เราใช้ในการทดสอบหรือไม่ เช่น รูปภาพสินค้าที่ใช้ทดสอบเป็นสินค้าน้ำหอม ซึ่งหมวดหมู่ที่เราคาดหวังว่าจะได้จากการค้นหาคือหมวดหมู่ “น้ำหอม” แต่เมื่อนำไปค้นหาแล้วแอปพลิเคชันไม่สามารถระบุได้ว่าเป็นหมวดหมู่น้ำหอม แต่สามารถแนะนำสินค้าในหมวดหมู่น้ำหอมได้ ก็จะถือว่าแอปพลิเคชันจำแนกได้ถูกต้อง

สำหรับขั้นตอนในการทดสอบค้นหาสินค้าด้วยภาพบนแอปพลิเคชันอีคอมเมิร์ซ Shopee สามารถทำได้ดังนี้ 1) เข้าแอปพลิเคชัน Shopee บนโทรศัพท์มือถือ 2) เข้าไปที่เมนูการค้นหาสินค้าและกดที่รูปกล้องถ่ายภาพ 3) เลือกรูปภาพสินค้าที่ต้องการทดสอบ 4) แอปพลิเคชันจะแสดงหมวดหมู่ของสินค้าและแนะนำสินค้าที่เกี่ยวข้องหรือเหมือนกับภาพสินค้าที่เราใช้ในการค้นหาดังตัวอย่างในรูปที่ 12



รูปที่ 5 ขั้นตอนในการทดสอบค้นหาสินค้าด้วยภาพบนแอปพลิเคชัน Shopee

4. ผลการทดลอง

4.1 ผลการทดลองและประสิทธิภาพของตัวแบบ

จากตารางที่ 2 ในภาพรวมตัวแบบ EfficientNet ทั้ง 3 ตัวแบบให้ค่า Accuracy ในการทดสอบกับชุดข้อมูลทดสอบและชุดข้อมูลทดสอบเพิ่มเติมดีกว่าตัวแบบ VGG19, ResNet50 และ ResNet152 อย่างเห็นได้ชัด โดยตัวแบบ VGG19, ResNet50 และ ResNet152 มีค่า Accuracy ในการทดสอบกับชุดข้อมูลทดสอบ อยู่ที่ 69% ถึง 77% และค่า Accuracy ในการทดสอบกับชุดข้อมูลทดสอบเพิ่มเติมอยู่ที่ 76% ถึง 82% ในขณะที่ตัวแบบ EfficientNet ทั้ง 3 ตัวแบบมีค่า Accuracy ในการทดสอบกับชุดข้อมูลทดสอบอยู่ที่ 80% ถึง 84% และค่า Accuracy ในการทดสอบกับชุดข้อมูลทดสอบเพิ่มเติมอยู่ที่ 89% ถึง 92% ซึ่งตัวแบบ EfficientNetB5 (ที่ใช้ Loss function ทั้ง 2 แบบ) เป็นตัวแบบที่มีประสิทธิภาพในการจำแนกดีที่สุดของงานวิจัยนี้ โดยมีค่า Accuracy 84% เมื่อทดสอบกับชุดข้อมูลทดสอบและ Accuracy 92% เมื่อทดสอบกับชุดข้อมูลทดสอบเพิ่มเติม โดยสามารถตรวจสอบค่า Accuracy และ Loss ระหว่างการฝึกสอนตัวแบบตามรูปที่ 6

ผู้วิจัยพบว่าการใช้ Focal loss ช่วยให้ค่า Accuracy ของตัวแบบ VGG19, ResNet50 และ ResNet152 ดีขึ้นเล็กน้อยเมื่อเทียบกับการใช้ Categorical Cross entropy แต่สำหรับตัวแบบ EfficientNet ทั้ง 3 ตัวแบบ Focal loss กลับไม่ช่วยทำให้ค่า Accuracy ของตัวแบบดีขึ้นเลย โดยผู้วิจัยคาดว่าเป็นเพราะตัวแบบ EfficientNet ทั้ง 3 ตัวแบบค่อนข้างมีประสิทธิภาพที่ดีอยู่แล้ว โดยในเรื่องประสิทธิภาพด้านเวลาที่ตัวแบบใช้ในการจำแนกหมวดหมู่ต่อ 1 ภาพ ตัวแบบ ResNet50 จะใช้เวลาน้อยที่สุดคือ 0.044 วินาทีต่อ 1 ภาพ และเรื่องขนาดของตัวแบบ (MB) ตัวแบบ EfficientNetB0 มีขนาดเล็กที่สุดคือ 38.7 MB

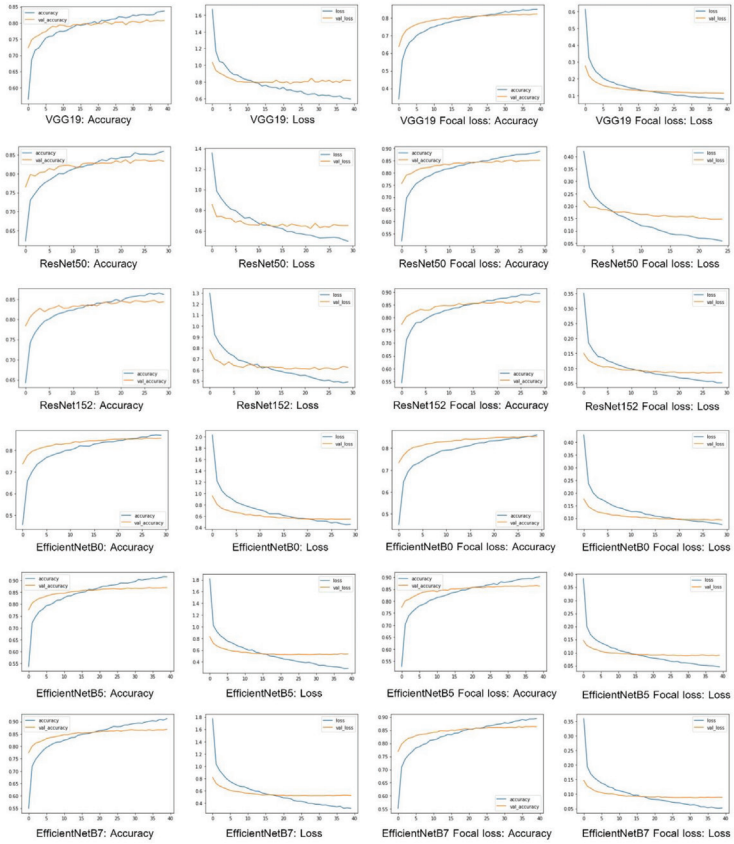
เพื่อให้เห็นความเปรียบต่างระหว่างแต่ละแบบจำลองที่ทำการทดลองและสามารถทำการคัดเลือกแบบจำลองที่เหมาะสมได้ดีขึ้น ผู้วิจัยนำข้อมูลจากตารางที่ 2 มาสร้างเป็นแผนภูมิในรูปที่ 7 และ 8 ที่ประกอบไปด้วยข้อมูล 3 แกน คือ (1) แกน X คือค่า Accuracy ความแม่นยำของแต่ละตัวแบบบนชุดข้อมูลทดสอบ (2) แกน Y คือเวลาที่แต่ละตัวแบบใช้ในการจำแนกรูปภาพ 1 รูป (หน่วยเป็นวินาที) และ (3) ขนาดของตัวแบบแทนด้วยขนาดของวงกลม

ตารางที่ 2 ประสิทธิภาพ เวลาที่ใช้ในการจำแนก และขนาดของตัวแบบที่ทำการทดลอง ทั้งนี้ค่าที่ถูกขีดเส้นใต้หมายถึงค่าที่ดีที่สุดในคอลัมน์นั้น ๆ

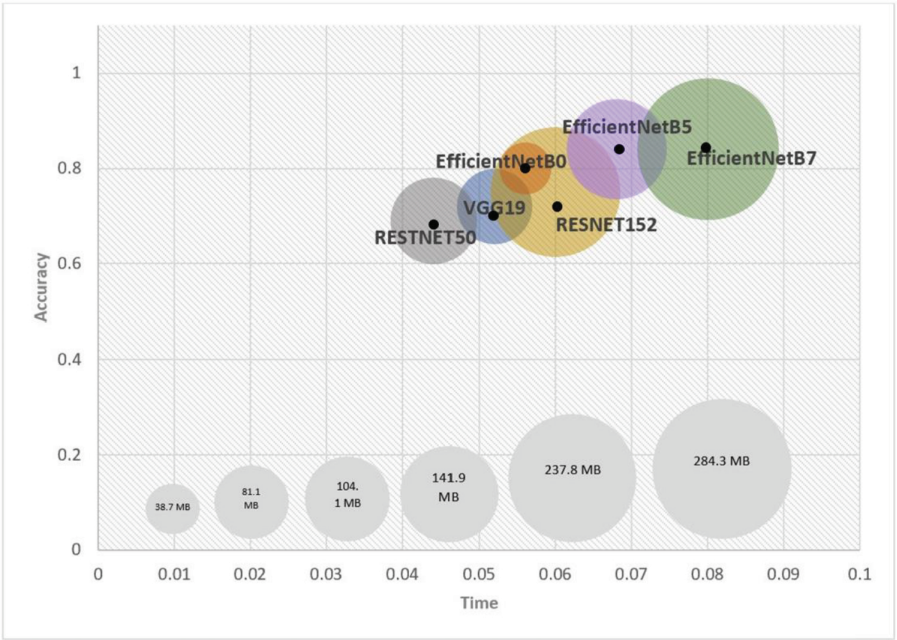
ตัวแบบ	ชุดข้อมูลประเมิน (validation set)		Accuracy ↑ ชุดข้อมูลทดสอบ (Test set)	Accuracy ↑ ชุดข้อมูลทดสอบเพิ่มเติม Additional Test set)	เวลาเฉลี่ยต่อการจำแนก 1 ภาพ ↓ (วินาที)	ขนาดตัวแบบ ↓ (MB)
	Accuracy ↑	Loss ↓				
VGG19	80%	0.81	72%	80%	0.052	81.1
VGG19 (FL)	82%	0.11	75%	82%	0.052	81.1
ResNet50	83%	0.65	69%	76%	0.044	104.1
ResNet50 (FL)	85%	0.09	72%	77%	0.044	104.1
ResNet152	84%	0.60	75%	81%	0.06	237.8
ResNet152 (FL)	85%	0.08	77%	81%	0.06	237.8
EfficientNetB0	85%	0.54	80%	90%	0.056	38.7
EfficientNetB0 (FL)	85%	0.09	80%	89%	0.056	38.7
EfficientNetB5	86%	0.52	84%	92%	0.068	141.9
EfficientNetB5 (FL)	86%	0.09	84%	92%	0.068	141.9
EfficientNetB7	86%	0.52	84%	91%	0.08	284.3
EfficientNetB7 (FL)	86%	0.08	83%	90%	0.08	284.3

(FL) = Focal loss

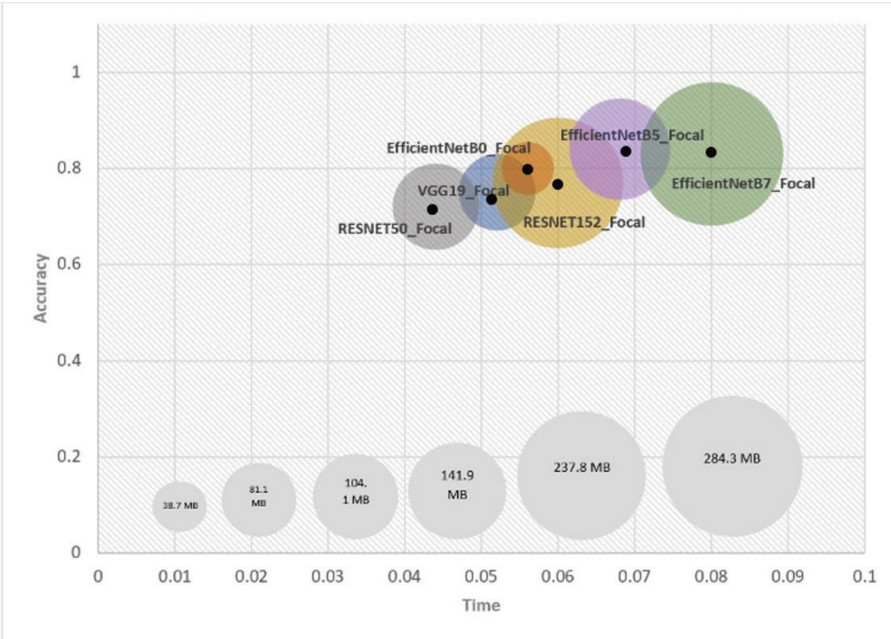
เวลาต่อการจำแนก 1 ภาพ หน่วยเป็นวินาที โดยใช้การทดสอบกับชุดข้อมูลทดสอบ (Test set) 14,862 รูป ขนาด 300x300 Pixel ใช้ GPU Tesla V100-SXM2-16GB



รูปที่ 6 ค่า Accuracy และ Loss ในการฝึกสอนของแต่ละตัวแบบ



รูปที่ 7 แผนภูมิเปรียบเทียบตัวแบบในการทดลองที่ใช้ Loss function แบบ Categorical Cross Entropy โดยเปรียบเทียบระหว่างประสิทธิภาพความแม่นยำบนชุดข้อมูลทดสอบ (แกนตั้ง) เวลาที่ใช้ในการจำแนกภาพหน่วยวินาที (แกนนอน) และขนาดของตัวแบบ (ขนาดของวงกลม)



รูปที่ 8 แผนภูมิเปรียบเทียบตัวแบบในการทดลองที่ใช้ Focal loss โดยเปรียบเทียบระหว่างประสิทธิภาพความแม่นยำบนชุดข้อมูลทดสอบ (แกนตั้ง) เวลาที่ใช้ในการจำแนกภาพหน่วยวินาที (แกนนอน) และขนาดของตัวแบบ (ขนาดของวงกลม)

จากรูปที่ 7 และ 8 ตัวแบบ EfficientNetB5 ทั้งแบบที่ใช้ Categorical Cross Entropy และ Focal loss ค่อนข้างเหมาะสมที่จะนำไปใช้งานจริงเมื่อพิจารณาจากตัววัดผล 3 ค่าประกอบกัน ซึ่งตัวแบบ EfficientNetB5 มีค่า Accuracy ที่ทดสอบกับชุดข้อมูลทดสอบเท่ากับ 84% และ ที่ทดสอบกับชุดข้อมูลทดสอบเพิ่มเติมเท่ากับ 92% ซึ่งเป็นค่าที่มากที่สุดของทุกตัวแบบในงานวิจัยนี้ ในด้านของขนาดตัวแบบ ตัวแบบ EfficientNetB5 มีขนาดที่เล็กกว่าตัวแบบ EfficientNetB7 และ ResNet152 และในเรื่องของเวลาในการจำแนก ตัวแบบ EfficientNetB5 ใช้เวลาในการจำแนก 0.068 วินาทีต่อรูป ซึ่งน้อยกว่าตัวแบบ EfficientNetB7 ที่ใช้เวลา 0.08 วินาทีต่อรูป อย่างไรก็ตามถ้าหากให้ความสำคัญในเรื่องของขนาดตัวแบบเป็นหลัก อาจจะเลือกพิจารณาตัวแบบ EfficientNetB0 เนื่องจากตัวแบบนี้มีขนาดที่เล็กที่สุด (38.7 MB) แต่จะได้ค่า Accuracy เพียง 80% เมื่อทดสอบกับชุดข้อมูลทดสอบและ 90% เมื่อทดสอบกับชุดข้อมูลทดสอบเพิ่มเติม ซึ่งก็ยังเป็นค่าที่ดีกว่าตัวแบบ VGG19, ResNet50 และ ResNet152 นอกจากนี้ในเรื่องของเวลา ตัวแบบ EfficientNetB0 ก็ใช้เวลาในการจำแนกรูปน้อยกว่าตัวแบบ EfficientNetB5 และ EfficientNetB7

นอกจากการพิจารณาค่า Accuracy โดยรวมของตัวแบบแล้ว ผู้วิจัยยังทำการพิจารณาค่า Precision, Recall และ F1-Score ประกอบเพิ่มเติมแสดงในตารางที่ 3 ซึ่งค่าในตารางจะเป็นค่า Precision, Recall และ F1-Score โดยเฉลี่ยของทุกหมวดหมู่ในแต่ละตัวแบบ นอกจากนี้ยังมีผลการทดลองในตารางที่ 4 ซึ่งแสดงหมวดหมู่ที่แต่ละตัวแบบจำแนกได้ไม่ค่อยแม่นยำหรือมีค่า Recall ในหมวดหมู่นั้นๆ น้อยกว่า 60% จากตารางที่ 4 จะเห็นว่า ตัวแบบในกลุ่ม EfficientNet ให้ผลที่ดีกว่าตัวแบบ VGG19, ResNet50 และ ResNet152 อย่างเห็นได้ชัด เนื่องจากตัวแบบ VGG19, ResNet50 และ ResNet152 มีหมวดหมู่ที่จำแนกได้ไม่แม่นยำหรือมีค่า Recall น้อยกว่า 60% อยู่ที่ 7 ถึง 9 หมวดหมู่ จากจำนวนหมวดหมู่ทั้งหมด 38 หมวดหมู่ ในขณะที่ตัวแบบในกลุ่ม EfficientNet มีหมวดหมู่ที่จำแนกได้ไม่แม่นยำหรือค่า Recall น้อยกว่า 60% อยู่เพียง 0 ถึง 3 หมวดหมู่เท่านั้น โดยเฉพาะตัวแบบ EfficientNetB5 ที่ทดสอบกับชุดข้อมูลทดสอบเพิ่มเติมแล้วไม่มีหมวดหมู่ใดที่ค่า Recall น้อยว่า 60% เลย

อ้างอิงจากผลการทดลองในตารางที่ 4 ตัวอย่างหมวดหมู่ที่ตัวแบบมักจะไม่สามารถจำแนกได้แม่นยำ เช่น

- หมวดหมู่ Alcohol gel เนื่องจากมีลักษณะสินค้าใกล้เคียงกับหมวดหมู่ Skincare และภาพสินค้าในหมวดหมู่ Alcohol gel มีความหลากหลายสูงทำให้ผลการจำแนกออกมาไม่ค่อยดี แม้แต่ตัวแบบ EfficientNetB5 และ EfficientNetB7 ที่มีประสิทธิภาพสูงก็ยังคงมีปัญหาในการจำแนกภาพสินค้าในหมวดหมู่นี้
- หมวดหมู่ Mobile และ Mobile case ทั้ง 2 หมวดหมู่มีลักษณะรูปภาพสินค้าที่ใกล้เคียงกันมาก ทำให้ตัวแบบจำแนก 2 หมวดหมู่นี้ออกจากกันได้ยาก โดนส่วนมากตัวแบบมักจะจำแนกหมวดหมู่ Mobile case ว่าเป็นหมวดหมู่ Mobile ทำให้ตัวแบบส่วนใหญ่จำแนกหมวดหมู่ Mobile case ได้ไม่ค่อยแม่นยำ แม้แต่ตัวแบบ EfficientNetB5 และ EfficientNetB7 ก็ไม่สามารถจำแนกหมวดหมู่นี้ได้แม่นยำเช่นเดียวกับกรณีของหมวดหมู่ Alcohol gel
- หมวดหมู่ Jeans ตัวแบบส่วนใหญ่มักจะจำแนกหมวดหมู่ Jeans (กางเกงยีนส์) ว่าเป็นหมวดหมู่ Trousers (กางเกงขายาว) เนื่องจากทั้งสองหมวดหมู่มีความคล้ายกันมาก โดยเฉพาะลักษณะรูปทรงที่เป็นทรงกางเกงเหมือนกันแตกต่างเพียงลวดลายและเนื้อผ้าของกางเกง ทำให้ตัวแบบส่วนมากจะจำแนกได้ไม่ค่อยแม่นยำ แต่ปัญหาดังกล่าวจะไม่พบในตัวแบบ EfficientNetB5 และ EfficientNetB7 ซึ่งตัวแบบทั้ง 2 ตัวสามารถจัดการกับปัญหาหรือข้อจำกัดนี้ได้
- หมวดหมู่ Coffee เป็นหมวดหมู่ของสินค้าประเภทเมล็ดกาแฟคั่วบดซึ่งมีบรรจุภัณฑ์ที่หลากหลาย ลักษณะของรูปภาพค่อนข้างมีความหลากหลายทำให้ผลการจำแนกออกมาไม่ค่อยดี แต่ปัญหาดังกล่าวจะไม่พบในตัวแบบ EfficientNetB5 และ EfficientNetB7

ตารางที่ 3 ค่าเฉลี่ย Precision, Recall และ F1-Score ของแต่ละตัวแบบ ทั้งนี้ค่าที่ขีดเส้นใต้หมายถึงค่าที่ดี ที่สุดในคอลัมน์นั้น ๆ (ค่ายิ่งสูงยิ่งดี)

ตัวแบบ	ชุดข้อมูลทดสอบ (Test set)			Additional test set)		
	ค่าเฉลี่ย Precision	ค่าเฉลี่ย Recall	ค่าเฉลี่ย F1-Score	ค่าเฉลี่ย Precision	ค่าเฉลี่ย Recall	ค่าเฉลี่ย F1-Score
VGG19	74%	72%	71%	82%	80%	80%
VGG19 (FL)	78%	75%	74%	85%	82%	82%
ResNet50	73%	69%	69%	79%	76%	75%
ResNet50 (FL)	75%	72%	71%	81%	77%	77%
ResNet152	76%	75%	74%	83%	81%	81%
ResNet152 (FL)	78%	78%	76%	83%	81%	81%
EfficientNetB0	83%	80%	80%	91%	90%	90%
EfficientNetB0 (FL)	82%	80%	80%	90%	89%	89%
EfficientNetB5	84%	84%	84%	93%	92%	92%
EfficientNetB5 (FL)	84%	84%	83%	92%	91%	91%
EfficientNetB7	84%	84%	84%	92%	91%	91%
EfficientNetB7 (FL)	84%	84%	83%	92%	91%	91%

ตารางที่ 4 หมวดหมู่ที่มีค่า Recall น้อยกว่า 60% ของในแต่ละตัวแบบ

ตัวแบบ	ชุดข้อมูลทดสอบ		ชุดข้อมูลทดสอบเพิ่มเติม	
	Categorical crossentropy	Focal loss	Categorical crossentropy	Focal loss
VGG19	Coffee Glove Jacket Mobile case Skincare Snack Sports equipment Jeans	Coffee Handbag Mobile case Ring Snack Jeans	Alcohol gel Jacket Mobile case Skincare Shorts	Alcohol gel Jeans Shorts
ResNet50	Alcohol gel Covid protection equipment Belt, Jacket Mobile case Skincare, Snack Sports equipment Jeans	Covid protection equipment Belt, Handbag Jacket Mobile case Skincare Snack Jeans	Alcohol gel Jeans Shorts	Alcohol gel Covid protection equipment Handbag Jacket Mobile, Mobile case Skincare, Snack Sports equipment Jeans, Short
ResNet152	Covid protection equipment Handbag Mobile case Skincare Snack Jeans	Covid protection equipment Handbag Mobile case Skincare Snack Jeans	Alcohol gel Covid protection equipment Mobile case Skincare Jeans Shorts	Alcohol gel Covid protection equipment Handbag Mobile Mobile case Skincare Shorts
EfficientNetB0	Coffee Mobile case Jeans	Coffee Mobile case Jeans	Coffee Mobile case Jeans	Mobile case Jeans
EfficientNetB5	-	Coffee Mobile case	Alcohol gel	Mobile case
EfficientNetB7	Mobile	Alcohol gel	Alcohol gel	Mobile case

4.2 ผลการทดลองเปรียบเทียบประสิทธิภาพระหว่างตัวแบบกับแอปพลิเคชัน Shopee

ผู้วิจัยนำตัวแบบ EfficientNetB5 Focal loss ซึ่งเป็นตัวแบบที่มีประสิทธิภาพในการจำแนกดีที่สุดของงานวิจัยนี้ (พิจารณาจากค่า Accuracy ตามที่ได้สรุปในหัวข้อที่ 4.1) และมีจำนวนหมวดหมู่ที่ตัวแบบจำแนกไม่แม่นยำเพียง 1 หมวดหมู่เท่านั้น (ตามตารางที่ 6) มาทำการทดสอบต่อยอด โดยจะเปรียบเทียบประสิทธิภาพความแม่นยำกับแอปพลิเคชัน Shopee ซึ่งเป็นแอปพลิเคชันที่มีการใช้งานจริงในปัจจุบัน ทั้งนี้การทดสอบจะแบ่งข้อมูลออกเป็น Cut-off point 2 แบบ คือ 1) แบ่งข้อมูลออกเป็น Cut-off point ที่ 12 รูป, 24 รูป, 37 รูป และ 50 รูป ในแต่ละหมวดหมู่ 2) แบ่งข้อมูลออกเป็น Cut-off point ที่ 25% ของรูปทั้งหมด, 50% ของรูปทั้งหมด, 75% ของรูปทั้งหมด และ 100% ของรูปทั้งหมด ในแต่ละหมวดหมู่ ซึ่งได้ผลการทดสอบตามตารางที่ 5 และ 6

ตารางที่ 5 ผลการทดสอบค่า Recall ของการจำแนกระหว่างตัวแบบ EfficientNetB5 Focal loss และ แอปพลิเคชัน Shopee โดยแบ่งข้อมูลออกเป็น Cut-off point ที่ 12 รูป, 24 รูป, 37 รูป และ 50 รูป ในแต่ละหมวดหมู่

หมวดหมู่	Cut-off point 12		Cut-off point 24		Cut-off point 37		Cut-off point 50	
	ตัวแบบ	Shopee	ตัวแบบ	Shopee	ตัวแบบ	Shopee	ตัวแบบ	Shopee
Air conditioner	100.00%	83.33%	100.00%	87.50%	97.30%	81.08%	98.00%	80.00%
Alcohol gel	83.33%	66.67%	75.00%	75.00%	75.68%	72.97%	76.00%	74.00%
Covid protection equipment	91.67%	66.67%	91.67%	75.00%	91.89%	75.68%	94.00%	80.00%
Baby bottle	100.00%	100.00%	100.00%	95.83%	100.00%	91.89%	100.00%	94.00%
Bathroom accessories	91.67%	66.67%	95.83%	75.00%	94.59%	78.38%	98.00%	82.00%
Belt	100.00%	83.33%	100.00%	79.17%	97.30%	70.27%	98.00%	72.00%
Bag	100.00%	91.67%	100.00%	95.83%	100.00%	91.89%	100.00%	94.00%
Canvas shoes	91.67%	100.00%	87.50%	100.00%	86.49%	97.30%	90.00%	94.00%
Chair	100.00%	100.00%	100.00%	100.00%	100.00%	97.30%	100.00%	96.00%
Coffee	75.00%	83.33%	75.00%	62.50%	67.57%	59.46%	74.00%	60.00%
Dishware	100.00%	91.67%	100.00%	79.17%	100.00%	72.97%	98.00%	74.00%
Earphones	91.67%	91.67%	95.83%	91.67%	91.89%	94.59%	94.00%	96.00%
Earring	100.00%	66.67%	100.00%	79.17%	100.00%	78.38%	98.00%	78.00%
Flash drive	91.67%	91.67%	95.83%	87.50%	91.89%	81.08%	90.00%	82.00%
Glove	83.33%	91.67%	87.50%	95.83%	91.89%	91.89%	86.00%	92.00%
Handbag	91.67%	100.00%	91.67%	95.83%	94.59%	97.30%	94.00%	98.00%
Hat	83.33%	100.00%	91.67%	100.00%	94.59%	100.00%	94.00%	100.00%
Helmet	100.00%	100.00%	100.00%	95.83%	100.00%	91.89%	100.00%	92.00%
Home accessories	100.00%	75.00%	95.83%	66.67%	97.30%	56.76%	98.00%	52.00%
Home clock	100.00%	100.00%	100.00%	95.83%	100.00%	97.30%	100.00%	94.00%
Islamic clothing	100.00%	75.00%	100.00%	66.67%	94.59%	75.68%	94.00%	76.00%
Jacket	100.00%	100.00%	100.00%	100.00%	94.59%	100.00%	96.00%	100.00%
Mobile	75.00%	66.67%	83.33%	58.33%	86.49%	56.76%	84.00%	52.00%
Mobile case	25.00%	100.00%	33.33%	95.83%	32.43%	97.30%	32.00%	98.00%
Notebook	100.00%	100.00%	95.83%	95.83%	94.59%	97.30%	96.00%	98.00%
Perfume	58.33%	66.67%	70.83%	79.17%	72.97%	81.08%	78.00%	80.00%
Rice cooker	100.00%	91.67%	100.00%	95.83%	97.30%	97.30%	98.00%	98.00%
Ring	91.67%	83.33%	87.50%	87.50%	86.49%	91.89%	88.00%	94.00%
Skincare	91.67%	91.67%	79.17%	87.50%	70.27%	81.08%	72.00%	80.00%
Snack	66.67%	58.33%	62.50%	62.50%	64.86%	70.27%	68.00%	70.00%
Sports equipment	100.00%	41.67%	100.00%	62.50%	97.30%	62.16%	94.00%	64.00%
T-shirt	91.67%	100.00%	95.83%	100.00%	97.30%	100.00%	98.00%	100.00%
Toy car	100.00%	100.00%	100.00%	100.00%	100.00%	97.30%	100.00%	98.00%
Watch	100.00%	91.67%	100.00	95.83%	97.30%	97.30%	98.00%	98.00%
Women shoes	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
Jeans	100.00%	91.67%	100.00%	95.83%	100.00%	97.30%	100.00%	98.00%
Trousers	100.00%	91.67%	100.00%	95.83%	100.00%	97.30%	100.00%	98.00%
Shorts	100.00%	100.00%	100.00%	100.00%	100.00%	97.30%	100.00%	96.00%

ตารางที่ 6 ผลการทดสอบค่า Recall ของการจำแนกระหว่างตัวแบบ EfficientNetB5 Focal loss และ แอปพลิเคชัน Shopee โดยการทดสอบแบบแบ่งข้อมูลออกเป็น Cut-off point ที่ 25% ของรูปทั้งหมด, 50% ของรูปทั้งหมด, 75% ของรูปทั้งหมด และ100% ของรูปทั้งหมด ในแต่ละหมวดหมู่

หมวดหมู่	Cut-off point 25%		Cut-off point 50%		Cut-off point 75%		Cut-off point 100%	
	ตัวแบบ	Shopee	ตัวแบบ	Shopee	Model	Shopee	ตัวแบบ	Shopee
Air conditioner	100.00%	83.33%	95.83%	83.33%	97.30%	81.08%	98.00%	80.00%
Alcohol gel	75.00%	91.67%	75.00%	75.00%	72.97%	72.97%	76.00%	74.00%
Covid protection equipment	100.00%	83.33%	100.00%	79.17%	91.89%	75.68%	94.00%	80.00%
Baby bottle	100.00%	91.67%	100.00%	87.50%	100.00%	91.89%	100.00%	94.00%
Bathroom accessories	100.00%	91.67%	100.00%	87.50%	100.00%	81.08%	98.00%	82.00%
Belt	91.67%	66.67%	100.00%	66.67%	97.30%	70.27%	98.00%	72.00%
Bag	100.00%	91.67%	100.00%	87.50%	100.00%	94.59%	100.00%	94.00%
Canvas shoes	91.67%	91.67%	95.83%	91.67%	89.19%	94.59%	90.00%	94.00%
Chair	100.00%	100.00%	100.00%	100.00%	100.00%	97.30%	100.00%	96.00%
Coffee	75.00%	58.33%	79.17%	50.00%	70.27%	48.65%	74.00%	60.00%
Dishware	91.67%	75.00%	100.00%	83.33%	70.27%	75.68%	98.00%	74.00%
Earphones	83.33%	100.00%	91.67%	95.83%	94.59%	97.30%	94.00%	96.00%
Earring	100.00%	83.33%	95.83%	83.33%	97.30%	81.08%	98.00%	78.00%
Flash drive	83.33%	83.33%	91.67%	75.00%	94.59%	78.38%	90.00%	82.00%
Glove	83.33%	91.67%	79.17%	87.50%	91.89%	97.30%	86.00%	92.00%
Handbag	100.00%	91.67%	95.83%	95.83%	94.59%	97.30%	94.00%	98.00%
Hat	83.33%	100.00%	91.67%	100.00%	91.89%	100.00%	94.00%	100.00%
Helmet	100.00%	91.67%	100.00%	95.83%	100.00%	89.19%	100.00%	92.00%
Home accessories	91.67%	75.00%	100.00%	54.17%	97.30%	45.95%	98.00%	52.00%
Home clock	100.00%	100.00%	100.00%	91.67%	100.00%	94.59%	100.00%	94.00%
Islamic clothing	100.00%	58.33%	95.83%	70.83%	94.59%	78.38%	94.00%	76.00%
Jacket	100.00%	100.00%	100.00%	100.00%	97.30%	100.00%	96.00%	100.00%
Mobile	91.67%	50.00%	75.00%	37.50%	81.08%	48.65%	84.00%	52.00%
Mobile case	33.33%	100.00%	25.00%	95.83%	27.03%	97.30%	32.00%	98.00%
Notebook	100.00%	100.00%	95.83%	100.00%	94.59%	97.30%	96.00%	98.00%
Perfume	66.67%	75.00%	70.83%	70.83%	75.68%	81.08%	78.00%	80.00%
Rice cooker	95.83%	100.00%	95.83%	95.83%	97.30%	97.30%	98.00%	98.00%
Ring	66.67%	83.33%	95.83%	95.83%	89.19%	94.59%	88.00%	94.00%
Skincare	66.67%	100.00%	75.00%	79.17%	70.27%	83.78%	72.00%	80.00%
Snack	100.00%	75.00%	62.50%	66.67%	70.27%	67.57%	68.00%	70.00%
Sports equipment	100.00%	58.33%	91.67%	62.50%	94.59%	56.76%	94.00%	64.00%
T-shirt	100.00%	100.00%	100.00%	100.00%	97.30%	100.00%	98.00%	100.00%
Toy car	100.00%	100.00%	100.00%	95.83%	100.00%	97.30%	100.00%	98.00%
Watch	100.00%	100.00%	95.83%	95.83%	100.00%	97.30%	98.00%	98.00%
Women shoes	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
Jeans	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	98.00%
Trousers	100.00%	100.00%	100.00%	100.00%	100.00%	97.30%	100.00%	98.00%
Shorts	100.00%	91.67%	100.00%	95.83%	100.00%	97.30%	100.00%	96.00%

จากตารางที่ 5 และ 6 จะเห็นว่าค่าเฉลี่ย Recall ในแต่ Cut-off มีค่าใกล้เคียงกัน โดยการทดสอบกับตัวแบบจะมีค่าเฉลี่ย Recall ประมาณ 91% ในขณะที่แอปพลิเคชัน Shopee ค่าเฉลี่ย Recall ประมาณ 86% โดยไม่ว่าจะมีการเพิ่มหรือลดจำนวนข้อมูล ผลการทดสอบส่วนมากในแต่ละหมวดหมู่ตามตารางที่ 5 และ 6 จะได้ผลลัพธ์เหมือนเดิม ยกเว้นบางหมวดหมู่ เช่น Alcohol gel ในตารางที่ 5 ค่าของ Cut-off point แรกที่จำนวนรูป 12 รูป จะให้ผลลัพธ์ว่าแอปพลิเคชัน Shopee จำแนกได้แย่กว่าตัวแบบ แต่เมื่อเพิ่มปริมาณข้อมูลรูปภาพทดสอบเพิ่มเติมใน Cut-off point ที่ 24 รูป, 37 รูป และ 50 รูป จะให้ผลลัพธ์ว่าแอปพลิเคชัน Shopee และตัวแบบ ให้ผลในการจำแนกได้พอๆกัน เช่นเดียวกันกับในตารางที่ 6 Cut-off point แรกที่จำนวนรูปภาพ 25% ของจำนวนรูปทั้งหมดจะให้ผลลัพธ์ว่าแอปพลิเคชัน Shopee จำแนกได้ดีกว่าตัวแบบ แต่เมื่อเพิ่มปริมาณข้อมูลรูปภาพทดสอบเพิ่มเติมใน Cut-off point ที่ 50% ของจำนวนรูปทั้งหมด, 75% ของจำนวนรูปทั้งหมด และ 100% ของจำนวนรูปทั้งหมด จะให้ผลลัพธ์ว่าแอปพลิเคชัน Shopee และตัวแบบ ให้ผลในการจำแนกได้พอๆ กัน ทั้งนี้ผู้วิจัยสังเกตว่าการทดสอบประสิทธิภาพของตัวแบบและแอปพลิเคชัน Shopee จำนวนข้อมูลไม่ได้ส่ง

ผลกับผลการทดสอบของหมวดหมู่ส่วนใหญ่ แต่ก็มีบางหมวดหมู่ถ้าใช้จำนวนข้อมูลน้อยอาจส่งผลต่อความผิดพลาดในการสรุปผลได้ ดังนั้นเพื่อลดข้อผิดพลาดในการสรุปผลการทดสอบควรใช้จำนวนข้อมูลตั้งแต่ 24 รูปภาพ หรือ 50% ของจำนวนรูปภาพทั้งหมดขึ้นไปในการทดสอบ

ตารางที่ 7 เปรียบเทียบประสิทธิภาพในการจำแนกแต่ละหมวดหมู่ของตัวแบบ EfficientNetB5 Focal loss และแอปพลิเคชันShopee

ตัวแบบแม่นยำกว่า	ผลใกล้เคียงกัน	Shopee แม่นยำกว่า
Air conditioner	Alcohol gel, Bag	Mobile case
Covid protection equipment	Chair, Glove	
Bathroom accessories	Hat, Home clock	
Belt	Notebook, Rice cooker	
Coffee	Skincare, T-shirt	
Dishware	Watch, Jeans	
Earring	Shorts, Baby bottle	
Flash drive	Canvas shoes	
Home accessories	Earphones, Handbag	
Islamic clothing	Helmet, Jacket	
Mobile	Perfume, Ring	
Sports equipment	Snack, Toy car	
	Women shoes	
	Trousers	

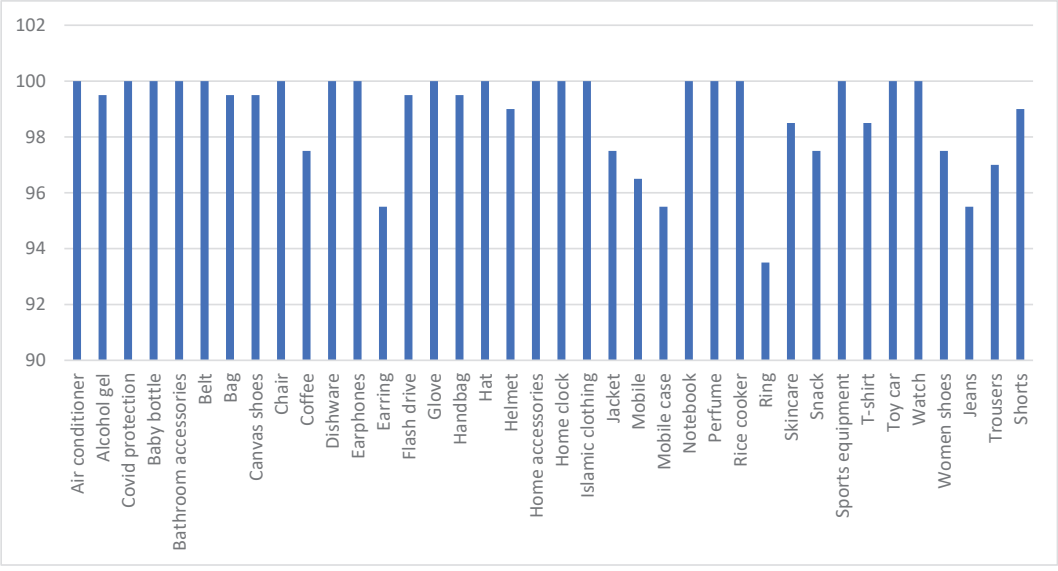
ตารางที่ 7 แสดงการเปรียบเทียบประสิทธิภาพความแม่นยำในการจำแนกในแต่ละหมวดหมู่ ระหว่างตัวแบบ EfficientNetB5 Focal loss และแอปพลิเคชัน Shopee ซึ่งจะเห็นว่าตัวแบบ EfficientNetB5 Focal loss มีความแม่นยำกว่าแอปพลิเคชัน Shopee อยู่ถึง 12 หมวดหมู่ (ค่า Recall ของตัวแบบมากกว่าแอปพลิเคชัน Shopee ตั้งแต่ 10% ขึ้นไป) มีประสิทธิภาพความแม่นยำในการจำแนกพอ ๆ กัน 25 หมวดหมู่ (ค่า Recall ของตัวแบบและแอปพลิเคชัน Shopee แตกต่างกันไม่เกิน 10%) และตัวแบบ EfficientNetB5 Focal loss มีความแม่นยำน้อยกว่าแอปพลิเคชัน Shopee เพียง 1 หมวดหมู่ (ค่า Recall ของตัวแบบน้อยกว่าแอปพลิเคชัน Shopee ตั้งแต่ 10% ขึ้นไป)

โดยหมวดหมู่ที่ตัวแบบ EfficientNetB5 Focal loss มีความแม่นยำในการจำแนกน้อยกว่าแอปพลิเคชัน Shopee คือ หมวดหมู่ Mobile case เนื่องจากลักษณะรูปภาพของหมวดหมู่ Mobile case มีความใกล้เคียงกับหมวดหมู่ Mobile มาก ทำให้ตัวแบบมักทำนายผิดว่ารูปที่เป็นหมวดหมู่ Mobile case เป็นหมวดหมู่ Mobile แต่ในขณะเดียวกันแอปพลิเคชัน Shopee ก็มักจะจำแนกผิดว่ารูปภาพหมวดหมู่ Mobile เป็นหมวดหมู่ Mobile case ทำให้ผลการทดลองออกมาว่าตัวแบบในงานวิจัยนี้จะจำแนกหมวดหมู่ Mobile ได้ดีกว่า แต่แอปพลิเคชัน Shopee จะจำแนกหมวดหมู่ Mobile case ได้ดีกว่า

สำหรับหมวดหมู่ Alcohol gel ที่เป็นหมวดหมู่เกี่ยวกับผลิตภัณฑ์เจลแอลกอฮอล์และน้ำยาฆ่าเชื้อโรค ซึ่งสินค้ามีบรรจุภัณฑ์ที่หลากหลายและบางรูปภาพค่อนข้างคล้ายกับสินค้าในหมวดหมู่ Skincare และ Perfume (ผลิตภัณฑ์เสริมความงามและน้ำหอม) ทำให้การจำแนกทำได้ค่อนข้างยากนั้น จากตารางที่ 6 และ 7 จะเห็นได้ว่ามีหลายตัวแบบที่จำแนกหมวดหมู่นี้ได้ไม่ค่อยแม่นยำ แต่เมื่อทดลองนำไปเทียบกับการทดลองกับแอปพลิเคชัน Shopee ตามตารางที่ 8 ก็พบ

ว่าตัวแบบและแอปพลิเคชัน Shopee ให้ผลในการจำแนกที่ใกล้เคียงกัน โดยตัวแบบมีค่า Recall เท่ากับ 76% ในขณะที่แอปพลิเคชัน Shopee ก็มีค่า Recall ที่พอ ๆ กันที่ 74% (การทดลองที่ Cut-off ที่ 50 รูป)

โดยในงานวิจัยนี้นอกจากจะทดลองกับตัวแบบและแอปพลิเคชัน Shopee แล้ว ผู้วิจัยยังทดลองให้คนจำแนกสินค้าจากภาพถ่ายอีกด้วย โดยใช้ชุดข้อมูลทดสอบเพิ่มเติมเช่นเดียวกับการทดลองของตัวแบบและแอปพลิเคชัน Shopee เพื่อทดสอบและเปรียบเทียบประสิทธิภาพในกรณีที่ใช้มนุษย์ในการจำแนก การทดสอบจะใช้ผู้ทดสอบจำนวน 4 คน ซึ่งเป็นผู้ที่ไม่มีปัญหาทางด้านการมองเห็นและไม่มีปัญหาในการใช้งานคอมพิวเตอร์ โดยมีขอบเขตในการทำการทดสอบดังนี้ (1) ใช้รูปภาพสินค้าจากชุดข้อมูลทดสอบเพิ่มเติมจำนวน 1,900 รูป 38 หมวดหมู่ (2) ทำการทดสอบบนคอมพิวเตอร์ โดยให้ผู้ทดสอบทำการย้ายรูปภาพสินค้าไปยัง Folder ของแต่ละหมวดหมู่ให้ถูกต้องและเร็วที่สุด (3) ใช้ผลที่ได้จากผู้ทดสอบทั้ง 4 คน ที่ประกอบไปด้วยค่าความถูกต้องในการจำแนกแต่ละหมวดหมู่และเวลาที่ใช้ในการจำแนก มาหาค่าเฉลี่ยตามรูปที่ 9 และสรุปผล



รูปที่ 9 ผลการทดลองจำแนกหมวดหมู่ของสินค้าจากรูปภาพด้วยมนุษย์ โดยแสดงค่าเฉลี่ย Recall ของแต่ละหมวดหมู่ (แกนตั้ง)

จากรูปที่ 9 การทดลองจำแนกหมวดหมู่ของสินค้าจากรูปภาพด้วยมนุษย์ พบว่าค่าเฉลี่ยของความถูกต้องในการจำแนกหรือค่าเฉลี่ย Recall ในแต่ละหมวดหมู่เท่ากับ 98.63% ซึ่งแม่นยำกว่าตัวแบบที่มีค่าเฉลี่ย 91.47% และแอปพลิเคชัน Shopee ที่มีค่าเฉลี่ย 86.37% แต่เมื่อเปรียบเทียบด้านเวลาที่ใช้ในการจำแนกพบว่าการทดลองจำแนกด้วยมนุษย์ใช้เวลาในการจำแนกเฉลี่ยประมาณ 55 นาที หรือคิดเป็น 1.74 วินาทีต่อรูป ในขณะที่ตัวแบบ EfficientNetB5 ซึ่งเป็นตัวแบบที่ดีที่สุดในงานวิจัยนี้ใช้เวลาเพียง 0.068 วินาทีต่อรูป ซึ่งตัวแบบ EfficientNetB5 ใช้เวลาน้อยกว่าถึงประมาณ 25.5 เท่า

5. สรุปผล

จากผลการทดลองทั้งหมดในงานวิจัยนี้พบว่าตัวแบบในกลุ่ม EfficientNet ให้ผลประสิทธิภาพในการจำแนกหมวดหมู่ของสินค้าดีที่สุด โดยเฉพาะตัวแบบ EfficientNetB5 ที่ใช้ Loss function ทั้ง 2 แบบ มีความเหมาะสมในการนำไปใช้งานจริงมากที่สุดเมื่อพิจารณาจากค่าประสิทธิภาพ Accuracy ขนาดของตัวแบบและเวลาที่ใช้ในการจำแนกประกอบกันตามรูปที่ 7 และ 8 โดยมีค่า Accuracy เมื่อทดสอบกับชุดข้อมูลทดสอบเท่ากับ 84% ค่า Accuracy เมื่อทดสอบกับชุดข้อมูลทดสอบเพิ่มเติมเท่ากับ 92% ใช้เวลาในการจำแนก 0.068 วินาทีต่อรูป

และขนาดของตัวแบบ 141.9 MB นอกจากนั้นแล้วเมื่อพิจารณาที่ค่า Recall ของแต่ละหมวดหมู่เมื่อทดสอบกับชุดข้อมูลทดสอบ พบว่าทุกหมวดหมู่มีค่า Recall มากกว่า 60% ซึ่งแสดงให้เห็นว่าตัวแบบ EfficientNetB5 ค่อนข้างจำแนกทุกหมวดหมู่ได้แม่นยำมากกว่าตัวแบบอื่นๆ และเมื่อทดลองเปรียบเทียบประสิทธิภาพในการจำแนกหมวดหมู่ระหว่างตัวแบบ EfficientNetB5 Focal loss และแอปพลิเคชัน Shopee พบว่าตัวแบบสามารถจำแนกได้ดีกว่ามากถึง 12 หมวดหมู่จากจำนวนทั้งหมด 38 หมวดหมู่ หรือคิดเป็น 31.5% โดยมีอีก 25 หมวดหมู่ที่มีประสิทธิภาพพอๆกันและมีเพียงแค่ 1 หมวดหมู่ที่ตัวแบบจำแนกได้แย่กว่า

อย่างไรก็ตามถึงแม้ว่าตัวแบบ EfficientNetB5 ที่ทดสอบกับชุดข้อมูลทดสอบจะมีค่า Recall มากกว่า 60% ในทุกหมวดหมู่ แต่เมื่อทดสอบกับชุดข้อมูลทดสอบเพิ่มเติมหรือเปลี่ยนไปใช้ Loss function แบบ Focal loss พบว่ายังมีบางหมวดหมู่ที่มีค่า Recall น้อยกว่า 60% เช่น Mobile, Mobile case, Alcohol gel และ Coffee ซึ่งเป็นหมวดหมู่ที่ตัวแบบส่วนมากมักจะทำนายได้ไม่ค่อยแม่นยำด้วยเช่นกัน เนื่องจากบางหมวดหมู่มีลักษณะใกล้เคียงกันมาก เช่น Mobile กับ Mobile case ทำให้ตัวแบบจำแนกระหว่าง 2 หมวดหมู่นี้ได้ยาก โดยการนำไปประยุกต์ใช้งานจริงอาจจะมีการเพิ่มขั้นตอนวิธีบางอย่างเข้าไปหรือเพิ่มตัวเลือกบางอย่างให้กับผู้ใช้งานเพื่อระบุความแน่ชัดของสินค้ามากขึ้นที่นอกเหนือจากรูปภาพปกติ เพื่อช่วยลดความสับสนของตัวแบบและสามารถจำแนกระหว่าง 2 หมวดหมู่ที่มีความใกล้เคียงกันได้แม่นยำขึ้น

6. เอกสารอ้างอิง

1. Electronic Transactions Development Agency, 2019, Value of e-Commerce: Survey in Thailand 2019 [Online], Available: http://blog.bru.ac.th/wp-content/uploads/bp-attachments/35694/เอกสารการแถลงผลการสำรวจมูลค่าพาณิชย์อิเล็กทรอนิกส์-ปี-2562.pdf. (In Thai) [1 August 2021]

2. Wunderman Thompson, 2021, "The Future Shopper Report 2021 [Online], Available: https://www.wundermanthompson.com/insight/the-future-shopper-report-2021. [1 August 2021]

3. Dan, C., 2014, "Consumer-To-Consumer (C2C) Electronic Commerce: The Recent Picture," *International Journal of Networks and Communications*, 4 (2), pp. 29-32.

4. Herdian, Kusuma, G.P. and Suharjito, 2019, "Classification of C2C e-Commerce Product Images Using Deep Learning Algorithm," *International Journal of Advanced Computer Science and Applications*, 10 (9), pp. 196-203.

5. Partalas, I. and Balikas, G., 2016, E-Commerce Product Classification: Our Participation At cDiscount 2015 Challenge [Online], Available: https://arxiv.org/abs/1606.02854. [1 August 2021]

6. Oyewole, S.A. and Olugbara, O.O., 2018, "Product Image Classification Using Eigen Colour Feature with Ensemble Machine Learning," *Egyptian Informatics Journal*, 19 (2), pp. 83-100.

7. Jha, B.K., Sivasankari, G.G. and Venugopal, K.R., 2021, "E-Commerce Product Image Classification Using Transfer Learning," *International Conference on Computing Methodologies and Communication (ICCMC)*, 8-10 April 2021, Erode, India, pp. 904-912. https://doi.org/ 10.1109/ICCMC51019.2021.9418371

8. Simonyan, K. and Zisserman, A., 2015, Very Deep Convolutional Networks for Large-Scale Image Recognition [Online], Available: https://arxiv.org/abs/1409.1556. [1 August 2021]

9. He, K., Zhang, X., Ren, S. and Sun, J., 2016, "Deep Residual Learning for Image Recognition," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 27-30 June 2016, Las Vegas, USA, pp. 770-778.

10.Tan, M. and Le, Q.V., 2019, EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks [Online], Available: https://arxiv.org/abs/1905.11946. [1 August 2021]

11. Li, F., Kant, S., Araki, S., Bangera, S. and Shukla, S.S., 2020, Neural Networks for Fashion Image Classification and Visual Search [Online], Available:

https://arxiv.org/abs/2005.08170 [1 August 2021]

12. Rajnoha, M., Burget, R. and Povoda, L., 2018, "Image Background Noise Impact on Convolutional Neural Network Training," *International Congress on Ultra-Modern Telecommunications and Control Systems and Workshops (ICUMT)*, 5-9 November 2018, Moscow, Russia, pp. 1-4.

13. Dodge, S. and Karam, L., 2016, "Understanding How *Image Quality Affects Deep Neural Networks," International Conference on Quality of Multimedia Experience (QoMEX)*, 6-8 June 2016, Lisbon, Portugal, pp. 1-6.

14. Lin, T.Y., Goyal, P., Girshick, R., He, K. and Dollár, P., 2017, "Focal Loss for Dense Object Detection," *IEEE International Conference on Computer Vision (ICCV)*, 22-29 October 2017, Venice, Italy, pp. 2999-3007.

15. Zhao, Y., Lin, F., Liu, S., Hu, Z., Li, H. and Bai, Y., 2019 "Constrained-Focal-Loss Based Deep Learning for Segmentation of Spores," *IEEE Access*, 7, pp. 165029-165038.

16. Bitirim, Y., Bitirim, S., Ertugrul, D.C. and Toygar, O., 2020, "An Evaluation of Reverse Image Search Performance of Google," *IEEE Annual Computers, Software, and Applications Conference (COMPSAC)*, 13-17 July 2020, Madrid, Spain, pp. 1368-1372.