

บทความวิจัย (Research Article)

การเปรียบเทียบประสิทธิภาพการคัดแยกขยะขวดน้ำด้วยการเรียนรู้เชิงลึก

Comparison of deep learning approach for water bottle waste Classification

ธีรภัทร์ จรเชื้อ¹ และ แสงดาว นพพิทักษ์^{1*}

Thiraphat Chorakhe¹ and Sangdaow Noppitak^{1*}

¹คณะวิทยาศาสตร์ มหาวิทยาลัยราชภัฏบุรีรัมย์

¹Faculty of Science, Buriram Rajabhat University

*Corresponding author email sangdaow.np@bru.ac.th

วันที่รับบทความ (Received)

15 มีนาคม 2566

วันที่ได้รับบทความฉบับแก้ไข (Revised)

10 พฤษภาคม 2566

วันที่ตอบรับบทความ (Accepted)

29 มิถุนายน 2566

บทคัดย่อ

ปัญหาขยะขวดน้ำเป็นปัญหาสำคัญส่งผลกระทบต่อระบบนิเวศ งานวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพการคัดแยกขยะขวดน้ำ โดยใช้การเรียนรู้เชิงลึก (Deep learning) ด้วยเทคนิคโครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolution neural network) จากชุดข้อมูลขยะขวดน้ำ (Water bottle waste dataset) จำนวน 6 คลาส จำนวนทั้งสิ้น 841 ภาพ ขั้นตอนแรกเป็นการเปรียบเทียบประสิทธิภาพเพื่อหาโมเดลที่ดีที่สุดจากโครงข่ายประสาทเทียมแบบคอนโวลูชัน 3 สถาปัตยกรรม ได้แก่ EfficientNetV2B0, EfficientNetV2B1 และ EfficientNetV2B3 พบว่า EfficientNetV2B1 มีค่าความถูกต้อง (Accuracy) สูงที่สุด 98.30% จากนั้นปรับปรุงประสิทธิภาพด้วยการเพิ่มเลเยอร์ในโครงข่ายที่มีการเชื่อมถึงกันหมด (Fully connected layer) กับสถาปัตยกรรมทั้ง 3 แบบ พบว่า ค่าความผิดพลาด (Loss) ของทั้ง 3 สถาปัตยกรรมลดลงเมื่อเปรียบเทียบกับวิธีที่ไม่เพิ่มเลเยอร์ ซึ่ง EfficientNetV2B3 ด้วยวิธีเพิ่มเลเยอร์มีค่าความถูกต้องมากที่สุด 98.59% และขั้นตอนสุดท้ายได้นำโมเดลที่ดีที่สุดคือ EfficientNetV2B3 ไปทดสอบกับข้อมูลทดสอบ (Testing set) ได้ค่าความถูกต้อง 97.16% และค่าความผิดพลาด 0.08

คำสำคัญ : ขยะขวดน้ำ, โครงข่ายประสาทเทียมแบบคอนโวลูชัน, การเรียนรู้เชิงลึก

Abstract

An used plastic water bottle is an important problem affect to the ecosystem. The aim of this research is to compare the efficiency in classification from datasets of the used plastic water bottle using Convolution Neural Network (CNN). The water bottle waste dataset consisted of 6 classes with a total of 841 images. First, a performance comparison was performed to find the best model from the three CNNs (EfficientNetV2B0, EfficientNetV2B1 and Efficient NetV2B3). EfficientNetV2B1 architecture has the highest accuracy of 98.30%. Then, it improves the three

architectures' performance using adding layers to the fully connected layer. The experimental results showed that the error of all three CNNs decreased compared to the method without adding layers. The highest accuracy achieves 98.59% from EfficientNetV2B3. Furthermore, EfficientNetV2B3 that the best model from the last experiment test on the testing dataset. The result of the highest accuracy was 97.16%, and the error was 0.08

Keywords: Water bottle waste, Convolution neural network, Deep learning

บทนำ

มลพิษของเสียเป็นปัญหาสิ่งแวดล้อมที่สำคัญที่สุดในโลกสมัยใหม่ [1] ซึ่งปัญหาสิ่งแวดล้อมหลัก ๆ เกิดจากขวดพลาสติกที่ส่งผลกระทบต่อสภาพแวดล้อม การนำขวดพลาสติกมารีไซเคิลนั้นเกิดขึ้นได้น้อย เนื่องจากการทิ้งขวดน้ำไม่เป็นที่และไม่มีมาตรการคัดแยกขวดพลาสติกในถังขยะ จึงทำให้ขวดน้ำพลาสติกเหล่านั้นไม่ถูกนำมารีไซเคิล ส่งผลกระทบต่อสภาพแวดล้อมเป็นอย่างมาก [2] หากมีเครื่องมือที่ช่วยให้การคัดแยกขยะขวดพลาสติกถูกต้องและสะดวก เช่น ตู้คัดแยกขยะขวดพลาสติกนอกจากจะช่วยลดปัญหาขยะแล้ว ยังสามารถสร้างรายได้ให้กับผู้คนจากการนำขวดพลาสติกมาคัดแยกที่ตู้คัดแยกขยะขวดน้ำ [3]

ถึงแม้ปัจจุบันการคัดแยกขยะทำโดยมนุษย์ แต่ด้านการพัฒนางานวิจัยมีนักวิจัยส่วนหนึ่งที่มุ่งเน้นพัฒนาวิธีการจำแนกขยะอัตโนมัติ โดยใช้แนวทางการเรียนรู้ของเครื่อง (Machine learning) เช่น การเปรียบเทียบแนวทางการเรียนรู้ของเครื่องสำหรับการจำแนกขยะขวดน้ำ โดยใช้เทคนิคโครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolution neural network) และซัพพอร์ตเวกเตอร์แมชชีน (Support vector machine) พัฒนาโมเดลคัดแยกขยะขวดน้ำจากชุดข้อมูลขยะขวดน้ำพลาสติก จัดเก็บจากการถ่ายภาพซึ่งมีขนาด $2,352 \times 4,160$ พิกเซล ภายในชุดข้อมูลแบ่งเป็น 10 คลาส จำนวน 1,100 ภาพ ทั้งสองเทคนิคมีข้อดีข้อเสียแตกต่างกัน แต่เทคนิคโครงข่ายประสาทเทียมแบบคอนโวลูชันมีความถูกต้องมากกว่าเทคนิคซัพพอร์ตเวกเตอร์แมชชีน โดยมีค่าความถูกต้องที่ 99% และ 94% ตามลำดับ [4] อีกทั้งงานวิจัยของ Chazhoor et al. (2022) ได้เปรียบเทียบประสิทธิภาพการจำแนกขยะจากโครงข่ายประสาทเทียมแบบคอนโวลูชัน 6 สถาปัตยกรรม ได้แก่ Resnet-50, ResNeXt, MobileNet_v2, DenseNet, ShuffleNet และ AlexNet ฝึกการเรียนรู้จากโมเดลที่ผ่านการสอนแล้ว (Pre-trained model) ด้วย ImageNet ทดลองกับชุดข้อมูล WaDaBa plastic ประกอบด้วย 5 คลาส จำนวนทั้งหมด 4,000 ภาพ จาก 6 สถาปัตยกรรม ResNeXt มีประสิทธิภาพการจำแนกสูงสุด [5] จากงานวิจัยดังกล่าวแสดงให้เห็นว่า การเรียนรู้เชิงลึก (Deep Learning) เป็นวิธีการที่นิยมและมีประสิทธิภาพในการจำแนกขยะ [6] และจากการเปรียบเทียบประสิทธิภาพโครงข่ายประสาทเทียมแบบคอนโวลูชันสำหรับงานด้านการจำแนก เป็นการเปรียบเทียบ 5 สถาปัตยกรรม ได้แก่ ResNet, DenseNet, EfficientNetV2, Vision Transformer, และ VGGNet ผลการจำแนกทุกสถาปัตยกรรมมีประสิทธิภาพใกล้เคียงกัน แต่ขนาดของโมเดลของสถาปัตยกรรม EfficientNetV2 มีขนาดเล็กที่สุด ซึ่งทำให้การเรียนรู้รวดเร็วกว่าสถาปัตยกรรมแบบอื่น ๆ [7]

ดังนั้นงานวิจัยนี้ มุ่งเน้นเปรียบเทียบประสิทธิภาพการจำแนกขยะขวดน้ำจากชุดข้อมูลขยะขวดน้ำ (Water bottle waste dataset) ด้วยสถาปัตยกรรม EfficientNetV2 [7] และปรับปรุงประสิทธิภาพด้วยวิธีเพิ่มเลเยอร์ใน

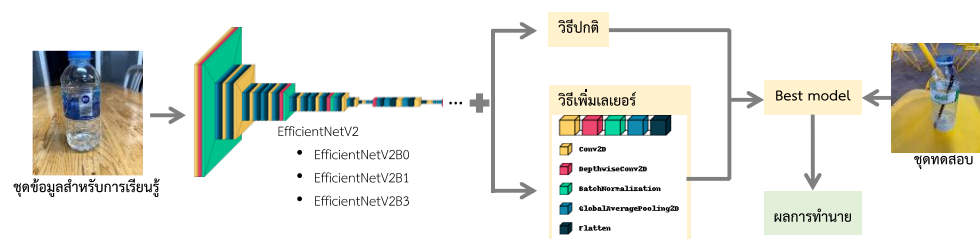
โครงข่ายที่มีการเชื่อมถึงกันหมด (Fully connected layer) เพื่อหาโมเดลที่มีประสิทธิภาพสูงสุดที่สามารถนำไปพัฒนาเป็นเครื่องคัดแยกขยะขวดน้ำต่อไป

วัตถุประสงค์ของการวิจัย

เพื่อเปรียบเทียบประสิทธิภาพการคัดแยกขยะขวดน้ำ โดยการเรียนรู้เชิงลึกด้วยเทคนิคโครงข่ายประสาทเทียมแบบคอนโวลูชัน

วิธีดำเนินการวิจัย

งานวิจัยนี้ เป็นการเปรียบเทียบประสิทธิภาพการจำแนกขยะขวดน้ำจากชุดข้อมูลขยะขวดน้ำ (Water bottle waste dataset) ที่สร้างขึ้นแบ่งเป็น 6 คลาส จำแนกด้วยสถาปัตยกรรม EfficientNetV2 ขั้นแรกเป็นการเปรียบเทียบประสิทธิภาพโครงข่ายประสาทเทียมแบบคอนโวลูชัน 3 สถาปัตยกรรมเพื่อหาโมเดลที่มีประสิทธิภาพที่สุด จากนั้นปรับปรุงประสิทธิภาพโดยการเพิ่มเลเยอร์ในโครงข่ายที่มีการเชื่อมถึงกันหมด และทดสอบกับข้อมูลทดสอบ ดังภาพที่ 1 กรอบการวิจัยการคัดแยกขยะขวดน้ำด้วยการเรียนรู้เชิงลึก



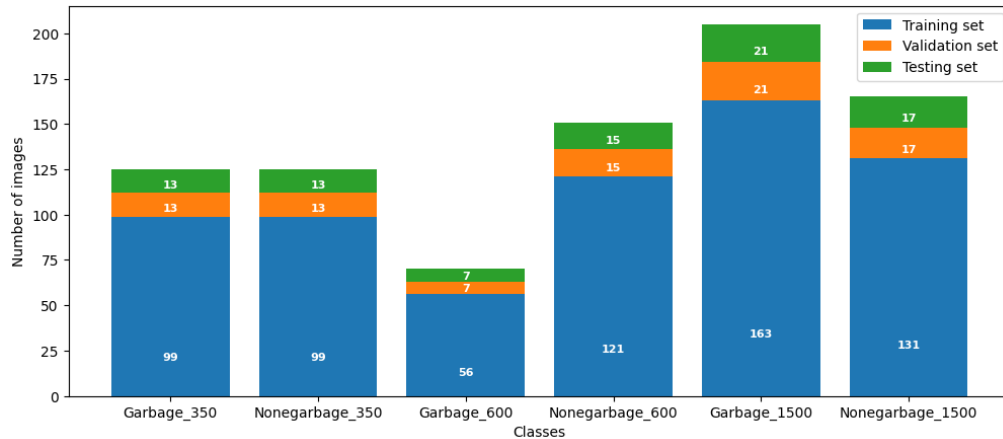
ภาพที่ 1: กรอบการวิจัยการคัดแยกขยะขวดน้ำด้วยการเรียนรู้เชิงลึก

1. สร้างชุดข้อมูล

งานวิจัยครั้งนี้ ได้สร้างชุดข้อมูลขยะขวดน้ำ (Water bottle waste dataset) จากการถ่ายภาพขวดน้ำพลาสติก โดยมีขนาดภาพ 1080 x 1920 พิกเซล และสับคั่นจากอินเทอร์เน็ต แบ่งออกเป็น 6 คลาส ได้แก่ คลาสที่เป็นขยะขวดน้ำขนาด 350 มิลลิลิตร (Garbage_350) คลาสที่ไม่เป็นขยะขวดน้ำขนาด 350 มิลลิลิตร (Nonegarbage_350) คลาสที่เป็นขยะขวดน้ำขนาด 600 มิลลิลิตร (Garbage_600) คลาสที่ไม่เป็นขยะขวดน้ำขนาด 600 มิลลิลิตร (Nonegarbage_600) คลาสที่เป็นขยะขวดน้ำขนาด 1500 มิลลิลิตร (Garbage_1500) และคลาที่ไม่เป็นขยะขวดน้ำขนาด 1500 มิลลิลิตร (Nonegarbage_1500) ดังภาพที่ 2 ก) รวมจำนวนทั้งหมด 841 ภาพโดยแบ่งเป็นชุดการเรียนรู้ (Training set) 80% หรือเท่ากับ 669 ภาพ และชุดสำหรับตรวจสอบ (Validation set) 10% หรือเท่ากับ 86 ภาพ และชุดทดสอบ (Testing set) 10% หรือเท่ากับ 86 ภาพ ดังภาพที่ 2 ข) การกำหนดคลาสที่เป็นขยะขวดน้ำ ใช้ภาพที่เป็นขวดเปล่าหรือมีน้ำเหลือที่ก้นขวดเพียงเล็กน้อย และคลาสที่ไม่เป็นขยะขวดน้ำเป็นภาพที่มีน้ำเหลือในขวดปริมาณมากหรือมีหลอดภายในขวด



ก)



ข)

ภาพที่ 2: ชุดข้อมูลขยะขวดน้ำ ก) ตัวอย่างภาพในแต่ละคลาส และ ข) จำนวนภาพในแต่ละคลาส

2. การวัดประสิทธิภาพของโมเดล

2.1 ค่าความถูกต้อง

งานด้านการจำแนกภาพสามารถใช้การวัดค่าความถูกต้อง (Accuracy) ซึ่งเป็นการแสดงความถูกต้องของการจำแนกข้อมูลที่มีความสอดคล้องตรงกันระหว่างจุดตรวจสอบกับข้อมูลอ้างอิง กล่าวคือพิจารณาโดยรวมของทุกชั้นข้อมูลที่จำแนกแล้วแสดงเป็นค่าความถูกต้องเดียว [8] ดังสมการ (1)

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

โดย True positive (TP) คือ จำนวนข้อมูลที่โมเดลทำนายได้อย่างถูกต้อง
 True negative (TN) คือ จำนวนข้อมูลที่โมเดลนั้นทำนายว่าไม่ถูกต้อง
 False positive (FP) คือ จำนวนข้อมูลที่ไม่ถูกต้องแต่โมเดลทำนายว่าข้อมูลนั้นถูกต้อง
 False negative (FN) คือจำนวนข้อมูลที่ถูกต้องแต่โมเดลทำนายว่าข้อมูลนั้นไม่ถูกต้อง

2.2 ค่าความผิดพลาด

ค่าความผิดพลาด (Loss) คือ ค่าคลาดเคลื่อนที่เกิดขึ้นโดยคำนวณ ได้จากการหาผลต่างของค่าที่เป็นจริงกับค่าที่ได้จากการพยากรณ์ที่เป็นจริง [9] แสดงดังสมการ (2)

$$Loss = \frac{1}{n} \sum_{i=1}^n (A - P)^2 \quad (2)$$

โดย A คือ ค่าของข้อมูลขยะขวดน้ำที่เป็นจริง

3. ทฤษฎีและเทคโนโลยีที่ใช้

การเรียนรู้เชิงลึก (Deep learning) เป็นวิธีการเรียนรู้ที่มีต้นแบบมาจากโครงข่ายประสาทของมนุษย์ที่สามารถเรียนรู้แบบอัตโนมัติ เป็นการนำโครงข่ายประสาทเทียมมาซ้อนกันหลาย ๆ ชั้น (Layer) เป็นรูปแบบหนึ่งของการเรียนรู้ของเครื่องแบบมีผู้สอน ใช้สำหรับงานด้านการจำแนกและตรวจจับภาพ เช่น โครงข่ายประสาทเทียมแบบคอนโวลูชัน [10]

โครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolution neural network: CNN) เป็นสถาปัตยกรรมการเรียนรู้เชิงลึกประเภทหนึ่ง ใช้สำหรับการจัดกลุ่มหรือจำแนกข้อมูลประเภทรูปภาพ [4] ซึ่งลักษณะเด่นของสถาปัตยกรรมโครงข่ายประสาทเทียมแบบคอนโวลูชัน คือ วิธีการคำนวณแบบคอนโวลูชันและจำนวนชั้น (Layer) ที่ใช้ในการคำนวณ ตัวอย่างเช่น VGG ที่ถูกออกแบบให้มีจำนวนชั้นมากถึง 16 และ 19 ชั้น ตามลำดับ [11] ในงานวิจัยนี้เลือกใช้สถาปัตยกรรม EfficientNetV2 [7] ซึ่งมีโครงข่ายของสถาปัตยกรรมดังภาพที่ 3

ภาพที่ 3: สถาปัตยกรรมโครงข่ายประสาทเทียมแบบคอนโวลูชันของ EfficientNetB2 [12]

โครงข่ายที่มีการเชื่อมถึงกันหมด (Fully connected layer) อยู่ในส่วนท้ายของโครงข่ายประสาทเทียมแบบคอนโวลูชัน มีการเชื่อมต่อแบบเต็มรูปแบบกับฟังก์ชันกระตุ้นทั้งหมดในชั้นก่อนหน้า แล้วถูกเชื่อมต่อกับชั้นสุดท้ายด้วยชั้นฟังก์ชัน Softmax เพื่อแสดงผลความน่าจะเป็นของข้อมูลขาออก [3]

เป็นภาษาระดับสูงที่รองรับเครื่องมือมากมายซึ่งทำให้ง่ายต่อการใช้งานและการอ่านและเขียนโปรแกรมภาษาไพธอนง่ายกว่ามากเมื่อเทียบกับภาษาอื่น ๆ เช่น C++, Java และ C# โดยภาษาไพธอนช่วยให้การเขียนโปรแกรมนั้นมุ่งเน้นไปที่โซลูชันของการแก้ปัญหา มากกว่าไวยากรณ์ของภาษาไพธอน [13]

3.5 TensorFlow

TensorFlow คือ ไลบรารีสำหรับการสร้างโมเดลการเรียนรู้เชิงลึกที่พัฒนาโดยบริษัทกูเกิ้ล รองรับการทำงานแบบประมวลผลร่วมกันหลาย ๆ เครื่องและสามารถใช้ GPU ในการประมวลผลได้ มีอัลกอริทึมสำหรับการสร้างโครงข่ายประสาทเทียมและการเรียนรู้เชิงลึกที่หลากหลาย โดยมีการนำไปประยุกต์ใช้กับงานวิจัยในหลายสาขาวิชา เช่น การรู้จำเสียง การประมวลผลภาพ และสร้างหุ่นยนต์ เป็นต้น ซึ่งการทำงานของ TensorFlow จะมีการสร้างกราฟสำหรับการประมวลผล [14]

3.6 Keras

Keras คือ ส่วนประกอบซอฟต์แวร์สองส่วนที่สามารถสื่อสารกันได้โดยใช้โปรโตคอลเดียวกัน รองรับการสร้างโมเดลการเรียนรู้ของเครื่องในรูปแบบโครงข่ายประสาทเทียม ตลอดจนการสร้างโมเดลการเรียนรู้เชิงลึก โดยมีข้อดี คือ ง่ายต่อการเขียนและทำความเข้าใจ มีการทำงานแยกเป็นส่วน ๆ สามารถแยกส่วนประกอบต่าง ๆ ในการสร้างโมเดล เช่น Neural layers, Cost functions, Optimizers และ Activation functions ประกอบกันเป็นโมเดลใหม่ได้ โดยการพัฒนาทั้งหมดต้องใช้ภาษาไพธอนในการพัฒนา [14]

3.7 Google colab

เป็นบริการจาก Google Research เป็นเครื่องมือที่ช่วยในการพัฒนาโปรแกรมที่อนุญาตให้ผู้ใช้สามารถเขียนซอร์สโค้ด แก้ไข และเรียกใช้จากเบราว์เซอร์ได้ ซึ่งรองรับภาษาการเขียนโปรแกรมด้วยภาษาไพธอน และเน้นงานด้านปัญญาประดิษฐ์ ซึ่งการทดลองของงานวิจัยนี้ ใช้บริการของ Google colab pro โดยมีหน่วยประมวลผลกลาง คือ Intel(R) Xeon(R) CPU @ 2.20 GHz หน่วยความจำขนาด 32 GB และมีหน่วยประมวลผลภาพ Tesla T4 [15]

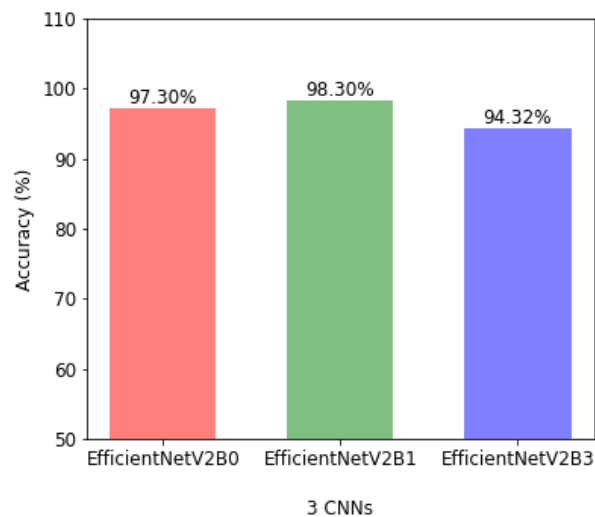
ผลการวิจัยและอภิปรายผล

การเปรียบเทียบประสิทธิภาพการคัดแยกขยะขวดน้ำด้วยการเรียนรู้เชิงลึก มีขั้นตอนการทดลอง 3 ขั้นตอน เริ่มจากการหาโมเดลที่มีประสิทธิภาพมากที่สุดจากการเปรียบเทียบโครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolution neural network: CNN) 3 สถาปัตยกรรม จากนั้นปรับปรุงประสิทธิภาพด้วยการเพิ่มเลเยอร์ในโครงข่ายที่มีการเชื่อมถึงกันหมด (Fully connected layer) ของสถาปัตยกรรมทั้ง 3 แบบ และขั้นตอนสุดท้ายนำโมเดลที่ดีที่สุดไปทดสอบกับข้อมูลทดสอบ

1. ผลการเปรียบเทียบประสิทธิภาพโมเดล

การทดลองในขั้นตอนนี้ ใช้บริการของ Google Colab [15] เพื่อเปรียบเทียบประสิทธิภาพโครงข่ายประสาทเทียมแบบคอนโวลูชัน 3 สถาปัตยกรรม ได้แก่ EfficientNetV2B0, EfficientNetV2B1 และ EfficientNetV2B3 ในการฝึกฝนโมเดลกับชุดข้อมูลขยะขวดน้ำ (Water bottle waste dataset) ด้วยชุดการเรียนรู้ (Training set) 80% และชุดสำหรับตรวจสอบ (Validation set) 10% ผ่านกระบวนการถ่ายทอดการเรียนรู้ (Transfer Learning) ด้วย ImageNet ซึ่งเป็นโมเดลที่ผ่านการสอนแล้ว (Pre-trained model) การกำหนดไฮเปอร์พารามิเตอร์สำหรับการเรียนรู้ ได้แก่ Optimizer เป็น Adam อัตราการเรียนรู้ (Learning rate) เท่ากับ 0.0001 และ Batch size มีขนาดเท่ากับ 8 และจำนวนในการฝึกฝนโมเดล (Epoch) 100 รอบ ฟังก์ชันกัณาคำนวณค่าความสูญเสีย (Loss function) กำหนดเป็น Categorical cross-entropy และฟังก์ชันกระตุ้น (Activation function) เป็น Softmax

จากการฝึกฝนโมเดลและประเมินค่าความถูกต้องกับชุดสำหรับตรวจสอบ พบว่า EfficientNetV2B1 มีค่าความถูกต้องสูงสุดที่ 98.30% ลำดับถัดมาเป็น EfficientNetV2B0 เท่ากับ 97.30% และ EfficientNetV2B3 มีค่าความถูกต้องต่ำสุด 94.32% ดังภาพที่ 4 และจากการเปรียบเทียบจำนวนพารามิเตอร์ พบว่า EfficientNetV2B3 มีจำนวนพารามิเตอร์มากที่สุด 12,930,622 ล้านพารามิเตอร์ ส่วน EfficientNetV2B0 และ EfficientNetV2B1 มีจำนวนพารามิเตอร์เท่ากันคือ 5,919,312 ล้านพารามิเตอร์ ดังตารางที่ 1



ภาพที่ 4: ประสิทธิภาพการคัดแยกขยะชนิดน้ำของ EfficientNetV2B0, EfficientNetV2B1 และ EfficientNetV2B3 ด้วยวิธีปกติ

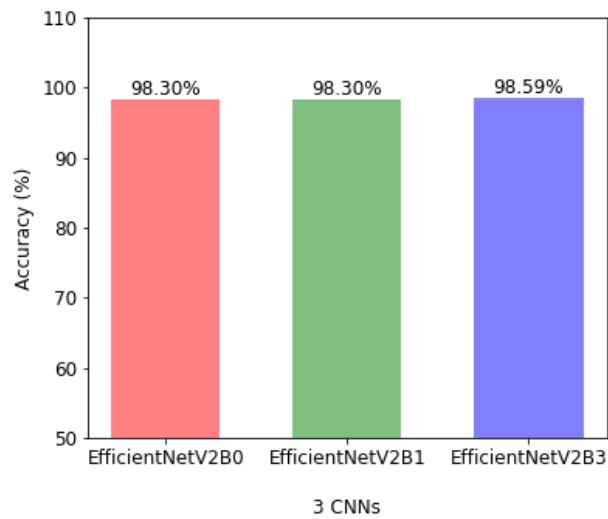
ตารางที่ 1: จำนวนพารามิเตอร์ของ EfficientNetV2B0, EfficientNetV2B1 และ EfficientNetV2B3 ด้วยวิธีปกติ

CNNs	จำนวนพารามิเตอร์ (M)
EfficientNetV2B0	5,919,312
EfficientNetV2B1	5,919,312
EfficientNetV2B3	12,930,622

2. ผลการเพิ่มประสิทธิภาพโมเดล

การทดลองในขั้นตอนนี้ เป็นการเพิ่มประสิทธิภาพโมเดลโดยการเพิ่มเลเยอร์ในโครงข่ายที่มีการเชื่อมถึงกันหมดของสถาปัตยกรรมทั้ง 3 แบบ เลเยอร์ที่ถูกเพิ่ม ได้แก่ Conv2D, BatchNormalization layer, DepthwiseConv2D, GlobalAveragePooling2D และ Flatten ในการฝึกฝนโมเดลใช้ภาพจากชุดการเรียนรู้ ชุดสำหรับตรวจสอบ และการกำหนดไฮเปอร์พารามิเตอร์สำหรับการเรียนรู้เหมือนกับขั้นตอนแรก

จากการฝึกฝนโมเดลและประเมินค่าความถูกต้องกับชุดสำหรับตรวจสอบ พบว่า EfficientNetV2B3 มีค่าความถูกต้องมากที่สุด 98.59% ส่วน EfficientNetV2B0 และ EfficientNetV2B1 ค่าความถูกต้องเท่ากัน 98.30% ดังภาพที่ 5 และจากการเปรียบเทียบจำนวนพารามิเตอร์ พบว่า EfficientNetV2B3 มีจำนวนพารามิเตอร์มากที่สุด 14,159,600 ล้านพารามิเตอร์ ส่วน EfficientNetV2B0 และ EfficientNetV2B1 มีจำนวนพารามิเตอร์เท่ากันคือ 6,943,490 ล้านพารามิเตอร์ ดังตารางที่ 2



ภาพที่ 5: ประสิทธิภาพการคัดแยกขยะขวดน้ำของ EfficientNetV2B0, EfficientNetV2B1 และ EfficientNetVB3 โดยการเพิ่มเลเยอร์

ตารางที่ 2: จำนวนพารามิเตอร์ของ EfficientNetV2B0, EfficientNetV2B1 และ EfficientNetVB3 ด้วยวิธีเพิ่มเลเยอร์

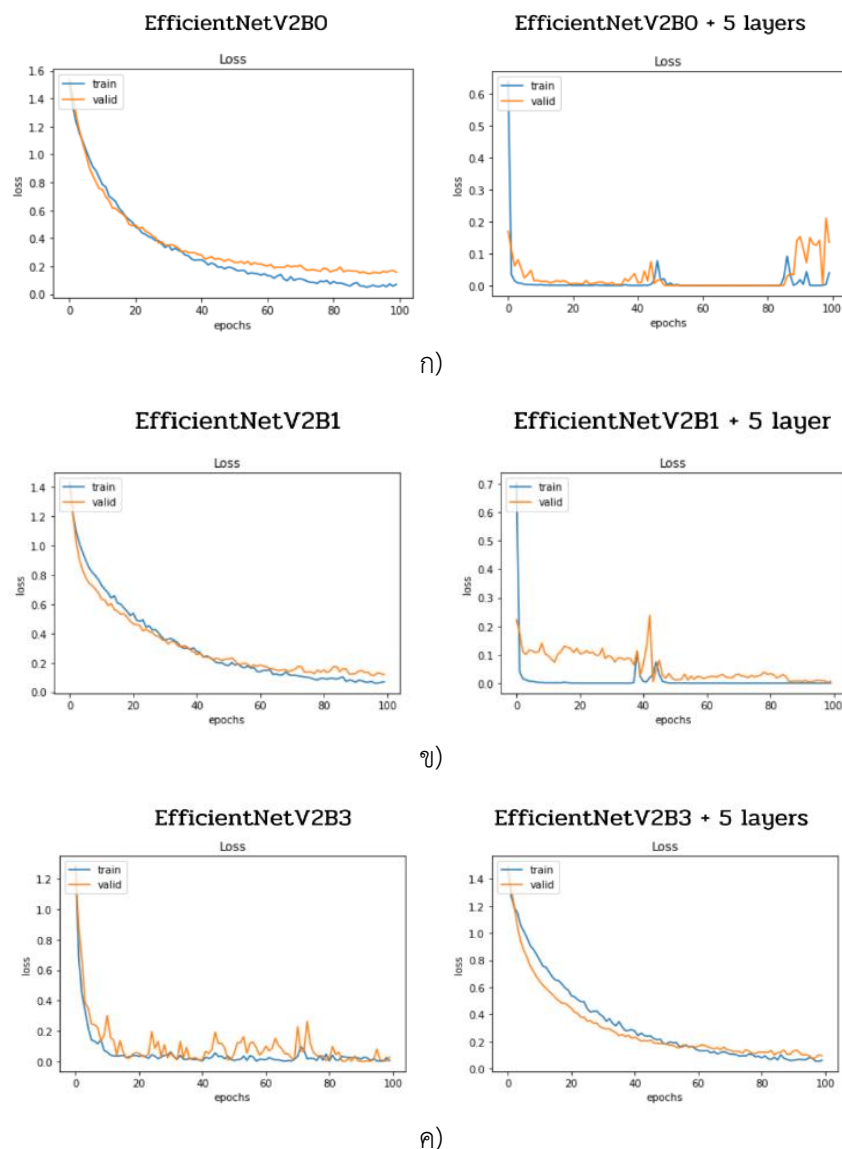
CNNs	จำนวนพารามิเตอร์ (M)
EfficientNetV2B0	6,943,490
EfficientNetV2B1	6,943,490
EfficientNetV2B3	14,159,600

ผลการทดลองด้วยวิธีการเพิ่มเลเยอร์ในโครงข่ายที่มีการเชื่อมถึงกันหมด 5 แบบ พบว่า ค่าความผิดพลาด (Loss) ของทั้ง 3 โมเดลลดลงเมื่อเปรียบเทียบกับวิธีปกติ โดยค่าความผิดพลาดของ EfficientNetV2B0 จาก 0.08 ลดลงเหลือ 0.07 ค่าความผิดพลาดของ EfficientNetV2B1 จาก 0.13 ลดลงเหลือ 0.09 และค่าความผิดพลาดของ EfficientNetV2B3 จาก 0.23 ลดลงเหลือ 0.10 การใช้วิธีเพิ่มเลเยอร์ด้วย EfficientNetV2B0 มีค่าความผิดพลาดที่ต่ำสุดคือ 0.07 ดังตารางที่ 3

ตารางที่ 3: ค่าความผิดพลาดระหว่างวิธีปกติและวิธีเพิ่มเลเยอร์ของ EfficientNetV2B0, EfficientNetV2B1 และ EfficientNetVB3

CNNs	ค่าความผิดพลาด	
	วิธีปกติ	วิธีเพิ่มเลเยอร์
EfficientNetV2B0	0.08	0.07
EfficientNetV2B1	0.13	0.09
EfficientNetV2B3	0.23	0.10

กราฟแสดงค่าความผิดพลาดจากการฝึกฝนโมเดลกับชุดข้อมูลการเรียนรู้และชุดข้อมูลสำหรับตรวจสอบของ 3 CNNs (EfficientNetV2B0, EfficientNetV2B1 และ EfficientNetV2B3) จากการทดลอง 2 วิธี ได้แก่ วิธีปกติ และวิธีเพิ่มเลเยอร์ในโครงข่ายที่มีการเชื่อมถึงกันหมด 5 แบบ พบว่า EfficientNetV2B0 มีค่าความผิดพลาดของวิธีเพิ่มเลเยอร์ลดลงอย่างรวดเร็วเมื่อเปรียบเทียบกับวิธีปกติ และแนวโน้มค่าความผิดพลาดของทั้ง 2 วิธีเพิ่มขึ้นเมื่อรอบการฝึกฝนเข้าใกล้ 100 ดังภาพที่ 6 ก) สำหรับ EfficientNetV2B1 มีค่าความผิดพลาดของวิธีเพิ่มเลเยอร์ลดลงอย่างรวดเร็วเมื่อเปรียบเทียบกับวิธีปกติ และค่าความผิดพลาดของวิธีเพิ่มเลเยอร์มีแนวโน้มลดลงเมื่อรอบการฝึกฝนเข้าใกล้ 100 ดังภาพที่ 6 ข) และ EfficientNetV2B3 มีค่าความผิดพลาดของวิธีปกติลดลงอย่างรวดเร็วเมื่อเปรียบเทียบกับวิธีเพิ่มเลเยอร์ และค่าความผิดพลาดของวิธีปกติมีแนวโน้มลดลงเมื่อรอบการฝึกฝนเข้าใกล้ 100 ดังภาพที่ 6 ค)



ก)

ข)

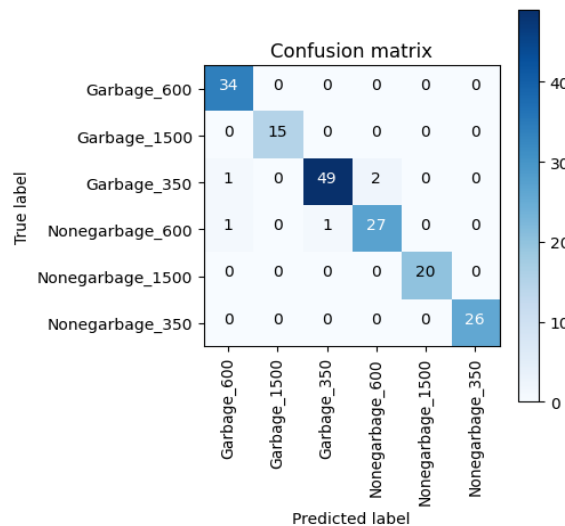
ค)

ภาพที่ 6: ค่าความผิดพลาดระหว่างวิธีปกติและวิธีเพิ่มเลเยอร์ของ EfficientNetV2B0, EfficientNetV2B1 และ EfficientNetV2B3

3. ผลการประเมินประสิทธิภาพโมเดล

ผลจากการทดลองระหว่างวิธีปกติและวิธีเพิ่มเลเยอร์ในโครงข่ายที่มีการเชื่อมถึงกันหมด 5 แบบ พบว่า EfficientNetV2B3 ด้วยวิธีเพิ่มเลเยอร์มีค่าความถูกต้องมากที่สุด คือ 98.59% เมื่อนำไปทดสอบกับชุดข้อมูลทดสอบ (Testing set) มีค่าความถูกต้อง 97.16% และค่าความผิดพลาด 0.08

แสดงผลลัพธ์การทำนายด้วย Confusion matrix ดังภาพที่ 7 พบว่า มี 2 คลาสที่ทำนายผิด คือ คลาส Nonegarbage_600 (ไม่เป็นขยะขวดน้ำขนาด 600 มิลลิลิตร) ทำนายเป็น คลาส Garbage_600 (ขยะขวดน้ำขนาด 600 มิลลิลิตร) 1 ภาพ คลาส Nonegarbage_600 (ไม่เป็นขยะขวดน้ำขนาด 600 มิลลิลิตร) ทำนายเป็นคลาส Garbage_350 (ขยะขวดน้ำขนาด 350 มิลลิลิตร) 1 ภาพ คลาส Garbage_350 (ขยะขวดน้ำขนาด 350 มิลลิลิตร) ทำนายเป็นคลาส Garbage_600 (ขยะขวดน้ำขนาด 600 มิลลิลิตร) 1 ภาพ และคลาส Garbage_350 (ขยะขวดน้ำขนาด 350 มิลลิลิตร) ทำนายเป็นคลาส Nonegarbage_600 (ไม่เป็นขยะขวดน้ำขนาด 600 มิลลิลิตร) 2 ภาพ เนื่องจากระดับน้ำในขวดน้ำมีการสะท้อนแสงและเงาในภาพ ทำให้เกิดความคลาดเคลื่อนไปจากรูปแบบเดิมของภาพ จึงทำให้รูปแบบของขยะขวดน้ำในแต่ละคลาสมีความคล้ายคลึงกัน



ภาพที่ 7: ประสิทธิภาพการคัดแยกขยะขวดน้ำของ EfficientNetV2B3

สรุปผล

งานวิจัยการเปรียบเทียบประสิทธิภาพการคัดแยกขยะขวดน้ำด้วยการเรียนรู้เชิงลึก มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพการคัดแยกขยะขวดน้ำกับชุดข้อมูลขยะขวดน้ำ (Water bottle waste datasets) จำนวน 6 คลาส จากภาพถ่ายจำนวนทั้งสิ้น 841 ภาพ โดยใช้การเรียนรู้เชิงลึกด้วยเทคนิคโครงข่ายประสาทเทียมแบบคอนโวลูชัน มีขั้นตอนการทดลอง 3 ขั้นตอน ได้แก่ เปรียบเทียบประสิทธิภาพโมเดล เพิ่มประสิทธิภาพโมเดล และการประเมินประสิทธิภาพโมเดล ขั้นแรกเป็นการเปรียบเทียบประสิทธิภาพโมเดลของโครงข่ายประสาทเทียมแบบคอนโวลูชัน 3 สถาปัตยกรรม ได้แก่ EfficientNetV2B0, EfficientNetV2B1 และ EfficientNetV2B3 ผลการทดลอง พบว่า EfficientNetV2B1 มีค่าความถูกต้องสูงที่สุด 98.30% และ EfficientNetV2B3 มีจำนวนพารามิเตอร์มากที่สุด 12,930,622 ล้านพารามิเตอร์ จากนั้นเป็นการเพิ่มประสิทธิภาพโมเดลโดยการเพิ่มเลเยอร์ในโครงข่ายที่มีการเชื่อมถึง

กันหมดของสถาปัตยกรรมทั้ง 3 แบบ เลเยอร์ที่ถูกเพิ่ม ได้แก่ Conv2D, BatchNormalization layer, DepthwiseConv2D, GlobalAveragePooling2D และ Flatten จากการทดลองพบว่า EfficientNetV2B3 มีค่าความถูกต้องมากที่สุด 98.59% และ EfficientNetV2B3 มีจำนวนพารามิเตอร์มากที่สุด 14,159,600 ล้านพารามิเตอร์จากการผลการทดลองของทั้ง 2 วิธี พบว่า ค่าความถูกต้องสูงสุดของทั้ง 2 วิธีใกล้เคียงกัน โดยวิธีปกติ 98.30% (EfficientNetV2B1) และวิธีเพิ่มเลเยอร์ 98.59% (EfficientNetV2B3) แต่จากการเปรียบเทียบค่าความผิดพลาดของทั้ง 2 วิธี พบว่า สถาปัตยกรรมทั้ง 3 แบบที่ใช้วิธีเพิ่มเลเยอร์มีค่าความผิดพลาดน้อยกว่าวิธีปกติ ซึ่ง EfficientNetV2B3 มีค่าความผิดพลาดลดลงจากวิธีปกติถึง 0.13 และ EfficientNetV2B0 ที่ใช้วิธีเพิ่มเลเยอร์มีค่าความผิดพลาดต่ำสุด 0.07 แสดงให้เห็นว่าวิธีเพิ่มเลเยอร์ไม่เหมาะสมกับทุกสถาปัตยกรรมแต่เหมาะกับสถาปัตยกรรมขนาดใหญ่ เช่น EfficientNetV2B3 แต่วิธีเพิ่มเลเยอร์ก็ทำให้จำนวนพารามิเตอร์ของทั้ง 3 สถาปัตยกรรมเพิ่มจากวิธีปกติ 1 ล้านพารามิเตอร์โดยประมาณ และในขั้นตอนสุดท้ายเป็นการประเมินประสิทธิภาพโมเดล เมื่อนำ EfficientNetV2B3 ไปทดสอบกับชุดข้อมูลทดสอบ (Testing set) มีค่าความถูกต้อง 97.16% และค่าความผิดพลาด 0.08 จากผลลัพธ์การทำนายด้วย Confusion matrix ผลการทำนายที่ผิดพลาดเกิดจากภาพขวดน้ำพลาสติกมีการสะท้อนของแสงและเงาทำให้ขนาดและลักษณะของภาพนั้นเกิดความคล้ายกันระหว่างคลาส

ข้อเสนอแนะ

ผลการวิจัยนี้สามารถคัดแยกขยะขวดน้ำกับขวดน้ำพลาสติก 3 ขนาด ได้แก่ ขนาด 350 มิลลิลิตร ขนาด 600 มิลลิลิตร และขนาด 1,500 มิลลิลิตร ไม่ได้ครอบคลุมขวดพลาสติกขนาดอื่นและขยะประเภทอื่น และเพื่อแก้ปัญหาผลการทำนายผิดพลาด งานวิจัยถัดไปอาจใช้เทคนิคการเพิ่มภาพ (Data augmentation) [16] เพื่อปรับแสงเงา, BatchNormalization [17], Weight Decay [18] และ Dropout [19]

เอกสารอ้างอิง

- 1 Majchrowska S, Mikołajczyk A, Ferlin M, Klawikowska Z, Plantykowski MA, Kwasigroch A, et al. Deep learning-based waste detection in natural and urban environments. Waste Management 2022;138:274–84. <https://doi.org/10.1016/j.wasman.2021.12.001>.
- 2 Armitage S, Awty-Carroll K, Clewley D, Martinez-Vicente V. Detection and classification of floating plastic litter using a vessel-mounted video camera and deep learning. Remote Sens (Basel) 2022;14:3425. <https://doi.org/10.3390/rs14143425>.
- 3 Sai Susanth G, Jenila Livingston LM, Agnel Livingston LGX. Garbage waste segregation using deep learning techniques. IOP Conf Ser Mater Sci Eng 2021;1012:012040. <https://doi.org/10.1088/1757-899X/1012/1/012040>.
- 4 Fadlil A, Umar R, Sunardi ., Nugroho AS. Comparison of machine learning approach for waste bottle classification. Emerging Science Journal 2022;6:1075–85. <https://doi.org/10.28991/ESJ-2022-06-05-011>.

- 5 Chazhoor AAP, Ho ESL, Gao B, Woo WL. Deep transfer learning benchmark for plastic waste classification. *Intelligence & Robotics* 2022;2:1–19. <https://doi.org/10.20517/ir.2021.15>.
- 6 Liu W, Ouyang H, Liu Q, Cai S, Wang C, Xie J, et al. Image recognition for garbage classification based on transfer learning and model fusion. *Math Probl Eng* 2022;2022:1–12. <https://doi.org/10.1155/2022/4793555>.
- 7 Tan M, Le Q V. EfficientNetV2: Smaller models and faster training. *Proceedings of the 38th International Conference on Machine Learning (PMLR)*, 2021;10096–10096.
- 8 ชลรวี หาญลำยวง. การคาดการณ์ดัชนีราคาหุ้นตลาดหลักทรัพย์แห่งประเทศไทยโดยใช้มะชินเลิร์นนิ่งเทคนิค. (สารนิพนธ์ปริญญาการจัดการมหาบัณฑิต). วิทยาลัยการจัดการ มหาวิทยาลัยมหิดล, 2562.
- 9 ภาณุพงษ์ ร่องอ้อ. ผลกระทบของการทำการสกัดคุณลักษณะของข้อมูลและการแทนที่ข้อมูลกับแบบจำลองพยากรณ์ฝุ่นละออง PM2.5 สำหรับบริเวณกรุงเทพมหานคร. (ปริญญาานิพนธ์วิทยาศาสตรมหาบัณฑิต). สาขาวิชาเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ, 2562.
- 10 Zhao Z-Q, Zheng P, Xu S, Wu X. Object detection with deep learning: A review. *IEEE Trans Neural Netw Learn Syst* 2019;30:3212–32.
- 11 Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *ArXiv Preprint ArXiv:14091556* 2014.
- 12 Tan M, Le Q. Efficientnet: Rethinking model scaling for convolutional neural networks. *International conference on machine learning*, PMLR; 2019;6105–14.
- 13 Srinath KR. Python—the fastest growing programming language. *International Research Journal of Engineering and Technology* 2017;4:354–7.
- 14 เขียวศักดิ์ พลาคัยเลิศ, ธนิตา นุ่มนนท์. การเปรียบเทียบโมเดลการเรียนรู้ของเครื่องแบบต่าง ๆ สำหรับการทำนายราคาบิตคอยน์. *วารสารเทคโนโลยีสารสนเทศลาดกระบัง* 2561;6:1–9.
- 15 Arantes PR, Polêto MD, Pedebos C, Ligabue-Braun R. Making it rain: Cloud-based molecular Simulations for everyone. *J Chem Inf Model* 2021;61:4852–6. <https://doi.org/10.1021/acs.jcim.1c00998>.
- 16 Lei C, Hu B, Wang D, Zhang S, Chen Z. A preliminary study on data augmentation of deep learning for image classification. *Proceedings of the 11th Asia-Pacific Symposium on Internetworking*, New York, NY, USA: ACM; 2019;1–6. <https://doi.org/10.1145/3361242.3361259>.
- 17 Zhu Q, He Z, Zhang T, Cui W. Improving classification performance of softmax loss function based on scalable batch-normalization. *Applied Sciences* 2020;10:2950. <https://doi.org/10.3390/app10082950>.
- 18 Bjorck J, Weinberger KQ, Gomes C. Understanding Decoupled and Early Weight Decay. *Proceedings of the AAAI Conference on Artificial Intelligence* 2021;35:6777–85. <https://doi.org/10.1609/aaai.v35i8.16837>.

- 19 Garbin C, Zhu X, Marques O. Dropout vs. batch normalization: an empirical study of their impact to deep learning. *Multimed Tools Appl* 2020;79:12777–815. <https://doi.org/10.1007/s11042-019-08453-9>.