



KKU SCIENCE JOURNAL

Journal Home Page : <https://ph01.tci-thaijo.org/index.php/KKUSciJ>

Published by the Faculty of Science, Khon Kaen University, Thailand



การตรวจจับการสวมหน้ากากอนามัยโดยเทคนิคการส่งถ่ายความรู้ Face Mask Detection using Transfer Learning Technique

สุขสันต์ พรหมบุญเรือง^{1*}, ธรรมรัตน์ บุญรอด¹, สุธิพงษ์ พลทอง¹ และ ธนา โชติประเดิม¹
Suksun Promboonruang^{1*}, Thummarat Boonrod¹, Suttipong Phontong¹
and Thana Chotpraderm¹

¹สาขาวิชาเทคโนโลยีดิจิทัล คณะบริหารศาสตร์ มหาวิทยาลัยกาฬสินธุ์ จังหวัดกาฬสินธุ์ 46000

¹Digital Technology, Faculty of Administrative Science, Kalasin University, Kalasin, 46000, Thailand

บทคัดย่อ

การตรวจจับการสวมหน้ากากอนามัยโดยเทคนิคการส่งถ่ายความรู้มีวัตถุประสงค์เพื่อสร้างโมเดลการตรวจจับการสวมหน้ากากอนามัยโดยเทคนิคการส่งถ่ายความรู้ และเปรียบเทียบประสิทธิภาพโมเดลการส่งถ่ายความรู้ เครื่องมือที่ใช้ภาษาไพธอนร่วมกับไลบรารี Keras สำหรับสร้างแบบจำลองโมเดลการเรียนรู้เชิงลึก และไลบรารี Open CV สำหรับการประมวลผลภาพ กลุ่มตัวอย่างข้อมูลใบหน้าสวมใส่หน้ากากและไม่สวมใส่หน้ากากอนามัยใช้ได้จากเว็บไซต์ kaggle โดยแต่ละภาพมีขนาด 150 x 150 พิกเซล จำนวน 3,833 ภาพ แบ่งเป็นภาพใบหน้าสวมใส่หน้ากาก จำนวน 1,915 ภาพ และภาพใบหน้าไม่สวมใส่หน้ากาก จำนวน 1,918 ภาพ วิธีการดำเนินการแบ่งออก 3 ขั้นตอน ได้แก่ การเตรียมข้อมูล การสร้างแบบจำลอง การวัดประสิทธิภาพโดยใช้ค่าความถูกต้อง (accuracy) และค่าความผิดพลาด (loss) โดยการเปรียบเทียบโมเดลแบบจำลอง 3 แบบ ได้แก่ VGG16 Inceptionresnetv2 และ ResNet152V2 จากการทดลองแสดงให้เห็นว่าโมเดลแบบจำลอง ResNet152V2 มีค่าความถูกต้องมากที่สุด 98% และค่า Loss 0.08%

ABSTRACT

The face mask detection using transfer learning technique aims to create model for the face mask detection using transfer learning technique and compare the efficiency the models. The tools used are Python language with Keras and Open CV libraries for image processing. The dataset was downloaded from kaggle website, which contained 3,833 images of 150 x 150 pixels image resolution. The images were divided into 1,915 images of faces masks and 1,918 images of faces without masks. The methodology was divided into 3 steps: 1. Data preparation, 2. model generation, and 3. Efficiency evaluation of the 3 models including VGG16, Inceptionresnetv2, and ResNet152V2, using accuracy and loss assessment. The experimental results show that the ResNet152V2 model had the highest accuracy of 98% and the loss of 0.08%.

คำสำคัญ: การประมวลผลภาพ การเรียนรู้เชิงลึก การตรวจจับการสวมหน้ากากอนามัย

Keywords: Image Processing, Deep Learning, Mask Detection

*Corresponding Author, E-mail: suksun.pr@ksu.ac.th

บทนำ

การระบาดของเชื้อไวรัสโควิด-19 เริ่มระบาดในปี พ.ศ. 2562 และแพร่ระบาดไปทั่วโลก และเชื้อไวรัสมีความรุนแรงต่อระบบทางเดินหายใจส่งผลให้มีจำนวนผู้เสียชีวิตเป็นจำนวนมาก ส่งผลกระทบต่อสุขภาพ สังคม และเศรษฐกิจของโลก รวมไปถึงการใช้ชีวิตประจำวัน เช่น การเดินทางไปทำงาน การท่องเที่ยว เป็นต้น ต้องมีการป้องกันตัวเองเพื่อไม่ให้เสี่ยงต่อการติดเชื้อไวรัสโควิด-19 โดยวิธีการสวมใส่หน้ากากอนามัย สำหรับสถานการณ์การระบาดและการติดเชื้อไวรัสโควิด-19 ในประเทศไทย กระทรวงสาธารณสุขได้กำหนดมาตรการเพื่อลดการแพร่ระบาดของเชื้อไวรัสโควิด-19 แต่อย่างไรก็ตามการควบคุมการแพร่ระบาดของไวรัสโควิด-19 ไม่สามารถควบคุมการแพร่ระบาดได้ ส่งผลให้เกิดการแพร่ระบาดไปยังจังหวัดต่างๆ จังหวัดกาฬสินธุ์เป็นจังหวัดหนึ่งที่มีการติดเชื้อไวรัสโควิด-19 และเฝ้าระวังติดตามผู้ที่เดินทางกลับมาจากประเทศหรือพื้นที่กลุ่มเสี่ยงที่มีการระบาดของเชื้อไวรัสโควิด-19 โดยความร่วมมือทั้งจากฝ่ายปกครอง สาธารณสุข กำนัน ผู้ใหญ่บ้าน และอาสาสมัคร (อสม.) เฝ้าระวังในพื้นที่อย่างใกล้ชิดและต่อเนื่อง ซึ่งจังหวัดกาฬสินธุ์ได้จัดอยู่ในเขตพื้นที่ควบคุมสูง ทำให้ส่งผลกระทบต่อภาครัฐกิจ สถานที่ท่องเที่ยว โรงเรียนและมหาวิทยาลัยที่กำลังเปิดภาคเรียน ส่งผลให้ไม่สามารถไปซื้อสินค้าไปท่องเที่ยว และไปเรียนได้ เนื่องจากการติดเชื้อไวรัสโควิด-19 (กรมควบคุมโรค, 2565)

วิธีป้องกันการติดเชื้อ และการลดโอกาสติดเชื้อไวรัสโควิด-19 วิธีที่ง่ายและสะดวกที่สุดได้แก่ การสวมใส่หน้ากากอนามัย เมื่อมีการสวมใส่หน้ากากอนามัยทำให้ระบบตรวจสอบบุคคลทำได้ยาก เนื่องจากใบหน้าถูกบังด้วยหน้ากากอนามัย ซึ่งทำให้เป็นช่องทางให้กลุ่มมิจฉาชีพอาศัยการสวมใส่หน้ากากอนามัยในการลงมือกระทำความผิด เพราะการจดจำใบหน้าทำได้ยาก ผู้ที่พบเห็นคนร้ายไม่สามารถระบุตัวบุคคลได้ การแก้ไขปัญหาดังกล่าวสามารถนำวิธีการทางด้านการประมวลผลภาพมาใช้ในการวิเคราะห์ใบหน้าบุคคลเพื่อระบุตัวบุคคล ซึ่งมีหลายวิธี เช่น เทคนิคการเรียนรู้เชิงลึก (deep learning) ซึ่งเป็นการใช้โมเดลเชิงลึก (deep neural network) โดยคณะผู้วิจัยได้ใช้วิธีการแก้ไขปัญหาดังกล่าวด้วยการตรวจจับการสวมหน้ากากอนามัยโดยเทคนิคการส่งถ่ายความรู้ เทคนิคการส่งถ่ายความรู้เป็นเทคนิคที่ใช้ความรู้ที่ได้จากโมเดลที่เรียนรู้ไว้ในการเรียนรู้โมเดลใหม่ช่วยให้โมเดลใหม่สามารถเรียนรู้ภายในเวลาน้อยกว่าโดยไม่ต้องเรียนรู้จากข้อมูลทั้งหมดใหม่ ดังนั้นการเรียนรู้เพื่อสร้างโมเดลใหม่จะทำให้รวดเร็วขึ้น (Chellapandi *et al.*, 2021; Rajan and Noumida, 2021)

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

1. การประมวลผลภาพ (Image Processing)

การประมวลผลภาพเกิดขึ้นครั้งแรกในปี ค.ศ. 1964 ที่ห้องแลป Jet propulsion รัฐแคลิฟอร์เนีย ประเทศสหรัฐอเมริกา โดยงานชิ้นแรกคือการพิจารณาภาพถ่ายดาวเทียมของดวงจันทร์ ซึ่งได้มาจากดาวเทียม Ranger 7 ดังรูปที่ 1 หลังจากนั้นมีการจัดตั้งสาขาทางวิทยาศาสตร์ ชื่อการประมวลผลภาพ หรือ Digital image processing ทำให้งานด้านการประมวลผลภาพเป็นที่น่าสนใจเป็นอย่างมาก และมีการนำไปประยุกต์ใช้กับงานในหลายๆ ด้าน เช่น การสื่อสารทางโทรศัพท์ ทางด้านการพิมพ์ภาพ การแพทย์ และงานด้านวิทยาศาสตร์ (Gonzalez and Woods, 2006)



รูปที่ 1 ภาพถ่ายดาวเทียมของดวงจันทร์จากดาวเทียม Ranger 7 ที่มา: Gonzalez และ Woods (2006)

ชนิดของภาพ แบ่งออกได้เป็น 3 ประเภท ดังนี้

1.1 ภาพขาวดำ (Binary image) ภาพขาวดำเป็นภาพที่ใช้เนื้อที่เพียง 1 บิต ต่อหนึ่งพิกเซล และเป็นภาพที่ได้รับการ Quantised ค่าเป็นสองค่า ซึ่งจะแทนค่าเป็น 0 (สีดำ) และ 1 (สีขาว) ภาพขาวดำมีการใช้ในงานจำนวนมากเนื่องจากง่ายต่อการดำเนินการ (Owens, 2022)

1.2 ภาพระดับสีเทา (Grayscale image) เป็นภาพเฉดสีเทา จะถูกเก็บเป็นจำนวนเต็ม 8 บิต เก็บค่าระดับสีเทา ตั้งแต่ 0 ถึง 255 เฉดสี (0 คือระดับเข้ม 255 คือระดับความสว่าง) ภาพระดับสีเทาเป็นภาพที่เพียงพอสำหรับงานการประมวลผลภาพ และไม่มี ความซับซ้อน

1.3 ภาพสี (Color image) สีเป็นคุณลักษณะที่ใช้อธิบายวัตถุได้ดี หรือใช้ในการแบ่งกลุ่มวัตถุ ซึ่งระบบสีที่มีการใช้งานด้านการประมวลผลภาพ ได้แก่ ระบบสี RGB (R คือ สีแดง G คือ สีเขียว B คือ สีน้ำเงิน) แต่การรับรู้สี RGB ไม่คงที่ต่อการเปลี่ยนแปลงของสภาพแสง ระบบสี HSV เป็นระบบสีที่มีรูปแบบเป็นเอกลักษณ์ ประกอบด้วย 3 องค์ประกอบ ได้แก่ H (Hue) คือค่าของเนื้อ สีจะอยู่ระหว่าง 0 - 255 ซึ่งถ้า Hue มีค่าเท่ากับ 0 จะแทนด้วยสีแดง และเมื่อ Hue มีค่ามากขึ้นเรื่อยๆ สีจะเปลี่ยนแปลงไปตามสเปกตรัมของสีจนถึงค่า 255 จะกลับมาเป็นสีแดงอีกครั้ง ซึ่งสามารถแทนด้วยองศาได้ดังนี้ สีแดง 0 องศา สีเขียว 120 องศา และสีน้ำเงิน 240 องศา Saturation คือค่าความบริสุทธิ์ของสี ซึ่งถ้า Saturation มีค่าเท่ากับ 0 แล้วสีที่ได้จะไม่มี Hue ซึ่งจะได้สีออกมาเป็นสีขาวล้วน ถ้าค่าเท่ากับ 255 แสดงว่าไม่มีแสงสีขาวผสมอยู่เลย Value คือความสว่างของสี ซึ่งสามารถวัดได้โดยค่าความเข้มของความสว่าง (Color, 2022)

2. การจำแนกข้อมูล (Image Classification)

การจำแนกรูปภาพเป็นเทคนิคการจัดประเภทภาพโดยใช้โมเดลประสาทเทียม โดยโมเดลจะเรียนรู้จากข้อมูลที่ได้รับมาและทำนายค่าให้กับภาพใหม่ที่ได้รับมา ซึ่งจะมีการใช้ Convolutional Neural Network (CNN) เป็นหลัก โดย CNN จะใช้การคำนวณหลายครั้งบนภาพโดยใช้เครื่องมือที่เรียกว่า Convolution และ Pooling เพื่อหาคุณลักษณะของภาพ และจะต้องปรับปรุงค่าพารามิเตอร์ในโมเดลให้เหมาะสม ด้วยการใช้ข้อมูลฝึกหัด เพื่อให้โมเดลสามารถจัดประเภทภาพได้เป็นอย่างดี (Akshai and Anitha, 2021)

3. การส่งถ่ายความรู้ (Transfer Learning)

Transfer learning เป็นการใช้ความรู้ที่ได้จากโมเดลที่เรียนรู้ไว้ในการเรียนรู้โมเดลใหม่ ซึ่งจะช่วยให้โมเดลใหม่สามารถเรียนรู้ภายในเวลาน้อยกว่าโดยไม่ต้องเรียนรู้จากข้อมูลใหม่เสมอ ในกรณีที่ต้องการใช้โมเดลประมวลผลภาพในการจำแนกรูปภาพ หากมีโมเดลประสาทเทียมที่เรียนรู้แล้วในการจำแนกรูปภาพประเภทต่างๆ สามารถใช้ความรู้ของโมเดลนั้นในการเรียนรู้โมเดลใหม่ โดยไม่ต้องเรียนรู้จากข้อมูลใหม่เสมอ (Chellapandi *et al.*, 2021; Rajan and Noumida, 2021)

VGG16 เป็นสถาปัตยกรรมโครงข่ายประสาทเทียมแบบ Convolutional (CNN) ที่พัฒนาโดย Visual Geometry Group (VGG) ที่มหาวิทยาลัยออกซฟอร์ด สถาปัตยกรรมเป็นที่รู้จักในด้านความเรียบง่ายและการใช้สถาปัตยกรรมที่สอดคล้องกันในทุกเลเยอร์ซึ่งช่วยให้โมเดลได้รับการฝึกฝนและใช้งานจดจำภาพได้อย่างง่าย สถาปัตยกรรมของ VGG16 ประกอบด้วย 16 ชั้น ซึ่งแบ่งออกเป็นสองส่วนหลัก ชั้นคอนโวลูชันและชั้นที่เชื่อมต่ออย่างเต็มที่ เลเยอร์ที่ซับซ้อนมีหน้าที่ในการแยกคุณสมบัติจากภาพอินพุตในขณะที่เลเยอร์ที่เชื่อมต่ออย่างสมบูรณ์จะใช้สำหรับการจำแนกประเภท (Akshai and Anitha, 2021)

ResNet152V2 เป็นส่วนหนึ่งของสถาปัตยกรรม ResNet152 ซึ่งเป็นโครงข่ายประสาทเทียมแบบ Deep Convolutional (CNN) ที่ออกแบบมาสำหรับงานการจำแนกภาพ "V2" ในชื่อระบุว่าเป็นรุ่นที่สองของสถาปัตยกรรม ResNet152 (Akshai and Anitha, 2021)

ความแตกต่างที่สำคัญระหว่าง ResNet152V2 และ ResNet152 ตั้งเดิมคือรุ่น V2 ใช้บล็อก Bottleneck ซึ่งลดจำนวนช่องสัญญาณในเครือข่ายในขณะที่ยังเป็นตัวแทน ช่วยให้เครือข่ายลึกขึ้นด้วยพารามิเตอร์ที่น้อยลง และการคำนวณน้อยลงทำให้การเรียนรู้และปรับใช้มีประสิทธิภาพมากขึ้น

ResNet152V2 เป็นเครือข่ายลึก 152 ชั้น ในการคอนโวลูชันและ 2 ชั้นที่เชื่อมต่ออย่างสมบูรณ์ โมเดลนี้ได้รับการเรียนรู้เกี่ยวกับชุดข้อมูล Imagenet ซึ่งมีรูปภาพมากกว่า 14 ล้านภาพ และคลาส 1,000 คลาส (Chellapandi *et al.*, 2021; Rajan and Noumida, 2021)

4. โครงข่ายประสาทเทียมแบบสังวัตนาการ (Convolutional Neural Network: CNN) (Yann *et al.*, 1998)

เป็นโครงข่ายประสาทเทียมที่ใช้ในการประมวลผลภาพและตรวจจับวัตถุ เพื่อค้นหาคุณลักษณะและทำซ้ำหลายรอบจนกว่าจะได้รับความแม่นยำของการคัดแยกคุณลักษณะ โดยพิจารณาจากความสัมพันธ์ของคุณลักษณะกับผลลัพธ์มากที่สุด มีวิธีการทำงาน ดังนี้

4.1 Convolutional Layer เป็นชั้นแรกของโครงข่าย จะใช้เครื่องมือหรือตัวแปรที่เรียกว่า ตัวกรอง (Filter) ในการคำนวณค่าของภาพตามพื้นที่ที่กำหนด ซึ่งจะเป็นการค้นหาคุณลักษณะที่สำคัญในภาพ

4.2 ReLU (Rectified Linear Unit) Layer เป็นชั้นที่จะทำการแปลงค่าของภาพให้มีค่าเป็น 0 หรือ 1 โดยใช้ฟังก์ชันตรวจสอบค่าต่างๆ ของภาพ

4.3 Pooling Layer เป็นชั้นที่ทำการลดขนาดของภาพโดยการเลือกค่าสูงสุดหรือเฉลี่ยของพื้นที่

การจำแนกรูปภาพโดยใช้เทคนิคคอนโวลูชันเป็นวิธีการแยกคุณลักษณะของรูปภาพ (Feature extraction) ใช้การคำนวณคอนโวลูชัน ในการจำแนกรูปภาพ ซึ่งเป็นวิธีการแยกคุณลักษณะของรูปภาพ โดยใช้เทคนิคคอนโวลูชันเป็นวิธีการคำนวณ โดยใช้เทคนิคประมวลผลภาพแบบคอนโวลูชัน โครงข่ายจะแยกคุณลักษณะของรูปภาพต่างๆ โดยการตรวจจับคุณลักษณะ (Feature detector) เพื่อเรียกดูคุณลักษณะของรูปภาพในระดับที่ละเอียดขึ้น จากนั้นจะนำคุณลักษณะที่ได้ไปประมวลผลเพื่อเป็นการจำแนกรูปภาพ

งานวิจัยที่เกี่ยวข้อง

Akshai และ Anitha (2021) การเกษตรมีบทบาทสำคัญในเศรษฐกิจประเทศอินเดีย การตรวจหาโรคพืชตั้งแต่เนิ่นๆ เป็นสิ่งสำคัญมากในการป้องกันการสูญเสียพืชและการแพร่กระจายของโรคต่อไป พืชส่วนใหญ่ เช่น แอปเปิ้ล มะเขือเทศ เชอร์รี่ องุ่น แสดงอาการของโรคที่มองเห็นได้ผ่านทางใบ รูปแบบที่มองเห็นได้เหล่านี้สามารถระบุได้เพื่อทำนายโรคได้อย่างถูกต้องและดำเนินการตั้งแต่เนิ่นๆ เพื่อป้องกัน วิธีการทั่วไปคือเกษตรกรหรือนักพยาธิวิทยาพืช สังเกตใบพืชด้วยตนเอง และระบุประเภทของโรค ในงานวิจัยชิ้นนี้นำเสนอโมเดลการเรียนรู้เชิงลึกที่ได้รับการฝึกฝนเพื่อจำแนกโรคพืชต่างๆ แบบจำลองโครงข่ายประสาทเทียมแบบ Convolutional (CNN) ถูกนำมาใช้เนื่องจากความสำเร็จอย่างมากในการจำแนกประเภทตามภาพ แบบจำลองการเรียนรู้เชิงลึกให้การคาดการณ์ที่รวดเร็วและแม่นยำกว่าการสังเกตใบพืชด้วยตนเอง ในงานนี้โมเดล CNN และแบบจำลองที่ผ่านการเรียนรู้แบบ Supervised learning ล่วงหน้า เช่น โมเดล VGG ResNet และ DenseNet ได้รับการฝึกฝนโดยใช้ชุดข้อมูล ผลการวิจัยพบว่าโมเดล DenseNet มีความแม่นยำสูงสุด

จักรินทร์ (2562) การประยุกต์ใช้เทคนิคการเรียนรู้เชิงลึกสำหรับการรู้จำพรรณไม้ที่อยู่ในสิ่งแวดล้อมทางธรรมชาติ วัตถุประสงค์เพื่อเปรียบเทียบโครงข่ายประสาทเทียมแบบคอนโวลูชันจำนวน 4 โครงสร้าง ประกอบด้วย LeNet-5 AlexNet GoogLeNet และ VGGNet ข้อมูลพรรณไม้ที่นำมาใช้ในการทดสอบมีจำนวน 3 ชุดข้อมูล คือ PNE 102 Flower และ Folio สำหรับชุดข้อมูล PNE และ 102 Flower เป็นรูปภาพที่อยู่สิ่งแวดล้อมทางธรรมชาติทำให้มีพื้นที่พื้นหลังซับซ้อน สำหรับชุด Folio เป็นรูปภาพใบไม้ที่ถ่ายในห้องทดลองโดยกำหนดให้พื้นหลังของภาพเป็นสีขาว ผลการเปรียบเทียบระหว่างโครงข่ายประสาทเทียมแบบคอนโวลูชันเชิงลึกโดยใช้โครงสร้าง GoogLeNet และ VGGNet พบว่าโครงสร้างแบบ GoogLeNet มี

ประสิทธิภาพสูงสุดในชุดข้อมูล PNE และ 102 Flower และยังใช้เวลาในการเรียนรู้ที่เร็วกว่าเมื่อเทียบกับโครงสร้างแบบ VGGNet โดยใช้จำนวนรอบในการเรียนรู้ 10,000 รอบ แต่ทั้งนี้โครงสร้างแบบ VGGNet มีอัตราความถูกต้องสูงที่สุดในชุดข้อมูล Folio ซึ่งเป็นรูปภาพที่ถ่ายในห้องทดลองที่มีพื้นหลังสีขาว และใช้จำนวนรอบในการเรียนรู้เพียง 1,000 รอบ จึงสรุปได้ว่าหากต้องการจะสร้างโมเดลเพื่อนำไปใช้ในการจำแนกพรรณไม้ที่อยู่ในสิ่งแวดล้อมทางธรรมชาติ ควรใช้โครงสร้างแบบ GoogLeNet

นพรุจ และณัฐภูมิ (2563) การประมวลผลสำหรับการจำแนกคุณภาพมะม่วงพันธุ์โชคอนันต์โดยการจำลองการมองเห็นของมนุษย์ด้วยวิธีการเรียนรู้เชิงลึก โดยดำเนินการเก็บรวบรวมข้อมูลภาพถ่ายมะม่วงพันธุ์โชคอนันต์และเก็บข้อมูลการจำแนกคุณภาพกำหนดระดับคุณภาพออกเป็น 4 ระดับ ประกอบด้วยคุณภาพระดับเกรด A คุณภาพระดับเกรด B คุณภาพระดับเกรด C และคุณภาพระดับเกรด D คือมะม่วงเน่า ผลของการวิจัยแสดงให้เห็นว่าการใช้วิธีการเรียนรู้เชิงลึกด้วยอัลกอริทึมโครงข่ายประสาทเทียมแบบสังวัตนาการ (Convolutional Neural Network: CNN) ในการประมวลผลภาพเพื่อสร้างแบบจำลองแสดงค่าความแม่นยำสูงสุดคือ 99.79% จากนั้นนำแบบจำลองพัฒนาเป็นระบบต้นแบบสำหรับจำแนกคุณภาพมะม่วงพันธุ์โชคอนันต์พบว่าอัตราความสำเร็จในการจำแนกคือ 100%

วีระยุทธ และคณะ (2564) การจำแนกความสูงของกล้วยฉาบด้วยเทคนิคการประมวลผลภาพ โดยมีวัตถุประสงค์เพื่อนำเสนอวิธีการจำแนกกล้วยฉาบจากภาพถ่าย โดยวิเคราะห์จากคุณลักษณะของสี ได้แก่ แบบจำลอง RGB และ HSV ซึ่งเป็นประโยชน์อย่างมากสำหรับเครื่องทอดกล้วยกึ่งอัตโนมัติ ที่ขาดการนำระบบอัตโนมัติเข้าไปใช้ในการจำแนกชนิดของสีกล้วยฉาบ ผลงานวิจัยพบว่า ความถูกต้องในการตรวจจับสี 86.27% ค่าความครบถ้วน (recall) 83.06% และค่าความแม่นยำ (precision) 84.06% และทดสอบประสิทธิภาพการจำแนกกล้วยฉาบโดยใช้เทคนิคการประมวลผลภาพ

Chellapandi *et al.* (2021) ปัญญาประดิษฐ์ได้รับการพิสูจน์แล้วว่าเป็นประโยชน์อย่างมากในเกือบทุกภาคส่วนของอุตสาหกรรม ในช่วงไม่กี่ครั้งที่ผ่านมามีความต้องการอาหารเพิ่มขึ้นในขณะที่อุปทานยังขาดอยู่ เพื่อตอบสนองความต้องการที่เพิ่มขึ้นเหล่านี้ การป้องกันและตรวจหาโรคพืชในระยะเริ่มต้นเป็นมาตรการบางอย่างที่ต้องปลูกฝังในการทำฟาร์มเกษตรเพื่อดูแลพืชในระยะแรกและจะช่วยลดการสูญเสียโดยรวม ในงานวิจัยนี้ผู้วิจัยได้ใช้แบบจำลองการเรียนรู้เชิงลึกและถ่ายทอดแบบจำลองตามการเรียนรู้ เพื่อจำแนกภาพใบพืชที่เป็นโรคออกเป็นโรคพืชจำนวน 38 ประเภท และใช้ Transfer learning ทั้งหมด 8 โมเดล ที่ผ่านการฝึกอบรมแล้ว ได้แก่ VGG16 VGG19 ResNet50 InceptionV3 InceptionResnetV2 MobileNet MobileNetV2 และ DenseNet along ด้วยแบบจำลองที่สร้างขึ้นเองแบบเดียวถูกนำมาใช้ในการศึกษางานวิจัยนี้พบว่า DenseNet ให้ผลลัพธ์ที่ดีที่สุดในการทดสอบด้วยความแม่นยำ 99%

Rajan และ Noumida (2021) การกักตุนการระบุนกหลายสายพันธุ์จากการบันทึกเสียง วิธีการที่เสนอใช้หนึ่งในเครือข่ายประสาทเทียม Deep Convolutional ที่ได้เรียนรู้แบบ Supervised learning มาล่วงหน้า (DCNN) แบบจำลอง VGG-16 เพื่อเรียนรู้การแปลงเสียงของนกผ่านการวิเคราะห์ Sliding window บน Melspectrogram งานวิจัยได้ใช้กลยุทธ์การรวมเพื่อตัดสินใจเกี่ยวกับไฟล์ทดสอบที่เอาต์พุต Sigmoid ถูกรวบรวมและทำให้เป็นมาตรฐาน นกที่มีคะแนนความน่าจะเป็นสูงสุดจะถือว่าเป็นนกที่มีอยู่ในการบันทึกเสียง วิธีการที่เสนอได้รับการประเมินในฐานข้อมูลเสียงนก Xeno-canto ผู้วิจัยใช้การเรียกนกจาก 10 สายพันธุ์ที่แตกต่างกัน Mel-spectrograms (คุณสมบัติภาพ) ที่สร้างขึ้นจากการเรียกนกถูกใช้เป็นอินพุตสำหรับ VGG-16 ประสิทธิภาพยังถูกเปรียบเทียบกับแนวทาง MFCC-DNN ผลการวิจัยได้คะแนน F1 เฉลี่ย 0.65 และมีประสิทธิภาพดีกว่าแนวทาง MFCC-DNN ที่ใช้ Acoustic cue-based

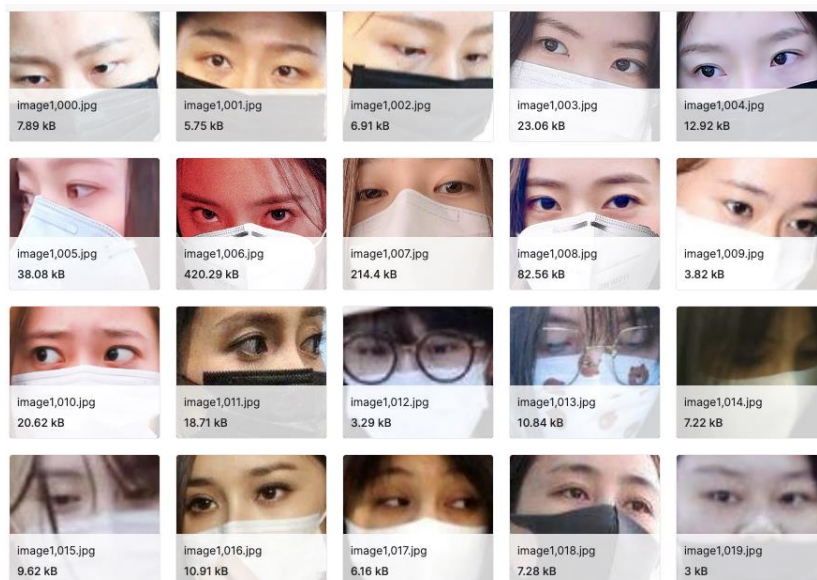
Shrivastava และ Pradhan (2020) ศึกษาวิจัยเรื่องข้าวเป็นอาหารหลักและมีบทบาทสำคัญในการสร้างความมั่นคงทางอาหารทั่วโลก อย่างไรก็ตามโรคหลายชนิดส่งผลกระทบต่อพืชชนิดนี้ซึ่งช่วยลดการผลิตและคุณภาพลงอย่างมาก ดังนั้นการตรวจหาโรคตั้งแต่เนิ่นๆ จึงเป็นงานที่จำเป็นมากในการป้องกันการแพร่กระจายของโรค ดังนั้น เป็นที่ต้องการอย่างมากในการพัฒนาระบบอัตโนมัติซึ่งจะช่วยให้นักปฐพีวิทยา นักพยาธิวิทยาและแม้แต่เกษตรกร ในการวินิจฉัยโรคข้าวอย่างมีประสิทธิภาพ

ประสิทธิภาพมากขึ้น และใช้มาตรการป้องกันได้ทันเวลา ในยุคปัจจุบันของปัญญาประดิษฐ์ขั้นสูงมีการสำรวจเทคนิคการเรียนรู้ที่หลากหลายสำหรับการจำแนกโรคพืชข้าว ในเทคนิคการเรียนรู้ของเครื่องต่างๆ การเรียนรู้เชิงลึกถูกนำไปใช้อย่างกว้างขวางในโดเมนต่างๆ ของคอมพิวเตอร์วิทัศน์และการวิเคราะห์ภาพ เมื่อเร็วๆ นี้ ประสบความสำเร็จในการให้ความถูกต้องที่มีแนวโน้มและมีศักยภาพสูง อย่างไรก็ตามการฝึกอบรมโมเดลการเรียนรู้เชิงลึกตั้งแต่เริ่มต้นต้องใช้ข้อมูลที่มีป้ายกำกับจำนวนมาก และการรวบรวมข้อมูลที่มีป้ายกำกับขนาดใหญ่ที่มีต้นทุนสูง กระบวนการที่ลำบากและใช้เวลา การถ่ายโอนการเรียนรู้ของรูปแบบการเรียนรู้เชิงลึกที่ได้รับการฝึกฝนล่วงหน้าเป็นเทคนิคในการเอาชนะปัญหาดังกล่าว และได้สำรวจประสิทธิภาพของโมเดล CNN เชิงลึกที่ได้รับการฝึกฝนมาเป็นอย่างดี เช่น AlexNet Vgg16 ResNet152V2 InceptionV3 InceptionResNetV2 Xception MobileNet DenseNet169 NasNetMobile และ NasNetLarge สำหรับการจำแนกโรคพืชข้าวตามภาพชุดข้อมูลที่ใช้ในงานวิจัยนี้ประกอบด้วยภาพโรคพืชข้าว จำนวน 1,216 ภาพ และภาพเหล่านี้ถูกรวบรวมจากพื้นที่เกษตรกรรมจริงที่มี จำนวน 7 คลาส ดังนี้ Rice blast Bacterial leaf blight Brown spot Sheath blight Sheath rot False smut และ Healthy leaves รุ่น Vgg16 ให้ความแม่นยำในการจำแนกประเภทสูงสุด 93.11% ผลลัพธ์ของแบบจำลองสามารถใช้เป็นที่ปรึกษาและเป็นเครื่องมือตรวจจับเบื้องต้นในโดเมนการเกษตรจริง

วิธีการดำเนินการวิจัย

1. การเตรียมข้อมูล

งานวิจัยได้ดำเนินการเก็บรวบรวมข้อมูลใบหน้าสวมใส่หน้ากากและไม่สวมใส่หน้ากากอนามัยใช้ดาต้าเซตจากเว็บไซต์ <https://www.kaggle.com/datasets/aneerbranchakraborty/face-mask-detection-data> โดยแต่ละภาพจะมีขนาด 150 x 150 พิกเซล จำนวน 3,833 ภาพ แบ่งเป็นภาพใบหน้าสวมใส่หน้ากาก จำนวน 1,915 ภาพ และภาพใบหน้าไม่สวมใส่หน้ากาก จำนวน 1,918 ภาพ จากนั้นได้ทำการประมวลผลภาพแล้วทำการติดฉลากจำแนกประเภท โดยใช้งานไลบรารี Open CV2 ปรับขนาดรูปภาพแล้วดำเนินการแปลงรูปภาพเป็นสี RGB 3 ช่องสัญญาณและจัดเก็บข้อมูลในรูปแบบ Matrix ดังรูปที่ 2 และรูปที่ 3



รูปที่ 2 ภาพใบหน้าสวมใส่หน้ากาก

ผลการวิจัยและวิจารณ์ผล

1. การศึกษาและรวบรวมข้อมูล

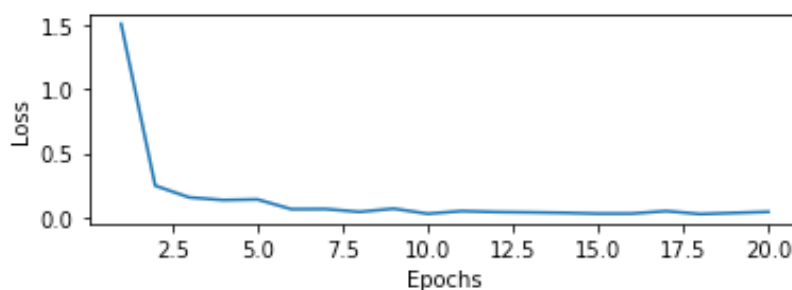
งานวิจัยนี้ได้ศึกษาและรวบรวมข้อมูลใบหน้าคนสวมหน้ากาก และไม่สวมหน้ากากอนามัยโดยใช้เทคนิคการประมวลผลภาพแล้วทำการติดฉลากจำแนกประเภท โดยผู้วิจัยได้ใช้งานไลบรารี Open CV2 ปรับขนาดรูปภาพแล้วดำเนินการแปลงรูปภาพเป็นสี RGB และจัดเก็บข้อมูลในรูปแบบ Matrix และกำหนดให้ภาพมีขนาด 150 x 150 พิกเซล จำนวน 3,833 ภาพ แบ่งเป็นภาพใบหน้าสวมใส่หน้ากาก จำนวน 1,915 ภาพ และภาพใบหน้าไม่สวมใส่หน้ากาก จำนวน 1,918 ภาพ

2. ผลการเปรียบเทียบประสิทธิภาพโมเดลส่งถ่ายความรู้

ผลการเปรียบเทียบประสิทธิภาพโมเดลส่งถ่ายความรู้ ทั้ง 3 โมเดล ผู้วิจัยได้ทำการเรียนรู้จำนวน 20 รอบ พบว่าโมเดล Vgg16 ให้ผลลัพธ์ค่าความถูกต้อง 96% และได้ค่า Loss 0.07% และโมเดล Inceptionresnetv2 ให้ผลลัพธ์ค่าความถูกต้อง 97% และได้ค่า Loss 0.08% และพบว่าโมเดล ResNet152V2 มีค่าความถูกต้องมากที่สุด 98% และค่า Loss 0.08% ดังตารางที่ 1 และแสดงโครงสร้างสถาปัตยกรรมของ ResNet152V2 ดังรูปที่ 4 กราฟแสดงผลค่าผิดพลาด (Loss) ที่เกิดจากการสร้างแบบจำลอง ResNet152V2 ดังรูปที่ 7

ตารางที่ 1 ประสิทธิภาพของโมเดลส่งถ่ายความรู้

Deep Transfer Learning	จำนวน Epoch	Loss (%)	Accuracy (%)
Vgg16	20	<u>0.07</u>	96
Inceptionresnetv2	20	0.08	97
<u>ResNet152V2</u>	20	0.08	<u>98</u>



รูปที่ 7 แสดงผลการวัดค่า Loss ของแบบจำลองโมเดล ResNet152V2

3. การนำแบบจำลองไปใช้งาน

จากการทดสอบแบบจำลองโมเดลการส่งถ่ายความรู้จาก ImageNet 3 โมเดล ได้แก่ Vgg16 Inceptionresnetv2 และ ResNet152V2 ซึ่งงานวิจัยนี้ได้นำเสนอโมเดลที่ให้ผลลัพธ์ดีที่สุดคือ ResNet152V2 โดยใช้ภาษา Python เรียกใช้ไลบรารีจาก Tensorflow และ Keras จากนั้นได้นำโมเดลแบบจำลองไปทดสอบกับชุดข้อมูลทดสอบภาพคนไม่สวมหน้ากากอนามัย จำนวน 833 ภาพ และพบว่าโมเดล ResNet152V2 มีค่าความถูกต้องมากที่สุด 98% และค่า Loss 0.08% ดังตารางที่ 1

อภิปรายผล

งานวิจัยนี้ได้ทำการเตรียมข้อมูลเรียนรู้และชุดข้อมูลทดสอบโดยใช้เทคนิคการประมวลผลภาพแล้วทำการติดฉลากจำแนกประเภท โดยผู้วิจัยได้ใช้งานไลบรารี Open CV2 ปรับขนาดรูปภาพแล้วดำเนินการแปลงรูปภาพเป็นสี RGB จำนวน 3,833 ภาพ แบ่งเป็นภาพใบหน้าสวมใส่หน้ากาก จำนวน 1,915 ภาพ และภาพใบหน้าไม่สวมใส่หน้ากาก จำนวน 1,918 ภาพ

ผู้วิจัยได้ทดลองเปรียบเทียบประสิทธิภาพโมเดลส่งถ่ายความรู้ทั้ง 3 โมเดล ผู้วิจัยได้ทำการเรียนรู้จำนวน 20 รอบ พบว่า โมเดล Vgg16 ให้ผลลัพธ์ค่าความถูกต้อง 96% และได้ค่า Loss 0.07% และโมเดล Inceptionresnetv2 ให้ผลลัพธ์ค่าความถูกต้อง 97% และได้ค่า Loss 0.08% และพบว่าโมเดล ResNet152V2 มีค่าความถูกต้องมากที่สุด 98% และค่า Loss 0.08% ดังตารางที่ 1

จากผลการทดลองในงานวิจัยครั้งนี้ผู้วิจัยได้ใช้ โครงข่ายประสาทเทียมส่งถ่ายความรู้ เป็นการใช้ความรู้ที่ได้จากโมเดลที่เรียนรู้ไว้ในการเรียนรู้โมเดลใหม่ ซึ่งจะช่วยให้สามารถเรียนรู้ภายในเวลาน้อยกว่าโดยไม่ต้องเรียนรู้จากข้อมูลใหม่เสมอในกรณีที่ต้องการใช้โมเดลประมวลผลภาพในการจำแนกรูปภาพ และสามารถนำมาสร้างแบบจำลองโมเดลเพื่อคัดแยกและจำแนกข้อมูล โดยผู้วิจัยได้ดำเนินการเก็บรวบรวมข้อมูลภาพข้อมูลใบหน้าคนสวมหน้ากากและไม่สวมหน้ากากอนามัย ดังรูปที่ 2 และรูปที่ 3 จากนั้นได้ใช้เทคนิคการประมวลผลภาพแล้วทำการติดฉลากจำแนกประเภท โดยผู้วิจัยได้ใช้ไลบรารี Open CV2 ปรับขนาดรูปภาพแล้วดำเนินการแปลงรูปภาพเป็นสี RGB และจัดเก็บข้อมูลในรูปแบบ Matrix เพื่อสะดวกในการประมวลผลภาพ

จากนั้นผู้วิจัยได้ทดลองเปรียบเทียบประสิทธิภาพโมเดลส่งถ่ายความรู้ทั้ง 3 โมเดล ได้แก่ Vgg16 Inceptionresnetv2 และ ResNet152V2 ผู้วิจัยได้ทำการเรียนรู้จำนวน 20 รอบ พบว่าโมเดล ResNet152V2 มีค่าความถูกต้องมากที่สุด 98% และค่า Loss 0.08% ดังตารางที่ 1 แสดงให้เห็นถึงประสิทธิภาพของแบบจำลองโมเดลถ่ายโอนความรู้ ResNet152V2 สามารถจำแนกการสวมหน้ากากอนามัยโดยเทคนิคการส่งถ่ายความรู้ โดยสามารถตรวจจับใบหน้าได้ถูกต้อง 98%

ข้อเสนอแนะในการวิจัยครั้งต่อไป

- 1) ควรเตรียมข้อมูลเรียนรู้ให้มีความหลากหลายมากยิ่งขึ้น
- 2) ควรมีการเปรียบเทียบแบบจำลองโมเดลจากหลากหลายโครงข่ายเพื่อให้เห็นประสิทธิภาพของโมเดลมากยิ่งขึ้น
- 3) สามารถนำแบบจำลองโมเดลไปพัฒนาต่อในด้านการตรวจจับ Mask แบบเรียลไทม์ด้วยอุปกรณ์อินเทอร์เน็ทของสรรพสิ่ง

เอกสารอ้างอิง

- กรมควบคุมโรค. (2565). โรคติดเชื้อไวรัสโคโรนาสายพันธุ์ใหม่ 2019. แหล่งข้อมูล: <https://ddc.moph.go.th/viralpneumonia>. ค้นเมื่อวันที่ 29 กันยายน 2565.
- จักรินทร์ สุนกแสน. (2562). การประยุกต์ใช้เทคนิคการเรียนรู้เชิงลึกสำหรับการรู้จำพรรณไม้ที่อยู่ในสิ่งแวดล้อมทางธรรมชาติ. วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต, มหาวิทยาลัยมหาสารคาม. มหาสารคาม.
- นพรุจ พัฒนสาร และณัฐภูมิ ศรีวิบูลย์. (2563). การประมวลผลภาพสำหรับการจำแนกคุณภาพมะม่วงพันธุ์คอนันต์โดยการจำลองการมองเห็นของมนุษย์ด้วยวิธีการเรียนรู้เชิงลึก. วารสารวิทยาการและเทคโนโลยีสารสนเทศ 10(1): 24 – 29.
- วีระยุทธ สุริคำ, นพพล อ่ำดี และชัยธำรง พงษ์พัฒนศิริ. (2564). การจำแนกความสูงของกล้วยฉาบด้วยเทคนิคการประมวลผลภาพ. วารสารวิชาการเทคโนโลยีอุตสาหกรรมและวิศวกรรม มหาวิทยาลัยราชภัฏวชิรวิทยาดง 3(1): 55 – 67.

- Akshai, K. and Anitha, J. (2021). Plant disease classification using deep learning. In: 2021 3rd International Conference on Signal Processing and Communication (ICPSC), Coimbatore. doi: 10.1109/icpsc51351.2021.94.
- Chellapandi, B., Vijayalakshmi, M. and Chopra, S. (2021). Comparison of Pre-Trained Models Using Transfer Learning for Detecting Plant Disease. In: 2021 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), Greater Noida doi: 10.1109/icccis51004.2021.9397098.
- Color. (2022). Color. Source: <http://msdn.microsoft.com/en-us/library/windows/desktop/aa511283.aspx>. Retrieved from 4 February 2022.
- Gonzalez R.C. and Woods R.E. (2006). Digital Image Processing (3rd Edition). New Jersey: Prentice-Hall, Inc.
- Owens R. (2022). Binary Images. Source: http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/OWENS/LECT2/node3.html. Retrieved from 20 February 2022.
- Rajan, R. and Noumida, A. (2021). Multi-label Bird Species Classification Using Transfer Learning. In: 2021 International Conference on Communication. Control and Information Sciences (ICCISc), Idukki. doi: 10.1109/iccisc52257.2021.9484858.
- Shrivastava, V.K. and Pradhan, M.K. (2020). Rice plant disease classification using color features: a machine learning paradigm. Journal of Plant Pathology 103:17 - 26.
- Yann, L. C., Bottou, L., Bengio, Y. and Haffner, P. (1998) Gradient-Based Learning Applied to Document Recognition. Proceedings of the IEEE 86(1): 2278 - 23. doi: 10.1109/5.726791.

