

ระบบตรวจจับวัตถุและจำแนกแยกแยะประเภทของสินค้าจากภาพถ่ายด้วยการยืนยันผ่าน
เทคนิค OCR ร่วมกับเทคนิคการเรียนรู้ข้อมูลที่ไม่สมดุล
Object Detection and Classification System for Product Differentiation Using
OCR and Imbalanced Data Learning Techniques

ฐากร พุทธรักษ์¹ และ นพวรรณ ชื่นอารมณ²

¹สาขาวิชาวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยสุโขทัยธรรมาธิราช, E-Mail takorn.pre@stou.ac.th

²สาขาวิชาวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยสุโขทัยธรรมาธิราช, E-Mail noppawan.chu@stou.ac.th

Takorn Prexawanprasut¹ and Noppawan Chuenarom²

¹School of Science and Technology, Sukhothai Thammathirat Open University, E-Mail
takorn.pre@stou.ac.th

²School of Science and Technology, Sukhothai Thammathirat Open University, E-Mail
noppawan.chu@stou.ac.th

วันที่รับบทความ 27 พฤษภาคม 2567
วันที่แก้ไขบทความ 7 มิถุนายน 2567
วันที่ตอบรับบทความ 14 มิถุนายน 2567

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาแนวทางและเพิ่มประสิทธิภาพของการจำแนกและตรวจจับวัตถุในตู้คอนเทนเนอร์ระหว่างพิธีการทางศุลกากรโดยใช้ภาพถ่ายบรรจุภัณฑ์ จากรูปภาพ 5,113 รูปที่เก็บจากบริษัทขนส่งสินค้าระหว่างประเทศ 3 แห่งในกรุงเทพฯ โดยใช้แบบจำลองการเรียนรู้ของเครื่องเพื่อจำแนกสินค้าเป็น 5 ประเภท พบว่าอัลกอริทึม YOLOv4 มีความแม่นยำสูงสุดเมื่อไม่ใช้เทคนิค Resampling ในขณะที่เทคนิค Resampling เช่น SMOTE และ Borderline SMOTE ช่วยเพิ่มประสิทธิภาพการจำแนกในกรณีข้อมูลไม่สมดุล เทคนิค Undersampling ไม่เพิ่มความแม่นยำเนื่องจากทำให้ข้อมูลสำคัญสูญหาย การผสมผสานการตรวจจับวัตถุและ OCR ช่วยให้ระบบสามารถระบุประเภทสินค้าจากภาพถ่ายได้แม่นยำและรวดเร็วมากขึ้น กลไกที่เสนอในงานวิจัยนี้จะช่วยลดความเสี่ยงในการลักลอบนำเข้าสินค้าผิดกฎหมายและเพิ่มประสิทธิภาพในการดำเนินการกระบวนการพิธีการทางศุลกากร

คำสำคัญ: การจำแนกวัตถุจากรูปภาพ, การตรวจจับวัตถุจากรูปภาพ, การเรียนรู้ของเครื่อง, การเรียนรู้จากข้อมูลที่ไม่สมดุล

ABSTRACT

This research develops a method for classifying and detecting objects within container shipments during customs procedures using package images. Using 5,113 images from three international freight forwarding companies in Bangkok, we employed machine learning models to classify products into five categories. The study found that the YOLOv4 algorithm achieved the highest accuracy without resampling techniques. However, resampling techniques such as SMOTE and Borderline SMOTE significantly improved classification performance in cases of imbalanced data. Undersampling did not enhance accuracy, as it led to the loss of crucial information. The integration of object detection and OCR verification enabled the system to accurately and quickly identify product types from images. The proposed method can reduce the risk of illegal imports and enhance efficiency in customs procedures.

KEYWORDS: Object classification, Object recognition, Machine learning, Imbalanced learning

1. บทนำ

ปัจจุบันประเทศไทยพึ่งพาสินค้านำเข้าจากประเทศจีนเป็นจำนวนมาก เนื่องจากจีนเป็นหนึ่งในคู่ค้าหลักและเป็นแหล่งผลิตสินค้าที่มีราคาถูกและหลากหลาย การนำเข้าสินค้าจากจีนครอบคลุมตั้งแต่สินค้าทั่วไปไปจนถึงอุปกรณ์อิเล็กทรอนิกส์และสินค้าที่เป็นวัตถุดิบทางอุตสาหกรรม การจัดการด้านพิธีการทางศุลกากรของสินค้านำเข้าจากจีนจึงเป็นเรื่องสำคัญและมีความซับซ้อนสูง ซึ่งประกอบไปด้วยขั้นตอนที่หลายหลายและต้องการความร่วมมือระหว่างหลายฝ่าย รวมถึงเจ้าหน้าที่ศุลกากร ผู้นำเข้า ผู้ประกอบการขนส่ง และตัวแทนศุลกากร กระบวนการเหล่านี้กระทำเพื่อให้แน่ใจว่าสินค้านำเข้ามานั้นถูกต้องตามกฎหมาย ไม่มีการลักลอบนำเข้าสินค้าผิดกฎหมายหรือสินค้าที่อาจเป็นอันตราย ซึ่งเป็นส่วนสำคัญในการรักษาความปลอดภัยและการปกป้องระบบเศรษฐกิจภายในประเทศ

กระบวนการด้านพิธีการทางศุลกากรและการนำสินค้าออกจากท่าเรือเริ่มจากการแจ้งเตือนการมาถึงของสินค้าเมื่อเรือบรรทุกสินค้ามาถึงท่าเรือ ผู้ประกอบการขนส่งหรือสายการบินจะต้องแจ้งเตือนหน่วยงานศุลกากรถึงการมาถึงของสินค้าและรายละเอียดต่างๆ จากนั้นจะมีการตรวจสอบเอกสารที่จำเป็น เช่น ใบขนสินค้านำเข้า ใบกำกับสินค้า และรายการบรรจุภัณฑ์ เพื่อให้แน่ใจว่าข้อมูลถูกต้องและครบถ้วน ต่อมาเจ้าหน้าที่ศุลกากรจะทำการตรวจสอบสินค้า ซึ่งอาจมีการสุ่มตรวจหรือการตรวจสอบอย่างละเอียดเพื่อให้แน่ใจว่าสินค้าตรงกับเอกสาร การตรวจสอบนี้อาจรวมถึงการใช้เครื่องเอกซเรย์หรือเครื่องสแกนสินค้า หลังจากนั้นผู้นำเข้าจะต้องชำระภาษีและอากรที่เกี่ยวข้องตามประเภทและมูลค่าของสินค้า เมื่อทุกขั้นตอนเสร็จสิ้นและภาษีถูกชำระเรียบร้อยแล้ว ศุลกากรจะออกใบปล่อยสินค้า ซึ่งอนุญาตให้ผู้นำเข้าหรือผู้ประกอบการขนส่งนำสินค้าออกจากท่าเรือได้

โดยปกติ บริษัทที่ทำหน้าที่เป็นตัวแทนการขนส่งสินค้านานาชาติระหว่างประเทศ (Freight Forwarder) จะมีเจ้าหน้าที่ประจำอยู่ที่ท่าเรือเพื่อสังเกตการณ์ตู้สินค้าที่กำลังจะถูกเปิดเพื่อตรวจสอบ และทำการตรวจสอบความถูกต้องของสินค้าที่อยู่ในตู้ให้ถูกต้องตรงกับข้อมูลที่บริษัทแจ้งมา ซึ่งอาจทำได้โดยการเปิดหีบห่อและนับจำนวน หรือใช้ประสบการณ์ของเจ้าหน้าที่ในการทำนายสินค้าที่อยู่ในหีบห่อ การกระทำดังกล่าวเป็นการยืนยันความถูกต้องก่อนที่เจ้าหน้าที่ของบริษัทจะนำเอกสารต่างๆ ไปดำเนินการในขั้นตอนถัดไปเพื่อให้พิธีการทางศุลกากรเป็นไปอย่างราบรื่น อย่างไรก็ตาม กระบวนการตรวจสอบสินค้านี้ใช้เวลาอย่างมาก ซึ่งอาจส่งผลให้ไม่สามารถนำสินค้าออกจากท่าเรือได้ตรงตามเวลาที่ลูกค้ากำหนด

ผู้วิจัยได้ดำเนินการศึกษาข้อมูลการนำเข้าสินค้าจากบริษัทตัวแทนการขนส่งสินค้านานาชาติระหว่างประเทศจำนวน 3 บริษัทพบว่าในปัจจุบัน ทั้ง 3 บริษัทมีระบบที่สามารถวิเคราะห์สินค้าจากภาพถ่ายของหีบห่อสินค้าที่อยู่ในตู้ด้วยการใช้เทคโนโลยีในการตรวจสอบและวิเคราะห์ข้อมูล ซึ่งสามารถเพิ่มประสิทธิภาพและความแม่นยำในการตรวจสอบสินค้าและทำให้กระบวนการนำสินค้าออกจากท่าเรือให้มีความรวดเร็วและปลอดภัยมากยิ่งขึ้น ช่วยลดปริมาณเวลาการทำงานและลดความผิดพลาดที่อาจเกิดจากการใช้ดุลพินิจของเจ้าหน้าที่ที่ทำงานได้อย่างมีนัยสำคัญ อย่างไรก็ตาม ประสิทธิภาพของแบบจำลองจะด้อยลงสำหรับกรณีสินค้าที่มีข้อมูลจำนวนน้อย อัลกอริทึมการเรียนรู้ของเครื่องจะเผชิญปัญหาการเรียนรู้แบบไม่สมดุล (Imbalanced Learning) ทำให้ไม่สามารถสร้างแบบจำลองที่เข้าใจแนวคิดของปัญหาของสินค้านำเข้าได้อย่างเต็มที่ การวิจัยขั้นนี้จึงมีวัตถุประสงค์เพื่อเพิ่มความสามารถของแบบจำลองการทำนายประเภทสินค้าในตู้คอนเทนเนอร์จากการวิเคราะห์รูปภาพของหีบห่อในกรณีที่อัลกอริทึมการเรียนรู้เผชิญกับปัญหาการเรียนรู้ที่ไม่สมดุล ซึ่งจะช่วยให้การนำสินค้าออกจากท่าเรือมีความสะดวกและมีประสิทธิภาพมากยิ่งขึ้นในทุกๆ ชนิดและประเภทสินค้า ลดความเสี่ยงในการลักลอบนำเข้าสินค้าผิดกฎหมายและเพิ่มความเร็วในการตรวจสอบและปล่อยสินค้าได้

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 เทคนิคและวิธีการที่เกี่ยวข้องกับการประมวลผลรูปภาพ

Prexawanprasut et al.(2566) พัฒนาเทคนิคในการระบุผลิตภัณฑ์โดยใช้ภาพบรรจุภัณฑ์เป็นพื้นฐาน โดยแบ่งชนิดของผลิตภัณฑ์เป็น 5 หมวดหมู่ ซึ่งรวมถึงสินค้าทางการแพทย์ อุปกรณ์ก่อสร้าง เอกสาร อุปกรณ์ไฟฟ้า และสินค้าเพิ่มเติม งานวิจัยดังกล่าวได้นำรูปภาพพัสดุทั้งหมด 5,675 รูปมาใช้สร้างแบบจำลองในการแยกแยะลักษณะพื้นฐานและการทดสอบเรียนรู้ของเครื่อง โดยทุกภาพถูกกำหนดชื่อหมวดหมู่โดยเจ้าหน้าที่ไว้ล่วงหน้า เพื่อระบุว่าสิ่งของที่อยู่ภายในอยู่ในหมวดหมู่ใด การพัฒนาแบบจำลองดังกล่าวใช้เทคนิค Convolutional Neural Network (CNN) ซึ่งเป็นโครงข่ายประสาทเทียมที่นิยมใช้ในการจำแนกข้อมูลประเภทรูปภาพ ไฟล์วิดีโอ และไฟล์เสียง มาเป็นเทคนิคพื้นฐานที่ใช้ในกระบวนการจัดหมวดหมู่ในงานวิจัย

นอกจากนั้น การทดลองยังใช้กระบวนการการทดสอบแบบ 10-fold cross-validation เพื่อช่วยยืนยันผลการทดลอง และใช้เทคนิค OCR เพื่อช่วยในการตรวจสอบยืนยันข้อมูลในกรณีที่เทคนิค CNN ไม่สามารถจำแนกแยกแยะได้

การสร้างแบบจำลอง Convolutional Neural Network (CNN) (LeCun et al., 1998) ที่มีชั้นเอาต์พุตที่เชื่อมต่อเต็มรูปแบบ (fully connected) นั้นเป็นเรื่องที่ยาก เนื่องจากความยาวของชั้นเอาต์พุตไม่ได้ถูกกำหนดไว้ล่วงหน้าตามกล่องขอบเขต (bounding boxes) ที่แทนสิ่งที่น่าสนใจต่างๆ ในรูปภาพ วิธีการที่ง่ายที่สุดในการแก้ปัญหานี้คือการสกัดพื้นที่ที่สนใจจากภาพและจัดประเภทวัตถุที่อยู่ภายในด้วยเทคนิค CNN อย่างไรก็ตาม ปัญหาที่เกิดขึ้นกับวิธีการนี้คือวัตถุที่น่าสนใจภายในภาพอาจมีตำแหน่งที่ตั้งและอัตราส่วนของพื้นที่ที่แตกต่างกัน อัลกอริทึมเช่น R-CNN, YOLO จึงถูกพัฒนาขึ้นมาเพื่อแก้ไขปัญหานี้ โดย R-CNN ย่อมาจาก Region-based Convolutional Neural Network คือ การใช้ CNN ในการระบุและจำแนกพื้นที่ที่มีสิ่งที่น่าสนใจภายในภาพ (Region of Interest) จากนั้นใช้ CNN ในการจำแนกวัตถุภายในแต่ละพื้นที่ วิธีการนี้สามารถใช้ในการตรวจจับวัตถุและจำแนกวัตถุในภาพได้อย่างมีประสิทธิภาพ ในขณะที่เทคนิค YOLO (You Only Look Once) เป็นแบบจำลองการตรวจจับวัตถุที่พัฒนาขึ้นเพื่อระบุวัตถุในภาพโดยการคาดการณ์พื้นที่และประเภทของวัตถุไปพร้อมกัน โดย YOLO จะสามารถทำงานได้อย่างรวดเร็วและมีประสิทธิภาพสูงในการตรวจจับวัตถุในภาพที่มีการเปลี่ยนแปลงอย่างรวดเร็ว

R-CNN (Girshick et al., 2014) เป็นอัลกอริทึมที่ใช้ในการตรวจจับตำแหน่งของวัตถุ และแบ่งแยกออกจากกัน (segmentation) โดยใช้เทคโนโลยี Convolutional Neural Networks (CNN) เป็นพื้นฐาน อัลกอริทึม R-CNN ประกอบด้วยการทำงานสามส่วนคือ 1) สร้างและสกัดพื้นที่ที่อยู่ในความสนใจที่เป็นอิสระต่อหมวดหมู่ และสร้างกล่องขอบเขตที่เป็นไปได้ของวัตถุขึ้นนั้นๆ 2) สกัดลักษณะสำคัญจากแต่ละพื้นที่ที่เป็นไปได้ โดยใช้ Convolutional Neural Network เชิงลึก และ 3) จำแนกลักษณะสำคัญของคลาสที่นำมาเรียนรู้ โดยใช้แบบจำลอง Linear SVM classifier อย่างไรก็ตาม R-CNN ต้องการการพัฒนาและการทำงานของแบบจำลองที่แตกต่างกันอย่างชัดเจน นอกจากนี้ การเรียนรู้เชิงลึกบนหลายพื้นที่เป็นกระบวนการที่ใช้เวลานาน ดังนั้น R-CNN จึงลดประสิทธิภาพลงเมื่อใช้กับชุดข้อมูลที่มีขนาดใหญ่มาก

Fast R-CNN (Fast Region-based Convolutional Neural Network) (Ren et al., 2015) เป็นอัลกอริทึมที่ใช้ในการตรวจจับวัตถุในภาพ โดยพัฒนาต่อมาจาก R-CNN เพื่อเพิ่มประสิทธิภาพและลดเวลาในการประมวลผล โดย Fast R-CNN ทำงานโดยการรับภาพและ Region Proposals ที่สร้างจากขั้นตอนการตรวจจับเบื้องต้นเป็นข้อมูลนำเข้า จากนั้นใช้เครือข่าย CNN เชิงลึกเพื่อสกัดคุณลักษณะ (features) จากภาพ หลังจากนั้นมีการใช้เลเยอร์พิเศษที่เรียกว่า RoI (Region of Interest) Pooling Layer เพื่อสกัดคุณลักษณะที่เฉพาะเจาะจงสำหรับแต่ละ Region Proposal ต่อจากนั้นคุณลักษณะเหล่านี้จะถูกส่งผ่าน Fully Connected Layers เพื่อจำแนกประเภทของวัตถุและกำหนดตำแหน่งกล่องขอบเขต (bounding box) ของวัตถุ Fast R-CNN สามารถเรียนรู้และให้ผลลัพธ์ได้เร็วกว่า R-CNN แบบดั้งเดิม เนื่องจากไม่ต้องทำการคำนวณ CNN ซ้ำๆ สำหรับแต่ละ Region Proposal นอกจากนี้ยังใช้การฝึกแบบ End-to-End ซึ่งทำให้แบบจำลองสามารถปรับตัวได้ดีขึ้นและมีความแม่นยำในการตรวจจับวัตถุที่สูงขึ้น Fast R-CNN ได้รับการพัฒนาต่อเนื่องเพื่อแก้ปัญหาค่าการประมวลผลที่ซับซ้อนของ R-CNN โดยการรวมขั้นตอนการสกัดคุณลักษณะและการจำแนกประเภทเข้าด้วยกัน ทำให้การทำงานมีประสิทธิภาพมากยิ่งขึ้น

Faster R-CNN (Faster Region-based Convolutional Neural Network) (Ren et al., 2017) เป็นอัลกอริทึมที่ใช้สำหรับการตรวจจับวัตถุในภาพอย่างรวดเร็วและมีประสิทธิภาพ โดยได้รับการพัฒนาเพื่อลดข้อจำกัดของ R-CNN และ Fast R-CNN ด้วยการรวมขั้นตอนการสร้าง Region Proposals และการจำแนกวัตถุเข้าด้วยกันในแบบจำลองเดียวกัน ทำให้การตรวจจับวัตถุทำได้เร็วขึ้นมาก Faster R-CNN ประกอบด้วยสองส่วนหลัก คือ Region Proposal Network (RPN) ที่ใช้ในการสร้าง Region Proposals ซึ่งเป็นขั้นตอนการระบุบริเวณที่น่าจะมีวัตถุอยู่ในภาพ โดย RPN จะสแกนภาพและสร้าง anchor boxes ที่มีขนาดและรูปร่างต่างกัน และให้คะแนนความมั่นใจว่าวัตถุอยู่ใน anchor boxes เหล่านี้ และ Object Detection Network ที่ใช้ Convolutional Neural Network (CNN) สำหรับการตรวจจับวัตถุและการจำแนกประเภท โดยใช้ Convolutional Layers เพื่อสกัดคุณลักษณะ (features) จากภาพ และ Fully Connected Layers เพื่อจำแนกประเภทของวัตถุและกำหนดตำแหน่งกล่องขอบเขต Faster R-CNN ถูกออกแบบมาให้สามารถฝึกทั้งสองเครือข่ายย่อย (RPN และ Object Detection Network) พร้อมกันโดยใช้วิธีการฝึกที่สลับกัน ทำให้สามารถปรับปรุงการทำงานของโมเดลได้อย่างมีประสิทธิภาพ โมเดลนี้สามารถตรวจจับวัตถุได้เร็วขึ้นและแม่นยำกว่าวิธีการก่อนหน้านี้ เช่น R-CNN และ Fast R-CNN ซึ่งจำเป็นต้องแยกขั้นตอนการสร้าง Region Proposals และการจำแนกวัตถุออกจากกัน

YOLO (You Only Look Once) (Redmon et al., 2016) เป็นอัลกอริทึมที่ใช้สำหรับการตรวจจับวัตถุในภาพแบบเรียลไทม์ โดยมีความสามารถในการตรวจจับวัตถุในภาพได้อย่างรวดเร็วและมีความแม่นยำ YOLO แตกต่างจากอัลกอริทึมอื่น

ๆ เช่น R-CNN ตรงที่มันประมวลผลภาพเพียงครั้งเดียว (single forward pass) ผ่านเครือข่าย neural network เพื่อทำนายตำแหน่งและชนิดของวัตถุในภาพทั้งหมดพร้อมกัน YOLOv2 (Redmon et al., 2017) ได้รับการปรับปรุงจาก YOLO เดิมโดยใช้ anchor boxes ซึ่งเป็น bounding boxes ที่กำหนดล่วงหน้าด้วยรูปร่างและขนาดที่หลากหลาย เพื่อเพิ่มความแม่นยำในการทำนายตำแหน่งของวัตถุ นอกจากนี้ยังมีการใช้ batch normalization และ high-resolution classifier เพื่อเพิ่มประสิทธิภาพในการตรวจจับวัตถุ และมีการปรับปรุงโครงสร้างของเครือข่ายเพื่อให้สามารถตรวจจับวัตถุที่มีขนาดเล็กได้ดีขึ้น YOLOv3 (Redmon et al., 2018) มีการพัฒนาเพิ่มเติมจาก YOLOv2 โดยเพิ่มความลึกของเครือข่ายเพื่อเพิ่มความสามารถในการสกัดคุณลักษณะ (features) ของวัตถุ นอกจากนี้ยังมีการปรับปรุงให้สามารถทำนาย bounding boxes ที่มีขนาดต่างกันได้ในระดับต่าง ๆ ของเครือข่าย (multi-scale predictions) ซึ่งช่วยให้สามารถตรวจจับวัตถุได้ดีขึ้นในทุกขนาดและระยะห่าง นอกจากนี้ YOLOv3 ยังใช้การประมวลผลแบบ logistic regression ในการทำนาย class scores ซึ่งทำให้สามารถทำนายวัตถุหลายชนิดใน bounding box เดียวกันได้ YOLOv4 (Bochkovskiy et al., 2020) เป็นการพัฒนาเพิ่มเติมจาก YOLOv3 โดยมีการใช้เทคนิคขั้นสูงต่าง ๆ เช่น CSPNet (Cross-Stage Partial Network) และ PANet (Path Aggregation Network) เพื่อเพิ่มประสิทธิภาพในการสกัดคุณลักษณะและการรวมคุณลักษณะจากหลาย ๆ ชั้นของเครือข่าย นอกจากนี้ยังมีการปรับปรุงกระบวนการฝึกอบรมโดยใช้ self-adversarial training และการเพิ่มข้อมูลเสริม (data augmentation) เพื่อเพิ่มความแข็งแกร่งของโมเดล YOLOv4 จึงมีความสามารถในการตรวจจับวัตถุได้ดีขึ้นทั้งในด้านความเร็วและความแม่นยำ ทั้งหมดนี้ทำให้ YOLO และเวอร์ชันต่าง ๆ ของมันเป็นเครื่องมือที่มีประสิทธิภาพสูงสำหรับการตรวจจับวัตถุในภาพที่ต้องการความเร็วและความแม่นยำในเวลาเดียวกัน

2.2 เทคนิคและวิธีการที่เกี่ยวข้องกับการเรียนรู้ของเครื่องแบบไม่สมดุล

การเรียนรู้ของเครื่องแบบไม่สมดุล (Imbalanced Learning) คือปัญหาที่เกิดขึ้นเมื่อข้อมูลที่ใช้ในการเรียนรู้และสร้างแบบจำลองมีการกระจายตัวของคลาสไม่สมดุลกัน กล่าวคือ ข้อมูลในบางคลาสมีจำนวนมากกว่าอีกคลาสหนึ่งอย่างมีนัยสำคัญ ตัวอย่างเช่น ในการจำแนกอีเมลเป็นอีเมลขยะ (spam) หรืออีเมลปกติ (non-spam) อาจมีอีเมลปกติเป็นจำนวนมากกว่าอีเมลขยะมาก ปัญหานี้ทำให้โมเดลที่ฝึกอบรมจากข้อมูลดังกล่าวมีแนวโน้มที่จะทำนายคลาสที่มีจำนวนตัวอย่างมากกว่าได้แม่นยำกว่า ซึ่งอาจทำให้เกิดการละเลยคลาสที่มีจำนวนตัวอย่างน้อยกว่า

การเรียนรู้ของเครื่องแบบไม่สมดุลเป็นปัญหาที่พบได้บ่อยในหลาย ๆ ด้าน เช่น การตรวจจับการฉ้อโกง การวินิจฉัยโรค และการตรวจจับข้อบกพร่องของผลิตภัณฑ์ วิธีการแก้ปัญหานี้มีหลายวิธี รวมถึงการสุ่มเพิ่มจำนวนข้อมูลในคลาสที่มีจำนวนน้อย (oversampling) การสุ่มลดจำนวนข้อมูลในคลาสที่มีจำนวนมาก (undersampling) และการใช้เทคนิคการสร้างข้อมูลใหม่ เช่น การใช้เทคนิค Synthetic Minority Over-sampling Technique (SMOTE) นอกจากนี้ยังมีการใช้วิธีการปรับแต่งโมเดล เช่น การปรับค่าความเสียหาย (cost-sensitive learning) เพื่อให้โมเดลให้ความสำคัญกับคลาสที่มีจำนวนน้อยมากขึ้น

SMOTE (Synthetic Minority Over-sampling Technique) (Chawla et al., 2002) เป็นเทคนิคที่ใช้แก้ปัญหาการเรียนรู้ของเครื่องแบบไม่สมดุลโดยการสร้างตัวอย่างใหม่ในคลาสที่มีจำนวนน้อยเพื่อให้ข้อมูลมีการกระจายตัวของคลาสที่สมดุลมากขึ้น วิธีการทำงานของ SMOTE คือ สำหรับแต่ละตัวอย่างในคลาสที่มีจำนวนน้อยจะเลือกตัวอย่างที่อยู่ใกล้เคียงจำนวนหนึ่ง จากนั้นสุ่มเลือกหนึ่งในตัวอย่างที่ใกล้เคียงมา แล้วคำนวณค่ากลางระหว่างตัวอย่างเดิมกับตัวอย่างที่เลือกมาและสร้างตัวอย่างใหม่โดยการสุ่มตำแหน่งบนเส้นตรงระหว่างตัวอย่างเดิมกับตัวอย่างที่เลือกมา การใช้ SMOTE ช่วยเพิ่มจำนวนตัวอย่างในคลาสที่มีจำนวนน้อยโดยไม่ทำให้ข้อมูลเกิดการซ้ำซ้อนเหมือนกับการสุ่มเพิ่มจำนวนข้อมูลแบบธรรมดา และช่วยให้โมเดลมีความสามารถในการเรียนรู้คุณลักษณะของคลาสที่มีจำนวนน้อยได้ดีขึ้น ซึ่งช่วยเพิ่มประสิทธิภาพในการทำนายคลาสที่มีจำนวนน้อย อย่างไรก็ตาม การใช้ SMOTE ต้องระวังเรื่องการสร้างตัวอย่างที่อาจไม่สอดคล้องกับข้อมูลจริงและการเพิ่มความซับซ้อนของข้อมูลอาจทำให้การฝึกโมเดลมีความยากลำบากมากขึ้น

Borderline-SMOTE (Han et al., 2005) เป็นเทคนิคที่ใช้ในการแก้ปัญหาความไม่สมดุลของข้อมูลในการฝึกโมเดล เป็นการปรับปรุงมาจากเทคนิค SMOTE (Synthetic Minority Over-sampling Technique) ซึ่งมักถูกใช้ในงานที่มีคลาสข้อมูลหลักขาดแค่บางคลาสเท่านั้น เช่น การจำแนกภาพด้วยคอมพิวเตอร์ที่มีภาพของคลาสหนึ่งน้อยมาก หรือการทำนายความเสี่ยงในการป่วยที่มีผู้ป่วยหนึ่งๆ เพียงไม่กี่คน วิธีการทำงานของ Borderline-SMOTE คือการระบุตัวอย่างข้อมูลที่อยู่บนขอบเขต (borderline examples) ของคลาสน้อย หลังจากนั้นจะทำการสร้างตัวอย่างข้อมูลเสริมใหม่ โดยใช้เทคนิคเดียวกับ SMOTE แต่จะมีการสร้างตัวอย่างเฉพาะในพื้นที่ที่เสี่ยงที่สุดที่จะถูกจำแนกผิดพลาด นั่นหมายถึงการสร้างตัวอย่างใหม่ในบริเวณที่คลาสน้อยอาจจะถูกพลิกเป็นคลาสหลักได้

การลดจำนวนตัวอย่างในกลุ่มข้อมูลหลัก เพื่อให้มีจำนวนตัวอย่างในกลุ่มข้อมูลน้อยเท่ากับหรือใกล้เคียงกับจำนวนตัวอย่างในกลุ่มข้อมูลหลัก เรียกว่า under-sampling (He et al., 2009) หรือการสุ่มตัวอย่างในกลุ่มข้อมูลหลักเพื่อลดจำนวนตัวอย่างในกลุ่มข้อมูลน้อยลง เพื่อให้คลาสมีความสมดุลกันมากขึ้น การใช้เทคนิค under-sampling มักถูกนำมาใช้เมื่อกลุ่มข้อมูลหลักมีจำนวนตัวอย่างมากมายเมื่อเปรียบเทียบกับกลุ่มข้อมูลน้อย โดยการลดจำนวนตัวอย่างในกลุ่มข้อมูลหลักอาจทำให้

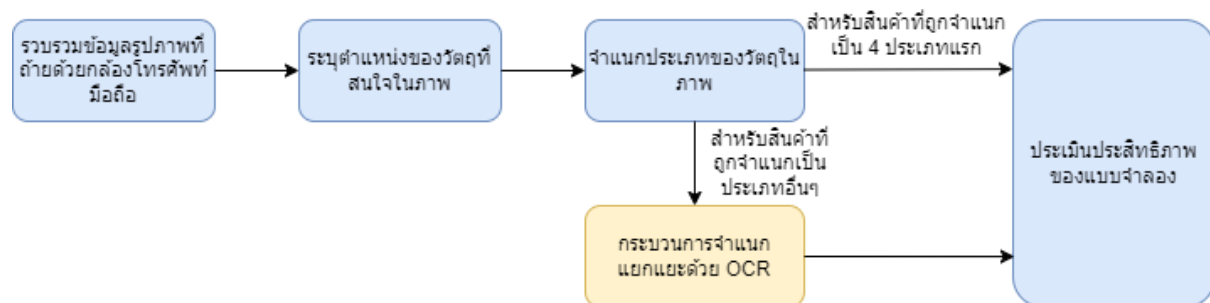
โมเดลไม่ได้รับข้อมูลที่เพียงพอเพื่อให้มั่นใจในการจำแนกหรือทำนายสำหรับกลุ่มข้อมูลหลัก แต่การลดจำนวนตัวอย่างในกลุ่มข้อมูลหลักอาจช่วยลดปัญหาการเรียนรู้ที่ไม่สมดุล และลดโอกาสในการเกิด overfitting ของโมเดลได้

3. ระเบียบวิธีวิจัย

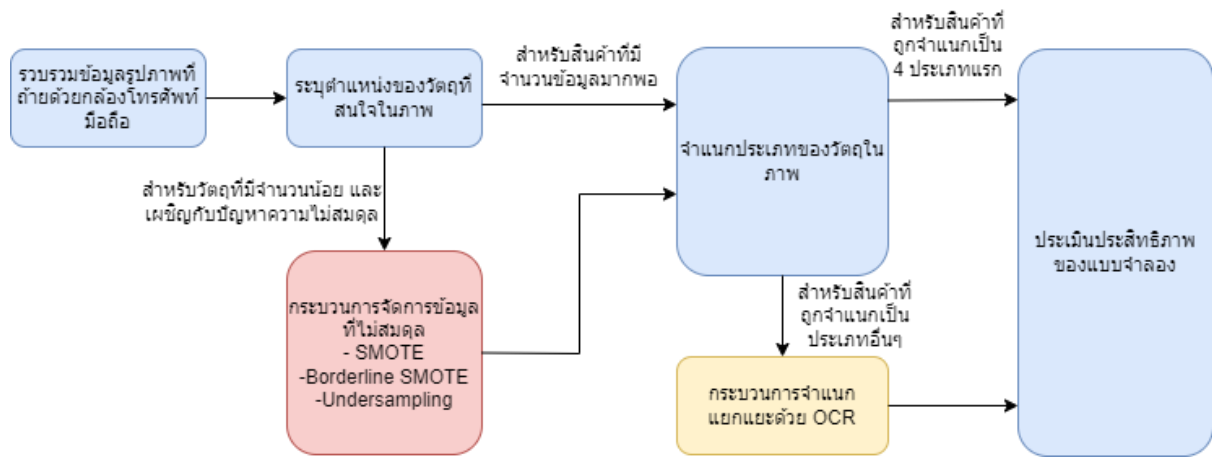
งานวิจัยนี้มีวัตถุประสงค์เพื่อเสนอแนวทางการจำแนกแยกแยะประเภทของสินค้าจากรูปภาพของกล่องพัสดุที่ถ่ายจากตู้คอนเทนเนอร์ในระหว่างพิธีการทางศุลกากร โดยใช้ข้อมูลรูปภาพจากบริษัทตัวแทนการขนส่งสินค้าระหว่างประเทศในกรุงเทพมหานครจำนวน 3 บริษัท โดยรูปภาพทั้งหมด 5,113 รูปจากการถ่ายด้วยกล้องโทรศัพท์มือถือของเจ้าหน้าที่พนักงานที่ทำเรื่องกรุงเทพจะถูกนำมาวิเคราะห์และสร้างแบบจำลองการเรียนรู้ของเครื่องเพื่อทำนายและจำแนกแยกแยะประเภทของสินค้าออกเป็น 5 ประเภทคือ 1) กลุ่มอุปกรณ์ทางการแพทย์ 2) กลุ่มเครื่องมือก่อสร้าง 3) กลุ่มเอกสารและสิ่งพิมพ์ 4) กลุ่มเครื่องใช้ไฟฟ้า 5) กลุ่มอื่นๆ

กระบวนการจะเริ่มต้นด้วยขั้นตอนการระบุวัตถุในรูปภาพ เมื่อพบวัตถุที่เป็นไปได้ จะมีการจำแนกประเภทของวัตถุดังกล่าวโดยใช้เทคโนโลยีการเรียนรู้ของเครื่อง เช่น Convolutional Neural Network (CNN), Region-based Convolutional Neural Network (R-CNN), Fast R-CNN, Faster R-CNN, และ You Only Look Once (YOLO) เพื่อช่วยในการจำแนกประเภทของวัตถุ สำหรับวัตถุที่ระบบสามารถจำแนกแยกแยะเป็น 4 ประเภทแรก คือ 1) กลุ่มอุปกรณ์ทางการแพทย์ 2) กลุ่มเครื่องมือก่อสร้าง 3) กลุ่มเอกสารและสิ่งพิมพ์ และ 4) กลุ่มเครื่องใช้ไฟฟ้านั้น ระบบจะทำการประเมินประสิทธิภาพของแบบจำลองและรายงานผลความถูกต้อง แต่สำหรับวัตถุที่แบบจำลองจำแนกออกเป็นกลุ่มอื่นๆ วัตถุชิ้นนั้นจะถูกนำเข้ากระบวนการจำแนกแยกแยะด้วย Optical Character Recognition (OCR) (Tappert et al., 1990) เพื่ออ่านข้อความจากบรรจุภัณฑ์ ข้อมูลที่ได้จาก OCR นี้จะถูกนำมาใช้ในแบบจำลองการเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning) เช่น การจัดกลุ่ม (Clustering) เพื่อทำการจัดหมวดหมู่เพิ่มเติมและระบุประเภทของผลิตภัณฑ์ที่ไม่สามารถจัดหมวดหมู่ได้อย่างชัดเจนในขั้นตอนก่อนหน้านี้ จากนั้น ระบบจะนับจำนวนวัตถุในรูปภาพเพื่อให้เจ้าหน้าที่สามารถประมาณการปริมาณสินค้าที่จะต้องผ่านการขนส่งที่ท่าเรือได้

สำหรับข้อมูลสินค้าที่มีจำนวนน้อยและเผชิญกับปัญหาข้อมูลไม่สมดุล ผู้วิจัยวางแผนที่จะใช้เทคนิคการเพิ่มข้อมูล (Data Augmentation) และการสุ่มข้อมูล (Data Sampling Techniques) เพื่อปรับปรุงปัญหาความไม่สมดุลในข้อมูล ด้วยเทคนิค Synthetic Minority Over-sampling Technique (SMOTE) เพื่อสร้างตัวอย่างข้อมูลใหม่ในกลุ่มข้อมูลสินค้าที่มีจำนวนน้อย โดยการสร้างตัวอย่างใหม่จากการรวมข้อมูลของตัวอย่างในกลุ่มข้อมูลน้อย ซึ่งการสร้างตัวอย่างใหม่จะใช้เทคนิคการสร้างข้อมูลจำลองในพื้นที่ระหว่างตัวอย่างของกลุ่มข้อมูลน้อย นอกจากนี้ เทคนิค Borderline SMOTE ซึ่งเป็นการสร้างตัวอย่างข้อมูลของคลาสน้อยโดยการคำนวณค่าความใกล้เคียงของตัวอย่างในกลุ่มข้อมูลน้อยและสร้างตัวอย่างใหม่โดยใช้ตัวอย่างที่มีค่าความใกล้เคียงต่ำที่สุด และเพื่อแสดงการเปรียบเทียบผลของประสิทธิภาพของการจำแนกแยกแยะข้อมูลที่ไม่สมดุลมากยิ่งขึ้นเรายังใช้เทคนิค Undersampling โดยการลดจำนวนตัวอย่างของคลาสมากเพื่อใช้ในการอ้างอิงผลลัพธ์อีกด้วย



ภาพที่ 1 กระบวนการจำแนกแยกแยะวัตถุแบบไม่มีการจัดการข้อมูลที่ไม่สมดุล



ภาพที่ 2 กระบวนการจำแนกแยกแยะวัตถุแบบที่มีการจัดการข้อมูลที่ไม่สมดุลรวมอยู่ด้วย

4. ผลการวิจัยและอภิปรายผลการวิจัย

ค่า Accuracy ของแบบจำลองที่ใช้เทคนิคต่างๆ ในการตรวจจับและจำแนกวัตถุแสดงให้เห็นว่า YOLOv4 ให้ค่า Accuracy สูงสุดในทุกประเภทสินค้า ทั้งเครื่องมือการแพทย์ เครื่องมือก่อสร้าง เอกสารและสิ่งพิมพ์ และเครื่องใช้ไฟฟ้า โดยมีค่าเฉลี่ยรวมอยู่ที่ 86.36% ซึ่งสูงที่สุดเมื่อเทียบกับเทคนิคอื่นๆ การที่ค่า Accuracy ของกลุ่มเครื่องมือทางการแพทย์มีค่าต่ำสุดอาจเกิดจากความซับซ้อนของรูปร่างและลักษณะของหีบห่อของเครื่องมือการแพทย์ที่มีความหลากหลายและมีรายละเอียดมาก ทำให้การตรวจจับและจำแนกทำได้ยากกว่าสินค้าชนิดอื่น นอกจากนั้น เครื่องมือการแพทย์บางชนิดอาจมีการติดฉลากหรือลายพิมพ์ที่ไม่ชัดเจนหรือไม่เลย ทำให้เทคนิค OCR หรือการตรวจจับด้วยภาพไม่สามารถทำงานได้อย่างมีประสิทธิภาพ ตารางที่ 1 แสดงค่า Accuracy จากการทำนายผลลัพธ์ของแบบจำลองแยกตามเทคนิคการตรวจจับฯ และประเภทของสินค้า

ตารางที่ 1 ค่า Accuracy(%) ของแบบจำลอง แยกตามเทคนิคการตรวจจับฯ และประเภทของสินค้า

เทคนิคตรวจจับและจำแนกวัตถุ	เครื่องมือการแพทย์	เครื่องมือก่อสร้าง	เอกสารและสิ่งพิมพ์	เครื่องใช้ไฟฟ้า
R-CNN	65.22	83.21	82.44	79.54
Fast R-CNN	69.45	88.64	85.45	79.65
Faster R-CNN	70.25	82.5	88.73	82.45
YOLOv2	71.11	85.72	86.54	85.49
YOLOv3	70.36	85.14	88.05	89.54
YOLOv4	73.58	89.72	88.12	90.12
ค่าเฉลี่ย	70.00	85.82	86.56	84.47

ตารางที่ 2 แสดงค่า Precision (%) และ Recall (%) ของการจำแนกสินค้ากลุ่มเครื่องมือแพทย์ โดยเปรียบเทียบระหว่างการไม่ใช้ Resampling (without Resampling) กับการใช้เทคนิค Resampling ต่าง ๆ ได้แก่ SMOTE, BorderlineSMOTE และ Undersampling จากผลลัพธ์จะเห็นได้ว่าเทคนิคการใช้ SMOTE และ BorderlineSMOTE ช่วยเพิ่มค่า Precision และ Recall ได้ในเกือบทุกกรณี โดยเฉพาะเทคนิค YOLOv4 กับ SMOTE ที่มีค่า Precision สูงสุดถึง 81.10% และ Recall สูงสุดถึง 68.44% ขณะที่การใช้ Undersampling มีค่า Precision และ Recall ต่ำสุดเมื่อเทียบกับเทคนิค Resampling อื่น ๆ ซึ่งแสดงให้เห็นว่า SMOTE และ BorderlineSMOTE มีประสิทธิภาพในการปรับปรุงความแม่นยำในการจำแนกประเภทของเครื่องมือแพทย์มากกว่า Undersampling

ตารางที่ 2 ค่า Precision(%) และ Recall(%) ของการจำแนกสินค้ากลุ่มเครื่องมือแพทย์แยกตามเทคนิคการตรวจจำแนก และเทคนิค Resampling (SMOTE, BorderlineSMOTE และ Undersampling)

	กลุ่มเครื่องมือการแพทย์							
	without Resampling		with SMOTE		with BoderlineSMOTE		with Undersampling	
	Precision	Recall	Precision	Recall	Precision	Recall	Precision	Recall
R-CNN	74.55	49.87	75.88	50.89	73.21	51.02	72.11	45.25
Fast R-CNN	75.44	52.23	76.92	55.88	74.59	53.21	70.45	48.77
Faster R-CNN	40.21	55.52	50.28	60.35	50.69	55.22	35.21	52.68
YOLOv2	70.25	52.54	80.55	65.32	75.45	55.89	60.58	52.44
YOLOv3	71.88	59.84	80.34	63.28	75.92	59.63	62.47	55.74
YOLOv4	72.78	61.87	81.1	68.44	76.22	62.47	65.89	57.47
Average	67.52	55.31	74.18	60.69	71.01	56.24	61.12	52.06

การใช้เทคนิค Undersampling ในการจำแนกและการตรวจจำแนกวัตถุในกลุ่มเครื่องมือการแพทย์มักจะมีผลไม่ดี เนื่องจาก การลดจำนวนตัวอย่างของกลุ่มข้อมูลส่วนมาก (Majority Class) อาจทำให้ข้อมูลสำคัญหายไป ซึ่งส่งผลให้แบบจำลองทำนายผลลัพธ์ได้ไม่แม่นยำ นอกจากนี้ ข้อมูลที่หลงเหลืออยู่หลังจากการลบหน่วยตัวอย่างทิ้งไปอาจไม่ครอบคลุมทุกกรณีของประเภทข้อมูล ซึ่งอาจทำให้แบบจำลองไม่สามารถจำแนกวัตถุที่มีความแตกต่างกันได้อย่างแม่นยำ

ตารางที่ 3, 4 และ 5 แสดงค่า Precision และ Recall ของการจำแนกสินค้ากลุ่มเครื่องมือก่อสร้าง กลุ่มเอกสารและสิ่งพิมพ์ และกลุ่มเครื่องใช้ไฟฟ้า ตามลำดับ ซึ่งผลการทดลองสอดคล้องกับผลการทดลองของกลุ่มเครื่องมือแพทย์ กล่าวคือ เทคนิคการใช้ SMOTE และ BorderlineSMOTE ช่วยเพิ่มค่า Precision และ Recall ได้ในเกือบทุกกรณี

ตารางที่ 3 ค่า Precision(%) และ Recall(%) ของการจำแนกสินค้ากลุ่มเครื่องมือก่อสร้างแยกตามเทคนิคการตรวจจำแนก และเทคนิค Resampling (SMOTE, BorderlineSMOTE และ Undersampling)

	กลุ่มเครื่องมือก่อสร้าง							
	without Resampling		with SMOTE		with BoderlineSMOTE		with Undersampling	
	Precision	Recall	Precision	Recall	Precision	Recall	Precision	Recall
R-CNN	75.82	76.55	79.42	75.48	78.26	74.21	72.12	50.87
Fast R-CNN	79.72	82.41	83.77	82.44	81.1	82.24	75.44	54.78
Faster R-CNN	75.68	83.83	80.55	81.02	81.15	80.1	73.47	63.41
YOLOv2	75.49	80.26	83.47	82.98	82.47	81.3	72.98	67.92
YOLOv3	74.12	80.88	82.98	82.45	80.14	81.47	71.47	60.96
YOLOv4	76.45	88.04	81.92	81.89	80.47	80.87	74.14	60.68
Average	76.21	82.00	82.02	81.04	80.60	80.03	73.27	59.77

ตารางที่ 4 ค่า Precision(%) และ Recall(%) ของการจำแนกสินค้ากลุ่มเอกสารและสิ่งพิมพ์แยกตามเทคนิคการตรวจจับๆ และเทคนิค Resampling (SMOTE, BorderlineSMOTE และ Undersampling)

	กลุ่มเอกสารและสิ่งพิมพ์							
	without Resampling		with SMOTE		with BoderlineSMOTE		with Undersampling	
	Precision	Recall	Precision	Recall	Precision	Recall	Precision	Recall
R-CNN	74.22	80.21	78.93	80.47	76.66	79.98	70.44	70.41
Fast R-CNN	72.54	82.45	77.99	81.42	72.45	79.45	70.14	72.55
Faster R-CNN	75.12	84.68	78.54	83.24	75.22	79.47	70.47	73.83
YOLOv2	75.69	85.72	78.21	84.72	75.41	81.24	69.58	78.14
YOLOv3	78.66	85.23	80.2	84.21	76.32	81.23	69.54	77.55
YOLOv4	79.63	85.78	80.14	84.63	77.45	83.35	75.87	75.89
Average	75.98	84.01	79.00	83.12	75.59	80.79	71.01	74.73

ตารางที่ 5 ค่า Precision(%) และ Recall(%) ของการจำแนกสินค้ากลุ่มเครื่องใช้ไฟฟ้าแยกตามเทคนิคการตรวจจับๆ และเทคนิค Resampling (SMOTE, BorderlineSMOTE และ Undersampling)

	กลุ่มเครื่องใช้ไฟฟ้า							
	without Resampling		with SMOTE		with BoderlineSMOTE		with Undersampling	
	Precision	Recall	Precision	Recall	Precision	Recall	Precision	Recall
R-CNN	72.45	78.63	74.87	77.21	72.44	75.44	65.45	50.24
Fast R-CNN	72.04	75.25	74.89	75.01	72.92	75.41	65.74	52.45
Faster R-CNN	75.88	75.75	76.55	75.88	75.45	75.98	63.72	54.47
YOLOv2	74.92	78.72	76.98	78.87	74.43	76.54	60.21	55.47
YOLOv3	75.65	80.12	77.99	81.42	76.55	80.11	61.24	55.74
YOLOv4	78.97	82.98	79.82	82.55	78.5	81.95	62.35	54.78
Average	74.99	78.58	76.85	78.49	75.05	77.57	63.12	53.86

5. สรุปผลการวิจัย

การทดลองในงานวิจัยชิ้นนี้ช่วยในการทำความเข้าใจถึงประสิทธิภาพของระบบจำแนกและตรวจจับวัตถุที่เป็นสินค้าในตู้คอนเทนเนอร์ เพื่อระบุประเภทของสินค้าให้ตรงกับเอกสารและแจ้งเตือนเจ้าหน้าที่ของบริษัทตัวแทนด้านการนำเข้าและส่งออกถึงการมาถึงของสินค้าในล็อตนั้นๆ ซึ่งพบว่าอัลกอริทึม YOLOv4 มีความแม่นยำสูงสุดเมื่อไม่มีการใช้เทคนิค Resampling เมื่อเปรียบเทียบกับอัลกอริทึมอื่นๆ นอกจากนั้น การใช้เทคนิค Resampling เช่น SMOTE และ BoderlineSMOTE สามารถช่วยเพิ่มประสิทธิภาพของการจำแนกแยกแยะให้มีความถูกต้องมากขึ้น การใช้เทคนิค Undersampling ร่วมกับการจำแนกและการตรวจจับวัตถุจะไม่ช่วยเพิ่มความถูกต้องของผลลัพธ์ เนื่องจากการลดจำนวนตัวอย่างอาจทำให้ข้อมูลสำคัญสูญหายไป ข้อมูลที่หลงเหลืออยู่อาจไม่ครอบคลุมแนวคิดในทุกกรณีของประเภทข้อมูล และทำให้แบบจำลองไม่สามารถจำแนกวัตถุที่มีความแตกต่างกันได้อย่างแม่นยำ

Precision และ Recall เป็นตัววัดสำคัญที่ช่วยในการทำความเข้าใจเรื่องความแม่นยำและความสามารถในการจำแนกและตรวจจับวัตถุของระบบ ค่า Precision และ Recall ที่สูงทำให้มั่นใจในความถูกต้องและความสามารถของระบบ การเลือกใช้โมเดลและเทคนิค Resampling ที่เหมาะสมสำหรับงานที่กำลังดำเนินการมีความสำคัญอย่างมาก โดยควรพิจารณาความต้องการและลักษณะของข้อมูลเพื่อเลือกใช้แบบจำลองและเทคนิคที่มีประสิทธิภาพสูงที่สุดในงานนั้น

ระบบตรวจจับวัตถุและจำแนกแยกแยะประเภทของสินค้าด้วยการยืนยันผ่านเทคนิค OCR เป็นระบบที่สามารถใช้ระบุประเภทของสินค้าผ่านการวิเคราะห์ภาพถ่ายบรรจุภัณฑ์ที่เจ้าหน้าที่หน้างานถ่ายภาพจากตู้คอนเทนเนอร์ในวันที่สินค้ามาถึง

และกำลังดำเนินพิธีการทางศุลกากร และเมื่อประมวลผลข้อมูลร่วมกับเทคนิคการเรียนรู้ข้อมูลที่ไม่สมดุล จะทำให้แบบจำลองมีความแม่นยำมากขึ้นในกรณีที่เป็นสินค้าชิ้นใหม่ และมีข้อมูลไม่มากพอที่แบบจำลองจะเรียนรู้รูปแบบได้อย่างเต็มที่

6. กิตติกรรมประกาศ

ผู้วิจัยขอขอบคุณผู้บริหารระดับสูงของบริษัทตัวแทนบริการนำเข้าส่งออกทั้งสามบริษัท โดยเฉพาะคุณณัฐริชา ผลจันทร์ ที่ให้ข้อมูลและรูปภาพที่ใช้ในงานวิจัยนี้ ผู้วิจัยตระหนักดีว่าข้อมูลที่เก็บรวบรวมสำหรับการศึกษาในครั้งนี้เป็นข้อมูลที่อ่อนไหว และเป็นความลับ ข้อมูลที่มีคุณค่าทั้งหมดจะถูกใช้อย่างระมัดระวังและใช้เพื่อวัตถุประสงค์ในการวิจัยเท่านั้น

7. เอกสารอ้างอิง

- Prexawanprasut, T., Santiworarak, L., Nurarak, P., & Juasiripukdee, P. (2023). Employing Machine Learning and an OCR Validation Technique to Identify Product Category Based on Visible Packaging Features. In Proceedings of the 2023 6th International Conference on Machine Vision and Applications (ICMVA '23) (pp. 114–119). Association for Computing Machinery. <https://doi.org/10.1145/3589572.3589589>.
- LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278–2324. <https://doi.org/10.1109/5.726791>.
- Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
- Ren, S., He, K., Girshick, R., and Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks, Vol. 28. Advances in neural information processing systems, ISBN: 9781510825024.
- Ren, S., He, K., Girshick, R., & Sun, J. (2017). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI).
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR) (pp. 779-788).
- Redmon, J., & Farhadi, A. (2017). YOLO9000: Better, faster, stronger. In Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR) (pp. 7263-7271).
- Redmon, J., & Farhadi, A. (2018). YOLOv3: An incremental improvement. arXiv preprint arXiv:1804.02767.
- Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection. arXiv preprint arXiv:2004.10934.
- Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. Journal of artificial intelligence research, 16, 321-357.
- Han, H., Wang, W. Y., & Mao, B. H. (2005). Borderline-SMOTE: A new over-sampling method in imbalanced data sets learning. In International conference on intelligent computing (pp. 878-887). Springer, Berlin, Heidelberg.
- He, H., & Garcia, E. A. (2009). Learning from imbalanced data. IEEE Transactions on Knowledge and Data Engineering, 21(9), 1263-1284.
- Tappert, C. C., Suen, C. Y., & Wakahara, T. (1990). The state of the art in online handwriting recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 12(8), 787-808.