



การแนะนำบรรจุภัณฑ์อาหารที่เหมาะสมตามรูปภาพของของหวานโดยใช้โมเดลการจำแนกประเภท  
หลายเลเบลและการค้นหาภาพสำหรับแพลตฟอร์มอีคอมเมิร์ซ



โดย  
นายอันดามัน นพนภาพร

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรมหาบัณฑิต

สาขาวิชาเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล แผน ก แบบ ก 2

ภาควิชาคอมพิวเตอร์

มหาวิทยาลัยศิลปากร

ปีการศึกษา 2567

ลิขสิทธิ์ของมหาวิทยาลัยศิลปากร

การแนะนำบรรจุภัณฑ์อาหารที่เหมาะสมตามรูปภาพของของหวานโดยใช้โมเดลการ  
จำแนกประเภทหลายเลเบลและการค้นหาภาพสำหรับแพลตฟอร์มอีคอมเมิร์ซ



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรมหาบัณฑิต

สาขาวิชาเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล แผน ก แบบ ก 2

ภาควิชาคอมพิวเตอร์

มหาวิทยาลัยศิลปากร

ปีการศึกษา 2567

ลิขสิทธิ์ของมหาวิทยาลัยศิลปากร



By

MR. Undaman NOPNAPORN

A Thesis Submitted in Partial Fulfillment of the Requirements  
for Master of Science INFORMATION TECHNOLOGY AND DIGITAL INNOVATION  
Department of COMPUTER SCIENCE  
Academic Year 2024  
Copyright of Silpakorn University

|                      |                                                                                                                                |
|----------------------|--------------------------------------------------------------------------------------------------------------------------------|
| หัวข้อ               | การแนะนำบรรจุภัณฑ์อาหารที่เหมาะสมตามรูปภาพของของหวานโดยใช้โมเดลการจำแนกประเภทหลายเลเบลและการค้นหาภาพสำหรับแพลตฟอร์มอีคอมเมิร์ซ |
| โดย                  | นายอันดามัน นพนภาพร                                                                                                            |
| สาขาวิชา             | เทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล แผน ก แบบ ก 2                                                                              |
| อาจารย์ที่ปรึกษาหลัก | ผู้ช่วยศาสตราจารย์ ดร. ณัฐโชติ พรหมฤทธิ์                                                                                       |
| อาจารย์ที่ปรึกษาร่วม | อาจารย์ ดร. สัจจาภรณ์ ไวจรรยา                                                                                                  |

คณะวิทยาศาสตร์ มหาวิทยาลัยศิลปากร ได้รับพิจารณาอนุมัติให้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรมหาบัณฑิต

.....คณบดีคณะวิทยาศาสตร์  
(ผู้ช่วยศาสตราจารย์ ดร. นรงค์ นิมพาลี)

พิจารณาเห็นชอบโดย

.....ประธานกรรมการ  
(รองศาสตราจารย์ ดร. อนิราช มิ่งขวัญ)  
.....อาจารย์ที่ปรึกษาหลัก  
(ผู้ช่วยศาสตราจารย์ ดร. ณัฐโชติ พรหมฤทธิ์)  
.....อาจารย์ที่ปรึกษาร่วม  
(อาจารย์ ดร. สัจจาภรณ์ ไวจรรยา)

.....ผู้ทรงคุณวุฒิภายใน  
(ผู้ช่วยศาสตราจารย์ ดร. อรวรรณ เชาวลิต)  
.....ผู้ทรงคุณวุฒิภายนอก  
(รองศาสตราจารย์ ดร. โอฬาริก สุรินตะ)

650720059 : เทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล แผน ก แบบ ก 2

คำสำคัญ : การแนะนำ, การจำแนกประเภทหลายเลเบล, การแบ่งส่วนภาพ, การค้นหาภาพ, การเรียนรู้เชิงลึก

นาย อ้นดามัน นพนภาพร: การแนะนำบรรจุภัณฑ์อาหารที่เหมาะสมตามรูปภาพของของหวานโดยใช้โมเดลการจำแนกประเภทหลายเลเบลและการค้นหาภาพสำหรับแพลตฟอร์มอีคอมเมิร์ซ อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก : ผู้ช่วยศาสตราจารย์ ดร. ณัฐโชติ พรหมฤทธิ์

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนากระบวนการแนะนำบรรจุภัณฑ์อาหารบนแพลตฟอร์มอีคอมเมิร์ซด้วยเทคนิคจำแนกประเภทหลายเลเบล โดยใช้แบบจำลองการเรียนรู้เชิงลึก และภาพที่เป็นอินพุตของแบบจำลองคือภาพขนมหวานที่ผ่านกระบวนการแบ่งส่วนภาพด้วย YOLOv8 เพื่อแบ่งส่วนที่เป็นเบเกอรี่และตัดส่วนอื่นที่ไม่ต้องการออก จากนั้นนำผลลัพธ์ที่ได้เข้าสู่กระบวนการสกัดคุณลักษณะ โดยใช้แบบจำลองสำเร็จรูป InceptionV3 และจำแนกประเภทด้วยแบบจำลองจำแนกประเภทหลายเลเบล จากนั้นนำคุณลักษณะของภาพบรรจุภัณฑ์ที่ได้จากการแนะนำโดยแบบจำลองจำแนกประเภทหลายเลเบล จับคู่กับคุณลักษณะของบรรจุภัณฑ์ในคลังภาพสำหรับการทดสอบทั้งหมด 2,243 ภาพ ด้วยการคำนวณค่าความคล้ายคลึงของโคไซน์ เพื่อค้นหาภาพอื่นๆ ที่มีลักษณะใกล้เคียงซึ่งอาจทำให้ได้ภาพบรรจุภัณฑ์อื่น ๆ ที่มีลักษณะใกล้เคียงแต่ยังไม่เคยถูกฝึกฝนมาก่อน ผลการทดลองแสดงให้เห็นว่ากระบวนการที่พัฒนาขึ้นสามารถทำการแนะนำบรรจุภัณฑ์ได้โดยอ้างอิงจากความคล้ายคลึงของลักษณะภาพผลลัพธ์ดังกล่าวสามารถนำไปใช้เพื่อสนับสนุนการตัดสินใจเลือกบรรจุภัณฑ์สำหรับสินค้าเบเกอรี่ได้ในบริบทของแพลตฟอร์มอีคอมเมิร์ซได้มากยิ่งขึ้น



650720059 : Major INFORMATION TECHNOLOGY AND DIGITAL INNOVATION

Keyword : Recommendation, Multi-Label Classification, Image Segmentation, Image Search, Deep Learning

MR. Undaman NOPNAPAPORN : Recommending Suitable Food Packaging According to Dessert Images Using Multi-Label Classification Model and Image Search for an E-commerce Platform Thesis advisor : Assistant Professor Nuttachot Promrit, Ph.D.

*This research aims to develop a process for recommending food packaging on e-commerce platforms using multi-label classification techniques and deep learning models. The input images for the model are bakery images that have undergone segmentation using YOLOv8 to isolate the bakery from unwanted elements. The resulting images are then processed for feature extraction using a pre-trained model InceptionV3, and classified using a multi-label classification model. The packaging features derived from the model are matched with the features of packaging images in a dataset containing 2,243 test images by calculating cosine similarity. This process identifies other packaging images with similar characteristics, even those that have not been previously trained. The experimental results indicate that the developed process can recommend packaging based on image feature similarity. These results can be applied to support decision-making for selecting packaging for bakery products in the context of e-commerce platforms, making the selection process more effective.*



## กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้ได้อย่างเรียบร้อย เนื่องด้วยความอนุเคราะห์อย่างยิ่งของผู้ช่วยศาสตราจารย์ ดร.ณัฐโชติ พรหมฤทธิ์ อาจารย์ที่ปรึกษา ซึ่งได้ให้คำแนะนำสนับสนุนทางวิชาการอย่างเต็มกำลัง ขอขอบคุณอาจารย์ ดร.สัจจาภรณ์ ไวจรรยา อาจารย์ที่ปรึกษาร่วม ที่ได้โปรดให้ความคิดเห็นและข้อเสนอแนะต่างๆอันมีค่าในการทำงานวิจัยขึ้นนี้ให้สำเร็จลุล่วงไปได้ด้วยดี ขอขอบคุณรองศาสตราจารย์ ดร. อนิราช มิ่งขวัญ ประธานกรรมการสอบวิทยานิพนธ์ ที่สละเวลา และได้ใช้ความเชี่ยวชาญของท่านในการให้แนวคิดต่างๆจนได้งานวิจัยที่มีคุณภาพ รวมถึงคณะกรรมการซึ่งได้แก่ ผู้ช่วยศาสตราจารย์ ดร. อรรธรณ เชาวลิต และรองศาสตราจารย์ ดร. โอฬาริก สุรินตะ ทั้งสองท่านมีส่วนสำคัญในการให้ข้อเสนอแนะที่ทรงคุณค่าเพื่อนำมาใช้ในการปรับปรุงและพัฒนางานวิจัยนี้ให้สมบูรณ์แบบมากยิ่งขึ้น ขอขอบคุณ คุณมนู มากมณี ซึ่งเป็น CEO และเจ้าของ Box Journey รวมถึงทีมงาน Box Journey ทุกคน ที่ให้การสนับสนุนและความช่วยเหลือในด้านข้อมูล ตลอดจนข้อแนะนำที่สำคัญ และมีคุณค่าต่อกระบวนการวิจัย นอกจากนี้ขอขอบคุณคณะเจ้าหน้าที่บุคลากรของคณะวิทยาศาสตร์ และบัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร สำหรับการสนับสนุนและให้บริการที่เป็นเลิศในด้านธุรการ ผลงานวิจัยขึ้นนี้จะสำเร็จลงมิได้หากขาดทุกท่านที่กล่าวมานี้

อันดามัน นพนภาพร

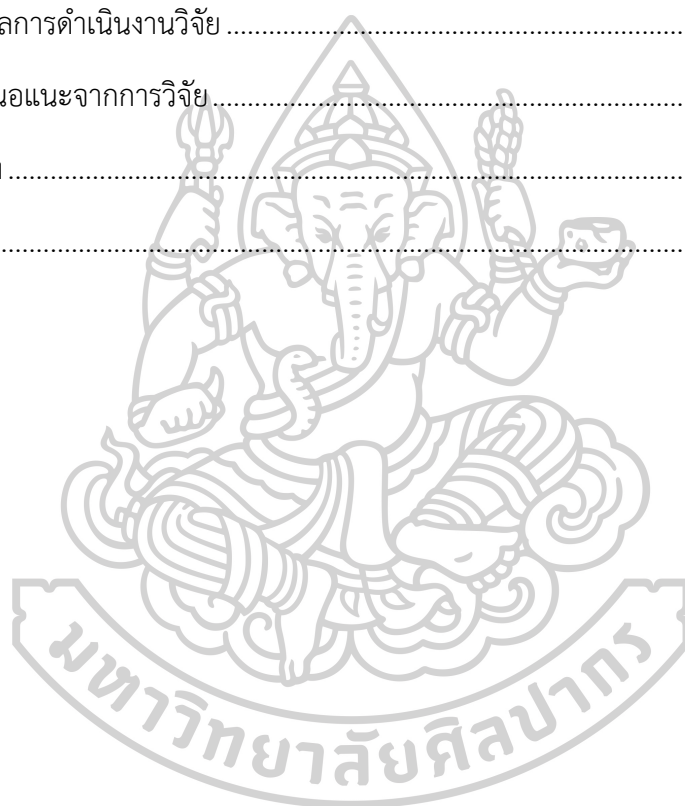


## สารบัญ

|                                                                        | หน้า |
|------------------------------------------------------------------------|------|
| บทคัดย่อภาษาไทย.....                                                   | ง    |
| บทคัดย่อภาษาอังกฤษ.....                                                | จ    |
| กิตติกรรมประกาศ.....                                                   | ฉ    |
| สารบัญ.....                                                            | ช    |
| สารบัญตาราง.....                                                       | ญ    |
| สารบัญภาพ .....                                                        | ฎ    |
| บทที่ 1 บทนำ .....                                                     | 1    |
| 1.1 ที่มาและความสำคัญ.....                                             | 1    |
| 1.2 วัตถุประสงค์ของการวิจัย .....                                      | 2    |
| 1.3 ขอบเขตการวิจัย.....                                                | 2    |
| 1.4 ขั้นตอนการดำเนินการวิจัย.....                                      | 6    |
| 1.5 ประโยชน์ที่คาดว่าจะได้รับ.....                                     | 6    |
| บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง .....                            | 8    |
| 2.1 การประมวลผลภาพ (Image Processing).....                             | 8    |
| 2.2 การเรียนรู้เชิงลึก (Deep Learning).....                            | 9    |
| 2.3 โครงข่ายประสาทแบบคอนโวลูชัน (Convolution Neural Network: CNN)..... | 10   |
| 2.4 การตรวจจับวัตถุด้วย YOLO (You Only Look Once) .....                | 11   |
| 2.5 ความคล้ายคลึงของโคไซน์ (Cosine Similarity) .....                   | 13   |
| 2.6 การขยายข้อมูล (Data Augmentation) .....                            | 14   |
| 2.7 Transfer Learning .....                                            | 15   |
| 2.8 Docker.....                                                        | 16   |

|                                                          |    |
|----------------------------------------------------------|----|
| 2.9 งานวิจัยที่เกี่ยวข้อง.....                           | 18 |
| 2.9.1 Kumar และคณะ .....                                 | 18 |
| 2.9.2 Mohammadi และคณะ.....                              | 19 |
| 2.9.3 Nayak และคณะ .....                                 | 19 |
| 2.9.4 Kang และคณะ .....                                  | 20 |
| 2.9.5 Fadhilla และคณะ.....                               | 20 |
| 2.9.6 Chang และคณะ.....                                  | 21 |
| 2.9.7 Islam และคณะ .....                                 | 21 |
| 2.9.8 Sharma และคณะ .....                                | 21 |
| บทที่ 3 ขั้นตอนและวิธีดำเนินการวิจัย.....                | 24 |
| 3.1 ศึกษางานวิจัยที่เกี่ยวข้อง .....                     | 24 |
| 3.2 ออกแบบโครงสร้างระบบ.....                             | 25 |
| 3.3 ออกแบบและฝึกสอนโมเดล Image Segmentation.....         | 27 |
| 3.4 การเพิ่มข้อมูลด้วยเทคนิค Data Augmentation.....      | 32 |
| 3.5 ออกแบบและฝึกสอนโมเดล Multi-label Classification..... | 33 |
| 3.6 กระบวนการแยกส่วนและลบพื้นหลังภาพ.....                | 41 |
| 3.7 การสกัดพีเจอร์ของภาพบรรจุภัณฑ์.....                  | 42 |
| 3.8 ระบบการกรองบรรจุภัณฑ์ตามขนาดและประเภท .....          | 43 |
| 3.9. กระบวนการค้นหาภาพ.....                              | 45 |
| 3.10. ประยุกต์ใช้โมเดลบนระบบ API.....                    | 46 |
| 3.11. พัฒนาระบบ API Gateway.....                         | 50 |
| 3.12. ประเมินผลการทำงานระบบแนะนำบรรจุภัณฑ์ .....         | 51 |
| บทที่ 4 ผลการดำเนินงานวิจัย.....                         | 52 |
| 4.1. การพัฒนาเว็บแอปพลิเคชันและเก็บผลข้อมูลเลเบล.....    | 52 |

|                                                                                            |    |
|--------------------------------------------------------------------------------------------|----|
| 4.2. ผลการทดสอบประสิทธิภาพโมเดลคัดแยกพื้นหลังออกจากวัตถุ (Image Segmentation).             | 54 |
| 4.3. ผลการฝึกสอนและทดสอบโมเดลการจำแนกประเภทแบบหลายเลเบล (Multi-label Classification) ..... | 56 |
| 4.4. การค้นหาบรรทัดฐานที่คล้ายกันด้วย Cosine Similarity .....                              | 59 |
| 4.5. ผลการเปรียบเทียบโมเดลที่ฝึกสอนด้วยข้อมูลจากนักศึกษาและผู้เชี่ยวชาญ .....              | 60 |
| บทที่ 5 สรุปผลการดำเนินงานวิจัยและข้อเสนอแนะ .....                                         | 61 |
| 5.1. สรุปผลการดำเนินงานวิจัย .....                                                         | 61 |
| 5.2. ข้อเสนอแนะจากการวิจัย .....                                                           | 63 |
| รายการอ้างอิง .....                                                                        | 65 |
| ประวัติผู้เขียน .....                                                                      | 71 |



## สารบัญตาราง

### หน้า

|                                                                                                      |    |
|------------------------------------------------------------------------------------------------------|----|
| ตารางที่ 1 ตารางเปรียบเทียบของแบบจำลองการเรียนรู้เชิงลึกสำหรับการแนะนำสินค้าและการวิเคราะห์ภาพ ..... | 22 |
| ตารางที่ 2 ตารางเทคนิคการวัดความคล้ายคลึง .....                                                      | 23 |
| ตารางที่ 3 ตารางเปรียบเทียบโมเดล YOLOv8 Instance Segmentation (docs.ultralytics.com, 2024).....      | 30 |
| ตารางที่ 4 ค่าพารามิเตอร์ระหว่างการฝึกสอนโมเดล .....                                                 | 31 |
| ตารางที่ 5 ตัวอย่างข้อมูล Label Matrix .....                                                         | 38 |
| ตารางที่ 6 ตัวอย่างข้อมูล Label Matrix หลังการปรับเปลี่ยนด้วยวิธี Majority Voting .....              | 39 |
| ตารางที่ 7 ผลการวิเคราะห์ข้อมูลการเลเบลโดยผู้เชี่ยวชาญ.....                                          | 54 |
| ตารางที่ 8 ตารางเปรียบเทียบประสิทธิภาพโมเดลหลังจากฝึกสอนด้วยชุดข้อมูล FoodSeg103 .....               | 54 |
| ตารางที่ 9 เวลาในการทำนายแต่ละโมเดล.....                                                             | 56 |
| ตารางที่ 10 ผลการประเมินประสิทธิภาพโมเดลการจำแนกแบบหลายเลเบล (ก่อน Data Augmentation).....           | 57 |
| ตารางที่ 11 ผลการประเมินประสิทธิภาพโมเดลการจำแนกแบบหลายเลเบล (หลัง Data Augmentation).....           | 58 |
| ตารางที่ 12 การเปรียบเทียบข้อมูลเลเบลและประสิทธิภาพของโมเดล.....                                     | 60 |

## สารบัญภาพ

|                                                                                                 | หน้า |
|-------------------------------------------------------------------------------------------------|------|
| ภาพที่ 1 ตัวอย่างชุดข้อมูลผลิตภัณฑ์เบเกอรี่.....                                                | 4    |
| ภาพที่ 2 การตัดพื้นหลัง (Background Removal).....                                               | 8    |
| ภาพที่ 3 Deep Neural Network [3].....                                                           | 9    |
| ภาพที่ 4 การเรียนรู้เชิงลึก (Deep Learning).....                                                | 10   |
| ภาพที่ 5 โครงข่ายประสาทเทียมแบบคอนโวลูชัน (CNN).....                                            | 11   |
| ภาพที่ 6 Transfer Learning [28].....                                                            | 15   |
| ภาพที่ 7 Architecture of Docker [30].....                                                       | 16   |
| ภาพที่ 8 กรอบการดำเนินการวิจัย.....                                                             | 25   |
| ภาพที่ 9 ขั้นตอนการทำงานของระบบ.....                                                            | 27   |
| ภาพที่ 10 ตัวอย่างชุดข้อมูล FoodSeg103.....                                                     | 29   |
| ภาพที่ 11 ขั้นตอนการฝึกสอนโมเดล Image Segmentation.....                                         | 29   |
| ภาพที่ 12 Data Augmentation Example.....                                                        | 33   |
| ภาพที่ 13 ขั้นตอนการฝึกสอนโมเดล Multi-label Classification.....                                 | 33   |
| ภาพที่ 14 เว็บไซต์แหล่งข้อมูลภาพออนไลน์ <a href="https://images.cv">https://images.cv</a> ..... | 34   |
| ภาพที่ 15 ตัวอย่างโครงสร้างข้อมูลบรรจุภัณฑ์.....                                                | 34   |
| ภาพที่ 16 ตัวอย่างข้อมูลภาพบรรจุภัณฑ์.....                                                      | 35   |
| ภาพที่ 17 เว็บไซต์เลเบลข้อมูลความเหมาะสมระหว่างบรรจุภัณฑ์และเบเกอรี่.....                       | 36   |
| ภาพที่ 18 ตัวอย่างผลการเลือกบรรจุภัณฑ์ กับภาพเบเกอรี่.....                                      | 37   |
| ภาพที่ 19 รูปแบบข้อมูลการเลเบลรูปแบบ JSON.....                                                  | 37   |
| ภาพที่ 20 Strategies of Fine-Tuning Model.....                                                  | 40   |
| ภาพที่ 21 ผลลัพธ์การตัดพื้นหลังด้วยโมเดล YOLOv8 Image Segmentation.....                         | 41   |

|                                                                                           |    |
|-------------------------------------------------------------------------------------------|----|
| ภาพที่ 22 ตัวอย่างกระบวนการกรองบรรจุภัณฑ์ด้วยขนาดเบเกอร์                                  | 44 |
| ภาพที่ 23 ขั้นตอนการกรองบรรจุภัณฑ์ด้วย Metadata (type, size)                              | 45 |
| ภาพที่ 24 กระบวนการค้นหา Nearest Neighbors ด้วยเทคนิค Cosine similarity                   | 46 |
| ภาพที่ 25 ข้อมูลภายในไฟล์ docker-compose-package.yml                                      | 46 |
| ภาพที่ 26 ผลลัพธ์การค้นหา Mask ด้วยโมเดล YOLOv8 Segmentation                              | 47 |
| ภาพที่ 27 ผลการปรับค่าสีแต่ละ pixel                                                       | 47 |
| ภาพที่ 28 กระบวนการตัดพื้นหลังด้วยโมเดล YOLOv8 Segmentation                               | 48 |
| ภาพที่ 29 การทำ Feature Extraction ด้วยโมเดล InceptionV3                                  | 49 |
| ภาพที่ 30 ระบบแนะนำสินค้าด้วยรูปภาพจากเว็บไซต์ www.boxjourney.com                         | 49 |
| ภาพที่ 31 ตัวอย่างไฟล์ docker-compose-kong.yml                                            | 50 |
| ภาพที่ 32 Konga Dashboard                                                                 | 51 |
| ภาพที่ 33 หน้าต่าง Login เข้าสู่ระบบสำหรับผู้เลเบลข้อมูล                                  | 52 |
| ภาพที่ 34 ตัวอย่างบันทึกผลลัพธ์ข้อมูลเลเบล                                                | 53 |
| ภาพที่ 35 ตัวอย่างการเลือกบรรจุภัณฑ์ผ่านเว็บแอปพลิเคชัน                                   | 53 |
| ภาพที่ 36 ตัวอย่างผลลัพธ์การแยกวัตถุอาหารจากพื้นหลังด้วยโมเดล YOLOv8x-seg                 | 54 |
| ภาพที่ 37 ผลลัพธ์การค้นหาภาพจากค้นหาความคล้ายคลึงของพีเจอร์ภาพ                            | 59 |
| ภาพที่ 38 การเปรียบเทียบผลลัพธ์ที่ทำนายโดยโมเดลที่ฝึกด้วยข้อมูลจากนักศึกษาและผู้เชี่ยวชาญ | 60 |

## บทที่ 1

### บทนำ

#### 1.1 ที่มาและความสำคัญ

ในยุคที่เทคโนโลยีและอินเทอร์เน็ตมีบทบาทสำคัญในชีวิตประจำวันของคนทั่วโลก การซื้อขายสินค้าและบริการผ่านระบบ E-commerce [1] เป็นที่นิยมและพบเห็นอย่างแพร่หลาย การเปิดร้านค้าออนไลน์หรือเปิดธุรกิจ E-commerce มีผู้ประกอบการเพิ่มขึ้น และเกิดการแข่งขันทางธุรกิจมากขึ้น ซึ่งเป็นประโยชน์ต่อผู้บริโภคในการเลือกสรรสินค้าและบริการตามความต้องการ รวมถึงสร้างกำไรให้แก่ผู้ประกอบการผลิต

การเข้าถึงสินค้าและบริการผ่านระบบออนไลน์ช่วยสร้างความสะดวกสบายและลดเวลาในการซื้อขาย สำหรับลูกค้าที่มีความต้องการสินค้าหลากหลายหมวดหมู่ การค้นหาสินค้าที่ตรงตามความต้องการ บางครั้งอาจเป็นภาระหนึ่งในกระบวนการซื้อขาย ดังนั้น การพัฒนาระบบที่ช่วยให้ลูกค้าสามารถค้นหาสินค้าได้สะดวกและรวดเร็วมีความสำคัญอย่างยิ่ง และลดต้นทุนให้กับผู้ประกอบการผลิต

เพื่อตอบสนองความต้องการของการแข่งขันทางธุรกิจในยุคอีคอมเมิร์ซทำให้ผู้ประกอบการจำเป็นต้องมีการพัฒนาระบบให้ทันสมัยและตรงต่อความต้องการของลูกค้า ซึ่งจุดประสงค์สำคัญในการพัฒนาระบบคือ การทำระบบ Product matching [2] เพื่อให้ลูกค้าสามารถค้นหาสินค้าได้สะดวกและรวดเร็ว ระบบดังกล่าวจะช่วยให้ผู้ใช้งานสามารถค้นหาและเลือกซื้อสินค้าตามความต้องการได้อย่างสะดวกและรวดเร็ว นอกจากนี้ สามารถเสนอสินค้าที่เกี่ยวข้องหรือสินค้าที่คล้ายกันให้กับลูกค้าเพื่อเสริมสร้างประสบการณ์ในการเลือกซื้อสินค้าให้กับลูกค้าจากการใช้งานระบบ E-commerce อย่างเป็นระบบ ทำให้ลูกค้ามีความพึงพอใจในกระบวนการซื้อขายและนอกจากผู้ประกอบการสามารถสร้างข้อได้เปรียบทางการค้าที่เหนือกว่าคู่แข่ง ผู้ประกอบการสามารถลดต้นทุนโดยสามารถเพิ่มโอกาสในการขายสินค้ามากยิ่งขึ้น

งานวิจัยนี้ได้รับความอนุเคราะห์จากคุณมนู มากมณี ผู้บริหารบริษัท บราเธอร์ แอนด์ บราเธอร์ จำกัด ผู้ผลิตและจำหน่ายบรรจุภัณฑ์ภายใต้แบรนด์ Boxjourney ซึ่งเป็นผู้สนับสนุนทุนการศึกษา และให้ข้อมูลบรรจุภัณฑ์สำหรับใส่อาหารและขนมที่ใช้ในการฝึกสอนโมเดลการแนะนำบรรจุภัณฑ์ในการทำวิทยานิพนธ์ชิ้นนี้ ทางบริษัทได้จัดเตรียมชุดข้อมูลภาพบรรจุภัณฑ์ที่มีความหลากหลายทั้งรูปแบบและขนาด โดยข้อมูลเหล่านี้เป็นองค์ประกอบสำคัญในการสร้างฐานข้อมูลที่ใช้ในการแนะนำบรรจุภัณฑ์ที่เหมาะสมกับอาหารและขนมที่ลูกค้าต้องการ

จากการสัมภาษณ์ผู้บริหารของบริษัท บราเธอร์ แอนด์ บราเธอร์ จำกัด และ Admin ที่ทำหน้าที่ตอบคำถามและรับ Order จากลูกค้าทาง Chat พบว่าลูกค้าหลายรายมีความต้องการซื้อบรรจุภัณฑ์สำหรับเบเกอร์ แต่ยังไม่แน่ใจว่าจะเลือกซื้อบรรจุภัณฑ์แบบใดให้เหมาะสมกับสินค้าที่ต้องการใส่ด้วยเหตุนี้ การออกแบบและพัฒนาระบบแนะนำบรรจุภัณฑ์จึงมีความสำคัญเพื่อช่วยแก้ไขปัญหาในการตัดสินใจของลูกค้า และช่วยเพิ่มประสิทธิภาพในการขายสินค้าของบริษัท งานวิจัยนี้จึงต้องการนำเสนอระบบการแนะนำบรรจุภัณฑ์อาหารบนแพลตฟอร์มอีคอมเมิร์ซโดยใช้โมเดลวิเคราะห์ภาพอาหารที่ลูกค้านำเข้าระบบ และแสดงผลการแนะนำบรรจุภัณฑ์ที่เหมาะสม อีกทั้งยังสามารถค้นหาภาพบรรจุภัณฑ์อื่นๆ ที่มีลักษณะใกล้เคียงจากฐานข้อมูลที่มีบรรจุภัณฑ์อาหารจำนวน 2,243 ภาพเพื่อเพิ่มตัวเลือกในการตัดสินใจของลูกค้า ระบบจะประเมินความเข้ากันได้ระหว่างบรรจุภัณฑ์ที่แนะนำกับรูปภาพอาหารผ่านการสอบถามความพึงพอใจจากผู้ใช้งาน

## 1.2 วัตถุประสงค์ของการวิจัย

- 1.2.1. เพื่อศึกษาและพัฒนาเทคนิคการเรียนรู้เชิงลึก สำหรับระบบแนะนำบรรจุภัณฑ์เบเกอร์บนแพลตฟอร์มอีคอมเมิร์ซ
- 1.2.2. พัฒนาโมเดลการแบ่งส่วนภาพเพื่อตัดพื้นหลังของภาพเบเกอร์ออก
- 1.2.3. พัฒนาระบบการกรองข้อมูลจากขนาดของเบเกอร์
- 1.2.4. พัฒนาโมเดลจำแนกประเภทหลายเลเวลเพื่อแนะนำบรรจุภัณฑ์เบเกอร์โดยจะนำภาพที่ผ่านการตัดพื้นหลังออกเป็นข้อมูลเข้า
- 1.2.5. พัฒนาโมเดลการค้นหาภาพเพื่อนำภาพจากผลลัพธ์จากโมเดลจำแนกประเภทหลายเลเวลนำไปค้นหาภาพบรรจุภัณฑ์เบเกอร์อื่นที่มีความใกล้เคียง

## 1.3 ขอบเขตการวิจัย

วิทยานิพนธ์นี้กำหนดแนวคิดการสร้างต้นแบบโมเดลแนะนำบรรจุภัณฑ์จากภาพเบเกอร์ และการเรียนรู้คุณสมบัติของภาพ โดยกำหนดขอบเขตของการวิจัยแบ่งเป็น ขอบเขตของกระบวนการในงานวิจัย ขอบเขตของโมเดล ขอบเขตการทดลอง และขอบเขตการวัดและประเมินผล อธิบายได้ดังนี้

### 1.3.1 ขอบเขตของกระบวนการในงานวิจัย

1) ศึกษาและออกแบบระบบแนะนำบรรจุภัณฑ์เบเกอร์โดยใช้โมเดลการจำแนกประเภทหลายเลเบล (Multi-label Classification) และการค้นหาภาพสำหรับแพลตฟอร์มอีคอมเมิร์ซ

2) พัฒนาระบบการกรองข้อมูลด้วยขนาดของบรรจุภัณฑ์ก่อนการทำนายด้วยโมเดลการจำแนกแบบหลายเลเบล และการค้นหาภาพ

3) พัฒนาระบบการจัดเตรียมข้อมูลสำหรับการแนะนำบรรจุภัณฑ์เบเกอร์

4) พัฒนาโมเดลการเรียนรู้เชิงลึก โดยการคัดแยกวัตถุกับพื้นหลัง

### 1.3.2 ขอบเขตของโมเดล

1) โมเดลที่ใช้ในการวิจัยนี้เป็นโมเดลการเรียนรู้เชิงลึก (Deep Learning) ซึ่งประกอบด้วยโมเดลที่ผ่านการฝึกสอนล่วงหน้าบนชุดข้อมูล ImageNet ดังนี้ VGG19, VGG16, ResNet50 และ InceptionV3 โดยโมเดลเหล่านี้ถูกนำมาใช้ในการจำแนกประเภทหลายเลเบล (Multi-label Classification) สำหรับการแนะนำบรรจุภัณฑ์เบเกอร์

2) โมเดลการจำแนกประเภทหลายเลเบล (Multi-label Classification) ที่พัฒนาขึ้นจะถูกออกแบบให้สามารถรองรับคลาส (Classes) ของบรรจุภัณฑ์ที่แตกต่างกันได้มากถึง 2,243 คลาส โดยมีการแยกเลเยอร์ที่เหมาะสมสำหรับการฝึกสอนใหม่ (Fine-Tuning) ตามกลยุทธ์ที่แตกต่างกัน เพื่อปรับปรุงความแม่นยำของโมเดล

3) โมเดลการเรียนรู้เชิงลึกที่ใช้ในการวิจัยนี้จะมีการประยุกต์ใช้ Transfer Learning โดยแช่แข็ง (Freeze) เลเยอร์ที่ฝึกสอนมาแล้วบางส่วน และเพิ่มเลเยอร์ใหม่สำหรับการจำแนกบรรจุภัณฑ์เบเกอร์ เพื่อให้ระบบสามารถทำนายที่เหมาะสมกับคุณสมบัติของภาพที่ได้รับ

4) โมเดลการเรียนรู้เชิงลึกสำหรับการแยกวัตถุออกจากพื้นหลัง (Object Segmentation) จะถูกพัฒนาโดยใช้โมเดล YOLOv8 ซึ่งเป็นโมเดลที่ถูกออกแบบมาเพื่อการประมวลผลภาพแบบเรียลไทม์ (Real-time Image Processing) โดยโมเดลนี้จะถูกนำไปใช้เพื่อสกัดฟีเจอร์สำคัญจากภาพเบเกอร์

### 1.3.3 ขอบเขตการทดลอง

1) การทดลองในวิทยานิพนธ์นี้จะดำเนินการโดยใช้ชุดข้อมูลภาพของผลิตภัณฑ์เบเกอรี่หลากหลายประเภท โดยไม่มีการจำกัดเฉพาะประเภท เช่น ขนมปัง, เค้ก, คุกกี้, คัพเค้ก, โดนัท หรือมาการอง เป็นต้น ดังภาพที่ 1 ทั้งนี้ ภาพเหล่านี้จะถูกจัดเตรียมและปรับแต่งด้วยเทคนิคการแยกวัตถุออกจากพื้นหลัง เพื่อให้แบบจำลองสามารถสกัดคุณลักษณะจากภาพได้อย่างมีประสิทธิภาพ



ภาพที่ 1 ตัวอย่างชุดข้อมูลผลิตภัณฑ์เบเกอรี่

2) การทดลองจะใช้โมเดลการเรียนรู้เชิงลึก (Deep Learning) ดังนี้ VGG19, VGG16, ResNet50 และ InceptionV3 โดยฝึกสอนและทดสอบโมเดลเพื่อประเมินความสามารถในการจำแนกประเภทหลายเลเบล ของบรรจุภัณฑ์ที่เหมาะสมกับผลิตภัณฑ์เบเกอรี่

3) การทดลองจะถูกแบ่งออกเป็นหลายขั้นตอน เริ่มจากการฝึกสอนโมเดลโดยใช้ข้อมูลที่ผ่านการจัดเตรียมแล้ว ตามด้วยการทำ Fine-Tuning เพื่อปรับปรุงประสิทธิภาพของโมเดล โดยจะมีการทดสอบโมเดลด้วยชุดข้อมูลทดสอบที่ไม่เคยถูกใช้ในการฝึกสอน

4) การทดลองจะครอบคลุมถึงการประเมินผลการทำนายของโมเดล โดยใช้เมตริกต่างๆ เช่น Loss และ Confusion Matrix รวมถึงการวัดประสิทธิภาพของโมเดลในการจำแนกประเภทหลายเลเบล

5) ในการทดลองจะมีการเปรียบเทียบประสิทธิภาพระหว่างโมเดลต่างๆ เพื่อหาว่าโมเดลใดให้ผลลัพธ์ที่ดีที่สุดในบริบทของการแนะนำบรรจุภัณฑ์สำหรับผลิตภัณฑ์เบเกอรี่

6) การทดลองจะถูกออกแบบให้รองรับการขยายขอบเขตในอนาคต เช่น การนำโมเดลไปใช้กับชุดข้อมูลที่ใหญ่ขึ้น หรือการทดสอบในสถานการณ์การใช้งานจริง เพื่อให้ได้ผลลัพธ์ที่สอดคล้องกับการใช้งานในระบบอีคอมเมิร์ซ

#### 1.3.4 ขอบเขตการวัดและประเมินผล

1) การวัดและประเมินผลในวิทยานิพนธ์นี้จะมุ่งเน้นไปที่การประเมินประสิทธิภาพของโมเดลการจำแนกประเภทหลายเลเบล โดยใช้เมตริกได้แก่ F1-Score, Precision, Recall และ mAP50 เพื่อวัดความถูกต้องและประสิทธิภาพในการทำนายบรรจุกฎที่เหมاسبกับผลิตภัณฑ์เบเกอรี่

2) การประเมินผลจะรวมถึงการสร้างและวิเคราะห์ Confusion Matrix สำหรับการจำแนกประเภทหลายเลเบล เพื่อทำความเข้าใจถึงความสามารถของโมเดลในการแยกแยะบรรจุกฎที่คล้ายคลึงกัน รวมถึงการวัดประสิทธิภาพในแต่ละคลาสของบรรจุกฎ

3) การวัดและประเมินผลจะครอบคลุมถึงการใช้เทคนิค Cross-Validation เพื่อให้แน่ใจว่าผลการประเมินมีความน่าเชื่อถือและไม่ขึ้นอยู่กับการแบ่งชุดข้อมูลเพียงครั้งเดียว โดยจะมีการทดสอบโมเดลด้วยชุดข้อมูลที่ไม่เคยถูกใช้ในการฝึกสอน

4) การประเมินผลจะทำในหลายระดับ รวมถึงการประเมินผลโดยรวมของโมเดล (Overall Model Performance) การประเมินผลในระดับคลาส (Class-wise Performance) และการประเมินผลในระดับตัวอย่าง (Sample-wise Performance) เพื่อให้ได้ภาพรวมที่ชัดเจนของประสิทธิภาพโมเดล

5) การวัดผลจะรวมถึงการใช้ Learning Curve และ Loss Curve เพื่อติดตามพัฒนาการของโมเดลระหว่างการฝึกสอน ซึ่งจะช่วยให้ผู้วิจัยสามารถตัดสินใจว่าจะหยุดการฝึกสอนเมื่อใด หรือควรปรับค่า Hyperparameters อย่างไรเพื่อเพิ่มประสิทธิภาพของโมเดล

6) การทำ Fine-Tuning ของโมเดลจะถูกประเมินผลด้วยการเปรียบเทียบผลลัพธ์ก่อนและหลังการปรับแต่ง เพื่อวัดผลกระทบของการปรับแต่งเลเยอร์ในโมเดล และตรวจสอบว่าโมเดลที่ปรับแต่งแล้วสามารถเพิ่มความแม่นยำได้อย่างมีนัยสำคัญหรือไม่

## 1.4 ขั้นตอนการดำเนินการวิจัย

1.4.1. ศึกษางานวิจัยที่เกี่ยวข้องและเทคนิคทางด้านการเรียนรู้เชิงลึก การประมวลผลภาพ และการจำแนกประเภทหลายเลเบล ภายใต้ขอบเขตการวิจัยเพื่อพัฒนาโมเดลแนะนำบรรจุภัณฑ์ที่เหมาะสมสำหรับผลิตภัณฑ์เบเกอร์รี่บนแพลตฟอร์มอีคอมเมิร์ซ

1.4.2. สัมภาษณ์ผู้เชี่ยวชาญในด้านการพัฒนาโมเดลการเรียนรู้เชิงลึก และการประมวลผลภาพ เพื่อศึกษาแนวทางการออกแบบระบบแนะนำบรรจุภัณฑ์ที่มีประสิทธิภาพ รวมถึงการนำเสนอเทคนิคการปรับปรุงโมเดลจากการวิจัยที่เกี่ยวข้อง

1.4.3. ออกแบบและพัฒนาระบบแนะนำบรรจุภัณฑ์เบเกอร์รี่ โดยใช้โมเดลการจำแนกประเภทหลายเลเบล ซึ่งประกอบด้วยการสร้างโมเดลจากฐานข้อมูลภาพเบเกอร์รี่ และบรรจุภัณฑ์ที่ผ่านการสกัดฟีเจอร์จากโมเดล Image Segmentation

1.4.4. พัฒนาระบบการกรองข้อมูลโดยใช้ข้อมูลขนาดของบรรจุภัณฑ์ เพื่อช่วยในการประมวลผลภาพและปรับปรุงความแม่นยำของการทำนายบรรจุภัณฑ์ที่เหมาะสม

1.4.5. ทดลองใช้งานระบบแนะนำบรรจุภัณฑ์กับข้อมูลชุดภาพใหม่ๆ ที่เข้ามาในระบบ โดยนำข้อมูลที่ผ่านการจัดเตรียมเข้าสู่กระบวนการฝึกสอนโมเดลและทำนายผลการแนะนำบรรจุภัณฑ์

1.4.6. ทดสอบโมเดลโดยเปรียบเทียบประสิทธิภาพของโมเดลหลายประเภท เช่น VGG19, ResNet50 และ InceptionV3 รวมถึงปรับ Fine-tuning โมเดลเพื่อเพิ่มประสิทธิภาพในการทำนาย

1.4.7. นำโมเดลที่ได้รับการปรับแต่ง (Fine-tuned) มาประยุกต์ใช้ในระบบแนะนำบรรจุภัณฑ์ ซึ่งทำให้ระบบสามารถทำนายบรรจุภัณฑ์ที่เหมาะสมสำหรับผลิตภัณฑ์เบเกอร์รี่ใหม่ๆ ได้อย่างมีประสิทธิภาพ

1.4.8. ประเมินผลการทำงานของระบบโดยใช้ข้อมูลจริงจากแพลตฟอร์มอีคอมเมิร์ซ โดยตรวจสอบประสิทธิภาพการทำนายและความพึงพอใจของผู้ใช้งานที่เกี่ยวข้อง

1.4.9. สรุปผลการวิจัย โดยนำผลการทดลองและการประเมินโมเดลมาวิเคราะห์เพื่อสร้างข้อเสนอแนะในการพัฒนาระบบในอนาคต

## 1.5 ประโยชน์ที่คาดว่าจะได้รับ

1.5.1. การพัฒนาระบบแนะนำบรรจุภัณฑ์ที่เหมาะสมสำหรับผลิตภัณฑ์เบเกอร์รี่ จะช่วยเพิ่มประสิทธิภาพในการดำเนินงานของแพลตฟอร์มอีคอมเมิร์ซ ซึ่งระบบนี้สามารถช่วยลดเวลาและค่าใช้จ่ายในการเลือกบรรจุภัณฑ์ที่เหมาะสม และเพิ่มความพึงพอใจให้กับลูกค้า

1.5.2. โมเดลการจำแนกประเภทหลายเลเบล ที่พัฒนาขึ้นจะสามารถนำไปใช้ประโยชน์ในการจัดการข้อมูลที่ซับซ้อนและมีความหลากหลายมากขึ้น เช่น การแนะนำผลิตภัณฑ์ในอุตสาหกรรมอื่นๆ ที่มีความต้องการลักษณะคล้ายคลึงกัน เช่น อุตสาหกรรมอาหารและเครื่องดื่ม อุตสาหกรรมเสื้อผ้า หรืออุตสาหกรรมเครื่องสำอาง



## บทที่ 2

### ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

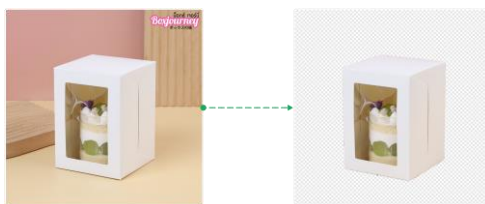
ในงานวิจัยนี้ การพัฒนาระบบแนะนำบรรจุภัณฑ์สำหรับผลิตภัณฑ์เบเกอรี่มีความจำเป็นอย่างยิ่งที่จะต้องใช้เทคนิคต่างๆ ที่เกี่ยวข้องในหลายด้าน ทั้งด้านการประมวลผลภาพ (Image Processing) เพื่อเตรียมข้อมูลภาพ การสกัดฟีเจอร์ (Feature Extraction) จากภาพสินค้า การเรียนรู้เชิงลึก (Deep Learning) เพื่อสร้างโมเดลในการจำแนกประเภทแบบหลายเลเบล (Multi-label Classification) และเทคนิคการวัดความคล้ายคลึง (Similarity Measurement) เพื่อจับคู่ฟีเจอร์ของบรรจุภัณฑ์ที่ใกล้เคียงกับสินค้าที่ได้รับการแนะนำ นอกจากนี้ยังมีการใช้เทคนิคการเรียนรู้แบบถ่ายโอน (Transfer Learning) เพื่อลดความซับซ้อนในการฝึกสอนโมเดลใหม่และเพิ่มประสิทธิภาพของการจำแนกประเภท การประยุกต์ใช้เทคนิคต่างๆ เหล่านี้ร่วมกันเป็นสิ่งที่จำเป็นในการวิจัยระบบแนะนำบรรจุภัณฑ์สำหรับผลิตภัณฑ์เบเกอรี่โดยมีรายละเอียดดังนี้

#### 2.1 การประมวลผลภาพ (Image Processing)

เป็นขั้นตอนสำคัญในการพัฒนาระบบแนะนำบรรจุภัณฑ์สำหรับผลิตภัณฑ์เบเกอรี่ ซึ่งภาพเบเกอรี่ที่ได้จากผู้ใช้งานจะถูกนำมาประมวลผลเพื่อลบพื้นหลังออก ทำให้ระบบสามารถตรวจจับและแยกวัตถุอาหารออกจากพื้นหลังและสิ่งรบกวนอื่นๆ ในภาพได้อย่างแม่นยำและมีประสิทธิภาพ โดยทฤษฎีการประมวลผลภาพที่ศึกษามีดังนี้

##### 2.1.1. การตัดพื้นหลัง (Background Removal)

การตัดพื้นหลังเป็นขั้นตอนแรกของการประมวลผลภาพ ซึ่งมีวัตถุประสงค์เพื่อแยกวัตถุหลักออกจากพื้นหลังที่ไม่จำเป็น ในงานวิจัยนี้ใช้โมเดล YOLOv8 สำหรับการแบ่งส่วนภาพและลบพื้นหลังของผลิตภัณฑ์เบเกอรี่ เมื่อผ่านกระบวนการนี้ จะได้ภาพที่มีเพียงผลิตภัณฑ์อยู่เพียงลำพังโดยไม่มีองค์ประกอบพื้นหลังที่อาจทำให้โมเดลเข้าใจผิด ซึ่งช่วยเพิ่มความแม่นยำในการสกัดฟีเจอร์และการจำแนกประเภทดังภาพที่ 2



ภาพที่ 2 การตัดพื้นหลัง (Background Removal)

### 2.1.2. การปรับแต่งสี (Color Adjustment)

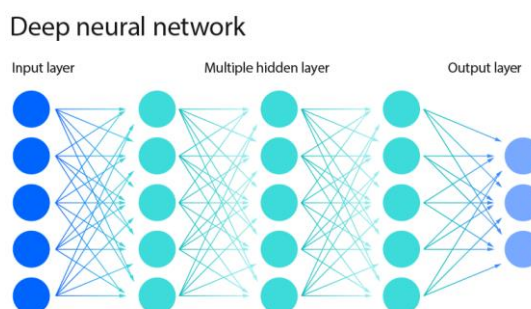
หลังจากการตัดพื้นหลังเสร็จสิ้น การปรับแต่งสีภาพถือเป็นอีกขั้นตอนสำคัญในงานวิจัยนี้ การปรับแต่งสีอาจประกอบด้วยการปรับสมดุลของสี (Color Balance), การแก้ไขความสว่าง (Brightness Correction), หรือการปรับค่า Contrast ของภาพ เพื่อให้ภาพมีความชัดเจนและสอดคล้องกับภาพอื่นๆ ที่มีการนำเข้ามายังระบบ กระบวนการนี้ช่วยให้โมเดลสามารถจำแนกรูปภาพได้แม่นยำมากขึ้น เนื่องจากสีและความชัดเจนของภาพมีผลต่อการสกัดฟีเจอร์ (Feature Extraction)

### 2.1.3. การแก้ไขภาพเบื่องต้น

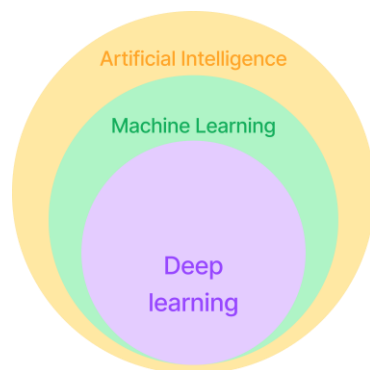
ก่อนที่จะนำภาพเข้าสู่โมเดลการจำแนกประเภทหรือการสกัดฟีเจอร์ ภาพจะต้องถูกปรับขนาดและแปลงให้อยู่ในรูปแบบที่เหมาะสมสำหรับการฝึกสอนหรือทำนาย ในงานวิจัยนี้ใช้การย่อขนาดของภาพให้มีขนาดคงที่ (เช่น 224x224 พิกเซล หรือ 299x299 พิกเซล ขึ้นอยู่กับโมเดลที่ใช้) และการแปลงภาพให้อยู่ในรูปแบบของอาร์เรย์ที่โมเดลสามารถเข้าใจได้ การปรับขนาดและแปลงภาพเบื้องต้นนี้จะทำให้ภาพทุกภาพมีรูปแบบที่สอดคล้องกัน ช่วยลดความแตกต่างในกระบวนการฝึกสอนโมเดล

## 2.2 การเรียนรู้เชิงลึก (Deep Learning)

การเรียนรู้เชิงลึก เป็นเทคนิคหนึ่งในสาขาการเรียนรู้ของเครื่อง (Machine Learning) ที่มีการใช้งานอย่างแพร่หลายในหลากหลายด้าน เช่น การประมวลผลผลภาพ การจดจำเสียง และการประมวลผลภาษาธรรมชาติ โดย Deep Learning เป็นวิธีการฝึกสอนโมเดลที่จำลองการทำงานของสมองมนุษย์ด้วยโครงข่ายประสาทเทียม (Deep Neural Networks) ที่มีโครงสร้างซับซ้อนซึ่งมีหลายชั้นต่อกัน ดังภาพที่ 3 โดย Deep Learning เป็นเซตย่อยของ Machine Learning ซึ่งเป็นหนึ่งในฟังก์ชันของ Artificial Intelligence (AI) ดังภาพที่ 4



ภาพที่ 3 Deep Neural Network [3]



ภาพที่ 4 การเรียนรู้เชิงลึก (Deep Learning)

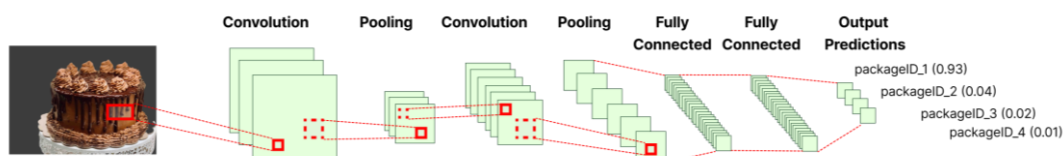
## 2.3 โครงข่ายประสาทแบบคอนโวลูชัน (Convolution Neural Network: CNN)

โครงข่ายประสาทแบบคอนโวลูชัน [4] เป็นรูปแบบหนึ่งของโครงข่ายประสาทเทียมที่ได้รับความนิยมในการประมวลผลภาพ เนื่องจาก CNN มีความสามารถในการสกัดฟีเจอร์จากข้อมูลภาพที่ซับซ้อนได้อย่างมีประสิทธิภาพ โครงสร้างพื้นฐานของ CNN ประกอบด้วยเลเยอร์หลายชั้นที่ทำหน้าที่สกัดฟีเจอร์จากข้อมูลนำเข้า ซึ่งช่วยให้ระบบสามารถจดจำลักษณะเฉพาะของภาพ เช่น ขอบเขต รูปทรง และรูปแบบต่าง ๆ ในภาพดังภาพที่ 5 ซึ่งโครงสร้างพื้นฐานของ CNN ประกอบด้วยชั้นหลัก ๆ ดังนี้

2.3.1. Convolutional Layer ชั้นคอนโวลูชันเป็นชั้นแรกที่สำคัญใน CNN โดยทำหน้าที่สกัดฟีเจอร์จากภาพผ่านการคอนโวลูชัน (Convolution) ซึ่งเป็นกระบวนการที่ใช้ Kernel หรือ Filter เลื่อนผ่านภาพเพื่อคำนวณและสกัดฟีเจอร์ต่าง ๆ เช่น ขอบของวัตถุในภาพ ข้อมูลที่ได้จาก Convolutional Layer จะเป็น Feature Map ซึ่งแสดงข้อมูลของฟีเจอร์ที่ถูกสกัดออกมา

2.3.2. Pooling Layer ชั้น Pooling เป็นชั้นที่ลดขนาดของ Feature Map ที่ได้จาก Convolutional Layer เพื่อช่วยลดความซับซ้อนและจำนวนพารามิเตอร์ในโมเดล กระบวนการที่ใช้บ่อยคือ Max Pooling ซึ่งจะเลือกค่าที่มีมากที่สุดในแต่ละพื้นที่ย่อยของ Feature Map การลดขนาดนี้ช่วยลดการคำนวณและเพิ่มประสิทธิภาพของโมเดล โดยยังคงลักษณะสำคัญของฟีเจอร์เอาไว้

2.3.3. Fully Connected Layer หลังจากผ่านชั้นคอนโวลูชันและชั้น Pooling แล้ว ข้อมูลจะถูกส่งต่อไปยัง Fully Connected Layer ซึ่งเป็นชั้นที่เชื่อมโยงกันทั้งหมด โดยทำหน้าที่รวมฟีเจอร์ที่ได้จากชั้นก่อนหน้าและนำไปใช้ในการทำนายผล เช่น การจำแนกประเภทของภาพ



ภาพที่ 5 โครงข่ายประสาทเทียมแบบคอนโวลูชัน (CNN)

โดยงานวิจัยที่เกี่ยวข้องกับการประมวลผลภาพและการใช้โมเดล CNN ในการจำแนกภาพ ดังตัวอย่างคุณ Saranya และ Durga [5] ได้นำเสนอการตรวจสอบสิ่งเจือปนในภาพผลิตภัณฑ์อาหารด้วย CNN โดย Zou และ Liu [6] ได้เสนอวิธีการประมวลผลภาพเพื่อการจดจำวัตถุ โดยได้ใช้เทคนิคการสกัดฟีเจอร์ของภาพจากการประมวลผลเบื้องต้น เช่น การตรวจจับขอบ การแยกสี RGB ซึ่งจากงานวิจัยที่เกี่ยวข้องมีการใช้โมเดลที่สามารถใช้สกัดฟีเจอร์จากภาพและนำไปใช้ในการจำแนกประเภทได้อย่างแม่นยำ

ซึ่งจะเห็นได้ว่า CNN ถูกนำไปใช้ในงานประมวลผลภาพที่หลากหลาย เช่น การจำแนกรูปภาพ การตรวจจับวัตถุ และการแบ่งส่วนภาพ ในงานวิจัยนี้ CNN ถูกใช้ในการสกัดฟีเจอร์จากภาพเบเกอรี่และบรรจุภัณฑ์เพื่อใช้ในการแนะนำบรรจุภัณฑ์ที่เหมาะสมกับผลิตภัณฑ์ CNN ช่วยให้สามารถสกัดฟีเจอร์ที่ซับซ้อนและลึกซึ่งในภาพได้ โดยฟีเจอร์ที่สกัดได้จะถูกนำไปใช้ในการทำ Multi-label Classification และวัดความคล้ายคลึงด้วยเทคนิค Cosine Similarity เพื่อให้ได้บรรจุภัณฑ์ที่มีความใกล้เคียงกัน

## 2.4 การตรวจจับวัตถุด้วย YOLO (You Only Look Once)

YOLO [7] เป็นหนึ่งในเทคนิคที่ได้รับความนิยมสำหรับการตรวจจับวัตถุ (Object Detection) ด้วยความสามารถที่โดดเด่นในด้านความเร็วและความแม่นยำ YOLO ถูกพัฒนาในปี 2015 โดย Redmon, Farhadi และคณะ [8] YOLO มีคุณสมบัติในการตรวจจับวัตถุแบบเรียลไทม์ได้อย่างมีประสิทธิภาพ ซึ่งทำให้เหมาะสมสำหรับงานด้านการตรวจจับวัตถุในหลายแอปพลิเคชัน เช่น การวิเคราะห์วิดีโอ การควบคุมหุ่นยนต์ และการประมวลผลภาพในระบบตรวจจับต่างๆ การตรวจจับวัตถุ (Object Detection) เป็นเทคนิคหนึ่งที่สำคัญในงานด้านการประมวลผลภาพ โดยมีวัตถุประสงค์ในการระบุวัตถุในภาพ พร้อมทั้งหาพิกัดขอบเขต (Bounding Box) ของวัตถุเหล่านั้น โดยเทคนิคที่ได้รับความนิยมและมีประสิทธิภาพสูงในการตรวจจับวัตถุ

เนื่องจากสามารถตรวจจับวัตถุหลายชิ้นในภาพพร้อมกันได้ในขั้นตอนเดียว โมเดลนี้ได้รับการพัฒนาอย่างต่อเนื่องในหลายเวอร์ชัน ซึ่งแต่ละเวอร์ชันมีการปรับปรุงให้ดีขึ้นเรื่อย ๆ ในด้านต่าง ๆ

เช่น ความเร็ว ความแม่นยำ และความสามารถในการใช้งานในหลายบริบทของการประมวลผลภาพ ซึ่งในปัจจุบัน ได้มีการปรับปรุงจนถึงปัจจุบันคือเวอร์ชัน YOLO11 [9] โดยมีประวัติการพัฒนา

2.4.1. YOLOv1 (2015) [8] เปิดตัวครั้งแรกในปี 2015 เน้นการตรวจจับวัตถุด้วยความเร็วสูง จากการประมวลผลภาพในขั้นตอนเดียว ซึ่งทำให้ YOLO มีความได้เปรียบในด้านความเร็วเมื่อเทียบกับเทคนิคอื่นๆในช่วงเวลานั้น

2.4.2. YOLOv2 (2016) [10] มีการปรับปรุงด้วยการใช้ Batch Normalization, Anchor Boxes, และ Dimension Clusters ทำให้มีความแม่นยำสูงขึ้น

2.4.3. YOLOv3 (2018) [11] เพิ่มการใช้ Backbone Network ที่มีประสิทธิภาพมากขึ้น รองรับ Anchor หลายตัว และใช้ Spatial Pyramid Pooling ช่วยเพิ่มความสามารถในการตรวจจับวัตถุในหลายขนาด

2.4.4. YOLOv4 (2020) [12] แนะนำการใช้ Mosaic Data Augmentation, การใช้ Anchor-Free Detection Head, และ Loss Function แบบใหม่

2.4.5. YOLOv5 (2021) [13] ปรับปรุงด้านประสิทธิภาพและเพิ่มคุณสมบัติต่าง ๆ เช่น Hyperparameter Optimization, การติดตามผลการทดลอง, และการ Export อัตโนมัติในหลายรูปแบบ

2.4.6. YOLOv6 (2022) [14] พัฒนาโดย Meituan และใช้งานในหุ่นยนต์ส่งของอัตโนมัติ

2.4.7. YOLOv7 (2022) [15] เพิ่มความสามารถใหม่ ๆ เช่นการตรวจจับ Pose จาก Dataset ของ COCO Keypoints [16]

2.4.8. YOLOv8 (2024) [17] พัฒนาโดย Ultralytics, YOLOv8 ได้เพิ่มฟีเจอร์และการปรับปรุงเพื่อเพิ่มประสิทธิภาพและความยืดหยุ่นในการใช้งาน รองรับการประมวลผลหลายภารกิจ เช่น การตรวจจับ, การแยกส่วน, การจำแนกประเภท และการคาดคะเนการเคลื่อนไหว (Pose Estimation)

2.4.9. YOLOv9 [18] แนะนำเทคนิค Programmable Gradient Information (PGI) และ Generalized Efficient Layer Aggregation Network (GELAN) เพื่อเพิ่มความแม่นยำและประสิทธิภาพ

2.4.10. YOLOv10 [19] พัฒนาโดยนักวิจัยจากมหาวิทยาลัย Tsinghua ซึ่งแนะนำ End-to-End Head ที่สามารถตัดขั้นตอน Non-Maximum Suppression (NMS) [20] ออกได้

2.4.11. YOLOv11 [21] เวอร์ชันล่าสุดจาก Ultralytics พัฒนาสำหรับการตรวจจับ การแยกส่วนภาพ และการประมวลผลหลายงานพร้อมกัน มีความสามารถสูงสุดในการตรวจจับวัตถุ การจำแนกประเภท การติดตาม และการคาดคะเนการเคลื่อนไหว

ในงานวิจัยนี้ โมเดล YOLOv8 ถูกนำมาใช้ในกระบวนการ Image Segmentation ดังภาพที่ 5 เพื่อตรวจจับและแยกวัตถุออกจากพื้นหลัง ก่อนที่จะนำภาพไปผ่านการสกัดฟีเจอร์ (Feature Extraction) โดยการแยกวัตถุด้วย YOLO นี้ช่วยให้สามารถประมวลผลภาพอาหารได้ชัดเจนมากขึ้น ไม่ถูกรบกวนจากองค์ประกอบอื่นๆ เช่น พื้นหลัง ทำให้การแนะนำบรรจุภัณฑ์ที่เหมาะสมมีประสิทธิภาพและแม่นยำมากยิ่งขึ้น

ตัวอย่างงานวิจัยของ Cao, Su, Geng และ Wang [22] ได้นำเสนอการใช้ YOLO-SF ซึ่งเป็นการผสมผสานเทคโนโลยีการแยกวัตถุเชิงอินสแตนซ์กับอัลกอริธึมการตรวจจับวัตถุ YOLOv7-Tiny เพื่อเพิ่มความแม่นยำ โดยได้ปรับปรุง Back Bone ด้วยโมดูล MobileViTv2 เพื่อลดจำนวนพารามิเตอร์ และเพิ่มความสามารถในการดึงฟีเจอร์ของโมเดล อีกทั้งยังได้แนะนำการใช้ Varifocal Loss เพื่อแก้ไขปัญหาการวางตำแหน่งวัตถุที่ไม่ถูกต้องในภาพที่มีขอบวัตถุ และในงานวิจัยของ Jindal, Raj, Saranya และ Sundarabalan [23] ได้แสดงให้เห็นถึงศักยภาพของแบบจำลอง YOLO ในการตรวจจับวัตถุ โดยเฉพาะเครื่องบินขนาดเล็กจากภาพถ่ายดาวเทียมที่มีความซับซ้อนของพื้นหลังและความคมชัดที่ไม่สมบูรณ์ ในงานวิจัยดังกล่าว ได้นำเสนอการเปรียบเทียบประสิทธิภาพของแบบจำลอง YOLOv4, YOLOv4 Tiny และ YOLOv5 โดยใช้ชุดข้อมูลภาพเครื่องบินจากดาวเทียม (CGI Planes in Satellite Imagery) พร้อมการปรับปรุงข้อมูลด้วยเทคนิคการเพิ่มข้อมูลภาพ (Image Augmentation) เพื่อแก้ไขปัญหาข้อมูลไม่เพียงพอ ซึ่งเป็นแนวทางที่สามารถประยุกต์ใช้ในงานวิจัยนี้ ที่เน้นการตรวจจับวัตถุขนาดเล็กและพื้นที่ซับซ้อนเช่นกัน

## 2.5 ความคล้ายคลึงของโคไซน์ (Cosine Similarity)

Cosine Similarity [24] เป็นเทคนิคที่ใช้วัดความคล้ายคลึงกันระหว่างเวกเตอร์สองเวกเตอร์ โดยพิจารณาจากมุมที่เกิดขึ้นระหว่างเวกเตอร์ทั้งสอง มากกว่าที่จะใช้ระยะทางระหว่างจุดในเวกเตอร์ นั่นๆ วิธีนี้ถูกนำมาใช้แพร่หลายในงานประมวลผลข้อมูลเชิงเวกเตอร์ เช่น การค้นหาความคล้ายคลึงของข้อความ ภาพ หรือฟีเจอร์ต่างๆ

Cosine Similarity คำนวณได้ดังสมการที่ (1)

$$\text{cosine similarity} = S_c(A, B) := \cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2 \cdot \sum_{i=1}^n B_i^2}} \quad (1)$$

โดยที่

- $A$  และ  $B$  คือเวกเตอร์ที่ต้องการเปรียบเทียบ
- $A \cdot B$  คือการคูณเชิงจุด (Dot Product) ของเวกเตอร์ทั้งสอง
- $\|A\|$  และ  $\|B\|$  คือขนาดความยาว (Magnitude) ของเวกเตอร์ทั้งสอง

ค่าของ Cosine Similarity จะอยู่ในช่วงระหว่าง -1 ถึง 1

- ค่า 1 หมายถึงเวกเตอร์ทั้งสองมีทิศทางเดียวกันหรือคล้ายกันมาก
- ค่า 0 หมายถึงเวกเตอร์ทั้งสองมีมุมฉากกันหรือไม่สัมพันธ์กันเลย
- ค่า -1 หมายถึงเวกเตอร์ทั้งสองมีทิศทางตรงกันข้ามกัน

ในบางกรณีเช่น การวิเคราะห์ข้อมูลที่ต้องการค่าที่ไม่เป็นลบ ค่า Cosine Similarity จะถูกจำกัดอยู่ในช่วงระหว่าง 0 ถึง 1 โดยการนำ Cosine Similarity มาใช้ในการวิจัยนี้ถูกนำมาใช้ในการคำนวณความคล้ายคลึงกันระหว่างฟีเจอร์ของบรรจุภัณฑ์ที่ได้จากการสกัดฟีเจอร์ (Feature Extraction) โดยใช้โมเดล InceptionV3 โมเดลจะแปลงรูปภาพของบรรจุภัณฑ์ให้อยู่ในรูปแบบเวกเตอร์ แล้วคำนวณ Cosine Similarity เพื่อเปรียบเทียบฟีเจอร์ของบรรจุภัณฑ์แต่ละชิ้นกับบรรจุภัณฑ์ในฐานข้อมูล ผลลัพธ์จะเป็นคะแนนความคล้ายคลึงที่แสดงว่าบรรจุภัณฑ์ไหนมีลักษณะที่ใกล้เคียงกับบรรจุภัณฑ์ที่ผู้ใช้ต้องการมากที่สุด

## 2.6 การขยายข้อมูล (Data Augmentation)

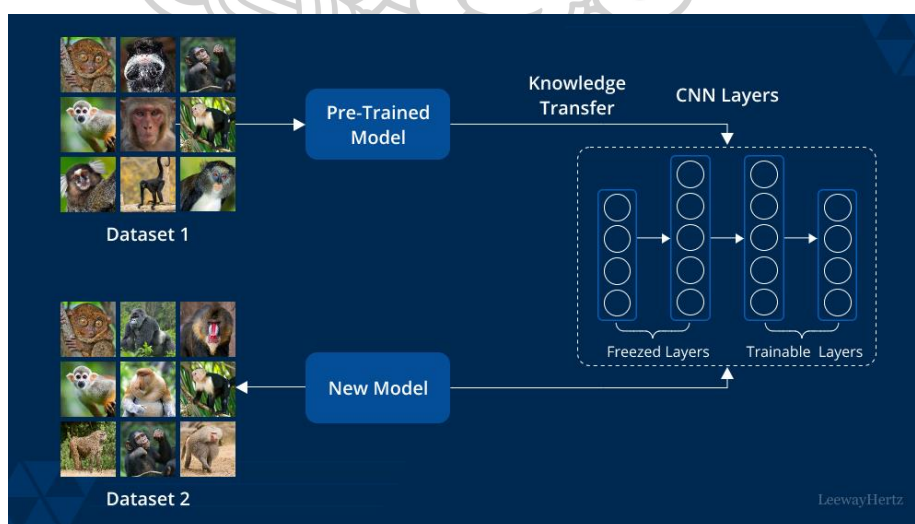
การทำ Data Augmentation [25] หรือ การขยายข้อมูล เป็นเทคนิคหนึ่งที่ถูกใช้ในงานด้านการเรียนรู้เชิงลึก (Deep Learning) เพื่อเพิ่มความหลากหลายของข้อมูลในชุดข้อมูลฝึกสอนโดยไม่ต้องรวบรวมข้อมูลใหม่ การทำ Augmentation ช่วยลดปัญหาความไม่สมดุลของข้อมูล (Data Imbalance) และเพิ่มความสามารถของโมเดลในการรับมือกับข้อมูลที่หลากหลายมากขึ้น ในงานวิจัยนี้ การทำ Augmentation ถูกนำมาใช้เพื่อเพิ่มชุดข้อมูลที่มีอยู่โดยการปรับแต่งภาพในลักษณะต่าง ๆ เช่น การหมุนภาพ (Rotation) และ การยืดและหดภาพ (Scaling)

## 2.7 Transfer Learning

เป็นแนวทางที่มีความสำคัญอย่างยิ่งในด้านการเรียนรู้ของเครื่อง โดยเฉพาะในงานที่ข้อมูลมีจำนวนจำกัด หรือที่เรียกว่า "Small Data Learning [26]" ซึ่งการเรียนรู้แบบนี้ช่วยให้ระบบปัญญาประดิษฐ์สามารถปรับใช้ความรู้จากงานหนึ่งไปสู่อีกงานหนึ่งได้โดยไม่ต้องเริ่มฝึกสอนโมเดลใหม่จากข้อมูลขนาดใหญ่เหมือนในกรณีของการเรียนรู้แบบดั้งเดิม

ตั้งแต่การเสนอแนวคิดเรื่อง Transfer Learning โดย Stevo Bozinovski และ Ante Fulgosi ในปี 1976 [27] แนวทางนี้ได้รับความสนใจอย่างกว้างขวางในวงการวิชาการ Transfer Learning ช่วยแก้ปัญหาคriticalของการเรียนรู้ของเครื่องที่ต้องการข้อมูลจำนวนมาก โดยทำให้โมเดลสามารถปรับตัวและเรียนรู้จากข้อมูลใหม่ที่มีลักษณะคล้ายคลึงกับงานที่เคยฝึกสอนมาก่อน การประยุกต์ใช้ Transfer Learning ทำให้โมเดลสามารถเรียนรู้จากข้อมูลน้อยลง แต่ยังคงมีความแม่นยำ

เป็นแนวคิดที่ถูกใช้ในงานด้าน Deep Learning โดยเฉพาะในกรณีที่มีข้อมูลจำกัด การทำ Transfer Learning ช่วยให้สามารถนำความรู้ที่โมเดลได้รับจากการฝึกสอนบนชุดข้อมูลขนาดใหญ่มาแล้ว (Pre-trained Model) เช่น ImageNet มาปรับใช้กับงานใหม่ได้โดยไม่ต้องเริ่มต้นฝึกโมเดลใหม่ทั้งหมด ซึ่งช่วยประหยัดเวลาและทรัพยากรในการฝึกสอน และเพิ่มประสิทธิภาพการเรียนรู้ของโมเดล ดังภาพที่ 6



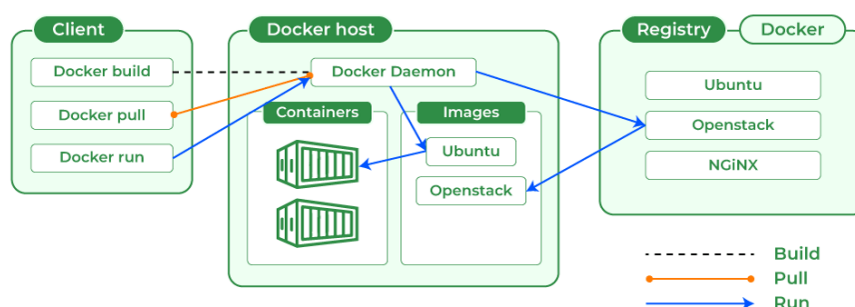
ภาพที่ 6 Transfer Learning [28]

## 2.8 Docker

Docker [29] เป็นหนึ่งในเทคโนโลยีที่ช่วยแก้ปัญหาการทำงานที่แตกต่างกันของซอฟต์แวร์บนเครื่องที่แตกต่างกัน หรือที่เรียกว่า "it works on my machine" problem ซึ่งเป็นปัญหาที่พบบ่อยในระหว่างการพัฒนาแอปพลิเคชัน Docker ทำให้สามารถสร้างสภาพแวดล้อมสำหรับแอปพลิเคชันได้อย่างครบถ้วนในรูปของ container โดยไม่ต้องกังวลเกี่ยวกับความแตกต่างของระบบปฏิบัติการหรือไลบรารีที่ติดตั้งอยู่ในแต่ละเครื่อง

### 2.8.1. สถาปัตยกรรมของ Docker

Docker ใช้สถาปัตยกรรมแบบ client-server โดยที่ Docker client ทำหน้าที่สื่อสารกับ Docker daemon ซึ่งจะจัดการกระบวนการสร้าง (build) รัน (run) และจัดส่ง (distribute) คอนเทนเนอร์ โดย Docker client และ Docker daemon สามารถรันบนระบบเดียวกัน หรือสามารถเชื่อมต่อกันได้จากระยะไกลผ่าน REST API ผ่าน UNIX socket หรือเครือข่าย ดังภาพที่ 7



ภาพที่ 7 Architecture of Docker [30]

โดยส่วนประกอบหลักในสถาปัตยกรรมของ Docker มี

- 1) Docker Daemon ทำหน้าที่จัดการบริการทั้งหมดที่เกี่ยวข้องกับ Docker เช่นการจัดการ objects ต่างๆ อย่างเช่น images, containers, networks, และ volumes โดยใช้คำขอ (API requests) จาก Docker client
- 2) Docker client เป็นส่วนที่ผู้ใช้ Docker ใช้ในการโต้ตอบกับระบบ โดยผู้ใช้สามารถใช้คำสั่งต่างๆ เช่น docker build, docker pull, และ docker run เพื่อติดต่อกับ daemon ผ่าน Docker API
- 3) Docker host คือเครื่องที่รันคอนเทนเนอร์หลายตัว ประกอบด้วย Docker daemon, images, containers, networks, และ storage

- 4) Docker registry เป็นสถานที่จัดเก็บ Docker images ทั้งหมด โดยมี Docker Hub เป็น public registry ที่ผู้ใช้ทุกคนสามารถเข้าถึงได้ นอกจากนี้ยังสามารถสร้าง private registry ได้อีกด้วย การดึง images จาก registry จะใช้คำสั่ง docker pull และการส่ง images เข้า registry ใช้คำสั่ง docker push

## 2.8.2. Docker Objects

Docker มี object หลายชนิดที่ใช้ในการสร้างและจัดการ container ดังนี้

- 1) Docker Images เป็นเทมเพลตแบบอ่านอย่างเดียว (read-only template) ที่มีคำสั่งสำหรับการสร้าง container การทำงานร่วมกันระหว่างนักพัฒนาใน Docker ทำได้สะดวกยิ่งขึ้นด้วยการใช้ images
- 2) Docker Containers เป็น instance ของ Docker images ที่สามารถเริ่มต้น หยุด ลบ หรือย้าย container ได้ผ่าน Docker API หรือ CLI คอนเทนเนอร์สามารถเข้าถึงทรัพยากรที่ถูกกำหนดใน image เท่านั้น ยกเว้นมีการกำหนดการเข้าถึงเพิ่มเติม
- 3) Docker Storage สามารถเก็บข้อมูลในชั้นเขียนของคอนเทนเนอร์ (writable layer) ได้ โดย Docker ใช้ storage driver เพื่อควบคุมและจัดการการจัดเก็บข้อมูล

Docker จึงมีบทบาทสำคัญในการพัฒนาแอปพลิเคชันโดยเฉพาะในด้านการจัดการสภาพแวดล้อมที่มีความหลากหลายและการพัฒนาในลักษณะที่ต้องการความสอดคล้องกับสภาพแวดล้อมจริง (production environment) นักพัฒนาสามารถใช้ Docker เพื่อสร้างและทดสอบแอปพลิเคชันในสภาพแวดล้อมจำลองที่เหมือนกับ production ได้อย่างรวดเร็วและง่ายดาย ซึ่งช่วยลดปัญหาความไม่เข้ากันระหว่างระบบพัฒนาและระบบจริงได้อย่างมาก

นอกจากนี้ Docker ยังถูกนำมาใช้ในกระบวนการ Continuous Integration/Continuous Deployment (CI/CD) โดยช่วยสร้าง pipeline ที่ทำงานได้อย่างอัตโนมัติและมีประสิทธิภาพ นักพัฒนาสามารถทดสอบการเปลี่ยนแปลงของโค้ดใน container ที่มีความสอดคล้องกับ production ก่อนการนำไปใช้งานจริงได้ ส่งผลให้สามารถลดเวลาในการปรับปรุงระบบและลดความเสี่ยงในการ deploy

และ Docker ยังเป็นเครื่องมือที่สำคัญในกระบวนการขยายระบบ (scalability) โดยเฉพาะในโครงสร้างของ microservices ซึ่งแต่ละบริการสามารถถูกสร้างขึ้นเป็นคอนเทนเนอร์แยกจากกัน การใช้ Docker ในการพัฒนาแอปพลิเคชันในลักษณะนี้ช่วยให้การจัดการ การขยาย และการปรับเปลี่ยนบริการต่างๆ เป็นไปได้อย่างคล่องตัวและมีประสิทธิภาพ

## 2.9 งานวิจัยที่เกี่ยวข้อง

ในปัจจุบัน การใช้เทคโนโลยีการเรียนรู้เชิงลึก (Deep Learning) และโครงข่ายประสาทเทียม (Convolutional Neural Networks: CNN) ได้กลายเป็นวิธีการหลักในการพัฒนาระบบแนะนำสินค้า ระบบค้นหาสินค้าแบบอิงเนื้อหา และการจำแนกประเภทสินค้าภาพโดยใช้ฟีเจอร์ของภาพ งานวิจัยที่เกี่ยวข้องจำนวนมากได้เสนอวิธีการต่างๆ ในการประยุกต์ใช้เทคโนโลยีเหล่านี้เพื่อให้ได้ผลลัพธ์ที่แม่นยำ รวดเร็ว และตอบสนองต่อความต้องการของผู้ใช้

2.9.1 Kumar และคณะ [31] งานวิจัยนี้มุ่งเน้นไปที่การสร้างระบบแนะนำสินค้าบนพื้นฐานของรูปภาพ โดยใช้เทคนิคการเรียนรู้เชิงลึก (Deep Learning) และการเรียนรู้แบบถ่ายโอน (Transfer Learning) ผ่านโมเดล ResNet-50 ซึ่งเป็นโครงข่ายประสาทเทียมเชิงลึก (Convolutional Neural Network: CNN) ที่มีประสิทธิภาพสูงในการสกัดฟีเจอร์จากภาพ วิธีการที่นำเสนอในงานวิจัยนี้ประกอบด้วยการสกัดฟีเจอร์จากภาพสินค้า แล้วเปรียบเทียบความคล้ายคลึงระหว่างภาพโดยใช้เทคนิคต่างๆ เช่น Euclidean distance และ Cosine similarity เพื่อค้นหาสินค้าที่ใกล้เคียงกับที่ผู้ใช้งานเข้ามา ระบบจะแนะนำผลิตภัณฑ์ที่มีความคล้ายคลึงกัน โดยงานวิจัยนี้แสดงถึงประสิทธิภาพของ ResNet-50 ในการสร้างระบบแนะนำสินค้าบนพื้นฐานของรูปภาพ โดย ResNet-50 สามารถสกัดฟีเจอร์ที่สำคัญจากภาพได้อย่างแม่นยำ อย่างไรก็ตาม ผู้วิจัยได้เสนอข้อแนะนำเพิ่มเติม เช่น การนำการปรับแต่งโดยรวมความสนใจของผู้ใช้ (User Preferences) มาช่วยเพิ่มประสิทธิภาพในการแนะนำสินค้า และการพัฒนาโมเดลที่มีความซับซ้อนมากขึ้นเพื่อปรับปรุงความสามารถในการขยายตัว (Scalability) ของระบบ โดยกระบวนการของระบบเริ่มจากการนำโมเดล ResNet-50 ที่ฝึกสอนล่วงหน้าไปใช้สำหรับการสกัดฟีเจอร์จากภาพ หลังจากที่ใช้ส่งภาพเข้าสู่ระบบ โมเดลจะสกัดฟีเจอร์แล้วเปรียบเทียบความคล้ายคลึงของฟีเจอร์ระหว่างภาพผู้ใช้งานกับภาพสินค้าในฐานข้อมูลด้วยเทคนิค Cosine similarity เพื่อนำเสนอสินค้าที่ตรงกับความต้องการของผู้ใช้ได้อย่างมีประสิทธิภาพ ผลการทดลองชี้ให้เห็นว่าโมเดล ResNet-50 เป็นเครื่องมือที่มีประสิทธิภาพในการสกัดฟีเจอร์จากภาพเพื่อนำไปใช้ในระบบแนะนำสินค้า โดยให้ผลลัพธ์ที่แม่นยำและมีประสิทธิภาพในการแนะนำสินค้าโดยอิงจากความคล้ายคลึงของภาพ

2.9.2 Mohammadi และคณะ [32] งานวิจัยนี้นำเสนอระบบแนะนำสินค้าแฟชั่นที่ใช้การประมวลผลแบบ content-based โดยใช้ ResNet-50 ในการสกัดฟีเจอร์จากภาพสินค้าแฟชั่นแต่ละชิ้น จากนั้นเปรียบเทียบความคล้ายคลึงของฟีเจอร์ระหว่างภาพสินค้าที่ผู้ใช้งานค้นหากับสินค้าที่มีอยู่ในร้านค้า เพื่อแนะนำสินค้าที่คล้ายกัน ระบบนี้ได้รับการออกแบบมาเพื่อใช้งานในแพลตฟอร์ม E-commerce ที่ต้องการแนะนำสินค้าที่คล้ายคลึงกันภายในร้านเดียวกัน โดยมุ่งเน้นไปที่การค้นหาสินค้าที่อยู่ในประเภทเดียวกัน (Single-Item Recommender) การพัฒนาระบบในงานวิจัยนี้ได้ปรับเปลี่ยนโครงสร้างจาก ResNet-101 มาเป็น ResNet-50 เพื่อเพิ่มประสิทธิภาพโดยไม่กระทบต่อผลลัพธ์สุดท้าย ระบบนี้สามารถจัดการกับข้อมูลภาพจากร้านค้าที่เป็นพื้นหลังธรรมดา และยังมีการปรับแต่งเพื่อให้สามารถทำงานได้กับชุดข้อมูล Street-to-Shop [33] โดยการเพิ่มเทคนิคการ Augmentation [25] ที่นำแฟชั่นไอเท็มไปวางบนพื้นหลังที่ซับซ้อนมากขึ้น เพื่อให้ระบบสามารถโฟกัสไปที่ตัวสินค้าแทนที่จะเป็นพื้นหลังของภาพ เทคนิคนี้ช่วยให้ระบบสามารถใช้งานได้กับภาพที่อยู่ในชีวิตจริง โดยไม่จำเป็นต้องใช้ข้อมูลจากหลายโดเมน นอกจากนี้ งานวิจัยยังมีการปรับแต่งระบบเพื่อให้สามารถปรับแต่งผลลัพธ์ตามความต้องการของผู้ใช้ โดยใช้ข้อมูลฟีเจอร์ของสินค้าที่ผู้ใช้เคยเลือกซื้อมาแนะนำสินค้าเพิ่มเติมที่ตรงกับความต้องการของผู้ใช้ โดยใช้เทคนิคการปรับน้ำหนักฟีเจอร์ (Weighted Feature) ของสินค้าที่เคยซื้อเพื่อสร้างคำแนะนำที่เป็นแบบเฉพาะบุคคล ระบบนี้มีการใช้เทคนิค Cosine similarity ในการคำนวณความคล้ายคลึงระหว่างภาพสินค้า ซึ่งสามารถประมวลผลได้อย่างรวดเร็วและให้ผลลัพธ์ที่แม่นยำ ในการประเมินผล ระบบสามารถแนะนำสินค้าแฟชั่นได้ทั้งจากภาพสินค้าที่มาจากร้านและจากภาพถนนจริง ซึ่งทำให้การใช้งานระบบนี้มีความยืดหยุ่นมากขึ้นในด้านการแนะนำสินค้าจากโดเมนต่างๆ

2.9.3 Nayak และคณะ [34] งานวิจัยนี้นำเสนอวิธีการใช้เทคนิคการเรียนรู้เชิงลึก (Deep Learning) และการใช้ Transfer Learning ในการจำแนกภาพเสื้อผ้าแฟชั่นแบบละเอียด โดยมีเป้าหมายเพื่อเพิ่มประสิทธิภาพในการค้นหาและแนะนำสินค้าที่เกี่ยวข้องกับภาพเสื้อผ้าที่ผู้ใช้งานใช้งานงานวิจัยนี้ได้นำเสนอการประยุกต์ใช้ Fine-grained Image Classification เพื่อให้สามารถจำแนกคุณสมบัติต่างๆ ของเสื้อผ้าได้อย่างละเอียด เช่น สี ลวดลาย และวัสดุ ซึ่งช่วยให้ผู้ใช้สามารถค้นหาสินค้าที่ตรงกับความต้องการโดยไม่ต้องพิมพ์คำค้นหา (Text Input) ข้อมูลที่ใช้ในการประเมินระบบนี้มาจาก Kaggle iMaterialist 2020 dataset ซึ่งมีภาพสินค้าแฟชั่นจำนวน 45,600 ภาพในการฝึกสอน และ 3,200 ภาพสำหรับการทดสอบ โดยมีการจัดหมวดหมู่เสื้อผ้าออกเป็น 46 หมวดหลัก และมีฟีเจอร์ย่อยเพิ่มเติม 294 ฟีเจอร์เพื่อช่วยให้การจำแนกมีความละเอียดมากขึ้น สถาปัตยกรรม CNN ที่

นำมาใช้ งานวิจัยนี้ได้เปรียบเทียบโครงสร้าง CNN ที่ใช้ในการจำแนกคุณสมบัติของเสื้อผ้า โดยประกอบด้วย SpineNet143, Mask R-CNN, และ Feature Pyramid Network (FPN) ซึ่งเป็นโครงสร้างที่สามารถจัดการกับปัญหาการแบ่งส่วนวัตถุ (Instance Segmentation) และการจำแนกวัตถุได้อย่างแม่นยำ ซึ่งในงานวิจัยนี้มีการใช้ โครงข่าย SpineNet143 ซึ่งเป็น Meta-Architecture ที่พัฒนาโดย Google ใช้ในการจำแนกและจัดกลุ่มขนาดของวัตถุต่างๆ โดยสามารถรักษาข้อมูลเชิงพื้นที่ของภาพไว้ในกระบวนการจำแนกได้ดี ซึ่งช่วยให้สามารถตรวจจับวัตถุที่มีขนาดเล็กและใหญ่ได้อย่างแม่นยำในระดับต่างๆ ส่วนโครงข่าย Mask R-CNN ถูกออกแบบมาเพื่อตรวจจับวัตถุและแบ่งส่วนภาพในระดับ Instance Segmentation โดยมีโครงสร้าง Backbone ที่เป็น Deep Neural Network ในสถาปัตยกรรม FPN ใช้ในการดึงฟีเจอร์จากภาพและสร้าง Feature map ที่สามารถปรับขนาดได้ตามวัตถุที่ตรวจจับได้ นอกจากนี้ยังใช้เทคนิค Cosine Similarity ในการเปรียบเทียบฟีเจอร์ของสินค้าเพื่อแนะนำสินค้าเพิ่มเติมจากภาพสินค้าที่ผู้ใช้นำเข้ามา

2.9.4 Kang และคณะ [35] งานวิจัยนี้นำเสนอวิธีการแนะนำผลิตภัณฑ์เสริมที่เข้ากันได้กับภาพจากสถานการณ์จริง (scene image) โดยแนวคิดนี้เรียกว่า "Complete the Look" ซึ่งระบบจะแนะนำผลิตภัณฑ์ที่เข้ากันได้โดยใช้ข้อมูลจากภาพของสถานการณ์จริง เช่น การแต่งตัวของบุคคลในรูปภาพที่ถ่ายบนท้องถนน (Street2Shop) เพื่อนำมาแนะนำสินค้าที่คล้ายคลึงกันที่สามารถหาซื้อได้ในร้านค้าออนไลน์ งานวิจัยนี้มุ่งเน้นไปที่การใช้ภาพจริงที่มีบริบทซับซ้อนมากกว่าภาพที่ถ่ายในสตูดิโอ ซึ่งในภาพจริงนี้อาจมีแสงเงาไม่สม่ำเสมอ องค์การถ่ายภาพไม่แน่นอน หรือมีสิ่งของรอบข้างอื่นๆ แต่สิ่งเหล่านี้จะช่วยเพิ่มข้อมูลที่สำคัญ เช่น ข้อมูลรูปร่างของผู้ใช้งาน หรือฤดูกาล ซึ่งสามารถช่วยให้การแนะนำสินค้าแม่นยำยิ่งขึ้น ในงานวิจัยนี้ใช้โมเดล ResNet-50 เพื่อสกัดฟีเจอร์จากภาพของสถานการณ์จริงและผลิตภัณฑ์ จากนั้นจึงนำฟีเจอร์เหล่านี้ไปคำนวณความเข้ากันได้โดยใช้ระยะห่าง  $\ell_2$  (Euclidean distance) เพื่อวัดความเข้ากันระหว่างฟีเจอร์ของภาพสถานการณ์จริงและฟีเจอร์ของผลิตภัณฑ์ที่คาดการณ์ว่าเข้ากันได้

2.9.5 Fadhilla และคณะ [36] งานวิจัยนี้เสนอวิธีการค้นหาภาพเสื้อผ้าที่มาจากมุมมองต่างๆ โดยใช้เทคนิคโครงข่ายประสาทเชิงลึก (Deep Neural Network) ระบบนี้ใช้ข้อมูลของประเภทและคุณสมบัติของเสื้อผ้าในการสร้างเครือข่ายเพื่อค้นหาภาพเสื้อผ้าที่คล้ายคลึงกัน การสกัดฟีเจอร์จากภาพเสื้อผ้าใช้โครงข่ายประสาทเชิงลึก ResNet50 ร่วมกับเลเยอร์ Fully Connected (FC) โดยมีการใช้เทคนิคการวัดความคล้ายคลึงแบบ Cosine Similarity เพื่อหาภาพเสื้อผ้าที่มีความคล้ายคลึงกันในชุดข้อมูลหลายมุมมองของเสื้อผ้า

2.9.6 Chang และคณะ [37] งานวิจัยนี้เสนอการพัฒนาระบบค้นหาสินค้าบนแพลตฟอร์ม E-commerce โดยใช้วิธีการค้นหาภาพที่อิงจากเนื้อหา (Content-based Image Retrieval) ซึ่งเป็นการนำเทคนิคโครงข่ายประสาทเชิงลึก (Deep Neural Network) มาใช้ในการจัดหมวดหมู่และค้นหาภาพสินค้า โดยเฉพาะการใช้โมเดล DenseNet ซึ่งมีประสิทธิภาพสูงในการสกัดฟีเจอร์จากภาพและการวัดความคล้ายคลึงระหว่างภาพด้วยวิธี Cosine Similarity นอกจากนี้ ระบบยังถูกออกแบบมาเพื่อให้สามารถค้นหาผลิตภัณฑ์ที่มีความคล้ายคลึงกันในฐานะข้อมูลได้อย่างแม่นยำ โดย

2.9.7 Islam และคณะ [38] งานวิจัยนี้นำเสนอการสำรวจการใช้เทคนิค Deep Transfer Learning สำหรับการจำแนกภาพอาหาร โดยใช้โครงข่ายประสาทเทียมเชิงลึกที่ฝึกมาล่วงหน้า (เช่น AlexNet, GoogLeNet, และ ResNet-50) ซึ่งเคยถูกฝึกมาแล้วบนชุดข้อมูล ImageNet ผู้วิจัยนำโมเดลเหล่านี้มาฝึกสอนใหม่ด้วยภาพจากฐานข้อมูลอาหาร และใช้ฟีเจอร์ที่สกัดจากชั้นกลางของ DCNNs โดยไม่ฝึกสอนเพิ่มเติม จากนั้นนำฟีเจอร์ที่ได้ไปฝึกสอนโมเดลดั้งเดิม เช่น Artificial Neural Networks (ANNs) และ Support Vector Machines (SVMs) งานวิจัยนี้ชี้ให้เห็นว่าเทคนิค Transfer Learning และการใช้ฟีเจอร์ที่สกัดจากโครงข่ายประสาทเทียมเชิงลึกที่ฝึกมาล่วงหน้ามีประสิทธิภาพในการจำแนกภาพอาหารได้อย่างมีประสิทธิภาพ และมีศักยภาพสูงในการใช้งานในแอปพลิเคชันที่เกี่ยวข้องกับการจำแนกอาหาร

2.9.8 Sharma และคณะ [39] งานวิจัยนี้กล่าวถึงการใช้เทคนิค Transfer Learning ในการสกัดฟีเจอร์หลายเลเบล จากภาพสินค้าในกลุ่มสินค้าแฟชั่น งานวิจัยนี้เน้นการสกัดฟีเจอร์จากภาพสินค้าละเอียดสูง เช่น ข้อมูลสี เนื้อผ้า ลวดลาย และลักษณะการตัดเย็บ โดยใช้โมเดลโครงข่ายประสาทเชิงลึก (CNN) ที่ได้รับการฝึกสอนมาก่อนหน้านี้ งานวิจัยนี้ชี้ให้เห็นว่า Transfer Learning ช่วยลดเวลาที่ใช้ในการฝึกสอนโมเดลและเพิ่มความสามารถในการสกัดฟีเจอร์จากข้อมูลที่มีอยู่จำนวนน้อย โดยงานวิจัยใช้โมเดลการเรียนรู้เชิงลึกที่ได้รับการฝึกสอนล่วงหน้ามาก่อน ได้แก่ InceptionV3, EfficientNetB4, Xception, NasnetLarge, และ FAttNet โมเดลเหล่านี้ถูกนำมาใช้ในการสกัดฟีเจอร์ เช่น วัสดุ สี ประเภทของแขนเสื้อ และลวดลาย งานวิจัยนี้ชี้ให้เห็นถึงความสำคัญของการใช้เทคนิคการเรียนรู้เชิงลึกและการสกัดฟีเจอร์ในการพัฒนาแอป

ผลจากการศึกษางานวิจัยที่เกี่ยวข้องดังกล่าวข้างต้น เห็นได้ว่าผู้วิจัยสามารถนำสินค้าที่ต้องการ มาผ่านระบบการค้นหาที่มีประสิทธิภาพ และนำมาประยุกต์ใช้งานได้จริง โดยในงานวิจัย

ที่ศึกษามีเทคนิคการพัฒนาระบบการค้นหามีความใกล้เคียงกันและเป็นไปในทิศทางเดียวกันกับ  
หัวข้องานวิจัยที่ผู้วิจัยศึกษาอยู่ ซึ่งงานวิจัยที่กล่าวมาสรุปได้ดังตารางที่ 1 และตารางที่ 2

ตารางที่ 1 ตารางเปรียบเทียบของแบบจำลองการเรียนรู้เชิงลึกสำหรับการแนะนำสินค้าและการ  
วิเคราะห์ภาพ

| Literature                                                                           | ResNet50 | DenseNet121 | Inceptionv3 | GoogLeNet | EfficientNetB4 | NasNetLarge | Mask R-CNN | SpineNet |
|--------------------------------------------------------------------------------------|----------|-------------|-------------|-----------|----------------|-------------|------------|----------|
| A Deep Learning Approach for Product Recommendation using ResNet-50 CNN Model [31]   | ✓        | -           | -           | -         | -              | -           | -          | -        |
| Single-Item Fashion Recommender: Towards Cross-Domain Recommendations [32]           | ✓        | -           | -           | -         | -              | -           | -          | -        |
| Fine-grained Fashion Clothing Image Classification and Recommendation [34]           | ✓        | -           | -           | -         | -              | -           | ✓          | ✓        |
| Complete the Look: Scene-Based Complementary Product Recommendation [35]             | ✓        | -           | -           | -         | -              | -           | -          | -        |
| Multi-View Clothing Image Searching Based on Deep Neural Network [36]                | ✓        | ✓           | -           | -         | -              | -           | -          | -        |
| Content-based Commodity Search Engine Using Convolutional Image Retrieval [37]       | -        | ✓           | -           | -         | -              | -           | -          | -        |
| An Exploration of Deep Transfer Learning for Food Image Classification [38]          | ✓        | -           | ✓           | ✓         | -              | -           | -          | -        |
| Transfer Learning Approach for Multilabel Fine Grain Image Attribute Extraction [39] | -        | -           | ✓           | -         | ✓              | ✓           | -          | -        |
| Our research                                                                         | ✓        | -           | ✓           | -         | -              | -           | -          | -        |

ตารางที่ 2 ตารางเทคนิคการวัดความคล้ายคลึง

| Literature                                                                           | Cosine Similarity | Euclidean Distance |
|--------------------------------------------------------------------------------------|-------------------|--------------------|
| A Deep Learning Approach for Product Recommendation using ResNet-50 CNN Model [31]   | ✓                 | -                  |
| Single-Item Fashion Recommender: Towards Cross-Domain Recommendations [32]           | ✓                 | -                  |
| Fine-grained Fashion Clothing Image Classification and Recommendation [34]           | -                 | ✓                  |
| Complete the Look: Scene-Based Complementary Product Recommendation [35]             | -                 | ✓                  |
| Multi-View Clothing Image Searching Based on Deep Neural Network [36]                | ✓                 | -                  |
| Content-based Commodity Search Engine Using Convolutional Image Retrieval [37]       | ✓                 | -                  |
| An Exploration of Deep Transfer Learning for Food Image Classification [38]          | ✓                 | -                  |
| Transfer Learning Approach for Multilabel Fine Grain Image Attribute Extraction [39] | ✓                 | -                  |
| <b>Our research</b>                                                                  | <b>✓</b>          | <b>-</b>           |



### บทที่ 3

#### ขั้นตอนและวิธีดำเนินการวิจัย

การวิจัยครั้งนี้มุ่งเน้นการพัฒนาระบบที่สามารถแนะนำบรรจุภัณฑ์เบเกอร์ที่เหมาะสม โดยใช้ภาพของผลิตภัณฑ์เบเกอร์เป็นข้อมูลตั้งต้น และเนื่องจากการวิจัยในด้านการแนะนำผลิตภัณฑ์ที่เหมาะสมกับผลิตภัณฑ์ชนิดอื่นมีความแตกต่างกับการแนะนำผลิตภัณฑ์ประเภทเดียวกัน จึงมีความท้าทายและต้องการความเข้าใจที่ลึกซึ้งสำหรับสร้างความสัมพันธ์ระหว่างผลิตภัณฑ์ที่แตกต่างกัน การวิจัยนี้จึงมุ่งเน้นการออกแบบระบบที่สามารถเรียนรู้และทำความเข้าใจความเหมาะสมระหว่างสินค้าได้ นอกจากนี้ยังรวมถึงการประยุกต์ใช้เทคนิคการเรียนรู้เชิงลึกในการประมวลผลภาพสำหรับแยกวัตถุออกจากพื้นหลัง เพื่อเพิ่มประสิทธิภาพการทำนาย รวมถึงการจัดการกับข้อมูลใหม่ที่เพิ่มเข้ามาอย่างต่อเนื่องซึ่งทำให้การเลเบล และฝึกสอนโมเดลใหม่ทำได้ยาก จึงมีการออกแบบระบบค้นหาภาพซึ่งจะทำให้ระบบสามารถแนะนำภาพที่เข้ามาใหม่ได้ รวมถึงการพัฒนาเทคนิคในการประเมินผลและปรับปรุงคุณภาพของผลลัพธ์ การวิจัยนี้จึงได้แบ่งขั้นตอนการดำเนินการวิจัยออกเป็น 12 ขั้นตอน คือ

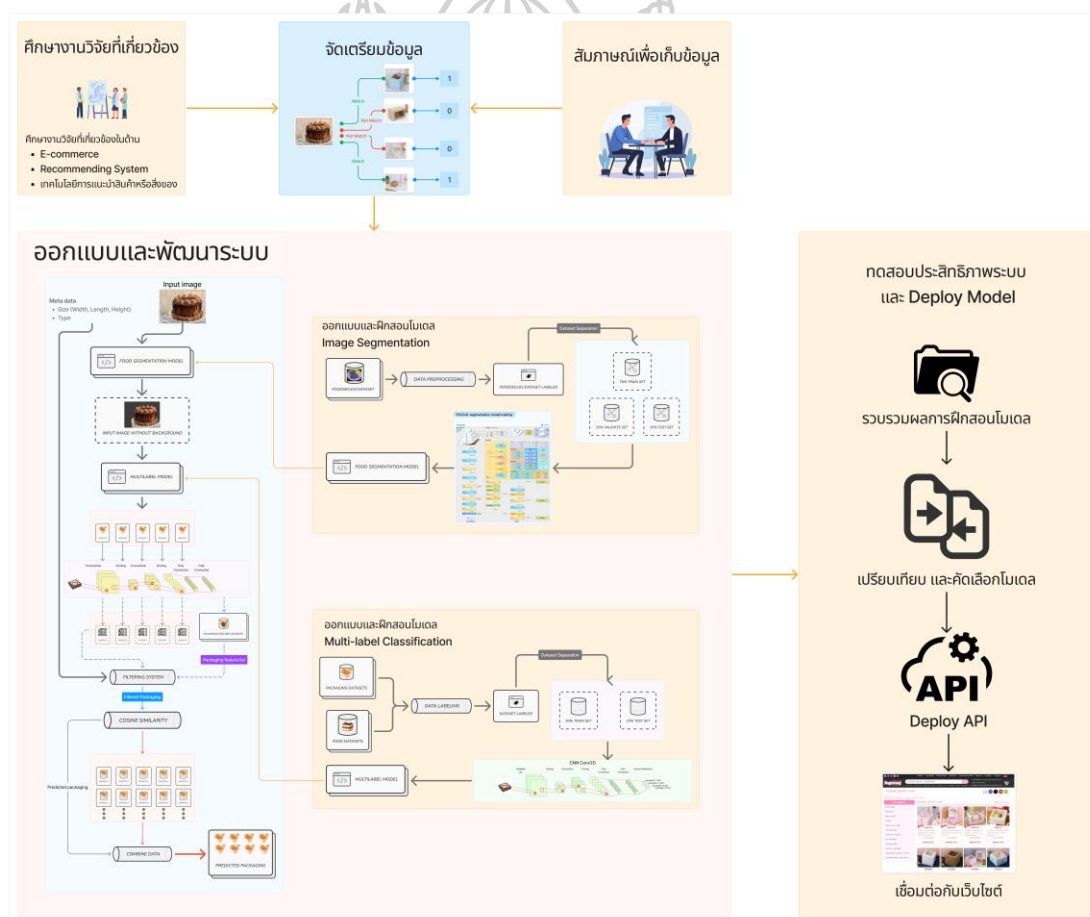
- 3.1 ศึกษางานวิจัยที่เกี่ยวข้อง
- 3.2 ออกแบบโครงสร้างระบบ
- 3.3 ออกแบบและฝึกสอนโมเดล Image Segmentation
- 3.4 การเพิ่มข้อมูลด้วยเทคนิค Data Augmentation
- 3.5 ออกแบบและฝึกสอนโมเดล Multi-label Classification
- 3.6 กระบวนการแยกส่วนและลบพื้นหลังภาพ
- 3.7 การสกัดฟีเจอร์ของภาพบรรจุภัณฑ์
- 3.8 กระบวนการคัดกรองบรรจุภัณฑ์
- 3.9 กระบวนการค้นหาภาพ
- 3.10 ประยุกต์ใช้โมเดลบนระบบ API
- 3.11 พัฒนาระบบ API Gateway
- 3.12 ประเมินผลการทำงานของระบบแนะนำบรรจุภัณฑ์

#### 3.1 ศึกษางานวิจัยที่เกี่ยวข้อง

การทบทวนและศึกษางานวิจัยที่เกี่ยวข้องกับเทคนิคการแนะนำผลิตภัณฑ์ การประมวลผลภาพ และการเรียนรู้เชิงลึก เพื่อค้นหาแนวทางที่มีความเหมาะสมสำหรับการสร้างระบบที่สามารถ

แนะนำบรรจุภัณฑ์ที่เหมาะสมกับผลิตภัณฑ์เบเกอรี่ได้อย่างแม่นยำ การศึกษานี้ครอบคลุมถึงบทความวิชาการ งานวิจัย เอกสารวิชาการ และหนังสือตำราที่เกี่ยวข้อง เพื่อระบุประเด็นสำคัญที่เป็นช่องว่างทางวิชาการ และแนวทางที่สามารถนำมาประยุกต์ใช้ได้อย่างเหมาะสมในบริบทของการพัฒนาโมเดลแบบหลายเลเบลสำหรับการแนะนำบรรจุภัณฑ์

จากการศึกษางานวิจัยที่เกี่ยวข้องนี้ กระบวนการวิจัยประกอบไปด้วยขั้นตอนต่างๆ ตามรายละเอียดในภาพที่ 8 ซึ่งแสดงถึงขั้นตอนหลักในการพัฒนาระบบแนะนำบรรจุภัณฑ์ตั้งแต่การศึกษา งานวิจัยที่เกี่ยวข้อง การออกแบบและพัฒนาระบบ ไปจนถึงขั้นตอนการนำโมเดลขึ้นสู่ระบบเพื่อทดสอบประสิทธิภาพ



ภาพที่ 8 กรอบการดำเนินการวิจัย

### 3.2 ออกแบบโครงสร้างระบบ

การออกแบบโครงสร้างระบบ Packaging Recommendation ประกอบด้วยสองโมเดลหลัก ซึ่งทำหน้าที่ร่วมกันเพื่อพัฒนากระบวนการแนะนำบรรจุภัณฑ์ที่เหมาะสมกับผลิตภัณฑ์เบเกอรี่

3.2.1. โมเดลที่ 1 Image Segmentation โมเดลแรกทำหน้าที่ในการตัดพื้นหลังที่ไม่เกี่ยวข้องออกจากภาพของเบเกอร์ เพื่อช่วยให้การสกัดฟีเจอร์จากภาพมีประสิทธิภาพมากขึ้น โดยในขั้นตอนนี้ ระบบจะใช้ YOLOv8 Segmentation ซึ่งเป็นโมเดลที่มีความสามารถสูงในการตรวจจับวัตถุและแยกวัตถุออกจากพื้นหลังได้อย่างรวดเร็วและแม่นยำ ข้อมูลที่ได้จากการประมวลผลด้วยโมเดลนี้จะถูกส่งไปยังขั้นตอนถัดไปในการวิเคราะห์ลักษณะเฉพาะของ เบเกอร์

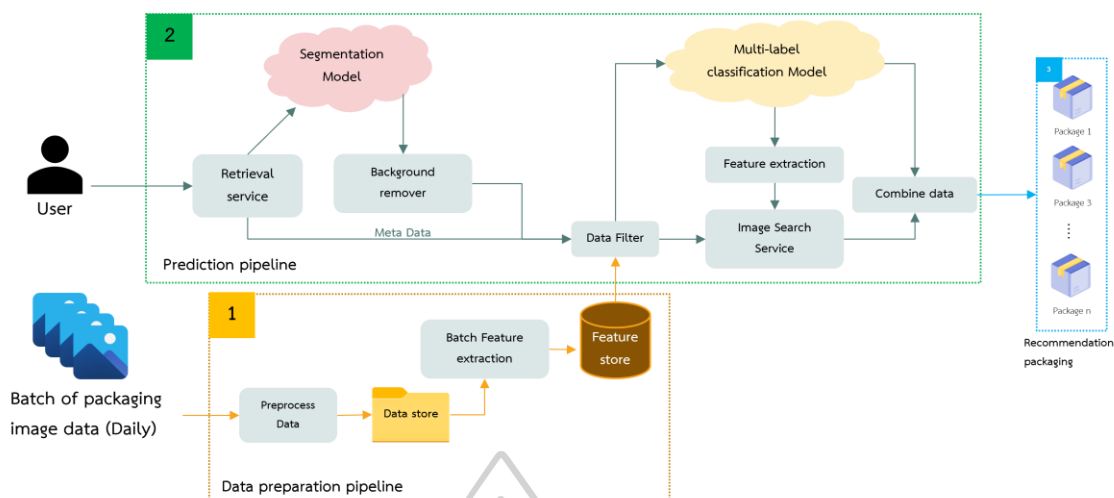
3.2.2. โมเดลที่ 2 Multi-label Classification โมเดลที่สองคือ โมเดลการจำแนกประเภทแบบหลายเลเบล ซึ่งจะทำนายผลิตภัณฑ์ที่เหมาะสมโดยใช้ข้อมูลจากการสกัดฟีเจอร์ที่ได้จากภาพที่ผ่านการตัดพื้นหลังแล้ว โมเดลนี้จะประเมินว่าบรรจุภัณฑ์ประเภทใดที่มีความเข้ากับผลิตภัณฑ์เบเกอรี่นั้นๆ ซึ่งจะช่วยในการแนะนำบรรจุภัณฑ์ที่เหมาะสมที่สุด

หลังจากที่โมเดลจำแนกประเภทแบบหลายเลเบลทำนายบรรจุภัณฑ์เสร็จสิ้น ข้อมูลที่ได้รับจะถูกนำไปวิเคราะห์เพิ่มเติม โดยใช้การสกัดฟีเจอร์ เพื่อคำนวณความคล้ายคลึงของฟีเจอร์ (Cosine Similarity [24]) ของผลิตภัณฑ์กับฐานข้อมูลบรรจุภัณฑ์ที่มีอยู่ ซึ่งข้อมูลในฐานข้อมูลจะถูกกรองตามขนาดของเบเกอร์เพื่อให้ผลลัพธ์ที่ได้มีความเหมาะสมมากยิ่งขึ้น

แต่โดยพื้นฐานของระบบอีคอมเมิร์ซ จะมีสินค้าใหม่เข้าสู่ระบบอย่างสม่ำเสมอ ซึ่งทำให้ยากต่อการเลเบลทุกครั้งที่มีข้อมูลใหม่เข้า ดังนั้นเมื่อมีข้อมูลใหม่เข้ามาจะถูกนำไปเก็บไว้ยังคลังข้อมูลชั่วคราว และในทุกวันจะมีการจัดการกับข้อมูลใหม่ในคลังข้อมูลชั่วคราวโดยส่งเข้าสู่กระบวนการสกัดฟีเจอร์และเก็บลงในฐานข้อมูลฟีเจอร์ของบรรจุภัณฑ์ เพื่อให้สินค้าใหม่มีโอกาสที่จะถูกนำไปแนะนำในระบบผ่านการคำนวณความคล้ายคลึงของฟีเจอร์

ภาพรวมของขั้นตอนทั้งหมดนี้แสดงให้เห็นถึงความเชื่อมโยงกันของการประมวลผลภาพ การสกัดฟีเจอร์ และการทำนายบรรจุภัณฑ์ที่เหมาะสม ดังภาพที่ 9 โดยแบ่งเป็น 2 ส่วนหลักดังนี้

1. การจัดเตรียมข้อมูล โดยนำข้อมูลภาพบรรจุภัณฑ์ใหม่เข้าสู่ระบบ ขั้นตอนนี้จะเก็บข้อมูลภาพบรรจุภัณฑ์ไว้ในฐานข้อมูลชั่วคราว เพื่อเตรียมสำหรับนำข้อมูลภาพบรรจุภัณฑ์ไปสกัดฟีเจอร์ และเก็บเข้าสู่ฐานข้อมูลฟีเจอร์
2. การทำนายบรรจุภัณฑ์จากภาพเบเกอร์ เป็นขั้นตอนรับภาพเบเกอร์จากผู้ไปตัดพื้นหลัง และนำภาพเบเกอร์ที่ผ่านการตัดพื้นหลังเข้าสู่โมเดลการทำนายแบบหลายเลเบล และนำผลลัพธ์ซึ่งเป็นบรรจุภัณฑ์เข้าสู่โมเดลการค้นหาภาพเพื่อหาบรรจุภัณฑ์ที่มีลักษณะใกล้เคียงกันจากฐานข้อมูลคุณลักษณะ



ภาพที่ 9 ขั้นตอนการทำงานของระบบ

### 3.3 ออกแบบและฝึกสอนโมเดล Image Segmentation

การออกแบบและฝึกสอนโมเดล Image Segmentation เป็นขั้นตอนสำคัญในการประมวลผลภาพที่ใช้ในการลบพื้นหลังที่ไม่จำเป็นออกจากภาพเบเกอรี่ โดยการแบ่งแยกวัตถุ (Object Segmentation) จะช่วยให้ระบบสามารถโฟกัสเฉพาะฟีเจอร์ของผลิตภัณฑ์ได้อย่างแม่นยำและชัดเจนมากยิ่งขึ้น สำหรับขั้นตอนนี้ได้เลือกใช้โมเดล Ultralytics YOLOv8 ซึ่งเป็น State-of-the-Art Model ที่พัฒนาขึ้นจากความสำเร็จของ YOLO รุ่นก่อน ๆ โดยเพิ่มคุณสมบัติใหม่ ๆ และปรับปรุงประสิทธิภาพให้ดียิ่งขึ้น YOLOv8 ถูกออกแบบมาให้ทำงานได้อย่างรวดเร็ว แม่นยำ และใช้งานง่ายเหมาะสำหรับการตรวจจับวัตถุ การติดตามวัตถุ การแบ่งส่วนวัตถุ การจำแนกภาพ และการประมาณการเคลื่อนไหว (Pose Estimation) ในงานต่าง ๆ ได้อย่างหลากหลาย ด้วยการออกแบบที่มีความคล่องตัว โมเดลนี้จึงเหมาะสำหรับการใช้งานในแอปพลิเคชันต่าง ๆ และสามารถปรับให้เข้ากับแพลตฟอร์มฮาร์ดแวร์ที่หลากหลาย ตั้งแต่อุปกรณ์ต้นทาง (Edge Devices) ไปจนถึง Cloud APIs ซึ่งกระบวนการออกแบบและฝึกสอนโมเดลมีดังนี้

3.3.1. เก็บข้อมูล (Data Collection) ในการฝึกสอนโมเดล Image Segmentation เพื่อใช้ในการลบพื้นหลังของภาพเบเกอรี่ ข้อมูลภาพที่ใช้ในการฝึกสอนจะนำมาจาก ชุดข้อมูล FoodSeg103 [40] ซึ่งเป็นชุดข้อมูลที่ได้รับการออกแบบมาเพื่อใช้สำหรับการทำ Segmentation ของภาพอาหาร โดยเฉพาะ ชุดข้อมูลนี้ประกอบด้วยภาพอาหารหลากหลายชนิด ซึ่งมีรายละเอียดของวัตถุและพื้นหลังที่แตกต่างกันไป ชุดข้อมูลดังกล่าวได้รับการ Labeling อย่างละเอียด โดยการระบุขอบเขตของวัตถุอาหารในแต่ละภาพเพื่อใช้ในการฝึกสอนโมเดล Segmentation

3.3.2. เตรียมข้อมูล (Data Preparation) ในการเตรียมข้อมูลสำหรับการฝึกสอนโมเดล Image Segmentation ผู้วิจัยใช้ชุดข้อมูล FoodSeg103 ซึ่งประกอบด้วยภาพอาหารจำนวน 7,118 ภาพ (จากภาพที่ 10(a) ) โดยภาพทั้งหมดได้ผ่านการป้ายเลเบลสำหรับการแบ่งส่วน (Segmentation) ของวัตถุอาหารในภาพตามประเภทส่วนผสม อย่างไรก็ตาม เพื่อให้การฝึกสอนโมเดลเกิดประสิทธิภาพมากขึ้น ผู้วิจัยปรับชุดข้อมูลดังนี้

1) การรวมเลเบลวัตถุอาหาร เนื่องจากงานวิจัยนี้ต้องการให้โมเดลสามารถแยกวัตถุอาหารทั้งหมดออกจากพื้นหลังได้ ผู้วิจัยจึงปรับการป้ายเลเบลในภาพ โดยการรวมเลเบลของส่วนผสมต่างๆ (จากภาพที่ 10(b) ) ในภาพให้เป็นวัตถุเดียว ซึ่งจะช่วยให้โมเดลสามารถแยกวัตถุอาหารออกจากพื้นหลังได้อย่างแม่นยำมากขึ้น

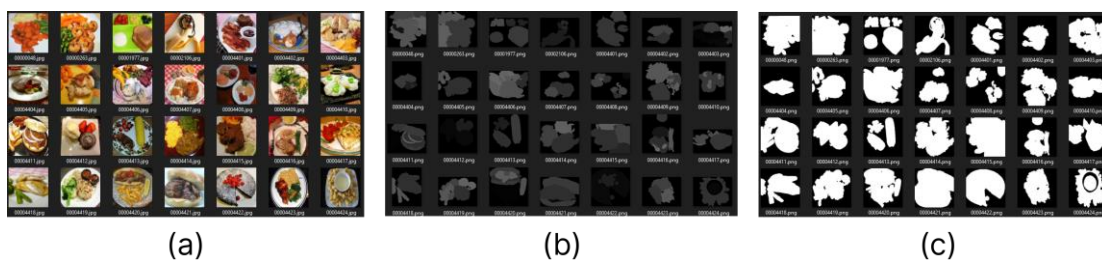
2) การปรับแต่งภาพด้วย OpenCV [41] ผู้วิจัยใช้ไลบรารี OpenCV เพื่อประมวลผลภาพให้ทุกส่วนของภาพที่ไม่ใช่พื้นหลัง (ค่าสีไม่เท่ากับ 0) ถูกเปลี่ยนเป็นสีขาว (255) ทำให้วัตถุอาหารในภาพถูกแยกออกอย่างชัดเจนจากพื้นหลังสีดำ (จากภาพที่ 10(c)) ซึ่งกระบวนการนี้ช่วยลดความซับซ้อนของข้อมูลนำเข้าและทำให้โมเดลสามารถเรียนรู้ลักษณะของวัตถุอาหารได้อย่างชัดเจนมากยิ่งขึ้น

3) การสร้างข้อมูลพิกัดขอบเขต (Bounding Coordinates) สำหรับโมเดล Ultralytics YOLOv8 ชุดข้อมูลจะถูกจัดรูปแบบใหม่ให้เหมาะสมกับการฝึกสอนของโมเดล YOLOv8 ผู้วิจัยใช้ไลบรารี Rasterio [42] เพื่อสร้างพิกัดขอบเขตจากชุดข้อมูล FoodSeg103 โดยมีการแปลงภาพในรูปแบบอาร์เรย์ให้เป็นเส้นตามพิกเซลที่มีค่าสีไม่เท่ากับศูนย์ กระบวนการนี้ทำให้สามารถสร้างข้อมูลฟีเจอร์ (attributes) ซึ่งถูกจัดเก็บในรูปแบบ

`<class-index> <x1> <y1> <x2> <y2> ... <xn> <yn>`

เพื่อใช้งานกับโมเดล YOLOv8 กระบวนการนี้ช่วยให้ระบบสามารถระบุพิกัดขอบเขตของวัตถุได้อย่างแม่นยำ

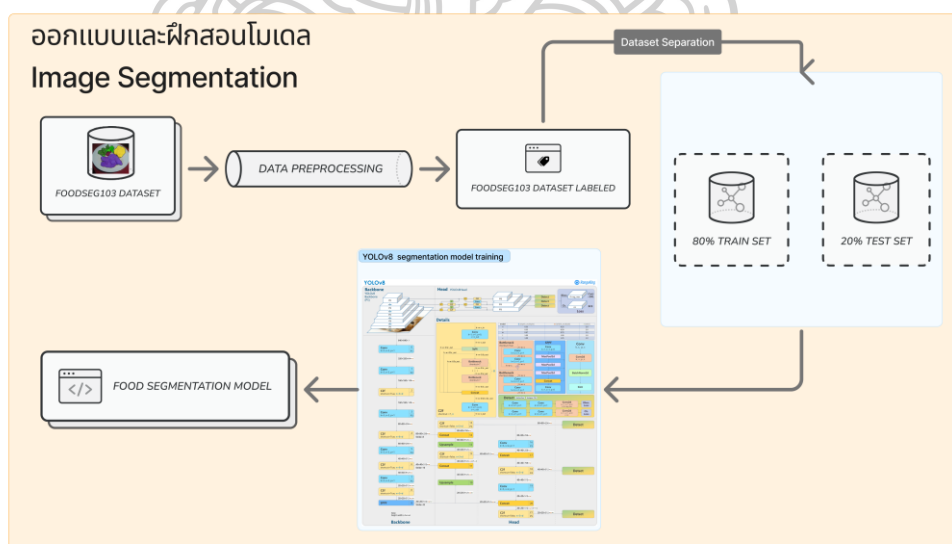
4) การเตรียมข้อมูลสำหรับการฝึกสอนโมเดล หลังจากการปรับเลเบลและการประมวลผลภาพแล้ว ข้อมูลที่ได้จะถูกแบ่งออกเป็นชุดข้อมูลฝึกสอน (Training Set) และชุดข้อมูลทดสอบ (Testing Set) โดยใช้สัดส่วน 80:20 เพื่อใช้ในการฝึกสอนและประเมินประสิทธิภาพของโมเดลในการแยกวัตถุออกจากพื้นหลัง



ภาพที่ 10 ตัวอย่างชุดข้อมูล FoodSeg103

3.3.3. การสร้างโมเดล (Model Design) เลือกใช้และโมเดล YOLOv8 Instance Segmentation ซึ่งเป็นโมเดลประมวลผลด้าน Instance Segmentation โดยไม่เพียงแค่ระบุตำแหน่งของวัตถุในภาพเท่านั้น แต่ยังสามารถระบุรูปร่างที่แน่นอนของวัตถุแต่ละชนิด ซึ่งในงานวิจัยนี้ได้มีการปรับปรุงโดยการทำ Fine-Tuning โดยโมเดลจะถูกฝึกสอนเพิ่มเติมด้วยชุดข้อมูล FoodSeg103 ที่ผ่านการเตรียมข้อมูลมาแล้ว เพื่อให้โมเดลสามารถประมวลผลวัตถุที่เกี่ยวข้องกับอาหารได้อย่างมีประสิทธิภาพ โดยปรับให้โมเดลตรวจจับเพียงคลาสเดียวที่สนใจ สำหรับการใช้งานในระบบแนะนำบรรจุภัณฑ์

ซึ่งขั้นตอนการฝึกสอนโมเดล Image Segmentation แสดงดังภาพที่ 11



ภาพที่ 11 ขั้นตอนการฝึกสอนโมเดล Image Segmentation

โมเดล YOLOv8 มีหลายขนาดและหลายเวอร์ชัน แต่ละเวอร์ชันจะมีความซับซ้อนและประสิทธิภาพที่ต่างกัน ทั้งในด้านความแม่นยำในการตรวจจับ (mAPbox และ mAPmask) และความเร็วในการประมวลผล (Speed) ซึ่งการเลือกโมเดลที่เหมาะสมจะพิจารณาจากปัจจัยต่าง ๆ

เช่น ความต้องการในการใช้งานจริง การประมวลผลที่รวดเร็ว และความแม่นยำสูง โดยในการวิจัยนี้จะเปรียบเทียบโมเดลดังตารางที่ 3

ตารางที่ 3 ตารางเปรียบเทียบโมเดล YOLOv8 Instance Segmentation (docs.ultralytics.com, 2024)

| Model       | Size (pixels) | $mAP_{50-95}^{box}$ | $mAP_{50-95}^{mask}$ | Speed CPU ONNX (ms) | Speed A100 TensorRT (ms) | Params (M) | FLOPs (B) |
|-------------|---------------|---------------------|----------------------|---------------------|--------------------------|------------|-----------|
| YOLOv8n-seg | 640           | 36.7                | 30.5                 | 96.1                | 1.21                     | 3.4        | 12.6      |
| YOLOv8s-seg | 640           | 44.6                | 36.8                 | 155.7               | 1.47                     | 11.8       | 42.6      |
| YOLOv8m-seg | 640           | 49.9                | 40.8                 | 317.0               | 2.18                     | 27.3       | 110.2     |
| YOLOv8l-seg | 640           | 52.3                | 42.6                 | 572.4               | 2.79                     | 46.0       | 220.5     |
| YOLOv8x-seg | 640           | 53.4                | 43.4                 | 712.1               | 4.02                     | 71.8       | 344.1     |

โดยแต่ละตัวแปรมีความหมายดังนี้

- Model เป็นชื่อของโมเดลที่ใช้ในการทำนายแต่ละขนาด YOLOv8 (Nano, Small, Medium, Large, และ Extra Large)
- Size (pixels) เป็นขนาดภาพที่ใช้ทำนายซึ่งกำหนดเป็น 640x640 pixels
- $mAP_{50-95}^{box}$  (%) เป็นค่าความแม่นยำเฉลี่ยในการตรวจจับวัตถุ (Bounding Box) โดยมี IoU ตั้งแต่ 50% ถึง 95%
- $mAP_{50-95}^{mask}$  (%) เป็นค่าความแม่นยำเฉลี่ยในการคาดการณ์รูปร่างของวัตถุ (Mask) โดยมี IoU ตั้งแต่ 50% ถึง 95%
- Speed CPU ONNX (ms) เป็นความเร็วในการประมวลผลบน CPU โดยใช้ ONNX วัดเป็นมิลลิวินาที (ms)
- Speed A100 TensorRT (ms) เป็นความเร็วในการประมวลผลบน GPU รุ่น A100 โดยใช้ TensorRT วัดเป็นมิลลิวินาที (ms)
- Params (M) เป็นจำนวนพารามิเตอร์ (Weights) ของโมเดลในล้านหน่วย (Million)

- FLOPs (B) เป็นจำนวนการคำนวณที่โมเดลทำ วัดเป็นพันล้านครั้ง (Billion)

3.3.4. การฝึกสอน (Training) การทดลองในการฝึกสอนโมเดล YOLOv8x-Seg ได้ดำเนินการบนระบบปฏิบัติการ Windows 10 Pro โดยใช้ CPU INTEL CORE I9-14900K - 24C 32T 6.0GHz และการ์ดจอ NVIDIA GeForce RTX 4070 Ti สถาปัตยกรรมการประมวลผลเชิงลึกที่ใช้คือ PyTorch 2.4.0 และ Compute Unified Device Architecture (CUDA) 12.1 โดยการใช้การตั้งค่าพารามิเตอร์ ดังแสดงในตารางที่ 4

ตารางที่ 4 ค่าพารามิเตอร์ระหว่างการฝึกสอนโมเดล

|                      |                |
|----------------------|----------------|
| learning rate        | 0.01           |
| batch size           | 16             |
| number of iterations | 300 epochs     |
| image size           | 224 224 pixels |

3.3.5. การประเมินผล (Evaluation) ในขั้นตอนนี้ ผู้วิจัยประเมินประสิทธิภาพของ YOLOv8 โดยใช้เกณฑ์การวัดผล ได้แก่ *Precision*, *Recall* และ *F1 – Score*

$$Precision = \frac{TP}{(TP + FP)} \times 100\% \quad (1)$$

$$Recall = \frac{TP}{(TP + FN)} \times 100\% \quad (2)$$

$$F1 - Score = \frac{2 \times (Precision \times Recall)}{Precision + Recall} \quad (3)$$

โดยที่

- *Precision* หมายถึง ค่าความแม่นยำ
- *Recall* หมายถึง ค่าความถูกต้อง
- TP (True Positive) หมายถึง ตัวอย่างที่เป็นจริงและถูกทำนายว่าเป็นจริง
- FP (False Positive) หมายถึง ตัวอย่างที่ไม่เป็นจริงแต่ถูกทำนายว่าเป็นจริง
- FN (False Negative) หมายถึง ตัวอย่างที่เป็นจริงแต่ถูกทำนายว่าไม่จริง
- *F1 – Score* หมายถึง ค่าเฉลี่ยแบบฮาร์โมนิก ระหว่างค่า *Precision* และค่า *Recall*

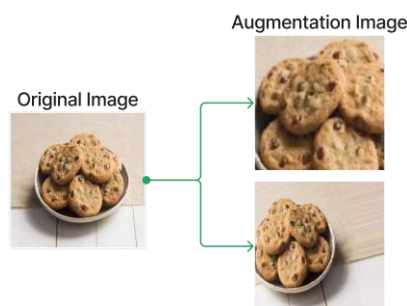
### 3.4 การเพิ่มข้อมูลด้วยเทคนิค Data Augmentation

เป็นเทคนิคที่ใช้ในการเพิ่มจำนวนข้อมูล โดยการสร้างชุดข้อมูลใหม่จากข้อมูลเดิมผ่านการปรับเปลี่ยนในรูปแบบต่างๆ เช่น การหมุนภาพ (Rotation), การยืดและย่อภาพ (Scaling), การและสะท้อนภาพ (Flipping) โดยเพิ่มจำนวนข้อมูลภาพเบเกอร์ที่ผ่านการ Segment แล้วจาก 300 ภาพ เป็น 1,000 ภาพ ซึ่งช่วยให้โมเดลสามารถเรียนรู้จากชุดข้อมูลที่หลากหลายมากขึ้น และลดความเสี่ยงในการ Overfitting เนื่องจากการฝึกสอนจากข้อมูลที่มีลักษณะคล้ายกันเกินไป

3.4.1. เทคนิคที่ใช้ใน Data Augmentation สำหรับภาพเบเกอร์ที่ถูกปรับแต่งเพื่อให้สอดคล้องกับลักษณะเฉพาะของภาพเบเกอร์ที่มีการ Label แล้ว โดยมุ่งเน้นการรักษาความถูกต้องของข้อมูล Label เอาไว้ แต่เพิ่มความหลากหลายในการแสดงภาพเหล่านั้น ด้วยวิธีการดังต่อไปนี้

- 1) การหมุนภาพเบเกอร์ที่มีการ Label แล้วในมุมต่างๆ เพื่อสร้างภาพที่มีมุมมองที่แตกต่างกันออกไป ซึ่งสามารถช่วยให้โมเดลสามารถเรียนรู้ได้จากภาพที่มีการจัดวางในมุมต่างๆ
- 2) การสะท้อนภาพในแนวนอนหรือแนวตั้ง เพื่อเพิ่มจำนวนภาพที่คล้ายคลึงกันแต่ไม่ซ้ำกัน ช่วยเพิ่มข้อมูลที่แตกต่างไปจากเดิมเล็กน้อยแต่ยังคงคุณสมบัติเดิมของเบเกอร์ไว้
- 3) การปรับขนาดของภาพเบเกอร์ให้ใหญ่ขึ้นหรือเล็กลง ซึ่งช่วยให้โมเดลมีความสามารถในการรับมือกับขนาดที่หลากหลายของเบเกอร์ได้มากยิ่งขึ้น

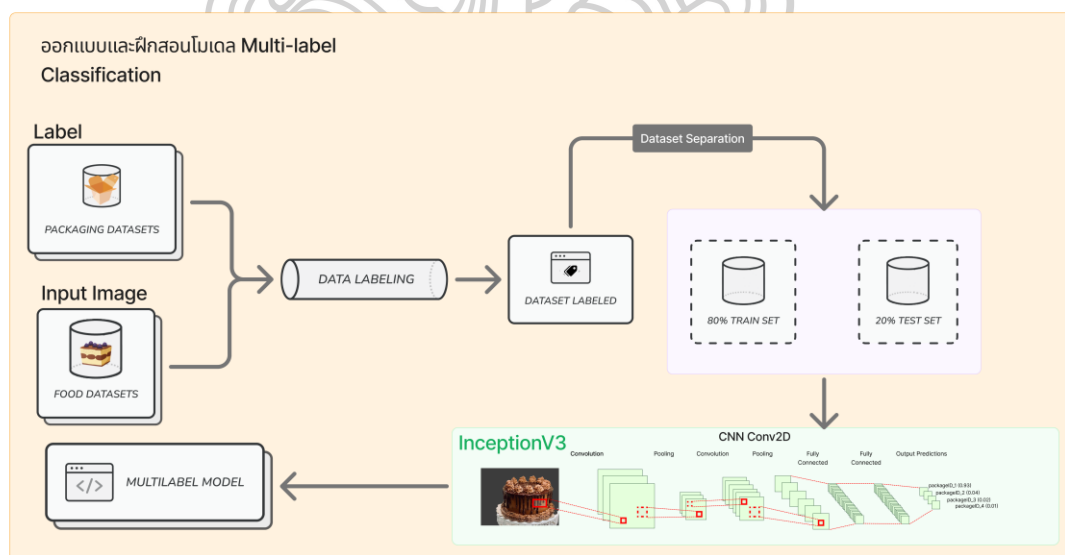
3.4.2. การนำ Data Augmentation ไปใช้ในกระบวนการฝึกสอนโมเดล โดยภาพที่ถูกสร้างขึ้นใหม่จาก Data Augmentation ถูกนำไปใช้ในการฝึกสอนโมเดล Multi-label Classification โดยภาพที่ได้รับการ Label แล้วในขั้นตอนนี้ถูกขยายจำนวนด้วยวิธีการ Augmentation ข้างต้น การฝึกสอนโมเดลด้วยข้อมูลที่มีความหลากหลายเช่นนี้ช่วยลดโอกาสการเกิด Overfitting ที่มักเกิดขึ้นเมื่อมีการฝึกสอนโมเดลจากข้อมูลที่มีจำนวนจำกัดและมีลักษณะคล้ายกันมากเกินไป นอกจากนี้ การเพิ่มข้อมูลภาพที่หลากหลายยังทำให้โมเดลสามารถจำแนกประเภทของบรรจุภัณฑ์ที่เหมาะสมกับเบเกอร์ได้อย่างมีประสิทธิภาพมากขึ้นดังตัวอย่างภาพที่ 12



ภาพที่ 12 Data Augmentation Example

### 3.5 ออกแบบและฝึกสอนโมเดล Multi-label Classification

ในการวิจัยนี้ ได้ออกแบบโมเดลการจำแนกประเภทแบบหลายเลเบล ซึ่งโมเดลนี้จะถูกใช้ในการทำนายบรรจุภัณฑ์ที่เหมาะสมกับภาพเบเกอรี่ในรูปแบบที่มากกว่าหนึ่งประเภท โดยโมเดลจะสามารถระบุบรรจุภัณฑ์ที่เหมาะสมได้หลายประเภทพร้อมกัน ซึ่งเป็นข้อจำเป็นในงานวิจัยนี้เพราะภาพเบเกอรี่แต่ละภาพสามารถถูกจัดอยู่ในบรรจุภัณฑ์หลายประเภทพร้อมกันได้ ตัวอย่างเช่น ขนมเค้กอาจถูกใส่ได้ทั้ง กล่องเค้กและ กล่องอาหารว่าง ในขั้นตอนนี้จะประกอบไปด้วยการเตรียมข้อมูล การออกแบบสถาปัตยกรรมของโมเดล การเลเบลข้อมูล การฝึกสอนโมเดล และการประเมินผล ซึ่งมีขั้นตอนการฝึกสอนโมเดลดังภาพที่ 13

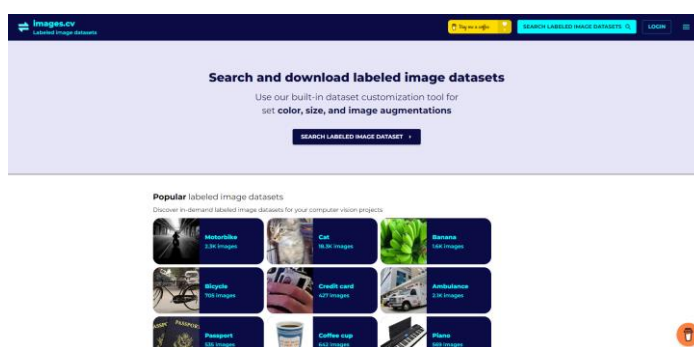


ภาพที่ 13 ขั้นตอนการฝึกสอนโมเดล Multi-label Classification

โดยแต่ละขั้นตอนมีรายละเอียดดังนี้

3.5.1. เก็บข้อมูล (Data Collection) ในการสร้างโมเดลการจำแนกประเภทแบบหลายเลเบล  
ในงานวิจัยนี้จำเป็นต้องมีชุดข้อมูลหลักทั้งหมดสองชุดข้อมูลคือ

1) ชุดข้อมูลภาพเบเกอร์รี่ ข้อมูลภาพเบเกอร์รี่ในงานวิจัยนี้ได้รับการรวบรวมจากเว็บไซต์ <https://images.cv> ซึ่งเป็นแหล่งข้อมูลภาพออนไลน์ที่มีภาพเบเกอร์รี่ ดังภาพที่ 14 หลากหลายประเภท เช่น ขนมปัง เค้ก คุกกี้ มาการอง และอื่นๆ ชุดข้อมูลดังกล่าวถูกนำมาใช้เป็นข้อมูลภาพสำหรับฝึกสอนโมเดลเพื่อจำแนกประเภทของบรรจุภัณฑ์ที่เหมาะสมกับแต่ละผลิตภัณฑ์เบเกอร์รี่



ภาพที่ 14 เว็บไซต์แหล่งข้อมูลภาพออนไลน์ <https://images.cv>

2) ชุดข้อมูลภาพบรรจุภัณฑ์ ชุดข้อมูลบรรจุภัณฑ์ได้รับการสนับสนุนจาก บริษัท บราเธอร์ แอนด์ บราเธอร์ จำกัด ผู้ผลิตและจำหน่ายบรรจุภัณฑ์สำหรับใส่ขนมและอาหาร ภายใต้แบรนด์ Boxjourney ข้อมูลบรรจุภัณฑ์เหล่านี้ครอบคลุมหลากหลายประเภทและขนาดที่ถูกใช้ในการบรรจุผลิตภัณฑ์เบเกอร์รี่ โดยมีรูปแบบข้อมูลถูกจัดเก็บเป็นโครงสร้าง JSON ที่มีการจัดกลุ่มบรรจุภัณฑ์ตามประเภทและขนาด ตัวอย่างโครงสร้างข้อมูลดังภาพที่ 15

```
{
  "FlatPastryBox": {
    "40.00x21.50x3.00": {
      "0104025": [
        "0104025-0004397-L.jpeg",
        "0104025-0004395-L.jpeg",
        "0104025-0008177-L.jpeg",
        "0104025-z8slV1mnr-L.jpeg",
        "0104025-0004396-L.jpeg",
        "0104025-0010285-L.jpeg"
      ]
    },
    "30.00x34.00x4.00": {
      "0111043": [ ... ],
      "0111042": [ ... ]
    },
    "21.50x23.00x2.00": {
      "0104159": [ ... ]
    },
    "31.20x29.70x0.80": {
      "0104196": [ ... ],
      "0104197": [ ... ],
      "0104195": [ ... ],
      "0104192": [ ... ],
      "0104194": [ ... ],
      "0104193": [ ... ]
    }
  },
}
```

ภาพที่ 15 ตัวอย่างโครงสร้างข้อมูลบรรจุภัณฑ์

ข้อมูลนี้ถูกจัดเก็บในรูปแบบโครงสร้างต้นไม้ (tree structure) ซึ่งมีการระบุประเภทบรรจุภัณฑ์ ขนาด (กว้าง x ยาว x สูง) และหมายเลขอ้างอิงของบรรจุภัณฑ์ (รหัสสินค้า) เพื่อใช้ในการค้นหาและระบุสินค้าตามเงื่อนไข โดยตัวอย่างข้อมูลภาพดังตัวอย่างภาพที่ 16



ภาพที่ 16 ตัวอย่างข้อมูลภาพบรรจุภัณฑ์

3.5.2. พัฒนาเครื่องมือเลเบล (Labeling Tool Development) เพื่อรองรับการสร้างชุดข้อมูลสำหรับการจำแนกประเภทแบบหลายเลเบล และการจัดการข้อมูลที่มีประสิทธิภาพ งานวิจัยนี้ได้พัฒนาเว็บไซต์ที่สามารถใช้งานได้ง่ายและสะดวก โดยใช้ React TS สำหรับ Frontend และ FastAPI สำหรับ Backend เพื่อให้สามารถจัดการข้อมูลและการเลเบลได้อย่างมีประสิทธิภาพ มีรายละเอียดการพัฒนาดังนี้

#### 1) Frontend Development

- เว็บไซต์ถูกพัฒนาโดยใช้ React TS ซึ่งเป็นเทคโนโลยีที่รองรับการทำงานแบบอินเทอร์แอกทีฟและทันสมัย ผู้ใช้สามารถเลือกและเลเบลบรรจุภัณฑ์ที่เหมาะสมกับเบเกอร์รี่ได้อย่างง่ายดาย
- การออกแบบ UI/UX ของเว็บไซต์มุ่งเน้นไปที่การใช้งานง่าย โดยมีการแสดงภาพเบเกอร์รี่จากชุดข้อมูล และมีตัวเลือกบรรจุภัณฑ์ที่ผู้ใช้สามารถเลือกได้อย่างรวดเร็ว

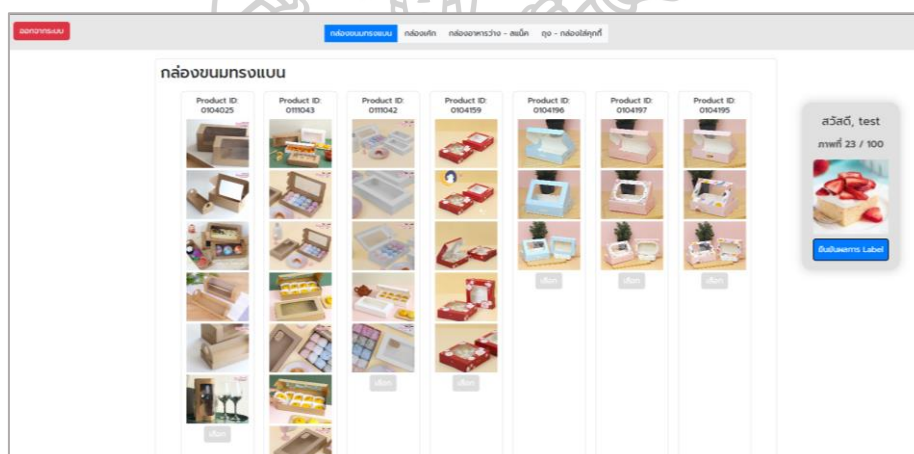
#### 2) Backend Development

- ระบบ Backend ถูกพัฒนาโดยใช้ FastAPI ซึ่งเป็นเฟรมเวิร์คที่รวดเร็วและมีประสิทธิภาพในการจัดการกับ API สำหรับการรับส่งข้อมูลการเลเบล โดยข้อมูลที่ผู้ใช้เลเบลจะถูกส่งไปยังระบบ Backend ผ่าน API

- ระบบยังรองรับการทำงานร่วมกับฐานข้อมูล MongoDB ซึ่งช่วยให้สามารถจัดเก็บข้อมูลการเลเบลในรูปแบบ JSON ที่มีโครงสร้างชัดเจน พร้อมสำหรับการนำไปใช้งานต่อในกระบวนการฝึกสอนโมเดล

3) การจัดการข้อมูล ข้อมูลภาพเบเกอร์ที่ใช้สำหรับการเลเบลจะถูกแสดงผลบนเว็บไซต์จากชุดข้อมูลที่เตรียมไว้ โดยระบบจะแสดงตัวเลือกภาพกล่องบรรจุภัณฑ์ที่เหมาะสมดังภาพที่ 17 ซึ่งแบ่งออกเป็น 4 ประเภทหลัก ได้แก่

- กล่องขนมทรงแบน
- กล่องเค้ก
- กล่องอาหารว่าง - สแน็ค
- ถู - กล่องใส่คุกกี้

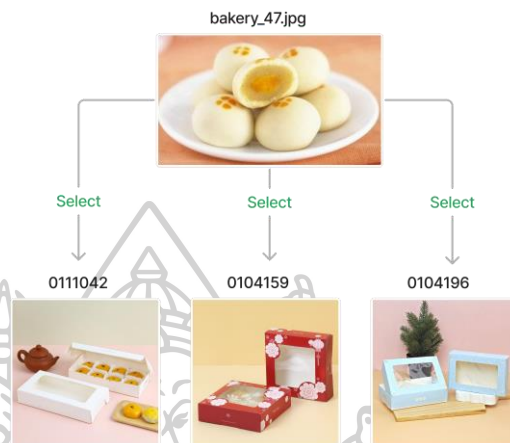


ภาพที่ 17 เว็บไซต์เลเบลข้อมูลความเหมาะสมระหว่างบรรจุภัณฑ์และเบเกอร์

#### 4) กระบวนการเลเบล

- ระบบจะแสดงภาพขนมหนึ่งภาพบนหน้าจอ และแสดงตัวเลือกกล่องบรรจุภัณฑ์ที่เป็นสินค้าเฉพาะประเภทกล่องเท่านั้น (ไม่แสดงสินค้าที่ไม่เกี่ยวข้อง เช่น ถู แผ่นรองเค้ก หรือ แก้วน้ำ เป็นต้น)
- ผู้ใช้งานสามารถเลือกกล่องบรรจุภัณฑ์ที่เหมาะสมกับขนมภาพนั้น โดยเลือกได้มากกว่า 1 ประเภทหากเห็นว่ากล่องหลากหลายประเภทสามารถใช้กับขนมนั้นได้ ซึ่งแสดงความสัมพันธ์ในการเลือกดังตัวอย่างภาพที่ 18

- เมื่อผู้ใช้งานเลือกเสร็จสิ้นและกดยืนยัน ผลการเลเบลของภาพขมนั้นจะถูกบันทึกลงในฐานข้อมูล
- ระบบจะแสดงภาพขมนภาพถัดไปโดยอัตโนมัติ และกระบวนการนี้จะทำซ้ำไปเรื่อย ๆ จนกว่าภาพขมนทั้งหมดจะถูกเลเบลครบ



ภาพที่ 18 ตัวอย่างผลการเลือกบรรจุภัณฑ์ กับภาพเบเกอรี่

5) เมื่อผู้ใช้งานเลเบลข้อมูลสำเร็จ ระบบจะบันทึกผลลัพธ์การเลเบลโดยใช้ฐานข้อมูล MongoDB ในรูปแบบ JSON ดังภาพที่ 19 เพื่อจัดเก็บข้อมูลที่มีโครงสร้างชัดเจนและพร้อมสำหรับการนำไปใช้งานในกระบวนการฝึกสอนโมเดลแบบหลายเลเบล ซึ่งจะมีโครงสร้างข้อมูลการบันทึกดังนี้

- "username": ระบุชื่อผู้ใช้งานที่เลเบล
- "images": รายการภาพเบเกอรี่ทั้งหมดที่ได้รับการเลเบล
- "current\_index": ตำแหน่งของภาพล่าสุดที่กำลังเลเบล
- "labels": ผลการเลเบลในรูปแบบคู่คีย์ระหว่างชื่อภาพเบเกอรี่และรหัสบรรจุภัณฑ์
- "last\_saved": เวลาที่ข้อมูลถูกบันทึกล่าสุด

```
{
  "username": "string",           // ชื่อผู้ใช้งานที่ทำการเลเบล
  "images": ["string"],           // รายการชื่อไฟล์ภาพเบเกอรี่ทั้งหมด
  "current_index": "integer",     // ตำแหน่งภาพที่กำลังเลเบลในปัจจุบัน
  "labels": {                     // ผลการเลเบลบรรจุภัณฑ์
    "string": ["string"]          // ชื่อภาพ (key) และรหัสบรรจุภัณฑ์ (array ของ string)
  },
  "last_saved": "string"         // เวลาที่ข้อมูลถูกบันทึกล่าสุดในรูปแบบ ISO 8601
}
```

ภาพที่ 19 รูปแบบข้อมูลการเลเบลรูปแบบ JSON

3.5.3. ข้อมูลที่ใช้ในการฝึกสอนโมเดลจำแนกแบบหลายเลเบล มาจากการเลเบลภาพเบเกอรี่หลายประเภทที่มีการจับคู่กับบรรจุภัณฑ์ที่เหมาะสม โดยแต่ละภาพเบเกอรี่จะถูกเชื่อมโยงกับ

หลายบรรจุกณ์ในรูปแบบหลายเลเบล ซึ่งมีการจับคู่โดยผู้เชี่ยวชาญด้านการแนะนำบรรจุกณ์เบเกอรี่ซึ่งเป็น Admin ของเว็บไซต์ Boxjourney จากทางบริษัท บราเธอร์ แอนด์ บราเธอร์ จำกัด ทั้งหมด 6 ท่าน โดยแต่ละผู้เลเบลจะเลเบลข้อมูลเบเกอรี่คนละ 100 ภาพ ซึ่งมีโอกาสได้ภาพเบเกอรี่ซ้ำกันได้เพื่อเพิ่มความน่าเชื่อถือและคุณภาพของข้อมูลการเลเบล โดยในกรณีที่ภาพเบเกอรี่มีการเลเบลซ้ำกันจากผู้เชี่ยวชาญหลายท่าน ระบบจะใช้วิธี Majority Voting ในการรวมผลลัพธ์การเลเบลให้ได้ฉันทามติ โดยค่าที่ถูกเลือกมากที่สุดจากผู้เลเบลทั้งหมดจะถูกกำหนดเป็นผลลัพธ์สุดท้าย (Ground Truth) ดังตารางที่ 5 แสดงตัวอย่างข้อมูลที่ผ่านมาการประมวลผลหลังจากการเลเบล โดยมีการบันทึกค่าการจับคู่ระหว่างภาพเบเกอรี่และบรรจุกณ์ในรูปแบบ Label Matrix ซึ่งประกอบด้วยค่าที่แสดงถึงความเหมาะสมของแต่ละบรรจุกณ์ต่อภาพเบเกอรี่ในรูปแบบตัวเลข

ตารางที่ 5 ตัวอย่างข้อมูล Label Matrix

| ID             | 0104025 | 0111043 | 0111042 | 0104159 | 0104196 | 0104197 | 0104195 | 0104192 | 0104194 |
|----------------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| bakery_47.jpg  | 0       | 1       | 2       | 2       | 2       | 2       | 2       | 2       | 2       |
| bakery_176.jpg | 0       | 0       | 0       | 0       | 1       | 1       | 1       | 1       | 1       |
| bakery_50.jpg  | 1       | 2       | 1       | 2       | 2       | 2       | 1       | 1       | 1       |
| bakery_84.jpg  | 0       | 1       | 1       | 1       | 1       | 1       | 1       | 1       | 1       |
| bakery_137.jpg | 0       | 0       | 0       | 0       | 0       | 0       | 0       | 0       | 0       |

- คอลัมน์แรก (ID) ระบุชื่อภาพเบเกอรี่
- แต่ละคอลัมน์ถัดไปแสดงรหัสของบรรจุกณ์ และค่าที่ปรากฏในแต่ละเซลล์ระบุความถี่ของการเลือกบรรจุกณ์นั้นโดยผู้เลเบล
- ตัวอย่างเช่น ในแถวของ "bakery\_47.jpg" ค่า "2" หมายถึงบรรจุกณ์นั้นถูกเลือกโดยผู้เลเบล 2 คน

หากพบว่าแถวใดไม่มีค่าที่มากกว่า 1 เลย ข้อมูลในแถวนั้นจะถูกใช้เป็น Ground Truth โดยตรง โดยไม่มีการเปลี่ยนแปลงใด ๆ ซึ่งวิธีนี้ช่วยให้สามารถรักษาค่าความจริงที่ผู้เลเบลเลือกไว้ได้อย่างครบถ้วน ในขณะเดียวกัน หากแถวใดมีค่าที่มากกว่า 1 ระบบจะใช้การประมวลผลด้วย Majority Voting ค่าที่มากกว่า 1 จะถูกแปลงเป็น "1" และค่า "1" เดิมจะถูกแปลงเป็น "0" เพื่อสร้าง Label Matrix สำหรับการฝึกสอนโมเดลที่รองรับ Multi-Label Classification ตารางที่ 6 แสดงข้อมูลหลังจากการปรับเปลี่ยน

ตารางที่ 6 ตัวอย่างข้อมูล Label Matrix หลังการปรับเปลี่ยนด้วยวิธี Majority Voting

| ID             | 0104025 | 0111043 | 0111042 | 0104159 | 0104196 | 0104197 | 0104195 | 0104192 | 0104194 |
|----------------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| bakery_47.jpg  | 0       | 0       | 1       | 1       | 1       | 1       | 1       | 1       | 1       |
| bakery_176.jpg | 0       | 0       | 0       | 0       | 1       | 1       | 1       | 1       | 1       |
| bakery_50.jpg  | 1       | 1       | 0       | 1       | 1       | 1       | 1       | 1       | 1       |
| bakery_84.jpg  | 0       | 1       | 1       | 1       | 1       | 1       | 1       | 1       | 1       |
| bakery_137.jpg | 0       | 0       | 0       | 0       | 0       | 0       | 0       | 0       | 0       |

3.5.4. การออกแบบสถาปัตยกรรมโมเดล (Model Design) ในขั้นตอนการออกแบบสถาปัตยกรรมของโมเดลที่ใช้สำหรับการจำแนกประเภทแบบหลายเลเบล (Multi-label Classification) ระบบที่พัฒนาในงานวิจัยนี้มีการใช้โมเดลหลายชนิด โดยโมเดลหลักที่เลือกใช้คือโมเดลที่มีประสิทธิภาพสูงในการสกัดฟีเจอร์จากภาพ และมีความเหมาะสมสำหรับการนำไปปรับใช้ในงานที่เกี่ยวข้องกับภาพอาหารและบรรจุภัณฑ์ ซึ่งได้แก่โมเดล VGG16, VGG19, ResNet50 และ InceptionV3 แต่ละโมเดลได้รับการปรับแต่งเพื่อให้เข้ากับงานจำแนกประเภทหลายเลเบลโดยในแต่ละโมเดลกำหนดจำนวนคลาสของผลลัพธ์เป็น 736 คลาส และมีรายละเอียดการปรับแต่งดังนี้

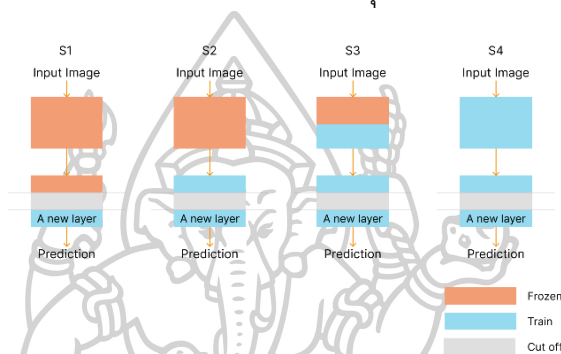
1) การปรับแต่งโมเดล (Fine-Tuning) เพื่อให้โมเดลสามารถจำแนกบรรจุภัณฑ์ที่เหมาะสมกับผลิตภัณฑ์เบเกอรี่ได้อย่างแม่นยำ โมเดลทุกตัวได้รับการปรับแต่งผ่านการแช่แข็งเลเยอร์บางส่วน (freeze layers) ดังภาพที่ 20 และปรับเลเยอร์ด้านบนสำหรับการทำนาย ดังนี้

**S1 Strategies** แช่แข็งทุกเลเยอร์ (Freeze All Layers) กลยุทธ์ S1 เริ่มต้นโดยการแช่แข็งทุกเลเยอร์ของโมเดลต้นฉบับที่ได้จากการฝึกสอนด้วยข้อมูล ImageNet และเพิ่มเลเยอร์ GlobalAveragePooling2D เพื่อลดขนาดข้อมูลตามแนวแกน และตามด้วยเลเยอร์ Dense ที่ประกอบไปด้วย 2,240 output nodes ซึ่งเป็นจำนวนคลาสของบรรจุภัณฑ์ที่ต้องการทำนาย โมเดลจะถูกฝึกสอนโดยใช้ค่า learning rate เริ่มต้นที่ 0.0001 พร้อมทั้งเทคนิคการลดค่า learning rate แบบลดลงอัตโนมัติเมื่อโมเดลเข้าสู่จุดที่ราบสูง (plateau)

**S2 Strategies** แช่แข็งทุกเลเยอร์ ยกเว้น fc2 และ fc1 ในกลยุทธ์ S2 จะแช่แข็งทุกเลเยอร์ของโมเดลยกเว้นเลเยอร์ fc2 และ fc1 (Fully Connected Layers) ซึ่งเป็นเลเยอร์สุดท้ายก่อนการทำนาย เพื่อให้โมเดลสามารถเรียนรู้ฟีเจอร์ใหม่ที่มีความเฉพาะตัวของบรรจุภัณฑ์ได้ และจากนั้นเพิ่มเลเยอร์ Dense เช่นเดียวกับใน S1

**S3 Strategies** แช่แข็งเลเยอร์เริ่มต้น และปลดล๊อคเลเยอร์ตั้งแต่เลเยอร์ที่ 10 เป็นต้นไป ในกลยุทธ์ S3 เลเยอร์ตั้งแต่เลเยอร์ที่ 10 ของโมเดลจะถูกปลดล๊อคและฝึกสอนใหม่ โดยเลเยอร์ก่อนหน้านี้จะยังคงถูกแช่แข็ง ทำให้โมเดลสามารถปรับแต่งฟีเจอร์ที่ซับซ้อนขึ้นได้สำหรับการจำแนกประเภทบรรจุกัญท์

**S4 Strategies** ปลดล๊อคทุกเลเยอร์ในโมเดล ในกลยุทธ์ S4 เลเยอร์ทั้งหมดของโมเดลจะถูกปลดล๊อคและฝึกสอนใหม่ ทำให้โมเดลสามารถเรียนรู้ฟีเจอร์จากข้อมูลใหม่ได้อย่างเต็มที่ ซึ่งช่วยเพิ่มความยืดหยุ่นและความสามารถในการจำแนกประเภทบรรจุกัญท์ได้ดีขึ้น



ภาพที่ 20 Strategies of Fine-Tuning Model

3.5.5. การฝึกสอนโมเดล (Model Training) ในการฝึกสอนโมเดลสำหรับการจำแนกประเภทแบบหลายเลเบล จะต้องใช้ชุดข้อมูลที่เตรียมไว้อย่างเหมาะสม โดยโมเดลที่ถูกใช้ในงานวิจัยนี้ประกอบไปด้วย VGG19, VGG16, ResNet50 และ InceptionV3 มีการกำหนดขนาดของอินพุต (Input Size) ที่แตกต่างกัน โดย VGG19, VGG16 และ ResNet50 ใช้ขนาดอินพุต 224x224 พิกเซล ในขณะที่ InceptionV3 ถูกออกแบบมาเพื่อรับอินพุตขนาด 299x299 พิกเซลโดยเฉพาะ การใช้ขนาดภาพที่ไม่เหมาะสม เช่น 224x224 สำหรับ InceptionV3 จะทำให้โครงสร้างโมเดลเกิดข้อผิดพลาดในกระบวนการประมวลผล จึงเป็นเหตุผลที่แต่ละโมเดลมีขนาดอินพุตที่ต่างกัน เพื่อให้เหมาะสมกับโครงสร้างภายในและการออกแบบเลเยอร์ของโมเดลแต่ละตัว ซึ่งถูกพัฒนาขึ้นให้สามารถประมวลผลข้อมูลอินพุตในขนาดที่กำหนดได้อย่างมีประสิทธิภาพที่สุด

#### 1) การกำหนดค่าพารามิเตอร์สำหรับการฝึกสอน

- Learning rate: 0.0001
- Optimizer: Adam
- Loss function: binary\_crossentropy
- Patience (Early Stopping): 10

- ReduceLR Factor: 0.1 (ลดค่า Learning Rate เมื่อ val\_loss คงที่)
- Input Size:
  - 224x224 สำหรับ VGG19, VGG16, ResNet50
  - 299x299 สำหรับ InceptionV3

## 2) การเตรียมข้อมูลการฝึกสอน

- ชุดข้อมูลถูกแบ่งออกเป็น 80% สำหรับการฝึกสอน (Training Set) และ 20% สำหรับการทดสอบ (Test Set)
- ในขั้นตอนการฝึกสอนมีการใช้เทคนิค Fine-tuning โดยการแช่แข็ง (Freezing) เลเยอร์บางส่วนตามกลยุทธ์ที่ออกแบบไว้ เช่น การแช่แข็งเลเยอร์ทั้งหมดใน S1 และค่อยๆ ปลดล๊อคเลเยอร์ตามกลยุทธ์ S2, S3, และ S4

โดยในแต่ละกลยุทธ์ทุกครั้งที่จะนำภาพเข้ามาฝึกสอนจะต้องผ่านการตัดพื้นหลังภาพเบเกอรี่ด้วยโมเดล YOLOv8x-Seg ที่ผ่านการฝึกสอนมาแล้วด้วยชุดข้อมูล FoodSeg103 ดังภาพที่ 21 ก่อนการนำไปสกัดฟีเจอร์ในกระบวนการฝึกสอนโมเดล



ภาพที่ 21 ผลลัพธ์การตัดพื้นหลังด้วยโมเดล YOLOv8 Image Segmentation

3.5.6. การประเมินผล (Evaluation) การประเมินผลโมเดลทำโดยการคำนวณค่า Precision, Recall, F1-score และค่า loss บน Test Set เพื่อเปรียบเทียบประสิทธิภาพของโมเดลในแต่ละแบบ นอกจากนี้ยังใช้การวัดค่า mAP (Mean Average Precision) เพื่อประเมินประสิทธิภาพของการทำนายหลายเลเบล

## 3.6 กระบวนการแยกส่วนและลบพื้นหลังภาพ

หลังจากการฝึกสอนโมเดล YOLOv8 แล้ว โมเดลถูกนำไปใช้ในกระบวนการแบ่งส่วนและลบพื้นหลังจากภาพวัตถุเบเกอรี่ เพื่อให้ระบบสามารถวิเคราะห์เฉพาะวัตถุได้โดยไม่ถูกรบกวนจากพื้นหลังที่ไม่เกี่ยวข้อง ผลลัพธ์ของการแบ่งส่วนนี้จะเป็นการสร้าง Mask ที่ระบุขอบเขตของวัตถุในภาพ ซึ่งขั้นตอนของการแบ่งส่วนและการลบพื้นหลังประกอบด้วย

3.6.1. การแบ่งส่วนภาพ เริ่มจากปรับขนาดของภาพเป็น (224,224) หรือ (299,299) ขึ้นอยู่กับโมเดลที่ใช้ จากนั้นภาพที่ผ่านการปรับขนาดจะถูกประมวลผลผ่านโมเดล YOLOv8 เพื่อสร้าง Mask ที่แบ่งแยกวัตถุออกจากกัน โดยโมเดลจะตรวจจับตำแหน่งของวัตถุและสร้างขอบเขตที่ครอบคลุมวัตถุทุกชิ้นในภาพ

3.6.2. การลบพื้นหลัง นำ Mask ที่สร้างจากการแบ่งส่วนมารวมเป็นภาพเดียวเพื่อแยกวัตถุออกจากพื้นหลัง กระบวนการนี้จะทำให้พื้นที่ที่อยู่นอกขอบเขตของ Mask ถูกทำให้เป็นสีดำ ซึ่งส่งผลให้ได้ภาพที่มีเพียงวัตถุเบเกอรี่เท่านั้น และพื้นหลังถูกลบออกไปทั้งหมดดังตัวอย่างในภาพที่ 21 ภาพที่ได้จากการลบพื้นหลังนี้จะถูกใช้ในการทำนายบรรจุภัณฑ์ในขั้นตอนถัดไป

### 3.7 การสกัดฟีเจอร์ของภาพบรรจุภัณฑ์

การสกัดฟีเจอร์ของภาพ เป็นขั้นตอนสำคัญในการวิเคราะห์และจำแนกภาพบรรจุภัณฑ์ที่เหมาะสมสำหรับผลิตภัณฑ์เบเกอรี่ กระบวนการนี้จะดึงเอาลักษณะเฉพาะจากภาพ ซึ่งประกอบไปด้วยรูปร่าง ขนาด สี และรายละเอียดเชิงโครงสร้างอื่น ๆ เพื่อนำไปใช้ในการเรียนรู้ของโมเดลสำหรับการจำแนกประเภทแบบหลายเลเบล การสกัดฟีเจอร์สามารถแบ่งออกได้ดังนี้

3.7.1. การเตรียมภาพสำหรับการสกัดฟีเจอร์ หลังจากการลบพื้นหลังออกจากภาพเบเกอรี่แล้ว ภาพที่ได้จะถูกปรับขนาดและปรับรูปแบบตามสถาปัตยกรรมของโมเดลที่ใช้ สำหรับโมเดล VGG16, VGG19, และ ResNet50 ภาพจะถูกปรับขนาดเป็น (224,224) ในขณะที่โมเดล InceptionV3 ใช้ขนาด (299,299) ซึ่งขั้นตอนนี้ช่วยให้โมเดลสามารถประมวลผลภาพได้อย่างมีประสิทธิภาพ

3.7.2 การสกัดฟีเจอร์ด้วยโมเดลที่ผ่านการฝึกสอน โมเดลที่ใช้สำหรับการสกัดฟีเจอร์ในงานวิจัยนี้ ได้แก่ VGG16, VGG19, ResNet50 และ InceptionV3 ซึ่งโมเดลเหล่านี้ได้รับการฝึกสอนเบื้องต้นด้วยชุดข้อมูล ImageNet ซึ่งเป็นฐานข้อมูลภาพขนาดใหญ่ที่ช่วยให้โมเดลสามารถเรียนรู้ฟีเจอร์พื้นฐานได้

โมเดลจะประมวลผลภาพผ่านเลเยอร์ต่าง ๆ เพื่อดึงฟีเจอร์ที่สำคัญ เช่น ขอบเขตของภาพ รูปแบบเชิงพื้นที่ และองค์ประกอบที่มีลักษณะเฉพาะสำหรับการจำแนกบรรจุภัณฑ์ ระบบจะนำฟีเจอร์ที่สกัดออกมาใช้ในการทำนายบรรจุภัณฑ์ต่อไป

### 3.8 ระบบการกรองบรรจุภัณฑ์ตามขนาดและประเภท

ระบบจะรับข้อมูล Metadata จากผู้ใช้ ซึ่งประกอบด้วยประเภทของบรรจุภัณฑ์และ ขนาดของสินค้าเพื่อใช้ในการกรองข้อมูลบรรจุภัณฑ์ก่อนเข้าสู่กระบวนการทำนายด้วยจำแนกหลายเลเบล และการค้นหาความคล้ายคลึงด้วยเทคนิค Cosine Similarity โดยกระบวนการกรองจะเป็นลำดับดังนี้

3.8.1. กรองตามประเภท (Type) ระบบจะเริ่มจากการตรวจสอบประเภทของบรรจุภัณฑ์ที่ผู้ใช้ระบุเข้ามาใน Metadata เช่น CakeBox หรือ SnackBox หากผู้ใช้ระบุประเภทไว้ ระบบจะกรองเฉพาะบรรจุภัณฑ์ที่ตรงกับประเภทดังกล่าวเพื่อนำไปสู่ขั้นตอนการกรองตามขนาดต่อไป ในกรณีที่ ไม่ได้ระบุประเภท ระบบจะใช้ข้อมูลบรรจุภัณฑ์ทุกประเภทในการพิจารณา

3.8.2. กรองตามขนาด (Size) หลังจากกรองตามประเภทแล้ว ระบบจะใช้ข้อมูลขนาดที่ผู้ใช้ระบุเข้ามา ช่วยให้ระบบแนะนำบรรจุภัณฑ์สามารถคัดเลือกบรรจุภัณฑ์ที่มีขนาดเหมาะสมกับสินค้าที่ผู้ใช้ป้อนเข้ามา โดยข้อมูลขนาดจะประกอบไปด้วยความกว้าง ความยาว และความสูงของสินค้าหรืออาหาร ซึ่งระบบจะนำขนาดที่ได้รับจากผู้ใช้ไปเปรียบเทียบกับขนาดของบรรจุภัณฑ์ในฐานข้อมูลเพื่อตัดสินใจว่าบรรจุภัณฑ์นั้นสามารถใช้ได้หรือไม่ โดยกระบวนการกรองขนาดทำงานตามหลักการดังนี้

1) รับข้อมูลขนาดจากผู้ใช้ โดยข้อมูลขนาดของสินค้าถูกส่งเข้ามาในรูปแบบของ string โดยแสดงค่าเป็นความกว้าง ความยาว และความสูง เช่น "01.00x01.00x03.00". ข้อมูลนี้เป็นการบอกขนาดของสินค้าที่ผู้ใช้ต้องการบรรจุลงในบรรจุภัณฑ์

2) การแยกค่าขนาดและแปลงเป็นตัวเลข

- ระบบจะทำการแยกข้อความขนาดออกเป็น 3 ส่วน (กว้าง, ยาว, สูง) โดยใช้ตัวอักษร "x" เป็นตัวแบ่งระหว่างค่าขนาด

- แต่ละค่าที่ได้จากการแยกจะถูกแปลงเป็นตัวเลขเชิงทศนิยม (float) เพื่อให้ง่ายต่อการนำไปคำนวณและเปรียบเทียบ

3) การเปรียบเทียบขนาดของสินค้าและบรรจุภัณฑ์

- เมื่อระบบมีขนาดของสินค้าแล้ว ระบบจะนำขนาดนี้ไปเปรียบเทียบกับขนาดของบรรจุภัณฑ์ที่เก็บอยู่ในฐานข้อมูล

- ในการเปรียบเทียบ ระบบจะตรวจสอบว่าทุกมิติของบรรจุภัณฑ์ (กว้าง, ยาว, สูง) มีค่าที่เท่ากันหรือใหญ่กว่าขนาดของสินค้า

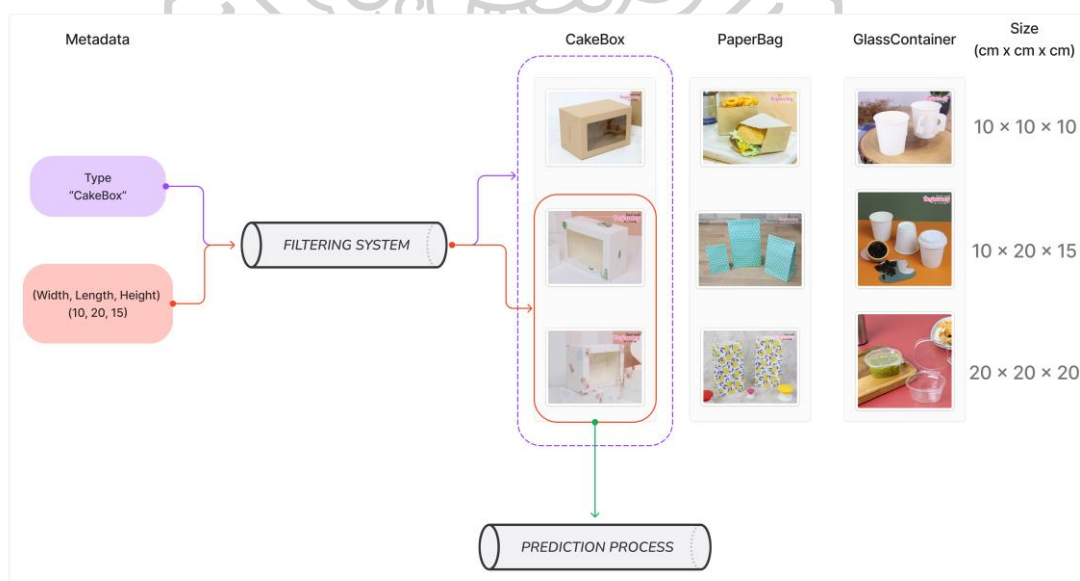
- การกรองจะตรวจสอบบรรจุภัณฑ์ทุกประเภทที่อยู่ในขอบเขตของขนาดที่เหมาะสม หากพบว่าบรรจุภัณฑ์ใดที่ขนาดเล็กกว่าขนาดของสินค้าที่ป้อนเข้ามา ระบบจะตัดบรรจุภัณฑ์นั้นออกจากการประมวลผลต่อไป

#### 4) ผลลัพธ์หลังการกรองขนาด

- หลังจากกระบวนการกรองตามขนาดเสร็จสิ้น ระบบจะเหลือเพียงบรรจุภัณฑ์ที่ขนาดเหมาะสมและสามารถบรรจุสินค้าได้

- บรรจุภัณฑ์เหล่านี้จะถูกนำไปใช้ในการทำนายด้วยโมเดลแบบหลายเลเบล และยังใช้ในการค้นหาความคล้ายคลึงระหว่างฟีเจอร์ (Feature) ของบรรจุภัณฑ์ด้วยวิธี Cosine Similarity ในขั้นตอนถัดไป

การทำงานของระบบการกรองบรรจุภัณฑ์ตามประเภทและขนาดดังภาพที่ 22 ซึ่งแสดงถึงกระบวนการตั้งแต่การรับข้อมูลนำเข้า การกรองบรรจุภัณฑ์ จนถึงการนำข้อมูลที่ผ่านการกรองไปประมวลผลในขั้นตอนการทำนายและค้นหาความคล้ายคลึงของบรรจุภัณฑ์ มีขั้นตอนการกรองดังภาพที่ 22 โดยกระบวนการกรองข้อมูลแสดงให้เห็นจากภาพที่ 23



ภาพที่ 22 ตัวอย่างกระบวนการกรองบรรจุภัณฑ์ด้วยขนาดเบเกอร์

```

1. IF size_from OR size_to OR packaging_input_type EXISTS:
2.   FOR each packaging_type, sizes IN features:
3.     # Check if packaging type matches the input type
4.     IF packaging_input_type EXISTS AND packaging_input_type != packaging_type:
5.       CONTINUE to the next packaging_type
6.
7.   FOR each product_size, products IN sizes:
8.     # Check if the product size is within the specified range
9.     IF NOT is_size_in_range(size_from, size_to, product_size):
10.      CONTINUE to the next product_size
11.
12.   FOR each product_id IN products:
13.     # Add the product to filtered_predictions if it exists in predictions_dict
14.     IF product_id EXISTS IN predictions_dict:
15.       filtered_predictions[product_id] <- predictions_dict[product_id]

```

ภาพที่ 23 ขั้นตอนการกรองบรรจุภัณฑ์ด้วย Metadata (type, size)

### 3.9. กระบวนการค้นหาภาพ

หลังจากผ่านกระบวนการการทำ Image Segmentation จนถึงขั้นตอนการทำนายแบบหลายเลเวลแล้ว นำผลลัพธ์ซึ่งเป็นรูปภาพบรรจุภัณฑ์ที่มีค่าความมั่นใจ 5 อันดับแรก นำไปค้นหาภาพที่มีลักษณะใกล้เคียงกับบรรจุภัณฑ์ละ 2 ภาพ โดยข้อมูลที่ใช้ในการค้นหาจะผ่านขั้นตอนการกรองบรรจุภัณฑ์ด้วยขนาดเบเกอรี่มาแล้ว เพื่อเพิ่มทางเลือกบรรจุภัณฑ์ที่มีลักษณะใกล้เคียงกับบรรจุภัณฑ์ที่เหมาะสมกับอาหาร

3.9.1 นำเข้าสู่ชุดข้อมูลฟีเจอร์ทั้งหมดของบรรจุภัณฑ์นำผลลัพธ์ซึ่งถูกดำเนินการสกัดฟีเจอร์แล้วเข้าสู่กระบวนการ

3.9.2 แปลงข้อมูลให้อยู่ในรูปแบบที่โมเดล InceptionV3 รองรับ จากนั้นนำผลลัพธ์บรรจุภัณฑ์ทั้ง 5 ภาพซึ่งมีค่าความมั่นใจ 5 อันดับแรก แปลงให้อยู่ในรูปแบบอาร์เรย์ และเพิ่มมิติให้เป็น (1, 299, 299, 3)

3.9.3 สกัดคุณสมบัติของภาพผลลัพธ์ด้วยโมเดล InceptionV3 โดยนำโมเดล InceptionV3 ซึ่งตัดเลเยอร์ชั้นบนสุดออก นำมาสกัดฟีเจอร์ของผลลัพธ์บรรจุภัณฑ์ทั้ง 5 ภาพ และเก็บคุณสมบัติทั้งหมดไว้ในรูปแบบ List

3.9.4 การค้นหาความใกล้เคียงหรือความคล้ายคลึง โดยนำคุณสมบัติของบรรจุภัณฑ์ทั้ง 5 ภาพนำมาค้นหาความใกล้เคียงระหว่างฟีเจอร์ โดยในงานวิจัยนี้เสนอการค้นหา Nearest Neighbors ด้วยเทคนิค Cosine similarity ซึ่งเป็นกระบวนการทางคณิตศาสตร์ที่ใช้ในการวัดความคล้ายคลึงระหว่างเวกเตอร์ของข้อมูล โดยกำหนดให้ค่า Nearest Neighbors n (n\_neighbors) เป็น 3 เพื่อตัดผลลัพธ์ลำดับแรกที่เป็นภาพบรรจุภัณฑ์ที่ใช้ค้นหาออก ดังการทำงานในภาพที่ 24

```

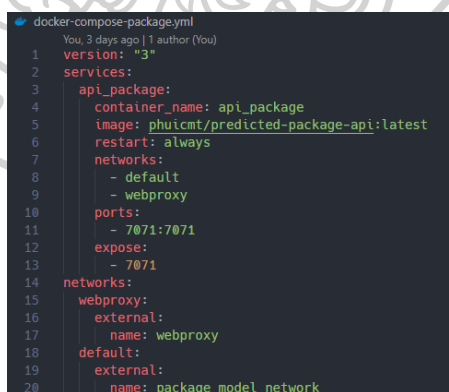
1. FUNCTION get_nearest_neighbors(product_features, features, packaging_ids, n_neighbors=2):
2.   # Step 1: Convert product_features to a NumPy array
3.   product_features <- Convert product_features to NumPy array
4.
5.   # Step 2: Extract all feature vectors from 'features'
6.   all_features <- Initialize an empty list
7.   FOR each sizes IN features:
8.     FOR each products IN sizes:
9.       FOR each feature IN products:
10.        Append feature to all_features
11.
12.   # Step 3: Reshape all_features into 2D array
13.   Convert all_features to NumPy array
14.   Reshape all_features to ensure it has 2 dimensions
15.
16.   # Step 4: Fit Nearest Neighbors model using cosine similarity
17.   nn <- Create NearestNeighbors model with:
18.     - n_neighbors = (n_neighbors + 1)
19.     - metric = "cosine"
20.   Fit nn with all_features
21.
22.   # Step 5: Find nearest neighbors for the product_features
23.   distances, indices <- Use nn to find k-nearest neighbors for product_features
24.
25.   # Step 6: Extract packaging IDs for nearest neighbors
26.   nearest_neighbors <- Initialize an empty list
27.   FOR each index IN indices[0]:
28.     IF index != 0 AND index < Length of packaging_ids:
29.       packaging_id <- packaging_ids[index]
30.       Append packaging_id to nearest_neighbors
31.   RETURN nearest_neighbors
32. END FUNCTION

```

ภาพที่ 24 กระบวนการค้นหา Nearest Neighbors ด้วยเทคนิค Cosine similarity

### 3.10. ประยุกต์ใช้โมเดลบนระบบ API

ผู้วิจัยได้เลือกใช้ Fast API: API Framework ร่วมกับ Docker ในการสร้างคอนเทนเนอร์สำหรับรวมโมเดลและการทำงานทั้งหมดไว้ เพื่อส่งออกโมเดลสำหรับทดลองใช้งานจริงซึ่ง โดยสร้างไฟล์ docker-compose-package.yml สำหรับจัดการคอนเทนเนอร์ API ดังภาพที่ 25



```

docker-compose-package.yml
You, 3 days ago | 1 author (You)
1 version: "3"
2 services:
3   api_package:
4     container_name: api_package
5     image: phuicmt/predicted-package-api:latest
6     restart: always
7     networks:
8       - default
9       - webproxy
10    ports:
11      - 7071:7071
12    expose:
13      - 7071
14    networks:
15      webproxy:
16        external:
17          name: webproxy
18      default:
19        external:
20          name: package_model_network

```

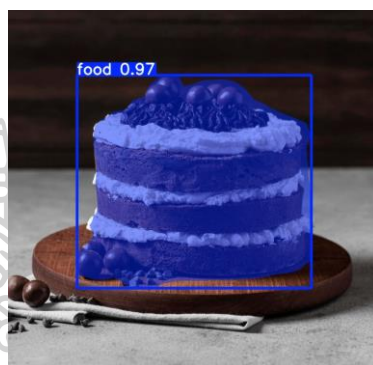
ภาพที่ 25 ข้อมูลภายในไฟล์ docker-compose-package.yml

โดยกระบวนการทำงานของระบบสามารถแบบได้ดังนี้

3.10.1 รับข้อมูลนำเข้าภาพของหวาน ระบบจะเริ่มต้นด้วยการรับภาพของหวานจากผู้ใช้ที่ส่งเข้ามาผ่านเว็บไซต์ ข้อมูลภาพจะถูกส่งผ่าน API และเข้าสู่ส่วน Backend โดยผู้ใช้งานสามารถระบุข้อมูล

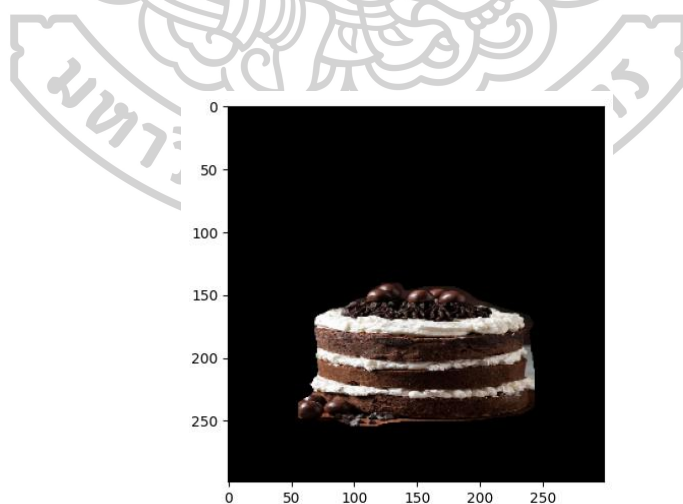
เพิ่มเติม เช่น ขนาดของบรรจุภัณฑ์ที่ต้องการ หรือประเภทของบรรจุภัณฑ์ (เช่น กล่องเค้กหรือถุงคุกกี้) เพื่อใช้ในการกรองบรรจุภัณฑ์ที่ตรงกับความต้องการ

3.10.2 ประมวลผลข้อมูลนำเข้า เมื่อระบบ Backend ได้รับภาพที่นำเข้า ภาพจะถูกส่งไปผ่านขั้นตอนการตัดพื้นหลังโดยใช้เทคนิค Image Segmentation ด้วยโมเดล YOLOv8 ที่ฝึกสอนมาแล้ว ซึ่งเมื่อนำภาพเข้าสู่โมเดล YOLOv8 Segmentation ผลลัพธ์ที่ได้จะอยู่ในรูปแบบ Mask ซึ่งเป็นค่าพื้นที่รอบวัตถุที่ตรวจจับได้ ดังภาพที่ 26



ภาพที่ 26 ผลลัพธ์การค้นหา Mask ด้วยโมเดล YOLOv8 Segmentation  
(ภาพต้นฉบับ [www.freepik.com](http://www.freepik.com), 2024)

จากนั้นคำนวณหาพื้นที่ส่วนนอกของ Mask และปรับสีให้เป็นสีดำด้วยวิธีการปรับค่าสีแต่ละ pixel ดังภาพที่ 27



ภาพที่ 27 ผลการปรับค่าสีแต่ละ pixel

ผลลัพธ์จากขั้นตอนนี้คือภาพที่ผ่านการตัดพื้นหลัง ซึ่งจะถูกละเปลี่ยนเป็นอาเรย์เพื่อใช้ในขั้นตอนต่อไปดังกระบวนการในภาพที่ 28

```

1. FUNCTION image_segmentation(image_path, target_size_dim)
2. # Step 1: Load and preprocess the input image
3.   image <- Load image from image_path with size target_size_dim
4.   image <- Convert image to array format
5. # Step 2: Predict segmentation masks using the model
6.   results <- Predict masks using model_seg
7. # Step 3: Initialize and combine masks
8.   mask_new <- NULL
9.   TRY
10.    FOR each result IN results
11.      FOR each mask IN result.masks
12.        IF mask_new IS NULL THEN
13.          mask_new <- Convert mask to numpy array
14.        ELSE
15.          mask_new <- Logical OR of mask_new and mask (numpy array)
16.    CATCH
17.      PRINT "No mask found"
18. # Step 4: Process the results to create food masks
19.   food_masks <- NULL
20.   FOR each result IN results
21.     query_image_rgb <- Convert result.orig_img from BGR to RGB
22.     food_masks <- Dictionary containing:
23.       - "query_image" = query_image_rgb
24.       - "mask" = mask_new
25. # Step 5: Apply the mask to the original image
26.   masked_image <- Copy of food_masks["query_image"]
27.   IF food_masks["mask"] IS NOT NULL THEN
28.     FOR each color channel (R, G, B)
29.       Multiply corresponding channel of masked_image with food_masks["mask"]
30. # Step 6: Return the final masked image
31.   RETURN masked_image
32. END FUNCTION

```

ภาพที่ 28 กระบวนการตัดพื้นหลังด้วยโมเดล YOLOv8 Segmentation

3.10.3 ทำนายด้วยโมเดลแบบหลายเลเบล ภาพในรูปแบบอาเรียจะถูกระมวลผลผ่านโมเดล ที่ฝึกสอนมาแล้วเพื่อทำนายบรรจุภัณฑ์ที่เหมาะสมกับเบเกอร์ที่นำเข้า ระบบจะกรองบรรจุภัณฑ์โดยพิจารณาจากขนาดของเบเกอร์หรือประเภทบรรจุภัณฑ์ที่ผู้ใช้ระบุ เพื่อจำกัดการค้นหาภายในขอบเขตของขนาดหรือประเภทที่ระบุไว้ ผลลัพธ์จะเป็นอาเรียที่แสดงความน่าจะเป็นของบรรจุภัณฑ์ที่เหมาะสมที่สุด 5 อันดับแรก ตามความมั่นใจที่ได้จากการทำนาย

3.10.4 สกัดฟีเจอร์บรรจุภัณฑ์ที่ได้จากการทำนายด้วยผลลัพธ์ 5 อันดับแรกของบรรจุภัณฑ์ที่ทำนายได้ นำเข้าสู่กระบวนการสกัดฟีเจอร์ ด้วยโมเดล InceptionV3 ที่ผ่านการฝึกสอนมาแล้วด้วย ImageNet ซึ่งปรับแต่ง โดยการตัดเลเยอร์ชั้นบนสุดออก เพื่อดึงข้อมูลฟีเจอร์สำคัญของบรรจุภัณฑ์ในรูปแบบอาเรียดังภาพที่ 29

```

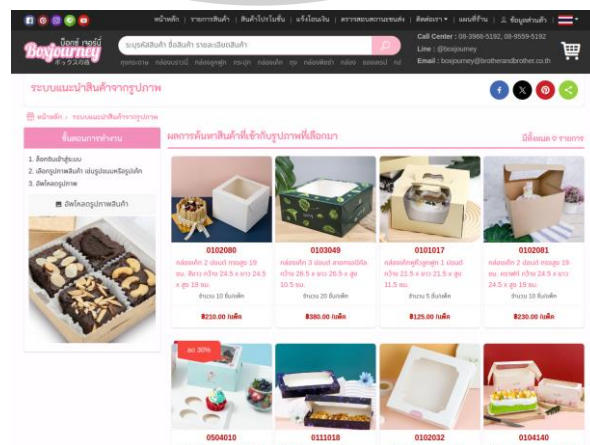
1. CLASS InceptionV3_Extraction:
2.   # Define the model using InceptionV3 pre-trained on ImageNet
3.   model <- LoadInceptionV3 model with:
4.     - include_top = False
5.     - weights = "imagenet"
6.     - pooling = "avg"
7.
8.   FUNCTION predict(image_path):
9.     # Step 1: Load the input image
10.    img <- Load image from image_path with target size (299, 299)
11.
12.    # Step 2: Resize the image to (299, 299) (if needed)
13.    img <- Resize img to size (299, 299)
14.
15.    # Step 3: Expand image dimensions to match model input
16.    img <- Expand dimensions of img to (1, 299, 299, 3)
17.
18.    # Step 4: Use the model to predict features
19.    features <- modelpredict(img)
20.
21.    RETURN features
22. END CLASS

```

ภาพที่ 29 การทำ Feature Extraction ด้วยโมเดล InceptionV3

3.10.5 ค้นหาบรรจุภัณฑ์ที่มีลักษณะใกล้เคียง ระบบจะนำพีเจอร์ของบรรจุภัณฑ์ที่สกัดได้มาเปรียบเทียบกับพีเจอร์ของบรรจุภัณฑ์ทั้งหมดในฐานข้อมูลโดยใช้เทคนิค Cosine Similarity เพื่อค้นหาบรรจุภัณฑ์ที่มีพีเจอร์ใกล้เคียงกันมากที่สุด ผลลัพธ์จะเป็นบรรจุภัณฑ์ 2 อันดับแรกที่มีพีเจอร์คล้ายคลึงกันมากที่สุดกับบรรจุภัณฑ์ที่ทำนายได้

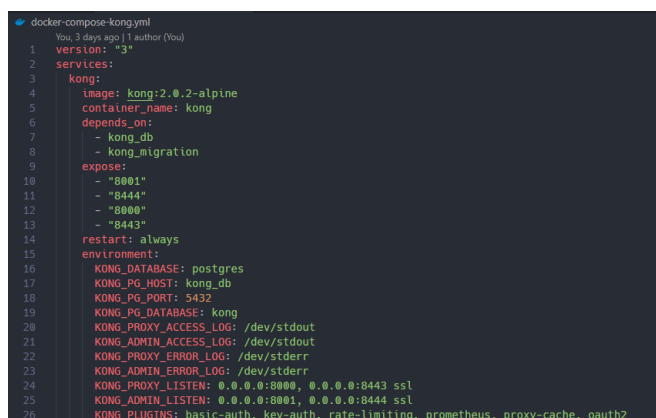
3.10.6 รวบรวมและส่งออกผลลัพธ์ ระบบจะรวบรวมผลลัพธ์บรรจุภัณฑ์ที่ได้จากการทำนายด้วยโมเดลหลายเลเบลและผลการเปรียบเทียบพีเจอร์ด้วย Nearest Neighbors จาก Cosine Similarity และส่งผลลัพธ์กลับไปยังผู้ใช้ในรูปแบบ JSON ซึ่งประกอบไปด้วยบรรจุภัณฑ์ที่เหมาะสมและใกล้เคียงกับเบเกอร์ที่ผู้ใช้ระบุต้องการทำงานในภาพที่ 30



ภาพที่ 30 ระบบแนะนำสินค้าด้วยรูปภาพจากเว็บไซต์ [www.boxjourney.com](http://www.boxjourney.com)

### 3.11. พัฒนาระบบ API Gateway

การใช้ Kong เพื่อจัดการการรับส่งคำร้องระหว่างผู้ใช้และระบบหลังบ้าน ซึ่งช่วยจัดการและควบคุมการเชื่อมต่อระหว่าง API ต่างๆ ในการทำงาน Kong API Gateway มีความยืดหยุ่นและปรับขยายได้ง่าย ช่วยให้ระบบสามารถรับมือกับการเรียกใช้งานในปริมาณมากได้ซึ่งได้ใช้ Kong มาใช้ในการจัดการระบบ ดังภาพที่ 32 โดยการนำระบบขึ้นสู่เซิร์ฟเวอร์จริงด้วย Docker Container ซึ่งสร้างไฟล์ docker-compose-kong.yml สำหรับจัดการ docker หลาย container ดังตัวอย่างภาพที่ 31



```

1 version: "3"
2 services:
3   kong:
4     image: kong:2.0.2-alpine
5     container_name: kong
6     depends_on:
7       - kong_db
8       - kong_migration
9     expose:
10      - "8001"
11      - "8444"
12      - "8000"
13      - "8443"
14     restart: always
15     environment:
16       KONG_DATABASE: postgres
17       KONG_PG_HOST: kong_db
18       KONG_PG_PORT: 5432
19       KONG_PG_DATABASE: kong
20       KONG_PROXY_ACCESS_LOG: /dev/stdout
21       KONG_ADMIN_ACCESS_LOG: /dev/stdout
22       KONG_PROXY_ERROR_LOG: /dev/stderr
23       KONG_ADMIN_ERROR_LOG: /dev/stderr
24       KONG_PROXY_LISTEN: 0.0.0.0:8000, 0.0.0.0:8443 ssl
25       KONG_ADMIN_LISTEN: 0.0.0.0:8001, 0.0.0.0:8444 ssl
26       KONG_PLUGINS: basic-auth, key-auth, rate-limiting, prometheus, proxy-cache, oauth2

```

ภาพที่ 31 ตัวอย่างไฟล์ docker-compose-kong.yml

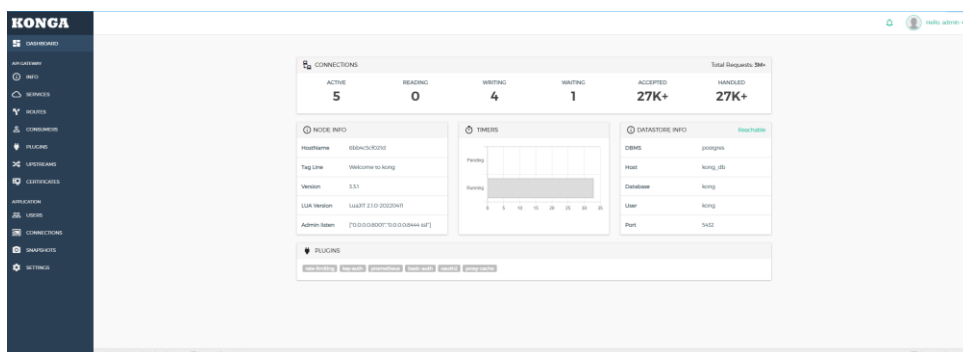
Kong API Gateway นอกจากจะช่วยควบคุมเส้นทางและการจัดการ Traffic แล้ว ยังถูกนำมาใช้ในการออกแบบระบบ Logging เพื่อบันทึกและวิเคราะห์ข้อมูลการทำงานของระบบ รวมถึงการวิเคราะห์ประสิทธิภาพของโมเดล Machine Learning อย่างไรก็ตาม เนื่องจากข้อจำกัดทางเทคนิคในกระบวนการ Logging ที่ไม่สามารถใช้งานในสภาพแวดล้อมปัจจุบันได้ จึงสามารถดำเนินการในขั้นตอนออกแบบโครงสร้าง Log และกำหนดแนวทางการพัฒนาต่อไปในอนาคตได้ ดังนี้

3.11.1. Monitoring Plugins สำหรับเก็บข้อมูลการเรียกใช้งาน API เช่น ระยะเวลาการตอบสนอง (Response Time) และอัตราความสำเร็จของการเรียกใช้งาน API แต่ละคำร้อง

3.11.2. Logging Plugins สำหรับบันทึกคำร้องและผลลัพธ์จากการทำนายของโมเดล เพื่อช่วยวิเคราะห์ประสิทธิภาพของโมเดลที่ถูกนำขึ้นระบบ

3.11.3. Rate Limiting และ Circuit Breaking ช่วยควบคุมปริมาณคำร้องที่เข้ามาและลดภาระการทำงานของระบบ

การใช้งาน Kong ทำให้สามารถติดตาม ประสิทธิภาพของโมเดล Machine Learning ได้อย่างต่อเนื่องและทำให้มั่นใจได้ว่าระบบจะทำงานอย่างมีประสิทธิภาพเมื่อถูกนำขึ้นใช้งานจริง



ภาพที่ 32 Konga Dashboard

### 3.12. ประเมินผลการทำงานระบบแนะนำบรรจุภัณฑ์

ในการประเมินผลการทำงานของระบบแนะนำบรรจุภัณฑ์ จะวัดประสิทธิภาพในหลายมิติ เพื่อให้มั่นใจว่าระบบสามารถแนะนำบรรจุภัณฑ์ที่เหมาะสมกับผลิตภัณฑ์เบเกอรี่ได้อย่างมีประสิทธิภาพ โดยการประเมินผลแบ่งออกเป็น 3 ส่วนหลัก ดังนี้

#### 3.12.1. ความแม่นยำของการแนะนำ

ระบบจะถูกประเมินโดยใช้ตัวชี้วัดทางสถิติ เช่น Precision, Recall, และ F1 Score เพื่อวัดความแม่นยำของการแนะนำบรรจุภัณฑ์ที่เหมาะสมกับภาพผลิตภัณฑ์เบเกอรี่ ความแม่นยำของระบบขึ้นอยู่กับความสามารถของโมเดลในการคาดการณ์บรรจุภัณฑ์ที่ตรงกับการใช้งานจริง

#### 3.12.2. Mean Average Precision (mAP@0.5)

การวัดค่า mAP@0.5 ถูกนำมาใช้เพื่อตรวจสอบประสิทธิภาพของโมเดลการจำแนกประเภทแบบหลายเลเบลในการแนะนำบรรจุภัณฑ์ ค่า mAP คำนวณจากความแม่นยำเฉลี่ยของโมเดลในการคาดการณ์ที่มีค่า IoU  $\geq 0.5$  ซึ่งช่วยในการประเมินว่าโมเดลสามารถจับคู่บรรจุภัณฑ์ได้แม่นยำเพียงใด

#### 3.12.3. ประสิทธิภาพของระบบ (System Performance)

ประสิทธิภาพการทำงานของระบบจะถูกประเมินผ่านการวัดเวลาในการตอบสนอง (Response Time) ของระบบ API โดย Kong API Gateway จะช่วยตรวจสอบและบันทึกข้อมูลเกี่ยวกับการทำงานของระบบ เช่น Latency และ Throughput เพื่อให้แน่ใจว่าระบบสามารถตอบสนองคำร้องในปริมาณมากได้ในเวลาที่เหมาะสม

ผลลัพธ์ที่ได้จากการประเมินนี้จะถูกนำมาใช้ในการปรับปรุงและปรับแต่งโมเดลเพิ่มเติม เพื่อเพิ่มประสิทธิภาพในการแนะนำบรรจุภัณฑ์และยกระดับประสบการณ์ของผู้ใช้งาน

## บทที่ 4

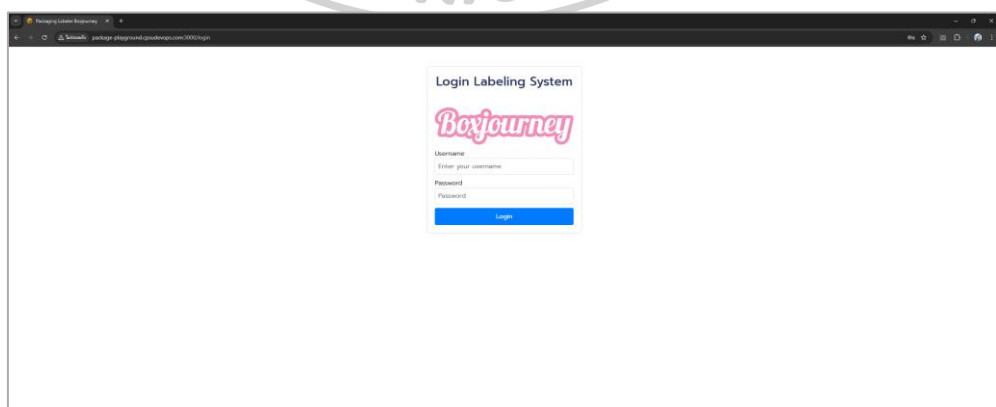
### ผลการดำเนินงานวิจัย

จากขั้นตอนการดำเนินงานที่ได้กล่าวถึงในบทที่ 3 วิทยานิพนธ์นี้ได้ออกแบบ และพัฒนา โมเดล และทดสอบระบบแนะนำบรรจุภัณฑ์โดยใช้เทคนิคการเรียนรู้เชิงลึก และการประมวลผลภาพ สำหรับผลิตภัณฑ์เบเกอรี่ อีกทั้งได้พัฒนาระบบ API Service สำหรับนำโมเดลไปปรับใช้ให้เข้ากับ สถานการณ์จริง จึงแบ่งส่วนของผลการวิจัย ดังนี้

#### 4.1. การพัฒนาเว็บแอปพลิเคชันและเก็บผลข้อมูลเลเบล

ในการพัฒนาระบบแนะนำบรรจุภัณฑ์อาหารเบเกอรี่ การเลเบลข้อมูลเป็นขั้นตอนสำคัญที่ต้องทำอย่างแม่นยำเพื่อให้โมเดลสามารถเรียนรู้และจำแนกประเภทได้ถูกต้อง ดังนั้น เว็บแอปพลิเคชันที่พัฒนาขึ้นในงานวิจัยนี้จึงมีจุดประสงค์เพื่ออำนวยความสะดวกในการเลเบลข้อมูลให้กับผู้ใช้และผู้วิจัย โดยสามารถทำงานผ่านอินเทอร์เน็ตเฟสที่ใช้งานง่ายและมีฟังก์ชันที่ครอบคลุมในกระบวนการเลเบลข้อมูล โดยรองรับการบันทึกผลการเลเบลลงในฐานข้อมูล MongoDB

เว็บแอปพลิเคชันนี้ถูกพัฒนาด้วย FastAPI สำหรับฝั่ง Backend และ React สำหรับฝั่ง Frontend โดยการออกแบบของแอปพลิเคชันมีการคำนึงถึงความสะดวกในการใช้งานโดยให้ผู้เลเบลกรอกรหัสประจำตัวเพื่อบันทึกผลการเลเบลดังภาพที่ 33 และให้ระบบบันทึกข้อมูลเลเบลที่ได้ลงในฐานข้อมูลพร้อมกับรหัสผู้ใช้ในทุกการบันทึก ซึ่งทำให้สามารถติดตามและระบุได้ว่าข้อมูลใดถูกเลเบลโดยผู้ใด และสามารถนำข้อมูลนั้นมาใช้เพื่อแสดงผลการเลเบลที่ค้างอยู่ของผู้ใช้นั้นได้ดังตัวอย่างข้อมูลในภาพที่ 34

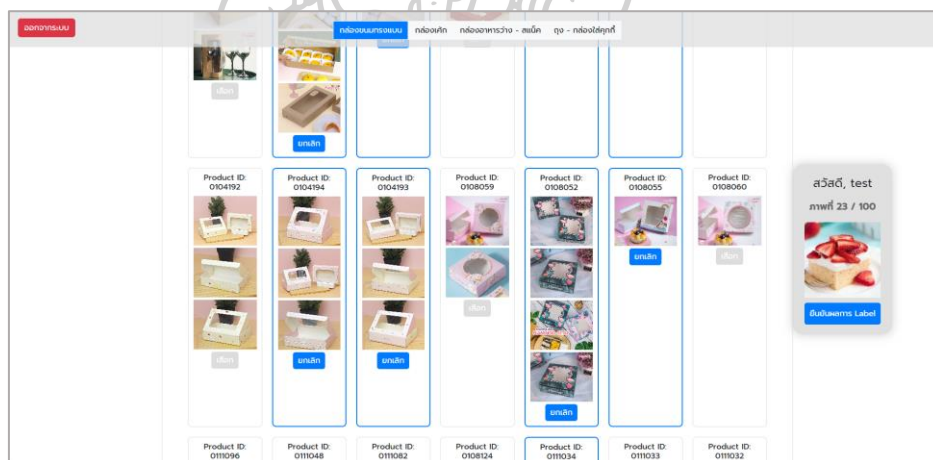


ภาพที่ 33 หน้าต่าง Login เข้าสู่ระบบสำหรับผู้เลเบลข้อมูล

```
{
  "username": "boxjourney1",
  "images": [
    "bakery_47.jpg",
    "bakery_176.jpg"
  ],
  "current_index": 2,
  "labels": {
    "bakery_47__dot__jpg": [
      "0111042",
      "0104159",
      "0104196"
    ],
    "bakery_176__dot__jpg": [
      "0102031",
      "0103064"
    ]
  },
  "last_saved": "2024-11-17 20:15:24"
}
```

ภาพที่ 34 ตัวอย่างบันทึกผลลัพธ์ข้อมูลเลเบล

โดยระบบสามารถแสดงภาพเบเกอร์รี่ให้ผู้ใช้งานเลือกบรรจุภัณฑ์ผ่านหน้าจอที่ใช้งานง่ายและสามารถเลือกบรรจุภัณฑ์ได้มากกว่า 1 ประเภท และกดยืนยันผลการเลเบล ระบบจะบันทึกผลในฐานข้อมูลโดยตรง ดังภาพที่ 35



ภาพที่ 35 ตัวอย่างการเลือกบรรจุภัณฑ์ผ่านเว็บแอปพลิเคชัน

โดยผลการเก็บข้อมูลการเลเบลจากผู้เชี่ยวชาญดังตารางที่ 7 พบว่า สินค้าในภาพเบเกอร์รี่แต่ละภาพถูกเลือกซ้ำโดยผู้เชี่ยวชาญหลายคน โดยผลการวิเคราะห์แสดงให้เห็นว่าบรรจุภัณฑ์ในภาพเบเกอร์รี่แต่ละภาพถูกเลือกซ้ำจากผู้เชี่ยวชาญโดยเฉลี่ย 29.81 รายการต่อภาพ ซึ่งผู้เชี่ยวชาญเลือกบรรจุภัณฑ์ในแต่ละภาพเบเกอร์รี่เฉลี่ย 108.63 รายการต่อภาพ แสดงให้เห็นว่าผู้เชี่ยวชาญมีแนวโน้มที่จะเลือกบรรจุภัณฑ์ที่คล้ายคลึงกันในภาพเดียวกัน ซึ่งสะท้อนถึงความสอดคล้องของการตัดสินใจในการเลือกบรรจุภัณฑ์ที่เหมาะสมกับเบเกอร์รี่ในแต่ละภาพ นอกจากนี้ ค่าเฉลี่ยของจำนวนรายการ

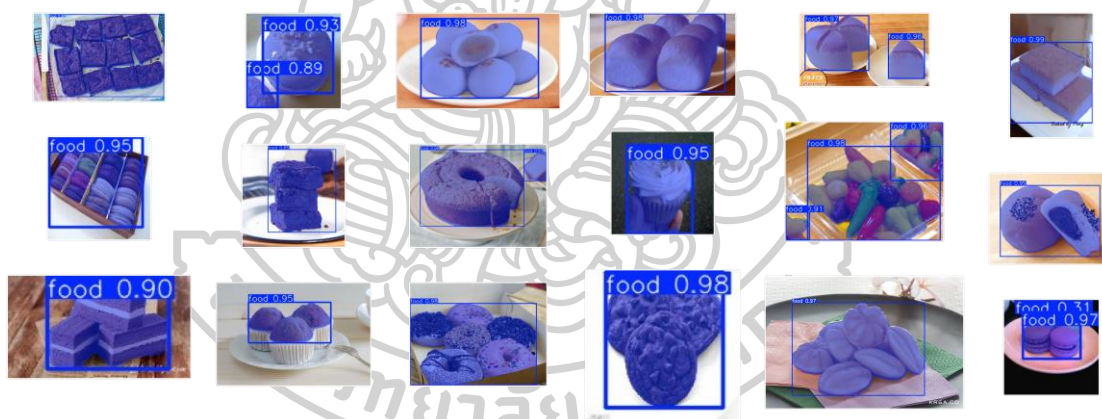
ที่เลือกทั้งหมดยังบ่งบอกถึงการพิจารณาที่รอบคอบและหลากหลาย โดยผู้เชี่ยวชาญได้ประเมินตัวเลือกบรรจุก่อนอย่างละเอียด เพื่อให้ครอบคลุมความเหมาะสมกับลักษณะของผลิตภัณฑ์เบเกอรี่

ตารางที่ 7 ผลการวิเคราะห์ข้อมูลการเลเบลโดยผู้เชี่ยวชาญ

| เกณฑ์                                                                   | ค่าเฉลี่ย |
|-------------------------------------------------------------------------|-----------|
| บรรจุก่อนที่ถูกเลือกซ้ำเฉลี่ยต่อ 1 ภาพเบเกอรี่ (รายการ)                 | 29.81     |
| บรรจุก่อนที่ถูกเลือกเฉลี่ยต่อ 1 ผู้เชี่ยวชาญ ต่อ 1 ภาพเบเกอรี่ (รายการ) | 108.63    |

#### 4.2. ผลการทดสอบประสิทธิภาพโมเดลคัดแยกพื้นหลังออกจากวัตถุ (Image Segmentation)

โมเดล YOLOv8 ถูกนำมาใช้ในการแบ่งส่วนพื้นหลังออกจากภาพอาหาร ผลลัพธ์จากการฝึกสอนด้วยชุดข้อมูล FoodSeg103 [40] แสดงให้เห็นว่าแต่ละโมเดลมีความสามารถในการแยกวัตถุอาหารจากพื้นหลังได้อย่างแม่นยำดังตัวอย่างผลลัพธ์ในภาพที่ 36 โดยวัดผลจากค่า Precision, Recall, mAP50, และ mAP50-95 ของแต่ละโมเดล โดยแบ่งค่าของ Bounding Box (B) และ Mask (M) สำหรับตรวจจับวัตถุและคัดแยกวัตถุ ดังตารางที่ 8



ภาพที่ 36 ตัวอย่างผลลัพธ์การแยกวัตถุอาหารจากพื้นหลังด้วยโมเดล YOLOv8x-seg

ตารางที่ 8 ตารางเปรียบเทียบประสิทธิภาพโมเดลหลังจากฝึกสอนด้วยชุดข้อมูล FoodSeg103

| Model       | Precision (B) (%) | Recall (B) (%) | mAP50 (B) (%) | mAP50-95 (B) (%) | Precision (M) (%) | Recall (M) (%) | mAP50 (M) (%) | mAP50-95 (M) (%) |
|-------------|-------------------|----------------|---------------|------------------|-------------------|----------------|---------------|------------------|
| YOLOv8n-seg | 90.8              | 85.5           | 91.9          | 82.9             | 91.6              | 85.9           | 92.4          | 76.6             |
| YOLOv8s-seg | 92.4              | 84.8           | 92.2          | 84.2             | 93.1              | 85.4           | 92.4          | 78.6             |
| YOLOv8m-seg | 91.8              | 85.9           | 92.4          | 84.9             | 92.6              | 86.6           | 92.6          | 80.3             |
| YOLOv8l-seg | 93.0              | 86.8           | 92.5          | 92.5             | 94.0              | 87.6           | 92.9          | 80.8             |
| YOLOv8x-seg | 93.9              | 86.2           | 93.1          | 87.0             | 94.6              | 86.7           | 93.4          | 82.3             |

หลังจากที่ได้ผลลัพธ์จากการฝึกสอนโมเดลด้วยชุดข้อมูล FoodSeg103 แล้ว ตารางที่ 8 แสดงผลการเปรียบเทียบประสิทธิภาพของแต่ละโมเดลโดยพิจารณาจากตัวชี้วัดหลัก เช่น Precision (B), Recall (B), mAP50 (B), mAP50-95 (B), รวมถึงตัวชี้วัดแบบ (M) ซึ่งมาจากการวัดผลทั้งกลุ่มของ Box (B) และ Mask (M) ตามลำดับ โมเดล YOLOv8 ทั้ง 5 รุ่นได้ถูกทดสอบเพื่อเปรียบเทียบความสามารถในการแยกวัตถุและการทำนายผลจากภาพ ซึ่งผลการทดสอบชี้ให้เห็นถึงประสิทธิภาพที่แตกต่างกันดังนี้

4.2.1. YOLOv8n-seg แม้ว่า เป็นโมเดลที่มีขนาดเล็กที่สุด แต่สามารถทำ Precision ได้สูงถึง 90.8% และ Recall ที่ 85.5% สำหรับการทำนายแบบ Box และมีค่า mAP50 สูงถึง 91.9% ซึ่งแสดงถึงความแม่นยำในการตรวจจับวัตถุที่อยู่ในกรอบได้ดี อย่างไรก็ตาม ค่า mAP50-95 ที่ 82.9% บ่งบอกถึงความยากลำบากในการตรวจจับวัตถุที่มีขนาดต่าง ๆ กัน ส่วนการทำนาย Mask ก็ได้ผลลัพธ์ที่ใกล้เคียงกัน โดยมีค่า Precision ที่ 91.6% และค่า mAP50-95 อยู่ที่ 76.6%

4.2.2. YOLOv8s-seg แม้ว่า YOLOv8s-seg จะเป็นโมเดลที่มีขนาดเล็กกว่า YOLOv8l-seg หรือ YOLOv8x-seg แต่ยังคงแสดงผลที่ดีในด้านการทำนายวัตถุ ค่า Precision สำหรับการทำนายแบบ Box อยู่ที่ 92.4% และ Recall อยู่ที่ 84.8% ซึ่งแสดงถึงความสามารถในการตรวจจับวัตถุในภาพได้อย่างมีประสิทธิภาพ นอกจากนี้ ค่า mAP50 สำหรับการทำนาย Box อยู่ที่ 92.2% และค่า mAP50-95 ที่ 84.2% ซึ่งบ่งบอกถึงความสามารถในการจัดการกับวัตถุที่มีขนาดและรูปร่างต่าง ๆ กันได้สำหรับการทำนาย Mask, YOLOv8s-seg ก็ให้ผลลัพธ์ที่ใกล้เคียงกัน โดยมีค่า Precision (M) อยู่ที่ 93.1% และ mAP50-95 (M) อยู่ที่ 78.6% ซึ่งทำให้เห็นว่าโมเดลนี้สามารถทำงานได้ดีทั้งในด้านการทำนายตำแหน่งและรูปร่างของวัตถุ

4.2.3. YOLOv8m-seg โมเดลนี้มีขนาดใหญ่กว่า YOLOv8n-seg ทำให้มีประสิทธิภาพที่ดีขึ้นเล็กน้อย Precision และ Recall สำหรับการทำนายแบบ Box อยู่ที่ 91.8% และ 85.9% ตามลำดับ และค่า mAP50 ที่ 92.4% สูงกว่าของ YOLOv8n-seg อย่างไรก็ตาม โมเดลนี้ยังคงทำงานได้ดีในด้าน Mask ด้วย Precision 92.6% และค่า mAP50-95 ที่ 80.3%

4.2.4. YOLOv8l-seg เป็นหนึ่งในโมเดลที่มีประสิทธิภาพสูงสุด Precision (B) 93.0% และ Recall (B) 86.8% รวมถึงค่า mAP50 (B) 92.5% และ mAP50-95 (B) 92.5% ที่แสดงถึงความสามารถในการตรวจจับและการทำนายวัตถุในภาพได้อย่างแม่นยำสูง โมเดลนี้ทำผลลัพธ์ได้ดีกับ Mask เช่นเดียวกัน โดยมีค่า Precision (M) 94.0% และ mAP50-95 (M) ที่ 80.8%

4.2.5. YOLOv8x-seg เป็นโมเดลที่ใหญ่ที่สุดในกลุ่มและให้ผลลัพธ์ที่ดีที่สุด Precision (B) อยู่ที่ 93.9% และค่า mAP50-95 (B) สูงถึง 87.0% ซึ่งทำให้การตรวจจับวัตถุมีความแม่นยำสูง นอกจากนี้ การทำนาย Mask ยังได้ค่า Precision สูงถึง 94.6% และค่า mAP50-95 ที่ 82.3% แสดงถึงประสิทธิภาพในการตรวจจับวัตถุขนาดต่าง ๆ ได้ดี

ตารางที่ 9 เวลาในการทำนายแต่ละโมเดล

| Model       | Prediction Time (ms) |
|-------------|----------------------|
| YOLOv8n-seg | 95                   |
| YOLOv8s-seg | 157                  |
| YOLOv8m-seg | 212                  |
| YOLOv8l-seg | 285                  |
| YOLOv8x-seg | 578                  |

จากตารางที่ 9 แสดงเวลาในการทำนายของแต่ละโมเดล พบว่าโมเดลที่มีขนาดเล็กกว่า เช่น YOLOv8n-seg และ YOLOv8s-seg ใช้เวลาทำนายบนการ์ดจอ NVIDIA GeForce RTX 4070 Ti น้อยกว่าโมเดลขนาดใหญ่ เช่น YOLOv8x-seg และ YOLOv8l-seg ซึ่งเป็นไปตามการออกแบบของโมเดลที่มีขนาดใหญ่ขึ้นและซับซ้อนมากขึ้น

โดยสรุปโมเดล YOLOv8 แต่ละรุ่นให้ประสิทธิภาพที่แตกต่างกันตามขนาดของโมเดล โดยโมเดลขนาดใหญ่เช่น YOLOv8l-seg และ YOLOv8x-seg แสดงผลลัพธ์ที่ดีที่สุดในด้าน Precision และ Recall อย่างไรก็ตาม โมเดลขนาดเล็กกว่า เช่น YOLOv8n-seg และ YOLOv8s-seg มีข้อได้เปรียบในด้านความเร็วในการทำนาย ซึ่งอาจเหมาะสมกับแอปพลิเคชันที่ต้องการความรวดเร็วในการประมวลผล

### 4.3. ผลการฝึกสอนและทดสอบโมเดลการจำแนกประเภทแบบหลายเลเบล (Multi-label Classification)

จากการฝึกสอนโมเดลจำแนกประเภทแบบหลายเลเบลที่ใช้โมเดล VGG19, VGG16, ResNet50 และ InceptionV3 ภายใต้กลยุทธ์การปรับแต่งโมเดล (Fine-Tuning) ทั้ง 4 แบบ (S1, S2, S3, S4) ได้ผลลัพธ์การประเมินประสิทธิภาพของโมเดลแต่ละตัว ซึ่งประกอบด้วยค่า Precision, Recall, F1-Score และ Mean Average Precision (mAP50) ซึ่งเป็นเกณฑ์ในการวัดประสิทธิภาพของการทำนายแบบหลายเลเบล

เนื่องจากในระหว่างการฝึกสอนโมเดล พบว่าข้อมูลที่มีอยู่มีจำนวนจำกัดและไม่เพียงพอต่อการฝึกสอนโมเดลที่มีประสิทธิภาพสูง นอกจากนี้ข้อมูลเบเกอร์ที่ได้มานั้นมีความไม่สมดุลกัน (Imbalanced Data) กล่าวคือ มีจำนวนข้อมูลของเบเกอร์บางประเภทที่มากกว่าประเภทอื่นๆ ทำให้การฝึกสอนโมเดลไม่สามารถเรียนรู้ได้อย่างเต็มที่และเกิดปัญหาความแม่นยำที่ต่ำลง ผลการทดสอบจากการฝึกสอนก่อนการใช้ Data Augmentation มีค่าที่ต่ำ ดังแสดงในตารางที่ 10

ตารางที่ 10 ผลการประเมินประสิทธิภาพโมเดลการจำแนกแบบหลายเลเบล (ก่อน Data Augmentation)

| Model       | Strategy | Precision (%) | Recall (%) | F1-Score (%) | mAP50 (%) |
|-------------|----------|---------------|------------|--------------|-----------|
| VGG19       | S1       | 3.61          | 1.89       | 1.95         | 15.02     |
|             | S2       | 3.71          | 2          | 2.14         | 15.09     |
|             | S3       | 6.25          | 8.11       | 6.24         | 15.3      |
|             | S4       | 6.46          | 5.33       | 4.89         | 15.29     |
| VGG16       | S1       | 3.81          | 1.83       | 2.06         | 15.11     |
|             | S2       | 4.54          | 1.87       | 2.13         | 15.14     |
|             | S3       | 11.33         | 10.04      | 9.14         | 16.09     |
|             | S4       | 5.36          | 5.94       | 4.64         | 15.1      |
| ResNet50    | S1       | 1.73          | 1.83       | 1.34         | 14.9      |
|             | S2       | 1.36          | 1.44       | 1.06         | 14.82     |
|             | S3       | 6.67          | 30.16      | 9.68         | 15.02     |
|             | S4       | 5.1           | 18.93      | 6.97         | 14.99     |
| InceptionV3 | S1       | 6.37          | 5.72       | 5.61         | 15.19     |
|             | S2       | 7.16          | 5.57       | 5.72         | 15.37     |
|             | S3       | 8.66          | 8.89       | 7.47         | 15.71     |
|             | S4       | 7.69          | 8.65       | 7.01         | 15.34     |

จากตารางที่ 10 สังเกตได้ว่าค่า Precision และ Recall ค่อนข้างต่ำ ซึ่งแสดงถึงความสามารถของโมเดลในการจำแนกประเภทที่ยังไม่แม่นยำพออีกทั้งค่า mAP50 และ mAP50-95 อยู่ในระดับต่ำเช่นกัน ซึ่งเป็นตัวชี้วัดสำคัญที่บ่งชี้ว่าโมเดลยังไม่สามารถจำแนกประเภทได้อย่างมีประสิทธิภาพ

จากปัญหาดังกล่าว การทำ Data Augmentation [25] เป็นสิ่งจำเป็นเพื่อเพิ่มข้อมูลในการฝึกสอนโมเดล เพิ่มความหลากหลายของข้อมูล และช่วยให้โมเดลสามารถเรียนรู้ได้ดีขึ้นจากข้อมูล

ที่มากขึ้น ผู้วิจัยจึงเพิ่มข้อมูลด้วย Data Augmentation โดยผลการฝึกสอนได้ผลลัพธ์ดังตารางที่ 10 ซึ่งสามารถสรุปได้ว่าโมเดลที่มีประสิทธิภาพสูงสุดในการจำแนกประเภทแบบหลายเลเบลคือ InceptionV3 ซึ่งให้ค่า Precision และ Recall สูงสุดในกลุ่มโมเดลที่ทดสอบ อีกทั้งการใช้ Data Augmentation ยังช่วยปรับปรุงประสิทธิภาพของโมเดลอย่างชัดเจน

ตารางที่ 11 ผลการประเมินประสิทธิภาพโมเดลการจำแนกแบบหลายเลเบล (หลัง Data Augmentation)

| Model       | Strategy | Precision (%) | Recall (%) | F1-Score (%) | mAP50 (%) |
|-------------|----------|---------------|------------|--------------|-----------|
| VGG19       | S1       | 18.28         | 5.24       | 6.71         | 14.67     |
|             | S2       | 9.1           | 2.89       | 3.46         | 13.83     |
|             | S3       | 85.85         | 67.34      | 74.62        | 62.11     |
|             | S4       | 2.78          | 1.53       | 1.29         | 13.42     |
| VGG16       | S1       | 14.3          | 4.02       | 4.9          | 14.23     |
|             | S2       | 16.87         | 5.69       | 7.13         | 14.74     |
|             | S3       | 86.44         | 76.9       | 80.98        | 69.72     |
|             | S4       | 1.57          | 1.97       | 1.32         | 13.4      |
| ResNet50    | S1       | 1.69          | 1.14       | 1            | 13.42     |
|             | S2       | 2.17          | 1.72       | 1.4          | 13.43     |
|             | S3       | 91.92         | 78.52      | 84.14        | 75.2      |
|             | S4       | 1.53          | 5          | 1.7          | 13.4      |
| InceptionV3 | S1       | 62.72         | 47.81      | 51.27        | 35.68     |
|             | S2       | 65.24         | 50.89      | 54.71        | 38.59     |
|             | S3       | 97.52         | 92.92      | 95.04        | 91.62     |
|             | S4       | 95.1          | 91.88      | 93.25        | 88.55     |

จากผลการฝึกสอนในตารางที่ 11 พบว่าโมเดล InceptionV3 ภายใต้กลยุทธ์ S3 ให้ผลลัพธ์ดีที่สุด โดยมีค่า Precision สูงถึง 97.52% ค่า Recall อยู่ที่ 92.92% ค่า F1-Score อยู่ที่ 95.04% และค่า mAP50 สูงถึง 91.62% นอกจากนี้ ผลการทดลองยังแสดงให้เห็นว่าโมเดลอื่นๆ เช่น VGG16 และ ResNet50 ภายใต้กลยุทธ์ S3 มีการปรับปรุงประสิทธิภาพขึ้นหลังการเพิ่มข้อมูล แต่ยังคงมีความแม่นยำน้อยกว่า InceptionV3 ในทุกตัวชี้วัด

สำหรับโมเดลในกลยุทธ์ S3, S4 โดยเฉพาะ InceptionV3 แสดงให้เห็นว่าเมื่ออนุญาตให้ปรับเลเยอร์เพิ่มเติมได้ ประสิทธิภาพของโมเดลจะสูงขึ้นอย่างมีนัยสำคัญ โดยค่า Precision และ

mAP@0.5 เพิ่มขึ้นอย่างต่อเนื่อง ซึ่งแสดงให้เห็นถึงความสำคัญของการ Fine-Tuning เลเยอร์ในกระบวนการจำแนกประเภทแบบหลายเลเบล

#### 4.4. การค้นหาบรรจุภัณฑ์ที่คล้ายกันด้วย Cosine Similarity

ในการค้นหาบรรจุภัณฑ์ที่คล้ายกัน ระบบที่พัฒนาขึ้นในงานวิจัยนี้ได้ใช้เทคนิค Cosine Similarity ซึ่งเป็นวิธีการที่นิยมใช้ในการวัดความคล้ายคลึงกันระหว่างเวกเตอร์ ในกรณีของการแนะนำบรรจุภัณฑ์ในงานวิจัยนี้ ภาพบรรจุภัณฑ์ถูกแปลงเป็นเวกเตอร์ของฟีเจอร์ผ่านกระบวนการสกัดฟีเจอร์จากโมเดล InceptionV3 โดยก่อนที่จะเข้าสู่ขั้นตอนการเปรียบเทียบกับ Cosine Similarity ระบบได้คัดกรองบรรจุภัณฑ์ตามขนาดก่อนเป็นลำดับแรก โดยระบบจะกรองเฉพาะบรรจุภัณฑ์ที่มีประเภทที่ตรงกันและ ขนาดเท่ากับหรือใหญ่กว่าสินค้าที่ต้องการบรรจุเท่านั้น วิธีนี้ช่วยจำกัดจำนวนบรรจุภัณฑ์ที่อาจไม่เหมาะสมออกไป และเพิ่มความแม่นยำในการแนะนำบรรจุภัณฑ์ที่เหมาะสมกับสินค้าที่ผู้ใช้สนใจ หลังจากนั้นฟีเจอร์เหล่านี้จะถูกนำไปคำนวณหาค่าความคล้ายคลึงกับเวกเตอร์ของฟีเจอร์ของบรรจุภัณฑ์ในฐานข้อมูลที่ผ่านมาแล้ว ซึ่งจะได้ผลลัพธ์ดังภาพที่ 37



ภาพที่ 37 ผลลัพธ์การค้นหาภาพจากค้นหาความคล้ายคลึงของฟีเจอร์ภาพ

Cosine Similarity มีข้อดีสำคัญคือไม่สนใจขนาดของเวกเตอร์ แต่พิจารณาเฉพาะมุมระหว่างเวกเตอร์เท่านั้น ทำให้เทคนิคนี้เหมาะสมสำหรับการประมวลผลภาพบรรจุภัณฑ์ที่มีขนาดแตกต่างกัน แต่มีลักษณะฟีเจอร์ที่คล้ายคลึงกัน นอกจากนี้ Cosine Similarity ยังมีความเร็วในการคำนวณสูงและมีความซับซ้อนต่ำ ซึ่งทำให้ระบบสามารถค้นหาบรรจุภัณฑ์ที่คล้ายกันได้

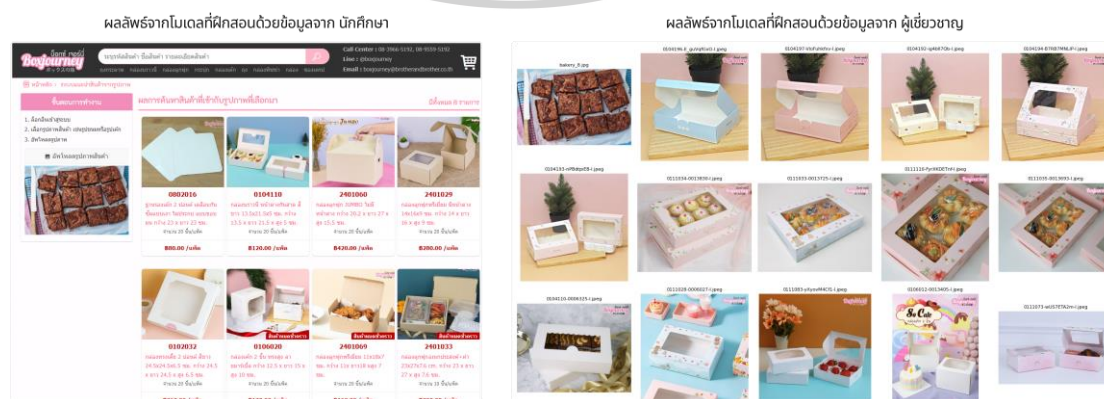
#### 4.5. ผลการเปรียบเทียบโมเดลที่ฝึกสอนด้วยข้อมูลจากนักศึกษาและผู้เชี่ยวชาญ

เพื่อวิเคราะห์ผลกระทบของกลุ่มผู้เลเบลที่มีต่อประสิทธิภาพของโมเดลแนะนำบรรจุภัณฑ์ งานวิจัยนี้ได้เปรียบเทียบผลลัพธ์ของโมเดลที่ฝึกสอนด้วยข้อมูลจากนักศึกษา 5 คน และ ผู้เชี่ยวชาญ 6 คน โดยทั้งสองกลุ่มเลเบลข้อมูลเบเกอร์รี่ชุดเดียวกัน แต่มีเงื่อนไขการเลเบลที่แตกต่างกันได้ว่า กลุ่มนักศึกษาแต่ละคนเลเบลข้อมูลภาพที่แตกต่างกันทั้งหมดโดยไม่มีการเลือกซ้ำกันส่งผลให้ข้อมูลที่ได้มีค่าเฉลี่ยของสินค้าที่เลือกต่อภาพต่อคนสูง โดยเฉลี่ยคนละ 236.98 รายการ แต่ในส่วนกลุ่มผู้เชี่ยวชาญมีการเลเบลภาพเดียวกันซ้ำในบางกรณี ส่งผลให้สามารถวัดความสอดคล้องในการตัดสินใจระหว่างผู้เลเบลได้ ค่าเฉลี่ยของสินค้าที่เลือกต่อภาพต่อคนลดลงเหลือเฉลี่ย 108.63 รายการ แต่มีค่าเฉลี่ยของสินค้าที่เลือกซ้ำในภาพเดียวกันสูงถึง 29.81 รายการดังตารางที่ 12

ตารางที่ 12 การเปรียบเทียบข้อมูลเลเบลและประสิทธิภาพของโมเดล

| ตัวชี้วัด                               | กลุ่มนักศึกษา (5 คน) | กลุ่มผู้เชี่ยวชาญ (6 คน) |
|-----------------------------------------|----------------------|--------------------------|
| ค่าเฉลี่ยสินค้าที่เลือกต่อภาพต่อคน      | 236.98               | 108.63                   |
| ค่าเฉลี่ยสินค้าที่เลือกซ้ำในภาพเดียวกัน | -                    | 29.81                    |
| Precision (%)                           | 95.30                | 97.52                    |
| Recall (%)                              | 77.94                | 92.92                    |
| F1-Score (%)                            | 84.87                | 95.04                    |

การเปรียบเทียบนี้แสดงให้เห็นว่าข้อมูลเลเบลจากกลุ่มผู้เชี่ยวชาญมีผลต่อการเพิ่มความแม่นยำและความครอบคลุมของโมเดลอย่างชัดเจน โดยเฉพาะในตัวชี้วัด Recall และ F1-Score ซึ่งแสดงถึงความสามารถของโมเดลในการเรียนรู้ความสัมพันธ์ระหว่างสินค้าและภาพเบเกอร์รี่ได้ดียิ่งขึ้น โดยผลลัพธ์การทำนายของโมเดลแสดงดังภาพที่ 38



ภาพที่ 38 การเปรียบเทียบผลลัพธ์ที่ทำนายโดยโมเดลที่ฝึกด้วยข้อมูลจากนักศึกษาและผู้เชี่ยวชาญ

## บทที่ 5

### สรุปผลการดำเนินงานวิจัยและข้อเสนอแนะ

#### 5.1. สรุปผลการดำเนินงานวิจัย

ในกระบวนการพัฒนาระบบ ได้มีการออกแบบและพัฒนาเครื่องมือสำหรับการเลเบลข้อมูล ซึ่งช่วยให้การจัดเก็บข้อมูลเป็นไปอย่างมีระเบียบและลดความซับซ้อนในกระบวนการทำงาน ระบบดังกล่าวถูกนำไปใช้งานจริงผ่านเว็บแอปพลิเคชันที่ออกแบบมาให้รองรับการเลเบลข้อมูลได้อย่างสะดวก รวดเร็ว และมีประสิทธิภาพสูง โดยข้อมูลที่ได้จากกระบวนการเลเบลถูกจัดเก็บในฐานข้อมูล MongoDB ซึ่งมีความยืดหยุ่นและเหมาะสำหรับการจัดเก็บข้อมูลที่มีโครงสร้างซับซ้อนในรูปแบบ JSON ทำให้ข้อมูลที่ได้จากผู้ใช้งานสามารถนำไปใช้ในการฝึกสอนโมเดลจำแนกประเภทแบบหลายเลเบลได้โดยตรง นอกจากนี้ ระบบยังช่วยแก้ปัญหาความซับซ้อนในการจัดการข้อมูลที่จำเป็นสำหรับการเรียนรู้ของโมเดล เช่น การจัดรูปแบบข้อมูล การตรวจสอบความถูกต้องของข้อมูลที่เลเบล และการเตรียมข้อมูลในรูปแบบที่พร้อมใช้งานในกระบวนการฝึกสอน (Training Pipeline) การออกแบบที่คำนึงถึงความสะดวกและประสิทธิภาพนี้ไม่เพียงช่วยลดภาระงานของผู้ใช้งาน แต่ยังช่วยเพิ่มคุณภาพของข้อมูลที่ใช้ในการฝึกสอนโมเดล ทำให้โมเดลสามารถเรียนรู้ได้อย่างมีประสิทธิภาพและแม่นยำยิ่งขึ้น นอกจากนี้ในกระบวนการเลเบลข้อมูลมีการเปรียบเทียบผลการเลเบลระหว่างกลุ่มผู้เลเบลที่เป็น นักศึกษา และ ผู้เชี่ยวชาญ ซึ่งเป็นส่วนสำคัญในการวิเคราะห์ผลกระทบของคุณภาพข้อมูลเลเบลที่มีต่อประสิทธิภาพของโมเดลแนะนำบรรจุภัณฑ์ โดยผลการเปรียบเทียบข้อมูลเลเบลจากกลุ่มนักศึกษาและผู้เชี่ยวชาญแสดงให้เห็นถึงความแตกต่างในแง่ของความละเอียดและความสอดคล้องของข้อมูล กลุ่มนักศึกษาเลเบลภาพโดยไม่มีการเลือกซ้ำกัน ส่งผลให้ค่าเฉลี่ยของสินค้าที่ถูกเลือกต่อภาพสูง แต่ข้อมูลจะไม่สามารถวัดความสอดคล้องระหว่างการตัดสินใจของแต่ละคนได้ในขณะที่กลุ่มผู้เชี่ยวชาญที่มีการเลเบลภาพซ้ำในบางกรณี สามารถเพิ่มความสอดคล้องในข้อมูลเลเบลได้อย่างชัดเจน โดยค่าเฉลี่ยของสินค้าที่ถูกเลือกต่อภาพต่อคนลดลง แต่มีความละเอียดและความน่าเชื่อถือที่สูงขึ้น

ในงานวิจัยนี้ได้ใช้ YOLOv8 ซึ่งเป็นโมเดลที่พัฒนาสำหรับงาน Image Segmentation ในการแยกวัตถุออกจากพื้นหลังสำหรับผลิตภัณฑ์เบเกอรี่ โดยใช้ชุดข้อมูล FoodSeg103 ที่ได้รับการออกแบบมาเพื่อการตรวจจับและแยกวัตถุอาหารโดยเฉพาะซึ่ง การฝึกสอน YOLOv8 แสดงให้เห็นว่าโมเดลสามารถแยกวัตถุอาหารออกจากพื้นหลังได้อย่างแม่นยำ แม้ในกรณีที่ภาพมีลักษณะซับซ้อน

เช่น พื้นหลังที่คล้ายกับวัตถุหรือวัตถุที่มีขนาดเล็ก โดยผลการทดสอบประสิทธิภาพโมเดล YOLOv8 แต่ละรุ่น (YOLOv8n-seg, YOLOv8s-seg, YOLOv8m-seg, YOLOv8l-seg และ YOLOv8x-seg) แสดงถึงความแตกต่างในด้าน ความแม่นยำ (Precision), ความครอบคลุม (Recall), และ ค่าเฉลี่ยความแม่นยำ (mAP) โดยแยกออกเป็นการวัดผลสำหรับ Bounding Box (B) และ Mask (M) โดยที่ YOLOv8x-seg เป็นโมเดลที่ใหญ่ที่สุด แสดงผลลัพธ์ที่ดีที่สุดในแง่ของความแม่นยำ ด้วยค่า Precision (B) ที่ 93.9% และ mAP50 (B) ที่ 93.1% สำหรับการทำนายวัตถุแบบกรอบ (Box) ในขณะที่ค่า Precision (M) สำหรับ Mask อยู่ที่ 94.6% และ mAP50 (M) สูงถึง 93.4% ซึ่งแสดงถึงความสามารถในการตรวจจับวัตถุและการแยกพื้นหลังได้อย่างแม่นยำสูง อย่างไรก็ตามการใช้ YOLOv8x-seg ในระบบจริงอาจมีข้อจำกัดในด้านความเร็วในการประมวลผล เนื่องจากโมเดลมีขนาดใหญ่และใช้เวลาในการทำนายถึง 578 มิลลิวินาทีต่อภาพ ในทางกลับกัน YOLOv8n-seg ซึ่งเป็นโมเดลขนาดเล็กที่สุด ใช้เวลาน้อยที่สุดในการทำนาย (เพียง 95 มิลลิวินาทีต่อภาพ) แต่ยังคงให้ค่า Precision (B) ที่ 90.8% และ mAP50 (B) ที่ 91.9% แม้ว่าในบางกรณีประสิทธิภาพจะต่ำกว่า YOLOv8x-seg แต่ข้อได้เปรียบในด้านความเร็วและการใช้ทรัพยากรที่น้อยทำให้ YOLOv8n-seg เป็นตัวเลือกที่เหมาะสมสำหรับระบบที่ต้องการประมวลผลภาพจำนวนมากในเวลาจำกัด

จากการทดสอบและเปรียบเทียบประสิทธิภาพของโมเดลจำแนกประเภทแบบหลายเลเวลระหว่าง VGG19, VGG16, ResNet50 และ InceptionV3 ผลการทดลองชี้ให้เห็นว่า InceptionV3 เป็นโมเดลที่มีความเหมาะสมที่สุดสำหรับการจำแนกประเภทบรรจุภัณฑ์เบเกอร์รี่ในระบบแนะนำบรรจุภัณฑ์ โดยเฉพาะเมื่อใช้กลยุทธ์การฝึกสอนแบบ Fine-Tuning (S3) ซึ่งช่วยเพิ่มความสามารถของโมเดลในการเรียนรู้จากข้อมูลได้อย่างมีประสิทธิภาพสูงสุด จากผลการประเมินแสดงให้เห็นว่า InceptionV3 ให้ค่า Precision สูงสุดถึง 97.52% และค่า F1-Score ที่ 95.04% ซึ่งดีกว่าโมเดลอื่นในทุกตัวชี้วัดดังนี้ Precision, Recall, และ mAP@0.5 ข้อมูลดังกล่าวสะท้อนให้เห็นถึงความสามารถในการจำแนกประเภทที่แม่นยำและครอบคลุม แม้ว่าโมเดลอื่น เช่น ResNet50 และ VGG จะมีประสิทธิภาพที่เหมาะสมในบางกรณี แต่เมื่อพิจารณาโดยรวมแล้ว InceptionV3 ให้ผลลัพธ์ที่ดีที่สุดและสาเหตุหนึ่งที่ทำให้ InceptionV3 ทำงานได้ดีขึ้นอย่างมีนัยสำคัญคือการเพิ่มข้อมูลผ่านกระบวนการ Data Augmentation ซึ่งเพิ่มจำนวนข้อมูลที่มีการเลเบลจาก 300 เป็น 1,000 ผลเลเบล ทำให้โมเดลสามารถเรียนรู้จากข้อมูลที่หลากหลายและลดผลกระทบจากปัญหาความไม่สมดุลของข้อมูลที่เกิดจากจำนวนตัวอย่างของบางประเภทบรรจุภัณฑ์มีจำนวนมากกว่าอีกประเภทอย่างชัดเจน เช่น กล่องขนมทรงแบนและกล่องเค้กที่มีจำนวนข้อมูลมากกว่าเมื่อเทียบกับกล่องอาหารว่าง

หรือถุงใส่คุกกี้ ซึ่งทำให้โมเดลมีแนวโน้มที่จะเรียนรู้จากประเภทที่มีตัวอย่างมากกว่า และอาจละเลยหรือให้ความสำคัญน้อยกับประเภทที่มีตัวอย่างน้อย การเพิ่มข้อมูลดังกล่าวช่วยให้โมเดลสามารถประมวลผลและจำแนกประเภทที่หลากหลายได้ดีขึ้น นอกจากนี้ การอนุญาตให้ปรับแต่งเลเยอร์ผ่านกระบวนการ Fine-Tuning ในกลยุทธ์ S3 ยังมีบทบาทสำคัญในการเพิ่มประสิทธิภาพของ InceptionV3 โดยช่วยให้โมเดลสามารถปรับการเรียนรู้ให้เข้ากับลักษณะเฉพาะของข้อมูลเบเกอร์รี่ได้ดียิ่งขึ้น เมื่อเปรียบเทียบกับกลยุทธ์ S1 และ S2 ที่ไม่ได้เปิดโอกาสให้ปรับแต่งเลเยอร์ การทำ Fine-Tuning ช่วยเพิ่มความแม่นยำในการจำแนกประเภทอย่างมีนัยสำคัญ ด้วยผลการทดลองที่ได้ สรุปได้ว่า InceptionV3 เป็นตัวเลือกที่เหมาะสมที่สุดสำหรับระบบแนะนำบรรจุภัณฑ์ในงานวิจัยนี้ เนื่องจากสามารถรองรับการจัดการข้อมูลที่หลากหลายและซับซ้อนได้อย่างแม่นยำ การใช้โมเดลนี้ช่วยให้ระบบสามารถทำงานได้อย่างมีประสิทธิภาพและเชื่อถือได้ ซึ่งเป็นปัจจัยสำคัญสำหรับการประยุกต์ใช้งานในสภาพแวดล้อมจริง

เนื่องจากระบบต้องรองรับการเพิ่มข้อมูลสินค้าที่เข้ามาใหม่ในทุกๆ วัน ซึ่งการใช้ โมเดลจำแนกประเภทแบบหลายเลเบลไม่สามารถรองรับการอัปเดตข้อมูลแบบเรียลไทม์ได้ เนื่องจากต้องผ่านกระบวนการเลเบลข้อมูลและฝึกสอนโมเดลใหม่ทุกครั้งที่มีข้อมูลใหม่เข้ามา ซึ่งเป็นกระบวนการที่ใช้ทรัพยากรมากและไม่เหมาะสมสำหรับระบบที่ต้องการความเร็วและเป็นปัจจุบัน การใช้ Cosine Similarity ช่วยเพิ่มความยืดหยุ่นให้กับระบบ โดยการเปรียบเทียบคุณลักษณะของสินค้าใหม่กับสินค้าที่มีอยู่ในฐานข้อมูลในรูปแบบเวกเตอร์ วิธีนี้ช่วยให้สามารถแนะนำบรรจุภัณฑ์ที่เหมาะสมได้โดยไม่ต้องฝึกสอนโมเดลใหม่ ซึ่งเป็นการลดขั้นตอนและเพิ่มประสิทธิภาพของระบบ นอกจากนี้ การใช้ Cosine Similarity ยังเพิ่มโอกาสให้สินค้าใหม่ๆ ที่มีคุณลักษณะใกล้เคียงกับสินค้าที่มีอยู่แล้วในระบบ ได้รับการแนะนำไปยังลูกค้าได้อย่างรวดเร็ว เนื่องจากสินค้าในระบบมักมีความคล้ายคลึงกันในแง่ของขนาดหรือประเภท การประมวลผลด้วยวิธีนี้จึงเหมาะสมสำหรับการจัดการข้อมูลที่เปลี่ยนแปลงบ่อย และช่วยให้ระบบสามารถตอบสนองความต้องการของลูกค้าได้

## 5.2. ข้อเสนอแนะจากการวิจัย

จากการดำเนินงานวิจัยและผลการทดลองที่ได้ ระบบที่พัฒนามีความสามารถในการแนะนำบรรจุภัณฑ์ที่เหมาะสมสำหรับผลิตภัณฑ์เบเกอร์รี่ และรองรับการเพิ่มข้อมูลสินค้าใหม่ในระบบได้อย่างมีประสิทธิภาพ อย่างไรก็ตาม ยังมีประเด็นที่ควรพิจารณาเพื่อปรับปรุงและพัฒนาต่อไปในอนาคต

หนึ่งในแนวทางการพัฒนาที่สำคัญคือการเพิ่มกระบวนการจัดเก็บข้อมูล Log เพื่อเก็บพฤติกรรมของลูกค้าเกี่ยวกับการตอบสนองต่อสินค้าที่ระบบแนะนำ เช่น การคลิกดูสินค้า หรือการซื้อสินค้า การบันทึกพฤติกรรมเหล่านี้ช่วยให้สามารถรวบรวมข้อมูลเชิงลึกเกี่ยวกับการเลือกสินค้าของลูกค้าได้ โดยข้อมูล Log ดังกล่าวจะถูกนำมาใช้เพื่อฝึกสอนโมเดลในรอบถัดไป ซึ่งจะช่วยให้โมเดลสามารถเรียนรู้และปรับตัวให้เข้ากับพฤติกรรมของลูกค้าได้อย่างต่อเนื่อง โดยเสนอรูปแบบการบันทึกข้อมูล Log ซึ่งแบ่งออกเป็น 2 รูปแบบหลักที่สำคัญ เพื่อใช้ในการปรับปรุงและพัฒนาระบบแนะนำบรรจุกู้ในอนาคต ดังนี้ข้อมูลสำหรับบันทึกสินค้าที่ระบบแนะนำให้ลูกค้าและข้อมูลพฤติกรรมของลูกค้าหลังจากได้รับคำแนะนำ โดยงานวิจัยนี้ควรพิจารณาการใช้เทคนิคการเรียนรู้ที่ทันสมัย เช่น Semi-supervised Learning [43] หรือ Self-supervised Learning [44] เพื่อช่วยให้โมเดลสามารถเรียนรู้ได้จากข้อมูลที่ถูก label เพียงบางส่วน โดยการบันทึกข้อมูล Log ที่เกี่ยวกับพฤติกรรมของลูกค้า เช่น การคลิกดูสินค้า การซื้อสินค้า หรือการปฏิเสธสินค้า สามารถนำมาประยุกต์ใช้เป็นข้อมูลแบบ Unlabeled Data ที่ช่วยเพิ่มประสิทธิภาพในการฝึกสอนโมเดล นอกจากนี้การประเมินประสิทธิภาพของโมเดลในสถานการณ์จริงยังเป็นสิ่งสำคัญ การบันทึกข้อมูล Log เช่นนี้ไม่เพียงช่วยให้สามารถปรับปรุงโมเดลให้มีความแม่นยำมากขึ้น แต่ยังช่วยให้ระบบสามารถปรับตัวได้โดยอัตโนมัติในสภาพแวดล้อมที่เปลี่ยนแปลงอยู่

ดังนั้นการประมวลผลในงานวิจัยนี้จึงถือเป็นการวางรากฐานที่สำคัญสำหรับการพัฒนาระบบแนะนำสินค้า และการเพิ่มกระบวนการบันทึก Log จะช่วยสร้างแนวทางในการปรับปรุงระบบในอนาคต ทำให้สามารถใช้งานได้อย่างกว้างขวาง ยืดหยุ่น และตอบสนองความต้องการของผู้ใช้งานได้อย่างมีประสิทธิภาพมากยิ่งขึ้น

## รายการอ้างอิง

- [1] H. Xin, "Research on E-Commerce Business Model Based on Intelligent Terminals Mobile," presented at the 2015 7th International Conference on Information Technology in Medicine and Education (ITME), 2015.
- [2] W.-C. Kang, E. Kim, J. Leskovec, C. Rosenberg, and J. McAuley, "Complete the Look: Scene-Based Complementary Product Recommendation," presented at the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019.
- [3] I. D. a. A. Team. "AI vs. machine learning vs. deep learning vs. neural networks: What's the difference?" IBM® watsonx.ai™. <https://www.ibm.com/topics/neural-networks> (accessed 30/11, 2024).
- [4] O. Yalçın, "Convolutional Neural Networks," 2021, pp. 145-160.
- [5] P. Saranya and R. Durga, "Food Safety Control Using CNN Model in Image Processing Technique," in *2023 International Conference on New Frontiers in Communication, Automation, Management and Security (ICCAMS)*, 27-28 Oct. 2023 2023, vol. 1, pp. 1-8, doi: 10.1109/ICCAMS60113.2023.10526092.
- [6] J. Zou, Z. Liu, S. Jiang, X. Li, Y. Dong, and C. Wang, "Research on Object Recognition Detection Based on Image Processing and Machine Learning," in *2024 IEEE 2nd International Conference on Control, Electronics and Computer Technology (ICCECT)*, 26-28 April 2024 2024, pp. 896-902, doi: 10.1109/ICCECT60629.2024.10546059.
- [7] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 27-30 June 2016 2016, pp. 779-788, doi: 10.1109/CVPR.2016.91.
- [8] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," presented at the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016. [Online]. Available: <https://doi.ieeecomputersociety.org/10.1109/CVPR.2016.91>.

- [9] YOLO by Ultralytics. (2023). [Online]. Available: <https://github.com/ultralytics/ultralytics>
- [10] J. Redmon and A. Farhadi, "YOLO9000: Better, Faster, Stronger," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 21-26 July 2017 2017, pp. 6517-6525, doi: 10.1109/CVPR.2017.690.
- [11] J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement," 04/08 2018, doi: 10.48550/arXiv.1804.02767.
- [12] A. Bochkovskiy, C.-Y. Wang, and H.-y. Liao, *YOLOv4: Optimal Speed and Accuracy of Object Detection*. 2020.
- [13] M. S, M. S, K. T, and S. P, "Image Detection and Segmentation using YOLO v5 for surveillance," *Applied and Computational Engineering*, vol. 8, pp. 160-165, 08/01 2023, doi: 10.54254/2755-2721/8/20230109.
- [14] C. Li *et al.*, "YOLOv6: A Single-Stage Object Detection Framework for Industrial Applications," *arXiv [cs.CV]*, 2022 2022. [Online]. Available: <http://arxiv.org/abs/2209.02976>.
- [15] C. Y. Wang, A. Bochkovskiy, and H. Y. M. Liao, "YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors," in *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 17-24 June 2023 2023, pp. 7464-7475, doi: 10.1109/CVPR52729.2023.00721.
- [16] M. M. Tsung-Yi Lin, Serge Belongie, Lubomir Bourdev, Ross Girshick, James Hays, Pietro Perona, Deva Ramanan, C. Lawrence Zitnick, Piotr Dollár, D. a. P. Fleet, Tomas and Schiele, Bernt and Tuytelaars, Tinne, Ed. *Microsoft COCO: Common Objects in Context* (Computer Vision – ECCV 2014). Springer International Publishing, 2014.
- [17] R. Varghese and M. S, "YOLOv8: A Novel Object Detection Algorithm with Enhanced Performance and Robustness," in *2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS)*, 18-19 April 2024 2024, pp. 1-6, doi: 10.1109/ADICS58448.2024.10533619.
- [18] C.-Y. Wang, I. H. Yeh, and H.-Y. M. Liao, "YOLOv9: Learning What You Want to Learn Using Programmable Gradient Information," *arXiv [cs.CV]*, 2024 2024. [Online]. Available: <http://arxiv.org/abs/2402.13616>.

- [19] A. Wang et al., *YOLOv10: Real-Time End-to-End Object Detection*. 2024.
- [20] A. Neubeck and L. V. Gool, "Efficient Non-Maximum Suppression," in *18th International Conference on Pattern Recognition (ICPR'06)*, 20-24 Aug. 2006 2006, vol. 3, pp. 850-855, doi: 10.1109/ICPR.2006.479.
- [21] *Ultralytics YOLO11*. (2024). [Online]. Available: <https://github.com/ultralytics/ultralytics>
- [22] X. Cao, Y. Su, X. Geng, and Y. Wang, "YOLO-SF: YOLO for Fire Segmentation Detection," *IEEE Access*, vol. 11, pp. 111079-111092, 2023, doi: 10.1109/ACCESS.2023.3322143.
- [23] M. Jindal, N. Raj, P. Saranya, and V. S, "Aircraft Detection from Remote Sensing Images using YOLOV5 Architecture," in *2022 6th International Conference on Devices, Circuits and Systems (ICDCS)*, 21-22 April 2022 2022, pp. 332-336, doi: 10.1109/ICDCS54290.2022.9780777.
- [24] F. Rahutomo, T. Kitasuka, and M. Aritsugi, *Semantic Cosine Similarity*. 2012.
- [25] A. Mikołajczyk and M. Grochowski, "Data augmentation for improving deep learning in image classification problem," in *2018 International Interdisciplinary PhD Workshop (IIPHDW)*, 9-12 May 2018 2018, pp. 117-122, doi: 10.1109/IIPHDW.2018.8388338.
- [26] X. Li, S. Deng, S. Wang, Z. Lv, and L. Wu, "Review of Small Data Learning Methods," in *2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC)*, 23-27 July 2018 2018, vol. 02, pp. 106-109, doi: 10.1109/COMPSAC.2018.10212.
- [27] S. Bozinovski and A. Fulgosi, "The influence of pattern similarity and transfer learning upon training of a base perceptron b2," in *Proceedings of symposium informatica*, 1976, vol. 3, pp. 121-126.
- [28] A. Takyar. "Accelerating AI model training with transfer learning." LeewayHertz A HACKETT GROUP COMPANY. <https://www.leewayhertz.com/transfer-learning/> (accessed 30/11, 2024).
- [29] D. Merkel, "Docker: Lightweight Linux Containers for Consistent Development and Deployment," *Linux J*, vol. 2014, no. 239, 2014. [Online]. Available: <http://dl.acm.org/citation.cfm?id=2600239.2600241>.

- [30] K. Pareek. "Architecture of Docker." GeeksforGeeks, Sanchhaya Education Private Limited. <https://www.geeksforgeeks.org/architecture-of-docker/> (accessed 30/11, 2024).
- [31] B. Kumar, A. K. Singh, and P. Banerjee, "A Deep Learning Approach for Product Recommendation using ResNet-50 CNN Model," in *2023 International Conference on Sustainable Computing and Smart Systems (ICSCSS)*, 14-16 June 2023 2023, pp. 604-610, doi: 10.1109/ICSCSS57650.2023.10169441.
- [32] S. O. Mohammadi, H. Bodaghi, and A. Kalhor, "Single-Item Fashion Recommender: Towards Cross-Domain Recommendations," in *2022 30th International Conference on Electrical Engineering (ICEE)*, 17-19 May 2022 2022, pp. 12-16, doi: 10.1109/ICEE55646.2022.9827421.
- [33] M. H. Kiapour, X. Han, S. Lazebnik, A. C. Berg, and T. L. Berg, "Where to Buy It: Matching Street Clothing Photos in Online Shops," in *2015 IEEE International Conference on Computer Vision (ICCV)*, 7-13 Dec. 2015 2015, pp. 3343-3351, doi: 10.1109/ICCV.2015.382.
- [34] A. Nayak, J. Shah, A. Kuruvilla, J. A, and J. S. B, "Fine-grained Fashion Clothing Image Classification and Recommendation," in *2021 2nd International Conference on Electronics, Communications and Information Technology (CECIT)*, 27-29 Dec. 2021 2021, pp. 600-606, doi: 10.1109/CECIT53797.2021.00111.
- [35] W. C. Kang, E. Kim, J. Leskovec, C. Rosenberg, and J. McAuley, "Complete the Look: Scene-Based Complementary Product Recommendation," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 15-20 June 2019 2019, pp. 10524-10533, doi: 10.1109/CVPR.2019.01078.
- [36] M. Fadhillah, J. Y. Lin, W. J. Chen, and G. S. Lin, "Multi-View Clothing Image Searching Based on Deep Neural Network," in *2020 IEEE International Conference on Consumer Electronics - Taiwan (ICCE-Taiwan)*, 28-30 Sept. 2020 2020, pp. 1-2, doi: 10.1109/ICCE-Taiwan49838.2020.9258235.
- [37] C. W. Chang and C. M. Lo, "Content-based Commodity Search Engine Using Convolutional Image Retrieval," in *2022 2nd International Conference on Social Sciences and Intelligence Management (SSIM)*, 24-26 Nov. 2022 2022, pp. 26-28,

doi: 10.1109/SSIM55504.2022.10047930.

- [38] K. T. Islam, S. Wijewickrema, M. Pervez, and S. O. Leary, "An Exploration of Deep Transfer Learning for Food Image Classification," in *2018 Digital Image Computing: Techniques and Applications (DICTA)*, 10-13 Dec. 2018 2018, pp. 1-5, doi: 10.1109/DICTA.2018.8615812.
- [39] G. Sharma, K. Abraham, and V. B, "Transfer Learning Approach for Multilabel Fine Grain Image Attribute Extraction," in *2022 International Conference on Computational Science and Computational Intelligence (CSCI)*, 14-16 Dec. 2022 2022, pp. 1521-1524, doi: 10.1109/CSCI58124.2022.00269.
- [40] X. a. F. Wu, Xin and Liu, Ying and Lim, Ee-Peng and Hoi, Steven and Sun, Qianru, "A Large-Scale Benchmark for Food Image Segmentation," presented at the Proceedings of the 29th ACM International Conference on Multimedia, 2021, 10.
- [41] G. Bradski, "The openCV library," *Dr. Dobb's Journal: Software Tools for the Professional Programmer*, vol. 25, no. 11, pp. 120-123, 2000.
- [42] *Rasterio: geospatial raster I/O for {Python} programmers*. (2013--). Accessed: 8 August 2023. [Online]. Available: <https://github.com/rasterio/rasterio>
- [43] X. Yang, Z. Song, I. King, and Z. Xu, "A Survey on Deep Semi-Supervised Learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 9, pp. 8934-8954, 2023, doi: 10.1109/TKDE.2022.3220219.
- [44] J. Gui *et al.*, "A Survey on Self-Supervised Learning: Algorithms, Applications, and Future Trends," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 46, no. 12, pp. 9052-9071, 2024, doi: 10.1109/TPAMI.2024.3415112.



## ประวัติผู้เขียน

ชื่อ-สกุล

อ้นตามัน นพนภาพร

วุฒิการศึกษา

วิทยาศาสตร์บัณฑิต (วท.บ.) วิทยาการคอมพิวเตอร์

