

แอปพลิชันรวมสินค้าไลฟ์สดด้วยการประมวลผลทางภาพ

application live shop with image processing

โครงงานปริญญานิพนธ์

ของ

นายกฤตเมธ บัวสิงห์

นายกิตติศักดิ์ มนพรหมมา

เสนอต่อมหาวิทยาลัยมหาสารคาม เพื่อเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

ปริญญาวิทยาศาสตรบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์

ปีการศึกษา 2565

ลิขสิทธิ์เป็นของมหาวิทยาลัยมหาสารคาม

คณะวิทยาการสารสนเทศ มหาวิทยาลัยมหาสารคาม

แอปพลิชันรวมสินค้าไลฟ์สดด้วยการประมวลผลทางภาพ

application live shop with image processing

โครงงานปริญญานิพนธ์

ของ

นายกฤตเมธ บัวสิงห์

นายกิตติศักดิ์ มนพรหมมา

เสนอต่อมหาวิทยาลัยมหาสารคาม เพื่อเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

ปริญญาวิทยาศาสตรบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์

ปีการศึกษา 2565

ลิขสิทธิ์เป็นของมหาวิทยาลัยมหาสารคาม

คณะวิทยาการสารสนเทศ มหาวิทยาลัยมหาสารคาม

รูปภาพประกอบด้วย ข้อความ, โคมไฟ

คำอธิบายที่สร้างโดยอัตโนมัติ

คณะกรรมการสอบโครงงานปริญญานิพนธ์ ได้พิจารณาปริญญานิพนธ์ของ นายกฤตเมธ บัวสิงห์ และนายกิตติศักดิ์ มนพรมมา  แล้วเห็นสมควรรับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาการสารสนเทศ ของมหาวิทยาลัยมหาสารคาม

      คณะกรรมการสอบโครงงานปริญญานิพนธ์

|  |  |
| --- | --- |
| **………………………………………..……**  **(…………………………………….)** | **ประธานสอบ** |
| **………………………………………..……**  **(……………………………………..)** | **กรรมการ** |
| รูปภาพประกอบด้วย ไม้แขวนเสื้อ  คำอธิบายที่สร้างโดยอัตโนมัติ  **………………………………………..……**  **(ดร.พัฒนพงษ์ ชมพูวิเศษ)** | **ที่ปรึกษาโครงงานปริญญานิพนธ์หลัก** |

หลักสูตรวิทยาการคอมพิวเตอร์อนุมัติให้รับโครงงานปริญญานิพนธ์ฉบับนี้  เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรบัณฑิต  สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาการสารสนเทศ มหาวิทยาลัยมหาสารคาม

|  |  |
| --- | --- |
| **…………………………….……………….…………**  **(อาจารย์พชระ  พฤกษะศรี)** | **…………………………….……………….…………**  **(ผู้ช่วยศาสตราจารย์พิมลรัตน์  อ้วนศรีเมือง)** |
| **อาจารย์ผู้ประสานงานวิชาโครงงานปริญญานิพนธ์**  **วันที่  เดือน  พ.ศ.** | |

สารบัญ

**หน้า**

[สารบัญ ก](#_Toc113308717)

[สารบัญตาราง ค](#_Toc113308718)

[สารบัญรูปภาพ ง](#_Toc113308719)

[บทที่ 1 บทนำ 1](#_Toc113308720)

[1.1 หลักการและเหตุผล 1](#_Toc113308721)

[1.2 วัตถุประสงค์ของโครงงาน 1](#_Toc113308722)

[1.3 ขอบเขตของโครงงาน 1](#_Toc113308723)

[1.3.1 ผู้ใช้งาน (ใช้งานผ่าน Mobile Application) 2](#_Toc113308724)

[1.3.2 ผู้ดูแลระบบ (ใช้งานผ่าน Desktop Application ที่พัฒนาด้วย Python) 3](#_Toc113308725)

[1.3.3 ระบบ (ระบบพัฒนาเป็น Module ด้วย Python) 3](#_Toc113308726)

[1.4 ภาพรวมของระบบ 5](#_Toc113308727)

[จากภาพรวมระบบจะการทำงานดังนี้ 5](#_Toc113308728)

[1.5 ประโยชน์ที่คาดว่าจะได้รับ 6](#_Toc113308729)

[1.5.1 ช่วยให้ผู้ใช้งานที่เป็นผู้ค้าขายมีช่องทางการขายสินค้ามากขึ้น 6](#_Toc113308730)

[1.5.2 ช่วยให้ผู้ใช้งานที่เป็นลูกค้ามีความสะดวกมากขึ้น 6](#_Toc113308731)

[1.6 อุปกรณ์และเครื่องมือที่ใช้ในการดำเนินงาน 6](#_Toc113308732)

[1.6.1 ฮาร์ดแวร์ 6](#_Toc113308733)

[1.6.2 ซอฟต์แวร์ 7](#_Toc113308734)

[1.7 แผนการดำเนินงาน 7](#_Toc113308735)

[1.8 ตัวอย่างโปรแกรม 8](#_Toc113308736)

[1.8.1 แอปพลิเคชั่นบนอุปกรณ์ไร้สายแบบเคลื่อนที่ 8](#_Toc113308737)

[1.8.2 แอปพลิเคชั่นบนคอมพิวเตอร์สำหรับผู้ดูแลระบบ 13](#_Toc113308738)

[บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง 16](#_Toc113308739)

[2.1 ทฤษฎีที่เกี่ยวข้อง 16](#_Toc113308740)

[2.1.1 Flutter 16](#_Toc113308741)

[2.1.2 Firebase 17](#_Toc113308742)

สารบัญ(ต่อ)

**หน้า**

[2.1.3 การประมวลผลภาพ (Image Processing) 19](#_Toc113308743)

[2.1.4 OCR หรือ Optical Character Recognition 21](#_Toc113308744)

[2.1.1 Object Detection 22](#_Toc113308745)

[2.1.1 YOLO (You Only Look Once) 22](#_Toc113308746)

[2.2 ระบบงานที่เกี่ยวข้อง 22](#_Toc113308747)

[2.2.1 Safety Helmet Detection Based on YOLOv5 22](#_Toc113308748)

[อ้างอิง 26](#_Toc113308749)

สารบัญตาราง

**หน้า**

[**ตารางที่ 1** แผนการดำเนินงาน 7](#_Toc112973879)

สารบัญรูปภาพ

**หน้า**

[ภาพประกอบที่ 1.1 ตัวอย่างข้อมูลที่นำมาแยก 4](#_Toc113109244)

[ภาพประกอบที่ 1.2 ภาพรวมของระบบ 5](#_Toc113109245)

[ภาพประกอบที่ 1.3 หน้าแรกของโปรแกรม 8](#_Toc113109246)

[ภาพประกอบที่ 1.4 หน้าล็อกอินเข้าสู่ระบบ 9](#_Toc113109247)

[ภาพประกอบที่ 1.5 หน้าหลักของแอปพลิเคชั่น 10](#_Toc113109248)

[ภาพประกอบที่ 1.6 หน้าค้นหาสินค้า 11](#_Toc113109249)

[ภาพประกอบที่ 1.7 หน้ารายละเอียดสินค้า 12](#_Toc113109250)

[ภาพประกอบที่ 1.8 หน้าแรกโปรแกรม 13](file:///C:\Users\ROG%20STRIX\Downloads\project03092.docx#_Toc113109251)

[ภาพประกอบที่ 1.9 หน้าเลือกพื้นที่ภาพหน้าจอ 14](file:///C:\Users\ROG%20STRIX\Downloads\project03092.docx#_Toc113109252)

[ภาพประกอบที่ 1.10 หน้าผลลัพธ์ในการถ่ายภาพหน้าจอ 14](#_Toc113109253)

[ภาพประกอบที่ 1.11 หน้ารายระเอียดข้อมูล 15](#_Toc113109254)

[ภาพประกอบที่ 2.1 ตัวอย่างโครงสร้าง Flutter 16](#_Toc113109255)

[ภาพประกอบที่ 2.2 การตรวจจับและประมวลผลใบหน้าจากภาพ [4] 20](#_Toc113109256)

[ภาพประกอบที่ 2.3 ขบวนการการทำงานของ Optical Character Recognition 21](#_Toc113109257)

[ภาพประกอบที่ 2.4 ภาพการตรวจจับหมวกนิรภัยของคนงาน [8] 23](file:///C:\Users\ROG%20STRIX\Downloads\project03092.docx#_Toc113109258)

[ภาพประกอบที่ 2.5 การแบ่งช่องของรูปภาพออกเป็น grid [9] 23](file:///C:\Users\ROG%20STRIX\Downloads\project03092.docx#_Toc113109259)

[ภาพประกอบที่ 2.6 ภาพของเทคนิค Anchor Box [9] 24](file:///C:\Users\ROG%20STRIX\Downloads\project03092.docx#_Toc113109260)

[ภาพประกอบที่ 2.7 การทำงานของขบวนการ IOU [9] 25](file:///C:\Users\ROG%20STRIX\Downloads\project03092.docx#_Toc113109261)

# บทนำ

## หลักการและเหตุผล

ปัจจุบันเทคโนโลยีอิเล็กทรอนิกส์ มีความสําคัญในการดํารงชีวิตของผู้คนในปัจจุบันมากขึ้น ผู้คนมากมายได้ทำการใช้เทคโนโลยีในการค้าขายของออนไลน์ผ่านการไลฟ์และได้เกิดการพัฒนาของ **ปัญญาประดิษฐ์  (AI : Artificial Intelligence)**  คือเครื่องจักรหรือเทคโนโลยีที่มีฟังก์ชันทีมีความสามารถในการทำความเข้าใจ เรียนรู้องค์ความรู้ต่างๆ อาทิเช่น การรับรู้ การเรียนรู้ การให้เหตุผล และการแก้ปัญหาต่างๆ เครื่องมือที่มีความสามารถเหล่านี้ก็ถือว่าเป็น ปัญญาประดิษฐ์ เพราะฉะนั้นจึงสามารถกล่าวได้ว่า AI ถือกำเนิดขึ้นเมื่อเครื่องจักรมีความสามารถที่จะเรียนรู้นั่นเอง

ปัจจุบันการที่จะขายพัสดุออนไลน์ในการไลฟ์สดนั้น ใช้เวลาที่ค่อนข้างมากที่จะต้องให้ลูกค้าเข้ามารับชมเพื่อเลือกสิ้นค้าทีละชิ้นและทำให้ลูกค้าบางรายที่ไม่สะดวกมารับชมไม่ได้เลือกซื้อสินค้าที่ต้องการจึงทำให้พวกเราได้ทำการพัฒนาแอปพลิชั่นรวมสินค้าไลฟ์สดขึ้นมาเพื่อแก้ไขปัญหาดังกล่าว

ดังนั้นเพื่อให้เกิดความสะดวกสบายพวกเราจึงได้พัฒนาแอปพลิเคชั่นเพื่อตอบโจทย์การใช้งานของแม่ค้าและลูกค้า ก็คือแอปพลิชั่นรวมสินค้าไลฟ์สดซึ่งจะทำการรวมเอาสินค้าต่างๆในไลฟ์สดของแม่ค้าแต่ละคนมารวมไว้ในแอปพลิเคชั่นเดียว โดยการใช้ image processing ในการตรวจจับข้อมูลสินค้าในไลฟ์สด และนำข้อมูลมาเก็บในฐานข้อมูลและนำไปแสดงผลในแอปพลิเคชั่น และเมื่อเลือกสินค้าที่ต้องการจะสามารถดูข้อมูลของสินค้าและไปสู่หน้าไลฟ์สดของแม่ค้าได้ เพื่อให้ลูกค้าได้เลือกสินค้ากันอย่างสะดวกและแม่ค้าก็ไม่เสียลูกค้าที่ไม่สะดวกรับชม

## วัตถุประสงค์ของโครงงาน

สร้างแอปพลิชั่นที่สามารถใช้ AI ในการตรวจสอบและดึงข้อมูลของสินค้าในไลฟ์สดขายของเข้ามาในแอปพลิชั่นได้

## ขอบเขตของโครงงาน

โครงงานแอปพลิชั่นรวมสินค้าไลฟ์สดทำงานบนระบบปฏิบัติการ android แบ่งการทํางานของระบบออกเป็น 3 ส่วน ประกอบไปด้วย

### ผู้ใช้งาน (ใช้งานผ่าน Mobile Application)

1. ผู้ใช้งานสามารถเข้าสู่ระบบผ่าน Facebook ได้
2. ผู้ใช้งานสามารถดูสถานะการไลฟ์สดของร้านแต่ละร้านค้าได้ ประกอบด้วย

- สถานะ online / offline

- รูปภาพในการไลฟ์สด ถ้าสถานะ online รูปจะเป็นสีปกติ และถ้าเป็นสถานะ offline รูปจะเป็นรูปภาพขาวดำ

1. ผู้ใช้งานสามารถดูรายการของสินค้าที่แสดงสินค้าของทุกร้านที่มีสถานนะ online ทุกร้าน ใน Mobile Application ได้ โดยรายการที่ปรากฏในไลฟ์สด ประกอบไปด้วย

- ชื่อร้านค้า

- รหัสสินค้า

- ชื่อสินค้า

- ราคาสินค้า

- วันเวลาที่ capture รูปภาพ

ตัวอย่าง เช่น ปีรันย่า บิวตี้ ช๊อป คือชื่อร้าน ,น39 คือรหัสสินค้า, ดีท๊อกญี่ปุ่น คือชื่อสินค้า ,39 คือราคาสินค้า,16/07/2022 16:30 คือ วันเวลาที่ capture รูปภาพ

1. ผู้ใช้งานสามารถเลือกที่สินค้าเพื่อดูรายละเอียดเพิ่มเติมได้โดยรายการที่ปรากฏในไลฟ์สด*ประกอบด้วย*

- ชื่อร้านค้า

- รหัสสินค้า

- ชื่อสินค้า

- ราคาสินค้า

- รูปภาพในการไลฟ์สด

- วันเวลาที่ capture รูปภาพ

ตัวอย่าง เช่น Aoom shop live v2 คือชื่อร้าน ,อ7 คือรหัสสินค้า, สครับ ขมิ้น คือชื่อสินค้า ,59 คือราคาสินค้า,14/06/2022 12:30 คือ วันเวลาที่ capture รูปภาพ

1. เมื่อผู้ใช้เลือกสินค้าผู้ใช้งานสามารถกดปุ่ม “chat”ในหน้ารายละเอียดสินค้าเพื่อลิ้งไปยังแชท ของร้านค้าได้
2. เมื่อผู้ใช้เลือกสินค้าผู้ใช้งานสามารถกดปุ่ม “buy”ในหน้ารายละเอียดสินค้าเพื่อลิ้งไปยังแชท ของร้านค้าและจะมีการนำรหัสสินค้า ชื่อสินค้า ราคาสินค้า ส่งไปให้ร้านค้าอัตโนมัติ
3. ผู้ใช้งานสามารถค้นหาสินค้าจากชื่อสินค้าหรือชื่อร้านค้าได้ ในช่วงที่เวลาร้านค้าไลฟ์สดตัวอย่างเช่น

* ชื่อสินค้า
* ชื่อร้านค้า

ตัวอย่าง เช่น ลิป KO(เบอร์ 7) คือชื่อสินค้า, Aoom shop live v2 คือชื่อร้าน

1. เมื่อร้านค้ามีสถานะ offline ผู้ใช้จะไม่สามารถมองเห็นสินค้าของร้านค้าที่ offline ได้

### ผู้ดูแลระบบ (ใช้งานผ่าน Desktop Application ที่พัฒนาด้วย Python)

1. ผู้ดูแลระบบสามารถเลือกร้านค้าที่ไลฟ์สดได้ 3 ร้านค้ามีร้านค้า

- ของฝากใต้ขายส่งราคาถูก

- ร้านปีรันย่า บิวตี้ ช๊อป

- Aoom shop live v2

1. ผู้ดูแลระบบสามารถหยุดร้านค้าที่กำลังประมวลผลในระบบและแจ้งเตือนไปยัง Application ของผู้ใช้งานได้
2. ผู้ดูแลระบบสามารถดูข้อมูลรายการสินค้า ของแต่ละร้านค้าได้ ประกอบด้วย

- รหัสสินค้า

- ชื่อสินค้า

- ราคาสินค้า

ตัวอย่าง เช่น อ61 คือรหัสสินค้า, ลิป KO(เบอร์ 8) คือชื่อสินค้า,59 คือราคาสินค้า

### ระบบ (ระบบพัฒนาเป็น Module ด้วย Python)

1. ระบบสามารถจับภาพหน้าจอไลฟ์สดได้ 3 ร้านค้ามีร้านค้า

#### - FIRST SHOP V2

- KANYA SHOP ขายถูกทุกอย่าง

#### - มหัศจรรย์"วันของAuuM

1. ระบบสามารถส่งสตรีมมิ่งภาพหน้าจอไลฟ์สดเพื่อนำไปประมวลผลและนำไปแสดงในแต่พื้นที่แสดงผลแต่ละพื้นที่ใน Desktop Application ***ของผู้ดูแลระบบ*** ได้ 3 ร้านค้ามีร้านค้า

#### - ของฝากใต้ขายส่งราคาถูก

#### - ร้านปีรันย่า บิวตี้ ช๊อป

- Aoom shop live v2

1. ระบบสามารถตรวจจับกรอบของตัวอักษร รหัสสินค้า ชื่อสินค้า ราคาสินค้า ที่เป็นรายการรวมสินค้าจากไลฟ์สดได้โดยที่ขนาดไม่เล็กเกินไป
2. ระบบสามารถตัดบรรทัดของข้อมูลในกรอบที่ตรวจจับมาได้
3. ระบบลบข้อมูลสินค้าของร้านค้าที่มีสถานะ offline ได้
4. ระบบสามารถแยกข้อมูลที่แปลงจากภาพเป็นตัวอักษรได้โดยรายการที่ปรากฏในไลฟ์สด *ประกอบด้วย*

- รหัสสินค้า

- ชื่อสินค้า

- ราคาสินค้า

ตัวอย่าง เช่น ป29 คือรหัสสินค้า, ดีท็อกญี่ปุ่นคือชื่อสินค้า,39 คือราคาสินค้า



ภาพประกอบที่ 1.1 ตัวอย่างข้อมูลที่นำมาแยก

1. ระบบสามารถนำข้อมูลไปจัดเก็บในฐานข้อมูลได้โดยรายการที่ปรากฏในไลฟ์สด*ประกอบด้วย*

- ชื่อร้านค้า

- รหัสสินค้า

- ชื่อสินค้า

- ราคาสินค้า

- รูปภาพในการไลฟ์สด

- วันเวลาที่ capture รูปภาพ

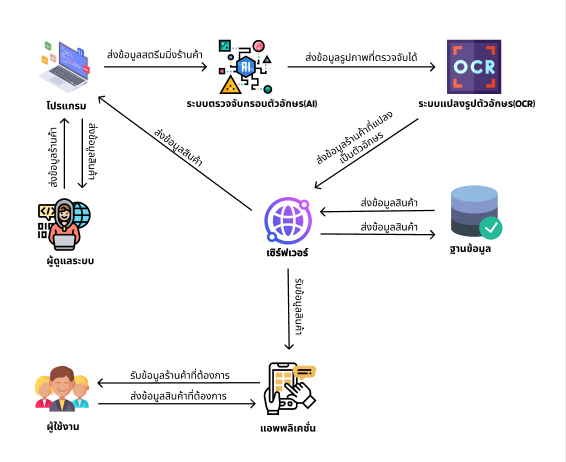
ตัวอย่าง เช่น ของฝากใต้ขายส่งราคาถูก คือชื่อร้าน,มม3 คือรหัสสินค้า,เม็ดมะม่วงหัก คือชื่อสินค้า,89บ คือราคาสินค้า,14/06/2022 14:30 คือ วันเวลาที่ capture รูปภาพ

1. ระบบจะทำงานแบบ real-time เป็น Database ของ Firebase
2. การวัดประสิทธิภาพประเมินจากความถูกผิดของการตรวจจับกรอบของข้อความและความถูกผิดของการทำ OCR(Optical Character Recognition) โดยวัดด้วย CER

(character error rate)

1. ระบบสามารถแจ้งเตือนผู้ใช้งานได้ เมื่อมีร้านค้าที่มีสถานะ online
2. ระบบสามารถตรวจจับข้อความของไลฟ์สดที่จบแล้ว และหยุดการประมวลผลของร้านค้าที่จบไลฟ์และแจ้งเตือนไปยัง Mobile Application ผุ้ใช้งานได้
3. ระบบไม่สามารถตรวจจับไลฟ์สดที่ไม่มีกรอบข้อความได้
4. ระบบไม่สามรถตรวจจับกรอบตัวอักษรและตัวอักษรที่มีสีเหมือนกันได้

## ภาพรวมของระบบ



ภาพประกอบที่ 1.2 ภาพรวมของระบบ

### จากภาพรวมระบบจะการทำงานดังนี้

1. ผู้ดูแลระบบต้องทำการเลือกร้านค้าที่ต้องการจำนวน 3 ร้านค้า ในโปรแกรมที่พัฒนาด้วย python ที่ปุ่ม“capture”ในตัวโปรแกรมเพื่อทำการครอบหน้าจอในส่วนที่จะนำไปสตรีม และสามารถกดปุ่ม “clear” เพื่อลบร้านค้าที่ต้องการได้

2. หลังจากได้หน้าจอสตรีมที่ต้องการระบบจะทำการตรวจจับกรอบตัวกรอบอักษรในสตรีมโดยใช้อัลกอริทึม Yolov5 และนำข้อมูลที่เป็นรูปภาพกรอบตัวอักษรมาตัดเป็นแถวๆโดยจะใช้เทคนิคในการตรวจ pixel ในหน้าจอทีละแถว เมื่อตรวจเจอช่องว่างระหว่างบรรทัด จะทำการตัดรูปเพื่อแยกข้อมูลออกจากกันและนำชุดข้อมูลสินค้าที่เป็นรูปภาพที่ได้ไปส่งต่อไปยังขบวนการ แปลงภาพเป็นตัวอักษร

3. ขบวนการแปลงภาพเป็นตัวอักษรจะทำการนำชุดข้อมูลที่เป็นภาพมาแปลง เป็นชุดข้อมูลที่ เป็นตัวอักษร โดยใช้ Tesseract ในการแปลง

4. หลังจากได้ชุดข้อมูลที่เป็นตัวอักษร จะนำข้อมูลไปเก็บไว้ใน Database ของ firebase ที่มีลักษณะการทำงานแบบ real-time ผ่านตัว server โดยข้อมูลที่ส่งไปจะมี ชื่อร้านค้า รหัสสินค้า ชื่อสินค้า ราคาสินค้า รูปภาพของไลฟ์สดที่ทำการบันทึกมา ตัวอย่างของข้อมูลชื่อร้านค้า คือ ปีรันย่า บิวตี้ ช๊อป รหัสสินค้าคือ ก22, ชื่อสินค้าคือ ดีทอกซ์ญี่ปุ่น, ราคาสินค้าคือ 39 บาท เป็นต้น

5. server จะทำการส่งข้อมูลสินค้าประกอบไปด้วย ชื่อร้านค้า รหัสสินค้า ชื่อสินค้า ราคาสินค้า และรูปภาพของไลฟ์สด ใน Database ไปแสดงผลใน Desktop Application ***ของผู้ดูแลระบบ***

6. server จะทำการส่งข้อมูลสินค้าประกอบไปด้วย ชื่อร้านค้า รหัสสินค้า ชื่อสินค้า ราคาสินค้า และรูปภาพของไลฟ์สด ใน Database ไปแสดงผลใน mobile application ของผู้ใช้งาน

7. ผู้ใช้ต้องเข้าสู่ระบบผ่านเฟสบุ๊คจึงจะสามารถเข้า application เพื่อเลือกสินค้าที่ต้องการและรอแจ้งเตือนร้านที่กำลังไลฟ์สดได้

8. เมื่อผู้ใช้งานกดปุ่ม “chat” ในหน้ารายละเอียดสินค้าเพื่อลิ้งไปยังแชท ของร้านค้าและผู้ใช้งานสามารถกดปุ่ม “buy” ในหน้ารายละเอียดสินค้าเพื่อลิ้งไปยังแชท ของร้านค้าและจะมีการนำรหัสสินค้า ชื่อสินค้า ราคาสินค้า ส่งไปให้ร้านค้าอัตโนมัติ

## ประโยชน์ที่คาดว่าจะได้รับ

### ช่วยให้ผู้ใช้งานที่เป็นผู้ค้าขายมีช่องทางการขายสินค้ามากขึ้น

### ช่วยให้ผู้ใช้งานที่เป็นลูกค้ามีความสะดวกมากขึ้น

## อุปกรณ์และเครื่องมือที่ใช้ในการดำเนินงาน

### ฮาร์ดแวร์

คอมพิวเตอร์ 2 เครื่อง มีคุณสมบัติดังนี้

1. Acer Nitro5

- ระบบปฏิบัติการ Windows 11

- AMD Ryzen 5 3550H with Radeon Vega Mobile Gfx 2.10 GHz

- Installed RAM 16.0 GB (13.9 GB usable)

1. ROG STRIX

- ระบบปฏิบัติการ Windows 11

- Intel(R) Core(TM) i7-9750H CPU @ 2.60GHz 2.59 GHz

- Installed RAM 16.0 GB (15.9 GB usable)

### ซอฟต์แวร์

1. Python ใช้สำหรับการพัฒนาโปรแกรม
2. Tesseract ใช้สำหรับการประมวลผลภาพ
3. firebase ใช้เป็นฐานข้อมูลในการเก็บข้อมูลที่ใช้ระหว่างพัฒนาโปรแกรม
4. Flutter เป็น Framework หลักสำหรับการพัฒนาแอปพลิเคชั่น
5. Android Studio ใช้สำหรับการสร้างแอปพลิเคชั่น

## แผนการดำเนินงาน

โครงงานปริญญานิพนธฉบับนี้ ดําเนินงาน ณ คณะวิทยาการสารสนเทศ

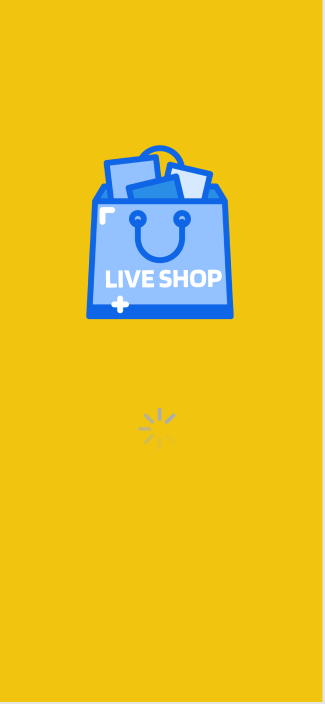
มหาวิทยาลัยมหาสารคาม ระหว่างเดือน พ.ค. 2565 ถึง ก.พ. 2566

ตารางที่ 1 แผนการดำเนินงาน

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| กิจกรรม | เดือน | | | | | | | | | |
| พ.ค. | มิ.ย. | ก.ค. | ส.ค. | ก.ย. | ต.ค. | พ.ย. | ธ.ค. | ม.ค. | ก.พ. |
| 1. ศึกษา และรวบรวมข้อมูล |  |  |  |  |  |  |  |  |  |  |
| 2. กำหนดขอบเขต |  |  |  |  |  |  |  |  |  |  |
| 3. วิเคราะห์และออกแบบ |  |  |  |  |  |  |  |  |  |  |
| 4. พัฒนาโปรแกรม |  |  |  |  |  |  |  |  |  |  |
| 5. ทดสอบการใช้งานโปรแกรม |  |  |  |  |  |  |  |  |  |  |
| 6. ทำรายงานสรุป |  |  |  |  |  |  |  |  |  |  |
| 7. นำเสนอโครงงาน |  |  |  |  |  |  |  |  |  |  |

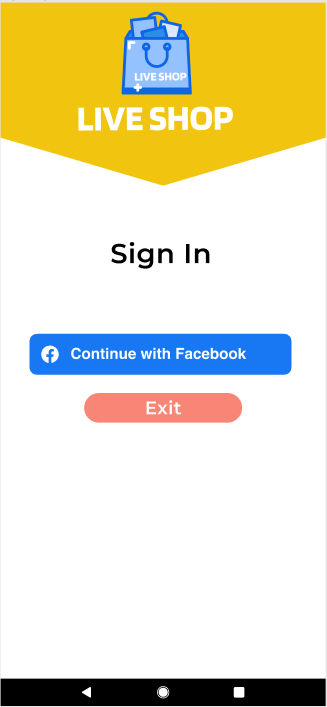
## ตัวอย่างโปรแกรม

### แอปพลิเคชั่นบนอุปกรณ์ไร้สายแบบเคลื่อนที่



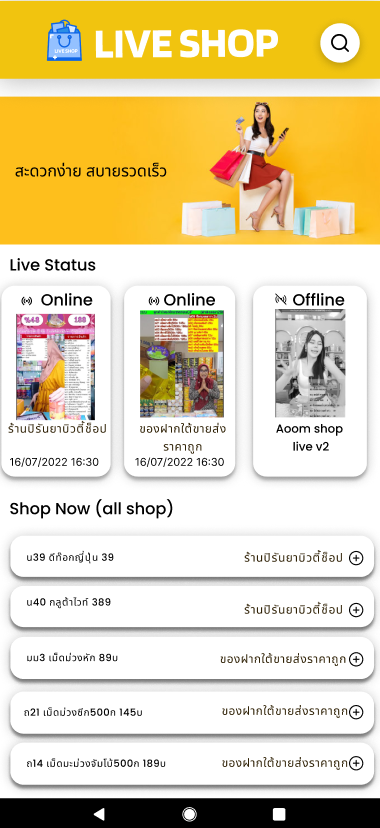
ภาพประกอบที่ . หน้าแรกของโปรแกรม

หน้าแรกของโปรแกรม จะมีการรอโหลดข้อมูลจาก Database เมื่อทำการโหลดข้อมูลเสร็จจะนำพาไปยังหน้าล็อกอินทันที



**ภาพประกอบที่ 1.4** หน้าล็อกอินเข้าสู่ระบบ

หน้าล็อกอินเข้าสู่ระบบ จะมีปุ่ม 2 ปุ่ม “Conitnue with Facebook” เพื่อเข้าสู่ระบบผ่านแอปพลิเคชั่นเฟสบุ๊คเมื่อเข้าสู่ระบบสำเร็จจะนำไปยังหน้าหลักของแอปพลิเคชั่นและ ปุ่ม “Exit” เพื่อออกจากแอปพลิเคชั่น



ภาพประกอบที่ . หน้าหลักของแอปพลิเคชั่น

***หน้าหลักของแอปพลิเคชั่น จะมีปุ่ม”ค้นหา” กดเพื่อไปยังหน้าค้นหาสินค้า และ จะมีแถบ “***Live Status***”แสดงสถานะการไลฟ์สดของร้านทั้งสามร้านว่าร้านไหนกำลังออนไลน์หรือออฟไลน์อยู่ และวันเวลาที่อัพเดทรูปภาพและจะมีแถบ”***Shop Now***” แสดงรายการสินค้าของร้านที่กำลังไลฟ์สดอยู่ทั้งหมด***

รูปภาพประกอบด้วย ข้อความ

คำอธิบายที่สร้างโดยอัตโนมัติ

ภาพประกอบที่ 1.6 หน้าค้นหาสินค้า

หน้าค้นหาสินค้า จะมีช่องเพื่อใส่ข้อมูลชื่อของสินค้าหรือชื่อร้านค้าที่ต้องการค้นหา และเมื่อใส่ข้อมูลลงไปจะมีส่วนในการแสดงผลลัพท์ ของการค้นหาเป็นรายการสินค้าขึ้นมาและเมื่อเลือกรายการสินค้าจะไปยังหน้ารายละเอียดสินค้าทันที



ภาพประกอบที่ 1.7 หน้ารายละเอียดสินค้า

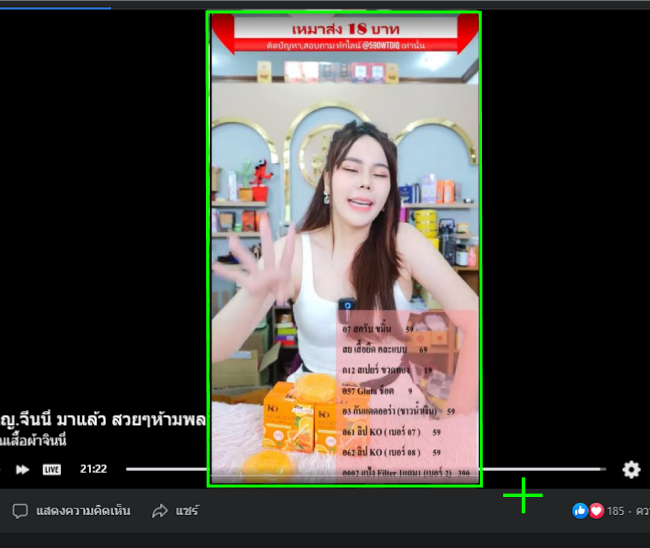
หน้ารายละเอียดสินค้า จะแสดงรูปภาพที่แคปมาจากไลฟ์สดและวันเวลาทอัพเดทรูปภาพและแสดงชื่อร้านค้า ข้อมูลของสินค้าประกอบไปด้วย รหัสสินค้า ชื่อสินค้า ราคา และจะมีปุ่ม 2 ปุ่ม “Chat”เพื่อลิ้งไปยังแชทของร้านค้าและ “Buy”เพื่อลิ้งไปยังแชทของร้านค้าและนำข้อมูลสินค้าส่งไปอัตโนมัติ

### แอปพลิเคชั่นบนคอมพิวเตอร์สำหรับผู้ดูแลระบบ



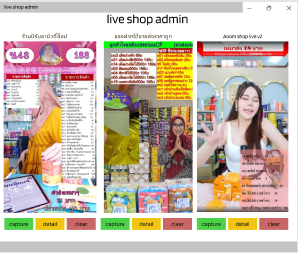
ภาพประกอบที่ 1.8 หน้าแรกโปรแกรม

หน้าแรกโปรแกรม มีพื้นที่ในการแสดงผลสำหรับสตรีม 3 ส่วน และ มีปุ่ม 3 ปุ่ม“capture”เพื่อเลือกครอบจากหน้าไลฟ์สดที่ต้องการนำมาประมวลผล ปุ่ม“clear”เพื่อทำการลบสตรีมที่ต้องการยกเลิกประมวลผล และปุ่ม“detail”เพื่อทำการแสดงข้อมูลรหัสสินค้า ชื่อสินค้า และราคาสินค้า ที่ระบบทำการประมวลผลแปลงรูปภาพเป็นข้อความออกมาโชว์เป็น Popup ที่บริเวรหน้าจอ



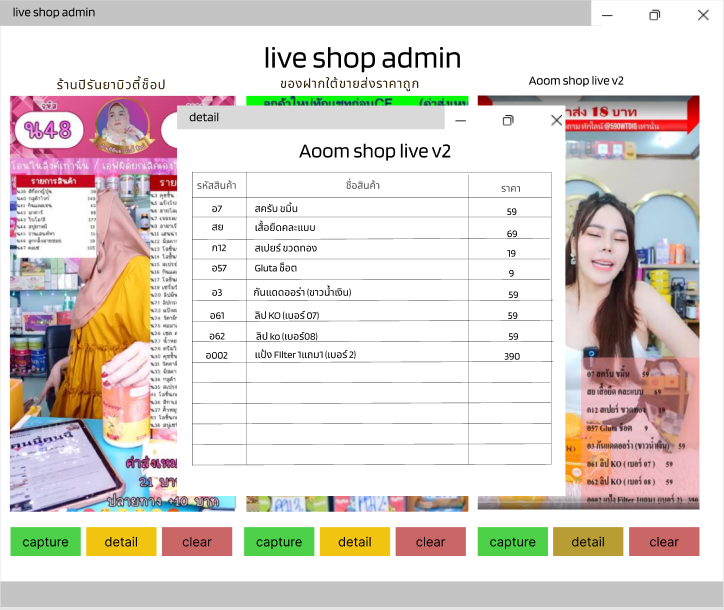
ภาพประกอบที่ 1.9 หน้าเลือกพื้นที่ภาพหน้าจอ

เมื่อกดปุ่ม “capture” แล้วต้องลาก cursor ครอบพื้นที่ที่ต้องการสตรีมจากไลฟ์สดเพื่อนำไปประมวลผล



ภาพประกอบที่ 1.10 หน้าผลลัพธ์ในการถ่ายภาพหน้าจอ

ผลลัพธ์ในการถ่ายภาพหน้าจอ ของปุ่ม Capture



ภาพประกอบที่ 1.11 หน้ารายระเอียดข้อมูล

หน้ารายระเอียดข้อมูล เมื่อกดที่ปุ่ม "detail" ในส่วนของสตรีมมิ่งนั้นๆจะทำการแสดงข้อมูลรหัสสินค้า ชื่อสินค้า และราคาสินค้า ที่ระบบทำการประมวลผลแปลงรูปภาพเป็นข้อความ ออกมาโชว์เป็น Popup ที่บริเวรหน้าจอ

# ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

## ทฤษฎีที่เกี่ยวข้อง

### Flutter

Flutter [1] คือ Framework ที่ใช้สร้าง UI สําหรับ mobile application ที่สามารถทํางานได้ทั้ง IOS และ Android ในเวลาเดียวกันโดยที่ใช้ source code ตัวเดียวกัน โดยภาษาที่ใช้ใน Flutter นั้นจะเป็นภาษา dart ซึ่งถูกพัฒนาโดย Google และ ยังเป็น open source ที่สามารถใช้งานได้ฟรี ตัวอย่าง syntax ของภาษา dart ที่ใช้ใน Flutter ซึ่งจะมีความคล้ายกับภาษา Java เนื่องจาก

dart เป็นภาษาที่รองรับ OOP และมีแนวคิด เช่นเดียวกับภาษา Java

รูปภาพประกอบด้วย ข้อความ

คำอธิบายที่สร้างโดยอัตโนมัติ

ภาพประกอบที่ 2.1 ตัวอย่างโครงสร้าง Flutter

ซึ่งหากสังเกตจาก **ภาพประกอบที่ 2.1** ตัวอย่างโครงสร้าง Flutterจะเห็นว่า Flutter นั้นจะมี Widget พื้นฐานมาให้ เพื่อทำให้การออกแบบ UI มีความง่าย และสะดวกยิ่งขึ้น โดย Widget พื้นฐานของ Flutter หลัก ๆ จะมีอยู่ 2 ชนิดคือ StatelessWidget และ StatefulWidget โดยที่ StatelessWidget จะใช้สร้าง Widget ที่ไม่มีการจัดการสถานะการทำงานใดหรือหน้านั้นๆจะไม่มีการเปลี่ยนแปลง เช่น การแสดงข้อความ,Iconหรือรูปภาพที่ไม่มี animation เข้ามาเกี่ยวข้องเป็นต้น ส่วน StatefulWidget จะใช้สร้าง Widget ที่มีการจัดการสถานการณ์ทำงานต่างๆหรือมีปุ่มที่มี action เช่น การสร้าง Icon ที่มีการใส่ animation ให้สามารถขยับไปมาได้, ปุ่มกดต่างๆ บนหน้า UI เป็นต้น

จุดเด่นหลัก ๆ ของ Flutter คือ ระบบ Hot Reload โดยเมื่อมีการทดสอบ, การสร้าง, การadd features หรือการกระทําต่าง ๆ กับ UI จะต้องมีการ reload เพื่อให้หน้า UI update ซึ่งระบบHot Reload จะเข้ามาช่วยในส่วนของการ reload โดยจุดเด่นของระบบนี้คือการย่นระยะเวลาที่ใช้ในการ reload ให้เหลือเพียงเสี้ยววินาทีเท่านั้น ทําให้การพัฒนา UI ของ application มีความรวดเร็วขึ้นอย่างมาก และยังมีจุดเด่นอื่น ๆ ที่ช่วยให้การพัฒนาเป็นไปได้ง่ายขึ้นไม่ว่าจะเป็น Build-In ที่ช่วยในการออกแบบ UI ให้มีความสวยงามยิ่งขึ้นอย่าง Material Design และ Cupertino (iOS-flavor), มีFramework ที่ช่วยให้การทํา animation ต่าง ๆ หรือ gesture ของ UI เป็นเรื่องง่ายยิ่งขึ้น และยังสามารถใช้งานร่วมกับ IDE ที่กําลังเป็นที่นิยมอยู่ในปัจจุบันอย่าง VS Code และ Android Studio ได้

ข้อเสียหลัก ๆ คือ การใช้ภาษา dart ในการเขียน ซึ่งคนส่วนใหญ่อาจจะยังไม่คุ้นเคยกับsyntax ของภาษา dart ประกอบกับ community ยังเล็กเนื่องจาก Flutter ยังเปิดตัวมาได้ไม่นานนักเมื่อเทียบกับ Framework ตัวอื่น ๆ อย่าง React Native ที่มี community ค่อนข้างใหญ่จึงทําให้ document ต่าง ๆ ยังไม่เยอะเท่าที่ควร ทําให้เวลามีปัญหาเกี่ยวกับการใช้งานอาจจะต้องใช้เวลาในการหาวิธีแก้

### Firebase

เป็นหนึ่งในผลิตภัณฑ์ของ Google โดย Firebase [2]คือ Platform ที่รวบรวมเครื่องมือต่าง ๆ สำหรับการจัดการในส่วนของการเก็บข้อมูล ซึ่งทำให้สามารถ สร้าง Mobile Application ได้อย่างมีประสิทธิภาพ และยังลดเวลาและค่าใช้จ่ายของการทำ Server side หรือการวิเคราะห์ข้อมูลให้อีกด้วย โดยมีทั้งเครื่องมือที่ฟรี และเครื่องมีที่มีค่าใช้จ่าย Firebase มีบริการให้ใช้หลายอย่าง สามารถแบ่งเป็นหมวดหมู่ดังนี้

1.) Better Apps

- Cloud Firestore  จัดเก็บและซิงค์ข้อมูลระหว่างผู้ใช้และอุปกรณ์ในระดับโลกโดยใช้ฐานข้อมูล NoSQL ที่โฮสต์บนคลาวด์ Cloud Firestore ให้การซิงโครไนซ์แบบสดและการสนับสนุนออฟไลน์พร้อมกับการสืบค้นข้อมูลที่มีประสิทธิภาพ การผสานรวมกับผลิตภัณฑ์ Firebase อื่นๆ ช่วยให้คุณสร้าง แอปแบบไร้เซิร์ฟเวอร์ได้อย่างแท้จริง

- Authentication  จัดการผู้ใช้ของคุณด้วยวิธีที่ง่ายและปลอดภัย Firebase Auth มีหลายวิธีในการตรวจสอบสิทธิ์รวมถึงEmailและpasswordผู้ให้บริการบุคคลที่สามเช่น Google หรือ Facebook และใช้ระบบบัญชีที่คุณมีอยู่โดยตรง สร้างอินเทอร์เฟซของคุณเองหรือใช้ประโยชน์จากโอเพ่นซอร์ส UI ที่ปรับแต่งได้อย่างเต็มที่

- Hosting  ลดความซับซ้อนของเว็บโฮสติ้งของคุณด้วยเครื่องมือที่สร้างขึ้นเฉพาะสำหรับเว็บแอปสมัยใหม่ เมื่อคุณอัปโหลดเนื้อหาเว็บของคุณเราจะส่งเนื้อหาเหล่านั้นไปยัง CDN ทั่วโลกของเราโดยอัตโนมัติและมอบใบรับรอง SSL ฟรีเพื่อให้ผู้ใช้ของคุณได้รับประสบการณ์ที่ปลอดภัยเชื่อถือได้และมีเวลาแฝงต่ำไม่ว่าจะอยู่ที่ใดก็ตาม

- Realtime Database  Realtime Database คือฐานข้อมูลดั้งเดิมของ Firebase เป็นโซลูชันที่มีประสิทธิภาพและมีเวลาแฝงต่ำสำหรับแอปบนอุปกรณ์เคลื่อนที่ที่ต้องการสถานะการซิงค์ระหว่างไคลเอนต์แบบเรียลไทม์ เราขอแนะนำ Cloud Firestore แทน Realtime Database สำหรับนักพัฒนาส่วนใหญ่ที่เริ่มโปรเจ็กต์ใหม่

2.) Improve app quality

- [Crashlytics](https://firebase.google.com/products/crashlytics)  ลดเวลาในการแก้ไขปัญหาของคุณด้วยการเปลี่ยนข้อขัดข้องจากหิมะถล่มให้เป็นรายการปัญหาที่จัดการได้ รับข้อมูลเชิงลึกที่ชัดเจนและนำไปปฏิบัติได้ว่าปัญหาใดที่ต้องจัดการก่อนโดยเห็นผลกระทบของผู้ใช้ในแดชบอร์ด Crashlytics การแจ้งเตือนแบบเรียลไทม์จะช่วยให้คุณมีความเสถียรแม้ในขณะเดินทาง Crashlytics เป็นตัวรายงานข้อขัดข้องหลักของ Firebase

- Performance Monitoring  วินิจฉัยปัญหาประสิทธิภาพของแอปที่เกิดขึ้นบนอุปกรณ์ของผู้ใช้ ใช้การติดตามเพื่อตรวจสอบประสิทธิภาพของบางส่วนของแอปและดูมุมมองสรุปในคอนโซล Firebase อยู่เหนือเวลาเริ่มต้นของแอปและตรวจสอบคำขอ HTTP โดยไม่ต้องเขียนโค้ดใด ๆ

- Test Lab  เรียกใช้การทดสอบอัตโนมัติและกำหนดเองสำหรับแอปของคุณบนอุปกรณ์เสมือนและจริงที่โฮสต์โดย Google ใช้ Firebase Test Lab ตลอดวงจรการพัฒนาของคุณเพื่อค้นหาจุดบกพร่องและความไม่สอดคล้องกันเพื่อให้คุณสามารถนำเสนอประสบการณ์ที่ยอดเยี่ยมบนอุปกรณ์หลากหลายประเภท

3.) Grow your business

- Google Analytics  วิเคราะห์คุณลักษณะและพฤติกรรมของผู้ใช้ในแดชบอร์ดเดียวเพื่อทำการตัดสินใจอย่างชาญฉลาดเกี่ยวกับแผนงานผลิตภัณฑ์ของคุณ รับข้อมูลเชิงลึกแบบเรียลไทม์จากรายงานหรือส่งออกข้อมูลเหตุการณ์ดิบไปยัง Google BigQuery สำหรับการวิเคราะห์ที่กำหนดเอง

- Remote Config  กำหนดวิธีการแสดงผลแอปของคุณสำหรับผู้ใช้แต่ละคน เปลี่ยนรูปลักษณ์เปิดตัวฟีเจอร์ทีละน้อยเรียกใช้การทดสอบ A / B ส่งมอบเนื้อหาที่กำหนดเองให้กับผู้ใช้บางรายหรือทำการอัปเดตอื่น ๆ โดยไม่ต้องปรับใช้เวอร์ชันใหม่ทั้งหมดนี้ทำได้จากคอนโซล Firebase ตรวจสอบผลกระทบของการเปลี่ยนแปลงของคุณและทำการปรับเปลี่ยนในเวลาไม่กี่นาที

- Cloud Messaging  ส่งข้อความและการแจ้งเตือนไปยังผู้ใช้ข้ามแพลตฟอร์มทั้ง Android, iOS และเว็บได้ฟรี สามารถส่งข้อความไปยังอุปกรณ์เดียวกลุ่มอุปกรณ์หรือหัวข้อเฉพาะหรือกลุ่มผู้ใช้ Firebase Cloud Messaging (FCM) ปรับขนาดเป็นแอปที่ใหญ่ที่สุดโดยส่งข้อความหลายแสนล้านข้อความต่อวัน

### การประมวลผลภาพ (Image Processing)

การประมวลผลภาพ (Image Processing) [3]เป็นการนำภาพมาประมวลผลหรือคิดคำนวณด้วย คอมพิวเตอร์แล้วใช้กรรมวิธีใดๆมากระทำกับข้อมูลภาพเพื่อให้ได้ภาพที่มีคุณสมบัติตามต้องการทั้งในเชิง คุณภาพและปริมาณ โดยมีขั้นตอนต่างๆที่สำคัญ คือ การทำให้ภาพมีความคมชัดมากขึ้น การกำจัดสัญญาณรบกวนออกจากภาพ การแบ่งส่วนของวัตถุที่สนใจออกมาจากภาพ เพื่อนำภาพวัตถุที่ได้ ไปวิเคราะห์หาข้อมูลเชิงปริมาณ เช่น ขนาด รูปร่าง และทิศทางการเคลื่อนของวัตถุในภาพ คอมพิวเตอร์มีความสามารถในการคำนวณและประมวลผลข้อมูลจำนวนมหาศาลได้ในเวลาอันสั้นจึงมีประโยชน์อย่างมากในการเพิ่มประสิทธิภาพการประมวลผลภาพและวิเคราะห์ข้อมูลที่ได้จากภาพในระบบ ลักษณะของการใช้งานจะมีเป็น

1. การรู้จำภาพ (Image Recognition)

การรู้จำภาพ (Image Recognition) การรู้จำภาพเป็นกระบวนการในการระบุและการ ตรวจจับวัตถุหรือคุณสมบัติที่มีอยู่ในระบบดิจิตอล ซึ่งได้จากภาพหรือวีดีโอแนวคิดนี้ได้ถูกนำมาใช้ใน แอพพลิเคชั่นมากมายที่เหมือนกันในระบบสำหรับการทำงานของระบบอัตโนมัติในโรงงานการ ตรวจสอบด่านเก็บเงินด้วยการรักษาความปลอดภัยและการรักษาความปลอดภัย อัลกอริทึมการรู้จำ ภาพทั่วไป

1. การสกัดคุณลักษณะเด่น (Feature Extraction)

การสกัดคุณลักษณะเด่น (Feature Extraction) เป็นการหาลักษณะเด่นที่ได้จากภาพ โดยผลลัพธ์ที่ได้จากการนำไปหาลักษณะเด่นจะบ่งบอกถึงลักษณะเช่น ความสูง ความยาว สีหรือเวคเตอร์ ที่เป็นคุณลักษณะ ที่สำคัญของรูปภาพ เป็นต้น นอกเหนือจากข้อมูลคุณลักษณะดังกล่าว ในปัจจุบันยัง มีวิธีการการสกัดหา ลักษณะเด่นที่นิยมใช้กันอย่างแพร่หลายที่ใช้ในการต่อภาพคือ การวางตัวของเส้นขอบ(Edge) การหามุม (Corner) และการหาพื้นที่น่าสนใจ (Blob) เป็นต้น



ภาพประกอบที่ . การตรวจจับและประมวลผลใบหน้าจากภาพ [4]

ตัวอย่างการ นำการประมวลผลภาพไปใช้งาน เช่น

- ระบบตรวจจับใบหน้าในกล้องดิจิตอล โดยกล้องจะมีระบบตรวจว่าส่วนไหนของภาพมี ลักษณะคล้ายใบหน้า แล้วกล้องก็จะท่าการโฟกัสตำแหน่งที่ตรวจจับเพื่อภาพมีความคมชัดมากขึ้น เช่น ระยะห่าง ระหว่างคิวมุมปาก โหนกแก้ม จมูก โครงหน้า

- ระบบตรวจจับการเคลื่อนไหวเพื่อรักษาความปลอดภัย โดยกล้องจะเปรียบเทียบภาพก่อน หน้าและภาพปัจจุบันถ้ามีส่วนใดส่วนใดเปลี่ยนแปลงระบบจะบันทึกเฉพาะภาพที่มีการเปลี่ยนแปลงทำให้ประหยัด เนื้อที่ในการเก็บภาพและสามารถตรวจสอบได้ภาพหลัง

- ระบบอ่านคิวบาร์โค้ด โดยระบบนี้จะอ่านรหัสจากบาร์โค้ด แล้วแปลงเป็นข้อมูลซึ่งสามารถ นำมาใช้งานได้ เช่น QR code, Microsoft tag ภาพที่นำมาประมวลผลนั้นเป็นได้ทั้งภาพนิ่งและ ภาพเคลื่อนไหวในการนำภาพนิ่งเข้ามาประมวลผลนั้น โปรแกรมจะอ่านไฟล์ภาพขึ้นมาแล้วถอดรหัสจาก ค่าสีของภาพแต่ละจุดเป็นตัวเลขเพื่อนำมาประมวลผลสำหรับการประมวลผลภาพเคลื่อนไหวนั้น โปรแกรมจะมองเห็นเป็นภาพนิ่งหลายภาพที่เรียงต่อกัน จึงมีประโยชน์อย่าง มากในการเพิ่ม ประสิทธิภาพการประมวลผลภาพและวิเคราะห์ข้อมูลได้จากภาพในระบบต่าง ๆดังกล่าวข้างต้น การประมวลผลภาพ สามารถจำแนกตาม วัตถุประสงค์ในการใช้งานดังนี้

- การปรับปรุงคุณภาพของภาพ (Image Enhancement and Restoration) เช่น ปรับภาพ ให้คมชัด มากขึ้นการปรับ Contrast หรือการปรับเน้นเส้นขอบภาพ (Edge Enhancement) หรือการ กรองสัญญารบกวน (Image Filtering) เพื่อกำจัดสัญญาณรบกวนได้

- การบีบอัดภาพ (Image Compression) เนื่องจากข้อมูลภาพนั้นมีขนาดใหญ่มากโดยเฉพาะ ภาพสี ซึ่งจะ ทำให้การจัดเก็บหรือรับส่งไฟล์ข้อมูลขนาดใหญ่เสียเวลามาก เพื่อทำให้ขนาดภาพเล็กลงจึง ทำให้การบีบอัด ข้อมูลภาพ เช่น JPEG หรือ GIF และตัวอย่างการบีบอัดข้อมูลวีดีโอ เช่น MPEG หรือ AVI เป็นต้น

- การวิเคราะห์ข้อมูลภาพ (Image Analysis) เป็นการสังเคราะห์ ข้อมูลที่มีความหมายจากภาพ แล้ว นำไปใช้งาน ซึ่งอาจต้องมีรู้จ่าวัตถุ (object recognition) การแยกส่วนวัตถุ 317 The Tenth National conference on Computing and Information Technology NCCIT4 (segmentation) ตรวจจับการ เคลื่อนไหวที่ของวัตถุ (motion detection)

### OCR หรือ Optical Character Recognition

**OCR หรือ Optical Character Recognition**[5] หรือ **"การรู้จำอักขระด้วยแสง"** เป็นเทคโนโลยีที่แปลงเอกสาร รูปภาพ หรือแม้แต่ [ลายเซ็นอิเล็กทรอนิกส์](https://tips.thaiware.com/1598.html)  ให้อยู่ในรูปแบบตัวอักษร หรือ ข้อความ (Plain Text) ไฟล์ดิจิทัลที่สามารถสืบค้นได้ด้วยคำค้นหา (Keyword) ยกตัวอย่าง ไฟล์เอกสารที่มาในรูปแบบ PDF มีข้อดีก็คือไม่สามารถปลอมแปลง แก้ไขได้ แต่ถ้าต้องการแก้ไขข้อความใด ๆ ในนั้นก็เป็นเรื่องยากลำบาก วิธีแก้ไขปัญหาก็คือ แปลงไฟล์ PDF ให้เป็น OCR ซึ่งวิธีการที่ทุกคนคุ้นเคยก็คือ การนำแผ่นเอกสาร มาสแกนข้อความ ตัวอักษร เส้นตารางและอื่น ๆ ให้เป็นไฟล์ Word หรือ PDF ด้วยเครื่องสแกนเนอร์ หรือกล้องถ่ายภาพ



ภาพประกอบที่ . ขบวนการการทำงานของ Optical Character Recognition

ส่วนการทำงานของ OCR นั้น หลัก ๆ คือ ใช้การจดจำรูปแบบ เพื่อกำหนดอักขระของแต่ละประเภทไฟล์ จากนั้น ซอฟต์แวร์จะทำการอ่านข้อความและอักขระ แล้วแปลงเป็นไฟล์ที่สืบค้นได้ นอกจากนี้ การทำงานของ OCR ยังขึ้นอยู่กับซอฟต์แวร์ที่ใช้ประมวลผล เพราะนอกเหนือจากการแปลงรูปภาพ อักขระใด ๆ เป็นข้อความแล้ว ซอฟต์แวร์บางตัวสามารถจัดวางรูปแบบข้อความในไฟล์ OCR ได้

### Object Detection

Object Detection [6]การตรวจจับวัตถุ คือ เทคโนโลยีในทางคอมพิวเตอร์ หลักการที่เกี่ยวกับ Computer Vision และ Image Processing ที่ใช้ในงาน [AI](https://www.bualabs.com/archives/19/machine-learning-%e0%b8%84%e0%b8%b7%e0%b8%ad%e0%b8%ad%e0%b8%b0%e0%b9%84%e0%b8%a3-%e0%b8%84%e0%b8%a7%e0%b8%b2%e0%b8%a1%e0%b8%ab%e0%b8%a1%e0%b8%b2%e0%b8%a2%e0%b8%82%e0%b8%ad%e0%b8%87-machine-learning/) ตรวจจับวัตถุชนิดที่กำหนด เช่น มนุษย์ รถยนต์ อาคาร ที่อยู่ในรูปภาพ หรือวิดีโองาน Object Detection การตรวจจับวัตถุในรูปภาพ สามารถเจาะลึกลงไปได้อีกหลายแขนง เช่น การทำ ตรวจจับหน้าคน ตรวจจับคนเดินถนน สามารถประยุกต์ใช้ได้หลากหลาย เช่น ใช้ในงานรักษาความปลอดภัย และรถยนต์ไร้คนขับ เป็นต้น

### YOLO (You Only Look Once)

เป็นอัลกอริทึม [7]ที่นําแนวความคิดของการทํานายตําแหน่งและขนาดของ กล่องจากความน่าจะเป็นที่กล่องนั้น จะเป็นกรอบล้อม วัตถุ แต่สิ่งที่ YOLO มีความสามารถและความเร็ว เหนือกว่าอัลกอริทึมอื่น เช่น Faster R-CNN ที่มีการ ทํางานในลักษณะที่จะทํานายตําแหน่งของกรอบล้อม วัตถุและค่อยนําวัตถุในกล่องนั้น ไปผ่านแบบจําลองเพื่อ ทํานายวัตถุในกล่อง แต่ YOLO นั้นจะทํานายทั้งกรอบ ล้อมวัตถุ และความน่าจะเป็นของวัตถุบางส่วนที่อยู่ใน กรอบออกมาพร้อมกันทีเดียว YOLO จัดว่าเป็นเทคนิคการตรวจจับวัตถุในภาพซึ่ง เป็นซอฟต์แวร์เปิดสําหรับงานปัญญาประดิษฐ์แบบ โครงข่ายใยประสาท (Neural Network) ที่พัฒนาด้วย โปรแกรมภาษา C++ และสามารถทํางานบนหน่วยประมวลผล CUDA ของ GPU ได้เป็นอย่างดี เหมาะกับ การประมวลผลภาพแบบ Real Time ภาพจากกล้องหรือ วิดีโอ ปัจจุบัน YOLO มี การพัฒนามาแล้ว 5 เวอร์ชั่นปัจจุบันคือ YOLO v5 โครงสร้างโครงข่ายประสาทเทียมของ YOLO แต่ละ เวอร์ชั่นจะมีConvolution Box ที่ แตกต่างกัน แต่หลักการโดยทั่วไปแล้ว YOLO จะแบ่งภาพออกเป็น Grid Cell เล็กๆ และแต่ละ Grid Cell จะถูกทํานายผ่าน แบบจําลอง เพื่อหาตําแหน่งจุดกึ่งกลางของวัตถุ และ ความน่าจะเป็นที่จะมีวัตถุใด ใน Grid Cell

## ระบบงานที่เกี่ยวข้อง

### Safety Helmet Detection Based on YOLOv5

"Safety Helmet Detection Based on YOLOv5" [8]เป็นงานวิจัยของ Fangbo Zhou, Huailin Zhao , Zhen Nie มหาวิทยาลัย Shanghai Institute of Technology [8]ที่พัฒนาระบบตรวจสอบความปลอดภัยของคนงาน ซึ่งระบบจะทำการตรวจจับภาพของคนงาน และ จะแสดงกรอบที่มีคำอธิบาย โดยได้นำรูปภาพจากอินเทอร์เน็ตจำนวน 6045 ภาพมาทำการทดสอบ โดยศีรษะของคนงานที่ไม่สวมหมวกนิรภัยจะมีคำอธิบายว่า"Alarm" และสำหรับศีรษะของคนงานที่สวมหมวกนิรภัยจะมีคำอธิบายว่า“Helmet” โดยใช้ algorithm ที่ทันสมัยที่สุดในการตรวจสอบ algorithm งานตรวจจับวัตถุถูกใช้อย่างแพร่หลายในความเป็นจริง เป้าหมายของการตรวจจับคือการค้นหาวัตถุทั้งหมดที่น่าสนใจในภาพ ซึ่งจะมี 2 งานย่อยคือการกำหนดหมวดหมู่และการระบุตำแหน่งของวัตถุ แม้ว่าอัลกอริธึมการตรวจจับแบบดั้งเดิมสามารถทำงานได้ผลลัพธ์ที่ดี ในบางสถานการณ์ในสภาพแวดล้อมที่แปรปรวน เช่น การเปลี่ยนแปลงของสภาพอากาศ การกระจายตัวของคนงานที่ไม่สม่ำเสมอ

และหมวกกันน็อคแบบต่างๆ ทำให้ความแม่นยำนั้นรับประกันได้ยากว่ามันสามารถทำงานได้ถูกต้อง จึงได้นำตัวYOLOv5มาใช้ซึ่งมีความเร็วและความแม่นยำสูง

ภาพประกอบที่ 2.4 ภาพการตรวจจับหมวกนิรภัยของคนงาน [8]

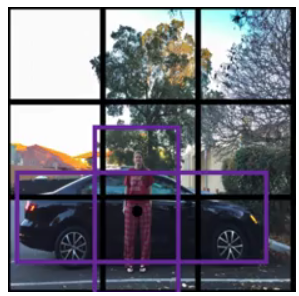
โดยการทำงานของYOLOv5 [9]คือจากรูป 1 รูปเต็มๆ จะทำการแบ่ง Grid cell ออกมาเป็น

รูปภาพประกอบด้วย ห้องญี่ปุ่น

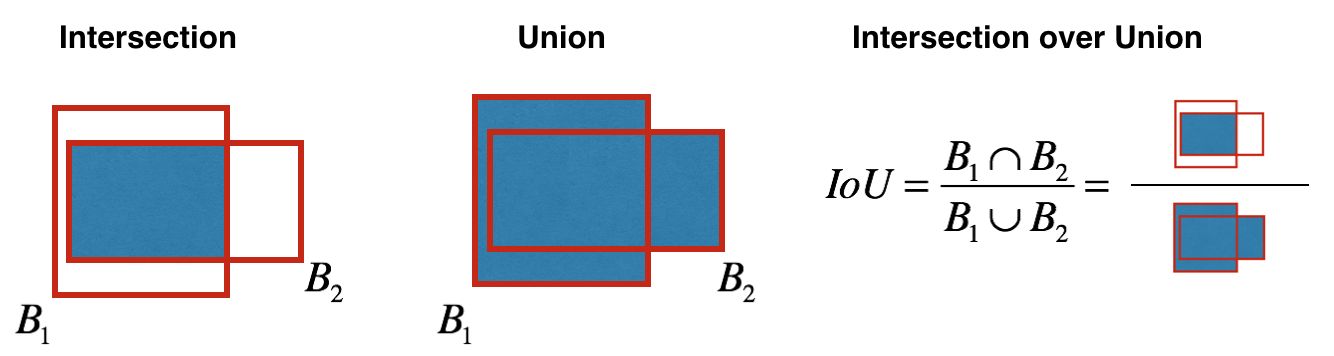
คำอธิบายที่สร้างโดยอัตโนมัติn x n grid ยิ่งแบ่งมากก็จะละเอียด trade off กับการคำนวณ

ภาพประกอบที่ 2.5 การแบ่งช่องของรูปภาพออกเป็น grid [9]

และในแต่ละ grid จะมี Label ทุกช่อง เช่น[Pc, bx, by, bh, bw, c1, c2 ,… , cn] โดยที่ Pc คือ ความน่าจะเป็นที่มีวัตถุอยู่ใน Grid นั้นๆ ถ้าไม่มีคือจะได้ค่า 0 ถ้ามีจะได้ค่า 1 bx, by คือตำแหน่งตรงกลางของ Object ว่าอยู่พิกัดไหนส่วน bh, bw คือขนาดความสูงและกว้างของ Object ว่าสูง, กว้างขนาดไหน และ c1,c2,..,cn คือ ผลลัพธ์ว่าเป็น class อะไร ถ้าโจทย์มีแค่ detect หมวกนิรภัย ก็จะมี class เดียว และ grid นั้นมีค่าเป็น 1 แต่ถ้ามีหลาย Object ก็จะมีเลขต่อๆไป



ภาพประกอบที่ 2.6 ภาพของเทคนิค Anchor Box [9]

จากภาพประกอบที่ 2.6 ภาพของเทคนิค Anchor Box [9] จะรับได้เฉพาะ 1 grid คือ 1 Object แต่กรณีถ้ามีหลาย Object เราจะใช้หลักการที่เรียกว่า Anchor Box โดยเราสามารถกำหนดจำนวน Box นี้ได้ ก็จะมี Label แบบด้านบน 2 อันใน 1 Gridได้ และตัว YOLO ก็จะคำนวณให้ว่ารูปนั้นใกล้ Anchor อันไหนสุดจากค่า IOU ก็จะถูกกำหนดไปที่ Anchor นั้น

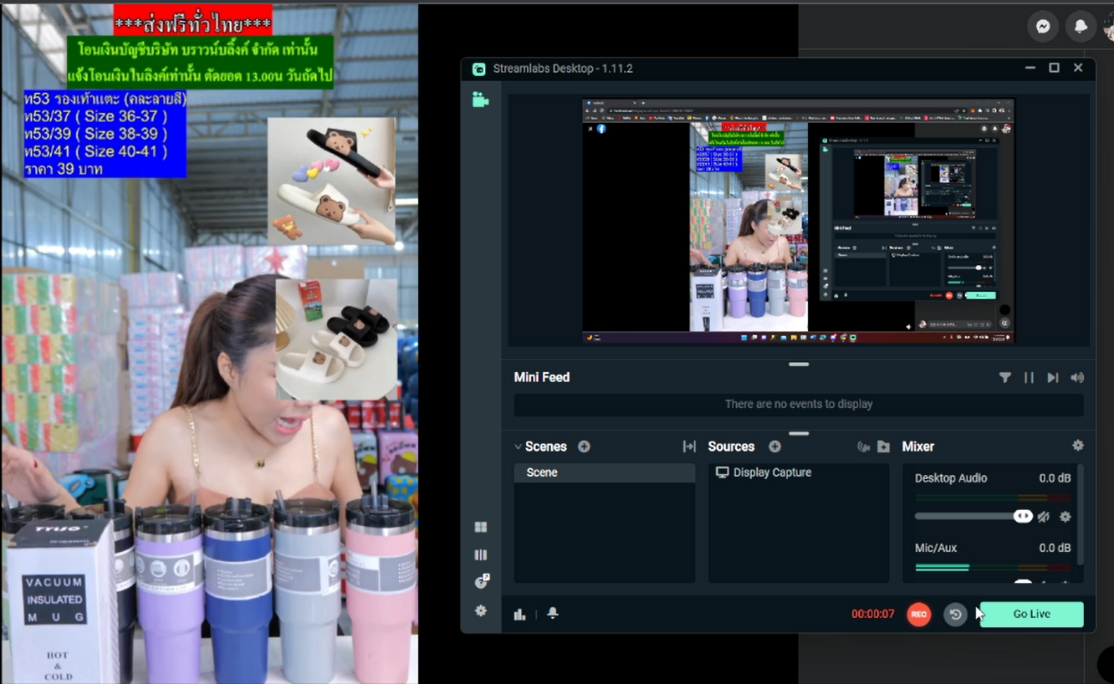
IOU จะมามีส่วนหลักๆหลายส่วนใน YOLO เช่นเวลาโมเดลทำนายออกมา มันอาจจะทำนายรูปรถเดียวกันแต่หลายกล่องได้ จึงต้องใช้ metric IOU ในเลือกกล่องเดียวเป็นตัวแทนของ Object โดยหลักการคำนวณคือ หาส่วนที่ Intersect หารส่วนที่ Union กัน ถ้ากล่องใดมีค่านี้สูง เกินเกณฑ์ที่กำหนดแสดงว่ามันคือ Object เดียวกัน

ภาพประกอบที่ 2.7 การทำงานของขบวนการ IOU [9]

**ขั้นตอนการดำเนินงาน**

สำหรับในบทนี้จะกล่าวถึงขั้นตอนในการดำเนินงานของโครงงานปริญญานิพนธ์ซึ่งจะทำให้ทราบถึงการวิเคราะห์และการออกแบบแอพพลิเคชั่นโดยละเอียดว่ามีแนวทางในการดำเนินงานหรือมีขั้นตอนในการทำงานของแอพพลิเคชั่นอย่างไรบ้างโดยขั้นตอนในการดำเนินงานมีรายละเอียดดังนี้

## เก็บรวบรวมข้อมูล

เก็บข้อมูลโดยการบันทึกคลิปวิดีโอโดยใช้ Streamlabs Desktop ในการบันทึกคลิปวิดีโอ แล้วนำวิดีโอมาเปิดและบันทึกเป็นภาพนิ่ง ทำการบันทึกวิดีโอจำนวน 3 ร้านค้า เงื่อนไขในการเก็บ โดยมีการเก็บวิดีโอเป็น 1 ร้านจะมี 3 วิดีโอ 1 วิดีโอจะแบ่งเป็น 6 คลิปวิดีโอ 1 คลิปวิดีโอ จะแบ่งเป็น คลิปละ 3 นาที

ภาพประกอบที่ 3.1 Streamlabs Desktop

## ตรวจจับกรอบข้อความ

### การ training

เตรียมข้อมูลในการ training โดยบันทึกรูปภาพจาก Video ไลฟ์สดที่เก็บมาได้โดยจะทำการบันทึกเฉพาะช่วงที่มีการเปลี่ยนแปลงของข้อความบนหน้าจอไลฟ์สดดัง โดยจะได้จาก

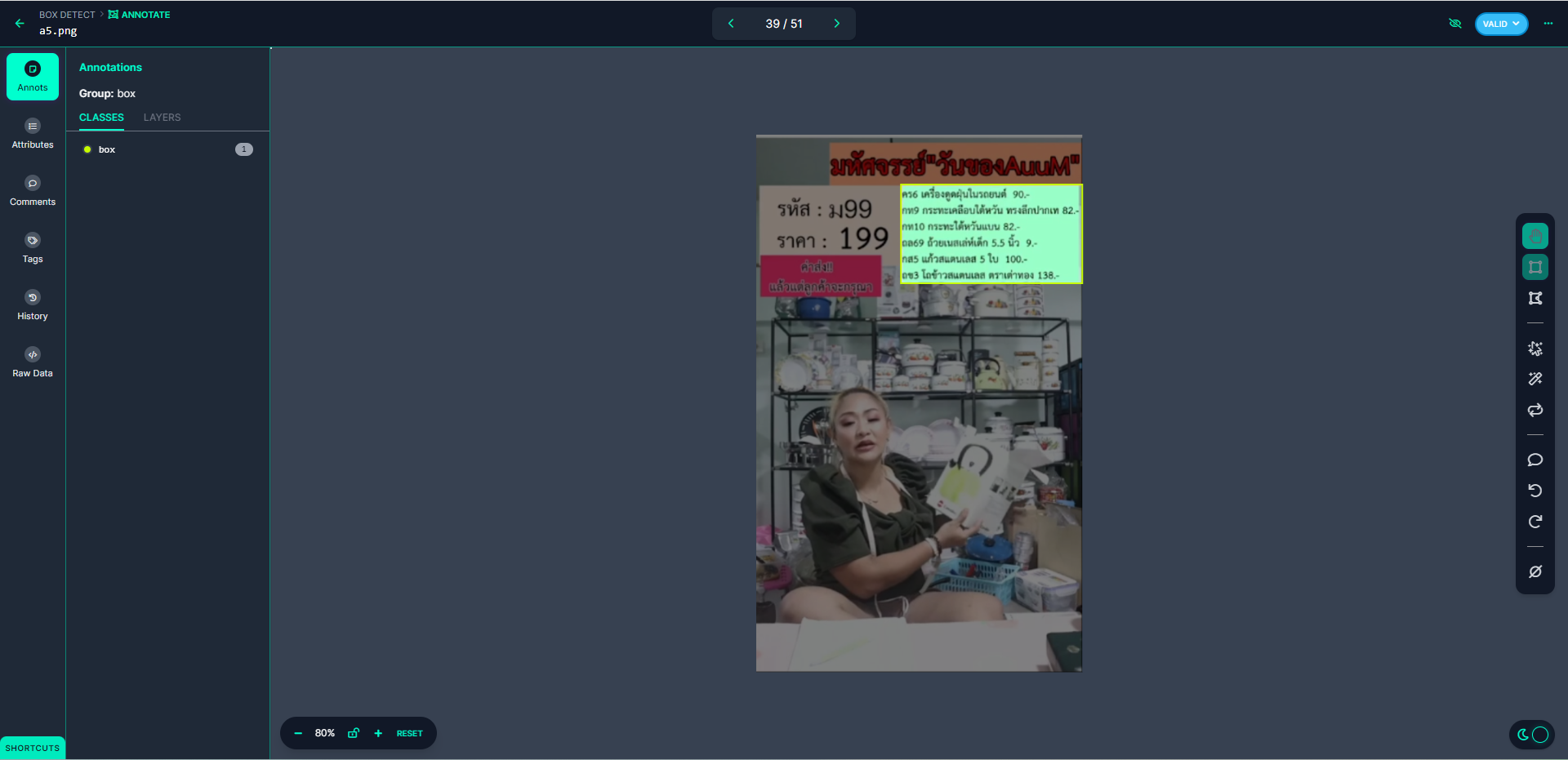
ร้าน FIRST SHOP V2 จำนวน 33 ภาพ จากร้าน KANYA SHOP ขายถูกทุกอย่าง จำนวน 28 ภาพ จากร้าน มหัศจรรย์"วันของAuuM" จำนวน 33 ภาพ รวมกันทั้งหมด 92 รูปภาพ



**ภาพประกอบที่** 3.2 ตัวอย่างรูปภาพที่ทำการเก็บ

### การวาดภาพผลเฉลย

ทำการ การวาดภาพผลเฉลยรูปภาพผ่าน roboflow framework สำหรับใช้จัดเก็บ เตรียมชุดข้อมูล และสร้างแบบจำลองต่างๆ ที่สามารถใช้งานผ่าน web browser โดยจะทำการวาดภาพผลเฉลยในรูปภาพที่เตรียมมาทั้งหมด



**ภาพประกอบที่** 3.3 การวาดภาพผลเฉลยรูปภาพ

เมื่อทำการวาดภาพผลเฉลยรูปภาพทั้งหมดจะทำการแบ่งส่วนของข้อมูลออกเป็นดังนี้

**ตารางที่** 2 การแบ่งส่วนของข้อมูล

|  |  |  |
| --- | --- | --- |
| Training Set | Validation Set | Testing Set |
| 62 | 17 | 13 |

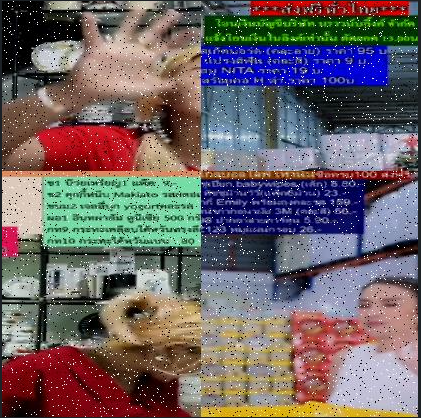
และจะนำข้อมูลในส่วนของ Training Set ไปทำการ Generate รูปภาพเพิ่มเพื่อเพิ่มจำนวนของข้อมูลที่จะทำไป train โดยจะทำการ หมุนภาพ 90 องศา,ทำให้เป็น ภาพสีเทา,เพิ่มแสงสดแสง,เพิ่มสิ่งรบกวนในรูปภาพและปรับขนาดของรูปภาพเป็น 416 x 416 pixel ทุกรูปภาพ



**ภาพประกอบที่** 3.4 การเพิ่มรูปภาพใน Traing Set

โดยจะได้รูปภาพจากการ Generate ทั้งหมด 186 ภาพ จะทำการ Generate ทั้งหมด 3 ครั้ง

จะมีรูปภาพที่ใช้ในการ train ทั้งหมด 558 โดยภาพที่ Generate ออกมาจะมีรูปแบบดังนี้



ภาพประกอบที่ 3.5 ตัวอย่างรูปภาพที่ Generate ออกมา

## อ่านข้อความ OCR

## วัดประสิทธิภาพ

อ้างอิง

|  |  |
| --- | --- |
| [1] | Hizoka, "Flutter," [Online]. Available: https://shorturl.asia/zPKJZ. [Accessed 15 07 2022]. |
| [2] | J. Saengow, "Firebase," [Online]. Available: https://shorturl.asia/g6H0J. [Accessed 07 16 2022]. |
| [3] | N. Jesadapatrakul, "Image Processing," [Online]. Available: https://shorturl.asia/4LirK. [Accessed 10 06 2022]. |
| [4] | E. BARRETT, "fortune," [Online]. Available: https://fortune.com/2018/10/28/in-china-facial-recognition-tech-is-watching-you/. [Accessed 14 07 2022]. |
| [5] | NUMKINGSTON, "Optical Character Recognition," [Online]. Available: https://shorturl.asia/nmq7h. [Accessed 12 07 2022]. |
| [6] | S. KANOKTIPSATHARPORN. [Online]. Available: https://shorturl.asia/twFLQ. [Accessed 14 07 2022]. |
| [7] | S. D. R. G. a. A. F. J. Redmon, "ieeexplore," [Online]. Available: https://ieeexplore.ieee.org/document/7780460. [Accessed 14 07 2022]. |
| [8] | H. Z. a. Z. N. F. Zhou, "Safety Helmet Detection Based on YOLOv5," [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9362711. [Accessed 07 15 2022]. |
| [9] | I. Kunakorntum, "medium," [Online]. Available: http://surl.li/ctfuc. [Accessed 15 07 2022]. |