

Table of contents









Methodology & Performance Metric



Core Functionality



Conclusion



Introduction

Overview of the problem & MVP

Introduction





Digital transformation is increasingly strong to increase work efficiency, technology must be applied to reduce labor.

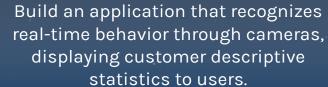






Purpose

From MVP, capture customer action trends in a grocery store context. Based on the results, users provide solutions for businesses.







Problems

Challenge of the problem

Problems



User experience

Items are not arranged properly, affecting the shopping experience



Product Management

Stores often waste time and effort in checking inventory



Abnormal behavior

Customers whose theft behavior is difficult to control. Items fall and broken





User experience

Safe behaviors for customers when shopping: tiptoeing, bending, reaching, kneeling, etc.

Product Management

Focus on actions: picking up items, viewing items, putting items in the cart, etc. so the application can count the number of items each customer buys.

Abnormal behavior

Unusual behaviors at the store: putting products in pants or shirt pockets, arbitrarily opening product packaging, returning goods almost every day, dropping items, etc.

SOLUTIONS

Using AI to identify the above behaviors





Methodology & Performance Metric

Method and metric



Method





Detect Model



Evaluation

Why YOLOv8?

This is the model used by learning bounding boxes. In addition, YOLOv8 also uses mosaic data enhancement technique. Helps the model learn about objects in an obscured state.

Dataset

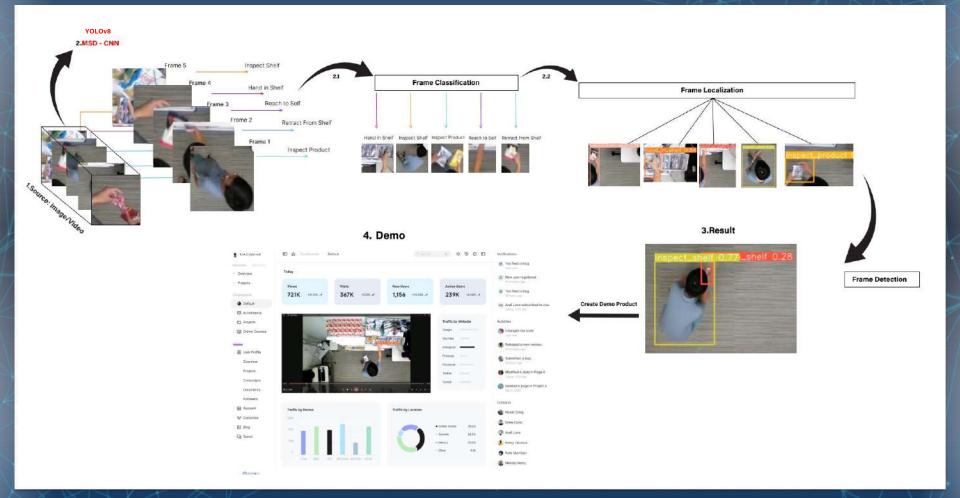
From the customer purchase videos, we will label each image frame belonging to the behavior analyzed above.

Use YOLOv8 To identify user behavior.

After training, there needs to be an evaluation method to determine whether the model achieves good results.

Intechnology, to mark labels, the team will use the RoboFlow framework, and the YOLOv8 model will be from the Ultralytics Framework.

Pipeline





What metrics?

The main Metric we will use in this project is mAP (Mean Average Precision) is mean of AP of all class in Detect task.

What is AP?

This is the area below the Precision Recall Curve. And It depends on the IoU (Intersection over Union) threshold. The larger the AP, the larger the area.

Prove that the model is good and vice versa





Core Functionality

Features and functions



Features



Identify behavior

Behavior classification

Real-time video

Does not cause large delays





Multi-source data

Integrate data from multiple sources.



Funtions



Statistic

Comprehensive report on user behavior



Warning when detecting unusual behavior





Manage privacy

The system is privacy compliant

Visualization

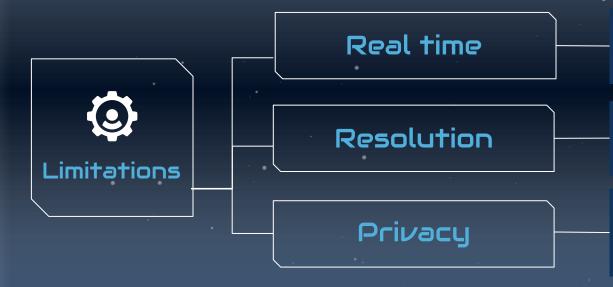
Display video combined with statistical charts







CONCLUSION



In real-time processing problems, there may be problems with delay in data transmission.

Data often faces challenges such as low lighting, and different viewing angles.

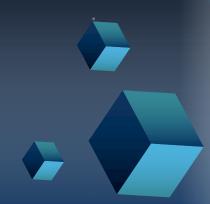
The use of data about human behavior raises issues of privacy and information security.

CONCLUSION



+++

THANKS!



I. Introduce

Overview of the problem

In the context of increasing competition and increasingly strong digital transformation momentum, businesses, especially grocery stores, face great pressure to improve efficiency and optimize business processes. The application of technology to reduce labor power and enhance productivity has become an indispensable factor to maintain and thrive in today's business environment. From the initial dataset, the organizers came up with several videos, each about 2 minutes long and recorded from a fixed overhead camera placed in the grocery store. In these videos, customers take five distinct actions: "Reach the shelf," "Pull out of the shelf," "Hand on the shelf," "Check the product," and "Check the shelf." These actions describe the different interactions that individuals typically make when shopping. The goal of the problem is to build an application that recognizes real-time behavior through the camera, displaying descriptive statistical information of customers to users. Users can observe through the application on the phone or computer connected directly to the camera. From there, give useful information to users.

Minimum Viable Product (MVP) and Its Purpose:

First, it is necessary to develop a system that can accurately identify and classify five specific actions performed by individuals in grocery store settings. Next, statistics on the frequency of actions (action type), other relevant information in the video and visualization of the results. From the above results, we can capture the trend of customer actions in the context of groceries. Review the distribution of actions in different areas of your store and identify outlier areas to find corrective causes (for example, inappropriate arrangements of items that result in multiple actions by customers). Or it can develop a model that classifies normal purchase behavior and abnormal behavior for cases where customers deviate from normal shopping behavior, supporting store control and security.

II. Problem statement

- Clearly identify the problem or challenge that the problem wants to solve

The problem: building an app that recognizes human behavior in a retail store environment. Exploit information from the problem results and draw solutions for businesses.

Describe weaknesses or inefficiencies related to the problem:

- Real-time processing capability: In the real-time transmission processing problem, there may be a delay problem in transmitting data to the device using the application, because the model problem is complex.
- Ambient changes: Data collected from retail environments often face challenges such as low lighting, blurring, and different viewing angles. This can reduce the accuracy of image recognition models.
- Security and Privacy: The collection and processing of data about human behavior raises privacy and information security issues. The team needs to ensure that its systems comply with data protection and privacy regulations.
- Describe the current context and positioning of competitors if applicable:

- E-commerce is growing: The explosion of e-commerce has created a strong competitive environment, placing high demands on the shopping experience and accompanying services.
- Growing interest in user experience: Detecting and understanding user behavior is an important factor for optimizing services and enhancing customer satisfaction.
- Technology integration: Enterprises are integrating advanced technology solutions such as artificial intelligence, machine learning, and natural language processing to provide intelligent and personalized services.

III. Solution Overview

Solution Overview:

The solution below proposes the use of artificial intelligence (AI) for the Minimum Viable Product (MVP), the goal of which is to revolutionize the management and analysis of customer interactions in the context of groceries. Using the knowledge of artificial intelligence and computer vision for the object detection problem, MVP focuses on accurately identifying and classifying five specific actions performed by the individual in the store: "Reach the shelf," "Withdraw from the shelf," "Hand to the shelf," "Product inspection," and "Check Shelf." From there, develop more accompanying proposals.

AI Techniques, Algorithms or Models:

1. Computer Vision for Object Detection:

- From a video, take the picture in the video in frames. After capturing the set of images separated from the video, label the 5 action types of the topic (do this by creating a separate detection pane for each action). From labeled photos, use the Yolov8 model for training. The result of the training was an initial video with a detect frame for 5 actions.
- It is proposed to add other action type labels such as: stooping, tiptoeing, kneeling, ... To be able to evaluate the areas of purchase, the location of items makes it difficult for customers to shop. From there, modify accordingly, minimize manipulation.

2. Statistical analysis and visualization:

 Visualization techniques are applied to present the results in an intuitive and understandable way. Evaluate the complexity of customer behavior, suggest new designs that are more suitable for the store.

3. Anomaly Detection and Classification Model:

- Support Vector Machine (SVM) or deep learning models can be used for binary classification of normal and abnormal buying behavior.
- To be able to implement this solution, video data needs to have unusual purchase behaviors (theft, on-site product use, product damage, ...).

Innovation and Uniqueness:

1. Real-time action recognition:

 Real-time action recognition allows businesses to understand customer behavior immediately. Contribute to proactive decision-making and reactive approach to store management.

2. Comprehensive analysis of the store:

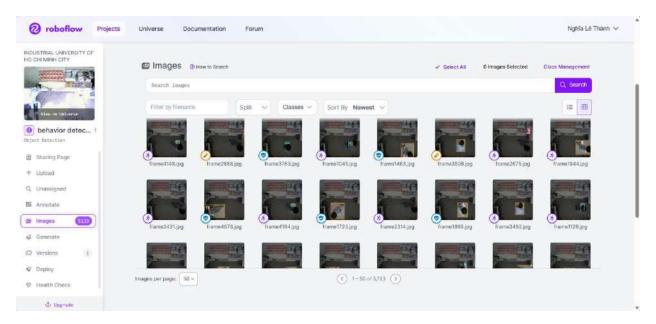
• The system is not only about recognizing behavior, but also analyzing their distribution in different areas of the store. Identifying outliers in specific areas helps identify potential problems.

3. Flexible security controls:

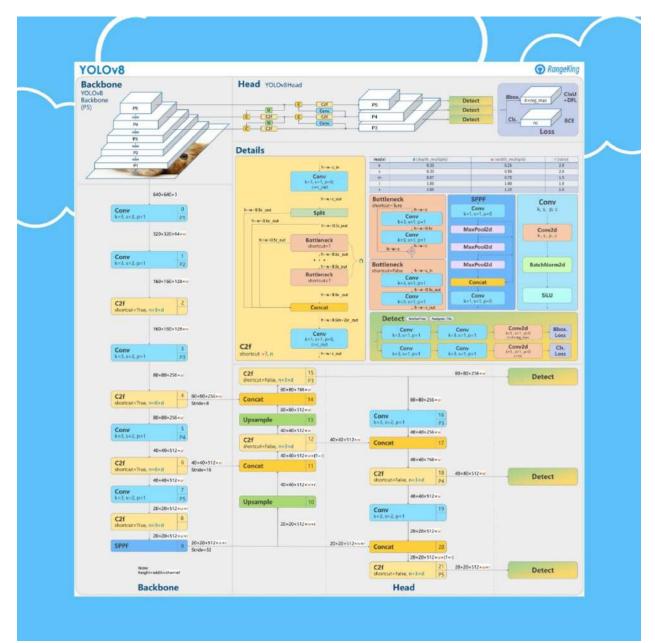
Ensure the system can react to abnormal operations, contributing to overall safety and loss prevention in the store.

IV. Methodologies

Dataset preparation method: the team will bining the box of actions/objects recognized by the AI model



The above dataset, the team will use YOLOv8. With components including Backbone and head. The architecture of the model consists of Conv and C2f blocks, in C2f blocks there are bottleneck and conv blocks. Overall, Convolutional Networks are the premise of the YOLOv8 model.

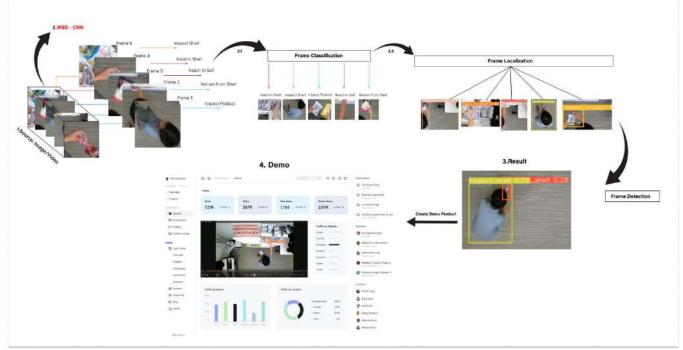


During training, YOLOv8 performs many techniques to enhance the training image. One of these techniques is mosaic data augmentation.

Mosaic data augmentation is a simple technique in which four different images are joined together and fed into the model as input. This helps the model learn real objects from different locations and in an obscured state.



In terms of technology, to mark the label, the team will use the RoboFlow framework, and the YOLOv8 model will be the Ultralytics framework.



PipeLine

V. Core Functionality

- Outline the main features and functions of the MVP:

Main features:

 Behavioral recognition: Categorizes and identifies actions such as "Reach to Self," "Retract From Shelf," "Hand in Shelf," "Inspect Product," and "Inspect Shelf."

- Real-time video processing: Live video processing to recognize and categorize user actions without causing major latency.
- Data from multiple sources: Integrate data from multiple sources such as cameras, motion sensors, and natural language data where available.
- Data visualization: Visualize results by displaying videos combined with timelines and statistical charts to help understand customer behavior.

Main functions:

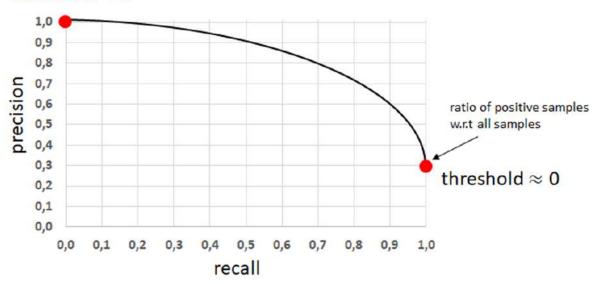
- Reports and statistics: Create consolidated reports on user behavior, frequency of actions, and descriptions of changes over time.
- Alerts and logging: Alerts when special behavior or exceptions are detected, and logs for each milestone.
- Privacy Management: Ensures that the system complies with user information privacy and security regulations.
- Configurable and scalable: Allows flexible configuration to customize to the specific needs of each specific retail environment and has the ability to scale as needed.

VI. Performance Metrics

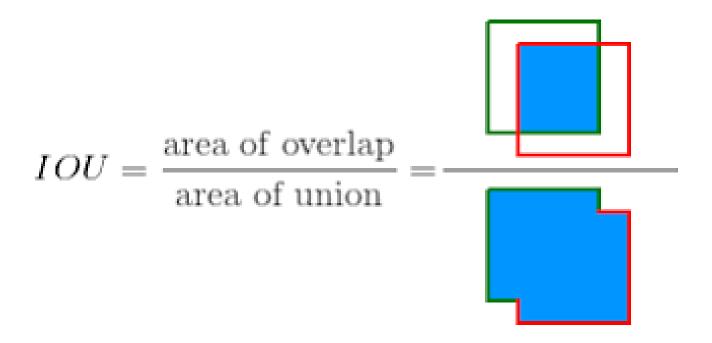
The main evaluation method in the Classification model will use the scale: mAP. The abbreviation for Mean Average Precision is the arithmetic average of different classes. So what is AP: here's the area under the Precision Recall Curve:

- AP is large if this area is large, inferring that the curve tends to be near the upper right corner and means that at different thresholds the Precision and Recall are both quite high. From this deduces the good model.
- Small APs, both Precision and Recall are quite low and the model is not good.

threshold ≈ 1



Threshold here will use IoU:



VII. Timeline and Roadmap

Roadmap and plan for MVP development and implementation:

Stage 1: Planning and choosing a topic

- Duration: 1 day
- Important milestones:
 - + Define the scope and goals.
 - + Create user stories and prioritize features.

Phase 2: Receive data and analyze it

- Duration: 2 days
- Important milestones:
 - + Read the data
 - + Idea proposal

Phase 3: Implementation

- Duration: 3 days
- Important milestones:
 - + Assign labels to each image frame
 - + Conduct model training

Phase 4: Deployment

- Duration: 1 day

- Important milestones:

VIII. User interface (UI) or Interaction (optional)

Our main interface will be a dashboard, the upper left position occupying 1/4 of the dashboard is a real-time video display frame from the camera to monitor and identify customer behavior. The remaining positions display a summary of the frequency of activity of each action type in the video, information about the trend of customer actions will be presented in the form of graphs and statistics. From that statistic, AI can make suggestions about unusual buying areas, from which we can reasonably adjust the position of items, arrange items accordingly, minimize behavioral manipulations that hinder customers. Or consider unusual behavior, loss or damage to items, and finally come up with an appropriate solution.

IX. Limitations and Future Enhancements (optional)

Limit:

- Data is limited, resulting in poor recognition.
- Performance and stability are not perfectly optimized, so small errors may appear when detecting actions in practice.
- Future improvements
- Expand the scope of data from a variety of customer actions to increase data diversity.
- Improve models for peak performance

X. Conclusion

In summary, in the context of the increasing shopping demand of users of businesses, stores tend to have the need to optimize business processes and enhance the shopping experience. The application of artificial intelligence to recognize customer behavior in the store has become important. The proposed Minimum Viable Product (MVP) uses computer vision and the Yolov8 model to accurately identify and classify five specific actions. This helps businesses track and understand customer behavior, thereby providing useful information to optimize your store and provide the best shopping experience. Key features of the system include the ability to recognize video processing real-time behaviors without causing large latency, integrating data from various sources, and visualizing results to help managers better understand customer behavior. Besides, functions such as reporting and statistics, alerting and logging, privacy management, flexible configuration, and scalability are all integrated to ensure flexibility and compatibility with multiple retail environments.