

# Deep learning model for abnormal behavior detection and attention behavior analysis in shopping centers using CCTV

Le Thi Ngoc Diem, Tran Tuyet Minh, Nguyen Thi Huyen Trang,  
Nguyen Thi Huynh Huong, Nguyen Nguyen Khuong

November 19, 2023

## 1 Introduction

In today's era, with the rapid development of technology and the robust growth of surveillance systems, the application of deep learning models in the field of abnormal behavior detection and attention behavior analysis on security camera systems (CCTV) has become a significant trend, especially in shopping centers.

The system for detecting abnormal behavior can enhance security and improve safety in public areas of shopping centers. By using deep learning models, the system can automatically learn to differentiate between normal human behaviors and simultaneously alert to signs of suspicious or potentially harmful actions. This allows businesses and security order management to take appropriate measures without adversely affecting customers. Additionally, this technology enhances the ability to respond quickly and effectively to emergency security situations in real-time.

Through the combination of deep learning and surveillance camera technology, the system for detecting abnormal behavior and analyzing attention behavior plays a crucial role not only in preventing threatening behavior but also in efficiently managing and optimizing security in shopping centers. Analyzing customer behavior also aids in directing the development of stores with the most accurate data, which we believe leads to high efficiency. Therefore, we aim to develop a System for Detecting Abnormal Behavior and Analyzing Customer Attention Behavior in shopping centers.

## 2 Problem Statement

However, the increased foot traffic in shopping centers has attracted the attention of criminals. Incidents such as thefts and even acts disrupting security, such as killings or terrorism, have occurred, posing a challenge for businesses in these

areas. This necessitates the implementation of preventive measures and monitoring to avoid potential adverse situations. The attack on Siam Paragon shopping center in Thailand on 3/10/2023 [3] highlights the emergence of criminal activities hidden within crowds. Importantly, authorities could only apprehend the troublemakers after unfortunate incidents occurred.

This concern is not only for governments but also for businesses and individuals. If similar situations persist, people will feel unsafe in these places, leading to a decline in revenue and impacting the business situation of these enterprises. This has prompted our team to work on building a machine-learning model capable of preventing malicious behavior before it occurs in crowded places. Additionally, targeting retail businesses as our customer base, our model should help them sell more products.

In this study, we focus on a specific service branch, which is the system for retail supermarkets nationwide. These supermarkets typically have two major needs when applying machine learning technology to their operations.

The first need is to improve security and order during customer shopping. They aim to eliminate all risks of theft, attacks, and even terrorism in the business area to provide customers with an excellent shopping experience. This is a significant challenge when the supermarkets welcome a high volume of people every day.

The second need is for businesses to gather data on the products that customers are interested in. By recording and analyzing customer shopping behavior, the business, marketing, and customer service departments can understand what customers need in terms of service, which products they like, and which products they are interested in but hesitant to purchase. Every business aims to maximize profits, and thus, these analyses can help expand revenue and fill their pockets.

## 3 Solution Overview

Building an advanced security system at the shopping center to monitor and safeguard the safety of customers and assets, while providing customer behavior analysis to optimize the shopping experience and ensure security.

### 3.1 Security Enhancement

#### 3.1.1 Utilizing Models

Object Detection: Apply the RetinaNet model for weapon detection to identify if anyone is carrying a weapon. (RetinaNet Model: RetinaNet is a model capable of handling small and important objects.) Detecting Violent Behavior: Utilize the OpenPose model in conjunction with the smart city CCTV violence detection dataset to identify violent behavior. (OpenPose Model: OpenPose can be used to recognize basic gestures.) Combining Models: Combine the results from the above models to determine overall behavior. Establish alert rules to trigger the system when threatening behavior is detected. Build an alert system

based on rules to trigger alerts when the level of threat is identified. Real-Time Alert System and Monitoring: When the system detects abnormal behavior, trigger alerts immediately so that security management can respond promptly. Use a real-time monitoring system to track and confirm events, facilitating quick responses.

## 3.2 Customer Behavior Analysis

### 3.2.1 Define Analysis Objectives

The goal is to predict the behavior of individuals in videos based on information about the location and posture of body parts (head, shoulders, arms, legs). Use OpenPose to extract features: Input includes video clips or frames from videos captured by security cameras, and OpenPose will extract information about the positions of body landmarks such as the head, shoulders, arms, and legs. Use labeled data: Input data consists of previously labeled videos where each frame indicates the behavior of the person. OpenPose will be used to extract features from these frames. Then, labeled data will be used to train a model to predict human behavior.

### 3.2.2 Data Preparation

**1. Weapon Detection** In public places such as shopping malls, carrying weapons is prohibited. However, at present, there are not effective measures to control customers' weapons in shopping malls. Detecting weapons can contribute to preventing unfortunate incidents before they occur.

We propose using the Weapon Detection Dataset[2] for real-time weapon detection. After a preliminary evaluation, we found that this dataset can be used for weapon detection on CCTV. Given recent cases of gun usage in criminal activities in Vietnam, detecting guns is crucial. Moreover, the dataset from the competition involves foreign CCTV.

**2. Detecting Abnormal Behavior:** Security in shopping malls needs to be detected and controlled quickly. We propose detecting abnormal behaviors such as violence, theft, robbery, etc., and label them as unusual. Using a single label makes labeling lighter and facilitates model development.

We also propose using two additional datasets: Smart-City CCTV Violence Detection Dataset (SCVD) [4] for detecting violent behavior and UCF Crime Dataset [1], which includes criminal behaviors such as robbery and theft.

**2. Attention Behavior Analyze:** In the retail industry, understanding and accurately responding to consumer behavior is crucial for business success. This includes shopping decisions, interactions, and customer choices when buying products and using services. Understanding and leveraging information about consumer behavior are not only keys to effective business strategies but also ways to create a unique shopping experience, attracting attention and trust.

For analyzing whether a store receives much attention from customers by evaluating customer behavior when passing by storefronts, the degree of attraction can be assessed through the ratio of time spent looking at the store, whether customers enter after looking, etc. This information can reveal factors related to space, time, location, etc., and help formulate the best business strategies.

We suggest using the Customer Behavior dataset from the competition, as it already contains suitable data for analyzing customer behavior, particularly attention behavior.

### 3.2.3 Building Model

Use a neural network model to learn from the data. Use Convolutional Neural Network (CNN) layers to process image data and Fully Connected layers to predict behavior. Innovation Aspect of the Two Systems:

The innovation of these two customer behavior analysis and security systems not only lies in the use of advanced models but also demonstrates flexible integration and rapid response to the increasingly diverse and complex needs of businesses.

**Technical Aspect:** Both systems use state-of-the-art detection and classification models in deep learning and computer vision. The combination of YOLO and RetinaNet for object detection not only provides accurate detection but also ensures high-performance processing, especially for small and important objects like weapons. The OpenPose model for behavior detection provides a unique capability in understanding basic gestures and violent behavior, expanding the application potential of the system. The innovation is also evident in the use of diverse and rich datasets from sources such as Kaggle. This not only helps model real-world challenges but also promotes creativity in model building, setting high standards for accuracy and stability.

**Business Demand Response:** These customer behavior analysis and security systems not only focus on identifying risks but are tightly integrated to meet the specific needs of businesses. The ability to combine information from multiple models helps comprehensively identify overall behavior, from detecting theft to predicting who is carrying a weapon.

The real-time alert and monitoring system establishes a robust infrastructure for security management. When abnormal behavior is detected, the system triggers alerts immediately, enabling timely and effective security management response.

In summary, the innovation of these two systems is not only in applying advanced technology but also in their flexibility, information integration, and quick response to the diverse requirements of today's business environment.

## 4 Methodologies

The application is built with AI technology to modernize the fields of security and retail data analysis by utilizing advanced technology in person and action recognition. The system architecture includes advanced deep learning models, with Convolutional Neural Networks (CNN) and OpenPose for accurate person identification and subsequent analysis of human actions.

Core values include real-time video processing modules, using computer vision techniques to detect and classify various actions. The system's core functionality focuses on the ability to identify potential crimes or abnormal activities in crowded areas, significantly enhancing security monitoring. Additionally, the application excels in analyzing customer behavior, using sophisticated algorithms to identify actions such as taking products from shelves.

The integrated technology stack includes leading image processing libraries and frameworks, ensuring the powerful performance of the application. This MVP aims to provide a smooth experience in crime prevention and retail optimization through strategic insights from artificial intelligence.

## 5 Core Functionality

The Minimum Viable Product (MVP) of this application features advanced capabilities at the intersection of security and retail data analysis. Firstly, the application excels in real-time person recognition, using Convolutional Neural Networks (CNN) and OpenPose to accurately identify individuals in a specific space. It not only limits itself to identification but can also analyze and classify human actions, aiding in the detection of potential crimes or unusual activities.

In the retail analysis domain, the MVP demonstrates strong capabilities in analyzing customer behavior. By employing sophisticated algorithms, the application can identify and track customer actions, such as taking products from shelves. This functionality not only contributes to enhancing security but also provides valuable insights for optimizing retail operations and improving customer service.

The MVP also allows users to access real-time information and alerts generated by artificial intelligence models. Additionally, the MVP supports easy integration with existing security and retail systems, ensuring compatibility and scalability.

Overall, the key features of the MVP include real-time person recognition, action analysis, crime prevention, customer behavior tracking, and seamless integration, making it a comprehensive solution to improve security and optimize retail processes.

## 6 Timeline and Roadmap

The timeline and roadmap are presented in Table 1.

Table 1: Project Milestones and Activities

<b>Milestone</b>	<b>Activity</b>	<b>Detailed Content</b>
Pre-processing	Build dataset	Collect data from various sources
Staging	Build models	Research deployed models, note pros/cons Develop high-accuracy machine learning models Document decision on selected model Source code of machine learning model
Development	Product testing	Test the product at several convenience stores Compare theoretical and actual results
Production	Deploy product in production environment	Deployment guide  Ensure system security and stability

## 7 Conclusion

Our team has proposed a model with significant potential for development. For the consumer behavior recognition feature, this system can be applied across all retail stores, including supermarkets, convenience stores, and shopping malls. Wherever businesses want to expand revenue and increase wallet share, this system will become a valuable supporting tool. For the dangerous behavior detection feature, the system’s scope of operation is even broader. Wherever there are crowds, there is a need to detect and prevent dangerous behavior early. At that point, our proposed system becomes a weapon against violent behavior to protect the community.

## References

- [1] SANSKAR HASIJA. Ucf crime dataset. <https://www.kaggle.com/datasets/odins0n/ucf-crime-dataset/data>, 2021.
- [2] Andalusian Research Institute in Data Science and Computational Intelligence. Weapon detection dataset. [https://github.com/ari-dasci?fbclid=IwAR3gprZh690EHT\\_YA\\_qL-vyQ54sqMq0R7HWEWP8dkRWpCPaDerVlebQwJv8](https://github.com/ari-dasci?fbclid=IwAR3gprZh690EHT_YA_qL-vyQ54sqMq0R7HWEWP8dkRWpCPaDerVlebQwJv8), 2021.
- [3] Jonathan Head, Thanyarat Dokson, and Tossapol Chaisamritpol. Bangkok: Two dead and 14-year-old held over siam paragon mall shooting. <https://www.bbc.com/news/world-asia-66994274>, 2023.

- [4] Li Zhiyuan, Reem Alameeri, and Toluwani Aremu. Smart-city cctv violence detection dataset (scvd). [https://www.kaggle.com/datasets/toluwaniaremu/smartcity-cctv-violence-detection-dataset-scvd?fbclid=IwAR25yV80I19x5kb29Ck\\_fntXrzmABejFy9l7th4y-6ieM534XZskBUx8fms](https://www.kaggle.com/datasets/toluwaniaremu/smartcity-cctv-violence-detection-dataset-scvd?fbclid=IwAR25yV80I19x5kb29Ck_fntXrzmABejFy9l7th4y-6ieM534XZskBUx8fms), 2023.