International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 06 | June - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

DR ASTAN: Data Report AI-Powered Shopper Tracking & Analysis Network

Meera Sawalkar, Gaurav Pandit, Prajwal Nivangune, Aryan Tikle, Anushka Yeole, Hemani Wankhedkar

[1]*Meera Sawalkar, Computer Department, JSPM NTC*
[2]*Gaurav Pandit, Computer Department, JSPM NTC*
[3]*Prajwal Nivangune, Computer Department, JSPM NTC*
[4]*Aryan Tikle, Computer Department, JSPM NTC*
[5]*Anushka Yeole, Computer Department, JSPM NTC*
[6]*Hemani Wankhedkar, Computer Department, JSPM NTC*

---***---

Abstract - This research paper introduces **Dr. ASTAN**, an advanced system for monitoring, analyzing, and visualizing shopper behavior in retail environments using state-of-the-art computer vision and deep learning techniques. The system integrates multi-camera inputs and leverages **YOLOv8** for real-time object detection, coupled with **DeepSORT** and the **Kalman Filter** for accurate multi-object tracking across predefined store zones. Behavioral data is processed using **Pandas** and visualized through **Power BI**, generating actionable insights into foot traffic patterns, zone-wise engagement, and consumer movement trajectories. **Dr. ASTAN** enables data-driven decision-making to optimize store layouts, improve customer experience, and increase sales, all while adhering to strict privacy standards and ethical data handling practices.

Key Words: DR. ASTAN, deep learning, YOLOv8, DeepSORT, Kalman Filter, consumer behavior, retail analytics, computer vision, zone-based analysis

1.INTRODUCTION

In today's competitive retail landscape, customer experience is one of the most crucial differentiators that drive loyalty and sales. Physical retail stores are constantly seeking innovative ways to understand how customers interact with their spaces—from the paths they take to the products they engage with. These insights enable data-driven decisions on store layouts, product placements, marketing campaigns, and staffing strategies. Traditional techniques such as surveys, manual observation, RFID tags, and beacon-based tracking provide limited insight due to their intrusiveness, requirement for customer cooperation, or inability to offer real-time granular data.

To address these limitations, the integration of computer vision (CV) and deep learning has opened up new possibilities for passive, accurate, and real-time tracking of customer behavior. Unlike older systems that require users to carry devices or wear tags, vision-based systems can capture anonymous behavioral data through strategically placed cameras, enabling continuous monitoring without disrupting the customer experience. However, implementing such systems comes with challenges, including identity switching, occlusions, crowded scenes, and scalability across large retail environments.

In response to these challenges, this paper presents **Dr. ASTAN (Advanced Shoppers Tracker and Analysis Network)**, an AI-powered system that leverages the capabilities of **YOLOv5mu** for efficient object detection and **DeepSORT** with Kalman filtering for robust multi-object tracking across predefined store zones. Dr. ASTAN is designed to monitor, analyze, and visualize shopper behavior in real-time using multiple video feeds installed in key sections of the store—namely, the Entry Zone, Main Shopping Zone, and Checkout Zone.

What differentiates Dr. ASTAN from existing systems is its **zone-based analytical approach**, which segments the store into logical regions to capture more detailed interactions and flow patterns. The system measures key behavioral metrics such as dwell time, transition frequency, and zone engagement. To capture the temporal dynamics of customer movement, **Recurrent Neural Networks (RNNs)** are employed for modeling sequential patterns in foot traffic and behavioral trends over time. This enables the system to forecast visitor flows, peak hours, and duration trends with higher accuracy—supporting predictive decision-making in store management and marketing.

Processed insights are managed using the **Pandas** data analysis library and visualized via dynamic dashboards built in **Power BI**. Outputs include heatmaps, traffic graphs, and temporal movement trajectories, offering store managers actionable insights to optimize layouts, improve staff allocation, and boost customer satisfaction.

To further enhance usability, Dr. ASTAN integrates an **interactive chatbot interface** that allows store personnel to query the system in natural language. The chatbot can respond to questions like *"How many visitors entered the main shopping zone today?"*, *"What was the busiest hour yesterday?"*, or *"How long did most people stay in the checkout zone?"*. By mapping these queries to pre-analyzed metrics, the chatbot makes insights accessible even to non-technical users, empowering data-driven actions without the need to navigate complex dashboards.

Moreover, Dr. ASTAN is built with privacy and ethical data handling at its core. It does not capture or store personally identifiable information (PII), adhering to global data protection regulations such as the General Data Protection Regulation (GDPR). All tracking is performed using anonymized data, ensuring that individual privacy is respected while still enabling valuable behavioral insights.

In summary, Dr. ASTAN aims to revolutionize the way retail analytics is performed by providing an end-to-end, real-time, and ethical system for customer behavior analysis. By combining deep learning, recurrent neural networks for sequential understanding, and an intuitive chatbot interface, it represents a significant advancement in applying AI and computer vision to practical, high-impact business problems in the retail sector.

2. Body of Paper

2.1 Literature Review

Deep Learning for Customer Behavior Analysis

Senarath et al. (2022) explored customer gaze estimation using deep learning techniques, demonstrating how eye-tracking can enhance the understanding of shopper engagement and attention in retail settings [1].

Complementing this, Quyen et al. (2023) proposed a hybrid approach combining shape fitting and heatmap regression for accurate facial landmark detection, which can be applied to improve person re-identification accuracy in multi-camera systems [2]. These techniques are fundamental for enhancing precision in identifying and tracking customers.

Multi-Camera Tracking and Heatmap Analytics

Narvilas et al. (2022) studied human behavior tracking in stores through multiple security cameras, emphasizing the importance of robust multi-view data fusion to mitigate challenges like occlusion and crowd density [3].

Kajabad and Ivanov (2019) combined movement detection with deep learning to identify high-interest store areas, effectively using heatmap analytics to highlight customer attraction zones [4].

Similarly, Siam and Biswas (2022) introduced a multi-camera deep learning-based detection and heatmap generation system for visualizing customer movement patterns [9].

Simşek and Tekbaş (2024) further customized the YOLO-DeepSORT framework specifically for in-store customer behavior analysis, producing detailed heatmaps to support retail decision-making [10].

Advanced Tracking Algorithms and Pose Estimation

Durve et al. (2023) benchmarked YOLOv5 and YOLOv7 models integrated with DeepSORT for high-accuracy object tracking, underscoring the relevance of this combination in crowded and complex environments similar to retail floors [5].

Maji et al. (2022) enhanced multi-person pose estimation by incorporating keypoint similarity loss in YOLO, facilitating more detailed behavioral analysis in dense crowds [6].

Mendes et al. (2023) focused on multi-camera person re-identification through trajectory data analysis, a critical component for maintaining identity consistency across different views in large retail settings [7].

Zhang et al. (2020) combined YOLO with CSRT trackers to tackle occlusion and clutter challenges, laying groundwork for robust tracking systems such as Dr. ASTAN [11].

Nasseri et al. (2021) improved tracking performance by using geometric cues to manage occlusions and identity switching, enhancing scalability in dense environments [12].

Sensor Fusion and Indoor Positioning

Huang et al. (2021) explored the fusion of motion sensors with camera feeds to boost tracking accuracy, particularly under occlusion, pointing towards hybrid approaches for more reliable retail analytics [13].

Ding et al. (2020) detailed common challenges in vision-based customer tracking such as identity switching and crowd occlusion, motivating the integration of DeepSORT and Kalman filters to maintain robustness in such scenarios [14].

From the indoor positioning perspective, Mackey et al. (2020) demonstrated how Bayesian filtering techniques could improve BLE beacon proximity estimation, highlighting the use of probabilistic models to achieve greater tracking precision with affordable hardware [15].

Reviews by Huang et al. (2023) and Ashraf et al. (2022) surveyed advances in indoor localization, emphasizing sensor fusion methodologies combining Wi-Fi, IMUs, and cameras to enhance accuracy and reliability in dynamic indoor environments, which is highly relevant for comprehensive retail tracking [16,17].

Big Data and Consumer Segmentation

Ehsani and Hosseini (2023) applied big data analytics for consumer segmentation based on location and timing, enabling personalized marketing and operational strategies in retail marketplaces [8].

Their work underscores the value of integrating spatial-temporal data for actionable insights, complementing visual tracking technologies.

2.2 Problem Statement

In today's competitive retail landscape, understanding customer behavior is essential for optimizing layouts, marketing, and customer experience. Traditional observation and survey methods are slow, error-prone, and lack real-time feedback. Retailers require efficient, accurate, and automated solutions to track and analyze customer movements for data-driven decision-making. Dr. ASTAN meets this need by employing deep learning and computer vision to deliver real-time, actionable insights into shopper behavior, supporting better store management and marketing optimization

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 06 | June - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

2.3 Methodology

The development of the Dr. ASTAN system followed a multi-stage pipeline that integrates computer vision, deep learning, data analytics, and natural language interaction components. The methodology is divided into the following key modules:

2.3.1 Data Acquisition and Preprocessing

- **Video Input:** Video feeds were collected from CCTV cameras strategically placed in three key zones: Entry, Main Shopping Area, and Checkout.
- **Frame Sampling**: Real-time frames were extracted at consistent intervals (e.g., 15 FPS) for object detection and tracking.
- **Privacy Handling**: All videos were anonymized and downsampled to ensure compliance with GDPR and ethical guidelines.

2.3.2 Object Detection and Tracking

- **Detection:** YOLOv5mu (a lightweight and fast variant) was used to detect human figures in each frame due to its high speed and accuracy in real-time applications.
- **Tracking:** Detected humans were tracked across frames using DeepSORT with Kalman Filtering. This ensured consistency of identity (track ID) even in occlusion or dense crowds.
- **Zone Mapping:** The store layout was digitally divided into predefined zones. A zone-detection logic was applied using coordinate thresholds to identify a person's presence within each zone.

2.3.3 Behavior Analysis and RNN Processing

- **Feature Extraction:** For each tracked individual, features such as dwell time, transition paths, zone entry/exit counts, and movement speed were logged.
- **Sequence Modeling with RNN:** To understand temporal patterns in user behavior (e.g., path prediction or repetitive habits), a Recurrent Neural Network (RNN) was trained on sequences of zone transitions per customer.
- **Pattern Detection**: RNN outputs were used to classify visitor types (e.g., quick shopper vs. explorer) and forecast possible zone visits for real-time insights.

2.3.4 Data Aggregation and Visualization

- **Data Logging:** Zone-wise analytics were computed using the Pandas library and stored in CSV or SQL databases.
- **Visualization:** Key insights like heatmaps, transition graphs, and visitor flow were displayed using Power BI dashboards for store managers.
- **Metrics Tracked:** Metrics included zone dwell time, total visits, transition matrices, and heatmap accumulations.

2.3.5 Chatbot Integration Using spaCy

- **Natural Language Understanding:** The chatbot uses the spaCy NLP library to process and interpret user queries. spaCy enables:
  - Named Entity Recognition (NER) for extracting entities like timeframes, zones, and visitor counts.
  - Dependency parsing to identify the intent of user queries (e.g., questions about "how many," "when," or "where").
- **Query Mapping:** Parsed queries are mapped to corresponding backend functions using custom rule-based patterns and intent classification.
- Supported Queries:
  - *"How many visitors entered today?"*
  - *"What was the foot traffic in the checkout zone yesterday?"*
  - *"Show the busiest hour of the day."*
- **Response Generation:** The chatbot returns visual or textual summaries based on filtered analytics from the Pandas-processed data or Power BI outputs.
- **Context Handling:** spaCy's tokenization and vector representations help maintain conversational flow for follow-up queries (e.g., *"And what about the main zone?"*).

2.3.6 System Deployment

- Real-Time Capability: The system was deployed using a Flask or FastAPI backend to handle real-time video stream processing.
- Multithreading: Used for efficient video frame handling and inference.
- Scalability: The architecture supports horizontal scaling for larger retail environments by parallelizing camera stream inputs.

2.4 Flow Chart of the Proposed Model

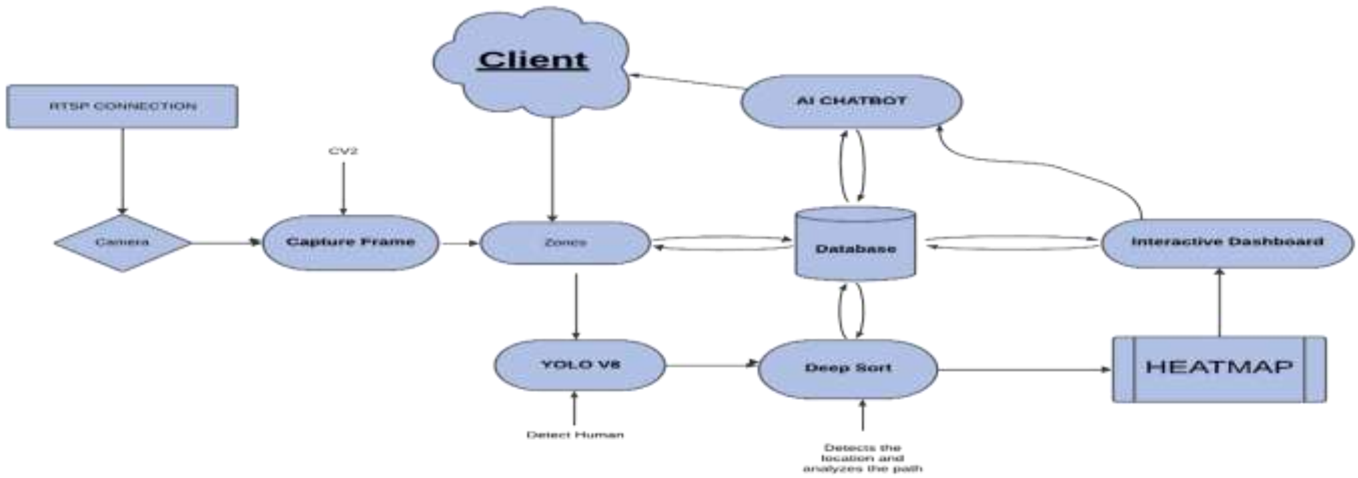


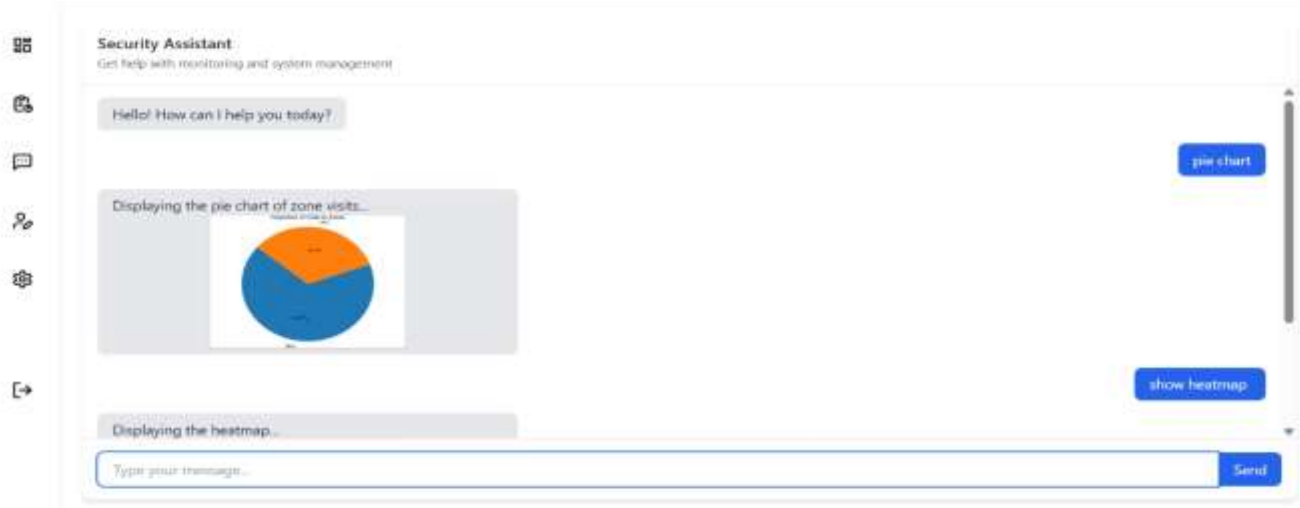
Figure 1: Flow Chart of Proposed Model

2.5 Results

Figure 2: Chatbot Integration

Figure 2: MAIN DASHBOARD

Figure 3: Heatmap Genration



2.6 Privacy Considerations in Video-Based Customer Analysis

This section presents the design of the customer behavior analysis system with a strong emphasis on **privacy protection** as a core component. To prevent the collection of **personally identifiable information (PII)**, the system implements stringent privacy protocols throughout the video capture and processing pipeline.
The video analytics architecture is powered by YOLOv8 and **DeepSORT**, which are configured specifically to detect and track **general human movement patterns**. These models are fine-tuned to exclude facial features and other biometric identifiers, ensuring that the data collected remains anonymous. Only **trajectory data**—which describes the paths and behaviors of individuals in aggregate form—is extracted and analyzed.
To further reinforce privacy, all video recordings undergo immediate **anonymization**. The system retains only non-identifiable spatial information, which is then used for **heatmap generation**. These heatmaps visually represent customer activity in the retail environment, with warmer colors indicating high-engagement zones and cooler colors highlighting low-traffic areas.
In compliance with international data protection standards such as the **General Data Protection Regulation (GDPR)** and similar legislations, the system ensures that no personal or sensitive data is stored or processed beyond what is strictly necessary for movement analysis. This non-intrusive approach allows businesses to gain actionable insights into store layout effectiveness and product placement, without compromising customer privacy.
By adhering to a privacy-by-design philosophy, the system demonstrates how ethical data practices can coexist with effective video-based customer behavior analytics.

3. CONCLUSIONS

Dr. ASTAN offers a novel, AI-driven solution for real-time retail analytics by combining advanced computer vision techniques with RNN-based temporal analysis and an interactive chatbot interface. Utilizing YOLOv5mu and DeepSORT for accurate multi-object tracking across store zones, the system provides detailed insights on customer behavior such as foot traffic, dwell time, and zone transitions, visualized through dynamic dashboards. The chatbot enables intuitive, natural language queries, making data accessible to store managers without technical expertise. Designed with strong privacy measures to ensure GDPR compliance, Dr. ASTAN represents a significant step forward in ethical, actionable, and user-friendly retail intelligence, helping businesses optimize operations and enhance customer experiences.

REFERENCES

1. Senarath, S., Pathirana, P., Meedeniya, D., & Jayarathna, S. (2022). Customer Gaze Estimation in Retail Using Deep Learning. *IEEE Access*, 10, 64904–64919.

2. Quyen, N.T.N., Linh, T.D., Phuc, V.H., & Nam, N.V. (2023). SHELF: Combination of Shape Fitting and Heatmap Regression for Landmark Detection in Human Face. *EAI Endorsed Transactions on Industrial Networks and Intelligent Systems*, 10, e3.

3. Narvilas, G., Urbonas, V., & Butkevičiūtė, E. (2022). Human's Behavior Tracking in a Store Using Multiple Security Cameras. *Baltic Journal of Modern Computing*, 10, 526–544.

4. Kajabad, E.N., & Ivanov, S.V. (2019). People Detection and Finding Attractive Areas by the Use of Movement Detection Analysis and Deep Learning Approach. *Procedia Computer Science*, 156, 327–337.

5. Durve, M., Orsini, S., Tiribocchi, A., Montessori, A., Tucny, J.-M., Lauricella, M., Camposeo, A., Pisignano, D., & Succi, S. (2023). Benchmarking YOLOv5 and YOLOv7 Models with DeepSORT for Droplet Tracking Applications. *European Physical Journal E*, 46, 32.

6. Maji, D., Nagori, S., Mathew, M., & Poddar, D. (2022). YOLO-Pose: Enhancing YOLO for Multi Person Pose

Estimation Using Object Keypoint Similarity Loss. In *Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops* (pp. 2636–2645).

7. Mendes, D., Correia, S., Jorge, P., Brandão, T., Arriaga, P., & Nunes, L. (2023). Multi-Camera Person Re-Identification Based on Trajectory Data. *Applied Sciences*, 13, 11578.

8. Ehsani, F., & Hosseini, M. (2023). Consumer Segmentation Based on Location and Timing Dimensions Using Big Data from Business-to-Customer Retailing Marketplaces. *Big Data*, 11, 5.

9. Siam, M.S.I., & Biswas, S. (2022). A Deep Learning Based Person Detection and Heatmap Generation Technique with a Multi-Camera System. In *Proceedings of the 12th International Conference on Electrical and Computer Engineering* (pp. 260–263).

10. Simşek, M., & Tekbaş, M.K. (2024). Heatmap Creation with YOLO-Deep SORT System Customized for In-Store Customer Behavior Analysis. *Communications Faculty of Science University of Ankara*, 66, 118–131.

11. Zhang, Y., et al. (2020). Integrating YOLO and CSRT for Enhanced Multi-Person Tracking in Crowded Retail Environments.

12. Nasseri, M.H., Moradi, H., Hosseini, R., & Babaee, M. (2021). Simple Online and Real-Time Tracking with Occlusion Handling.

13. Huang, L., et al. (2021). Combining Motion Sensors and Camera Feeds for Enhanced Retail Analytics.

14. Ding, J., et al. (2020). Challenges in Computer Vision-Based Customer Tracking: Occlusion and Crowding.

15. Mackey, A., Spachos, P., Song, L., & Plataniotis, K. (2020). Improving BLE Beacon Proximity Estimation Accuracy through Bayesian Filtering.

16. Huang, J., Junginger, S., Liu, H., & Thurow, K. (2023). Indoor Positioning Systems of Mobile Robots: A Review. *Robotics*.

17. Ashraf, I., Hur, S., & Park, Y. (2022). Recent Advancements in Indoor Positioning and Localization. *Electronics*.