World Journal of Information Technology
Print ISSN: 2959-9903
Online ISSN: 2959-9911
DOI: https://doi.org/10.61784/wjit3110

METHODOLOGY FOR SHELF PRODUCT RECOGNITION AND INTELLIGENT REPLENISHMENT DECISION-MAKING USING YOLOV11

Xuan Gao
Business College, Southwest University, Rongchang 402460, Chongqing, China.

Abstract: To address the issues of delayed inventory visibility and experience-driven retail shelf management, this paper proposes an intelligent replenishment decision-making framework that integrates computer vision and management science. First, a model for shelf item recognition and quantity counting is constructed using You Only Look Once version 11 (YOLOv11), enabling real-time inventory sensing. On this basis, inventory control theory is incorporated to establish a dynamic safety stock warning mechanism. Combined with exponential smoothing for demand forecasting, the dynamic replenishment quantity is calculated as "Replenishment Quantity = Forecasted Demand – Current Inventory + Safety Stock." Using a self-built shelf image dataset, this paper compares YOLOv11 models of various scales, among which the m model achieves the optimal balance between computational efficiency and performance, with an average product detection accuracy (mAP50) of 60.2%; Compared to other mainstream detection models (Faster R-CNN, RetinaNet, YOLOv8m), YOLOv11m achieved the highest accuracy with the fewest parameters, validating the reliability of the visual perception module. Theoretical analysis, case simulations, and ablation experiments demonstrate that the proposed dynamic replenishment model can more effectively address demand fluctuations and reduce the risks of stockouts and inventory buildup compared to fixed-threshold strategies and minimum-maximum inventory strategies; furthermore, safety stock and demand forecasting are complementary and indispensable. This study provides a feasible solution and theoretical reference for visual inventory management and scientific replenishment decision-making in smart retail.
Keywords: Shelf product recognition; Intelligent replenishment decisions; Object detection and counting; Demand forecasting; Inventory management

1 INTRODUCTION

Against the backdrop of the rapid development of smart retail and new retail, real-time monitoring of shelf inventory and scientific replenishment decisions have become core components of retail operations management. Stockouts directly lead to lost sales and customer churn, while excessive inventory ties up capital and increases costs. Studies indicate that the average stockout rate for retailers is around 8%, and approximately half of these stockouts resulting from untimely shelf replenishment [1]. Traditional shelf management relies on manual inventory counts and experience-based restocking, which suffer from poor timeliness, high error rates, and high costs [2]. Computer vision technology offers a new approach to achieving real-time inventory monitoring. By integrating visual recognition with inventory control and demand forecasting theories, it establishes a closed-loop system spanning from perception to decision-making.This holds significant theoretical value and practical significance [3].

Over the past decade or so, computer vision-based shelf item detection and recognition technologies have made significant strides. These methods can be broadly categorized into those based on traditional image processing, those based on object detection, and integrated approaches that incorporate shelf structure analysis. Early traditional methods often employed manually extracted features combined with classifiers. Varol et al. achieved item detection and classification with low time complexity by combining shape and color information [4]. Sun utilized projection histograms to segment shelf levels [5], extracted product color, shape, and position features, and performed product recognition using an SVM classifier. However, such methods exhibit poor robustness to lighting variations, occlusions, and densely stacked scenes, making them unsuitable for real-world retail environments. With the advancement of deep learning, object detection-based shelf product recognition has become the mainstream approach. Goldman et al. enhanced RetinaNet by adding a Soft IoU layer and an EM Merger module [6], significantly improving detection accuracy on a dense scene dataset containing 110,000 SKUs. Guo et al. proposed a lightweight detection algorithm based on YOLOv11 [7]; by refining modules and introducing an attention mechanism, they achieved a 0.6% improvement in mAP50 on a self-built dataset. Chen et al. proposed the RetailDet architecture [8], which integrates RGB and depth information to jointly identify products and empty spaces. Additionally, some researchers have specifically analyzed shelf-level structures. The ShelfManagement system developed by Pietrini et al. effectively achieved precise identification of shelf edges using the Deep Hough transform [9]. Alghaslan et al. proposed the RetailEye method, which employs supervised contrastive learning and compliance matching techniques to address product misplacement and out-of-stock issues on retail shelves [10].

The aforementioned studies have made significant progress in product detection, shelf structure analysis, and out-of-stock identification; however, they still have limitations: First, the vast majority of these studies stop at detecting products or identifying product categories, failing to translate visual recognition results into inventory management metrics (such as current inventory levels, safety stock thresholds, and replenishment quantity calculations). Even studies specifically focused on out-of-stock detection are limited to binary determinations of whether an item is out of stock, lacking quantitative analysis of the severity of stockouts and the required replenishment quantities. Second, existing methods generally treat replenishment as an independent manual rule or a simple fixed-threshold strategy, lacking systematic integration with demand forecasting models. Replenishment decisions should not only rely on current inventory but also incorporate historical sales trends to forecast future demand, achieving an optimal balance between inventory turnover and service levels. Third, from the perspective of management science and engineering, there is currently a lack of a closed-loop decision-making framework that integrates visual perception, inventory assessment, demand forecasting, and replenishment quantity calculation. Therefore, how to utilize object detection to obtain real-time inventory data and embed it into classical inventory control theory models is a research direction worthy of in-depth exploration.

To address the aforementioned issues, this paper proposes a framework for intelligent shelf product recognition and replenishment decision-making that integrates computer vision and management science. The overall approach is as follows: First, the YOLOv11 object detection algorithm is used to perform real-time recognition and quantity counting of products in shelf images. Then, based on the recognition results, the current inventory level is calculated and compared with the dynamic safety stock threshold to determine whether a replenishment alert should be triggered. Finally, by integrating a demand forecasting model, a formula for calculating dynamic replenishment quantities is derived, forming a complete closed-loop system from visual perception to replenishment decision-making. The main contributions of this paper are as follows:

(1) A YOLOv11-based method for shelf product recognition and quantity counting: a product detection dataset for multi-scenario shelf images was constructed, and a YOLOv11 model was trained to achieve automatic recognition and quantity counting of different product categories. Experiments validate the model’s detection accuracy and counting reliability in shelf scenarios, providing a data foundation for inventory sensing.

(2) An inventory status assessment and early warning mechanism based on visual recognition results: We establish mapping rules from detected product quantities to current inventory levels, design a method for determining dynamic safety stock thresholds, and define early warning trigger conditions, thereby achieving a transition from passive inventory counting to active sensing.

(3) A dynamic replenishment decision model that integrates demand forecasting and safety stock, using the core formula “Replenishment Quantity = Forecasted Demand − Current Inventory + Safety Stock” to construct a dynamic replenishment quantity calculation model. Through theoretical analysis and comparison, this model’s advantages over fixed-threshold replenishment strategies in reducing the risk of stockouts and inventory buildup are elucidated, forming a closed-loop management framework.

2 METHODS

The overall framework of the intelligent shelf product recognition and replenishment decision-making method proposed in this paper is shown in Figure 1. The entire system consists of two core modules connected in series: the product detection and counting module and the replenishment decision-making module. The product detection and counting module utilizes the YOLOv11 object detection algorithm to perform real-time identification and quantity counting of products in the input shelf images, outputting the current inventory levels for each product category. The replenishment decision module calculates the recommended replenishment quantity based on the detected current inventory, along with dynamic safety stock thresholds and demand forecasting models, thus forming a closed-loop process.

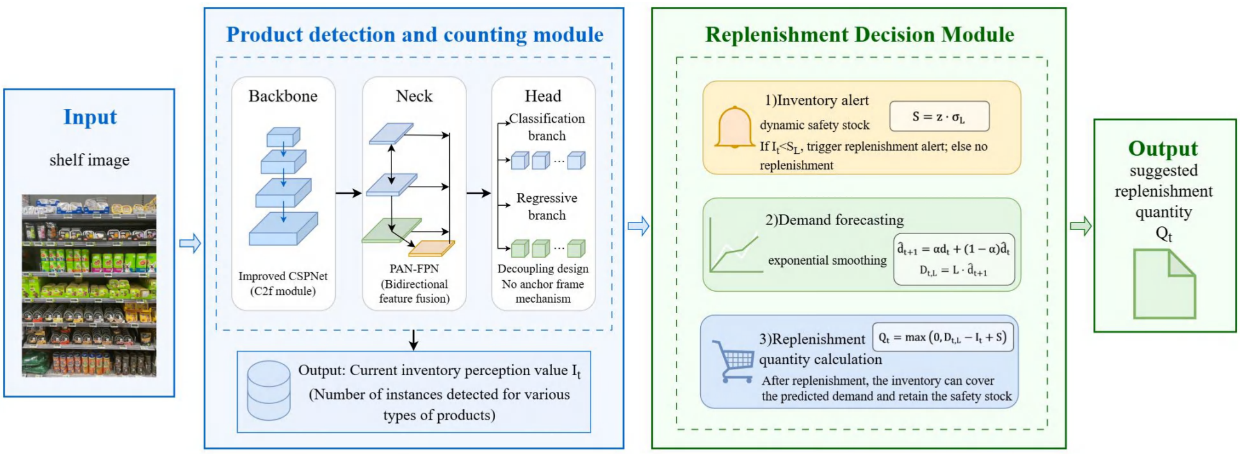

Figure 1 Framework of the Smart Shelf Product Recognition and Replenishment Decision System

2.1 Product Inspection and Counting

This module uses YOLOv11 as its core detector. YOLOv11 is a real-time, single-stage object detection model whose network architecture consists of three components: a backbone network, a neck network, and a detection head. The backbone network adopts an improved Cross Stage Partial Network (CSPNet) architecture, incorporating Cross Stage

Partial Bottleneck with 2 Convolutions(C2f) modules to enhance gradient flow and multi-scale feature representation. The neck network employs a Path Aggregation Network-Feature Pyramid Network (PAN-FPN) architecture, enabling bidirectional feature fusion from top-down and bottom-up, thereby improving small-object detection capabilities. The detection head utilizes a decoupled design and an anchor-free mechanism, outputting separate classification and regression branches, which simplifies hyperparameter tuning and improves localization accuracy in dense scenes.
For an input shelf image, the YOLOv11 model outputs the bounding boxes and class labels for each detected item. Let N denote the total number of bounding boxes detected by the model. For a specific item category c, the inventory perception value I_{detect} on the current shelf is defined as the number of instances detected for that category, as shown in Equation (1).

$$I_{detect}=\sum_{i=1}^{N}\chi(class_i=c) \tag{1}$$

Here, $\chi(\cdot)$ denotes the indicator function, which equals 1 if the condition holds (i.e., $class_i=c$) and 0 otherwise. This value represents the current inventory level I_t (where t denotes the current time) output by the vision module, and serves as the input to the replenishment decision module.

2.2 Replenishment Decision-Making Methods

Based on the current inventory level I_t output by the counting module, and drawing on inventory control theory from management science, this paper proposes a dynamic replenishment decision-making method comprising three stages: inventory status assessment, demand forecasting, and replenishment quantity calculation.

2.2.1 Inventory alert system

Suppose the daily demand for a product follows a distribution with a mean of μ_d and a standard deviation of σ_d, and the lead time (the time from order placement to delivery) is L days. Then, the total demand during the lead time follows a distribution with a mean of $L\mu_d$ and a standard deviation of $\sigma_L=\sqrt{L}\sigma_d$. This paper uses the service level method to determine the safety stock, as shown in Equation (2).

$$S=z\cdot\sigma_L \tag{2}$$

Here, z represents the service level coefficient; for example, a 95% service level corresponds to z=1.645. The inventory decision rule is as follows: if the current inventory meets the condition $I_t<S$, a replenishment alert is triggered; otherwise, no replenishment is initiated. This rule facilitates a shift from passive manual inventory counts to proactive, automated monitoring.

2.2.2 Demand forecast

To determine total demand over the forecast horizon, this paper employs simple exponential smoothing for short-term demand forecasting. Let the actual demand on day t be d_t and the forecast demand on day t be $\hat{d}_t$; the forecast for the following day is then updated using Equation (3).

$$\hat{d}_{t+1}=\alpha d_t+(1\text{-}\alpha)\hat{d}_t \tag{3}$$

Here, $\alpha\in(0,1)$ is the smoothing parameter, which controls the rate at which historical data is smoothed out. The closer α is to 1, the more sensitive the forecast is to recent changes in demand; the closer α is to 0, the smoother the forecast becomes. In practical applications, α can be calibrated to a fixed optimal value by minimizing the historical forecast error on offline data. To facilitate theoretical analysis and subsequent case studies, this paper adopts α=0.2 as a representative value under a stationary demand scenario. The total forecast demand for the next L days is then calculated using Equation (4).

$$D_{t,L}=L\cdot\hat{d}_{t+1} \tag{4}$$

2.2.3 Replenishment quantity calculation

Based on the current perceived inventory level I_t, the safety stock S, and the forecasted total demand $D_{t,L}$, this paper defines the formula for calculating the recommended replenishment quantity Q_t, as shown in (5).

$$Q_t=\max\left(0, D_{t,L}\text{-}I_t+S\right) \tag{5}$$

The core idea of this formula is that the inventory level after replenishment should be sufficient to cover the forecasted demand during the lead time, with an additional safety stock reserved to buffer against fluctuations. If the current inventory is already higher than the sum of the forecasted demand and the safety stock, then Q_t=0, indicating that no replenishment is needed. This formula directly incorporates visually detected real-time inventory into the classical inventory control model, thus achieving a quantitative transformation from visual counting to replenishment decisions.
The aforementioned replenishment decision model possesses several important theoretical properties. First, the model exhibits self-balancing of inventory levels. When I_t is high, the replenishment quantity automatically decreases or even becomes zero, avoiding overstocking; when I_t is low, the replenishment quantity rapidly increases, driving inventory levels toward the vicinity of the safety stock plus forecast demand. Second, the model ensures timely responsiveness to demand. The forecast demand $D_{t,L}$ is directly incorporated into the replenishment quantity calculation, enabling replenishment decisions to proactively address anticipated sales growth or seasonal fluctuations, rather than passively responding only after inventory falls below the reorder point, as in traditional fixed-threshold strategies. Third, the parameters possess clear managerial interpretability. The service level coefficient z in the safety stock formula $S=z\cdot\sigma_L$ directly reflects the manager's risk preference. Retailers can independently set z values for different categories based on product importance and profit margins, enabling differentiated inventory strategies. From a cybernetic perspective, this model constitutes a closed-loop feedback system: visual sensing provides the inventory measurement I_t, the decision module calculates the replenishment order Q_t, and after replenishment is executed, the inventory level rises, entering the

next sensing cycle. As long as the lead time L and the demand forecast error remain within reasonable ranges, the feedback system remains stable. Even if forecast deviations or detection errors occur, the safety stock S can serve as a buffer to prevent system oscillations.

3 EXPERIMENT

3.1 Dataset Construction

This study constructed a labeled dataset for shelf item recognition and counting based on the 'shelf_detection' subset of the 'ShelfManagement' dataset published by Pietrini et al. [9]. The dataset contains raw shelf images from multiple scenarios, covering various real-world conditions such as normal display, partial occlusion, stacking, and empty spaces. Given the annotation costs, this study selected 19 categories for annotation, including two types of empty spaces (k, k2) and 17 product categories (green1–green7, black1–black10). The Labelme tool was used to annotate bounding boxes in YOLO format for each image, resulting in a final dataset of 120 images with a total of 7,513 annotated bounding boxes. Figure 2 shows an example of a shelf product image annotated using Labelme.

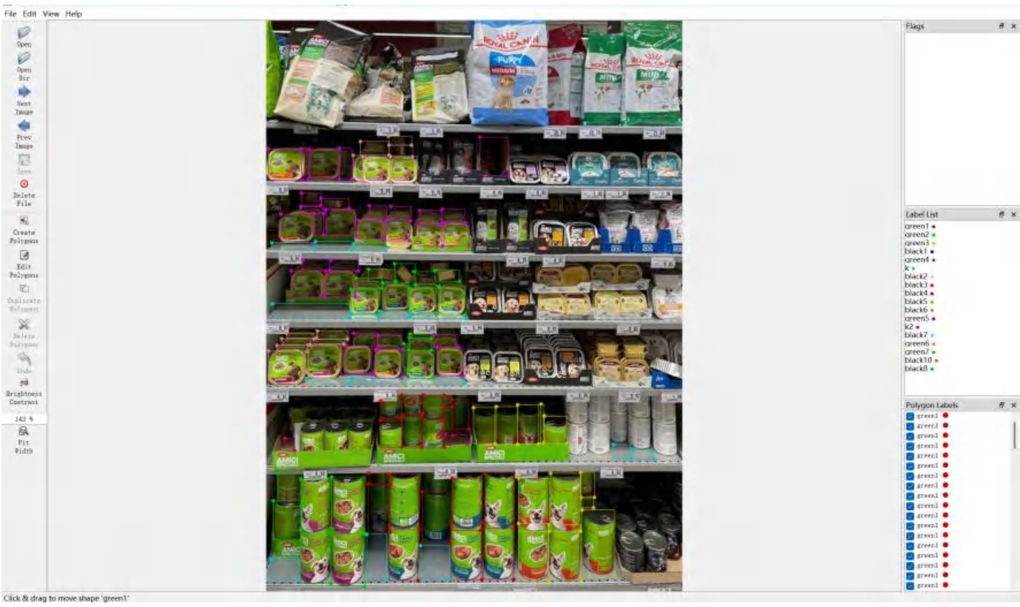

Figure 2 Annotation of Shelf Product Images Using Labelme

To expand the dataset, data augmentation was applied to the dataset in this paper. The augmentation strategies included keeping the original image, horizontal flipping, vertical flipping, 90° clockwise rotation, brightness increase (+35), brightness decrease (−35), and Gaussian blurring (with a kernel size of 5×5). All geometric transformations are applied with corresponding updates to the bounding box coordinates, and cropping is performed to ensure that labels remain within the valid region. After augmentation, the number of images is expanded to seven times that of the original dataset. Figure 3 illustrates the effect of data augmentation using a typical image as an example. Finally, the entire dataset is randomly split into a training set and a validation set in an 8:2 ratio; the training set is used for model parameter learning, and the validation set is used to evaluate model performance.

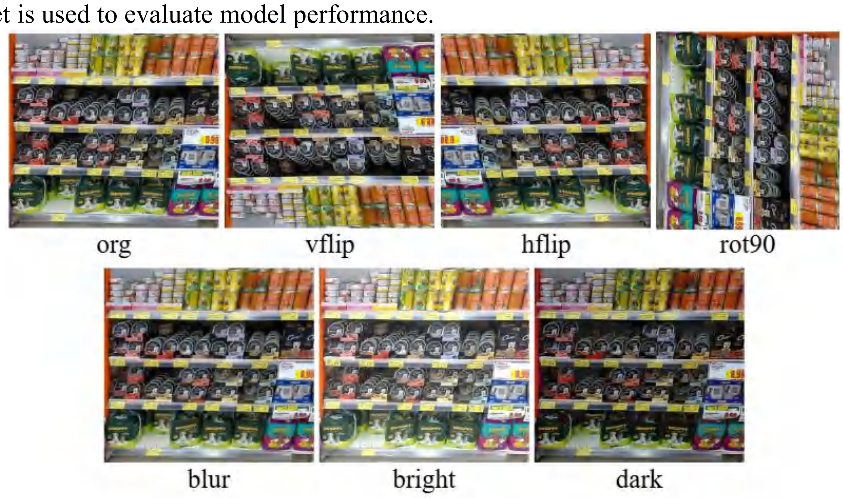

Figure 3 Visualization of Data Augmentation Results

3.2 Training and Analysis of a Shelf Product Recognition Model

This experiment was conducted on an RTX 4090D (24 GB) graphics card using PyTorch 2.0.0, Python 3.10, and CUDA 12.4. Five variants of YOLOv11 were selected for the experiment: YOLOv11n, YOLOv11s, YOLOv11m, YOLOv11l, and YOLOv11x. All models were initialized using ImageNet pre-trained weights, with the AdamW optimizer, an initial learning rate of 0.001, and a cosine-smoothed decay strategy. The batch size was set to 16, and training was conducted for 100 epochs.

Taking YOLOv11l as an example, its training loss curve and validation mAP curve are shown in Figure 4. The training loss decreases steadily from a high initial value, with a smooth convergence process and no significant fluctuations. The validation set mAP50 gradually increases from a low value, reaches a high peak, then fluctuates slightly before stabilizing. This indicates that the model has achieved effective learning and good convergence for the shelf item detection task.

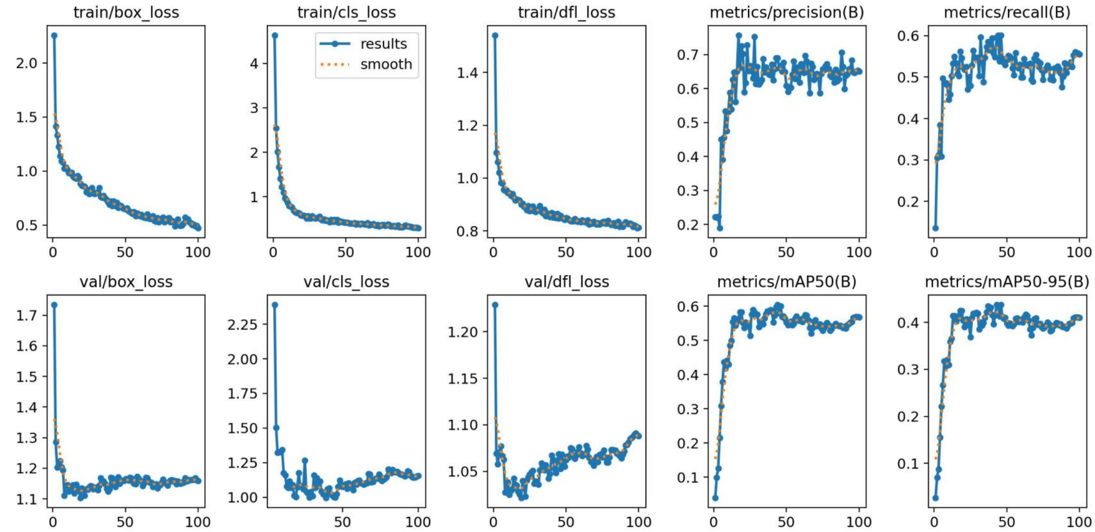

Figure 4 Training Loss Versus Validation mAP Curve (YOLOv11l)

Table 1 compares the best detection performance of five YOLOv11 models on the validation set. As shown in Table 1, detection accuracy generally increases as model capacity increases. Due to their limited feature extraction capabilities, YOLOv11n and YOLOv11s achieve mAP50 values of only 0.402 and 0.449, respectively; YOLOv11m, with its balanced depth-and-width design, achieves an mAP50 of 0.602 with 9.0M parameters, representing a significant improvement in accuracy; YOLOv11l has a slightly higher mAP50 (0.604), but its parameter count increases to 25.3M, approximately 2.8 times that of YOLOv11m; although YOLOv11x has the highest precision (0.737), its recall drops to 0.552, resulting in an mAP50 of 0.601—roughly on par with YOLOv11m—yet its parameter count reaches as high as 56.9M. This phenomenon indicates that overly large models may overfit on limited datasets, while also significantly increasing computational costs. Therefore, after comprehensively considering detection accuracy, the number of parameters, and inference efficiency, the subsequent restocking decision module selected YOLOv11m as the visual perception model.

Table 1 Comparison of the Best Detection Performance of Different YOLOv11 Models on the Validation Set

Model	Number of Features (M)	Precision	Recall	mAP50	mAP50-95
YOLOv11n	2.6	0.446	0.461	0.402	0.290
YOLOv11s	7.2	0.494	0.467	0.449	0.327
YOLOv11m	**9.0**	**0.648**	**0.590**	**0.602**	**0.442**
YOLOv11l	25.3	0.690	0.601	0.604	0.436
YOLOv11x	56.9	0.737	0.552	0.601	0.435

To visually demonstrate the counting performance of each model, Figure 5 present a comparison of the class-wise prediction results for the five models on the same typical shelf image. As shown in Figure 5, the distribution of predicted bounding boxes for YOLOv11m is closest to the ground truth (GT); YOLOv11l and YOLOv11x also perform well, while the lightweight models YOLOv11n and YOLOv11s exhibit a higher number of false positives and false negatives.

Figure 5 Comparison of Per-Category Product Predictions for Different Models Using the Same Shelf Image

3.3 Comparison with Other Visual Models

To further validate the effectiveness of the selected model, this paper conducts comparative experiments between YOLOv11m and classic object detection models—Faster R-CNN, RetinaNet, and YOLOv8m from the same series—using the same validation set. The results are shown in Table 2. As shown in Table 2, Faster R-CNN, as a classic two-stage detector, has a parameter count as high as 41.8M, but its mAP50 is only 0.548, indicating that it does not excel in either accuracy or efficiency; Although RetinaNet introduces Focal Loss to mitigate the issue of data imbalance, its mAP50 (0.571) and mAP50-95 (0.403) remain lower than those of YOLOv11m, and its parameter count is approximately four times that of YOLOv11m; YOLOv8m and YOLOv11m both belong to the YOLO series and are highly comparable; however, YOLOv11m outperforms YOLOv8m in both mAP50 and mAP50-95, while reducing the number of parameters by approximately 65%. In summary, YOLOv11m achieves optimal detection performance with the fewest parameters, offering clear advantages in both accuracy and efficiency, making it well-suited for the practical application requirements of the task described in this paper.

Table 2 Comparison of Performance with Other Visual Inspection Models

Model	Number of Features (M)	Precision	Recall	mAP50	mAP50-95
Faster R-CNN	41.8	0.574	0.523	0.548	0.376
RetinaNet	36.4	0.601	0.537	0.571	0.403
YOLOv8m	25.9	0.717	0.525	0.579	0.423
YOLOv11m (Ours)	**9.0**	**0.648**	**0.590**	**0.602**	**0.442**

3.4 Comparison with Other Replenishment Strategies

This section uses a numerical example to illustrate the dynamic reorder decision-making process proposed in this paper and compares it with the fixed-threshold strategy commonly used in retail practice.

Assume that a product has a historical average daily sales volume of 10 units, a standard deviation of 3 units, and a lead time for replenishment of L=3 days. The standard deviation of demand within the lead time, $\sigma_L=\sqrt{3}\times3\approx5.2$. Taking a service level coefficient z=1.645 (corresponding to a 95% service level), the safety stock $S=z\cdot\sigma_L\approx8.6$ units, rounded to 9 units.

During a visual shelf inspection, YOLOv11m detected that the current shelf inventory for this product is 6 units (i.e., I_t=6). Meanwhile, an exponential smoothing forecast based on recent sales data (with α=0.2) indicates that the predicted demand for the next day, $\widehat{d}_{t+1}$=12 units (higher than the historical average, possibly due to a promotional campaign). Based on this, the total forecast demand for the next 3-day lead time is calculated as $D_{t,L}$=3×12=36 units. Substituting this into the replenishment formula $Q_t=\max(0, D_{t,L}-I_t+S)$ yields a recommended replenishment quantity of Q_t=39 units.

By contrast, the fixed-threshold strategy sets a fixed reorder point $R=S$=9 units (to ensure a fair comparison, the reorder point is set at the safety stock level) and a fixed replenishment quantity $Q=L\cdot\mu_d$=30 units (average demand over the lead time). The current inventory of 6 units is below the reorder point, triggering a replenishment of 30 units, resulting in a post-replenishment inventory of 36 units.The minimum-maximum inventory strategy sets the minimum inventory level $S_{min}=S$=9 units and the maximum inventory level $S_{max}=L\cdot\mu_d+S$=30+9=39 units. When the current inventory is $\leq S_{min}$, replenish to S_{max}. At this point, the replenishment quantity $Q_t=S_{max}-I_t$=39-6=33 units, and the inventory after replenishment is 39 units.

The performance of the three strategies differed significantly, as shown in Table 3. The dynamic model used exponential smoothing to forecast and promptly capture the upward trend in demand, increasing the reorder quantity to 39 units, which better met the high demand during the promotional period; in contrast, the fixed-threshold strategy failed to detect changes in demand, and the reorder quantity of 30 units could only sustain inventory for about 2.5 days, potentially leading to another stockout during the lead time; Although the minimum-maximum strategy replenished inventory to 39 units, this was still below the dynamic model's target level of 45 units, and it similarly failed to account for demand trends. In terms of inventory fluctuations, the dynamic model maintains inventory near the target level of $D_{t,L}+S$=45 units, with the actual inventory after replenishment being 6+39=45 units; the fixed-threshold strategy results in an inventory of 6+30=36 units after replenishment, and the minimum-maximum strategy results in 39 units, both of which are below the target level. Consequently, subsequent inventory troughs are lower, and fluctuations are more pronounced.

Table 3 Comparison of Results for Different Replenishment Strategies

Strategy	Current Inventory (units)	Replenishment Quantity (units)	Post-Replenishment Inventory (units)	Consideration of Demand Trends	Risk of Stockouts during Lead Time	Risk of Excess Inventory
Dynamic Model in This Paper	**6**	**39**	**45**	**Yes**	**Low**	**Low**
Fixed Threshold Strategy	6	30	36	No	High	Medium
Min-Max Strategy	6	33	39	No	Medium	Medium

Looking at inventory changes during periods of demand fluctuation, when the initial inventory falls below the safety stock level, the dynamic model immediately triggers a replenishment, allowing inventory to quickly return to the target level; subsequently, during the rising demand phase, the forecast-driven replenishment mechanism maintains stable inventory levels, effectively preventing stockouts. In contrast, the fixed-threshold strategy and the minimum-maximum strategy results in a passive decline in inventory, which is prone to triggering repeated replenishments and may still lead to stockouts. During the demand decline phase, the dynamic model automatically reduces the reorder quantity to avoid overstocking; in contrast, the other two strategies continues to replenish at a fixed quantity, leading to excess inventory.

The above case study and analysis demonstrate that the replenishment decision-making model proposed in this paper—which integrates visual inventory sensing, dynamic safety stock, and demand forecasting—offers greater adaptability and superior inventory efficiency compared to fixed-threshold and minimum-maximum inventory strategies.

3.5 Ablation Experiment

This paper has validated the rationality and effectiveness of the selected vision model by comparing different sizes of YOLOv11 models and by contrasting YOLOv11 with other vision models. In this section, we further conduct ablation tests on the dynamic replenishment model to analyze the independent contributions of its components (safety stock and demand forecasting). We design ablation experiments to compare the replenishment performance of the following three strategies in the same scenario.

Strategy A (Full Model): Uses both safety stock and demand forecasting; replenishment quantity is $Q_t=\max(0, D_{t,L}-I+S)$.

Strategy B (Safety Stock Only): No demand forecasting; instead, historical average demand μ_d replaces the forecast demand. The replenishment quantity is $Q_t=\max(0, L\cdot\mu_d-I+S)$. This degenerates to the classic fixed reorder point and fixed order quantity, but retains dynamic safety stock threshold.

Strategy C (Demand Forecast Only): Sets the safety stock to zero and relies solely on demand forecasts for replenishment; the formula is $Q_t=\max(0, D_{t,L}-I)$.

Based on the classic principle in inventory management theory—that the average inventory level should be as low as possible while meeting a given service level—this paper employs two concise core metrics consistent with inventory management theory to evaluate each strategy. The first metric is service level achievement, reflected by the risk of stockouts. An ideal strategy should be able to keep the probability of stockouts close to a preset target service level (e.g., 95%) and avoid a significant increase in the stockout rate due to demand fluctuations or trend changes. The second metric is inventory efficiency, reflected by capital tied up and the risk of excess inventory. Under the premise of meeting the service level, the lower the average inventory level, the faster the inventory turnover, and the higher the capital efficiency. An ideal strategy should avoid excessive inventory buildup during periods of declining demand.

Using the case parameters from Section 3.4 (μ_d=10, σ_d=3, L=3, target service level 95%, safety stock S=9, current perceived inventory I_t=6, and forecast demand rising to $\hat{d}$=12), a comparison of the performance of the three strategies is shown in Table 4.

Table 4 Comparison of Ablation Experiment Strategies

Strategy	Demand Forecast	Safety Stock	Reorder Point Formula	Service Level Achievement	Inventory Efficiency
A	**Yes**	**Yes**	$\boldsymbol{Q_t=D_{t,L}-I+S}$	**Consistently meets target (≈95%)**	**High (close to target level, no excess)**
B	No	Yes	$Q_t=L\cdot\mu_d-I+S$	Volatile: High stockouts during upturns, excess inventory during downturns	Low (Poor inventory efficiency when trends change)
C	Yes	No	$Q_t=D_{t,L}-I$	Often below target (No buffer to handle errors)	Moderate (High stockout costs, but low excess inventory)

In Strategy B, safety stock can mitigate random fluctuations, bringing the long-term average stockout probability close to the target value, but it cannot respond to changes in demand trends. During periods of rising demand (e.g., when $\hat{d}$=12>μ_d in the case study), the reorder quantity is still calculated based on the historical average, leading to a significant increase in the stockout probability; conversely, during periods of declining demand, the reorder quantity is excessively large, resulting in excess inventory. Therefore, its service level is unstable, and inventory efficiency is low. Strategy C can promptly detect demand trends but lacks a safety stock buffer. Once forecasting errors occur or demand fluctuations exceed expectations, inventory is highly likely to be depleted within the lead time, resulting in a stockout probability far higher than the target service level—trading off service quality for limited improvements in excess inventory. In Strategy A, demand forecasting enables replenishment quantities to respond to trend changes in advance, preventing stockouts during upward trends; safety stock provides a buffer against forecasting errors and fluctuations, ensuring the probability of stockouts remains stable near the target level; simultaneously, it automatically reduces replenishment quantities during periods of declining demand to prevent excessive inventory buildup. Consequently, this strategy achieves both stable compliance with service levels and high inventory efficiency.

Ablation experiments demonstrate that demand forecasting and safety stock play complementary roles in dynamic replenishment decision-making and are both indispensable. Relying solely on safety stock cannot address changes in demand trends, while relying solely on demand forecasting cannot withstand uncertainty and forecasting errors. The comprehensive model proposed in this paper, by synergistically utilizing both components, significantly outperforms any simplified version under concise and reasonable evaluation criteria, thereby validating the necessity and superiority of the model design.

4 CONCLUSIONS AND OUTLOOK

This paper addresses the challenges of inventory perception lag and reliance on experience in replenishment decision-making within retail shelf management by proposing an intelligent replenishment decision-making framework that integrates computer vision and management science. The framework utilizes YOLOv11 as the core of its product recognition and counting module to enable automatic perception of current inventory levels from image data. It also incorporates inventory control theory to establish a closed-loop decision-making model that encompasses dynamic safety stock alerts, exponential smoothing demand forecasting, and adaptive replenishment quantity calculation. Five versions of YOLOv11 were evaluated on a self-built dataset comprising 19 categories and 7,513 instances. The results indicate that the YOLOv11m model achieves the optimal balance between detection accuracy and counting reliability, with an mAP50 of 0.602. It also demonstrates superiority over other mainstream models such as Faster R-CNN, RetinaNet, and YOLOv8m, providing an accurate data foundation for management decisions. Theoretical analysis, case simulations and ablation experiments further reveals that the proposed dynamic replenishment model can more effectively address demand fluctuations and reduce the risks of stockouts and inventory buildup compared to fixed-threshold strategy and minimum-maximum inventory strategy.At the same time, they validate the complementary roles of safety stock and demand forecasting in dynamic replenishment decisions, demonstrating that both are indispensable.

Although this work provides a theoretical framework and technical approach for visual inventory management in smart retail, there is still room for further development. First, the coupling method between the visual perception and decision-making models can be further optimized by exploring end-to-end joint learning mechanisms that directly incorporate feedback signals from the decision module (such as the cost of stockouts) into the training process of the detection network. Second, since the current model's decision logic is based on the assumption of independent items, future research could address issues such as multi-category joint replenishment optimization and shelf space allocation to comprehensively enhance retail operational efficiency. Additionally, future research plans to deploy this system in actual retail stores and conduct long-term field tests to verify its robustness and stability under real-world conditions—such as complex lighting and varying display densities—thereby advancing this framework from proof-of-concept to large-scale engineering applications.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Gruen T W, Corsten D S, Bharadwaj S. Retail out-of-stocks: A worldwide examination of extent, causes and consumer responses. Washington, DC: Grocery Manufacturers of America, 2002.
[2] Jiang Honghua. Several Suggestions for Improving the Supermarket Inventory Counting System. Times Finance, 2014 (4Z): 173–174.
[3] Kikuchi T, Takeuchi S. Self-supervised human-object interaction of complex scenes with context-aware mixing: Towards in-store consumer behavior analysis//Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2024: 744-751.
[4] Varol G, Kuzu R S. Toward retail product recognition on grocery shelves//Sixth International Conference on Graphic and Image Processing (ICGIP 2014), SPIE, 2015, 9443: 46-52.
[5] Sun Haixia. Study on segmentation and recognition methods of supermarket shelf. Master's thesis, Nanjing University of Technology and Engineering, 2014.
[6] Goldman E, Herzig R, Eisenschtat A, et al. Precise detection in densely packed scenes//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2019: 5227-5236.
[7] Guo Z, Wang G, Xie K. A lightweight YOLOv11-based commodity detector for dense smart retail cabinets. Scientific Reports, 2026.
[8] Chen B, Li L, Wang Y, et al. RetailDet: An Efficient Fusion Attention Network for Joint Product and Vacancy Identification in Smart Retail. IEEE Access, 2025.
[9] Pietrini R, Paolanti M, Mancini A, et al. Shelf Management: A deep learning-based system for shelf visual monitoring. Expert Systems with Applications, 2024, 255: 124635.
[10] Alghaslan M, Almutairy K, El-Alfy E S, et al. RetailEye: Supervised Contrastive Learning with Compliance Matching for Retail Shelf Monitoring//International Conference on Neural Information Processing, Singapore: Springer Nature Singapore, 2024: 248-263.