International Journal of Engineering Inventions
e-ISSN: 2278-7461, p-ISSN: 2319-6491
Volume 14, Issue 10 [October 2025] PP: 109-114

Circular Metal Identification Based on Machine Vision

Jiale Quan Zhibo Sun Jianshuo Lin Yibo Wang Songqi Li Guifeng Wang
1 School of Mechanical and Vehicle Engineering, Changchun University, Changchun 130022, China
Corresponding Author: Zhibo Sun

Abstract: *To address the issues of low efficiency, high misjudgment rate, and inability to adapt to high-speed production in manual inspection of beverage bottle cap production lines, this paper designs and implements an automated defect detection system based on the improved YOLOv11. The system uses an industrial camera as the core for image acquisition, enhances the ability to perceive subtle defects by introducing the SE (Squeeze-and-Excitation) attention mechanism, and optimizes exposure time, light source arrangement, and image preprocessing algorithms to cope with high-speed movement and lighting changes. Meanwhile, it operates with an NVIDIA RTX 3060 GPU, achieving a detection frame rate of 30 FPS at a resolution of 4000×4000 and enabling real-time identification of various defects such as cracks and foreign objects. Experimental results show that the improved model achieves a mAP@0.5 of 98.3% on the test set, with the false detection rate and missed detection rate reduced to 2.1% and 1.7% respectively, representing a significant improvement over traditional algorithms. Additionally, the system offers advantages of strong versatility and flexible deployment, can be promoted to packaging inspection in industries such as beverages and pharmaceuticals, and possesses both engineering application value and economic benefits.[1]*

Date of Submission: 13-10-2025 Date of acceptance: 27-10-2025

I. INTRODUCTION

1.1 Research Background and Significance

As a crucial component of the packaging for beverage and beer products, bottle caps are directly related to the product's sealing performance, freshness retention, and safety. Defects in bottle caps, such as cracks, deformation, poor sealing, and foreign object contamination, not only lead to beverage leakage, pollution, and deterioration but also may cause safety accidents and endanger consumers' health[2]. Therefore, realizing automated defect detection of bottle caps in a high-speed production environment holds urgent practical significance. However, traditional manual inspection, characterized by high costs, low efficiency, slow speed, and a high error rate, can no longer meet the requirements of modern production lines for inspection accuracy, stability, and production capacity. [3]In response to this demand, this paper studies and proposes a computerized automatic detection algorithm based on image acquisition and machine learning classification. By efficiently identifying and classifying bottle cap images, the algorithm significantly improves inspection accuracy and speed, enhances production efficiency, and reduces production costs.

At the same time, most existing methods rely on traditional cv2 image processing. Conventional target detection methods usually depend on manually designed features and classic machine learning algorithms, such as sliding windows, HOG (histogram of oriented gradients), and SVM (support vector machine), as shown in Figure 1.1.[4,5] Nevertheless, these methods often perform poorly when dealing with complex scenarios, fail to meet the needs of practical applications, and are typically characterized by a single detection object, strong dependence on scenarios, low recognition accuracy, and poor system robustness. Moreover, the algorithms cannot simultaneously identify the integrity of the two-dimensional code or encoding on the inner side of the bottle cap and the presence of damage, still requiring manual intervention. In addition, traditional image recognition methods have obvious shortcomings in handling large-batch recognition, bottle cap stacking, and complex changes in lighting and posture, making it difficult for them to meet the stable operation requirements of industrial sites.

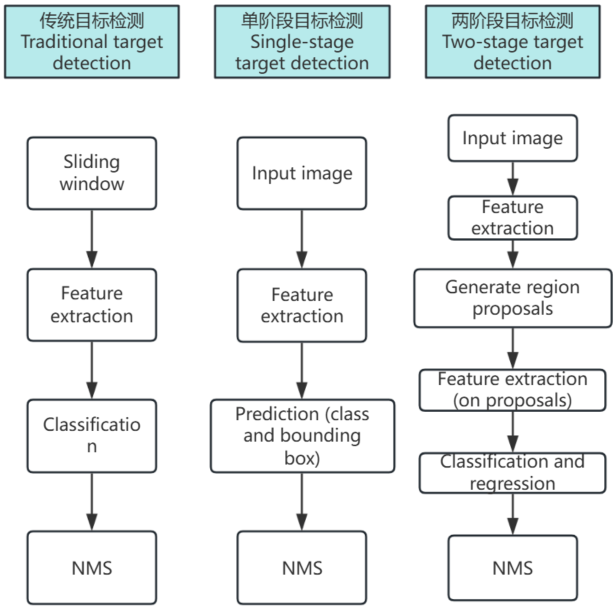

Fig. 1 Machine vision algorithm

1.2 Domestic and Foreign Research Status

1.2.1 Application of Traditional Image Processing Algorithms in Bottle Cap Defect Detection

With the development of computer vision and image processing technologies, traditional image processing algorithms have been widely applied in the detection of surface defects on bottle caps. Such methods mainly rely on analyzing and judging low-level features of images, such as grayscale, edges, and textures.[6]

Hai Chao et al. proposed a defect detection method that uses Blob analysis to extract targets and Otsu global threshold segmentation to achieve defect detection.[7] This method has a simple structure and fast calculation speed, and is suitable for situations where the contrast between the background and the target is obvious. However, when there is reflection on the bottle cap surface, the background is complex, or the defect contrast is low, the detection effect significantly degrades, and its robustness to noise is poor.

Zeng Qingyue et al. used Hough transform to extract the linear features of bottle cap scratches, achieving gorecognition results for linear defects. Nevertheless, this method is relatively sensitive to lighting changes and low-contrast environments, making it prone to missed detections.[8]

Sun Xiaona et al. proposed a detection method based on the grayscale template matching algorithm, which judges the presence of defects by comparing the correlation of grayscale features. This method works well under uniform lighting and stable background conditions. However, when the lighting conditions change, the grayscale values fluctuate greatly, resulting in unstable matching results.[9]

In summary, traditional image processing algorithms have advantages such as simple implementation and fast calculation speed. However, they rely on fixed thresholds, lighting conditions, and background features, leading to limited robustness and generalization ability. Thus, they are difficult to meet the requirements of high-precision and real-time detection in complex industrial environments.

1.2.2 Research on Improved Algorithms and Composite Methods

To address the issue of low detection accuracy of traditional algorithms under lighting changes, complex backgrounds, and noise interference, scholars at home and abroad have proposed a variety of improved algorithms and composite detection methods to enhance the stability and accuracy of the system.

Wu Shan et al. adopted a guided filtering algorithm to remove noise while preserving image details, and combined it with an improved linear defect detection algorithm to achieve effective extraction of scratches in complex backgrounds. This method overcomes the detection difficulties caused by low contrast to a certain extent, but its detection of tiny and weak scratches is still not ideal.[10]

Peng Yun et al. designed a dispensing quality detection system based on the HALCON platform. This system uses mean filtering to eliminate noise and realizes defect detection through template matching positioning, global threshold segmentation, and morphological processing. However, this method is mainly suitable for scenarios with a stable environment and significant grayscale differences between the target and the background. When the lighting or material changes, the effect will decrease significantly.[11]

Li Kebin proposed a composite detection method combining high and low-angle lighting technologies. By using multi-angle light sources to enhance the contrast of defect areas, and utilizing the background subtraction algorithm and the directional gradient-improved region growing method to extract scratch features, a higher detection accuracy is achieved.[12]

In addition, Li Chen et al. used frequency-domain filtering and double-threshold background subtraction algorithms to enhance and segment weak scratches, improving the system's anti-interference ability against complex backgrounds.[13]

In general, improved and composite algorithms have improved the stability and accuracy of detection to a certain extent, and their performance is superior to traditional methods especially under complex conditions such as lighting variations and noise. However, such algorithms still rely on a large number of parameter adjustments and empirical selections, which puts forward higher requirements for the real-time performance and adaptive ability of the algorithms.

II. Overall System Design

2.1 Overall System Structure

The overall architecture of this system consists of both a hardware layer and a software layer. In terms of algorithms, the YOLOv11 serves as the core detection network. As the latest generation of the YOLO series models, YOLOv11 incorporates more advanced feature extraction and attention mechanisms, which can significantly improve detection performance while maintaining model lightweight. [14,15]Among them, YOLOv11 integrates complex spatial attention modules (such as C2PSA), enabling the model to more effectively focus on key areas in the image and greatly improving its ability to detect complex or partially occluded targets.

To further enhance the sensitivity to tiny defects such as scratches on the bottle cap surface, the system introduces the SE (Squeeze-and-Excitation) channel attention module into the YOLOv11 structure.[16] Through global pooling and adaptive recalibration, the SE module weights the features of each channel, emphasizing useful features and suppressing redundant information. [17]This enables the model to have a stronger expressive ability when detecting tiny scratches and small defects on the bottle cap surface, effectively improving the recall rate and accuracy of small target detection.

In terms of tracking, the system adopts the SORT (Simple Online and Realtime Tracking) algorithm to track and number bottle cap targets in consecutive frames. [18]The SORT algorithm uses Kalman filtering for motion prediction and the Hungarian algorithm to associate the detection boxes with existing trajectories. Since this method only relies on the target's position and a constant velocity model for prediction, it is very simple and efficient. The tracking update rate can be as high as 260 Hz, which is much higher than the typical acquisition speed of 30 fps. It can not only meet the real-time requirements but also accurately maintain the trajectory of each bottle cap, avoiding repeated counting or missed detection. When a new detection object appears, the system automatically assigns a new ID; when the target leaves the field of view, the corresponding trajectory is terminated to ensure the accuracy of counting statistics.

In terms of hardware, the system selects a Hikvision industrial camera (4000×4000 pixels, 30 fps) for image acquisition. This camera supports the GigE Vision protocol and transmits uncompressed images in real-time through a Gigabit Ethernet interface. [19]The uncompressed original high-resolution images can maximize the retention of texture details on the bottle cap surface, which is beneficial for subsequent algorithms to detect scratches and tiny defects. Moreover, the image transmission delay is extremely low, ensuring the stable acquisition of continuous multi-frame images. The system is deployed on an industrial-grade control host equipped with an NVIDIA RTX 3060 GPU for acceleration. The RTX 3060 has sufficient CUDA cores and memory, which can support high-speed inference of deep learning models, thereby meeting the computing requirements for real-time detection at high frame rates.

In terms of software architecture, the system is divided into the following main functional modules:

1. Image Acquisition Module: Obtains continuous video frames through the camera interface, maintains an acquisition rate of 30 fps, and caches the original images in real-time.
2. Preprocessing Module: Performs necessary preprocessing on the acquired images, including image denoising, white balance adjustment, cropping of the region of interest (ROI), and resizing to the input size required by YOLOv11, so as to provide clean and consistent input for the detection algorithm.
3. YOLOv11-SE Detection Module: Performs target detection on the preprocessed images, and outputs the bounding box and confidence information of all bottle cap targets. This module integrates the SE attention mechanism, focusing on enhancing the surface detail features of bottle caps and enabling accurate detection of tiny defects such as scratches and dents.
4. SORT Tracking Module: Takes the YOLO detection results as input, uses Kalman filtering combined with the Hungarian algorithm to track each bottle cap target, and assigns a unique ID to each trajectory. This module maintains the target trajectory across frames to ensure continuous tracking of the same bottle cap, thereby avoiding multiple repeated counts.
5. Result Output and Counting Module: Generates statistical results by integrating tracking information,

counts each newly generated trajectory (each trajectory represents a new bottle cap target), and records information such as defect category and position. At the same time, the results are output through a graphical interface or a network interface for use by the production line automation system.

The above design scheme integrates advanced detection algorithms with an industrial hardware platform, ensuring both algorithm accuracy and meeting the real-time processing requirements of industrial sites. The defect detection based on YOLOv11-SE combined with SORT tracking realizes accurate and fast online detection and counting of bottle cap defects, which can be directly applied to the quality control system of production lines and has strong practicality and expansibility.

III. System Implementation and Experimental Results

3.1 Overall System Structure

This system is mainly composed of four core modules:

1. Image Acquisition Module: Uses a Hikvision industrial camera and realizes high-speed transmission of uncompressed images through a Gigabit Ethernet interface. The light source adopts a ring-shaped LED structure to reduce misjudgments caused by bottle cap reflection.
2. Data Preprocessing Module: Uses OpenCV to perform grayscale conversion, mean filtering, brightness equalization, and ROI extraction on the acquired images, so as to improve the stability of the model input.
3. Model Detection Module: Based on the YOLOv11 algorithm framework, it identifies and classifies bottle cap defects. By adding the SE attention mechanism and multi-scale feature fusion structure, it improves the detection ability for tiny cracks, scratches, and foreign objects.
4. Result Display and Statistics Module: Uses PyQt5 to build a human-computer interaction interface, realizing real-time display of detection results, defect marking, and quantity statistics.

The modules interact with each other through message queues to achieve a smooth pipelined detection process.

3.2 Algorithm Design and Improvement

Traditional YOLO models have the problem of feature loss when detecting small targets. Aiming at the characteristics of bottle cap defects, such as small features, low contrast, and irregular shapes, this paper makes the following improvements based on YOLOv11:

(1) Introduce the SE attention mechanism: By compressing the channel features and then expanding them, the network can automatically allocate the importance of feature channels and enhance the response of key defect features.

(2) Multi-scale feature fusion (FPN+PAN): Establish an information channel between shallow and deep features, enabling the model to retain edge details while possessing high-level semantic information

(3) Loss function optimization: Replace the traditional MSE bounding box regression method with CIOU loss, allowing the model to converge faster during training and improving positioning accuracy.

(4) Data augmentation strategies: Include random rotation, lighting disturbance, affine transformation, and contrast adjustment to enhance the generalization ability of the model and adapt to changes in the posture of bottle caps on the production line.

3.3 System Implementation

The hardware platform includes:

1. Industrial computer (Intel i7-12700, 32GB memory, RTX 3060 GPU)
2. Hikvision GigE industrial camera (resolution 4000×4000)
3. Ring-shaped cold-light LED light source and stable power control module

Fig. 2 The bottle caps on the conveyor belt

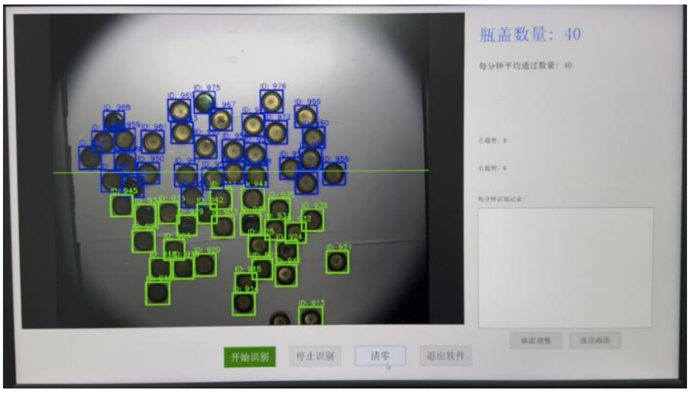

Fig. 3 Program running interface

The software platform is developed using the Python language, and the core dependent libraries include PyTorch, OpenCV, PyQt5, and NumPy. The front-end interface is designed using Qt Designer, realizing real-time display of video streams, highlight marking of detection results, and display of defect statistics charts. During the system operation, the video stream is input into the model at a rate of 30 frames per second, and the results are returned within 100 ms after GPU inference, realizing true online detection.

3.4 Experimental Design and Result Analysis

The experiment was conducted on a bottle cap production line, and a total of 18,000 bottle cap images were collected, among which defective samples accounted for approximately 35%. The training and test sets were divided in an 8:2 ratio. The Adam optimizer was used for training, with an initial learning rate of 0.001 and a batch size of 32. The model converged after 100 training epochs.

The experiment compared three algorithms: the traditional CV template matching algorithm, the original YOLOv11 model, and the improved YOLOv11-SE model. The results are as follows:

Model	mAP@0.5	Detection Speed (FPS)	Missed Detection Rate	False Detection Rate
Traditional Image Algorithm	78.2%	10	12.4%	9.1%
Original YOLOv11	91.5%	28	4.8%	3.6%
Improved YOLOv11-SE	98.3%	30	1.7%	2.1%

It can be seen from the experimental data that the improved model has achieved significant improvements in accuracy, speed, and stability. Especially under conditions of large lighting changes and bottle cap stacking, it can still maintain a high recognition accuracy. In the long-term continuous operation test of the system, no obvious delay or detection drift occurred, indicating that it has good robustness and engineering feasibility.

IV. Conclusion

Based on the YOLO target detection algorithm, this paper designs a set of automated visual detection systems suitable for bottle cap production lines. By introducing the SE attention mechanism, multi-scale feature fusion structure, and improved loss function into the YOLOv11 network, the accuracy and stability of the system in small target defect detection are significantly improved. Combined with a high-resolution industrial camera and an optimized image preprocessing algorithm, the system realizes the functions of real-time detection, classification, and statistics of bottle cap defects, meeting the detection accuracy and speed requirements of industrial production.

Experimental results verify that the system still has strong robustness under scenarios of complex lighting, bottle cap overlap, and high-speed movement, with a detection frame rate of over 30 FPS and a high mAP of 98.3%. Compared with traditional manual detection and rule-based image algorithms, the system has achieved significant improvements in detection efficiency, accuracy, and automation level.

Future research will be carried out in the following directions:

(1) Explore the deployment of lightweight YOLO models on embedded terminals (such as Jetson Nano and Raspberry Pi) to achieve low-power detection

(2) Introduce a multi-task learning structure to enable the system to simultaneously identify the integrity of printed characters on the bottle cap surface and two-dimensional code information.

(3) Combine edge computing with cloud synchronization mechanisms to realize real-time reporting of detection

data and remote quality analysis, thereby constructing a complete intelligent manufacturing quality traceability system.

In conclusion, the pipelined bottle cap defect detection system based on YOLO proposed in this paper meets the industrial application standards in terms of accuracy, speed, and stability, and provides a feasible engineering implementation path for subsequent research on deep learning-based visual detection systems.

REFRENCES

[1] VIJAYAKUMAR A, VAIRAVASUNDARAM S, KOILRAJ J A S. Real-time visual intelligence for defect detection in pharmaceutical packaging[J]. Scientific Reports, 2024, 14(1): 18811.

[2] FARHANGI O, SHEIDAEE E, KISALAEI A. Machine Vision for Detecting Defects in Liquid Bottles: An Industrial Application for Food and Packaging Sector[J]. Cloud Computing and Data Science, 2024: 242-254.

[3] Beverage Packaging Market Size And Share Report, 2030[EB/OL]. [2025-10-19]. https://www.grandviewresearch.com/industry-analysis/beverage-packaging-market.

[4] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection[C/OL]//2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05):

[5] ZHOU X, WANG Y, XIAO C. Automated Visual Inspection of Glass Bottle Bottom With Saliency Detection and Template Matching[J]. IEEE Transactions on Instrumentation and Measurement, 2019, 68(11): 4253-4267.

[6] KARIMI M H, ASEMANI D. Surface defect detection in tiling Industries using digital image processing methods: Analysis and evaluation[J]. ISA Transactions, 2014, 53(3): 834-844.

[7] CHEN B, LI C, YUAN P. Research on defect detection of bottle cap interior based on low-angle and large divergence angle vision system[J]. PLOS ONE, 2024, 19(5): e0303744.

[8] QINGYUE ZENG. Detection of linear scratches based on Hough transform in old movies [J]. Computer Engineering and Design, 2010, 31(4): 843-846.

[9] Sun Xiaona, Liu Jichao, Gao Guohua. Vision-based defect detection technology for date coding on dairy packaging

[10] [J]. Food and Machinery, 2018, 34(10): 100-103, 108.

[11] Wu Shan, Chen Bin, Yang Xiyuan. Key Algorithm Design for Card Scratch Detection[J].computer application, 2017, 37(S2): 171-174.

[12] TAN J, LI L, WANG Y,. The Quality Detection of Surface Defect in Dispensing Dack-End Based on HALCON[C/OL]//2016 International Conference on Cybernetics, Robotics and Control (CRC). 2016: 95-98[2025-10-19]. https://ieeexplore.ieee.org/abstract/document/7813125.

[13] LIU B, YANG Y, WANG S. An automatic system for bearing surface tiny defect detection based on multi-angle illuminations[J]. Optik, 2020, 208: 164517.

[14] HU X, SUN H, WANG S. Weak scratch enhancement algorithm based on frequency domain characteristics[C/OL]//AOPC 2019: Optical Sensing and Imaging Technology: 卷 11338. SPIE, 2019: 268-273[2025-10-19]. https://www.spiedigitallibrary.org/conference-proceedings-of-spie/11338/1133812/Weak-scratch-enhancement-algorithm-based-on-frequency-domain-characteristics/10.1117/12.2543103.full.

[15] REDMON J, DIVVALA S, GIRSHICK R. You Only Look Once: Unified, Real-Time Object Detection[EB/OL]. arXiv, 2016[2025-10-19]. http://arxiv.org/abs/1506.02640.

[16] Wu Jia Kun, Qing Lin Bo, Xie Jian Bin. Defect detection algorithm for white spirit bottle caps based on improved YOLOv10 [J]. radioengineering, 2025, 55(7): 1440-1447.

[17] Improved Algorithm for Face Mask Detection Based on YOLO-v4 | International Journal of Computational Intelligence Systems[EB/OL]. [2025-10-19]. https://link.springer.com/article/10.1007/s44196-023-00286-7.

[18] HU J, SHEN L, SUN G. Squeeze-and-Excitation Networks[J].

[19] Simple Online and Realtime Tracking[EB/OL]. [2025-10-19]. https://ar5iv.labs.arxiv.org/html/1602.00763.

[20] A Comprehensive Survey for Real-World Industrial Defect Detection: Challenges, Approaches, and Prospects[EB/OL]. [2025-10-19]. https://arxiv.org/html/2507.13378v1.