Automation and Machine Learning (2025)
Clausius Scientific Press, Canada

DOI: 10.23977/autml.2025.060207
ISSN 2516-5003 Vol. 6 Num. 1

A Review of the Application of Machine Vision in Food Quality Inspection

Peng Yin

School of Electronic and Information Engineering, University of Science and Technology Liaoning, Anshan, China
3473792229@qq.com

Keywords: Machine Vision; Food Quality Inspection; Appearance Defects; Packaging Integrity; Impurity Detection

Abstract: With the rise in consumer demands for food safety and quality, coupled with the development of automation in the food industry, traditional manual inspection has become insufficient. Machine vision technology, with advantages such as non-contact operation and high precision, has become a key detection method. This paper reviews the application of machine vision in food quality inspection, covering core scenarios such as appearance defect detection of fruits and vegetables, packaging integrity recognition, and impurity detection. It analyzes the technical principles and breakthroughs supported by datasets like Food-101 and Fruits-360, and discusses key challenges including complex shape detection, balancing speed and accuracy, and the cost of data labeling. Research shows that technical solutions integrating deep learning, multimodal fusion, and edge computing significantly enhance detection efficiency. In the future, lightweight models, self-supervised learning, and intelligent quality traceability are expected to become important development directions. This paper provides a technical reference for intelligent inspection in the food industry and promotes the deeper application of machine vision in full-chain quality control.

1. Introduction

With the continuous improvement of consumers' requirements for food safety and quality, as well as the rapid development of automated production in the food industry, traditional manual inspection methods have been difficult to meet the needs of modern food quality inspection due to low efficiency, strong subjectivity, and high missed detection rates. Machine vision technology, leveraging its advantages of non-contact, high precision, high speed, and automation, has become one of the core technologies in the field of food quality inspection. By image acquisition, processing, analysis, and understanding, this technology can accurately detect appearance defects, packaging integrity, and impurities of food, effectively improving the quality control level of food production. This paper discusses the application of machine vision in core scenarios such as the detection of appearance defects in fruits and vegetables, the identification of food packaging integrity, and impurity detection. It analyzes technical challenges combined with typical datasets such as Food-101[1] and Fruits-360, and looks forward to the future direction of rapid detection in production lines.

2. Technology and Application of Appearance Defect Detection in Fruits and Vegetables

2.1 Visual Inspection Principle for Insect Damage and Rot Defects

During growth, transportation, and storage, fruits and vegetables are prone to appearance defects such as insect damage and rot due to pests, diseases, or microbial erosion. These defects not only affect product value but also may pose food safety risks. The core process of machine vision inspection includes:

1) Image Acquisition: High-resolution CCD/CMOS cameras combined with annular light sources, backlighting, and other technologies are used to obtain images with clear surface details of fruits and vegetables, reducing the impact of uneven lighting on inspection.

2) Preprocessing and Feature Extraction: Noise is removed and contrast is enhanced through methods such as median filtering and histogram equalization. Texture features are extracted using gray-level co-occurrence matrices (GLCM) and local binary patterns (LBP), and defect regions are separated by color space conversion (e.g., RGB to HSV).

3) Defect Recognition and Classification: Traditional machine learning algorithms such as support vector machines (SVM) and random forests, or deep learning models such as convolutional neural networks (CNN) like ResNet and YOLO, are used to locate and classify irregular holes of insect damage and discolored patches of rot.

2.2 Typical Datasets and Technical Breakthroughs

The Food-101 and Fruits-360 datasets provide important support for the detection of fruit and vegetable defects. Food-101 contains color images of 101 types of food, covering various forms of fruits and vegetables, and is often used to train multi-category classification models. Fruits-360 focuses on fruit appearance, including multi-angle images of 120 types of fruits, with labeled defect types such as insect damage, scars, and rot, promoting the development of fine-grained defect recognition algorithms. In recent years, studies based on Fruits-360 have shown that the improved ResNet model combined with transfer learning can achieve a detection accuracy of 98.2% for rot defects, significantly outperforming the 85% missed detection rate of manual inspection.

2.3 Detection Challenges and Countermeasures Under Complex Morphologies

The irregularity of fruit and vegetable morphologies (such as changes in surface curvature and occlusion) makes it difficult for traditional algorithms to accurately segment defect regions. Research has shown that introducing 3D vision technology (such as structured light scanning) to obtain depth information, combined with semantic segmentation models (such as U-Net)[3] for multi-modal feature fusion, can increase the recall rate of apple insect damage detection from 89% to 96%. In addition, to address the missed detection of small-sized insect damage (diameter < 2mm), attention mechanisms[4] are used to enhance the model's sensitivity to local details, and data augmentation techniques (such as rotation, scaling, and Gaussian noise addition) are employed to improve model robustness, effectively solving the problem of detection accuracy under complex morphologies.

3. Technical System for Food Packaging Integrity Identification

3.1 Multi-Dimensional Inspection Content of Packaging Defects

Food packaging integrity inspection covers several key dimensions:

1) Appearance Defects: Including damage, stains, and sealing skew of packaging bags/boxes, which affect product sealing and visual experience.

2) Labels and Markings: Detecting label position offset, missing content (such as production date and shelf life), and blurred barcodes/QR codes to ensure compliance and traceability.

3) Sealing Performance: By inspecting the continuity of heat-sealing lines and residual air bubbles at the seal, preventing microbial invasion or content leakage.

3.2 Technical Implementation of Machine Vision in Packaging Inspection

Inspection systems usually integrate line-scan cameras or area-scan cameras, combined with high-speed image acquisition cards for dynamic capture. For transparent packaging (such as bottled beverages), backlighting is used to highlight internal defects; for metal or reflective packaging, diffused light sources are used to reduce reflection interference. At the algorithm level, template matching-based methods are used to detect label positions, edge detection algorithms analyze the integrity of seal contours, and optical character recognition (OCR) technology verifies printed information. For complex packaging textures, target detection algorithms in deep learning (such as Faster R-CNNand SSD[5])) can accurately locate tiny damages with a detection speed of 50 frames per second, meeting the real-time inspection needs of high-speed production lines (100 pieces per minute).

3.3 Practical Cases of Intelligent and Automated Inspection

A dairy enterprise adopted a machine vision system to inspect Tetra Pak packaging. Multi-camera arrays were used to obtain front, side, and top images of the packaging, and a multi-task learning model was utilized to simultaneously complete the detection of label alignment, seal heat-sealing line defects, and QR code recognition. The missed detection rate was less than 0.1%, and the efficiency was 300% higher than that of manual inspection. In addition, in the packaging inspection of quick-frozen food, infrared imaging technology was combined to identify residual ice slag inside the seal, making up for the shortcomings of visible light inspection and constructing a multi-modal fusion inspection system.

4. Technology and Application Scenarios of Food Impurity Detection

4.1 Foreign Body Types and Detection Technology Selection

Impurities in food mainly include inorganic impurities (stones, metal fragments), organic impurities (hair, insect residues), and other foreign bodies (plastic, glass shards). According to the characteristics of foreign bodies, detection technologies can be divided into:

1) Visible Light Detection: Suitable for foreign bodies with obvious color contrast (such as black hair on light-colored food surfaces), identified through threshold segmentation, edge detection, and other algorithms.

2) X-ray Detection: For stones, metals, and other substances with large density differences, X-ray penetration imaging is used to locate foreign bodies through grayscale value differences, which is widely used in the inspection of meat and baked food.

3) Hyperspectral Imaging: By combining spectral and image information and analyzing the differences in spectral reflection characteristics between foreign bodies and food, we can detect transparent or similar-colored foreign bodies (such as transparent plastic fragments).

4.2 Foreign Body Recognition Algorithms Under Complex Backgrounds

In food backgrounds with multi-texture and multi-color (such as mixed nuts and vegetable salads), traditional algorithms are prone to interference leading to false detection. Deep learning-based target detection frameworks (such as YOLOv5 and EfficientDet) fuse multi-scale information through feature pyramid networks (FPN), combined with attention mechanisms to suppress background noise, achieving a detection accuracy of 95% for hair larger than 0.5mm . In the impurity detection task of the Food-101 dataset, the improved Faster R-CNN model increased the recall rate of stone detection in complex food scenarios to 92%, significantly outperforming the traditional HOG+SVM algorithm (78%).

4.3 Engineering Challenges of Real-Time Inspection in Production Lines

High-speed production lines (such as potato chip production lines with a speed of 200 pieces per second) place strict requirements on the frame rate and computing power of inspection systems. Current solutions include: using GPU acceleration technology to compress single-frame processing time to less than 10ms, and combining model quantization and pruning technologies to reduce the number of parameters of neural networks, improving inference speed while maintaining accuracy. A real-time inspection system deployed by a puffed food enterprise achieved localized foreign body detection through edge computing nodes, with a delay controlled within 50ms, meeting the demand for real-time rejection of unqualified products in production lines.

5. Core Challenges of Machine Vision in Food Inspection

5.1 Detection Robustness Issues Caused by Irregular Food Morphologies

There are a wide variety of food types with diverse morphologies (such as the uneven surfaces of fruits and vegetables and the texture differences of meat products), resulting in insufficient generalization ability of the same detection algorithm across different categories. For example, defect detection models trained on planar images have a 10%-15% decrease in detection accuracy on curved fruits (such as oranges and apples). In addition, the random placement attitudes of food on conveyor belts (such as tilting and overlapping) increase the difficulty of image segmentation and feature matching. Traditional geometric feature descriptors (such as SIFT and ORB) have significantly reduced matching accuracy under complex attitudes.

5.2 The Dilemma of Balancing Detection Speed and Accuracy

High-speed production lines require inspection systems to have millisecond-level response capabilities, but high-precision deep learning models usually have high computational complexity and are difficult to meet real-time requirements. Taking the YOLOv3 model[2] as an example, the CPU inference time for foreign body detection on images with a resolution of 640×480 is about 200ms, which cannot meet the detection demand of 100 frames per second. In addition, under harsh environments such as low light and reflection, the decline in image quality leads to fluctuations in detection accuracy, further exacerbating the contradiction between speed and accuracy.

5.3 The Contradiction Between Data Annotation Cost and Model Generalization Ability

High-quality annotated data is the foundation for training high-performance models, but food defect samples (such as rot and insect damage) are diverse, and manual annotation is timeconsuming and labor-intensive. Taking the Fruits-360 dataset as an example, annotating 10,000 defect images requires more than 200 hours, and some rare defects (such as spots caused by specific pests) are scarce in samples, which is likely to cause model overfitting. In addition, the lighting conditions and shooting angles of different production lines vary greatly, and models need to be retrained for cross-scenario applications, increasing the cost of technology landing.

6. Future Outlook of Rapid Detection Technology in Production Lines

6.1 Deep Integration of Lightweight Models and Edge Computing

To meet real-time inspection needs, lightweight neural networks (such as MobileNet and ShuffleNet) are developed. Combined with knowledge distillation technology, the knowledge of large pre-trained models is transferred to lightweight models, increasing inference speed by 3-5 times while maintaining detection accuracy. At the same time, edge computing devices (such as industrial-grade GPU servers and embedded vision controllers) are deployed to realize localized operation of image acquisition, processing, and decision-making, reducing data transmission delay and constructing an intelligent inspection system with "terminal-edge-cloud" collaboration.

6.2 Application of Multi-Modal Fusion and 3D Vision Technology

Fusing multi-modal image information such as visible light, infrared, and hyperspectral, combined with 3D point cloud data to obtain spatial structure features of food, constructing a more comprehensive quality inspection model. For example, structured light 3D imaging technology is used to obtain height information of fruit and vegetable surfaces, and spectral data is combined to judge the degree of internal rot, realizing the joint inspection of appearance and internal quality. In addition, binocular vision or TOF (Time of Flight) cameras are introduced to obtain depth information, solving the problem of defect localization under irregular morphologies with the support of semantic segmentation technologies[6].

6.3 Self-Supervised Learning and Active Learning to Reduce Data Dependence

Self-supervised learning technologies (such as contrastive learning and masked image modeling) are adopted to pre-train models using a large amount of unannotated data, reducing dependence on manually annotated data. Combined with active learning algorithms, high-value samples are automatically selected for annotation, reducing annotation costs by more than 50%. For example, through uncertainty sampling strategies, fuzzy defect samples that are difficult for the model to judge are prioritized for annotation, improving data utilization efficiency.

6.4 Intelligent Quality Traceability and Production Optimization

Machine vision systems not only realize defect detection but also can build quality databases through long-term data accumulation, analyze the temporal, spatial, and type distribution rules of defects, and provide data support for production process optimization. For example, by counting the high-frequency occurrence periods of packaging seal defects, tracing back to the temperature fluctuation problem of heat-sealing equipment, and realizing preventive maintenance. Combined with digital twin technology, virtual simulation of the production line inspection process is carried out to verify the effectiveness of new algorithms in advance and shorten the technology landing cycle.

7. Conclusion

The application of machine vision technology in food quality inspection has evolved from single defect detection to a new stage of multi-scenario integration and intelligent inspection. By deeply integrating technologies such as deep learning, multi-modal perception, and edge computing, this technology has effectively solved core problems such as the detection of appearance defects in fruits and vegetables, packaging integrity, and impurity detection, significantly improving the quality control level of food production. However, facing challenges such as the diversity of food morphologies and the balance between detection speed and accuracy, continuous innovation is still needed in directions such as lightweight model design, multi-source data fusion, and self-supervised learning. In the future, with the continuous progress of technology, machine vision will be deeply coordinated with the Internet of Things, big data, and other technologies to build a full-process and full-chain intelligent inspection system, providing stronger technical support for food safety and quality assurance.

References

[1] Bossard, Lukas, Matthieu Guillaumin, and Luc Van Gool. "Food-101–mining discriminative components with random forests." European conference on computer vision. Cham: Springer International Publishing, 2014.

[2] Redmon J, Farhadi A. Yolov3: An incremental improvement[J]. arxiv preprint arxiv:1804.02767, 2018.

[3] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation[C]// International Conference on Medical image computing and computer-assisted intervention. Cham: Springer international publishing, 2015: 234-241.

[4] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[J]. Advances in neural information processing systems, 2017, 30.

[5] Liu W, Anguelov D, Erhan D, et al. Ssd: Single shot multibox detector[C]//European conference on computer vision. Cham: Springer International Publishing, 2016: 21-37.

[6] Garcia-Garcia A, Orts-Escolano S, Oprea S, et al. A review on deep learning techniques applied to semantic segmentation[J]. arxiv preprint arxiv:1704.06857, 2017.