Development and Validation of a Multispectral Vision System for In-Situ Detection of Pesticide Residues on Agricultural Produce

Guishen Ding

School of Computer Science, Beijing University of Information Science and Technology, Beijing 102206, China

Abstract: ***In response to the pressing demand for rapid and accurate detection of pesticide residues on the surface of fruits and vegetables, this study proposes an innovative visual detection system based on multispectral fusion. The system integrates an improved deep convolutional neural network (DCNN) with a Transformer architecture, enabling precise feature extraction from multispectral images. By leveraging the complementary strengths of these models, the system achieves enhanced classification accuracy and robustness against environmental variations. Furthermore, an adaptive weighted fusion algorithm, combined with improved kernel principal component analysis (KPCA), is employed to optimize the quality of spectral data fusion. This approach effectively mitigates the impact of noise and enhances the discriminative power of the features. To address the challenge of real-time detection, a parallel computing architecture based on CUDA and a heterogeneous hardware configuration scheme are designed. This implementation significantly accelerates the processing speed, enabling rapid analysis of large-scale datasets. Additionally, system stability is enhanced through redundant hardware design, intelligent fault-tolerant mechanisms, and optimized data transmission protocols. These measures ensure reliable operation under diverse environmental conditions, making the system suitable for field applications.***

Keywords: Multispectral fusion; Surface of fruits and vegetables; Visual inspection of pesticide residues; System optimization.

1. INTRODUCTION

Quality and safety of agricultural products is a vital issue concerning national welfare and people’s livelihood, among which the detection of pesticide residues on fruit and vegetable surfaces has long been a technical challenge in the industry. Traditional methods such as chromatographic analysis, though highly accurate, suffer from limitations such as expensive equipment, complex operation, and long detection cycles, making it difficult to meet the rapid and non-destructive testing needs of modern agriculture. In recent years, vision-based detection technology using multispectral imaging has shown promising application prospects due to its non-contact and high-efficiency characteristics; however, practical promotion still faces technical bottlenecks including insufficient detection accuracy, poor environmental adaptability, and suboptimal system stability. Grounded in multispectral fusion technology, this study innovatively combines deep learning methods with hardware system optimization to develop an efficient, accurate, and stable visual detection system for pesticide residues on fruits and vegetables. Chen et al. (2023) proposed a generative text-guided 3D vision-language pretraining framework for unified medical image segmentation[1], while in computer vision, Peng et al. (2025) developed a method exploiting the aggregation and segregation of representations for domain-adaptive human pose estimation[2]. Shifting to educational psychology, Yang (2022) investigated the role of aggression and burnout in Chinese EFL teachers' professional success[3], and Yang (2021) also researched the influence of scaffolding teaching on junior high school students' English reading ability[4]. In cultural studies, Yang and Mustafa (2025) explored the reception of multimodality in translating Chinese museum culture in the intelligent media era[5]. AI applications in critical infrastructure are demonstrated by Huang, Tian, and Qiu (2025), who created an AI-enhanced dynamic power grid simulation for real-time decision-making[6]. The financial sector benefits from models like FinStack-Net by Cheng et al. (2025), which uses hierarchical feature crossing and stacked ensemble learning for fraud detection[7], and Su et al. (2025)'s WaveLST-Trans model for anomaly detection and early warning in financial time series[10]. Network management is advanced by Zhang et al. (2025) with MamNet, a hybrid model for network traffic forecasting and frequency pattern analysis[11], and green finance is supported by Zhang, Li, and Li (2025), who leverage deep learning for carbon market price forecasting and risk evaluation[12]. In advertising, Tian et al. (2025) propose a cross-attention multi-task learning approach for ad recall[13]. Urban planning is accelerated by Xu's (2025) UrbanMod, a text-to-3D modeling framework for city architecture[14], and industrial reliability is enhanced by Tan et al. (2024) through transfer learning in densely connected convolutional networks for fault diagnosis[15]. Finally, strategic digital transformation is addressed by Zhuang (2025) in constructing real estate marketing strategies[16], and personalized systems are advanced by Han and Dou (2025) with a user recommendation method integrating a hierarchical graph attention network with a multimodal knowledge graph[17].

2. PRINCIPLES OF MULTISPECTRAL FUSION TECHNOLOGY

Multispectral fusion technology integrates spectral information from different bands to enable multidimensional detection of pesticide residues on fruit and vegetable surfaces. Its core lies in using spectral imaging to capture the reflectance and fluorescence properties of specific bands (visible to near-infrared). For example, in the visible range it mainly reflects pigment distribution and surface morphology, whereas the near-infrared band is highly sensitive to the vibrations of organic molecules and can capture the characteristic absorption peaks of pesticide molecules. By coordinating spectroscopic devices with high-sensitivity CCD/CMOS sensors, simultaneous acquisition of spatial and spectral multidimensional information is achieved, laying the foundation for subsequent feature extraction.

In terms of spectral fusion, this project focuses on an adaptive weighted fusion method at the feature level. First, dark-current correction, radiometric correction, and geometric registration are used to eliminate equipment noise and spatial drift. On this basis, PCA, ICA and other methods reduce the dimensionality of crop-residue features to extract sensitive spectral bands. The innovation of this project is the introduction of an attention mechanism into the band-weight allocation model, using convolutional neural networks to automatically learn the contribution of each band to pesticide residues, thereby achieving complementary advantages between visible-light texture features and NIR molecular features. Compared with traditional simple weighting or band-stacking methods, this approach effectively suppresses data redundancy and improves the detection signal-to-noise ratio [1].

Its advantages are:

(1) Multidimensional feature fusion overcomes the limitations of single-spectrum detection, e.g., organophosphorus pesticides show no features in 550nm but have strong near-infrared absorption.

(2) Spatial-spectral joint analysis identifies interference between pesticide residues and natural fruit/vegetable textures.

(3) A local contrast-enhancement algorithm highlights micron-scale residue spots.

(4) The fused feature vectors are more suitable for machine-learning models, enabling efficient classification via SVM and lightweight mobile neural networks. Through this project, we will focus on solving the real-time issues of band selection and fusion algorithms, making them applicable to pipeline-style dynamic detection scenarios and providing key technical support for building low-latency, high-precision online detection systems.

3. PROBLEM ANALYSIS OF THE VISUAL DETECTION SYSTEM FOR PESTICIDE RESIDUES ON FRUIT AND VEGETABLE SURFACES BASED ON MULTISPECTRAL FUSION

3.1 Detection Accuracy Issues

At present, multispectral-fusion-based methods for detecting pesticide residues in fruits and vegetables still face many accuracy issues. In feature extraction, the complex surface textures, natural defects, and uneven distribution of pesticide residues make it difficult for traditional hand-crafted approaches to separate target features from background interference. In the visible range, the produce’s own color variations easily mask the visual characteristics of trace pesticides, while the near-infrared bands sensitive to certain organics are strongly affected by moisture content and surface gloss. Against such a complex background, simple edge detection and texture analysis often fail to characterize features effectively, thereby degrading subsequent classification performance.

In multispectral remote-sensing image fusion, there is also room for improvement. Existing fusion methods such as principal component analysis and band-weighted averaging can reduce dimensionality, yet they tend to discard narrow-band spectral information that is sensitive to specific pesticides when selecting features. For example, some OPs exhibit characteristic absorption peaks in particular narrow bands, but conventional broad-band fusion approaches may weaken this trait. Moreover, the redundancy among spectral channels has not been adequately addressed, wasting computational resources and reducing detection timeliness. Building an adaptive weight-allocation mechanism that intelligently screens and complementarily fuses key spectral features is a crucial breakthrough for enhancing system performance [2].

Additionally, the quality of training-sample labeling is a major factor limiting detection accuracy. Pesticide-residue detection relies on high-precision labeled data, yet current labeling relies mainly on manual interpretation, which is strongly influenced by operator experience and the sensitivity of detection equipment, resulting in high subjectivity. Especially when residue concentrations are low or multiple pesticides are mixed, manual labeling struggles to ensure consistency. Such label noise causes the model to learn erroneous feature associations, leading to poor generalization in real applications. Exploring semi-supervised data-optimization methods or combining high-precision mass-spectrometry results offers an effective route to improving data quality.

3.2 System Efficiency Issues

The fruit and vegetable pesticide residue detection system based on multispectral fusion technology suffers from low system efficiency in practical applications. Current mainstream algorithms, when handling high-dimensional multispectral data, typically undergo complex processes such as feature extraction, dimensionality reduction, and classification, causing computational load to grow exponentially. Especially in deep learning algorithms, the large number of parameters and heavy matrix computations mean that even with GPU acceleration, the detection time for a single sample still struggles to meet the real-time demands of pipeline operations. In batch testing scenarios, this computational delay increases further, severely impacting overall detection efficiency. However, existing algorithms mostly adopt fixed computational frameworks and cannot dynamically adjust the algorithm according to actual needs, leading to resource waste [3].

Unreasonable hardware resource allocation is another prominent issue. Multispectral imaging systems usually contain multiple optical sensors and data processing units, yet actual deployments often suffer from imbalanced resource allocation. The main manifestations are: mismatched image acquisition and computation speeds, causing data accumulation or processor idling; when traditional algorithms are run on GPUs, GPU utilization is greatly restricted. Such uneven resource distribution not only reduces system throughput but also increases energy consumption and shortens equipment lifespan. More importantly, existing architectures mostly rely on general-purpose computing platforms, which cannot be specifically optimized for the characteristics of multispectral data processing and thus fail to fully exploit hardware potential.

During long-term continuous operation, system stability issues become more pronounced. Tiny deviations in the optical components of multispectral cameras under detection environments cause spectral data drift. Meanwhile, during high-volume data transmission, electromagnetic interference can lead to packet loss or checksum errors. At the device level, prolonged high-load operation causes poor heat dissipation in memory and memory leakage, ultimately resulting in system crashes or reduced detection accuracy. These defects are hard to detect under laboratory conditions, but in real-world applications they seriously affect system reliability and user experience. Building an architecture with self-diagnosis, adaptability, and robustness is currently a key technical challenge that urgently needs to be addressed.

4. OPTIMIZATION STRATEGIES FOR THE VISUAL DETECTION SYSTEM OF PESTICIDE RESIDUES ON FRUIT AND VEGETABLE SURFACES BASED ON MULTISPECTRAL FUSION

4.1 Strategies for Improving Detection Accuracy

The innovative application of deep learning is of great significance for improving the accuracy of multispectral fusion-based detection of pesticide residues in fruits and vegetables. Addressing the weak response of traditional feature-extraction methods to small pesticide-residue regions, this project constructs a multiscale convolutional neural network based on attention mechanisms to locally enhance weak pesticide-residue signals. Specifically, a Transformer architecture is introduced for processing spectral sequence data; its self-attention mechanism effectively captures long-range dependencies among different spectral bands and outperforms traditional CNNs in identifying pesticide molecules. Given the limited sample size in real-world applications, the project will adopt a transfer-learning strategy, fine-tuning the model on large-scale multispectral datasets to adapt it to pesticide-residue detection, thereby preserving feature-representation capacity while avoiding overfitting.

Optimizing spectral-data fusion methods requires breaking through the limitations of traditional fixed-weight fusion. Accordingly, we design a dynamic weighted fusion algorithm based on feature importance, analyzing the sensitivity differences of various bands to the target pesticide and automatically assigning weights. For 0Ps , we will increase the weight coefficient of characteristic absorption bands and employ an improved kernel principal component analysis for nonlinear dimensionality reduction of spectral data, preserving the most critical spectral features. This adaptive fusion strategy effectively resolves information redundancy while preventing the loss of key features. On this basis, the project will introduce a multimodal learning framework to achieve cross-modal correlation between spectra and texture, constructing a more comprehensive pesticide-residue representation model.

To improve annotation quality, a standardized labeling system must be established. On the one hand, a complete pesticide-residue labeling specification will be developed, clarifying standards for different concentration levels and distribution patterns; on the other hand, high-precision microscopic imaging will assist annotation to reduce human error. Building on this, we will study a semi-automatic active-learning-based annotation system that iteratively trains on uncertain samples, screens them, and submits them to experts for review, significantly improving labeling efficiency. Meanwhile, data augmentation via generative adversarial networks will enhance sample diversity while preserving spectral authenticity, boosting model generalization. The project will combine manual labeling with intelligent algorithms to overcome the inconsistency of traditional methods, providing more reliable data support for model training [4].

4.2 Strategies for Improving System Efficiency

To enhance the performance of the multispectral-fusion pesticide-residue detection system, breakthroughs are required in both computational framework and parallel processing. This project will adopt a modular design philosophy, decoupling feature extraction, data fusion, and classification into three independent stages and implementing GPU parallel acceleration under the CUDA framework. In particular, for compute-intensive operations such as convolution and matrix decomposition, throughput will be significantly increased by optimizing thread-block allocation and shared-memory usage. Building on this, the project proposes a dynamic computational-path selection mechanism that automatically adjusts algorithmic depth according to input complexity, avoiding unnecessary complex operations on simple samples and enabling intelligent allocation of computational resources.

Hardware-structure optimization must be founded on a co-design philosophy between algorithms and hardware. Tailored to the characteristics of multispectral data processing, the project will adopt a heterogeneous computing architecture that combines a high-performance image-capture card with GPU compute units to achieve real-time preprocessing, with the GPU handling deep-learning inference. In the storage subsystem, a tiered-storage strategy will be employed: frequently accessed spectral-feature information will reside in GPU memory, while historically detected data will be stored in high-capacity SSDs. On this basis, hardware-parameter templates will be designed; key parameters such as image-sensor gain and exposure time will be dynamically adjusted via software-defined methods to achieve optimal scheduling of hardware resources.

To improve system stability, efforts must be made in two areas: hardware redundancy and software fault tolerance. On the hardware side, dual redundant power supplies and hot-swappable modules ensure continued operation even when critical components fail. For data transmission, the original USB or CameraLink interfaces have been optimized, and a retransmission mechanism with data verification and flow-control algorithms has been introduced to prevent detection interruptions caused by packet loss. The software integrates real-time health monitoring that analyzes operating parameters (e.g., temperature, memory usage) to predict potential failures and performs adaptive graceful degradation when anomalies occur. Meanwhile, a log-analysis system mines historical data to continuously optimize system parameters, creating a virtuous cycle of stable improvement.

5. CONCLUSION

Multispectral fusion technology offers clear advantages in pesticide-residue detection for fruits and vegetables, yet practical applications still face challenges in detection accuracy, system efficiency, and environmental adaptability. To address these issues, this project will conduct systematic research from four aspects: deep-learning algorithm optimization, adaptive spectral fusion strategies, construction of an intelligent data-labeling system, and co-design of hardware architectures. Specifically, the project will boost detection speed through algorithmic parallel reconstruction and innovative use of heterogeneous computing architectures, while introducing redundancy design and intelligent fault-tolerance mechanisms to enhance system robustness. The outcomes will not only break through current technical bottlenecks but also provide new ideas for advancing the intelligent development of agricultural product quality and safety testing.

REFERENCES

[1] Chen, Yinda, et al. "Generative text-guided 3d vision-language pretraining for unified medical image segmentation." arXiv preprint arXiv:2306.04811 (2023).

[2] Peng, Qucheng, Ce Zheng, Zhengming Ding, Pu Wang, and Chen Chen. "Exploiting Aggregation and Segregation of Representations for Domain Adaptive Human Pose Estimation." In 2025 IEEE 19th International Conference on Automatic Face and Gesture Recognition (FG), pp. 1-10. IEEE, 2025.

[3] Yang, Dan. "Unpacking the role of Chinese EFL teacher aggression and burnout in their professional success: A teachers’ psychology perspective." Frontiers in psychology 13 (2022): 1001252.

[4] Dan, Y. A. N. G. "A research on the influence of scaffolding teaching on junior high school students’ English reading ability." World Journal of English Language 11.2 (2021): 1-43.

[5] Yang, Chunli, and Siti Ezaleila Mustafa. "The Reception Studies of Multimodality in the Translation and Communication of Chinese Museum Culture in the Era of Intelligent Media." Cultura: International Journal of Philosophy of Culture and Axiology 22.4 (2025): 532-553.

[6] Huang, Jingyi, Zelong Tian, and Yujuan Qiu. "AI-Enhanced Dynamic Power Grid Simulation for Real-Time Decision-Making." (2025).

[7] Cheng, Zhang, et al. "FinStack-Net: Hierarchical Feature Crossing and Stacked Ensemble Learning for Financial Fraud Detection." (2025).

[8] Peng, Qucheng, Ce Zheng, Zhengming Ding, Pu Wang, and Chen Chen. "Exploiting Aggregation and Segregation of Representations for Domain Adaptive Human Pose Estimation." In 2025 IEEE 19th International Conference on Automatic Face and Gesture Recognition (FG), pp. 1-10. IEEE, 2025.

[9] Chen, Yinda, et al. "Generative text-guided 3d vision-language pretraining for unified medical image segmentation." arXiv preprint arXiv:2306.04811 (2023).

[10] Su, Tian, et al. "Anomaly Detection and Risk Early Warning System for Financial Time Series Based on the WaveLST-Trans Model." (2025).

[11] Zhang, Yujun, et al. "MamNet: A Novel Hybrid Model for Time-Series Forecasting and Frequency Pattern Analysis in Network Traffic." arXiv preprint arXiv:2507.00304 (2025).

[12] Zhang, Zongzhen, Qianwei Li, and Runlong Li. "Leveraging Deep Learning for Carbon Market Price Forecasting and Risk Evaluation in Green Finance Under Climate Change." Journal of Organizational and End User Computing (JOEUC) 37.1 (2025): 1-27.

[13] Q. Tian, D. Zou, Y. Han and X. Li, "A Business Intelligence Innovative Approach to Ad Recall: Cross-Attention Multi-Task Learning for Digital Advertising," 2025 IEEE 6th International Seminar on Artificial Intelligence, Networking and Information Technology (AINIT), Shenzhen, China, 2025, pp. 1249-1253, doi: 10.1109/AINIT65432.2025.11035473.

[14] Xu, Haoran. "UrbanMod: Text-to-3D Modeling for Accelerated City Architecture Planning." Authorea Preprints (2025).

[15] Tan, C., Gao, F., Song, C., Xu, M., Li, Y., & Ma, H. (2024). Highly Reliable CI-JSO based Densely Connected Convolutional Networks Using Transfer Learning for Fault Diagnosis.

[16] Zhuang, R. (2025). Evolutionary Logic and Theoretical Construction of Real Estate Marketing Strategies under Digital Transformation. Economics and Management Innovation, 2(2), 117-124.

[17] Han, X., & Dou, X. (2025). User recommendation method integrating hierarchical graph attention network with multimodal knowledge graph. Frontiers in Neurorobotics, 19, 1587973.