Frontiers in Computing and Intelligent Systems
ISSN: 2832-6024 | Vol. 11, No. 1, 2025

Design and Implementation of an Automatic White Sugar Impurity Detection System based on Multispectral Imaging

Xin Zhong [1], Li Luo [1,*], Xiangyang Yu [2], Yefan Cai [3], Weibing Hong [3]

[1] School of Physics & Optoelectronic Engineering, Guangdong University of Technology, Guangzhou Guangdong, 510000, China
[2] School of Physics, Sun Yat-sen University, State Key Laboratory of Optoelectronic Materials and Technologies, Nanchang Research Institute, Guangzhou Guangdong, 510000, China
[3] Guangzhou Guangxin Technology Co., Ltd., Guangzhou Guangdong, 510000, China

*** Corresponding author:** Li Luo

Abstract: Traditional methods for removing impurities from sugar are not very inefficient and have limited accuracy. This study designed and implemented an automatic white sugar impurity detection system based on multi-modal image fusion technology to detect impurities in white sugar efficiently and accurately. The system utilizes 4 multi-spectral cameras and an RGB camera to acquire high-resolution images of the samples. By extending the input channels of the Attention U-Net model to thirteen, the recognition rate of impurity categories is improved by approximately 60%, significantly enhancing the detection performance of the model. The experimental results show that the system performs excellently in the image segmentation task, achieving an average intersection over union (mIoU) of 92.70% and an average pixel accuracy (mPA) of 93.74%. The accuracy for impurity categories reaches 95%. Finally, the model was applied to test the effectiveness in identifying impurities; the results indicate that the model can effectively identify various types of impurities and also perform well in recognizing unknown categories.

Keywords: Multi-modal Image Fusion; Attention Unet; Multispectral Images; Impurity Detection; Deep Learning.

1. Introduction

Food safety and quality control are critical global concerns. As food production processes and supply chains become increasingly complex, impurity detection is more important than ever. Traditional methods, such as physical and chemical laboratory analyses, are accurate but time-consuming and destructive, making them unsuitable for rapid and convenient applications [1].

As one of the most commonly used sweeteners [2], sugar is widely applied in the food and beverage industries. The sugar available on the market today is primarily derived from sugarcane or sugar beet. During production and storage, sugar is prone to contamination by impurities. Traditional impurity removal methods rely on the sugar's water solubility, requiring repeated processing with limited effectiveness. Remaining impurities often require manual sorting, resulting in high detection costs and inefficiency. Research on sugar has mainly focused on qualitative analysis, sugar content, and particle size measurement [3,4], with fewer studies addressing impurity detection.

Spectral vision, which combines imaging and spectroscopy, enables simultaneous acquisition of spatial and spectral information from samples. Depending on resolution, it can be categorized into hyperspectral and multispectral imaging. This technology not only provides detailed spatial information but also reveals the chemical composition of materials, making it highly suitable for food detection. For example, multispectral imaging has been used to detect mango chilling damage [5] and non-smoking materials in cigarettes [6], achieving high detection accuracy in complex scenarios.

Based on the background, this study proposes an automated sugar impurity detection system using multimodal image fusion technology. The system employs a multispectral and an RGB cameras for image capture, along with a deep learning model (Attention U-Net) for automated impurity detection. Multispectral imaging enables precise composition analysis, while RGB imaging captures object appearances. Compared to traditional methods, this system offers significant advantages in non-destructive testing and real-time analysis, improving detection efficiency and accuracy.

2. Overall Framework Design of the System

2.1. Functional Design of the System

This study develops an automated sugar impurity detection system based on spectral imaging and deep learning, achieving efficient, accurate, and real-time detection to meet the quality control demands in complex production environments. The system integrates multispectral image processing techniques and deep learning models, enabling rapid impurity identification under complex backgrounds while ensuring the stability and accuracy of detection results. The system is user-friendly, requiring no specialized knowledge for operation, and offers flexibility and scalability for future upgrades in line with technological advancements and market demands. The system comprises the following key modules, with the system architecture illustrated in Figure 1:

(1) Imaging System:

Before operation, the camera's exposure time and focal length need to be adjusted and set appropriately. The image capture process can be controlled via software, allowing the system to save the acquired image data. During system operation, real-time image acquisition from the conveyor belt is required. The software is capable of displaying both the captured images and the results of model analysis in real time, providing continuous monitoring and feedback throughout the detection process.

(2) Computer:

The computer serves as the master control unit, allowing software-based control of the conveyor belt system, image acquisition, and display of the model processing results. It manages the entire detection process, facilitating seamless coordination between the conveyor belt, image capture, and real-time presentation of analysis outcomes.

(3) PLC control system:

The Programmable Logic Controller (PLC) serves as the slave unit, communicating with the master control unit (such as a computer). It receives command signals from the master unit and executes the corresponding control functions based on these instructions, transmitting control signals to the subordinate devices.

(4) Conveyor Belt System:

To achieve automated and real-time impurity detection, the system must ensure that samples are detected while dynamically moving along the conveyor belt. Accurate control of the conveyor belt's speed and direction is crucial, as it not only affects data acquisition efficiency but also plays a key role in the subsequent impurity removal process. The conveyor belt parameters must be carefully set to meet the demands of high-speed, high-quality data collection while providing sufficient time and space for impurity selection, thereby improving the overall efficiency and accuracy of the production line.

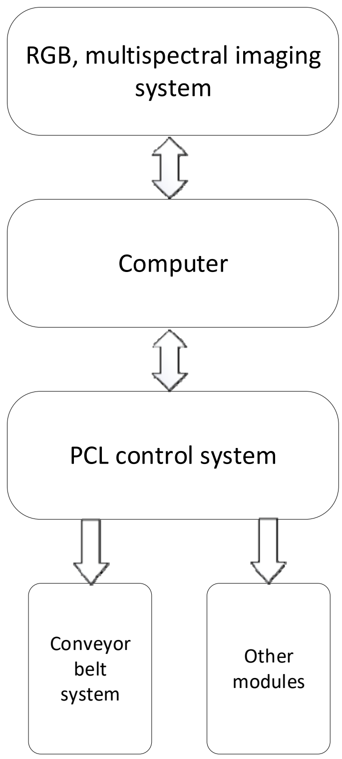


Fig 1. System Architecture Diagram

2.2. System Hardware Design

In this study, spectral image data of sugar and various impurities were collected using the spectral system shown in Figure 2. The system consists of four multispectral cameras and an RGB camera. The multispectral cameras are designed with ten bands, covering wavelengths from 713 nm to 920 nm, with a wavelength resolution of 23 nm. Their compact size, low cost, and high performance make them ideal for on-site rapid detection applications. The RGB camera is equipped with a wide-angle lens to capture image information over a larger field of view. The system also includes four 50W halogen lamps as the lighting source. The conveyor belt, made of white food-grade PU material, serves as the background for material transport and also functions as a reference whiteboard for spectral image data calibration.

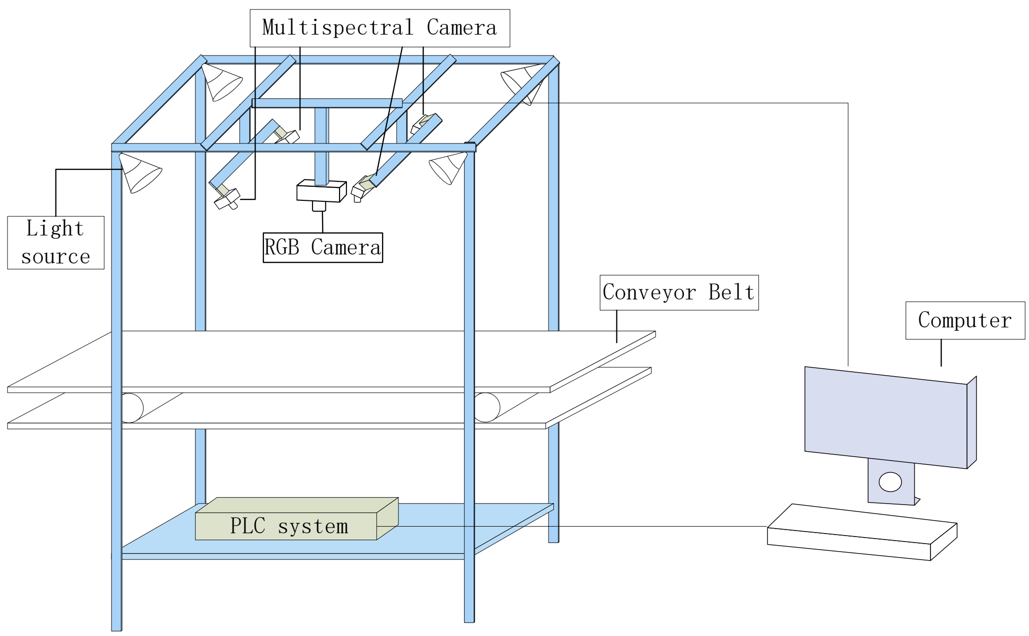


Fig 2. Hardware Architecture of the Sugar Impurity Detection System

The workflow of the system is shown in Figure 3. Initially, the hardware setup is debugged and adjusted, including the calibration of the cameras and light sources. Next, the whiteboard and calibration board are captured to obtain the whiteboard data and the geometric parameters of the camera. These data are used during the preprocessing stage to calibrate the acquired spectral images, ensuring both image quality and data consistency. After preprocessing, the data are used for model training. Finally, the results are displayed on the software interface, showing the processed images.

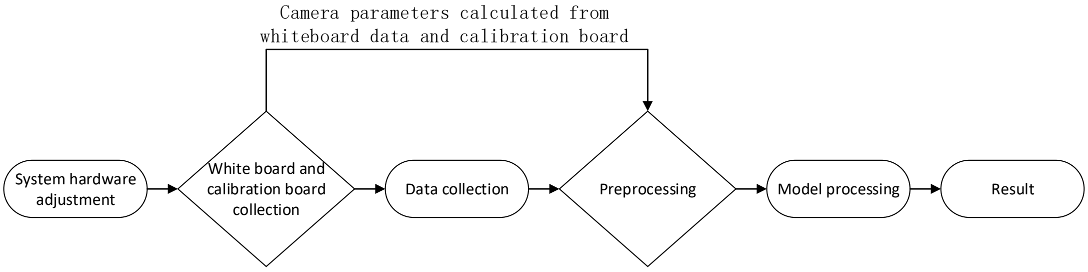


Fig 3. System Workflow Diagram

3. Data Preparation and Processing

3.1. Experimental Materials

White sugar and impurity samples sourced from a leading domestic producer were used to replicate real-world production conditions. Impurities included common foreign objects such as yellow sugar, hair strands, rubber fragments, packaging debris, black rubber, blue threads, and metals, providing diverse and challenging scenarios for testing. Some of these impurities are shown in Figure 4.

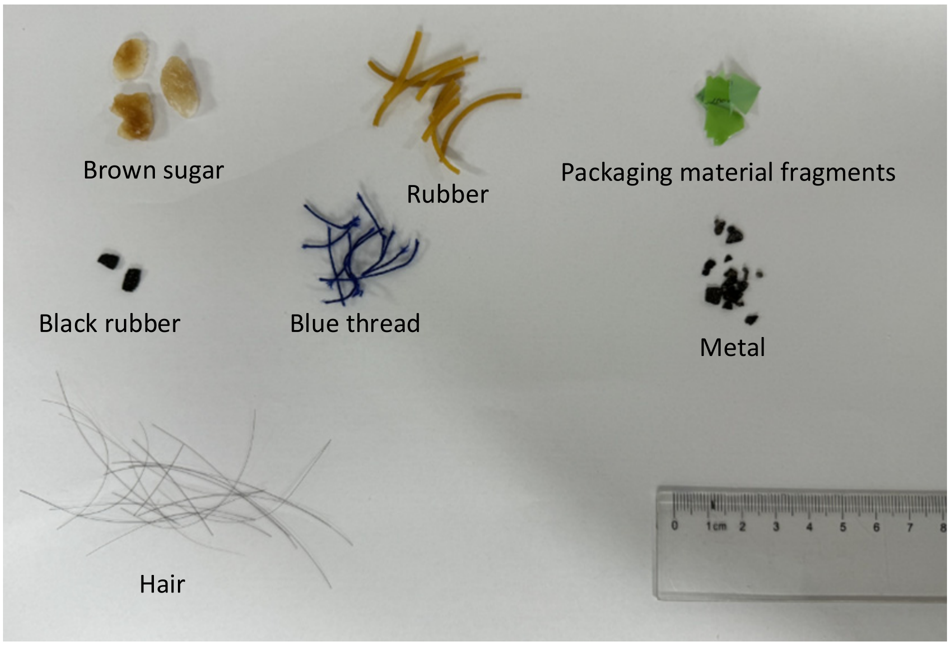


Fig 4. Color Images of Different Types of Impurities

3.2. Experimental Data Preparation

In this experiment, impurities were randomly selected and mixed with white sugar to simulate the randomness and uncontrollability of impurities in actual production. Although the impurity content in real-world production is typically lower than the settings used in this experiment, a higher impurity ratio was intentionally chosen to ensure the applicability of the detection method under varying impurity content conditions. This design helps to verify the robustness of the proposed method in complex scenarios, providing more reliable technical support and assurance for impurity detection in real-world production environments.

3.2.1. Data Collection

Before data collection, the zoom ring of the RGB camera's wide-angle lens must be adjusted to focus on the plane where the samples are located. Additionally, the exposure time for both the multispectral and RGB cameras needs to be fine-tuned to avoid overexposure or underexposure, which could affect the experimental results.

During the formal data collection process, the first step is to capture the data of a white reference board and a calibration plate for system calibration. This ensures the accuracy and consistency of the spectral data. After calibration, white sugar is mixed with impurities, and data collection begins. Once a set of samples is photographed, the samples must be promptly replaced for the next round of data collection. This process ensures that multiple different sample combinations are captured, improving the model's generalization ability and detection performance.

3.2.2. Spectral Data Preprocessing

During the imaging process, spectral data can be influenced by various factors, making data preprocessing particularly important. In this study, we performed several preprocessing steps on the spectral image data, including cropping, spectral calibration, image registration, standard normalization, and multiple scattering correction. Additionally, we utilized image fusion techniques to combine high-resolution images captured by the RGB camera with spectral images obtained from the multispectral camera.

The processed image data cube contains 13 channels: the first three channels represent the RGB image information, while the remaining 10 channels correspond to the multispectral image data at 10 different wavelengths in the near-infrared region. These images provide a solid foundation for subsequent analysis and model training, ensuring that the model can effectively learn and improve detection accuracy.

Spectral calibration is performed to eliminate the effects of sensor baseline signals and lighting factors, ensuring the accuracy of the spectral reflectance in the images. After each whiteboard data collection, the camera's built-in program calibrates the images based on the following formula (1) to obtain the corrected image:

$$\text{Reflenctance}=\frac{RawData^{t1}-Dark^{t1}}{(White^{t2}-Dark^{t2})/r}\times\frac{t2}{t1} \tag{1}$$

In the above equation, Reflectance represents the corrected multispectral data, RawData refers to the raw data before calibration, and Whiteis the standard reference board data. $t1$ and $t2$ are the exposure times for capturing the sample and the whiteboard, respectively. In this study, both $t1$ and $t2$ are set to 50 ms. Dark^{t1} and Dark^{t2} are the dark current values automatically recorded by the camera when capturing the sample and the whiteboard. The r is the reflectance of the standard whiteboard.

Image registration is necessary because different cameras may have variations in viewpoint and optical characteristics during imaging, which results in inconsistencies in the position, shape, and size of the same object across their respective images. The process is shown in Figure 5. In this study, the AprilGrid calibration board was used, as shown in Figure 6. The AprilGrid calibration board combines the features of traditional checkerboard grids and AprilTag markers, offering advantages such as high-precision positioning, strong robustness, ease of automation, and applicability in various scenarios.

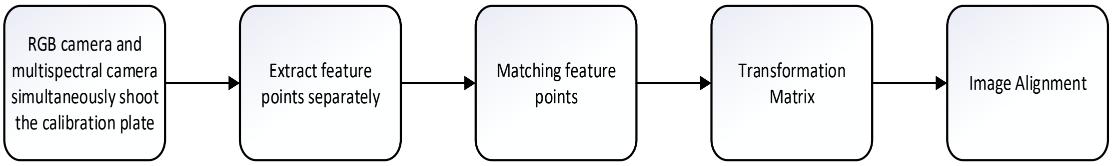


Fig 5. Image Registration Workflow Diagram

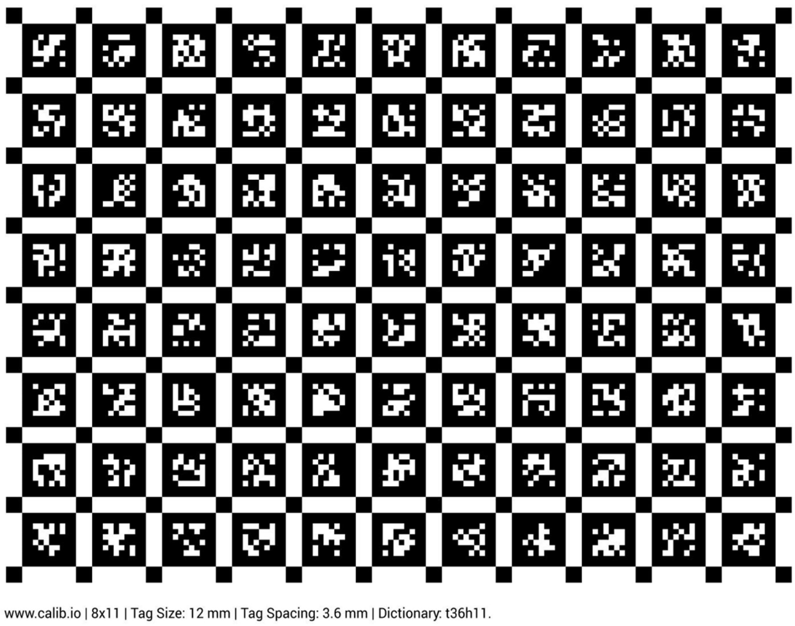


Fig 6. AprilGrid Calibration Board

3.2.3. Spectral Data Annotation

After the data preprocessing was completed, the images were annotated point by point using LabelMe software [7]. This tool provides flexible and efficient image annotation capabilities, helping us quickly and accurately perform image segmentation and labeling. Using LabelMe, JSON files containing the coordinates of impurity regions were generated. Additionally, scripts were used to process these JSON files, automatically generating a visual annotated image (visual.png), which provided an intuitive and precise reference for subsequent model training.

4. Introduction to the Attention Unet Model and Result Analysis

U-Net was originally proposed for segmenting neuronal structures by training an end-to-end fully convolutional network on electron microscope images. It is a classic image segmentation model. With its encoder-decoder architecture, U-Net is capable of capturing both global and local information at multiple scales and has been widely used in medical image segmentation and natural image segmentation tasks [8].

Attention U-Net is an improved version of the classic U-Net that incorporates an attention mechanism, as shown in Figure 7 [9]. The attention mechanism enables the network to automatically focus on more important feature regions, especially when the contrast between the target and background is low or the boundaries are unclear. This mechanism enhances the ability to capture critical areas, making it particularly suitable for organ segmentation in medical imaging [10].

In the task of white sugar impurity detection, the impurities are typically small and have a low contrast with the background (white sugar). The attention mechanism of Attention U-Net can more precisely focus on the impurity regions, effectively improving the accuracy and robustness of detection. Through this mechanism, the network is able to better ignore irrelevant background information and highlight the features of the impurities, significantly enhancing segmentation performance. This makes Attention U-Net more effective in recognizing and locating impurities in complex scenarios, providing reliable support for automated detection.

In the white sugar impurity detection task of this study, due to the introduction of a ten-band multispectral camera, the model was modified to handle a thirteen-channel multispectral image input, instead of the traditional single-channel grayscale or three-channel RGB images. This modification allows the model to fully utilize the multispectral image data, enhancing its ability to detect and identify impurities in white sugar more effectively.

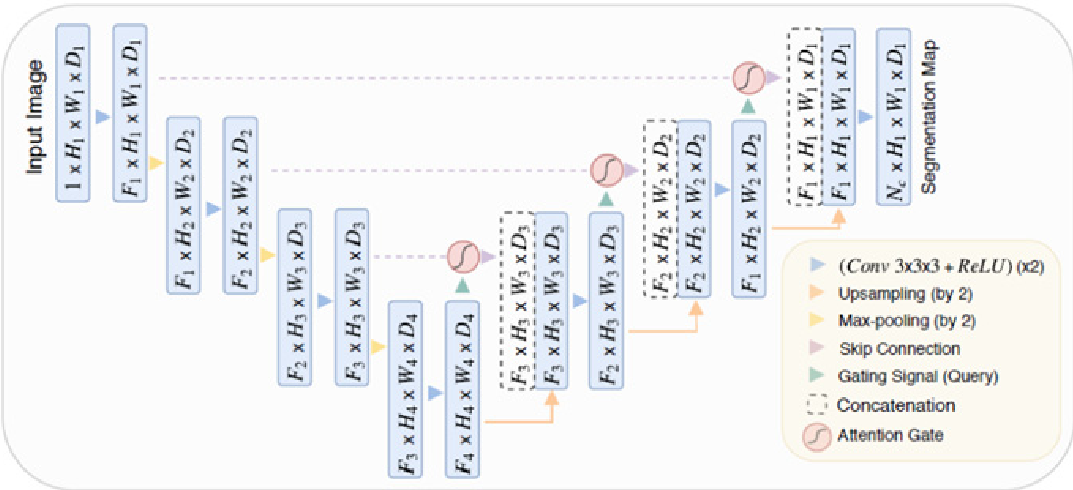


Fig 7. Attention Unet Model Structure

4.1. Model Evaluation Metrics

To comprehensively evaluate the model's performance in the sugar impurity detection task, this study uses several common segmentation evaluation metrics, including accuracy, mean pixel accuracy (MPA), mean intersection over union (MIoU), and dice coefficient.

4.2. Analysis of Model Comparison results

During the training process, we recorded the model's loss and Dice coefficient, as shown in Figure 8, to evaluate the model's performance. These metrics provide us with an intuitive understanding of the model's training effectiveness and help in further optimizing the model's parameters and training strategy.

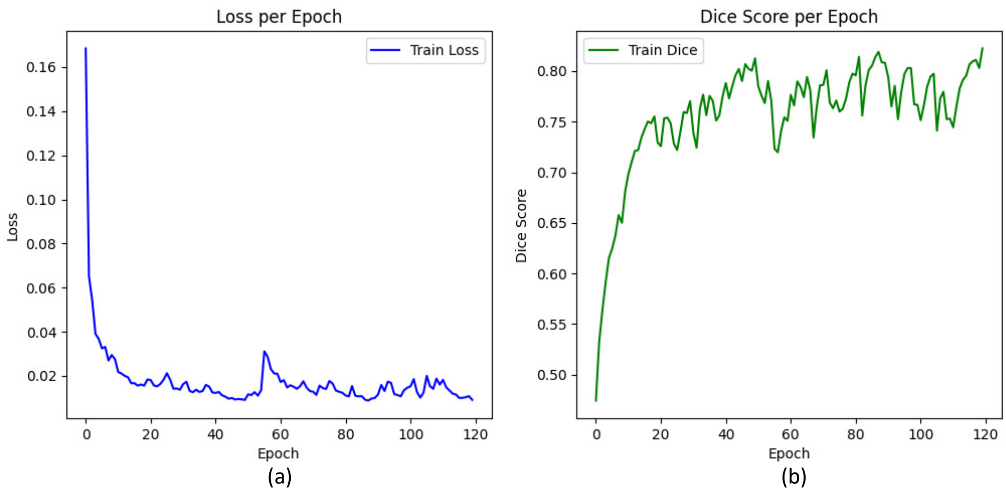


Fig 8. (a) Loss during the training process and (b) Real-time changes in the Dice coefficient

As seen from the figure above, the loss value gradually decreases during the training process, while the Dice coefficient steadily increases, indicating that the model exhibits good convergence and improving segmentation performance. These results suggest that after multiple training rounds, the model can effectively adapt to the data and enhance its ability to detect and segment impurities in white sugar.

In this experiment, we optimized the Attention Unet model. Traditional Attention Unet typically uses a single RGB image as input, whereas in this experiment, we modified the input channels to thirteen images to better suit our dataset. During the training process, the dataset was divided into training, validation, and test sets in a 60%, 20%, and 20% ratio, respectively, for model training, parameter tuning, and performance evaluation. Additionally, we performed data processing for both types of input and conducted a detailed analysis of the results for each category as well as the overall results, as shown in Table 1.

Table 1. Recognition results of different categories under models with different input image numbers (All refers to the overall results of the model for sugar and impurities)

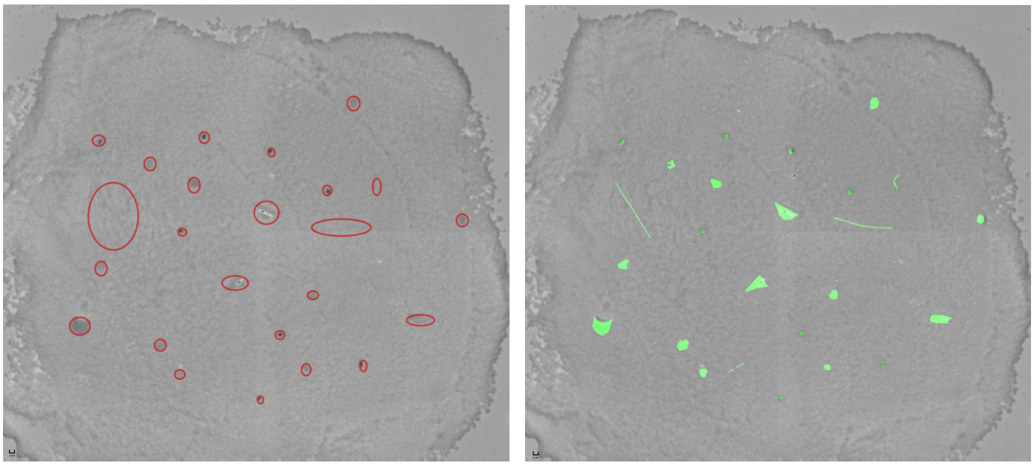
Input image category	Category	Mean Intersection over Union (mIoU) %	Mean Pixel Accuracy (mPA) %	Accuracy %
1 RGB image	Background	99.66	99.67	99.80
	Impurities	31.22	94.16	31.84
	All	66.44	96.91	99.66
R, G, B, and ten multispectral images	Background	99.97	99.97	99.97
	Impurities	95.42	87.50	95.17
	All	92.70	93.74	99.97

From the table above, it can be observed that when the input is a single RGB image, the model performs excellently in the background (i.e., white sugar), with the mean intersection over union (mIoU), mean pixel accuracy (MPA), and overall accuracy all exceeding 99%. However, the model's performance in impurity detection is unsatisfactory, with both the mean intersection over union and accuracy falling below 50%. Despite the high overall accuracy, this is primarily due to the much higher proportion of white sugar compared to impurities, which skews the overall accuracy.

In contrast, the improved Attention Unet model shows significant improvement in both background and impurity detection. Particularly in impurity detection, the mean intersection over union (mIoU) increased to 95.42%, and the overall accuracy reached 95.17%. This indicates that the introduction of multispectral images significantly enhanced the model's detection capability in complex scenarios, improving its ability to identify and segment impurities in white sugar.

4.3. Attention U_Net Training Results and Visualization Results

In this test, we not only evaluated the model's ability to recognize the impurity types present in the training dataset but also specifically examined its performance in detecting new impurities (i.e., impurities that were not present during the training phase). The results for the recognition of known impurities are shown in Figure 9, and the results for the detection of new impurities are shown in Figure 10.



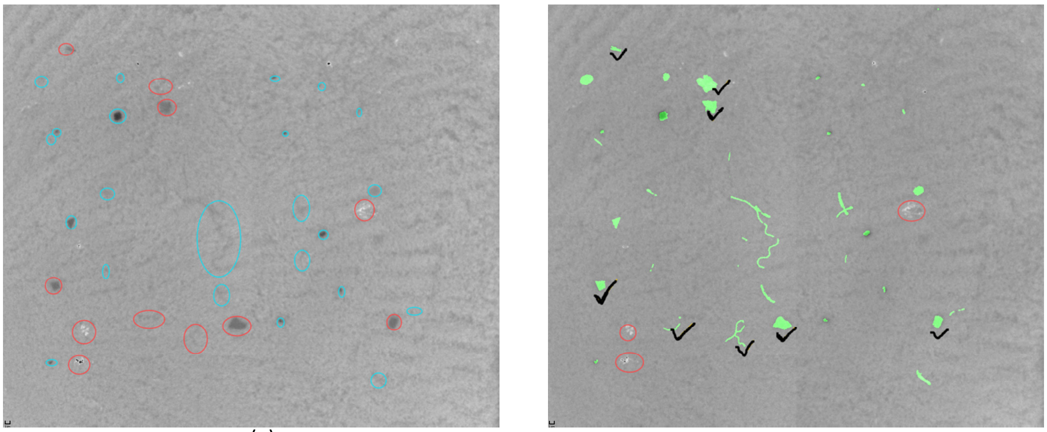
(a) (b)

Fig 9. (a) Original image of impurities from existing categories in the training set; (b) Predicted result image

As shown in the above figures, Figure 9(a) displays the original multispectral image, with the impurity locations marked by red circles. Figure 9(b) presents the model's prediction results for the image in Figure 9(a). A comparison reveals that most impurities are accurately identified, and the model performs well in distinguishing between sugar and impurities.

As shown in the above figure, the detection of previously unseen impurities during the model training process is demonstrated. Figure 10(a) displays the original image, with red circles marking the newly introduced impurities and cyan circles indicating the impurities that were present during training. Figure 10(b) shows the model's detection results for the image in Figure 10(a), with red circles indicating the impurities that were not detected and black markings representing correctly identified impurities. It is important to note that the undetected impurity is a transparent film, which has a high transmittance for incident light and a weak reflected spectral signal, making it difficult for the model to accurately recognize these impurities during detection.

The detection results indicate that the model performs well in the practical application of white sugar impurity detection, achieving efficient and accurate impurity identification. Additionally, the model also demonstrates strong performance when recognizing unfamiliar impurity categories.



(a) (b)

Fig 10. (a) Original image with added new impurities; (b) Model prediction result image

5. Summarize

In the sugar processing industry, existing impurity removal techniques fail to meet the demands of current production processes. Inadequate impurity detection not only affects product quality but also poses potential food safety risks. To address this, this study designs and proposes an impurity detection system for white sugar that combines multimodal image fusion technology, integrating both multispectral and RGB cameras, along with deep learning models, for automated, non-destructive impurity detection.

To better utilize multispectral data, the traditional Attention Unet model was optimized by expanding the input channels from the conventional single-channel grayscale or three-channel RGB images to thirteen channels, which are adapted to the system's needs. The multispectral images cover multiple bands from visible light to near-infrared, enabling the capture of both chemical and physical characteristics of sugar and its impurities, thereby significantly improving impurity detection accuracy.

Experimental results demonstrate a significant performance improvement in the recognition of impurity categories with the optimized model. This indicates that multispectral data helps enhance the precision and reliability of impurity detection. The system provides a new solution for impurity detection in sugar production, greatly improving detection efficiency, reducing reliance on manual intervention, and enabling efficient real-time monitoring. The system can substantially reduce errors caused by human factors in production, lowering labor costs while ensuring food safety during the production process.

Future research could explore detection methods for different types of impurities and attempt to apply the system to other food industries, expanding its application range.

Acknowledgments

National Natural Science Foundation of China (11574058).

References

[1] Müller, H., & Agustín, P. (2021). A review of traditional food quality control methods and their limitations. Food Chemistry, 332, 127373. DOI: 10.1016/j.foodchem.2020.127373.

[2] World Health Organization. Guideline: sugars intake for adults and children[M]. World Health Organization, 2015.

[3] Jethi G S, Sunori S K, Joshi P, et al. Application of Machine Learning in Assessing the Sugar Quality in Sugar Mills[C]// 2024 5th International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV). IEEE, 2024: 167-172.

[4] Wang, Y. H., Zhao, D. D., & Li, Q. T. (2014). Rapid analysis of molasses hammer degree, sucrose content, and reducing sugar content by near-infrared spectroscopy [in Chinese]. Food Science and Technology, 39(9), 284-288.

[5] HASHIM N, ONWUDE D I, OSMAN M S. Evaluation of chilling injury in mangoes using multispectral imaging[J]. Journal of Food Science, 2018, 83(5): 1271-1279.

[6] Huang, Z., Chen, J., Tan, G., et al. (2024). Hyperspectral non-tobacco material detection method based on multi-type convolutional mixing [in Chinese]. Journal of Image and Signal Processing, 13, 76.

[7] Russell, B. C., Torralba, A., Murphy, K. P., & Freeman, W. T. (200ˇ). LabelMe: A Database and Web-Based Tool for Image Annotation. International Journal of Computer Vision, 77(1-3), 157-173.

[8] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. Medical Image Computing and Computer-Assisted Intervention (MICCAI), pp. 234-241. Springer, Cham. DOI: 10. 1007/ 978-3-319-24574-4_28.

[9] Oktay, O., Schlemper, J., Folgoc, L. L., et al. (2018). Attention U-Net: Learning Where to Look for the Pancreas. arXiv preprint arXiv:1804.03999.

[10] Wang, X., & Guo, X. (2020). Attention U-Net Based on Residual Learning for Medical Image Segmentation. IEEE Access, 8, 55230-55239. DOI: 10.1109/ ACCESS. 2020. 298 2408.