DSN-BR-Based Online Inspection Method and Application for Surface Defects of Pharmaceutical Products in Aluminum-Plastic Blister Packages

Mingzhou Liu[1], Yu Gong[1], Xiaoqiao Wang[1]* , Conghu Liu[2] and Jing Hu[1]

Abstract

Ensuring high product quality is of paramount importance in pharmaceutical drug manufacturing, as it is subject to rigorous regulatory practices. This study presents a research focused on the development of an on-line detection method and system for identifying surface defects in pharmaceutical products packaged in aluminum-plastic blisters. Firstly, the aluminum-plastic blister packages exhibit multi-scale features and inter-class indistinction. To address this, the deep semantic network with boundary refinement (DSN-BR) model is proposed, which leverages semantic segmentation domain knowledge, to accurately segment the defects in pixel level. Additionally, a specialized image acquisition module that minimizes the impact of ambient light is established, ensuring high-quality image capture. Finally, the image acquisition module, image detection module, and data management module are designed to construct a comprehensive online surface defect detection system. To validate the effectiveness of our approach, we employ a real dataset for instance verification on the implemented system. The experimental results substantiate the outstanding performance of the DSN-BR, achieving the mean intersection over union (MIoU) of 90.5%. Furthermore, the proposed system achieves an inference speed of up to 14.12 f/s, while attaining an F1-Score of 98.25%. These results demonstrate that the system meets the actual needs of the enterprise and provides theoretical and methodological support for intelligent inspection of product surface quality. By standardizing the control process of pharmaceutical manufacturing and improving the management capability of the manufacturing process, our approach holds significant market application prospects.

Keywords Surface defect detection system, Deep learning, Semantic segmentation, Aluminum-plastic blister packages identification

1 Introduction

Product quality is a matter of utmost significance in pharmaceutical drug manufacturing, which adheres to stringent regulatory practices [1]. The quality of pharmaceuticals and their packaging can be adversely affected by variations in raw material properties and process disturbances. The presence of defective products in the market not only leads to substantial economic losses for enterprises but also poses a significant risk to patient health [2]. Consequently, after the completion of the regular production process, quality control becomes crucial to

*Correspondence:
Xiaoqiao Wang
wxq20061087@163.com
1 School of Mechanical Engineering, Hefei University of Technology, Hefei 230009, China
2 School of Mechanical and Electronic Engineering, Suzhou University, Suzhou 234000, China

Liu *et al. Chinese Journal of Mechanical Engineering*   (2024) 37:86

Page 2 of 21

ensure product quality [3]. It becomes imperative to gain a comprehensive understanding of the dynamic process and its impact on quality, particularly during start-up and shut-down phases. This entails identifying relevant variables and attributes and designing robust monitoring systems capable of accurately controlling product quality [4, 5]. Presently, within the realm of quality visual inspection, many pharmaceutical companies still rely on manual methods or sampling inspections, which present various undesirable aspects. Firstly, this labor-intensive approach entails substantial costs and inefficiencies, consuming significant human and material resources. Secondly, conducting visual inspections on large quantities of products not only intensifies the workload of workers but also reduces overall production efficiency. Furthermore, the sustained engagement in mechanical repetitive visual inspection tasks can lead to visual fatigue among workers, resulting in potential oversights in inspection and misjudgment of product quality. Consequently, unqualified products may flow into the consumer market, resulting in substantial economic losses and a decline in patient well-being [6].

With the rapid advancements in computer science and technology, the application of vision-based inspection systems has gained significant momentum across various industries, aligning with the concept of smart factories [7, 8]. Machine vision, as a real-time, efficient, and cost-effective detection method, has emerged as a powerful tool [9, 10]. It enables the swift and accurate inspection of surface defects, overcoming the limitations associated with manual inspection, such as human error, inspection omissions, and fatigue-induced detection issues [11]. Leveraging machine vision not only enhances detection accuracy but also substantially reduces the costs associated with manual inspection, leading to substantial economic and social benefits. Hence, the adoption of mechanized and intelligent artificial intelligence devices for pharmaceutical quality inspection represents the most viable alternative to manual visual inspection [12, 13]. This paradigm shift towards AI-driven technologies in manufacturing processes embodies the general trend towards unmanned and intelligent operations in the future [14].

Traditional machine vision-based methods employed for surface defect detection typically rely on conventional image processing algorithms or manually designed feature extraction techniques combined with classifiers [15]. In these approaches, different attributes of the inspected surface or defect are commonly utilized to design the imaging scheme [16]. Typically, a two-step process is followed, where a high-level representation of the defect is constructed using various feature extraction techniques, followed by classification using a dedicated classifier to determine the defect class [17]. For instance, in the work of Bay et al. [18], Hessian matrix-based detector measurements and distribution-based descriptors were employed, resulting in novel combinations of detection, description, and matching steps. In another study by Liu et al. [19], an automatic image segmentation method was proposed, wherein hand-designed features and depth features were integrated to train a structured random forest classifier. However, these methods necessitate the manual crafting of features specific to defects, which comes with inherent limitations [20]. Moreover, if an incorrect design is chosen, these methods may not be suitable for accurately identifying specific types of drugs.

Indeed, researchers are increasingly turning towards data-driven techniques to address the inherent complexity associated with image processing methods [21]. In recent years, the rapid advancements in deep learning techniques have led to the application of various algorithms for surface quality control tasks [22–24]. For instance, in the study of Zhang et al. [25], a cross-scale weighted feature fusion network was proposed to identify and locate surface defects in hot-rolled steel, achieving a mean average precision (mAP) of 86.8%. In another study by Hao et al. [26], a detection model for intelligent industrial monitoring was proposed to classify and locate multi-scale defects on steel surfaces, achieving a mAP of 80.5%. Deep learning methods exhibit distinct advantages over traditional machine vision approaches, as they possess the capability to learn features directly from raw data and demonstrate a higher capacity for representing intricate structures [27]. Consequently, these methods eliminate the need for manual feature design, replacing it with an automated learning process. This automated feature learning process aligns well with the requirements of flexible production lines [28], which necessitate efficient and adaptable product quality control mechanisms to accommodate rapid adaptation to new products.

In conclusion, the aforementioned findings highlight several prominent challenges in the field:

1. Traditional image processing algorithms exhibit significant drawbacks, including high labor and time costs. Moreover, these algorithms face substantial limitations due to their reliance on manually designed defect feature characterizations. Consequently, they possess limited error tolerance and fail to address the quality inspection requirements of multi-species and small-batch pharmaceuticals at a fundamental level.

2. Design detection methods based on deep learning are founded upon observed inputs. However, in real and complex environments characterized by diverse defect types and substantial noise interference in defect images, these methods struggle to achieve

Liu *et al. Chinese Journal of Mechanical Engineering*   (2024) 37:86

Page 3 of 21

detection performance that aligns with the practical needs of enterprises.

3. Existing detection methods primarily focus on enhancing the model itself, leading to difficulties in realizing a complete system spanning data collection to real-time detection. Furthermore, there remains ample room for improvement to facilitate the direct deployment of these methods in the actual enterprise environment.

In this study, we propose a comprehensive approach to tackle the aforementioned challenges by integrating an enhanced semantic segmentation model into a complete system. The proposed strategy encompasses four key stages, as follows: (1) Introducing a deep semantic network with boundary refinement (DSN-BR) learning model that leverages the power of deep networks for end-to-end online detection. (2) Developing a dedicated image acquisition module to ensure independent data acquisition regardless of ambient light conditions. (3) Designing the overall hardware and software architecture, constructing a system model, and building a demonstration system. (4) Implementing real-time quality state assessment and feedback adjustment to detect and adjust product conditions, providing timely feedback. The proposed system combines practical application for medicinal products with effectiveness constraints, supported by comprehensive experiments that validate its efficacy and feasibility. It outperforms current defect detection methods in terms of speed, efficiency, and cost. The contributions of this paper can be summarized as follows.

- Proposed a novel defect detection model based on semantic segmentation, enhancing the ability to detect multi-scale defects and accurately obtain defect boundary information for end-to-end defect detection in aluminum-plastic blister packages.
- Integrated modules to build a complete online defect detection system for surface defects of aluminum-plastic blister pharmaceuticals, including image acquisition, detection, and data management modules. Designed a software and hardware implementation scheme, combining deep learning methods with the product quality control process of a continuous manufacturing enterprise.
- Verified the feasibility and effectiveness of the system using actual data. Experimental results demonstrated improved performance and efficiency in defect detection during rapid production, enabling large-scale online quality control, enhanced production process management, and providing theoretical and technological support for comprehensive surface quality control in various fields of manufacturing.

The remainder of this paper is structured as follows. Section 2 presents the detailed architecture of the proposed method and system. In Section 3, validation experiments are conducted using a case study. Finally, Section 4 concludes the study and outlines potential avenues for future work.

2 Proposed Methodology and Integrated System

First, this section introduces the framework of the method, which mainly consists of three parts: Integrated learning model and algorithm, modules and system implementation. Subsequently, details of each part are described. In the next section, a system demo is built for instance validation.

2.1 Overall Framework

This study first designs the integrated learning model and algorithm, and constructs the deep learning model mainly by analyzing the difficulty of detecting aluminum-plastic materials. Then the image acquisition module is established to collect images. After training the model with datasets, the executed algorithm is deployed to the image detection module. Finally, the detection results and real-time information will be passed to the robotic arm and display terminal through the data management module, realizing the closed-loop connection of the whole system. The method framework is shown in Figure 1.

2.2 Proposed Learning Model

The core concept of the FCN-based method revolves around learning the mapping of pixel-level information in images, thus eliminating the need for extracting regions of interest. This breakthrough overcomes the limitations of traditional image segmentation techniques and paves the way for the development of end-to-end deep learning networks. In 2014, Long et al. [29] introduced the FCN network, which replaced the fully connected layers typically found at the end of conventional CNN networks with convolutional layers, enabling pixel-wise classification. This novel approach allowed for end-to-end classification of input images of arbitrary sizes, thereby establishing a foundational framework for addressing image semantic segmentation problems using deep networks. In this study, we propose the DSN-BR, which builds upon the FCN architecture and incorporates targeted structural improvements to better align with the data characteristics specific to our research environment. These enhancements are carefully designed to enhance the applicability and effectiveness of the model in handling the unique features present in the dataset under investigation.
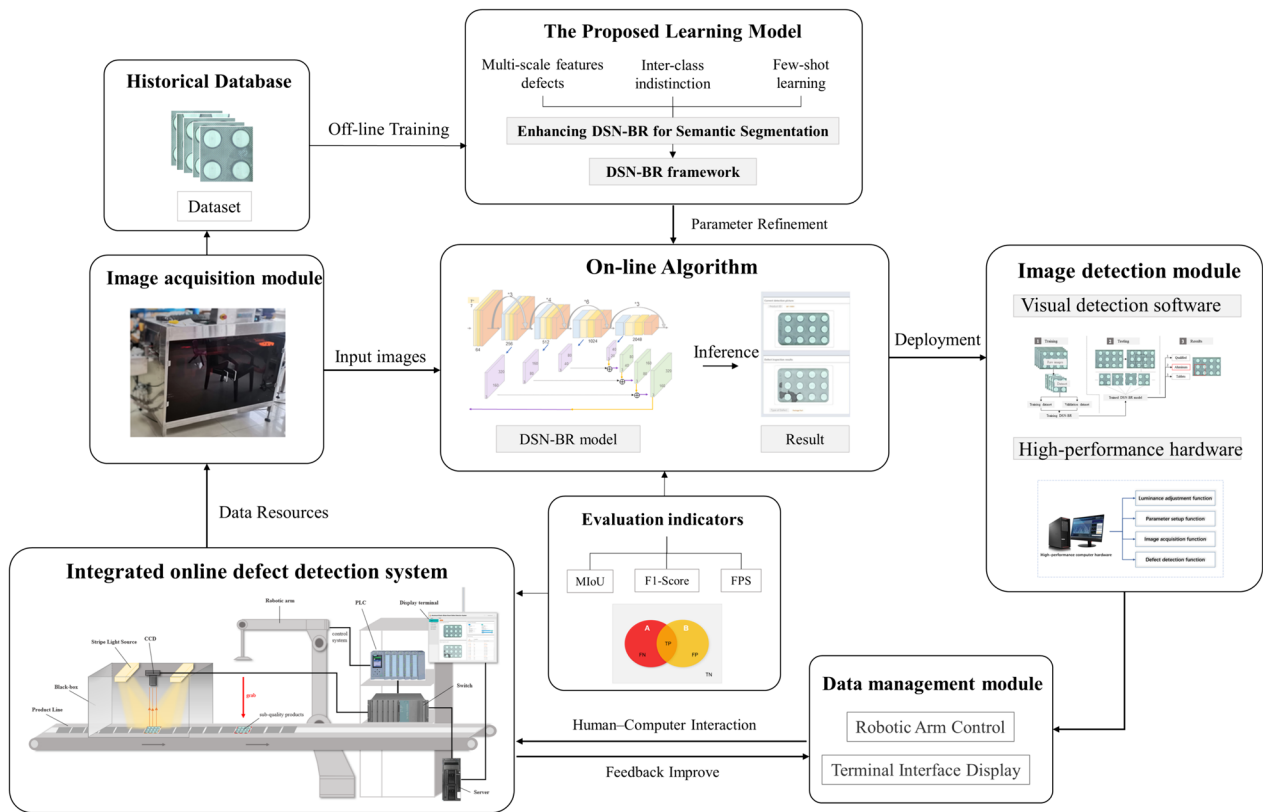
Figure 1 Overall framework of the online defect detection system

2.2.1 Problem Analysis

In this study, our objective is to achieve accurate pixel-level identification of aluminum-plastic blister packs through semantic segmentation, thereby detecting specific classes and their corresponding regions. Unlike existing methods, our proposed approach primarily relies on the integration of data and domain knowledge to improve overall segmentation performance while addressing several targeted features of importance.

(1) Multi-scale feature defects. Firstly, we need to tackle the challenge of detecting multi-scale defect features. This includes identifying both large-scale defects such as powder pressing of foam cover plates, lack of mesh and batch numbers, as well as small-scale defects like poor foam formation and drug-related anomalies (Figure 2). At the low-level stage, features capture more global information but struggle to capture the details of small defects due to the broad acceptance perspective. Conversely, at
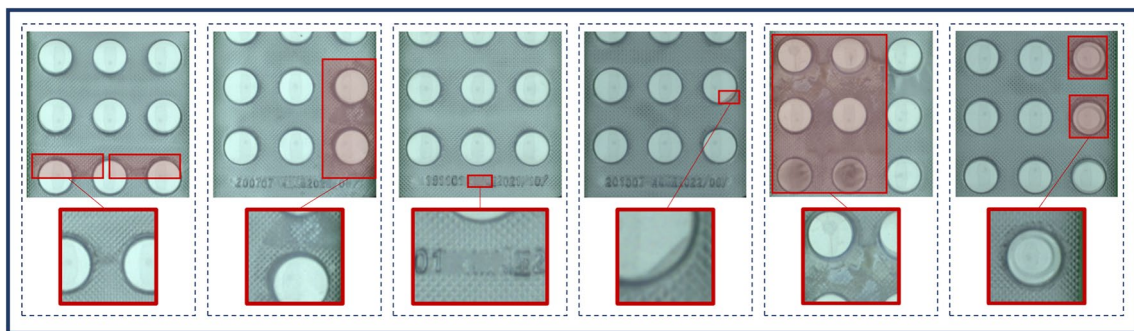

Figure 2 Multi-scale surface feature defects on the dataset (red borders represent defects)

Liu *et al. Chinese Journal of Mechanical Engineering*   (2024) 37:86

Page 5 of 21

the high-level stage, features lack the necessary contextual information to accurately distinguish large-scale defects. Additionally, the large image size used in this study (1280×640) exacerbates the discrepancy between large and small defects in terms of feature size. To address these issues, we employ a multi-scale approach that encodes both local and global context features, aiming to capture comprehensive defect information. However, the varying sizes of receptive fields pose challenges in terms of feature extraction, potentially leading to incorrect labels or conflicting results. Thus, it is essential to carefully select discriminative features for predicting the semantic labels of specific classes.

(2) Inter-class indistinction. Secondly, we address the issue of inter-class indistinction, which is often overlooked in semantic segmentation tasks. Existing methods typically treat this problem as a dense recognition task, aiming to differentiate adjacent patches with similar appearances but different semantic labels. However, they often neglect the inter-class relationships that are crucial for model performance [30]. For instance, in our study, defective features like fine lines and missing mesh patterns in blister packages exhibit similarities to patterns found in aluminum-plastic packaging (Figure 3). This similarity, especially when adjacent spatially, can result in confusion in predicted classes due to their resembling appearances. To overcome this challenge, we explicitly incorporate semantic boundaries to guide feature learning. By amplifying feature variations on both sides of the boundary, we can distinguish adjacent regions with similar appearances but different semantic labels, thereby

improving localization and prediction near defect edges.

(3) Few-shot learning. Lastly, we address the issue of few-shot learning. Traditional machine learning algorithms and tools have excelled in scenarios with large datasets, leveraging "big data" for improved model performance through automatic learning from experience [31]. However, when datasets are small, models are susceptible to overfitting, and challenges arising from data imbalance are difficult to overcome, ultimately hindering the performance of machine learning algorithms. Despite the satisfactory results achieved by existing models trained on large public datasets, the design of quality inspection processes specific to pharmaceutical manufacturing, particularly in the context of deep learning, remains insufficiently addressed. For instance, while previous methods have utilized public datasets with thousands of images to train millions of parameters, our dataset consists of a limited number of defective training samples (237 images with 100 defective samples), leading to a few-shot learning problem.

In summary, this study explores appropriate network architectures from the aforementioned three perspectives, conducting comparative improvement experiments, ultimately culminating in the development of DSN-BR. The specific modules for improvement are described in detail in the following sections.

2.2.2 Enhancing DSN-BR for Semantic Segmentation

1. Currently, most networks employ 3×3 sized convolutional kernels for feature extraction due to their
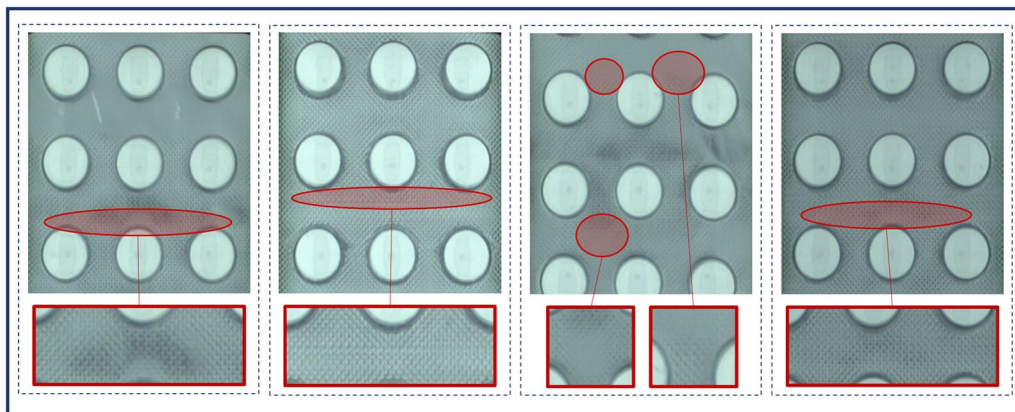

Figure 3 Defects with similar patterns on the aluminum-plastic surface, the inter-class distinction is not obvious on the dataset (red borders represent defects)

Liu *et al. Chinese Journal of Mechanical Engineering*   (2024) 37:86

Page 6 of 21

advantages, such as fewer parameters, lower computational requirements, and the ability to capture fine image details, thereby enhancing the model's capabilities. However, there are still limitations associated with small kernels. Their limited receptive fields may lead to the loss of global context information when applied to larger input feature maps. In our study, the image size is 1280×640, encompassing both small defects on the pill section and larger defects on the packaging board. Small kernels may struggle to capture large-scale features adequately and might overlook important details during information processing, consequently compromising model performance.

To address this issue, we propose increasing the kernel size of the first layer in the feature extraction network. The inclusion of larger kernels serves multiple purposes. Firstly, the enlarged kernel size expands the perceptual field, enabling the network to capture broader contextual information and global features. This allows the model to focus on larger local areas and extract more intricate details in subsequent deeper networks (Figure 4). Additionally, larger kernels possess the ability to learn more complex and abstract feature representations, facilitating the gradual extraction of higher-level features as multiple convolutional layers are stacked. This, in turn, enables the model to better comprehend semantic information for subsequent semantic segmentation tasks. For ease of implementation and efficient convolutional anchor point placement, it is customary for the kernel size to be an odd number. In the experiments conducted in Section 3.2.1,

we compare the performance of 3×3, 5×5, and 7×7 convolutional kernels. After considering the trade-off between model performance, the number of parameters, and computational requirements, we ultimately select 7×7 kernels for the first layer of the feature extraction network. This decision ensures dense connections over a larger area for feature mapping, thereby striking an optimal balance between model performance and computational efficiency.

2. Deep convolutional networks exhibit the capability to enhance feature representation by effectively leveraging the depth of layers to integrate low, medium, and high-level features in an end-to-end, multi-layer fashion. Ref. [32] has demonstrated that network depth plays a crucial role in visual recognition tasks. In the FCN, Vgg16 [33] is commonly utilized as the backbone network, comprising 13 convolutional layers and 3 fully connected layers. However, for our specific requirements, it is insufficient to solely rely on Vgg16, as we not only need multi-scale features for accurately recognizing each type of defect but also necessitate the selection of more discriminative features to predict semantic labels for specific classes. Consequently, we adopt Resnet [34] as the backbone network in our proposed method, which enables us to significantly increase the network depth and achieve outstanding accuracy. Additionally, the inclusion of residual modules in Resnet mitigates issues such as gradient explosion and vanishing gradients that may arise with deeper models. The replacement of the specific backbone structure with Resnet is depicted in Figure 5.
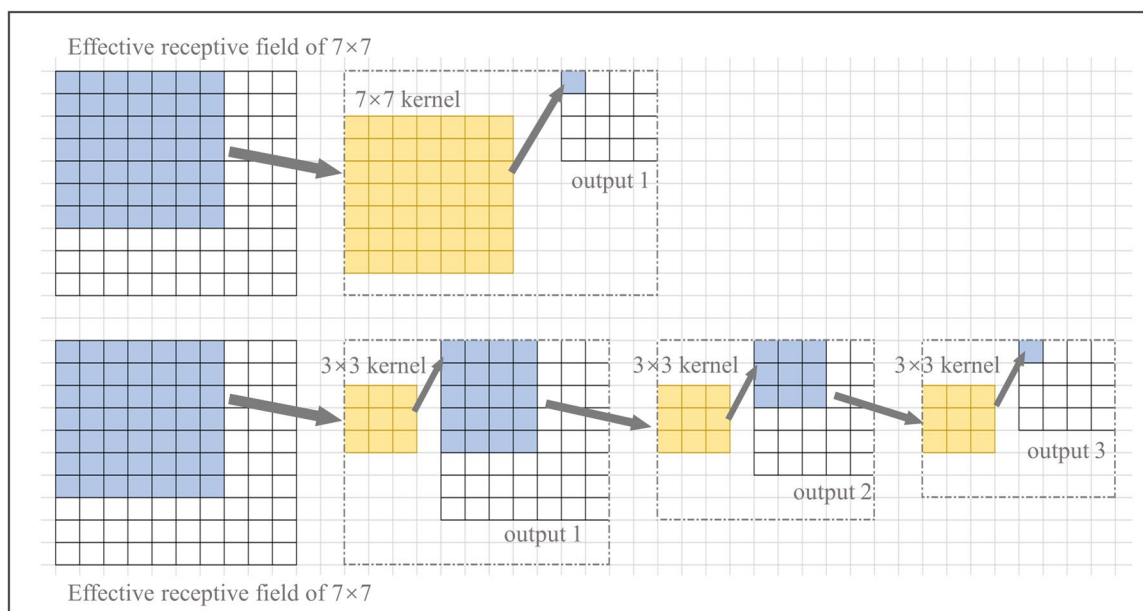

Figure 4 Differences in receptive fields with different sizes of convolutional kernels
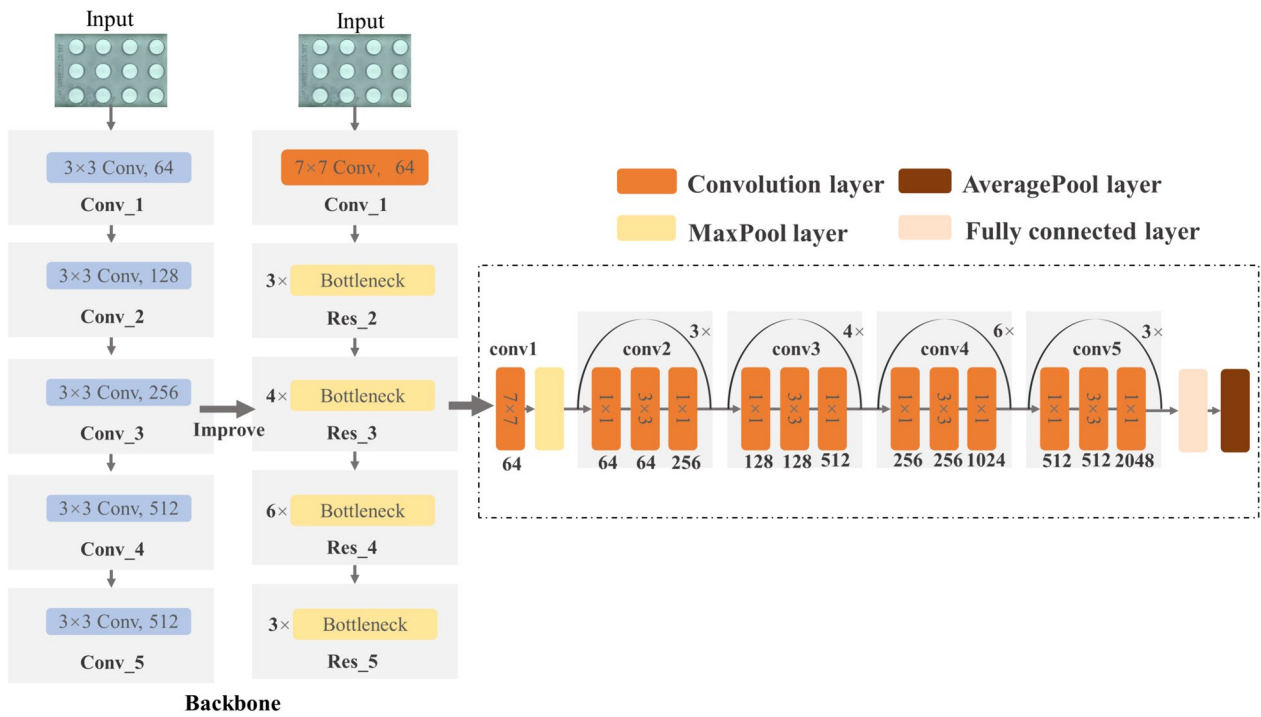
Figure 5 Improvement structure of feature extraction network

The model can be divided into five stages based on the size of the feature map. Each stage exhibits varying recognition capabilities, resulting in different performance levels (as depicted in Figure 6(a)). The lower stages encode finer contextual information, while the higher stages excel at capturing small defects with greater accuracy. Building upon this foundation, we leverage the strengths of both higher and lower stages to achieve optimal predictions. By combining the advantages of different stages, we enhance the overall prediction capability of the model.
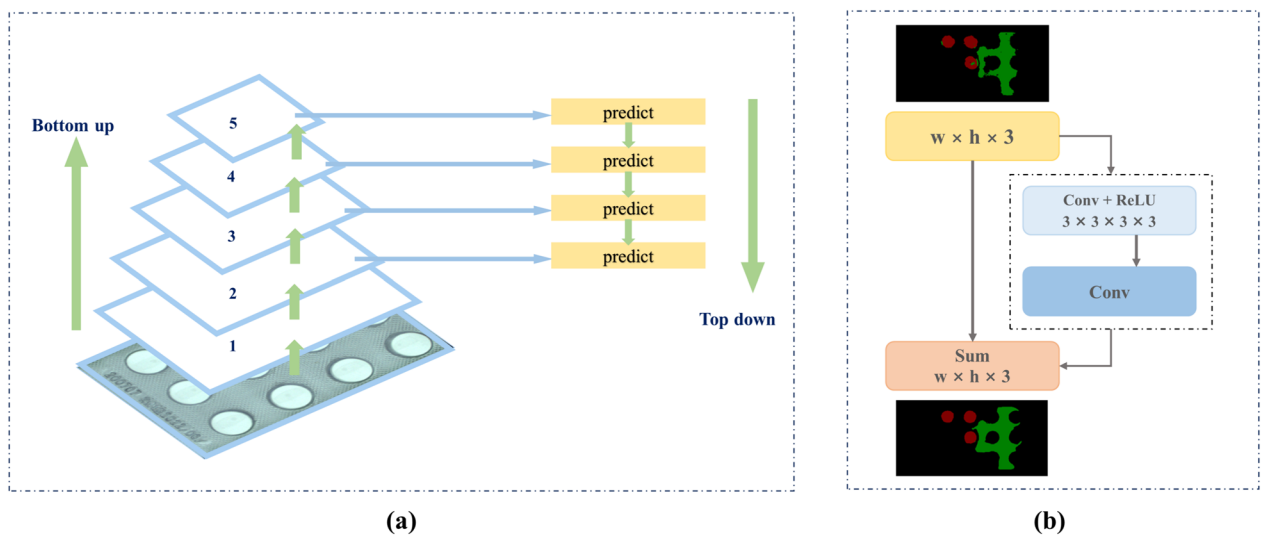

(a)

(b)

Figure 6 Schematic diagram of algorithm improvement: (**a**) Schematic representation of the different stages of the model, (**b**) Specific structure of BR module

Liu *et al. Chinese Journal of Mechanical Engineering*   (2024) 37:86

Page 8 of 21

3. To effectively distinguish between classes that exhibit similar appearances, it is essential to enhance the distinction between features and extract accurate semantic boundaries. To achieve this, we incorporate a semantic boundary loss during the training process, which bears a resemblance to the task of semantic boundary detection. This loss function enables the model to learn discriminative features that amplify the inter-class distinction, facilitating the distinguishability of features on both sides of the semantic boundary. We introduce the boundary refinement (BR) module derived from GCN [35] to learn boundary features. By leveraging this module, we can magnify the variations in features on both sides of the boundary, effectively differentiating adjacent areas that share similar appearances but possess different semantic labels. Consequently, this approach improves the localization and prediction accuracy near defect edges, addressing the challenges associated with these regions.

The specific operation is to add a BR module after the multi-scale feature layer, which is a small residual structure to refine the defect edges with residual branches. Assume that the size of the input feature map is $W \times H \times C$, which is the width, height and channels. For the residual branch on the right side, convolution is performed using the 3×3 convolution kernel of stride 1. After ReLU, another convolution is performed, finally the output is summed with the direct mapping layer. More specifically, we define $\tilde{F}$ as the refined score map $\tilde{F} = F + R(F)$, where F is the coarse score map and $R(F)$ is the residual branch. The details can be referred to Figure 6(b).

2.2.3 DSN-BR Framework

We propose a novel deep fully convolutional network called DSN-BR for semantic segmentation, aiming to improve the spatial accuracy of the segmentation output. The overall architecture of the DSN-BR model is depicted in Figure 7(a). To construct the model, we utilize Resnet50 as the feature extraction network and employ the FCN framework for segmentation. We initialize the backbone network with pre-trained weights from the PASCAL VOC 2011 dataset [36]. To capture both high and low-level features, we fuse feature maps from different stages of the Resnet50 backbone. To enhance the dense prediction, we incorporate network upsampling and pixel-level loss. Specifically, each node in the network generates a multiscale semantic feature map through the full convolutional network structure. Additionally, the BR module is employed to learn boundary features. We introduce branches between layers to fuse coarse, semantic, and local appearance information, achieving high-resolution feature map fusion through bilinear interpolation. The skip connections in the model

aim to refine the semantic and spatial accuracy of the output. At each fusion point, boundary learning is performed. Finally, through the last upsampling operation, a prediction image of the same size as the original image is generated. The flowchart of the proposed DSN-BR model is illustrated in Figure 7(b).

2.3 Modules

The image acquisition module plays a crucial role in obtaining high-quality surface images of the tested aluminum-plastic blister packages. It consists of components such as CCD, light source, and clamping device. Its primary function is to capture the product surface images accurately. The customized image acquisition module is essential as it directly influences the visualization of defects on the product surface. In this study, the tested aluminum-plastic blister packages undergo a process that involves molding a transparent plastic hard film, filling it with solid drugs such as tablets, pills, and capsules, and then heat bonding it with an adhesive-coated aluminum foil to form a sealed package. The aluminum foil used possesses reflective properties. However, due to variations in the angle between the ambient light and the surface of the packages (as depicted in Figure 8(a)), the illumination in the detection room becomes uneven and uncertain across different locations and surfaces. This irregular distribution of reflected light intensities greatly impacts the subsequent segmentation detection process. Additionally, certain curved surfaces on the plastic hard disk, formed by the plastic absorption process, exhibit strong reflections, leading to uneven local surface brightness. Therefore, achieving high-quality image acquisition for blister packages' surfaces under normal inspection environments proves to be a challenging task. To address the lighting conditions, surface characteristics of the blister packages, defect properties, and the detection objectives, we designed and constructed a specialized image acquisition module.

2.3.1 Image Acquisition Module

The image acquisition module plays a crucial role in obtaining high-quality surface images of the tested aluminum-plastic blister packages. It consists of components such as CCD, light source, and clamping device. Its primary function is to capture the product surface images accurately. The customized image acquisition module is essential as it directly influences the visualization of defects on the product surface. In this study, the tested aluminum-plastic blister packages undergo a process that involves molding a transparent plastic hard film, filling it with solid drugs such as tablets, pills, and capsules, and then heat bonding it with an adhesive-coated aluminum foil to form a sealed package. The aluminum
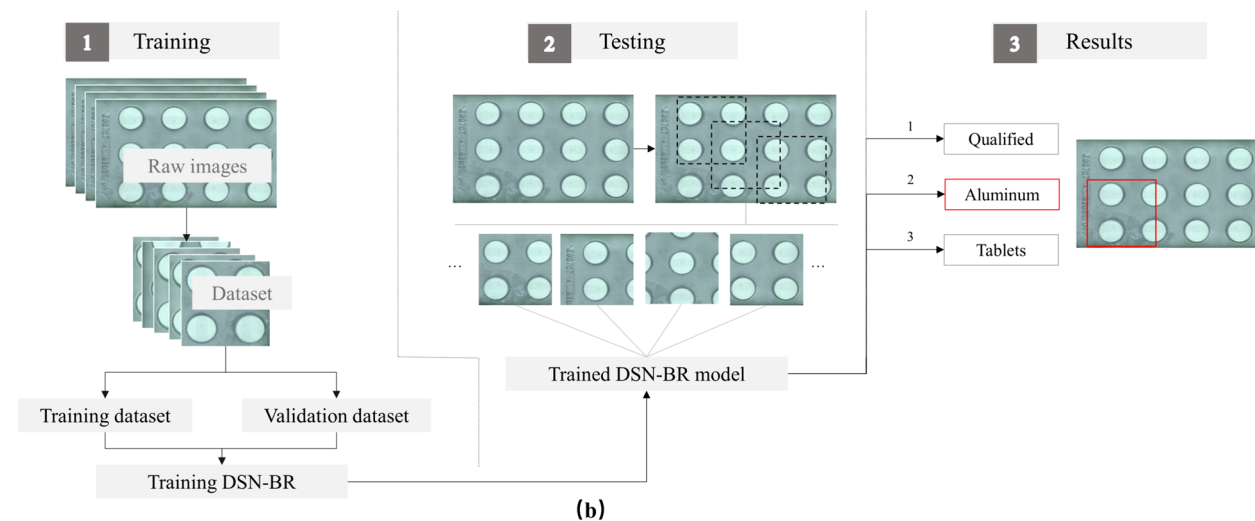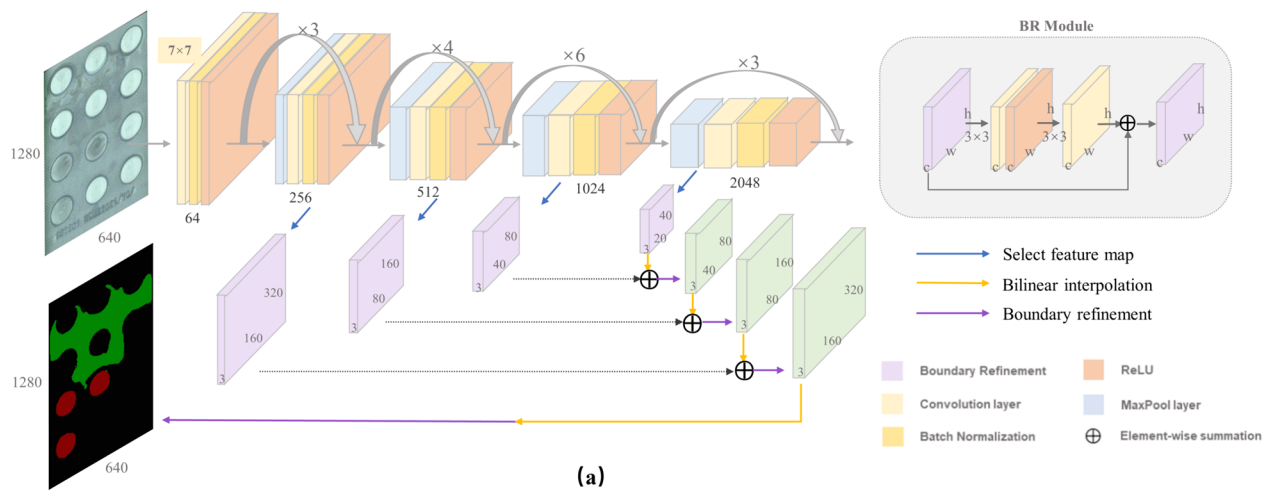
Figure 7 Detailed introduction of the DSN-BR model: (**a**) Specific architecture of DSN-BR, (**b**) Flowchart

foil used possesses reflective properties. However, due to variations in the angle between the ambient light and the surface of the packages (as depicted in Figure 8(a)), the illumination in the detection room becomes uneven and uncertain across different locations and surfaces. This irregular distribution of reflected light intensities greatly impacts the subsequent segmentation detection process. Additionally, certain curved surfaces on the plastic hard disk, formed by the plastic absorption process, exhibit strong reflections, leading to uneven local surface brightness. Therefore, achieving high-quality image acquisition for blister packages' surfaces under normal inspection environments proves to be a challenging task. To address the lighting conditions, surface characteristics of the blister packages, defect properties, and the detection objectives, we designed and constructed a specialized image acquisition module.

The image acquisition module utilizes a closed black box configuration to mitigate the influence of ambient light, as depicted in Figure 9(a). Within the black box, a controlled brightness strip light source is employed in conjunction with an industrial-grade CCD camera. The strip light source offers several advantages, including high light uniformity, ensuring that captured images are not affected by unevenly reflected light. Moreover, the strip light source exhibits high stability and provides adjustable angles for optimal illumination. The industrial CCD camera is responsible for acquiring images within the enclosed black box, which will be subsequently processed by the detection module. After the implementation of the designed image acquisition module, an experimental setup was established, as illustrated in Figure 9(b). The experimental results demonstrate that the images captured by the image
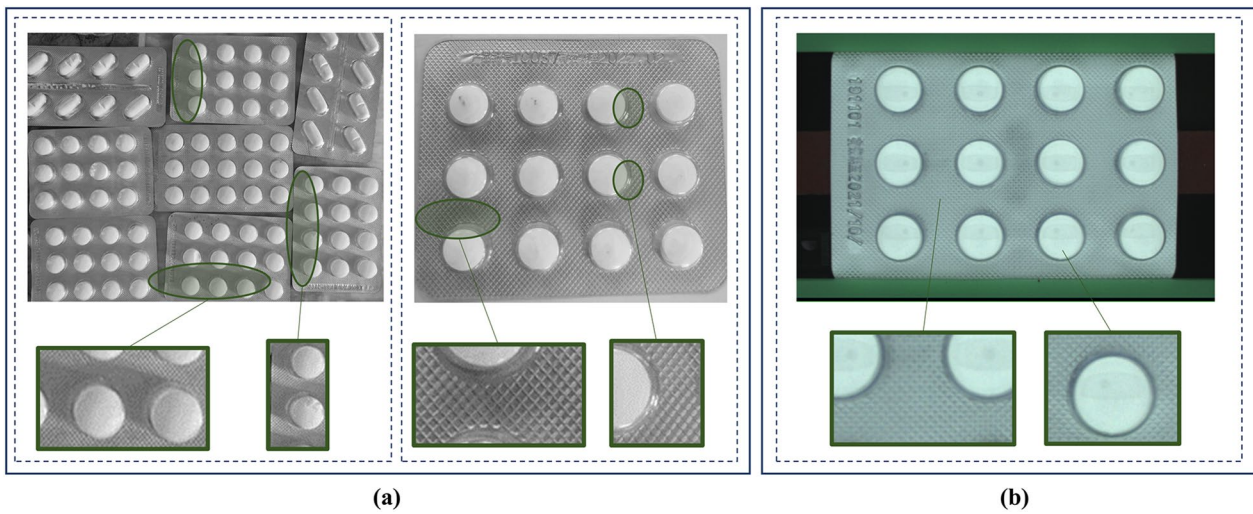
Liu *et al. Chinese Journal of Mechanical Engineering* (2024) 37:86

Page 10 of 21

Figure 8 Images collected before and after using the image acquisition module: (**a**) Images of aluminum blister pack taken under natural ambient light, (**b**) Image captured by the image acquisition module we built
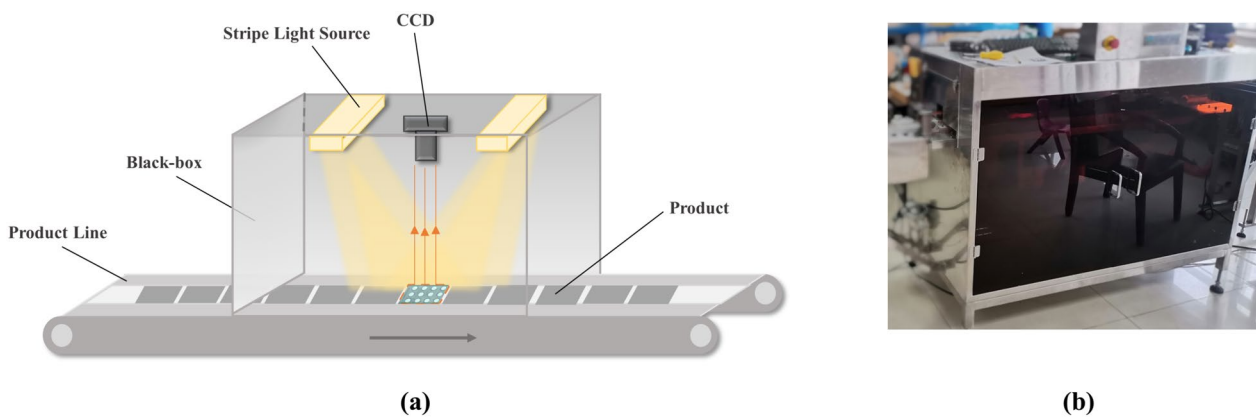

Figure 9 Details of the image acquisition module: (**a**) Design of image acquisition module, (**b**) Established experimental setup

acquisition module exhibit no discernible impact from reflected light or uneven brightness. This improvement in image quality significantly benefits subsequent image detection tasks.

2.3.2 Image Detection Module

Our image detection module encompasses both high-performance hardware and visual detection software, as depicted in Figure 10. The hardware configuration includes a Lenovo workstation equipped with a powerful CPU (Intel Xeon Silver 4210R) and GPU (NVIDIA Geforce RTX 3090) to handle the computational workload. Additionally, a 4K HD smart display terminal is utilized for efficient human-computer interaction. To ensure a comprehensive solution, the software module is designed to incorporate various functionalities, such as image processing, analysis, real-time parameter
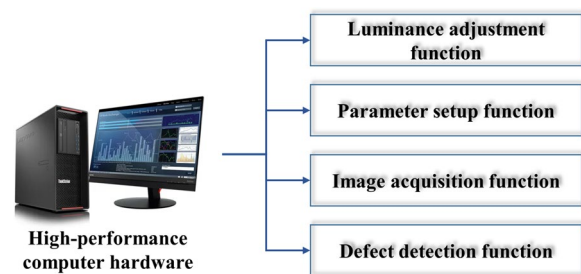

Figure 10 Composition of image detection module

adjustment, user-friendly operation, and scalability. The visual inspection software comprises several modules, including the image acquisition module, defect detection

module, light source adjustment module, and acquisition parameter setting module (Figure 10).

In operation, the image detection module first receives high frame rate images captured by industrial cameras. These images are then fed into the visual inspection software, where they undergo analysis and processing using an end-to-end integration algorithm. The resulting detection outcomes are output to the display terminal for visualization. Simultaneously, the PLC system receives the detection results to facilitate the control of the robotic arm, enabling it to retrieve the identified unqualified products. Moreover, the software interface of the display terminal offers functions such as automatic and manual adjustment of LED strip light source intensity for luminance control, as well as manual parameter setup for fine-tuning the system. Real-time updates of these functional parameters are displayed on the software interface, ensuring efficient monitoring and customization during operation.

2.3.3 Data Management Module

To effectively manage the dynamic information within the system, timely and efficient data management is crucial. The data management module plays a vital role in facilitating the easy management of product design processes, controlling product description data, and providing real-time mobility data information to authorized personnel. This module ensures the flexibility of the product data management system while maintaining robust information security. Moreover, it enables seamless communication of information between the robotic arm and the human-machine interface, ensuring the smooth operation of the system. In this section, we present a comprehensive overview of the system's operation process and design the specific steps to illustrate the flow

of data (as depicted in Figure 11). The entire operation process is divided into three main areas: The acquisition area, detection area, and interaction area. The specific steps involved are outlined below.

Step 1: Acquisition Area. Upon entering the black box from the production line, the sensor detects the presence of the product at the designated shooting station. If the product is not detected, it is redirected back to the beginning of the process. If the product is detected, the CCD camera initiates the image capture process, leading to Step 2.

Step 2: Detection Area. The captured image is fed into the image detection module, where the integrated algorithm predicts whether the product is qualified or not. If the product is deemed qualified, it is redirected back to the acquisition area to restart the process. If the product is determined to be unqualified, the system proceeds to Step 3.

Step 3: Interaction Area. In the interaction area, if there is information indicating the presence of an unqualified product, the PLC sends instructions to both the robotic arm and the display terminal. The robotic arm then proceeds to grasp the unqualified product at the designated station. Simultaneously, the display terminal provides real-time information regarding the unqualified product.

2.4 System Implementation

The proposed surface defect detection system operates through a series of well-defined steps. Initially, surface images of the products are captured using suitable light sources and image sensors, such as charge-coupled devices. Subsequently, the collected images undergo a series of operations, including localization, identification, classification, and statistical analysis of surface defects. The processed information is then stored, and based on the analysis, subsequent instructions are generated. The
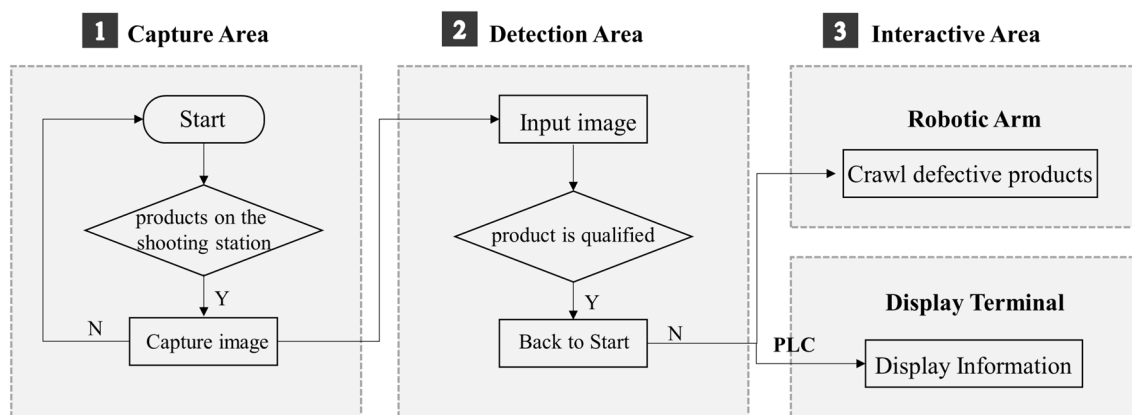

Figure 11 System running flow

key components of the system consist of the image acquisition module, image detection module, and data management module, which collectively ensure the smooth functioning of the system. In this section, we provide a detailed introduction to these modules, along with an overview of the system implementation, as depicted in Figure 12.

2.4.1 Hardware Platform Framework

Based on the theoretical foundations and experimental findings discussed earlier, we have constructed a comprehensive system, as depicted in Figure 13. The hardware components of the system encompass a product line, an image acquisition module, an image monitoring module, a robotic arm, PLC, a switch, a server, and a display terminal. The image acquisition module plays a pivotal role in ensuring the acquisition of high-quality images, even under challenging lighting conditions. The robotic arm operates in conjunction with the conveyor system, intercepting the identified unqualified products at specific stations based on the detection results. The PLC, connected to the switch, provides logical control over the

entire system's operations, while the high-speed server efficiently handles the processing tasks. Ultimately, the system's results are displayed in real-time on the display terminal, providing users with immediate feedback and insights.

2.4.2 Software Platform Framework

The software platform architecture, as illustrated in Figure 14, serves as a comprehensive solution for the detection and retrieval of defective products using an industrial robotic arm in conjunction with machine vision technology. This platform encompasses various components, including the encapsulated segmentation detection algorithm, communication interfaces with the robotic arm, and control mechanisms for robotic arm operations. It incorporates essential functionalities such as PLC monitoring and control, database services, automatic sorting, and defect detection capabilities. The features of the software platform are outlined as follows.

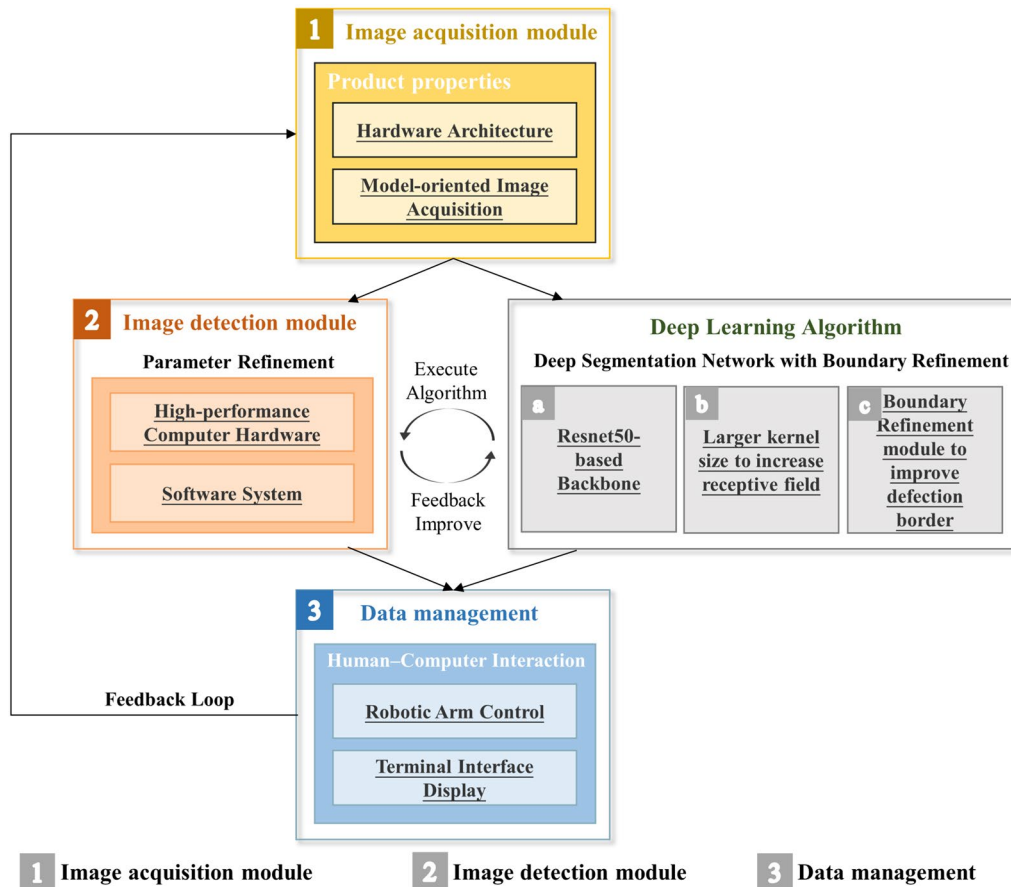(1) The monitoring and control aspect of the PLC is composed of Siemens' Step 7 professional soft-

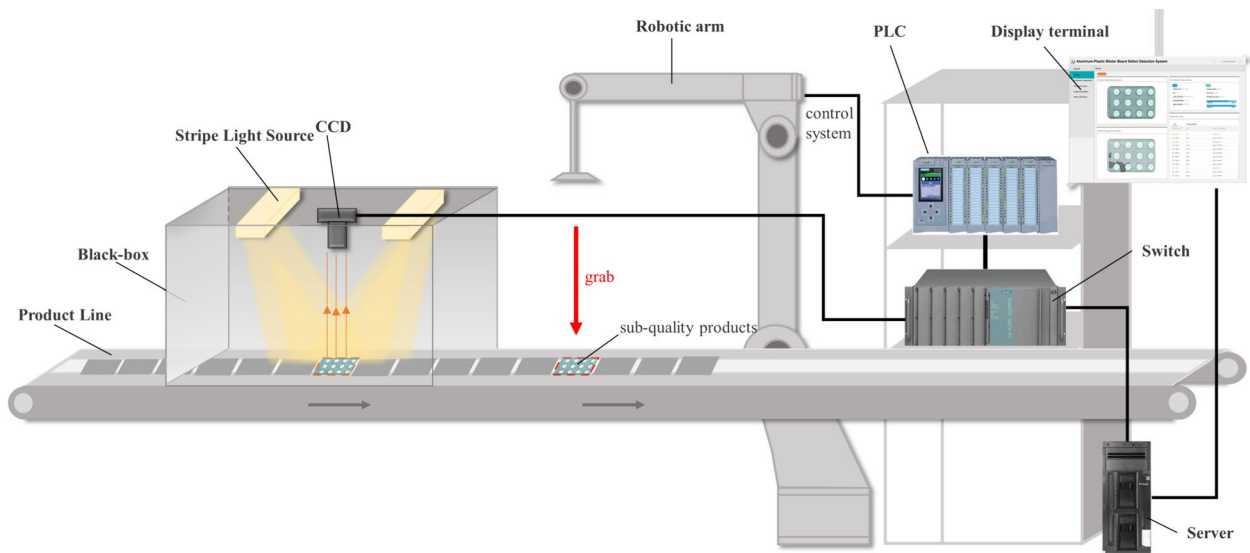**Figure 12** Composition of the visual detection system

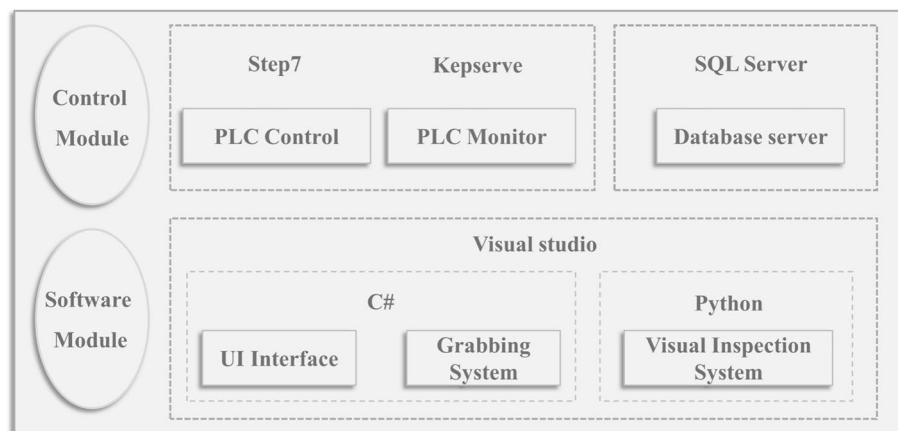**Figure 13** Integrated online defect detection system

Figure 14 System software module

ware, which includes the mobile control part of the robotic arm. It facilitates three-party communication between the robot, PLC, and server. Real-time monitoring of the robotic arm's status, fault detection, and analysis of the fault cause are accomplished through intermediary software, such as kepserver.

(2) The database service component primarily handles the recording of product quantities during the defect detection process, different types of defects, and inspection logs. This enables efficient detection supervision and real-time tracking of non-conforming products. SQL-based database software is utilized to facilitate this functionality.

(3) Communication between the industrial camera and the server enables image acquisition, as well as the retrieval, interpretation, and storage of images for symbolic reasoning purposes.

(4) In the image detection module, the trained segmentation model is deployed on the server. Captured images are retrieved and processed by the model for real-time end-to-end defect detection. The implementation is predominantly written in Python using PyTorch.

(5) The intelligent capture network involves retrieving key information of defective products through RFID. Commands are then transmitted to the robotic arm via the PLC, enabling the arm to carry out sorting and grasping tasks for non-conforming

products at specific workstations. Programming for this component is performed in the PyCharm environment.

(6) The UI design provides real-time access to inspection information, including current product details, defect images, types, batch accuracy, and more. The UI interface also facilitates the issuance of commands to the system, enabling users to operate each function effectively. The UI is designed to display variable text boxes, image results, work logs, and other relevant information. Python is employed to implement this component, and the primary UI interface is illustrated in Figure 15.

2.4.3 Evaluation Indicators

In this study, we employ various evaluation metrics to assess the performance of both the semantic segmentation model and the system. The mean intersection over union (MIoU) serves as the primary evaluation metric for the semantic segmentation model, while accuracy, recall, precision, F1 score, and frames per second (FPS) are used to evaluate the system.

To evaluate the semantic segmentation results, we analyze the intersection ratio between the true and predicted values of each class at the pixel level. The MIoU is then obtained by averaging the intersection ratios for all classes, using Eq. (1). Here, TP represents true positive, FP represents false positive, TN represents true negative, and FN represents false negative. The MIoU provides a quantitative measure of the overall performance of the semantic segmentation model. For evaluating the system performance, we consider the image as the smallest unit of analysis. If the model detects any type of defect in an image, we classify it as a defect image. Accuracy measures the proportion of correctly predicted samples, encompassing all categories (Eq. (2)). Recall reflects the model's ability to correctly predict true defect images (Eq. (3)). Precision measures the proportion of defect predictions that are truly defects (Eq. (4)). F1-score is a balanced mean of Recall and Precision that can be used to assess the overall performance of the model (Eq. (5)). Additionally, the FPS metric is employed to quantify the system's detection rate and verify its real-time performance.
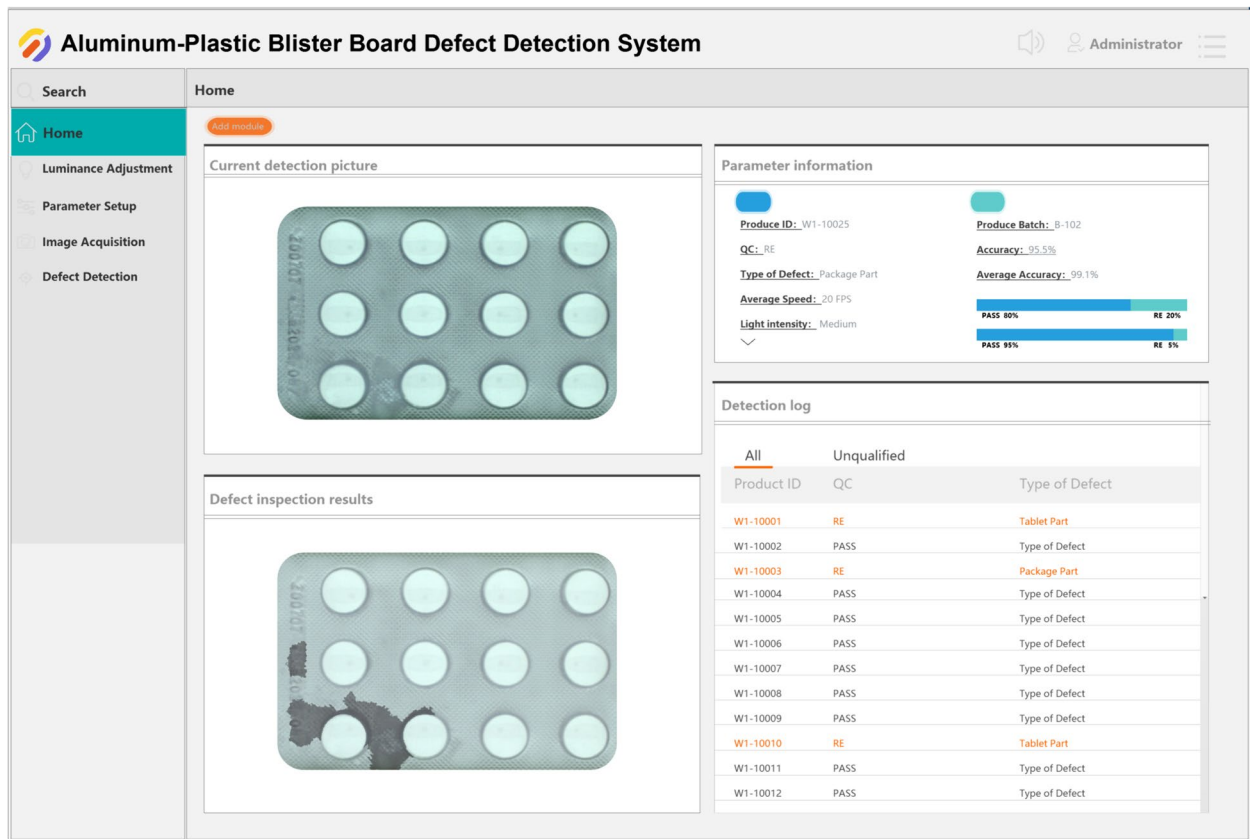

Figure 15 Main interface of the software platform

$$MIoU = 1/(1 + n) \sum_{i=0}^{n} TP/(FN + FP + TP), \qquad (1)$$

$$Accuracy = (TP + TN)/(TP + FP + TN + FN), \qquad (2)$$

$$Recall = TP/(TP + FN), \qquad (3)$$

$$Precision = TP/(TP + FP), \qquad (4)$$

$$F1\ score = 2 \times Recall \times Precision/(Recall + Precision). \qquad (5)$$

3 Case Study

3.1 Dataset Preparation

The manufacturing process of aluminum-plastic blister packages involves the conversion of plastic film into blisters, followed by heat pressure sealing and bonding methods to enclose tablets between the blister and the bottom plate. Defects in this process can be broadly categorized into two types: Aluminum-plastic package part defects and tablet part defects. Aluminum-plastic package part defects (Figure 16(b)) encompass various issues such as blister plate pressed powder, missing mesh, problems with cross grain, incomplete production batch number wording, and batch number omissions, among others. Tablet part defects (Figure 16(c)) include subpar bubble forming, missing drugs, drug leakage, and similar anomalies. These defects pose significant challenges to the quality control of aluminum-plastic blister packages, necessitating effective inspection and detection methods to ensure product integrity and compliance with quality standards.

To train and test our integrated algorithm, we collected 237 images of 1280×640 pixels in a batch of licorice tablets packed in aluminum-plastic blister through our own image acquisition module. Among them, 100 were qualified products and 137 were unqualified products. We performed several classical geometric deformation techniques on the dataset, including flipping, cropping and scaling, translation, and the addition of noise points. Then each image was finely labeled, with labels divided into three categories: Aluminum plastic packaging defect region, tablet defect region, and qualified region. The distribution of defect categories and the division of the dataset are shown in Table 1. All images and labels were saved in Pascal voc 2012 format.

3.2 Analysis of Experimental Results

In the preceding sections, we introduced the DSN-BR integrated algorithm designed for the detection of multiclass defects on the surface of pharmaceutical products packaged in aluminum-plastic blister packages. In this section, we will evaluate the performance of our algorithm and assess its overall effectiveness using a system example. To conduct the evaluation, we utilized a subset of 59 test images extracted from the dataset acquired by our image acquisition module. These test images consist of three classes: Aluminum-plastic part defects, pharmaceutical part defects, and qualified parts.

Table 1 Distribution of each category in dataset

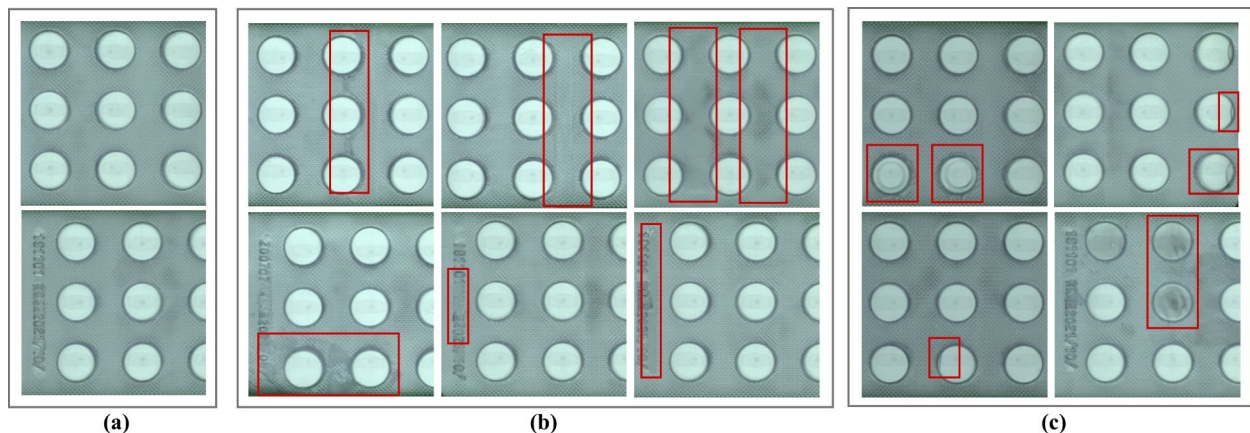| Category	Qualified sample	Package area	Tablets area	Total
Number/Train	60	90	5	155
Number/Validation	10	12	1	23
Number/Test	30	25	4	59
Total	100	127	10	237

Figure 16 Description of defect annotation: (**a**) Qualified sample, (**b**) Various types of defects in the aluminum-plastic part, (**c**) Various types of defects for the tablet component

During the training process, various image pre-processing techniques, such as center crop and horizontal flipping, were applied. We employed the Adam optimizer for training the model, with Softmax utilized to monitor the output. The NLLLoss was employed to optimize the model, and the MIoU was used as the performance metric. All experiments were conducted using the PyTorch framework. For the DSN-BR network, a total of 30 epochs were trained. During the initial seven iterations, the model was trained using an initial learning rate of 1×10^{-3}. Subsequently, the learning rate was reduced by 0.5 times in the following iterations. The batch size was set to 4, and the weights were updated using SGD. To ensure a fair comparison, all methods were tested on the same testing set.

3.2.1 DSN with Large Convolutional Kernel

The utilization of a backbone architecture in the front-end plays a crucial role in extracting image information and generating feature maps for subsequent segmentation networks. In our previous analysis, we postulated that employing a deeper backbone facilitates establishing dense connections between classifiers and features, thereby being more suitable for the specific task investigated in this study. Furthermore, we sought to enhance the feature extraction network by enlarging the kernel size. To validate this notion, we replaced various backbones and conducted comparative experiments using different convolutional kernel sizes denoted as '*k*'. Specifically, the semantic segmentation experiments were performed on Vgg16, Resnet50, and Resnet101, with kernel sizes of *k*=3, 5, 7 and 9, respectively.

Based on the obtained experimental results (refer to Figure 17), it is evident that employing a deeper feature extraction network noticeably improves the outcomes for the current task. Moreover, enlarging the size of the convolutional kernel in the initial layer of the feature network yields a marginal enhancement in model performance. Although the impact is not pronounced in the case of Vgg16, it proves effective in the Resnet series. Specifically, the performance of the resnet50 model improves by 0.9%, 1.9% and 0.5%, while the resnet101 model exhibits an improvement of 1.0%, 1.6%, and 0.6%. Considering the practical performance requirements of present-day enterprises, as well as the constraints imposed by parameter count and computational complexity, we opted for a balanced approach. Accordingly, we selected Resnet50 as the feature extraction network and set the kernel size of the initial layer to 7×7.

3.2.2 Semantic Boundary Learning - Boundary Refinement Module

In the previous sections, we highlighted the significance of learning semantic boundaries in enhancing inter-class distinction, particularly in distinguishing between aluminum composite panel meshes and defects in our study. By explicitly modeling the boundaries, the distinguishing characteristics of both sides become more pronounced, facilitating cross-class discrimination of features. In this section, we conduct ablation experiments to investigate the impact of incorporating BR modules. Specifically, we examine the effects of adding aligned residual structures before and after the feature maps at different scales. In our experiments (as presented in Table 2), we integrate

	Vgg16	Resnet50	Resnet101
■ k=3	85.8	87.7	89.1
■ k=5	86.1	88.6	90.1
■ k=7	86.7	90.5	91.7
■ k=9	86.9	91.0	92.3

Figure 17 Segmentation results of different backbones are tested on our dataset

Liu *et al. Chinese Journal of Mechanical Engineering*   (2024) 37:86

Page 17 of 21

Table 2 Comparisons of the detection results (IoU and MIoU (%)) on dataset

Model	Aluminum-plastic	Tablets part	Qualified sample	Mean-IoU
DSN (ResNet50)	91.8	81.1	98.0	90.3
DSN-BR (ResNet50)	91.9	81.4	98.2	90.5
DSN (ResNet101)	92.7	83.3	98.2	91.4
DSN-BR (ResNet101)	92.9	83.3	98.9	91.7

the BR module into different models and compare the MIoU obtained before and after its incorporation. This comparative analysis allows us to assess the performance improvement resulting from the inclusion of the BR module in the respective models.

The experimental results indicate that the inclusion of the BR module yields an increase of approximately 0.2% in the MIoU metric. This improvement demonstrates the positive effect of the BR module on the final segmentation performance. However, the overall gain achieved in the entire segmentation network is relatively modest. To provide a visual comparison of it, we present Figure 18, which showcases the output images before and after the addition of the BR module to the network. Upon observation, it becomes apparent that the BR module enhances the boundary learning capability of the network. Furthermore, it exhibits a certain level of smoothing effect on the boundary prediction of the objects. Based on these visual observations, we tentatively conclude that the BR module facilitates the model's ability to better focus on semantic boundaries and improve inter-class distinction.

After the above series of ablation experiments, we verified the segmentation capability of DSN-BR and the significant advantages of each module. Figure 19 shows the comparison of the final visual detection results of each type of defect on different models.

3.2.3 System Detection Performance

In this section, we assess the performance of our proposed framework at the system level. Our evaluation focuses on the detection performance and detection rate at the individual image level, aiming to verify whether the framework meets the practical requirements of enterprises. For evaluating the classification performance, we consider a pixel-level detection approach, where once the model detects a pixel belonging to a defect class, we classify the entire image as non-conforming.

Firstly, we compared our approach with Vgg16 and ResNet50. The testing set was uniformly divided as previously mentioned, consisting of 30 defect-free images and 29 defective images. The experimental results are illustrated in Figure 20. It is observed that Vgg16 showed slightly less effectiveness compared to the other methods. Resnet50 demonstrated significant improvements in detection results, achieving 94.92% accuracy and 94.74% F1-Score. Our method further improved upon these results, achieving 98.31% accuracy and 98.25% F1-Score.
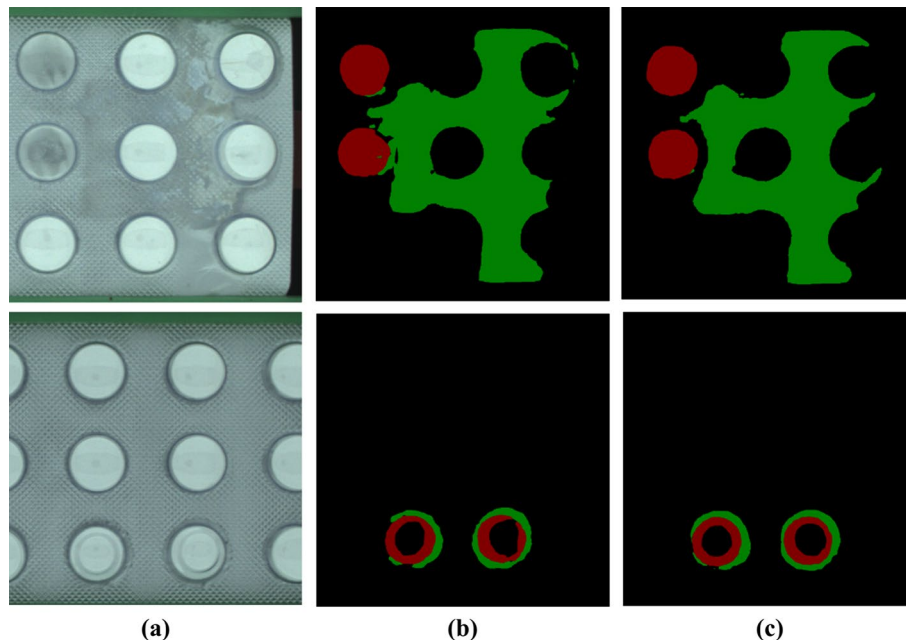
Figure 18 Results before and after using the BR module: (**a**) Image samples, (**b**) Output without the BR module, (**c**) Output after adding BR module

(a) Images   **(b) FCN**   **(c) FCN_Resnet50**   **(d) DSN-BR**
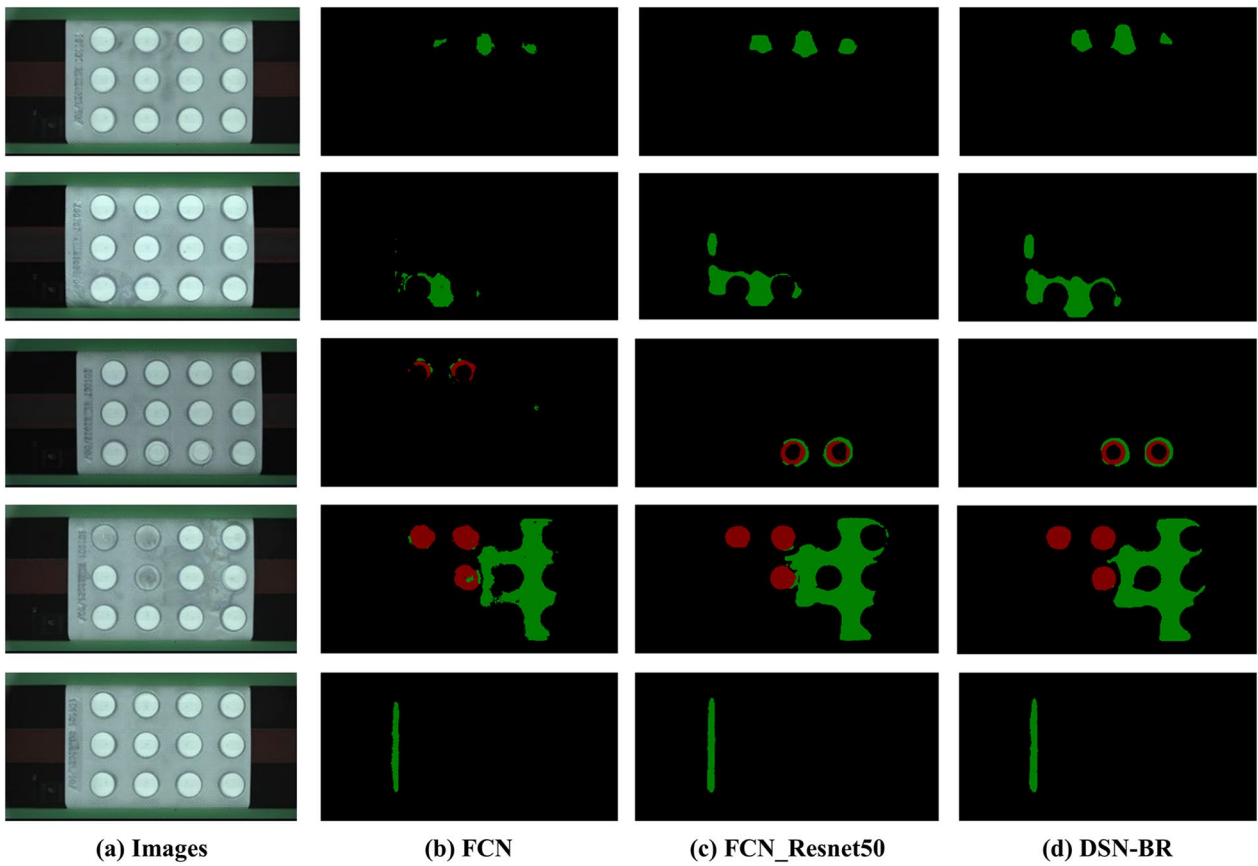
Figure 19 Examples of segmentation results on the testing set: (**a**) Input images, (**b**) FCN segmentation result, (**c**) FCN_ResNet50 segmentation result, (**d**) DSN-BR segmentation result

	Vgg16	Resnet50	DSN-BR
Accuracy	91.53	94.92	98.31
Recall	86.21	93.10	96.55
Precision	96.15	96.43	100.00
F1-Score	90.91	94.74	98.25

Figure 20 Defect detection results of Vgg16, Resnet50 and our method

Liu *et al. Chinese Journal of Mechanical Engineering*   (2024) 37:86

Page 19 of 21

Table 3 Comparison of inference speed when different batch sizes and models

Model	Batch size	Speed (f/s)	Model	Batch size	Speed (f/s)
DSN (Vgg16)	1	19.65	DSN-BR (Vgg16)	1	19.17
	4	20.14		4	20.02
DSN (ResNet50)	1	14.03	DSN-BR (ResNet50)	1	13.88
	4	14.96		4	**14.12***
DSN (ResNet101)	1	9.01	DSN-BR (ResNet101)	1	8.22
	4	9.85		4	8.93

*The bold data highlights the best results achieved by the model in the experiment

These performance metrics, especially the F1-Score, adequately meet the actual detection requirements.

Then we evaluated the processing speed of DSN-BR using a GPU machine equipped with RTX 3090. All images were of size 1280×640. Table 3 presents the detection rates of different networks used. Initially, we compared the detection rate before and after incorporating the BR module. It was observed that the BR module marginally reduced the detection rate, but the magnitude of this reduction was negligible. Hence, we chose to disregard it to obtain more accurate detection results. Subsequently, in terms of different backbone choices, Vgg16 achieved an advantage with a detection rate of 20.02 f/s. However, considering detection accuracy, we preferred using Resnet as the backbone. Finally, we selected DSN-BR with Resnet50 as the backbone for our detection model*, which achieved a detection rate of 14.12 f/s with a batch size of 4.

3.2.4 Discussion

In this section, we present the validation of our approach using actual data from aluminum-plastic blister-packed tablets. We conduct comprehensive comparison and ablation experiments at both the model and system levels. Regarding the proposed DSN-BR model, we first perform comparison experiments by selecting different kernel sizes at the first layer of various backbones. After carefully considering the trade-off between model performance and computational efficiency, we choose the resnet50 backbone with a 7×7 kernel size. Subsequently, we conduct ablation experiments on different networks to evaluate the impact of the BR module. Our results demonstrate that the BR module significantly improves semantic boundary segmentation and inter-class differentiation in terms of both model performance and visualization results. Ultimately, the DSN-BR model achieves a notable 90.5% mean intersection over union (MIoU). Regarding the designed and implemented online defect detection system, we commence by comparing the defect detection results of diverse networks using various evaluation metrics. Furthermore, we evaluate and compare the

detection rate of our method under different parameter settings. Ultimately, the system attains an outstanding 98.25% F1-Score and operates at a high speed of 14.12 f/s. We also investigate the current detection performance of a drug testing manufacturer, which demonstrates an average accuracy of approximately 95% and a detection rate of 6 f/s. Comparative analysis reveals that our system outperforms in terms of detection performance and efficiency, enabling reduced human resource allocation and alleviating workers' labor intensity. The intelligent approach facilitates efficient quality monitoring on the production line, supporting large-scale online quality control and providing essential technical assistance for comprehensive surface quality control in the manufacturing process.

4 Conclusions

In this study, we proposed an online detection method and system for surface defects of pharmaceutical products on aluminum-plastic blister packages based on deep learning. The study revealed the following conclusions:

(1) The system firstly attempts to integrate image features and semantic segmentation knowledge into the detection of aluminum-plastic blister packages.

(2) The system successfully addresses the challenges of large-scale online quality control using a data-driven approach. From the perspective of production practice, it improves the efficiency of quality monitoring in intelligent production lines, reducing the labor intensity of human resources and workers.

(3) From the management perspective, the system standardizes the control process of pharmaceutical manufacturing and enhances the management capabilities of the manufacturing process.

(4) The effectiveness and feasibility of the proposed method is verified through experiments, providing technical support for comprehensive control of surface quality in the manufacturing processes of products in various fields.

Liu *et al. Chinese Journal of Mechanical Engineering*   (2024) 37:86

Page 20 of 21

However, the actual manufacturing process may introduce time-varying phenomena and additional challenges, such as simultaneous detection of multiple drugs, stricter boundary accuracy requirements, and higher real-time detection rates for certain drugs. These uncertainties present more complex problems to solve. In future research, we will focus on exploring the adaptability of detection methods in uncertain environments, further exploring the potential characteristics of the data, and improving the system's flexibility and applicability. Additionally, efforts will be made to integrate the detection system into the production line to enable real-time observation and analysis of the packaging system's operations, facilitating timely troubleshooting, maintenance adjustments, and ensuring the continuous closed-loop operation of the automatic line. This will contribute to providing a comprehensive inspection solution for pharmaceutical applications.

Authors' Contributions
ML was in charge of the whole trial; YG and XW wrote the manuscript and response to laboratory analyses; CL and JH assisted with supervision. All authors read and approved the final manuscript.

Availability of Data and Materials
The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Competing Interests
The authors declare no competing financial interests.

References

[1] X Y Lawrence, J Woodcock. FDA pharmaceutical quality oversight. *Int. J. Pharmaceut.*, 2015, 491(1-2): 2-7.

[2] M S Firouz, K Mohi-Alden, M Omid. A critical review on intelligent and active packaging in the food industry: Research and development. *Food Res. Int.*, 2021, 141: 110113.

[3] C Zhuang, Z Liu, J Liu, et al. Digital twin-based quality management method for the assembly process of aerospace products with the grey-markov model and apriori algorithm. *Chin. J. Mech. Eng.*, 2022, 35: 105.

[4] N S Arden, A C Fisher, K Tyner, et al. Industry 4.0 for pharmaceutical manufacturing: Preparing for the smart factories of the future. *Int. J. Pharmaceut.*, 2021, 602: 120554.

[5] X Wang, M Liu, M Ge, et al. Research on assembly quality adaptive control system for complex mechanical products assembly process under uncertainty. *Comput. Ind.*, 2015, 74: 43-57.

[6] Y Gao, X Li, X V Wang, et al. A review on recent advances in vision-based defect recognition towards industrial intelligence. *J. Manuf. Syst.*, 2022, 62: 753-766.

[7] E Oztemel, S Gursev. Literature review of industry 4.0 and related technologies. *J. Intell. Manuf.*, 2020, 31: 127-182.

[8] P Osterrieder, L Budde, T Friedli. The smart factory as a key construct of industry 4.0: A systematic literature review. *Int. J. Prod. Econ.*, 2020, 221: 107476.

[9] S Minaee, Y Boykov, F Porikli, et al. Image segmentation using deep learning: A survey. *IEEE T. Pattern Anal.*, 2021, 44(7): 3523-3542.

[10] Z Li, F Liu, W Yang, et al. A survey of convolutional neural networks: Analysis, applications, and prospects. *IEEE T. Neur. Net. Lear.*, 2021, 33(12): 6999-7019.

[11] Y Fu, G Zhu, M Zhu, et al. Digital twin for integration of design-manufacturing-maintenance: An overview. *Chin. J. Mech. Eng.*, 2022, 35: 80.

[12] P Rajpurkar, E Chen, O Banerjee, et al. AI in health and medicine. *Nat. Med.*, 2022, 28(1): 31-38.

[13] Q Zhang, K Barri, S K Babanajad, et al. Real-time detection of cracks on concrete bridge decks using deep learning in the frequency domain. *Engineering*, 2021, 7(12): 1786-1796.

[14] X Wang, M Liu, C Liu, et al. Data-driven and knowledge-based predictive maintenance method for industrial robots for the production stability of intelligent manufacturing. *Expert Syst. Appl.*, 2023, 234: 121136.

[15] H Guan, M Liu. Domain adaptation for medical image analysis: A survey. *IEEE T. Bio-Med. Eng.*, 2021, 69(3): 1173-1185.

[16] J Zhang, C W Liu, F R Bi, et al. Fault feature extraction of diesel engine based on bispectrum image fractal dimension. *Chin. J. Mech. Eng.*, 2018, 31: 40.

[17] G Fu, P Sun, W Zhu, et al. A deep-learning-based approach for fast and robust steel surface defects classification. *Opt Laser Eng.*, 2019, 121: 397-405.

[18] H Bay, A Ess, T Tuytelaars, et al. Speeded-up robust features (SURF). *Comput. Vis. Image Und.*, 2008, 110(3): 346-359.

[19] X Liu, T Fu, Z Pan, et al. Automated layer segmentation of retinal optical coherence tomography images using a deep feature enhanced structured random forests classifier. *IEEE J. Biomed. Health*, 2018, 23(4): 1404-1416.

[20] J Zhang, R X Gao. Deep learning-driven data curation and model interpretation for smart manufacturing. *Chin. J. Mech. Eng.*, 2021, 34: 71.

[21] V Monga, Y Li, Y C Eldar. Algorithm unrolling: Interpretable, efficient deep learning for signal and image processing. *IEEE Signal Proc. Mag.*, 2021, 38(2): 18-44.

[22] Y Gong, M Liu, X Wang, et al. Research on surface defects detection method and system in manufacturing processes based on the fusion of multi-scale features and semantic segmentation for intelligent manufacturing. *J. Intell. Fuzzy Syst.*, 2023, 44(4): 6463-6481.

[23] S Niu, B Li, X Wang, et al. Defect image sample generation with GAN for improving defect recognition. *IEEE T Autom. Sci. Eng.*, 2020, 17(3): 1611-1622.

[24] R Chen, D Cai, X Hu, et al. Defect detection method of aluminum profile surface using deep self-attention mechanism under hybrid noise conditions. *IEEE T. Instrum. Meas.*, 2021, 70: 1-9.

[25] Y Zhang, W Wang, Z Li, et al. Development of a cross-scale weighted feature fusion network for hot-rolled steel surface defect detection. *Eng. Appl. Artif. Intel.*, 2023, 117: 105628.

[26] R Hao, B Lu, Y Cheng, et al. A steel surface defect inspection approach towards smart industrial monitoring. *J. Intell. Manuf.*, 2021, 32: 1833-1843.

[27] S Dong, P Wang, K Abbas. A survey on deep learning and its applications. *Computer Science Review*, 2021, 40: 100379.

[28] D Tabernik, S Šela, J Skvarč, et al. Segmentation-based deep-learning approach for surface-defect detection. *J. Intell. Manuf.*, 2020, 31(3): 759-776.

[29] J Long, E Shelhamer, T Darrell. Fully convolutional networks for semantic segmentation. *CVPR*, 2015: 3431-3440.

[30] C Yu, J Wang, C Peng, et al. Learning a discriminative feature network for semantic segmentation. *CVPR*, 2018: 1857-1866.

[31] M Mohri, A Rostamizadeh, A Talwalkar. *Foundations of machine learning*. MIT Press, 2018.

[32] T Bouwmans, S Javed, M Sultana, et al. Deep neural network concepts for background subtraction: A systematic review and comparative evaluation. *Neural Networks*, 2019, 117: 8-66.

[33] K Simonyan, A Zisserman. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, 2014.

Liu *et al. Chinese Journal of Mechanical Engineering*   *(2024) 37:86*

Page 21 of 21

[34] K He, X Zhang, S Ren, et al. Deep residual learning for image recognition. *CVPR*, 2016: 770-778.

[35] C Peng, X Zhang, G Yu, et al. Large kernel matters—improve semantic segmentation by global convolutional network. *CVPR*, 2017: 4353-4361.

[36] M Everingham, S M A Eslami, L Van Gool, et al. The pascal visual object classes challenge: A retrospective. *Int. J. Comput. Vision*, 2015, 111: 98-136.

Mingzhou Liu born in 1968, is currently a professor at *Hefei University of Technology, China.* He received his Ph.D. degree from *V.N. Karazin Kharkiv National University, Ukraine,* in 1999. His research interests include manufacturing process optimization operation theory, quality control, intelligent manufacturing, cloud manufacturing, scheduling, remanufacturing.

Yu Gong born in 1997, is currently a doctoral candidate at *School of Mechanical Engineering, Hefei University of Technology, China.*

Xiaoqiao Wang received the B.S. degree in industry engineering from *Hefei University of Technology, China,* in 2010, and the Ph.D. degree in industry engineering from *Hefei University of Technology, China*, in 2015. Since 2015, he has been a lecturer at *Mechanical Engineering Department, Hefei University of Technology, China*. His research interests include quality management, intelligent manufacturing system, artificial intelligence, and machine vision applications.

Conghu Liu received the Ph.D. degree in industry engineering from *Hefei University of Technology, China*, in 2016. He is currently a professor at *Suzhou University, China*. He worked as a postdoctor at *Shanghai Jiaotong University, China,* from 2016 to 2019, and is currently a visiting scholar at *Tsinghua University, China*. He presided over and participated in a number of national and provincial scientific research projects. To data, he has authored/coauthored more than 50 academic papers, and most of paper has been published in TOP Journal with some high citations. His research interests include remanufacturing, sustainable manufacturing, and management.

Jing Hu born in 1976, is currently a professor at *Hefei University of Technology, China*. Her research interests include human factors engineering, manufacturing process optimization operation theory, intelligent manufacturing and scheduling.