



Article

# Evaluation Method of Potato Storage External Defects Based on Improved U-Net

**Kaili Zhang** <sup>1</sup>, **Shaoxiang Wang** <sup>1</sup>, **Yaohua Hu** <sup>2,\*</sup>, **Huanbo Yang** <sup>1</sup>, **Taifeng Guo** <sup>1</sup> and **Xuemei Yi** <sup>1,\*</sup>

<sup>1</sup> College of Mechanical and Electronic Engineering, Northwest A&F University, Yangling 712100, China; kelly999@nwafu.edu.cn (K.Z.); wsx@nwafu.edu.cn (S.W.); yanghuanbo@nwafu.edu.cn (H.Y.); gtf2021@nwafu.edu.cn (T.G.)

<sup>2</sup> College of Optical, Mechanical and Electrical Engineering, Zhejiang A&F University, Hangzhou 311300, China

\* Correspondence: huyaohua@zafu.edu.cn (Y.H.); xuemei\_yi@nwsuaf.edu.cn (X.Y.); Tel.: +86-15291680166 (Y.H.); +86-15388627635 (X.Y.)

**Abstract:** The detection of potato surface defects is the key to ensuring potato storage quality. This research explores a method for detecting surface flaws in potatoes, which can promptly identify storage defects such as dry rot and the shriveling of potatoes. In order to assure the quality and safety of potatoes in storage, we used a closed keying method to obtain the pixel area of the mask image for a potato's surface. The improved U-Net realizes the segmentation and pixel area measurement of potato surface defects and enhances the feature extraction capability of the network model by adding a convolutional block attention module (CBAM) to the baseline network. Compared with the baseline network, the improved U-Net showed a much better performance with respect to MIoU (mean intersection over union), precision, and  $F_{\beta}$ , which were improved by 1.99%, 8.27%, and 7.35%, respectively. The effect and efficiency of the segmentation algorithm were also superior compared to other networks. Calculating the fraction of potato surface faults in potato mask images allows for the quantitative detection of potato surface problems. The experimental results show that the absolute accuracy of the quantitative potato evaluation method proposed in this study was greater than 97.55%, allowing it to quantitatively evaluate potato surface defects, provide methodological references for potato detection in the field of deep processing of potatoes, and provide a theoretical basis and technical references for the evaluation of potato surface defects under complex lighting conditions.



**Citation:** Zhang, K.; Wang, S.; Hu, Y.; Yang, H.; Guo, T.; Yi, X. Evaluation Method of Potato Storage External Defects Based on Improved U-Net. *Agronomy* **2023**, *13*, 2503. <https://doi.org/10.3390/agronomy13102503>

Academic Editor: Enrico Corrado Borgogno Mondino

Received: 11 August 2023

Revised: 12 September 2023

Accepted: 15 September 2023

Published: 28 September 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Potato is one of the four major food crops for human beings [1]. Because of its rich nutrition, it is also used as a staple food in many countries. However, potato tubers can have different degrees of surface defects during growth, harvesting, and transportation. These defects affect their nutritional value on the one hand and their economic value on the other. Therefore, it is of great significance to accurately detect defects and evaluate the quality of potato tubers [2].

Recent studies have focused on the quality of round fruits and vegetables such as dates, mangoes, apples, prunes, tomatoes, and cabbage [3]. At present, most enterprises and production bases prefer to use the human eye to grade and evaluate the quality of potatoes. However, this detection method is labor-intensive, and, as it mainly depends on individual a priori experiences, there is a certain rate of error. Therefore, some corresponding intelligent detection techniques have been proposed to qualitatively or quantitatively detect and evaluate potato defects. Referring to China's national potato grading standard 《NY/T 1066-2006》, the surface defects of potatoes mainly include the following: (a) rot, (b) dry rot, (c) greenish skin, (d) cracked seams, (e) surface bruises, (f) internal damage, (g)

growth cracks, (h) secondary growth, (i) scab, (j) hollow heart, (k) black heart, (l) brown spot, (m) insect eyes or rodent spots, (n) brown spots, and (o) insect eyes or rat bites [4].

The current intelligent detection technology mainly relies on the spectral technology inspection method and the inspection method based on machine vision. Spectral technology is generally used to analyze the optical principles of near-infrared spectroscopy (NIR) and uses hyperspectral technology to obtain the surface information of the object. Then, according to the optical principles of spectroscopy, the surface defects of potatoes are gradually analyzed and judged [5]. Rizaa et al. [6] achieved automation of postharvest grading of potatoes and improvement in the quality. The authors identified a variety of external defects; the UV-vis-NIR regional diffuse reflectance characteristics of various surface defects of potatoes were measured to categorize them for external defects with high accuracy, but the pre-treatment process was complicated. Deng et al. [7] carried out a principal component analysis of the external reflectance spectral images of potatoes with different defects, selected different feature bands, and built a support vector machine model for the corresponding spectral data, and the prediction accuracies of the prediction sets all achieved very satisfactory results. Spectral technology combined with traditional machine vision can achieve relatively good experimental results and can be better used in the field of potato surface defect detection; however, with the high levels of pre-processing and experimental processing needed, there is a large amount of subjective human influence, which can affect the group detection of the potato surface defects. Hassankhani et al. [8] used traditional machine vision to classify the acquired potato surface defects through MATLAB using color features as well as physical properties of defective potato surfaces, with a classification accuracy of up to 97.67%; however, the method was unable to achieve a quantitative evaluation or, in terms of efficiency, to meet the current demand for efficient detection.

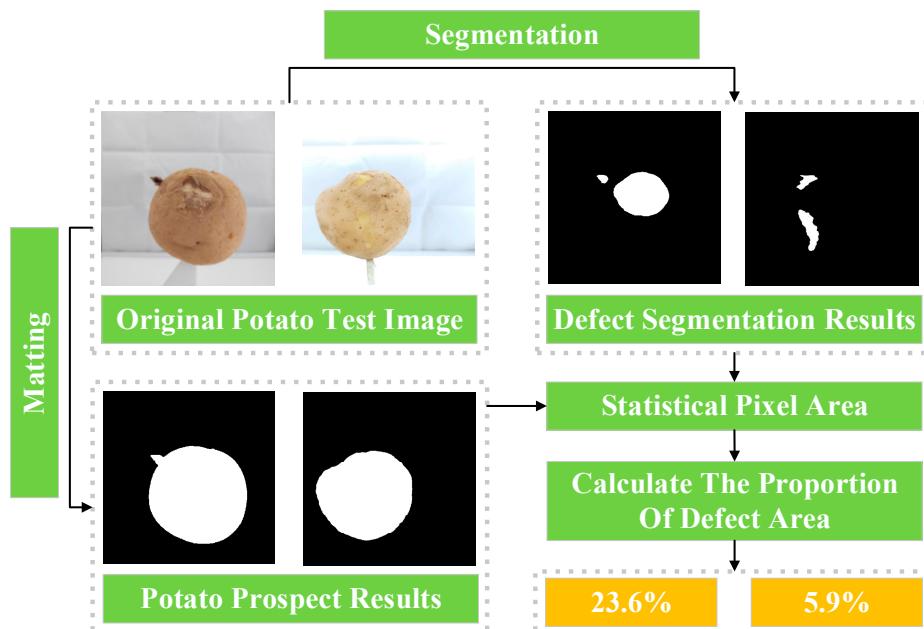
In recent years, machine learning and deep learning, in particular, have been applied to a variety of fields, most notably fruit and vegetable detection [9,10], medical image detection [11,12], industrial product detection [13,14], and other fields [15,16] as a result of the continuous improvement in computer performance. Among them, fruit and vegetable inspection mainly includes semantic segmentation, target detection, and image classification. For example, Qiao et al. [17] proposed a red date counting approach based on enhanced YOLOv5, which uses ShuffleNet V2 as the model's foundation to increase the model's detection capability and lighten the model's weight. A new data loading module, Stem, was also suggested, and PANet was replaced with BiFPN to increase the model's capability for feature fusion and increase its accuracy. To count red dates, the upgraded YOLOv5s detection model was employed. The experimental results showed that the model parameters were lowered while the accuracy rose by 4.3%, leading to improved experimental findings. To achieve the automatic detection of jujube crack, Zheng et al. [18] presented an attention feature fusion network (AFFU-Net) based on U-Net architecture and integrated it with the loss and residual mixing refinement module (RRM). To categorize the surface flaws (rot, cracks, wounds, and spots) of green plums, Zhou et al. [19] employed a WideResNet model using the WideResNet50 AdamW-Wce model, which has outstanding performance in terms of recall, precision, etc. A ConvNeXt-based, high-precision lightweight classification network was proposed by Jiang et al. [20], which greatly reduces the number of model parameters while still guaranteeing that the model precision criterion is satisfied. This also provides a helpful recommendation for upgrading the automatic detection system used by the kiwifruit sector. Nithya et al. [21] proposed a computer vision recognition system based on a convolutional neural network (CNN) to realize the automatic detection of mango in the field of mango surface defects detection. This system also offers some better references for the automatic detection of other round-like fruits. Additionally, Sun et al. [22] investigated citrus surface defect detection, combining all pertinent machine learning and image processing techniques, cutting the detection algorithm's average running time to 0.84097 seconds and increasing the accuracy of citrus area detection to 95.32%. Liang et al. [23] proposed a semantic segmentation approach based on the BiSeNet V2 deep

learning network to segregate the defective section of defective apples in order to meet the need for automation that is expected to continue to rise. Better experimental results were attained, and model pruning was used to optimize the YOLOv4 network's topology. The YOLOv4 network that had been pruned increased the accuracy of finding faulty areas in apple photos. In their investigation of potato surface defects, Wang et al. [24] used deep convolutional neural network (DCNN) models to find problems on the potato's surface. Three DCNN models—SSD Inception V2, RFCN ResNet101, and Faster RCNN ResNet101 base model—were each optimized through migration learning. Results of the tests showed accuracy percentages of 92.5%, 95.6%, and 98.7%, respectively. RFCN ResNet101 demonstrated the best overall performance in terms of detection speed and accuracy, showing greater all-around performance. A multi-type identification and classification system for potatoes was also developed by Yang et al. [25], which used improved YOLOv3 tiny models and multispectral (MS) images. By incorporating the Res2Net module into the YOLOv3 small net, the multi-type defect detection network (MDDNet) was developed to identify potatoes with multiple types of faults, considerably improving detection accuracy. At the moment, the classification element of potato surface defect detection is the main focus, and it is mostly focused on improving surface defect classification accuracy. From past research in the direction of image segmentation, the U-Net network model structure could be improved to have excellent performance in defect detection and image segmentation in various fields [26–28]. With reference to the latest potato storage testing protocol issued by China's Ministry of Agriculture and Rural Development on 3 February 2023, the current research direction does not meet the needs of the potato industry. In this research, a potato external defect evaluation method based on improved U-Net is proposed to realize the automatic detection and evaluation of potatoes. The main contributions are as follows:

- (1) To further evaluate the potato surface defects accurately, this research uses the attention mechanism to improve the U-Net, which improves the detection accuracy of potato surface defects and realizes the precise and accurate detection of potato surface defects under complex lighting conditions.
- (2) To realize the potato foreground extraction under complex lighting conditions, this research adopts a closed keying method to accurately acquire the potato surface mask image, which provides pre-preparation for the evaluation of potato surface defects.
- (3) To realize the accurate evaluation of potato surface defects under complex lighting conditions, this study combines the keying method with the improved U-net to propose a quantitative evaluation method of potato surface defects, which can accurately evaluate the percentage of potato surface defects.

## 2. Materials and Methods

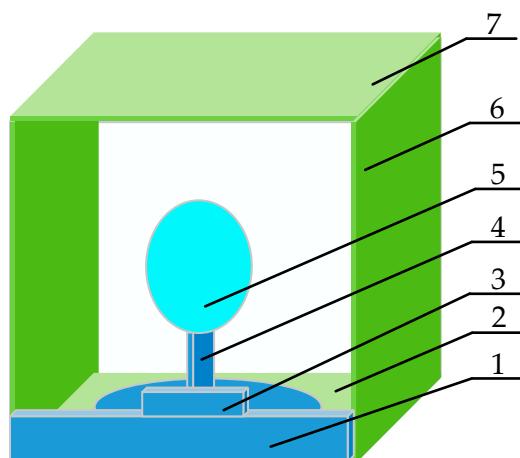
In this research, as shown in Figure 1, we propose a quantitative method for evaluating surface defects of potatoes. Firstly, a closed keying method was utilized to extract the pixel area of the complete foreground information of the potato, which provided support for the subsequent surface defects percentage of the potato. Then, the improved U-net was utilized for segmenting the surface defects of the potato to obtain the pixel area of the surface defects, and the pixel area percentage of the surface defects was solved using a division operation to realize quantitative evaluation of the surface defects of the potato.



**Figure 1.** Flowchart of the quantitative method for evaluating surface defects of potato.

### 2.1. Acquisition of Datasets

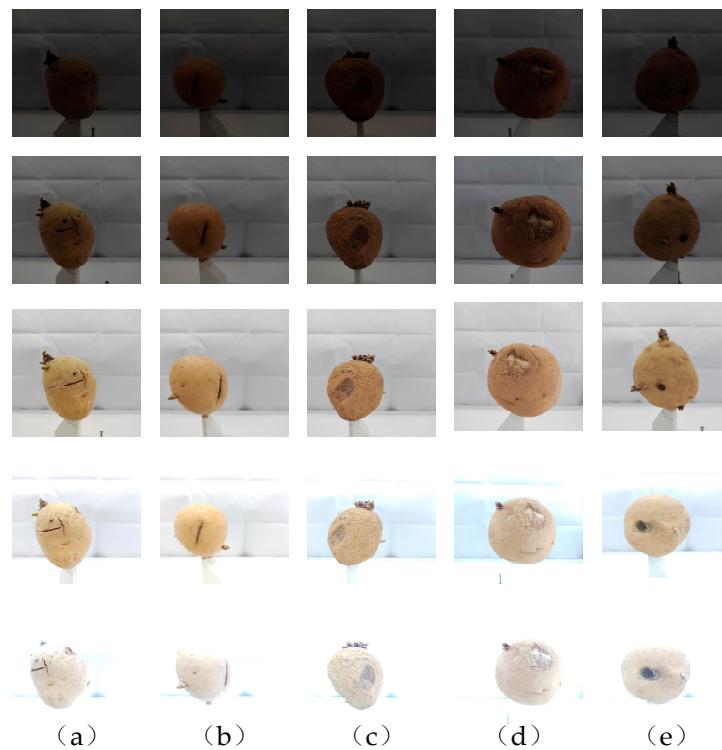
In order to simulate the experimental conditions of natural illumination, we built an image acquisition device, as shown in Figure 2. The luminous whiteboard was a  $38 \times 38$  cm shadowless lamp made in China, which was used to regulate the light intensity during the image acquisition process. We designed five gradient levels of light intensity to replace different ranges of natural light. The rotating base was used to obtain images of all surface defects of the potato, and the collection stand was used to support the handpiece. The overall image acquisition device was constructed with a rigid metal skeleton and covered by a photographic white cloth, and the whole device could realize the effect of shadowless image acquisition; a single variable control of light intensity could be realized.



**Figure 2.** Schematic diagram of the image acquisition device. 1—Mobile phone holder, 2—Luminous whiteboard, 3—Mobile phone, 4—Rotating support base, 5—Potato, 6—Luminous whiteboard, 7—Luminous whiteboard.

The experimental images shown in Figure 3 were collected in the laboratory of Northwest A&F University, Yang Ling District, Shaanxi Province, China, using the Pride Pro60 cell phone with a resolution of  $1920 \times 1080$  pixels and, to better adapt to the training of the network, our images were uniformly adjusted to  $640 \times 640$  pixels. Our dataset contained

1080 images of potato defects, of which 216 images of potato surface defects existed for each illumination level, and the images contained the potatoes themselves and their surface defects. One of the potato varieties was “Xisen 6”, independently bred in China, and the main types of surface defects were (a) cracks, (b) mechanical damage, (c) sprouting, (d) dry rot, and (e) insect eyes, with about 43 images of potatoes with each type of defect captured for use in the experiment.



**Figure 3.** Some samples of defective potatoes from our dataset under different light. (a) Cracks (b) Mechanical damage (c) Sprouting (d) Dry rot (e) Worm eyes.

## 2.2. A Closed-Form Matting Scheme for Natural Images

Extracting foreground items from an image based on little user input is a crucial problem in image and video processing since it is an interactive digital keying method [29]. From the perspective of computer vision, this task is extremely difficult because it is a pathological problem in which foreground and background colors must be estimated on each pixel, as well as the foreground opacity (alpha) of a single-color channel. The techniques used are to strictly limit the estimation to a small portion of the image, estimate the foreground and background colors based on the known pixel neighborhoods, or invert foreground and background colors iteratively. Following that, a cost function is created using the foreground and background colors:

$$J(\alpha) = \alpha^T L \alpha \quad (1)$$

A quadratic cost function can be obtained by eliminating foreground and background colors in the alpha channel in the obtained expression (1). This allows us to solve the equation for a sparse linear system to find the global opaque alpha blur. Next, for the closed equations, analyzing the eigenvectors of the sparse matrices allows the estimation of the features of the scheme, which are very close to the matrices in the spectral image style algorithm. With very little user input, high-quality keying can be obtained on natural images, resulting in highly accurate foreground images. In this study, we applied this closed keying scheme to obtain the surface mask image of a potato. The main process is shown in Figure 4. First, we applied different colors to distinguish the potato and its

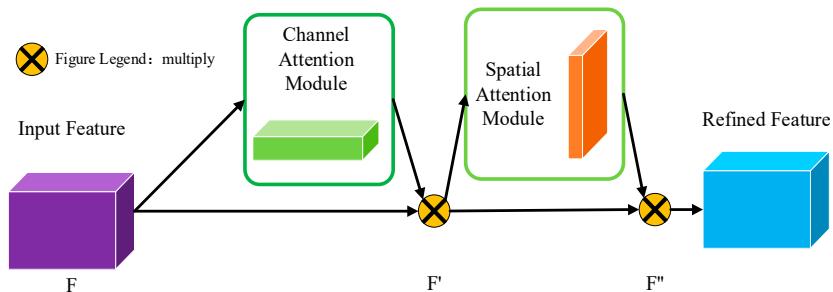
background. For this part of the operation, we simply needed to mark the two parts of the region. Then, we completed the algorithm's parameter settings and ran the closed keying algorithm to obtain the accurate surface mask image of the potato.



**Figure 4.** Details of the potato surface mask image acquisition technique.

### 2.3. CBAM Attention Mechanism

A convolutional attention module proposed by Yun et al. [30] in 2018 innovatively proposed an attention mechanism that fused channel attention with spatial attention, giving a lightweight and adaptable attention module for feed-forward convolutional neural networks. It can be easily integrated into any CNN network due to its simplicity and effectiveness; the authors' experiments have shown that there is sustained improvement in classification and detection performance for the various models. CBAM is a combinatorial model that combines channel attention and spatial attention in order to increase the model's expressive power. The overall structure of the CBAM module is shown in Figure 5.

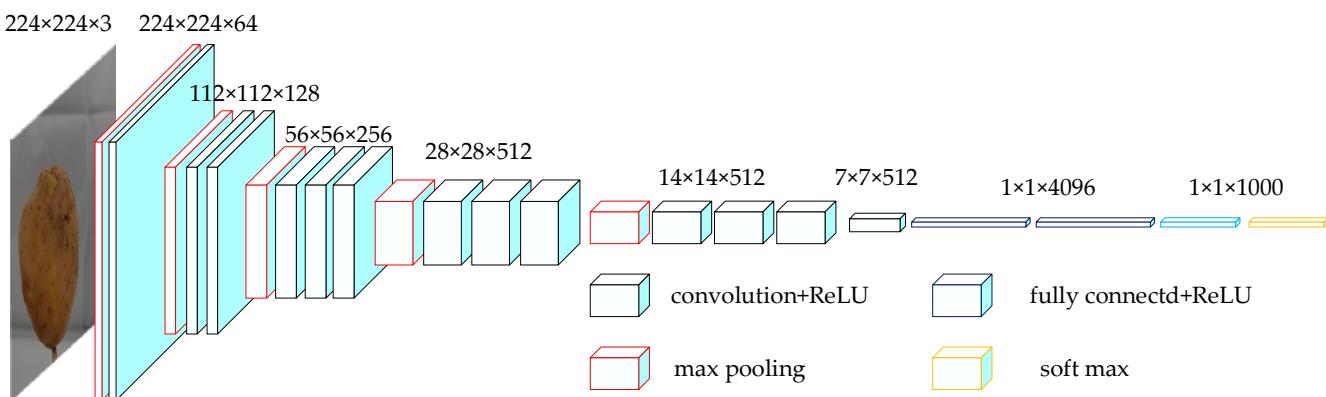


**Figure 5.** Overall structure of the CBAM module.

### 2.4. Improved U-Net Model

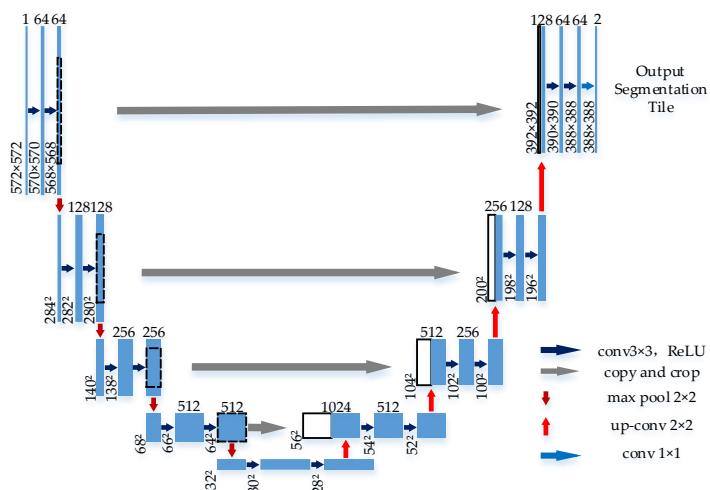
With the rapid development of various types of algorithms in the field of deep learning, convolutional neural networks (CNNs) are widely used for classification tasks, where the output is the class labeling of the entire image. In the biomedical field, where doctors need to pathologically analyze a patient's lesion area, a more advanced network model is needed, i.e., a network that can predict the class of a pixel point with a small set of training images and can color map the pixel point to form a more complex and rigorous judgment.

As a consequence, scientists from Google DeepMind and the Visual Geometry Group at the University of Oxford created a fresh deep convolutional neural network [31]: VGGNet. VGGNet investigates the relationship between a convolutional neural network's depth and performance, and it has successfully built a 16–19 layer deep convolutional neural network, demonstrating that expanding the network's depth can somewhat affect its final performance. This has led to a significant reduction in error rates while at the same time having very strong expandability and the generalization of the migration to other image data. The processing of each layer of VGGNet is shown in Figure 6.



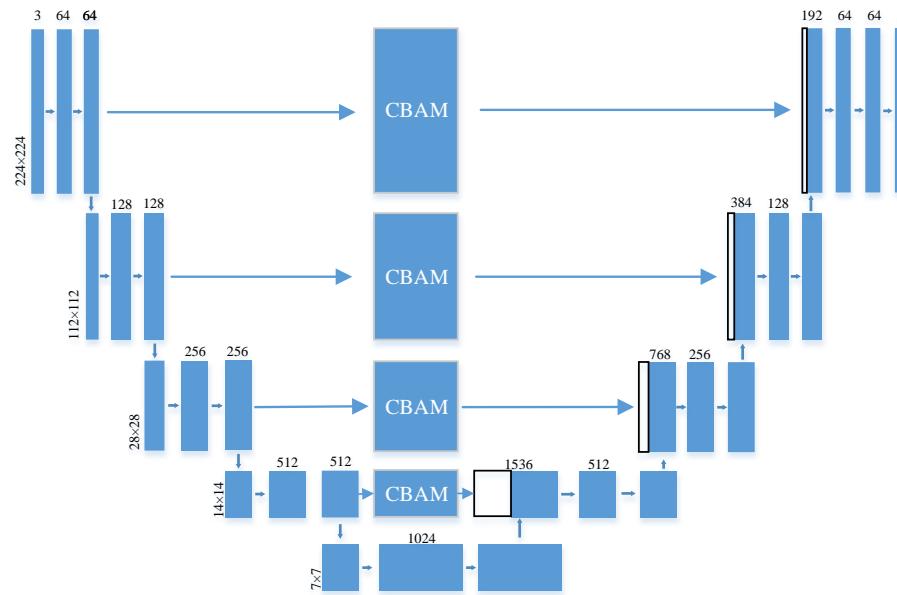
**Figure 6.** Structure of the VGG network model.

One year after the development of the VGGNet network was completed, the U-Net network was designed. The U-Net network structure was first proposed by Ronneberger et al. [32] in 2015. The core idea of this image is the introduction of jump connections, which makes the accuracy of image segmentation much better. The main structure of the U-Net network consists of three parts: decoder, encoder, and bottleneck layer. The process is illustrated in Figure 7.



**Figure 7.** Structure of the U-Net network model.

In this study, we adopted the same structure as VGG in the down-sampling part to enhance the feature extraction ability of the network and added the CBAM attention enhancement module in each layer of the down-sampling cropping and replication process to further enhance the feature expression ability of the network. The specific process is to first use the VGG network to down-sample the feature map of the loaded image. After down-sampling four times, resulting in feature maps of different sizes, the amount of data are gradually reduced in the process of each sampling, but will inevitably make the feature image of the features of the distortion; therefore, we will further use the U-Net to enhance the feature image in the up-sampling process to obtain a more accurate defect segmentation image. In the encoder part of the model, VGG is used to enhance the feature representation ability of the network; in the decoder part, U-Net is used to enhance the fine segmentation ability of the network, and the CBAM attention module is added in each cropping and copying process mainly to improve the comprehensive performance of the model. The specific structural framework of the scheme is shown in Figure 8.



**Figure 8.** Improved U-Net network structure diagram.

### 3. Experiments and Results

The hardware and software utilized in the potato surface defect evaluation method are introduced in this section. After that, the ablation experiments we designed to determine the impact of various modules on the performance of the model are described. In order to confirm the viability of the model, we first introduce the evaluation indices that were used in the experiment, examine the validity and viability of the results of the foreground extraction experiment, and then examine the segmentation effect and segmentation performance of the model that we created on the potato surface flaws. In order to confirm the viability of the approach developed in this work, the results of foreground extraction and defect segmentation were combined with the relevant calculations, and the error between the actual results and the experimental results was assessed.

All of our experiments were carried out with identical hardware and software to ensure fairness, and the model suggested in this study was based on the enhanced U-Net potato surface defect segmentation model, coded in Python, and put to the test by the Pytorch deep learning framework. The testing environment and hardware are shown in Table 1.

**Table 1.** Experimental environment.

| Configuration           | Parameter                     |
|-------------------------|-------------------------------|
| CPU                     | Intel (R) Core (TM) i7-10700K |
| GPU                     | NVIDIA GeForce RTX 3070       |
| Accelerated environment | CUDA11.1 CUDNN8.2.1           |
| Development environment | Pycharm 2021.3.2              |
| Operating system        | Ubuntu 18.04                  |

#### 3.1. Evaluation Indicators

The semantic segmentation model is based on pixel accuracy (PA), mean pixel accuracy (MPA), mean intersection over union (MIoU), precision, recall, and the  $F_\beta$  value that combines the measures of precision and recall [33].

The PA pixel precision is the ratio of the number of correctly classified pixels to the total number of pixels in an image, while the MPA average pixel precision is the average of all classes of pixel precision. IOU intersection over union is the ratio of the intersection of segmented and labeled maps to the concatenation of the two sets, which usually indicates the degree of overlap between segmented and labeled maps, and the MIoU average

intersection over union is the average of the cumulative pixel intersection and union ratios for each class of pixels. The precision ratio indicates the proportion of samples that are truly positive among those recognized as positive by the model. The recall ratio shows how many the classifier can predict among the actual positive samples. The formulas for pixel precision, average pixel precision, average intersection and merger ratio, precision, recall, and  $F_\beta$  are given in Equations (2)–(7):

$$PA = \frac{\sum_{i=0}^k p_{ii}}{\sum_{i=0}^k \sum_{j=0}^k p_{ij}} \times 100\% \quad (2)$$

$$MPA = \frac{1}{k+1} \sum_{i=0}^k \frac{p_{ii}}{\sum_{j=0}^k p_{ij}} \times 100\% \quad (3)$$

$$MIOU = \frac{1}{k+1} \sum_{i=0}^k \frac{p_{ii}}{\sum_{j=0}^k + \sum_{j=0}^k p_{ji} - p_{ii}} \times 100\% \quad (4)$$

$$F_\beta = (1 + \beta^2) \cdot \frac{Precision \cdot Recall}{\beta^2 \cdot Precision + Recall} \times 100\% \quad (5)$$

$$Precision = \frac{TP}{TP + FP} \times 100\% \quad (6)$$

$$Recall = \frac{TP}{TP + FN} \times 100\% \quad (7)$$

**True Positive (TP):** true class. The true class of the sample is positive and the result recognized by the model is also positive.

**False Negative (FN):** false negative class. The true class of the sample is a positive class, but the model recognizes it as a negative class.

**False Positive (FP):** false positive category. The true category of the sample is negative, but the model recognizes it as positive.

**True Negative (TN):** The true category of the sample is negative and the model recognizes it as negative.

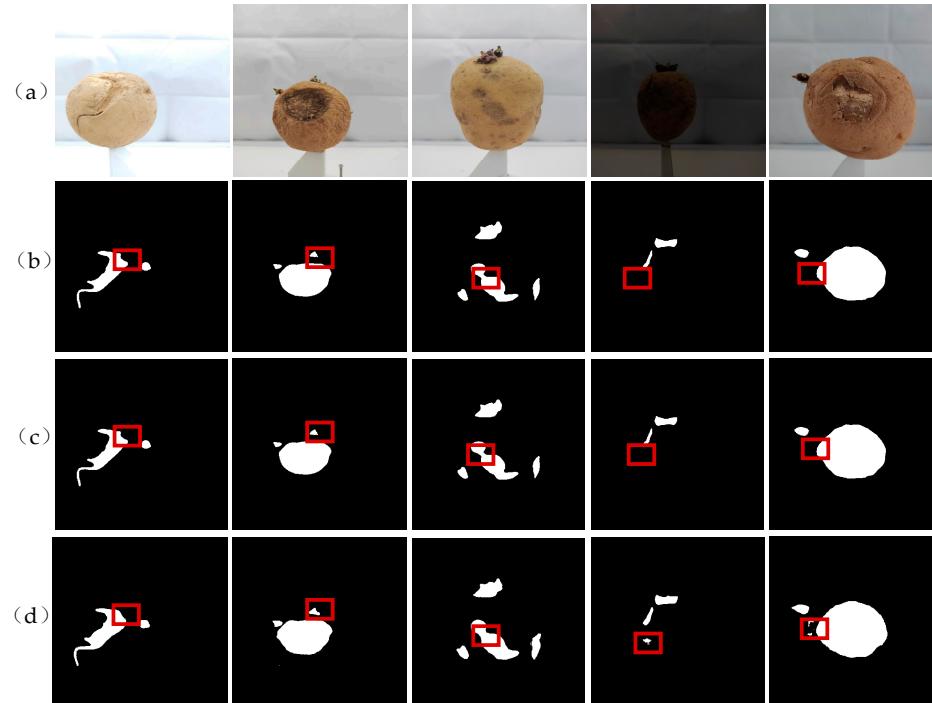
### 3.2. Ablation Experiments

In order to validate the effectiveness of our work, an ablation study was conducted in which three different architectures, including U-Net (Baseline), VGG+Baseline, and VGG+CBAM+Baseline, were used to obtain the best experimental results by comparing the structure of our proposed model.

As shown in Table 2, with the introduction of VGG and CBAM, the performance of the network improved accordingly. Specifically, MIoU and  $F_\beta$  improved, and MPA decreased. Compared with the baseline network, the improved U-Net had its MIoU and  $F_\beta$  improved by 1.99% and 7.35%, respectively, and the MPA decreased by 5.13%. In order to further validate the feasibility of the modeled backbone network improvements, further improvements were carried out using MobilenetV3 [34] and ShufflenetV2 [35] in the network architecture; our model structure improved MIoU by 3.86% and 9.8%, F1 by 3.42% and 5.8%, and reduced MPA by 1.06% and 2.54%, respectively. As shown in Figure 9, from the qualitative results, with the improved model structure, the improved U-Net network was closer to the true value of detection at the edges as well as in the details compared to the baseline network.

**Table 2.** Comparison of ablation experiment results.

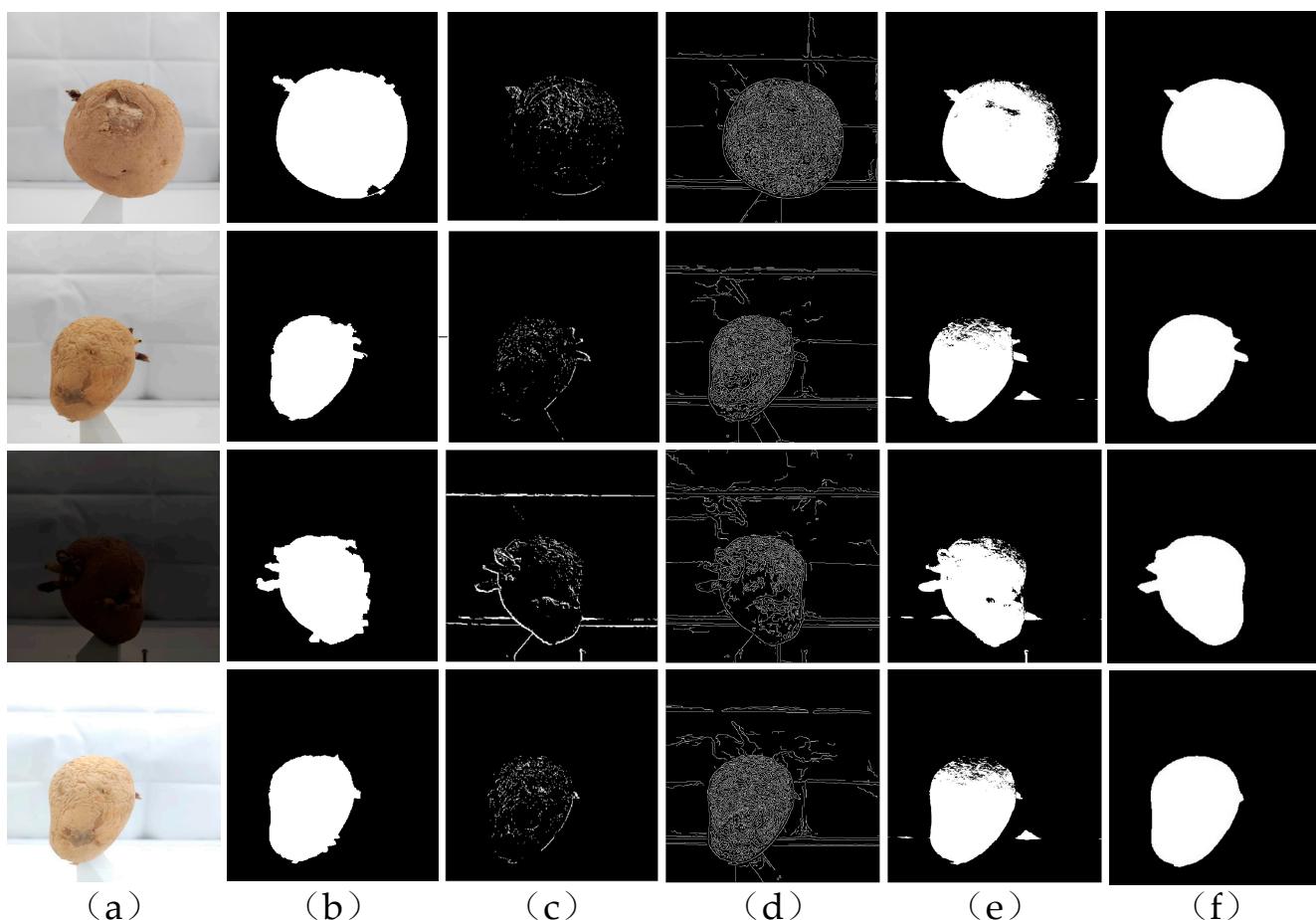
| Model                | MIoU/%       | MPA/%        | PA/%         | Precision/%  | Recall/%     | $F_\beta$ /% |
|----------------------|--------------|--------------|--------------|--------------|--------------|--------------|
| U-Net                | 89.03        | 97.24        | 99.33        | 82.13        | 95.03        | 83.06        |
| U-Net+VGG            | 90.75        | 94.84        | 99.54        | 90.26        | 89.91        | 90.23        |
| U-Net + ShuffletV2*1 | 87.16        | 96.18        | 99.20        | 79.65        | 93.00        | 80.61        |
| U-Net + MobilenetV3  | 87.60        | 92.58        | 99.29        | 85.49        | 75.91        | 84.61        |
| U-Net + CBAM + VGG   | <b>91.02</b> | <b>95.12</b> | <b>99.50</b> | <b>90.40</b> | <b>90.49</b> | <b>90.41</b> |

**Figure 9.** Comparison of ablation experiments. (a) Test Images (b) Our method (c) U-Net (d) Ground truth (GT).

### 3.3. Comparison of Potato Surface Mask Extraction

The extraction of the potato surface mask is an indispensable experimental step in the realization of a potato surface defect evaluation method. The accuracy of potato surface mask acquisition will directly affect the accuracy of the potato surface defect evaluation method; therefore, we chose a closed matting method to obtain a better potato surface mask image. As shown in Figure 10, four algorithms, Sobel [36], Canny [37], k-means clustering [38], and GrabCut [39], were selected to compare with our method.

The experimental results show that the difference between potato and background was a bit greater compared to the recognition and segmentation of potato defects. However, in this research experiment, with the change in light intensity, the difference between the potato and the background gradually lessened. Therefore, when we used Sobel and Canny for segmentation in the experimental process, we did not obtain the ideal experimental results; Sobel and Canny operators show that the experimental results have the problems of incomplete edge contour and generating a lot of noise. The main reason for this is that when the light intensity was too bright or too dark, the threshold between the potato itself image and the background made it difficult to set the k-means clustering algorithm. Although it could obtain better experimental results, there were still a large number of distortions in some of the image information at the edge of the clusters. However, our chosen closed keying algorithm did not involve the selection of the threshold value, and the mask image of the potato could be obtained directly, which could realize the optimal segmentation of the potato.



**Figure 10.** Comparison of potato surface mask extraction results (a) Test image (b) Grabcut (c) Sobel (d) Canny (e) K-means (f) Matting.

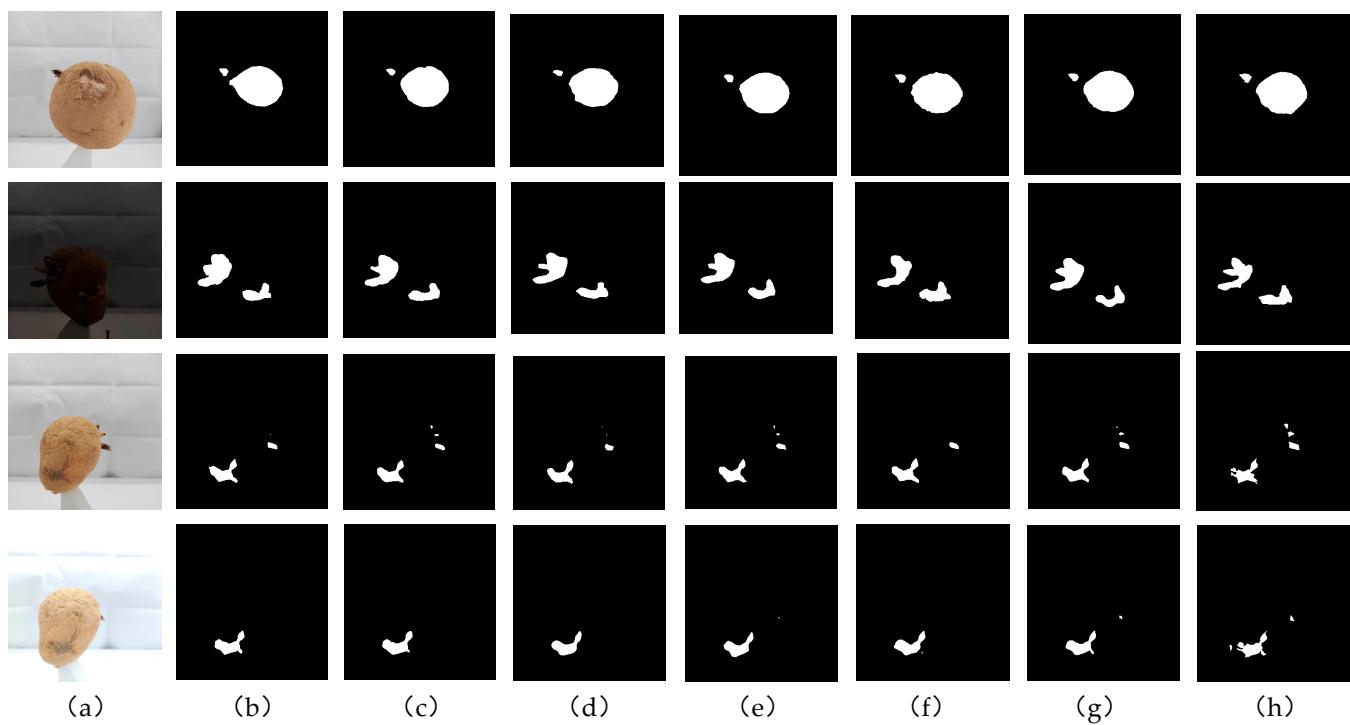
### 3.4. Comparison of Defect Segmentation Results

We compared our proposed improved U-Net model with five relatively state-of-the-art deep learning SOD models, including PSPNet\_RESNet50 [40], PSPNet\_MobileNetV2 [41], FCN [42], DeepLabv3\_plus [43], and SegNet [44]. In order to ensure the fairness of the experiments, all comparison experiments were conducted using the authors' default parameters, and all experiments were conducted on the same training set and test set.

**Qualitative Evaluation:** As seen in Figure 11, when compared to other cutting-edge models, our strategy produced good results. Our solution specifically addresses the issues of internal region inhomogeneity and hazy, noisy borders. The significant results from the various approaches were contrasted with the actual data. The method's forecasts, which had more defined boundaries and entire interior zones, were the ones that were closest to reality.

**Quantitative Evaluation:** In order to further validate the model's adaptability to complex lighting environments, we designed harsh environments with different lighting gradients during the dataset production process. Especially regarding  $F_\beta$ , the degree of improvement in our improved network model was, to some extent, much better than the other models.

**Running time:** In real-time quality inspection of potato production lines, detection effectiveness is a crucial consideration. The enhanced U-Net and the other five cutting-edge models' average detection times are displayed in Table 3. As can be seen, our suggested upgraded U-Net had equivalent detection efficiency to other approaches, such as SegNet, with an average detection time of 0.084 s/picture.



**Figure 11.** Comparison of segmentation results of different models **(a)** Test image **(b)** DeepLabv3\_plus **(c)** FCN **(d)** PSPNet\_MobilenetV2 **(e)** PSPNet\_RESNet50 **(f)** SegNet **(g)** Improved U-Net **(h)** Labeling.

**Table 3.** Results of evaluation indicators for different models.

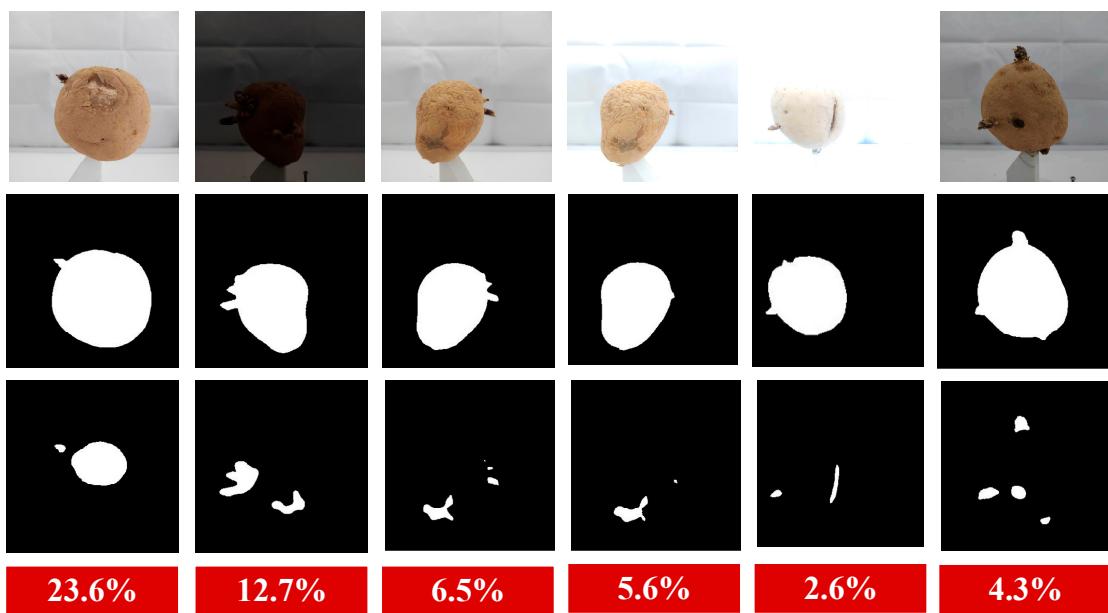
| Model                     | PA/%         | MPA/%        | MIoU/%       | Precision/%  | Recall/%     | $F_\beta$ /% | Running Time<br>(s/Sheet) |
|---------------------------|--------------|--------------|--------------|--------------|--------------|--------------|---------------------------|
| PSPNet_RESNet50           | 99.46        | 93.74        | 89.22        | 88.79        | 87.74        | 88.70        | 0.103                     |
| PSPNet_MobilenetV2        | 99.37        | 92.46        | 87.65        | 87.46        | 85.21        | 87.27        | 0.098                     |
| FCN                       | 99.47        | 93.70        | 89.36        | 89.23        | 87.65        | 89.10        | 0.178                     |
| DeepLabv3_plus            | 99.44        | 93.35        | 88.91        | 88.83        | 86.96        | 88.67        | 0.124                     |
| SegNet                    | 99.28        | 90.68        | 85.92        | 86.71        | 81.66        | 86.27        | 0.084                     |
| <b>U-Net + CBAM + VGG</b> | <b>99.50</b> | <b>95.12</b> | <b>91.02</b> | <b>90.40</b> | <b>90.49</b> | <b>90.41</b> | <b>0.084</b>              |

### 3.5. External Defect Evaluation Methods

In order to quantitatively evaluate the degree of damage of potato surface defects, inspired by the spot detection area of potato late blight in general [45], we introduced the damage ratio  $\alpha$  to reflect the degree of damage, which was defined as (Equation (8))

$$\alpha = \frac{S_{\text{defect}}}{S_{\text{potato}}} \quad (8)$$

where defect and potato are, respectively, the zones of defects and potatoes. We determined the defect's pixel count on the potato's surface and the amount of potato in the mask image, respectively. The process's visualization results are displayed in Figure 12. The outcomes of the experiment demonstrate that the approach can offer quantitative and qualitative evaluation indices for detecting potato surface defects. Table 3 shows the outcomes of evaluation indicators for several approaches.



**Figure 12.** Calculation results of potato surface defect evaluation methods.

In order to reflect the accuracy of the measurement method, absolute accuracy (AA) was used to evaluate the accuracy of the method in this study. The absolute accuracy is used to evaluate the accuracy of defect area segmentation [46], which is calculated as (Equations (9)–(11))

$$Ra = La / Sa \times 100\% \quad (9)$$

$$Ri = Li / Si \times 100\% \quad (10)$$

$$AA = (1 - |Ra - Ri|) \times 100\% \quad (11)$$

Among them,  $La$  and  $Sa$  are the results obtained from manual segmentation of the samples using the software as the true value of the number of pixel points of the target potato and the number of pixel points of the defects, i.e., the target potato and the defects were selected using the quick selection tool in the software and filled with different colors. Then, the counting tool was used to count the pixel points of the potato mask image and the pixel points of the defects.  $Ra$  is the calculated relative defect area as the true value,  $Li$  and  $Si$  are the pixel points of the target potato mask image and the defect area or the area of the target potato mask image and the defects using the method of the present study, and  $Ri$  is the relative defect area calculated by the algorithm of the present study. The larger the  $AA$  is, the greater the accuracy of the calculation method and the better the performance.

From Table 4, it is evident that the absolute accuracy of our experimental results declined when the light intensity was either too high or too low. The highest accuracy was achieved when the light intensity was close to natural light conditions.

**Table 4.** Statistics on the accuracy of research methods.

| Light Intensity Level | Level 1 | Level 2 | Level 3 | Level 4 | Level 5 |
|-----------------------|---------|---------|---------|---------|---------|
| Absolute accuracy (%) | 98.43   | 98.92   | 99.36   | 98.65   | 97.55   |
| Deviation (%)         | ±0.3    | ±0.09   | ±0.32   | ±0.07   | ±0.4    |

#### 4. Discussion

As shown in Figure 11, we compared the experimental results of the improved U-Net with the segmentation results of five state-of-the-art algorithms, PSPNet\_RESNet50 [38],

PSPNet\_MobilenetV2 [39], FCN [40], DeepLabv3\_plus [41], and SegNet [42]. The segmentation effect of our method was the best for the potato defective dataset. Furthermore, the ablation experiment showed that our improvement could further enhance the detection accuracy of the potato surface defect model.

In concrete terms, our network can satisfy real-time detection efficiency while assuring that the segmentation of potato surface flaws is better than that of existing state-of-the-art networks in the details. Our network performed very well in regard to the details of potato surface defects compared to other cutting-edge networks. Because the networks do not pay enough attention to the underlying characteristics that reflect the morphological information of the objects, the edge regions of the segmented pictures of PSPNet\_RESNet50 and FCN specifically were partially missing. Additionally, in order to convey rich feature information, PSPNet\_RESNet50 and FCN fused multi-scale features via complex jump connections; nonetheless, this had a similar impact on the network's detection efficiency, as evidenced by the running time in Table 3. As shown in Figure 11, although the models achieved higher segmentation results than those of PSPNet\_MobilenetV2 and DeepLabv3\_plus, some target details were lost because the features were not fused. Deep learning is now flourishing in the area of computer vision. SegNet, a representative work in the field of image segmentation, consists of a unique lightweight All-MLP decoder that combines local attention and global attention with a hierarchical converter encoder and performs exceptionally well on open datasets [47]. However, as observed in Figure 11g, SegNet's segmentation mask was imprecise and lacked local information, particularly at the edges. The segmentation of the image into small chunks for encoder coding was the cause of this issue, which led to inadequate acquisition of local features, inconsistent details, and subpar segmentation of the image at the edges. Our enhanced U-Net performed similarly to SegNet in terms of efficiency and greatly exceeded it at the edges and in the fine details.

Our improved U-Net network model showed good improvement in all quantitative evaluation indexes. Particularly in terms of MPA, Recall, and  $F_\beta$ , the specific data were increased to 95.12%, 90.49%, and 90.41%, respectively. In comparison to other quantitative indicator approaches, as shown in Table 3, our strategy could meet the demand for real-time detection and had the highest detection effectiveness among all methods except for SegNet. SegNet replaces time-consuming convolutional feature extraction with effective self-attention, which enhances detection speed to some extent. However, as shown in Table 3, our modified U-Net beat the impacts of other network models in several of the six evaluation categories. It perfectly satisfied the criteria of our evaluation system. As a result, it is clear that our improved network model is entirely dependable and viable.

Our method can also obtain the most accurate potato surface mask picture compared to that of other methods. It performed notably well when information was acquired from the edge of the potato surface. In contrast to previous mask image acquisition methods, spectral analysis, and traditional machine learning operators have been used by certain researchers to construct mask pictures of potato boundaries and defective regions for the evaluation methodology of potato surface flaws. After the construction, the damage ratio can be calculated by dividing it by the number of masks [48]. Using the usual Sobel [36] operator, which generates a lot of noise, it is difficult to precisely remove flaws from intricate spectral images. These segmentation techniques, along with other methods, such as the Canny operator [37], k-means clustering method [38], and Grabcut algorithm [39], are unable to generate adequate experimental results due to problems such as inappropriate threshold demarcation or insufficient clustering. Comparing the Grabcut technique to the keying algorithms we have employed on our dataset, there is also a certain amount of inaccuracy in the extraction results at the edges.

In contrast to earlier strategies for spotting potato surface problems, spectral imaging methods have recently been adopted by certain researchers. Our approach combines image spectroscopy with deep learning, and the accuracy is increased by 4.86% when compared to Yang et al. [25] and other researchers. Our detection efficiency is not highly efficient, but it still fully satisfies the requirements for real-time detection and can be quanti-

tatively achieved without a lot of pre-processing and processing conditions for the assessment of potato surface flaws. Although our detection accuracy is lower than that of certain studies, such as Wang et al. [24], who used deep transfer learning to recognize potato surface defects, our system has more detection types and a more accurate evaluation. There are numerous deep learning researchers who have conducted extensive studies on surface defect detection [48], but when compared to our method, these researchers' techniques mainly concentrate on the classification of potato surface defects, whereas our evaluation method has very good quantitative performance.

Since natural lighting conditions are the most common in our daily working environment, this study demonstrates the general feasibility of our method. In future work, more consideration should be given to the effects of the curvature of the potato surface defects during image acquisition as well as the planarization of the potato surface defects, where there is a large gap between the actual size of the image and the actual size of the defects in the edge portion of the image when the stereoscopic 3D object is placed in the plane. In addition, we consider adding three-dimensional curvature features to the acquired two-dimensional image, which could further improve the accuracy of the acquisition of the true value of potato surface defects. This could help improve the detection accuracy of the potato surface defect evaluation method.

## 5. Conclusions

This research proposes a potato surface defect evaluation method based on the enhanced U-Net, which employs the improved U-Net to acquire the pixel area of the potato surface flaws and then combines it with the potato surface mask image acquisition method to quantitatively evaluate potato surface defects. This method attempts to increase the economic benefits of the potato business by providing a more quantitative assessment of surface flaws in potatoes. The key conclusions are as follows:

- (1) The improved U-Net network has an excellent performance in all evaluation indices, allowing for more exact segmentation effects. The modified U-Net is employed as the backbone network in this study, and the model's feature expression ability is enhanced by the addition of the CBAM module.
- (2) The closed potato mask image acquisition approach employed in this work may produce more accurate potato surface mask images, especially in the finer details of the segmentation effect, as compared to other traditional methods and general machine learning algorithms.
- (3) The quantitative percentage of the potato surface defects was calculated in accordance with the pixel areas of the potato surface mask image and potato surface defects in order to realize the quantitative evaluation of potato surface defects. Based on the results of the absolute accuracy evaluation, the accuracy of this method can meet the requirements of real-time detection, and its performance is very good in the detection of potato surface defects.

With a wide variety of applications, the potato surface defect evaluation approach put forward in this study has a high detection accuracy and can be adjusted relative to the majority of difficult lighting settings. It can also give workers in the edible potato sector technical assistance and theoretical support to provide inspiration for researchers looking to identify surface flaws in rounded fruits.

**Author Contributions:** Data curation, K.Z., S.W., H.Y. and T.G.; funding acquisition, Y.H.; investigation, K.Z.; methodology, K.Z. and X.Y.; project administration, Y.H.; validation, Y.H., X.Y., S.W. and H.Y.; writing—original draft, K.Z.; writing—review and editing, Y.H., H.Y., S.W., T.G. and X.Y. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was financially supported by the National Natural Science Foundation of China (32171894 (C 0043619) and 31971787(C 0043628)).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** We are grateful to Yichen Qiao and Peng Zhang for data collection.

**Conflicts of Interest:** The authors declare that they have no conflict of interest.

## References

1. Zhang, H.; Fen, X.U.; Yu, W.U.; Hu, H.H.; Dai, X.F. Progress of potato staple food research and industry development in China. *J. Integr. Agric.* **2017**, *16*, 2924–2932. [[CrossRef](#)]
2. Sanchez PD, C.; Hashim, N.; Shamsudin, R.; Nor, M.Z.M. Applications of imaging and spectroscopy techniques for non-destructive quality evaluation of potatoes and sweet potatoes: A review—ScienceDirect. *Trends Food Sci. Technol.* **2020**, *96*, 208–221. [[CrossRef](#)]
3. Hasan, M.U.; Malik, A.U.; Ali, S.; Imtiaz, A.; Munir, A.; Amjad, W.; Anwar, R. Modern drying techniques in fruits and vegetables to overcome postharvest losses: A review. *J. Food Process. Preserv.* **2019**, *43*, e14280. [[CrossRef](#)]
4. Su, Q.; Kondo, N.; Li, M.; Sun, H.; Al Riza, D.F.; Habaragamuwa, H. Potato quality grading based on machine vision and 3D shape analysis. *Comput. Electron. Agric.* **2018**, *152*, 261–268. [[CrossRef](#)]
5. Shi, Y.; Wang, X.; Borhan, M.S.; Young, J.; Newman, D.; Berg, E.; Sun, X. A Review on Meat Quality Evaluation Methods Based on Non-Destructive Computer Vision and Artificial Intelligence Technologies. *Food Sci. Anim. Resour.* **2021**, *41*, 563–588. [[CrossRef](#)]
6. Al Riza, D.F.; Suzuki, T.; Ogawa, Y.; Kondo, N. Diffuse reflectance characteristic of potato surface for external defects discrimination. *Postharvest Biol. Technol.* **2017**, *133*, 12–19. [[CrossRef](#)]
7. Ji, Y.; Sun, L.; Li, Y.; Li, J.; Liu, S.; Xie, X.; Xu, Y. Non-destructive classification of defective potatoes based on hyperspectral imaging and support vector machine. *Infrared Phys. Technol.* **2019**, *99*, 71–79. [[CrossRef](#)]
8. Hassankhani, R. Potato surface defect detection in machine vision system. *Afr. J. Agric. Res.* **2012**, *7*, 844–850. [[CrossRef](#)]
9. Koirala, A.; Walsh, K.B.; Wang, Z.; McCarthy, C. Deep learning—Method overview and review of use for fruit detection and yield estimation. *Comput. Electron. Agric.* **2019**, *162*, 219–234. [[CrossRef](#)]
10. Zheng, Z.; Hu, Y.; Guo, T.; Qiao, Y.; He, Y.; Zhang, Y.; Huang, Y. AGHRNet: An attention ghost-HRNet for confirmation of catch-and-shake locations in jujube fruits vibration harvesting. *Comput. Electron. Agric.* **2023**, *210*, 107921. [[CrossRef](#)]
11. Zhao, M.; Jha, A.; Liu, Q.; Millis, B.A.; Mahadevan-Jansen, A.; Lu, L.; Landman, B.A.; Tyska, M.J.; Huo, Y. Faster Mean-shift: GPU-accelerated clustering for cosine embedding-based cell segmentation and tracking. *Med. Image Anal.* **2021**, *71*, 102048. [[CrossRef](#)] [[PubMed](#)]
12. Zhao, M.; Liu, Q.; Jha, A.; Deng, R.; Yao, T.; Mahadevan-Jansen, A.; Tyska, M.J.; Millis, B.A.; Huo, Y. VoxelEmbed: 3D instance segmentation and tracking with voxel embedding based deep learning. In Proceedings of the Machine Learning in Medical Imaging: 12th International Workshop, MLMI 2021, Held in Conjunction with MICCAI 2021, Proceedings 12, Strasbourg, France, 27 September 2021; pp. 437–446.
13. Zheng, Z.; Yang, H.; Zhou, L.; Yu, B.; Zhang, Y. HLU 2-Net: A residual U-structure embedded U-Net with hybrid loss for tire defect inspection. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 1–11.
14. Zhang, Y.; Wang, Y.; Jiang, Z.; Zheng, L.; Chen, J.; Lu, J. Subdomain adaptation network with category isolation strategy for tire defect detection. *Measurement* **2022**, *204*, 112046. [[CrossRef](#)]
15. You, L.; Jiang, H.; Hu, J.; Chang, C.H.; Chen, L.; Cui, X.; Zhao, M. GPU-accelerated Faster Mean Shift with euclidean distance metrics. In Proceedings of the 2022 IEEE 46th Annual Computers, Software, and Applications Conference (COMPSAC), Los Alamitos, CA, USA, 27 June–1 July 2022; pp. 211–216.
16. Alzubaidi, L.; Zhang, J.; Humaidi, A.J.; Al-Dujaili, A.; Duan, Y.; Al-Shamma, O.; Santamaría, J.; Fadhel, M.A.; Al-Amidie, M.; Farhan, L. Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *J. Big Data* **2021**, *8*, 53. [[CrossRef](#)] [[PubMed](#)]
17. Qiao, Y.; Hu, Y.; Zheng, Z.; Yang, H.; Zhang, K.; Hou, J.; Guo, J. A Counting Method of Red Jujube Based on Improved YOLOv5s. *Agriculture* **2022**, *12*, 2071. [[CrossRef](#)]
18. Zheng, Z.; Hu, Y.; Yang, H.; Qiao, Y.; He, Y.; Zhang, Y.; Huang, Y. AFFU-Net: Attention feature fusion U-Net with hybrid loss for winter jujube crack detection. *Comput. Electron. Agric.* **2022**, *198*, 107049. [[CrossRef](#)]
19. Zhou, C.; Wang, H.; Liu, Y.; Ni, X.; Liu, Y. Green Plums Surface Defect Detection Based on Deep Learning Methods. *IEEE Access* **2022**, *10*, 100397–100407. [[CrossRef](#)]
20. Yao, J.; Qi, J.; Zhang, J.; Shao, H.; Yang, J.; Li, X. A real-time detection algorithm for Kiwifruit defects based on YOLOv5. *Electronics* **2021**, *10*, 1711. [[CrossRef](#)]
21. Nithya, R.; Santhi, B.; Manikandan, R.; Rahimi, M.; Gandomi, A.H. Computer vision system for mango fruit defect detection using deep convolutional neural network. *Foods* **2022**, *11*, 3483. [[CrossRef](#)]
22. Sun, B.; Liu, K.; Feng, L.; Peng, H.; Yang, Z. The Surface Defects Detection of Citrus on Trees Based on a Support Vector Machine. *Agronomy* **2022**, *13*, 43. [[CrossRef](#)]
23. Liang, X.; Jia, X.; Huang, W.; He, X.; Li, L.; Fan, S.; Li, J.; Zhao, C.; Zhang, C. Real-Time grading of defect apples using semantic segmentation combination with a pruned YOLO V4 network. *Foods* **2022**, *11*, 3150. [[CrossRef](#)] [[PubMed](#)]

24. Wang, C.; Xiao, Z. Potato surface defect detection based on deep transfer learning. *Agriculture* **2021**, *11*, 863. [[CrossRef](#)]
25. Yang, Y.; Liu, Z.; Huang, M.; Zhu, Q.; Zhao, X. Automatic detection of multi-type defects on potatoes using multispectral imaging combined with a deep learning model. *J. Food Eng.* **2023**, *336*, 111213. [[CrossRef](#)]
26. Zhao, J.; Wang, J.; Qian, H.; Zhan, Y.; Lei, Y. Extraction of winter-wheat planting areas using a combination of U-Net and CBAM. *Agronomy* **2022**, *12*, 2965. [[CrossRef](#)]
27. Su, H.; Wang, X.; Han, T.; Wang, Z.; Zhao, Z.; Zhang, P. Research on a U-Net bridge crack identification and feature-calculation methods based on a CBAM attention mechanism. *Buildings* **2022**, *12*, 1561. [[CrossRef](#)]
28. Liu, L.; Liu, Y. Load image inpainting: An improved U-Net based load missing data recovery method. *Appl. Energy* **2022**, *327*, 119988. [[CrossRef](#)]
29. Levin, A.; Lischinski, D.; Weiss, Y. A closed-form solution to natural image matting. *IEEE Trans. Pattern Anal. Mach. Intell.* **2007**, *30*, 228–242. [[CrossRef](#)]
30. Woo, S.; Park, J.; Lee, J.-Y.; Kweon, I.S. Cbam: Convolutional block attention module. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 3–19.
31. Muhammad, U.; Wang, W.; Chattha, S.P.; Ali, S. Pre-trained VGGNet architecture for remote-sensing image scene classification. In Proceedings of the 2018 24th International Conference on Pattern Recognition (ICPR), Beijing, China, 20–24 August 2018; pp. 1622–1627.
32. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference, Proceedings, Part III 18, Munich, Germany, 5–9 October 2015; pp. 234–241.
33. Feng, M.; Lu, H.; Ding, E. Attentive feedback network for boundary-aware salient object detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 16–20 June 2019; pp. 1623–1632.
34. Koonce, B.; Koonce, B. MobileNetV3. In *Convolutional Neural Networks with Swift for Tensorflow: Image Recognition and Dataset Categorization*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 125–144.
35. Qian, H.; Zhou, Y.; Ding, P.; Feng, S. ConShuffleNet: An Efficient Convolutional Neural Network Based on ShuffleNetV2. In Proceedings of the International Conference on Guidance, Navigation and Control, Harbin, China, 5–7 August 2022; pp. 948–955.
36. Gao, W.; Zhang, X.; Yang, L.; Liu, H. An improved Sobel edge detection. In Proceedings of the 2010 3rd International Conference on Computer Science and Information Technology, Chengdu, China, 9–11 July 2010; pp. 67–71.
37. Rong, W.; Li, Z.; Zhang, W.; Sun, L. An improved CANNY edge detection algorithm. In Proceedings of the 2014 IEEE international conference on mechatronics and automation, Tianjin, China, 3–6 August 2014; pp. 577–582.
38. Patil, R.; Jondhale, K. Edge based technique to estimate number of clusters in k-means color image segmentation. In Proceedings of the 2010 3rd International Conference on Computer Science and Information Technology, Chengdu, China, 9–11 July 2010; pp. 117–121.
39. Chen, D.; Chen, B.; Mamic, G.; Fookes, C.; Sridharan, S. Improved grabcut segmentation via gmm optimisation. In Proceedings of the 2008 Digital Image Computing: Techniques and Applications, Canberra, Australia, 1–3 December 2008; pp. 39–45.
40. Liang, W.; Sheng, Y.; Zhou, Z.; Su, B.; Chen, J.; Lai, Y.; Lin, S.; Zhao, Z.; Ma, C. Multi-scale fusion based super-resolution underwater image segmentation network. In Proceedings of the AOPC 2022: Atmospheric and Environmental Optics, Beijing, China, 18–19 December 2022; pp. 98–103.
41. Liu, B.-Y.; Fan, K.-J.; Su, W.-H.; Peng, Y. Two-stage convolutional neural networks for diagnosing the severity of alternaria leaf blotch disease of the apple tree. *Remote Sens.* **2022**, *14*, 2519. [[CrossRef](#)]
42. Villa, M.; Dardenne, G.; Nasan, M.; Lettissier, H.; Hamitouche, C.; Stindel, E. FCN-based approach for the automatic segmentation of bone surfaces in ultrasound images. *Int. J. Comput. Assist. Radiol. Surg.* **2018**, *13*, 1707–1716. [[CrossRef](#)]
43. Sun, J.; Zhou, J.; He, Y.; Jia, H.; Liang, Z. RL-DeepLabv3+: A lightweight rice lodging semantic segmentation model for unmanned rice harvester. *Comput. Electron. Agric.* **2023**, *209*, 107823. [[CrossRef](#)]
44. Abdollahi, A.; Pradhan, B.; Alamri, A.M. An ensemble architecture of deep convolutional Segnet and Unet networks for building semantic segmentation from high-resolution aerial images. *Geocarto Int.* **2022**, *37*, 3355–3370. [[CrossRef](#)]
45. Lastochkina, O.; Pusenkova, L.; Garshina, D.; Kasnak, C.; Palamutoglu, R.; Shpirnaya, I.; Mardanshin, I.d.; Maksimov, I. Improving the biocontrol potential of endophytic bacteria *Bacillus subtilis* with salicylic acid against *Phytophthora infestans*-caused postharvest potato tuber late blight and impact on stored tubers quality. *Horticulturae* **2022**, *8*, 117. [[CrossRef](#)]
46. Oakley, S.P.; Portek, I.; Szomor, Z.; Turnbull, A.; Murrell, G.A.; Kirkham, B.W.; Lassere, M.N. Accuracy and reliability of arthroscopic estimates of cartilage lesion size in a plastic knee simulation model. *Arthrosc. J. Arthrosc. Relat. Surg.* **2003**, *19*, 282–289. [[CrossRef](#)]
47. Dai, Y.; Zheng, T.; Xue, C.; Zhou, L. SegMarsViT: Lightweight mars terrain segmentation network for autonomous driving in planetary exploration. *Remote Sens.* **2022**, *14*, 6297. [[CrossRef](#)]
48. Yang, J.; Li, S.; Wang, Z.; Dong, H.; Wang, J.; Tang, S. Using deep learning to detect defects in manufacturing: A comprehensive survey and current challenges. *Materials* **2020**, *13*, 5755. [[CrossRef](#)] [[PubMed](#)]