ISSN 2616-5910 Vol. 4, Issue 1: 30-42, DOI: 10.25236/AJALS.2023.040106

Defect Detection and Size Grading of Harvested Japanese Yam Using Computer Vision Technology

Cao Tianzhi[1,a,*], Hiroshi Okamoto[2,b,*]

[1]Graduate School of Agriculture, Hokkaido University, Sapporo, Hokkaido, 060-8589, Japan
[2]Research Faculty of Agriculture, Hokkaido University, Sapporo, Hokkaido, 060-8589, Japan
[a]sousou@bpe.agr.hokudai.ac.jp, [b]hiro@agr.hokudai.ac.jp
[*]Corresponding author

Abstract: *Root crop yield estimate and harvest technology, especially Japanese Yam harvest production, has a high demand for intelligent automation systems. This paper develops a Japanese Yam quality grading system based on computer vision and deep learning, which can automatically detect the shape and quality of Japanese Yam and grade the size of harvested. Specifically, based on Shuffle-Net and transfer learning, a lightweight deep learning model (CDD Net) was constructed to detect surface and shape defects of Japanese Yam. It's methods were also proposed based on minimum bounding rectangle (MBR) fitting and convex polygon approximation. Experimental results showed that the detection accuracy of the proposed CDD Net was 98.94% for binary classification (normal and defective) and 92.92% for multi-class classification (normal, curve, fork root, fracture), and demonstrated good performance both in time efficiency and detection accuracy. The size accuracy of MBR fitting and convex polygon approximation was 92.8% and 95.1% respectively. This study provides a practical method for defect detection and size grading, which has great potential for application in yield estimation of root crops.*

Keywords: *Computer Vision; Quality Grading System; Intelligent Automation System*

1. Introduction

Japanese Yam, also known as"nagaimo"in Japan, belongs to the Dioscorea genus and the Dioscoreaceae family. Dioscorea species are important sources of energy, micronutrients and phytochemicals with numerous health benefits, and they are widely cultivated in many parts of the world. In Japan Japanese Yam usually be grated and used as an ingredient in tororo udon/soba noodles. Quality assessment of Japanese Yam is a necessary process to detecting defects and grading size before entering the market. Quality inspection and graded sales can improve the quality and market competitiveness of Yam[1]. At present, Yam quality inspection and grading are mainly conducted manually and poses some problems such as low efficiency and accuracy, lack of consistency and uniformity, which greatly affect the quality and efficiency of Yam grading. Therefore, there is an urgent need for reliable and efficient quality processing methods[2]. As an emerging and ongoing technology, computer vision has provided an accurate, highly efficient and non-destructive method for grading agricultural products and has become one of the most popular methods for external quality inspection of Yams[3].Although traditional techniques have been successfully applied in the field of Yam defect detection, these traditional methods are complex as they require several different steps and have significant limitations in accuracy and flexibility, which greatly limits their application in practical situations[4]. To overcome these limitations, deep learning provided practical solutions to some of the most challenging problems in recent years. Deep learning can automatically extract features and has great advantages over traditional image processing methods[5-7]. It has been successfully used to detect defects in apples, cucumbers, and carrots[8]. Deep learning also used in the field of appearance quality inspection of root crops[9]. Although the performance of deep learning methods has been greatly improved, efficiency issues have also arisen. The detection efficiency of these methods makes challenging, to my knowledge, there is currently no deep learning based system for detecting Yam defects.In size grading research, previous studies mainly classified agricultural products into normal and defective products, without further grading normal products. In the actual processing of yam products, some workers first select defective yams, and then other workers further classify normal yams into different levels based on their size. Therefore, it is necessary to establish a specialized image acquisition system and develop effective methods for detecting and grading Yam defects, so that

ISSN 2616-5910 Vol. 4, Issue 1: 30-42, DOI: 10.25236/AJALS.2023.040106

emerging processing technologies can succeed and benefit the Japanese yam industry[10-12].

The purpose of this article is to develop an automated online system that can meet the requirements of accurate and real-time defect detection and automatic grading of yams. The innovation points of this article are as follows:

Develop a fully automated Japanese Yam defect detection and size grading system in farm field and before packing in factory; Develop a lightweight network (CDDNet) for Japanese Yam defect detection based on ShuffleNet and transfer learning; Propose a new Japanese Yam grading method based on minimum bounding rectangle (MBR) fitting and convex polygon approximation.

2. Materials and Method

2.1. Materials

Underground tuber is the main edible part of Japanese Yam. Two tractors are usually used when harvesting Yams. A plow on the first tractor scoops up the Yams, which are loaded into containers on the next tractor. Several people entered between the two tractors and manually removed the soil from the dug Yams and arranged them in the field. Several workers behind the second tractor conduct preliminary manual inspection and grading and manually load them into containers, picking out defective Yams, and then ordinary Yams are further divided into different grades according to size. In this paper, Yams of different sizes and shapes are selected as experimental samples, covering normal Yams and defective Yams (curve, fork roots, fracture, flat), as shown in Fig.1.

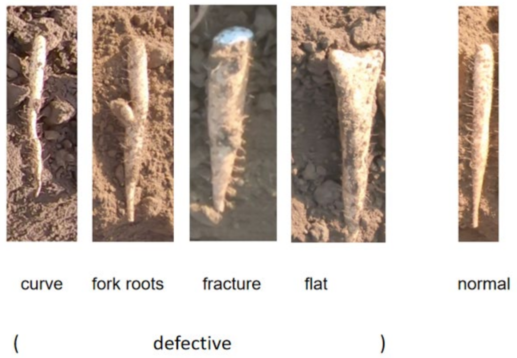


Figure 1: Defective Japanese Yams and Normal Japanese Yam.

2.2. Dataset collection and Image acquisition

The whole system process mainly included four steps: image acquisition, image preprocessing, defect detection and Yam grading, as presented in Fig. 2.

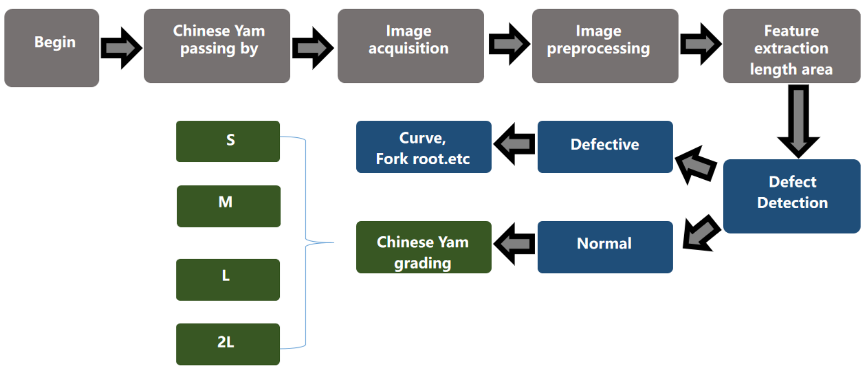


Figure 2: System technical flow chart

ISSN 2616-5910 Vol. 4, Issue 1: 30-42, DOI: 10.25236/AJALS.2023.040106

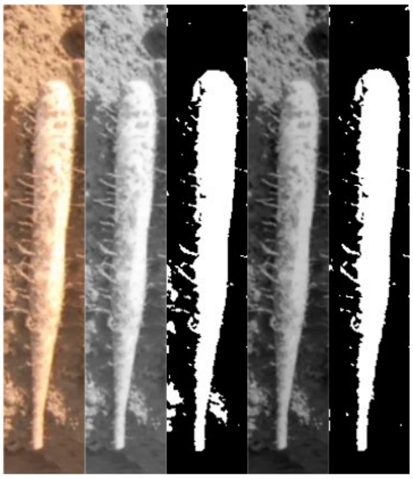
Figure 3: GoPro and dataset collection scenario

The dataset used in the experiment is mainly from the Michishita Hironaga Farm in Obihiro City, Hokkaido. The camera used is GoPro Hero 7 Black, and pixel count is 2.76 million pixels and shot at 24fps. The camera is equipped with image stabilization for mounting the tractor on unstable fields. Fix the GoPro camera on a tractor running in parallel and at constant speed, taking a bird's-eye view of the Japanese Yams placed on the field. During image acquisition, frame-triggered mode was used to capture images of Yams moving at a low speed, as in Fig.3. In order to achieve brightness equalization, referring to the Ohta color space, we used a new grayscale method calculated by Eq1.

$$I_{gray} = 2.5I_r - 2I_g - 0.5I_b \tag{1}$$

where I_r, I_g and I_b represent the red, green and blue components of the source image I_{gray}, respectively. The process of image preprocessing is illustrated in Fig.4.

(a)



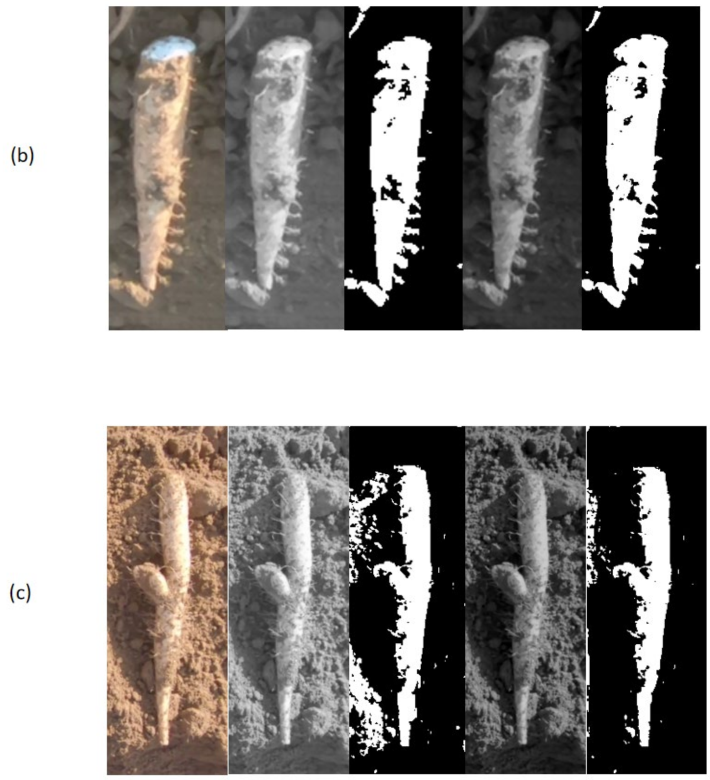


Figure 4: Illustration of image preprocessing. (a) Original Japanese Yam image of normal; gray image and its binary images; gray image calculated by Eq.1 and its binary images. (b) Original Japanese Yam image of fracture; gray image and its binary images; gray image calculated by Eq.1 and its binary images. (c) Original Japanese Yam image of fork root; gray image and its binary images; gray image calculated by Eq.1 and its binary images.

2.3. Japanese Yam defect detection model based on deep learning

Considering that ShuffleNet is a very lightweight network with very low complexity, and transfer learning provides a very effective technique for object recognition, especially when training data is limited, we constructed Yam defect detection network (CDDNet) based on ShuffleNet and transfer learning to realize online defect detection of Japanese Yams, as illustrated in Fig. 5.

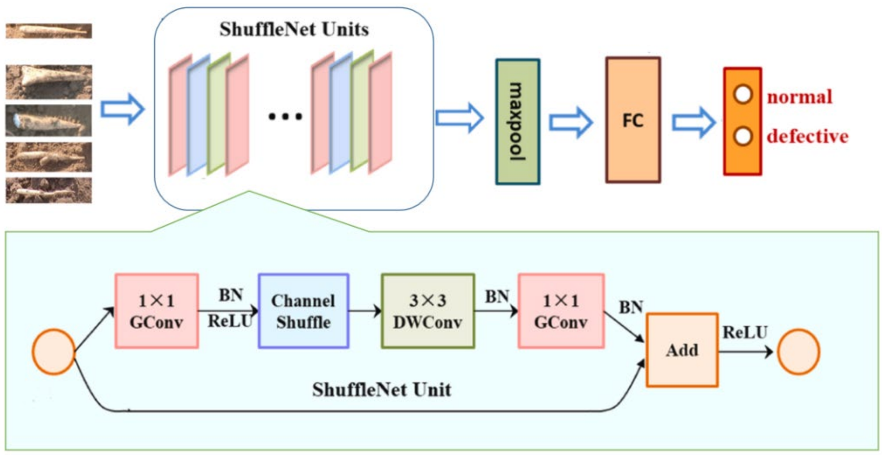


Figure 5: Model architecture of the proposed defect detection model. CDDNet model and its ShuffleNet unit.

ISSN 2616-5910 Vol. 4, Issue 1: 30-42, DOI: 10.25236/AJALS.2023.040106

2.4. Size grading methods based on MBR fitting and convex polygon approximation

After defect detection, normal Japanese Yams were further classified into different specifications according to their length. Therefore, it is necessary to establish the relationship between the pixel length and the actual Japanese Yam length. The pixel length can be measured by calculating the minimum bounding rectangle of the Japanese Yam region[13-14]. Therefore, a new regression method was proposed to predict the actual length from Yam images based on the minimum bounding rectangle approximation and linear fitting, which can be described as follows:

Step 1: Convert the source image (Fig. 6) into binary image.

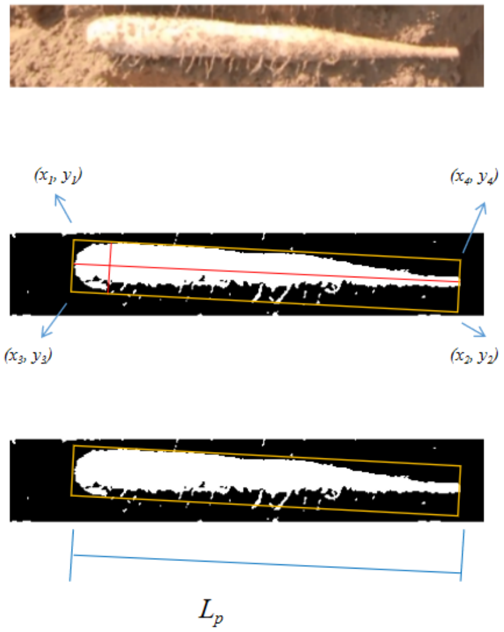


Figure 6: Fitting of Yam length: Source image; minimum bounding rectangle computation; calculation of L_p;

Step 2: Compute the minimum bounding rectangle of the Japanese Yam region (Fig. 6) based on approximation algorithm, and MBR[15]corner point positions *tl* (tl_x, tl_y), *tr (trx, try), bl (blx, bly), br (brx, bry)* can be obtained by the formula in Eqs. (2)–(5):

$$tl_x = \frac{x_1 \tan\theta + x_3 \cot\theta + y_3 - y_1}{\tan\theta + \cot\theta}$$
$$tl_y = \frac{y_1 \cot\theta + y_3 \tan\theta + x_3 - x_1}{\tan\theta + \cot\theta} \tag{2}$$

$$tr_x = \frac{x_1 \tan\theta + x_4 \cot\theta + y_4 - y_1}{\tan\theta + \cot\theta}$$
$$tr_y = \frac{y_1 \cot\theta + y_4 \tan\theta + x_4 - x_1}{\tan\theta + \cot\theta} \tag{3}$$

$$bl_x = \frac{x_2 \tan\theta + x_3 \cot\theta + y_3 - y_2}{\tan\theta + \cot\theta}$$
$$bl_y = \frac{y_2 \cot\theta + y_3 \tan\theta + x_3 - x_2}{\tan\theta + \cot\theta} \tag{4}$$

ISSN 2616-5910 Vol. 4, Issue 1: 30-42, DOI: 10.25236/AJALS.2023.040106

$$\mathrm{br}_x = \frac{x_2 \tan\theta + x_4 \cot\theta + y_4 - y_2}{\tan\theta + \cot\theta}$$
$$br_\mathrm{y} = \frac{y_2 \cot\theta + y_4 \tan\theta + x_4 - x_2}{\tan\theta + \cot\theta} \tag{5}$$

In the realm of bounding rectangle, where θ is the angel (radian) of the object orientation. (x1, y1) is the top edge point of Japanese Yam region, (x2, y2) is the bottom edge point, (x3, y3) is the left edge point, and (x4, y4) is the right edge point. These markers are used to define the four corners of the rectangle and thus determine the position and size of the bounding box. Through these coordinates, the position and size of the rectangle can be clearly specified for graphics processing and calculations.

Step 3: Calculate the pixel length L_p of Chinese Yam (Fig. 6).

Step 4: Establish the relationships between the actual length L_a (cm) and the pixel length L_p by linear fitting (Fig. 7):

$$\mathrm{L}_a = 0.1115 L_p - 0.3241$$

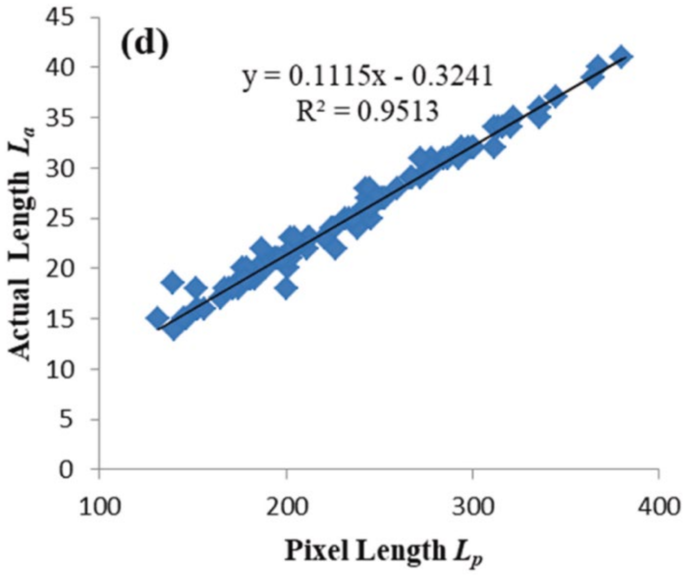


Figure 7: Linear fitting of L_a and L_p.

After obtaining the actual length L (cm), normal Japanese Yams can be divided into four size grades: S, M, L and 2L by:

Grade= { S:L_a≤30cm;

M: 30cm<L_a≤50cm;

L:50cm<L_a≤70cm;

2L: L_a >70cm;

According to the grading requirements for Japanese Yam sales, the normal shape of Japanese Yam should be natural and uniform. Based on the shape characteristics of Yams, a convex polygon approximation method is proposed, dividing each specification into two levels (Level 1 and Level 2) to maintain consistency and uniformity in appearance(Fig. 8). An external convex polygon[16] is generated to approximate the contour of the Japanese Yams, and shape regularity R_s was defined as:

$$\mathrm{R}_s = \frac{A_c}{A_p}$$

where A_c is the area of Japanese Yam region and A_p is the area of its external convex polygon. R_s is used to measure the regularity of Japanese Yam shape, and its value ranges from 0 to 1. The bigger the R_s, the more regular the Yam is. In addition, it is not affected by image size and Japanese Yam position,

ISSN 2616-5910 Vol. 4, Issue 1: 30-42, DOI: 10.25236/AJALS.2023.040106

and has good adaptability.

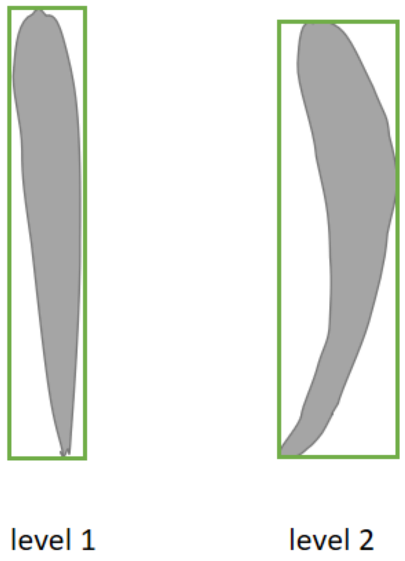


Figure 8: Two specifications divided with convex polygon.

A threshold T can be set to classify Japanese Yam into two levels (Fig.9):

$$level = \begin{cases} 1 & R_s \geq T \\ 2 & R_s < T \end{cases}$$

Figure 9: Two specifications divided into two levels

2.5. Algorithm of defect detection and grading

The algorithm of defect detection and grading is described as below:

ISSN 2616-5910 Vol. 4, Issue 1: 30-42, DOI: 10.25236/AJALS.2023.040106

Algorithm 1. Defect detection and grading

Algorithm 1: Defect detection and grading

Input:Source image I ;
Output:Decision signal s ;
Initialization s= -1 ;
for i **do**
Defect detection on image i by CDDNet: *dFlag*=CDDNet(i)
if(*dFlag*='defective')
S=0;
exit for
end if
end for
if(s=-1)
Calculate the actual length L_a of image I by MBR fitting method;
If($L_a \leq 30$)then classify I into two levels(s=1,2)by polygon Approximation method;
else if($L_a \leq 50$)then classify I into two levels(s=3,4);
else if($L_a \leq 70$)then classify I into two levels(s=5,6);
else classify I into two levels(s=7,8);
end if
end if
return *s;*

2.6. Evaluation standards

Five statistical parameters were calculated to evaluate the performance of the proposed model: Accuracy, Recall, Specificity, Precision and F1-Score, which are specified by equations (6), (7), (8), (9), (10):

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{6}$$

$$\mathrm{R}ecall = \frac{TP}{TP + FN} \tag{7}$$

$$\mathrm{S}pecificity = \frac{TN}{TN + FP} \tag{8}$$

$$\mathrm{P}recision = \frac{TP}{TP + FP} \tag{9}$$

$$\mathrm{F}_1 - Score = \frac{2 \times \mathrm{Pr}ecision \times \mathrm{Re}call}{\mathrm{Pr}ecision + \mathrm{Re}call} \tag{10}$$

Table 1: Description of four image datasets.

Dataset Name	Categories	Training size	Validation size	Total size
Dataset 1	Normal	1684	726	2439
	Defective	3579	1566	5227
Dataset 2	Normal	700	300	1000
	Fork roots	212	91	303
	Curve	409	75	478
	Fracture	988	422	1396
Dataset 3	S	-	238	238
	M	-	245	245
	L	-	211	211
	2L	-	211	211
Dataset 4	Level 1	-	649	649
	Level 2	-	217	217

where TP, TN, FP and FN represent the number of true positives, true negatives, false positives and

ISSN 2616-5910 Vol. 4, Issue 1: 30-42, DOI: 10.25236/AJALS.2023.040106

false negatives, respectively.Four image datasets were constructed (Table1) for different experimental purposes using image acquisition system.

3. Results and Discussion

3.1. Effects of model parameters on CDDNet

3.1.1. Effect of batch size

Batch size represents the number of samples used in each step of CNN training process. Batch size determines the direction of gradient decline. For small datasets, the batch size can choose the size of all datasets, but for large datasets, if the batch size is too large, the memory will be insufficient and the training will not continue. Therefore, a suitable batch size can speed up the training speed of the model. To explore its effect on the detection performance of CDDNet, different batch sizes (10, 20, 40, 80 and 160) were used to train and validate the proposed model on dataset 1.

Fig. 10 presents the validation accuracy and training time with different batch sizes for 10 epochs. Validation accuracy is the evaluation of model generalization ability. The validation accuracy first increases and then decreases with the increase of the batch size. The highest accuracy is achieved when the batch size is 40. The training time decreases first and then has no significant change after batch size 40. For the trade-off between model generalization ability and training time, the value of batch size was set to 40.

3.1.2. Effect of the learning rate

As an important parameter in deep learning, learning rate determines whether and when the objective function converges to the local minimum. In order to select an appropriate value for our model, CDDNet was trained and validated on dataset 1 with different learning rates (0.05, 0.01, 0.005, 0.001, 0.0005, 0.0001, 0.00005, and 0.00001), the validation accuracy for different learning ratesas shown in Fig.10, When the learning rate is too small (0.00001) or too large (0.05), the validation accuracy becomes relatively low. The best performance is achieved when the learning rate is 0.0005.

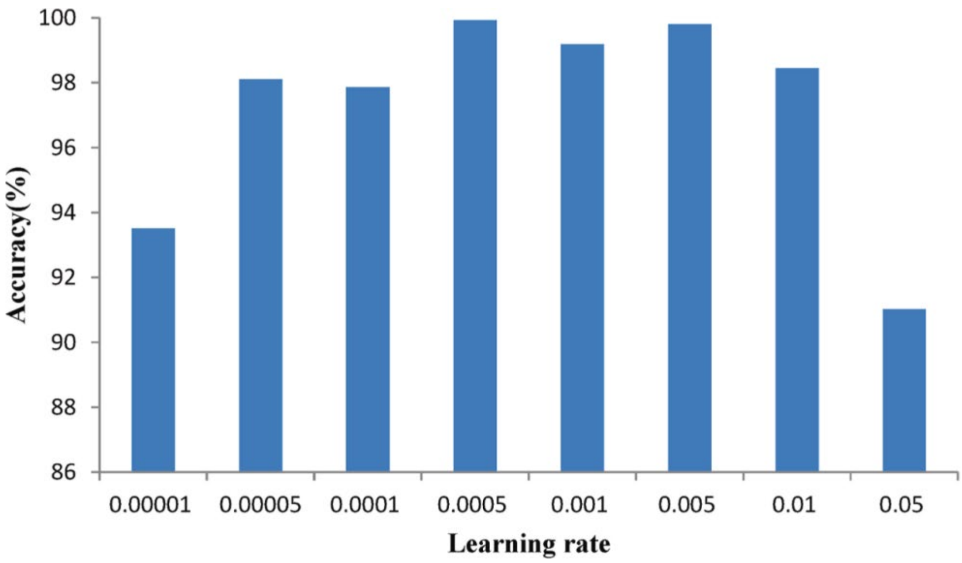


Figure 10: Validation accuracy of different learning rate.

3.2. Performance of CDDNet to detect defective Japanese Yams

In order to evaluate the performance of CDDNet in Yam defect detection[16], experiments were conducted on Dataset 1 (binary classification) and Dataset 2 (multiclass classification), and compared our method with manual methods by workers and some popular deep learning methods (including AlexNet) were compared with ResNet[17], two lightweight models: MobileNet v2 and ShuffleNet.The average values of the statistical parameters defined by Eqs. (10), (11), (12), (13), (14) were used to measure the final performance of the model[18-19]. All deep models were trained for 10 epochs using cross-entropy loss function and Adam optimizer with a learning rate of 0.0005. Experiments were

ISSN 2616-5910 Vol. 4, Issue 1: 30-42, DOI: 10.25236/AJALS.2023.040106

conducted on a computer (CPU: Intel(R) CoreTM i7-6700@3.20GHZ, RAM: 16G; OS: Win11).

3.2.1. Performance evaluation of CDDNet for binary classification

Table 1 lists the average values of the statistical parameters (accuracy, precision, specificity, recall and F1-Score), training time, processing time and model size for binary classification.The accuracy of the manual methods by workers was far lower than that of deep learning methodst[20]. Traditional methods included several distinct steps, such as removing background, detecting region of interest and designing features for each kind of defect, which were very complex and had poor flexibility. However, deep learning can automatically extract features of Japanese Yam defect, despite of little knowledge of the defect[21]. The deep learning model works well because it has a nearly perfect understanding of normal yams, which helps significantly reduce errors and has greater advantages than traditional methods.

Accuracy, precision, specificity, sensitivity and F1-Score of CDDNet were 98.94%, 99.02%, 98.93%, 98.76% and 98.87%, respectively, which were only slightly lower than ResNet (98.98%, 100%, 98.93%, 98.71% and 98.92%). However, CDDNet was superior to ResNet in training time, processing time and model size; its training time was only one-eighteenth of ResNet, and the model size was only one-thirtieth of ResNet. Compared with the original ShuffleNet, the model size of CDDNet was reduced, but it still kept good detection accuracy. Therefore, It is very necessary to select a suitable model and network structure for Japanese Yam defect detection. In general, CDDNet achieved excellent performance both in detection accuracy and model complexity.

Time efficiency is an important aspect to be considered in Japanese Yam grading. As can be seen from Table 2, CDDNet could make it possible to meet the requirements of real-time detection of Japanese Yam defects.

Table 2: Experimental results of the defect detection performance of binary classification.

Methods	AlexNet	ResNet50	MobileNetv2	ShuffleNet	CDDNet
Accuracy(%)	97.98±0.60	98.98±0.90	98.06±0.19	98.89±0.27	98.94±0.11
Recall(%)	98.10±0.76	100	100	99.08±0.06	99.02±0.10
Specificity(%)	97.92±0.97	98.93±0.11	97.56±0.29	98.80±0.32	98.93±0.19
Precision(%)	96.64±1.98	98.71±0.32	95.88±0.59	98.45±0.57	98.76±0.32
F1-Score(%)	97.35±0.91	98.92±0.13	97.47±0.30	98.76±0.33	98.87±0.17
Training time(min)	243.2	2033.2	245.7	114.0	107.5
Processing time(ms)	54.4	84.7	48.7	26.4	**25.0**
Model size(MB)	213.5	85.9	7.8	2.7	2.1

3.2.2. Performance evaluation of CDDNet for multiclass classification

Table 3 illustrates the comparison results among our method and some prevalent deep learning methods in multi-classification task. In this experiment, ResNet model achieved the highest detection accuracy of 94.19%, and the detection accuracy of CDDNet was 92.92%, but the training time of ResNet was much longer than CDDNet. In the tradeoff between time and accuracy, its model demonstrated an overall advantage over other models in multiclass classification task.

Table 3: Experimental results of the multiclass detection performance.

Methods	Handcrafted	AlexNet	ResNet50	MobileNetv2	ShuffleNet	CDDNet
Normal(%)	91.12	95.71±3.62	97.03±2.16	98.09±0.88	97.98±0.96	97.45±0.95
Curve(%)	93.84	90.31±9.80	95.29±1.76	91.59±2.91	92.08±4.80	96.93±2.36
Fork root(%)	95.43	93.17±3.41	90.59±8.41	90.81±8.30	95.66±2.88	91.05±7.22
Fracture(%)	81.09	84.25±11.14	88.43±3.10	80.89±6.91	83.04±7.84	76.87±6.94
Total(%)	90.92	91.19±4.10	**94**.19±0.60	91.56±1.80	92.54±1.46	92.92±1.60
Training time(min)	-	55	446	91	44	**43**
Processing time(ms)	-	54.4	84.7	48.7	26.4	**25.0**

There are significant differences in the detection accuracy of different types of defects. Generally speaking, the accuracy of ordinary Yam and curved Yam is higher; However, the detection accuracy of fork roots and fractures is relatively low, especially for fractures. There may be two reasons:

(1)Sample imbalance problem

There are significant differences in sample size among different categories; There are 138 curve samples, while there are only 29 fibrous root samples. The imbalanced sample leads to serious bias in the classification model. Some strategies can be used to address sample imbalance issues, such as expanding the dataset, oversampling and undersampling, and manually generating data samples. In the future, I will collect more Yam samples or use data augmentation techniques to amplify image samples to further improve the performance of CDDNet.

(2)Variety of defect appearance

There are many factors that cause other surface appearance defects, such as mechanical damage, decay, bruises, wormholes, etc., resulting in great differences in appearance. Due to the large variance of phenotype, it is difficult for CDDNet to extract effective features, and one possible solution is to recognize each possible situation as a separate category.

3.2.3. Comparison with previous studies

Comparing CDDNet with other studies is challenging due to the lack of public datasets to serve as benchmarks in the field. Published studies on defect detection in agricultural products often use their own datasets, making it difficult to compare the results of this paper with those of related work. In order to make a fair comparison, for all deep learning models, this paper uses the same model parameters to train all deep learning models and uses pre-trained AlexNet for experiments instead of directly comparing with experimental results. Therefore, in this reserach, we compared the proposed method with our previous work and some prevalent deep learning methods: AlexNet, ResNet and ShuffleNet on our own dataset. Although there was no comprehensive comparison, considering the limited research studies using computer vision for Japanese Yam defect detection, our research had some advantages.

The proposed CDDNet can perform defect detection well without manually defined features or any other preprocessing[22], so it can be applied to other agricultural products if sufficient dataset is available[23]. Therefore, our model can also be applied to defect detection of other root and tuber agricultural products and has good applicability.

3.3. Evaluation of Japanese Yam grading method

3.3.1. Evaluation of the MBR fitting method

Figure 11: Confusion matrix of (a) manual method by workers and (b) proposed method.

In this experiment, 419 Japanese Yams of four size grades (S, M, L, 2L) were selected as experimental samples on a farm in Obihiro, Hokkaido. Then, the computer vision system is used to obtain Yam images and the MBR fitting method is used for grading. The grading results were verified by measuring the actual length of Japanese Yam with a ruler, as shown in Fig.11.

Due to the strong subjectivity and poor stability of manual methods, the manual experimental results are only responsible for this experiment. Even for the same person, the results of repeated experiments may be different. Experimental results showed that overall accuracy of our method achieved was 92.8%, which was much higher than that of manual method (83.1%).The grading accuracy of manual method varied greatly among different specifications. The accuracy of S and M were only 78.4% and 69.9%, respectively, however, accuracy of L and 2L were both above 90% (90.7% and 95.8%). The accuracy of our method for S, M, L and 2L were 91.7%, 89.4%, 94.4%, and 96.3%, respectively, which demonstrated that our method had has higher reliability and accuracy than manual method.

To my knowledge, there is currently no relevant work on the fitting and grading of Japanese Yam size, so in this article, I will only compare my method with manual methods. Obviously, manual methods have drawbacks such as low accuracy and instability. Our method has advantages in classification accuracy and reliability, greatly improving the accuracy and efficiency of Yam classification.

3.4. Practicability of the proposed approach

A defect detection and image recognition system for Japanese Yam grading was constructed based on the actual size and grade specifications of Japanese Yam processing. In this system, CDDNet is first used to screen out defective Yams (curve, fork root, fracture, flat) and then a grading method is used to classify normal Japanese Yams into different specifications and grades. Compared with previous work, this system is more in line with the actual process and is easy to apply in practical situations.The proposed CDDNet can detect most of defects that affect the appearance quality of Japanese Yam.

At present, relevant research mainly focuses on surface defect detection before harvesting of root crops, while there are few reports on yam grading based on computer vision. This method is simple and feasible, providing an effective method for yam grading. There are some grading standards for the grade and specifications of yams, such as shape uniformity, surface smoothness, color uniformity, and maturity. But most of these standards are qualitative and difficult to quantify. In the future, we will further study the determination of quality grading standards for yams, taking into account more factors such as flatness, color uniformity, size, maturity, etc.

4. Conclusion

In summary, the combination of deep learning and computer vision for Japanese Yam root defect detection and automatic size grading has important research significance and practical application value. By overcoming the limitations of worker manual methods and using computer vision deep learning technology, accurate detection and automatic grading of Japanese Yam defects can be achieved, providing strong support for harvest yield estimation of Japanese Yam and other root crops. This article elaborates on the entire process of quality inspection from aspects such as image acquisition system, quality inspection methods, and grading standards. Based on ShuffleNet and transfer learning, a lightweight network (CDDNet) was developed to detect Yam defects, and Yam size grading method based on MBR fitting and convex polygon approximation was proposed. The experimental results show that the detection accuracy of the proposed CDDNet for binary and multi class classification is 98.94% and 92.92%, respectively. The classification accuracy of MBR fitting and convex polygon approximation is 92.8% and 95.1%, respectively. This can meet the requirements of this study and prove the feasibility of deep learning Japanese Yam defect detection, providing a practical method for defect detection and size grading of root crops such as Japanese Yam.

References

[1] Da Costa, A. Z., Figueroa, H. E. H., & Fracarolli, J. A. Computer vision based detection of external defects on tomatoes using deep learning. Biosystems Engineering.(2020). https://doi.org/ 10.1016/j.biosystemseng.2019.12.003
[2] Rong J C, Wang P B, Yang Q, Huang F. A field-tested harvesting robot for oyster mushroom in

ISSN 2616-5910 Vol. 4, Issue 1: 30-42, DOI: 10.25236/AJALS.2023.040106

greenhouse. Agronomy, (2021): 1210. doi: 10.3390/agronomy11061210.

[3] Zhang F, Chen Z J, Wang Y F, Bao R F, Chen X G, Fu S L, et al. Research on flexible end-effectors with humanoid grasp function for small spherical fruit picking. Agriculture, (2023): 123. doi: 10.3390/agriculture13010123.

[4] Nasiri, A., Omid, M., & Taheri-Garavand, A. An automatic sorting system for unwashed eggs using deep learning. Journal of Food Engineering, (2020). https://doi.org/ 10.1016/j.jfoodeng.2020.110036

[5] Kala, J. R., Viriri, S., & Tapamo, J. R. An approximation based algorithm for minimum bounding rectangle computation. In IEEE international conference on adaptive science and technology, ICAST, 2015-Janua. (2015). https://doi.org/10.1109/

[6] Rong J C, Wang P B, Wang T J, Ling H, Yuan T. Fruit pose recognition and directional orderly grasping strategies for tomato harvesting robots. Computers and Electronics in Agriculture, (2022); 202: 107430. doi: 10.1016/j.compag.2022.107430.

[7] Afroza A, Ambreen N, Baseerat A; Nigeena N, Ahmad S P, Azrah I S, Amreena S; Insha J, Majid R. Evaluation of Cherry Tomato (Solanum lycopersicum L. var. cerasiforme) Genotypes for Yield and Quality Traits. Journal of Community Mobilization and Sustainable Development, (2021): 72-76.

[8] Vasseghian, Y., Bahadori, A., Khataee, A., Dragoi, E. N., & Moradi, M. Modeling the interfacial tension of water-based binary and ternary systems at high pressures using a neuro-evolutive technique. ACS Omega, (2020). https://doi.org/10.1021/acsomega.9b03518

[9] Yamamoto K, Guo W, Ninomiya S S. Node detection and internode length estimation of tomato seedlings based on image analysis and machine learning. Sensors, (2016): 1044.

[10] Wu, J. S., Zhang, B., & Gao, Y. L. An effective flame segmentation method based on OHTA color space. Applied Mechanics andMaterials,485,7–11.(2012).https://doi.org/10. 4028/www. scientific.net/ AMR. 485.7.

[11] Wang Z L. Underwood J. Walsh KB. Machine vision assessment of mango orchard flowering. Computers and Electronics in Agriculture, (2018); 151: 501–511.

[12] Wu J G, Zhang B H, Zhou J, Xiong Y J, Gu B X, Yang X L. Automatic Recognition of Ripening Tomatoes by Combining Multi-Feature Fusion with a Bi-Layer Classification Strategy for Harvesting Robots. Sensors, (2019): 612- 612.

[13] Xiong J T, Lin R, Liu Z, He Z L, Tang L Y, Yang Z G Zou X J. The recognition of litchi clusters and the calculation of picking point in a nocturnal natural environment. Biosystems Engineering, (2018); 166: 44-57.

[14] Bechar A, Vigneault C. Agricultural robots for field operations: Concepts and components. Biosystems Engineering, (2016); 149: 94-111.

[15] Silwal A, Davidson J R, Karkee M, Mo C K, Zhang Q, Lewis K. Design, integration, and field evaluation of a robotic apple harvester. Journal of Field Robotics, (2017); 34(6: 1140-1159.

[16] Silwal A, Karkee M, Zhang Q. A hierarchical approach to apple identification for robotic harvesting. Transactions of the ASABE, (2016); 59(5): 1079-1086.

[17] Guo Y M, Liu Y, Oerlemans A, Lao S Y, Lew M S. Deep learning for visual understanding: A review. Neurocomputing, (2016); 187: 27-48.

[18] He K M, Zhang X Y, Ren S Q, Sun J. Deep residual learning for image recognition. (2016)IEEE Conference on Computer Vision and Pattern Recognition (CVPR), (2016); 770-778.

[19] Xu Z F, Jia R S, Liu Y B, Zhao C Y, Sun H M. Fast Method of Detecting Tomatoes in a Complex Scene for Picking Robots. IEEE Access, (2020); 8: 55289 - 55299.

[20] Ren S Q, He K M, Girshick R, Sun J. Faster R-CNN: Towards real-time object detection with region proposal networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, (2017); 39(6): 1137-1149.

[21] Thuyet, D. Q., Kobayashi, Y., & Matsuo, M. A robot system equipped with deep convolutional neural network for autonomous grading and sorting of root-trimmed garlics. Computers and Electronics in Agriculture,(2020). https://doi.org/10.1016/j. compag.2020.105727.

[22] Wang C Y, Liao H M, Wu Y H, Chen P Y, Hsieh J W, Yeh I H. CSPNet: A new backbone that can enhance learning capability of CNN. In: (2020) IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), (2020); pp.1571-1580.

[23] Woo S, Park J C, Lee J, Lweon I. CBAM: Convolutional Block Attention Module. In: Computer Vision - ECCV, (2018); pp. 3-19.