

## Application of Image Big Data in the Ecological Agricultural Product Supply Chain

Ruixue Zhang 

School of Economics and Management, Dalian Minzu University, Dalian 116600, China

Corresponding Author Email: [20101021@dlmu.edu.cn](mailto:20101021@dlmu.edu.cn)

Copyright: ©2025 The author. This article is published by IIETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/ts.420128>

### ABSTRACT

**Received:** 3 June 2024

**Revised:** 28 November 2024

**Accepted:** 5 December 2024

**Available online:** 28 February 2025

#### **Keywords:**

*ecological agricultural products, image big data, deep learning, surface disparity estimation, convolutional neural network (CNN), stereo matching, supply chain management*

With the rapid development of ecological agriculture and the increasing demand for agricultural product supply chain management, effectively monitoring the quality and circulation status of agricultural products has become an urgent issue. Image big data technologies, particularly advancements in deep learning and computer vision, offer innovative solutions for surface quality detection, analysis, and traceability of agricultural products. By precisely estimating surface disparity and analyzing quality, these technologies not only improve the efficiency of quality control but also enhance supply chain transparency, ensuring the stability of product quality. However, existing image analysis methods face significant limitations when dealing with minor surface defects, lighting variations, and complex textures. Traditional image processing techniques are less effective in these areas, and the application of deep learning is still in the exploratory phase. To address these issues, this study proposes a deep learning-based method for surface disparity estimation of agricultural products and designs three innovative models: 1) a Convolutional Neural Network (CNN) for surface disparity estimation of agricultural products, 2) an end-to-end deep learning stereo matching model for surface disparity estimation, and 3) a deep learning pyramid stereo matching network model for surface disparity estimation of agricultural products. These models aim to overcome the shortcomings of current methods and enhance the precision and stability of agricultural product image analysis, providing more efficient and intelligent technical means for quality control in the agricultural product supply chain.

## 1. INTRODUCTION

With the transformation of global agricultural production models and the rise of ecological agriculture concepts, ecological agricultural products, as an environmentally friendly and healthy food choice, have gradually become an important part of the market [1-3]. At the same time, the complexity of agricultural product supply chains and the high demands for quality control have become increasingly prominent. How to efficiently and accurately identify and trace the quality and circulation of agricultural products has become an urgent issue in modern agriculture [4-7]. Image big data technologies, especially advancements in computer vision and deep learning, provide new ideas for quality detection, monitoring, and traceability of ecological agricultural products [8, 9]. Through precise analysis of the surface quality of agricultural products, it is possible to effectively improve the transparency and trustworthiness of agricultural products in the supply chain, thus promoting the healthy development of the ecological agricultural product market.

The research on the application of image big data in the ecological agricultural product supply chain has significant academic and practical value. First, image-based agricultural product quality evaluation systems can greatly improve detection efficiency and accuracy, reduce labor costs, and promote agricultural modernization [10, 11]. Second, image

big data not only helps enhance the transparency of supply chain management but also effectively prevents quality problems during transportation and storage, ensuring the consistency and stability of product quality. Finally, the integration of deep learning and computer vision technologies, especially their innovative applications in agricultural product surface defects, quality assessment, and other fields, has high research value and can provide theoretical support and technical guarantees for the intelligent transformation of the agricultural sector [12-16].

Although some progress has been made in related fields in recent years, existing agricultural product image analysis methods still have certain limitations. Currently, most methods rely mainly on traditional image processing technologies, which, although capable of performing certain recognition tasks for agricultural products' appearance, still show obvious shortcomings when dealing with complex backgrounds, surface disparity, and varying scales [17-20]. Especially in identifying minor surface defects, lighting changes, and texture complexity in agricultural products, traditional methods often fail to provide accurate and stable results [21, 22]. The application of deep learning in this field is still in its early stages, lacking targeted end-to-end models and refined deep learning architectures, which limits the further development of agricultural product image analysis.

This paper mainly addresses the technical bottlenecks in the

current agricultural product surface disparity estimation and proposes a new image analysis method based on deep learning. Specifically, this study focuses on three core areas: First, a CNN for surface disparity estimation of agricultural products is proposed, which can effectively extract detailed features from the surfaces of agricultural products; second, an end-to-end deep learning stereo matching model is designed to further improve the accuracy of detecting surface defects in agricultural products; and third, a pyramid stereo matching network model is proposed, which enhances the robustness and adaptability of the model under different resolutions by integrating multi-scale features. Through these innovative methods, this paper not only provides a new technical pathway for agricultural product quality detection but also lays the theoretical foundation and technical support for the intelligent management of ecological agricultural product supply chains.

## 2. CNN FOR SURFACE DISPARITY ESTIMATION OF AGRICULTURAL PRODUCTS

With the rapid development of ecological agriculture, the issue of agricultural product quality detection has become increasingly complex and important. Ecological agricultural products, due to their natural and pollution-free characteristics, usually have high market value and consumer recognition. However, the quality standards for ecological agricultural products often lack unified specifications, and small defects, blemishes, or damages on the surface of agricultural products are difficult to efficiently and accurately assess using traditional manual detection methods. As the supply chain becomes increasingly globalized, agricultural products are easily affected by external factors during transportation, storage, and sales, leading to inconsistencies in product appearance. Therefore, using efficient technical means for quality monitoring, especially through image big data technologies for precise surface analysis of agricultural products, can allow real-time and comprehensive understanding of the quality status of agricultural products and ensure that each link meets quality standards. This image big data-based quality detection method can effectively improve detection efficiency, reduce human errors, and ensure the stability and consistency of product quality, thus enhancing the competitiveness of ecological agricultural products in the market.

Surface disparity estimation, as an important computer vision technology, can precisely measure the details of agricultural product surfaces, capturing surface defects, damages, and other quality issues. Methods based on disparity estimation can effectively simulate and reconstruct the three-dimensional structure of agricultural product surfaces, which is of great significance in determining their surface quality and defects. Compared with traditional image processing methods, deep learning-based surface disparity estimation methods can maintain high detection accuracy even under complex backgrounds and changing lighting conditions. Deep learning technology can automatically learn surface features of agricultural products in different states and types from a large number of agricultural product images, further improving the adaptability and robustness of the detection model.

### 2.1 Convolutional layer

CNNs have powerful feature extraction capabilities in the field of computer vision, allowing them to automatically learn

the complex textures, color changes, and depth information of agricultural product surfaces through multiple convolutional operations, capturing subtle morphological features. Particularly for ecological agricultural products, their surfaces may exhibit natural inhomogeneity, such as the skin texture, roughness, and small-scale concave-convex structures of fruits and vegetables. CNNs can gradually learn surface information at different scales through hierarchical feature extraction mechanisms, and, in combination with disparity estimation algorithms, construct an accurate three-dimensional surface model of agricultural products. Figure 1 shows the schematic diagram of the CNN architecture.

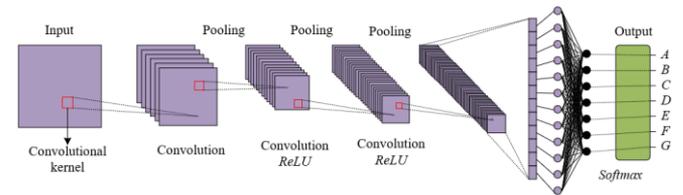


Figure 1. Schematic diagram of CNN architecture

For the task of surface disparity estimation of agricultural products, the CNN architecture needs to possess key characteristics such as multi-channel feature extraction, hierarchical depth matching, and adaptive feature fusion. Multi-channel feature extraction ensures that the model can capture depth information of agricultural product surfaces from different angles and under varying lighting conditions, avoiding misjudgments caused by lighting changes in traditional 2D image analysis. Hierarchical depth matching, achieved by constructing multi-scale convolutional layers, allows the model to capture large-scale surface curvature while retaining fine local disparity information, enabling accurate identification of minor surface defects of agricultural products. Adaptive features allow the model to dynamically adjust the feature information at different layers, enhancing the model's adaptability to complex surface forms, so that the detection system can be applied to ecological agricultural products of different types and production environments. Specifically, suppose the calculated pixel at position  $(u, k)$  in the  $f$ -th output feature layer of the  $m$ -th convolutional layer is represented by  $T_f^m$ , with  $l$  and  $v$  being the size of the convolution kernel and  $y_f^m$  being the bias at the current layer. The convolution operation output can be expressed as:

$$T_f^l(u, k, j) = d \left( \sum_{l,v} U^{m-1}(u+l, k+v) \times J_{jf}^m(l, v) + y_f^m \right) \quad (1)$$

### 2.2 Pooling layer

For the task of surface disparity estimation of agricultural products, the main goal of the pooling layer is to improve the model's sensitivity to important features by removing unnecessary details, reducing noise and redundant information in the image. Since the surface quality differences of agricultural products may be caused by small texture changes or slight surface flaws, the pooling layer helps the model ignore subtle surface variations and focus on more representative regional features, thus avoiding overfitting and improving the model's adaptability to image data in complex environments. Suppose the resampling factor is represented by  $\theta_f^m$ , the downsampling factor is represented by  $SU(\cdot)$ , and the

bias is represented by  $y_f^m$ . The output of the pooling layer can be expressed as:

$$o_f^m(u, k, j) = d(\theta_f^m SUT_f^m(l, v, j) + y_f^m) \quad (2)$$

The pooling layer further enhances the performance of the surface disparity estimation model for agricultural products by using either max pooling or average pooling strategies. Max pooling retains the most significant features in the image by selecting the maximum value within the pooling window, which is particularly effective for capturing prominent defects on the agricultural product surface, such as cracks, stains, or damage. On the other hand, average pooling calculates the average value within the pooling region, integrating more regional information to improve the model's stability and robustness to local features. In surface disparity estimation for agricultural products, this characteristic of the pooling layer is particularly important because quality issues on the surface of agricultural products often involve changes in local areas. Through pooling, the network can better adapt to variations under different scales and surface conditions, thereby enhancing its overall ability to assess the quality of agricultural products.

### 2.3 Fully connected layer

The surface features extracted progressively through the convolutional and pooling layers are mapped into a higher-dimensional space when they enter the fully connected layer, allowing the network to capture complex feature interactions. For the task of agricultural product surface disparity estimation, the role of the fully connected layer is not just to perform a simple weighted sum of features but to integrate features from different scales and regions through connections with multiple neurons, extracting the correlated features between surface depth information and quality.

To improve the model's stability and generalization performance, the fully connected layer typically employs Dropout strategies and regularization methods to prevent overfitting when handling complex agricultural product surface data. Since surface disparity estimation of agricultural products involves dealing with various environmental lighting and surface texture changes, Dropout helps to effectively prevent the network from relying too heavily on specific neurons' feature representations during the training process by randomly discarding a certain proportion of hidden layer neurons. This enhances the model's robustness to different agricultural product samples.

### 2.4 Regression layer

The regression layer, as the final layer, is primarily responsible for predicting the surface depth information of agricultural products based on the features extracted and integrated by the preceding layers, and outputs numerical values that can be used for quality assessment. Since the goal of surface disparity estimation for agricultural products is to acquire precise surface depth information based on image data for identifying surface defects, the regression layer must be capable of handling continuous output problems. By continuously adjusting the model parameters, the regression layer learns the patterns of surface depth distribution of agricultural products, improving the sensitivity to small surface undulations, depressions, or protrusions, thus

providing more reliable numerical data for subsequent quality assessment.

In specific agricultural product quality grading or defect classification tasks, the regression layer can also combine activation functions such as Softmax or Sigmoid for discretization, transforming the continuous depth estimation results into discrete category labels. For example, the surface quality of agricultural products can be categorized into multiple levels based on disparity information, such as "Normal," "Minor Defect," "Severe Defect," etc. In this case, a Softmax layer can be used to compute the probabilities of different categories and select the category with the highest probability as the final predicted result. Softmax ensures that the sum of the probabilities of all categories is 1 and provides clear classification boundaries, making the results of agricultural product quality detection more intuitive and easier to interpret. Suppose the number of labels is represented by  $j$ , and the Softmax function can be expressed as:

$$d_u(b) = \frac{\exp(b_u)}{\sum_j \exp(b_u)} \quad (3)$$

Given the input  $a_u$  and parameter  $q$ , the probability of output label  $k(b_u=k)$  can be calculated as follows:

$$o(b_u = k | a_u; q) = \frac{\exp(q_k \bullet a_u)}{\sum_{v=1}^j \exp(q_k \bullet a_u)} \quad (4)$$

### 2.5 Gradient descent and backpropagation

Since the quality detection of agricultural products involves complex image features such as texture changes, cracks, stains, etc., optimizing the loss function using gradient descent can effectively minimize the difference between the predicted values and the actual values. In this process, the loss function is used to measure the deviation between the model's predicted value and the actual value, i.e., the disparity between the depth estimation result and the actual surface depth. The goal of the network is to reduce this gap through gradient descent. Specifically, gradient descent updates the trainable parameters of the network in the opposite direction of the loss function gradient, enabling the model to gradually adjust its weights and biases to more accurately predict the surface depth information of agricultural products. The constructed  $M_1$  loss function is represented as:

$$M_1(\hat{b}, b) = \sum_{u=0}^l |b^{(u)} - \hat{b}^{(u)}| \quad (5)$$

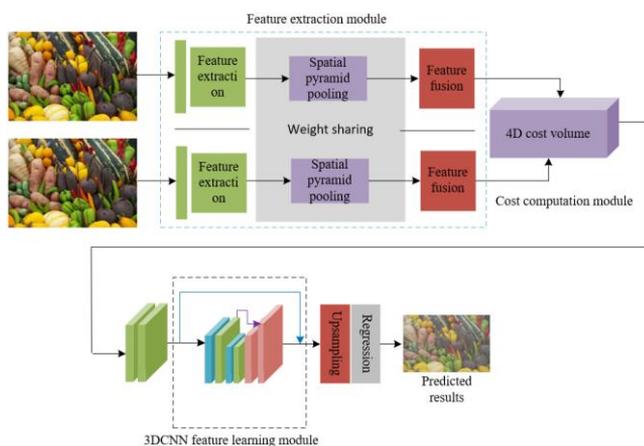
Backpropagation plays a crucial role in the gradient descent process. It uses the chain rule to compute the gradient of each neuron layer by layer and propagates the loss backward from the output layer to the input layer. This process helps compute the gradient of each layer's parameters so that the optimizer can adjust the weights and biases in the network, reducing the output error. In the task of surface disparity estimation for agricultural products, backpropagation efficiently updates the parameters of various network layers such as convolutional layers, pooling layers, and fully connected layers, enabling the model to learn from errors and improve with each training iteration, gradually enhancing sensitivity to small surface defects of agricultural products. The total discrepancy

calculated by backpropagation can be expressed as:

$$E_{\tau_0} = \sum \frac{1}{2} (\text{Target value} - \text{Calculated output value})^2 \quad (6)$$

### 3. END-TO-END DEEP LEARNING STEREO MATCHING MODEL FOR AGRICULTURAL PRODUCT SURFACE DISPARITY ESTIMATION

Ecological agricultural products often exhibit differences in appearance, such as lighting, shadows, stains, etc., which can affect the extraction of surface features. Traditional methods often fail to handle these complex situations. The end-to-end deep learning stereo matching model can be trained using large-scale image data, providing significant advantages in surface disparity estimation for agricultural products. Moreover, the stereo matching model can simultaneously analyze the structural differences of agricultural product surfaces from two perspectives, providing more accurate depth estimation results. Figure 2 shows the grid structure of the constructed end-to-end deep learning stereo matching model.



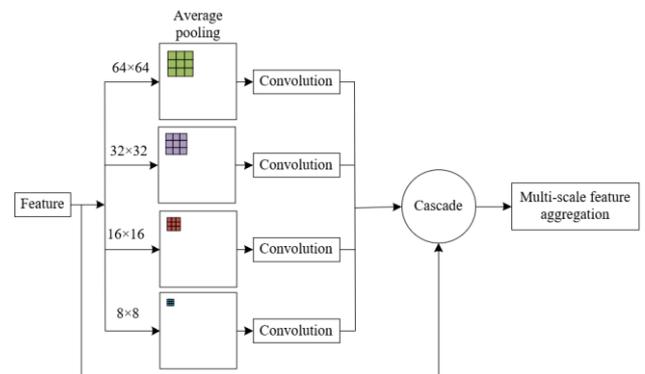
**Figure 2.** Grid structure of the constructed end-to-end deep learning stereo matching model

#### 3.1 Feature extraction

In the end-to-end deep learning stereo matching model for agricultural product surface disparity estimation, the principle of feature extraction mainly relies on a twin network structure with weight sharing. This structure effectively extracts the depth information of agricultural product surfaces through the feature sharing channels of the left and right views. The principle is shown in Figure 3. Specifically, the input consists of the observed left and right images of the agricultural product surface. To better capture the disparity information, the network uses a CNN for feature extraction. Since the surfaces of agricultural products exhibit various complex geometric variations, textures, and fine defects, traditional convolution operations may not be able to capture these local details well. Therefore, the feature extraction process adopts dilated convolutions and feature pyramid networks to expand the receptive field of the convolutional kernels, allowing the network to focus on both large-scale background information and small-scale surface details.

To further enhance the robustness of feature extraction and reduce the interference of background noise, spatial pyramid pooling is applied during the feature extraction process.

Spatial pyramid pooling extracts features at different spatial scales through pooling operations of varying sizes, which helps capture local variations and overall structures on the agricultural product surface. In addition, although using lightweight structures (such as depthwise separable convolutions) helps improve computational efficiency, it may also lead to a loss in feature resolution. Therefore, when designing the network, a balance between lightweight architecture and feature extraction capability must be considered to ensure the network provides sufficient resolution to accurately capture the fine defects and imperfections on the agricultural product surface while maintaining efficiency. This, in turn, improves the final disparity estimation accuracy.



**Figure 3.** Principle of feature extraction in the constructed end-to-end deep learning stereo matching model

#### 3.2 Cost volume construction

In traditional stereo matching, the cost volume is usually constructed by matching features from the left and right images within different disparity ranges and calculating the cost value for each pixel. For agricultural product quality detection, since the surface has various subtle differences, efficiently handling these differences and generating accurate depth maps is especially important. In deep learning models, the construction of the cost volume typically involves extracting high-dimensional features from the images and then misaligning these features along the disparity dimension, ultimately forming a high-dimensional cost volume. This type of cost volume can preserve rich contextual and geometric information at each pixel location, thus improving sensitivity to complex surface features. Specifically, let the dot product be represented by  $IN$ , the extracted left image features be represented by  $d_m$ , the extracted right image features be represented by  $d_e$ , the number of feature channels be represented by  $V_z$ , the pixel location in pixel coordinates be represented by  $(a, b)$ , and the corresponding disparity value for the pixel location be represented by  $f$ . The final cost volume can be expressed as a  $G \times Q \times F$  cost volume, where  $G$  represents the length of the cost volume,  $Q$  represents the width of the cost volume, and  $F$  represents the maximum disparity value. The cost volume based on layers is represented by the following formula:

$$Z_{co}(a, b, f) = \frac{1}{V_z} IN \langle d_m(a, b), d_e(a - f, b) \rangle \quad (7)$$

Let the cascading operation be represented by  $CC$ , and the cost volume by  $DN$ . The 3D cost volume expression is given by the following formula:

$$Z_{co}(a,b,f,\bullet) = CC \langle d_m(a,b), d_e(a-f,b) \rangle \quad (8)$$

In practical applications, the deep learning stereo matching method using a 4D cost volume is more suitable for agricultural product surface disparity estimation tasks. In this method, the cost volume not only considers the spatial features of the left and right images but also misaligns and concatenates these features along the disparity dimension, thereby obtaining more accurate depth estimation. The 4D cost volume can effectively preserve more geometric information, especially when the agricultural product surface has irregular undulations, texture changes, and occlusions. The 4D cost volume can provide more contextual information about surface details. For example, small protrusions or depressions on the agricultural product surface can be accurately matched using this high-dimensional cost volume, improving the accuracy of disparity regression. Through end-to-end training, the model can automatically learn from large-scale agricultural product image data and gradually optimize the construction of the cost volume to better identify different types of defects in quality detection.

### 3.3 Disparity regression

In the constructed model, disparity regression is set to accurately estimate the depth information of input image data and generate detailed disparity maps. The disparity map reflects the depth information of each pixel in the image, as the surface of agricultural products often exhibits subtle undulations and variations, which directly impact quality assessment. There are two common methods for disparity regression: *argmin* operation and *softargmin* operation. In the traditional *argmin* operation, the model selects the disparity with the minimum cost as the final output, using a "winner-takes-all" strategy. However, this method may not handle small variations on complex surfaces well, as it only returns the disparity value with the smallest cost, ignoring other potential disparity candidates. This is not precise enough for estimating the fine details of agricultural product surfaces. Let the predicted disparity be represented by  $\hat{f}$ , the cost volume be represented by  $Z(a,b,f)$ , and the disparity by  $f$ , then the expression is:

$$\hat{f} = ARGMIN_f Z(a,b,f) \quad (9)$$

In contrast, the *softargmin* operation has been widely used in deep learning stereo matching tasks, especially for agricultural product surface disparity estimation. It first calculates the probability distribution of different disparity values through the *softmax* operation, which is between 0 and 1, representing the likelihood of each disparity value. Then, it uses these probability values to perform a weighted sum of the disparities, resulting in the final disparity output. The advantage of this method is that it not only considers the disparity with the minimum cost but also integrates the possibilities of other candidate disparities, preserving more information about the geometry and texture of the agricultural product surface. The final output disparity value can be obtained by performing the weighted sum of the disparity values and their corresponding probabilities:

$$\tilde{f} = \sum_{f=0}^{f_{MAX}} f \times SOFTMAX(Z(a,b,f)) \quad (10)$$

$$SOFTMAX(a_u) = \frac{e^{a_u}}{\sum_{k=0}^K e^{a_k}} \quad (11)$$

## 4. DEEP LEARNING PYRAMID STEREO MATCHING NETWORK MODEL FOR AGRICULTURAL PRODUCT SURFACE DISPARITY ESTIMATION

End-to-end deep learning stereo matching models are typically optimized directly from raw images to disparity map outputs through end-to-end learning. This type of model has strong global feature learning capabilities and can automatically extract and optimize relevant features from the images. However, when faced with surface details of agricultural products, especially in situations with complex lighting changes, surface textures, and small-scale defects, challenges may arise. This is because end-to-end models tend to optimize performance globally, potentially neglecting certain details, which can lead to insufficient accuracy when processing fine surface defects. On the other hand, pyramid networks provide stronger adaptability in multi-scale feature processing. Particularly when dealing with the diverse defects of agricultural product surfaces, the pyramid structure can refine the localization and estimation of small surface changes through disparity maps at different scales, effectively improving the robustness and accuracy of agricultural product surface quality detection. The loss function defined for the model is as follows:

$$M(f, \hat{f}) = \frac{1}{V} \sum SO_{M_i}(f_u - \hat{f}_u) \quad (12)$$

In the disparity computation process, the pyramid stereo matching network uses 3D CNNs to regularize the 4D cost volume, improving the stability of disparity matching. Unlike fixed 3D convolution layers and hourglass structures, the pyramid stereo matching network can adjust features on the cost volume at different scales, making the model more flexible in handling complex backgrounds, lighting changes, and local occlusions. For example, when detecting diseases on the surface of ecological agricultural products, some subtle lesions may be difficult to identify on a disparity map at a single scale. However, the pyramid model can leverage multi-scale information, making the depth information of the diseased areas clearer. During the disparity regression phase, the pyramid network typically employs the *softargmin* method, calculating the probability distribution of different disparity values through the *softmax* operation and performing a weighted sum to obtain a smoother and more accurate disparity map. The disparity estimation results can be expressed as:

$$\hat{f} = \sum_{f=0}^{f_{MAX}} f \times \delta(-z_f) \quad (13)$$

## 5. EXPERIMENTAL RESULTS AND ANALYSIS

In this study, Figure 4 shows the training performance of four different models under different loss function values and accuracies. First, by observing the changes in the loss function values, the model in this paper and the *PointNet* model exhibit similar decreasing trends in loss function values. Both models

show a gradual decrease in the loss function values during training, eventually reaching relatively low values, around 0.11 to 0.12. This indicates that they optimize well and maintain stability during the training process. In contrast, the *YOLOv9-DSM* model shows a relatively smooth decrease in the loss function, but the improvement slows down in the later stages of training, and its loss value remains relatively high, reaching a minimum of 0.19. The *StereoNet* model shows a slower decrease in loss value and fluctuates significantly throughout the training process. Regarding accuracy, both the *PointNet* and the model in this paper performed excellently in the later stages of training, maintaining a high and stable accuracy around 0.98-0.99, indicating their strong performance in agricultural product surface disparity estimation. The accuracy of the *YOLOv9-DSM* model and the *StereoNet* model fluctuated. Although they could also achieve high accuracy during training, their stability was not as strong as that of the model in this paper and *PointNet*.

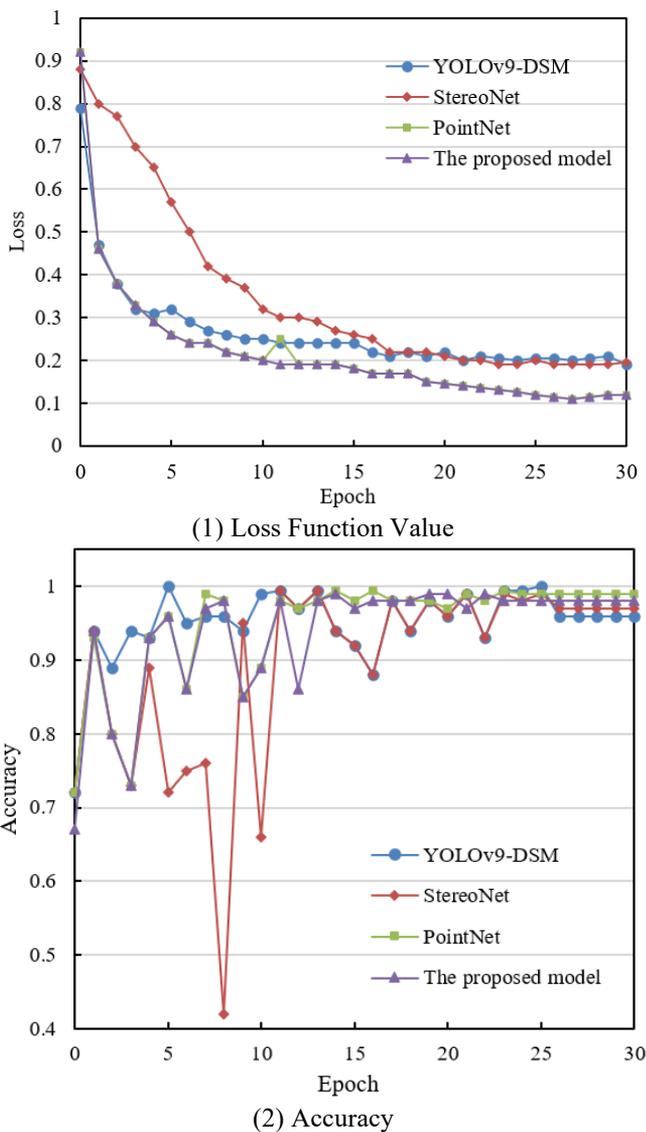


Figure 4. Training performance of different models

In Table 1, the performance differences between the model in this paper and the other three comparison models on the test set are significant. First, the model in this paper has a substantially higher number of layers, reaching 158 layers, far exceeding the 8 layers of *YOLOv9-DSM*, 15 layers of *StereoNet*, and 17 layers of *PointNet*, indicating a more

complex and deeper network structure. Despite this, the model in this paper shows a relatively low test set loss of 0.168, second only to *PointNet* at 0.142, while achieving a test set accuracy of 0.986, close to the highest value of 0.988 for *StereoNet*. Moreover, the number of trainable parameters in this model is similar to *StereoNet*, and is only a quarter of *YOLOv9-DSM*, while its *FLOPs* are significantly lower than the other models, particularly much lower than *PointNet* and *StereoNet*. This indicates that, despite its complex structure, the proposed model is relatively efficient in terms of computational resource consumption.

Table 1. Grid parameters

Grid Type	<i>YOLOv9-DSM</i>	<i>StereoNet</i>	<i>PointNet</i>	The Proposed Model
Number of Layers	8	15	17	158
Test Set Loss	0.178	0.163	0.142	0.168
Test Set Accuracy	0.951	0.988	0.978	0.986
Trainable Parameters Count	$6.2 \times 10^7$	$1.5 \times 10^7$	$1.1 \times 10^7$	$1.5 \times 10^7$
FLOPs	$3.6 \times 10^9$	$8 \times 10^9$	$9.4 \times 10^9$	$1.9 \times 10^9$

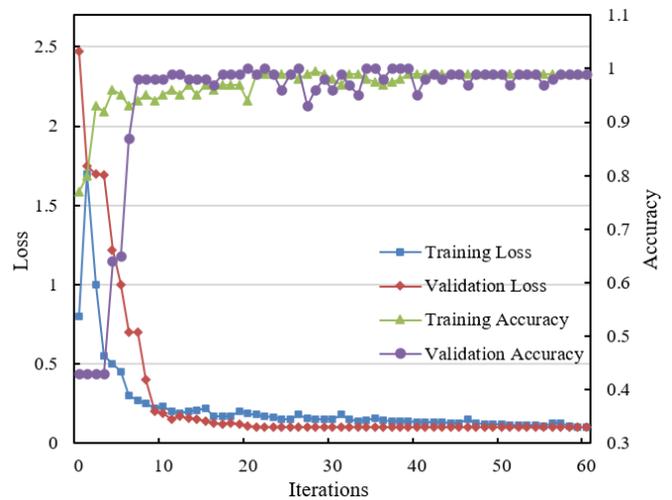


Figure 5. Loss and recognition accuracy of the model in training and validation over iterations

As shown in Figure 5, during the training and validation process, the model in this paper exhibits stable and significant improvement in both training set loss and accuracy. Regarding the change in training set loss, as the number of iterations increases, the loss value gradually decreases, quickly dropping from 0.8 to around 0.14. This indicates that the model is continuously optimizing and gradually converging during training. For the training set accuracy, as training progresses, the accuracy continuously increases from an initial 77% to over 99%, demonstrating the model's good fitting ability on the training set. The performance on the validation set is also stable, with the validation set loss decreasing from 2.47 to near 0.1, and the validation set accuracy increasing from 43% to nearly 100%, indicating that the model's generalization ability on the validation set gradually strengthens and ultimately achieves near-perfect recognition performance. Especially when the validation set accuracy reaches 1.0, this shows the model's high accuracy in agricultural product surface disparity estimation tasks.

In the experiments in this paper, the performance of different models in the agricultural product surface disparity estimation task was compared, validating the effectiveness and

reliability of the proposed method. First, the training and validation results show that the integrated method based on CNNs, deep learning stereo matching models, and pyramid stereo matching networks significantly improves the model's learning and generalization capabilities. The training set loss decreased from an initial 0.8 to around 0.14, and the validation set loss dropped from 2.47 to near 0.1, indicating good convergence during both the training and validation phases. Meanwhile, the training set accuracy increased from 77% to over 99%, and the validation set accuracy rose from 43% to nearly 100%, demonstrating the model's high accuracy and stability in agricultural product surface disparity estimation tasks. Furthermore, the visualization results in Figure 6 show that the model mapped high-dimensional semantic information of the four agricultural product quality states in

the test set, clearly demonstrating the separability of the data samples. This result further verifies that the disparity sample data for different agricultural product quality states is separable and has distinct class features before being input into the classification network for quality state classification. This not only proves the feasibility and reliability of using surface depth information for quality state recognition of agricultural products but also provides insights for further model optimization by analyzing typical error samples. This indicates that the deep learning-based image analysis method proposed in this paper has significant application potential and practical value in agricultural product surface disparity estimation and quality state classification tasks. Figure 7 gives some agricultural product surface disparity estimation examples.

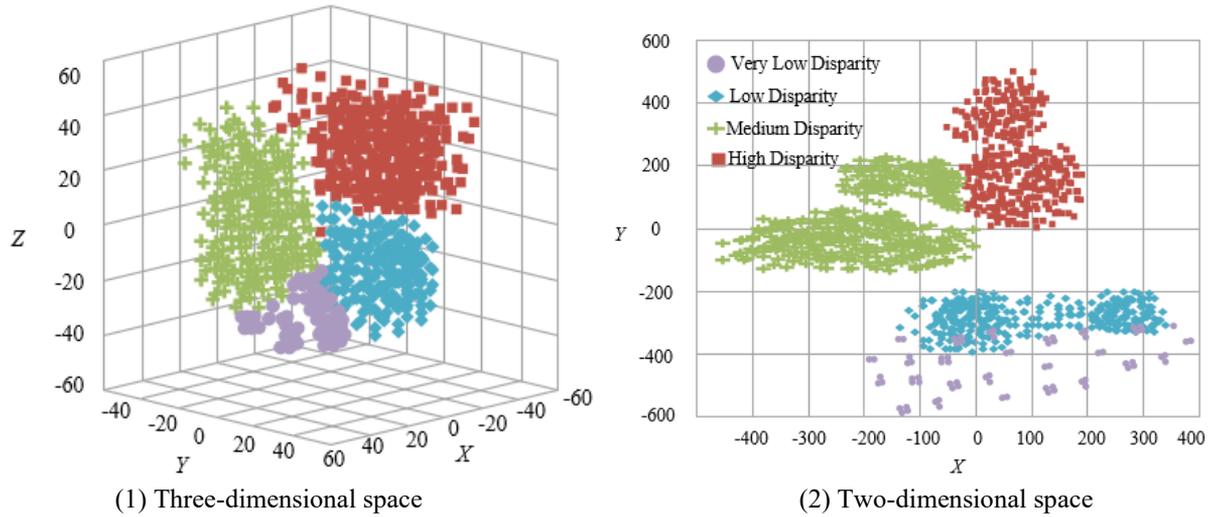


Figure 6. Dimensionality reduction visualization of agricultural product surface disparity estimation

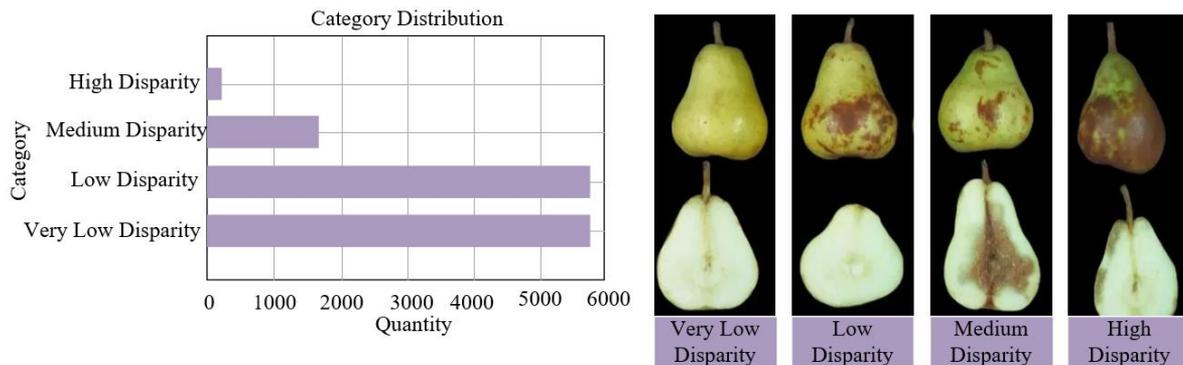


Figure 7. Examples of agricultural product surface disparity estimation

## 6. CONCLUSION

This paper addresses the technical bottlenecks in agricultural product surface disparity estimation by proposing a new image analysis method based on deep learning, aimed at improving the detection accuracy and robustness of agricultural product surface defects. The research mainly includes three aspects: First, a CNN for agricultural product surface disparity estimation is proposed, which efficiently extracts detailed features of the agricultural product surface, providing strong support for subsequent defect detection. Second, an end-to-end deep learning stereo matching model is designed, which further enhances the detection capability of

agricultural product surface defects and optimizes the accuracy of disparity estimation in traditional methods. Finally, a pyramid stereo matching network model is proposed, which, through the fusion of multi-scale features, enhances the model's adaptability and robustness at different resolutions, effectively solving the problem of feature extraction in low-resolution images. Experimental results show that the proposed method demonstrates significant performance improvements on both the training and validation sets, especially in the accuracy of agricultural product surface disparity estimation, reaching near-perfect levels, validating the method's effectiveness and feasibility.

Overall, the research in this paper has important application

value, particularly in the agricultural field for agricultural product quality inspection and automated sorting. Through high-precision surface defect detection and quality state classification, this method is expected to play a key role in improving agricultural product production efficiency, reducing human intervention, and enhancing the accuracy of quality inspection. However, there are certain limitations in this study. First, although the model performs well on the experimental dataset, its generalization ability still needs further validation, especially in diversified and complex real agricultural environments. Second, the computational resources required for model training are relatively large, which may pose some cost pressure for practical applications. Future research directions can focus on the following aspects: First, exploring the model's adaptability and robustness on more diverse datasets to enhance its generalization ability in real production environments. Second, optimizing the model's computational efficiency and reducing reliance on hardware resources to enable application on lower-cost devices. Finally, multimodal fusion with other types of sensor data can be considered to further improve defect detection and quality assessment accuracy and reliability.

## ACKNOWLEDGEMENTS

This paper was supported by Research project of Economic and Social Development of Liaoning of China (Project No.: 2025lslybwzzkt-055).

## REFERENCES

- [1] Hu, H., Wang, C., Chen, M. (2024). The influence of a regional public brand on consumers' purchase intention and behavior toward eco-agricultural products: A Chinese national park case. *Sustainability*, 16(21): 9253. <https://doi.org/10.3390/su16219253>
- [2] Xiang, R., Xiang, C.S., Li, Y. (2019). Influence of eco-agricultural tourism relationship marketing on customer loyalty based on product brand trust. *Journal of Environmental Protection and Ecology*, 20: 41-48. <https://doi.org/10.5555/20230203761>
- [3] Cheng, W.X., Li, J.N., Chen, H. (2016). Behavior of antibiotics and antibiotic resistance genes in eco-agricultural system: A case study. *Journal of Hazardous Materials*, 304: 18-25. <https://doi.org/10.1016/j.jhazmat.2015.10.037>
- [4] Chen, J. (2021). Multidimensional analysis model of agricultural product supply chain competition based on mean fuzzy. *Journal of Intelligent & Fuzzy Systems*, 41(2): 3591-3602. <https://doi.org/10.3233/JIFS-210962>
- [5] Lei, J. (2024). Research on the improvement path of international competitiveness of China's agricultural product supply chain from the perspective of machine learning. *Expert Systems*, 41(5): e12935. <https://doi.org/10.1111/exsy.12935>
- [6] Yu, Z., Rehman Khan, S.A. (2022). Evolutionary game analysis of green agricultural product supply chain financing system: COVID-19 pandemic. *International Journal of Logistics Research and Applications*, 25(7): 1115-1135. <https://doi.org/10.1080/13675567.2021.1879752>
- [7] Shen, L., Li, F., Li, C., Wang, Y., Qian, X., Feng, T., Wang, C. (2020). Inventory optimization of fresh agricultural products supply chain based on agricultural superdocking. *Journal of Advanced Transportation*, 2020(1): 2724164. <https://doi.org/10.1155/2020/2724164>
- [8] Singh Gill, H., Singh Khehra, B. (2020). Efficient image classification technique for weather degraded fruit images. *IET Image Processing*, 14(14): 3463-3470. <https://doi.org/10.1049/iet-ipr.2018.5310>
- [9] Gan, H., Lee, W.S., Alchanatis, V., Ehsani, R., Schueller, J.K. (2018). Immature green citrus fruit detection using color and thermal images. *Computers and Electronics in Agriculture*, 152: 117-125. <https://doi.org/10.1016/j.compag.2018.07.011>
- [10] Mamat, N., Othman, M.F., Abdulghafor, R., Alwan, A.A., Gulzar, Y. (2023). Enhancing image annotation technique of fruit classification using a deep learning approach. *Sustainability*, 15(2): 901. <https://doi.org/10.3390/su15020901>
- [11] Feng, J., Zeng, L., He, L. (2019). Apple fruit recognition algorithm based on multi-spectral dynamic image analysis. *Sensors*, 19(4): 949. <https://doi.org/10.3390/s19040949>
- [12] Wang, J., Feng, T. (2023). Supply chain ethical leadership and green supply chain integration: A moderated mediation analysis. *International Journal of Logistics Research and Applications*, 26(9): 1145-1171. <https://doi.org/10.1080/13675567.2021.2022640>
- [13] Li, D., Zhang, Z., Xu, R. (2025). The impact of environmental regulation on green image management of supply chain: Evidence from China. *Finance Research Letters*, 74: 106723. <https://doi.org/10.1016/j.frl.2024.106723>
- [14] Kumar, A., Agrawal, S. (2024). A quality-based sustainable supply chain architecture for perishable products using image processing in the era of Industry 4.0. *Journal of Cleaner Production*, 450: 141910. <https://doi.org/10.1016/j.jclepro.2024.141910>
- [15] Niu, B., Shen, Z., Xie, F. (2021). The value of blockchain and agricultural supply chain parties' participation confronting random bacteria pollution. *Journal of Cleaner Production*, 319: 128579. <https://doi.org/10.1016/j.jclepro.2021.128579>
- [16] Khan, M.I., Khalid, S., Zaman, U., José, A.E., Ferreira, P. (2021). Green paradox in emerging tourism supply chains: Achieving green consumption behavior through strategic green marketing orientation, brand social responsibility, and green image. *International Journal of Environmental Research and Public Health*, 18(18): 9626. <https://doi.org/10.3390/ijerph18189626>
- [17] Lahri, V., Shaw, K., Ishizaka, A. (2021). Sustainable supply chain network design problem: Using the integrated BWM, TOPSIS, possibilistic programming, and  $\epsilon$ -constrained methods. *Expert Systems with Applications*, 168: 114373. <https://doi.org/10.1016/j.eswa.2020.114373>
- [18] Zhang, R., Ying, Y., Rao, X., Li, J. (2012). Quality and safety assessment of food and agricultural products by hyperspectral fluorescence imaging. *Journal of the Science of Food and Agriculture*, 92(12): 2397-2408. <https://doi.org/10.1002/jsfa.5702>
- [19] Dağtekin, M., Beyaz, A. (2017). Use of different types of sensors for size measurement of some agricultural products (fruits) grown in Mediterranean climatic

- conditions. *Environmental Engineering & Management Journal (EEMJ)*, 16(5): 1129-1136.
- [20] Patel, K.K., Kar, A., Jha, S.N., Khan, M.A. (2012). Machine vision system: A tool for quality inspection of food and agricultural products. *Journal of Food Science and Technology*, 49: 123-141. <https://doi.org/10.1007/s13197-011-0321-4>
- [21] Costa, C., Antonucci, F., Pallottino, F., Aguzzi, J., Sun, D.W., Menesatti, P. (2011). Shape analysis of agricultural products: A review of recent research advances and potential application to computer vision. *Food and Bioprocess Technology*, 4(5): 673-692. <https://doi.org/10.1007/s11947-011-0556-0>
- [22] Li, J.B., Rao, X.Q., Ying, Y.B. (2011). Advance on application of hyperspectral imaging to nondestructive detection of agricultural products external quality. *Spectroscopy and Spectral Analysis*, 31(8): 2021-2026. [https://doi.org/10.3964/j.issn.1000-0593\(2011\)08-2021-06](https://doi.org/10.3964/j.issn.1000-0593(2011)08-2021-06)