

# Appearance quality identification and environmental factors tracing of *Lyophyllum decastes* for precise environment control using knowledge graph

Kai Zhou<sup>a</sup>, Junyuan Yu<sup>a</sup>, Haotong Shi<sup>a</sup>, Rui Hou<sup>b,\*</sup>, Huarui Wu<sup>c,\*</sup>, Jialin Hou<sup>a</sup>

<sup>a</sup> College of Mechanical and Electronic Engineering, Shandong Agricultural University, Tai'an 271018, China

<sup>b</sup> Department of Artificial Intelligence, Beijing University of Posts and Telecommunications, Beijing 100876, China

<sup>c</sup> National Engineering Research Center for Information Technology in Agriculture, Beijing 100097, China



## ARTICLE INFO

### Keywords:

*Lyophyllum decastes*  
Appearance identification  
Residual neural network  
Knowledge graph  
Graph attention network

## ABSTRACT

In the factory production of *Lyophyllum decastes*, inappropriate cultivation environments can lead to appearance quality issues, which in turn affect both yield and quality. However, the appearance characteristics of *Lyophyllum decastes* influenced by environmental factors share similarities, and the environmental factors that cause appearance quality problems exhibit coupling and complexity. Therefore, the identification of appearance characteristics and tracing of environmental factors present significant challenges. To address this issue, this paper proposes a multimodal learning network, DCRes-GAT, which integrates an improved Residual Neural Network (DCResNet) and a Graph Attention Network (GAT) to accurately identify the features of *Lyophyllum decastes*, while simultaneously tracing environmental factors and providing control recommendations. First, a knowledge graph based on the prior knowledge of quality and environmental factors is constructed, mapping this information to a point space and extracting key features. Next, DCResNet is employed to extract optical features from *Lyophyllum decastes* images. In addition, the receptive field is expanded through dilated convolutions, while pixel-level details are preserved, and a Convolutional Block Attention Module (CBAM) is incorporated to identify subtle visual differences. Finally, a dot product operation fuses point-space features with visual features, achieving accurate identification of characteristics and providing suggestions. Experimental results demonstrate that the DCRes-GAT model performs excellently, with a feature identification accuracy of 99.45%, and can precisely diagnose key environmental factors that cause appearance quality problems, achieving a diagnostic accuracy of 99.84%. This provides a basis for the precise control of the cultivation environment of *Lyophyllum decastes*.

## 1. Introduction

Edible fungi are high-value macrofungi that are rich in high-quality proteins, carbohydrates, various vitamins, and other nutrients. They are characterized by their high protein content, low fat, and low cholesterol levels (Zhang et al., 2021). Among these, *Lyophyllum decastes* is widely favored by consumers for its delectable taste and texture (Ke et al., 2023). Recent studies have highlighted its significant health benefits. Zhang et al. (2023) isolated a novel polysaccharide LDSP60-A from *Lyophyllum decastes* with significant antioxidant activity, supporting its use in functional foods and pharmaceuticals, as well as antioxidant, anti-hyperglycemic, lipid-lowering, and hepatoprotective activities (Zhang et al., 2022). *Lyophyllum decastes* and other edible fungi have demonstrated anti-cancer properties and various therapeutic effects, making

them valuable resources for both nutrition and medicine (Ba et al., 2021). These attributes underscore the dual role of *Lyophyllum decastes* as an edible and medicinal fungus, presenting significant opportunities for further research and development.

The cultivation of edible fungi involves a complex biological process influenced by various environmental factors, such as temperature, humidity, ventilation, and light intensity. To achieve both high yield and premium quality in production, environmental conditions necessitate precise adjustments based on the *Lyophyllum decastes* appearance during growth. However, the current monitoring and assessment of environmental factors rely on traditional manual records, and the identification of *Lyophyllum decastes* characteristics mainly depends on the experience-based judgment of production personnel. Subjective evaluations are easily influenced by individual experience and lack standardized

\* Corresponding authors.

E-mail addresses: [zhoukai2017@sdau.edu.cn](mailto:zhoukai2017@sdau.edu.cn) (K. Zhou), [rui.hou@bupt.edu.cn](mailto:rui.hou@bupt.edu.cn) (R. Hou), [wuhr@nercita.org.cn](mailto:wuhr@nercita.org.cn) (H. Wu).

criteria, which may lead to misjudgments of growth abnormalities or diseases, thereby affecting yield and quality, presenting significant limitations (Wu et al., 2022).

Artificial intelligence and machine vision technologies have been increasingly developed in research on edible fungi. However, current applications primarily address tasks such as the identification of poisonous mushrooms according to their forms, the plucking of cultivated mushrooms covered by soil, and the mechanized grading of mushrooms (Yin et al., 2022). Deng et al. (2022) developed a deep-learning-based wireless visual sensor system for shiitake mushroom sorting. The system achieved an accuracy of 98.53 % during actual mushroom-sorting tasks. Tao et al. (2024) proposed the ReYOLO-MSM evaluation method, which improves the ReYOLO model by incorporating a rotating ellipsoid framework. This method, based on monocular vision and the growth relationship of mushroom logs, achieves accurate, real-time, and cost-effective assessment and harvesting of mushrooms. Erdal Özbay et al. (2024) developed an innovative mushroom classification method combining Grad-CAM, LIME, and Heatmap techniques with a residual block Convolutional Neural Network (CNN). Using the ASO algorithm, they optimized the feature map from a size of 6714 samples × 9000 features to 6714 samples × 600 features. Through multi-feature fusion and *meta*-heuristic optimization, the model achieved a classification accuracy of 95.45 %, enhancing recognition precision and interpretability.

Although deep learning technologies have been utilized in the identification of edible fungi, single deep learning networks are limited to extracting representational features of the target, resulting in partial and inconsistent analysis. To achieve finer-grained characterization and accurately trace the environmental factors influencing the states of edible fungi, it is essential to integrate prior knowledge with advanced reasoning methodologies, facilitating precise traceability.

Knowledge graphs (KGs) are an effective knowledge management technology in Natural Language Processing (NLP), with applications across various domains. For instance, Mishra and Shridevi (2024) proposed a KG-driven medical recommendation system based on longitudinal medical records. By constructing clinical and pharmaceutical KGs, the study leveraged graph neural networks (GNNs) and recurrent neural networks (RNNs) to learn embeddings and temporal features, employing attention mechanisms to generate personalized recommendations. In the agricultural sector, KGs have demonstrated their potential to enhance production efficiency, improve decision-making accuracy, and increase the accessibility of information. For example, Yates et al. (2024) developed a weather analysis system based on natural language generation, offering planting recommendations and growers by analyzing local weather data and crop types. Similarly, Murali and Anuncia (2022) proposed a method to mining agricultural information, for creation of ontologies for agricultural filed. The incremental mining method was used to extract knowledge from multiple ontologies, which provided the possibility for the construction of agricultural KG. Wang et al. (2023) constructed a soybean pest KG and demonstrated their potential use in automated reasoning and recommendations for pest prevention solutions. Lv et al. (2024) proposed an agricultural multimodal KG with pretrained language models on cabbage and maize. Additionally, Guan et al. (2021) integrated KGs, representation learning, and deep neural networks to construct a diagnostic model for orchard pests and diseases, which significantly improved the diagnostic accuracy and efficiency. These examples collectively proved the robustness of logical reasoning and processing capabilities with NLP in critical agricultural applications including decision analysis, intelligent question answering, pest and disease analysis, and traceability of environmental factor.

As an important tool in knowledge management, KGs involve diverse forms of knowledge, such as images, text, audio, and video. Despite their potential, prior research on edible fungi appearance quality problems identification has predominantly focused on image-based information, which contrasts with principles of efficient smart agriculture and

knowledge management. Compared to single-modal data, multimodal provides richer semantic content and a more comprehensive understanding. However, integrating multimodal knowledge into the specific domain of edible fungi remains a significant challenge due to the complexity of combining diverse data types, such as images, environmental factors, and textual information, while maintaining accurate identification and traceability. To address these issues, this study proposes the following advances:

First, the raw data is systematically cleaned and integrated with external knowledge and expert annotations to construct a high-quality, multi-dimensional dataset. Based on this, a knowledge graph incorporating quality science and environmental factors is built by combining prior knowledge with text pattern information. Through graph structure mapping, unstructured information is projected into the point cloud space to enhance the model's understanding and reasoning capabilities under complex growth conditions. Second, an innovative multimodal inference learning network architecture, DCRes-GAT, is designed, which combines CNNs and GAT to extract features from both the point space of the knowledge graph and images of *Lyophyllum decastes*, followed by feature fusion. Finally, a framework is established to accurately track key environmental factors related to the appearance quality of *Lyophyllum decastes*. This framework integrates multiple environmental variables to better understand the mushroom's growth conditions, enabling precise feature recognition and traceability analysis while providing corresponding control recommendations.

## 2. Materials and methods

### 2.1. Data and modules

#### 2.1.1. Collection of prior knowledge and dataset establishment for *Lyophyllum decastes*

*Lyophyllum decastes* is highly sensitive to environmental conditions, with its quality and yield being directly influenced by key environmental factors. However, research investigating how these environmental factors affect the growth, yield, and quality of *Lyophyllum decastes* remains limited, posing significant challenges to the optimization of cultivation practices and production outcomes.

To address this gap, a comprehensive dataset on *Lyophyllum decastes* has been independently constructed, given the lack of publicly available datasets both domestically and internationally. This effort involved conducting an in-depth field survey at Shandong Agriculture and Mushroom Industry Co., Ltd. in China and reviewing extensive literature to investigate the environmental factors affecting its growth. The images in the constructed dataset have a resolution of 4096 × 3072 to ensure high-quality visual feature extraction. All images are captured in a standardized mushroom cultivation room setting with good lighting conditions, using a Huawei P40 Pro smartphone. The images are selected based on typical disease characteristics at different growth stages to ensure the dataset's representativeness. All images are independently annotated by three experienced edible mushroom cultivation experts, and annotations are considered correct only when all three experts agree. In cases of disagreement, the annotations are discussed to minimize the possibility of errors. Our study aims to demonstrate how these factors influence the quality and yield of *Lyophyllum decastes*, thus providing a foundation for advancing cultivation practices and optimizing production.

The appearance characteristics of factory-cultivated *Lyophyllum decastes* are classified into six categories, "Yellow Dry", "White Spot", "Disordered Budding", "Slow Growth", "Normal Growth", and "Incomplete Uneven" based on observable traits that reflect growth abnormalities or normal development under varying environmental conditions. The key environmental factors influencing the growth of *Lyophyllum decastes* including temperature, humidity, light, and ventilation. To facilitate the assessment of these factors, the study categorizes them into eight specific conditions: "High Temperature (HT)", "High

Humidity (HH)", "Adequate Light (HL)", "Adequate Ventilation (AV)", "Low Temperature (LT)", "Low Humidity (LH)", "Lack of Light (LL)", and "Lack of Ventilation (SV)". These categories help evaluate the impact of different environmental conditions on the growth characteristics of *Lyophyllum decastes* and support the optimization of cultivation strategies. As part of the survey, a diverse set of photos showcasing the six characteristics of *Lyophyllum decastes* was collected and used to construct the dataset. Examples of this dataset are shown in Fig. 1.

Various environmental factors influencing the growth of *Lyophyllum decastes* and their specific impacts on its characterization traits are shown in Table 1. The correlation rate reflects the extent to which each factor affects the characteristics of *Lyophyllum decastes*; a higher correlation rate signifies a greater impact of the respective factor. A higher correlation rate indicates a stronger impact of the respective factor.

### 2.1.2. Creation of *Lyophyllum decastes* KG

KG, a key technology in NLP, represent knowledge through graph structures, with nodes for entities or concepts and edges for semantic relationships. This representation facilitates the integration and analysis of multi-dimensional knowledge. GNNs are frequently used to process such graph-structured data, enabling the extraction of complex relationships between nodes and edges (Ji et al., 2022). Zhu et al. (2023) constructed a fruit pest detection and identification system using KG, achieving an accuracy of 94.9 % in pest recognition tasks. Zhao et al. (2022) integrated KG into the moment-preserving segmentation method to achieve automatic identification of crop types during the growing season.

For *Lyophyllum decastes*, a KG that links appearance quality traits with key environmental factors were developed, as shown in Fig. 2. KG offers significant advantages over traditional relational databases, particularly in their ability to construct highly flexible point-to-point semantic spaces. By visually representing prior knowledge and uncovering deep relationships within the data, the graph provides an intuitive and flexible framework for analyzing and understanding the complex interactions influencing the growth and quality of *Lyophyllum decastes*.

Furthermore, this KG can be continuously enriched through regular updates and expansions, enhancing its utility for network learning and future research. Fig. 3 illustrates the constructed KG, where each node (entity) represents an appearance quality feature of *Lyophyllum decastes* or an environmental factor, while each edge (relationship) reveals the inherent logic and dependencies between these entities. By mining and analyzing the KG, the critical environmental factors affecting the appearance quality of *Lyophyllum decastes* can not only be visualized, but

also how these factors interact with one another can be understood. For instance, how factors such as temperature, humidity, light intensity, and ventilation strength influence the color, shape, and size of the mushrooms. Moreover, the correlation probabilities between entities provide a quantitative assessment of each factor's influence, forming a robust scientific basis for optimizing environment control and enhancing production quality.

## 2.2. Multi-modal learning network DCRes-GAT for characterization and environmental traceability of *Lyophyllum decastes*

### 2.2.1. The structure of the DCRes-GAT network

Fig. 4 illustrates the DCRes-GAT architecture, which integrates two key learning processes: graph learning and representation learning. Specifically, the graph learning focuses on mining environmental factor attributes, while representation learning emphasizes analyzing the characteristic features of *Lyophyllum decastes*. In the graph learning phase, a KG is implemented to encapsulate the prior relationships between the appearance quality of *Lyophyllum decastes* and environmental factors. A GAT serves as the foundation for capturing and modeling the relationships between nodes in the KG. By introducing an attention mechanism, the GAT selectively prioritizes key nodes while minimizing the influence of less important ones, thereby improving both the efficiency and accuracy of feature extraction. Additionally, the parameterized weights within the attention mechanism enables dynamic structural adjustments during the learning process, enhancing the network's capacity to handle complex spatial dimensions.

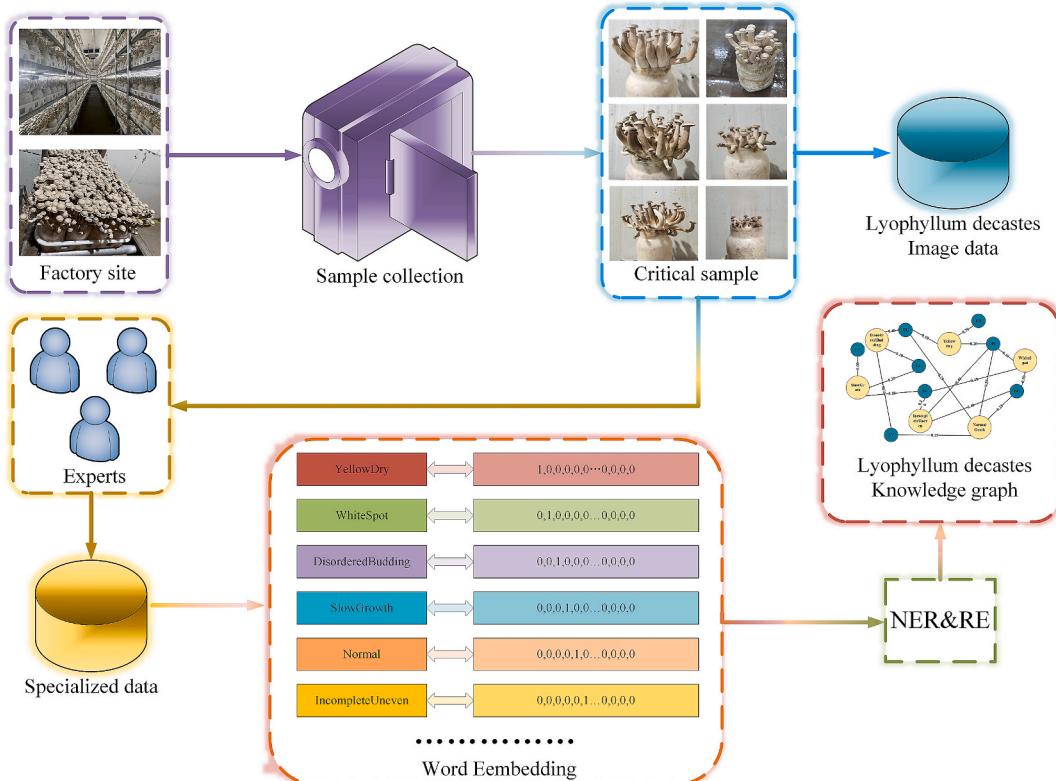
In the representation learning phase, an enhanced ResNet-34 network is developed to extract primary features from the input images. To further refine feature extraction, Dilated Convolution Networks are incorporated to dynamically adjusting the receptive field size so that sufficient pixel information can be retained. This design effectively excavates multi-scale features of *Lyophyllum decastes* while minimizing information loss. To improve sensitivity to subtle visual distinctions, CBAM are integrated into the architecture, enhancing the network's ability to focus on subtle visual differences in the identified images. The outputs from the graph learning and representation learning phases are then fused using a dot product method. Unlike simple concatenation methods, the dot product approach effectively simulates the complex interactions between visual and textual data, uncovering latent information which could significantly enhance the overall performance. ResNet34's conv1 has 64 filters, each with a  $7 \times 7$  (width  $\times$  height  $\times$  input channels) convolutional kernel. In the



**Fig. 1.** Six different characteristics of *Lyophyllum decastes*' appearance.

**Table 1**Appearance quality characterizations of *Lyophyllum decastes* and probability of environmental factors association.

Characterization	Probability of Association between Environmental Factors and Characterization							
	HT	HH	HL	AV	LT	LH	LL	SV
Yellow Dry	0.20	0	0.10	0	0	0.70	0	0
White Spot	0.10	0.50	0	0	0	0	0	0.40
Disordered Budding	0	0	0.40	0.50	0	0	0.10	0
Slow Growth	0	0	0	0	0.60	0	0.30	0.10
Normal Growth	0.25	0.25	0.25	0.25	0	0	0	0
Incomplete Uneven	0.40	0.40	0	0	0	0	0	0.20

**Fig. 2.** The establishment of the array library: The upper section focuses on image acquisition and the construction of image dataset, the lower section deals with the collection of expert knowledge embedding and the construction of domain-specific KGs.

residual layers, ResNet34 uses BasicBlock, which consists of two consecutive  $3 \times 3$  convolutional layers, each followed by batch normalization (BN) and ReLU activation. The input dimension for the global average pooling and fully connected layer is 512, and the output dimension is 14. The graph convolutional layer gat1 has in\_features of 300 and out\_features of 1024; gat2 has in\_features of 1024 and out\_features of 512. The activation function used is LeakyReLU with a negative slope of 0.2.

With rapid advancements in computer technology, image digital processing and deep learning have significantly improved, with neural network models becoming deeper. However, increasing depth initially boosts model performance but eventually leads to saturation and rapid decline, causing issues like gradient vanishing and failing to further enhance model performance (Jafar and Lee, 2021).

To address this issue, ResNet was proposed (He et al., 2016), which introduce residual blocks to effectively skip layers that fail to learn or perform poorly by approximating their weight parameters to zero. This design helps maintain or improve performance as network depth increases, avoiding the problem of gradient vanishing.

Nevertheless, deep network connectivity is both costly and complex, making it unsuitable for agricultural cultivation areas that demand

straightforward deployment and efficient operations. To select the most suitable model that optimally balances performance and efficiency, models' performance were evaluated by using the ImageNet-1 K image classification dataset. During the evaluation process, the Top-1 error rate and Top-5 error rate were calculated carefully, which are used to measure the accuracy of the model as well as its generalization ability and fault tolerance. Also, the time required to complete one training epoch and the number of parameters required by the model to evaluate the efficiency of model training and its consumption of computational resources were calculated. The computer system used for testing is Windows 10, with 24.0 GB of RAM, an Intel(R) Core(TM) i7-10700 @ 2.90 GHz CPU, and an NVIDIA GTX1660 SUPER GPU with 6.0 GB of dedicated GPU memory.

The experimental results are shown in Table 2. When trained using the standard training set ImageNet, deeper ResNet networks consume a significant amount of computational resources and training time. The 50-layer, 101-layer, and 152-layer ResNet networks have a training accuracy difference of less than 2 percentage points compared to the 18-layer and 34-layer networks. However, in terms of training time, the 50-layer, 101-layer, and 152-layer networks require much more time compared to the 34-layer network. Although deeper models such as

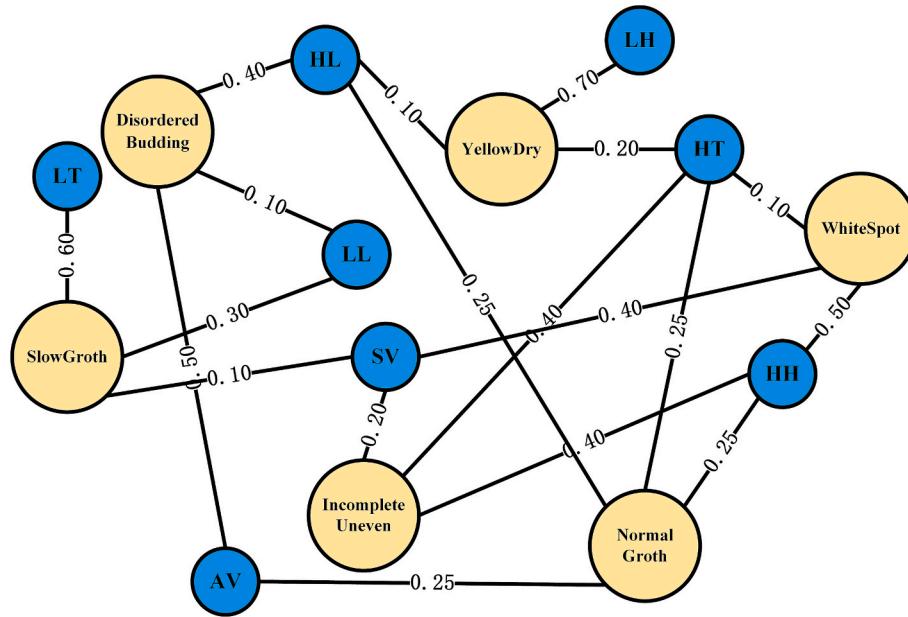


Fig. 3. KG of appearance quality characteristics and environmental factors.

ResNet-101 can capture more complex features, they require significantly higher computational resources (such as GPU memory and training time). For the specific task of *Lyophyllum decastes* grading and environmental factor traceability, ResNet-34 strikes an ideal balance between accuracy and efficiency. This study focuses on integrating the GCN with the backbone network, and ResNet-34 provides a lightweight and stable feature base for this purpose.

Therefore, ResNet34 was chosen, which consumes fewer resources, takes less time, and has relatively high accuracy, for building the fusion network.

#### 2.2.2. Graph attention network

Graph Convolutional Networks (GCNs) are a foundational type of GNNs designed to process graph-structured data by leveraging convolutional operations to extract meaningful features. They enable efficient feature learning for various tasks such as node classification, edge prediction, and overall graph classification (Valem et al., 2023). As GCNs lack the ability to differentiate the varying importance of neighboring nodes during feature aggregation, GAT extends the GCN architecture by incorporating attention mechanisms. This enhancement allows GAT to dynamically assign weights to neighboring nodes based on their features, enabling the graph attention layer, a core component of GAT, to generate refined feature vectors through weighted processing. By learning a shared weight matrix  $W$  and applying self-attention (Garcia-Garcia et al., 2018), the system calculates attention coefficients as follows:

$$e_{ij} = a(\vec{Wh}_i, \vec{Wh}_j) \quad (1)$$

In this process, the vectors  $\vec{h}_i$  and  $\vec{h}_j$  denote the feature vectors of nodes  $i$  and  $j$  respectively, while  $a$  represents a predefined function. This expression quantifies the relative importance of node  $j$  to node  $i$ . To ensure that attention is restricted to the neighborhood set  $N_i$  of node, masked attention is applied to the graph structure. Additionally, the SoftMax function is used to normalize the attention coefficients of all neighboring nodes  $j$  of node  $i$ , facilitating comparability and simplifying calculation, as is shown in equation (2).

$$\alpha_{ij} = softmax_i(e_{ij}) = \frac{\exp(e_{ij})}{\sum_{k \in N_i} \exp(e_{ik})} \quad (2)$$

Combining the above two formulas, the complete attention

mechanism is expressed as:

$$\alpha_{ij} = \frac{\exp\left(\text{LeakyReLU}\left(\vec{a}^T [\vec{Wh}_i || \vec{Wh}_j]\right)\right)}{\sum_{k \in N_i} \exp\left(\text{LeakyReLU}\left(\vec{a}^T [\vec{Wh}_i || \vec{Wh}_k]\right)\right)} \quad (3)$$

where  $\vec{a}$  represents the weight matrix used for connecting the various layers of the network. Additionally, a *LeakyReLU* activation function is introduced in the output layer of the model to enhance the network's nonlinear expressive capability. The symbol  $T$  denotes the transposition operation, and  $||$  indicates the concatenation of feature vectors  $\vec{Wh}_i$  and  $\vec{Wh}_k$ . These operations yield normalized attention coefficients, representing the relative importance of each neighboring node. These attention coefficients guide the aggregation of neighborhood information to estimate the output feature vector for each node.

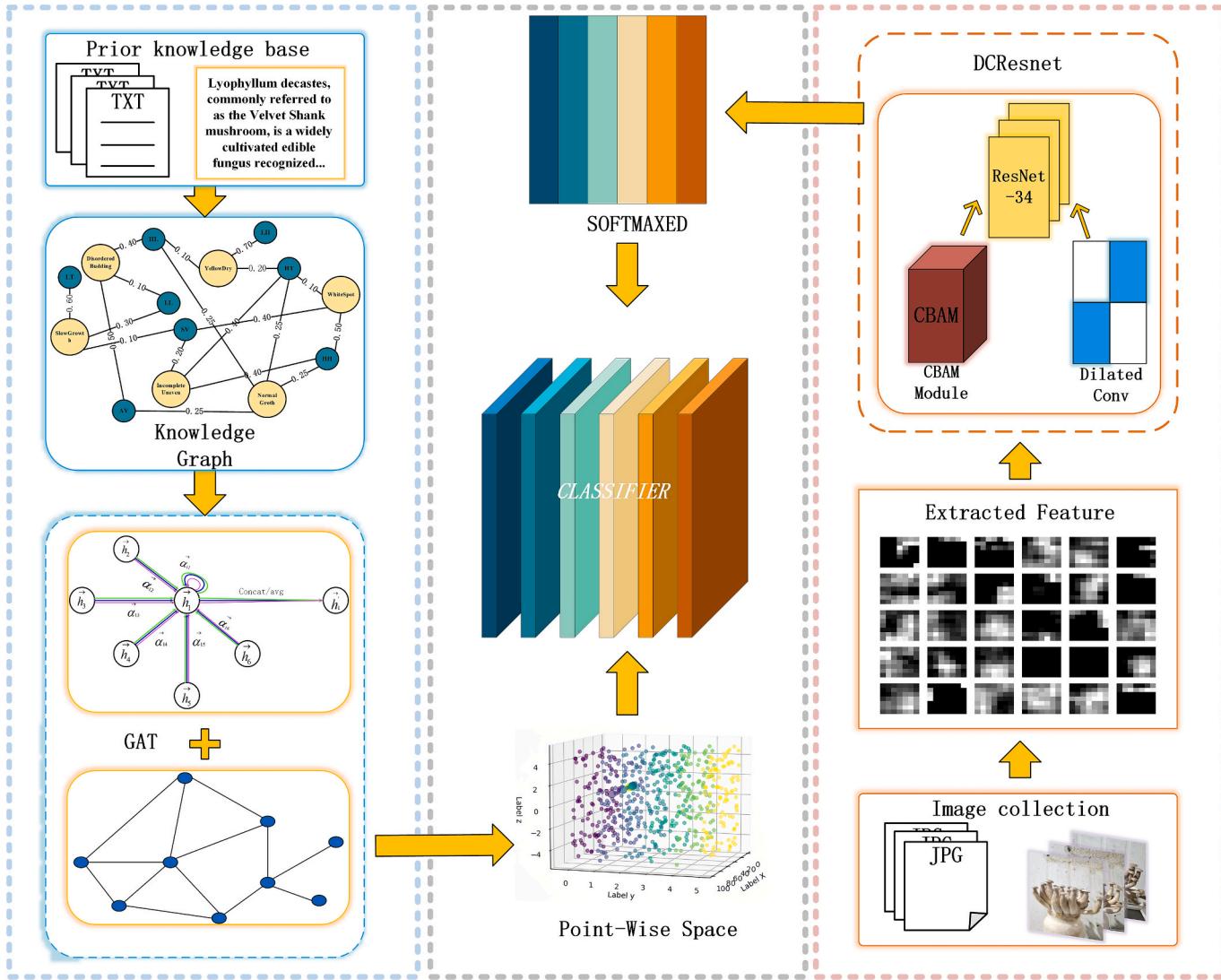
$$\vec{h}_i = \sigma\left(\sum_{j \in N_i} \alpha_{ij} \vec{Wh}_j\right) \quad (4)$$

Here,  $\vec{h}_i$  represents the feature vector output by node  $i$ , with  $W$  serving as the weight matrix for transforming the input features. Additionally,  $a$  denotes the attention coefficients obtained from the aforementioned computation, and  $\sigma$  is the activation function used to introduce non-linearity. To enhance the stability of the self-attention learning process, the model incorporates a multi-head attention mechanism, which forms the foundation of GAT:

$$\vec{h}'_i = \sigma\left(\frac{1}{K} \sum_{k=1}^K \sum_{j \in N_i} \alpha_{ij}^k W^k \vec{h}_j\right) \quad (5)$$

In equation (5), the model integrates  $K$  different attention mechanisms, each represented by  $a^k$ . For each mechanism  $k$ , the input features are transformed through a linear operation using the corresponding weight matrix  $W^k$ .

The GAT architecture with its multi-head attention mechanism, dynamically prioritizes important nodes and reduces the influence of less significant ones. This multi-head attention mechanism not only enhances flexibility in handling diverse graph structures but also strengthens the model's ability to represent complex relationships, laying a robust foundation for advanced graph-based learning and



**Fig. 4.** The architecture of the DCRes-GAT network.

**Table 2**  
Performance evaluation of different ResNet architectures.

Architectures	Top-1 Error (%)	Top-5 Error (%)	Param number ( per Epo )	Training Speed (min)
Resnet18	27.28	10.56	12,018,260	1.08
Resnet34	24.20	7.42	22,126,420	1.48
Resnet50	22.83	6.75	24,380,500	19.0
Resnet101	21.73	6.01	43,372,628	58.19
Resnet152	21.40	5.65	125,780,621	99.44

analysis.

### 2.2.3. Improved residual network DCResNet

#### (1) Amplification of receptive field.

In extracting visual features from images of *Lyophyllum decastes*, representation learning is essential to address the challenges posed by variations in size, shape, and color due to different shooting angles and positions. Traditional CNNs, with their fixed receptive field sizes, struggle to adapt to such variations, often resulting in the omission of critical feature information and reduced identification accuracy. To overcome this issue, dilated convolution operations were integrated into the third and fourth residual block groups of ResNet-34, enhancing the

network's ability to capture multi-scale features. Both dilated convolution and standard convolution have the same number of parameters and theoretical floating-point operations (FLOPs) because they use the same kernel size. The key difference is that dilated convolution introduces spacing within the kernel, thus expanding the receptive field. However, this does not increase the actual multiplication and accumulation operations within the kernel, so theoretically, the computational complexity is almost identical to that of standard convolution. In contrast, other methods, such as self-attention and deformable convolution, although improving feature extraction capabilities, significantly increase computational complexity.

In standard convolution operations, the input feature map  $X$ , representing the visual features of *Lyophyllum decastes*, is assumed to have dimensions  $H \times W \times C$ , where  $H$  and  $W$  denote the height and width of the image, and  $C$  represents the number of channels in the feature map. A convolution kernel  $K$ , designed to extract local features, has dimensions  $k_h \times k_w \times C \times C'$ , where  $k_h$  and  $k_w$  are the height and width of the kernel, and  $C'$  is the number of channels in the output feature map.

$$Y_{dilated}[m, n] = \sum_i \sum_j X[m + i \times d, n + j \times d] \times K[i, j] \quad (6)$$

$Y_{dilated}[m, n]$  represents the value of the output feature map from the dilated convolution at position  $[m, n]$ .  $X[m + i \times d, n + j \times d]$  denotes the value of the input feature map  $X$  at position  $[m + i \times d, n + j \times d]$ , where

$d$  is the dilation rate.  $K[i, j]$  represents the value of the convolution kernel  $K$  at position  $[i, j]$ .

Dilated convolution enables the model to capture a broader range of contextual information in the image, which is critical for addressing variations in size and morphology caused by different shooting angles of *Lyophyllum decastes*. In the ResNet-34 model, dilated convolutions were introduced to enhance the network's capability to perceive features of *Lyophyllum decastes*. The outputs of dilated convolutions are fused with those from standard convolutions via residual connections. The fused computation formula is expressed as:

$$Y_{\text{fused}} = \text{Residual}(X) + \text{DilatedConv}(X, K_{\text{dilated}}, d) \quad (7)$$

Here,  $X$  is the input feature map,  $\text{Residual}(X)$  represents the residual connection (adding the input  $X$  directly to the output), and  $\text{DilatedConv}(X, K_{\text{dilated}}, d)$  denotes the dilated convolution operation. The residual connection helps preserve local detail information in the image, while the dilated convolution expands the receptive field to capture global features of *Lyophyllum decastes*. Additionally, dilated convolutions with varying dilation rates ( $d_1, d_2, \dots, d_n$ ) are applied to extract multi-scale features from the image. This process can be mathematically represented as:

$$Y_{\text{multi-scale}} = \sum_{i=1}^n \text{DilatedConv}(X, K_{\text{dilated}}^i, d_i) \quad (8)$$

Here,  $X$  represents the input feature map, and  $K_{\text{dilated}}^i$  denotes the convolution kernel with a specific dilation rate  $d_i$ . By integrating features from dilated convolutions at multiple scales, the model adapts more effectively to variations in the features of *Lyophyllum decastes* caused by different shooting angles, background interference, and size changes.

Dilated convolution enhances the receptive field by introducing gaps between the elements of the convolution kernel, allowing the model to capture a broader range of contextual information without increasing the number of parameters or computational complexity. This method significantly improves the model's ability to perceive features in *Lyophyllum decastes* images, especially under variations in object size and morphology. By adaptively extracting both global and local detailed features across multiple scales, dilated convolution enhances the model's robustness, enabling better identification of *Lyophyllum decastes* under varying angles and positions. Consequently, it improves identification accuracy and reliability.

#### (2) Application of CBAM.

CBAM enhances the performance of CNN by combining channel attention and spatial attention mechanisms. The channel attention mechanism prioritizes critical feature channels, improving the network's ability to discriminate key features, while spatial attention mechanism focuses on decisive regions of the image. Together, these mechanisms enable CBAM to reconstruct intermediate feature mappings, emphasizing essential information and suppressing irrelevant details for adaptive feature optimization. The structure of CBAM is shown in Fig. 5, where  $H$ ,  $W$ , and  $C$  denote the height, width, and the

number of channels, respectively.

To enhance the network's focus on critical features of *Lyophyllum decastes* fruiting bodies, the CBAM module was integrated after each residual block group. This addition enables more precise feature extraction from targeted regions of interest. The feature computation after convolution is expressed as:

$$F_n = \text{Conv}(F_m) \quad (9)$$

Here,  $F_n$  represents the final residual convolution layer integrated with the CBAM module, and  $F_m$  is the input feature map. After residual convolution processing, the feature map is adjusted by applying channel attention and spatial attention. The features first pass through the Channel Attention (CA) module and then the Spatial Attention (SA) module. In the CA module, the data flow is processed as follows:

$$u_{\text{avg}} = \text{AvgPool}(F) \quad (10)$$

$$u_{\text{max}} = \text{MaxPool}(F) \quad (11)$$

$$z_{\text{avg}} = \text{MLP}(u_{\text{avg}}) = W_2(W_1(u_{\text{avg}})) \quad (12)$$

$$z_{\text{max}} = \text{MLP}(u_{\text{max}}) = W_2(W_1(u_{\text{max}})) \quad (13)$$

$$M_c = \sigma(z_{\text{avg}} + z_{\text{max}}) \quad (14)$$

$$F' = F \times M_c \quad (15)$$

Here, different spatial attention maps,  $u_{\text{max}}$  and  $u_{\text{avg}}$ , are obtained by the employment of max pooling and average pooling operations, respectively.  $\text{MLP}$  represents a Multi-Layer Perceptron, composed of two fully connected layers,  $W_1$  and  $W_2$ , where  $W_1$  reduces the number of channels, and  $W_2$  expands it back to the original number of channels.  $M_c$  is the final channel attention weight map.

In the SA module the processed data was dealt in the same way, through average and max pooling, perform data concatenation and convolution, then the feature map  $F''$  can be calculated.

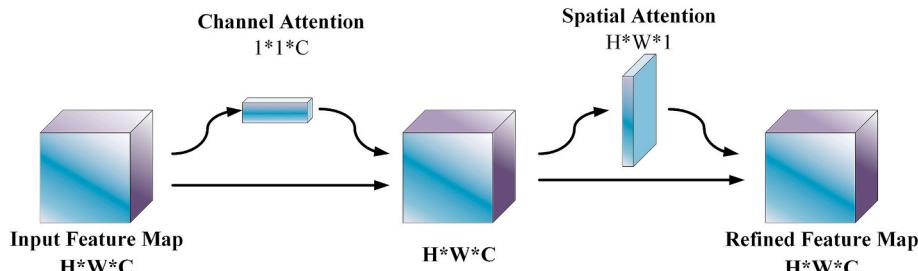
The pooled feature maps are concatenated along the channel dimension, and the spatial attention weights are then applied to obtain the final feature map.

$$F_{\text{out}} = F' + F_m \quad (16)$$

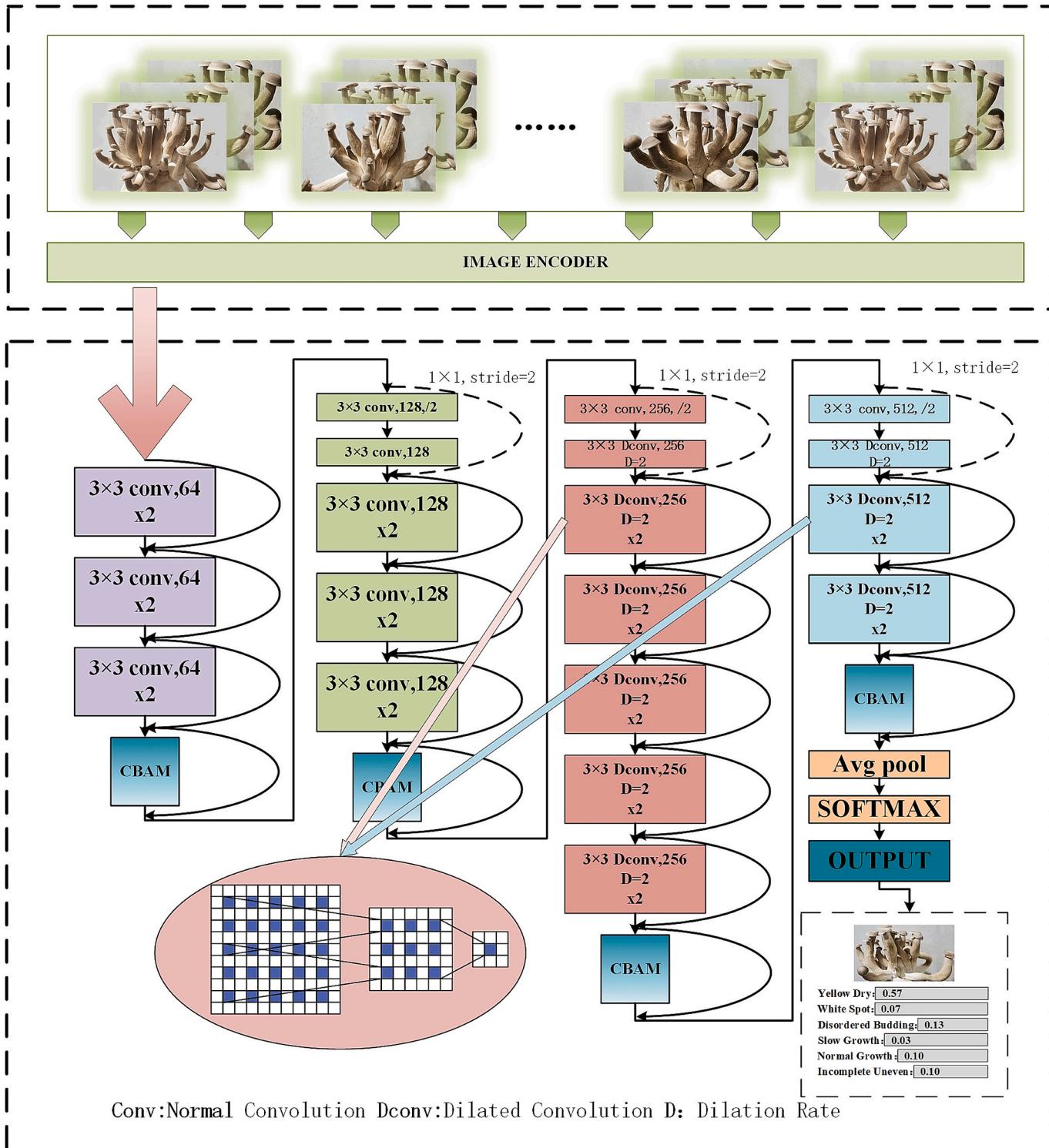
Here,  $F_{\text{out}}$  is the final output of the residual block, which is obtained by adding the CBAM-processed feature map  $F'$  to the original input  $F_m$ . The improved network architecture is shown in Fig. 6.

#### 2.3. Data sample and experimental setup

To validate the advanced nature of the proposed multimodal learning network, an experiment was conducted on the grading and environmental factors traceability of *Lyophyllum decastes*. The network was compared with mainstream machine vision grading algorithms and its pre-improvement version. Each algorithm was tested 10 times, with results averaged to eliminate random errors. The appearance of



**Fig. 5.** The structure of CBAM: The Channel Attention Module and the Spatial Attention Module, ultimately enhancing the important features while suppressing less useful ones.



**Fig. 6.** The structure of improved ResNet-34: Integrated the CBAM at the conclusion of various convolutional layers to enhance feature representation. Also, dilated convolutions were introduced in the final two stages to effectively capture multi-scale contextual information without significantly increasing the number of parameters.

*Lyophyllum decastes* was categorized into six categories based on its visual characteristics: dry yellow, white dot fungi, redundant budding, slow growth, normal, and residual limbs, as illustrated in Fig. 1. The data used in this experiment was sourced from Shandong Agriculture and Mushroom Industry Co., Ltd., with images annotated by expert staff from the company.

Each image was labeled with a 14-dimensional vector  $a = (a_1, a_2, \dots,$

$a_{14})$ , where the first six components represent the appearance quality of *Lyophyllum decastes*, and the last eight components represent the environmental factors influencing the growth of *Lyophyllum decastes*. During annotation, one of the first six components corresponding to the appearance quality was set to 1, with the others set to 0. Similarly, one of the last eight components corresponding to the environmental factors was set to 1, with the others set to 0. Considering the workload of expert

annotations and the resources involved, *Lyophyllum decastes* from the same mushroom house were selected to form the training set. Additionally, due to the varying probabilities of occurrence for different types of diseases, differences in sample sizes for each class are observed.

Since the experimental dataset contains 11,097 images of *Lyophyllum decastes*, and given that a single random split may result in an uneven distribution of key feature samples (such as minority classes or edge cases) between the training/validation sets, which could introduce significant random error, cross-validation was adopted to reduce the error and compare the results. In the preprocessing process, the image size is standardized through scaling and cropping. Additionally, data augmentation techniques based on random sampling are applied, including random rotation, horizontal flipping, Gaussian blur, and motion blur. All networks are implemented using Python 3.7 in the PyTorch framework and run on an NVIDIA 1660 SUPER GPU.

A 5-fold cross-validation approach is employed to divide the training and validation sets. The dataset is randomly divided into  $k$  equal-sized subsets (folds), with  $k$  chosen as 5. In each iteration,  $k-1$  folds are used as the training set, and the remaining fold is used as the test set. This process is repeated  $k$  times to ensure that each subset is given the opportunity to be used as the test set. During each training session, the same model parameters and training algorithm are applied, and the performance of the model is evaluated using the corresponding test set, with evaluation metrics recorded.

The accuracy reported in this study is the overall classification accuracy. Precision, recall, and F1 score are calculated for each category separately, and then averaged using the macro-average method. The detailed evaluation metrics are described as follows:

$$\text{Accuracy} = \frac{x}{y} \quad (17)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (18)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (19)$$

$$F1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (20)$$

where  $x$  represents the number of *Lyophyllum decastes* images correctly predicted, and  $y$  the total number of *Lyophyllum decastes* images.  $TP$  (*True Positives*) refers to the number of samples where the model correctly detects the corresponding features of *Lyophyllum decastes*.  $FP$  (*False Positives*) refers to the number of samples that are incorrectly identified as non-corresponding features or where the model fails to provide the correct solution.  $FN$  (*False Negatives*) refers to the number of samples that are incorrectly predicted as non-corresponding features.

Through 5-fold cross-validation, the model achieved an average accuracy of 99.11 % on the test set. The specific results are shown in Table 3:

As seen in Table 3, the model performed consistently across different folds, with all metrics maintaining a high level. This indicates that the knowledge graph enhanced the model's generalization ability. Compared to the single split, the model's accuracy showed minimal fluctuation. To reduce the load and improve computational efficiency,

**Table 3**  
Comparison of 5-fold cross-validation results and single split results.

Verification Method	Accuracy	Precision	Recall	F1
1	0.9856	0.9934	0.9958	0.9836
2	0.9785	0.9876	0.9867	0.9962
3	0.9987	0.9957	0.9936	0.9861
4	0.9968	0.9768	0.9954	0.9956
5	0.9960	0.9832	0.9836	0.9876
Average	0.99112	0.98734	0.99102	0.98982
Single Split	0.9945	0.9854	0.9942	0.9898

we adopted a single split approach to validate the model's superiority.

The dataset, consisting of 11,097 images, was randomly split into training, validation, and testing sets in an 8: 1: 1 ratio, with details shown in Table 4. Parameters in the loss function are set so that higher weights are assigned to samples from minority classes during the loss calculation, thereby balancing the impact of different classes on the model training.

The model was optimized by stochastic gradient descent (SGD) with a momentum of 0.9, and Multi-Label Soft Margin Loss served as the loss function. The initial learning rate was set to be 0.001, with a decay factor of by  $10^{-5}$  applied every 40 epochs. Batch size was set to be 12. The GAT uses Xavier initialization to ensure a reasonable distribution of the initial weights, promoting stable gradient updates. Training is automatically terminated after 500 epochs to prevent overfitting. The software versions used for the experiments are Python 3.7 and PyTorch 1.7.1. The output for each image from the model is a 14-dimensional vector  $b = (b_1, b_2, \dots, b_{14})$ , representing the predicted characterization of *Lyophyllum decastes* and the associated environmental factors influencing its growth. Model performance was evaluated using accuracy, precision, recall and F1 score.

### 3. Results and discussion

#### (1) Comparison of different models.

The performance metrics of different networks trained on the same dataset are shown in Table 5.

To validate the stability and effectiveness of the DCRes-GAT model, a comprehensive comparison test of its performance against several Cost-effective and efficient models in the vision field was conducted, including VGG, GoogLeNet, InceptionV3. This evaluation was performed across various training stages to assess how each model's accuracy evolves over time. The results are illustrated in Fig. 7, which presents a detailed analysis of the performance metrics throughout the entire training process. The findings revealed that the DCRes-GAT model consistently outperformed the other models at every stage of training. Specifically, from the early epochs to the later stages, DCRes-GAT demonstrated superior accuracy and stability. Notably, by the 50th epoch, DCRes-GAT achieved an impressive accuracy of 90 %, surpassing the performance of VGG, GoogLeNet, and InceptionV3, whose accuracies remained below 85 %. Moreover, the consistency of DCRes-GAT performance underscores its robustness and reliability. Unlike the other models, which exhibited fluctuations in accuracy during training, DCRes-GAT maintained a steady improvement trajectory. This stability is particularly crucial for practical applications where consistent performance is essential for reliable outcomes.

#### (2) Comparison of model training results.

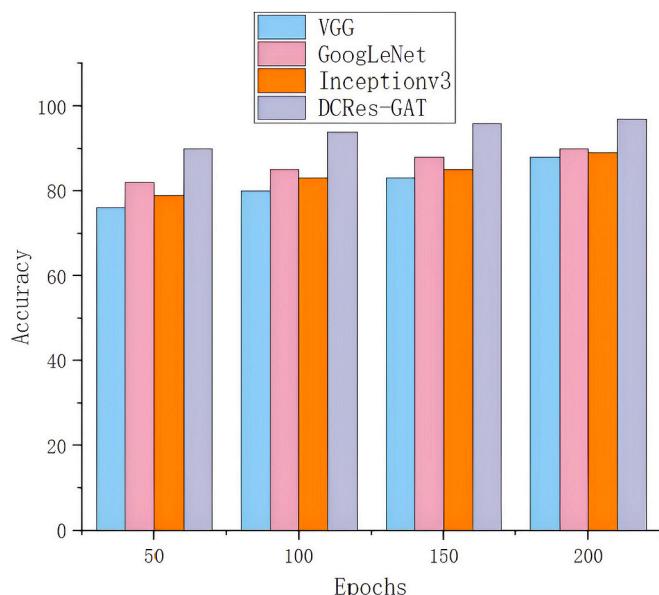
Fig. 8 illustrated the loss and accuracy changes during the training process of DCRes-GAT and ResNet-34. The data indicated that DCRes-GAT outperformed ResNet-34 in terms of the rate of loss reduction, and the convergence fluctuations were also notably smaller than those of ResNet-34. After 200 training epochs, the training loss of DCRes-GAT decreased to below 0.1, with an accuracy improvement to 99.45 %, while ResNet-34 only reduced the loss to 0.2, with an accuracy of only 94.58 %. Furthermore, ResNet-34 exhibited significant fluctuation and

**Table 4**  
Data set division.

Characteristic Features Types	Number of samples			Total
	Training set	Validation set	Test set	
Yellow Dry	1920	241	241	2402
White Spot	1952	243	243	2438
Disordered Budding	1586	199	199	1984
Slow Growth	1856	232	232	2320
Normal	1160	145	145	1450
Incomplete Uneven	405	50	50	505
Total	8877	1110	1110	11,097

**Table 5**  
Model evaluation metrics comparison.

Model	Accuracy	Precision	Recall	F1
DCRes-GAT	0.9945	0.9854	0.9942	0.9898
ResNet34	0.9458	0.9543	0.9361	0.9451
VGG19	0.9191	0.9075	0.8996	0.9035
GoogleNet	0.9156	0.9123	0.9088	0.9105
InceptionV3	0.8896	0.8745	0.8768	0.8756



**Fig. 7.** Comparison of accuracy of different algorithms: DCRes-GAT demonstrates superior performance over multiple epochs when compared to established architectures such as VGG, GoogLeNet, and InceptionV3.

instability in loss and accuracy during the validation phase, further demonstrating the superiority of DCRes-GAT in model fitting. This study systematically evaluated the performance differences among five deep learning models (DCRes-GAT, ResNet34, etc.) in the *Lyophyllum decastes* image classification task using an “ANOVA + Tukey HSD” two-tier statistical validation framework. The Shapiro-Wilk normality test confirmed that the Accuracy data from ten independent training runs for each model met the assumption of normality, and Levene’s test for homogeneity of variance showed no significant variance differences between groups, satisfying the prerequisites for ANOVA.

One-way ANOVA revealed a significant global performance difference among the models, with model selection explaining over 99 % of the variance in results, reaching a high effect size level in the deep learning field. Tukey HSD post-hoc tests confirmed that all pairwise model comparisons showed statistically significant differences ( $p < 0.001$ ). DCRes-GAT demonstrated a notable advantage: compared to the second-best model, ResNet34, its Accuracy improved by 4.89 % (95 % CI [4.83 %, 4.94 %]); when compared to the baseline model, InceptionV3, the Accuracy improvement was even more pronounced, reaching 10.51 % (95 % CI [10.45 %, 10.57 %]).

Considering these factors, the lower loss metric indicated less information loss, suggesting that DCRes-GAT significantly surpassed ResNet-34 in terms of learning capability.

### (3) Improved attention effect.

As illustrated in Fig. 9, to visually compare the performance of different network components, two optimal networks from the training process were selected for further experiments. Heatmaps illustrating the network’s identification of the “White Spot” characteristic were generated. The ResNet34 allocated more attention to less important boundaries and exhibited lower brightness for mushroom bodies with

issues. In contrast, DCRes-GAT focused more effectively on the areas of interest, demonstrating superior feature learning for mushroom characteristics.

### (4) Effectiveness of dilated convolution

As illustrated in Fig. 10, it is evident that the increased network depth enhances the receptive field, leading to more intricate filter structures. Consequently, the features extracted by dilated convolution become more abstract, which is reflected in the improved hierarchical representation of the image. It can be observed that the utilization of dilated convolution enables the network to extract more complex features and significantly enhances its capability to capture subtle characteristics.

### (5) Model capability validation

To verify the model’s actual accuracy and additional traceability capabilities, two experiments were organized to test each of these abilities separately.

The ResNet-34, which performed best in “Comparison of Different Models” was selected for comparison. To validate the practical effectiveness of the trained network across diverse conditions, a new test set was created by collecting additional images at Shandong Agriculture and Mushroom Industry. This test set, incorporating variations in representations and environmental factors, was designed to assess the network’s effectiveness, as shown in Fig. 11, with incorrect predictions highlighted using red boxes. It could be observed that DCRes-GAT achieved error-free identification, whereas the ResNet-34 network had two identification errors, demonstrating that DCRes-GAT had higher accuracy and superior versatility compared to ResNet-34.

To validate the network’s ability to trace critical regulatory factors influencing characterization under various environmental conditions, a controlled variable experiment was conducted. Since ResNet-34 lacks the capability to trace key environmental factors, only the DCRes-GAT model was used to validate its traceability capability. In this experiment, the characterization was limited to the “white-spot” appearance quality issues, and key factors associated with different characterizations were identified using expert knowledge from industry professionals. The three different levels of *Lyophyllum decastes* characterization are shown in Fig. 12.

The number of datasets labeled by experts is shown in Table 6.

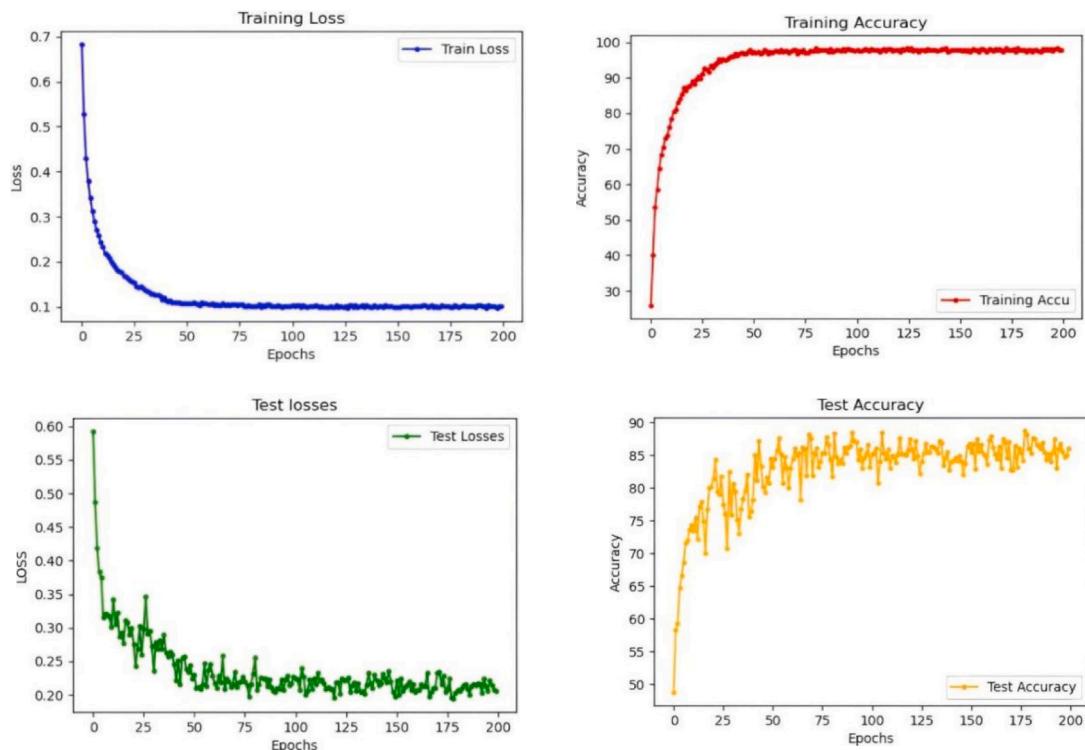
Using the trained model, a traceability analysis of these critical regulatory factors was conducted. The experimental results are shown in Table 7, and the traceability accuracy achieved is 99.84.

Moreover, as illustrated in the figure, the model successfully traced the key factors for three different levels of white spot fungus severity. This capability provides valuable insights for making informed adjustments in production processes, offering precise environment control guidance for enterprise operations.

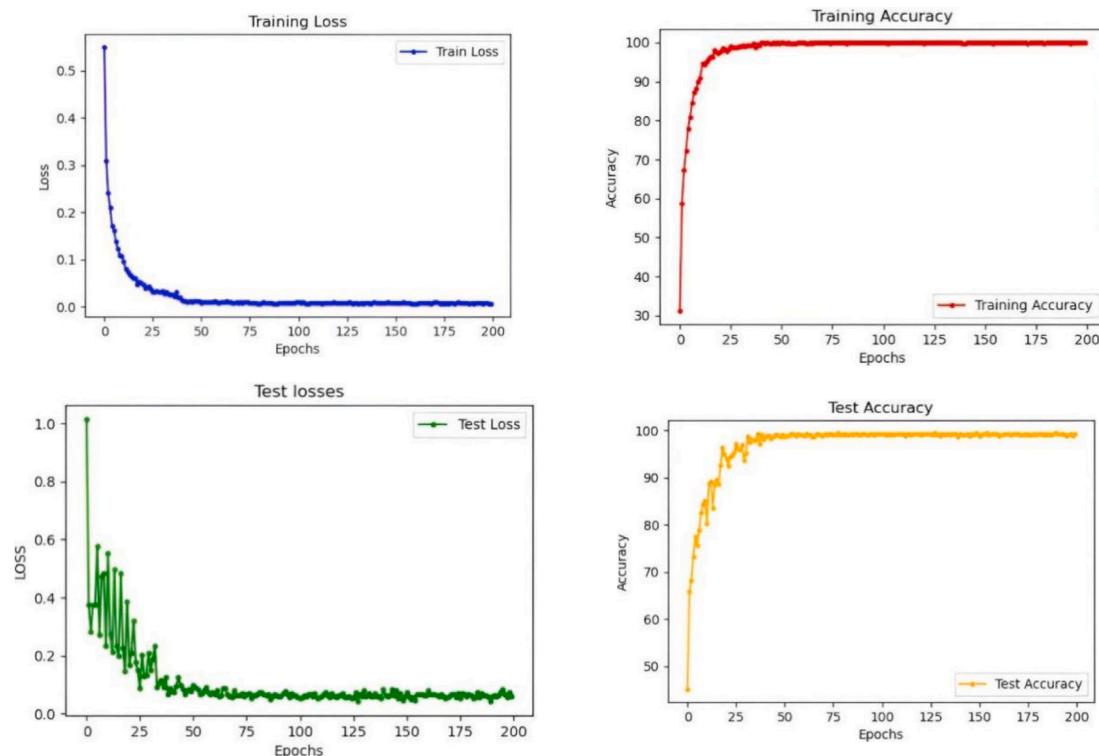
## 4. Application testing

To verify the network’s accuracy in real-world environments and assess its real-time performance and accessibility, an experimental platform was conducted to simulate characterization and environmental traceability tests for *Lyophyllum decastes*. The experimental workflow is shown in Fig. 13.

In the designed experiment, bottled *Lyophyllum decastes* with various characteristics, provided by Shandong Agriculture and Mushroom Industry, were used for detection. Three frames were sampled for each characteristic, with 16 bottles per frame, totaling 48 bottles for each type of *Ganoderma lucidum*. After statistical analysis, the samples were shuffled. Each mushroom cultivation unit is equipped with an NVIDIA Jetson Xavier NX, which has 8.0 GB of RAM and 128.0 GB of ROM, used for real-time monitoring of the quality characteristics and environmental factors of *Lyophyllum decastes*. The system recorded relevant quality metrics, recommended control methods, and QR code information, which were printed and affixed to the bottles. In order to verify the capability of the model in practical use, a test system that integrating



(a)

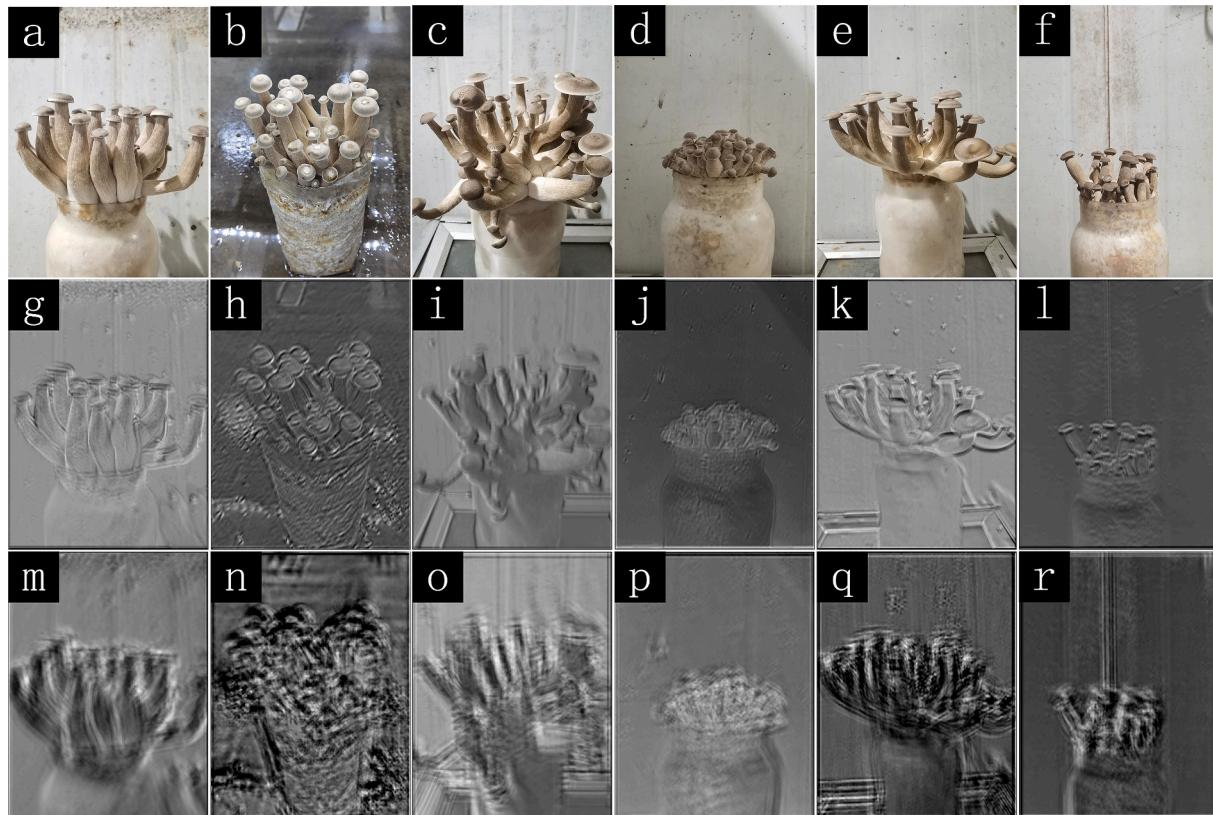


(b)

**Fig. 8.** Performances of ResNet-34 and DCRes-GAT on the *Lyophyllum decastes* dataset, (a) is the loss and accuracy of the ResNet34 network and (b) is the loss and accuracy of the DCRes-GAT network.



**Fig. 9.** Comparison of class activation mapping heatmaps for two networks: (a) shows the original image fed into the network, (b) shows the heatmap learned by the ResNet-34 network, and (c) shows the heatmap learned by the DCRes-GAT network.

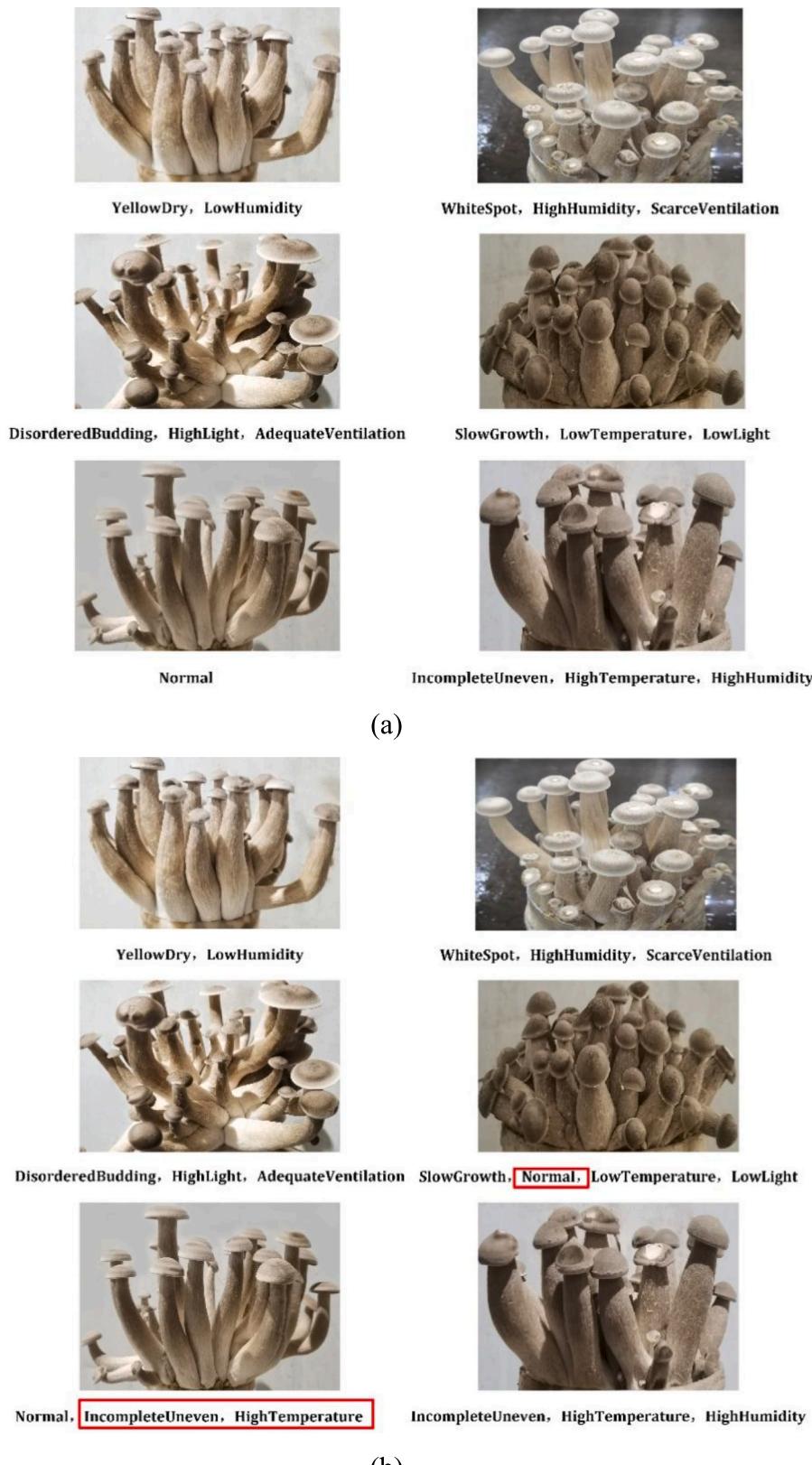


**Fig. 10.** Comparison of dilated convolution effects: (a-f) are original images, (g-l) are the effect images when the network uses ordinary convolution throughout, and (m-r) are the effect images when the last two layers of the network use dilated convolution.

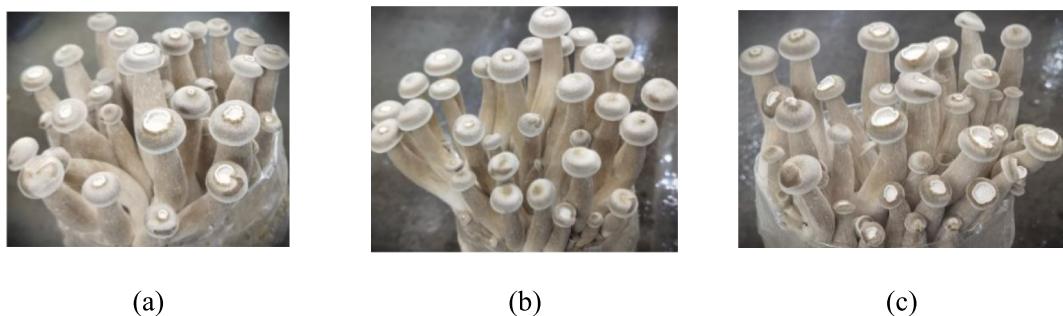
quality identification, critical factor traceability and defective mushroom elimination was developed. The accuracy of the system depends on whether the quality of *Lyophyllum decastes* can be correctly identified and the critical factor traceability can be carried out at the same time. To accommodate the single-bottle transportation requirement in factory production, a conveyor belt system was used for *Lyophyllum decastes* transportation. The direction and speed of the conveyor belt were controlled using an Altai motion control card, which adjusted motor pulse frequencies and signal levels. Additionally, a Delta robotic arm, equipped with visual positioning, was integrated to identify bottles predicted by the network to have "Slow Growth" or "Incomplete Unevenness" characteristics. These defective bottles were removed from the production line, ensuring only high-quality products were packaged, thereby maintaining economic efficiency and production continuity. The hardware parameters of the built physical experimental platform are as follows: the Delta robotic arm is from Xiaocong Robotics, model XCR-45-600; the motion control card is the Altai USB1020 four-axis controller; the conveyor belt has specifications of  $1\text{ m} \times 0.3\text{ m} \times 0.5\text{ m}$

; the stepper motor driver model is DM3522; and the camera used is the Intel RealSense D435i, which features a 20-megapixel RGB camera and a 3D sensor capable of providing a resolution of  $1280 \times 720$  at 30 frames per second.

For ease of observation during the experiment, the conveyor belt speed was set to 0.05 m/s, with a spacing of 0.2 m between each mushroom. Once the program was initiated, the QR code recognizer identified instances of "Slow Growth" and "Incomplete Unevenness" information, triggering the robotic arm to remove the corresponding bottled mushrooms. The software interface of the experimental platform and mushroom house design are illustrated in Fig. 14. At the same time, in order to verify the advantages of our model and test platform, A comparative test was conducted with two employees of Shandong Agriculture and Mushroom Industry factory on the evaluation of the same samples, the network and artificial prediction results within the mushroom house are summarized in Table 8. To evaluate the network's identification and critical environmental factor traceability ability performance based on the experimental results, the misjudgment rate



**Fig. 11.** Comparison of network identification performance, the results enclosed in the red box are misclassifications. (a) shows the identification performance of the DCRes-GAT network, (b) shows the identification performance of the ResNet-34 network. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



**Fig. 12.** Three different levels of *Lyophyllum decastes* “white-spot” appearance problem characterization, among them, (a) Larger percentage of “white-spot” with smaller caps is mainly caused by SV, (b) Smaller percentage of “white-spot” with smaller caps is mainly caused by HH, (c) Higher percentage of “white-spot” along with excessive limbs and thin branches is mainly caused by HT.

**Table 6**

The key factors of “white-spot” annotated by experts (only the factor with the greatest influence among the three factors in each category is taken as the traceable factor of this category).

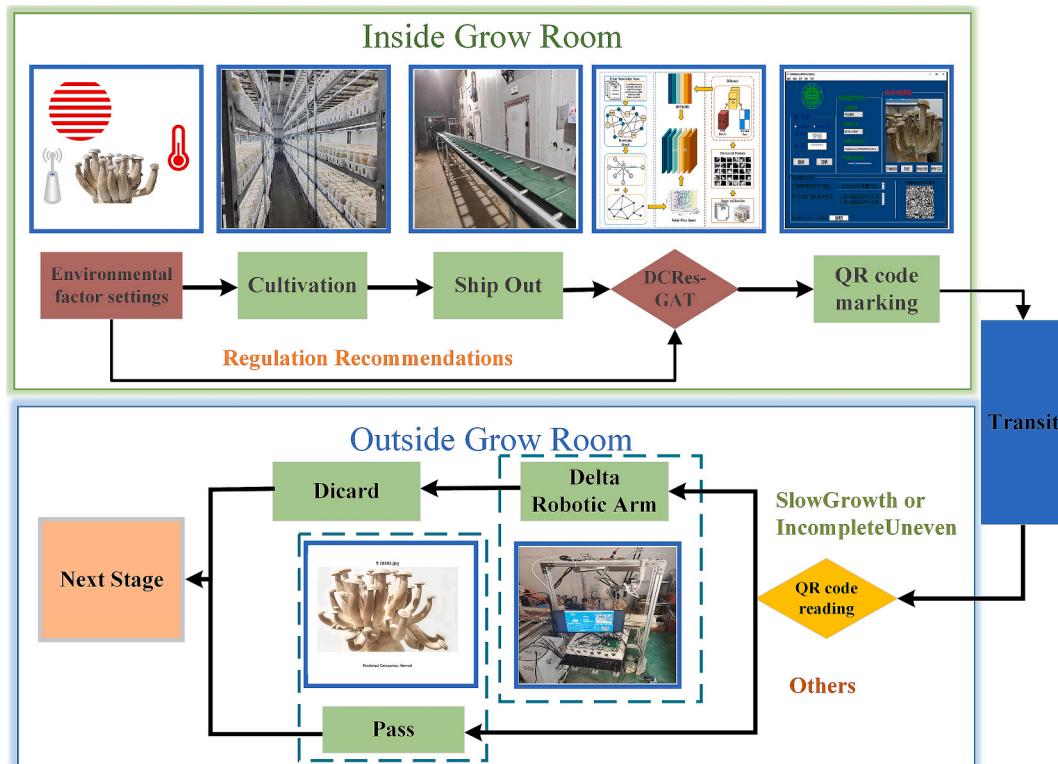
Factors	Critical factors calibrated by experts		
	SV	HH	HT
Numbers	1000	1200	238
Percentage	41 %	49 %	10 %

**Table 7**  
Model traceability accuracy.

Model	Accuracy	Precision	Recall	F1
DCRes-GAT	0.9984	0.9950	0.9922	0.9936

reflects how often incorrect judgments occur. It was calculated by determining the proportion of all judgments that were correct, across several categories or situations, and then subtracting this proportion from 1 to find the rate of incorrect judgments. The exclusion rates for categories labeled as “Slow Growth” and “Incomplete Uneven” were calculated based on the average frequency of occurrences associated with these two conditions. Specifically, the total instances identified as “Slow Growth” and those marked as “Incomplete Uneven” were added together. This sum was then divided by twice the total number of evaluated instances, to find the average proportion of cases that met either of these criteria for exclusion.

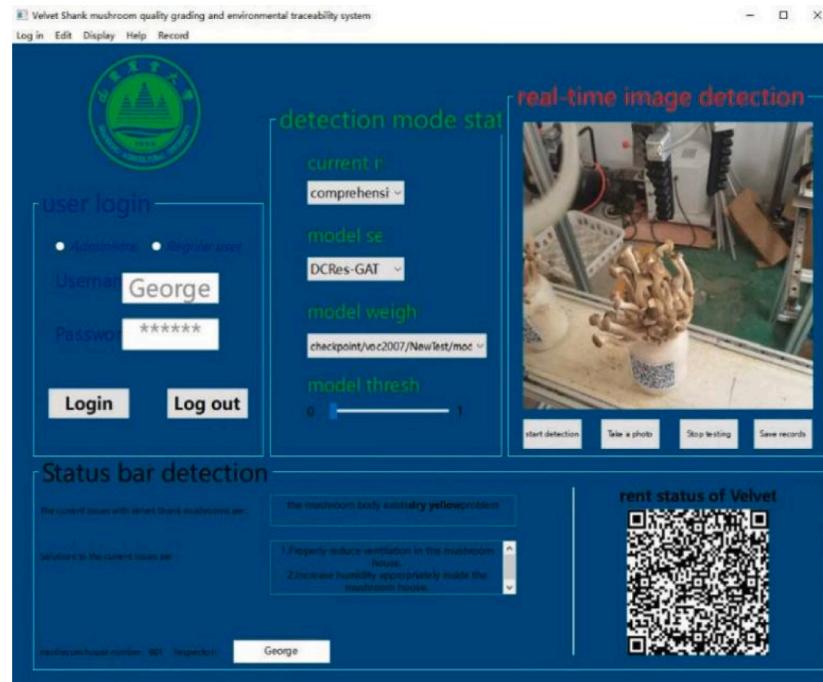
The automatic detection system developed in this study achieves real-time monitoring of the quality characteristics of bottled *Lyophyllum decastes* and defect removal, effectively replacing traditional manual detection methods, which significantly reduces dependence on human operation. This directly lowers labor costs and minimizes the risk of misjudgments caused by human error. Experimental data shows that the



**Fig. 13.** Experimental procedure diagram: In the grow room, the *Lyophyllum decastes* undergo environment control, cultivation, and ship out. They are then marked with a QR code through the proposed network before being transported out of the grow room. Outside the grow room, the QR codes are scanned to determine if any need to be discarded. If not, they proceed to the next stage.



(a)



(b)

**Fig. 14.** Experimental platform, (a) is structure of the platform and (b) is the interface of the software.

**Table 8**The network's identification performance on different appearance problems of *Lyophyllum decastes*.

Predicted results	Yellow Dry	White Spot	Disordered Budding	Slow Growth	Incomplete Uneven	Normal Growth	Misjudgment Rate
Total	48	48	48	48	48	48	
Employee W	46	47	48	42	40	41	8.33 %
Employee L	45	44	47	45	43	44	6.90 %
DCRes-GAT	48	48	46	47	48	48	1.04 %

misjudgment rate of the DCRes-GAT model is only 1.04 %, much lower than manual detection (employee misjudgment rates of 8.33 % and 6.90 %, respectively). This indicates that the system can not only identify quality issues more accurately but also maintain stable detection performance, improving the overall product quality control level. Its application is expected to bring significant economic benefits, as improvements in quality and optimized resource utilization during the production process can further enhance market competitiveness and increase enterprise profits.

In the experimental testing environment, the system meets the real-time detection requirements; however, in a high-throughput industrial environment with longer continuous working hours and larger data processing volumes, it is necessary to improve the performance of processors, robotic arms, and other hardware to meet the requirements of the factory production environment.

## 5. Conclusions

This paper introduces the DCRes-GAT method, which integrates NLP and CV for the appearance quality identification of *Lyophyllum decastes* and tracing its environmental impact factors. The method addresses the limitations of existing mushroom identification technologies, such as reliance on single identification approaches and lower network accuracy. By utilizing KG, the relationship between the characterization quality of *Lyophyllum decastes* and environmental factors is mapped to a graph node space, where GAT extract key features. The method incorporates an improved ResNet architecture, using dilated convolutions to expand the receptive field and CBAM modules to enhance the extraction of subtle image features. A dot product operation fuses the graph node features with visual features, enabling effective classification of *Lyophyllum decastes*. Experimental results demonstrate that the DCRes-GAT method achieves a fast convergence rate and a prediction accuracy of up to 99.45 %, outperforming previous improvements and existing networks. Additionally, in a typical experiment to verify the model's traceability capability, the model demonstrated its ability to accurately trace and identify critical influencing factors for different levels of appearance quality problems, achieving an accuracy of up to 99.84 %. Practical application experiments reveal an error rate of only 1.04 %, confirming the model's robustness and higher accuracy. Furthermore, defect removal experiments under simulated factory production conditions achieved a 100 % removal rate for defective mushrooms, indicating that the overall performance meets the requirements of industrial production. According to the research conducted by Nongfa Company, approximately 5 % of the total production of *Lyophyllum decastes* is affected by appearance quality issues caused by environmental factors, which significantly impacts the company's production efficiency. The method proposed in this paper provides a foundational approach for precise environmental regulation of *Lyophyllum decastes*, enabling improvements in both yield and quality, and enhancing the economic and social benefits of the company.

Although the dataset used in this study contains a large number of samples, its geographical origin is limited to a single region. This means that, while the model performs well in this specific geographical area, its performance and conclusions may not be directly generalized to data from other regions or globally. Future work will explore the relationship between *Lyophyllum decastes* characteristics and other environmental factors, such as pest and disease factors. In this study, although adjusting

the weights in the loss function partially alleviates the issue of sample imbalance, future work can improve the recognition capability of the model for minority classes by balancing the training dataset through oversampling the minority class or undersampling the majority class. This would further enhance the effectiveness of the proposed method in practical applications.

## CRediT authorship contribution statement

**Kai Zhou:** Writing – review & editing, Writing – original draft, Validation, Software, Methodology. **Junyuan Yu:** Methodology, Data curation. **Haotong Shi:** Writing – original draft, Software. **Rui Hou:** Writing – review & editing, Validation, Formal analysis. **Huarui Wu:** Data curation, Conceptualization. **Jialin Hou:** Validation, Funding acquisition, Formal analysis.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgments

This work was supported by the National Key Research and Development Program of China (2024YFD2000800), the Higher Education Youth Innovation Team Project of Shandong Province (2022KJ244), the Natural Science Foundation of Shandong Province (ZR2022QE103), the Key Research and Development of Shandong Province (2022CXGC010610).

## Data availability

Data will be made available on request.

## References

- Ba, D.M., Ssentongo, P., Beelman, R.B., Muscat, J., Gao, X., Richie, J.P., 2021. Higher mushroom consumption is associated with lower risk of cancer: a systematic review and meta-analysis of observational studies. *Adv. Nutrit.* 12, 1691–1704. <https://doi.org/10.1093/advances/nmab015>.
- Deng, J.W., Liu, Y.H., Xiao, X.Q., 2022. Deep-Learning-Based Wireless Visual Sensor System for Shiitake Mushroom Sorting. *Sensors.* 22 (12), 4606. <https://doi.org/10.3390/s22124606>.
- Garcia-Garcia, A., Orts-Escalona, S., Oprea, S., Villena-Martinez, V., Martinez-Gonzalez, P., Garcia-Rodriguez, J., 2018. A survey on deep learning techniques for image and video semantic segmentation. *Appl. Soft. Comput.* 70, 41–65. <https://doi.org/10.1016/j.asoc.2018.05.018>.
- Guan, L., Zhang, J., Geng, C., 2021. Diagnosis of fruit tree diseases and pests based on agricultural knowledge graph. *J. Phys.: Conf. Ser.* 1865, 042052. <https://doi.org/10.1088/1742-6596/1865/4/042052>.
- He, K.M., Zhang, X.Y., Ren, S.Q., Sun, J., 2016. Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. IEEE, New York, pp. 770–778. doi: 10.1109/CVPR.2016.90.
- Jafar, A., Lee, M., 2021. High-speed hyperparameter optimization for deep resnet models in image recognition. *Clust. Comput.* 26, 2605–2613. <https://doi.org/10.1007/s10586-021-03284-6>.
- Ji, S., Pan, S., Cambria, E., Marttinen, P., Yu, P.S., 2022. A survey on knowledgegraphs: Representation, acquisition, and applications. *IEEE Trans. Neural Netw. Learn. Syst.* 33, 494–514. <https://doi.org/10.1109/TNNLS.2021.3070843>.
- Ke, S.W., Ding, L.Q., Niu, X., Shan, H.J., Song, L.R., Xi, Y.L., Feng, J.H., Wei, S.L., Liang, Q.Q., 2023. Comparative transcriptome analysis on candidate genes associated with fruiting body growth and development in *Lyophyllum decastes*. *PeerJ* 11, e16288. <https://doi.org/10.7717/peerj.16288>.

- Lv, B.W., Wu, H.R., Chen, W.B., Chen, C., Miao, Y.S., Zhao, C.J., 2024. VEG-MMKG: Multimodal knowledge graph construction for vegetables based on pre-trained model extraction. Comput. Electron. Agric. 226, 109398. <https://doi.org/10.1016/j.compag.2024.109398>.
- Murali, E., Anuncia, S.M., 2022. An Ontology-based Knowledge Mining Model for Effective Exploitation of Agro Information. IETE J. Res. 69 (11), 7856–7873. <https://doi.org/10.1080/03772063.2022.2058629>.
- Mishra, R., Shridevi, S., 2024. Knowledge graph driven medicine recommendation system using graph neural networks on longitudinal medical records. Sci. Rep. 14, 25449. <https://doi.org/10.1038/s41598-024-75784-5>.
- Özbay, E., Özbay, F.A., Gharehchopogh, F.S., 2024. Visualization and classification of mushroom species with multi-feature fusion of metaheuristics-based convolutional neural network model. Appl. Soft. Comput. 164, 111936. <https://doi.org/10.1016/j.asoc.2024.111936>.
- Tao, K., Liu, J., Wang, Z., Yuan, J., Liu, L., Liu, X.M., 2024. ReYOLO-MSM: A novel evaluation method of mushroom stick for selective harvesting of shiitake mushroom sticks. Comput. Electron. Agric. 225, 109292. <https://doi.org/10.1016/j.compag.2024.109292>.
- Valem, L.P., Pedronette, D.C.G., Latecki, L.J., 2023. Graph Convolutional Networks based on manifold learning for semi-supervised image classification. Comput. vis. Image Underst. 227, 103618. <https://doi.org/10.1016/j.cviu.2022.103618>.
- Wang, P.X., Zhang, C., Wang, D.Q., Zhang, S.H., Wang, J., Wang, X.Z., Huang, L., 2023. Relation extraction for knowledge graph generation in the agriculture domain: a case Study on soybean pests and disease. Appl. Eng. Agric. 39, 215–224. <https://doi.org/10.13031/aea.15124>.
- Wu, Y.Q., Sun, Y.B., Zhang, S.Q., Liu, X., Zhou, K., Hou, J.L., 2022. A Size-Grading Method of Antler Mushrooms Using YOLOv5 and PSPNet. Agronomy 12, 2601. <https://doi.org/10.3390/agronomy12112601>.
- Yates, D., Blanchard, C., Clarke, A., Rehman, S., Islam, M.Z., Ford, R., Walsh, R., 2024. Combined location online weather data: easy-to-use targeted weather analysis for agriculture. Climatic Change 177, 138. <https://doi.org/10.1007/s10584-024-03793-4>.
- Yin, H., Yi, W.L., Hu, D.M., 2022. Computer vision and machine learning applied in the mushroom industry: A critical review. Comput. Electron. Agric. 198, 107015. <https://doi.org/10.1016/j.compag.2022.107015>.
- Zhu, D.J., Xie, L.Z., Chen, B.X., Tan, J.B., Deng, R.F., Zheng, Y.Z., Hu, Q., Mustafa, Rashed., Chen, W.S., Yi, S., Y, K.L., Andrew, W. H. 2023. Knowledge graph and deep learning based pest detection and identification system for fruit quality. Internet of Things. 21, 100649. doi: 10.1016/j.iot.2022.100649.
- Zhang, F.P., Xu, H., Yuan, Y., Huang, H.C., Wu, X.P., Zhang, J.L., Fu, J.S., 2022. *Lyophyllum decastes* fruiting body polysaccharide alleviates acute liver injury by activating the Nrf2 signaling pathway. Food Funct. 13, 2057–2067. <https://doi.org/10.1039/DFO001701B>.
- Zhang, G.P., Wang, Y.N., Qin, C.Q., Ye, S.M., Zhang, F.M., Linhardt, R.J., Zhang, A.Q., 2023. Structural characterization of an antioxidant polysaccharide isolated from the fruiting bodies of *Lyophyllum decastes*. J. Mol. Struct. 1285, 135507. <https://doi.org/10.1016/j.molstruc.2023.135507>.
- Zhao, L.C., Li, Q.Z., Chang, Q.R., Shang, J.L., Du, X., Liu, J.G., Dong, T.F., 2022. In-season crop type identification using optimal feature knowledge graph. ISPRS J. Photogramm. Remote Sens. 194, 250–266. <https://doi.org/10.1016/j.isprsjprs.2022.10.017>.
- Zhang, Y.R., Wang, D.W., Chen, Y.T., Liu, T.T., Zhang, S.S., Fan, H.X., Liu, H.C., Li, Y., 2021. Healthy function and high valued utilization of edible fungi. Food Sci. Hum. Wellness. 10 (4), 408–420. <https://doi.org/10.1016/j.fshw.2021.04.003>.