

Article

A Progressive Feature Learning Network for *Cordyceps sinensis* Image Recognition

Shangdong Liu ¹, Wenxiang Wu ¹, Haijun Chen ², Shuai You ³, Jiahuan Lu ³, Lin Mao ², Fan Zhang ⁴ and Yimu Ji ^{1,*}

¹ School of Computer Science, Nanjing University of Posts and Telecommunications, Nanjing 210023, China; lsd@njupt.edu.cn (S.L.); 1223045119@njupt.edu.cn (W.W.)

² Newlixon TechGroup.Co., Ltd., Nanjing 210012, China; chenhaijun@newlixon.com (H.C.); maolin@newlixon.com (L.M.)

³ School of Internet of Things, Nanjing University of Posts and Telecommunications, Nanjing 210023, China; youshuai_666@163.com (S.Y.); 2023070801@njupt.edu.cn (J.L.)

⁴ State Key Laboratory of Ocean Sensing, Ocean College, Zhejiang University, Zhoushan 316021, China; f.zhang@zju.edu.cn

* Correspondence: jiym@njupt.edu.cn; Tel.: +86-18-9189-6436

Abstract

Cordyceps sinensis (*C. sinensis*) is a valuable herbal medicine with wide-ranging applications. However, automating *C. sinensis* recognition is challenging due to the high morphological similarity and limited phenotypic variation among its subspecies. In this paper, we propose a novel approach called Progressive Feature Learning Network (PFL-Net) that mines multiple biological features to recognize different subspecies. Firstly, to comprehensively capture multi-scale discriminative features of *C. sinensis*, we propose the Spatial-aware Semantic Refinement Module (SSRM), which constructs discriminative feature groups by utilizing relative positions to model the intrinsic feature relations. Secondly, the Multi-scale Collaborative Perception Module (MCPM) avoids isolated biological features during modeling by establishing relations between different feature groups to enhance the recognition integrity of *C. sinensis*. Furthermore, to prevent the model from focusing on the same discriminative regions of *C. sinensis*, we propose a Channel Decouple (CD) loss that decouples features along the channel dimension, enhancing the diversity of *C. sinensis* discriminative features. In addition, we construct a *C. sinensis* dataset (CSD) to facilitate the application of biometric recognition, representing the first study focused on fine-grained *C. sinensis* recognition. Extensive experiments conducted on the CSD and three benchmark datasets validate the effectiveness of our proposed method, achieving a top-1 accuracy of 94.43% on the CSD dataset, which surpasses all existing approaches.

Keywords: *Cordyceps sinensis*; feature mining; progressive learning; image recognition



Academic Editor: Bogdan Smolka

Received: 7 September 2025

Revised: 21 October 2025

Accepted: 11 November 2025

Published: 20 November 2025

Citation: Liu, S.; Wu, W.; Chen, H.; You, S.; Lu, J.; Mao, L.; Zhang, F.; Ji, Y. A Progressive Feature Learning Network for *Cordyceps sinensis* Image Recognition. *Sensors* **2025**, *25*, 7082. <https://doi.org/10.3390/s25227082>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Cordyceps sinensis (*C. sinensis*) is a rare and expensive medicinal herb, abundant in bioactive substances such as nucleotides, amino acids, polysaccharides, flavonoids, and sterols. Cordycepin, the primary constituent of authentic *C. sinensis*, is an alkaloid with antioxidant properties and exhibits pharmacological effects, including immune enhancement, antifatigue, and antitumor activities [1–3]. In Figure 1a, we present authentic and counterfeit *C. sinensis* samples and six core discriminative regions. Recognition of authentic *C. sinensis* from counterfeits poses a significant challenge, resulting in a proliferation

of counterfeit products that pose a potential risk to public health. Existing recognition methods fall into two main categories. The first relies on expert experience, while the second employs multi-spectral and chemical analyses, which are both time-consuming and costly. With the advancement of computer vision [4–6], automatic image classification provides a new way to recognize *C. sinensis* accurately. To the best of our knowledge, this is the first study on image-based recognition of *C. sinensis*, which holds great significance for biomedicine.

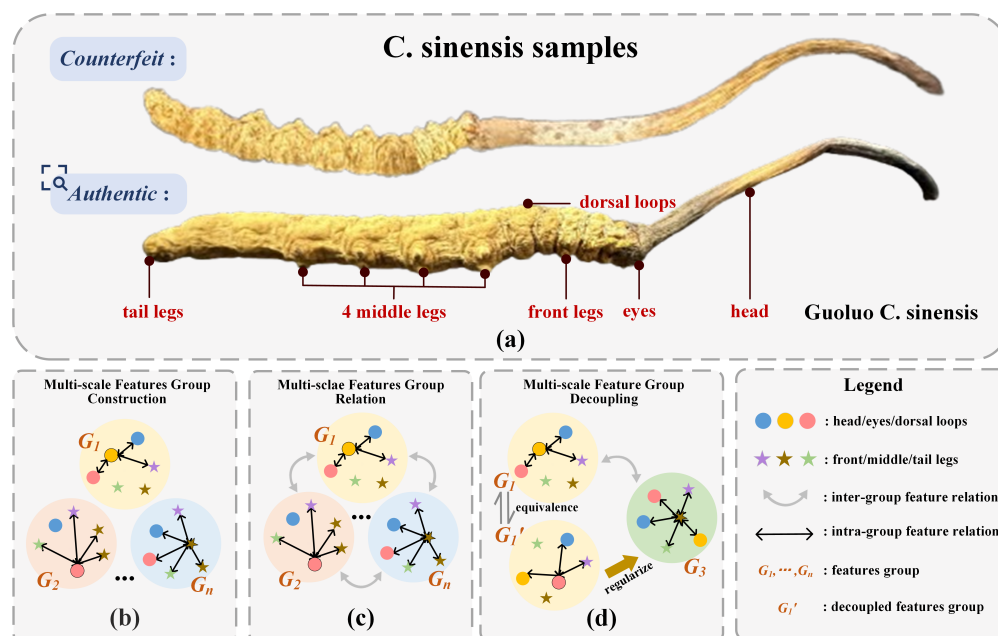


Figure 1. (a) illustrates authentic and counterfeit *C. sinensis* samples, highlighting six core discriminative features, while (b–d) detail the analysis process of multi-scale features of *C. sinensis*.

C. sinensis possesses six primary discriminative features: the head, eyes, dorsal loops, front legs, middle legs, and tail legs. The recognition of *C. sinensis* based on expert experience requires consideration of all six biological features. Constructing multi-scale feature groups to establish discriminative regions is essential. As shown in Figure 1b, where each feature group represents a distinct discriminative region. After establishing multi-scale feature groups following expert experience, these feature groups need to be related. In Figure 1c, we establish relations between feature groups to ensure the integration of discriminative regions. This manner fully utilizes multiple regions of *C. sinensis* and enhances feature diversity. Different discriminative feature groups may focus on the same region, which could lead to redundancy in feature learning and hinder the ability to capture diverse and unique characteristics completely. As shown in Figure 1d, G_1 and G_1' both represent the “Eyes–Head–Dorsal loops–Front legs” features, which means they belong to the same discriminative region. Therefore, it is necessary to guide the network to focus on the six core biometric features of *C. sinensis*.

To address this issue, we propose a Progressive Feature Learning Network (PFL-Net), including the Spatial-aware Semantic Refinement Module (SSRM), the Multi-scale Collaborative Perception Module (MCPM), and the Channel Decouple (CD) loss. The SSRM constructs multi-scale feature groups G_1 to G_n . G_1 represents the relation between local features such as “Eyes–Head–Dorsal loops–Front legs”. G_n represents the relation between larger scale features such as “Eyes–Head–Dorsal loops–Front legs–Middle legs–Tail legs”. The extraction of discriminative features is enhanced by modeling the relations among the six biometric features. The MCPM is designed to establish effective relations between different feature groups. This manner prevents information loss and ensures the integrity

of discriminative features. The CD operates on each feature group to further decouple features within the group, alleviating the feature coupling problem in the modeling process and guiding the network to focus on the six core biometric features of *C. sinensis*.

Notably, to facilitate the recognition of *C. sinensis*, we constructed the first *C. sinensis* dataset (CSD) with 17k images, significantly surpassing the scale of existing fine-grained public datasets. CSD includes 27 categories of *C. sinensis*, covering samples from six different regions and counterfeit. In addition, we collect images from multiple perspectives to overcome the limitations of single-angle views in capturing the *C. sinensis* features. In summary, the main contributions of our work are:

- (1) We propose PFL-Net, capable of accurately extracting and relating discriminative features. To the best of our knowledge, PFL-Net is the first study on recognizing *C. sinensis*.
- (2) The SSRM is designed to model spatial contextual, mining multi-scale discriminative features of *C. sinensis*. The MCPM relates the feature extracted at multiple scales to avoid the loss of *C. sinensis* features.
- (3) The CD loss decouples the features within the channel dimension and guides the network to focus on the *C. sinensis* features.
- (4) *C. sinensis* has significant medicinal value, we construct the first dataset for CSD. We perform extensive experiments on CSD and three fine-grained classification benchmarks demonstrating the superior performance of PFL-Net.

2. Related Work

2.1. General Image Classification

In recent years, CNN demonstrated significant potential in image classification tasks, giving rise to several representative CNN architectures, including Res2Net [7], EfficientNetV2 [8], RepVGG [9], ConvNext [10], and ConvNextV2 [11]. These models focused on classification accuracy and network scalability. Res2Net [7] improved optimization by using identity mappings for signal propagation and multi-scale convolutions to expand the receptive field. EfficientNetV2 [8] introduced a progressive learning method that improved training speed and accuracy. RepVGG [9] used reparameterization techniques to balance classification precision and performance. ConvNext [10] added independent downsampling layers to improve model stability in image classification. ConvNextV2 [11] optimized self-supervised learning for image classification by integrating neural architecture design with masked autoencoders. These methods performed well in image classification tasks but are unsuitable for fine-grained problems with large intra-class variance and small inter-class differences.

2.2. Fine-Grained Image Classification

Fine-grained image recognition aims to recognize subtle differences between objects within a supercategory, such as different subspecies of animals or cars. Existing methods are divided into three categories: bounding box annotation-based, local-based [12–14], and attention-based methods. First, bounding box annotation-based methods were fully supervised and required bounding box annotations in both the training and testing stages. They focused on learning discriminative mid-level features and performed well in bird species recognition and face verification. Second, local-based methods focused on learning local embeddings by predicting masked or erased regions in the image. For example, MaskCOV [12] used a covariance matrix to capture the mutual information between each quarter of randomly masked and shuffled image patches. Third, motivated by the success of attention in visual classification, TransFG [15] enhanced global discrimination by paying attention to important markers, alleviating the challenge of category variation. Notably,

existing methods demonstrate strong performance on public datasets, they are suboptimal for recognizing *C. sinensis*, which exhibits complex biological characteristics and structurally similar variants.

3. Method

3.1. Overview Architecture

In Figure 2a, the proposed Progressive Feature Learning Network (PFL-Net) is composed of a backbone ResNet50 [4], SSRM, and MCPM. Let F represent our backbone feature extractor, which consists of N stages. The output feature map from any intermediate stage is denoted as F^n , with $n = \{1, 2, \dots, N\}$.

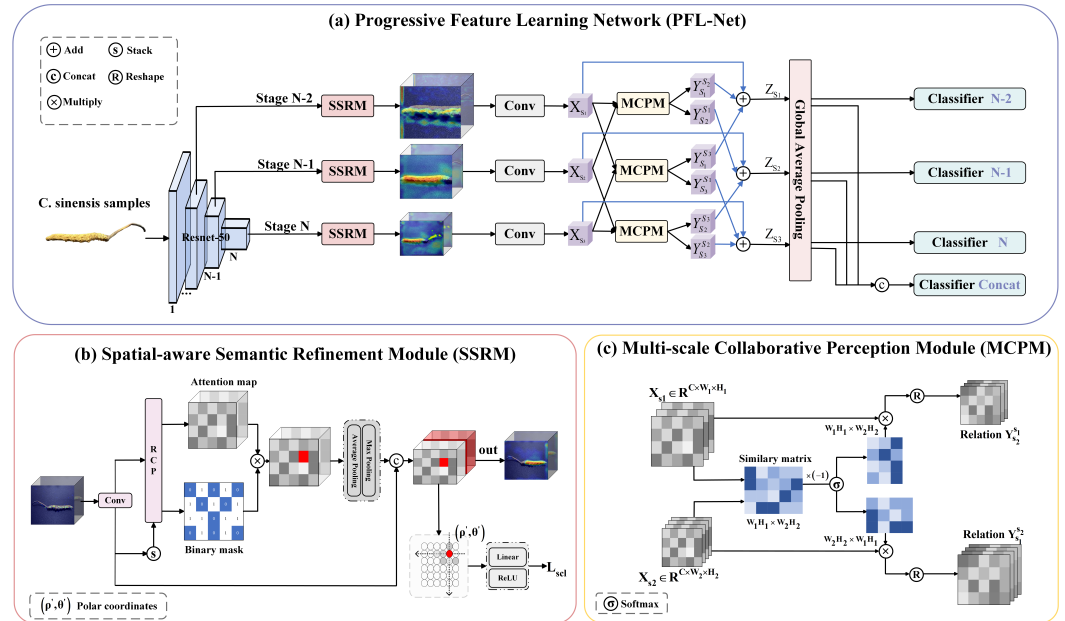


Figure 2. (a) The overall architecture of PFL-Net includes a feature extractor, the Spatial-aware Semantic Refinement Module(SSRM), and the Multi-scale Collaborative Perception Module(SSRM). “N” represents the number of stages in the backbone network. (b) shows the structure of the SSRM. (c) illustrates the MCPM.

The SSRM applies at each intermediate stage of the backbone to extract feature maps. We use the output F^n of the n -th intermediate stage of the backbone as the input of the SSRM module, denoted as $V^n = \text{SSRM}(F^n)$. The output vector V^n of each stage then passes through MCPM, denoted as $\{X^n, X^{n-1}\} = \text{MCPM}(V^n)$.

The classification module f_{cls}^n , consisting of two fully connected layers, predicts the probability distribution over classes for the n -th stage, denoted as $y^n = f_{cls}^n(X^n)$. The features extracted at the shallow level cannot fully demonstrate the biological features of *Cordyceps sinensis*, and the mining ability is limited. Therefore, we consider the final 3 stages: $n = n, n - 1, n - 2$. Lastly, we concatenate the outputs from the last three stages as follows [4,16]:

$$y^{cat} = f_{cls}^{cat}(\text{concat}[X^{n-2}, X^{n-1}, X^n]) \quad (1)$$

3.2. Spatial-Aware Semantic Refinement Module

CNN extracts structural features through convolution and gradually integrates features across layers to form a multi-level representation. Spatial structural information is crucial for *C. sinensis* image recognition. As shown in Figure 2b, we propose a Spatial-aware Semantic Refinement Module (SSRM) that builds feature groups by learning the spatial

context information of the object, thereby enhancing the feature representation capability of the backbone network.

For the input image I , SSRM directly applies to the feature map $f(I)$ extracted by the backbone. First, convolution is performed on the feature map to obtain $h(I) \in R^{C \times N \times N}$, which represents the spatial information of different features. The spatial relation of various features can be obtained by modeling the structural information between different parts of *C. sinensis*.

This paper uses polar coordinates to measure spatial relations between different regions. The traditional Cartesian coordinate system effectively measures linear distances but is sensitive to rotation and scale changes when modeling spatial relationships within an object [17]. In contrast, the polar coordinate system (r, θ) represents spatial layouts using radial distance and angle relative to a central point, offering inherent rotation invariance and scale adaptability. In biological image recognition, where structures such as the head, body, and tail of *Cordyceps sinensis* exhibit axial or radial organization, polar coordinates provide a more robust means of capturing geometric and topological relationships, ensuring stable modeling of structural coherence and spatial dependencies [18]. Given a reference region $R_0 = R_{x,y}$, indexed at (x, y) on the $N \times N$ plane, and a reference horizontal direction, the polar coordinates of the region $R_{i,j}$ can be written as $(r_{i,j}, \theta_{i,j})$:

$$r_{i,j} = \frac{1}{N} \sqrt{\frac{(x-i)^2 + (y-j)^2}{2}} \quad (2)$$

$$\theta_{i,j} = \frac{(\text{atan2}(y-j, x-i) + \pi)}{2\pi} \quad (3)$$

where $r_{i,j}$ measures the relative distance. $\theta_{i,j}$ measures the polar angle of $R_{i,j}$ relative to the horizontal direction. $\text{atan2}(\cdot)$ is a function that calculates the angle between two points. In this paper, we select the region with the maximum response $m(I)$ as the reference region:

$$R_0 = R_{x,y}, (x, y) = \arg \max_{1 \leq x, y \leq N} m'(I)_{i,j} \quad (4)$$

The SSRM utilizes polar coordinates to guide the module in learning and identifying the features of *C. sinensis*. Specifically, the SSRM module calculates the polar coordinates of the region $R_{i,j}$ by analyzing the discriminant information of the target region $R_{i,j}$ and the reference region R_0 . This process involves channel-wise concatenation of the feature map $h(I)$ with $h(I)_{x,y}$, followed by a fully connected layer to generate the predicted polar coordinates $(r'_{i,j}, \theta'_{i,j})$. The SSRM module models the spatial structure between different parts of an object and integrates the object region mask $m'(I)$ learned from the backbone network.

The SSRM first measures the relative distance differences between all regions and the object:

$$\mathcal{L}_d = \sum_{I \in \mathcal{I}} \sqrt{\frac{\sum_{1 \leq i, j \leq N} m'(I)_{i,j} (r'_{i,j} - r_{i,j})^2}{\sum m'(I)}} \quad (5)$$

The second measures the angular difference between regions within the object. The structural information of an object should be rotationally invariant and robust to various appearances and poses. Therefore, we calculate the angle loss $\mathcal{L}_a$ using the standard deviation of the difference between the predicted and true polar angles, as follows:

$$\mathcal{L}_a = \sum_{I \in \mathcal{I}} \sqrt{\frac{\sum_{1 \leq i, j \leq N} m'(I)_{i,j} (\theta_{\Delta_{i,j}} - \bar{\theta}_{\Delta})^2}{\sum m'(I)}} \quad (6)$$

$$\theta_{\Delta_{i,j}} = \begin{cases} \theta'_{i,j} - \theta_{i,j}, & \text{if } \theta'_{i,j} - \theta_{i,j} \geq 0 \\ 1 + \theta'_{i,j} - \theta_{i,j}, & \text{otherwise,} \end{cases} \quad (7)$$

$$\bar{\theta}_{\Delta} = \frac{1}{\sum m'(I)} \sum_{1 \leq i,j \leq N} m'(I)_{i,j} \theta_{\Delta_{i,j}} \quad (8)$$

where $\bar{\theta}_{\Delta}$ represents the average difference between the predicted polar angle and the actual polar angle. SSRM models the relative structure between object parts. During polar coordinate regression, the predicted semantic mask $m'(I)$ filters out irrelevant visual information outside the main object. In general, the loss function of SSRM can be expressed as:

$$\mathcal{L}_{ssrm} = \mathcal{L}_d + \mathcal{L}_a \quad (9)$$

The polar-coordinate-based losses supervise the network to learn structural consistency between predicted and ground-truth spatial relations. Specifically, the distance loss $\mathcal{L}_d$ penalizes deviations in the relative radial distances between regions, ensuring scale consistency. The angular loss $\mathcal{L}_a$ constrains the predicted angular relations to maintain rotational invariance and correct part ordering. Jointly minimizing $\mathcal{L}_d + \mathcal{L}_a$ enforces the backbone to encode spatial geometry of *C. sinensis* parts in a biologically meaningful manner, improving robustness to pose and orientation variations. Through SSRM, the backbone learns and recognizes the structural features of the object, enabling the backbone network to model the spatial dependencies between the parts of the object and mine features.

3.3. Multi-Scale Collaborative Perception Module

Given the importance of learning discriminative and diverse features in Fine-grained Image Classification (FGIC) [18–21], we propose the Multi-scale Collaborative Perception Module (MCPM). The module enhances the expressiveness of local features by aggregating complementary information from different scales, improving feature discriminability and diversity.

Figure 2c briefly illustrates the structure of the MCPM. To demonstrate the Effectiveness of the approach, we denote two different scale feature representations as $X_{s_1} \in R^{C \times W_1 H_1}$ and $X_{s_2} \in R^{C \times W_2 H_2}$. The subscript s_i indicates that X_{s_i} focuses on the i th part of the object. We regard the feature vector at each spatial location in the channel dimension as a pixel, as follows:

$$px(X, i) = (X_{1,i}, \dots, X_{C,i})^T \quad (10)$$

where px stands for pixels. We first compute the similarity between the pixels in X_{s_1} and those in X_{s_2} :

$$M = f(X_{p_1}, X_{p_2}), \quad f(X, Y) = X^T Y \quad (11)$$

We use the inner product to compute the similarity. The $M_{i,j}$ represents the similarity between the i th pixel of X_{s_1} and the j th pixel of X_{s_2} . The lower the similarity between two pixels, the greater their complementarity. Therefore, we use M as the complementarity matrix. We then normalize M along both the row and column directions, as follows:

$$\begin{aligned} A_{s_1}^{s_2} &= -\mathbf{\alpha}(M^T) \in [0, 1] \\ A_{s_2}^{s_1} &= -\mathbf{\alpha}(M) \in [0, 1] \end{aligned} \quad (12)$$

where $\mathbf{\alpha}$ is softmax, the operation is applied column-wise. This allows us to obtain the complementary information:

$$\begin{aligned} Y_{s_1}^{s_2} &= X_{s_2} A_{s_1}^{s_2} \in R^{C \times W_1 H_1} \\ Y_{s_2}^{s_1} &= X_{s_1} A_{s_2}^{s_1} \in R^{C \times W_2 H_2} \end{aligned} \quad (13)$$

where $Y_{s_j}^{s_i}$ represents the complementary information of X_{s_i} with respect to X_{s_j} . Each pixel of $Y_{s_2}^{s_1}$ can be written as:

$$px(Y_{s_1}^{s_2}, i) = \sum_{j \in [1, W_2 H_2]} (A_{s_1}^{s_2})_{i,j} \times px(X_{s_2}, j) \quad (14)$$

Each pixel in $Y_{s_1}^{s_2}$ uses all pixels of X_{s_2} as a reference. The higher the complementarity between $px(X_{s_1}, i)$ and $px(X_{s_2}, j)$, the greater the contribution of $px(X_{s_2}, j)$ to $px(Y_{s_1}^{s_2}, i)$. Thus, each pixel within these scale features can capture semantically complementary information from other pixels. In the normal case, given a set of part-specific features $S = \{X_{s_1}, X_{s_2}, X_{s_3}, \dots, X_{s_n}\}$, the complementary information of X_{s_i} is:

$$Y_{s_i} = \sum_{X_{s_j} \in P \wedge i \neq j} Y_{s_i}^{s_j} \quad (15)$$

The $Y_{s_1}^{s_2}$ can be obtained by applying X_{s_i} and X_{s_j} to Equations (13), (14) and (16). We can compute both $Y_{s_i}^{s_j}$ and $Y_{s_j}^{s_i}$ simultaneously. This results in the enhanced object features:

$$Z_{s_i} = X_{s_i} + \gamma \times Y_{s_i} \quad (16)$$

where γ is a hyperparameter controlling the diversification degree, default set to 2.

3.4. Loss Function

During feature learning, excessive focus on the same discriminative regions can occur. To address this, we propose a Channel Decouple (CD) loss, which decouples features and mitigates feature coupling. After inputting an image into the network, we extract the feature map, denoted as $F \in R^{C \times W \times H}$, with height H , width W , and number of channels C . We need to set the value of C equal to $c \times k$, where c and k indicate the number of classes in a dataset and the number of feature channels used to represent each class. The n th vector feature channel of F is represented as $F_n \in R^{WH}$, $n = 1, 2, \dots, N$. Each channel matrix of F is reshaped into a vector of size WH . The grouped feature channels corresponding to the i th class are represented as $F_i \in R^{k \times WH}$, $i = 0, 1, \dots, c - 1$. This can be expressed as:

$$F_i = F_{i \times k + 1}, F_{i \times k + 2}, \dots, F_{i \times k + k} \quad (17)$$

The feature group $F = \{F_0, F_1, \dots, F_{c-1}\}$ is processed through two parallel streams in the network, each designed with a distinct sub-loss tailored for two different objectives. In the cross-entropy stream, F is treated as the input to a fully connected layer with the traditional CE [22]. The CE encourages the network to extract informative features focused primarily on global discriminative regions. On the other hand, the CD stream supervises the network to highlight different local discriminative regions. A specific number of grouped feature channels is representative of each class. The discriminative component enforces class alignment among the feature channels, ensuring that each feature channel corresponding to a specific class has sufficient discriminative power. $\mathcal{L}_{CD}$ can be expressed as:

$$\mathcal{L}_{CD}(F) = \mathcal{L}_{CE}(y, \frac{1}{\sum_{i=0}^{c-1} e^{g(F_i)}} [e^{g(F_0)}, \dots, e^{g(F_{c-1})}]^T) \quad (18)$$

where $g(\cdot)$ is defined as:

$$g(F_i) = \frac{1}{WH} \sum_{k=1}^{WH} \max_{j=1,2,\dots,\tilde{c}} [M_i \cdot F_{i,j,k}] \quad (19)$$

where M_i is a random mask between 0 and 1. The cross-entropy (CE) and Channel Decouple (CD) losses are optimized jointly in a parallel manner. The backbone features are simultaneously fed into two loss branches: one for global discrimination (CE stream) and one for local feature diversification (CD stream). During training, the gradients from both loss functions are backpropagated to the shared feature extractor, with the total loss formulated as

$$Loss(F) = \mathcal{L}_{CE}(F) + \mu \times \mathcal{L}_{CD}(F) \quad (20)$$

The weighting coefficient μ controls the relative influence of the CD loss. Empirically, we set $\mu = 0.3$ to balance global classification accuracy and feature diversity. This joint optimization ensures that the model simultaneously enhances discriminability and suppresses channel redundancy.

4. Data Collection and Construction

4.1. Material Preparation

We purchased and collected 654 *Cordyceps sinensis* (*C. sinensis*) samples, covering six major production regions and including counterfeit samples. The pharmacological efficacy of *C. sinensis* is strongly influenced by its place of origin, with the six major production regions being Yushu, Guoluo, Haixi, Hainan, Haibei, and Huangnan. As shown in Figure 3, we present two sample images for each specification of *C. sinensis* from each production area. Counterfeit samples refer to artificially synthesized *C. sinensis*, which have no medicinal value and may harm health. Each region's *C. sinensis* samples vary in size, defined by the number of specimens per 500 g. Smaller sizes correspond to larger individual specimens, which generally have higher medicinal value. We then proceeded to capture images of these samples.

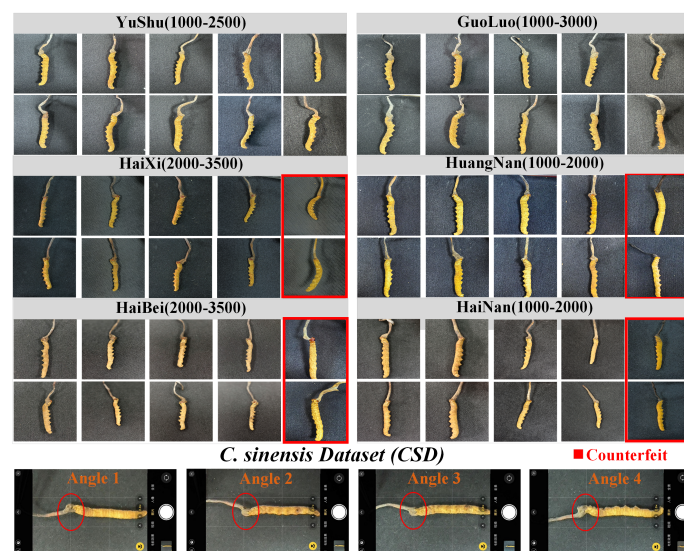


Figure 3. The *C. sinensis* dataset consists of *Cordyceps sinensis* samples from six major production regions, along with counterfeit samples. Dataset shooting rules are shown on the last line.

4.2. Data Collection and Annotation

Images are collected using eight smart devices, including both Android and iOS devices, covering a wide range of popular devices. The images of *C. sinensis* were taken from four angles: the back, the foot, the left side, and the right side. As shown in Figure 3 the last row. To ensure data consistency and minimize background interference, all images were captured using the camera's nine-grid layout, with the *C. sinensis* positioned within the central three grids and the eyes aligned along the first horizontal line. A ring light was

placed above the specimens to optimize imaging conditions and capture fine details, while a standardized black background was used to reduce background noise. The final dataset comprises over 17k images. Table 1 presents the number of *C. sinensis* images collected for each specification from different production regions.

Table 1. Distribution of *C. sinensis* products by origin, specification (number of pieces per 500 g), and number of images.

Origin	Specification	Image	Origin	Specification	Image
Yushu	1000	636	Guoluo	1000	600
	1200	626		1200	640
	1500	620		1500	640
	2000	568		2000	592
	2500	616		3000	576
Haixi	2000	636	Huangnan	1000	616
	2500	640		1200	584
	3000	616		1500	636
	3500	616		2000	640
Haibei	2000	616	Hainan	1000	640
	2500	640		1200	624
	3000	636		1500	600
	3500	640		2000	640
Counterfeit	/	3167			

To ensure reproducibility, the CSD dataset was randomly divided into five balanced subsets using a fixed random seed (seed = 1024). Each subset preserves the class distribution across all 27 categories. During 5-fold cross-validation, four subsets were used for training and one for testing in each iteration, ensuring no overlap between training and test samples. This partitioning principle guarantees consistency across repeated experiments.

Environmental Considerations and Optical Stability

Although the dataset was collected under controlled illumination using a ring light and standardized background, the quality of captured images may still be influenced by subtle environmental factors such as air turbulence, humidity, and temperature fluctuations during shooting. These conditions can induce wavefront aberrations and minor defocus effects that alter the sharpness and spatial coherence of fine-grained texture details. Previous optical studies [23] have demonstrated that turbulent environments can distort the wavefront propagation of light, leading to random phase errors and reduced contrast in microscopic structures. While our imaging setup minimizes such disturbances, potential optical aberrations may still affect the precision of spatial feature extraction, especially for thin or reflective specimens of *C. sinensis*. In future work, incorporating adaptive optics or turbulence-aware image correction techniques could further improve the robustness of fine-grained feature acquisition under non-ideal environmental conditions.

5. Experiments

5.1. Datasets and Settings

Data. In the CSD, we use 70% of the images for training and the remaining 30% for testing, consistent with the public datasets. To verify the generalization ability of our proposed PFL-Net, we conducted comprehensive experiments on three public fine-grained recognition benchmarks: CUB-200-2011 (CUB) [24], Stanford Cars (CAR) [25], and FGVC-Aircraft (AIR) [26]. In Table 2, we followed the standard procedure for splitting the training and testing images and used top-1 accuracy as the evaluation metric.

Table 2. Overview of the public datasets used in our experiments, including the number of classes, training, and test samples for each dataset.

Dataset	Classes	Training	Testing
CUB-200-2011 [24]	200	5994	5794
Stanford Cars [25]	196	8144	8041
FGVC-Aircraft [26]	100	6667	3333

Implementation Details. We used ResNet50 [4] as backbone networks. In the experiment, we only used image labels for supervised training, random cropping, horizontal flipping augmentation during training, and center cropping during testing. The input image size is 448×448 . The SGD optimizer is used with a momentum of 0.9 and weight decay 5×10^{-4} . The learning rate of the backbone layer is 0.0002, and the learning rate of the new layer is 0.002, both adjusted according to the cosine annealing strategy [27]. The learning rate of the auxiliary classifier is kept constant at 0.01. The model is trained for 200 rounds and is trained end-to-end on Nvidia RTX 3090 (Nvidia, Santa Clara, CA, USA) based on PyTorch 2.1.1.

5.2. Comparison with State of the Arts

Table 3 presents the experimental results of the comparative evaluation on the CSD, CUB-200-2011, Stanford Cars, and FGVC-Aircraft. We proposed that PFL-Net consistently outperforms state-of-the-art methods on both the CSD dataset and the three widely used benchmarks. The enhancement in performance is due to the design of our network and the efficiency of its components.

Table 3. The top-1 accuracy (%) comparison with state-of-the-art methods on the CSD, CUB-200-2011, Stanford Cars, and FGVC-Aircraft datasets is presented. Bold values indicate the best results, and underlined values represent the second-best results.

Method	Venue	<i>C. sinensis</i> Dataset	CUB-200-2011	Stanford Cars	FGVC-Aircraft
ISDA [28]	NeurIPS19	81.4	85.3	91.7	93.2
C2-Net [29]	AAAI24	88.5	84.6	-	88.9
Bi-FRN [30]	AAAI23	89.9	85.4	-	88.4
LearnableISDA [31]	TIP24	90.2	86.7	92.7	94.3
iSICE [32]	CVPR23	90.2	85.9	93.5	92.7
AP-CNN [33]	TIP21	90.5	87.2	92.2	93.6
ACNet [34]	CVPR20	90.6	88.1	92.4	94.6
API-Net [35]	AAAI20	91.0	87.7	93.0	94.8
S3Ns [36]	ICCV19	91.2	88.5	92.8	94.7
PMG [37]	ECCV20	92.1	88.9	92.8	95.0
DCL [38]	CVPR19	92.8	87.8	93.0	94.5
ViT [39]	ICLR21	93.1	90.3	94.2	94.8
P2P-Net [40]	CVPR22	93.2	90.2	94.9	94.2
TransFG [15]	AAAI22	<u>93.7</u>	91.7	<u>94.8</u>	-
PFL-Net (our)	-	94.4	<u>91.2</u>	94.9	95.1

PFL-Net surpasses methods that leverage implicit data augmentation, such as DCL [38], ISDA [28], and LearnableISDA [31], demonstrating its superior ability to capture fine-grained details. In addition, PFL-Net outperforms attention-based methods like S3Ns [36], ACNet [34], AP-CNN [33], and P2P-Net [40]. While PMG [37] achieved promising results by aggregating images, PFL-Net further enhances the learning of inter-class differences by mining multi-scale features and fusing information across scales via the MCPM. PFL-Net significantly outperforms PMG [37] in classification accuracy. Compared to API-Net [35], PFL-Net achieves superior classification performance without the need to construct specific image pairs. Furthermore, PFL-Net outperforms Bi-FRN [30] and C2-Net [29] by eliminating the need to construct support-query sample pairs. It places

greater emphasis on the relationships between features, resulting in generally superior classification performance.

5.3. Ablation Studies

To evaluate the effectiveness of the key components in our proposed approach, we conducted ablation experiments on CSD and CUB [24]. The results are shown in Table 4. The introduction of SSRM improves accuracy by +1.16% on CSD, demonstrating the benefit of spatial context modeling. MCPM further enhances accuracy by +1.39% through cross-scale feature collaboration. Finally, incorporating the CD loss provides an additional +1.39% gain by encouraging channel diversity and reducing redundant focus. Overall, PFL-Net achieves a total improvement of +2.79% over the baseline, validating the effectiveness of each component.

Specifically, as shown in Figure 4, SSRM provides a spatial context modeling manner, which helps to mine the discriminative features of *C. sinensis*. In addition, MCPM can associate the discriminative regions of *C. sinensis* to avoid information loss. The overall performance can be further improved owing to the complementary nature of SSRM and MCPM. Furthermore, we use CD loss to decouple features and guide the network to focus on the core discriminative features of *C. sinensis*. The results show that our Progressive Feature Learning Network can fully identify each feature region of *C. sinensis*.

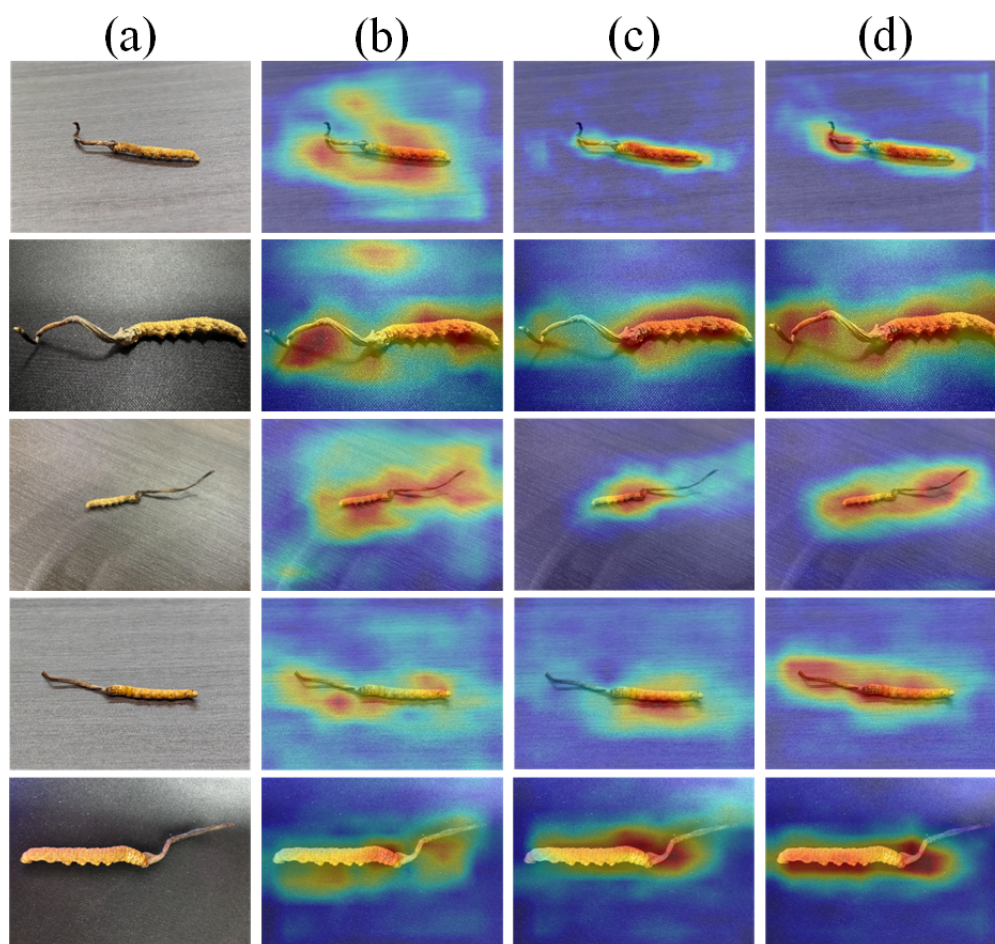


Figure 4. Visualization of ablation experiments. (a) input, (b) Baseline, (c) Baseline + SSRM + MCPM, and (d) PFL-Net.

Table 4. Ablation studies on the CSD and CUB (using original images) are conducted with the baseline (ResNet50) to evaluate the impact of different modules.

Index	Component				Accuracy (%)	
	Baseline	SSRM	MCPM	CD	CSD	CUB
0	✓				91.64	88.88
1	✓	✓			92.80	89.02
2	✓		✓		91.87	89.93
3	✓			✓	93.03	89.41
4	✓	✓	✓		93.34	90.95
5	✓	✓		✓	94.19	90.55
6	✓		✓	✓	93.75	90.41
7	✓	✓	✓	✓	94.43	91.26

5.4. Visualizations

Class Activation Map—We further apply Grad-CAM [41] to the final convolutional layer to provide intuitive visualizations. Figure 5 displays the activation maps for four datasets. Compared to the baseline model, PFL-Net focuses more on discriminative regions of the target object, such as the body of an airplane, the main contours of a car, the head of a bird, and the morphological features of *C. sinensis*. In addition, PFL-Net exhibits significantly lower activation in background areas, demonstrating its effectiveness in suppressing noise and irrelevant information. Visualizations across multiple datasets reveal that PFL-Net effectively extracts more category-specific regions, creating clearer boundaries between the object and background.

Parts Location—The discriminative regions identified by our PFL-Net are shown in Figure 6. Columns 1 to 3 display bounding boxes for only the first two parts, while columns 4 and 5 show all four primary discriminative regions. As illustrated, the first two regions exhibit notable similarities: for aircraft, they are the wings and tail; for birds, the head and body; for cars, the front and roof; for *C. sinensis*, they are the head and eyes. PFL-Net is capable of accurately and efficiently recognizing these key discriminative regions, demonstrating high robustness.

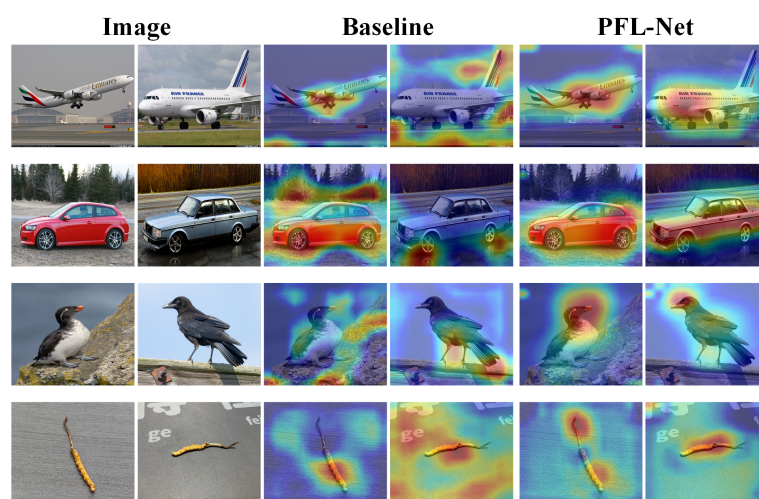


Figure 5. Class Activation Map on four datasets.

Feature Visualization—To intuitively evaluate the separability of features learned by the proposed PFL-Net, we employ the t-distributed Stochastic Neighbor Embedding (t-SNE) [42–44] to project high-dimensional features into a two-dimensional space for visualization, as shown in Figure 7. t-SNE is a nonlinear dimensionality reduction method that preserves the local neighborhood structure of data by modeling pairwise similarities between samples in both high-dimensional and low-dimensional spaces. Given two sam-

ples x_i and x_j in the high-dimensional feature space, their similarity is modeled by a conditional probability:

$$p_{ji} = \frac{\exp(-\|x_i - x_j\|^2/2\sigma_i^2)}{\sum_{k \neq i} \exp(-\|x_i - x_k\|^2/2\sigma_i^2)}, \quad (21)$$

where σ_i controls the variance of the Gaussian kernel centered at x_i . In the low-dimensional embedding, the similarity between points y_i and y_j is defined as:

$$q_{ij} = \frac{(1 + \|y_i - y_j\|^2)^{-1}}{\sum_{k \neq i} (1 + \|y_i - y_k\|^2)^{-1}}. \quad (22)$$

t-SNE minimizes the Kullback–Leibler (KL) divergence between these two distributions:

$$\mathcal{L}_{t\text{-SNE}} = \sum_i \sum_j p_{ij} \log \frac{p_{ij}}{q_{ij}}, \quad (23)$$

ensuring that nearby points in the high-dimensional space remain close in the low-dimensional visualization, while distant points are pushed apart. This results in an embedding that effectively preserves local similarities and reveals cluster structures. As illustrated in Figure 7, the features extracted by PFL-Net form more compact intra-class clusters and exhibit clearer inter-class boundaries compared to the baseline, indicating superior discriminative feature representation and robustness in fine-grained classification tasks.

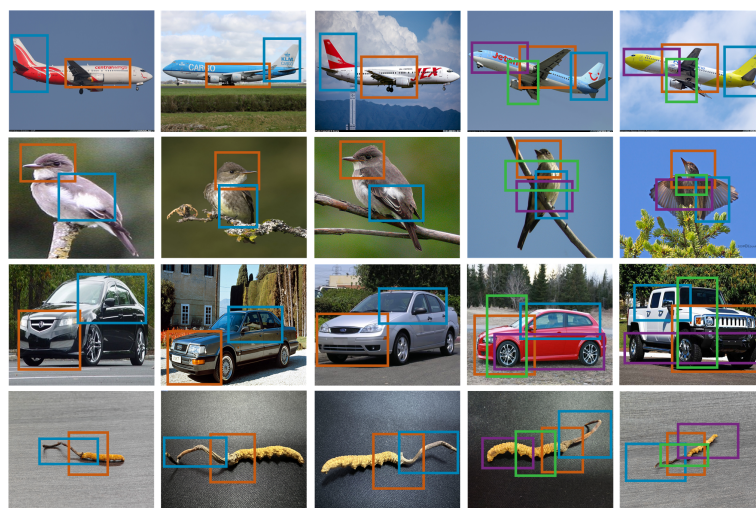


Figure 6. Discriminative regions detected by our PFL-Net.

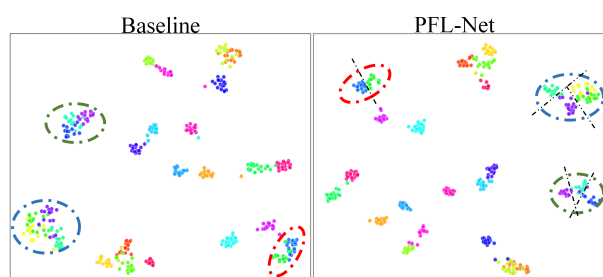


Figure 7. A t-SNE plot of learned representations on CSD.

Confusion Matrix—As shown in Figure 8, the left matrix illustrates the confusion matrix of the baseline model, where the diagonal elements are scattered and many off-diagonal entries exhibit noticeable values, indicating frequent misclassifications among visually similar subspecies. This suggests that the baseline model struggles to learn discriminative

representations for categories with subtle morphological differences. In contrast, the right panel presents the confusion matrix of PFL-Net, where nearly all the values are concentrated along the main diagonal, and off-diagonal elements are close to zero. This demonstrates that PFL-Net can accurately distinguish among all 27 *C. sinensis* categories, including those with highly similar appearances. The improved performance results from the integration of the Spatial-aware Semantic Refinement Module (SSRM) and Multi-scale Collaborative Perception Module (MCPM), which enable the model to better capture fine-grained spatial relationships and maintain feature consistency across scales. Overall, PFL-Net achieves clearer classification boundaries, lower confusion rates, and stronger generalization across subspecies.

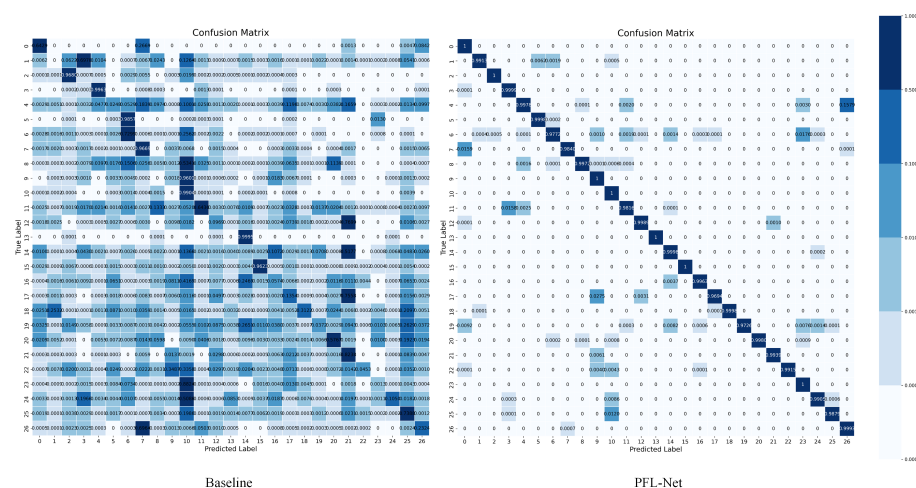


Figure 8. Confusion matrix on CSD

6. Conclusions

This paper is the first to study the recognition task of *Cordyceps sinensis*, constructing a dedicated dataset and proposing a Progressive Feature Learning Network (PFL-Net) to promote the development of recognition in the biomedical field. By effectively combining the Spatial-aware Semantic Refinement Module (SSRM) and Multi-scale Collaborative Perception Module (MCPM), PFL-Net enhances the relation between local and global features, improving its ability to recognize complex structures. The Channel Decoupling (CD) loss further optimizes the network to extract diverse features. Experiments on *Cordyceps sinensis* dataset and three fine-grained classification benchmarks indicate that PFL-Net achieves state-of-the-art performance, and visualization results demonstrate its effectiveness. In the future, PFL-Net has the potential to be applied to other biological recognition scenarios.

Limitations and Future Work

Although PFL-Net achieves promising results, several limitations remain. First, the model relies on high-quality annotations and balanced category distributions; significant noise or imbalance in the dataset may affect its stability. Second, while the Spatial-aware Semantic Refinement Module (SSRM) provides spatial robustness, extreme variations in illumination or occlusion may still degrade feature consistency. Third, the multi-scale collaborative design slightly increases computational cost compared to lightweight architectures. Future work will focus on optimizing model efficiency and extending PFL-Net to other biological recognition domains with more diverse imaging conditions.

Author Contributions: Conceptualization, S.Y.; Data curation, W.W.; Formal analysis, S.L.; Funding acquisition, S.L., H.C., L.M. and Y.J.; Investigation, S.L.; Methodology, S.L.; Project administration, W.W., H.C. and F.Z.; Supervision, H.C., S.Y., L.M., F.Z. and Y.J.; Validation, L.M.; Visualization, S.L. and W.W.; Writing—original draft, W.W.; Writing—review & editing, S.Y. and J.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by Jiangsu Newlixon TechGroup.Co., Ltd Project (KJ20240106), Jiangsu Key Development Planning Project (BE2023004-2).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: We would like to express our gratitude to Newlixon for their financial support and the provision of equipment used in our experiments. Their generous contributions were instrumental in the successful completion of this research.

Conflicts of Interest: Yimu Ji's supervision, the affiliation does not provide any financial support or funding. Authors Haijun Chen and Lin Mao were employed by the company Newlixon TechGroup.Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. The funding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Liang, J.; Li, X.; Chen, J.; Tang, C.; Wang, T.; Li, Y. Suitability and regionalization of Chinese cordyceps in Qinghai Province, Northwest China. *Mycosystema* **2022**, *41*, 1772–1785.
2. Liu, W.; Gao, Y.; Zhou, Y.; Yu, F.; Li, X.; Zhang, N. Mechanism of cordyceps sinensis and its extracts in the treatment of diabetic kidney disease: A review. *Front. Pharmacol.* **2022**, *13*, 881835. [[CrossRef](#)] [[PubMed](#)]
3. Krishna, K.V.; Ulhas, R.S.; Malaviya, A. Bioactive compounds from Cordyceps and their therapeutic potential. *Crit. Rev. Biotechnol.* **2024**, *44*, 753–773. [[CrossRef](#)]
4. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the CVPR, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 770–778.
5. Du, F.; Yang, P.; Jia, Q.; Nan, F.; Chen, X.; Yang, Y. Global and local mixture consistency cumulative learning for long-tailed visual recognitions. In Proceedings of the CVPR, Vancouver, BC, Canada, 10–22 June 2023; pp. 15814–15823.
6. Fang, F.; Liu, Y.; Xu, Q. Localizing discriminative regions for fine-grained visual recognition: One could be better than many. *Neurocomputing* **2024**, *610*, 128611. [[CrossRef](#)]
7. Gao, S.H.; Cheng, M.M.; Zhao, K.; Zhang, X.Y.; Yang, M.H.; Torr, P. Res2net: A new multi-scale backbone architecture. *IEEE Trans. Pattern Anal. Mach. Intell.* **2019**, *43*, 652–662. [[CrossRef](#)]
8. Tan, M.; Le, Q. Efficientnetv2: Smaller models and faster training. In Proceedings of the ICML, Online, 18–24 July 2021; pp. 10096–10106.
9. Ding, X.; Zhang, X.; Ma, N.; Han, J.; Ding, G.; Sun, J. Repvgg: Making vgg-style convnets great again. In Proceedings of the CVPR, Online, 19–25 June 2021; pp. 13733–13742.
10. Liu, Z.; Mao, H.; Wu, C.Y.; Feichtenhofer, C.; Darrell, T.; Xie, S. A convnet for the 2020s. In Proceedings of the CVPR, New Orleans, LS, USA, 21–24 June 2022; pp. 11976–11986.
11. Woo, S.; Debnath, S.; Hu, R.; Chen, X.; Liu, Z.; Kweon, I.S.; Xie, S. Convnext v2: Co-designing and scaling convnets with masked autoencoders. In Proceedings of the CVPR, Vancouver, BC, Canada, 10–22 June 2023; pp. 16133–16142.
12. Yu, X.; Zhao, Y.; Gao, Y.; Xiong, S. Maskcov: A random mask covariance network for ultra-fine-grained visual categorization. *Pattern Recognit.* **2021**, *119*, 108067. [[CrossRef](#)]
13. Song, J.; Yang, R. Feature boosting, suppression, and diversification for fine-grained visual classification. In Proceedings of the IJCNN, Shenzhen, China 18–22 July 2021; pp. 1–8.
14. Yu, X.; Zhao, Y.; Gao, Y. SPARE: Self-supervised part erasing for ultra-fine-grained visual categorization. *Pattern Recognit.* **2022**, *128*, 108691. [[CrossRef](#)]

15. He, J.; Chen, J.N.; Liu, S.; Kortylewski, A.; Yang, C.; Bai, Y.; Wang, C. Transfg: A transformer architecture for fine-grained recognition. In Proceedings of the AAAI, Online, 22 February–1 March 2022; Volume 36, pp. 852–860.
16. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [[CrossRef](#)] [[PubMed](#)]
17. Xie, E.; Wang, W.; Ding, M.; Zhang, R.; Luo, P. Polarmask++: Enhanced polar representation for single-shot instance segmentation and beyond. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *44*, 5385–5400. [[CrossRef](#)] [[PubMed](#)]
18. Behera, A.; Wharton, Z.; Hewage, P.R.; Bera, A. Context-aware attentional pooling (cap) for fine-grained visual classification. In Proceedings of the AAAI, Online, 2–9 February 2021; Volume 35, pp. 929–937.
19. Liu, C.; Xie, H.; Zha, Z.J.; Ma, L.; Yu, L.; Zhang, Y. Filtration and distillation: Enhancing region attention for fine-grained visual categorization. In Proceedings of the AAAI, New York, NY, USA, 7–12 February 2020; Volume 34, pp. 11555–11562.
20. Zhao, Y.; Li, J.; Chen, X.; Tian, Y. Part-guided relational transformers for fine-grained visual recognition. *IEEE Trans. Image Process.* **2021**, *30*, 9470–9481. [[CrossRef](#)] [[PubMed](#)]
21. Sun, G.; Cholakkal, H.; Khan, S.; Khan, F.; Shao, L. Fine-grained recognition: Accounting for subtle differences between similar classes. In Proceedings of the AAAI, New York, NY, USA, 7–12 February 2020; Volume 34, pp. 12047–12054.
22. Murphy, K.P. *Probabilistic Machine Learning: An Introduction*; MIT Press: Cambridge, MA, USA, 2022.
23. Khorin, P.; Dzyuba, A.; Khonina, S. Optical wavefront aberration: Detection, recognition, and compensation techniques—A comprehensive review. *Opt. Laser Technol.* **2025**, *191*, 113342. [[CrossRef](#)]
24. Welinder, P.; Branson, S.; Mita, T.; Wah, C.; Schroff, F.; Belongie, S.; Perona, P. *Caltech-UCSD Birds 200*; Caltech: Pasadena, CA, USA, 2010.
25. Dataset, E. Novel datasets for fine-grained image categorization. In Proceedings of the CVPR, Colorado Springs, CO, USA, 20–25 June 2011; Volume 5, p. 2.
26. Maji, S.; Rahtu, E.; Kannala, J.; Blaschko, M.; Vedaldi, A. Fine-grained visual classification of aircraft. *arXiv* **2013**, arXiv:1306.5151. [[CrossRef](#)]
27. He, T.; Zhang, Z.; Zhang, H.; Zhang, Z.; Xie, J.; Li, M. Bag of tricks for image classification with convolutional neural networks. In Proceedings of the CVPR, Long Beach, CA, USA, 16–20 June 2019; pp. 558–567.
28. Wang, Y.; Pan, X.; Song, S.; Zhang, H.; Huang, G.; Wu, C. Implicit semantic data augmentation for deep networks. *NeurIPS* **2019**, *32*, 12614–12623.
29. Ma, Z.X.; Chen, Z.D.; Zhao, L.J.; Zhang, Z.C.; Luo, X.; Xu, X.S. Cross-Layer and Cross-Sample Feature Optimization Network for Few-Shot Fine-Grained Image Classification. In Proceedings of the AAAI, Vancouver, BC, USA, 20–28 February 2024; Volume 38; pp. 4136–4144.
30. Wu, J.; Chang, D.; Sain, A.; Li, X.; Ma, Z.; Cao, J.; Guo, J.; Song, Y.Z. Bi-directional feature reconstruction network for fine-grained few-shot image classification. In Proceedings of the AAAI, Washington, DC, USA, 7–15 February 2023; Volume 37, pp. 2821–2829.
31. Pu, Y.; Han, Y.; Wang, Y.; Feng, J.; Deng, C.; Huang, G. Fine-grained recognition with learnable semantic data augmentation. *IEEE Trans. Image Process.* **2024**, *33*, 3130–3144. [[CrossRef](#)] [[PubMed](#)]
32. Rahman, S.; Koniusz, P.; Wang, L.; Zhou, L.; Moghadam, P.; Sun, C. Learning partial correlation based deep visual representation for image classification. In Proceedings of the CVPR, Vancouver, BC, USA, 18–22 June 2023; pp. 6231–6240.
33. Ding, Y.; Ma, Z.; Wen, S.; Xie, J.; Chang, D.; Si, Z.; Wu, M.; Ling, H. AP-CNN: Weakly supervised attention pyramid convolutional neural network for fine-grained visual classification. *IEEE Trans. Image Process.* **2021**, *30*, 2826–2836. [[CrossRef](#)] [[PubMed](#)]
34. Ji, R.; Wen, L.; Zhang, L.; Du, D.; Wu, Y.; Zhao, C.; Liu, X.; Huang, F. Attention convolutional binary neural tree for fine-grained visual categorization. In Proceedings of the CVPR, Seattle, WA, USA, 13–19 June 2020; pp. 10468–10477.
35. Zhuang, P.; Wang, Y.; Qiao, Y. Learning attentive pairwise interaction for fine-grained classification. In Proceedings of the AAAI, New York, NY, USA, 7–12 February 2020; Volume 34; pp. 13130–13137.
36. Ding, Y.; Zhou, Y.; Zhu, Y.; Ye, Q.; Jiao, J. Selective sparse sampling for fine-grained image recognition. In Proceedings of the ICCV, Seoul, Republic of Korea, 27 October–2 November 2019; pp. 6599–6608.
37. Du, R.; Chang, D.; Bhunia, A.K.; Xie, J.; Ma, Z.; Song, Y.Z.; Guo, J. Fine-grained visual classification via progressive multi-granularity training of jigsaw patches. In Proceedings of the ECCV, Online, 23–28 August 2020; pp. 153–168.
38. Chen, Y.; Bai, Y.; Zhang, W.; Mei, T. Destruction and construction learning for fine-grained image recognition. In Proceedings of the CVPR, Long Beach, CA, USA, 16–20 June 2019; pp. 5157–5166.
39. Alexey, D. An image is worth 16×16 words: Transformers for image recognition at scale. In Proceedings of the ICLR, Virtual Event, 3–7 May 2021; p. 11929.
40. Yang, X.; Wang, Y.; Chen, K.; Xu, Y.; Tian, Y. Fine-grained object classification via self-supervised pose alignment. In Proceedings of the CVPR, New Orleans, LS, USA, 19–23 June 2022; pp. 7399–7408.
41. Selvaraju, R.R.; Cogswell, M.; Das, A.; Vedantam, R.; Parikh, D.; Batra, D. Grad-cam: Visual explanations from deep networks via gradient-based localization. In Proceedings of the ICCV, Nicosia, Cyprus, 24–25 April 2017; pp. 618–626.
42. Van der Maaten, L.; Hinton, G. Visualizing data using t-SNE. *J. Mach. Learn. Res.* **2008**, *9*, 2579–2605.

43. Wattenberg, M.; Viégas, F.; Johnson, I. How to Use t-SNE Effectively. *Distill* **2016**, *1*, e2. [[CrossRef](#)]
44. Belkina, A.C.; Ciccolella, C.O.; Anno, R.; Spidlen, J.; Snyder-Cappione, J.E. Automated optimal parameters for t-distributed stochastic neighbor embedding improve visualization and allow analysis of large datasets. *Nat. Commun.* **2019**, *10*, 5415. [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.