Method for the classification of tea diseases via weighted sampling and hierarchical classification learning

Rujia Li[1], Weibo Qin[2], Yiting He[1], Yadong Li[1], Rongbiao Ji[1], Yehui Wu[1], Jiaojiao Chen[1], Jianping Yang[1*]

(1. *College of Big Data, Yunnan Agricultural University, Kunming 650201, China*; 2. *College of Plant Protection, Jilin Agricultural University, Changchun 130118, China*)

Abstract: This study proposed a weighted sampling hierarchical classification learning method based on an efficient backbone network model to address the problems of high costs, low accuracy, and time-consuming traditional tea disease recognition methods. This method enhances the feature extraction ability by conducting hierarchical classification learning based on the EfficientNet model, effectively alleviating the impact of high similarity between tea diseases on the model's classification performance. To better solve the problem of few and unevenly distributed tea disease samples, this study introduced a weighted sampling scheme to optimize data processing, which not only alleviates the overfitting effect caused by too few sample data but also balances the probability of extracting imbalanced classification data. The experimental results show that the proposed method was significant in identifying both healthy tea leaves and four common leaf diseases of tea (tea algal spot disease, tea white spot disease, tea anthracnose disease, and tea leaf blight disease). After applying the "weighted sampling hierarchical classification learning method" to train 7 different efficient backbone networks, most of their accuracies have improved. The EfficientNet-B1 model proposed in this study achieved an accuracy rate of 99.21% after adopting this learning method, which is higher than EfficientNet-b2 (98.82%) and MobileNet-V3 (98.43%). In addition, to better apply the results of identifying tea diseases, this study developed a mini-program that operates on WeChat. Users can quickly obtain accurate identification results and corresponding disease descriptions and prevention methods through simple operations. This intelligent tool for identifying tea diseases can serve as an auxiliary tool for farmers, consumers, and related scientific researchers and has certain practical value.

Keywords: tea diseases, hierarchical classification learning, weighted sampling, classification method, EfficientNet, mini-program

DOI: 10.25165/j.ijabe.20241703.8236

Citation: Li R J, Qin W B, He Y T, Li Y D, Ji R B, Wu Y H, et al. Method for the classification of tea diseases via weighted sampling and hierarchical classification learning. Int J Agric & Biol Eng, 2024; 17(3): 211–221.

1 Introduction

With the continuous development of artificial intelligence and deep learning technologies, increasing research fields intersect with artificial intelligence, gradually replacing traditional methods as the mainstream. Over the past decade, the Chinese government has focused on agriculture, rural areas, and farmers, frequently mentioning the concept of intelligent agriculture and implementing policies to strengthen agriculture using technology[1].

As the world's largest producer and consumer of tea, China's tea industry is a crucial component of the economy. Different types of diseases can affect tea plants, resulting in quality issues and economic losses for tea farmers. Existing tea plant diagnosis methods rely mainly on tea farmers' experience and plant protection experts' pathological knowledge. These methods involve subjective and fuzzy judgments, lacking objective evaluation. Even experienced experts often need to correct mistakes when diagnosing tea plant diseases.

With the continuous development of machine learning, Picon et al.[2] used convolutional neural networks to detect wheat diseases with an average accuracy of 87% on a self-constructed dataset. Su et al.[3] a network model based on VGG-16 and transfer learning was proposed, enabling high accuracy in grape leaf disease recognition. Lu et al.[4] proposed A-ResNet50 and A-ResNet101 methods based on the soft attention mechanism, which achieved an accuracy of 93.27% and 93.11% on the test set respectively. Li et al.[5] proposed a migration learning-based SE-DenseNet-FL tea disease recognition method, which achieved an accuracy of 92.66% in the case of small and inhomogeneous sample distribution. Jiang et al.[6] proposed an improved ResNet18 apple leaf disease multiclassification algorithm to solve the problems of low accuracy and inefficiency of traditional methods. The disease classification accuracy of the ResNet18-CBAM-RC1 model in this algorithm reached 98.25%. Wang Dongfang et al.[7] proposed a migration learning-based crop disease classification model TL-SE-ResNeXt-101 for complex agricultural production environments with an average accuracy of 98%. Based

Received date: 2023-03-14 **Accepted date:** 2023-09-22

Biographies: Rujia Li, MS candidate, research interest: image processing, Email: 747192680@qq.com; **Weibo Qin**, Doctoral candidate, research interest: image processing identification of pests and diseases, Email: qinweibo@mails.jlau.edu.cn; **Yiting He**, MS candidate, research interest: image processing, Email: 790194799@qq.com; **Yadong Li**, MS candidate, research interest: agricultural informatization and data mining, Email: 14787825720@163.com; **Rongbiao Ji**, MS candidate, research interest: data mining, Email: 2205520365@qq.com; **Yehui Wu**, MS candidate, research interest: image processing and natural language processing, Email: 1617899733@qq.com; **Jiaojiao Chen**, MS candidate, research interest: image processing, Email: 1045979105@qq.com.

***Corresponding author: Jianping Yang**, PhD, Professor, Dean, research interest: machine learning and agricultural information technology. College of Big Data, Yunnan Agricultural University, Kunming 650201, China. Tel: +86-13529089218, Email: yangjpyn@163.com.

on attention and residual concepts, Xu et al.[8] suggested a deep convolutional network for categorizing potato leaf diseases. The findings demonstrated the deep convolutional network's excellent classification accuracy. Ji et al.[9] proposed a BR-CNN network to simultaneously implement crop variety identification, crop disease classification, and disease severity estimation. The results show that BR-CNN outperforms LP-CNN and MLP-CNN in overall performance.

Using conventional machine learning techniques, illness detection has advanced significantly, but there are still certain problems that need to be handled in accordance with the guidelines and norms of scientific research. In general, it is easy to distinguish between two classes of samples, where most samples have low disease feature similarity. However, it is difficult to distinguish between two classes of samples with high disease feature similarity, and only a small number of such samples exist. This is because the model's learning attention and recognition ability towards the latter is reduced due to excessive samples for the former[10]. For these problems to be solved, we propose a learning method called "weighted stratified classification learning." This learning method consists of two key techniques: 1) building a backbone network-based hierarchical classification model that is effective; 2) employing a weighted sampling strategy for processing the data. Our learning method is based on the hierarchical learning scheme of EfficientNet[11,12], which achieves a balance between training speed and network depth and has better feature extraction capabilities through neural networks. In addition, it has a fast convergence speed and is widely used in classification tasks. Therefore, the EfficientNet model was chosen as the backbone network in this study to ensure the ability to extract disease features while giving the network faster training and inference speed.

Furthermore, this study introduced a hierarchical learning method to alleviate the problem of reduced recognition accuracy due to the high similarity between disease features and a small number of samples. In the initial level of hierarchical learning, categories with similar disease characteristics are treated as a whole. As it progressed to the next level, further discrimination of highly similar disease combinations is achieved by utilizing sub-models. Meanwhile, to alleviate the problem of imbalanced data, weighted sampling and data augmentation were used in the pre-processing of sub-models to increase the diversity of samples. Finally, based on transfer learning, the pre-trained weights were used to train on the Imagenet dataset in PyTorch to accelerate model convergence, improve model accuracy and robustness, and have strong feature extraction capabilities in the initial stage. In addition, the development of a WeChat mini-program for a tea disease recognition system has specific practical value.

2 Materials and methods

2.1 Deep learning model EfficientNet

EfficientNet[13] is a convolutional neural network model proposed by the Google Brain team in 2019, known for its outstanding performance in image classification tasks. It balances performance and efficiency through Compound Scaling, adjusting network depth, width, and resolution to create various model sizes like EfficientNet-B0, B1, B2, etc[14].

The model's backbone relies on depth-wise separable convolution, which significantly reduces parameters and computations. Additionally, EfficientNet introduces the Swish activation function, offering improved non-linear representation capabilities while maintaining computational efficiency. This enables achieving comparable or even superior performance to larger models with smaller sizes and computational requirements. Figure 1 shows the structure of EfficientNet-B1.

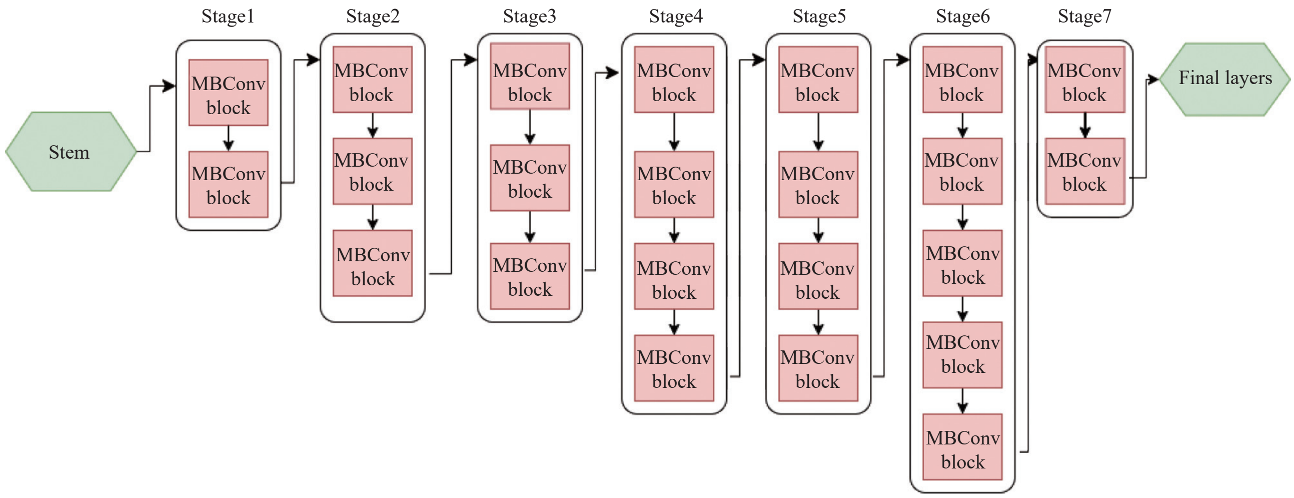


Figure 1 EfficientNet-B1 structure with an efficient backbone

2.2 Hierarchical learning

In specific image classification tasks, the small dissimilarity between diseases may lead to erroneous cross-category classification of certain samples during the classification process[15]. Hierarchical learning is a practical approach to address such problems by classifying the data from easy to difficult.

This study suggested an innovative two-layer hierarchical learning[16,17] strategy (Figure 2) to streamline the issue. All difficult-to-distinguish classes are merged into a single composite class in the first layer, where it was trained alongside other straightforward classes. The composite class is then separated into multiple distinct classes in the second layer.

In particular, in the tea disease dataset, tea algae spots and white tea star diseases have a common feature of oval or round spots, while tea anthracnose and tea leaf blight have similar spots, colors, and symptoms, making it challenging to distinguish them. These common characteristics make it difficult to distinguish tea anthracnose from tea leaf blight. However, the color of tea anthracnose and tea leaf blight is distinct from that of tea algae spots and tea white star disease, making it easy to differentiate between them. Thus, these three broad categories are easily distinguishable from each other.

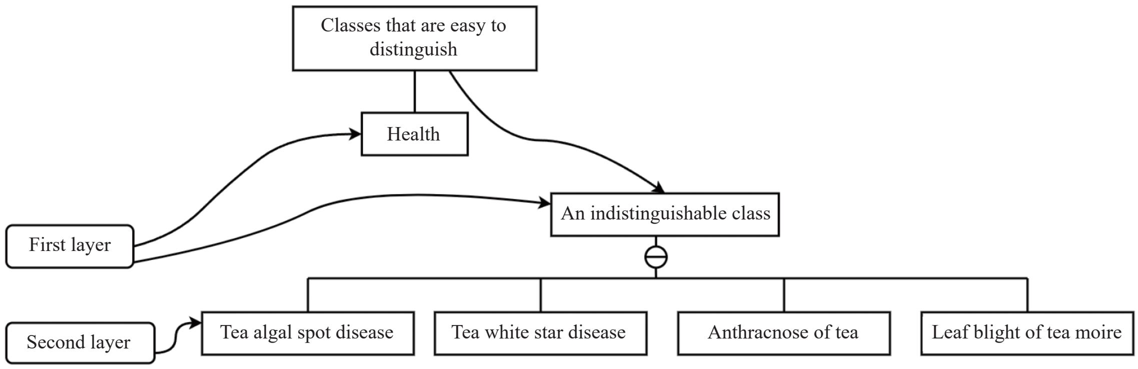


Figure 2 Two-level classification learning method for five types of tea diseases

In the training set, tea anthracnose had the fewest samples (237), while Tea white star disease was relatively few (363), and the samples from the other three categories were all over 1171. The tea algae spot disease and the tea white star disease were initially taken into account as a combination class using the concept of hierarchical learning (1078 samples total). Then, the tea anthracnose and tea leaf blight were regarded as a second combined class (with a sample number of 992), which can be easily distinguished from tea health. Therefore, the samples were divided into three categories at the first level: composite Class I, composite Class II, and health. This approach facilitates the model's learning of the differences between classes and mitigates the imbalance between tea anthracnose and tea leaf blight. In the second layer, composite class 1 (tea algae spot and tea white star disease) and composite class 2 (tea anthracnose and tea leaf blight) were further subdivided. Weighted sampling was also used for the three types of training. Finally, five classes are distinguished using the three models that were trained before.

2.3 Weighted sampling

Data that was unbalanced has a large disparity in sample sizes across its many categories. Unbalanced data have large sample size differences in their many categories. Facing the problem of category imbalance was usually handled by under-sampling and over-sampling[18,19], but these two methods often lead to fitting problems, and in order to effectively mitigate the effects of imbalance, a weighted sampling method was proposed in this study. To effectively mitigate the imbalance impact, this research suggests a weighted sampling approach. Applying a weight that is the reciprocal of the number of samples in this course to each sample of imbalanced data. In other words, samples from the minority class are given larger weights than samples from the majority class. There are fewer samples in the same class but a larger possibility of repeated sampling for the samples with higher weights. Although there are more samples in the same category, there is a smaller likelihood that samples with lower weights will be sampled. Because of this, the number of samples for each class produced using weighted sampling is about equal. The following sentence captures this concept:

$$w_i = \frac{1}{N_i} \tag{1}$$

$$p_i = \frac{w_i}{\sum_{j=1}^{k} w_i \times N_j} \tag{2}$$

$$E(C = c_i, M = m) = p_i \times N_i \times m = \frac{m}{k} \tag{3}$$

where, k represents the number of classes, N_i is the number of samples in class; w_i is the weight of each sample in class c_i; p_i is the probability of each sample in class c_i being sampled; m is the expected number of samples; E is the expected value of the sample size drawn from any category; C=c_i represents the category.

The expected sample number for any class is $E(C=c_i, M=m)=\frac{m}{k}$. As a result, when the dataset is calculated, it is balanced. They are plans for weighted sampling. Tea anthracnose and tea leaf blight were entered into two classification models in the second layer of hierarchical learning, and weighted sampling was added as a result of the data imbalance between the two classes. Each sample of tea anthracnose acquired by weighted sampling during the training period weighted 1/237 while samples of tea leaf blight weighed 1/409. The odds of getting chosen have an equivalent probability of 1/474 and 1/818. This implies that there is a higher likelihood of sampling for each sample of tea anthracnose. However, since tea leaf blight is the majority class (409) in the classification problem, the sampling probability between both classes is 1/2. Therefore, this strategy solves the problem of data imbalance.

3 Experiment

3.1 Overall design of the experiment

In this study's experimental section, experiments were conducted using tea disease datasets to explore the effectiveness of our proposed new method and compare it with other classification models. Prior to experimentation, the dataset was analyzed and the following characteristics were identified: 1) Tea algae spot and tea white star disease share the common feature of having oval or round disease spots. In contrast, tea anthracnose and tea leaf blight are difficult to distinguish due to almost identical disease spots, colors, and symptoms; 2) The data volume of tea anthracnose is lower than that of the other four categories. A hierarchical learning method was adopted to address the data imbalance issue to learn the differences between similar categories and weighted sampling to handle the imbalance problem.

The experiment's overall design is illustrated in Figure 3. First, The original sample (3510 photographs) was cleaned to obtain 2530 clear images, which were then divided into training, validation, and test sets (ratio 7:2:1). In the training phase, the data were classified into the combined categories of healthy, tea algal spot disease and tea white star disease, and tea anthracnose and tea leaf blight. After data preprocessing and the adoption of training strategies, the data were fed into a three-level model with EfficientNet-B1 as the main training dataset. Two classification models were also trained to deal with difficult-to-distinguish cases. During validation or testing, the samples were classified by the tertiary models, and the final predictions were generated either directly or by the classification models for the combined categories.

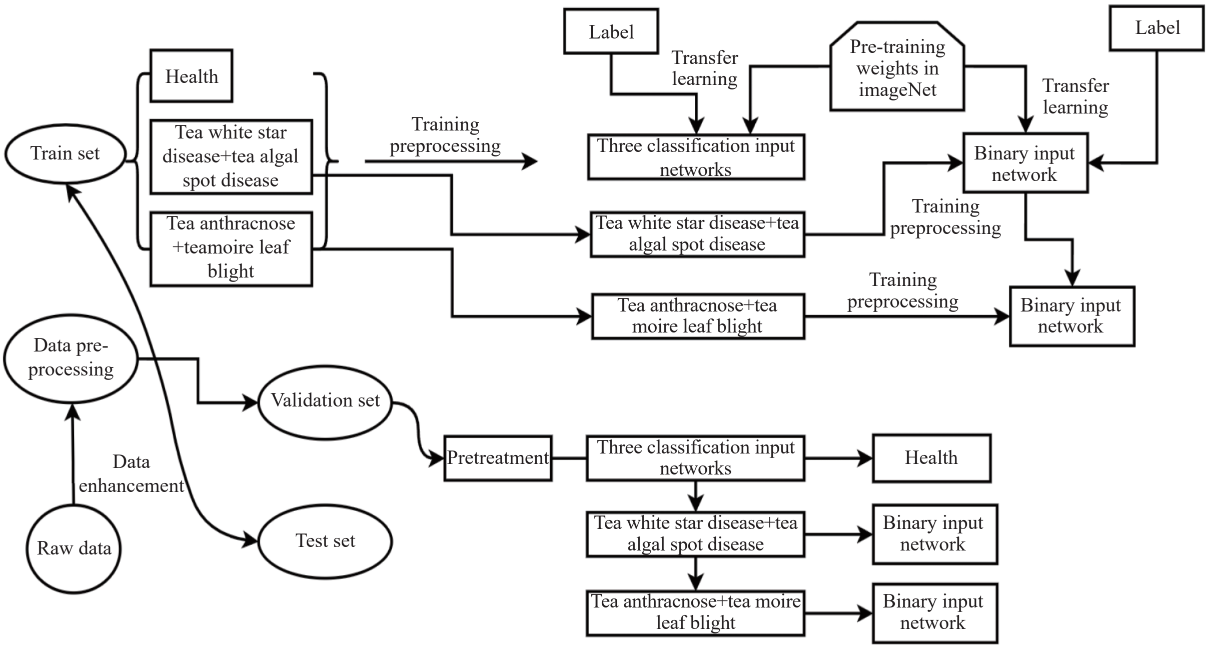


Figure 3 Overall design of the experiment of the classification of tea diseases

3.2 Data increase

3.2.1 Self-built data set

In this study, the experimental data was collected from Yunnan Agricultural University's tea plantation in the Panlong District of Kunming City, Yunnan Province (longitude: 102.744 546, latitude: 25.127 789), China. The plantation grows over 10 varieties of tea. The data were collected during three time periods, March 28-30, May 1-2, and December 1-2 of 2022, which coincided with the high incidence periods of tea diseases, making it favorable for collecting disease image data.

The EOS 70D (W) is a high-performance digital single-lens reflex camera with a high-precision 20.2 million effective pixels CMOS image sensor. Each image captured by the camera is 5472 pixels×3648 pixels. Four tea disease images and healthy tea images were captured, including tea algae spot disease, tea white star disease, tea anthracnose disease, and tea leaf blight. A total of 1169 disease images were collected, among them, healthy ones were 167, Tea algal spot disease was 196, Tea white star disease was 257, Tea anthracnose was 137, and Tea leaf blight was 412.

3.2.2 Image data expansion

Due to the limited and highly uneven distribution of collected disease images, data augmentation is employed to improve the classification accuracy and robustness of the model. In this research, data enhancement methods were mainly used, such as flipping, noise addition, rotation, and luminance[20]. After augmentation, a total of 2530 sample images were obtained. Examples of augmented images are shown in Figure 4:

Figure 4 The augmented dataset of different tea diseases

A 7:2:1 ratio was used to split the expanded dataset into training, validation, and test sets. The figures listed in Table 1 show how many examples of each class there are in each dataset.

The labels for tea health, tea algae spot disease, tea white star disease, tea anthracnose disease, and tea leaf blight were labeled as 0, 1, 2, 3, and 4, respectively. After introducing stratified learning, the labels in the first-level three-class model were adjusted to 0, 1, and 2 by changing the original labels 2 to 1, 3 to 2, and 4 to 2. In the second-level two-class model, the original labels 1 and 2 were changed to 0 and 1, and 3 and 4 were also changed to 0 and 1.

3.2.3 Data enhancement

More than 1771[21] training samples are required to train a robust deep-learning model. Therefore, As a result, the training data ought to be preprocessed using data augmentation approaches. Due to the diversity and effectiveness of most image lesions in this dataset, operations were added to the sample data, such as center cropping and rotation, scaling, horizontal flipping, and vertical flipping Figure 5.

Table 1 Division of training set, verification set, and test set

Dataset	Tea health	Tea algal spot disease	Tea white star disease	Tea anthracnose	Tea leaf blight
Filtering the data set	530	559	519	338	584
Training data	371	391	363	237	409
Validation set	106	112	104	67	116
Test set	53	56	52	34	59

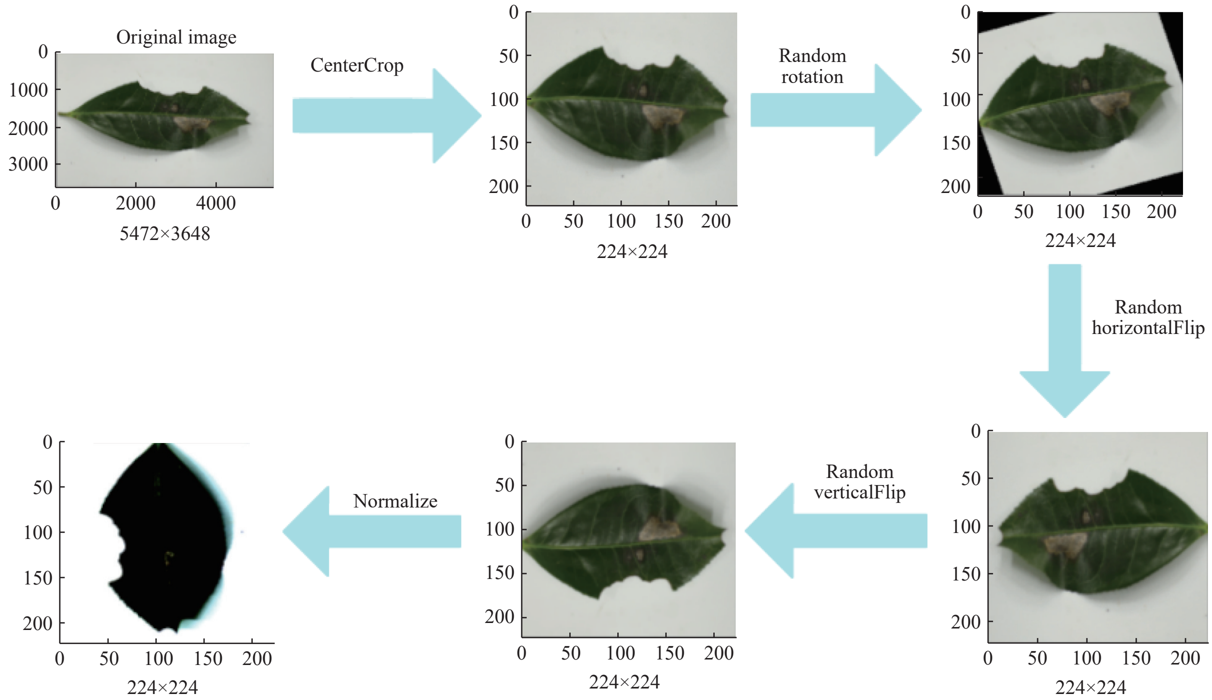


Figure 5 Data augmentation during the training phase

Assuming that the original training set is represented as T_t={(x_i, y_i), i=1, 2, 3, …, 1771}, where, 1771 represents the number of training samples, x_i represents i tea disease image, and y_i represents the category of x_i's tea disease. Iterative optimization is the learning strategy utilized in this model. Each full repetition of the training set is referred to as an epoch. The model parameters in each epoch are updated in accordance with the approximation data set of the training sample set T_t. The study's classification model's training procedure, assuming that it has experienced Ne iterations of learning cycles overall, may be summed up as follows:

Create an approximate dataset using the 1 to N sample set T_{tn}.

The data augmentation approach is used (as seen in Figure 5). T_{tn}={(z_i, y_i), i=1, 2, 3, …, 1771}. The training set T_t and test set T_{tn} both include the same number of samples, and samples z_i and x_i belong to the same class, y_i. Using the data augmentation approach shown in Figure 5, the sample z_i is generated from the sample x_i as follows:

Step 1: Sample x_i is center cropped (224×224) to generate image $V_i(1)$;

Step 2: Image $V_i(1)$ is randomly rotated (−25, +25) to generate image $V_i(2)$;

Step 3: Image $V_i(2)$ is horizontally flipped (Random Horizontal Flip) with a probability of p=0.5 to generate image $V_i(3)$;

Step 4: Image $V_i(3)$ is vertically flipped (Random Vertical Flip) with a probability of p=0.5 to generate image $V_i(4)$;

Step 5: Image $V_i(4)$ is normalized to generate image z_i.

Model training using dataset T_{tn}. Data augmentation Steps 3 and 4 have a chance of occurrence of 0.5, whereas other operations have a probability of occurrence of 1. T_{tn} is created independently by a data augmentation mechanism throughout the aforementioned learning process during each iteration.

This type of data improvement is known as online data enhancement; prior to iterative training, a certain quantity of enhanced data set T_{tn} is created from the training set T_t, and T_{tn} is utilized for computation in each iteration of learning. Throughout the iterative learning process, Ne×1771=177 100 samples are independently generated using the online data augmentation method. The diversity of training data is significantly increased, and the generalizability of the learning outcomes is guaranteed. In addition, each cycle of training uses just 1771 samples, which is alike the number of samples in the initial training set T_{tn}, assuring a manageable computing overhead.

In the three-category model, T_{tn} is generated by non-replacement sampling from the original training set T_t: each sample in T_t produces a unique augmented sample added to T_{tn}. However, the typical feature of the two-class models for tea anthracnose and tea leaf blight is imbalanced data, with a data ratio of 237:409≈1:1.72. Therefore, when generating the augmented data set T_{tn}, (tea anthracnose−tea leaf blight)={(z_i, y_i), i=1, 2, 3, …, 646}from the original training set T_t, (tea anthracnose−tea leaf blight)={(x_i, y_i), i=1, 2, 3, …, 646}, replacement sampling was used: for any i=1, 2, 3, …, 646, a sample x was randomly selected from T_t, (tea anthracnose-tea leaf blight) using the weighted sampling method in section Hierarchical learning assuming its class is y. Then, Steps 1-5 of the data were used to add x to generate an image z, making z_i=z and y_i=y(z_i, y_i) is i sample of T_{tn} (tea anthracnose-tea leaf blight), forming the augmented data set.

3.3 Training strategy and experimental parameters

3.3.1 Training strategy

In this study, a strategy was proposed for model optimization and initialization that uses one three-class model and two two-class models. Each sub-model is initialized by migration learning[22,23] and trained on the basis of pre-training weights obtained from PyTorch' s official Imagenet training set.

In order to select the optimal model, the three models were retained with the highest accuracy and finally formed the final model. Let Ne denote the total number of iterations in the model, N3e denote the total number of iterations for the three class models, and N2-1e and N2-2e denote the total number of iterations for the two two class models, respectively. In the final selection, the top three models were retained with the highest accuracy in each category. For each iteration, the accuracy and loss rate of the classifiers were calculated. After Ne iterations (a maximum value of 1000 was preset due to overfitting and underfitting considerations), the training process was stopped when the loss value of the validation set would no longer decrease and the validation accuracy would no longer increase. In summary, N3e, N2-1e, and N2-2e were finally determined to be 100.

Three sub-classifiers made up this study's optimized model: one three-class model and two two-class models. When establishing the system, the combination effect among sub-classifiers should be considered. Therefore, the top three accurate three-class models and the top three accurate two-class models among 100 candidates will be retained when selecting the best model. In 100 epochs, the validation accuracies of the three-class models were 99.60% (59th epoch), 99.60% (41st epoch), and 99.40% (42nd epoch), respectively. In 100 epochs, the validation accuracies of the two-class models of N2-1e were 98.61% (61st epoch), 99.53% (37th epoch), and 98.89% (33rd epoch), respectively. The validation accuracies of the two-class models of N2-2e were 98.36% (90th epoch), 99.45% (58th epoch), and 98.90% (43rd epoch), respectively. Finally, 7 models were selected and incorporated into the proposed learning method, and the final model was the model with the highest validation accuracy.

3.3.2 Experimental parameter setting

Different parameter selections frequently result in various experimental findings. The following were the experimental parameters employed in this study: PyTorch deep learning framework was employed with an RTX 3090 GPU and 6 GB memory. The batch size was set to 64.

For both the baseline model, the 3 classification models with the proposed learning method in this study, and the binary classification model, transfer learning with 100 epochs of iteration was employed. With a learning rate of 0.005, the learning rate decay strategy of cosineAnnealing was adopted[24], the batch data size was 24, and a cosine annealing learning rate decay strategy to prevent overfitting during the training process.

Another critical issue is determining the appropriate number of training epochs, i.e., when to stop the training process. The training loss was investigated as well as validation accuracy versus epoch number in Figure 6. As shown in Figure 6, during the training process, the loss function of the submodel keeps decreasing and tends to a lower value. However, some sub-models occasionally exhibited significant fluctuations during the convergence process. Whereas throughout the convergence process, some submodels experience fluctuations occasionally. However, the three-classification model exhibited a significant drop at the 29th epoch in Figure 6, while other models (ResNet50, ResNet152, ShuffNet-V2) also exhibited one or several significant drops before eventually stabilizing at a higher value.

With stable loss and validation accuracy, the sub-models had high data fitting and accuracy. All sub-models of the different models converged to complete or near-complete convergence after approximately 70 epochs. Therefore, 100 epochs were determined as the most appropriate number for the model.

3.3.3 Evaluation metrics

The evaluation metrics used in this experiment are Accuracy Equation (4), Recall Equation (5), Precision Equation (6), and F1-Score Equation (7). The proportion of samples that were accurately predicted to all samples is known as Accuracy.

$$\text{Accuracy} = \frac{N_{\text{TP}} + N_{\text{TN}}}{N_{\text{TP}} + N_{\text{TN}} + N_{\text{FP}} + N_{\text{FN}}} \times 100\% \tag{4}$$

where, N_{TP} is the number of true positives, N_{TN} is the number of true negatives, N_{FP} is the number of false positives, and N_{FN} is the number of false negatives. Recall is a statistic that assesses the percentage of samples in a category that were properly predicted out of all the samples in that category.

$$\text{Recall} = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FN}}} \times 100\% \tag{5}$$

Precision is defined as the ratio of the total number of predicted samples to the number of samples in a category that were accurately predicted.

$$\text{Precision} = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FP}}} \times 100\% \tag{6}$$

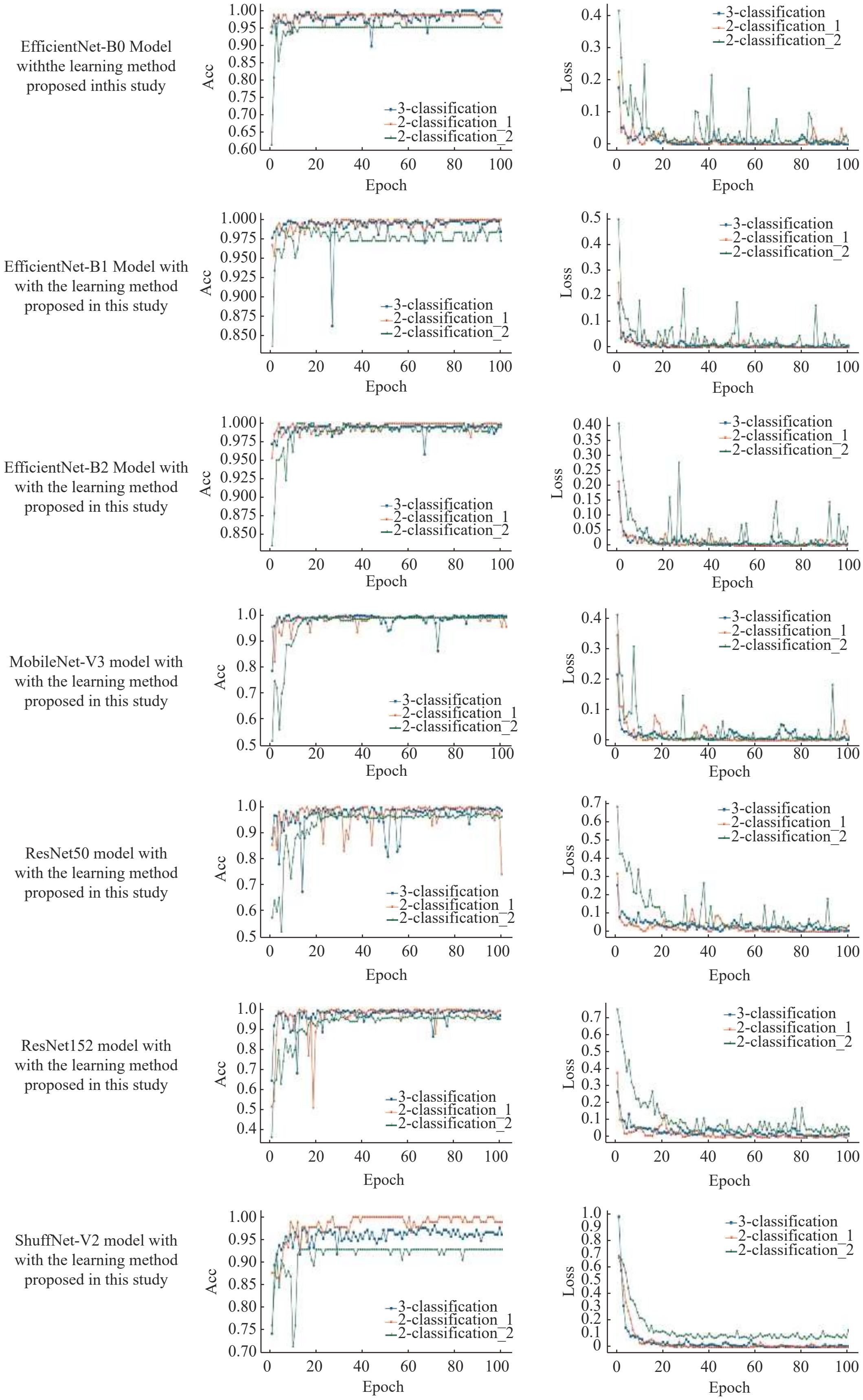
The emphasis is different for recall and precision. Additionally, the F1-Score is determined by the reconciled mean of Precision and Recall, an assessment tool that balances these 2 criteria.

$$F1-\text{Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{7}$$

4 Results and discussion

4.1 Experimental results

The effectiveness of the proposed learning method on the tea disease dataset has been validated. The method was applied to seven classic classification models, namely EfficientNet-B0, EfficientNet-B1, EfficientNet-B2, MobileNet-V3, ResNet152, ResNet50, and ShuffNet-V2, resulting in significant performance improvements (as shown in Table 2). The achieved test accuracy exceeded 91%, with EfficientNet-B1 exhibiting the best performance by achieving a relative increase of 2.76% in accuracy compared to the original model. The ShuffNet-V2 model demonstrated the greatest improvement with a 12.21% increase in test accuracy. Experimental data analysis confirms that our learning method significantly enhances the recognition accuracy of the models. This improvement can be attributed to the proposed hierarchical classification model, which addresses the challenge of distinguishing between highly similar tea diseases, and the weighted sampling method, which



a. Accuracy b. Loss

Note: The blue line represents the changes in accuracy and loss for the first layer of the three-class classification: healthy, composite class 1 (tea algae spot and tea white star disease), and composite class 2 (tea anthracnose and tea leaf blight). The orange line represents the changes in accuracy and loss for the second layer binary classification of composite class 1 (tea algae spot and tea white star disease). The green line represents the changes in accuracy and loss for the second layer binary classification of composite class 2 (tea anthracnose and tea leaf blight).

Figure 6 Loss and validation accuracy curves for each model during the training phase mitigates the issue of data imbalance. However, it should be noted that the learning method in this study exhibited a slight decrease in accuracy for certain models, such as EfficientNet-B0 and ResNet50, indicating the need for further improvements and research.

Table 2 Comparison of the overall classification performance of each model on the test set

Method	Test Accuracy/%	Test Precision/%	Test Recall/%	Test F1-Score/%
EfficientNet-B0	95.27	94.80	95.07	94.91
EfficientNet-B0 Added the learning methods in this article	91.37	93.46	89.51	90.26
EfficientNet-B1	96.45	97.04	95.61	96.18
EfficientNet-B1 Added the learning methods in this article	99.21	99.26	99.32	99.28
EfficientNet-B2	94.09	93.57	94.76	93.93
EfficientNet-B2 Added the learning methods in this article	98.82	98.91	98.98	98.92
MobileNet-V3	95.66	95.41	95.91	95.63
MobileNet-V3 Added the learning methods in this article	98.43	98.34	98.63	98.45
ResNet152	95.66	95.22	95.68	95.42
ResNet152 Added the learning methods in this article	96.46	96.35	96.41	96.30
ResNet50	98.03	97.64	98.25	97.89
ResNet50 Added the learning methods in this article	97.24	96.71	97.29	96.94
ShuffNet-V2	83.07	83.56	81.57	82.28
ShuffNet-V2 Added the learning methods in this article	95.28	94.54	95.43	94.79

4.2 Confusion matrix of transfer learning

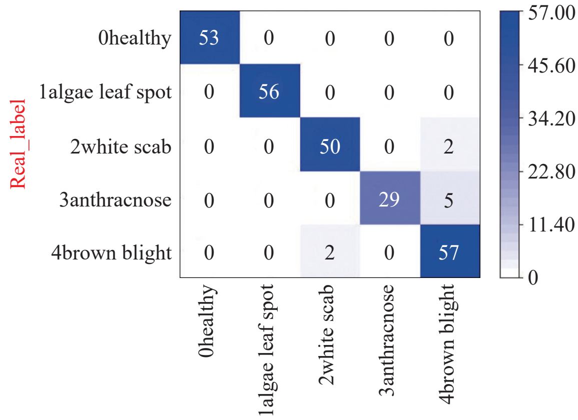
The confusion matrix is one of the methods used to evaluate the performance of a classification model. To verify the impact of the proposed learning method on the classification model's performance, We have tested the EfficientNet-B1 model separately before and after the incorporation of the proposed learning method, using a dataset of healthy tea leaves and four types of tea diseases. Confusion matrices were constructed to illustrate the misclassification of each model in recognizing different disease categories. The rows of the confusion matrix represent the true values of the predicted images, while the columns represent the predicted values of the disease images.

From Figure 7, According to Figure 7a, it can be observed that the features between the tea anthracnose and the tea leaf blight are extremely similar, causing the original EfficientNet-B1 model to misclassify more tea anthracnose as tea leaf blight, and tea white star disease as tea leaf blight. However, the improved model shown in Figure 7b greatly improves the confusion between the recognition of tea anthracnose as tea leaf blight, reducing the misclassification of tea anthracnose as tea leaf blight from 5 images to 0 images, effectively improving the accuracy of tea disease identification. This shows that the suggested learning approach has an advantageous effect on the EfficientNet-B1 model's classification performance.

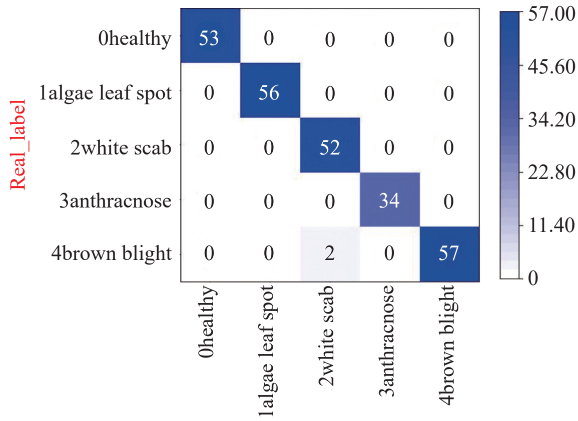
4.3 Visualization of model features

To investigate the characteristics of model feature extraction, this study conducted a visualization analysis of model features. The extracted features were obtained from the EfficientNet-B1 model and the EfficientNet-B1 model with the proposed learning algorithm introduced in this paper. In the experiment, a random selection of 1000 samples (including 254 samples from the test set) was visualized for analysis.

The chosen training or test data are fed directly into the EfficientNet-B1 model. We use flattening, dimensionality



a. EfficientNet-B1

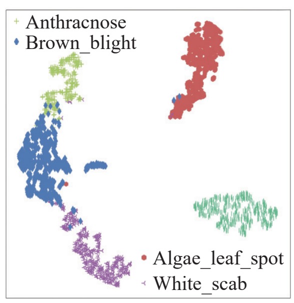


b. Join the learning method EfficientNet-B1 in this study

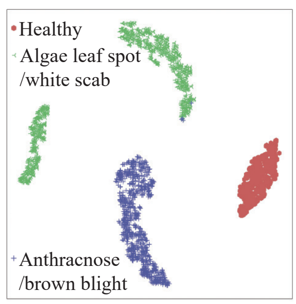
Figure 7 Hybrid matrix EfficientNet-B1 models

reduction, and visualization techniques in turn after getting the penultimate layer's output. The results are shown in Figure 8.

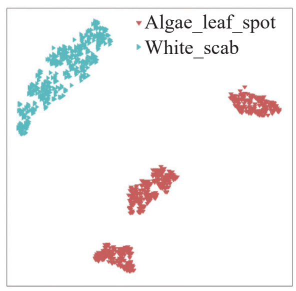
Figure 8 presents the experimental results. In the feature space of EfficientNet-B1 Figure 8a, samples of the same class are sparse,



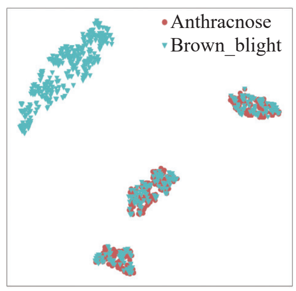
a. visualization of the original model



b. Visualization of three classification models



c. Visualization of binary classification model



d. Visualization of binary classification model

Figure 8 Visualization of the features of EfficientNet-B1 model while samples of different diseases are close to each other and entangled, leading to unclear class boundaries. In contrast, the feature space of the model using the proposed learning algorithm Figure 8b divides the four diseases and healthy leaves into three distinct categories. Samples of the same class are more densely clustered, with wider inter-class distances and clearer class boundaries. Figure 8c further refines the visualization, showing a significant reduction in entanglement between tea algal spot disease and tea white star disease, resulting in clearer class boundaries. However, due to the similarity between tea anthracnose and tea leaf blight, there is still class overlap in Figure 8d, although the inter-class distance is significant. Overall, the EfficientNet-B1 model with the proposed learning algorithm demonstrates superior performance in disease feature recognition.

4.4 Model attention visualization

Deep learning networks are often regarded as black boxes with limited interpretability. They use the gradients of the target flowing into the final convolutional layer to produce crude localization maps that emphasize critical areas in the input picture for prediction. Visualizing these localization maps as heatmaps facilitates human understanding and analysis of the classification basis of deep neural network models. To learn more about the picture areas that the model utilizes to classify tea sickness and to provide explanations for the classification results, we employed GRAD-CAM to visualize heatmaps for each of the five classes. The results, depicted in Figure 9, demonstrate that after applying the learning method proposed in this study, the EfficientNet-B1 model exhibits clearer visualization of disease regions and captures more local information related to tea disease features, leading to an improved recognition rate with an accuracy of 99.21%. These findings provide evidence for the effectiveness and practicality of the learning method proposed in this study.

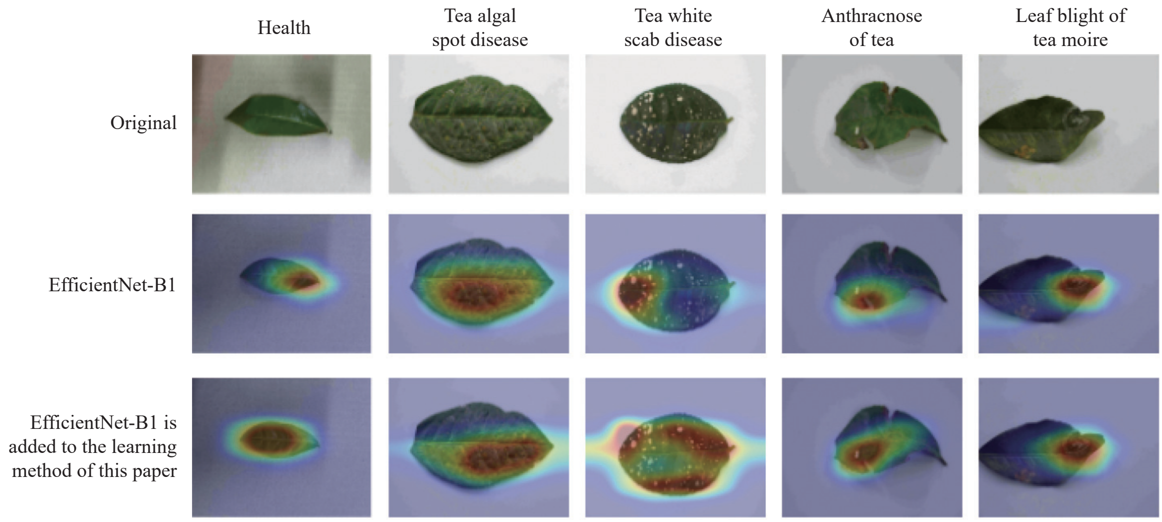
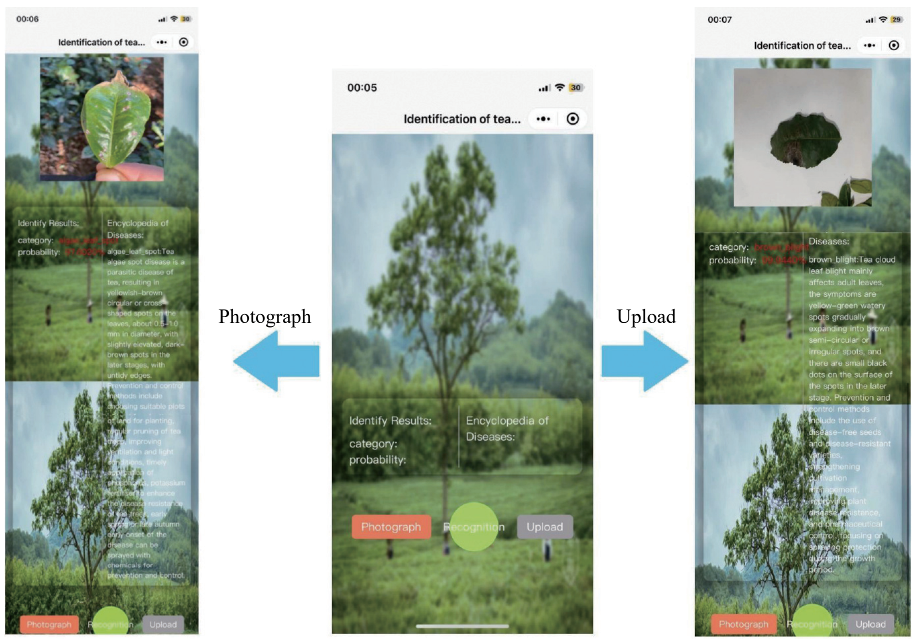


Figure 9 GRAD-CAM method for tea disease classsification

5 Tea pests and diseases small program

A WeChat mini-program for identifying tea diseases and pests was created, as shown in Figure 10, to confirm the method's applicability and put more of a focus on actual tea production. The mini-program was developed using HBuilderX as the software for



a.Photo recognition result b. Primary interface c. Upload recognition results

Figure 10 Schematic diagram of tea disease classification mini-program developing the mini-program, and the WeChat developer tool was used for debugging. Pycharm was used for server-side development, and data communication was done through HTTP. The core function of the program is to identify tea diseases and pests by taking photos and uploading images and to provide feedback on the identification results (category, probability, and disease encyclopedia) to the front end. On the front end, users can upload tea images with diseases by either taking photos or uploading images, and the server side will use the EfficientNet-B1 model trained with the proposed method to identify tea diseases and pests. Figure 10a demonstrates the page accessed by clicking the camera icon on the left-hand side of the interface, allowing users to capture real-time images of tea plants in the actual tea gardens. Figure 10b represents the homepage of a mobile application for the identification of tea diseases and pests, displaying identification results including classification, probability (recognition rate), and a compendium of diseases. By clicking the upload button on the right side of the interface, users can upload tea leaf images for disease and pest identification, as depicted in Figure 10c. This system was field-tested in the tea gardens behind Yunnan Agricultural University and made available for local tea farmers to use. Throughout the testing and practical usage, we observed that due to the diverse environmental conditions, variations in image quality, and the complexities of identifying tea leaf diseases and pests, the accuracy of the system was slightly lower by 3%-5% compared to the original test results. However, it still fully satisfies the daily operational needs of the tea farmers.

6 Conclusions

This article suggests a learning technique for categorizing tea diseases based on the EfficientNet family of superior backbone models, called the "weighted sampling hierarchical classification learning method." This method consists of two key technologies:

1) Building a backbone network-based hierarchical classification model that is effective;

2) employing a weighted sampling strategy for processing the data. The article conducted a series of experiments on a self-built tea disease dataset as the target dataset and obtained the following conclusions:

(1) This study selected 7 efficient models including EfficientNet, MobileNet, and ResNet as backbone models and trained them using the proposed learning strategy. The performance of these models in tea disease classification was evaluated using accuracy and precision metrics. The results indicated that the recognition rates of all seven models improved after applying the proposed learning method. This suggests that the learning method can effectively enhance the recognition rate of tea diseases;

(2) This study demonstrated the classification of tea diseases using methods such as transfer learning confusion matrix, model feature visualization, and Grad-CAM visualization. The results indicate that the proposed learning method effectively classifies tea diseases. The hierarchical classification model performs better in extracting disease features, enhancing the local receptive field of feature maps, and mitigating the problem of low recognition rates caused by high similarity between diseases. The weighted sampling scheme employed in data processing handles the impact of data imbalance on model performance through data augmentation.

(3) An application for WeChat that can identify tea diseases was created as part of this study.

The weighted sampling hierarchical classification learning method proposed in this article has sound effects on tea disease recognition in scenarios with small and unevenly distributed samples and high similarity between tea diseases. For farmers and the whole tea production process, the research results of this study have great value for tea growers and tea garden managers to diagnose tea diseases in daily work, save production costs, and improve production efficiency. However, the recognition rates of EfficientNet-B0 and ResNet50 models decreased by 3.54% and 0.79%, respectively, after applying this learning method to the 7 models. This indicates that this learning method still has some limitations. Therefore, future work will continue to improve this learning method, incorporate label smoothing, and combine it with morphological features to assign different meanings to classes with different degrees of similarity on the labels. Additionally, more efforts can be made to develop the WeChat mini-program for tea disease recognition to increase its practical value. In addition, in future research, we can focus on the actual application environment of the research results, in order to achieve greater application value of the research results.

Acknowledgements

The authors would like to express their sincere gratitude for the financial support provided by the Major Project of Yunnan Science and Technology, under Project No. 202302AE09002003, entitled "Research on the Integration of Key Technologies in Smart Agriculture."

[References]

[1] Zhao C J, Li J, Feng X. Developnt strategy of smart agriculture for 2035 in China. Stategic Study of CAE, 2021; 23(4): 1–9. (in Chinese)

[2] Picon A, Alvarez-Gila A, Seitz M, Ortiz-Barredo A, Echazarra J, Johannes A. Deep convolutional neural networks for mobile capture device-based crop disease classification in the wild. Computers and Electronics in Agriculture, 2019; 161: 280–290.

[3] Su S F, Qiao Y, Rao Y. Recognition of grape leaf diseases and mobile application based on transfer learning. Transactions of the CSAE, 2021; 37(10): 127–134. (in Chinese)

[4] Lu Y N, Chen B C. Identification of hops pests and diseases in small samples based on attentional mechanisms. Journal of Chinese Agricultural Mechanization, 2021; 42(3): 189–196. (in Chinese)

[5] Li Z M, Xu J, Zheng L, Tie J, Yu S. Small sample recognition method of tea disease based on improved DenseNet. Transactions of the CSAE, 2022; 38(10): 182–190. (in Chinese)

[6] Jiang H H, Yang X H, Ding R R, Wang D W, Mao W H, Qiao Y L. Identification of apple leaf diseases based on improved ResNet18. Transactions of the CSAM, 2023; 54(4): 295–303. (in Chinese)

[7] Wang D F, Wang J. Crop disease classification with transfer learning and residual networks. Transactions of the CSAE, 2021; 37(4): 199–207. (in Chinese)

[8] Xu Y, Li X Z, Wu Z H, Gao Z, Liu L. Potato leaf disease recognition via residual attention network. Journal of Shandong University of Science and Technology (Natural Science Edition), 2021; 40(2): 76–83. (in Chinese)

[9] Ji M M, Zhang K L, Wu Q F, Deng Z. Multi-label learning for crop leaf diseases recognition and severity estimation based on convolutional neural networks. Soft Computing: A fusion of foundations, methodologies and applications, 2020; 24(20): 15327–15340.

[10] Arnal Barbedo J G. Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification. Computers and Electronics in Agriculture, 2018; 153: 46–53.

[11] Sun K, He M J, He Z C, Liu H Y, Pi X T. EfficientNet embedded with spatial attention for recognition of multi-label fundus disease from color fundus photographs. Biomedical Signal Processing and Control, 2022; 77: 103768.

[12] Freeman, I, Roese-Koerner L, Kummert A. Effnet: An efficient structure for convolutional neural networks. In: 2018 25th IEEE International Conference on Image Processing (ICIP), Athens: IEEE, 2018; pp.6–10.

[13] Mimi A, Zohura S F T, Ibrahim M, Haque R R, Farrok O, Jabid T, et al. Identifying selected diseases of leaves using deep learning and transfer

learning models. Machine Graphics & Vision, 2023; 32(1): 55–71.

[14] Amin B, Samir R S, Tarek Y, Ahmed M, Ibrahim R, Ahmed M, et al. Brain tumor multi classification and segmentation in MRO images using deep learning. arXiv Preprint, arXiv: 2304.10039.

[15] Pateria S, Subagdja B, Tan A, Quek C. Hierarchical reinforcement learning: A comprehensive survey. ACM Computing Surveys, 2021; 54(5): 1–35.

[16] Pourvali M, Meng Y, Sheng C, Du Y Z. TaxoKnow: Taxonomy as prior knowledge in the loss function of multi-class classification. arXiv Preprint arXiv: 2305.16341, 2023.

[17] Lynch N, Mallmann-Trenn F. Learning hierarchically-structured concepts II: Overlapping concepts, and networks with feedback. In: International Colloquium on Structural Information and Communication Complexity, Cham: Springer Nature Switzerland, 2023; pp.46–86.

[18] Mohammed R, Rawashdeh J, Abdullah M. Machine Learning with oversampling and undersampling techniques: Overview study and ecperimental results. In: 2020 11th International Conference on Information and Communication Systems (ICICS), 2020; pp.243–248.

[19] Drummond C, Holte R C. C4. 5, class imbalance, and cost sensitivity: why under-sampling beats over-sampling. In: Workshop on learning from imbalanced datasets II, Washinton, 2003; 11: 1–8.

[20] Trabucco B, Doherty K, Gurinas M, Salakhutdinov R. Effective data augmentation with diffusion models. arXiv preprint, arXiv: 2302.07944.

[21] Ma X H, Li X R, Luo A, Zhang J Q, Li H. Galaxy image classification using hierarchical data learning with weighted sampling and label smoothing. Monthly Notices of the Royal Astronomical Society, 2023; 519(3): 4765–4779.

[22] Zhang W, Zhang P Y, Zhang B, Wang X X, Wang D. A collaborative transfer learning framework for cross-domain recommendation. arXiv Preprint, arXiv: 2306.16425.

[23] Li M, Li Y L, Lin M. A review of transfer learning for named entity recognition. Journal of Computer Science and Technology, 2021; 15(2): 206–218. (in Chinese)

[24] Wan Z W, Liu C, Zhang M, Fu J, Wang B Y, Cheng S B, et al. Med-UniC: Unifying cross-lingual medical vision-language pre-training by diminishing bias. arXiv Preprint, arXiv: 2305.19894.