

Deep Migration Learning-based Recognition of Diseases and Insect Pests in Yunnan Tea under Complex Environments

Zhaowen Li

Yunnan Agricultural University

Jihong Sun

Yunnan Agricultural University

Yingming Shen

Yunnan Provincial Academy of Science and Technology

Ying Yang

Yunnan Agricultural University

Xijin Wang

CMCC (Shanghai) Information and Communication Technology Co

Xinrui Wang

Yunnan Agricultural University

Peng Tian

Yunnan Agricultural University

Ye Qian

`sjh918a@163.com`

Yunnan Agricultural University

Research Article

Keywords: Convolutional neural network, Big-leaf sun-dried green tea, Identification of diseases and pests, Transfer learning, Complex environment

Posted Date: March 29th, 2024

DOI: <https://doi.org/10.21203/rs.3.rs-4170221/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Background

The occurrence, development, and outbreak of tea diseases and pests pose a significant challenge to the quality and yield of tea, necessitating prompt identification and control measures. Given the vast array of tea diseases and pests, coupled with the intricacies of the tea planting environment, accurate and rapid diagnosis remains elusive. In addressing this issue, the present study investigates the utilization of transfer learning convolution neural networks for the identification of tea diseases and pests. Our objective is to facilitate the accurate and expeditious detection of diseases and pests affecting the Yunnan big-leaf sun-dried green tea within its complex ecological niche.

Results

Initially, we gathered 1878 image data encompassing 10 prevalent types of tea diseases and pests from complex environments within tea plantations, compiling a comprehensive dataset. Additionally, we employed data augmentation techniques to enrich the sample diversity. Leveraging the ImageNet pre-trained model, we conducted a comprehensive evaluation and identified the Xception architecture as the most effective model. Notably, the integration of an attention mechanism within the Xception model did not yield improvements in recognition performance. Subsequently, through transfer learning and the freezing core strategy, we achieved a test accuracy rate of 99.17% and a verification accuracy rate of 96.3889%.

Conclusions

These outcomes signify a significant stride towards accurate and timely detection, holding promise for enhancing the sustainability and productivity of Yunnan tea. Our findings provide a theoretical foundation and technical guidance for the development of online detection technologies for tea diseases and pests in Yunnan.

Introduction

China, renowned for its vast tea plantations and robust consumption, boasts a planting area of 43.956 million *mu*, playing a pivotal role in propelling agricultural growth within the country. Yunnan, a prime tea-producing region, also serves as a global hub for Pu'er tea cultivation and trade. Statistics reveal that in 2020, Yunnan's tea plantations spanned 479,800 hectares, generating an output of 466,000 tons, topping the nation in both area and production. With an average annual yield of 970.35 kg per hectare and an annual output value of 39,000 yuan per hectare[1], tea has emerged as a cornerstone industry in Yunnan, generating over 100 billion yuan in revenue. Pu'er tea, the epitome of Yunnan's tea heritage, primarily relies on big-leaf sun-dried green tea as its raw material. Consequently, the quality and yield of Pu'er tea are intimately linked to the characteristics of this specific tea variety, thereby exerting a profound influence on Yunnan's economic development and the livelihoods of farmers and related enterprises. Given that tea diseases and pests are detrimental to the growth and development of tea trees, leading to a decline in quality and yield, it is imperative to address these challenges effectively. Traditional methods of pest control, while sometimes effective, often have inherent limitations. Therefore, this research aims to delve into the strengths and weaknesses of these traditional approaches and, leveraging artificial intelligence technology, proposes an intelligent model that can circumvent their shortcomings. This model strives to achieve intelligent, rapid, and accurate identification of tea diseases and pests in the complex planting environment, paving the way for more sustainable and efficient tea production in China.

Machine learning techniques have the capability to significantly enhance model performance by exploring the intricate nonlinear relationships between prediction factors and targets [2, 3]. Consequently, they have gained widespread adoption

in the realm of crop diseases and pests, encompassing diverse crops such as tobacco [4], rice [5, 6], corn [7], wheat [8], sugarcane [9], panax notoginseng [10], and cucumber [11]. The primary research approach involves constructing feature vectors based on characteristics like color, shape, and texture associated with the diseases. Subsequently, algorithms such as random forest [12] and support vector machine (SVM) [13] are employed for classification. Nevertheless, these traditional methods are fraught with challenges such as inconsistent feature extraction, limited adaptability, sensitivity to missing data, and overfitting issues in complex classification or regression tasks [14]. For instance, Liang Jiahui et al. [15] developed a disease spot image recognition model for wheat leaf disease, leveraging various traditional machine learning algorithms. Although the SVM model demonstrated the highest recognition accuracy of 91.31%, the traditional image recognition process involves numerous steps and methods, making it challenging to select the most suitable approach and limiting the ability to achieve automatic feature extraction. Similarly, Vimal K. Shrivastava et al. [16] introduced a rice disease classification model utilizing only color features from 619 image datasets, achieving a classification accuracy of 94.65%. However, this approach still relies on manual feature extraction, which can be both tedious and prone to human error. Kumar K Kishore et al. and other researchers have also explored the application of machine learning in crop disease detection, but they have encountered similar limitations associated with traditional methods. Given these challenges, there is a pressing need to explore more advanced techniques, particularly those leveraging artificial intelligence, to overcome these limitations and enhance the accuracy and efficiency of crop disease detection and classification.

With the rapid advancement of machine learning technologies, deep learning has emerged as a focal point of research in computer vision and image recognition [18]. Among the various deep learning techniques, the Convolutional Neural Network (CNN) model has garnered significant attention. This multi-layer neural network, developed in recent years, has found widespread application in diverse fields such as image processing and natural language processing (NLP). Its forward neural network architecture and deep learning capabilities enable it to achieve optimized performance and mitigate overfitting through techniques such as weight sharing, local connections, and pooling. Furthermore, CNN's ability to learn characteristics across different fields, scenes, and scales facilitates end-to-end detection, making it a valuable tool in crop pest identification, where its application is gradually becoming more widespread [19–23]. For instance, Lin Sen et al. [24] proposed a fine-grained pest identification method utilizing a Graph Pyramid Attention Convolutional Neural Network (GPA-Net). Comparative experiments conducted on cassava leaves, AI Challenger, and IP102 pest datasets demonstrate that the GPA-Net outperforms existing models, achieving accuracy rates of 99.0%, 97.0%, and 56.9%, respectively. Similarly, Lin Yuke et al. [25] employed neural networks to segment disease spots in citrus images, achieving accurate recognition results through feature matching. Their comparative experimental results reveal that this method improves the recognition rate by approximately 11.9% compared to traditional recognition methods, exhibiting superior performance. Nigam Sapna et al. [26] introduced a wheat disease identification model based on the EfficientNet architecture, designed to automatically detect major wheat rust. The fine-tuned EfficientNet B4 model achieved a remarkable test accuracy of 99.35%. Moreover, Guangsheng Liu et al. [27] proposed a lightweight model known as the Selective Kernel Mobile Network (SK MobileNet). This model achieves an accuracy rate of 99.28% on public datasets while being sufficiently lightweight to significantly reduce computing costs when deployed on servers. These advancements demonstrate the potential of deep learning techniques, particularly CNNs, in enhancing the accuracy and efficiency of crop diseases and pests detection.

In the realm of tea disease and pest recognition, Santhana Krishnan Jayapal et al. [28] introduced an image retrieval model specifically tailored for tea disease recognition. This model, leveraging depth hash and an integrated self-encoder, exhibits superior MAP (Mean Average Precision) scores compared to state-of-the-art methods. On the other hand, Lanjewar Madhusudan G [29] and his team proposed the development and implementation of a real-time disease prediction system, employing convolutional neural networks (CNN) on a Platform as a Service (PaaS) cloud. Li Zimao and colleagues [30] introduced an SE DenseNet FL approach for tea disease recognition, grounded in transfer learning. Notably, this method achieved a remarkable recognition accuracy rate of 92.66% even under challenging conditions of limited and unevenly distributed samples. Mukhopadhyay Somnath et al. [31] put forth an image clustering method leveraging a non-dominated sorting genetic algorithm (NSGA-II) for detecting tea disease areas. Experimental results demonstrated the algorithm's

ability to detect persistent disease types in tea, achieving an average accuracy of 83%. Furthermore, Gensheng Hu and his team [32] proposed a tea disease recognition method based on an improved deep convolutional neural network (CNN), achieving an average recognition accuracy of 92.5%. Lu Bing et al. [33] introduced a spectral extraction method focused on the region of interest within the speckle region. This approach enabled the neural network structure-based ELM model to attain a classification accuracy of 95.77%. These advancements not only demonstrate the significant progress made in the field of tea disease and pest recognition but also highlight the potential of deep learning techniques in enhancing the accuracy and efficiency of such recognition systems.

Nevertheless, there are some limitations in the application of convolutional neural networks (CNN) in the field of crop disease and pest recognition. While a large amount of training data is essential to train an accurate and efficient recognition model, the collection of crop disease and pest image data is difficult, time-consuming and labor-intensive. Additionally, the demand for training data and computational resources during the development process is quite large, which may be difficult to fulfill in some application scenarios. Therefore, improving the efficiency of CNN development has become an urgent problem currently. Based on the previous research results and the current development of China's tea industry, this study aims to combine deep learning algorithms to develop an automatic classification model of disease and pest digital images for small samples of Yunnan big-leaf sun-dried green tea in complex environments, and at the same time, explore ways to improve the development efficiency of CNNs in order to expand the feasibility of their application in small agricultural environments.

Experimental materials and methods

During the tea planting process, an extensive collection of over 1800 original field images of tea leaves was assembled. These images were subsequently categorized into 10 distinct disease and pest groups, thus establishing a comprehensive dataset specific to Yunnan big-leaf sun-dried green tea. Given the limited quantity and unbalanced distribution of the data, we employed data augmentation techniques to broaden the dataset's scope. Prior to training, the model was pre-trained on the ImageNet dataset, and the shared parameters obtained from this pre-training were then transferred to our tea disease and pest identification model for further fine-tuning and optimization. To assess the performance of various classification approaches, we explored and compared six high-performing image classifiers: Xception [34], VGG-16 [35], ResNet-50 [36], Inception V3 [37], MobileNetV2 [38], and DenseNet201 [39] within the context of transfer learning. Additionally, we analyzed the impact of diverse fine-tuning strategies, varying learning rates, and various attention mechanisms on the model's classification accuracy. After a thorough evaluation, the model exhibiting the most optimal performance was chosen to develop a disease and pest classification model tailored for the limited samples of Yunnan big-leaf sun-dried green tea in complex environments. This model was further enhanced through the integration of transfer learning techniques and an embedded attention mechanism. Finally, the effectiveness of the optimized model was rigorously verified.

Test data collection

Gathering a comprehensive tea disease dataset in a natural environment serves as the foundation for accurate disease and pest recognition in complex settings. On June 2, 2022, the images of tea diseases used in this experiment were meticulously captured from the tea garden planting base of Longrun Pu'er Tea College of Yunnan Agricultural University. To assess the precision of capturing images in the natural environment using portable devices, a mobile phone (specifically, an Iphone13 with 12 megapixels and a Huawei P50pro with 50 megapixels) was employed as the shooting equipment. The camera was positioned 10 ~ 15cm away from diseased tea leaves and 20 ~ 35cm away from insect pests. Varied shooting angles, including front, overhead, and side shots, were utilized to capture the diseases and pests. The imaging background naturally included interfering elements such as tea branches, soil, and other insects, ensuring the authenticity of the disease data, as illustrated in Fig. 1. In total, 97 photos of tea algal leaf spot, 322 photos of tea white spot, 75 photos of tea wheel spot disease, 157 photos of tea round red spot disease, 385 photos of powdery mildew of tea, and 261 photos of healthy tea leaf were collected. Additionally, 101 photos of *Empoasca pirusuga* Matumura, 339 photos of oil tea seed bug,

42 photos of stink bug, and 93 photos of neahnv were captured to document the insect pests. This comprehensive dataset provides a robust foundation for further analysis and the development of accurate recognition models in the realm of artificial intelligence.

Test data enhancement and data set production

Convolutional neural networks require an adequate number of samples to learn image features effectively, thereby enhancing the recognition accuracy of the target images. When the sample size is insufficient, it can significantly impact the model's training outcome, potentially leading to issues such as overfitting [40] and diminished generalization capabilities, ultimately reducing the model's practical utility. To enhance the generalization capacity of our network, we leveraged the ImageDataGenerator tool from the Keras framework in conjunction with Python. This approach allowed us to compress the image resolution to 640 × 640 pixels while simultaneously expanding the sample set through various data augmentation techniques. These techniques included random rotations of up to 30°, random horizontal and vertical shifts with an amplitude of 0.1, shear transformations with an intensity of 0.2, random zooming with an intensity of 0.2, and random horizontal and vertical flips. As depicted in Fig. 2, the images were labeled and categorized accordingly. The labels included TALS (tea algal spot disease), TWSD (tea white spot disease), TWS (tea ring spot disease), TRRS (tea red spot disease), LRT (tea rust), HL (healthy tea), EPM (tea green leafhopper), TP Lycopteris, N (Sarcophaga), and TSB (oil tea broad shield bug). To construct a comprehensive database of tea diseases and pests, the labeled images were imported into a computer in jpeg format. This database was then used to train the model on various diseases and pests affecting the Yunnan big-leaf sun-dried green tea. After applying the image augmentation techniques, the resulting sample set consisted of 5650 images. These images were subsequently divided into a training set (4520 images), a test set (565 images), and a validation set (565 images), using an 8:1:1 ratio. The distribution of these data is summarized in Table 1.

Table 1
Data distribution

Classes	Original data set	Enhanced data sets	Training set	Test set	Validation set
TALS	97	300	240	30	30
TWSD	322	900	720	90	90
TWS	75	250	200	25	25
TRRS	157	450	360	45	45
HL	261	800	640	80	80
PM	385	1200	960	120	120
EPM	101	300	240	30	30
SB	42	150	120	15	15
N	99	300	240	30	30
TSB	339	1000	800	100	100
Total	1878	5650	4520	565	565

Sample normalization

Firstly, the image is resized to 640 × 640 pixels and subsequently converted to the float32 datatype. This step is crucial as the original pixel values, ranging from 0 to 255, are excessively large for efficient computer processing. To normalize the image, we divide each pixel value by 255, resulting in a normalized range of 0 to 1, which is more suitable for computational purposes. Additionally, the dataset is thoroughly shuffled to ensure randomness and prevent any biases that

may arise from the order of the samples. Furthermore, the employment of a dynamic learning rate scheduling strategy effectively mitigates the potential learning imbalance caused by an over-concentration of specific features. This approach ensures the accuracy of the learning direction, preventing any deviations and promoting a more robust and generalized model.

Transfer learning

Migration learning, a subset of transfer learning, segregates data into two categories: source data and target data. Source data typically comprises large samples unrelated to the specific research task, whereas target data comprises smaller samples that are directly pertinent to the task. The essence of transfer learning lies in leveraging knowledge derived from source data and existing models to facilitate knowledge sharing. This approach not only maximizes the utilization of source data but also enhances the performance of models on target data [41]. In the context of this study, addressing the challenges posed by the limited number of samples and uneven distribution within the tea disease and insect pest dataset, we adopt a two-stage training process. Initially, the model is pre-trained on the ImageNet dataset [42], with the pre-training weight parameters subsequently saved. Subsequently, these parameters are transferred to the tea disease and insect pest dataset for further fine-tuning. Among the six convolutional neural network classification models considered, the one exhibiting the best performance is selected. To further optimize the model's performance, we employ a strategy known as "low-level freezing, high-level training." This approach involves freezing the parameters and structure of the lower layers of the model, while utilizing the pre-processed tea data for training the higher layers. This allows the model to learn the deeper characteristics associated with different diseases and pests, thereby enhancing its discriminative ability and overall performance.

Attention mechanism

The attention mechanism, which mimics the workings of biological vision, has garnered widespread attention in image recognition tasks due to its computational efficiency and profound understanding of image content [43]. Central to its operation is the ability to weight the output of neurons not solely based on the output of the preceding layer's neurons, but also according to distinct sections of the input data. This approach enables the model to prioritize key information within the input sequence, thereby enhancing both accuracy and efficiency. In the realm of image processing, three primary attention mechanisms stand out: SENet, ECANet, and CBAM. The SE attention mechanism introduces an attention mechanism in the channel dimension. By leveraging automatic learning, a supplementary neural network is employed to assess the significance of each channel in the feature map. Subsequently, this significance is utilized to assign weight values to individual features, allowing the neural network to concentrate on particular feature channels [44]. The ECA attention mechanism module incorporates a 1x1 convolution layer directly after the global average pooling layer, eliminating the need for a fully connected layer. This design choice circumvents dimension reduction and effectively captures cross-channel interactions. Notably, ECANet achieves impressive results with minimal parameters [45]. CBAM operates by initially passing the input feature map through the channel attention mechanism. The resulting channel weights are then multiplied with the input feature map, and the combined output is forwarded to the spatial attention mechanism. Here, the normalized spatial weights are multiplied with the input feature map of the spatial attention mechanism to yield the final weighted feature map [46]. In this study, we sought to assess the performance of various attention mechanisms on a challenging tea dataset with limited samples and a complex environmental backdrop. To achieve this, we integrated the three attention mechanisms into the best-performing model and evaluated their effectiveness.

Model design

Training

Utilizing the transfer learning approach, we fine-tuned the weights of a pre-trained basic CNN architecture, originally trained on the ImageNet dataset, for application on the tea image dataset. We conducted a comparative analysis of various common CNN architectures, evaluating their performance across different model parameters and data augmentation techniques. Based on this analysis, we designed a unified model architecture as depicted in Fig. 3. Initially, the pre-processed and standardized dataset is fed into the model, excluding the top layer, enabling feature extraction through the convolutional layers. Subsequently, global average pooling is employed to achieve feature dimensionality reduction. Finally, a softmax layer is utilized for classifying tea diseases and pests. This approach ensures that each feature map is directly associated with the classification probability, eliminating the need for a fully connected layer and thus mitigating the generation of additional training parameters. This reduction in parameters minimizes the risk of overfitting [47].

The computer operating system used for model training and testing is Windows11 64 bit system with 16GB memory and AMD Ryzen 7 7735H CPU@3.20GHz Processor, NVIDIA GeForce RTX4060 graphics card. The IDE is PyCharm2022, with Python 3.8.16, tensorflow-gpu2.6.0, and keras2.6.0 configured.

Network super parameter setting

The terminal output layer of the model comprises 10 neurons, corresponding to 10 distinct categories, and employs the softmax function as its activation mechanism. For both model training and testing, a GPU processor is utilized, while the Adam optimizer is chosen as the model optimizer. Additionally, a batch training approach is adopted, with each batch containing 64 samples. Overall, 100 iterations are performed to ensure comprehensive model training. In the fine-tuning phase of transfer learning and the integration of the embedded attention mechanism, a learning rate of 0.001 is employed. Furthermore, a dynamic learning rate scheduling strategy is implemented in the subsequent optimization stage, aiming to expedite model convergence and enhance overall performance.

Evaluation index

The final model was tested and validated using 565 pictures from the validation set, and evaluated using quantitative indicators derived from the confusion matrix. These indicators include accuracy rate, accuracy rate, recall rate and F-1 score. Accuracy is a common performance indicator in classification problems, which measures the overall performance of the model on all samples:

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

1

Accuracy is used to evaluate the accuracy of prediction by comparing the number of correctly predicted category images with the total number of the category images predicted by the model:

$$Precision = \frac{TP}{TP + FP}$$

2

The recall rate mainly measures the proportion of predicted positive samples among the actual positive samples:

$$Recall = \frac{TP}{TP + FN}$$

3

The F1 score is the harmonic average of the accuracy rate and recall rate, which takes into account the proportion of the real samples in the model prediction results and the proportion of the predicted positive samples in the actual positive samples:

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

4

TP (True Positives) is the number of samples correctly predicted by the model as positive examples; FP (False Positives) is the number of samples incorrectly predicted by the model as positive examples; TN (True Negatives) is the number of samples correctly predicted by the model as negative cases; FN (False negatives) is the number of samples whose model incorrectly predicts negative cases.

Experimental results and analysis

Model performance comparison

Under identical parameter settings and experimental conditions, Fig. 4 presents the training outcomes of various models—Xception, VGG19, ResNet50, Inception V3, MobileNetV2, and DenseNet201—on the designated test set. As evident in Fig. 4(a), the accuracy of each model demonstrates a gradual increase with the growing number of iterations. Notably, after 50 epochs of training, the accuracy of VGG19 experienced significant fluctuations, whereas the accuracy curves of the remaining models converged and stabilized. The peak recognition accuracies achieved by these six models were 90.62%, 90.2%, 87.78%, 83.38%, 82.81%, and 77.13%, respectively. Figure 4(b) illustrates the loss trends during training. Notably, the VGG19 model continued to exhibit significant fluctuations throughout the training process. In contrast, the loss values of the other models exhibited a steep decline and fluctuations during the initial 60 epochs, subsequently diminishing gradually, and ultimately attaining exceptionally low values. Among them, the Xception model exhibited smoother loss curves and lower loss values compared to the other five models.

Table 2 comprehensively outlines the comparative performance of six transfer learning architectures, encompassing evaluation metrics such as average precision, average loss, step time, round time, total parameters, trainable parameter quantities, and untrainable parameter quantities. Among these models, MobileNetV2 emerges as the top performer with an average precision of 90.05%, closely followed by Xception, DenseNet201, and ResNet50, each achieving 90.04%. When considering the loss function value, the Xception model distinguishes itself with the lowest and most stable loss of 0.4304, indicating its superiority in terms of model convergence and prediction accuracy.

In summary, the Xception model exhibits outstanding performance in terms of both accuracy and loss value. This is attributed to its lightweight design, shallow network structure, and balanced parameterization. Given the limited quantity and uneven distribution of samples in the tea disease and pest dataset, the Xception model demonstrates exceptional performance, making it the preferred choice for addressing this specific task.

Table 2

Comparison of model performance (average precision and average loss refer to the average after model convergence, both after 50 epochs)

Model	Average precision(%)	Average loss	Time of each step(ms)	Time of each epoch(s)	Depth	Total Parameters	Trainable Parameters	Non-trainable Parameters
Xception	90.04	0.4304	372	26	81	20881970	20490	20861480
VGG19	74.35	1.0884	225	16	19	20029514	5130	20024384
ResNet50	84.48	0.7094	180	13	107	23608202	20490	23578712
InceptionV3	82.49	1.2836	220	15	189	21823274	20490	21802784
MobileNetV2	90.05	0.6932	100	7	105	2270794	12810	2257984
DenseNet201	87.72	0.7255	270	20	402	18341194	19210	18321984

Analysis on the improvement of recognition performance of tea diseases and pests after embedding the attention mechanism into the Xception based model

The attention mechanism effectively enhances the activation weight of tea disease and pest-affected areas, mitigating the interference from background information. This allows the network to precisely identify the disease and pest regions that have the most significant impact on recognition outcomes. Consequently, prior to integrating SeNet, ECANet, and CBAM into the ultimate global average pooling layer of the Xception model, Fig. 5 illustrates the training outcomes achieved by employing these three methodologies.

As evident from Fig. 5, the introduction of the attention mechanism has indeed enhanced the recognition accuracy of each model: Xception_SeNet achieved the highest accuracy rate of 95.8%, while Xception_ECA reached 91.81%, and NetXception_CBAM attained 91.43%. However, it is noteworthy that the models exhibited significant fluctuations and were unable to converge stably, resulting in a severe overfitting phenomenon. Based on our analysis of the dataset, we speculate that the following issues may have contributed to this: (1) Prior to incorporating the attention mechanism, the model was already in a state of overfitting. The addition of further parameters exacerbated this issue, leading to a decline in performance. (2) Excessively deep layers and numerous channels are prone to causing overfitting. (3) Given the small size of the feature map, improper operations during convolution may introduce a significant amount of non-pixel information. (4) The proximity of the attention mechanism to the classification layer renders its effect highly sensitive to classification outcomes, potentially influencing decision-making at the classification level.

Result analysis of "low level freezing and high level training"

After comparing the Xception model's performance with the integration of the attention mechanism, we adopted a different approach to further optimize the experiment: the fine-tuning strategy of "freezing lower levels, training higher levels." Initially, we unfroze all the convolution layers in the Xception model. Subsequently, we focused on fine-tuning specific layers while freezing the remaining ones. More specifically, we trained the top 5, 10, 16, 20, and 30 layers of the unfrozen model, and the results are presented in Fig. 6. The experimental outcomes reveal that the fine-tuning strategy significantly enhances the model's precision. The accuracy rates, from highest to lowest, are: 98.71% for the unfrozen 16 layers, 97.16% for the unfrozen 20 layers, 97.02% for the unfrozen 30 layers, 96.16% for the unfrozen 10 layers, and 92.05% for the unfrozen 5 layers. This improvement can be attributed to the strong correlation between the specific features extracted from the deeper layers of the neural network and the target task. By reinitializing and retraining these network parameters, we can better adapt them to new tasks.

However, during the thawing process of the 16 layers, we observed significant fluctuations in the training results. We hypothesized that the use of a fixed learning rate might have led to the later parameters oscillating around the optimal solution. Initially, a larger learning rate is essential to quickly approach the optimal solution. Subsequently, a reduced learning rate is required to fine-tune and obtain a more precise optimal solution. Therefore, we subsequently employed a callback function to dynamically adjust the learning rate during training. We set the detection criterion as the accuracy on the test set, with an initial learning rate of 0.001 and a decay factor of 0.5. If the detection value fails to improve after 8 epochs, the learning rate is halved until it reaches 0.0001. The experimental results, presented in Fig. 7, demonstrate that the model achieved a maximum accuracy of 99.17% with a minimum loss value of 0.0353. Although the model exhibited some oscillations before 15 epochs, it converged and stabilized after 20 epochs, indicating the effectiveness of our dynamic learning rate scheduling approach.

Verification and analysis of disease and pest identification model of big-leaf sun-dried green tea

To assess the recognition performance and classification outcomes objectively and comprehensively using the Xception model, this study evaluated the model using a validation set of 565 images. A confusion matrix (illustrated in Fig. 8) was constructed based on the classification results, providing detailed insights into the prediction outcomes of the validation dataset (Table 3). Upon analyzing the confusion matrix, the model exhibited remarkable accuracy in disease recognition tasks, achieving a specific value of 0.968142. Furthermore, the model attained an average accuracy rate of 0.963889, an average recall rate of 0.973022, and an average F1 score of 0.967384. Based on these experimental findings, we confidently conclude that this approach demonstrates excellent performance on existing datasets and holds significant potential for effectively classifying crop diseases and insect pests in practical farmland environments.

Table 3
Prediction results for the validation dataset

Class	Precision	Recall	F1 Score	Accuracy
EPM	1.000000	0.967742	0.983607	
HL	0.975000	1.000000	0.987342	
PM	1.000000	0.983607	0.991736	
SB	0.933333	1.000000	0.965517	
TSB	0.900000	0.989011	0.942408	
TALS	0.900000	1.000000	0.947368	
TWS	1.000000	1.000000	1.000000	
TWSB	0.988889	0.936842	0.962162	
TRRSB	0.977778	0.880000	0.926316	
N	1.000000	0.937500	0.967742	
Average	0.9675	0.9694702	0.9674198	0.968142

Conclusion

Given the intricate characteristics of tea leaves, including their texture, shape, and color, as well as the numerous environmental interferences, accurate detection of diseases and pests poses a significant challenge. To address these

complexities, this study employed a comprehensive approach that integrates deep transfer learning methods. Initially, data augmentation techniques were utilized to curate a comprehensive dataset specifically tailored for diseases and pests affecting Yunnan big-leaf sun-dried green tea. Subsequently, leveraging existing convolutional neural network models and transfer learning principles, we developed a digital image-based automatic classification model specifically designed for the identification of diseases and pests in Yunnan big-leaf sun-dried green tea, even under challenging environmental conditions. This model was further optimized through transfer learning techniques and the integration of an attention mechanism. The key findings of this study are summarized as follows:

- (1) Employing the transfer learning technique of "low-level freezing and high-level training" significantly enhances the model's accuracy. Through meticulous experimentation with freezing different layers, we discovered that unfreezing 16 layers and adopting a dynamic scheduling strategy for the learning rate yields optimal results. Specifically, the Xception model exhibited the highest accuracy, achieving a training accuracy rate of 99.17%, a recognition accuracy rate of 0.968142%, an average accuracy rate of 0.963889, an average recall rate of 0.973022, and an average F1 score of 0.967384. These results demonstrate the effectiveness of our approach in accurately classifying diseases and pests in Yunnan big-leaf sun-dried green tea, even under challenging environmental conditions.
- (2) When faced with the challenge of identifying diseases and pests in complex environments, the integration of an attention mechanism does not always guarantee an improvement in recognition accuracy. This can be attributed to several factors. Firstly, if the model has already reached a satisfactory level of fitting before introducing the attention mechanism, the addition of extra parameters may exacerbate the overfitting problem, ultimately reducing the model's performance. Secondly, in cases where the feature map size is small or the convolution operation is inappropriate, the attention mechanism may inadvertently introduce a significant amount of non-essential information, thereby hindering accurate recognition. Lastly, if the attention mechanism is positioned too closely to the classification level, it may become overly sensitive to variations in feature extraction, potentially compromising the classification decisions and leading to a decrease in recognition accuracy. Therefore, when incorporating an attention mechanism, it is crucial to carefully consider these factors to ensure its effectiveness in complex environmental conditions.
- (3) Addressing the challenges posed by limited data scale, intricate environmental conditions, and the difficulties associated with data collection in agricultural microenvironments, the utilization of specific basic model architectures and frozen core training methods can significantly enhance the efficiency of CNN development. When confronted with relatively small and unevenly distributed training datasets, the key to achieving high model accuracy lies in the employment of data augmentation techniques and transfer learning strategies. By leveraging these techniques, we can overcome the limitations of limited data and complex environments, enabling us to develop robust and accurate CNN models for agricultural applications.

In this study, we successfully enhanced the model's accuracy and F1 scores by meticulously adjusting training super parameters and integrating the ImageNet dataset with our unique training parameters. However, it's crucial to acknowledge that our test datasets are somewhat limited in size and exhibit uneven sample distribution, which, to a certain extent, constrains the comprehensiveness and generalizability of our findings. The datasets utilized in this paper was primarily collected during early morning hours under sunny weather conditions. As we move forward, our research will focus on gathering a more diverse set of tea images, encompassing various time periods, light intensities, shading, and other complex backgrounds. This will enrich our big-leaf sun-dried green tea dataset, broaden the model's capability to identify a wider range of diseases and pests, and potentially facilitate its deployment on mobile devices. By doing so, we aim to enhance the model's popularity and application, addressing critical disease prevention and control issues, and ultimately providing robust support for the sustainable development of the tea industry.

Declarations

Authors' contributions

All authors read and approved the final manuscript.

Ethics approval and consent to participate

None.

Conflict of interest

Z.L. conducted the experiment, collected and organised the data, provided images for the article and wrote the manuscript. Y.S. modified the language of the paper. Y.Y. carried out part of the experimental work. X.W. provided some useful suggestions. X.W. and P. T. carried out the data collection and processing work. J.S. and Y.Q. participated in the design of the experiment and directed the study. The authors declare that they have no conflict of interest.

Funding This research work was funded by the Open Fund of Yunnan Key Laboratory of Crop Production and Smart Agriculture. This work has been supported by the Major Project of Science and Technology of Yunnan Province under Grant No. 202302AE090020, No. 202002AE090010 , No.202002AD080002 and Scholarship for Academic Leader of Yunnan Province funding No. 202405AC350108.

Author details

¹College of Big Data, Yunnan Agricultural University, Kunming650201, Yunnan, China.²Yunnan Key Laboratory of Crop Production and Smart Agriculture, Yunnan Agricultural University, Kunming650201, Yunnan, China.³College of Agronomy and Biotechnology, Yunnan Agricultural University, Kunming650201, Yunnan, China.⁴International Cooperation Office, Yunnan Provincial Academy of Science and Technology, Kunming650201, Yunnan, China.⁵College of Zoology, Yunnan Agricultural University, Kunming650201, Yunnan, China.⁶CMCC (Shanghai) Information and Communication Technology Co, Pudong New District200120, Shanghai, China

References

1. Chengxiu F, Yunfeng D, Yuan Y, et al. Comparative analysis of the present situation and export competitiveness of Yunnan tea industry. *Tropical Agricultural Sciences*.2022; 42 (12): 7.
2. FENG P, WANG B, LIU D L, et al. Dynamic wheat yield forecasts are improved by a hybrid approach using a biophysical model and machine learning technique. *Agricultural and Forest Meteorology*. 2020; 285-286(C).
<https://doi.org/10.1016/j.agrformet.2020.107922>
3. JUAN C, ZHAO Z, FULU T, et al. Integrating multi-source data for rice yield prediction across China using machine learning and deep learning approaches. *Agricultural and Forest Meteorology*.2021; 297.
<https://doi.org/10.1016/j.agrformet.2020.108275>
4. Yong Y, Yunwei Z, Jing W, Dalong W, Yanjun W, Jun B. Tobacco disease recognition based on computer vision. *Computer Engineering and applications*. 2015; 51(20): 167-71.
5. Chao Z. Master of embedded Image acquisition and diagnosis system for rice diseases and insect pests. Zhejiang University of Technology. 2017.
6. Yuan Y, Lei C, Na W, et al. Study on image recognition and processing method of rice sheath blight. *Agricultural mechanization research*. 2016; 38(06): 84-7+92.
7. Yandong M, He G. Maize disease recognition based on the fusion of SVM and DS evidence theory. *Chinese Journal of Agricultural Mechanization*. 2020; 41(04): 152-7.

8. Weiwei H, Wu Z, Lianzhong L. Image recognition of wheat leaf diseases based on variance-SFFS. *Journal of Hunan Agricultural University (Natural Science Edition)*. 2018; 44(02): 225-8.
9. ATHIRAJA A, RAGHAVENDER K V, LAKSHMI C B N, et al. Expert system for smart farming for diagnosis of sugarcane diseases using machine learning. *Computers and Electrical Engineering*. 2023; 109(PA).
<https://doi.org/10.1016/j.compeleceng.2023.108739>
10. Kai X, Qiliang Y, Chunxi Y, et al. Prediction of disease incidence of *Panax notoginseng* based on meteorological factors in high incidence period. *Journal of Agricultural Engineering*. 2020, 36(24): 170-6.
11. G. Y, R. M, P. P. Fractional weighted nuclear norm based two dimensional linear discriminant features for cucumber leaf disease recognition. *Multimedia Tools and Applications*. 2022; 81(27).
12. BREIMAN L. Random Forests. *Machine Learning*. 2001; 45(1). <https://doi.org/10.1007/s11042-022-13013-9>
13. Xuegong, Z.: About statistical learning theory and support vector machine. *Journal of Automation*. 2000; (01): 36-46.
14. LI H. Research review of feature extraction and classification recognition of rice disease images based on computer vision technology. *Journal of Physics Conference Series*. 2020; 1544(1). <https://doi.org/10.1088/1742-6596/1544/1/012116>
15. Jiahui L, Qian J, Hongli W, et al. Single spot image recognition of wheat leaf diseases from different sources based on machine learning. *Journal of Plant Pathology*. 2023; 53 (05): 1-22.
16. SHRIVASTAVA V K, PRADHAN M K. Rice plant disease classification using color features: a machine learning paradigm. *Journal of Plant Pathology*(prepublish). 2020 .<https://doi.org/10.1007/s42161-020-00683-3>
17. KISHORE K K, KANNAN E. Detection of rice plant disease using AdaBoostSVM classifier. *Agronomy Journal*. 2022; 114(4).<https://doi.org/10.1002/agj2.21070>
18. Yushuang H, Zhuo W, Xiangping W, et al. Research progress of deep learning in crop disease image recognition. *Chinese Journal of Agricultural Mechanization*. 2023; 44(02): 148-55.
19. JIAPING J, SHUFEI L, CHEN Q, et al. A tomato disease identification method based on leaf image automatic labeling algorithm and improved YOLOv5 model. *Journal of the science of food and agriculture*. 2023.
<https://doi.org/10.2139/ssrn.4314566>
20. LILI L, SHUJUAN Z, BIN W. Apple leaf disease identification with a small and imbalanced dataset based on lightweight convolutional networks. *Sensors*. 2021; 22(1).
21. YAN Q, YANG B, WANG W, et al. Apple leaf diseases recognition based on an improved convolutional neural network. *Sensors*. 2020; 20(12).<https://doi.org/10.3390/s20123535>
22. Jing Q, Jirong L, Zhiyong C, et al. Rice disease image recognition based on convolution neural network. *Journal of Yunnan Agricultural University (Natural Science)*. 2019; 34(05): 884-8.
23. PRAGYA H, PRASAD S M. A lightweight convolutional neural network for disease detection of fruit leaves. *Neural Computing and Applications*. 2023; 35(20).
24. SEN L, YUCHENG X, JIANLEI K, et al. An effective eyramid neural network based on graph-related attentions structure for fine-grained disease and pest identification in intelligent agriculture. *Agriculture*. 2023; 13(3).
<https://doi.org/10.3390/agriculture13030567>
25. YUKE L, JIN X, YING Z. Identification method of citrus aurantium diseases and pests based on deep convolutional neural network. *Computational Intelligence and Neuroscience*, 2022; 2022.
26. SAPNA N, RAJNI J, SUDEEP M, et al. Deep transfer learning model for disease identification in wheat crop. *Ecological Informatics*. 2023; 75.<https://doi.org/10.1016/j.ecoinf.2023.102068>
27. LIU G, PENG J, EL-LATIF A A A. SK-MobileNet: A lightweight adaptive network based on complex deep transfer learning for plant disease recognition. *Arabian journal for science and engineering*. 2023. <https://doi.org/10.1007/s13369-022-06987-z>

28. JAYAPAL S K, PORURAN S. Enhanced Disease Identification model for tea plant using deep learning. *Intelligent Automation & Soft Computing*. 2023; 35(1).<https://doi.org/10.32604/iasc.2023.026564>
29. G. L M, G. P K. Convolutional neural network based tea leaf disease prediction system on smart phone using paas cloud. *Neural Computing and Applications*. 2022; 35(3).<https://doi.org/10.1007/s00521-022-07743-y>
30. Zimao L, Jie X, Lu Z, et al. A small sample recognition method for tea diseases based on improved DenseNet. *Journal of Agricultural Engineering*. 2022; 38(10): 182-90.
31. SOMNATH M, MUNTI P, RAMEN P, et al. Tea leaf disease detection using multi-objective image segmentation. *Multimedia Tools and Applications*. 2020; 80(1).<https://doi.org/10.1007/s11042-020-09567-1>
32. HU G, YANG X, ZHANG Y, et al. Identification of tea leaf diseases by using an improved deep convolutional neural network. *Sustainable Computing: Informatics and Systems*. 2019; 24(C).
<https://doi.org/10.1016/j.suscom.2019.100353>
33. BING L, SUN J, YANG N, et al. Identification of tea white star disease and anthrax based on hyperspectral image information. *Journal of Food Process Engineering*. 2020; 44(1). <https://doi.org/10.1111/jfpe.13584>
34. CHOLLET, FRANCOIS. Xception: Deep learning with depthwise separable convolutions. *IEEE*. 2017.
<https://doi.org/10.1109/CVPR.2017.195>
35. SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition . *CoRR*. 2014; abs/1409.1556.
36. HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition. *CoRR*. 2015; abs/1512.03385
37. SZEGEDY C, VANHOUCKE V, IOFFE S, et al. Rethinking the inception architecture for computer vision. *IEEE*. 2016; 2818-2826.<https://doi.org/10.1109/CVPR.2016.308>
38. SANDLER M, HOWARD A, ZHU M, et al. MobileNetV2: Inverted residuals and linear bottlenecks. *proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2018.<https://doi.org/10.1109/CVPR.2018.00474>
39. HUANG G, LIU Z, LAURENS V D M, et al. Densely connected convolutional networks . *IEEE Computer Society*. 2016.
<https://doi.org/10.1109/CVPR.2017.243>
40. M. H D. The problem of overfitting. *Journal of Chemical Information and Computer Sciences*. 2004; 44(1).
<https://doi.org/10.1021/ci0342472>
41. Kailin Z, Xiaolong J, Yuanzhuo W. Summary of small sample learning research. *Journal of Software*. 2021.
42. ATES G C, GORGULUARSLAN R M. Convolutional encoder–decoder network using transfer learning for topology optimization. *Neural Computing and Applications*. 2024; 36(8): 4435-50.
43. ZHANG Y, LI M, ZHOU B, et al. Automatic cotton spider mites' damage grading algorithm for depth residual network based on transfer learning . *Computers and Electronics in Agriculture*. 2023; 212:108070.<https://doi.org/10.1016/j.compag.2023.108070>
44. HU J, SHEN L, ALBANIE S, et al. Squeeze-and-Excitation networks. *IEEE transactions on pattern analysis and machine intelligence*. 2020; 42(8): 2011-23.<https://doi.org/10.1109/TPAMI.2019.2913372>
45. WANG Q, WU B, ZHU P, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks. *proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2020.
<https://doi.org/10.1109/CVPR42600.2020.01155>
46. WOO S, PARK J, LEE J-Y, et al. Cbam: Convolutional block attention module; *proceedings of the Proceedings of the European conference on computer vision (ECCV)*. 2018.https://doi.org/10.1007/978-3-030-01234-2_1
47. GOUSIA H, SHAIMA Q. GAPCNN with HyPar: Global Average Pooling convolutional neural network with novel NNLU activation function and HYBRID parallelism. *Frontiers in Computational Neuroscience*. 2022.

Figures



Figure 1

Samples of diseases and pests in large-leaf sun-dried green tea with large leaves



Figure 2

Data enhancement of sample image

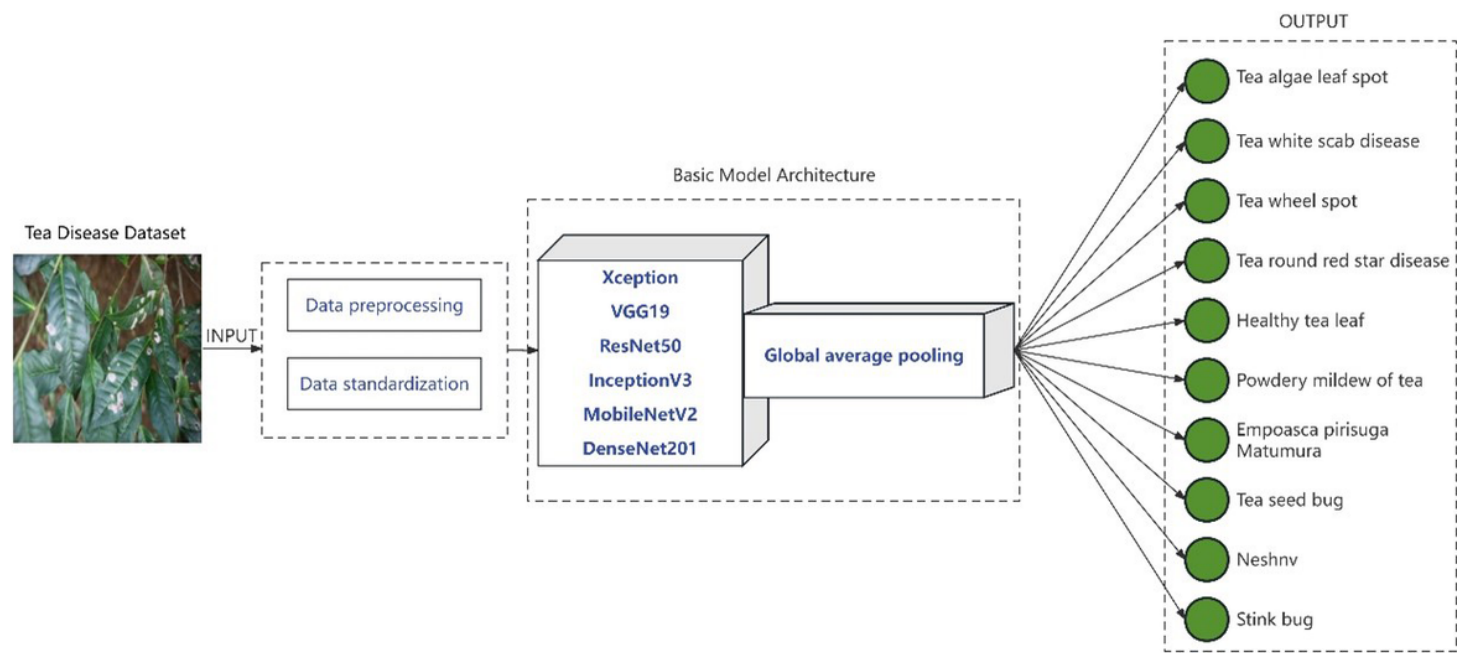


Figure 3

Convolutional neural network model architecture

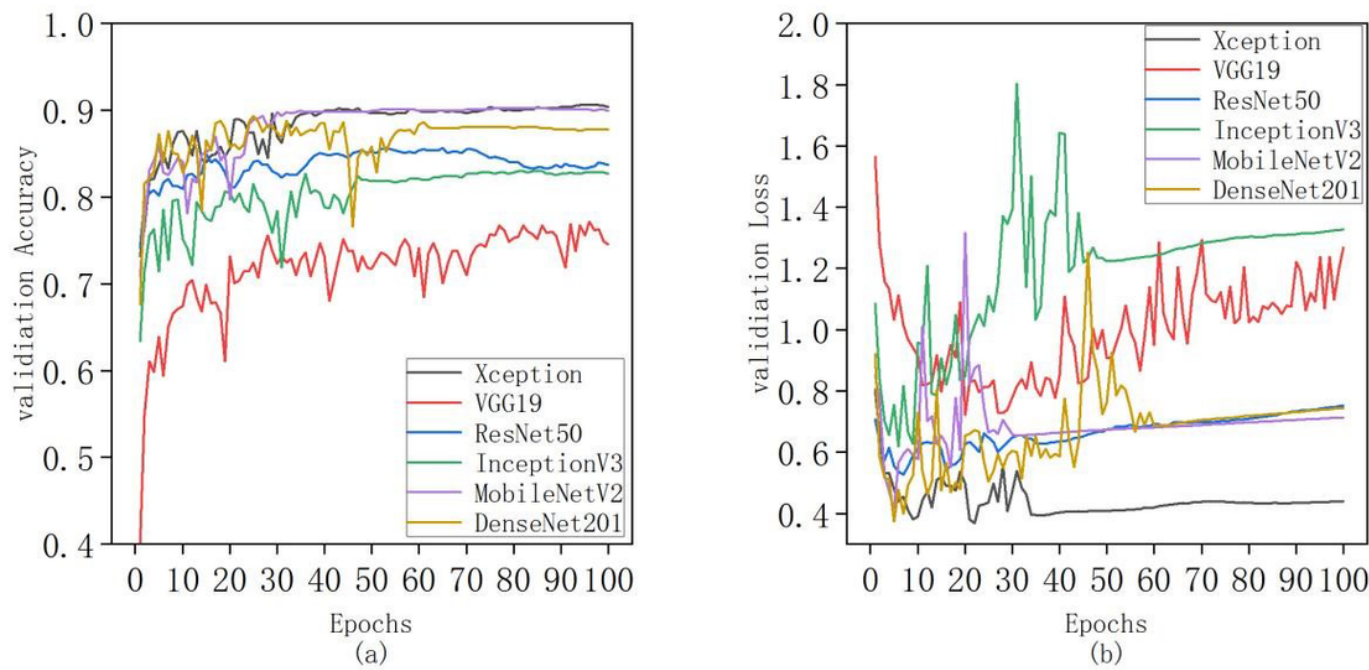


Figure 4

Comparison of accuracy and loss function values: (a) Curve of change in accuracy for the test set; (b) Curve of change in loss function values for the test set

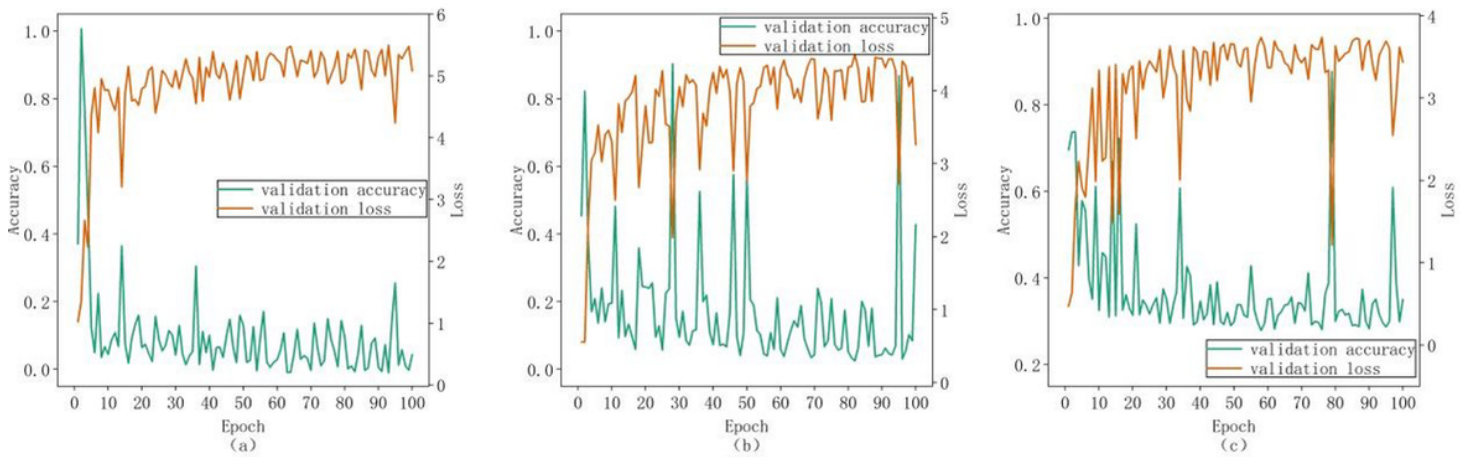


Figure 5

Comparison of accuracy and loss under different attention mechanism approaches: (a) Xception_SeNet; (b) Xception_ECANet (c) Xception_CBAM

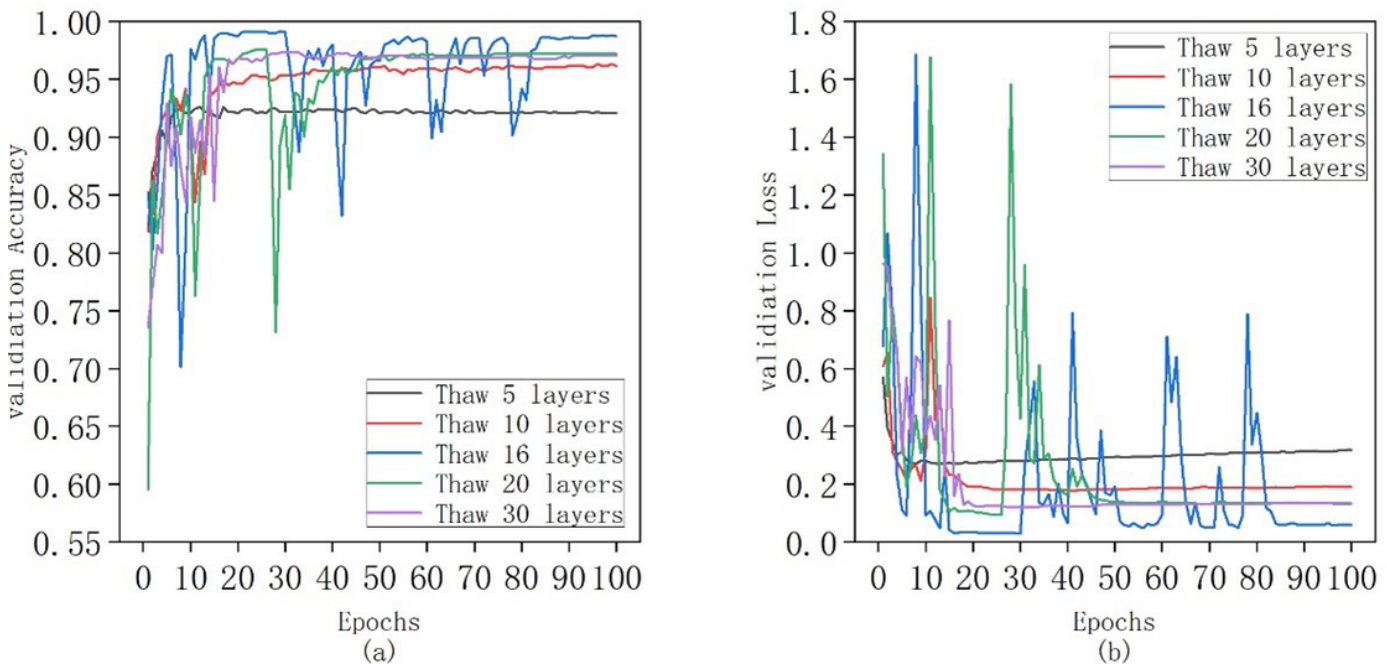


Figure 6

Comparison of model accuracy and loss function values under different fine-tuning strategies: (a) Curve of change in accuracy of test set; (b) Curve of change in loss function value of test set

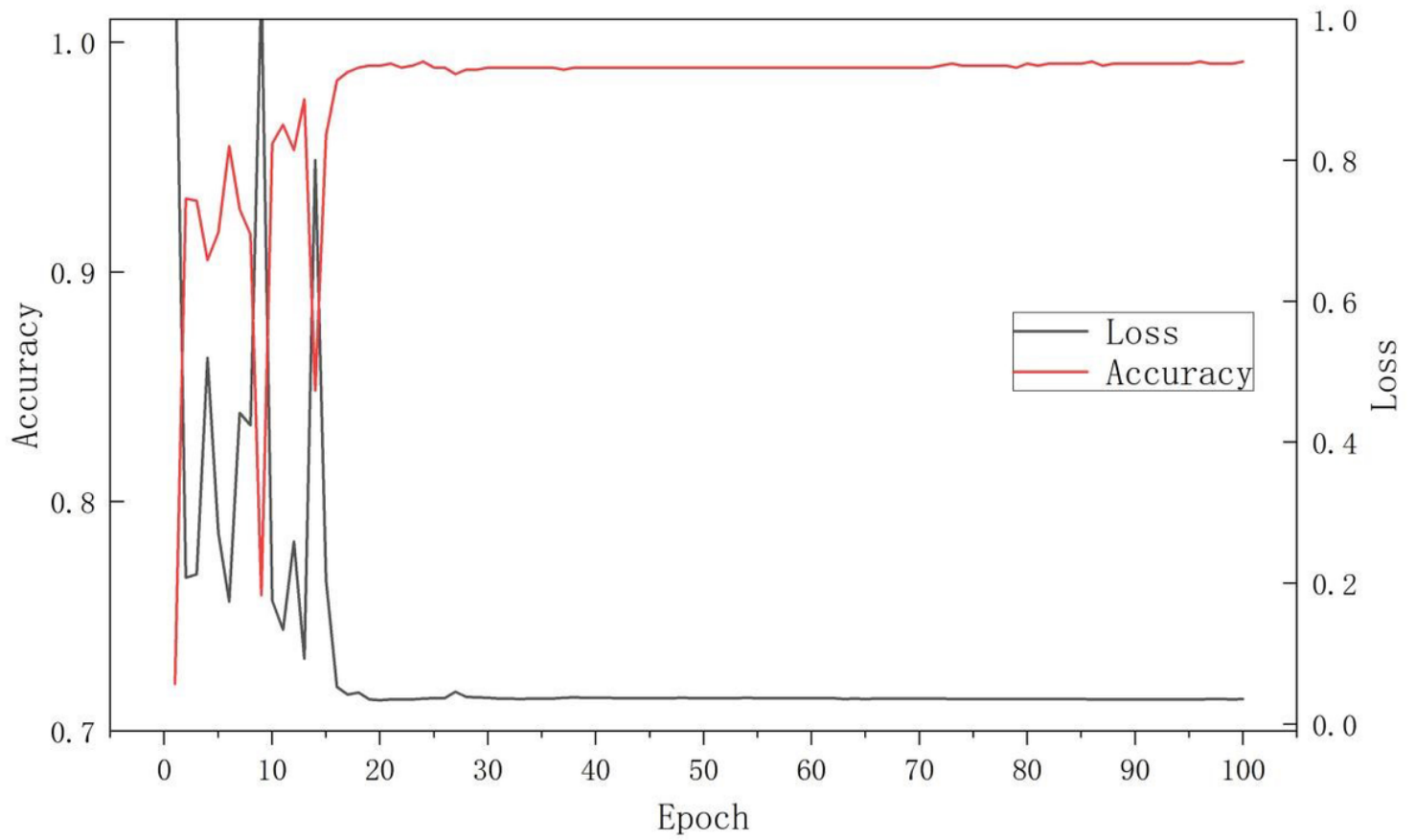


Figure 7

Accuracy and loss function values of the model with 16 layers unfrozen and a learning rate of 0.001 to 0.0001 dynamically descending selected

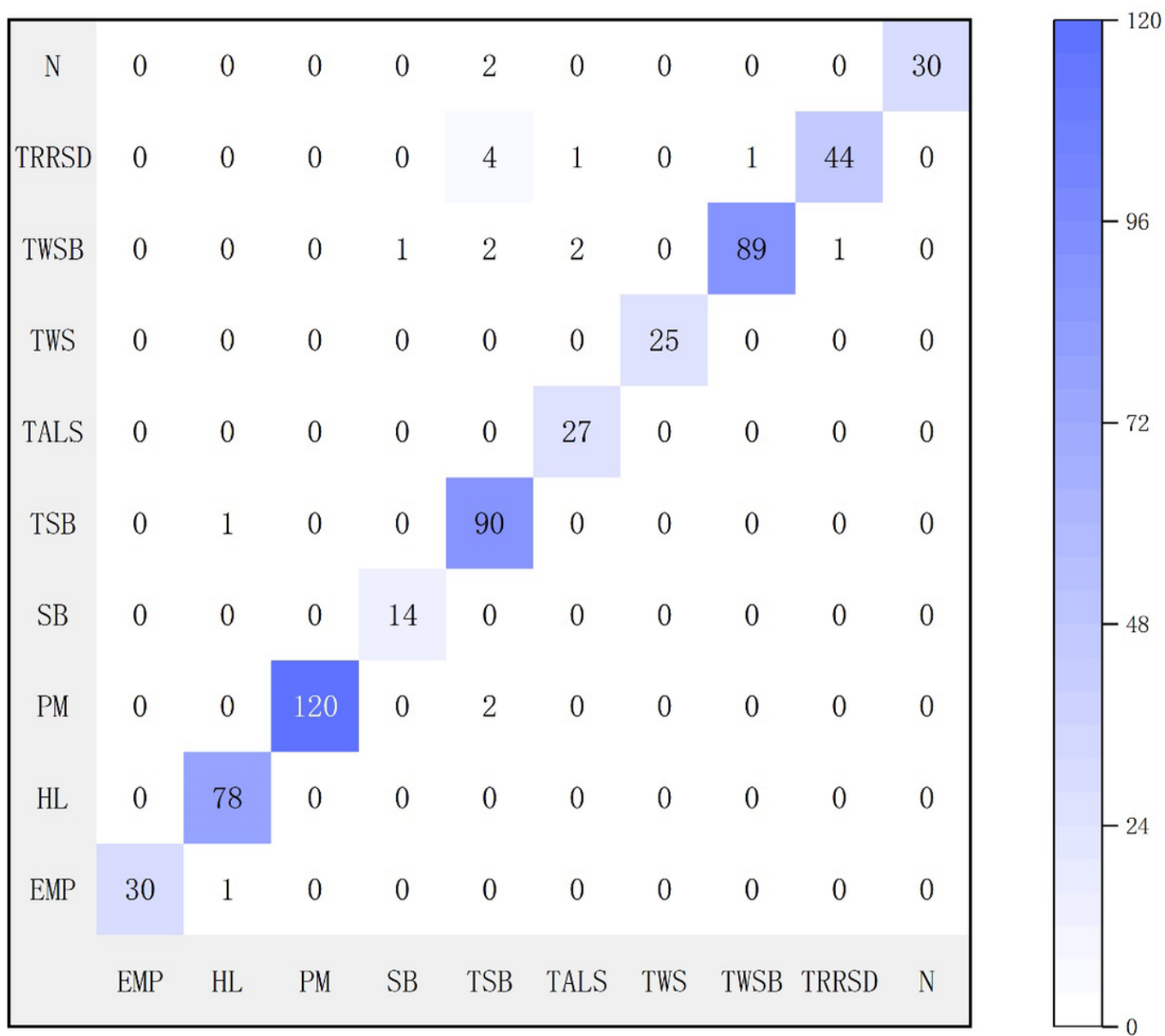


Figure 8

Confusion matrix for Xception model with 16 layers of fine-tuning