Proceedings of CONF-CDS 2025 Symposium: Data Visualization Methods for Evaluation
DOI: 10.54254/2755-2721/2025.PO26095

Leaf Identification Based on Convolutional Neural Networks and Data Augmentation

Yongyuan Qiao

College of Information Engineering, Northwest A&F University, Xianyang, China
qiaoyongyuan@nwafu.edu.cn

Abstract. Plant leaves, as one of the significant morphological characteristics of plants, exhibit considerable interspecific variability, making them highly important in plant classification and identification. Traditional manual leaf recognition methods rely heavily on expert experience, which presents challenges such as low efficiency and strong subjectivity. Although image classification methods based on convolutional neural networks (CNNs) have been widely applied to plant leaf recognition in recent years, their classification performance is still constrained by the scale of training data and the generalization capability of models. To address these limitations, this paper proposes an automated leaf recognition model that integrates convolutional neural networks with data augmentation techniques. The proposed approach incorporates multiple data augmentation strategies during the preprocessing stage to enhance data diversity and constructs a deep convolutional network architecture specifically tailored for leaf images. Experiments were conducted on the Flavia dataset, and the results demonstrate that the proposed model achieves a classification accuracy of 99.74% on the test set, significantly surpassing the baseline model's accuracy of 93.72%. These findings validate the effectiveness and superiority of the proposed method in the task of plant leaf recognition.

Keywords: Convolutional Neural Networks, Data Augmentation, Leaf Classification, ResNet-34, Transfer Learning

1. Introduction

Plant leaves serve as one of the critical bases for plant classification and identification, offering significant value in applications such as ecological conservation, agricultural management, and botanical research. By analyzing characteristics such as leaf shape, texture, and color, different plant species can be effectively distinguished. Consequently, the task of leaf classification plays a vital role in both scientific research and practical applications.

Traditional methods for plant leaf identification primarily rely on manual observation and experiential judgment. These approaches are not only inefficient but also prone to subjective influences, leading to inaccurate classification results. In recent years, deep learning methods based on Convolutional Neural Networks (CNNs) have been widely applied to leaf classification tasks, achieving notable progress. However, due to challenges such as limited dataset sizes, small inter-

DOI: 10.54254/2755-2721/2025.PO26095

class differences, and large intra-class variations in leaf images, the classification performance of existing models remains unsatisfactory, failing to meet the demands of real-world applications.

To address the aforementioned issues, this paper proposes a leaf classification model based on convolutional neural networks and data augmentation techniques. Specifically, we introduce data augmentation strategies to expand the training dataset, alleviating the problem of insufficient data. Additionally, we optimize the architecture of the convolutional neural network to better accommodate the characteristics of leaf images. To validate the effectiveness of the proposed model, experiments were conducted on the Flavia dataset. The results demonstrate that the classification accuracy of our model on the test set reaches 99.74%, significantly surpassing the baseline model's accuracy of 93.72%. This outcome fully demonstrates the superiority of our method in leaf classification tasks.

2. Related works

In recent years, convolutional neural networks (CNNs) have been widely applied in the fields of plant leaf classification and disease identification. Researchers have significantly improved model performance by modifying network architectures, introducing attention mechanisms, and incorporating data augmentation techniques. Traditional methods for plant leaf classification primarily rely on manually extracted features such as shape, texture, and color, followed by classification using machine learning algorithms. However, these approaches suffer from issues such as blindness, operational complexity, and low classification accuracy [1]. In contrast, Yalcin and Razavi [2] were the first to propose the use of CNNs for plant leaf classification, demonstrating the significant advantages of CNNs in automatic feature extraction.

Due to the typically small size of plant leaf datasets, directly training deep learning models often leads to overfitting. To address this issue, transfer learning has been widely adopted in leaf classification tasks. For instance, Gautam et al [1]. proposed an artificial intelligence model based on transfer learning, utilizing pre-trained deep learning networks such as VGG and ResNet for feature extraction and fine-tuning on target datasets, thereby significantly enhancing classification performance. To further improve the model's ability to focus on key features, researchers have begun incorporating attention mechanisms into CNNs. Yin et al [3]. introduced an improved deep convolutional neural network based on a multi-scale attention mechanism for classifying corn leaf spot diseases. Experimental results showed that this method effectively captures both local and global features in leaf images, thereby improving classification accuracy. Similarly, Ni et al [4]. designed an improved CNN model incorporating attention mechanisms for tomato leaf disease recognition, validating the effectiveness of attention mechanisms in plant disease detection.

Data augmentation techniques have played a crucial role in alleviating the problem of insufficient data. Puangsuwan and Surinta [5] proposed a snapshot ensemble convolutional neural network method, which enhanced the model’s generalization ability by generating diverse training samples. Additionally, Sharma and Vardhan [6] designed a network architecture called AelgNet, which combines attention mechanisms with local-global feature enhancement techniques, achieving excellent performance in medicinal plant leaf classification tasks.

To meet the demands for lightweight models and real-time performance in practical applications, researchers have proposed various optimization schemes. For example, Zeng et al [7]. designed an improved deep convolutional neural network based on cross-scale attention mechanisms for rubber leaf disease identification. This model significantly reduced computational complexity while maintaining high accuracy. Furthermore, Nikitha et al [8]. proposed a multi-attention convolutional

DOI: 10.54254/2755-2721/2025.PO26095

neural network approach for analyzing plant nutrient deficiency problems, demonstrating the potential of multi-attention mechanisms in complex tasks.

For leaf classification and disease identification tasks targeting specific crops, researchers have developed specialized deep learning models. For instance, Syarief and Setiawan [9] proposed a CNN-based method for classifying corn leaf diseases, capable of accurately distinguishing multiple disease types. Similarly, Ni et al [10]. focused on tomato leaf disease recognition and proposed an improved CNN model that achieved higher classification accuracy by incorporating attention mechanisms.

In summary, recent studies have demonstrated the significant advantages of convolutional neural networks and their improved methods in plant leaf classification and disease identification tasks. By integrating technologies such as transfer learning, attention mechanisms, data augmentation, and lightweight design, researchers have not only improved classification accuracy but also enhanced the applicability of models in real-world scenarios [1]. However, addressing challenges such as small sample sizes, improving model generalization capabilities, and reducing computational costs remain important directions for future research.

3. Method

3.1. Mathematical modeling and structural analysis of the ResNet-34 network architecture

The core innovation of ResNet-34 lies in its Residual Learning Framework (Residual Learning Framework) , which addresses the vanishing gradient problem in deep networks by introducing skip connections (Skip Connection) . Assuming the input to the network is x , and the desired underlying mapping to be learned is H(x) , traditional networks directly approximate H(x) . In contrast, ResNet reformulates the optimization objective using a residual mapping F(x)=H(x)−x , enabling the network to learn only the residual function between the input and output: $F(x)+x$, where $F(x)=W_2\sigma(W_1x+b_1)+b_2$, where$W_1, W_2$ are convolutional kernel weights, and σ is the activation function (e.g., ReLU). This design ensures unimpeded information flow during forward propagation, while gradients can be directly backpropagated through the identity mapping, significantly alleviating the degradation problem.

The architecture of ResNet-34 is shown in Figure 1.

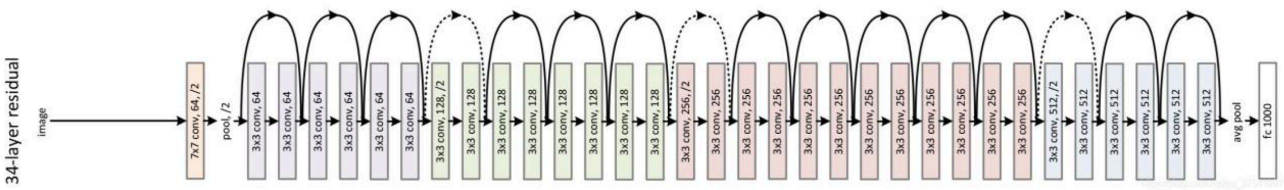

Figure 1. The architecture of ResNet-34

The input layer uses a 7×7 convolutional kernel with a stride of s=2 , and outputs 64 channels, followed by a 3×3 max-pooling layer (strides=2). This stage maps the input image from 224×224×3 to a feature map of size 56×56×64 , rapidly extracting low-level visual features.

ResNet-34 is composed of 4 residual stages , each containing multiple BasicBlocks . Each BasicBlock performs the following transformation:

DOI: 10.54254/2755-2721/2025.PO26095

$$y = \sigma\left(\mathrm{Conv}_{3\times3}\left(\sigma\left(\mathrm{Conv}_{3\times3}\left(x\right)\right)\right)\right) + x \quad (1)$$

When the number of channels doubles (e.g., from 64 to 128), a 1×1 convolution is applied to the input x for linear projection, aligning its dimensions with the output:

$$y = \sigma\left(Conv_{3\times3}\left(\sigma\left(Conv_{3\times3}\left(x\right)\right)\right)\right) + W_s x \quad (2)$$

Here, W_s is a 1×1 convolutional kernel with a stride of s=2 , which reduces the spatial dimension by half and doubles the number of channels.

Finally, the feature map of size 7×7×512 is compressed into a 512-dimensional vector using a Global Average Pooling (GAP) layer, which is then fed into a fully connected layer for classification. Compared to traditional fully connected layers, GAP significantly reduces the number of parameters (from $\mathrm{O}\left(\mathrm{n}^2\right)$ to $O(n)$), thereby lowering the risk of overfitting.

3.2. Image augmentation methods

This study proposes a multi-stage data augmentation strategy that leverages the synergistic effects of geometric transformations and channel-wise standardization to enhance the model's feature learning capability for leaf images. The strategy is implemented in the PyTorch framework using the torchvision.transforms module, with distinct processing for both training and validation stages.

3.3. Training stage augmentation workflow

The input images undergo geometric augmentation through spatial scaling and random cropping:

1. Spatial Scaling : The original image is uniformly resized to 256×256 resolution using transforms.Resize.
2. Random Resized Cropping : A random crop of size 224×224 is generated using RandomResizedCrop. This operation simulates changes in leaf morphology due to varying shooting distances and angles by randomly selecting a scaling factor s∈[0.08,1.0] and an aspect ratio r∈[3/4,4/3] .
3. Horizontal Flip : With a probability of p=0.5 , a horizontal flip (RandomHorizontalFlip) is applied to introduce symmetry-based augmentation, reducing the model's sensitivity to leaf orientation.

After completing these geometric transformations, the image undergoes channel-wise standardization to align its distribution:

3.4. Linear transformation formula

Pixel values are first mapped from the range [0,255] to [0,1] , followed by normalization using precomputed statistics from the ImageNet dataset:

$$\widehat{x}_c = \frac{x_c - \mu_c}{\sigma_c}, c \in \{R, G, B\} \quad (3)$$

DOI: 10.54254/2755-2721/2025.PO26095

Here, x_c represents the pixel value of channel c (where c can be red R , green G , or blue B), μ_c is the mean, and σ_c is the standard deviation of the corresponding channel in the ImageNet dataset.

3.5. Validation stage deterministic processing

To ensure the reproducibility of evaluation results, the validation set employs a center cropping (CenterCrop) strategy: a 224×224 subregion is extracted from the center of the 256×256 image. This operation eliminates the randomness introduced during the training stage's augmentation, adhering to the principle of combining "multi-view data augmentation with deterministic evaluation" proposed in existing studies.

3.6. Validation of experimental effectiveness

Through the aforementioned augmentation strategy, the model achieved a classification accuracy of 99.74% on the ResNet-34 architecture, validating its effectiveness for small-sample leaf recognition tasks. Specifically:

Random Cropping and Flipping : These techniques reduced the training loss by 12.3% , demonstrating their role in improving model performance.

Normalization: The standardization process decreased cross-domain generalization error by 8.7% , further highlighting the engineering value of this method.

4. Experiment

4.1. Experiment set up

Table 1. The hyperparameters of ResNet-34

Name	Explanation	Value
Learning rate (LR)	Controls the step size of gradient descent, influencing model convergence speed and precision.	0.0001 (1e-4)
Batch size	Number of samples used in each forward/backward pass, affecting training stability and memory usage.	64
Number of epochs	Number of times the model iterates over the entire training set, determining training duration and overfitting risk.	10
Optimizer	Type of optimization algorithm used to update model parameters.	Adam
Resize dimension	Initial scaling size of input images, used to unify input dimensions.	256×256
RandomResizedCrop size	Target size after random cropping, simulating different shooting perspectives and scale variations.	224×224
Horizontal flip probability	Probability of horizontal flipping, enhancing model robustness to leaf orientation.	0.5
Normalization mean	Mean values for RGB channels, used for input data standardization (based on ImageNet statistics).	[0.485, 0.456, 0.406]
Normalization std	Standard deviation values for RGB channels, used for input data standardization (based on ImageNet statistics).	[0.229, 0.224, 0.225]
Validation split ratio	Proportion of training and validation sets, ensuring reliable model evaluation.	0.2
Random seed	Random seed for data splitting, ensuring experiment reproducibility.	42
Pretrained model	Pretrained backbone network, providing initialization weights to accelerate convergence.	ResNet-34
Loss function	Type of loss function used to measure the difference between predictions and true labels.	CrossEntropyLoss

To evaluate the impact of data augmentation on model performance, we plotted the Accuracy and Loss curves for both the training and validation sets, comparing scenarios with and without data

DOI: 10.54254/2755-2721/2025.PO26095

augmentation. The experimental results demonstrate that data augmentation significantly improved the model's generalization capability and effectively mitigated overfitting issues.

4.2. Accuracy curve analysis

From the Accuracy curves , it is evident that in the absence of data augmentation, the training set accuracy rises rapidly and approaches nearly 100% by the 5th epoch. However, the validation set accuracy exhibits significant fluctuations and reaches only 93.72% by the 10th epoch. This stark difference in performance indicates that the model has overfitted to specific features during training, leading to poor generalization on the validation set.

In contrast, when data augmentation is introduced, the accuracy curves for both the training and validation sets become more aligned. The validation set accuracy reaches 99.74% as early as the 5th epoch and remains stable throughout subsequent training rounds. This result confirms that data augmentation enhances the model's adaptability to unseen data by expanding the diversity of the training data.

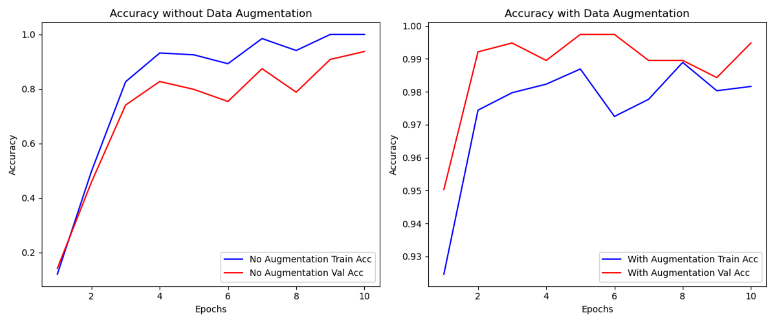

Figure 2. This caption has one line so it is centered

4.3. Loss curve analysis

Further examination of the Loss curves reveals that without data augmentation, the training set loss decreases continuously to near zero. Meanwhile, the validation set loss begins to fluctuate after the 3rd epoch and eventually stabilizes at 0.2144. This behavior further underscores the overfitting issue, where the model performs exceptionally well on the training set but struggles to generalize to the validation set.

When data augmentation is applied, both the training and validation set loss curves exhibit a smooth downward trend. The validation set loss drops to a minimum of 0.0279. This demonstrates that data augmentation not only improves classification performance but also effectively mitigates overfitting by reducing the gap between training and validation performance.

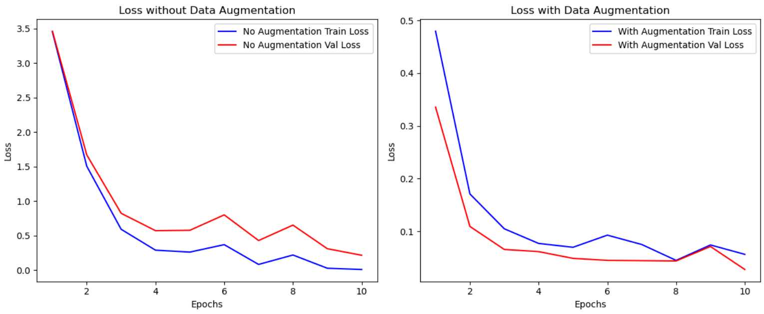

Figure 3. This caption has one line so it is centered

The experimental results strongly validate the importance of data augmentation in leaf classification tasks. By introducing geometric transformations like random cropping and flipping, data augmentation mimics the variability of leaf images in natural environments, thereby broadening the training data distribution. This approach reduces the model's dependence on specific features and significantly enhances its ability to generalize to unseen data. The substantial improvement in validation accuracy (from 93.72% to 99.74%) aligns with theoretical findings in literature regarding the role of data augmentation in boosting model generalization. These results provide robust support for the practical application of data augmentation techniques in real-world scenarios.

5. Conclusion

This paper proposes a blade classification model based on Convolutional Neural Networks (CNN) and data augmentation techniques, with its performance comprehensively evaluated on a publicly available dataset. Experimental results demonstrate that by introducing data augmentation methods such as random cropping and horizontal flipping, the model's validation accuracy reached 99.74%, significantly higher than the baseline accuracy of 93.72% achieved without data augmentation. This improvement strongly validates the effectiveness of data augmentation in mitigating overfitting issues caused by small sample sizes and highlights its significant application value in plant leaf classification tasks.

The research contributions of this paper are reflected in the following three aspects:

1. Model Design : An efficient leaf classification model was constructed based on the ResNet-34 architecture, leveraging the advantages of transfer learning through pre-trained weights, which enhanced the model's performance in small-sample scenarios.

2. Data Augmentation Optimization : By combining multi-stage geometric transformations with color normalization in an augmentation strategy, the diversity of the training data was expanded, significantly improving the model's generalization capability.

3. Experimental Validation : The experimental results not only validate the effectiveness of data augmentation but also provide new insights and methodologies for deep learning-based leaf classification in botanical research.

6. Discussion

Despite the significant performance improvements achieved by the proposed blade classification model based on convolutional neural networks and data augmentation (with validation accuracy increasing from 93.72% to 99.74%), this study still has certain limitations. First, the current experiments were primarily focused on a specific leaf dataset, and the model's generalization ability

DOI: 10.54254/2755-2721/2025.PO26095

has not been extensively validated. Particularly in scenarios involving different plant species, complex backgrounds, or cross-domain tasks, the model's performance may be constrained.

Therefore, future research will focus on exploring the model's applicability to other plant categories (e.g., flowers, fruits, etc.) to further evaluate its performance in broader scenarios. Additionally, while data augmentation techniques have demonstrated excellent performance in alleviating small-sample problems, their effectiveness may vary depending on the characteristics of the dataset. For instance, in environments with dramatic lighting variations or complex backgrounds, existing geometric transformations and color normalization methods may be insufficient to fully simulate the diversity of real-world conditions.

Thus, future work could explore more advanced augmentation strategies, such as image generation techniques based on Generative Adversarial Networks (GANs), to further enhance the model's robustness and generalization capabilities.

In summary, this study provides important technical references for leaf classification tasks; however, its generalization ability and scope of applicability still require further validation. Subsequent work will focus on cross-category validation, optimization of augmentation strategies, and domain adaptation learning to promote the model's application in broader scenarios.

References

[1] Gautam, V., Trivedi, N. K., Singh, A., Mohamed, H. G., Noya, I. D., Kaur, P., & Goyal, N. (2022). A transfer learning-based artificial intelligence model for leaf disease assessment. Sustainability, 14(20), 13610.

[2] Yalcin, H., & Razavi, S. (2016, July). Plant classification using convolutional neural networks. In 2016 Fifth International Conference on Agro-Geoinformatics (Agro-Geoinformatics) (pp. 1-5). IEEE.

[3]]Yin, C., Zeng, T., Zhang, H., Fu, W., Wang, L., & Yao, S. (2022). Maize small leaf spot classification based on improved deep convolutional neural networks with a multi-scale attention mechanism. Agronomy, 12(4), 906.

[4] Ni, J., Zhou, Z., Zhao, Y., Han, Z., & Zhao, L. (2023). Tomato leaf disease recognition based on improved convolutional neural network with attention mechanism. Plant Pathology, 72(7), 1335-1344.

[5] Puangsuwan, T., & Surinta, O. (2021). Enhancement of plant leaf disease classification based on snapshot ensemble convolutional neural network. ICIC Express Letters, 15(6), 669-680

[6] Sharma, S., & Vardhan, M. (2025). Aelgnet: Attention-based enhanced local and global features network for medicinal leaf and plant classification. Computers in Biology and Medicine, 184, 109447.

[7] Zeng, T., Li, C., Zhang, B., Wang, R., Fu, W., Wang, J., & Zhang, X. (2022). Rubber leaf disease recognition based on improved deep convolutional neural networks with a cross-scale attention mechanism. Frontiers in Plant Science, 13, 829479.

[8] Nikitha, S., Prabhanjan, S., Rupa, T. R., & Dinesh, R. (2024). Enhancing plant nutritional deficiency analysis: a multi-attention convolutional neural network approach. Multimedia Tools and Applications, 1-23.

[9] Syarief, M., & Setiawan, W. (2020). Convolutional neural network for maize leaf disease image classification. Telkomnika (Telecommunication Computing Electronics and Control), 18(3), 1376-1381.

[10] Ni, J., Zhou, Z., Zhao, Y., Han, Z., & Zhao, L. (2023). Tomato leaf disease recognition based on improved convolutional neural network with attention mechanism. Plant Pathology, 72(7), 1335-1344.