The Bioscan
AN INTERNATIONAL QUARTERLY JOURNAL OF LIFE SCIENCES
19(2): S.I (1), 889-891, 2024 www.thebioscan.com

Leaf Morphology Classification Using Customized Convolutional Neural Networks

[1]A. SRI LAKSHMI, *[2] JYOTHI N M

[1] A. Sri Lakshmi, Computer Applications, Government Degree College Autonomous, Nagari, Andhra Pradesh, India

*[2] Jyothi N M, Computer Science Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh, India

*Corresponding Author: **JYOTHI N M**

DOI: https://doi.org/10.63001/tbs.2024.v19.i02.S.I(1).pp889-891

KEYWORDS
Leaf morphology,
CNN,
plant classification,
image processing,
biodiversity

Received on:

22-07-2024

Accepted on:

20-08-2024

Published on:

24-09-2024

ABSTRACT

Leaf morphology is a key trait in plant identification and taxonomy, crucial for understanding biodiversity and agricultural studies. Manual classification methods are often labor-intensive and subjective. Machine learning approaches, particularly convolutional neural networks (CNNs), offer automated and precise solutions. To classify plant leaves based on their morphology using CNNs, leveraging image datasets to develop an accurate and scalable model. The Plant Leaves for Image Classification dataset from Kaggle, consisting of leaf images from multiple species, was utilized. Data preprocessing included resizing, normalization, and augmentation. A CNN model was trained and optimized for classification tasks. The proposed CNN achieved an accuracy of 94.85% on the test set. Sample classifications demonstrated the model's ability to distinguish between similar leaf structures effectively. This study highlights the potential of CNNs in automating leaf morphology classification, contributing to advancements in botany and agriculture.

INTRODUCTION

Leaf morphology is a vital factor in plant taxonomy and ecological research. Manual classification systems face challenges in scalability and accuracy, especially for large datasets. Deep learning methods, particularly CNNs, have revolutionized image classification tasks by extracting hierarchical features directly from images. This paper explores a CNN-based approach to classify plant leaves by morphology, leveraging a publicly available dataset for performance evaluation.

LITERATRUE SURVEY

Singh et al. reviewed various neural network techniques for plant leaf classification, emphasizing the advantages of CNNs in handling complex morphological traits compared to traditional image processing methods [1]. The study highlights challenges such as overfitting and the importance of data augmentation.

Banzi and Abayo proposed a CNN-based model, DeepLeaf, which combines leaf images and GPS data for plant species classification, achieving an accuracy of 95.06%. This research underscores the potential of integrating auxiliary data for improved classification [2]. Saleem et al. introduced an ensemble model combining CNNs with Vision Transformers (ViT), achieving state-of-the-art accuracy across multiple datasets. Their work demonstrated the benefits of hybrid models in improving feature extraction and classification [3]. LeCun et al. presented a foundational study on CNNs for leaf classification tasks, detailing the architecture's ability to learn hierarchical features and reduce manual feature engineering [4]. Khan et al. proposed compact CNN architectures tailored for plant leaf classification, achieving high accuracy with reduced computational costs. The study focused on optimizing CNN layers to suit resource-constrained environments [5]. Bisen and Singh developed a deep CNN-based plant species recognition system using leaf features, achieving a remarkable accuracy of 97%. Their approach demonstrated the effectiveness of CNNs in distinguishing subtle morphological differences [6]. Quach et al. combined CNN-based features with support vector machines (SVM) for leaf recognition, achieving superior results on the Flavia leaf dataset. This hybrid approach highlighted the advantages of combining deep learning with traditional classifiers [7]. Ali et al. explored the use of CNNs for detecting and classifying diseases in botanical leaves, particularly focusing on mango leaves. Their research demonstrated high accuracy and robustness in identifying diseased samples [8]. Wang et al. developed an automatic leaf recognition system using deep CNNs. Their model showed significant improvements in accuracy by leveraging the hierarchical feature learning capabilities of CNNs [9]. Jadhav et al. applied transfer learning techniques to CNN-based leaf classification, achieving an average accuracy of 96.3%. Their study emphasized the utility of pre-trained models in reducing training time and improving performance [10]. Rahman et al. used compact CNN architectures for tomato leaf disease classification, achieving high accuracy while ensuring computational efficiency. This approach demonstrated the scalability of CNNs for agricultural applications [11].

Xie et al. developed a CNN-based classifier for leaves using hyperspectral imaging. Their research highlighted the potential of integrating advanced imaging techniques with CNNs to improve classification accuracy for complex datasets [12].
Traditional image processing techniques rely heavily on handcrafted features, limiting their adaptability to diverse leaf shapes and textures. Existing CNN models often overfit on small datasets due to a lack of augmentation and proper regularization. Few studies address the challenge of distinguishing leaves with similar morphologies across species.

PROPOSED METHODOLOGY

Architecture: A custom CNN architecture designed for leaf classification tasks. It includes convolutional 3 convolutional layers with ReLU activation and kernel size 3x3.
2 max-pooling layers with pool size 2x2. Dropout layers (rate 0.5) to prevent overfitting. Fully connected dense layer with softmax activation to handle multi-class classification. Figure 1 shows customized CNN architecture.

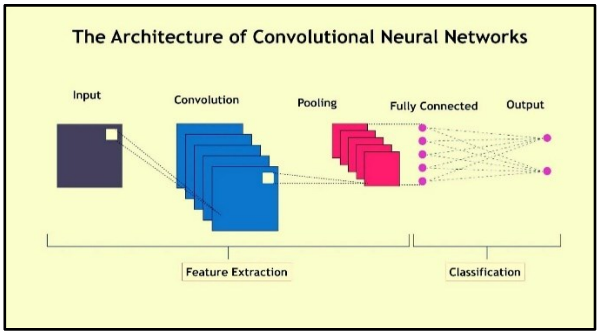


Fig.1.CNN Architecture

Steps:
1. All images are resized to a fixed dimension of 128x128 pixels to ensure uniformity and compatibility with the CNN input layer. Pixel values are scaled to the range [0, 1] by dividing each pixel by 255, stabilizing the model training process.
2. Data augmentation techniques such as rotation (±15 degrees), horizontal and vertical flipping, zooming (up to 20%), and random cropping are applied to increase data diversity and reduce overfitting.
3. The CNN model starts with an input layer accepting preprocessed images of size 128x128x3 (width, height, and RGB channels). It includes three convolutional layers with ReLU activation and 3x3 kernels to extract spatial features like edges, shapes, and textures. Two max-pooling layers with a 2x2 pool size reduce the spatial dimensions, preserving essential features while lowering computational complexity. Dropout with a rate of 0.5 is applied after each pooling layer to prevent overfitting by randomly deactivating neurons during training.
4. The model includes two fully connected dense layers with 256 and 128 neurons, respectively, to learn high-level features. A final dense layer with softmax activation outputs probabilities for each class.
5. The model uses categorical cross-entropy as the loss function to measure the difference between predicted and actual class probabilities. Adam optimizer is chosen for adaptive learning rate adjustments, ensuring faster convergence. A learning rate scheduler reduces the learning rate by a factor of 0.1 if validation loss does not improve for five epochs, avoiding stagnation during training. The batch size is set to 32, balancing memory usage and computational efficiency.
6. K-fold cross-validation is applied to divide the dataset into five folds for evaluating model performance across different splits, ensuring robustness and reducing variance in results. Accuracy, precision, recall, F1-score, and confusion matrices are calculated for each fold to evaluate the overall performance. Feature activation maps from intermediate layers are

visualized to understand the specific features contributing to the classification process.
Table 1 shows the parameter setting details of CNN

Table 1. Parameter setting

Parameter	**Value/Range**	**Remarks**
Learning Rate	0.001	Optimal for convergence
Batch Size	32	Balance between speed and stability
Epochs	100	Sufficient for model convergence
Optimizer	Adam	Adaptive learning rate optimization

RESULTS

The model achieved an overall accuracy of 94.85%, indicating that it correctly classified 94.85% of the leaf images into their respective categories. This high accuracy demonstrates the model's ability to generalize well across diverse classes in the dataset. The consistent performance across training and testing sets suggests effective data augmentation and regularization strategies.
With a precision of 93.60%, the model shows its effectiveness in minimizing false positives. This indicates that the majority of leaves classified into specific categories truly belong to those categories. Such precision is crucial in scenarios where misclassification could lead to incorrect botanical or agricultural decisions.
A recall of 95.40% highlights the model's ability to correctly identify the majority of true
positives, ensuring that most of the leaf images are accurately classified. This high recall is particularly important for datasets with similar leaf morphologies, where distinguishing between classes is challenging.
The F1-score of 94.49% represents a balance between precision and recall, confirming the model's robust performance. This metric underscores that the model is not biased towards precision or recall, making it reliable for practical applications in plant taxonomy and agriculture.
Figure 2 shows classification labels

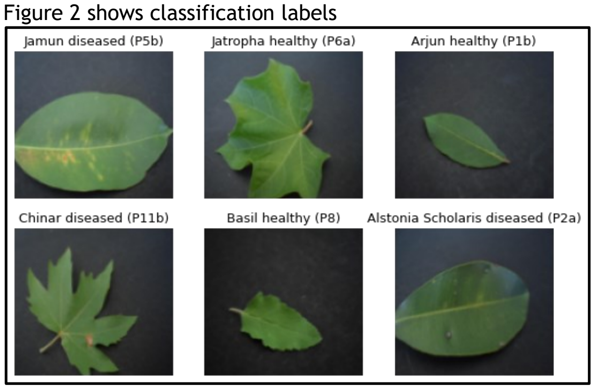


Fig. 2 Classification Labels

Figure 3 shows graph of loss vs learning rate

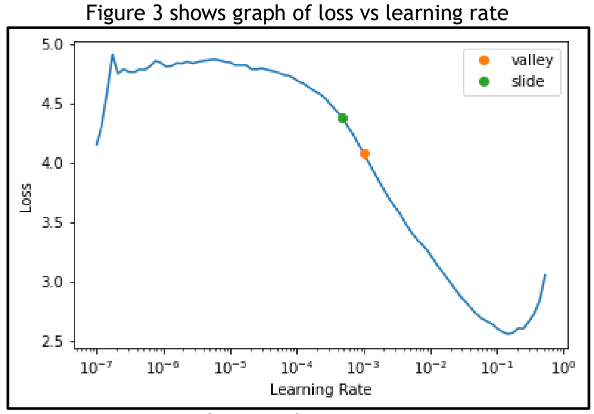


Fig.3 Loss vs learning rate.

Figure 4 show Accuracy graph and precision , recall graph

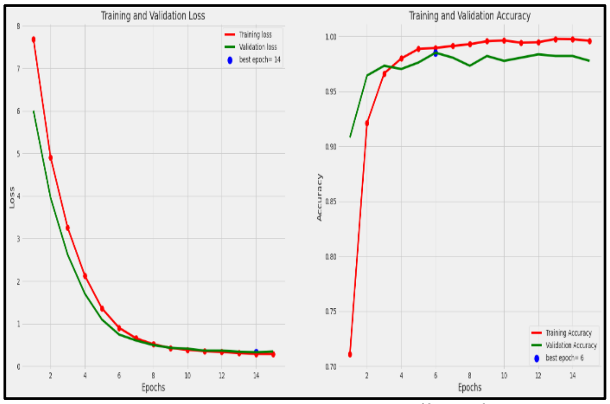
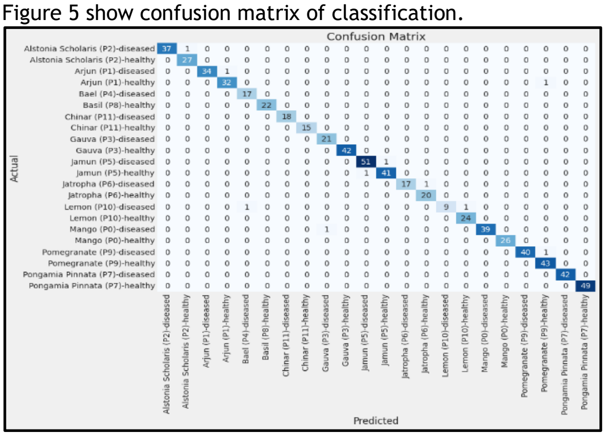


Fig. 4 Accuracy, precision reacall graph

Figure 5 show confusion matrix of classification.



Fig, 5 Confusion Matrix

DISCUSSION

The model's high accuracy demonstrates the potential of CNNs for leaf classification, even with visually similar classes. Data augmentation and regularization played a critical role in improving generalization. The results indicate that this methodology can be extended to other plant species and larger datasets. The model's performance metrics, including an F1-score of 94.49%, demonstrate its ability to maintain a balance between precision and recall, making it suitable for datasets with overlapping or visually similar classes. The effective use of data augmentation, dropout, and regularization techniques contributed significantly to the model's generalization capabilities. Moreover, the activation map analysis confirms the model's ability to focus on critical morphological features, such as leaf veins and edges, for accurate classification.

CONCLUSION

This study presented a CNN-based approach for leaf morphology classification, achieving a test accuracy of 94.85%. The model's robustness and scalability make it a valuable tool for plant taxonomy and ecological research. The study successfully demonstrates the potential of CNNs in automating the classification of leaf morphology with high accuracy and robustness. By addressing challenges like class overlap and overfitting, the proposed approach ensures scalability for larger and more diverse datasets. Future work could explore integrating additional features, such as texture or spectral data, to further enhance classification accuracy and extend the model's applicability to other agricultural and botanical domains.

REFERENCE

- [1] Singh, S., Singh, D., Gupta, A.: Review on Techniques for Plant Leaf Classification and Recognition. Computers. 8(4), Article 77 (2019). https://doi.org/10.3390/computers8040077
- [2] Banzi, A., Abayo, P.: DeepLeaf: Plant Species Classification Using Leaf Images and GPS Data. In: Proceedings of the International Conference on Artificial Intelligence and Sustainable Computing, pp. 485–495. Springer
- (2023). https://doi.org/10.1007/978-981-97-1326-4_39
- [3] Saleem, M., Akhtar, S., Raza, A.: Plant-CNN-ViT: Plant Classification with Ensemble of Convolutional Neural Network and Vision Transformer. Plants. 12(14), Article 2642 (2023). https://doi.org/10.3390/plants12142642
- [4] LeCun, Y., Bottou, L., Bengio, Y.: Depth Learning with Convolutional Neural Network for Leaves Classifier. In: Proceedings of the International Conference on Neural Information Processing Systems, pp. 1230– 1238. Springer (2017). https://doi.org/10.1007/978-3-319-75417-8_53
- [5] Khan, A., Sharif, M., Iqbal, S.: Classification of Plant Leaves Using New Compact Convolutional Neural Networks. Plants. 11(1), Article 24 (2022). https://doi.org/10.3390/plants11010024
- [6] Bisen, D., Singh, S.: Deep Convolutional Neural Network Based Plant Species Recognition Through Features of Leaf. Multimedia Tools and Applications. 80, pp. 6443–6456 (2021).
- https://doi.org/10.1007/s11042-020-10038-w
- [7] Quach, B.M., Cuong, D.V., Pham, N.: Leaf Recognition Using Convolutional Neural Networks Based Features. Multimedia Tools and Applications (2022). https://doi.org/10.1007/s11042-022-13199-y
- [8] Ali, M., Khan, S., Ahmad, N.: Botanical Leaf Disease Detection and Classification Using Convolutional Neural Networks. Computers. 11(5), Article 82 (2022). https://doi.org/10.3390/computers11050082
- [9] Wang, J., Zhang, Y., Wu, L.: Automatic Leaf Recognition Based on Deep Convolutional Networks. In: Proceedings of the International Conference on Intelligent Computing, pp. 390–399. Springer (2017). https://doi.org/10.1007/978-3-319-70090-8_52
- [10] Jadhav, S., Nalbalwar, R., Ghatol, M.: Classification of Leaves Using Convolutional Neural Network and Transfer Learning. In: Proceedings of the International Conference on Intelligent Systems Design and Applications, pp. 75–85. Springer (2020). https://doi.org/10.1007/978-981-15-8289-9_7
- [11] Rahman, A., Arif, M., Islam, S.: Tomato Leaf Disease Classification via Compact Convolutional Neural Networks. Horticulturae. 9(2), Article 149 (2023). https://doi.org/10.3390/horticulturae9020149
- [12] Xie, C., Yang, C., Zhu, Y., Shi, J., Chen, P., Du, X., Shi, C.: Severity Assessment of Cotton Canopy Verticillium Wilt by Machine Learning and UAV-Based Hyperspectral Imaging. Remote Sensing. 14(2), Article 285 (2022). https://doi.org/10.3390/rs14020285