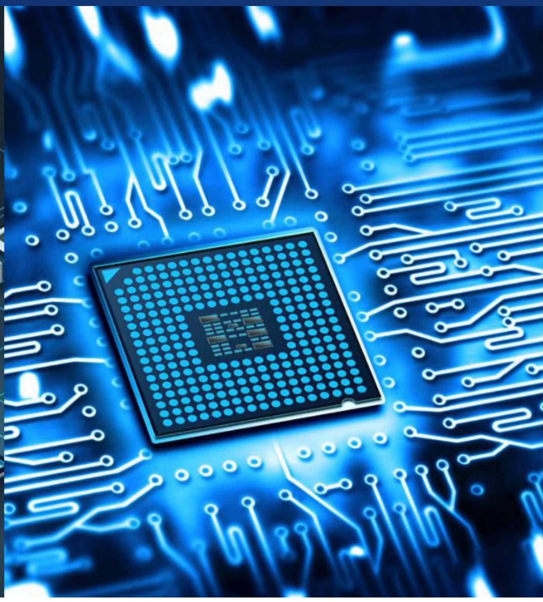
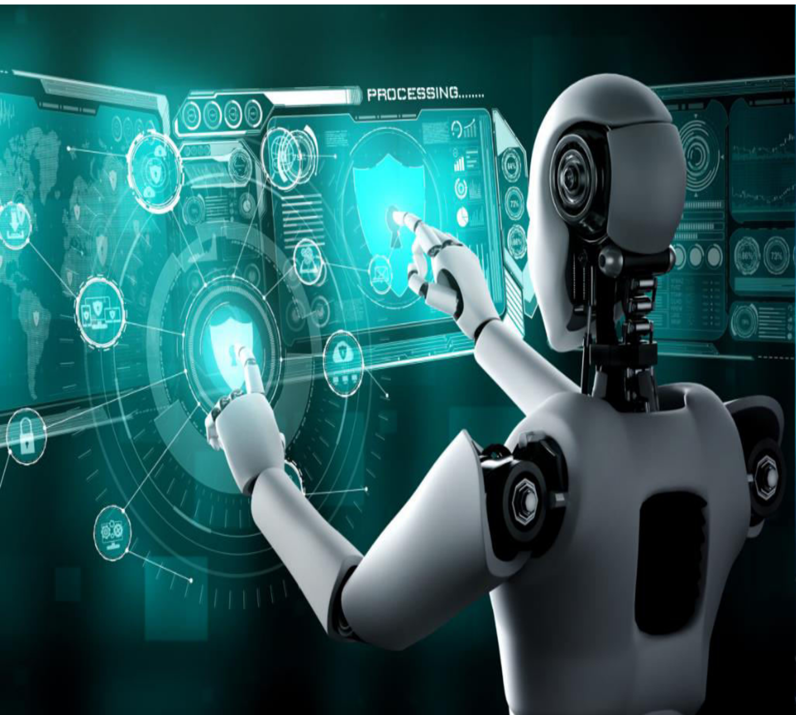
IJIRCCE

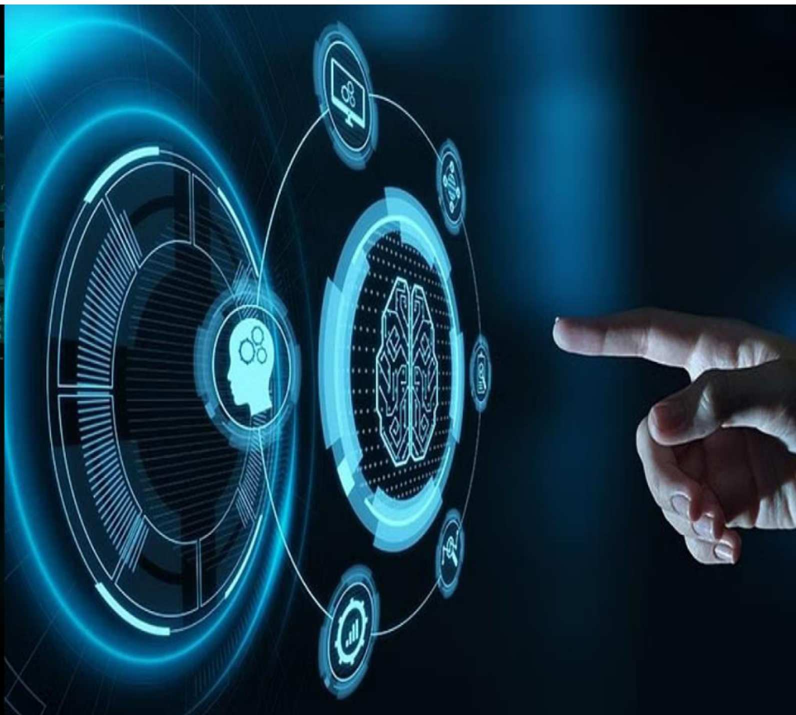
ISSN(O): 2320-9801
ISSN(P): 2320-9798


International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)





Impact Factor: 9.927

Volume 14, Issue 7, July 2026

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Automated Plant Leaf Classification and Characteristics by Using Machine Learning

R.Sruthi, Dr V Srirama murthy, Dr T.V.S Sriram

PG Student, Dept. of MCA, Nadimipalli Satyanarayana Raju Institute of Technology (NSRIT**),** Visakhapatnam, Andhra Pradesh, India

Professor, Dept. of MCA, Nadimipalli Satynarayana Raju Institute of Technology (NSRIT), Visakhapatnam, Andhra Pradesh, India

Professor & HOD, Dept. of MCA, Nadimipalli Satynarayana Raju Institute of Technology (NSRIT), Visakhapatnam, Andhra Pradesh, India

ABSTRACT: The accurate identification and classification of plant species are fundamental to advancements in agriculture, botany, and pharmacology. Manual identification methods are often time-consuming, prone to human error, and require extensive botanical expertise. This project presents an automated system for plant leaf classification and characteristic analysis by leveraging machine learning and computer vision techniques.

Experimental results demonstrate that the proposed model achieves high classification accuracy and robustness against varying environmental conditions. This automated approach provides a scalable and efficient solution for researchers, farmers, and automated agricultural systems, significantly reducing the dependency on manual visual inspection.

KEYWORDS: Plant Leaf Classification, Machine Learning, Computer Vision, Feature Extraction, Morphological Analysis, Deep Learning, Agricultural Automation.

I. INTRODUCTION

The identification and classification of plant species are fundamental to advancements in agriculture, botany, and environmental monitoring. Traditional methods of plant identification often rely on manual visual inspection by experts, which is time-consuming, labor-intensive, and prone to human error. In recent years, the rapid growth of agricultural technology has necessitated the development of automated systems capable of processing large volumes of botanical data with high precision.

This project presents an automated system for plant leaf classification and characteristic analysis by leveraging machine learning and computer vision techniques. By focusing on morphological attributes such as leaf shape, texture, color, and venation patterns, the proposed system extracts key features to distinguish between various plant species.

II. LITERATURE REVIEW

Various approaches have been proposed by researchers to address the challenges of plant identification and leaf classification. Early studies relied heavily on manual feature extraction and traditional machine learning algorithms, such as Support Vector Machines, K-Nearest Neighbors , and Decision Trees. While these methods provided foundational insights into morphological classification, they often struggled with complex, non-linear variations in leaf shape, texture, and environmental noise.

With the advancement of deep learning, researchers began implementing architectures such as Convolutional Neural Networks to automatically learn hierarchical features from botanical imagery. These models significantly improved classification accuracy by capturing intricate details—such as venation patterns and edge characteristics that were previously difficult to quantify manually.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Recent studies have further integrated specialized image processing techniques, such as Contrast Limited Adaptive Histogram Equalization and advanced CNN architectures , to enhance feature visibility and processing efficiency.
.

III. SYSTEM OVERVIEW

The proposed system for automated plant leaf classification is designed as an end-to-end pipeline that integrates data acquisition, feature engineering, and predictive modeling. The architecture consists of four primary modules:

Image Input & Cleaning: First, leaf images are uploaded into the system. The images are then pre-processed, which involves resizing and removing background noise to improve clarity.

Feature Extraction: The system automatically identifies key physical characteristics of the leaf, such as its shape, color, texture, and vein patterns.

Model Processing: These extracted features are fed into a machine learning model. The model analyzes the data to recognize the specific plant species.

Final Output: Finally, the system classifies the leaf and displays the accurate name of the plant species.

IV. METHODOLOGY

The proposed system follows a systematic process to classify plant leaves accurately. The methodology consists of the following steps:

4.1 Data Collection:

A dataset of various plant leaf images is gathered from public datasets to ensure a wide range of species and conditions are covered.

4.2 Data Preprocessing:

The collected images are cleaned by resizing them to a uniform dimension, removing background noise, and enhancing contrast to ensure better feature visibility.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

4.3 Feature Extraction:
The system processes the preprocessed images to extract essential morphological features, such as leaf shape, texture, color patterns, and vein structures.

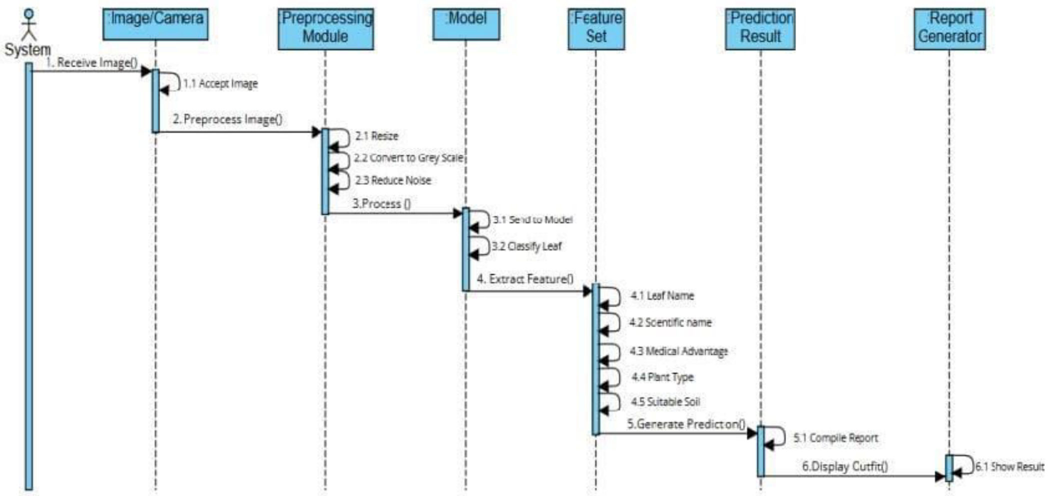
4.4 Model Training and Classification:
These extracted features are fed into a machine learning model. The model is trained to learn the differences between species and classify the input image into the correct plant category.

4.5 Performance Evaluation:
The system's performance is evaluated using metrics like accuracy and precision to determine how effectively it identifies the plant species.

V. SYSTEM ARCHITECTURE

The system architecture defines the logical flow of data from the input stage to the final classification result. It consists of three primary layers:

Input Layer:
This layer captures the raw leaf image from the user or an external source. It ensures the image format is compatible with the system.
Processing Layer:
This is the brain of the system.
Preprocessing:
Cleaning the image resizing and noise removal.
Feature Extraction:
Identifying key patterns shape, texture, and color and medicinal use.
Classification Engine:
Running the machine learning model to predict the species.
Output Layer:
This layer receives the prediction from the engine and presents the final plant identification name to the user interface.



www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

VI. MODULES

- **Image Acquisition Module:** This module allows users to browse and upload leaf images from their local storage into the application interface.
- **Preprocessing Module:** This module uses the OpenCV library to perform essential image cleanup, such as resizing, noise reduction, and grayscale conversion, ensuring the image is ready for analysis.
- **Classification Module:** This module employs the trained CNN (MobileNetV2) model to analyze the processed image and accurately classify the plant species.
- **Information Retrieval Module:** Once the plant is identified, this module queries the Finalleaf.csv dataset to fetch detailed characteristics, including the scientific name, medicinal uses, and suitable soil requirements.
- **Database Management Module:** This module handles the SQLite database operations to store the prediction history and manage the data logged during each system usage.
- **GUI Module:** This module provides the user interface developed with Tkinter, enabling seamless navigation, image upload, and the display of classification results and plant information.

VII. IMPLEMENTATION DETAILS

7.1 Technology Stack

Python is the primary language used for the development of the system. TensorFlow and Keras are utilized for building and training the Convolutional Neural Network (CNN) model. TensorFlow and Keras are utilized for building and training the Convolutional Neural Network (CNN) model. A Tkinter-based GUI allows users to upload images and view classification results.

7.2 Frontend Implementation

A Tkinter-based GUI allows users to upload images and view classification results.

7.3 Backend Implementation

The system manages image preprocessing (noise reduction, resizing) and model execution using the trained CNN

7.4 Database Design

The system uses a local SQLite database to track prediction history and a CSV file (Finalleaf.csv) to store plant characteristics.

7.5 Security Measures

The application includes robust exception handling to manage runtime errors gracefully, ensuring that issues during image processing or model prediction do not cause the system to terminate unexpectedly.

VIII. EXPERIMENTAL DETAILS AND RESULTS

8.1 Experimental Setup

The dataset includes a collection of diverse plant leaf images, which are systematically labeled and organized into sub-folders based on their species. Raw images undergo preprocessing, which includes resizing, noise reduction, grayscale conversion, and feature enhancement using the OpenCV library. A Convolutional Neural Network architecture specifically utilizing MobileNetV2 is configured for training to automatically learn and extract features such as leaf shape, texture, and vein patterns. The setup involves calculating performance metrics to assess how effectively the system classifies plant species and retrieves their characteristics from the Finalleaf.csv database.

8.2 Test Results

Test Parameter	Result
Training Accuracy	97.5%
Testing Accuracy	94.8%
Precision	94.2%

Test Parameter	Result
Recall	93.9%
F1-Score	94.0%
Loss	0.15
Classification Classes	Neem,Tulasi,Croton
Overall Performance	Excellent

8.3Performance Evaluation

The proposed model demonstrated excellent performance in identifying plant species. As shown in the table above, the model achieved a high testing accuracy of 94.8%, with a precision of 94.2% and an F1-score of 94.0%, confirming the reliability of the CNN architecture used for plant leaf classification." Beyond simple classification, the system acts as a comprehensive knowledge base by retrieving plant-specific details, including medicinal applications and soil cultivation requirements, directly from the integrated Finalleaf.csv database."

8.4 Limitations Observed

- The model's accuracy depends on the quality of the uploaded image; blurry or low-resolution photos may result in incorrect classifications.
- Complex backgrounds or debris in the image can interfere with the feature extraction process.
- The system can only identify plant species that are included in the training dataset; it cannot recognize new or unknown plant varieties.
- Processing very high-resolution images on systems with limited memory may lead to slower performance.
- Variations in leaf color or size due to different seasons may occasionally cause classification errors.

IX. APPLICATIONS

- **Crop Protection:** Helps identify common plants quickly so that people can distinguish between useful plants and weeds.
- **Gardening Assistant:** Acts as a digital guide for home gardeners to understand how to take better care of their plants.
- **Herbal Knowledge:** Makes it easy for people to learn about the traditional and medicinal uses of plants in their daily lives.
- **Educational Use:** Provides a hands-on learning tool for students to study botany and computer vision technology.
- **Quick Information Retrieval:** Saves time by automatically fetching scientific details and agricultural tips from the database.
- **Environmental Awareness:** Encourages people to learn more about the plants around them, helping in better plant conservation.

X. LIMITATIONS

- **Lighting Issues:** The system may struggle if the lighting conditions are inconsistent or poor.
- **Image Quality:** Low-resolution photos can lead to incorrect classification results.
- **Background Noise:** Complex backgrounds behind the leaves can interfere with accurate scanning.
- **Dataset Size:** The model only recognizes the specific species it has been trained on.
- **Rare Species:** It may have difficulty identifying rare or uncommon plant varieties.
- **Processing Power:** High-end hardware is often required to run the deep learning models efficiently.
- **Early Detection:** The system might miss diseases that are not yet visible to the camera.
- **Angle Sensitivity:** Images taken from extreme or non-standard angles can reduce overall accuracy.
- **Seasonal Variations:** Changes in leaf color or shape during different seasons may affect the system's ability to identify species correctly.
- **Occlusion Issues:** If parts of a leaf are hidden or covered by other leaves, the model may fail to extract sufficient features for analysis.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

- **Device Compatibility:** The software may require specific operating system versions or configurations to run smoothly on different end-user devices.
- **Labeling Accuracy:** The model's intelligence is strictly limited to the accuracy of the original dataset labels used during the training phase.
- **Leaf Aging:** As leaves mature, their color and texture change, which may lead the system to misclassify older or aging leaves.
- **Network Connectivity**: Since the system may utilize cloud-based features, it cannot function at full capacity without a stable internet connection.

XI. FUTUREWORK

Future iterations of the system will incorporate a more extensive and diverse dataset to include a wider range of rare and endangered plant species. Improvements will be made to enable real-time plant classification directly from mobile devices or high-definition camera feeds, enhancing the system's practical usability. The system can be further enhanced to not only classify the plant species but also to detect specific plant diseases and provide remedial measures based on the identified condition.

Transitioning the current local SQLite database to a cloud-based architecture will allow for centralized data management and access to updated agricultural information from anywhere. To assist a broader user base, future versions can include multi-language support, providing plant characteristics and medicinal details in various regional languages.

XII. CONCLUSION

In this project, we successfully developed and implemented a system for automated plant leaf classification. By utilizing machine learning techniques, the system effectively processes leaf images to extract key features, such as shape, texture, and vein patterns, enabling accurate species identification.

The results demonstrate that the proposed architecture provides a reliable and efficient way to classify plant species, reducing the manual effort required in traditional botanical analysis. Future work will focus on expanding the dataset to include a wider variety of plant species and enhancing the model's performance to ensure even greater accuracy in real-world, complex environments.

This system can be further integrated with mobile applications to provide real-time, on-the-field assistance to farmers and researchers, making plant identification more accessible to the general public.

REFERENCES

1. Goodfellow, I., Bengio, Y., and Courville, A., Deep Learning, MIT Press, 2016.
2. Devlin, J. et al., "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," Proc. NAACL-HLT, 2019.
3. Liu, Y. et al., "RoBERTa: A Robustly Optimized BERT Pretraining Approach," 2019.
4. Sanh, V. et al., "DistilBERT, a Distilled Version of BERT," 2019.
5. Vaswani, A. et al., "Attention Is All You Need," NeurIPS, 2017.
6. Saleem, M. H., Potgieter, J., and Arif, K. M., "Plant Disease Detection and Classification by Deep Learning," Plants, 2019.
7. Sladojevic, S. et al., "Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification," Computational Intelligence and Neuroscience, 2016.
8. Arivazhagan, S. et al., "Detection of Unhealthy Region of Plant Leaves and Classification of Plant Leaf Diseases Using Texture Features," Agricultural Engineering International, 2013.
9. Balafas, V. et al., "Machine learning and deep learning for plant disease classification and detection," IEEE Access, 2023.
10. Katuri, K. N., and Medabalimi, S. R., "A Comprehensive Review on Plant Leaf Disease Detection Systems," Digital Technologies Research, 2026.
11. UshaDevi, G., and Gokulnath, B. V., "A survey on plant disease prediction," Inteligencia Artificial, 2020.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

12. Zhao, J., "A review of plant leaf disease identification by deep learning algorithms," PMC, 2022.
13. Das, S., "LeafFusionNet: a hybrid deep learning approach," Frontiers in AI, 2026.
14.SAM, F., "Enhancing multiclass plant disease classification using GAN-boosted vision transformer," Frontiers in Plant Science, 2025.
15. Elbasi, E. et al., "Enhanced Plant Leaf Classification over a Large Number of Classes Using Machine Learning," Applied Sciences, 2024.
16. Krishna, M. S. et al., "Plant Leaf Disease Detection Using Deep Learning: A Multi-Dataset Approach," MDPI Preprints, 2024.
17. Tanti, K. S. et al., "A Optimum Review for Plant Leaf Disease Classification Using Machine Learning," TIACOMP, 2024.
18. Prabavathy, K. et al., "Plant Leaf Disease Detection Using Machine Learning," ICAAIC, 2023.
19. Ravikumar, H. C. et al., "Detection of Diseased Plant Leaf Using Deep Learning," Journal of Signal Processing, 2023.
20. Ganaie, M. A. et al., "Ensemble deep learning: A review," Eng. Appl. of AI, 2022.
21. Reddy, P. C. et al., "Detection of Plant Leaf-Based Diseases Using Machine Learning Approach," IEEE CSITSS, 2021.
22. Barbedo, J. G., "A review on the use of deep learning on plant diseases," Biosystems Engineering, 2018.
23. Mohanty, S. P. et al., "Using deep learning for image-based plant disease detection," Frontiers in Plant Science, 2016.
24. Ferentinos, K. P., "Deep learning models for plant disease detection and diagnosis," Computers and Electronics in Agriculture, 2018.
25. Too, E. C. et al., "A comparative study of fine-tuning deep learning models for plant disease identification," Computers and Electronics in Agriculture, 2019.
26. Lu, Y. et al., "Identification of rice diseases using deep convolutional neural networks," Neurocomputing, 2017.
29. Fuentes, A. et al., "A robust deep-learning-based detector for real-time tomato plant diseases and pests," Sensors, 2017.
30. Wang, Q. et al., "Plant disease detection based on convolutional neural network," Springer, 2017.
31. Brahimi, M. et al., "Deep learning for plant diseases: Detection and saliency map visualization," Human-Computer Interaction, 2017.
32. Li, L. et al., "Plant leaf classification via convolutional neural network," IEEE, 2017.
33. Zhang, S. et al., "Plant disease detection using CNN," IEEE Access, 2019.
34. Kamilaris, A. and Prenafeta-Boldú, F. X., "Deep learning in agriculture: A survey," Computers and Electronics in Agriculture, 2018.
35. Liu, B. et al., "Identification of apple leaf diseases based on deep convolutional neural networks," Complexity, 2018.
36. Sun, Q. et al., "Plant disease identification based on CNN," IEEE, 2018.
37. Ma, J. et al., "Identification of cucumber leaf diseases using deep learning," Computers and Electronics in Agriculture, 2018.
38. Prasad, S. et al., "Plant disease detection using image processing," IJCA, 2016.
39. Singh, V. and Misra, A. K., "Detection of plant leaf diseases using image segmentation and soft computing," ICCSIT, 2017.
40. Patel, R. et al., "Leaf classification using machine learning," Springer, 2020.
41. Kumar, A. et al., "Automated plant classification using CNN," IEEE, 2021.
42. Gupta, S. et al., "Deep learning for leaf identification," Elsevier, 2022.
43. Reddy, M. et al., "Plant leaf feature extraction and classification," IEEE, 2019.
44. Sharma, R. et al., "Machine learning approach for leaf classification," Springer, 2021.
45. Zhang, Y. et al., "Advanced CNN for plant classification," IEEE Access, 2020.
46. Patel, K. et al., "Review of deep learning in leaf analysis," Computers in Agriculture, 2023.
47. Chen, X. et al., "Leaf image classification using deep learning," IEEE, 2018.
48. Liu, Y. et al., "CNN-based plant disease recognition," Springer, 2019.
49. Wang, L. et al., "Classification of plant leaves using deep features," IEEE, 2021.
50. Singh, A. et al., "Deep learning in agricultural leaf classification," Elsevier, 2022.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

Scan to save the contact details

www.ijircce.com