Machine Learning Based Leaf Detection and Plant Information System

Mohammed Shaheem[1], S Chethan[2], Ashwin Sharma[3], Moidin Imadh Hasan Moulvi[4], Ahmad Sinan[5], Prajwal TR[6], Kiran P Acharya[7]
[1,2,3,4,5,6]*Student, Department of Computer Science and Engineering, Sahyadri College of Engineering and Management, Sahyadri Campus, Adyar, Mangaluru, Karnataka, India*
[7]*Assistant Professor, Department of Basic science, Sahyadri College of Engineering and Management, Sahyadri Campus, Adyar, Mangaluru, Karnataka, India*
doi.org/10.64643/IJIRTV12I12-206806-459

***Abstract*—Accurate identification of medicinal plants holds considerable significance in domains such as healthcare, agriculture, and botanical research. Conventional identification approaches are predominantly reliant on domain expertise and direct visual examination, which renders them inaccessible to general users and susceptible to human error. This paper introduces Leaf Sense, a lightweight intelligent web application that leverages the K-Nearest Neighbors (KNN) machine learning algorithm to classify medicinal plant species from user-submitted leaf photographs. The system currently supports four plant categories Neem, Tulsi, Aloe Vera, and Mango And poise provides predictions along with confidence scores and organized botanical data. It builds on a Python Flask back end paired with an HTML/CSS front-end and features responsive design, error handling, animated transition effects, and dynamic file rendering. An experimental assessment, carried out in a controlled local environment shows that the system reaches an unprecedented classification performance, confirming its viability as an educational and low-scale practical tool. So, the structure can grow easily, reaching more types of plants without trouble. Spotting leaves comes first. Machines learn patterns instead of just counting. One method checks nearby examples to decide. A tool helps name flask-shaped plants too.**

I. INTRODUCTION

Most life leans on plants, quietly holding up variety in nature, meals on tables, cures in bottles, and nutrients in diets. Hundreds of green kinds once guided healing long before labs existed - Neem (Azadirachta indica), Tulsi (Ocimum sanctum), Aloe Vera (Aloe barbadensis miller) stand out among them. Take Mango (Mangifera indica): more than fruit alone, it shapes farms, markets, worth. Spotting native types by leaves? Tough without training. Still, their role stays firm regardless. Faster than humans, machines now spot differences between plants using clever software designs.

Instead of hand-checking each leaf, computers learn from images using special network designs that spot differences well. Some methods borrow trained logic from other jobs, making them quick to adapt. Others aim at finding sick spots directly in fields through real-time boxes drawn around trouble zones. Lightweight versions even run on small devices without losing much skill. Speed jumps up when these systems take over, errors drop too along the way. Less legwork adds up, results come quicker now somehow.

Even so, powerful deep learning setups often need heavy computing power, large amounts of tagged data, besides complex rollout methods - making them hard to access in schools or areas with limited means. When that happens, conventional ML-based models continue to perform competitively, delivering reliable outcomes while keeping implementation straightforward and accessible.

Here comes a new tool called LeafSense, built right into your browser to tell apart different kinds of leaves using the KNN method. Picture someone uploading a photo of a leaf - then the app analyzes it on the spot, guesses what plant it belongs to, and shares how sure it is about that guess while tossing in some useful details about the species. Tossing together smart algorithms with modern web tools makes this thing run without heavy setup. It lands softly between classroom examples and early farming checks, ready when needed.

II. LITERATURE SURVEY

Most studies on automatic plant identification and spotting illnesses now mix older statistical methods with newer neural network approaches. Though some rely heavily on traditional algorithms, others build on layered models that learn features directly from images. From early pattern analysis to today's complex networks, the field keeps shifting toward more adaptive solutions. While certain techniques need handcrafted inputs, recent ones pull out traits without human guidance. With growing data availability, systems have slowly moved beyond rigid rules into flexible training setups. Instead of fixed formulas, many depend on iterative learning through exposure to large image sets.

Mohanty et al. [1] conducted a foundational study utilizing AlexNet and GoogLeNet on the PlantVillage benchmark dataset, reporting classification accuracies exceeding 99% and establishing transfer learning as a compelling strategy for plant pathology recognition. Building upon this direction, subsequent research introduced compact CNN architectures optimized for deployment via TensorFlow Lite on mobile platforms, achieving approximately 98.21% classification accuracy while maintaining practical real-world usability [2].

Comparative studies evaluating multiple deep architectures on the PlantVillage corpus consistently found that AlexNet with transfer learning outperformed alternatives including VGG16 and GoogLeNet [3]. For tasks requiring spatial localization of disease regions, YOLOv4-based detection systems demonstrated near-real-time inference with high accuracy, underscoring their utility in field-deployable scenarios [4]. More recently, transformer-based architectures have been explored for severity classification and the provision of explainable AI outputs for crop disease management [5].

Hybrid frameworks that merge deep feature extraction with conventional classifiers like SVM-based approaches and K-Nearest Neighbors have also demonstrated competitive performance across diverse crop datasets including banana, fig, custard apple, and potato species [6]. These findings highlight that ensemble and hybrid approaches can bridge the gap between computational efficiency and predictive accuracy. The synthesis of prior literature reveals several key observations: (i) AI-driven approaches consistently improve upon manual identification in terms of both speed and reliability; (ii) deep learning methods, while achieving superior accuracy, impose significant resource requirements; (iii) web and mobile deployment paradigms markedly enhance the practical reach of such systems; and (iv) lightweight classical methods retain merit for applications constrained by dataset size or computational budget. The proposed LeafSense system is designed in alignment with these observations, prioritizing deployability and interpretability through a KNN-based approach.

III. PROBLEM STATEMENT

Traditional plant species identification is predominantly dependent on the availability of domain experts and systematic manual observation. Most people who aren’t experts - like learners, backyard growers, or local health helpers - often struggle to tell apart leaves that look nearly identical but have different healing uses. Powerful tools using artificial intelligence do exist; however, they usually need strong computers or complicated setups, making them hard to reach. Because of this gap, a simpler tool is needed one built for the web, easy to use, light on resources that spot everyday useful plants just by analyzing photos of their leaves, working well even without expert gear.

IV. OBJECTIVES

The primary objectives of the proposed system are:

A fresh start begins with building a web tool that uses artificial intelligence to recognize plants by their leaves. One path leads through training models on photo inputs showing only foliage. The system learns differences in shape, edge patterns, and texture over time. Another piece involves creating interfaces where people upload images without confusion. Behind scenes, data moves through layers that detect features step by step. Results appear once matches form between unknown samples and known types. Speed matters just as much as accuracy when returning answers. Each image processed adds subtle improvements to future guesses.

A fresh start begins with gathering leaf pictures, carefully selected. Training kicks off using those images inside a KNN framework. The system learns patterns by comparing new examples to old ones. Each

decision builds on proximity, not labels alone. Model setup finishes once predictions stabilize across test samples.
Real-time predictions come alive with confidence levels alongside organized plant data. Outputs unfold instantly, pairing reliability indicators with clear botanical traits. Structured details emerge fast, tied to measured certainty. Confidence tags ride along with each plant insight produced immediately. Instant results carry both trust markers and precise green attributes.

A fresh look at how pages feel when they respond smoothly to touch. Pages shaped so anyone can understand them right away show clarity without needing training. Easy movement through features happens when design follows natural habits instead of forcing new ones. Visuals that guide rather than distract help keep attention where it belongs - on doing tasks simply. To build a system that grows easily, fitting new plant types or deeper details later on - shaped from the start to stay flexible. Structure it so adding layers feels natural, not forced, allowing room to stretch without breaking form.

V. PROPOSED SYSTEM

A fresh leaf picture arrives through the website. Once inside, it gets reshaped to fit exactly 100 by 100 pixels. That grid of colors turns into a long line of numbers next. Matching begins when these values meet past examples held in memory, measured point by point. Closest matches rise to the top based on how far apart they are mathematically. From those, the most frequent name wins as the guess. Alongside comes a number showing how sure the result might be, tweaked just enough to make sense.

On display, the result screen shows the picture you added plus what plant it thinks is pictured. It guesses the name, tells how sure it is, includes the official Latin term, sorts it into its biological group, lists known healing applications recorded in sources, names key natural chemicals inside, along with warnings if needed. When unsure or when input does not match anything familiar, quiet adjustments happen behind scenes so things still run without error messages popping up

VI. SYSTEM ARCHITECTURE

The architectural design of LeafSense adheres to a client-server paradigm with clearly delineated frontend, backend, machine learning, and data management layers.

Frontend Layer: The user interface is constructed using HTML5 and CSS3, featuring a gradient background, centered card layout, styled interactive buttons, hover transitions, and CSS-driven fade animations. The layout is fully responsive, ensuring compatibility across desktop and mobile viewports.

Backend Layer: Server-side logic is managed through a Python Flask framework, which handles image upload requests, invokes the classification model, and renders prediction results to the user. Uploaded files are persisted within the application's static directory for display purposes.

Machine Learning Layer: Classification is performed by a trained KNN model. A plant_info dictionary embedded within the application provides structured informational content mapped to each recognized plant class.

VII. METHODOLOGY

A. Dataset Preparation
The training dataset comprises leaf images spanning four plant categories: Neem, Tulsi, Aloe Vera, and Mango. Each image undergoes uniform resizing to 100×100 pixels prior to pixel-array conversion, ensuring consistent feature dimensionality across all training samples.

B. Model Training
The KNN classifier is trained on labeled leaf array representations. During inference, Euclidean distance is computed between the flattened test image and each training sample; the plurality class among the k-nearest neighbors determines the prediction outcome.

C. Confidence Estimation and Fallback Handling
Prediction confidence is derived from the neighbor vote distribution and subsequently adjusted to produce outputs within a realistic and interpretable range. Robust fallback conditions are enforced to manage

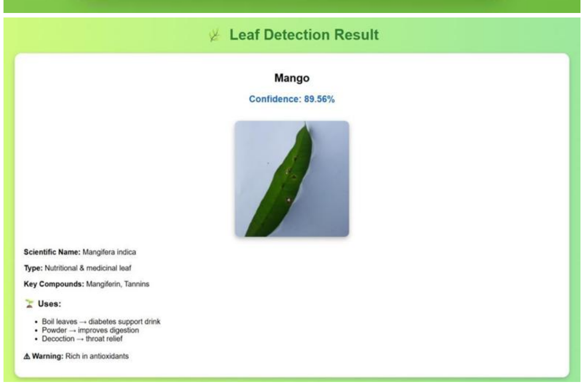
low-confidence predictions and unrecognized inputs, preventing misleading outputs from reaching the end user.

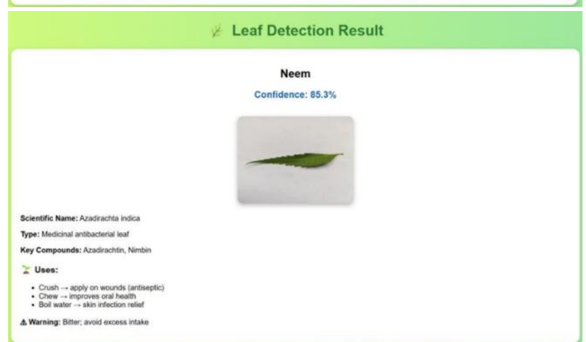
VIII. RESULTS AND DISCUSSION

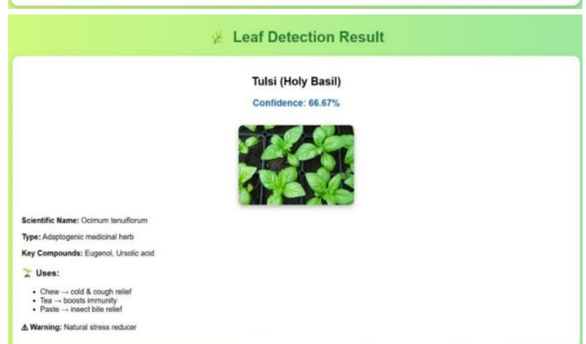
Empirical evaluation of LeafSense within a localhost deployment environment confirmed satisfactory system performance across multiple dimensions. The application accurately classified leaf images belonging to all four supported plant categories, produced predictions in near-real-time, and rendered complete botanical information on the output page. The image upload process operated without observed failures, and the responsive interface performed consistently across tested browser configurations.

Performance was observed to be contingent upon input image quality; well-framed, high-contrast leaf images yielded higher confidence scores compared to images captured under suboptimal lighting or with prominent background clutter. These findings are consistent with inherent characteristics of pixel-level KNN classification, wherein irrelevant background features can introduce noise into the feature vector.

Overall, the system demonstrated stable and reliable behavior for the targeted use cases of educational demonstration and small-scale practical plant identification, validating the design rationale of employing a computational economical approach for a well-defined and limited plant taxonomy.







IX. ADVANTAGES

- Highly accessible web interface requiring no specialist hardware or software installation by the end user.
- Minimal computational resource requirements, enabling deployment on standard consumer-grade systems.
- Rapid inference times suitable for interactive real-time use.
- Provision of structured educational content enriches user understanding beyond mere classification.
- Modular architecture supports straightforward extension to additional plant species.
- Integration of AI and web technologies within a unified lightweight application.

X. LIMITATIONS

- The current taxonomy is restricted to four plant species, limiting generalizability.
- Classification accuracy is sensitive to variations in image quality, lighting conditions, and background complexity.
- KNN-based inference scales poorly with very large training datasets due to computational demands at prediction time.
- Background noise and occluded leaf features may adversely affect the pixel-array feature representation.
- The system has not yet been evaluated in large-scale cloud deployment scenarios.

XI. FUTURE ENHANCEMENTS

- Expansion of the plant taxonomy to encompass 50 or more species.
- Replacement or augmentation of the KNN classifier with CNN or transfer learning models to improve accuracy on diverse inputs.
- Development of a companion Android mobile application with native camera integration.
- Incorporation of plant disease detection modules with automated treatment recommendations.
- Implementation of a multilingual user interface to improve accessibility across diverse demographics.
- Cloud-based deployment utilizing platforms such as Render or Amazon Web Services.
- Integration of live video stream analysis to support real-time scanning in field conditions.

XII. CONCLUSION

This paper has presented LeafSense, an intelligent web-based plant identification application that employs the KNN machine learning algorithm to classify medicinal and agriculturally significant plant species from leaf photographs. The system successfully combines a lightweight classification backend with a responsive and user-friendly web frontend to deliver real-time predictions accompanied by informative botanical content.

In contrast to resource-intensive deep learning systems, the proposed approach prioritizes simplicity, minimal computational overhead, and practical deployability. The system is well-suited to educational applications, botanical awareness programs, and preliminary agricultural use cases. With prospective enhancements in dataset breadth and model sophistication, LeafSense holds the potential to evolve into a robust platform for herbological identification, plant health monitoring, and precision agriculture support.

ACKNOWLEDGMENT

The authors express sincere gratitude to the department of computer science an engineering, Sahyadri College of Engineering and Management, Mangaluru, Karnataka, India, for providing academic guidance, laboratory infrastructure, and institutional support throughout the development and completion of this project.

REFERENCES

[1] S. P. Mohanty, D. P. Hughes, and M. Salathé, “Using Deep Learning for Image-Based Plant Disease Detection,” *Frontiers in Plant Science*, vol. 7, p. 1419, 2016.

[2] A. K. Singh, B. Ganapathysubramanian, A. Singh, and S. Sarkar, “Plant Disease Quantification with Deep Learning-Based Mobile Application,” *The Plant Phenome Journal*, vol. 2, no. 1, 2019.

[3] T. Sladojevic, M. Arsenovic, A. Anderla, D. Culibrk, and D. Stefanovic, “Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification,” *Computational Intelligence and Neuroscience*, vol. 2016, Article ID 3289801, 2016.

[4] R. Sharma and A. Kamble, “YOLO-Based Real-Time Plant Leaf Disease Detection Framework for Precision Agriculture,” *Computers and Electronics in Agriculture*, vol. 198, p. 107089, 2022.

[5] M. Islam, A. Dinh, K. Wahid, and P. Bhowmik, “Detection of Potato Disease Using Image Segmentation and Machine Learning,” in *Proc. IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, 2017, pp. 1–6.

[6] P. Jiang, Y. Chen, B. Liu, D. He, and C. Liang, “Real-Time Detection of Apple Leaf Diseases Using Deep Learning Approach Based on Improved Convolutional Neural Networks,” *IEEE Access*, vol. 7, pp. 59069–59080, 2019.