Volume XVII Number 4, 2025

A Highly Effective Deep Learning Tool for Identifying Plant Leaves

Adel Abdelhadi, Ouahab Kadri

Department of Computer Science, University of Batna 2, Algeria

Abstract

This work addresses pattern recognition in the agronomic domain, with a particular emphasis on identifying plant leaves using an adaptive neural network technique. We introduce a tool designed for two primary groups: botany researchers and a broader range of scientists applying it to plant identification and classification. We delve into the capabilities of Deep Learning, focusing on generalization abilities that enable accurate predictions on unseen data, which is essential for handling the variation in leaf shapes, sizes, and structures across species. The implementation details of these neural networks are described, including data preprocessing, network architecture design, training strategies, and evaluation techniques to ensure robustness and reliability in real-world applications.

Keywords

Pattern recognition, plant leaves, deep learning, classification, analysis, image processing, neural networks.

Abdelhadi, A. and Kadri, O. (2025) "A Highly Effective Deep Learning Tool for Identifying Plant Leaves", *AGRIS on-line Papers in Economics and Informatics*, Vol. 17, No. 4, pp. 3-9. ISSN 1804-1930. DOI 10.7160/aol.2025.170401.

Introduction

The problem of classifying plant leaves is closely related to the broader challenge of form recognition (FR), a fundamental step in human-machine communication. Image classification of plant leaves is complex due to several factors. First, forms exist in the physical world, and their digital transcription is often complicated by sensor limitations. Additionally, the nature and appearance of shapes can vary significantly between samples, even within the same plant family, increasing the complexity of the performance space and requiring more time for decisions and data analysis (Mufeng et al., 2023).

A key task in form recognition is accurately characterizing a shape based solely on its digital representation. This involves finding a description that distinguishes it from similar shapes and reduces misclassification risk. The description should capture features that allow comparison with neighboring shapes, ensuring only similar characteristics are grouped—a process commonly referred to as "learning" (Mufeng et al., 2023).

Another critical aspect is the decision-making process, which involves evaluating the input form and assigning it to the correct category by labeling it with a specific family or classification name. The system aims to find the best-matching family, typically by maximizing a similarity function between the input form's description and various family descriptions.

The complexity of form recognition tasks often requires processing vast amounts of data at high speed, challenging conventional methods. Traditional approaches struggle with the required processing power and intricate recognition algorithms. A promising research direction involves using artificial systems that autonomously adjust parameters, enabling generalization and adaptation to various input conditions—embodied in Deep Learning (Mufeng et al., 2023).

Deep Learning offers effective solutions for pattern recognition problems. Its learning capabilities reduce the need for human effort in research and comparison, while its adaptability enables recognizing patterns not explicitly trained on, making it powerful for complex problems like plant leaf classification.

Although Deep Learning for pattern recognition isn't new, its application to plant leaf classification is relatively novel. Previous research faced challenges leading to abandonment of the approach, but renewed interest in neural networks, especially with advancements in learning and generalization, has made them suitable for overcoming these challenges. This work ensures a clear presentation of ideas and concepts, contributing to improved plant leaf classification using Deep Learning.

Related work

Most existing plant identification work doesn't focus on automation, particularly in determining plant type by analyzing leaf properties, largely due to absent conventional classification criteria. Traditional approaches often rely on manual methods where experts use physical attributes like leaf shape, size, and fruit characteristics—effective but time-consuming, subjective, and not easily scalable. Recent comprehensive reviews, such as that by Patil and Shirdhonkar (2022), have cataloged the rapid advancement of deep learning approaches in this domain, highlighting a clear shift towards automated, data-driven solutions.

For example, Konstantinos et al. (2018) developed a reference manual of approximately 170 plant species categorized by fruit edibility and leaf shape, valuable but relying on individual expertise and manual comparison. Similarly, Arif Wani et al. (2019) published a book listing 250 tree types categorized by leaf characteristics, limited by textual descriptions and manual identification requirements.

Our work aims to bypass these manual procedures by providing an efficient, automated tool for accurate, quick plant species identification using leaf samples through modern computational methods, specifically neural networks and pattern recognition (Jackulin et al., 2022). This aligns with the broader trend of applying improved deep learning approaches for complex plant recognition tasks, as demonstrated in works on disease localization by Alqahtani et al. (2023).

Materials and methods

Deep Learning Applications for Leaf Pattern Recognition

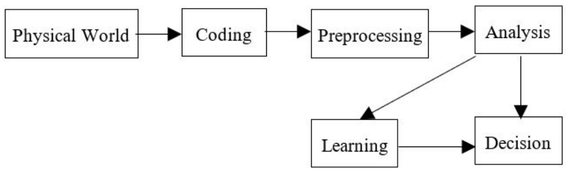
Deep learning has emerged as a powerful tool for pattern recognition, particularly in classifying and identifying plant species based on leaf characteristics using Convolutional Neural Networks (CNNs) to learn complex patterns from large leaf image datasets (Aanis et al., 2023). Its effectiveness is evident in various agricultural applications, including the automated identification of specific diseases like Northern Leaf Blight in maize from field imagery, as successfully shown by Chad et al. (2017). Recent architectural innovations, such as multi-scale feature fusion networks (Li et al., 2023), have further enhanced the ability of CNNs to capture discriminative features at various levels of abstraction, from fine textures to global shapes.

The classical approach to recognition problems involves several key steps (Guoqiang et al., 2016), shown in Figure 1:

1. **Preprocessing:** Cleaning and preparing raw data through noise reduction, image normalization, or enhancement.
2. **Feature Extraction:** Extracting relevant features that distinguish between classes/patterns.
3. **Classification:** Using algorithms to assign input to specific categories.
4. **Post-processing:** Refining results for improved accuracy.

Figure 1 illustrates this process from input data to final classification results.

Physical World → Coding → Preprocessing → Analysis → Learning → Decision



Source: Guoqiang et al, 2016

Figure 1: Overall diagram of the Pattern Recognition System.

We explain each step in our recognition process:

- **Physical World:** Modeled as infinite-dimensional "form space" with objects described through various properties.
- **Coding:** Conversion to digital representation (representation space).
- **Pre-treatment:** Selecting relevant information by eliminating noise, standardizing data, and removing redundancy.
- **Analysis:** Computing characteristics/parameters using recognition techniques.
- **Learning:** Refining decision-making by incorporating prior knowledge about shapes.
- **Decision:** Actual recognition using knowledge acquired during learning.

Both learning and decision phases involve neural networks—specifically a multilayer network with:

- Inputs/outputs corresponding to plant leaves and their species
- Error backpropagation learning algorithm

- Input number based on analysis phase parameters
- Output number corresponding to plant species classes (Guoqiang et al., 2016).

Characteristics of forms

Forms possess distinct features providing valuable information for pattern recognition, including perimeter, area, and other characteristics that describe shape and structure. These can be derived from both binary images (simplified black/white representations) and original images (detailed color information).

Characteristics are grouped into two categories:

1. **Boundary Features:** Describe outer contour/ perimeter
    - Perimeter: Boundary outline length
    - Compactness: Area/perimeter ratio
    - Convexity: Boundary convex/concave degree
    - Smoothness: Boundary regularity
2. **Interior Features:** Describe form interior
    - Area: Total covered space
    - Centroid: Center of mass
    - Shape descriptors: Elongation, aspect ratio, roundness (Ahmed et al., 2012).

Parameter extraction from plant leaves

Parameter extraction focuses on characteristics describing both shape and color variations, categorized as:

1. **Shape parameters:** Geometric features (perimeter, area, etc.)
2. **Color Parameters:** Color properties and variations

Single parameters may be insufficient for distinguishing between species.

Shape parameters

Essential for analyzing leaf shapes and environmental adaptations:

- **Compactness (C)**

$$C = \frac{Su}{Pe} \quad (1)$$

Where

Su = area (pixels within boundary), Pe = perimeter (pixels along boundary) demonstrates perimeter-area relationship (Figures 2-3.

Source: Authors

Figure 2: Efficiency of compactness (1) - Larger perimeter enclosing larger area.

Source: Authors

Figure 3: Efficiency of Compactness (2) - Same perimeter enclosing smaller area.

- **Elongation (Elong)**

$$Elong = \frac{Leng}{Width} \quad (2)$$

Where Leng = longest dimension (Ext1-Ext2), Width = broadest dimension (Ext3-Ext4) distinguishes broad-leaved vs. narrow-leaved types (Figures 4-5)

Source: Authors

Figure 4: Elongation of the leaf.

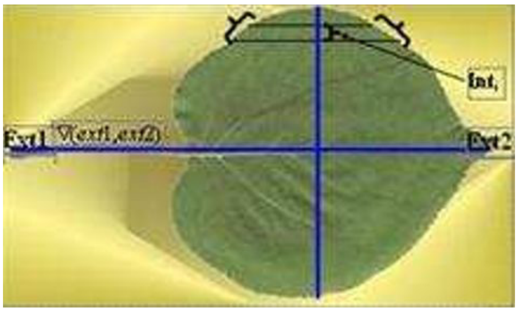
Source: Authors

Figure 5: Elongation measurement for leaves with similar shapes.

- **Level parabolic (LP)**

$$LP = \frac{Su}{Srect} \quad (3)$$

Where Srect = rectangle area around leaf using extreme points quantifies leaf curvature (Figure 6).



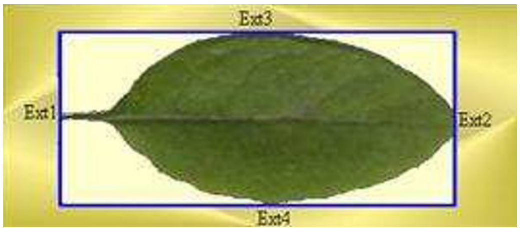
Source: Authors

Figure 6: Level parabolic shape.

- **Relativity (R)**

$$R = \frac{D1}{D2} \quad (4)$$

Where D1 = distance between Ext3 projection and Ext1, D2 = total leaf length distinguishes area distribution near supports (Figures 7-8).



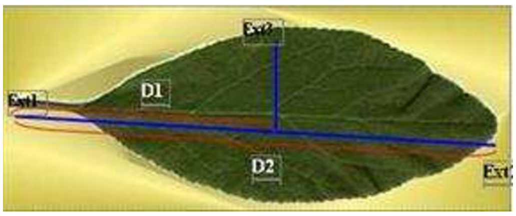
Source: Authors

Figure 7: Relativity.

Source: Authors

Figure 8: Efficiency of LP.

- **Concentration of Intervals of Distance (CID):** Ten sub-parameters representing contour points within distance intervals from leaf end Correlates with LP shape (Figure 9).



Source: Authors

Figure 9: Concentration of distance intervals.

Color parameters

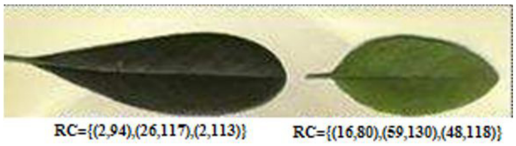
Provide essential insights into plant morphology, health, and adaptation:

- **Gray Level Medium (GLM):** Three sub-parameters (R-CM, CM-G, CM-B) representing average color values quantifies color variations (Figure 10)

Source: Authors

Figure 10: Efficiency of GLM.

- **Range of Color (RC):** Indicates color variation range within leaf, assesses color change across surface (Figure 11).



Source: Authors

Figure 10: Efficiency of GLM.

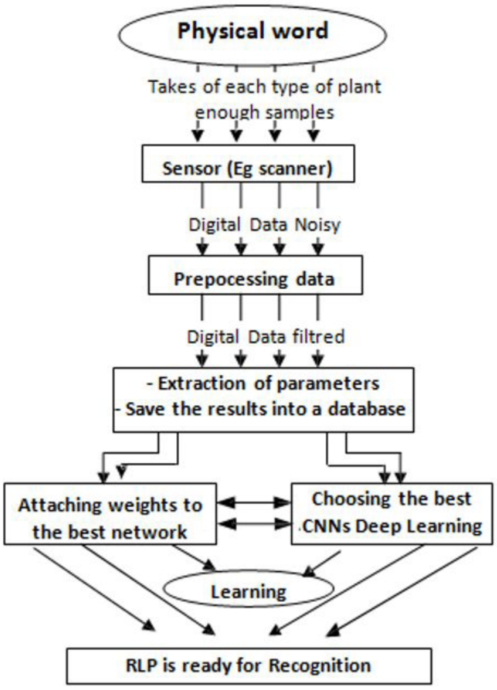
Results and discussion

PLR Tool (Plant Leaf Recognition Tool)

The PLR Tool simplifies plant species identification for researchers working in diverse environments with unfamiliar species. Unlike traditional time-consuming methods, it uses image recognition technology—users input leaf images, and the tool analyzes key features (shape, color, texture) against an extensive database for automatic identification. This automation saves time and increases accuracy, benefiting plant biologists, ecologists, and other researchers, particularly during fieldwork where quick, reliable identification is crucial. The tool also helps build and update plant databases.

General architecture of PLR

The PLR architecture combines data acquisition, image preprocessing, deep learning for feature extraction, classification, and output generation for accurate, efficient plant species recognition (Figure 12).



Source: Authors

Figure 12: General Architecture of RLP.

Sample characteristics

For optimal results, samples must meet criteria:

- Representative of plant types (collected at different growth stages)
- Exclude pathological cases with atypical characteristics
- Include leaf support (stem/petiole) for accurate identification

CNN selection for deep learning

Optimal CNN selection involves:

1. Determining iteration number
2. Validating network usage
3. Training all parameter combinations
4. Selecting network with smallest error

Application example

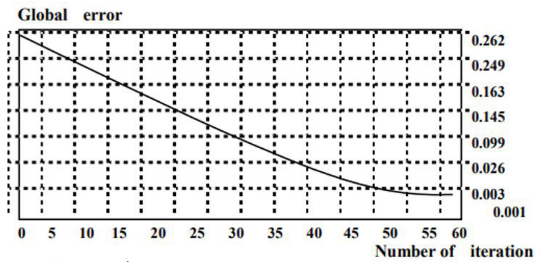
We tested the tool on 100 plant species, divided into:

- **Learning class:** 4-8 leaves/plant for CNN training
- **Test class:** 4 leaves/plant for evaluation CNN showed good performance, especially with consistent shapes. Semi-distributed coding improved convergence speed approximately threefold versus conventional coding.

Results

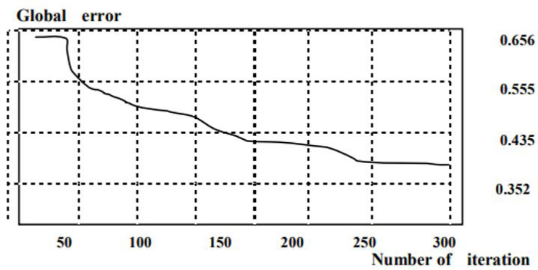
During training (100,000 iterations), CNN showed significant error reduction, especially with semi-distributed coding (error dropped from 0.7 to 10^{-8}). With semi-distributed coding, CNN correctly identified 83/100 species (85% success) versus 70% with conventional coding. This comparative result aligns with findings in the literature, where the choice of encoding and feature representation has been shown to significantly impact model performance in plant classification tasks (Sethy et al., 2024). The demonstration of improved accuracy, faster convergence, and better data handling underscores the efficacy of our approach.

Furthermore, the performance boost from semi-distributed coding suggests a pathway for future work. Integrating more sophisticated mechanisms, such as self-attention or lightweight architectures with dedicated attention modules (Zhang et al., 2022; Mbouembe and Ko, 2024), could potentially build upon this foundation to achieve even higher accuracy and efficiency.



Source: Authors

Figure 13: Graph of error (semi distributed coding).



Source: Authors

Figure 14: Graph of error (conventional coding).

The CNN achieved a strong performance using semi-distributed coding, accurately classifying 83 out of the 100 plant species, yielding a success rate of 85%. In contrast, the conventional coding approach achieved a notably lower accuracy of 70%. These results indicate that semi-distributed coding not only enhanced the model's generalization and classification capabilities but also facilitated faster convergence and more robust data processing compared to the conventional method.

Conclusion

Our objective was to create a comprehensive tool for botany researchers and scientists across disciplines. We identified mathematical parameters based on leaf shape measurements that enable plant classification into distinct categories, automated through image processing without user intervention. Using Deep Learning for recognition/classification with careful network selection, we achieved over 85% success rate with semi-distributed coding versus 70% with conventional coding. While satisfactory, this could potentially be improved by incorporating additional parameters like chemical and cellular characteristics.

Corresponding author:
Adel Abdelhadi
Department of Computer Science, Batna 2 University
53, Constantine Road, Fesdis, Batna 05078, Algeria
Phone: +213 559 364 312 / +213 657 267 599, Email: a.abdelhadi@univ-batna2.dz

References

[1] Aanis, A., Dharmendra, S. and Aly, E. (2023) "A survey on using deep learning techniques for plant disease diagnosis and recommendations for development of appropriate tools", S*mart Agricultural Technology*, Vol. 3., p. 100083. ISSN 2666-0159. DOI 10.1016/j.atech.2022.100083.

[2] Ahmed, A. A. and Gopireddy, H. R. (2021) "A Mobile-Based System for Detecting Plant Leaf Diseases Using Deep Learning", *AgriEngineering*, Vol. 3, No. 3, pp. 478-493. ISSN 2673-1906. DOI 10.3390/agriengineering3030032.

[3] Alqahtani, M., Nawaz, Y., Nazir, T., Javed, A., Jeribi, F. and Tahir, A. (2023) "An improved deep learning approach for localization and recognition of plant leaf diseases", *Expert Systems with Applications*, Vol. 230, p. 120717. ISSN 0957-4174. DOI 10.1016/j.eswa.2023.120717.

[4] Chad, D., Tyr, W., Siyuan, C., Ethan, L., Jason, Y., Michael, A., Rebecca, J. and Hod, L. (2017) "Automated Identification of Northern Leaf Blight-Infected Maize Plants from Field Imagery Using Deep Learning", *Phytopathology*, Vol. 107, No. 11. ISSN 0031-949X. DOI 10.1094/PHYTO-11-16-0417-R.

[5] Guoqiang, Z., Lina, W., Xiao, L. and Junyu, D. (2016) "An overview on data representation learning: From traditional feature learning to recent deep learning", *The Journal of Finance and Data Science*, Vol. 2, No. 4, pp. 265-278. ISSN 2405-9188. DOI 10.1016/j.jfds.2017.05.001.

[6] Jackulin, C. and Murugavalli, S. (2022) "A comprehensive review on detection of plant disease using machine learning and deep learning approaches", *Measurement: Sensors*, Vol. 24. p. 100441. ISSN 2665-9099. DOI 10.1016/j.measen.2022.100441.

[7] Konstantinos, P. F. (2018) "Deep learning models for plant disease detection and diagnosis", *Computers and Electronics in Agriculture*, Vol. 145, pp. 311-318. ISSN 0168-1699. DOI 10.1016/j.compag.2018.01.009.

[8] Lee, S. H., Chan, C. S., Mayo, S. J. and Remagnino, P. (2017) "How deep learning extracts and learns leaf features for plant classification", *Pattern Recognition*, Vol. 71, pp. 1-13. ISSN 0031-3203. https://doi.org/10.1016/j.patcog.2017.05.015.

[9] Li, W., Sun, Y., Li, J., Wang, H. and Zhang, C. (2023) "A multi-scale feature fusion convolutional neural network for high-performance plant leaf recognition", *Computers and Electronics in Agriculture*, Vol. 214, p. 108294. ISSN 0168-1699. DOI 10.1016/j.compag.2023.108294.

[10] Mufeng, T., Tommaso, S., Beren, M., Yuhang, S., Thomas, L. and Rafal, B. (2023) "Recurrent predictive coding models for associative memory employing covariance learning", *PLoS Computational Biology*, Vol. 19, No. 3, p. e1010719. ISSN 1553-7358. DOI 10.1371/journal.pcbi.1010719.

[11] Nazir, T., Iqbal, M. M., Jabbar, S., Hussain, A. and Albathan, M. (2023) "EfficientPNet—An Optimized and Efficient Deep Learning Approach for Classifying Disease of Potato Plant Leaves", *Agriculture*, Vol. 13, No. 4, p. 841. ISSN 2227-7072. DOI 10.3390/agriculture13040841.

[12] Wani, M. A., Ahmad Bhat, M. A., Saduf, F. and Asif Iqbal, K. A. (2019) "Basics of Supervised Deep Learning", In: *Advances in Deep Learning*, Chapter 2, pp. 13-29. ISBN-13 978-9811367939. DOI 10.1007/978-981-13-6794-6_2.