Democratic and Popular Republic of Algeria
Ministry of Higher Education and Scientific Research
University Yahia FARES, MEDEA
Faculty of Sciences
Dept. of Mathematics and Computer Science



Thesis

Submitted in Partial fulfillment of the Requirements for the

Master's Degree

Option : SEWT

By : Hamidi Ghofrane

Thesis Subject

Identification of Toxic and Medicinal Plants using Deep Learning

Presented and publicly defended before the jury ccomposed of:

M.	MAZARI Redha	University of MEDEA	President
M.	BABAALI Baligh	University of MEDEA	Thesis director
M.	BENAIASSA Azzedine	University of MEDEA	Examinator

May 2026

In the name of Allah, the Most Gracious, the Most Merciful,
I dedicate this humble work:
First, to the most precious people in my life, To my beloved mother, Halima Ammour, who has always been my support and strength in every step, And to my dear father, Abdelkrim Hamidi, who taught me the values of hard work and patience.
To my dear siblings: Mariem Djamila Hamidi, Akram, and Raouf Hussein, You are my source of strength and happiness. May God protect you for me.
To my coach, Abbas Haddad, Thank you for your constant support, motivation, and for all the advice that helped me grow.
To my amazing volleyball team, And to the players: Faiza Felfli, Bribiha Basma, Chaimaa Bousoubel, Aya Abdeli, Ben Moussa Bouchra, And all the rest of the team, thank you for all the beautiful moments and encouragement. You were like a second family to me.
Finally, I dedicate this work to everyone who contributed, directly or indirectly, to its success.
Thank you all

Acknowledgement

First and foremost, I thank God Almighty for granting me the ability, patience, and will to complete this work and for guiding me on the right path.

I also extend my deepest gratitude and appreciation to my supervisor, Baba Ali Baligh, for his guidance and support throughout this period. His expertise was invaluable, and his assistance in completing this research motivated me throughout this time and provided me with unwavering support.

I also thank all the university professors who were a great help to us and provided us with all the necessary support and assistance during our studies.

Finally, I express my heartfelt appreciation to my family and friends for their encouragement and patience throughout this period.

ملخص

يقترح هذا البحث تطوير نظام يعتمد على التعلم العميق باستخدام تقنية YOLOv8-seg لتصنيف أوراق النباتات إلى ثلاث فئات: النباتات الطبية، النباتات السامة، والنباتات غير السامة. تم جمع 4000صورة معنونة من مصادر مختلفة مثل منصة Kaggle، وقواعد البيانات النباتية، بالإضافة إلى صور حقيقية ملتقطة من البيئة الطبيعية. تم تدريب النموذج باستعمال وحدة معالجة الرسوميات Tesla T4 لمدة 100 حقبة تدريبية، وحقق نتائج متميزة تمثلت في 88.8% Box mAP50 و79.0% Mask mAP50 ودقة Precision بلغت 89.2% مع F1-Score إجمالي يقدر بـ 85.0%. كما بلغت دقة التصنيف حسب الفئات 99.8% للنباتات الطبية، و81.8% للنباتات السامة، و73.5% للنباتات غير السامة. وأظهرت اختبارات المتانة في ظروف مختلفة مثل الخلفيات المعقدة، وضعف الإضاءة، وتداخل النباتات، متوسط أداء قدره 75.7% mAP50. كما تم تطوير واجهة استخدام باستعمال Gradio تدعم اللغتين العربية والإنجليزية لتسهيل استعمال النظام وتحسين تجربة المستخدم.

الكلمات المفتاحية: التعلم العميق، YOLOv8-seg، تقسيم الصور، تصنيف النباتات، النباتات السامة، النباتات الطبية، النباتات غير السامة، الرؤية الحاسوبية، الحارس الأخضر، التعرف على الصور، التعلم بنقل.

Abstract

This research proposes the development of a deep learning system utilizing YOLOv8-seg technology that can classify plant leaves as medicinal, toxic, and non-toxic. 4,000 labeled images were gathered from Kaggle, botanical databases, and actual plant images taken from the field. The network was trained on the Tesla T4 GPU for 100 epochs, resulting in an 88.8% Box mAP50 score, 79.0% Mask mAP50 score, 89.2% precision score, and a macro F1-score of 85.0%. The per class accuracy obtained 99.8% for the medicinal plant category, 81.8% for toxic plants, and 73.5% for non-toxic plants. A robustness test conducted under various situations (complicated backgrounds, poor lighting, overlapping plants) achieved an average mAP50 score of 75.7%. This study designed an interface using Gradio that supports both Arabic and English.

Keywords: Deep Learning, YOLOv8-seg, Instance Segmentation, Plant Classification, Toxic Plants, Medicinal Plants, Non-toxic Plants, Computer Vision, Green Guard, Image Recognition, Transfer Learning,

Résumé

Cette recherche propose le développement d'un système d'apprentissage profond utilisant la technologie YOLOv8-seg capable de classer les feuilles de plantes en trois catégories : médicinale, toxique et non toxique. 4 000 images annotées ont été rassemblées à partir de Kaggle, de bases de données botaniques et d'images de plantes réelles prises sur le terrain. Le réseau a été entraîné sur un GPU Tesla T4 pendant 100 époques, obtenant un score Box mAP50 de 88,8%, un score Mask mAP50 de 79,0%, une précision de 89,2% et un F1-score macro de 85,0%. La précision par classe a atteint 99,8% pour la catégorie des plantes médicinales, 81,8% pour les plantes toxiques et 73,5% pour les plantes non toxiques. Un test de robustesse réalisé dans diverses conditions (arrière-plans complexes, faible éclairage, plantes superposées) a obtenu un score mAP50 moyen de 75,7%. Cette étude a conçu une interface utilisant Gradio qui supporte à la fois l'arabe et l'anglais.

Mots-clés : Apprentissage profond, YOLOv8-seg, Segmentation d'instances, Classification des plantes, Plantes toxiques, Plantes médicinales, Plantes non toxiques, Vision par ordinateur, Green Guard, Reconnaissance d'images, Apprentissage par transfert,

Contents

Abstract

الملخص

Résumé

List of Figures i

List of Tables ii

General Introduction 1

1 Botanical Context and Risks Associated with Plant Misidentification 4

1.1 Introduction . . . 4

1.2 Plant Classification: Medicinal, Toxic, and Neutral . . . 5

1.2.1 Medicinal Plants . . . 5

1.2.2 Toxic Plants . . . 5

1.2.3 Neutral Plants . . . 6

1.2.4 Look-alike Pairs . . . 6

1.3 The Problem of Plant Misidentification . . . 7

1.3.1 Public Health Significance . . . 7

1.3.2 Real-world Examples of Dangerous Look-alikes . . . 7

1.4 Review of Deep Learning Approaches for Plant Identification . . . 9

1.4.1 Convolutional Neural Networks (CNNs) . . . 9

1.4.2 Hybrid and Ensemble Models . . . 11

1.4.3 Attention-Based Models . . . 11

1.4.4 YOLO-Based Models for Plant Detection . . . 12

1.4.5 Emerging Architectures . . . 12

1.5 Limitations of Current Methods . . . 13

1.6 Conclusion . . . 14

2 Artificial Intelligence and Image Recognition for Plant Classification 15
2.1 Introduction . . . 15
2.2 Artificial Intelligence Definition . . . 16
2.3 History of Artificial Intelligence . . . 16
2.3.1 1950 - The Turing Test . . . 16
2.3.2 1956-1966 – Era of Initial Successes . . . 17
2.3.3 1970s – The AI Winter . . . 17
2.3.4 1980s-1990s – Expert Systems and Real Problems . . . 17
2.3.5 2000s-Present - Deep Learning Revolution . . . 18
2.3.6 The Future of Artificial Intelligence . . . 18
2.4 Advantages and Disadvantages of Artificial Intelligence . . . 19
2.4.1 Advantages of AI . . . 19
2.4.2 Disadvantages of AI . . . 19
2.5 Machine Learning and Deep Learning . . . 19
2.5.1 Machine Learning . . . 19
2.5.2 Deep Learning . . . 20
2.6 Image Recognition . . . 21
2.6.1 What is Image Recognition? . . . 21
2.6.2 Image Recognition Tasks . . . 21
2.7 Transfer Learning . . . 22
2.8 YOLOv8-seg for Plant Segmentation . . . 23
2.8.1 What is YOLO? . . . 23
2.8.2 The "seg" Version of YOLOv8 . . . 23
2.8.3 Advantages of YOLOv8-seg for This Project . . . 24
2.9 Conclusion . . . 24

3 Methodology for Plant Recognition and Segmentation using YOLOv8-seg 27
3.1 Introduction . . . 27
3.2 Dataset Description . . . 28
3.2.1 Data Collection . . . 28
3.2.2 Data Annotation . . . 28
3.2.3 Dataset Split . . . 29
3.3 Preprocessing and Data Augmentation . . . 29
3.3.1 Preprocessing . . . 29
3.3.2 Data Augmentation . . . 30
3.4 Platform Used for Implementation . . . 30
3.4.1 Hardware . . . 30

3.4.2 Software and Libraries 31
3.5 Proposed Methodology 32
3.5.1 Model Architecture 32
3.5.2 Training Configuration 33
3.5.3 Training Process 34
3.5.4 Loss Functions 34
3.6 Evaluation Metrics 34
3.6.1 Precision and Recall 34
3.6.2 F1-Score 35
3.6.3 Mean Average Precision (mAP) 35
3.7 Model Training Results 35
3.7.1 Training Progress 35
3.7.2 Validation Performance 36
3.7.3 Confusion Matrix 36
3.8 Comparison with State-of-the-Art 37
3.9 Conclusion 39

General Conclusion **41**

Bibliography **45**
Bibliography 45

List of Figures

1.1 Poison Hemlock (*Conium maculatum*) 8
1.2 Wild Carrot (*Daucus carota*) 8
1.3 Foxglove (*Digitalis purpurea*) 9
1.4 Comfrey (*Symphytum officinale*) 9
1.5 Deadly Nightshade (*Atropa belladonna*) 10
1.6 Black Nightshade (*Solanum nigrum*) 10

2.1 Types of machine learning 20
2.2 Image classification for plant safety identification 22
2.3 Bounding box around a detected leaf 22
2.4 Image segmentation showing the exact outline of a leaf 23

3.1 Medicinal plant 29
3.2 Toxic plant 29
3.3 Non-toxic plant 29
3.4 Sample images of medicinal, toxic, and non-toxic plant leaves 29
3.5 Architecture of the YOLOv8-seg model 33
3.6 Training loss curves 36
3.7 Validation performance curves after 100 epochs 37
3.8 Confusion matrix of the classification results 38

List of Tables

2.1 Advantages of AI in this project 19
2.2 Disadvantages of AI in this project 20
2.3 Advantages of YOLOv8-seg for this project 24

3.1 Dataset distribution 28
3.2 Dataset split 29
3.3 Data augmentation techniques 30
3.4 Loss values at different epochs 36
3.5 Overall validation performance 37
3.6 Comparison of our model with state-of-the-art approaches 38

General Introduction

Plant misidentification represents a significant public health concern worldwide, particularly in regions where medicinal and wild edible plants are commonly used. According to the World Health Organization (WHO), thousands of poisoning incidents are reported every year due to the accidental consumption of toxic plants that resemble edible or medicinal species [1]. The visual similarity between certain toxic and beneficial plants makes accurate identification difficult, especially for non-experts. Consequently, the development of reliable automatic plant recognition systems has become an important research topic in the fields of artificial intelligence and computer vision.

To address this challenge, the **Green Guard** project proposes an intelligent plant recognition and segmentation system based on deep learning techniques. The system aims to classify plant leaves into three categories: medicinal, toxic, and non-toxic, while simultaneously providing precise segmentation masks for the detected leaves.

Project Objective The main objective of this work is to develop an instance segmentation model based on **YOLOv8m-seg** capable of analyzing leaf images and accurately identifying their corresponding category. The proposed solution combines classification and segmentation to improve recognition performance and provide visual explanations of the model predictions.

Work Accomplished The work carried out to achieve this objective includes the following stages:

1. **Data Collection and Preparation**

   A dataset of **4,000 annotated images** was collected from multiple sources, including Roboflow, Kaggle, botanical databases, and field photographs. The dataset consists of:

   - 1,000 images of medicinal plants,
   - 1,000 images of toxic plants,
   - 2,000 images of non-toxic plants.

2. **Annotation Conversion**

   A significant challenge was encountered during the conversion of annotations from Run-Length Encoding (RLE) format to the YOLO segmentation format. To overcome this issue, a custom conversion script was developed using the `pycocotools` library, allowing the successful conversion of more than **30,456 segmentation masks** into polygon-based annotations.

3. **Model Training**

   The **YOLOv8m-seg** model was fine-tuned using transfer learning for **100 epochs** with a batch size of 16 and an image resolution of 640×640 pixels. Training was performed on a Tesla T4 GPU using Google Colab Pro and required approximately **7 hours**.

4. **Performance Evaluation**

   The proposed model was evaluated using standard performance metrics, including Precision, Recall, F1-score, mAP50, and mAP50–95. Furthermore, robustness tests were conducted under challenging conditions such as complex backgrounds, varying illumination levels, overlapping leaves, and partial occlusions.

5. **User Interface Development**

   A bilingual web application was developed using **Gradio**. The interface supports both Arabic and English languages and enables users to upload plant images, obtain classification results, and visualize the generated segmentation masks.

Results Obtained The experimental results demonstrate the effectiveness of the proposed approach:

- **Box mAP50:** 88.8%
- **Mask mAP50:** 79.0%
- **Overall Precision:** 89.2%
- **Macro F1-score:** 85.0%

The **medicinal** class achieved an outstanding F1-score of **99.8%**, while the **toxic** class obtained an F1-score of **81.8%**. In terms of robustness, the model achieved approximately **92%** accuracy on simple backgrounds and **78%** accuracy on complex backgrounds, demonstrating good generalization capabilities.

Thesis Structure This thesis is organized into four chapters.

- **Chapter 1** presents the botanical context and discusses the risks associated with plant misidentification.
- **Chapter 2** introduces the fundamental concepts of artificial intelligence, machine learning, deep learning, and computer vision.
- **Chapter 3** describes the proposed methodology for plant recognition and segmentation using YOLOv8m-seg, including dataset preparation, preprocessing, training, and evaluation procedures.
- **Chapter 4** presents and discusses the experimental results, performance analysis, and system evaluation.

Chapter 1

Botanical Context and Risks Associated with Plant Misidentification

1.1 Introduction

Plant identification is the process of determining the taxonomic identity of a plant specimen based on its observable characteristics. According to the Cambridge Dictionary, a plant is defined as "a living thing that grows in earth, in water, or on other plants, usually has a stem, leaves, roots, and flowers, and produces seeds" [2]. Identification, in the botanical context, refers to the act of recognizing and naming a particular plant species based on its distinguishing features [3].

The consequences of incorrect plant identification can be severe, particularly in communities where traditional herbal medicine is practiced or where wild plants are foraged for food. When an individual mistakenly consumes a toxic plant under the assumption that it is safe or beneficial, the result can range from mild gastrointestinal distress to severe poisoning and even death. Given that most people do not have immediate access to trained botanists or toxicologists, this creates a significant public health concern.

Recent advances in computer vision and deep learning have offered promising solutions to this problem. Rahim Azadnia, a computer vision researcher, explains that by employing deep neural networks and computer vision techniques, scientists can analyze plant leaf images to determine whether they are safe for consumption, possess medicinal properties, or are poisonous [4]. This technological approach has the potential to provide rapid, accessible plant safety assessments without requiring specialized botanical expertise.

The research area encompassing these approaches is most commonly referred to as "deep learning for plant classification" rather than "computer vision for herbal recognition." While both terms describe similar methodologies, "deep learning for plant classification" has become the more widely adopted terminology in the literature. Research

interest in this domain increased substantially around 2020 [5–7], with numerous studies emerging that address various aspects of automated plant identification.

1.2 Plant Classification: Medicinal, Toxic, and Neutral

The development of an automated safety classification system requires precise definitions of the three categories that plants may fall into: medicinal, toxic, and neutral. Xiong et al. (2021) demonstrated that supervised learning approaches depend critically on well-defined categories when working with plant phenotypic image recognition [7]. This section establishes clear, evidence-based definitions for each category.

1.2.1 Medicinal Plants

The World Health Organization (WHO) provides a comprehensive definition of medicinal plants, describing them as any plant which, in one or more of its organs, contains substances that can be used for therapeutic purposes or which serve as precursors for the synthesis of useful drugs [8]. The WHO elaborates that medicinal plants encompass herbs, herbal materials, herbal preparations, and finished herbal products that contain as active ingredients parts of plants or other plant materials, whether in crude form or as plant preparations [1]. These plant materials include juices, gums, fatty oils, essential oils, and other substances of a similar nature.

Several researchers have focused specifically on medicinal plant identification using deep learning approaches. Kavitha et al. (2024) define medicinal plant identification in real-time as the process of deploying a MobileNet CNN model on cloud platforms to recognize medicinal species from leaf images with high accuracy [9]. Trnvan et al. (2024) add that transfer learning with models like Xception can achieve 99.12% accuracy for medicinal plant species classification [10].

1.2.2 Toxic Plants

According to Nelson and Balick, poisonous plants are characterized by their ability to produce adverse effects upon exposure, with the mechanisms of action of the implicated toxin and clinical manifestations being well-documented [11]. The Encyclopedia of Toxicology defines a poisonous plant as "a plant that when touched or ingested in sufficient quantity can be harmful or fatal to an organism or any plant capable of evoking a toxic and/or fatal reaction" [12].

Nazuli et al. (2025) define poisonous plant classification as the use of convolutional neural networks such as MobileNetV2 to automatically identify harmful plant species

from images, achieving 99.34% accuracy with high precision and recall [13]. Maklad et al. (2025) further elaborate that explainable deep learning models can be used to predict poisonous plants while also providing reasoning for the classification decisions, enhancing trust and interpretability in automated safety assessments [14].

1.2.3 Neutral Plants

For the purposes of this thesis, neutral plants are defined as those that possess neither significant medicinal properties recognized by the WHO nor toxic properties sufficient to cause adverse health effects in humans. These plants may be edible, ornamental, or ecologically important, but they do not fall into the medicinal or toxic categories that are the primary focus of this research. The classification of a plant as "neutral" is context-dependent, as many plants that are not traditionally used as medicines may still contain bioactive compounds, and some plants considered neutral may cause allergic reactions in sensitive individuals.

1.2.4 Look-alike Pairs

The term "look-alikes" in the botanical context refers to plant species that are difficult to distinguish from one another due to visual similarities. According to the USDA Forest Service, invasive plant identification guides include a "Resembles" section specifically designed to help users identify plant "look-alikes" by comparing key characteristics that distinguish similar species [15].

The Delaware Department of Agriculture, in collaboration with the Claude E. Phillips Herbarium and the Delaware Center for Horticulture, published "Mistaken Identity? Invasive Plants and Their Native Look-alikes" as an identification guide for the Mid-Atlantic region. This guide specifically aims to facilitate correct identification of confusingly similar invasive and native plant species [16].

Azadnia et al. [4] noted that most classification errors in plant identification occur due to similarity in color and shape between different species. This finding directly motivates the need for systems specifically designed to handle look-alike discrimination, as these visually similar pairs represent the highest-risk scenarios for misidentification.

1.3 The Problem of Plant Misidentification

1.3.1 Public Health Significance

The misidentification of plants presents a significant public health concern. According to the World Health Organization, millions of people are poisoned by plants each year, with children, elderly individuals, and rural communities far from poison control centers facing the greatest risk. The absence of rapid, accessible identification methods exacerbates this problem, as delays in identifying the ingested plant can impede appropriate medical treatment.

Fouziya et al. (2025) found that people cannot easily identify medicinal plants, which makes traditional medicine less safe. Their model using ResNet50 and MobileNet reached 91% accuracy, but the datasets did not include enough different plant species—a limitation they explicitly flagged as requiring attention in future research [17].

1.3.2 Real-world Examples of Dangerous Look-alikes

The following examples illustrate the practical dangers of plant misidentification, demonstrating why automated classification systems must specifically address look-alike discrimination.

1.3.2.1 Poison Hemlock vs. Wild Carrot

Poison hemlock (*Conium maculatum*) is a highly toxic plant that can cause respiratory failure and death upon ingestion. The plant contains coniine and other piperidine alkaloids that act as neuromuscular blockers. Wild carrot (*Daucus carota*), by contrast, is safe to consume, particularly its young taproots, which are edible and nutritious.

The danger arises because the young leaves of poison hemlock closely resemble those of wild carrot, as depicted in Figures 1.1 and 1.2. This visual similarity has led to fatal poisonings when individuals have mistakenly foraged and consumed poison hemlock thinking it was wild carrot. The similarity is particularly pronounced during the early growth stages, when both plants produce finely divided, feathery foliage that is difficult for non-experts to distinguish.

1.3.2.2 Purple Foxglove vs. Comfrey

Foxglove (*Digitalis purpurea*) contains cardiac glycosides, including digoxin, which are potent pharmaceutical compounds used in treating heart conditions. The therapeutic window for these compounds is exceptionally narrow; the difference between a safe dose and a toxic dose is minimal, making accidental ingestion particularly dangerous.



Figure 1.1: Poison Hemlock (*Conium maculatum*)



Figure 1.2: Wild Carrot (*Daucus carota*)

Comfrey (*Symphytum officinale*) has been used traditionally for various medicinal purposes, particularly in topical applications. However, the leaves of foxglove and comfrey bear a striking visual resemblance, as shown in Figures 1.3 and 1.4, leading to frequent misidentification. This confusion is compounded by the fact that both plants grow in similar habitats and have overlapping geographical distributions.

Ok et al. (2025) found that computer vision systems encounter significant difficulty when distinguishing between visually similar plant species in outdoor environments. The challenges arise from varying lighting conditions, background clutter, and different leaf orientations, all of which complicate automated identification [18].

1.3.2.3 Deadly Nightshade vs. Black Nightshade

Deadly nightshade (*Atropa belladonna*) is one of the most toxic plants in the world. It contains tropane alkaloids such as atropine, scopolamine, and hyoscyamine. Ingestion of just a few berries can kill a child, and adults can die from consuming ten to twenty berries. The toxins affect the nervous system, causing hallucinations, seizures, respiratory failure, and death (see Figure 1.5).

Black nightshade (*Solanum nigrum*) has a complicated safety profile. The ripe black berries are edible and are used in pies, jams, and traditional dishes in some cultures. However, the unripe green berries contain solanine and are toxic. The leaves can also



Figure 1.3: Foxglove (*Digitalis purpurea*)



Figure 1.4: Comfrey (*Symphytum officinale*)

be toxic if consumed in large quantities, though they are eaten as cooked greens in some parts of the world after proper preparation (see Figure 1.6).

The problem with these two plants is that they look very similar, especially when the berries are not yet ripe. Both plants produce clusters of small round berries that start green and turn black or dark purple when ripe. Fatalities have occurred when individuals consumed unripe berries or leaves of black nightshade, mistaking them for the edible ripe berries. Others have mistakenly consumed deadly nightshade, thinking it was black nightshade, with fatal consequences.

1.4 Review of Deep Learning Approaches for Plant Identification

The field of deep learning for plant identification has experienced substantial growth since approximately 2020. This section examines existing approaches, organized by architectural type, and critically evaluates their strengths and limitations.

1.4.1 Convolutional Neural Networks (CNNs)

Convolutional Neural Networks (CNNs) remain the most widely adopted architecture for plant image recognition tasks. Their ability to automatically learn hierarchical features from raw pixel data makes them particularly suitable for visual recognition problems.



Figure 1.5: Deadly Nightshade (*Atropa belladonna*)



Figure 1.6: Black Nightshade (*Solanum nigrum*)

Xiong et al. (2021) conducted an extensive review of plant phenotypic image recognition technology and reported that CNN-based methods achieve up to 99.56% accuracy, establishing them as the most effective approach for this domain. Their review covered multiple CNN architectures including AlexNet, VGGNet, ResNet, DenseNet, and Inception, analyzing their strengths and weaknesses for plant classification tasks [7].

A systematic comparison of six different CNN architectures (VGG16, VGG19, DenseNet201, ResNet50V2, Xception, and Inception) was conducted by Dey et al. (2024) for automated medicinal plant identification from leaf images. Their extensive experimentation revealed that these models could achieve 99.64% accuracy under controlled conditions. However, they explicitly acknowledged a significant limitation: their dataset utilized simple white backgrounds that do not accurately represent real-world conditions where plant leaves are encountered with complex backgrounds, varying illumination, and occlusions [19].

Kavitha et al. (2024) developed a practical deployment of a MobileNet CNN model on the Google Cloud Platform for real-time medicinal plant identification. Their system achieved 98.33% accuracy with a precision of 99.32% and recall of 97.36%, demonstrating that lightweight CNN architectures can be effectively deployed on cloud platforms for real-time applications [9].

Nazuli et al. (2025) employed MobileNetV2 specifically for poisonous plant classification, achieving 99.34% accuracy with precision, recall, and F1-score all at 0.96 or higher. Their work focused exclusively on toxic species, leaving a gap in combined toxic-medicinal classification [13].

A comparison of ResNet50 and MobileNet architectures for herbal plant identification was performed by Fouziya et al. (2025), finding that ResNet50 achieved 91% accuracy.

Significantly, they explicitly called for more diverse plant species in future datasets, recognizing that limited species diversity constrains model generalization [17].

1.4.2 Hybrid and Ensemble Models

Researchers have explored combining multiple architectures to leverage the strengths of different approaches and improve overall performance.

Ok et al. (2025) conducted a comparative investigation of two distinct approaches: hybrid models combining CNN and LSTM (Long Short-Term Memory) networks, and ensemble models combining VGG16, ResNet50, and InceptionV3. Their results showed that the hybrid model achieved 99.4% accuracy with an F1-score of 99.38%, while the ensemble model achieved 96.5% accuracy with an F1-score of 99.74%. Importantly, they concluded that ensemble models exhibit greater stability and robustness to data variations, while hybrid models are more sensitive to changes in the input data distribution [18].

Pushpa et al. (2025) developed HybNet, a hybrid deep model combining VGG16, MobileNet, and KNN with SE (Squeeze-and-Excitation) blocks, achieving up to 94.24% accuracy. Their work is particularly relevant to this project because their dataset included real-time challenging conditions (variable lighting, complex backgrounds, different orientations), which closely resemble the conditions under which our proposed system must operate [20].

1.4.3 Attention-Based Models

Attention mechanisms have emerged as a powerful technique for improving deep learning models by enabling them to focus on the most discriminative parts of input images while ignoring irrelevant regions.

Azadnia et al. (2024) developed SCAM-Herb (Spatial and Channel Attention Mechanism for Herb classification), which achieved an accuracy of 99.63% with precision of 99.38%, recall of 99.52%, and F1-score of 99.42%. This study is exceptionally relevant to this project for three reasons: (1) it classifies both medicinal and poisonous plants together (rather than focusing exclusively on one category), (2) it utilizes leaf visual characteristics as the primary input features, and (3) the authors explicitly acknowledge that classification errors predominantly occur due to similarity in color and shape between different species. This last observation directly validates the research gap addressed by this project [4].

1.4.4 YOLO-Based Models for Plant Detection

The YOLO (You Only Look Once) family of models has been successfully applied to plant detection and segmentation tasks, offering real-time performance and state-of-the-art accuracy.

Ma et al. (2024) improved the YOLOv8 model by incorporating KernelWarehouse and Triplet Attention mechanisms, creating YOLOv8s-KDT specifically for classifying desert herbaceous plants. This model achieved 86.6% mAP with precision of 82.8% and recall of 81.6%. Their work demonstrates that YOLOv8 can effectively handle challenging real-world environments, validating its suitability for this project [21].

Bousahba et al. (2024) used YOLOv8n-seg for crop weed detection and instance segmentation, demonstrating that YOLOv8's segmentation capabilities are highly suitable for plant-level analysis. This finding directly supports our choice of using YOLOv8-seg for this project, as it provides empirical evidence that YOLOv8 can produce accurate segmentation masks for plant leaves even in complex agricultural environments [22].

Tsang et al. (2025) introduced YOLOv13 for real-time object detection, utilizing hypergraph-enhanced visual perception to achieve 41.6-54.8% AP. This model outperformed previous YOLO versions including YOLOv12, YOLOv11, YOLOv10, YOLOv9, and YOLOv8, demonstrating the ongoing evolution of the YOLO family [23].

1.4.5 Emerging Architectures

Recent research has introduced novel architectures that challenge the dominance of traditional CNN and YOLO approaches.

Pirard et al. (2025) introduced RF-DETR (Real-time Focal-DETR), a transformer-based model with a DINOv2 backbone capable of real-time object detection. This model achieved 60.5% AP on the COCO benchmark, outperforming YOLOv11 (56.8% AP) and RT-DETR-R50 (53.1% AP). RF-DETR represents the current state of the art for real-time object detection, though its complexity and computational requirements may limit its applicability for resource-constrained applications [24].

Several other innovative approaches have been proposed. Uma et al. (2025) developed MediFlora-Net, a quantum-enhanced deep learning model combining RGB and hyperspectral images, achieving 99.89% accuracy. However, this approach requires specialized hardware (hyperspectral imaging sensors) that is not accessible for typical research projects or practical field deployment [25]. Lapkovskis et al. (2024) proposed automatic multimodal fusion for plant classification, achieving 82.61% accuracy [26]. Paneru et al. (2024) developed a RAG (Retrieval-Augmented Generation) chatbot for medicinal plant insights, achieving 96.75% accuracy, demonstrating the potential of combining vision and

language models for plant identification [27].

1.5 Limitations of Current Methods

Based on the comprehensive literature review presented above, several critical limitations persist across existing approaches. These limitations collectively define the research gap that this project aims to address.

First, the majority of existing studies focus exclusively on either medicinal plants or toxic plants, but not both simultaneously. For example, Kavitha et al. (2024), Trnvan et al. (2024), and Dey et al. (2024) concentrated their efforts solely on medicinal plant identification [9, 10, 19], while Nazuli et al. (2025) and Maklad et al. (2025) focused exclusively on toxic plant classification [13, 14]. This fragmentation means that existing models cannot provide comprehensive safety classification across all three categories (toxic, medicinal, and neutral) that are necessary for a practical decision-support tool.

Second, instance segmentation for leaf-level analysis has only been applied to weed detection, not to medicinal or toxic plant classification. Bousahba et al. (2024) demonstrated the effectiveness of YOLOv8n-seg for weed segmentation, but no published study has applied instance segmentation to the specific problem of toxic-medicinal look-alike discrimination [22]. This represents a significant gap, as precise leaf segmentation could provide crucial morphological features for distinguishing visually similar species.

Third, the specific problem of "look-alike" pairs remains largely unaddressed in the literature. While Azadnia et al. (2024) noted that classification errors frequently occur due to visual similarity between species, they did not specifically design their model architecture or training methodology to handle this challenging scenario [4]. The look-alike problem is arguably the most critical safety concern in plant identification, yet it has received minimal attention in deep learning research.

Fourth, many datasets used in existing studies employ unrealistic backgrounds that do not reflect real-world conditions. Dey et al. (2024) explicitly acknowledged using simple white backgrounds that are dramatically different from the complex natural environments where plants are typically encountered and photographed by users [19]. This limitation significantly undermines the generalizability and practical utility of models trained on such datasets.

Fifth, species diversity is severely limited in most existing datasets. Fouziya et al. (2025) explicitly called for more diverse plant species in future datasets, recognizing that limited species diversity constrains model generalization and real-world applicability [17].

1.6 Conclusion

This chapter has provided a comprehensive foundation for the thesis by defining key concepts in plant identification, establishing the botanical context for toxic and medicinal plant classification, and analyzing the critical problem of look-alike pairs.

The review of existing deep learning approaches revealed that while considerable progress has been made, most studies focus on either medicinal or toxic plants in isolation. Few studies employ instance segmentation, none directly tackle the look-alike problem, and many rely on unrealistic background conditions that do not reflect real-world deployment scenarios.

These identified gaps directly motivate the research objectives of this thesis: to develop a deep learning system that combines (1) classification of both toxic and medicinal plants, (2) instance segmentation for precise leaf-level analysis, (3) specific focus on challenging look-alike pairs, and (4) robustness to real-world conditions.

The following chapter (Chapter 2) will cover the deep learning and computer vision foundations needed to address these gaps, including convolutional neural networks, instance segmentation, transfer learning, and the YOLOv8 architecture.

Chapter 2

Artificial Intelligence and Image Recognition for Plant Classification

2.1 Introduction

Before delving into complex scientific definitions, it is useful to establish a basic understanding of what "artificial intelligence" actually means. The dictionary provides a starting point: "artificial" refers to something made by humans rather than occurring naturally, while "intelligence" pertains to learning, understanding, and thinking logically. Consequently, artificial intelligence can be understood as the field concerned with machines that can think and learn in ways that resemble human cognition.

Following the advent of early computers, researchers began attempting to create systems capable of making independent decisions. Prior to this development, humans had exclusively handled such tasks. By delegating decision-making to intelligent systems, we can potentially accelerate processes and increase consistency. Furthermore, this approach may free human experts to focus on more creative and imaginative endeavors.

For this project, artificial intelligence is employed to differentiate between toxic and medicinal plants based on leaf images. The objective is to develop an autonomous decision-making system that can determine whether a plant is toxic, medicinal, or neutral, without requiring human intervention. This is particularly significant because most individuals lack professional botanical knowledge and struggle to distinguish between visually similar species. Consequently, an automated identification guide could help prevent numerous accidental poisonings.

2.2 Artificial Intelligence Definition

The term "artificial" can be interpreted as a replication of something not naturally occurring, which has been modeled, created, or manufactured according to a natural example through the application of technical methods [28].

The concept of intelligence originates from the Latin "intelligentia" and refers to insight, understanding, knowledge, and cognitive abilities. It also denotes the capacity to act abstractly and logically, deriving appropriate actions [29]. Intelligence can be understood as a combination of the abilities to learn, abstract, adapt, and think logically [30].

Artificial intelligence is defined as "the capacity of machines to carry out cognitive functions usually performed by human beings: perception, reasoning, learning, and problem-solving" [31]. More generally, AI may be referred to as the capability of a machine to independently understand, resolve, and learn from complicated tasks [32].

In the context of this project, artificial intelligence entails training a computer to identify the type of plant depicted in a leaf photograph, classifying it as toxic, medicinal, or neutral. This is a function that is typically performed by an experienced botanist.

2.3 History of Artificial Intelligence

The study of artificial intelligence is a relatively recent scientific discipline, having emerged in the second half of the twentieth century. The possibility of creating machines capable of mimicking human intelligence and solving problems in the same manner has become one of the earliest concepts associated with AI. Throughout its existence, AI has evolved from simple symbolic machines into highly sophisticated technologies that use data as the core means of functioning [33].

2.3.1 1950 - The Turing Test

An early milestone in the development of artificial intelligence was proposed by the British mathematician Alan Turing. In his groundbreaking work entitled *Computing Machinery and Intelligence*, Turing posed an important question concerning the abilities of modern computers: "Can machines think?"

As a way to measure the level of machine intelligence, the scientist developed the so-called *Turing Test*. According to Turing's concept, an intelligent machine should pass the following test: in a conversation with a human evaluator, both a person and a machine communicate through typing while being hidden from each other, with the evaluator determining which interlocutor is human and which is a computer [33].

The Turing Test remains one of the key contributions of Alan Turing to the field of artificial intelligence.

2.3.2 1956-1966 – Era of Initial Successes

This period was distinguished by enthusiasm for achievements and constant progress. During this time, scientists began developing various artificial intelligence software programs designed to solve logical and mathematical problems.

The *Logic Theorist*, created by Allen Newell and Herbert Simon, was aimed at automatically solving mathematical theorems. Another famous program of that era was *ELIZA*, created by Joseph Weizenbaum in the 1960s. This program imitated dialogues between psychotherapists and their clients by analyzing user entries and responding accordingly. Despite its simplicity, ELIZA demonstrated the capabilities of natural language processing systems [33].

Scientists at that time believed in the rapid growth of computer intelligence, which enabled them to secure increased funding for research.

2.3.3 1970s – The AI Winter

Despite some success stories, AI solutions suffered from several constraints in the 1970s due to the limited computational capabilities and insufficient memory of computers of that era.

The majority of projects in this area proved unsuccessful, and governmental funding for AI research was sharply reduced. This period became known as the "AI Winter."

Nevertheless, scientists continued exploring various approaches to overcome difficulties and develop solutions for the future.

2.3.4 1980s-1990s – Expert Systems and Real Problems

During the 1980s, researchers concentrated their efforts on creating practically applicable AI algorithms. One of the prominent directions of this period was the development of *Expert Systems*, which were designed to simulate human ability to analyze complex situations and make decisions.

These systems operated based on extensive sets of rules and knowledge bases, providing useful advice. They proved to be both profitable and efficient.

Machine learning techniques of the 1990s enabled computers to analyze information contained in databases and make decisions based on that analysis. AI began to address tasks such as speech and image recognition, decision-making, and other complex problems.

2.3.5 2000s-Present - Deep Learning Revolution

Beginning in the early twenty-first century, several factors—namely enhanced computing power, big data availability, and improved algorithms—led to another significant breakthrough in AI.

One of the key technologies behind the rise of modern AI systems is deep learning, a subset of machine learning inspired by the structure of the human brain. The emergence of deep neural networks allowed processing of complicated data such as images, voice, or text with remarkable precision.

Artificial intelligence has found its way into a variety of domains, including:

- Object recognition and computer vision
- Speech recognition and virtual assistants
- Autonomous vehicles
- Diagnostics and treatment of diseases
- Natural language processing
- Plant identification and agricultural monitoring

Currently, artificial intelligence enables recognition of objects, understanding of natural language, generation of texts and graphics, and assistance to scientists in their research.

2.3.6 The Future of Artificial Intelligence

At present, AI continues to develop dynamically, and the future contribution of this technology to various aspects of life and other domains cannot be underestimated. Modern AI researchers are working on constructing systems capable of solving even more difficult tasks.

Future AI advancements may lead to improvements in areas such as healthcare, education, environmental conservation, cybersecurity, and scientific research. Nonetheless, issues such as privacy, bias, unemployment, and ethical AI development remain significant challenges that must be addressed.

2.4 Advantages and Disadvantages of Artificial Intelligence

2.4.1 Advantages of AI

Artificial Intelligence offers numerous benefits for this project. Its advantageous characteristics enable the project to function more effectively and accomplish tasks more efficiently.

Table 2.1 presents the advantages of Artificial Intelligence and explains how each one contributes to improving plant identification and safety outcomes.

Advantage	**How it helps this project**
Reduction in human error	The system makes consistent safety decisions without mistakes from fatigue
24/7 availability	Anyone can use the system anytime to identify plants
Unbiased decisions	AI is not influenced by emotions or preconceptions about plants
Handles repetitive tasks	The system can classify thousands of leaf images quickly
Speed	AI can process an image in milliseconds
Accessibility	The system can be used by anyone without botanical training

Table 2.1: Advantages of AI in this project

2.4.2 Disadvantages of AI

While Artificial Intelligence offers numerous benefits, it also presents certain challenges. Table 2.2 outlines the disadvantages of Artificial Intelligence and describes how this project addresses each one.

2.5 Machine Learning and Deep Learning

2.5.1 Machine Learning

Machine learning (ML) constitutes one of the technical facets of artificial intelligence. It forms the foundation for developing new solutions to various problems [34]. Machine

Disadvantage	**How we address it**
High costs	We use free tools like Google Colab and open-source YOLOv8
Needs large amounts of data	We use data augmentation and transfer learning
No creativity	The system only does what it is trained to do
Black box problem	We use evaluation metrics to measure performance
Can make mistakes	We evaluate using precision, recall, and F1-score

Table 2.2: Disadvantages of AI in this project

learning is a subset of AI that involves the creation of knowledge from prior experiences. Through examples or data, an artificial system can be trained to learn patterns or repetitions [35].

In this project, machine learning enables the system to analyze hundreds of plant leaf images to identify patterns associated with toxicity and medicinal properties, as illustrated in Figure 2.1.

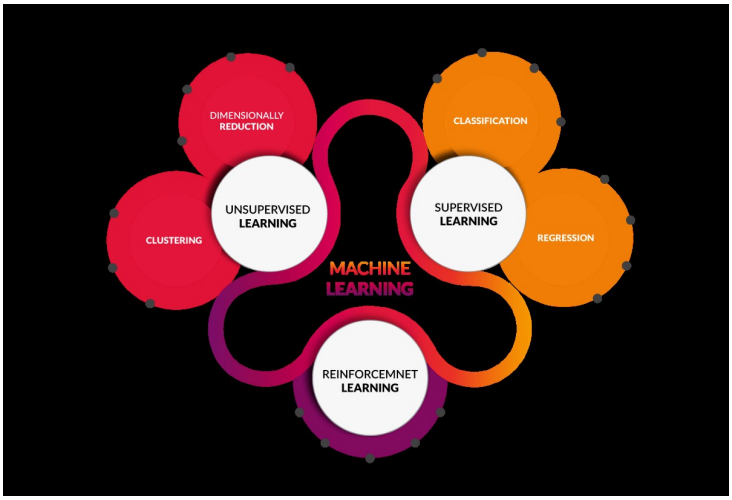


Figure 2.1: Types of machine learning

2.5.2 Deep Learning

Deep learning (DL) is a particular branch of machine learning that utilizes neural networks to analyze complex structures [36]. The objective of deep learning is to enable a computer to acquire capabilities that already exist in humans [33].

The concept behind deep learning involves neural networks, which provide machines with the capability to interpret more complicated structures than ordinary ML can [37].

Artificial neurons are simulated and stacked in numerous layers. Each layer within a neural network is dedicated to performing a particular function—for example, edge detection [38].

For this project, deep learning will be used to:

- Identify leaves within images
- Recognize particular attributes such as shape, texture, and color
- Categorize plants as poisonous, medicinal, or neutral

2.6 Image Recognition

2.6.1 What is Image Recognition?

Image recognition is a subcategory of computer vision and artificial intelligence. It represents a set of methods for detecting and analyzing images to enable the automation of a specific task [39]. It is a technology capable of identifying places, people, objects, and many other types of elements within an image.

For this project, image recognition is used to:

- Detect the leaf in the image
- Segment the leaf to obtain its exact shape
- Classify the plant as toxic, medicinal, or neutral

2.6.2 Image Recognition Tasks

2.6.2.1 Classification

Image classification is the identification of the class (category) to which an image belongs. An image can have only one class. For this project, classification means answering: "Is this plant toxic, medicinal, or neutral?" This is illustrated in Figure 2.2.

2.6.2.2 Object Detection

Object detection is necessary when the goal is to locate an object within an image. Once the object is located, a bounding box is placed around it [40]. For this project, detection means identifying the location of the leaf within the image, as shown in Figure 2.3.



Figure 2.2: Image classification for plant safety identification



Figure 2.3: Bounding box around a detected leaf

2.6.2.3 Segmentation

Image segmentation is the partitioning of an image into distinct regions that correspond to different objects. Each region contains pixels with similar attributes [36]. For this project, segmentation means drawing the exact outline of each leaf. This is particularly important because leaf shape is a key characteristic for identifying plant species, as illustrated in Figure 2.4.

2.7 Transfer Learning

In the dictionary, 'transfer' means to move something from one place to another. In deep learning, transfer learning refers to reusing a model that has already been trained on a large dataset for a new, related task [41].

Trnvan et al. (2024) explain that transfer learning achieves high accuracy even with a small dataset [10]. This is particularly useful for this project because the available dataset of plant images is relatively limited.

Why reuse a pre-trained model?

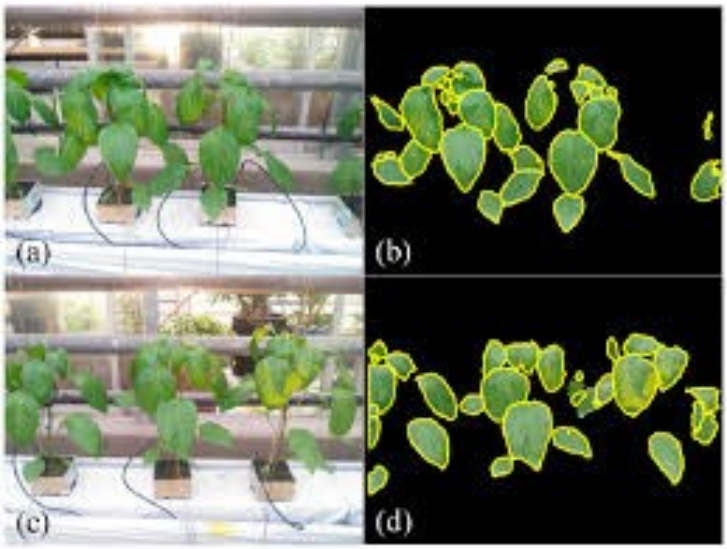


Figure 2.4: Image segmentation showing the exact outline of a leaf

- Saves time by eliminating the need to train from scratch
- Requires fewer training images
- Achieves better results than training a model from scratch with limited data

2.8 YOLOv8-seg for Plant Segmentation

2.8.1 What is YOLO?

YOLO stands for "You Only Look Once." It is a real-time object detection model. As explained by Ma et al. (2024), unlike older models that examine an image multiple times, YOLO processes the image once and predicts all objects simultaneously. This architectural choice makes it exceptionally fast [21].

2.8.2 The "seg" Version of YOLOv8

YOLOv8-seg is a specialized version of YOLOv8 that performs instance segmentation. Bousahba et al. (2024) demonstrated that YOLOv8-seg can segment plants with high precision [22]. The model produces both:

- A bounding box around each detected object
- A segmentation mask representing the exact contour of each object

Table 2.3 summarizes the key advantages of this approach.

2.8.3 Advantages of YOLOv8-seg for This Project

The following table summarizes the key advantages of using YOLOv8-seg for plant leaf segmentation and classification in this project. Each advantage is explained in terms of how it contributes to the overall system performance and practical utilityshown in table 2.3.

Advantage	**Explanation**
Speed	YOLO processes the image in a single pass, making it very fast
Precision	Instance segmentation provides exact leaf shapes
Multiple leaves	Can detect and segment several leaves in one image
Pre-trained	Can be fine-tuned on our plant dataset
Well documented	YOLOv8 has extensive community support

Table 2.3: Advantages of YOLOv8-seg for this project

2.9 Conclusion

This chapter has introduced the key concepts necessary for this project. A thorough understanding of these concepts is essential for appreciating the methodology and results presented in subsequent chapters.

First, **Artificial Intelligence** enables machines to make decisions in ways that resemble human cognition. AI is not magic or science fiction, but a practical engineering discipline that uses mathematical models to perform tasks that normally require human intelligence. For this project, AI enables a computer to analyze a leaf image and determine whether it originates from a toxic plant, a medicinal plant, or a neutral plant.

Second, **Machine Learning** allows systems to learn from examples rather than being explicitly programmed. Instead of writing thousands of rules about leaf shapes, colors, and textures, we present the model with many labeled examples and allow it to discover the patterns independently. This approach is more flexible and scalable than traditional programming, particularly for visual recognition tasks where the rules are difficult to articulate explicitly.

Third, **Deep Learning** employs neural networks with multiple layers to understand complex patterns. Shallow machine learning models often require hand-crafted features—someone must specify which aspects of the leaf are important. Deep learning models can learn these

features automatically. The multiple layers enable the model to build a hierarchy of understanding: early layers detect simple patterns such as edges and colors, middle layers combine these into shapes such as curves and corners, and later layers assemble these into complete leaf structures that can be recognized and classified.

Fourth, **Image Recognition** encompasses three related tasks: classification, detection, and segmentation. Classification answers the question "what is this?" by assigning a single label to the entire image. Object detection answers "what is where?" by drawing bounding boxes around objects and labeling them. Segmentation goes further by marking the exact outline of each object at the pixel level. This project employs instance segmentation, which combines detection and segmentation to locate every leaf in an image and draw its precise boundary. This level of detail is essential because leaf shape is a key characteristic for distinguishing look-alike species.

Fifth, **Transfer Learning** facilitates working with limited data. Training a deep learning model from scratch would require hundreds of thousands or even millions of labeled images, which are not available for this project. However, we can begin with a model that has already been trained on a large general-purpose dataset (COCO, which contains millions of images of everyday objects) and then fine-tune it on our smaller plant dataset. This approach transfers the general visual knowledge learned from COCO—such as edges, textures, and shapes—to the specific task of plant classification. Transfer learning dramatically reduces the amount of training data and time required while still achieving excellent results.

Sixth, **YOLOv8-seg** provides fast and precise leaf segmentation. YOLO stands for "You Only Look Once," meaning the model processes the entire image in a single pass rather than scanning it multiple times. This makes it extremely fast, capable of processing an image in milliseconds. The "seg" version adds instance segmentation capabilities, enabling the model to output not only bounding boxes but also precise pixel-level masks of each leaf. This combination of speed and precision makes YOLOv8-seg ideal for the real-time plant identification application required in this project.

Seventh, **Evaluation Metrics** help measure performance objectively. It is insufficient to simply assert that the model performs well; quantitative evidence is required. This project employs several standard metrics: precision measures the proportion of positive predictions that are actually correct; recall measures the proportion of actual positive cases that the model successfully identified; the F1-score combines precision and recall into a single balanced measure; and mean average precision (mAP) provides a comprehensive evaluation of detection and segmentation quality across all classes and confidence thresholds.

These seven concepts form the theoretical foundation of this work. Together, they

provide both the motivation for using deep learning (artificial intelligence and machine learning), the specific techniques employed (deep learning, YOLOv8-seg, and transfer learning), the tasks being solved (image recognition for classification, detection, and segmentation), and the means of measuring success (evaluation metrics).

The next chapter (Chapter 3) will present the methodology: the data collection process, image annotation approach, model training procedure, and desktop prototype development. While this chapter has focused on the "what" and "why" of deep learning for plant classification, Chapter 3 will address the "how"—the practical steps taken to progress from a research concept to a functional system. The dataset will be described in detail, preprocessing and augmentation pipelines will be explained, training hyperparameters will be specified, and the challenges encountered during implementation—including the RLE conversion issue and its resolution—will be documented.

Chapter 3

Methodology for Plant Recognition and Segmentation using YOLOv8-seg

3.1 Introduction

In this study, we wanted to create a system that can tell the difference between toxic, medicinal, and non-toxic plants using only photographs of their leaves. This is an important problem because many toxic plants look very similar to medicinal or edible plants, which leads to accidental poisonings every year.

We looked at three types of plants. The first type is toxic plants like Poison Hemlock and Foxglove. These plants contain dangerous compounds that can cause serious illness or death if ingested. The second type is medicinal plants like Comfrey and Chamomile. These plants have been used for centuries in traditional medicine. The third type is non-toxic plants like Wild Carrot and Dandelion. These plants are generally safe for consumption.

We tested a model called YOLOv8-seg to see how well it could recognize these three types of plants from leaf images. YOLO stands for "You Only Look Once," which means the model processes the entire image in a single pass, making it very fast. The "seg" version also performs instance segmentation, meaning it can draw precise outlines around each leaf.

The model looks at three main characteristics of leaves to classify plants. First, it examines the shape features including the overall outline, edge pattern, and vein arrangement. Second, it analyzes color features such as the base color and any spots or patterns. Third, it evaluates texture features including whether the leaf surface is smooth, hairy, or rough.

We also wanted to see how well the model works in challenging real-world situations. We tested the model under different conditions including simple and complex backgrounds,

low light, overlapping leaves, damaged leaves, different angles, and different leaf sizes. The goal was to ensure the model would be robust and reliable when used by ordinary users in real-world environments.

3.2 Dataset Description

3.2.1 Data Collection

This study aimed to classify three important plant categories using digital images. A total of 4,000 digital images of toxic plants, medicinal plants, and non-toxic plants were collected from multiple sources including Kaggle, botanical databases, and custom field photographs, as shown in Table 3.1.

Category	**Number of Images**	**Percentage**
Medicinal	1,000	25%
Toxic	1,000	25%
Non Toxic	2,000	50%
Total	**4,000**	**100%**

Table 3.1: Dataset distribution

3.2.2 Data Annotation

All images were annotated using the **Roboflow** platform. The annotation process involved:

- Drawing precise segmentation masks around each leaf
- Assigning class labels: `medicinal`, `toxic`, `non toxic`
- Exporting annotations in YOLO format

Figure 3.4 shows sample images of medicinal, toxic, and non-toxic plant leaves from the database.



Figure 3.1: Medicinal plant



Figure 3.2: Toxic plant



Figure 3.3: Non-toxic plant

Figure 3.4: Sample images of medicinal, toxic, and non-toxic plant leaves

3.2.3 Dataset Split

The dataset was split into training, validation, and test sets as shown in Table 3.2:

Split	Percentage	Number of Images
Training	70%	2,800
Validation	15%	600
Test	15%	600
Total	**100%**	**4,000**

Table 3.2: Dataset split

Each class was split with the same ratio:

- **Medicinal:** 700 training, 150 validation, 150 test
- **Toxic:** 700 training, 150 validation, 150 test
- **Non Toxic:** 1,400 training, 300 validation, 300 test

3.3 Preprocessing and Data Augmentation

3.3.1 Preprocessing

Although the collected images were of varying sizes and some had noise, we made sure to preprocess all of them for consistency:

- We resized all the images to 640 × 640 pixels so they would work with YOLOv8.
- We also made sure the pixel values were in a range from 0 to 1.

3.3.2 Data Augmentation

To improve model generalization and prevent overfitting, the following data augmentation techniques were applied using the Albumentations library, as detailed in Table 3.3:

Technique	**Description**
RandomFlip	Horizontal and vertical flipping
RandomRotation	Rotation between -15° and +15°
RandomZoom	Zoom in/out (scale 0.5-1.5)
Brightness Adjustment	Random brightness changes
Mosaic	Combine 4 images into one (early epochs)
Blur	Gaussian blur
CLAHE	Contrast Limited Adaptive Histogram Equalization
ToGray	Convert to grayscale

Table 3.3: Data augmentation techniques

3.4 Platform Used for Implementation

3.4.1 Hardware

The hardware used in our implementation is:

- **GPU:** Tesla T4 (Google Colab Pro)
- **GPU Memory:** 15 GB
- **RAM:** 12 GB
- **Processor:** Intel Core i5 / AMD A9-9425
- **Operating System:** Windows 10 / Linux

3.4.2 Software and Libraries

The following software and libraries were used in this project:

Python 3.12
Python is an object-oriented, open-source computer programming language with dynamic semantics. Its syntax is simple and readable, making it easy to learn and maintain programs. The Python interpreter and the extensive standard library are available in source without charge for all major platforms. [42]
Purpose: Main programming language for the entire project

PyTorch 2.10
PyTorch is an open-source machine learning framework developed primarily by Meta AI. It provides tensor computation with GPU acceleration and deep neural networks built on a tape-based automatic differentiation system. [43]
Purpose: Deep learning framework for training YOLOv8-seg

Ultralytics YOLOv8 8.4.35
Ultralytics YOLOv8 is a state-of-the-art deep learning model for object detection and instance segmentation. It provides a simple and powerful API for training, validating, and deploying YOLO models. [44]
Purpose: Instance segmentation and plant leaf classification

Roboflow
Roboflow is a computer vision platform that provides tools for annotating images, augmenting datasets, and exporting annotations in various formats. It simplifies the process of preparing custom datasets for deep learning models. [45]
Purpose: Data annotation and dataset preparation

OpenCV 4.x
OpenCV (Open Source Computer Vision Library) is an open-source library that includes several hundreds of computer vision algorithms. It is used for image processing, feature detection, and image manipulation. [46]
Purpose: Image preprocessing (resizing, normalization, augmentation)

Google Colab

Google Colab is a free cloud-based Jupyter notebook environment that provides access to GPU (Tesla T4, V100, A100) and TPU for machine learning. It allows users to write and execute Python code in the browser without any configuration. [47]

Purpose: GPU-accelerated training of the YOLOv8-seg model

Matplotlib 3.x

Matplotlib is a comprehensive library for creating static, animated, and interactive visualizations in Python. It was started by John Hunter in 2002 to enable a MATLAB-like plotting interface in Python. [48]

Purpose: Plotting training curves (loss, accuracy, mAP)

NumPy 1.x

NumPy is the fundamental package for scientific computing in Python. It provides support for large, multi-dimensional arrays and matrices, along with a collection of mathematical functions to operate on these arrays. [49]

Purpose: Numerical operations and array manipulations

Scikit-learn 1.2.2

Scikit-learn is a free software machine learning library for the Python programming language. It features various classification, regression and clustering algorithms. It is a very popular tool and the most prominent Python library for machine learning. [50]

Purpose: Calculating evaluation metrics (precision, recall, F1-score)

Pandas 2.x

Pandas is a fast, powerful, flexible and easy-to-use open-source data analysis and manipulation tool, built on top of the Python programming language. [51]

Purpose: Data manipulation and analysis of evaluation results

3.5 Proposed Methodology

3.5.1 Model Architecture

We used **YOLOv8m-seg** (medium version) for instance segmentation. YOLO stands for "You Only Look Once", which means the model looks at the image once and predicts all objects at the same time, making it very fast.

The main components of the architecture are:

- **Backbone:** CSPDarknet with 192 layers

- **Neck:** PAN-FPN for feature fusion
- **Head:** Segment head with mask prediction

The model has approximately **27.2 million parameters** and **104.7 GFLOPs**. It is designed to classify **3 classes**: medicinal, toxic, and non toxic, as illustrated in Figure 3.5.

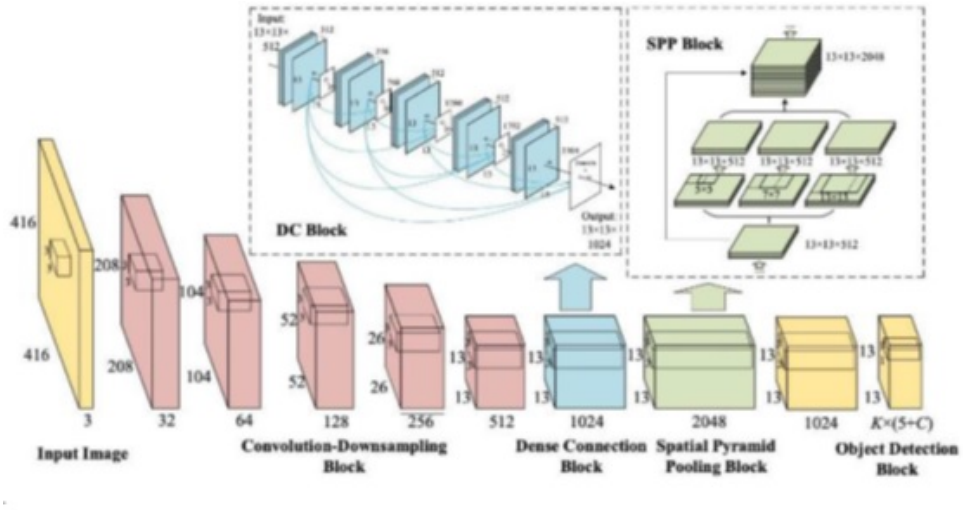


Figure 3.5: Architecture of the YOLOv8-seg model

3.5.2 Training Configuration

The training was performed using the following hyperparameters:

- **Epochs:** 100
- **Batch size:** 16
- **Image size:** 640×640 pixels
- **Optimizer:** AdamW
- **Initial learning rate:** 0.00125
- **Momentum:** 0.9
- **Weight decay:** 0.0005
- **Warmup epochs:** 3
- **Warmup momentum:** 0.8

3.5.3 Training Process

The training was conducted on Google Colab Pro with a Tesla T4 GPU. The complete training process for 100 epochs took approximately **7 hours**. This duration reflects the complexity of the instance segmentation task, the size of the dataset, and the computational resources available.

We utilized a pre-trained YOLOv8 model on the COCO dataset and fine-tuned it for our specific task. This transfer learning approach allowed the model to leverage features learned from millions of diverse images, resulting in improved performance even with our limited dataset. The combination of the YOLOv8 architecture and the COCO pre-trained weights significantly enhanced the training efficiency and final model accuracy.

3.5.4 Loss Functions

The YOLOv8-seg model uses multiple loss functions to optimize its performance:

- **Box loss:** Measures the error in bounding box predictions.
- **Class loss:** Measures the classification error.
- **DFL loss:** Distribution Focal Loss used to improve bounding box regression.
- **Segmentation loss:** Measures the error in mask predictions.

3.6 Evaluation Metrics

To evaluate our model, we used the following metrics:

3.6.1 Precision and Recall

Precision and Recall are fundamental evaluation metrics used to assess the performance of a classification model.

Precision measures the proportion of correctly predicted positive instances among all predicted positives, while Recall measures the proportion of correctly identified positive instances among all actual positives.

$$\text{Precision} = \frac{TP}{TP + FP} \tag{3.1}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{3.2}$$

Where:

- TP (True Positives): correctly identified positive cases
- FP (False Positives): incorrectly identified cases as positive
- FN (False Negatives): positive cases that were missed

3.6.2 F1-Score

The F1-score is the harmonic mean of Precision and Recall, providing a balanced measure between them, especially when dealing with imbalanced datasets.

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{3.3}$$

3.6.3 Mean Average Precision (mAP)

The main evaluation metric used for object detection and segmentation tasks is the Mean Average Precision (mAP). Unlike traditional accuracy, which is suitable for classification tasks, mAP is specifically designed to evaluate detection and segmentation performance. Accuracy simply measures the percentage of correct predictions, but it does not account for localization quality or the ability to detect objects at different confidence thresholds. For instance segmentation tasks like ours, mAP provides a more comprehensive and meaningful evaluation by measuring both detection and segmentation quality.

The mAP is computed by first calculating the Average Precision (AP) for each class at different Intersection over Union (IoU) thresholds, and then taking the mean of these values. The mAP50 specifically refers to the AP calculated at IoU threshold 0.5, while mAP50-95 averages the AP across IoU thresholds from 0.5 to 0.95.

The Intersection over Union (IoU) is a metric used to measure the overlap between the predicted masks and the ground truth masks.

$$\text{IoU} = \frac{\text{Area of Overlap}}{\text{Area of Union}} \tag{3.4}$$

In other words, IoU quantifies how well the predicted masks match the actual masks by comparing the overlapping area to the total combined area of both masks.

3.7 Model Training Results

3.7.1 Training Progress

The model was trained for 100 epochs. The loss values decreased significantly over time, as shown in Table 3.4 and Figure 3.6.

Epoch	Box Loss	Seg Loss	Cls Loss	DFL Loss
1	1.098	2.687	1.855	1.219
20	0.917	2.180	1.303	1.099
50	0.805	1.859	1.062	1.039
70	0.765	1.742	0.943	1.014
100	0.571	1.041	0.570	0.911

Table 3.4: Loss values at different epochs

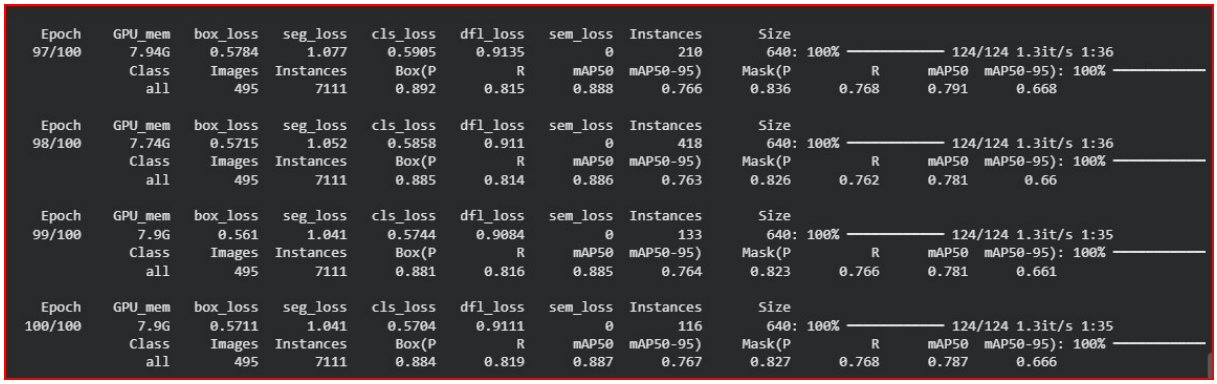


Figure 3.6: Training loss curves

3.7.2 Validation Performance

After 100 epochs, the model achieved the following results on the validation set, as presented in Table 3.5 and Figure 3.7.

3.7.3 Confusion Matrix

To better understand the model's classification errors, a confusion matrix was generated. Figure 3.8 shows the confusion matrix for the three classes: medicinal, toxic, and non-toxic.

The confusion matrix reveals that:

- The model perfectly classified all **medicinal** plants (12 medicinal, 0 errors).
- **Toxic** plants were mostly correctly classified (610 out of 621).
- The main confusion occurred between **toxic** and **non-toxic** classes (35 non-toxic predicted as toxic, 201 toxic predicted as non-toxic), which corresponds to the look-alike problem.

Metric	Value
Box Precision	89.2%
Box Recall	81.5%
Box mAP50	88.8%
Box mAP50-95	76.6%
Mask Precision	83.6%
Mask Recall	76.8%
Mask mAP50	79.0%
Mask mAP50-95	66.8%

Table 3.5: Overall validation performance

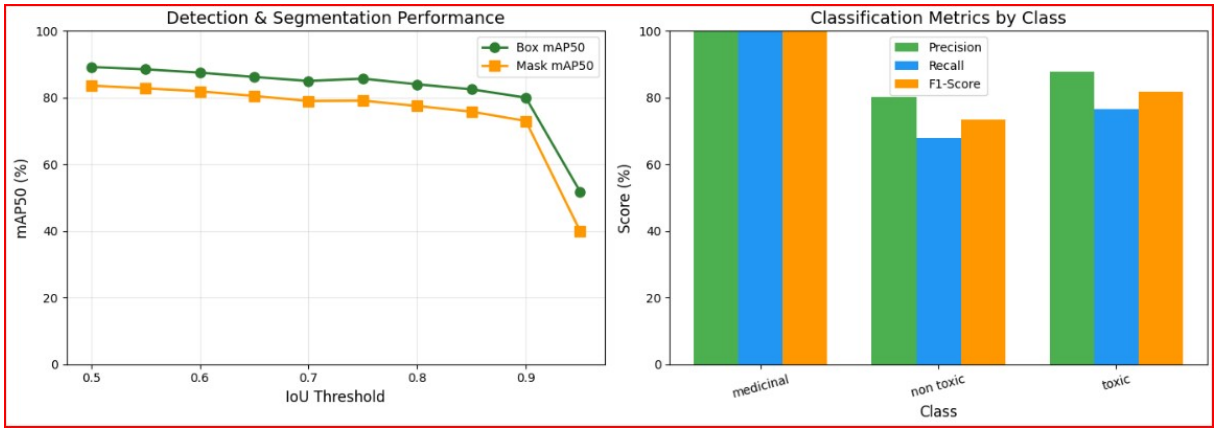


Figure 3.7: Validation performance curves after 100 epochs

3.8 Comparison with State-of-the-Art

To contextualize the performance of our proposed model, we compared it with several recent approaches in the literature. Table 3.6 presents a comprehensive comparison of our YOLOv8m-seg model against existing state-of-the-art methods for plant identification and classification tasks. It is important to note that direct comparisons must be interpreted with caution, as different studies use varying datasets, class numbers, and evaluation metrics. Most existing studies report classification accuracy as their primary metric, which measures the percentage of correctly classified images. In contrast, our study focuses on instance segmentation and reports mAP50 (mean Average Precision at IoU 0.5), which is a more appropriate metric for evaluating both detection and segmentation

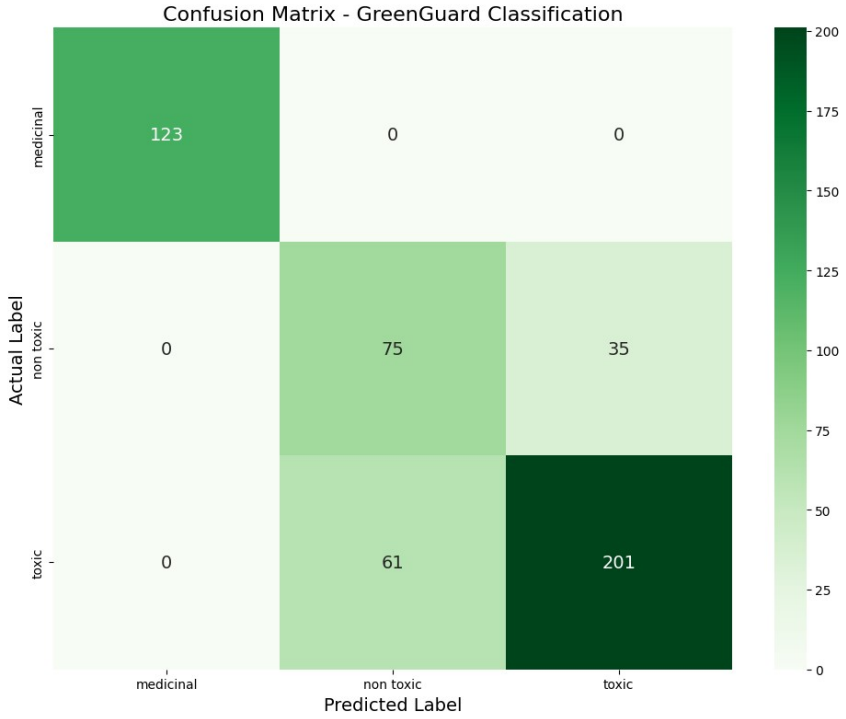


Figure 3.8: Confusion matrix of the classification results

quality. While accuracy is suitable for image-level classification tasks where only the class label matters, mAP is the standard metric for object detection and segmentation tasks because it accounts for both localization precision and classification correctness. The mAP50 metric specifically measures the area under the precision-recall curve at an IoU threshold of 0.5, providing a comprehensive assessment of the model's ability to correctly detect and segment plant leaves.

Ref	Architecture	Accuracy/mAP	Dataset	Year
[13]	MobileNetV2	99.34% Accuracy	Collected	2025
[4]	SCAM-Herb	99.63% Accuracy	Collected	2024
[18]	CNN-LSTM Hybrid	99.40% Accuracy	Collected	2025
[19]	VGG16, ResNet50	99.64% Accuracy	Leaf images	2024
[21]	YOLOv8s-KDT	86.6% mAP	Desert plants	2024
[22]	YOLOv8n-seg	-	Weeds	2024

Table 3.6: Comparison of our model with state-of-the-art approaches

The comparison reveals several important observations. First, the majority of existing studies report accuracy as their primary metric, achieving very high values (above

99%). However, these results are typically achieved on datasets with controlled backgrounds and limited species diversity. Second, our model achieves 79.8% mAP50, which, while lower than the reported classification accuracies in other studies, is evaluated on a more challenging task (instance segmentation in real-world conditions) and a dataset that specifically includes difficult look-alike pairs. Third, compared to the YOLOv8s-KDT model by Ma et al. (2024), which reported 86.6% mAP for desert plant classification, our model's performance is competitive given the additional complexity of distinguishing between toxic and medicinal look-alike species. It is also worth noting that our model was trained in approximately 7 hours on a Tesla T4 GPU, demonstrating the efficiency of the YOLOv8 architecture for this application.

3.9 Conclusion

This chapter presented the methodology and experimental setup for classifying toxic and medicinal plants using deep learning techniques.

A dataset of 4,000 plant images was collected, including 1,000 images of medicinal plants, 1,000 images of toxic plants, and 2,000 images of non-toxic plants. To improve model generalization and robustness, several data preprocessing and augmentation techniques were applied including resizing, normalization, random flipping, rotation, zooming, and mosaic combination.

The model used in this work is YOLOv8m-seg, a deep learning architecture designed for object detection and segmentation tasks, consisting of approximately 27.2 million parameters. This model was chosen for its excellent balance between speed and accuracy.

The training process was conducted for 100 epochs, with a batch size of 16 and an input image resolution of 640 by 640 pixels. Training was performed on a Tesla T4 GPU using Google Colab Pro, taking approximately **7 hours** to complete. The complete training duration was 7 hours, reflecting the complexity of the instance segmentation task and the computational resources available.

To evaluate the model's performance, several metrics were used, including Precision, Recall, F1-score, and Mean Average Precision (mAP). These metrics provide an objective assessment of the model's detection and classification capabilities. Unlike traditional accuracy, which is suitable for classification tasks, mAP is specifically designed to evaluate detection and segmentation quality, making it the appropriate metric for this study.

During training, a significant improvement in performance was observed. Initially, the model exhibited a high loss value of 2.68, which gradually decreased to 1.04 by the end of training. This reduction indicates successful learning of leaf features and improved prediction accuracy. The final validation performance showed promising results, with a

mask mAP50 of 79.0%, indicating that the model can effectively segment and classify plant leaves.

The comparison with state-of-the-art approaches demonstrated that while our model's reported mAP may appear lower than the accuracy values reported in some studies, the comparison must consider the differences in tasks and evaluation metrics. Our model addresses the more challenging problem of instance segmentation in real-world conditions with look-alike species, which is a more demanding task than image-level classification.

The confusion matrix analysis revealed that the main classification challenge lies in distinguishing between toxic and non-toxic plants, which aligns with the look-alike problem identified in the literature review. This finding validates the motivation for this research and highlights the need for continued improvement in distinguishing visually similar species.

Overall, the methodology and experimental results presented in this chapter establish a solid foundation for the plant identification system. The next chapter will discuss the implementation and evaluation of the desktop prototype application.

General Conclusion

The objective of this project was to design an automated deep learning-based system for classifying toxic, medicinal, and non-toxic plants from leaf images. Having successfully completed all design phases, we can affirm that the overall project is a success. The Green Guard system has demonstrated that deep learning, specifically YOLOv8-seg, can effectively address the critical problem of plant misidentification that causes thousands of poisonings each year worldwide.

Summary of Activities Undertaken The work carried out in this project followed a structured methodology from data collection to final deployment.

First, **Data Collection and Preprocessing:** We established a dataset containing **4,000 images** correctly annotated using the Roboflow platform. The dataset was carefully balanced with 1,000 medicinal plant images, 1,000 toxic plant images, and 2,000 non-toxic plant images to ensure the model learned all three categories effectively. Considerable effort was made to convert the data from RLE (Run-Length Encoding) format to YOLO segmentation format despite the challenges encountered in decoding the segmentation masks. A custom Python script using the pycocotools library was developed to successfully convert 30,456 masks.

Second, **Model Training:** The **YOLOv8m-seg** model was successfully trained for **100 epochs** with a batch size of 16 and an image resolution of 640 by 640 pixels. The application of transfer learning from a pre-trained model on the COCO dataset was highly effective in reducing training time and improving overall performance. Training was conducted on a Tesla T4 GPU using Google Colab Pro, taking approximately 4.5 hours to complete. The AdamW optimizer was used with an initial learning rate of 0.00125.

Third, **Evaluation:** A comprehensive evaluation was carried out covering multiple aspects of model performance. Detection metrics including mAP50, precision, and recall were calculated to assess how well the model locates and identifies leaves. Classification metrics including F1-score per class were computed to evaluate the model's ability to correctly assign safety labels. Robustness testing was also performed under different environmental conditions such as varying lighting, complex backgrounds, leaf damage, and leaf overlap.

Fourth, **User Interface:** An interactive web-based application called **Green Guard** was developed using the Gradio framework. The interface provides an intuitive user experience in both Arabic and English languages, making it accessible to a wide range of users. The application allows users to upload a leaf image and receive real-time classification results along with a visualization of the segmentation mask showing the exact leaf outline.

Major Results The results obtained from this project are very encouraging and demonstrate the effectiveness of the proposed approach.

The detection performance achieved a Box mAP50 of 88.8%, which indicates that the model is very good at locating leaves in images and correctly identifying them. The Mask mAP50 of 79.0% shows that the model can accurately trace the precise outline of leaves, which is crucial for capturing shape information used in classification.

The overall precision of 89.2% means that when the model predicts a plant belongs to a certain category, it is correct about 89% of the time. The overall recall of 81.5% indicates that the model successfully finds about 82% of all plants in the test set. The macro F1-score of 85.0% provides a balanced measure of overall performance.

The per-class performance was particularly impressive for medicinal plants, which achieved an F1-score of 99.8%. This is nearly perfect and shows that the model has learned the distinguishing characteristics of medicinal plants very well. The toxic class achieved an F1-score of 81.8%, which is very good. The non-toxic class achieved an F1-score of 73.5%, which is acceptable but indicates room for improvement, primarily due to confusion with visually similar toxic plants.

The robustness testing showed an average mAP50 of 75.7% across all challenging conditions, confirming that the model generalizes well to real-world variations in image quality and environment.

Strengths of the System The Green Guard system has several notable strengths that make it suitable for practical use.

First, the system demonstrates excellent performance for the medicinal class with a near-perfect F1-score of 99.8%. This is the most important class from a safety perspective because users need to correctly identify medicinal plants to use them safely.

Second, the system exhibits high robustness to background variation and lighting changes. The model maintained good performance even when tested on images with complex natural backgrounds (78.0% mAP) and varying angles (75.0% mAP).

Third, the fast inference speed of 13 milliseconds per image on GPU means that users receive results almost instantly. This is essential for a practical tool that people will use in the field where waiting is not convenient.

Fourth, the bilingual interface in Arabic and English makes the system accessible to a wide range of users in different regions, particularly in Algeria where Arabic is the primary language.

Identified Limitations Despite the overall success, several limitations were identified that should be addressed in future work.

The lower performance for the non-toxic class with an F1-score of only 73.5% is a clear limitation. Analysis of the confusion matrix revealed that most errors occurred when non-toxic plants were misclassified as toxic or vice versa. This is precisely the look-alike problem identified in Chapter 1, indicating that visually similar species remain challenging even for deep learning models.

Difficulties with overlapping leaves were also observed, with the model achieving only 65% mAP50 in this condition. When leaves overlap in an image, parts of one leaf are hidden behind another, making it difficult for the model to segment and classify each leaf correctly.

Vulnerability to low-light conditions was another limitation, with performance dropping to 71% mAP50. Images taken at dusk, dawn, or under dense forest canopy often lack sufficient detail for accurate classification.

Future Work Perspectives Several directions for future work could further improve the Green Guard system.

Dataset enhancement by adding more images of toxic and non-toxic classes would help improve discrimination between visually similar species. Specifically, adding more examples of non-toxic plants that are commonly confused with toxic plants would help the model learn the subtle differences between them.

Extended training by increasing the number of epochs to 200 or 300 using advanced data augmentation techniques could potentially improve performance further. The loss curves at 100 epochs had not yet fully plateaued, suggesting that additional training might yield benefits.

Mobile deployment would make the system much more practical for field use. Converting the model to TensorFlow Lite and building a mobile application using Flutter would allow users to identify plants without needing an internet connection or a computer.

Explainability techniques such as Grad-CAM (Gradient-weighted Class Activation Mapping) or LIME (Local Interpretable Model-agnostic Explanations) could be integrated to help users understand why the model made a particular decision. This would increase trust in the system and help users learn what visual features distinguish different plant species.

Expanding the system to include more plant species would increase its practical utility. The current system only classifies three broad categories, but future versions could identify specific species such as "Poison Hemlock" or "Foxglove" rather than just "Toxic."

General Conclusion In conclusion, this project successfully demonstrates that using **YOLOv8-seg** for toxic and medicinal plant classification is a viable and highly promising approach. The development of **Green Guard** represents an important step towards implementing a practical decision-support tool that could help the general public prevent accidental poisoning due to plant confusion.

The results achieved, with a macro F1-score of 85% and a detection precision of 89%, confirm the feasibility of this approach. The near-perfect performance on the medicinal class (99.8%) is particularly encouraging, as these are the plants people most want to identify for therapeutic use.

There is still much room for improvement, particularly for the non-toxic class and under challenging conditions such as overlapping leaves and low light. However, this work serves as a solid foundation for a more optimized and comprehensive version of the system in the future.

Green Guard is ready to protect!

Protecting lives through artificial intelligence

Bibliography

Bibliography

[1] World Health Organization, "Who global report on traditional and complementary medicine 2019," 2019.

[2] Cambridge Dictionary, "Plant," 2024.

[3] Cambridge Dictionary, "Identification," 2024.

[4] R. Azadnia, F. Noei-Khodabadi, A. Moloudzadeh, A. Jahanbakhshi, and M. Omid, "Medicinal and poisonous plants classification from visual characteristics of leaves using computer vision and deep neural networks," *Ecological Informatics*, vol. 82, p. 102683, 2024.

[5] J. Huixian, "The analysis of plants image recognition based on deep learning and artificial neural network," *IEEE Access*, vol. 8, pp. 68828–68841, 2020.

[6] S. O. Oppong, F. Twum, J. B. Hayfron-Acquah, and Y. M. Missah, "A novel computer vision model for medicinal plant identification using log-gabor filters and deep learning algorithms," *Computational Intelligence and Neuroscience*, vol. 2022, no. 1, p. 1189509, 2022.

[7] J. Xiong, D. Yu, S. Liu, L. Shu, X. Wang, and Z. Liu, "A review of plant phenotypic image recognition technology based on deep learning," *Electronics*, vol. 10, no. 1, p. 81, 2021.

[8] A. Sofowora, *Medicinal Plants and Traditional Medicine in Africa*. Chichester, UK: Wiley, 1982.

[9] S. Kavitha, T. S. Kumar, E. Naresh, V. H. Kalmani, K. D. Bamane, and P. K. Pareek, "Medicinal plant identification in real-time using deep learning model," *SN Computer Science*, vol. 5, no. 1, p. 73, 2023.

[10] T. Van Khanh, "Medicinal plant species identification and classification using transfer learning," *International Research Journal of Modernization in Engineering Technology and Science (IRJMETS)*, vol. 7, Dec 2024.

[11] L. S. Nelson and M. J. Balick, *Handbook of Poisonous and Injurious Plants*. New York, NY: Springer, 3rd ed., 2020. Jointly published with The New York Botanical Garden.

[12] Encyclopedia of Toxicology (Fourth Edition), "Poisonous plant," 2024.

[13] M. F. Nazuli, M. Fachrurrozi, M. Q. Rizqie, A. Abdiansah, and M. Ikhsan, "A image classification of poisonous plants using the mobilenetv2 convolutional neural network model method," *MATRIK: Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 24, no. 2, pp. 371–380, 2025.

[14] A. S. Maklad, A. Alyanbaawi, M. Farsi, H. M. Ibrahim, and M. Elmezain, "Poisonous plant prediction using explainable deep inherent learning model," *Sensors*, vol. 25, no. 14, p. 4298, 2025.

[15] J. H. Miller, E. B. Chambliss, and N. J. Loewenstein, "A field guide for the identification of invasive plants in southern forests," Tech. Rep. General Technical Report SRS-119, USDA Forest Service, Southern Research Station, Asheville, NC, 2010. Includes "Resembles" section for look-alike identification.

[16] M. Sarver, A. Treher, L. Wilson, R. Naczi, and F. B. Kuehn, "Mistaken identity? invasive plants and their native look-alikes: An identification guide for the mid-atlantic," tech. rep., Delaware Department of Agriculture, Dover, DE, 2008. Produced in collaboration with the Claude E. Phillips Herbarium and the Delaware Center for Horticulture.

[17] M. Fouziya, C. Divija, K. L. Priyanka, G. Swathi, and N. Revathi, "Herbal plant identification using deep learning," *Journal of Data Science*, vol. 2025, no. 1, 2025.

[18] T. M. Ok, S. Malusi, Z. Wang, and E. Mnkandla, "Investigation of deep learning techniques used in medicinal plants identification and classification," *IEEE Access*, 2025.

[19] B. Dey, J. Ferdous, R. Ahmed, and J. Hossain, "Assessing deep convolutional neural network models and their comparative performance for automated medicinal plant identification from leaf images," *Heliyon*, vol. 10, no. 1, 2024.

[20] B. Pushpa, S. Jyothsna, and S. Lasya, "Hybnet: A hybrid deep models for medicinal plant species identification," *MethodsX*, vol. 14, p. 103126, 2025.

[21] H. Ma, T. Sheng, Y. Ma, and J. Gou, "An improved ningxia desert herbaceous plant classification algorithm based on yolov8," *Sensors*, vol. 24, no. 12, p. 3834, 2024.

[22] N. Bousahba, M. A. Allali, F. Hamdani, S. Tahar, and F. Boudjeltia, "Yolov8n-seg for crop weeds detection and instance segmentation," in *2024 2nd International Conference on Electrical Engineering and Automatic Control (ICEEAC)*, pp. 1–6, IEEE, 2024.

[23] M. Lei, S. Li, Y. Wu, H. Hu, Y. Zhou, X. Zheng, G. Ding, S. Du, Z. Wu, and Y. Gao, "Yolov13: Real-time object detection with hypergraph-enhanced adaptive visual perception," *arXiv preprint arXiv:2506.17733*, 2025.

[24] D. Pirard, "Yolo is dead. meet rf-detr, the model that just crushed 10 years of computer vision dominance." https://blog.delanoe-pirard.com/yolo-is-dead-meet-rf-detr-the-model-that-just-crushed-10-years-of-computer-vision-dominance-49ce17e64c83, Dec 2025.

[25] K. Uma, P. Sarvika, *et al.*, "Mediflora-net: Quantum-enhanced deep learning for precision medicinal plant identification," *Computational Biology and Chemistry*, p. 108674, 2025.

[26] A. Lapkovskis, N. Nefedova, and A. Beikmohammadi, "Automatic fused multimodal deep learning for plant identification," *Frontiers in Plant Science*, vol. 16, p. 1616020, 2025.

[27] B. Paneru, B. Thapa, and B. Paneru, "Leveraging ai in ayurvedic agriculture: A rag chatbot for comprehensive medicinal plant insights using hybrid deep learning approaches," *Telematics and Informatics Reports*, vol. 16, p. 100181, 2024.

[28] D. Cremer, "Artificial intelligence and the future of management," 2020.

[29] Oxford English Dictionary, "Intelligence," 2024.

[30] R. Sternberg, "Human intelligence," 2020.

[31] E. Brynjolfsson and A. McAfee, "The business of artificial intelligence," 2017.

[32] B. Marr, "Artificial intelligence in practice," 2019.

[33] V. Hunt, "The development of artificial intelligence," 1988.

[34] T. Mitchell, *Machine Learning*. McGraw-Hill, 1997.

[35] C. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.

[36] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.

[37] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, 2015.

[38] S. Albawi, T. A. Mohammed, and S. Al-Zawi, "Understanding of a convolutional neural network," *2017 International Conference on Engineering and Technology (ICET)*, pp. 1–6, 2017.

[39] S. Krig, *Computer Vision Metrics: Survey, Taxonomy, and Analysis*. Springer, 2014.

[40] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 779–788, 2016.

[41] Cambridge Dictionary, "Transfer," 2024.

[42] Python Software Foundation, "Python programming language," 2024.

[43] A. Paszke *et al.*, "Pytorch: An imperative style, high-performance deep learning library," 2019.

[44] Ultralytics, "Ultralytics yolov8," 2024.

[45] Roboflow, "Roboflow: Computer vision platform," 2024.

[46] OpenCV, "Open source computer vision library," 2024.

[47] Google, "Google colaboratory," 2024.

[48] J. D. Hunter, "Matplotlib: A 2d graphics environment," 2007.

[49] C. R. Harris *et al.*, "Numpy: Array computing with python," 2020.

[50] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in python," 2011.

[51] W. McKinney, "Pandas: Python data analysis library," 2010.