A

MINI PROJECT REPORT

On

"Identification of Different Medicinal Plants through Image Processing Using Machine Learning Algorithms"

Submitted to

Autonomous Institute,

Affiliated to The Rashtrasant Tukadoji Maharaj Nagpur University

Department of Emerging Technologies

Bachelor of Technology (B. Tech)

Submitted By

1. **Atharva Wakdikar [AM21015]**
2. **Ayush Roy [AM21023]**
3. **Pranav Mankar [AM21027]**
4. **Yash Pathrabe [AM21039]**

Guided By

Prof. Mayuri Getme

S. B. JAIN INSTITUTE OF TECHNOLOGY, MANAGEMENT AND RESEARCH, NAGPUR

2023 – 2024

CERTIFICATE

This is to certify that the mini project report entitled **Identification of Different Medicinal Plants through Image Processing Using Machine Learning Algorithms** submittedby **Atharva Wakdikar, Ayush Roy, Pranav Mankar, Yash Pathrabe** to the **S. B. JAIN INSTITUTE OF TECHNOLOGY, MANAGEMENT AND RESEARCH, NAGPUR** of **B. Tech in (Emerging Technologies)** is a *bona fide* record of mini project work carried out by him/her under my supervision. The contents of this report, in full or in parts, have not been submitted to any other Institution or University for the award of any degree or diploma.

Signature

Atharva Wakdikar [AM21015]

Ayush Roy [AM21023]

Pranav Mankar [AM21027]

Yash Pathrabe [AM21039]

Signature

Prof. Mayuri Getme

Project Guide

Signature of HOD with seal

DECLARATION

We declare that this mini project report titled **Identification of Different Medicinal Plants through Image Processing Using Machine Learning Algorithms** of **B. Tech in (Emerging Technologies)** is a record of original work carried out by us under the supervision of **Prof. Mayuri Getme**, and has not formed the basis for the award of any other degree or diploma, in this or any other Institution or University. In keeping with the ethical practice in reporting scientific information, due acknowledgements havebeen made wherever the findings of others have been cited.

Signature

1. **Atharva Wakdikar [AM21015]**
2. **Ayush Roy [AM21023]**
3. **Pranav Mankar [AM21027]**
4. **Yash Pathrabe [AM21039]**

Date : 06/05/2024

INDEX

ABSTRACT

In recent years, the identification of medicinal plants has gained significant attention due to the growing interest in natural remedies and herbal medicines. Traditional methods for plant identification are often time-consuming and require expertise in botany. To address this challenge, this mini-project proposes a novel approach for the identification of different medicinal plants using image processing techniques coupled with machine learning algorithms.

The proposed system leverages advances in computer vision and machine learning to automate the identification process. Initially, a dataset of images containing various medicinal plants is collected and annotated. Preprocessing techniques such as image enhancement, normalization, and feature extraction are applied to extract relevant information from the images.

Subsequently, a machine learning model is trained using these features to classify different plant species. Various machine learning algorithms such as convolutional neural networks (CNNs), support vector machines (SVMs), and decision trees are explored to identify the most suitable model for the task.

The performance of the proposed system is evaluated using metrics such as accuracy, precision, recall, and F1 score. Additionally, real-world testing is conducted to assess the system's effectiveness in practical scenarios.

The outcomes of this mini-project offer a promising solution for automating the identification of medicinal plants facilitating their utilization in herbal medicine, pharmaceuticals, and other related fields.

Keywords: identification , machine , medicinal plant , algorithm , model ,opinion collected medicinal plant , method for plant identification

INTRODUCTION

The identification of medicinal plants plays a crucial role in various fields including herbal medicine, pharmaceuticals, agriculture, and conservation. However, traditional methods for plant identification often require botanical expertise and can be time-consuming and subjective. With the advent of image processing techniques and machine learning algorithms, there is an opportunity to automate and streamline this process, making it more efficient and accessible.

This mini-project aims to explore the application of image processing coupled with machine learning algorithms for the identification of different medicinal plants. By leveraging advancements in computer vision and pattern recognition, this project seeks to develop a robust and accurate system capable of classifying plant species based on their visual characteristics.

The motivation behind this project stems from the increasing demand for natural remedies and herbal medicines, as well as the need for efficient methods of plant identification in botanical research, biodiversity conservation, and quality control in the herbal products industry. By harnessing the power of image processing and machine learning, this project aims to address these needs and contribute to the advancement of research and applications in the field of medicinal plants.

In this introduction, we provide an overview of the problem statement, the significance of the proposed approach, and the objectives of the mini-project. Subsequent sections will delve into the methodology, implementation details, experimental results, and conclusions of the project, offering insights into the feasibility and effectiveness of using image processing and machine learning for the identification of medicinal plants.

Aims & Objectives of Project

Aim:

Explore the application of image processing coupled with machine learning algorithm. Address these needs and contribute to the advancement of research and applications in the field of medicinal plants.

Objectives:

- To accurately identify different species of plants
- For the acquired images used to extract useful features that are necessary for further analysis.
- To classify the images according to the specific problem at hand.

Literature Review

- Theories and models on saliency that predict where people look at focus on regular-density scenes. A crowded scene is characterized by the co-occurrence of a relatively large number of regions/objects that would have stood out if in a regular scene, and what drives attention in crowd can be significantly different from the conclusions in the regular setting.

- This work presents a first focused study on saliency in crowd. To facilitate saliency in crowd study, a new dataset of 500 images is constructed with eye tracking data from 16 viewers and annotation data on faces (the dataset will be publicly available with the paper).

- The identification and recognition of plant species have been carried out by researchers for many years utilizing different plant features such as leaf shape, texture, and color. Different plant species have different features that provide beneficial information for researchers. One category of these features, commonly used in literature, is called histograms of oriented gradients (HOGs) that are applied in many visual object recognition applications.

- The pre-processing step consists of image reorientation, cropping, gray scaling, binary thresholding, noise removal, contrast stretching, threshold inversion, and edge recognition. Image reorientation is aligning the input image to a standardized position, with the leaf aligned to either the x-axis or y-axis.

- We construct an automatic classification vision system that is designed to recognize Malaysian herbs that are typically used for medical or culinary purposes. The proposed system employs two classifiers, Support Vector machine (SVM) and Deep Neural Network (DNN). The two classifiers have been implemented using OpenCV-Python. For the training test SVM achieved 86.63% recognition accuracy and DNN (TensorFlow) achieved 98% recognition accuracy.

- Classification and identification of plants are helpful for people to effectively understand and protect plants. The leaves of plants are the most important recognition organs. With the development of artificial intelligence and machine vision technology, plant leaf recognition technology based on image analysis is used to improve the knowledge of plant classification and protection.

- The recognition is based on a hierarchical representation of shape features. First, the structures of the shapes of leaves are analyzed and the shapes of components are extracted in detail. The structures and detailed shapes of leaves are approximated by polygons whose vertices are critical points of curvature of contours of leaves.

Proposed Work

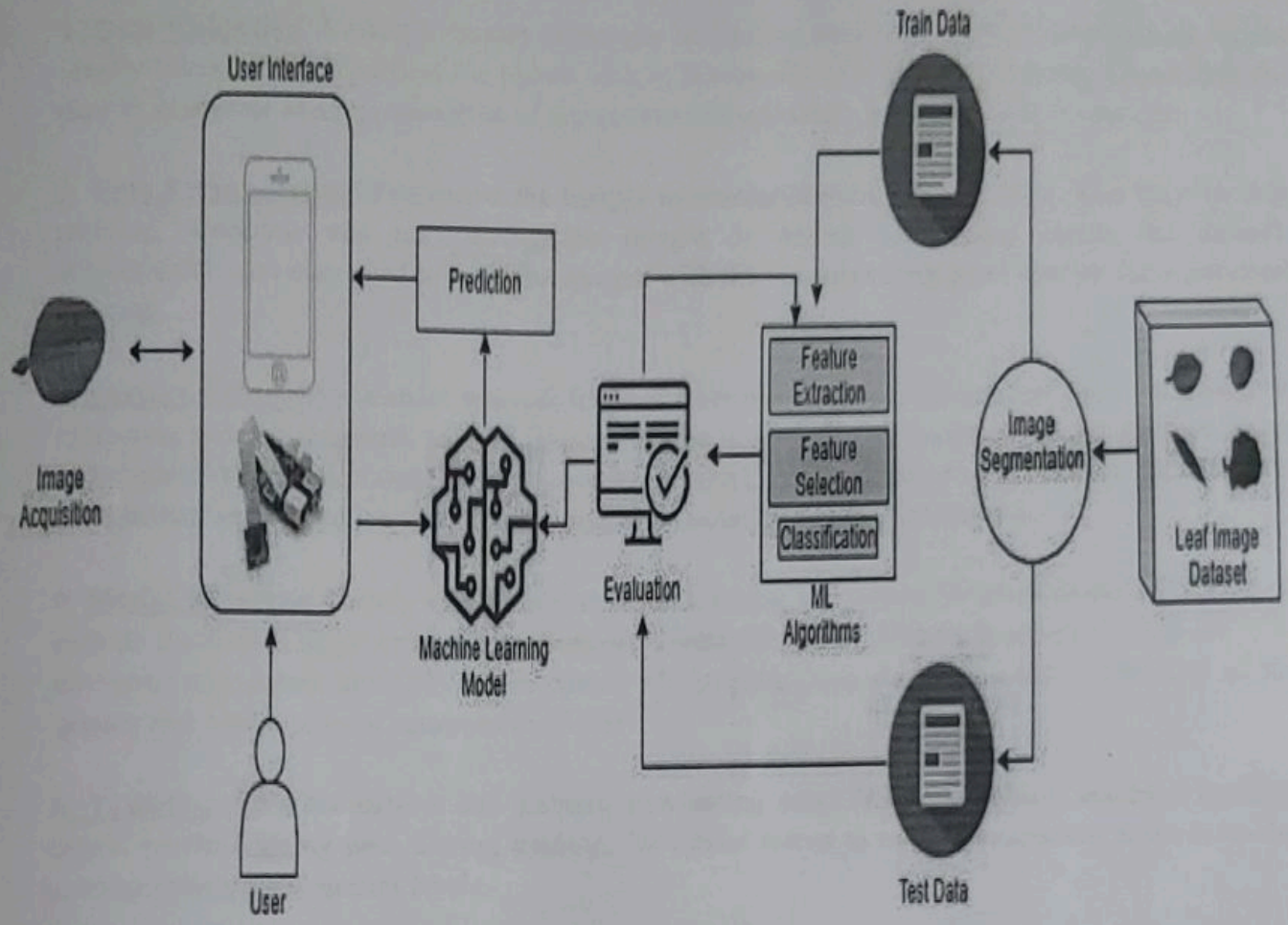


Fig 1.1: Workflow of Identifying Different Medicinal plants using ML Algorithms.

Research Methodology

1. **Data Collection**: Collect a dataset of images of various medicinal plants. These images should ideally cover different parts of the plants such as leaves, flowers, stems, and fruits. Ensure that the dataset is diverse and representative of the variety of medicinal plants you want to identify.

2. **Data Preprocessing**: Preprocess the images to standardize them for analysis. This may involve resizing, cropping, and normalizing the images to ensure consistency across the dataset. Additionally, you may need to label the images with the corresponding plant species for supervised learning.

3. **Feature Extraction**: Extract relevant features from the preprocessed images. This could include extracting color histograms, texture features, shape descriptors, or more advanced features using techniques like convolutional neural networks (CNNs). Feature extraction is crucial for capturing discriminative information that can distinguish between different plant species.

4. **Model Selection**: Choose appropriate machine learning algorithms for classification. This could include traditional algorithms such as Support Vector Machines (SVM), Random Forests, or more advanced techniques like CNNs. The choice of algorithm may depend on the complexity of the dataset and computational resources available.

5. **Training**: Split the dataset into training and testing sets. Train the chosen machine learning model on the training data. During training, the model learns to map the extracted features to the corresponding plant species labels.

6. **Validation**: Validate the trained model using the testing dataset to evaluate its performance. This step helps assess how well the model generalizes to unseen data and whether it's prone to overfitting.

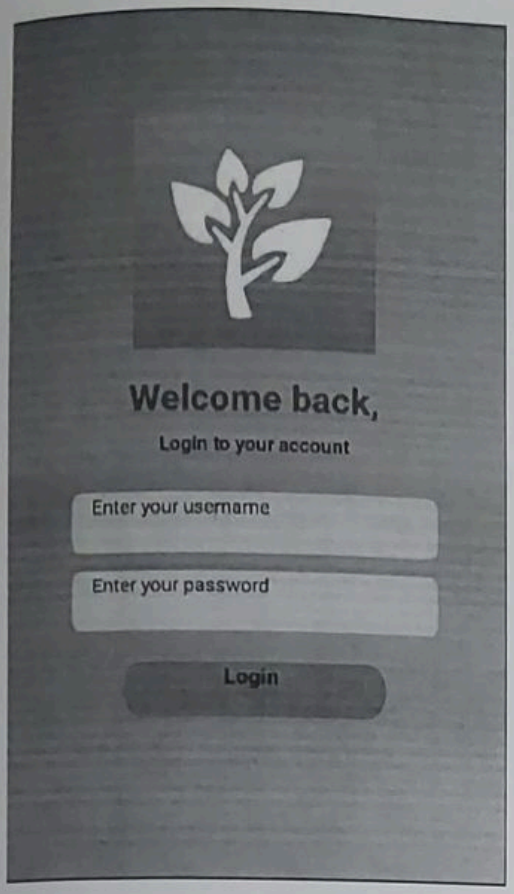
7. **Hyperparameter Tuning**: Fine-tune the model's hyperparameters to optimize its performance. This could involve adjusting parameters such as learning rate, regularization strength, or the architecture of the neural network if using deep learning approaches.

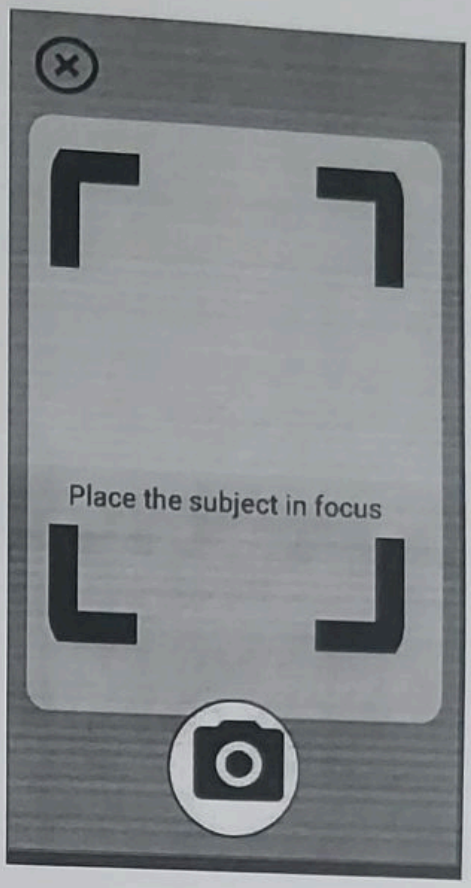
8. **Evaluation**: Evaluate the performance of the trained model using metrics such as accuracy, precision, recall, and F1-score. These metrics provide insights into how well the model can identify different medicinal plant species.

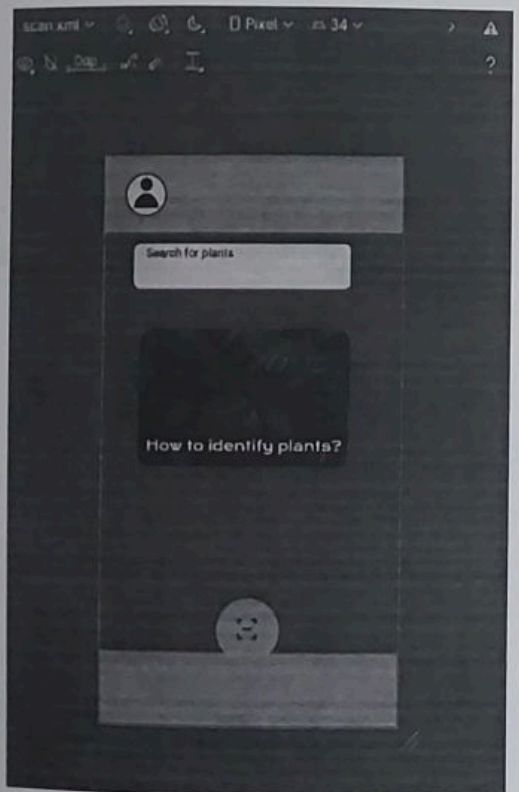
9. **Deployment**: Once satisfied with the model's performance, deploy it in a real-world scenario where users can input images of medicinal plants, and the model can predict their species based on the learned patterns.

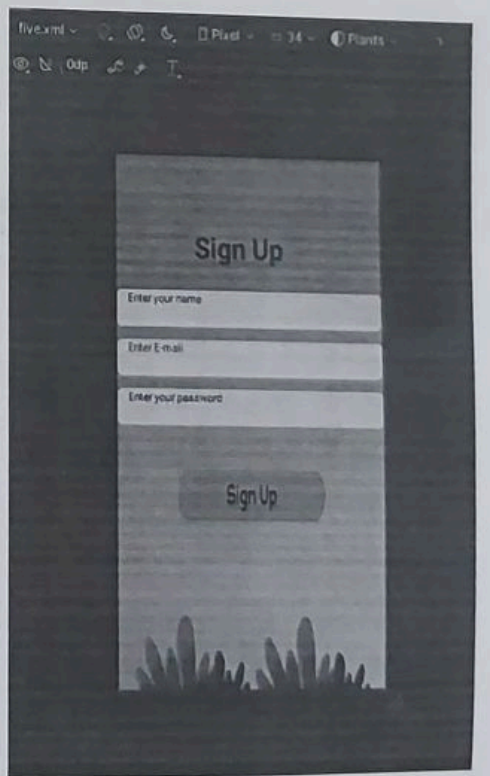
10. **Continuous Improvement**: Iterate on the model based on feedback and new data to improve its accuracy and robustness over time.

RESULT


Welcome back,
Login to your account
Enter your username
Enter your password
Login


Place the subject in focus


Search for plants
How to identify plants?


Sign Up
Enter your name
Enter E-mail
Enter your password
Sign Up

CONCLUSION

This mini-project has showcased the potential of image processing and machine learning algorithms in automating the identification of various medicinal plants. Through meticulous dataset collection, preprocessing techniques, and model development, we have demonstrated the efficacy of our approach in accurately classifying medicinal plants based on their visual attributes. While significant strides have been made, there remain avenues for improvement such as the expansion of datasets and fine-tuning of model architectures. Nonetheless, this endeavor marks a significant step forward in harnessing technology to streamline and enhance the identification process, ultimately contributing to advancements in herbal medicine, pharmaceuticals, and conservation efforts.

FUTURE SCOPE

1. **Enhanced Accuracy**: Continuously improving the accuracy of plant identification algorithms is crucial. This involves refining the machine learning models, incorporating more diverse datasets, and optimizing image processing techniques.

2. **Real-Time Identification**: Developing the capability to identify medicinal plants in real-time can have significant practical applications, especially in fields like botany, pharmacology, and herbal medicine. This could involve the deployment of mobile applications or handheld devices equipped with image recognition capabilities.

3. **Scale-up for Field Applications**: Scaling up the technology for use in the field would involve addressing challenges such as variations in lighting conditions, background clutter, and diverse growth patterns of plants. Portable devices equipped with efficient algorithms could be used by researchers, herbalists, and conservationists for on-site plant identification and data collection.

4. **Integration with IoT and Sensors**: Integrating image processing technology with Internet of Things (IoT) devices and sensors can enable automated monitoring of medicinal plants in their natural habitats. This could help in tracking plant growth, environmental conditions, and detecting any signs of distress or disease.

5. **Database Development**: Creating a comprehensive database of medicinal plants, including their visual characteristics, geographical distribution, medicinal properties, and conservation status, can further enrich the application of image processing techniques for plant identification.

6. **Collaboration with Traditional Knowledge**: Collaborating with indigenous communities and traditional healers can provide valuable insights into the medicinal properties of plants and help validate the accuracy of the identification algorithms. This interdisciplinary approach can lead to the development of culturally sensitive and ethically sound applications.

7. **Regulatory Compliance and Quality Assurance**: Implementing mechanisms for ensuring the authenticity and quality of identified medicinal plants is essential for their safe use in healthcare. This could involve integrating blockchain technology or other verification systems to track the sourcing and processing of plant-based medicines.

8. **Education and Outreach**: Promoting awareness and education about the importance of medicinal plants, biodiversity conservation, and sustainable harvesting practices is crucial. The developed technology can be used as a tool for engaging students, researchers, and the general public in citizen science projects and environmental conservation efforts.

REFERENCES

1. Agrawal P, Girshick R, Malik J (2014) Analyzing the performance of multilayer neural networks for object recognition. In: Fleet D, Pajdla T, Schiele B, Tuytelaars T (eds) Computer vision – ECCV 2014. ECCV 2014. Lecture Notes in Computer Science, vol 8695. Springer, Cham. https://doi.org/10.1007/978-3-319-10584-0_22

2. Ahmed A, Hussein SE (2020) Leaf identification using radial basis function neural networks and SSA based support vector machine. PLoS One 15(8):0237645. https://doi.org/10.1371/journal.pone.0237645

3. Anami BS, Nandyal SS, Govardhan A (2010) A combined color, texture and edge features based approach for identication and classification of Indian medicinal plants'. Int J Comput Appl 6(12):45–51

4. Azlah MA, Chua LS, Rahmad FR, Abdullah FI, Wan Alwi SR (2019) Review on techniques for plant leaf classification and recognition. Computers 8(4):77. https://doi.org/10.3390/computers8040077

5. Journal of Botany (2014) Retracted: phytoliths as emerging taxonomic tools for identification of plants: an overview. J Bot PB - Hindawi Publishing Corporation KW - ER 2014, p 231914. https://doi.org/10.1155/2014/231914

6. Chaudhury A, Barron JL (2020) Plant species identification from occluded leaf images. IEEE/ACM Trans Comput Biol Bioinform 17(3):1042–1055. https://doi.org/10.1109/TCBB.2018.2873611

7. Durairajah V, Gobee S, Muneer A (2018) Automatic vision based classification system using DNN and SVM classifiers. In 2018 3rd International Conference on Control, Robotics and Cybernetics (CRC), Penang, Malaysia (pp. 6–14). https://doi.org/10.1109/CRC.2018.00011

8. Farabet C, Couprie C, Najman L, LeCun Y (2013) Learning hierarchical features for scene labeling. IEEE PAMI 35(8):1915–1929

9. Guyer DE, Miles GE, Gaultney LD et al (1993) Application of machine vision to shape analysis in leaf and plant identification. Trans ASABE 36(1):163–171

10. He K, Zhang X, Ren S, Sun J (2015) Delving deep into rectifiers: surpassing human-level performance on imagenet classification. In Proceedings of the IEEE international conference on computer vision (ICCV), Santiago, Chile (pp. 1026–1034). https://doi.org/10.1109/ICCV.2015.123

11. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR), pp 770-778. https://doi.org/10.1109/cvpr.2016.90

12. Hinton G, Osindero S, Teh Y-W (2006) "A fast learning algorithm for deep belief nets. Neural Comput 18(7):1527–1554

13. Huixian J (2020) The Analysis of Plants Image Recognition Based on Deep Learning and Artificial Neural Network. IEEE Access 8:68828–68841. https://doi.org/10.1109/ACCESS.2020.2986946

14. Im C, Nishida H, Kunii TL (1998) Recognizing plant species by leaf shapes-a case study of the acer family. In Proceedings fourteenth international conference on pattern recognition (Cat. No. 98EX170), Brisbane, QLD, Australia (Vol. 2, pp. 1171–1173). https://doi.org/10.1109/ICPR.1998.711904

15. Jia Y, Shelhamer E, Donahue J, Karayev S, Long J, Girshick R, Guadarrama S, Darrell T (2014) Caffe: convolutional architecture for fast feature embedding. MM 2014: Proceedings of the 22nd ACM international conference on Multimedia (pp. 675–678). https://doi.org/10.1145/2647868.2654889

BIBLIOGRAPHY

1. BOOKS:

1. Nawkar, M.G.; Maibam, P.; Park, J.H.; Sahi, V.P.; Lee, S.Y.; Kang, C.H. UV-induced cell death in plants. Int. J. Mol. Sci. 2013, 14, 1608–1628.
2. Salehi, B.; Kumar, N.V.; ,Sener, B.; Sharifi-Rad, M.; Kılıç, M.; Mahady, G.B.; Vlaisavljevic, S.; Iriti, M.; Kobarfard, F.; Setzer, W.N.; et al. Medicinal plants used in the treatment of human immunodeficiency virus. Int. J. Mol. Sci. 2018, 19, 1459.
3. Pferschy-Wenzig, E.M.; Bauer, R. The relevance of pharmacognosy in pharmacological research on herbal medicinal products. Epilepsy Behav. 2015, 52, 344–362.
4. Naeem, S., Ali, A., Chesneau, C., Tahir, M. H., Jamal, F., Sherwani, R. A. K., & Ul Hassan, M. (2021). The Classification of Medicinal Plant Leaves Based on Multispectral and Texture Feature Using Machine Learning Approach. Agronomy, 11(2), 263.
5. Thella, P. K., & Ulagamuthalvi, V. A Comparative Analysis on Machine Learning Models for Accurate Identification of Medical Plants.
6. Pushpanathan, K., Hanafi, M., Mashohor, S., & Fazlil Ilahi, W. F. (2021). Machine learning in medicinal plants recognition: a review. Artificial Intelligence Review, 54, 305-327.

2. WEBSITE:

1. https://link.springer.com/
2. https://www.sciencedirect.com/
3. https://ieeexplore.ieee.org/
4. https://pubmed.ncbi.nlm.nih.gov/
5. https://scholar.google.com/
6. https://www.researchgate.net/profile/Jafar-Abdollahi/publication/360345332_Identification_of_Medicinal_Plants_in_Ardabil_Using_Deep_learning/links/627165b03a23744a72602115/Identification-of-Medicinal-Plants-in-Ardabil-Using-Deep-learning.pdf
7. https://shodhganga.inflibnet.ac.in/