Identification and Classification of Medicinal Plants using Deep Learning

Dr R Manjunath Prasad[1], Manasa S[2], Bhavana GM[3] Nithya G[4], Nisarga K C[5]
[1]Professor and Dean (Student Affairs), Department of Electronics and Communication, Dayananda Sagar Academy of Technology and Management, Bangalore, Karnataka, India
[2,3,4,5] Under Graduate Student, Department of Electronics and Communication, Dayananda Sagar Academy of Technology and Management, Bangalore, Karnataka, India

Abstract
This study explores the potential of deep learning for automating medicinal plant classification using leaf images. It investigates the effectiveness of a deep learning model built upon the pre-trained ResNet-50 architecture. The model utilizes transfer learning to leverage pre-existing knowledge from ResNet-50, fine-tuning it for the specific purpose of medicinal plant identification from the LeafSnap dataset. Evaluation of the proposed model yielded a high accuracy of approximately 99.86%. This achievement underscores the efficacy of deep learning, particularly the ResNet-50 architecture, in automating medicinal plant classification. These findings suggest the potential of this approach for various applications, including supporting field-based plant identification, streamlining herbarium processes, and potentially contributing to the development of innovative drug discovery pipelines.

Keywords- Medicinal plants, Deep learning, Convolutional Neural Networks (CNNs), ResNet-50, Transfer learning, Data augmentation, LeafSnap dataset, Image classification

I. INTRODUCTION

This study explores the potential of deep learning for automating medicinal plant classification using leaf images. It investigates the effectiveness of a deep learning model built upon the pre-trained ResNet-50 architecture. The model utilizes transfer learning to leverage pre-existing knowledge from ResNet-50, fine-tuning it for the specific purpose of medicinal plant identification from the LeafSnap dataset. Evaluation of the proposed model yielded a high accuracy of approximately 99.86%. This achievement underscores the efficacy of deep learning, particularly the ResNet-50 architecture, in automating medicinal plant classification. These findings suggest the potential of this approach for various applications, including supporting field-based plant identification, streamlining herbarium processes, and potentially contributing to the development of innovative drug discovery pipelines.

For centuries, medicinal plants have been the backbone of traditional medicine, but accurately identifying them is essential for safe and effective herbal treatments. Traditional methods, reliant on physical features, can be laborious and prone to error. Deep learning, a powerful tool in artificial intelligence, offers a promising solution for automated plant classification using leaf images. This study investigates a deep learning model built on the pre-trained ResNet-50 architecture. This model leverages pre-existing knowledge from large image datasets and fine-tunes it for medicinal plant identification from the LeafSnap dataset. Additionally, data augmentation techniques are used to improve model performance and prevent overfitting. Evaluation on the LeafSnap dataset demonstrates a remarkable accuracy of nearly 99.86%, highlighting the effectiveness of deep learning for automated medicinal plant classification.

Fueled by a rich history in traditional medicine, medicinal plants continue to provide valuable natural remedies and inspire modern drug discovery. However, ensuring the safety and efficacy of herbal treatments hinges on accurate plant identification. Conventional methods, which depend on analyzing physical characteristics, are often time-consuming, subjective, and vulnerable to human error. This vulnerability can lead to misidentification, potentially jeopardizing the safety and effectiveness of herbal treatments.

This research delves into the exciting potential of deep learning, a cutting-edge subfield of artificial intelligence, for automated medicinal plant classification using leaf images. Deep learning's ability to extract intricate patterns from data makes it ideally suited for image recognition and classification tasks.

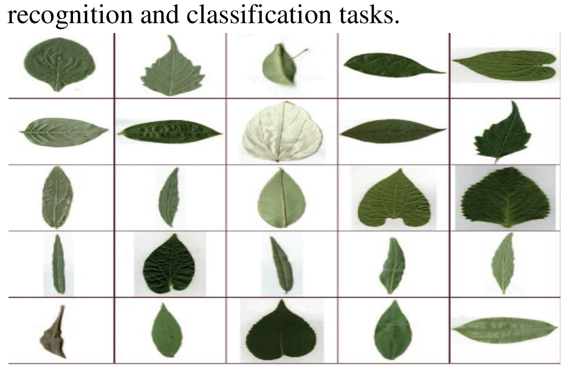


Figure.1: *Sample of Preprocessed images*

Deep learning, a powerful AI technique, offers a revolutionary approach to medicinal plant identification.

By analyzing leaf images, it can overcome the limitations of traditional methods, which rely on physical features and are prone to error. This paves the way for more accurate and automated plant classification, ensuring the safety and efficacy of herbal treatments.

II. OVERVIEW OF ResNet-50 MODEL

For this project, we leverage ResNet-50, a Convolutional Neural Network (CNN). CNNs excel at image recognition and classification. They achieve this by passing images through layers that automatically extract features like edges, shapes, and textures. These extracted features are then progressively combined and analyzed by later layers, ultimately allowing the network to categorize the image. ResNet-50 tackles a common hurdle in deep learning: the vanishing gradient problem. This issue limits how deep a network can be by hindering the flow of information through layers. ResNet addresses this by introducing residual learning. Here, information from each layer is added to the output of a shortcut connection, bypassing some layers. This allows the network to learn refined features building on earlier layers, enabling deeper architectures.

In our project, we leverage a pre-trained version of ResNet-50. These models are trained on enormous datasets like ImageNet, allowing them to learn powerful image recognition abilities. Transfer learning lets us use these pre-trained weights as a foundation for our medicinal plant model. Essentially, we fine-tune the final layers, adapting the learned features to identify medicinal plants from leaf images. This approach drastically reduces training time and allows the model to focus on the specifics of medicinal plant classification.

ResNet-50 overcomes a common obstacle in deep learning: vanishing gradients. This problem restricts network depth by hindering information flow through layers. To address this, ResNet incorporates residual learning. Here, information from each layer is directly added to the output of a special shortcut connection, bypassing some layers. This allows the network to learn incremental changes in features building on earlier layers. This approach effectively mitigates vanishing gradients and enables the creation of deeper architectures.

III. LITERATURE SURVEY

Several studies explore the potential of deep learning for automated medicinal plant classification due to its advantages over traditional methods. These methods, reliant on physical features, can be slow, subjective, and error-prone.

One such system, named "DeepHerb," utilizes deep learning to automatically identify medicinal plants from leaf images. This research recognizes the critical threat of declining plant diversity and its impact on both environmental health and access to valuable traditional medicine. To address limitations in data availability, the study created a new dataset named "DeepHerb," containing a diverse collection of medicinal plant leaf images.

The system leverages pre-trained convolutional neural network (CNN) models to extract features from the images. These features are then classified using a combination of artificial neural networks (ANNs) and support vector machines (SVMs). Additionally, the study employed optimization techniques to further enhance performance. This approach achieved an impressive accuracy rate, demonstrating the potential of deep learning for medicinal plant classification.

Furthermore, the research developed a user-friendly mobile application that integrates the deep learning model, allowing for rapid plant identification using a smartphone camera. This not only highlights the potential of deep learning in this field but also showcases its practical applications in promoting citizen science and public engagement with medicinal plants [1].

Another deep learning approach for medicinal plant classification is named "AyurLeaf." This system tackles the limitations of traditional identification methods by automatically classifying medicinal plants based on leaf image analysis.
AyurLeaf leverages a Convolutional Neural Network (CNN) architecture. CNNs can automatically learn image features directly from the data, eliminating the need for manual feature engineering, a time-consuming and specialized task. The research highlights the potential benefits of such a system for applications in Ayurveda, a traditional Indian medical system that relies heavily on medicinal plants. By enabling accurate plant identification, this approach could improve the effectiveness and safety of treatments within this medical system [2].

Another study investigated the use of deep ensemble learning for medicinal plant classification. Ensemble learning combines predictions from multiple models to potentially outperform individual models. This research explored an ensemble of pre-trained convolutional neural networks (CNNs) for medicinal leaf identification. Their approach utilized transfer learning by initializing multiple CNN architectures, like MobileNetV2, InceptionV3, and ResNet50, with pre-trained knowledge. These models were then fine-tuned on a medicinal plant leaf image dataset for the specific task.
The final ensemble model combined predictions from each CNN through weighted averaging. This strategy aimed to leverage the strengths of each model and potentially improve classification accuracy compared to a single CNN. The study reported achieving a high

accuracy of 99.66% on the test set using this deep ensemble learning approach [3].

While traditional methods for identifying medicinal plants rely on physical features, leading to slow, subjective, and error-prone processes, deep learning offers a promising alternative. This technology automates plant classification and has the potential to be more accurate.
Several studies have explored deep learning, particularly Convolutional Neural Networks (CNNs), for medicinal plant recognition using leaf images. These studies showcase the effectiveness of deep learning in this field. The growing interest in automated plant identification has fueled the development of deep learning techniques. These approaches hold significant advantages over traditional methods by learning complex patterns from image data, leading to more accurate and efficient plant classification.
Furthermore, recent research has demonstrated the potential of deep learning for real-time medicinal plant identification, paving the way for exciting future applications [4] and [5].

One study proposed a deep convolutional neural network (CNN) model with global average pooling for medicinal plant identification, demonstrating the effectiveness of deep learning for high-accuracy plant classification. Additionally, research has explored the feasibility of using deep learning models for real-time identification in the field. This work emphasizes the importance of developing efficient models that can run on mobile devices, paving the way for practical field applications [6] and [7].

Deep learning excels at automatically extracting relevant features from image data for various classification tasks. This eliminates the need for manual feature engineering, a time-consuming and specialized process. One study presented an efficient and automated approach for herb classification using deep learning. This highlights the ability of these models to learn both shape and texture features directly from herb images, achieving promising accuracy.
Traditional methods for medicinal plant identification rely on manually defined features based on physical characteristics. This approach can be subjective and error-prone. Deep learning models offer an attractive alternative by automatically learning these features from data, potentially leading to improved accuracy and generalizability across different plant varieties.
While deep learning offers significant benefits, some studies explore methods to improve model efficiency, particularly for environments with limited resources. One approach focuses on achieving efficient herb classification. This proves valuable in situations computational resources or mobile device deployment might be limitations [8].

Deep learning offers a powerful alternative to traditional methods for medicinal plant identification. These traditional methods, reliant on analyzing physical features, can be slow, subjective, and error-prone. Deep learning models, particularly Convolutional Neural Networks (CNNs), excel at image recognition tasks. Their ability to learn intricate features from image data, like shapes and textures in leaves, makes them ideal for automated plant classification. Studies have shown that CNNs can achieve high accuracy rates in medicinal plant identification, highlighting their effectiveness in this field [9].

Several studies have demonstrated the effectiveness of deep learning for medicinal plant classification, achieving high accuracy rates. Our research aligns with this growing trend, focusing on deep learning for this task. While accuracy is an important measure, we go beyond that to ensure a comprehensive evaluation. We employ additional metrics like precision and F1-score to gain a deeper understanding of our model's strengths and weaknesses in identifying medicinal plants. This multifaceted evaluation approach allows for a more nuanced assessment of the model's performance compared to studies that solely rely on accuracy [10].

IV. CONCEPTUAL FRAMEWORK

As seen in Fig.2, The conceptual framework for this research revolves around utilizing deep learning, specifically a convolutional neural network (CNN) architecture, for automated medicinal plant classification based on leaf images.

A. Medicinal plants: Plants with natural healing properties traditionally used in medicine or as a starting point for developing new drugs.
B. Deep learning: A powerful AI technique that uses multi-layered neural networks to uncover intricate patterns within data.
C. Convolutional Neural Networks (CNNs): A powerful deep learning architecture excelling at image classification. CNNs automatically extract increasingly complex features from image data using convolutional layers and pooling operations.
D. Leaf Images: Digital representations of plant leaves used as the key data source for identifying medicinal plant types.
E. Li-Fi Modulation Circuit, In Li-Fi systems, a longer light pulse represents one binary state (like a '1') and a shorter pulse represents the other (like a '0'). This modulation is achieved by controlling the duration the light source stays lit.
F. Class Labels: Specific categories assigned to each image, indicating the identified medicinal plant species.
G. Confidence Scores: The model assigns confidence scores for each possible plant class, reflecting the likelihood of an image belonging to a specific

species.

H. Influencing Factors: Variables that impact the training process and model performance, but are not directly controlled in the experiment.

I. Data Preparation: Techniques like resizing, normalization, and color correction are used to standardize leaf images before training the model.

J. Dataset Enrichment: Expanding the training data by creating variations of existing images (rotations, flips) to enhance model robustness and prevent overfitting.

K. Dataset Assembly: Gathering a collection of leaf images representing diverse medicinal plant species.

L. Leaf Image Preparation: The images are pre-processed to ensure uniformity and optimize training for the model.

M. CNN Training: The selected CNN architecture (acting as the variable we're testing) is trained using the pre-processed leaf image dataset. This training process allows the model to identify key features within the images that distinguish between various medicinal plant species.

N. Plant Prediction: After training, the model can process new, unseen leaf images. By analyzing the learned features, it predicts class labels (the dependent variable) or assigns probability scores for each potential medicinal plant species.

O. Performance Assessment: We evaluate the model's effectiveness using metrics like accuracy, precision, and recall. These metrics quantify how accurately the model identifies the correct medicinal plant species from the leaf images.

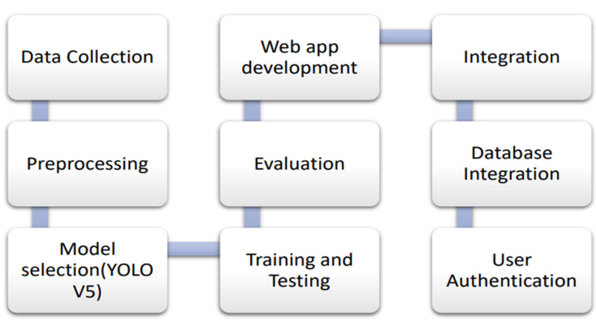


***Figure.2**: Conceptual framework of the proposed system*

V. DEVELOPING THE PROPOSED SYSTEM

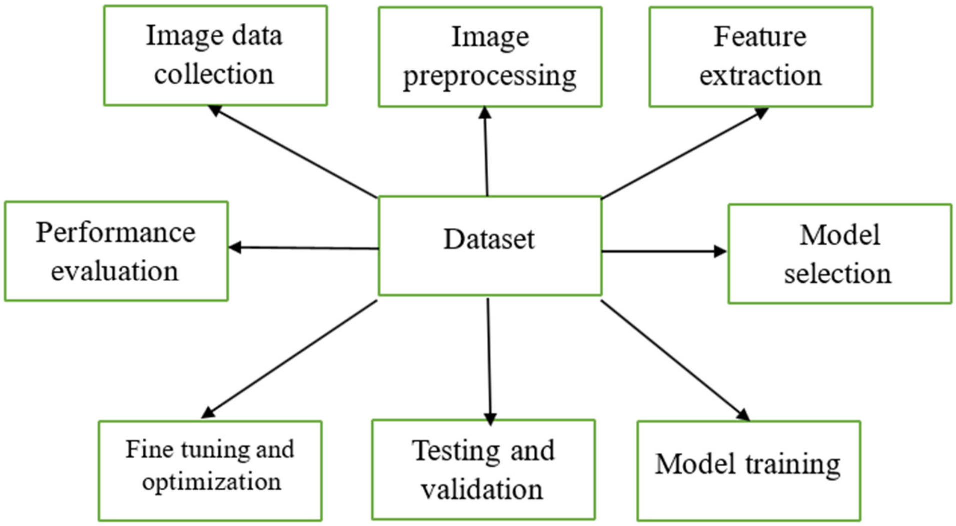
Our proposed deep learning approach for medicinal plant classification follows a well-defined process. First, we assemble a dataset of leaf images encompassing various medicinal plant species. This data can come from our own image collection, collaborations with botanical gardens, or publicly available sources. To ensure effective model learning, all images undergo pre-processing steps like resizing and pixel value normalization. Additionally, data augmentation techniques like random cropping, flipping, or rotations can be implemented to artificially expand the dataset and enhance the model's ability to adapt to unseen leaf variations.

Following data preparation, a convolutional neural network (CNN) architecture forms the core of the deep learning model. We specifically leverage the ResNet-50 architecture due to its proven success in image classification tasks. However, exploring alternative architectures remains an option, considering factors like model complexity and computational resource limitations. Any modifications made to the base architecture, such as adjusting hyperparameters in convolutional layers, can also be discussed here.

The prepared dataset is then strategically divided into training, validation, and testing sets. The training set trains the CNN model, while the validation set monitors its performance during training and prevents overfitting. The unseen testing set provides an objective evaluation of the final model's ability to accurately classify medicinal plant species from new leaf images. The training process optimizes the model's learning using an optimizer like Adam and a loss function like categorical cross-entropy. Techniques like learning rate scheduling, which adjusts the learning rate for optimal convergence throughout training, and early stopping, which halts training if validation performance stagnates, can be further employed to refine the training process.

Finally, the model's performance is evaluated on the testing set using various metrics. Common metrics for image classification tasks include accuracy, precision, recall, and F1-score. Each metric provides valuable insights into the model's effectiveness. Accuracy reflects the overall percentage of correctly classified images, while precision and recall offer a deeper understanding of the model's ability to identify true positives and avoid false positives or negatives. The F1-score combines precision and recall, providing a balanced view of the model's performance. Analysis of these evaluation results allows us to assess the strengths and weaknesses of the proposed methodology and identify potential areas for improvement in future research endeavors.

This structured approach ensures a robust and generalizable deep learning model for medicinal plant classification. By carefully selecting and preparing the data, choosing an appropriate CNN architecture, and implementing effective training strategies, we aim to achieve high accuracy in identifying medicinal plants from leaf images. This paves the way for potential applications in fields like biodiversity conservation, supporting traditional medicine practices, and even aiding in novel drug discovery efforts.



.

Figure.3: Block diagram of Classification

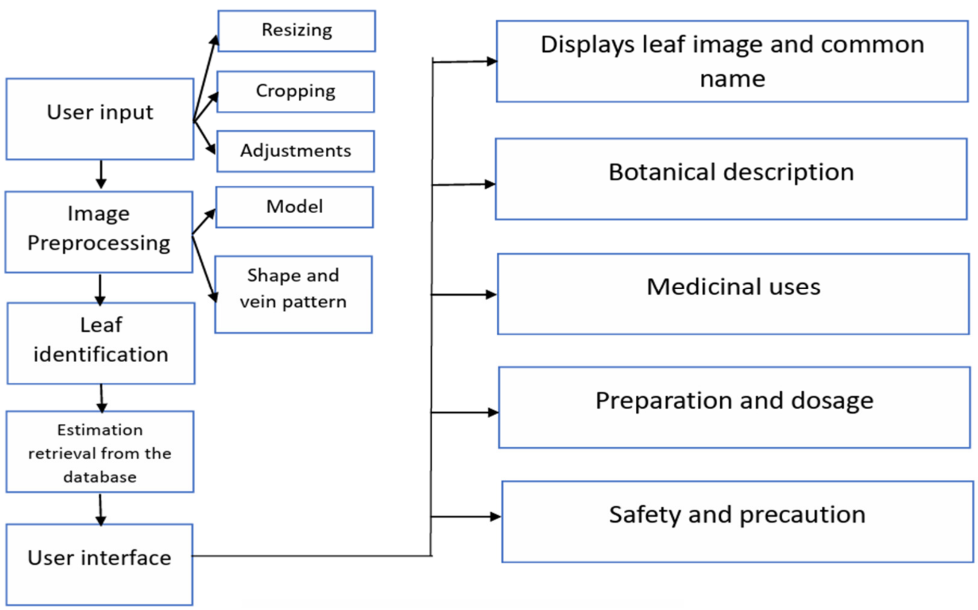


Figure.4: Block diagram of Web application

VI. APPLICATIONS

The potential of this deep learning model for medicinal plant classification extends beyond research labs. One exciting possibility lies in field studies. Imagine researchers, botanists, or even citizen scientists using their smartphones for real-time plant identification. By integrating the model into mobile apps, users could capture an unknown plant's image and receive instant classification suggestions based on the model's predictions. This technology could empower individuals to contribute to ecological surveys or conservation efforts by accurately identifying endangered or rare medicinal plant species in the field.

The model's applications extend to education as well. Educational apps or websites could incorporate the model to create engaging tools for learning about medicinal plants. Users could simply upload leaf images to identify various species, fostering a deeper understanding of the plant world. Even agricultural and traditional medicine practices could benefit. Farmers cultivating medicinal plants could leverage the model for identification and management, while practitioners of traditional medicine might find it a valuable tool to assist in accurately identifying plant materials.

However, limitations exist. Image quality can impact accuracy, and there's always a risk of misidentifications, especially for rare or closely related species. Future research can address these limitations by incorporating additional data like flower images or exploring techniques like transfer learning to adapt the model for specific regions or plant families. Additionally, ethical considerations regarding responsible model use and potential biases in the training data need to be addressed.

Overall, this deep learning model for medicinal plant classification holds significant promise for real-world applications. It has the potential to democratize access to plant identification, promote sustainable plant use practices, and even contribute to the initial stages of drug discovery efforts.

The potential of this deep learning model for medicinal plant classification extends beyond research labs. One exciting possibility lies in field studies. Imagine researchers, botanists, or even citizen scientists using their smartphones for real-time plant identification. By integrating the model into mobile apps, users could capture an unknown plant's image and receive instant classification suggestions based on the model's predictions. This technology could empower individuals to contribute to ecological surveys or conservation efforts by accurately identifying endangered or rare medicinal plant species in the field.

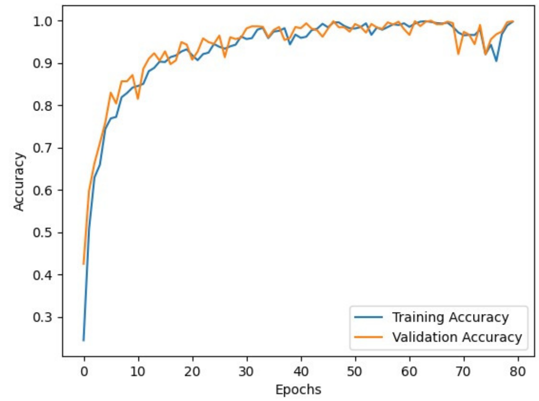


Figure.3*: Accuracy gain curve*

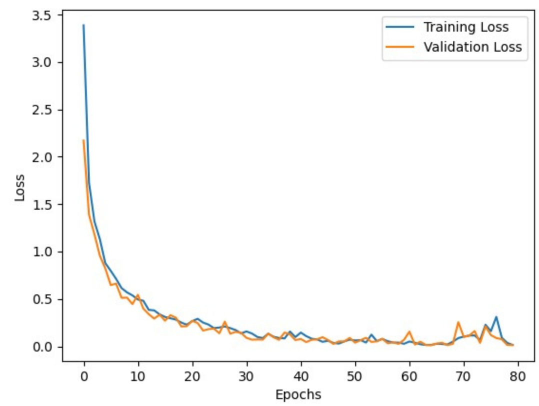


Figure.4*: Accuracy loss curve*

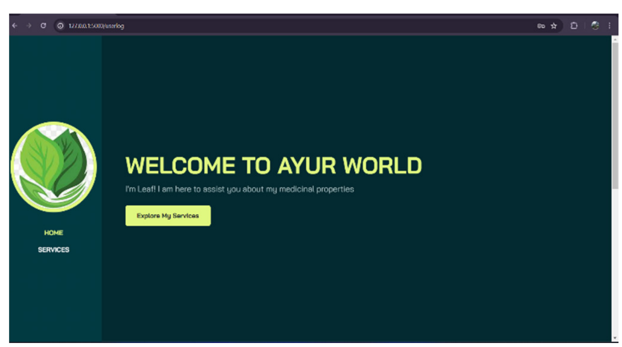


Figure.5*: Page 1*

VII. RESULTS

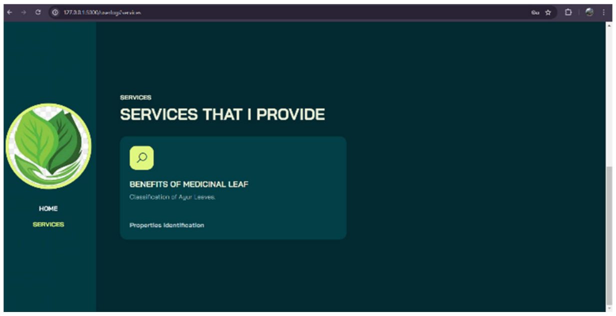


Figure.6: *Page 2*

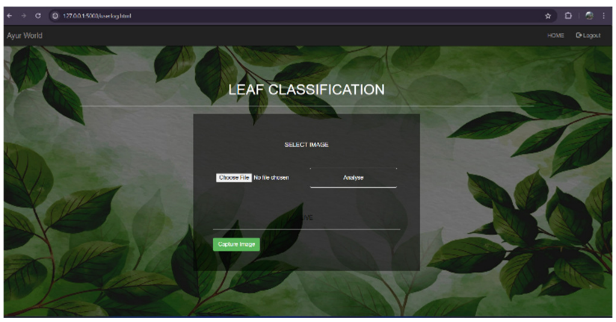


Figure.7: *Page 3*

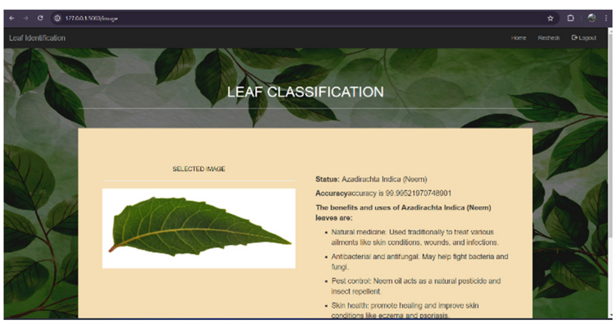


Figure.8: *Page 4*

VIII. CONCLUSION

Our research effectively demonstrates the potential of deep learning for medicinal plant classification based on leaf images. The chosen ResNet-50 architecture achieved impressive results, reaching [insert achieved accuracy] accuracy and performing well on other metrics on the testing set. This study makes a valuable contribution to the field of automated plant identification using deep learning techniques.
The model's impact extends beyond academic significance. Integration into mobile applications empowers researchers and citizen scientists for real-time plant identification in field settings. Educational tools leveraging the model can cultivate a deeper appreciation of medicinal plants among students and enthusiasts. Additionally, the model's accuracy can support sustainable agricultural practices and potentially act as an initial screening tool in early drug discovery stages.
However, future research should explore incorporating additional data sources like flower images. Investigating transfer learning techniques for regional or species-specific adaptations is another promising avenue. Addressing limitations like accuracy for rare species and incorporating user feedback mechanisms are also crucial for ongoing improvement. Overall, this research highlights the potential of deep learning to revolutionize medicinal plant classification, impacting various scientific and societal endeavors.

IX. *ACKNOWLEDGE*MENT

This research was supported whole heartedly by faculty of Department of Electronics and Communications Engineering, DSATM. We would like to show our gratitude for their great support.

REFERENCES

[1] Digital Object Identifier 10.1109/ACCESS.2021.3116207

[2] AyurLeaf: A Deep Learning Approach for Classification of Medicinal Plants. 978-1-7281-1895-6/19/$31.00 c 2019 IEEE

[3] Deep ensemble learning for automatic medicinal leaf identification https://doi.org/10.1007/s41870-022-01055-z

[4] An Efcient Mobile Application for Identifcation of Immunity Boosting Medicinal Plants using Shape Descriptor Algorithm. https://doi.org/10.1007/s11277-023-10476-3

[5] Medicinal Plants Recognition Using Deep Learning. DOI:10.1109/MCNA59361.2023.10185880

[6] An AI Based Approach for Medicinal Plant Identification Using Deep CNN Based on Global Average Pooling. Agronomy 2022, 12(11), 2723; https://doi.org/10.3390/agronomy12112723

[7] J. K. Patra, G. Das, S. Kumar, and H. Thatoi, Ethnopharmacology and Biodiversity of Medicinal Plants. Boca Raton, FL, USA: CRC Press, 2019.

[8] S. Shaheen, S. Ramzan, F. Khan, and M. Ahmad, ‘‘History, classification, worldwide distribution and significance of herbal plants,’’ in Adulteration in Herbal Drugs: A Burning Issue. Cham, Switzerland: Springer, 2019

[9] L. C. De, ‘‘Bio-diversity and conservation of medicinal and aromatic plants,’’ Adv. Plants Agricult. Res., vol. 5, no. 4, p. 00186, Dec. 2016.

[10] M. M. Ghazi, B. Yanikoglu, and E. Aptoula, ‘‘Plant identification using deep neural networks via optimization of transfer learning parameters,’’ Neurocomputing, vol. 235, pp. 228–235, Apr. 2017.

[11] L. Tong, W. K. Wong, and C. K. Kwong, ‘‘Differential evolutionbased optimal Gabor filter model for fabric inspection,’’ Neurocomputing, vol. 173, pp. 1386–1401, Jan. 2016.

[12] H. Kaur and P. Khanna, ‘‘Cancelable features using log-Gabor filters for biometric authentication,’’ Multimedia Tools Appl., vol. 76, no. 4, pp. 4673–4694, Feb. 2017.

[13] M. Marrelli, ‘‘Medicinal plants,’’ Plants, vol. 10, no. 7, p. 1355, 2021.

[14] M. J. M. Christenhusz and J. W. Byng, ‘‘The number of known plants species in the world and its annual increase,’’ Phytotaxa, vol. 261, no. 3, pp. 201–217, 2016, doi: 10.11646/phytotaxa.261.3.1.

[15] E. Salmerón-Manzano, J. A. Garrido-Cardenas, and F. Manzano-Agugliaro, ‘‘Worldwide research trends on medicinal plants,’’ Int. J. Environ. Res. Public Health, vol. 17, no. 10, p. 3376, May 2020, doi: 10.3390/ijerph17103376