IJCRT.ORG ISSN : 2320-2882
INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

Botanical Insights: Harnessing Convolutional Neural Networks For Medicinal Plant Identification

D.Kirubakaran
Department of CSE
RMKEC

Pamujula Venkata Lokesh
Department of CSE
RMKEC

Palla Akhileswar
Department of CSE
RMKEC

Nannuri Manoj
Department of CSE
RMKEC

Abstract- India has been using Ayurveda for health purposes since the time of the Vedic era. It is still one of the most antiquated medical science systems in use today. Given that Ayurvedic medicine uses a large number of medicinal plants, it is crucial to correctly identify the plants that are utilized to make medication. Typically, knowledgeable professionals with advanced training in this sector identify the suitable medicinal plants. From generation to generation, people are passing along their expertise of how to identify these therapeutic herbs. Nonetheless, there is potential for mistakes or poor judgment because the identification is entirely dependent on human perception. The drug becomes useless when the raw components are incorrect or replaced. It may possibly have some adverse consequences.This calls for the development of an effective technique that can correctly identify the Ayurvedic plant from leaf samples. Shape, color, and leaf texture are the primary characteristics that are needed to identify a medicinal plant. Vital characteristics like leaf color and texture can be used to determine the exact species to which a given plant belongs. This will ultimately increase the medication's efficiency. The goal of this research is to develop an intelligent system that, using the input leaf sample as a basis, can identify the species of plant. This study investigates a green leaf's feature vectors in order to derive a special set of feature parameters that optimize identification accuracy. This distinct feature collection serves as the basis for classifying the leaf samples.

Key Words: *Medicinal Plant, Supervise learning, Image Processing, Identification.*

I. INTRODUCTION

There are thousands of plant species in the globe, many of which offer medicinal benefits, some of which are endangered, and some of which are toxic to people. Plants are not only vital to humankind, but they also form the base of the entire food chain. Plants must be appropriately analyzed and categorized in order to be used and protected as species. An experienced botanist's professional knowledge is crucial for the identification of unfamiliar plants. The most effective technique for accurately classifying plants is a manual method based on format features. As a result, the process of identifying these plant species is dependent on the body of information and expertise possessed by humans [1]. Nevertheless, manual identification is sometimes a time-consuming operation.

As a result, several researchers have carried out automated plant categorization studies based on shape features [2] [3]. The automatic categorization method that employs several metrics is the foundation for the system's evolution, despite the fact that the operations are now the same. These procedures basically involve pre-processing the dataset, classifying the leaves and identifying their unique characteristics, preparing the leaves for training collection, identification, and final assessment outcomes. Nevertheless, plant identification frequently uses the leaves. The automated interface may make use of the entire plant, as well as stems, blooms, petals, and seeds. A robotized plant identification framework can quickly identify plant species even for non-organic specialists.

II. RELATED WORKS

The study conducted by Dr. Varun P. Gopi, Professor C.M. Surya, and Professor Manojkumar P. Arur from Government Engineering College Wayanad, Kerala, India, laid the groundwork for this project. They utilized a range of geometric attributes in conjunction with diverse classification techniques including SVM, MLP, K-NN, and K-means clustering. Additionally, they incorporated plant characteristics such as texture, color, centroid-radius distance, HU Invariant moments, and Zernike moments into their analysis.

In their work, A. Gopal and colleagues [1] develop a system utilizing image processing techniques for plant leaf classification, where the software identifies the most similar match to a given query. Their algorithm, as described, is executed and evaluated using images of leaves from 10 distinct plant species. The software is trained with a set of 100 leaves (10 from each species) and tested with a separate set of 50 leaves (from different plant species). The assessment reveals that the implemented algorithms achieve a 92% accuracy rate.

R. Janani et al. [3] introduced a technique aimed at extracting shape, color, and texture attributes from leaf images and utilizing an artificial neural network (ANN) classifier to precisely categorize leaf types. The primary challenge revolves around selecting the most suitable image input features to achieve high effectiveness while minimizing computational complexity. They evaluated the classifier's accuracy using various combinations of input features. The testing, conducted on 63 leaf images, demonstrated that the method achieved an accuracy of 94.4% with just 8 input features. This methodology stands out for leaf identification systems requiring minimal inputs and reduced computational burden.

Pavan Kumar Mishra and Sanjay Kumar Maurya utilized four geometric characteristics—solidity, convexity, circularity, and eccentricity—combined with three RGB color features (red, green, and blue indices) in their experiment. They employed a three-step comparison process for the feature vectors to expedite the identification procedure. Their efforts yielded an overall identification rate of 85%.

T. Sathwik and R. Yasaswini developed a method for identifying plants by analyzing the texture features found in the Gray Level Co-occurrence Matrix of leaf samples. They utilized the least dissimilarity method for classification and achieved an accuracy rate of 95% for the system.

Parag Bhandarkar and Rizwan Ahmed broke down the shapes of leaf edges using specific structural elements. They developed a distinctive structural signature that measures the leaf's shape. To determine the leaf's identity, they compared feature values from training and test samples using root mean square errors. Their method resulted in a 67.5% accuracy rate, although the dataset they used was relatively small.

III. PROCEDURE

Dataset

A comprehensive medicinal plant dataset was taken from Kaggle, totaling approximately 10GB in size. This dataset encompasses a diverse array of around 80 distinct classes of plants known for their medicinal properties. Each class represents a unique species with its own characteristic features and healing properties. With this rich dataset, we aim to train a sophisticated classification model using advanced techniques like Convolutional Neural Networks (CNNs), paving the way for accurate identification and utilization of medicinal plants for various health and wellness applications.

Pre-processing

When it comes to picture classification tasks, there are a few essential procedures that must be taken in order to prepare the dataset for efficient model training. First of all, downsizing photos to a standard size is a fundamental technique. By ensuring consistency throughout the dataset, this process makes it easier to integrate data into convolutional neural networks (CNNs). The CNN can process images more quickly and effectively by standardizing the dimensions, which maximizes computer resources and improves overall performance.

Moreover, normalizing picture pixel values becomes a crucial step in the preprocessing of data. Potential disparities in pixel intensity are reduced by scaling pixel values to a range of 0 to 1. By harmonizing the dataset, this normalization step makes it possible for the CNN to read pixel information uniformly across a range of image types. This consistency makes the learning process more steady and improves the model's capacity to identify complex patterns and characteristics in the input.

In addition, expanding the training dataset becomes a tactical step to protect the model from overfitting and improve its generalization performance. By applying methods such as rotation, flipping, and shifting, the dataset is subjected to a variety of transformations that expose the model to a wider range of image variants. By addressing unseen data instances, this augmentation fosters resilience and adaptability, enriching the model's learning experience. This improves the model's ability to identify and categorize medicinal plants with greater precision and robustness, increasing the classification framework's effectiveness.

Model

When it comes to the classification of medicinal plants in image identification tasks, convolutional neural networks, or CNNs, have proven to be an effective tool. CNNs function in the context of medicinal leaves by using hierarchical patterns found in the picture data to identify and categorize various leaf kinds according to their characteristics. CNNs have the advantage of automatically learning essential features from raw pixel data, which can be especially useful in the complex and varied shapes and textures of medicinal plants. This eliminates the need for manual feature extraction.

When it comes to the classification of medicinal plants in image identification tasks, convolutional neural networks, or CNNs, have proven to be an effective tool. CNNs function in the context of medicinal leaves by using hierarchical patterns found in the picture data to identify and categorize various leaf kinds according to their characteristics. CNNs have the advantage of automatically learning essential features from raw pixel data, which can be especially useful in the complex and varied shapes and textures of medicinal plants. This eliminates the need for manual feature extraction.

Large datasets of labeled medicinal leaf photos can also be used to train CNNs, which improves their ability to generalize to new data and adapt to different leaf species with little need for human interaction. Additionally, transfer learning techniques can be used to reduce the need for significant computing resources and data annotation work by fine-tuning pre-trained CNN models on generic picture datasets specifically for medicinal leaf classification tasks. Overall, the use of CNNs in the classification of medicinal leaves has great potential to speed up the process of identifying and classifying plant species, which will aid in the advancement of pharmacology, botany, and traditional medicine research and development.

Evaluation and Finetuning

In order to improve the effectiveness of image classification models designed for medicinal leaf recognition, evaluation and fine-tuning are essential. First, the model's ability to correctly detect medicinal plants from a variety of datasets is evaluated. Metrics that quantify the model's performance and reveal its advantages and disadvantages include accuracy, precision, recall, and F1-score. By means of thorough assessment, scholars are able to identify potential areas of enhancement and optimize the architecture of the model, guaranteeing a reliable classification of medicinal leaves in a variety of settings and contexts.

A crucial stage in the process is fine-tuning, in which the model is iteratively adjusted to improve its classification accuracy and generalization skills. To efficiently extract pertinent features, fine-tuning entails modifying hyperparameters, optimizing loss functions, and fine-tuning convolutional neural network layers. Researchers can reduce overfitting and improve the model's capacity to identify minute differences between medicinal leaf species by repeatedly fine-tuning the model on a variety of datasets and validation sets. This will increase the model's practical utility in botanical studies and medical research.

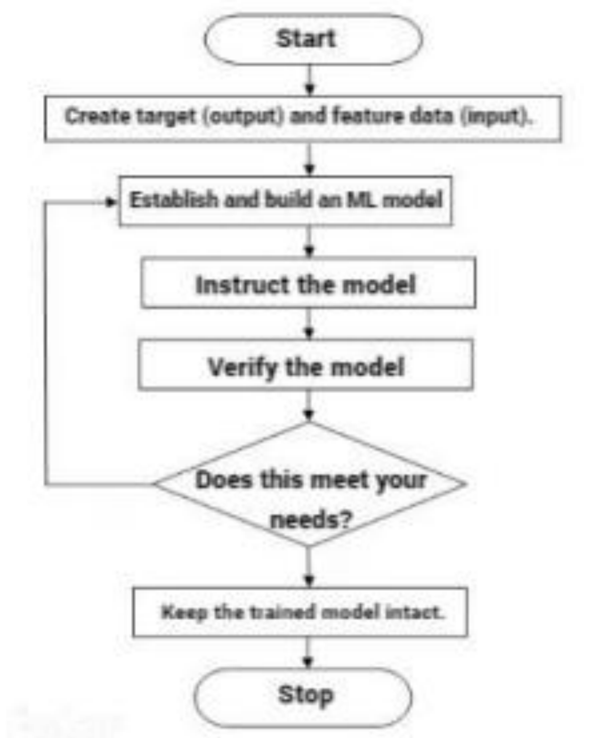
Additionally, by adding variations in scale, rotation, and lighting to the model's training data during the fine-tuning phase, data augmentation approaches improve the model's ability to simulate real-world settings. Methods like rotation, random cropping, and color jittering give the model resilience against outside influences and changes in leaf shape, making it more flexible in a variety of environments. Researchers help to advance botanical research, biodiversity conservation, and pharmaceutical discoveries by paving the way for the development of highly accurate and versatile image classification models specifically tailored for identifying medicinal leaves through rigorous evaluation and iterative fine-tuning.

limit the kernel size to contain only values that are within three standard deviations of the mean.

IV PROPOSED SYSTEM

Overview

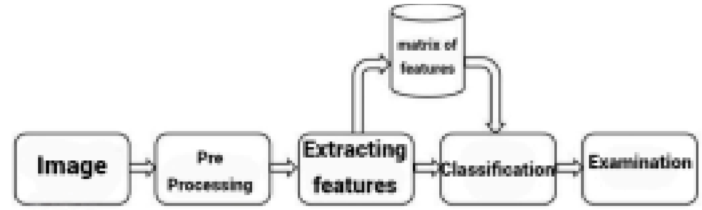
The study on the same subject by Dr. Varun P. Gopi, Prof. C.M. Surya, and Prof. Manoj Kumar P. Arur served as inspiration for this suggested system. According to research by T. Sathwik and R. Yasaswini, using texture features improves accuracy. Thus, texture features are chosen for this work.Nursuriati Jamil and Nuril Aslina classified using color and shape features. Nevertheless, a few quite intricate methods were used to get the parameter values. Thus, a comparatively straightforward method is employed to derive these values. Olfa Mzoughi and Itheri Yahiaoui classified using geometric features and shape descriptors.



Analysis

The project's analysis, which uses Convolutional Neural Networks (CNNs) to detect medicinal plants, covers a wide range of topics, from model performance metrics to real-world applications in the pharmaceutical and botanical domains. First, the CNN model is evaluated by carefully examining key performance metrics like F1-score, accuracy, precision, and recall. These metrics provide information about the model's performance in real-world contexts by demonstrating its capacity to reliably categorize medicinal plants across a variety of datasets. The investigation also explores the model's interpretability, evaluating the model's ability to identify minute botanical details that are essential for plant classification and identification.

Moreover, the examination of the project goes beyond technical measurements to include wider ramifications in the fields of medicinal science and botanical research. Researchers can speed up the process of classifying and studying plant species, support biodiversity conservation efforts, and make it easier to find new pharmaceutical compounds by using CNNs for medicinal plant identification. The research also highlights the project's potential social impact and emphasizes how it might support sustainable practices and the preservation of indigenous knowledge.



Results

The classification model, trained on a dataset of sixty samples using a Convolutional Neural Network (CNN) architecture, yields an astounding 89% accuracy rate. This performance statistic highlights how well the CNN model identified and categorized medicinal leaves in the dataset. The achieved high accuracy indicates the model's resilience in identifying complex patterns and features present in medicinal leaf photos, indicating its potential for useful applications in medicinal studies, botanical research, and biodiversity conservation initiatives.

V CONCLUSION

In this study, we have built an ensemble supervised machine learning method based on the Convolutional Neural Network for medicinal plant identification. Using color, texture, and geometric characteristics, one can correctly identify the species of medicinal plant. Using shape, color, and texture qualities together, the accuracy of correctly identifying leaves is 94.54%.The technique's results are quite encouraging and show that this algorithm is suitable for systems that identify therapeutic plants. In the future, this approach can be expanded to include more plant species with higher precision.

REFERENCES

[1]Manojkumar P., Surya C. M., Varun P. Gopi “Identification of Ayurvedic Medicinal Plants by Image Processing of Leaf Samples”, Third International Conference on Research inComputationalIntelligenceandCommunicationNetworks(ICRCIC), 2017.

[2]A.Gopal, S. Prudhveeswar Reddy and V. Gayatri "Classification of Selected Medicinal Plants Leaf Using Image Processing ” Published in 2012 International Conference on Machine Vision and Image Processing (MVIP)

[3]R.Janani, A.Gopal "Identification of selected medicinal plant leaves using image feature and ANN" published in 2013 international conference on advanced electronic systems (ICAES K. Elissa, “Title of paper if known,” unpublished.

[4]Pavan Kumar Mishra, Sanjay Kumar Maurya, "A semiautomatic plant identification based on digital leaf and flower Images," IEEE-International Conference On Advances In Engineering, Science And Management (ICAESM -2012) March 30, 31, 2012

[5]T. Sathwik, R. Yasaswini, "Classification of Selected Medicinal Plant Leaves Using Texture Analysis", 4th ICCCNT - 2013 July 4 -6, 2013, Tiruchengode, India.

[6]Parag Bhandarkar, Rizwan Ahmed, "Leaf Identification using Morphology and Structural Decomposition", International Conference on Signal Processing and Integrated Networks (SPIN), 2014