e-ISSN:2582-7219

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

Volume 7, Issue 7, July 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.521

6381 907 438 6381 907 438 ijmrset@gmail.com www.ijmrset.com

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 7, July 2024 |

| DOI:10.15680/IJMRSET.2024.0707180 |

Medicinal Leaf Identification Using SVM

Sravanthi Kalal , Veeranna S Dandgund

Assistant Professor, Department of MCA AMC Engineering College, Bengaluru, India

Department of MCA AMC Engineering College, Bengaluru, India

ABSTRACT: There are thousands of plant species in the globe, and many of them have therapeutic uses. The majority of the applications for medicinal plants are in herbal, ayurveda, and folk medicine production. Herbal plants are those that can be utilized as natural alternatives to treat illnesses. There are various methods to identify a variety of herbal plants based on leaves. The form, color, and texture of a leaf are the primary characteristics needed to identify a medicinal plant. The need for mass production has made it crucial to automatically identify these plants. The existing image processing technique should be resilient to changes in illumination intensity. As a result, a new technique requires the identification of certain therapeutic plants. Benefits can be utilized to identify herbal plants in order to avoid contamination and maintain high standards of product performance and safety in particular. In this project, a support vector machine method for medicinal plant identification has been used.

KEYWORDS*:* SVM, Ayurveda, Herbal, therapeutic plants, Medicinal

I. INTRODUCTION

India's Ayurvedic medicine is an age-old practise that dates back to the Vedic era, which was about 5000 years ago. The primary ingredients in ayurveda remedies are plant leaves, along with other plant parts like bark, roots, and so forth. It has been discovered that over 8000 plants with Indian origins have therapeutic properties.

Hindi Name	English Name	Botanical Name	Uses
Adusa/ Vasaka	Malabar Nut	Adhatoda vasicaNees	Cough, Asthma, Bronchitis
Ananas	Pineapple	Ananas comosus	Sore Throat, Diabetes, Heart Disease, Obesity
Babool	Indian Gum	Acacia arabica Willd	Oral Care, Bleeding Gums, Wounds

TABLE I : MEDICINAL NAMES

II. RELATED WORK

A.Gopal et.al [1] implement a system using image processing with images of the plant leaves as a basis of classification. The software returns the closest match to the query. The proposed algorithm is implemented and the efficiency of the system is found by testing it on 10 different plant species. The software is trained with100 (10 number of each plant species) leaves and tested with 50 (tested with different plant species) leaves. The efficiency of the implementation of the proposed algorithms is found to be 92%.

Umme Habiba et.al [2] In this paper, for automatically classifying medicinal plants, they present a Multichannel Modified Local Gradient Pattern (MCMLGP), a new texture-based feature descriptor that uses different channels of

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 7, July 2024 |

| DOI:10.15680/IJMRSET.2024.0707180 |

color images for extracting more significant features to improve the performance of classification. Author have trained their proposed approach using SVM classifier with various kernels such as linear, polynomial and HI. In addition, used different feature descriptors for comparative experimental analysis with MCMLGP by conducting the rigorous experiment on our own medicinal plants dataset. The proposed approach gain higher accuracy (96.11%) than other techniques, and significantly valuable for exploration and evolution of medicinal plants classification.

R. Janani et.al[3] have proposed a method for the extraction of shape, color and texture features from leaf images and training an artificial neural network (ANN) classifier to identify the exact leaf class. The key issue lies in the selection of proper image input feature to attend high efficiency with less computational
S. complexity. They tested the accuracy of network with different combination of input feature. The test result on 63 leaf images reviles that this method gives 94.4% accuracy with a minimum of 8 input features. This approach is more prominent for leaf identification system that have minimum input and demand less computational time.

Vijayashree.T et.al [4] has created database with 127 herbal leaves. For creating a database 11 texture parameters are taken into account. The parameters are Sum of Variance, Inverse Difference Moment, Aspect ratio, Correlation, Sum Entropy, Mean, and Sum Average. Gray level co-occurrence matrix (GLCM) is used for determining the parameters like entropy,

III. METHODOLOGY

SVM ALGORITHM

Support Vector Machine (SVM) is a powerful machine learning algorithm used for linear or nonlinear classification, regression, and even outlier detection tasks. The main objective of the SVM algorithm is to find the optimal hyperplane in an N-dimensional space that can separate the data points in different classes in the feature space. The hyperplane tries that the margin between the closest points of different classes should be as maximum as possible.

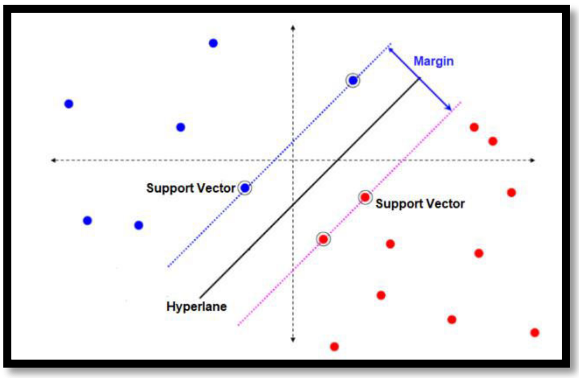


Figure 1.SVM

1) Support Vector Machine Terminology-

1. Hyperplane: Hyperplane is the decision boundary that is used to separate the data points of different classes in a feature space.
2. Support Vectors: Support vectors are the closest data points to the hyperplane, which makes a critical role in deciding the hyperplane and margin.
3. Margin: Margin is the distance between the support vector and hyperplane. The main objective of the support vector machine algorithm is to maximize the margin. The wider margin indicates better classification performance.
4. Kernel: Kernel is the mathematical function, which is used in SVM to map the original input data points into high-dimensional feature spaces, so, that the hyperplane can be easily found out even if the data points are not linearly separable in the original input space. Some of the common kernel functions are linear, polynomial, radial basis function(RBF), and sigmoid.

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 7, July 2024 |

| DOI:10.15680/IJMRSET.2024.0707180 |

5. Hard Margin: The maximum-margin hyperplane or the hard margin hyperplane is a hyperplane that properly separates the data points of different categories without any misclassifications.
6. Soft Margin: When the data is not perfectly separable or contains outliers, SVM permits a soft margin technique. Each data point has a slack variable introduced by the soft-margin SVM formulation, which softens the strict margin requirement and permits certain misclassifications or violations. It discovers a compromise between increasing the margin and reducing violations.
7. C: Margin maximization and misclassification fines are balanced by the regularization parameter C in SVM. The penalty for going over the margin or misclassifying data items is decided by it. A stricter penalty is imposed with a greater value of C, which results in a smaller margin and perhaps fewer misclassifications.

(2) Support Vector Machines – Implementation in Python

In Python, an SVM classifier can be developed using the sklearn library. The SVM algorithm steps include the following:

Step 1: Load the important libraries

- import pandas as pd
- import numpy as np
- import sklearn
- from sklearn import svm
- from sklearn.model_selection import train_test_split
- from sklearn import metrics

Step 2: Import dataset and extract the X variables and Y separately.

```
df = pd.read_csv("mydataset.csv")
X=df.loc[:,['Var_X1','Var_X2','Var_X3','Var_X4']]
Y = df[['Var_Y']]
```

Step 3: Divide the dataset into train and test

```
X_train, X_test, y_train, y_test = train_test_split(X, Y, test_size = 0.3, random_state=123)
```

Step 4: Initializing the SVM classifier model

```
svm_clf = svm.SVC(kernel = 'linear')
```

Step 5: Fitting the SVM classifier model

```
svm_clf.fit(X_train, y_train)
```

Step 6: Coming up with predictions

```
y_pred_test = svm_clf.predict(X_test)
```

Step 7: Evaluating model's performance

```
metrics.accuracy(y_test, y_pred_test)
metrics.precision(y_test, y_pred_test)
metrics.recall(y_test, y_pred_test)
```

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 7, July 2024 |

| DOI:10.15680/IJMRSET.2024.0707180 |

IV. SYSTEM ARCHITECTURE

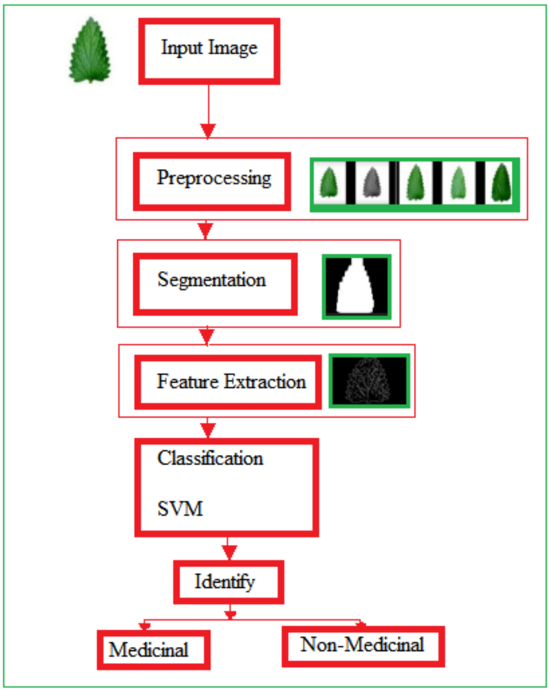


Figure 2: System Architecture

The architecture mentioned above receives the input image and proceeds with preprocessing; preprocessing data is a step in the preparation of data. It is the most important and initial phase in building our model. Prior to training and testing the data with several prototypes, the data needs to be preprocessed, sorted, and exposed in accordance with specific limitations. By taking this step, the data will get smaller and the unwanted data will be removed. Images captured with a camera have shadows on them. If the image's shadow is not eliminated, it will affect the calculations as a whole. Therefore, in order to remove shadows, the photos must first be converted to the HSV format and then split into their respective color channels, with only the saturation channel remaining. The generated image is subjected to a median blur filter with a window size of 25 in order to reduce noise. In addition, a thresholding process is used to turn an image into a binary image. The Otsu thresholding approach is applied in this case. After that, an opening procedure is performed on the photographs.

This illustrates how erosion functions following dilation. White pixels are impacted by erosion in a smaller size and by dilation in a larger size. This stage is essential for eliminating the image's numerous tiny noisy pixels, which are artifacts of the thresholding process. Additionally, the SVM classifier is utilized for classification after installing all required libraries.

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 7, July 2024 |

| DOI:10.15680/IJMRSET.2024.0707180 |

V. RESULTS AND DISCUSSIONS

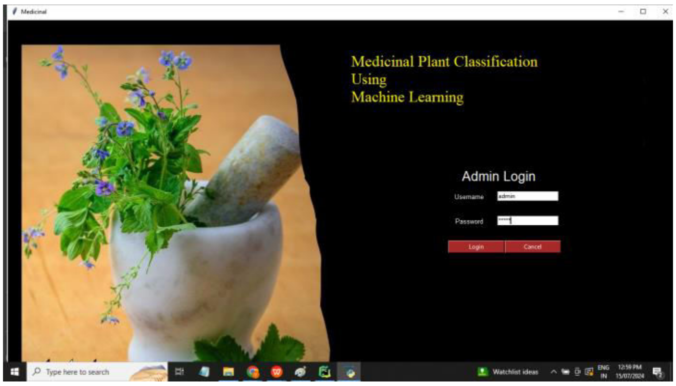


Figure 3: Main screen

This is a main screen of our project where we can interact with further operations through this.

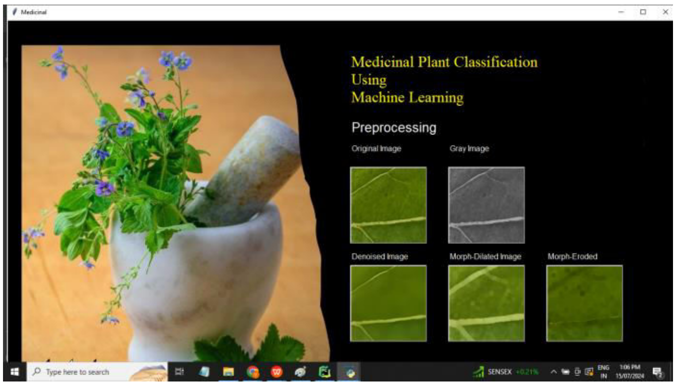


Figure 4: Pre-processing

The actions done to format images prior to their usage in model training and inference are known as image preprocessing. This covers resizing, aligning, and color adjustments, among other things. Preprocessing images can help speed up model inference and reduce training time.

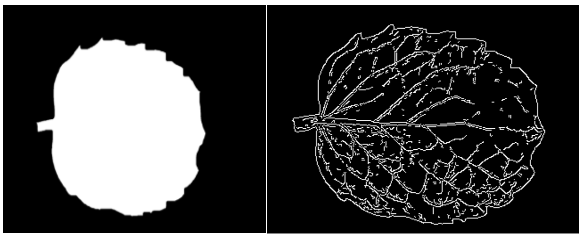


Figure 5. (a) Segmentation(b) Feature Extraction

Image segmentation involves converting an image into a collection of regions of pixels that are represented by a mask or a labeled image. By dividing an image into segments, can process only the important segments of the image instead of processing the entire image.

1. Thresholding

Is a very popular segmentation technique, used for separating an object considered as a foreground from its background. cv2.THRESH_BINARY_INV method gives max value tillit does not crosses threshold.

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 7, July 2024 |

| DOI:10.15680/IJMRSET.2024.0707180 |

In Otsu Thresholding, a value of the threshold isn't chosen but is determined automatically. A bimodal image (two distinct image values) is considered. The histogram generated contains two peaks. So, a generic condition would be to choose a threshold value that lies in the middle of both the histogram peak values.

2. Distance Transformation

Distance transformation is a technique used to calculate the distance of each pixel in an image from the nearest non-zero pixel. It is often used in image segmentation and object recognition tasks, as it can help identify the boundaries of objects in an image.

- The distance transform function in OpenCV, cv2.distanceTransform(), takes in a binary image and returns two arrays: the distance image and the label image.
- The distance image contains the distance values of each pixel from the nearest non-zero pixel.
- The label image contains the labels of the nearest non-zero pixels.

The distance transform function also takes in two optional arguments: the distance type and the mask size. The distance type can be specified using constants such as cv2.DIST_L1 (for the Manhattan distance) or cv2.DIST_L2 (for the Euclidean distance). The mask size determines the size of the mask used to calculate the distances, with larger values resulting in more accurate but slower calculations.Feature extraction

Feature extraction is a part of the dimensionality reduction process, in which, an initial set of the raw data is divided and reduced to more manageable groups. The Canny edge detector is an edge detection operator that uses a multi-stage algorithm to detect a wide range of edges inimages. It was developed by John F. Canny in 1986.

The Canny edge detection algorithm is composed of 5 steps:

1. Noise reduction using Gaussian filter
   Gradient calculation along the horizontal and vertical axis
   Non-Maximum suppression of false edges
   Double thresholding for segregating strong and weak edges
   Edge tracking by hysteresis.

	precision	recall	f1-score	support
Medicinal	1.00	1.00	1.00	23
NonMedicinal	1.00	1.00	1.00	5
accuracy			1.00	28
macro avg	1.00	1.00	1.00	28
weighted avg	1.00	1.00	1.00	28

The predicted image is : Medicinal

Figure 6: SVM Algorithm Report

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 7, July 2024 |

| DOI:10.15680/IJMRSET.2024.0707180 |

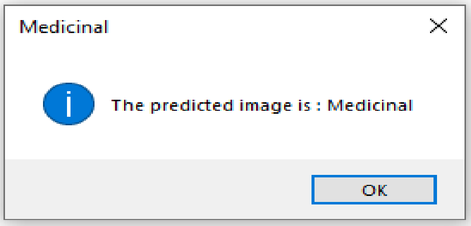
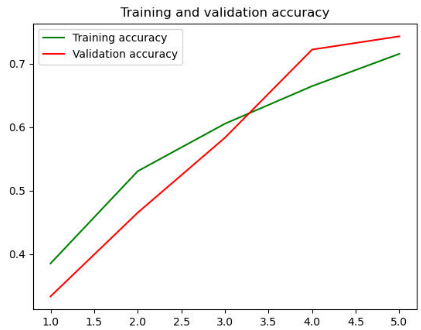


Figure 7: Leaf Identification



Graph 1 : Accuracy graph

VI. CONCLUSION

In this project, implemented a method for identifying medicinal plants using SVM, an ensemble supervised machine learning algorithm, and the well-known multistage feature extraction algorithm Canny Edge Detector, which uses geometrical, color, and texture features to identify the correct species of medicinal plant. The technique's very good results suggest that this algorithm is suitable for medicinal plant identification systems. It was discovered that the model has 80% accuracy. In the future, this study can be expanded with better accuracy to a greater number of plant species.

.

REFERENCES

[1]. 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)

[2]. Umme Habiba, Md. Rasel Howlader, Md. Aminul Islam, Rahat Hossain Faisal and Md. Mostafijur Rahman"Automatic "Medicinal Plants Classification using Multi- channel Modified Local Gradient Pattern with SVM Classifier" published in Joint 2019 8th International Conference on Informatics, Electronics & Vision (ICIEV) & 3rd International Conference on Imaging, Vision & Pattern Recognition (IVPR)

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

[4]. Vijayashree.T, Dr. A. Gopal "Leaf Identification for the Extraction of Medicinal Qualities Using Image Processing Algorithm" published in 2017 International Conference on Intelligent Computing and Control (I2C2)

[5]. D Venkataraman, Mangayarkarasi N "Support Vector Machine Based Classification of Medicinal Plants Using Leaf Features" published in 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 7, July 2024 |

| DOI:10.15680/IJMRSET.2024.0707180 |

[6]. Shitala Prasad, Pankaj P Singh"Medicinal Plant Leaf Information Extraction Using Deep Features" published in Proc. of the 2017 IEEE Region 10 Conference (TENCON), Malaysia, November 5-8, 2017

[7]. Dileep M.R., Pournami P.N."AyurLeaf: A Deep Learning Approach for Classification of Medicinal Plants" published in TENCON 2019 - 2019 IEEE Region 10 Conference(TENCON)

[8]. C.Amuthalingeswaran, Mr. M.Sivakumar, Dr. P.Renuga"IDENTIFICATION OF MEDICINAL PLANT"S AND THEIR USAGE BY USING DEEP LEARNING"
Published in Proceedings of the Third International Conference on Trends in Electronics and Informatics (ICOEI 2019) IEEE Xplore Part Number: CFP19J32-
ART; ISBN: 978-1-5386-9439-8

[9]. Manojkumar P., Surya C. M., Varun P. Gopi "Identification of Ayurvedic Medicinal Plants by Image Processing of Leaf Samples" Published in 2017 Third international conference on research in computational Intelligence and communication network (ICRCICN)

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com