Paper ID: 502

A Study on the Ayurveda Plant Recognition for Remedial Medications Using Image Processing Techniques

K. Anuradha#, D.P.M. Lakshan and R.P.S. Kathriarachchi

Faculty of Computing, General Sir John Kotelawala Defence University, Ratmalana, Sri Lanka
#anuradhakurunayaka@gmail.com

Abstract: Plants are considered an essential part of our ecosystem and Sri Lanka has a long history of using plants as a source of medicines in Ayurveda. In addition to some herbaceous plants serving as a food source, have medicinal values. In the Ayurveda medicinal industry, it is very important to identify the correct herbs that help in the preparation of remedial medicines. The identification of these suitable herbaceous plants is often done by skilled specialists. However the problem is since identification is based on human cognition, it can lead to misjudgment. So it a waste that humankind couldn't use the herbal power of remedial medications. To address this question the paper proposes a simple and effectual methodology for identification of Ayurveda's herbaria, using mobile devices in the android platform by implementing image processing techniques. The main characteristics required to identify a medicinal herb are the shape, color, and texture of the leaf. The color and texture of the leaf cover vital parameters that are unique to a particular plant. Preprocessing, feature extraction, and classification are the three major phases in the suggested methodology. In order to train neural networks, images of herbal plant leaves were captured under the supervision of an Ayurveda doctor. For all the images backgrounds are removed and resized before applying classification techniques. According to the methodology, the leaf images are trained and the result can be shown through the mobile application. The study got 94% of accuracy for the proposed methodology.

Keywords: Ayurveda herbs, Image processing, Feature Extraction

Introduction

Ayurveda is an ancient form of medicine that specialized in India and dates back to the Vedic period, roughly 5,000 years ago. Ayurveda is considered to be the oldest healing science in medicine. Ayurveda in Sanskrit means "The Science of Life". Ayurveda is widely used in its land and around the world, despite being subdued to living abroad for many years and also especially in Sri Lanka. It has still followed by countless people as it is purely natural and has no side effects. It is much applicable from ancient times to this most modern time because of its power to cure chronic

diseases. It inspires the use of herbal and non-toxic plants. It is known that about 8000 plants have medicinal properties. The principles built on this Ayurveda system are always true for all ages and can be easily followed from generation to generation.

The plants around us play a huge role in shaping the ecosystem. Some of the food we can eat is above or below the ground and there are a variety of plants that have been found to be used in medicine called Ayurveda. The main ingredients of Ayurveda medicines are plant leaves, roots, bark, fruits, seeds, etc. In the past, Ayurveda physicians themselves selected herbs and prepared medicines for patients. In those days, people were aware of the remedial values and uses of the medicinal plants found in their surrounding environment. But many people in this new generation do not have any

knowledge about these herbs. According to Ayurveda, every plant on earth has some medicinal value, so it is important to protect the plant and identify its medicinal values. The existence of our planet's ecology is impossible without plants. So, herbals are a great advantage that gains from plants. On the other hand, studies have proved that consuming so much of allopathic medicines may lead to side effects as it carries out many chemical reactions within the body. A general fact about western medicine is that once it is taken, western medicine requires taking another medicine to cure the side effects which has happened due to the previous medicine. In general, the process of consuming medicines will not end. Allopathic treatments are inevitable to Treat the Symptoms of a disease whereas Ayurveda treats the root of the disease.

These plants are usually collected by tribal people who are not professionally trained in the task of identifying the correct plants. Today the plants are collected by women and children from the forest areas. They do not have professional training to identify the correct medicinal plants. Production units often get the wrong or substitute medicinal plants. Besides, there is confusion over changes in local names. Some plants come in dry form, which makes manual recognition a more difficult matter. Misuse of medicinal plants makes Ayurveda medicine ineffective. It can also produce an unexpected side effect. For this purpose, creating and maintain a plant database is a critical step towards the protection of the earth's biosphere. There are countless species around the world. To handle such volumes of data, improvement of a brisk and effective system for storing and maintain plant data for later recognition is required.

Everything except the leaf is three-dimensional objects and increases the complexity of the analysis by computer. However, plant leaves are 2D objects and carry sufficient information to identify the plant. The leaves can be easily collected, and images can be taken using expensive digital cameras, cell phones, or document scanners. As the leaves grow, they acquire a specific color, texture, and shape, but these differences are relatively insignificant. Leaf-based plant identification depends on specific descriptions and feature vector extraction. The feature vectors of the training samples are then compared with the feature vectors of the test sample to find the degree of similarity using an appropriate classifier. Pattern recognition tasks performed specifically on leaves are called leaf recognition. Leaf pictures could be classified as "identified" as wel1 as "not identified" by evaluating leaf with images stored in the database. One can send an image or picture of a recognizable leaf to a computer, and the computer can identify the leaf if its data is in the database using a texture extraction method or other image processing methods. The input for any image processing method can be an image, photos, or frames in a video. The output can be another image or set of characteristics of an input image. Pictures represented in visual information are called images.

Background and Motivation

Because of the increased commercialization of the Ayurveda sector, several issues regarding the raw material quality used for their preparation need to be focused. The Ayurveda plants are usually collected by tribal masses that are not professionally trained in the work of identifying the correct plants. Even the manufacturing units, sometimes, receive improper or substituted medicinal plants. Most of these manufacturing units do not have proper quality control mechanisms that can screen these plants. It is difficult to identify a plant species through a photograph because of its complex three-dimensional structure which cannot be captured through cameras, but it is

possible if the leaf can be identified. Luckily most of the leaves are two- dimensional and it is possible to automate the identification of a plant species through its leaf morphology. In order to invent an automatic recognition system of remedial medications, it is necessary that medicinal plants be identified by common people. So the motivation of this study is to take the leverage of technology in identifying and classifying the plants.

Literature Review

A. Problem

By conducting an interview, researches, questionnaires, and case studies from with the researches botanists Ayurveda doctors and medical students, in the sample local areas of Matara and Galle it was proofed that a system for recognition of Ayurveda leaves would be a great help in the arena of improvement of the remedial medications in Sri Lanka with the busy schedules of the lifestyle of the people.

B. Previous Researches

(Dhingra et al., 2018)In this paper, an identification system has been developed using imagery of foliage. A mobile app was also developed to allow a user to take pictures of leaves and upload them to a server. To obtain a potential match, the client performs pre-processing and feature extraction techniques on the image before comparing the information extracted from this image with the information in the database.

(Gwo and Wei, 2013)This article proposes a simple and efficient method for classifying Ayurveda plants using digital image processing and machine vision technology.

The three main stages of the proposed methodology are preprocessing feature extraction, and classification. Many of these methods use a combination of many parameters, such as color, shape, and texture. The proposed method was tested with 208 sample leaflets of 26 different species and in many cases a positive response.

(Venkataraman and Mangayarkarasi, 2016) Identifying herbal remedies for home remedies for the general public will be helpful. The purpose of this project is to provide a free app for everyone and the most common mobile platform, Android, and can be used for personal or research and learning purposes. The detection is done by a texture extraction algorithm called the gray-level co-occurrence matrix (GLCM). A working prototype of the project was made and tested with samples that showed satisfactory results.

(Munisami et al., 2015)A vision-based approach is used to create an automated system that identifies plants and values their values. This article discusses how to create a feature set that is an important step in identifying any species.

(Pramanik et al., 2010)The most reliable solution to overcome the botanical taxonomy gap, which receives considerable attention from both the botanical and computer community, is the identification of automated plant images. The first mobile device was acquired by the BJFU100 dataset, which contains 10,000 images of 100 plant species, providing data pillars for further plant identification studies. The proposed model ResNet26 results show an accuracy of 91.78% in the test set, indicating that the deep environment is a good technology for large-scale classification of the natural environment.

(Mzoughi et al., 2013)This paper has described a working computer vision system that helps identify plant species. When a user takes an isolated leaf on an empty background, the system takes the shape of the leaf and fits it into the green shape of the known species. In just a few seconds, the

system displays the highest matching species with text descriptions and additional images.

(Belhumeur et al., 2008)This study proposes a feature extraction method for green contours, which describes the lines between the centroid and each contour point on an image. A long histogram is made to represent the distal distribution of the leaf contour. Then, a statistical model is used to calculate the fit of the template and the query form.

(Sun et al., 2017) This paper explores the vectors and morphological features of the front and back of a green leave. Scanned images of the front and back of the most widely used Ayurvedic medicinal plant leaves create a database of plant herbal leaves. The leaves are classically based on a unique combination of elements. Up to 99% detection rates have been obtained when testing a wide range of classifiers.

(Kumar et al., 2017)The article has addressed a comprehensive study of disease detection and leaf classification using image processing techniques. In addition to unequal processing times, it includes a specialty in the field of phytopathology. Therefore, image processing has been applied to identify plant diseases. A comprehensive discussion of disease detection and classification performance is presented. The proposed method is based on the local representation of leaf fragments. Semi-based decomposition is defined and is generally used by botanical. A global image query is a collection of partial sub-image queries. Experiments on leaf imagery in the real world.

(Br, 2016)This approach is useful in the classification of MLPs, which helps to identify different plants based on their botanical or genetic characteristics. Since the leaf structure of different crops varies, the replicas of the present work clearly show the specific characteristics of each variety.

Methodology

A. Requirement analysis

In Sri Lanka, there are no standard databases of Ayurveda medicinal plant leaves which is available for conducting the research experiments. Leaf images set of medicinal plants were collected from a private botanical garden. 15 leaves were collected in a random style from 30 different plant species used for remedial medications. The remedial leaves were collected from their natural habitat and they were selected in quite random.

Scanned images are obtained in high resolution (1,200 dpi) and RGB format. Images were rendered in the TIFF format to maintain the original quality of the images. Live leaves were collected here. Captured images were taken with a high-quality camera. And also a few leaf images were taken with cell phones. When capturing images, both still images and videos of herbaceous plants were collected. And then they are converted to images using a video editing software

B. Design and Implementation

The methodology for identifying Ayurveda leaf samples consists of five steps namely, Image acquisition, pre-processing, feature extraction, classification techniques, and testing. The proposed system will focus on identifying herbs especially Ayurveda plants. Image processing techniques will be used to identify the plant species and it would be identified based on leaf image processing.

When a leaf image of a specific plant is fed into the system and the system will pre-process the image to diminish the noise present in it and to get grayscale, binary, and edge for future extraction. In the feature extraction phase, arithmetic means on a color image, the standard deviation on color as well as, Entropy on grayscale image, Solidity, Extent, Eccentricity, and Equivalent

diameter would be calculated. After the feature extraction phase, the leaf factor of the particular leaf would be calculated using a suitable machine learning algorithm like Weka (Weka is a collection of different machine learning algorithms that are written in Java and it is open-source software) And calculations on that different samples of that plant type would consider and the average leaf factor would be calculated which is unique for a specific leaf type and its value would be stored in the database. When a new leaf is fed into the system for recognition, the leaf factor of that particular leaf would be calculated, and it would compare with the leaf factor which is stored in the database and the most matching leaf would be returned as the output.

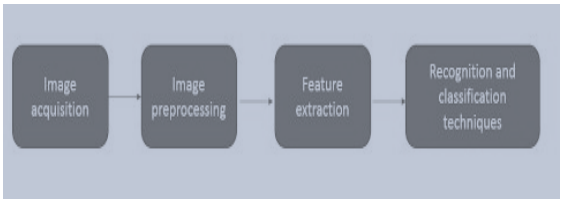

Figure 1. Steps in Image Processing
Source:(Amara et al., n.d.)

A. Image Acquisition

Datasets are collected in such a way that leaves are captured against a white background using a digital camera or through a scanner. The experiments would be carried over the datasets collected.

B. Pre-Processing

The main purpose of pre-processing the data is to enhance the visual appearance of the image and to improve the manipulation of datasets. This is achieved by removing unwanted noise, reconstructing the image, enhancing its quality, etc.

Pre-processing the image is an important step as it increases the probability of getting the desired output in the future steps of image processing. The input image is converted to grayscale and the binary image to keep the pixel values as either 1 or 0, so as the feature extraction operations can be made simpler as well as the image gets stored as lower sized binary images. The input image is smoothened to reduce the noise in the image. Smoothening reduces the number of pixels in the image and it helps in detecting the edges in an image.

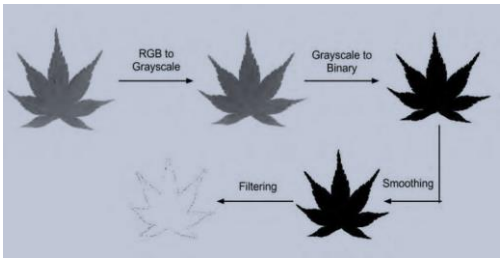

Figure 2. Steps in Pre-processing
Source: ("Preprocessing steps performed on an Acer Palmatum leaf image. | Download Scientific Diagram," n.d.)

C. Feature Extraction

Image processing techniques are used to extract a set of features that characterize or represent the image. The values of the extracted features represent the information in the image and those values are as Mean, Standard Deviation, Solidity, Extent, Eccentricity, and Equivalent diameter.

$$\text{Equivalent diameter} = \sqrt{\frac{4 * area}{\pi}}$$

Convex area specifies the total number of white pixels that are present in the Convex Image. A convex image is a binary image that specifies the smallest convex polygon that contains the region in which all the pixels are filled in within the polygon. Solidity can be calculated as;

$$\text{Solidity} = \frac{Convex\ area}{Original\ area}$$

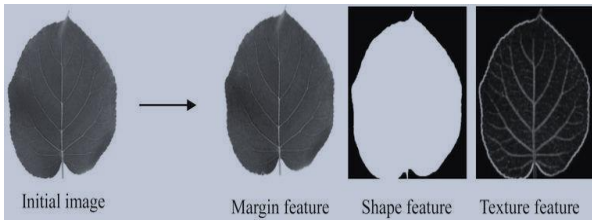

Figure 3. Image texture

Source:("Preprocessing steps of leaf images: a) original image; b) greyscale… | Download Scientific Diagram," n.d.)

Eccentricity is a characteristic of a conic section. It can be calculated as,

$$\text{Eccentricity} = \sqrt{1 - \left(\frac{Minor\ axis\ length^2}{Major\ axis\ length^2}\right)}$$

Extent is the ratio of the total number of pixels in the region to the number of pixels in the border-box. It can be calculated from the below formula.

$$\text{Extent} = \frac{No.\ of\ pixels\ in\ region}{No.\ of\ pixels\ in\ the\ bounding\ box}$$

The color element is an important feature in representing an image. This feature is most important because it does not change rotation, transformation, and scale. The key elements involved in color feature extraction are color space, similarity measurement, and calibration. Color instances are mainly defined by mean and standard deviations.

$$\text{Mean } (\mu) = \sum_{i=0}^{M}\sum_{j=0}^{N}\frac{P(i,j)}{MN}$$

$$\text{Standard Deviation } (\sigma) = \sqrt{\sum_{i=0}^{M}\sum_{j=0}^{N}\frac{(P(I,j)-\mu)^2}{MN}}$$

D. Classification techniques

To classify a leaf according to their species, the values that are extracted from the pre-processing stage are considered and such as Mean, Standard Deviation, Convex hull ratio, Isoperimetric Quotient, Eccentricity, and entropy. After the feature extraction phase, the leaf factor of the particular leaf is calculated using a suitable machine learning algorithm like Weka (Weka is a collection of different machine learning algorithms that are written in Java and it is opensource software)

E. The Overall System Architecture

In the proposed system server-side admin can log in with his credentials and can save images of leaves. From the user side, the user can upload the leaf images and find the details of the Ayurveda herb. The uploaded and server-side images are compared to identify. The development will be done using Android Studio. Any android phone with an internet connection will be able to run the app. and machine learning algorithms like CNN will be used. By using this application software not only benefits the common people, but also it would be very helpful for experts in fields like the cosmetic industry, botanical gardening, and the medical industry as well.

1) CNN Algorithm

CNN is very simple and almost the same as regular Neural Networks which has some specific weights and bias. Everything that we do for learning an ordinary neural network applies to a convolutional neural network. But the change is that CNN takes an assumption that the inputs are images and thus this allows us to make changes to its architecture. A CNN transforms its neurons in a three-dimension (height, width, depth) form

2) Android Studio

Android Studio is a software that offers more features that enhance the productivity of the mobile application when building Android apps, such as A flexible Gradle-based build system. It is a fast and feature-rich emulator.

And it is a unified environment where users can develop for all Android devices.

3) SQL Database

A database in SQL Server is made up of a collection of tables that stores a specific set of structured data. A table in a SQL database contains a collection of rows, also referred to as records or tuples, and columns also referred to as attributes.

Discussion

The paper has demonstrated an approach to classify plants into their appropriate species using images of their leaves. Ayurveda encourages the use of herbal and non-toxic plants for treatment and the creation of different remedial medications and it requires more understanding about plants and its species. A unique combination of geometry, color, and texture has been identified to maximize identification accuracy with image processing techniques. Features can be calculated for different types of herbaceous leaves and stored in the database which are the values of the training results. With obtained results, the study can notice that both the convolutional neural networks transcended the multi-layered perceptions in the process of validating and testing datasets. The study can evaluate, CNN trained with RGB images performs better than the CNN model trained with grayscale images. The testing accuracy can be considered as the highest accuracy obtained for the classification of Ayurveda herb leaves compared with prior carried out studies. The study used a limited set of data for this study with limited variation in scanned leaf images.

Identification of Medicinal herbs can be very helpful for common people to help them make home remedial medicines and for professionals and students a more effective way of research. The project aims at making an application that is available for free for everyone and on a more common mobile platform, Android, and can be used for personal or research and study purposes.

References

Amara, J., Bouaziz, B., Algergawy, A., n.d. A Deep Learningbased Approach for Banana Leaf Diseases Classification 10.

Belhumeur, P.N., Chen, D., Feiner, S., Jacobs, D.W., Kress, W.J.,Ling, H., Lopez, I., Ramamoorthi, R., Sheorey, S., White, S., Zhang, L., 2008. Searching the World's Herbaria: A System for Visual Identification of Plant Species, in Forsyth, D., Torr, P., Zisserman, A. (Eds.), Computer Vision – ECCV 2008. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 116–129. https://doi.org/10.1007/978-3-540-88693-8_9

Br, P., 2016. Ayurvedic Plant Species Recognition using Statistical Parameters on Leaf Images 11, 7.

Dhingra, G., Kumar, V., Joshi, H.D., 2018. Study of digital image processing techniques for leaf disease detection and classification. Multimed. Tools Appl. 77, 19951–20000. https://doi.org/10.1007/s11 042-0 17- 5445-8

Gwo, C.-Y., Wei, C.-H., 2013. Plant Identification through Images: Using Feature Extraction of Key Points on Leaf Contours. Appl. Plant Sci. 1, 1200005. https://doi.org/10.3732/apps.1200005

Kumar, P.M., Surya, C.M., Gopi, V.P., 2017. Identification of ayurvedic medicinal plants by image processing of leaf samples, in: 2017 Third International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN). Presented at the 2017 Third International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN), IEEE, Kolkata, pp. 231–238.

https://doi.org/10.1109/ICRCICN.2017.8234512

Mzoughi, O., Yahiaoui, I., Boujemaa, N., Zagrouba, E., 2013. Advanced tree species identification using multiple leaf parts image queries, in: 2013 IEEE International Conference on Image Processing. Presented at the 2013 20th IEEE International Conference on Image Processing (ICIP), IEEE, Melbourne, Australia, pp. 3967–3971.https://doi.org/10.1109/ICIP.2013.6738817

Pramanik, S., Bandyopadhyay, S.K., Bhattacharyya, D., Kim, T., 2010. Identification of Plant Using Leaf Image Analysis, in: Kim, T., Pal, S.K., Grosky, W.I., Pissinou, N., Shih, T.K., Ślęzak, D. (Eds.), Signal Processing and Multimedia. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 291–303. https://doi.org/10.1007/978 -3-642-17641-8_35

Preprocessing steps of leaf images: a) original image; b) greyscale... | Download Scientific Diagram [WWW Document], n.d. URL https://www.researchgate.net/figure/Preprocessingsteps-of-leaf-images-a-original-image-b-greyscale-image-cafter_fig1_320290546 (accessed 10.31.19).

Preprocessing steps performed on an Acer Palmatum leaf image. | Download Scientific Diagram [WWW Document], n.d. URL https://www.researchgate.net/figure/Preprocessing-stepsperformed-on-an-Acer-Palmatum-leaf-image_fig1_282148402 (accessed 10.31.19).

Sun, Y., Liu, Y., Wang, G., Zhang, H., 2017. Deep Learning for Plant Identification in Natural Environment. Comput. Intell. Neurosci. 2017, 1–6. https://doi.org/10.1 155/2017/7361042 Venkataraman, D., Mangayarkarasi, N., 2016. Computer vision based feature extraction of leaves for identification of medicinal values of plants, in: 2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC). Presented at the 2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC),

IEEE, Chennai, India, pp. 1–5.

https://doi.org/10.1109/ICCIC.2016.7919637

Acknowledgement

The authors would like to express the heartfelt gratitude to the research supervisor Professor R.P.S. Kathriarachchi for his valuable and constructive suggestions during the planning and development of this research and his willingness to give the time so generously is very much appreciated. Also would like to thank the staff of the Department of Information Technology, Faculty of Computing at KDU.

Authors Biography

Anuradha Kurunayaka is an undergraduate of Intake 34, Faculty of Computing, and Department of Information Technology at General Sir John Kotelawala Defence University.

D.P.M. Lakshan is an undergraduate of Intake 34, Faculty of Computing, and Department of Information Technology at General Sir John Kotelawala Defence University.

RPS Kathriarachchi obtained his BSc (Hons) in computer networks from the University of Wolverhampton UK and his master's degree in MIT in the same university. He also a CISCO certified network associate and having more than 15 years of IT and IS systems. His current research interests include the Internet of Things, Machine Learning, SWARM technologies. He has produced over ten peer-reviewed publications under his name.