Indian Herbal Leaf Identification using Artificial Intelligence Approach

1Roopashree S, 2Anitha J

Abstract

Inflation has alleged in side effects, deficiency in handling limited prolonged diseases, towering cost of the modern drugs has reintroduced the public interest in complementary and alternative drugs that is herbal remedies. Many foreign countries are fruitful in endorsing its own traditional therapies using herbs which have brought tremendous awareness among its people. India is still not tapping the full potential of its herbal remedies. Ayurveda, Traditional Indian Medicine (TIM) though active from long ago lacks in systematic research. Usage of the precise plant extracts is vital in Ayurveda treatment but many species are at risk of extinction due to lack of awareness and environmental degradation. Even today, identification and classifications of Indian medicinal plants are performed manually by experts who are few in number. Computer vision, patter recognition and image processing algorithms along with the availability of inexpensive camera offer various combinations of feature detection methods with different classifiers that can be used in building an automatic identification system for herbal leaves using leaf images and reveal its associated information to gain knowledge. While a botanist could be content with this system for recording a plant species discovered in its natural habitat by retrieving its associated information, the system aims in providing knowledge to every common man including hikers, campers, etc.

Keywords: Ayurveda Medicinal System, Computer Vision, Artificial Intelligence, Pattern Recognition, Digital Image Processing

Introduction

The mark of growing working population affluence with daily life and lessen of affordability in illness has increased the heed for traditional medicine. India is a widespread repository for the medicinal plants which harvests traditional medicine. In Traditional Indian Medicine (TIM), Ayurveda is long standing yet a living tradition. Primarily, Ayurveda objectives in promotion of health and quality of life.

Around 5000 years ago sprang in the pristine land of India, Ayurveda, the science of life and longevity is the ancient healthcare scheme in the world and it chains the intense thoughts of medicine and philosophy. Most of the Ayurveda medicines are primed using plant leaves and several of these plant species are in compromise.

Even today, proof of identity and classifications of Indian medicinal plants are performed manually by experts who are insufficient. Manual identification of herbs is time consuming and imprecise too. As plant recognition is an important task for botanist, drug designers, medicinal factory etc. using computer vision and pattern recognition the process can be automated.

One of the research topics in computer science is building an automatic recognition system of plants through the leaf images. Because of the availability of inexpensive camera and enormous study of computer vision, heaps of approaches provides a worthy response to leaf recognition when carefully chosen stable features have good capacity to distinguish individual leaves.
Figure 1: Proposed System Flow

Figure 1 explains the fundamental of recognition and classification system by Computer Vision using a leaf image.

The drawbacks of few existing methods are that some features are frequently preselected, but the properties of distinct features and blends of feature selection are not well explored in development of leaf identification. Subsequently, these preselected features do not equally or positively promote to the performance of accurate identification. Moreover, the leaf pattern varies with region, age of plant etc. Henceforward, a novel feature extraction algorithm is to be well-thought-out.

Leaf classification is a technique where leaf is categorized based on its diverse morphological features. There are innumerable effective classification techniques such as Support Vector Machine, k-Nearest Neighbor Classifier, Probabilistic Neural Network, Genetic Algorithm and Principal Component Analysis. Leaf classifications have wide-ranging claims in several fields such as botany, Ayurveda, agriculture etc. The observant study for the blend of feature extraction algorithm and classifier upholds the automated medicinal leaf identification.

Literature Review

Huge interest in the evolution of the old-style medicine has been amended by taking efforts in regulating and monitoring of the traditional plants. When correlated with China, it has been successful in endorsing its traditional therapies with agreeable scientific research and evidence base. But, Ayurveda needs more systematic research and evidence base. [1] Gives overview of elementary principles and commonalities of Traditional Indian Medicine (TIM) and Traditional Chinese Medicine (TCM). China has confidently supported its precise therapies over the world with a science-based methodology. This exploration springs a framework for research and development of Indian traditional plants by building a strong herbal database along with its associated information which are key source of India traditional medicine.

Automated Recognition of leaf images [2] using 12 Digital Morphological Feature extraction and Probabilistic Neural Network as classifier. Database of only 20 plants species grouping was prepared. Statistical Pattern Recognition approaches like noise was not well-thought-out. Accuracy more than 90% was attained.

Morphological features were extracted using MMC hypersphere Classifier [3] where the extracted features are from the leaf contours. The digital morphology features (DMF) mostly consist of geometrical features and invariable moment features. Accuracy was better shown in comparison to KNN Classifier. Fifteen features were used to categorize twenty species of plant. But, features like scaling, rotation and illumination were not thought-out.

Morphological features were extracted using SVM – BDT classifier [4] which showed superior performance in comparison with PNN and Fourier Moment based classifier. SVM-BDT were trained with 40 leaves from each 32 classes for training and 10 leaves from each 32 classes of plants from flavia database were used to test for accuracy. Rotation and illumination features and herbal leaves were not considered.

Herbal leaf identification [5] was accomplished using texture analyses. The investigation includes grey textures, Grey Tone Spatial Dependency Matrices (GTSDM) and Local Binary Pattern (LBP) texture combinations. Author concluded on a point that classification accuracy was in height in medicinal plants without pre-processing the images. The classification performance of Stochastic Gradient Descent, Decision Tree, and kNN classifiers were highlighted. As a whole, straight application of feature extraction without pre-processing with different classifiers produced better performance with flavia dataset.

Deterministic feature sequences are mined using GA technique [6] and trained by SVM classifier for 50 leaf images. General features, colour features, boundary and ripple features were
extracted. Genetic Algorithm (GA) an optimal approach for feature extraction and selection for classification of leaves was proposed. From the extracted features, the deterministic sequence is selected using GA technique and then further forwarded to be trained by Support Vector Machine (SVM). The experimental outcomes proved computationally operative in using GA for feature subset selection and SVM as a classifier. In contrast with kNN, SVM exhibited improved performance.

David G Lowe highlighted that Scale Invariant Feature Transform (SIFT) [7] is an object recognition system where new classes of local image features are incorporated. The features are invariant to image scaling, translation and rotation and partially invariant to illumination changes and affine or 3D projection. Author emphasized on SIFT as a solid object recognition which can be achieved in disordered and partially- obstructed images.

Data mining [8] plays a dynamic role in extraction of key information to help in decision making of a decision support system. Artificial Intelligence approaches improve the quality of decision support. Classification is one of the chief and treasured tasks of Data mining. Author made a detail comparison on SVM, SVM-ABoostM1 and Genetic Algorithm. Their projecting accuracy independent of data set size and domain. Underlined point was that SVM does not accomplish when class attribute in a dataset comprises of big number of diverse values. In large datasets training time is noteworthy. Genetic Algorithm would be the first choice when the selection criteria is accuracy and comprehensibility and decision tree when considering training time. SVM, the primary select when in view of predictive accuracy and training time.

ISIS (Integrated Scientific Information System) based Traditional Chinese Medicine (TCM) [9] has flexible desktop management system and also provides form-based customization. Original data stored in Microsoft Word is transformed using XML and Microsoft Visual Studio using C#. It provides Molecular information, Compound reference information, natural sources of compounds including Latin name, English name etc. It’s paid software of $185.

Recognition of leaves [12] is robust when selections of features are stable and have solid capability to distinguish the individual leaves. On this base, a methodology for recognizing plant leaf using domain related visual features were extracted and artificial neural network was projected.

**Proposed Approach**

From the above literature assessment, it’s easy to settle on that heaps of work had been accomplished for foreign country plant recognition but a very minor quantity of work has been through for Indian herbal leaves. A sensitive learning for the fusion of feature extraction algorithm and a classifier encourages the automated Indian herbal leaf identification system. A skilled study should comprise of seeing both theoretical and experimental features and checking for usefulness and efficiency. The surveyed TCM database integrates TCM plant and compound data. This when passed upon TIM enables the user to browse the botanical information and chemical structures of every herbal leaves in India to construct an effective and sound TIM database.

Ayurveda, arose in India is a medical system. As plant identification loads extensive consciousness and complex terms, even skilful botanists entail large time in this field for the mastery of the subject. Plant leaves are typically observed as the useful characteristics for species identification. In Ayurveda, leaf recognition over its images is considered as a vital research subject for Indian herbal leaf identification.

One of the straightforward feature for the design and development of the automated identification of Indian medicinal plant is to protect them from the commercial growth and also to accomplish an affirmative wakefulness of its evidence and its interrelated features. The conceptual diagram of the proposed approach is shown in Figure. 2.
Recent Advances in Technology and Engineering, Department of Computer Science and Engineering, T. John Institute of Technology, Bangalore, Published By: International Journal of Science, Engineering and Technology, ISSN (O): 2348-4098, ISSN (P): 2395-4752

**Figure 2:** Conceptual Diagram

**Image Procurement**

Leaf image to be recognized is learned by using a mobile camera or a digital camera which is fed to the identification system as shown in Figure 3. A complete database to be constructed by storing every detail information of every Indian herbal leaf.

**Figure 3:** Image Procurement

**Image Pre-Processing**

Image pre-processing take place earlier to feature extraction. This phase includes image cropping, image conversion and image enhancement. This phase augments the image quality by eradicating noise.

**Feature Extraction**

David Lowe [13] recommended a feature extraction algorithm - Scale Invariant Feature Transform. The algorithm extracts interest point features from the images which is used for matching of the same in changed views. Identification of features consists of staged filtering methodology. The features extracted are consistent to any changes in scale, affine transformation and illumination. It attains improved accuracy in verification and orientation when in par with indexing approaches. SIFT feature extraction algorithm comprises four stages such as Scale space extrema detection, Key point localization and filtering, Orientation assignment and Key point descriptor.

**Classification**

The classification objective is to arrange images by utilizing the logical data. TIM Leaves can be classified to their respective species based on the classification results.

Wu et al., projected a method [2] for leaf classification using Principal Component Analysis (PCA) to present original data as the linear combination of certain linear irrelevant variables. PCA is used to form an input vector for PNN. Probabilistic neural network is derived from Radial Basis Function (RBF) Network which is an Artificial Neural Network. RBF is a bell shaped function that measures the variable non-linearly. Several benefits has been perceived in PNN such as faster training speed when paralleled with Back Propagation network, robust to noisy input, simple in structure and training and also easy to train the network.

Support Vector Machine (SVM) [14] is a supervised learning ideal in recognition and classification process for their favorable performance. A lot of real-world problems can be resolved by SVM multiclass classification capably with precise theoretical model for multiclass dataset. Their performance can further be improved on the parameter selection and kernel design. SVM maximizes geometric margin at the same time it can diminish the empirical classification error. Implementing both the classifiers to provide a comparison on the herbal leaf dataset would be considered as a detailed contrast between the two.

**Novel Combination of Feature Extraction and Classifier**

Based on the above survey on the feature extraction and classifiers, a novel combination of the best classifier with SIFT feature extraction to be discovered for effective TIM leaves identification and
classification. As a result, which would eventually be very beneficial for any common man in identification process. This ensures to gain knowledge and also preserve TIM leaves which are very much essential for its protection. This competent study embraces sighting both theoretical and experimental aspects and checking for usefulness and efficiency of Traditional Indian Medicine.

TIM Scientific Database

TIM, a health care scheme to be designed. A database management system which would largely be used for storing, retrieving and inspecting of chemical structure and the corresponding scientific data of the medicinal plants. Added to this, maintain compound discovery, research and development used by pharmaceuticals, biotech, agrochemical and chemical industries world-wide. Some of the advantages of TIM database: (1) Better understanding & knowledge of traditional medicine. (2) Outshine the shooting accuracy and the searching procedure. (3) Well-organized in drug design or any herbal product design plant.

The TIM database consisting of the many Ayurveda plants along with their pertinent statistics can be presented in many types of customized forms and also in many smartphone platforms.

Conclusion

It can be concluded that loads of work has been carried out on foreign country plant recognition but not for Traditional Indian Medicinal herbs In Ayurveda, the right plant extracts is key for most treatments. For identification and classification, a combination of different artificial intelligence approach which is fast in execution and easy in implementation is to be incorporated. This work is a determination to accomplish an automatic identification system for Indian herbal leaves and providing the statistical nature of leaf features using TIM database.

One of the clear-cut features for the design and development of the automated identification of TIM is to safeguard them from the commercial expansion and also to attain positive awareness of its evidence and correlated features. Their significance lies in medicines, food/cosmetic products etc. and would chiefly be benefited by horticulturists, trekkers, botanists and most importantly herbal leaf

knowledge can be updated easily to every single common man in the world. As per the survey, robust and effective automated identification and classification of Ayurveda leaf would be extraction of SIFT feature and implementing SVM classifier for better accuracy. Also, portability can be achieved through different smartphone platforms.

References


[10]. Sandeep Kumar E, “Leaf Color, Area and Edge Features Based Approach for Identification of Indian
Medicinal Plants”; Indian Journal of Computer Science and Engineering (IJCSE).


Author’s details

1. Ms. Roopashree S, Assistant Professor, Department of Computer Science and Engineering, T. John Institute of Technology, Bengaluru, Karnataka, India, roopashaily@gmail.com

2. Dr. J Anitha, Professor, Department of Computer Science & Engg, Dayananda Sagar Academy of Technology and Management, Karnataka, India, anitha.jayapalan@gmail.com