Chapter Number : 1
Chapter Name : Introduction

Contents

1.1 Objective
1.2 Problem Definition
1.3 Strategy Planned
1.4 Biological Background
1.5 Introduction
1.6 Need and Significance
1.7 Objectives of work
1.8 Proposed Scheme
1.9 Flow Diagram of Proposed Scheme
1.1 Objective

The field of digital image processing and analysis of images can be merge with interdisciplinary context like mathematics, physics, chemistry, biology, computer science and many more. This research study has interdisciplinary relevance of digital image processing with botany. So finding of this study is useful to botanist and taxonomist. In this chapter I have presented description of plat body and structure of leaves in simple fashion. The perspective and social motivation to conserve the plant species is described in this chapter. The aim of this chapter is to give brief information about study of research, strategy planned, its biological background, significance of the study in the context of current status, its interdisciplinary relevance and organization of thesis.
1.2 Problem definition

1.2.1 Thesis title

A Study of Plant Species Classification based on Digital Morphological Features of Leaf

1.2.2 Problem statement

A study of various available plant species to recognize desired plant from plant kingdom. Geometrical and digital morphological features of leaf are extracted by introducing digital image processing and data processing theory into the numerical taxonomy in botany. Artificial intelligence approach is used to facilitate fast classifying the plant and understanding and managing them with better efficiency and effectiveness of result in order in protecting earth’s ecosystem and atmosphere.

1.2.3 Category

Image and data processing

1.2.4 Subject descriptor

Pattern classification

1.2.5 General terms

- Algorithm
- Design
- Experimentation
- Performance

1.2.6 Keywords

- Feature extraction
- Plant species
- Classification
- Leaf shape
- Morphological features
- Geometrical features
1.3 Strategy planned

1.3.1 Vision

To improve the effectiveness of classification using feature extraction.

1.3.2 Mission

To design a system which can classify the input leaf of particular plant species by using probabilistic neural network (PNN) algorithm based on digital morphological features of leaf so as to save or conserve the plant species.

1.3.3 Goal

The aim of the research is to implement a system which can select or extract features from large datasets depending on the problem at hand and which will be combined with probabilistic neural network classifier in order to provide effective means of classification for the application in the field of botany.
1.4 Biological background

Plant taxonomy is the field of science which discovers, identifies, describes, classifies and names plants. Classification is termed as grouping of particular object such that all object belong to that group have certain common features. One of the important application of plant species classification is one can make a decision for conservation, research, agriculture for purpose of particular plant. Typically plant body contains the root system which is usually underground and the shoot system which is usually above ground. The root system anchors the plant in the soil and absorbs water and nutrients. The shoot system includes leaves and other reproductive organs. This research is focus on plant species classification based on leaf.

![The Plant Body](image)

**Figure 1.1 : The Plant Body**

1.4.1 Structure of leaves

Leaves are the principle structure of plant. They are attached at the node of stem. Leaves cell contain chloroplasts due to which generally leaves are green in colour. Chloroplasts contains green pigment called chlorophyll. As chloroplast is green, so ultimately leaves get green colour. A leaf is divided into two parts. One part is called blade and other one petiole. The petiole is also called leaf stalk which is thin and stick like body between the blade and the the node. Sometimes
at the start of petiole there is a small leaf like structure called the stipule. If petiole is absent then leaf is said to be sessile.

![Figure 1.2: Structure of Leaf](image)

1.4.2 Arrangement of leaves

Leaves are arranged on stem in one of three ways- alternate, opposite and whorled. In alternate arrangement leaves are arranged in staggered fashion along stem. In opposite arrangement pair of leaves are organized such that they are opposite to each other and lastly in whorled arrangement leaves are arranged in ring fashion or three or more leaves are attached at one node.

![Figure 1.3: Arrangement of Leaves](image)

1.4.3 Shapes of leaf

This research describes use of shape tool for plant species classification. Generally leaves are broad, thin and flat in shape. The shape and structure of leaf varies from plant to plant. It depends on environmental condition or availability of light. Normally a leaf can have oval, lanceolate, obovate, elliptical, spatulate, cordate, oblanceolate, obcordate, oblong, linear, peltate, cuneate, reniform or hastate in shape. Some leaves have straight margins and some have uneven margins. The following figure shows various shapes of leaf
Figure 1.4: Various Shapes of Leaf
1.5 Introduction

Plants are useful for human as well as animals too. Plants provide food, medicine, oxygen and much more. Plants play important role in natural cycle of life. It is necessary to know which plants are useful and which are harmful to human being, so as to save the life of living being. It is very difficult as well as important too to recognize plant species on earth as large number of variable plant species are available all over the world. To conserve the plant species, it is necessary to perform classification and protect them before endangered. The botanists follow traditional method which is time consuming and involves large amount of their efforts. Plant species classification from body parts of the plant is one of the most present research work going in the field of image processing and computer vision. However, in order to recognize plant species there have been many recent studies on computer vision field which focus substantially on plant identification based on features of leaf. One can identify plant species from the parts of plant like flower, fruit, bark, seed or leaf. This is one of the most popular approach in the area of image processing. It is very difficult to analyze the shape of flower, seedling and fruit because of their 3 dimensional structure and it increases complexity too. Plant identification based on flower or fruit involve knowledge of morphological features such as number of stamens in flower and number of ovaries in fruits. Identifying plant species using such keys is really a very time consuming process. In addition to these, there are various other drawbacks in identifying plants using these features such as flowers can only be obtained during blooming season and fruits are also available only in specific seasons. Not every plant is having flower or fruit but every plant is having leaf. So these parts are not suitable for classification. One way to recognize the plant is from their leaves because every plant species leaf shape is different from any of the other one. Leaves can be found everywhere at all seasons. As compare with flower, seed and fruit of plant, leaves are two dimensional in nature. Thus, they are most suitable for classification. Each plant leaf carries large amount of information about plant. The differentiating parameters can be shape, colour, texture or margin of leaf. These differentiating characteristics can be used for plant species classification. The colour of most of the plant leaves is green. Due to nutritional value in leaf and changes in atmosphere and season and soil-water relation causes variety of shades of green colour for leaves of plant. So the colour feature has less reliability for classification purpose. Shape is easiest to extract from images. Thus shape characteristic of leaf is more reliable for research work. It serve as a tool to plant biologists and botanists for distinguishing plant species. Classification of plants based on leaves is the fastest and simpler
way to identify a plant. These are the most visible and variable aspect of a plant from trees. They can be easily found everywhere. So it is necessary to establish database by information technology as soon as possible. Sampling leaves and photoing them is expensive in cost. We can easily transfer the leaf image to a computer. The work performed by many researchers’ based on shape features extraction are reviewed. The proposed scheme commence with image preprocessing on input leaf image. Finally it determines outline of shape. An algorithm is required to compare the image of a leaf with a database and identify the leaf based on its features irrespective of its size, orientation and its background. The plant species classification system depends on how well the edges of the leaf from an input image have been detected from its background. The computer can extract necessary features automatically using image processing techniques and subsequently can recognize the plant leaf using machine learning.

1.5.1 Significance of the study in the context of current status

The outcome of the proposed work will be useful in application concerned with plant identification and classification which further useful for botanists and taxonomist.

1.5.2 Interdisciplinary relevance

The proposed work is related to a discipline of botany, computer vision and image processing.

1.5.3 Organization of thesis

This thesis includes five chapters namely introduction, literature review, methodology experimental result, conclusion and publications.

In chapter 1, I have explain plant shoot system and further I presented the perspective and motivation for the study of research work. After an extensive but not exhaustive necessary brief description about flow of research work and proposed scheme is mentioned.

Chapter 2 gives detail information of previous work done by researchers in the context of proposed theme of research. This chapter also describes detail study of various techniques used previously to solve the same problem. This chapter also proposes various shape methods. Finally it gives brief overview of comparative study of methods used previously.

Chapter 3 discusses image pre-processing methodology used to sharpen the images and introduce how twelve geometrical and digital morphological features are extracted from leaf image. This chapter also gives detail information about principal component analysis which is used to orthogonalized the extracted features and architecture of probabilistic neural network to classify plant species.
Chapter 4 gives detail information about dataset used and experimental results of the research study. This chapter describes finding of the feature extraction process and results of applied classification strategy. This chapter gives experimental results of image preprocessing, geometrical feature extraction, digital morphological feature extraction, principal component analysis, probabilistic neural network and results of classification. This chapter brief about comparison of results of two dataset and comparison of results with other algorithm.

Chapter 5 concludes the outcome of research study and proposes the scope for future work. Finally it gives detail information about publication in context with research study.
1.6 Need and significance

It is well known that plants play a crucial role in preserving earth’s ecology and environment by maintaining a healthy atmosphere and providing sustenance and shelter to innumerable insect and animal species. Plants are also important for their medicinal properties, as alternative energy sources like bio-fuel and for meeting our various domestic requirements like timber, clothing, food and cosmetics. All over the world, there are currently about 310000–420000 known plant species, and many are still unknown yet. At present, plant taxonomy usually adopts traditional classification method. And so far many other classification methods, such as morphologic anatomy, cell biology, molecule biology, phytochemistry have also been used. These methods have great relation to do with biology and chemistry. With the deterioration of environments, more and more rare plant species are at the margin of extinction. Many of rare plants have died out. Nowadays, there are about 22–47% plant species of all known plants that are endangered, of which, there are about 100000–150000 plants that probably will die out in short time. Plants are the backbone of every living life on earth.

1.6.1 Plants used as food source

Most food has its origin in plants. Some food is obtained from plants, but even animals that are used as food sources are raised by feeding them food derived from plants. So either indirectly or directly, human and animal food is dependent on plants. Human eat corn, rice, wheat and other cereals, which are gain from plant. Human food also includes spices, nuts, fruit, flower, tea and coffee etc. are gain from plant. We obtain honey from flower, sugar from sugar cane, oil from seed or flower etc. Animals feed on grasses and cereal plants.

1.6.2 Plants provide fresh air

Humans breath out carbon dioxide which is take in by plants and in place of it they give oxygen. Human and animal need oxygen in the air to breath. Oxygen is produced by plants. Oxygen is also a by-product of photosynthesis. Plants stored carbon which is produced from burning fossil fuels. If this carbon is not absorbed by plant then it is out into the environment and finally it gets polluted. In this way plants improve air quality. Plants help in preventing soil erosion. They also regulate water cycle. As human and animal can not live without breathing so we can’t imagine life on earth without plant.
1.6.3 Plants used in medicine

Most of medicines given for various disease are come directly from plants. Many people rely on plants for primary healthcare. There are many herbal medicines too. There are various plants having medicinal substance like neem, aloe vera, cucumber, lemon, tulsi etc. Nowadays plants are also used in research of finding useful drug to treat cancer.

1.6.4 Plants used in industry

Plants used in industry to manufacture many useful products which are useful for human being. Mainly plants are use to manufacture paper because most paper comes from softwood trees such as pines. The industrial organic chemicals were derived from plants. Perfume oils are manufacture from the flowers and leaves of certain plants.

1.6.5 Plants used to prevent soil erosion

Soil contains nutritional elements and useful bacteria. Soil erosion caused by rain and winds. Soil gets polluted due to overuse of chemical fertilizers, release of polluted water, chemical and slag from factories into the soil. This can be prevented by planting of trees, preventing felling of trees.

1.6.7 Plants used to maintain eco balance

Plants create food and shelter for wild life which gives every species sources to sustain. This is ensure all members of ecology will sustain on the earth.

So to conserve plant species, their identification is the first step. The traditional method includes manual labeling to plant which is time consuming and not reliable. It is highly essential to have an object recognition system to identify various species and protect them from being endangered. The development in the field of computer vision can be utilized for this purpose. So, it’s necessary and urgent for us to establish bio-diversity database by information technology as soon as possible. However, it is an important and difficult task to recognize plant species on earth. Designing a convenient and automatic recognition system of plants is necessary and useful since it can facilitate fast classifying plants and understanding and managing them.
1.7 Objectives of work

The objectives of the proposed research work has been listed below:

1) To preprocess the query image
2) To extract geometrical and digital morphological features
3) To facilitate fast classifying plant and understanding and managing them
4) To demonstrate efficiency and effectiveness of proposed method
1.8 Proposed scheme

This research work is focus on introducing plant recognition methodology by using digital image processing theory. The research work is organized as follows: Firstly, introducing the leaf image preprocessing method. Secondly, designing digital morphological feature extraction method. The features are extracted from the contours of leaf. It generally includes geometrical features and digital morphological features. Third, Implementation of principal component analysis. Fourth, Classification and recognition of plant species using probabilistic neural network.

For the first phase we have considered the dataset provided by Intelligent Computing Laboratory, Institute of Intelligent Machine, Chinese Academy of sciences. Firstly we have converted the 2 dimensional image of plant leaf to grey level image. for each conversion of a color image to grayscale is not unique, different weighting of the color channels effectively represent the effect of shooting black-and-white film with different strategy is to match the luminance of the grayscale image to the luminance of the color image. To convert plant leaf color image to a grayscale representation of its luminance, Firstly we have obtained the values of its red, green and blue(RGB)primaries. Finally only gray component for each is computed from the color image by

\[ \text{Gray} = 0.2999 \times R + 0.578 \times G + 0.114 \times B \]

Where R,G,B correspond to the color of the pixel, respectively. This image is then transformed into a binary image. Further, smoothing filter is applied to the binary image to reduce the noise. The boundaries are gained by applying a laplacian filter.

Original Image(RGB)⇒Gray⇒Binary⇒Smoothing⇒Contours of Leaf

The next step in plant leaf recognition is the digital morphological feature extraction. The aim behind this step is to remove redundancy from the image and to represent by a set of numerical features. These features are extracted from the contours of leaf. The digital morphological features generally include geometrical features and invariable moment features. The digital morphological features contain basic geometric features and morphological features. Our system extract 5 basic geometric features i.e longest diameter, physiological width, leaf area and leaf perimeter. Our system also extract 12 digital morphological feature are as follows smooth factor, aspect ratio, form factor, rectangularity, narrow factor, perimeter ratio of...
diameter, perimeter ratio of physiological length and physiological width and 5 vein features: \(f_1, f_2, \ldots, f_5\).

In the third step, we want to reduce the dimension of input vector of neural network. This is achieved by combining features. Linear combinations are attractive because they are simple to compute and analytically tractable. In effect linear methods project the high dimensional data onto a lower dimensional space. Principal component analysis is used to orthogonalize 12 features. PCA transforms the data to a new coordinate system such that the greatest variance by any projection of the data comes to lie on the first coordinate, the second greatest variance on the second coordinate, and so on. Each coordinate is called a principal component. In this research we have adopted 5 principal components.

Once the five principal components extracted from 12 digital morphological features then probabilistic neural network is used for classification purpose. PNN(probabilistic neural network) has 3 layers input layer, radial basis layer and the competitive layer. When 5 principal component of a leaf are given to input layer, radial basis layer find the distances from the input to the training input and gives result whose elements are close the input is to training input. These distances are scaled by radial basis function. Finally competitive layer finds the shortest distance among them and find the training pattern closest to the input pattern based on their distance and classify the given plant leaf.
1.9 Flow diagram of proposed scheme

Figure 1.5: Workflow of System