CHAPTER 1

INTRODUCTION
1.1 Overview

The human beings by the grace of God have dexterity in clearly demarcating between the voices of a human being from that of the sound of a violin, they have the ability to identify the handwritten numerals written, their ability to sniff and distinguish the odors of a rose from that of the onion, sensing the environment and taking an action according to the observations made during the event like recognizing the old friend out a worn and tattered school photograph, the ability of the pharmacist to read the prescription written by the doctor, comparing the fresh food from the stale are just a few remarkable abilities displayed by almost all the human beings. The scientists expect the similar behavior be imitated by the machines. This ability of human beings is owed to the perception skills acquired which are based upon identifying peculiar patterns, striking features, color etc. which could serve as a cue for identification or information retrieval at the hour of the need.

According to Gutierrez (2013), the concept of recognition by humans or machines involves the use of a collection of set of peculiar features unique to the object of interest. The term feature can be defined as any distinctive aspect, quality or characteristic which, may be symbolic (i.e., color) or numeric (i.e., height). The combination of these $d$ features is represented as a $d$-dimensional column vector called a feature vector. The $d$-dimensional space defined by the feature vector is called feature space. The Figure 1.1 shows that a feature vector with several elements identified from the object of interest. A feature vector is a vector of observations and $x$ is a point in feature space $X$ and can be represented as $x \in X$.

The word pattern is an entity, which can be defined, named or heard. It can be the sound track of the of song sung by Lata Mangeshkar, liked for the sweetness of the soothing sound, impression of the fingerprint, a human face, pattern sequence of DNA or the signature etc. The articulation produces sound waves which the ear conveys to the brain for processing. The human being is able to distinguish amongst the speakers due to their vocal range, vocal quality and accent. A pattern is defined as composite of features that are characteristic of an individual.
In classification, a pattern is a pair of variables \( \{x, C\} \) where \( x \) is a collection of observations or features (feature vector) and \( C \) is the concept behind the observation (label). The quality of a feature vector is related to its ability to discriminate examples from different classes. The examples from the same class should have similar feature values and should be different from the other classes.

According to Gutierrez (2013), in machines it is the algorithmic steps that act as a classifier, while in the humans, it is the brain, but the basic work remains the same and that is to arrange the objects into classes according to the patterns identified, which has been shown in the Figure 1.2. The whole feature space has been divided into different regions depending upon the patterns available. The boundary separating the feature space can be linear or wavy. The generation of algorithms for pattern recognition is greatly influenced by human effort for recognizing the things. The different features are extracted from the different regions of interest. These features should be unique and it
is expected that the features have definite space and the boundary of one feature set should be clearly separable from the other. Therefore the good features have a clear cut boundary and they belong to the same class with same class values, whereas the features from different classes have different feature values. The features extracted can be good features if they form a clear cut boundary and are separable completely and on the other hand the features can be bad ones, if the values or properties of the features are crossing the boundary or overlapping, which has been shown in the Figure 1.3.

![Figure 1.3](image) Types of features (a) Good (b) Bad

As the existing work deals purely with digital images, therefore it becomes quite imperative to study the human visual system, though the study of the complexity of the human visual system cannot be elaborated in few paragraphs.

### 1.2 Human Visual System

The human visual system gives our bodies the ability to see the physical world. During this process, there is a communication of signals between the major sensory organ that is the eye and the central nervous system (CNS) and the external stimuli is converted into a visual scene. There is a large difference between images that are displayed on the monitors and the image that is perceived by the human eye. There is a large difference between the luminance of the pixel and the brightness of the same pixel that is perceived by the eye. By increasing the pixel luminance value, the brightness does not increase as
brightness is dependent upon the contrast of the neighboring region of the pixel of interest.

The human visual system has two major functional parts, eye and the brain. The brain does all the image processing tasks and the eyes are the biological equivalent of the camera (a physical device). Figure 1.4 shows the detailed view of the eye with all important parts as described by Szeliski (2011).

![Figure 1.4 Internal structure of an eye](image)

The Figure 1.4 shows the details of the human visual system. Our eyes perceive the light rays emitted or reflected from the scene when in the range of visible spectrum (300 to 700nm), the eye reacts to such a ray of light and then sends an electric signal to to optic nerve. When the light ray hits the eye, it will pass through cornea and then through the aqueous humor, iris, the lens and the vitreous humor and then finally reaches the retina. The cornea is a transparent protective layer, which acts as a lens and refracts the light. The iris forms a round aperture that can vary in size and so determines the amount of light that can pass through. When the person is standing in the dark, the iris is wide open, and lets through as much light as possible. In normal daylight, the iris constricts to a small hole. The lens can vary its shape to focus the perceived image onto the retina. In the retina, the light rays are detected and converted to electrical signals by
photoreceptors. The eye has two types of photoreceptors: rods and cones, named after their approximate shape. There are some 100 million rods and are spread around retina except for fovea where there are no rod cells. The fovea is the area of the retina where the vision is the sharpest. There are no photoreceptors around the optic nerve and is called blind spot of the eye. The rods are more responsive to light under dark circumstances, the rods cannot perceive color, but they can perceive the different shades of gray and the phenomena is called scotopic or night vision. Under the day light circumstances the cones are active and we experience photopic or day vision. In the regions where there is dim light, it causes intermediate stage and both the rods and cones are active during this period and this is called mesopic vision.

The different types of color is perceived due to three different types of cones available in our eyes. The first of the color identifying cones is able to detect the wavelength of 400 to 500 nm (blue-Violet Spectrum). The other two are color sensors help in viewing the color in the cyan-red range.

Figure 1.5 shows the three types of cones used by humans to identify the three colors as described by Kriegman (2011). The blue color cones sense the blue color at lower wavelengths, the green color cones sense the middle wavelengths and the red cones sense the longer wavelengths.
According to Curcio et al. (1990), Figure 1.6 shows the distribution of rods and cones in the human visual systems. Consider a light source that emits light rays with energy $E(\lambda)$, where $\lambda$ is the wavelength of the emitted light, then the light is reflected from a certain object can be written as Eq. (1.1).

$$I(\lambda) = \rho(\lambda)E(\lambda)$$

(1.1)

where $\rho(\lambda)$ is the reflectivity of that object in Eq. (1.1). The reflectivity is a function that takes values between 0 and 1. Suppose $\rho(\lambda_1) = 1$, then this means that all of the light with wavelength $\lambda_1$ is reflected. Alternatively, if $\rho(\lambda_2) = 0$, this means none of the light at wavelength $\lambda_2$ is reflected, i.e., all of the light at this wavelength is absorbed.
by the object. Effectively, the reflectivity function $\rho(\lambda)$ determines the color of an object. For example: if a certain object reflects only light with wavelengths $\lambda$ of about 650 nm (i.e., $\rho$ is zero for other values of $\lambda$), it denotes red color.

According to Szeliski (2011), the luminance $L$ of an object is defined by Eq. (1.2).

$$L = \int_0^\infty I(\lambda)V(\lambda)d\lambda$$

(1.2)

where $V(\lambda)$ is the luminous efficiency function of a visual system. This function $V(\lambda)$ tells about how well a visual system is able to detect light of a certain wavelength. Every creature on the earth be it humans, vultures or insects have their own luminance efficiency function. Therefore, the image formed or seen by the rabbits, snakes or camera is different from the human visual system due to differential reaction to different light wavelengths. The difference in the visual image formed is due to the presence of different varieties of cones present in the visual system; the humans have three different varieties of cones, the turtles have two varieties of cones with four different kinds of colored oil, the pigeons have three different kinds of cones with five different kinds of colored oil. It has been observed that pigeon can differentiate four more shades of red than human beings. It has been observed that higher contrast is required to detect the spots of light in high luminance area in the background.

According to Szeliski (2011), Figure 1.8 (a) shows the Mach band effect where each bar is uniformly grey, but our visual system enhances each contrast jump. The Mach band effect shows us that our visual system sharpens the edges of the objects we perceive by adding a little contrast. Figure 1.8 (b) shows the simultaneous contrast effect which is another effect that shows us that the perceived brightness depends on the contrast. In Figure 1.8 (b), even though the center squares are of equal luminance, they appear to be brighter if the background is darker and the one with the largest contrast to the background appears darkest. Depth perception refers to our ability to see the world in three dimensions. With this ability, we can interact with the physical world by accurately gauging the distance to a given object. While depth perception is often attributed to binocular vision (vision from two eyes), it also relies heavily on monocular
cues (cues from only one eye) to function properly. These cues range from the convergence of our eyes and accommodation of the lens to optical flow and motion. Due to the presence of the receptors, the human visual system is quite complex and needs a great attention in order to develop automatic recognition systems for the humans based on the human visual system. Figures 1.8(a) and 1.8(b) show that contrast and luminance have a very important role to play in the design and development of automatic pattern recognition systems based purely on the human perception of vision.

![Figure 1.8](image)

**Figure 1.8** Effects of (a) Mach band (b) Simultaneous contrast

### 1.3 Overview of Computer Vision and Machine Learning

Our world that we live in is full of images and videos, be it a public image or videos captured by the surveillance cameras installed at the airports or railway stations or personal images and videos. The vision is an amazing feat of natural intelligence. There are two terms that are often interchangeably used and they are computer vision and machine vision. These two terms seem to be the same but in fact are different. The term computer vision represents a methodology in which the images are captured and the image analysis is performed automatically. In computer vision the stress is on image analysis because of the wide range of applications.

Computer vision is mainly used for checking, automation and recognition. Particular examples can be found in product packaging (deformation checking), automation of
material distribution (assembly lines) and recognition of color, smoke or recognition of vital functions in senior citizens. The computer vision technology can analyze and measure almost anything that is captured by a camera, and based on its assessment the technology automatically sends instructions for further requested actions.

Typical feats achieved by computer vision as described by Rao (2009):

- The remarkable story of Sharbat Gula, who was first photographed in 1984 aged 12 in a refugee camp in Pakistan by National Geographic photographer Steve McCurry as shown in Figures 1.9(a) and 1.9(b) and traced 18 years later to a remote part of Afghanistan where she was again photographed by McCurry, is told by National Geographic in their magazine (April 2002 issue) and on their website as shown in Figure 1.9 (c). The problem was solved using computer vision based iris recognition systems. This is one of the popular examples of problems solved through technology.

- Smart cars: The vision systems have been adopted in high-end BMW, GM, Volvo models as shown in Figure 1.10. The specialty of such cars is their sensing ability, advanced warning system and can provide more safety to pedestrians and road vehicles.

- Google cars: The car is a project of Google as shown in Figure 1.11, which has been working in secret, but in plain view on vehicles that can drive themselves, using artificial-intelligence software that can sense anything near the car and mimic the decisions made by a human driver as shown in Figure 1.11 (b). A self-driving car developed and outfitted by Google, with device on roof, cruising along recently on Highway 101 in Mountain View, California.

![Figure 1.9](image)

Figure 1.9  Face identification through IRIS recognition: Picture taken in (a), (b) 1984 and appeared in National Geographic (c) 2002
Figure 1.10  Smart Cars

Figure 1.11  Google Cars (a) In a traffic (b) With a driver

Figure 1.12  Foundation of computer vision
Figure 1.12 shows the foundation of computer vision as depicted by Li (2008). The elements of computer vision are machine learning, information retrieval, cognitive sciences and neurosciences. In addition to the above mentioned branches of learning control robotics, speech have a contributory relationship with computer vision.

Machine learning has evolved from the study of pattern recognition and computational learning theory in artificial intelligence. According to Samuel (1959) defined machine learning as a "Field of study that gives computers the ability to learn without being explicitly programmed". Machine learning explores the study and construction of algorithms that can learn from and make predictions on data. Such algorithms operate by building a model from example inputs in order to make data-driven predictions or decisions, rather than following strictly static program instructions.

1.3.1 Categories of machine learning tasks

Machine learning tasks are typically classified into three broad categories, depending on the nature of the learning "signal" or "feedback" available to a learning system as depicted in Figure 1.13 and described by Gutierrez (2013).

![Figure 1.13](image)

**Figure 1.13** Machine learning tasks (a) Supervised learning (b) Unsupervised learning (c) Semi-supervised learning
• **Supervised learning:** The computer is presented with example inputs and their desired outputs, given by a "teacher", and the goal is to learn a general rule that maps inputs to outputs. Simply, it can be interpreted as a kind of learning in which input data is called training data and has a known label or result such as spam/not-spam or a stock price at a time. A model is prepared through a training process where it is required to make predictions and is corrected when those predictions are wrong. The training process continues until the model achieves a desired level of accuracy on the training data as shown in Figure 1.13(a). Example problems which are a part of this learning are classification and regression and the popular algorithms are Logistic Regression and the Back Propagation Neural Network.

• **Unsupervised learning:** In this approach, no labels are given to the learning algorithm, leaving it on its own to find structure in its input. Unsupervised learning can be a goal in itself (discovering hidden patterns in data) or a means towards an end (feature learning). Simply it can be interpreted as a kind of learning in which input data is not labelled and does not have a known result. A model is prepared by deducing structures present in the input data. This may be to extract general rules. It may through a mathematical process to systematically reduce redundancy, or it may be to organize data by similarity as shown in Figure 1.13 (b). The examples of the problems of this learning are clustering, dimensionality reduction and association rule learning and the algorithms are Apriori algorithm and k-Means.

• **Semi-supervised learning:** A computer program interacts with a dynamic environment in which it must perform a certain goal (such as driving a vehicle), without a teacher explicitly telling it whether it has come close to its goal. Another example is learning to play a game by playing against an opponent. Simply, it can be interpreted as a kind of learning in which input data is a mixture of labelled and un-labelled examples. There is a desired prediction problem, but the model must learn the structures to organize the data as well as make predictions as shown in Figure 1.13 (c). The example problems of this learning
are classification and regression and algorithms are extensions to other flexible methods that make assumptions about how to model the un-labelled data.

1.3.2 Machine learning tools

The tools mentioned in this section are the latest in the state of the art in measuring important properties of the leaves of the plants. The unique feature of these instruments is the portability to the fields and battery operability. The results for area, accumulated area, mean area, diseased leaf area, length, width and shape factor can be instantly computed upon placing the leaf under study on the glass of the scanner.

- **AM350**: The AM350 which is a product of ADC Bio scientific, is a compact and portable leaf area meter suitable for the accurate, non-destructive measurement of leaf area and associated parameters. The AM350 consists of a high resolution scanner and scan board with integral data analysis and image storage. A multi-position carrying handle is provided for increased portability and ease of use. Measurement of leaves can be carried out on the scan-board or an independent plain surface. The large, high contrast liquid crystal display offers visual confirmation of scanned leaf area together with the measured leaf parameters. Operation is by menu driven software and measurements may be displayed in mm, cm or inches. An adjustable contrast control makes the AM350 suitable for damaged, discoloured or diseased leaf applications. Using a high speed USB port, the stored image can be downloaded, in bmp or tif formats, into commercially available image analysis software.

An increased understanding of leaf area development is important in a number of fields: in food and non-food crops, for example short rotation forestry as a biofuels feedstock, leaf area is intricately linked to biomass productivity; in paleontology leaf shape characteristics are used to reconstruct paleoclimate history. Such fields require measurement of large collections of leaves, with
resulting conclusions being highly influenced by the accuracy of the phenotypic measurement process. The AM350 has been shown in Figure 1.14 (a).

![AM350 Portable Leaf Area Meter and Leafsnap](image)

**Figure 1.14** Machine learning tools (a) AM350 Portable Leaf Area Meter (b) Leafsnap

- **Leafsnap**: It is the first in a series of electronic field guides being developed by researchers from Columbia University, the University of Maryland, and the Smithsonian Institution. This free mobile app uses visual recognition software to help identify tree species from photographs of their leaves. The Leafsnap has been shown in Figure 1.14 (b) as described by Smithsonian.

![Lamina, cropped image with dimension measurement and serration detection, cropped image with holes, regression analysis, and principle component analysis](image)

**Figure 1.15** (a) Screenshot of Lamina (b) Example of cropped image with dimension measurement and serration detection (c) Cropped image with holes (marked in green) and serration (marked in blue) (d) Regression analysis to compare data generated (e) Principle component analysis with loadings
• **LAMINA (Leaf shApe deterMINAtion):** According to Bylesjo et al. (2008), it is a tool for rapid quantification of leaf size and shape parameters. LAMINA has been designed to provide classical indicators of leaf shape (blade dimensions) and size (area), which are typically required for correlation analysis to biomass productivity, as well as measures that indicate asymmetry in leaf shape, leaf serration traits, and measures of herbivory damage (missing leaf area).

The main computations that can be performed using Lamina are thresholding, segmentation and filtering of the digital images. The object boundaries, cavities, serrations and indents can also be obtained using Lamina. After processing, LAMINA outputs cropped image files representing the identified objects after thresholding and segmentation. This allows the user to have a record of the results of the image analysis process. A number of quantitative measurements of the leaves are generated. This includes the leaf area, height, width, circularity, number of serrations, indent widths and depths as well as the boundary coordinates (normalised against leaf centre). For parameters that summarise several measurements, the output includes the mean, median and standard deviation. The Lamina has been shown in Figure 1.15, which displays its unique functions.

### 1.4 Overview of Pattern Recognition

The process of pattern recognition involves identification or interpretation of patterns found in the images. The pattern recognition technique involves the allocation of a label or a class to a given input instance according to some pattern recognition algorithm. Pattern recognition provides the solution to various problems from speech recognition, face recognition to classification of handwritten characters and medical diagnosis. The various application areas of pattern recognition are bioinformatics, document classification, image analysis, data mining, industrial automation, biometric recognition, remote sensing, handwritten text analysis, medical diagnosis, speech
recognition, GIS and many more. The striking similarity between all these applications is that for a solution-finding approach, features have to be extracted and then analyzed for recognition and classification purpose. There are three major processes that take place in a pattern recognition task. The first task involves the acquisition of data for the problem in question. The data acquisition process involves converting data from one form (speech, character, pictures etc.) into another form which is well understood and acceptable to the computing device for further processing. The data acquisition is generally performed by sensors, digitizing machine and scanners. Now, the data which has been acquired must be brought to the next stage called the data analysis stage. The data analysis stage involves the in-depth learning about the data, which involves studying the various patterns available in the data per class. This information or knowledge about the data is used for further processing. The third step used for pattern recognition is the classification process. Its purpose is to decide the category of new data on the basis of knowledge received from data analysis process. The data set presented to a pattern recognition system is divided into two sets: training set and testing set. The system learns from the training set and the efficiency of the system is checked by presenting a testing set to it.

To sum up, the pattern recognition process involves data acquisition from the source, its preprocessing to churn out requisite data, extracting unique features suitably representing the characteristics features of the entire dataset, selecting a few unique representatives of the extracted features to enhance the computational process involved in the final stage called classification, and this process has been represented through Figure 1.16.

The results obtained from the pattern recognition process are the average predictive accuracy results for the test data. These results are greatly influenced by the techniques involved in feature extraction process, the number of features extracted and the classification technique used. The density of the dataset greatly hampers the computational time as well as the average predictive accuracy values.
1.4.1 Pattern recognition tasks

The process of pattern recognition involves associating the extracted patterns, mapping and grouping of the extracted patterns, finding the variations in a group of patterns etc. and all the tasks have been diagrammatically represented in Figure 1.17 as described by Yegnanarayana (1993).

- **Pattern association**: The problem of pattern association involves storing a set of patterns or a set of input-output pattern pairs in such a way that when test data are presented, the pattern or pattern pair corresponding to the data is recalled. This is purely a memory function to be performed for patterns and pattern pairs.
• **Pattern mapping:** The objective of pattern mapping is to capture the implicit relationship between the patterns and the output labels that are given, so that when a test input is given, the corresponding output pattern or the class label is retrieved.

• **Pattern grouping:** In the case of pattern grouping, given a set of patterns, the problem is to identify the subset of patterns possessing similar distinct features and group them together. Since the number of groups and the features of each group are not explicitly stated, this problem belongs to the category of unsupervised learning or pattern clustering. In pattern grouping on the other hand, patterns belonging to several groups are given, and the system has to resolve the groups.

• **Feature mapping:** In several patterns the features are not unambiguous. In fact the features vary over a continuum, and hence it is difficult to form groups of patterns having some distinct features. In such cases, it is desirable to display the feature changes in the patterns directly. This again belongs to the unsupervised learning category. In this case what is learnt is the feature map of a pattern and not the group or class to which the pattern may belong.

• **Pattern variability:** In this task, there are many situations when the features in the pattern undergo unspecified distortions each time the pattern is generated by the system. This can be easily seen in the normal handwritten cursive script. Human beings are able to recognize them due to some implicit interrelations among the features, which themselves cannot be articulated precisely. Classification of such patterns falls into the category of pattern variability task.

• **Temporal pattern:** In this task, human beings are able to capture effortlessly the dynamic features present in a sequence of patterns. This is true, for example, in speech where the changes in the resonance characteristics of the vocal tract
system (e.g. formant contours) capture the significant information about the speech message.

- **Stability plasticity dilemma:** If the system is allowed to change its categorization continuously, based on new input patterns, it cannot be used for any application such as pattern classification or clustering, as it is not stable. This is called stability plasticity dilemma in pattern recognition.

### 1.4.2 Pattern recognition techniques

The techniques adopted for the pattern recognition process are categorized depending upon the methodology adopted for the data analysis and classification. The different techniques used for pattern recognition task have been shown through Figure 1.18.

![Pattern Recognition Techniques](image)

- **Statistical pattern recognition techniques (SPRT):** According to researchers [Fazal et al.(2011); Devroye et al.(1996) and Duda et al.(1973)], statistical pattern recognition is a term used to cover all stages of an investigation from problem formulation and data collection through to discrimination and classification, assessment of results and interpretation. The process of SPRT involves understanding the features available and then representing them in terms of ‘p’ feature vectors. In this process, the whole problem can be thought as a ‘p’ dimensional space. The goal is to choose those features that allow pattern vectors...
belonging to different categories to occupy compact and disjoint regions in a p-dimensional feature space. The arrangement is said to be effective if all the patterns belonging to different class labels be separated amicably without any boundary conflict. The decision boundary separates the features belonging to different classes. In the statistical decision theoretic approach, the decision boundaries are determined by the probability distributions of the patterns belonging to each class.

According to Duda et al. (1973), there are various decision rules to determine decision boundary like, Bayes Decision Rule, Optimal Bayes Decision Rule, The Maximum Likelihood Rule, Neyman-Pearson rule and MAP rule. Depending upon whether the method opted is supervised or unsupervised, the statistical technique can be categorized as: Discriminant Analysis and Principal Component Analysis.

- **Structural pattern recognition techniques:** The object of interest may be made up of inherent structures and in that case the statistical pattern recognition technique may not provide with accurate results as it is not able capture the vital information related to the inherent substructures.

In structural approach of pattern recognition a collection of complex patterns are described by a number of sub-patterns and the grammatical rules with which these sub patterns are associated with each other. According to Fu (1974) and Pavlidis (1997), the structural pattern recognition technique is concerned with structure and attempts to recognize a pattern from its general form. The language which provides structural description of patterns in terms of pattern primitives and their composition is termed as pattern description language. If the language used has increased descriptive power, it leads to increased complexity of syntax analysis system. Structural pattern recognition assumes that pattern structure is quantifiable and extractable so that structural similarity of patterns can be assessed. This structure quantification and description are mainly done using Formal grammars and Relational descriptions (principally graphs).
• **Template matching techniques:** The term template means anything fashioned, shaped, or designed to serve as a model from which something is to be made: a model, design, plan, or outline; or it can be something formed after a model or prototype, a copy; a likeness, a similitude; or it can be an example, an instance; especially a typical model or a representative instance. The term matching means to compare in respect of similarity. The template matching technique is the simplest and most primitive amongst all pattern recognition models. It is used to determine the similarity between two samples, pixels or curves. The pattern to be recognized is matched with the stored templates while assuming that template has undergone rotational or scalar transformations. The efficiency of this model depends upon the stored templates. The correlation function is taken as a recognition function and is optimized depending on the available training set. The shortcoming of this approach is that, it does not work efficiently in the presence of distorted patterns.

According to Bajcsy et al. (1989), template matching technique, especially in two dimensional cases, has many applications in object tracking, image compression, stereo correspondence, and other computer vision applications. Out of the several matching methods available, the Normalized Cross Correlation (NCC) and square root of Sum of Square Differences have been used as the measure for computing similarity. Many important computer vision tasks can be solved with template matching techniques like object detection or recognition, object comparison and depth computation. Template matching involves techniques for image processing, the concept of probability and statistics and signal processing.

Normalized cross correlation is a well-liked method for finding 2D patterns in images. A \((2h+1) \times (2w+1)\) template \(t\) is correlated in opposition to an image \(x\). At the image location \((u, v)\), the normalized cross correlation is, given by Eq. (1.3).
\[
c(u,v) = \frac{\sum_{i=h}^b \sum_{j=w}^u X(i, j)T(i, j)}{\sqrt{\sum_{i=h}^b \sum_{j=w}^u X(i, j)^2} \sqrt{\sum_{i=h}^b \sum_{j=w}^u T(i, j)^2}}
\]

(1.3)

where

\[
X(i, j) = x(u + i, v + j) - \bar{x}
\]

\[
T(i, j) = t(h + i, w + j) - \bar{t}
\]

- **Neural network based pattern recognition techniques:** According to Joshi et al. (1997), the concept of artificial neural networks can be viewed as computing models inspired by the structure and function of the biological neural network. These models are expected to deal with problem solving in a manner different from conventional computing. A distinction is made between pattern and data to emphasize the need for developing pattern processing systems to address pattern recognition tasks. Scientists are hoping that computing models inspired by biological neural networks may provide new directions to solving problems arising in natural tasks. In particular, it is hoped that neural networks would extract the relevant features from input data and perform the pattern recognition task by learning from examples, without explicitly stating the rules for performing the task.

The neural networks are the massively parallel structures composed of “neuron” like subunits. According to Yegnanarayana (1994), the biological neurons are as shown in the Figure 1.19. The neural networks provide efficient results for classification of patterns. Its property of changing its weight iteratively and learning, gives it an edge over other techniques for recognition process. The perceptron is a primitive neuron model. It is a two layer structure. If output function of perceptron is step, then it performs classification problems, if it is linear than it perform regression problems. The most commonly used family of
neural networks for pattern classification is the feed forward networks like MLP and RBF. It includes dendrites, the cell body and a single axon. Synapses connect the axons of neurons to various parts of other neurons.

![Schematic drawing of a typical neuron or nerve cell.](image)

Figure 1.19  Schematic drawing of a typical neuron or nerve cell.

- **Fuzzy logic based pattern recognition techniques:** According to Kandel (1982), the importance of fuzzy sets in pattern recognition lies in modeling forms of uncertainty that cannot be fully understood by the use of probability theory. In a very fundamental way, the intimate relation between theory of fuzzy sets and theory of pattern recognition and classification rests on the fact that most real world classes are fuzzy in nature. Syntactic techniques are utilized when the pattern sought is related to the formal structure of language. Semantic techniques are used when fuzzy partitions of data sets are to be produced. Then a similarity measure based on weighted distance is used to obtain similarity degree between the fuzzy description of unknown shape and reference shape.

- **Hybrid techniques:** In most of the emerging applications, it is clear that a single model used for classification doesn’t behave efficiently, so multiple methods have to be combined together giving result to hybrid models. The primitive approaches to design a pattern recognition system which aims at utilizing a best individual classifier have some drawbacks. It is very difficult to identify a best
classifier unless deep prior knowledge is available at hand. The statistical and structural models can be combined together to solve hybrid problems. In such cases statistical approach is utilized to recognize pattern primitives and syntactic approach is then used for the recognition of sub-patterns and pattern itself. Fu (1976) gave the concept of attributed grammars which unifies statistical and structural pattern recognition approach. To enhance system performance one can use a set of individual classifiers and combiner to make the final decision. Tumer et al. (1996) experimentally proved that using a linear combiner or order statistics combiner minimize the variance of actual decision boundaries around the optimal boundary. Multiple classifiers can be used in several ways to enhance the system performance. Each classifier can be trained in a different region of feature space or in other way, each classifier can provide probability estimate and decision can be made upon analyzing individual results. Methods utilizing classifier ensemble design generate a set of mutually complementary classifiers that achieve optimal accuracy using a fixed decision function. Those methods which utilize combination function design tend to find an optimal combination of decisions from a set of classifiers. To achieve optimum results, a large set of combination functions of increasing complexity, ranging from simple voting rules through trainable combination functions is available to designer.

1.5 Pattern Recognition in Botanical Sciences

Every object that is perceivable by the human eye has its own characteristic features. These characteristics features get stored in the brain. The human brain remembers these characteristic cues for future reference and recognizes the object in its next occurrence. The pattern recognition is the study of how machines can observe the environment, learn to distinguish patterns of interest from their background, and make sound and reasonable decisions about the categories of the patterns. In engineering and scientific disciplines such as biology, psychology, medicine, marketing, computer vision, artificial intelligence, and remote sensing, the process of automatic (machine) recognition, description, classification, and grouping of patterns have an important role
to play. The term pattern here refers to a fingerprint image, a handwritten cursive word, a human face, or a speech signal, portion of the plant leaf or flower etc. For any given pattern, the classification process involves comparing the given pattern understudy with the existing patterns and then labelling it; on the other hand if the classes are not known then the given pattern could be clustered along with the unknown class.

Since the ages, the sustenance of human life on this planet earth depends upon the natural vegetation. The vegetation provides food, shelter and clothing. There has been a constant effort in knowing the vegetation since ages and pattern recognition techniques are the recent human efforts in knowing the plants in an automatic manner. The human efforts in botanical sciences through the use of pattern recognition science is commendable. Therefore, the role of pattern recognition in botanical sciences can be grouped in two ways: one involving human efforts alone and the other involving the machine approach for recognition and the same has been shown in Figure 1.20.

### 1.5.1 Human approach

The human problem solving is more or less a pattern processing problem and not a data processing problem. In any pattern recognition task humans perceive patterns in the input data and manipulate the pattern directly. The main difference between human and machine intelligence comes from the fact that humans perceive everything as a pattern, whereas for a machine all are data. Even in routine data consisting of integer numbers (like telephone numbers, bank account numbers, car numbers), human beings
tend to see a pattern in the series of numbers. The data is recalled as a stored set of patterns. If there is no pattern, then it is very difficult for a human being to remember and reproduce the data later. Thus storage and recall operations in humans and machines are performed by different mechanisms. The pattern nature in storage and recall automatically gives robustness and fault tolerance for a human system. Moreover, typically far fewer patterns than the estimated capacity of human memory systems are stored, which makes the brain a unique asset.

The human beings are capable of making mental patterns in their brain system from the input data given in the form of numbers, text, pictures, sounds etc., using their sensory mechanisms of vision, sound, touch, smell and taste. These mental patterns are formed even when the data are noisy, or deformed due to variations such as translation, rotation and scaling. The patterns are also formed from a temporal sequence of data as in the case of speech and motion pictures. Humans have the ability to recall the stored patterns even when the input information is noisy or partial (incomplete) or mixed with information pertaining to other patterns. This ability makes the human beings unique. In the case of recognition of plants human beings observe the external features like flowers, leaves, stem, roots and fruits. The human eye observes the placement of leaves, types of leaves, types of seeds and flowers. The human approach to botanical science has been described in this section.

The human approach involves studying the plant morphology which involves thoroughly studying the following parts viz.: Vegetative parts which includes the leaves, stems and roots; Flower their types and parts; the seed morphology which involves the seeds and seedlings; and finally the fruits and their types. Leaves are the powerhouse of plants. In most plants, leaves are the major site of food production for the plant.

A leaf is an expanded, photosynthetic organ (blade or lamina) attached to the stem with a petiole just below a bud. This attachment location where the bud is located is called a node and the space between nodes is called an internode. The details about a leaf have been discussed in the next subsections as described by Geneve (2008).
1.5.1.1 Parts of leaves

Most leaves have two main parts: the blade and the petiole, or leafstalk. The leaves of some kinds of plants also have a third part, called the stipules as shown in the Figure 1.21 which shows the parts of the leaves of a plant.

**The Blade**, or lamina, is the broad, flat part of the leaf. Photosynthesis occurs in the blade, which has many green food-making cells. Leaf blades differ from one another in several ways: the types of edges, the patterns of the veins, and the number of blades per leaf.

![Figure 1.21 Parts of leaves](image)

**The Petiole** is the stem like part of the leaf that joins the blade to the stem. Within a petiole are tiny tubes that connect with the veins in the blade. Some of the tubes carry water into the leaf. Others carry away food that the leaf has made. In many trees and shrubs, the petioles bend in such a way that the blades receive the most sunlight, thus assuring that few leaves are shaded by other leaves. The petiole also provides a flexible "handle" that enables the blade to twist in the wind and so avoid damage as shown in Figures 1.21 (a) and 1.21 (b).

In some plants, the petioles are much larger than the stems to which they are attached. For example, the parts we eat of celery and rhubarb plants are petioles. In contrast, the leaves of some soft-stemmed plants, particularly grasses, have no petioles.
The Stipules are two small flaps that grow at the base of the petiole of some plants. In some plants, the stipules grow quickly, enclosing and protecting the young blade as it develops. Some stipules, such as those of willows and certain cherry trees, produce substances that prevent insects from attacking the developing leaf. In many plants the stipules drop off after the blade has developed, but garden peas and a few other kinds of plants have large stipules that serve as an extra food-producing part of the leaf.

1.5.1.2 Types of leaves

- **Simple leaves**: Simple leaves have a single leaf blade (lamina) attached to a petiole. They may be entire (not lobed) or lobed. Entire leaves consist of a petiole and leaf blade (lamina) without any lobes or leaflets as shown in the Figure 1.22 (a). A lobed leaf has the leaf blade (lamina) with rounded or pointed lobes separated by a sinuses. Lobes may be deep or shallow depending on the depth of the sinuses as shown in the Figure 1.22 (b).

![Figure 1.22 Types of simple leaves](image)

- **Compound leaves**: A compound leaf is divided into two or more leaflets attached to a main petiole axis. The main types of compound leaves are variations of the pinnate form with leaflets on opposite sides of a central axis (rachis) and
palmate where the leaflets attach at a single location at the apex of the petiole. There are six varieties of these compound leaves as shown in the Figure 1.23.

![Types of compound leaves](image)

Figure 1.23 Types of compound leaves

Ternate (Trifoliate) leaf, three leaflets arise from a central location at the petiole. A biternate leaf is doubly ternate. The leaf divided into three sets of trifoliates each with their own three leaflets. In a plamate leaf, the leaflets arise from a single location at the petiole. Leaflet number can vary between 5 and 9. A bipinnate leaf is a twice branched pinnate leaf. There is a central rachis the branches to form rachilla (subbranches) where each leaflet arises. In a tripinnate leaf, the pinnate leaf is divided three times. In pines, leaves are usually produced in bundles (fascicles) of 2, 3, or 5.

1.5.1.3 Shapes of leaves

The different varieties of leaf shapes have been shown in the Figure 1.24. Acerose (Needle-like) leaves are found in many conifers, including pines, spruce, and firs. The leaves may be grouped in bundles as in pines or attached singly. Linear leaves are thin and elongated with parallel sides and no lobes. An oblong-shaped leaf is an elongated, non-lobed leaf that is at least twice as long as wide. A Gladiate and Ensiform leaf is an elongated, non-lobed leaf that is sword-shaped. Falcate refers to
a leaf that is hooked or curled in shape. A lanceolate leaf is an elongated, non-lobed leaf that is lance-shaped. The widest part of the blade is below the middle of the leaf. Spatulate refers to spoon-shaped leaf that has a rounded top and a tapering base. An oblanceolate leaf is the inverse of being lanceolate (lance-shaped) with the attachment to the petiole being narrower than the top. In an oval-shaped leaf, the middle portion of the leaf is wider than half the leaf’s length. An elliptic-shaped leaf has a narrow oval shape with the center of the leaf boarder than the ends. Each end tapers and are approximately the same size. An ovate or egg-shaped leaf blade is narrower at the top than the base. It is the inverse of an obovate leaf. Obovate is an egg-shaped leaf blade that is broader at the top than the base. It is the inverse of ovate. In an Orbicular leaf, the leaf blade is a near circle in outline. In a cordate leaf, the blade is heart-shaped with the wide part of the blade attached at the base. It is the inverse of obcordate. Oxcordate is a heart-shaped leaf, but with the narrow part of the leaf at the base. It is the inverse of a cordate leaf. In case of reniform, the leaf blade is kidney-shaped. A Deltoid is a triangular-shaped leaf. A rhomboid leaf is diamond-shaped. Hastate refers to an arrow-shaped leaf with lobes that point outward. Sagittate refers to an arrow-shaped leaf with lobes that point downward. Flabellate is a fan-shaped leaf. In a peltate leaf, the petiole is attached on the back (abaxial) surface of the leaf blade rather than along the blade edge. Subulate refers to modified leaves found in some conifers. These are often referred to as awl-shaped or scale-like.

Figure 1.24  Different shapes of plant leaves
1.5.1.4 Leaf phyllotaxy

The spiral arrangement of leaves on the stem is called phyllotaxy. Leaves spiral around the stem in a regular pattern. It can be described as a fraction with the numerator being the number of turns around the stem and the denominator being the number of leaves it takes to return to original leaf position.

The vertical rank of leaves in a spiral are said to have the same orthostichy. Leaf Arrangements Smoke tree (Cotinus) Leaves 0 and 8 have the same orthostichy after three spirals around the axis (3/8). Phyllotaxy indices (like many other spirals in nature) follow a Fibonacci series. The Fibonacci series consists of a series of numbers that remarkably describe many spirals observed in nature. The previous pair of numbers sum to the next number in the series. The series is 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89 …. Phyllotaxy follows a numerator and denominator set of offset Fibonacci numbers. 1/2, 1/3, 2/5, 3/8, 5/13, 8/21, 13/34 …

The Figure 1.25 shows the different leaf phyllotaxy with the areolas and spines (modified leaves) of the cactus Mammillaria, there is actually one set of spirals in a clockwise direction and one set in a counter-clockwise direction. If you start to look for these types of patterns, you will see the clockwise and counter-clockwise spirals in many different plant structures. Phyllotaxy can also be expressed by these sets of opposing spirals. Each set of spirals is called a parastichies. Each parastichies occurs at a regular interval, but opposite spirals occur in a different number and a different interval.

Figure 1.25   Leaf Phylotaxy
1.5.1.5 Arrangements of leaves

There are 7 different types of arrangements as shown in the Figure 1.26.

- **Alternate- single leaf per node**: In the alternate arrangement, a single leaf is attached at each node. The leaves may be arranged in straight rows or spiral around the stem. The spiral arrangement of leaves on the stem is called phyllotaxy. There are a few plants that have leaves that occur only on one side of the stem. Usually the leaves are held on a gently curved stem. This pattern is called monostichous. Traveler’s palm (Ravenala madagascariensis) has a spectacular distichous leaf form. Several types of succulent plants display distichous leaves.

- **Opposite – two leaves per node**: Opposite is where a pair of leaves occur at each node. Subopposite is when some of the leaves are not perfectly paired but the distance between the nodes are not far enough apart to be considered alternate in arrangement.

- **Whorled**: Three or more leaves per node In the whorled arrangement, three or more leaves arising from the same node to form whorls of leaves along the stem.
- **Decussate**: Decussate leaves are arranged opposite at each node, but each pair of leaves is oriented at right angles to the pair at the next node.

- **Basal**: In a basal arrangement, all the leaves arise from the base (crown) of the plant.

- **Equitant**: Equitant leaves are overlapping as is typical in some Iris.

- **Rosette**: Rosette leaves arranged in a dense, radiating cluster. Rosettes usually form near the base of the plant.

### 1.5.1.6 Leaf tips

There are 14 different varieties of Leaf tips as shown in the Figure 1.27. Acute type of leaf tip is gradually tapering to a point. In case of acuminate type of leaf tip, the leaf blade has rounded shoulders leading to a pointed tip. In case of apiculate leaf tip, the leaf blade ends in a short tip. The caudate tip resembles a tail. The Cuspidate tip forms a short, narrow point (cusp). The cirrhose or cirrose tip forms a tendril. The aristate tip has a bristle at the tip. The mucronate tip forms a short, sharp point. The obtuse tip is a narrow, rounded tip. In case of rounded leaf tip, the leaf blade is rounded at the tip. In case of obcordate, the apex of the leaf blade has a deep notch. In case of retuse leaf tip, the apex of the leaf blade has a shallow notch. The mucronulate tip is a sharp point called a mucro. In case of truncate type of leaf tip, the leaf blade is square at the tip.

### 1.5.1.7 Leaf bases

There are 10 different types of leaf bases as shown in the Figure 1.28. An attenuate leaf blade gradually tapers to a narrow base. A cuneate leaf blade tapers to a narrow wedge-shaped base. An acute leaf blade tapers to a sharp triangular-shaped base. A
A cordate leaf base has a heart-shaped blade with a gently lobed base. A rounded leaf blade has a rounded base. In an oblique leaf base each side of the leaf blade attaches at a different point on the petiole. A truncate leaf blade is square at the base. An auriculate leaf blade has lobes resembling ears. A hastate leaf base occurs in arrow-shaped leaves with lobes that taper away from the petiole. Sagittate leaf bases occur in arrow-shaped leaves with lobes that taper downward.

Figure 1.27 Different varieties of leaf tips

Figure 1.28 Different types of leaf bases
1.5.1.8 Leaf margins

There are 20 different varieties of leaf margins as shown in the Figure 1.29. Entire leaves have a continuous margin without lobes, teeth or spines. Crisped is a wavy, crinkled margin. Sinuate is a wavy margin that borders on being lobed. Undulate is a wavy margin that is not as deeply notched as a sinuate margin. Lobed is a margin with distinct lobes with sinuses that are less than half way to the midrib. Pinnatifid is a margin with deep lobes with sinuses that extend more than half way to the midrib. Pinnatisect is a margin the extends all the way to the midrib. Palmatifid or Palmatisect is a palmately lobed leaf, while a palmatisect leaf is palmately divided to the midrib or petiole creating leaflets. Crenate is a scalloped margin with small rounded teeth. Serrate is a common margin type. The small, sharp teeth point upward. The teeth may be the same size (serrate) or vary in size (doubly serrate or erose). Serrulate is a margin that is composed of tiny serrate teeth. A dentate margin has teeth along the margin that point outward rather than forward as seen in a serrate margin. Ciliate is a leaf margin that contains small hairs. Spinose is having spines on the leaf margin. Runcinate is where the lobes are deeply cut and point downward. Incised leaves have a margin with deep irregularly-shaped lobes. Laciniate has a margin that is cut into narrow, irregular lobes. Dissected is a deeply divided leaf with irregular lobe margins.

![Figure 1.29 Different types of leaf margins](image)
1.5.1.9 Leaf venation

There are 6 different varieties of leaf venation patterns as shown in Figure 1.30. In pinnate leaf venation, the veins are produced on either side of the central main vein (midrib), which extends from the petiole to the leaf edge. In plamate venation, the main veins radiate from a central point at the petiole. Each main vein extends from the petiole to the tip of a lobe. Reticulate or net-veined leaves have many branched minor veins. Dichotomous venation occurs in Gingko. Numerous veins radiate from the base of the leaf that branch near the upper leaf surface to form a Y. Leaves with parallel venation are diagnostic for monocots. The veins extend parallel to the outer leaf edge and each other. In triplinerved venation, there are two lateral veins that branch from the midrib in the upper portion of the leaf blade.

![Figure 1.30: Different types of leaf venation patterns](image)

1.5.1.10 Leaf attachment

There are 7 different types of leaf attachments as shown in the Figure 1.31. Petiolate leaves are attached to the stem by a simple petiole. In sessile type, the leaf blade is attached to the stem without a petiole. Stipules are leaf-like appendages that occur at the base of the petiole. When they are present the attachment is termed stipulate. Clasping is when the base of the leaf blade wraps around the stem. When the leaf
base only wraps around the stem, it is called amplexicaul. When the leaf base or petiole extends down the stem, it is termed decurrent. Sheathing is where the leaf base or the stipule forms a sheath that surrounds the stem. Connate is where the bases of two opposite leaves fuse around the stem. Perfoliate is where the leaf surrounds the stem, which passes through the base of the leaf.

![Different types of leaf attachments](image)

**Figure 1.31 Different types of leaf attachments**

### 1.5.2 Machine learning approach

The process of machine learning involves the extraction of features from the images and then refine the extracted parts/components for further processing. Figure 1.32 shows the techniques for feature extraction.

![Types of features and their extraction methodologies](image)

**Figure 1.32 Types of features and their extraction methodologies**
1.5.2.1 Types of features

Feature extraction is a general term for methods of constructing combinations of the variables to get around the problems while still describing the data with sufficient accuracy. Feature extraction involves the image features to a distinguishable extent.

- **Color feature extraction:** According to Kodituwakku et al. (2010), color is one of the most important features of images. Color features are defined subject to a particular color space or model. There are a number of color spaces that can be, and they are: RGB, LUV, HSV and HMMD. Once the color space is specified, color feature can be extracted from images or regions. There are a number of important color features have been used frequently in many research practices and they are: color histogram, color moments (CM), color coherence vector (CCV) and color correlogram, average RGB etc. Among them, CM is one of the simplest yet very effective features. Usually they are calculated for each color channels (components) separately. There-fore, nine features form the feature vector. These features are useful when they are calculated for region or object. However, the moments are not enough to represent all the color information of an image. The common moments are mean, standard deviation and skewness, the corresponding calculation can be defined through Eqs. (1.4), (1.5) and (1.6).

\[
\mu_i = \frac{1}{N} \sum_{j=1}^{N} f_{ij}
\]  
\[\sigma_i = \left( \frac{1}{N} \sum_{j=1}^{N} (f_{ij} - \mu_j)^2 \right)^{1/2} \]
\[\gamma_i = \left( \frac{1}{N} \sum_{j=1}^{N} (f_{ij} - \mu_j)^3 \right)^{1/3} \]

Where \(f_{ij}\) is the color value of the \(i\)-th color component of the \(j\)-th image pixel and \(N\) is the total number of pixels in the image. \(\mu_i, \sigma_i, \gamma_i\) \((i=1,2,3)\) denote the
mean, standard deviation and skewness of each channel of an image respectively. The color histogram describes the color distribution of an image. It quantizes a color space into different bins and counts the frequency of pixels belonging to each color bin. This feature is robust to translation and rotation changes. However, a color histogram does not tell about pixels’ spatial information. Therefore, visually different images can have similar color histograms. In addition, the dimension of a histogram is usually very high.

The CCV incorporates spatial information into the basic color histogram. It divides each histogram bin into two components: coherent and non-coherent parts. The coherent component includes those pixels which are spatially connected. The non-coherent component includes those pixels that are isolated. As CCV captures spatial information, it usually performs better than a color histogram. However, the dimension of a CCV is twice of a conventional histogram.

According to Zhang et al.(2012), a color correlogram is the color version of grey level co-occurrence matrix (GLCM). It characterizes the distribution of color pairs in an image. A color correlogram can be treated as a 3D histogram where the first two dimensions represent the colors of any pixel pair and the third dimension is their spatial distance. Thus, in a correlogram, each bin \((i, j, k)\) represents the number of color pair \((i, j)\) at a distance \(k\). The color correlogram is calculated for horizontal distance. Correlograms for other distances can be similarly calculated. The other popular standard features are used for MPEG7 and they are: dominant color descriptor (DCD), color layout descriptor (CLD), color structure descriptor (CSD), and scalable color descriptor (SCD).

- **Texture features:** According to Materka et al. (1998), texture is a very useful characterization for a wide range of image. It is generally believed that human visual systems use texture for recognition and interpretation. In general, color is usually a pixel property while texture can only be measured from a group of
pixels. In texture based techniques, there are two methods: Spatial Texture Feature extraction methods and Spectral Texture feature extraction methods.

In spatial approach, texture features are extracted by computing the pixel statistics or finding the local pixel structures in original image domain. The spatial texture feature extraction techniques can be further classified as structural, statistical and model based.

Structural techniques describe textures using a set of texture primitives (texon or texture elements) and their placement rules. Textons are organized into a string descriptor and syntactical pattern recognition techniques are used to find similarity of two descriptors.

Statistical texture feature characterizes texture as a measure of low level statistics of grey level images. The common spatial domain statistical features are moments, Tamura texture features and features derived from grey level co-occurrence matrix (GLCM).

In model based techniques, texture is interpreted using stochastic (random) or generative models. Model parameters characterize the underlying texture property of the image. Popular texture models are Markov random field (MRF), simultaneous auto-regressive (SAR) mode, fractal dimension (FD) etc. As these models involve optimization, they are usually computationally expensive.

In spectral texture feature extraction techniques, an image is transformed into frequency domain and then feature is calculated from the transformed image. The common spectral techniques include Fourier transform (FT), discrete cosine transform (DCT), wavelet, and Gabor filters. FT and DCT are very fast to compute but are not scale and rotation invariant. Wavelet is both efficient and robust, but it only captures horizontal and vertical features. Among them, Gabor features are most robust because it captures image features in multi-orientations and multi-scales.
As the most common method for texture feature extraction, Gabor filter has been widely used in image texture feature extraction. To be specific, Gabor filter is designed to sample the entire frequency domain of an image by characterizing the center frequency and orientation parameters. The image is filtered with a bank of Gabor filters or Gabor wavelets of different preferred spatial frequencies and orientations. Each wavelet captures energy at a specific frequency and direction which provide a localized frequency as a feature vector. Thus, texture features can be extracted from this group of energy distributions. Given an input image \( I(x, y) \), Gabor wavelet transform convolves \( I(x, y) \) with a set of Gabor filters of different spatial frequencies and orientations. A two-dimensional Gabor function \( g(x, y) \) can be defined through Eq. (1.7).

\[
g(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp \left[ -\frac{1}{2} \left( \frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right) \right] + 2\pi j W_x
\]  

(1.7)

Where \( \sigma_x \) and \( \sigma_y \) are the scaling parameters of the filter (the standard deviations of the Gaussian envelopes), \( W \) is the center frequency, and \( \theta \) determines the orientation of the filter.

- **Shape features**: According to Yang et al. (2008), shape is known as an important cue for human beings to identify and recognize the real-world objects, whose purpose is to encode simple geometrical forms such as straight lines in different directions.

Shape feature extraction techniques can be broadly classified into two groups, viz., contour based and region based methods. Region-based methods use the whole area of an object for shape description, while contour-based methods use only the information present in the contour of an object. Local space-time features have recently become a popular representation for action recognition and visual detection. Local space-time features capture characteristic salient and motion patterns in video and provide relatively independent representation of events with respect to their spatio-temporal shifts and scales as well as
background clutter and multiple motions in the scene. Figure 1.33 shows the various techniques for extracting shape features.

Figure 1.33 Various shape description techniques
1.5.2.2 Types of machine learning algorithms

In machine learning, the algorithms are often grouped by similarity in terms of their function. The machine learning approach as described by Brownlee (2013) uses several algorithms and they have been grouped into different categories as depicted in Figure 1.34.

- **Regression based algorithm**: Regression is concerned with modelling the relationship between variables that is iteratively refined using a measure of error in the predictions made by the model. Regression methods are a workhorse of statistics and have been cooped into statistical machine learning. This may be confusing because we can use regression to refer to the class of problem and the class of algorithm. Really, regression is a process. The most popular regression algorithms are: Ordinary Least Squares Regression (OLSR), Linear Regression, Logistic Regression, Stepwise Regression, Multivariate Adaptive Regression Splines (MARS), Locally Estimated Scatterplot Smoothing (LOESS).

- **Instance based algorithm**: Instance based learning model a decision problem with instances or examples of training data that are deemed important or required to the model. Such methods typically build up a database of example data and compare new data to the database using a similarity measure in order to find the best match and make a prediction. For this reason, instance-based methods are also called winner-take-all methods and memory-based learning. Focus is put on representation of the stored instances and similarity measures used between instances. The most popular instance-based algorithms are: k-Nearest Neighbour (kNN), Learning Vector Quantization (LVQ), Self-Organizing Map (SOM), Locally Weighted Learning (LWL).

- **Regularization algorithm**: An extension made to another method (typically regression methods) that penalizes models based on their complexity, favoring simpler models that are also better at generalizing. The regularization
algorithms have been separately mentioned here because they are popular, powerful and generally simple modifications made to other methods. The most popular regularization algorithms are: Ridge Regression, Least Absolute Shrinkage and Selection Operator (LASSO), Elastic Net, Least-Angle Regression (LARS).

- **Decision tree based algorithm:** Decision tree methods construct a model of decisions made based on actual values of attributes in the data. Decisions fork in tree structures until a prediction decision is made for a given record. Decision trees are trained on data for classification and regression problems. Decision trees are often fast and accurate and a big favorite in machine learning. The most popular decision tree algorithms are: Classification and Regression Tree (CART), Iterative Dichotomiser 3 (ID3), C4.5 and C5.0 (different versions of a powerful approach), Chi-squared Automatic Interaction Detection (CHAID), Decision Stump, Conditional Decision Trees.

- **Bayesian algorithm:** Bayesian methods are those that are explicitly apply Bayes’ Theorem for problems such as classification and regression. The most popular Bayesian algorithms are: Naive Bayes, Gaussian Naive Bayes, Multinomial Naive Bayes, Averaged One-Dependence Estimators (AODE), Bayesian Belief Network (BBN), Bayesian Network (BN).

- **Clustering algorithm:** Clustering, like regression describes the class of problem and the class of methods. Clustering methods are typically organized by the modelling approaches such as centroid-based and hierarchal. All methods are concerned with using the inherent structures in the data to best organize the data into groups of maximum commonality. The most popular clustering algorithms are: k-Means, k-Medians, Expectation Maximisation (EM), Hierarchical Clustering.

- **Association rule learning algorithm:** Association rule learning are methods that extract rules that best explain observed relationships between variables in data. These rules can discover important and commercially useful associations in large
multidimensional datasets that can be exploited by an organisation. The most popular association rule learning algorithms are: Apriori algorithm, Eclat algorithm.

Figure 1.34 Different types of machine learning algorithms
• **ANN based algorithm:** Artificial Neural Networks are models that are inspired by the structure and/or function of biological neural networks. They are a class of pattern matching that are commonly used for regression and classification problems but are really an enormous subfield comprised of hundreds of algorithms and variations for all manner of problem types. The deep learning algorithms have been separated from neural networks because of the massive growth and popularity in the field. Here the major concern is with the more classical methods. The most popular artificial neural network algorithms are: Perceptron, Back-Propagation, Hopfield Network, Radial Basis Function Network (RBFN).

• **Deep learning algorithm:** They are concerned with building much larger and more complex neural networks, and as commented above, many methods are concerned with semi-supervised learning problems where large datasets contain very little labelled data. The most popular deep learning algorithms are: Deep Boltzmann Machine (DBM), Deep Belief Networks (DBN), Convolutional Neural Network (CNN), Stacked Auto-Encoders.

• **Dimensionality reduction algorithms:** Like clustering methods, dimensionality reduction seek and exploit the inherent structure in the data, but in this case in an unsupervised manner or order to summarize or describe data using less information. This can be useful to visualize dimensional data or to simplify data which can then be used in a supervised learning method. Many of these methods can be adapted for use in classification and regression: Principal Component Analysis (PCA), Principal Component Regression (PCR), Partial Least Squares Regression (PLSR), Sammon Mapping, Multidimensional Scaling (MDS), Projection Pursuit, Linear Discriminant Analysis (LDA), Mixture Discriminant Analysis (MDA), Quadratic Discriminant Analysis (QDA), Flexible Discriminant Analysis (FDA).
1.6 Thesis Outline

**Chapter 1** is an overview about the introduction to pattern recognition, its important tasks and techniques adopted, role pattern recognition in botanical sciences and the latest state of art instruments available for computer vision techniques.

**Chapter 2** is a detailed literature survey for application of computer vision based techniques in plant discrimination. It is essential to delve deeply into the topic of interest in a particular subject under study, there is a need to read, analyze, evaluate and summarize the scholarly material. The result of this literature review is a report that serves the part of this thesis. The research survey includes research work carried out by different researchers from 1973 to 2015. This detailed literature survey provide us deep insight of the present status of the work and what are the limitations and open challenges.

In **Chapter 3**, presents the results of our study to see the effects of feature selection algorithm on the predictive classification accuracy of algorithms used for discriminating the different plant leaf images. The process involves extracting the important texture features from the digital images and then subjecting them to feature selection and further classification process. The leaf image features have been extracted by using Gabor filter and these Gabor features are subjected to feature selection algorithm for extracting important texture features.

In **Chapter 4**, describes an alternative for the dorsal leaf image classification. As of now, only the dorsal side of the leaf image is considered for this purpose. This work proposes to utilize the texture features available on the ventral side of the leaf image for classification purpose using Gabor filter based texture features.

**Chapter 5**, proposes to improve the classification accuracy of the leaf images by extracting texture and statistical features by utilizing the presence of striking features on the dorsal and ventral sides of the leaves, which on other types of objects may not be that prominent. The texture features have been extracted from dorsal, ventral and a
combination of dorsal-ventral sides of leaf images using Gray level co-occurrence matrix. In addition to this, this chapter also uses certain general statistical features for discriminating them into various classes. The feature selection work has been performed separately for the dorsal, ventral and combined data sets (for both texture and statistical features) using the most common feature selection algorithms. After selecting the relevant features, the classification has been done using the classification algorithms. The classification accuracy has been calculated and compared to find which side of the leaf image (dorsal or ventral) gives better results with which type of features (texture or statistical).

In Chapter 6, presents the strategy and result of our study of the statistical feature set obtained from a digital leaf image for its dorsal and ventral sides and to estimate its effectiveness in accessing its discriminating power in classifying plant species. This chapter further delves into finding the effect of fusion of two different kind of features (statistical and directional) and then subjecting them to classification process. The chapter also studies whether the ventral side of the digital leaf image can be a suitable alternative for classification of leaf image data set or not.

In Chapter 7, we present the result of our work to identify a best set of features from the complete feature set, by applying PSO-CFS based feature selection optimization technique on the texture features extracted through Gabor based technique. This chapter further comparatively summarizes the effectiveness of ventral features over the dorsal ones in discriminating the plant species.

The Chapter 8 summarizes the major output of the present research work with its future scope in the field of plant discrimination using digital images of the leaves.