RESEARCH ARTICLE

LEAF CLASSIFICATION USING IMAGE PROCESSING: A REVIEW.

Aby Joseph and Dr. Anita H. B.

Dept. of Computer Science, Christ University, Bengaluru, India.

Manuscript Info	*Abstract*
	Earth consists of unicellular and multicellular organisms. Unicellular organisms like bacterium and multi cellular organisms like flora and fauna. Since plant species cannot survive on their own from artificial disasters, it is the responsibility of human beings to protect plant species from such calamities. In order to meet that goal, it is necessary to classify the different plant species and prevent them from extinction. Even though there are huge numbers of databases existing for plant species, there are some species which are still unknown to humans. So it is essential to distinguish the different plant species and classify them into proper categories and protect them. Earlier, manual databases were there for getting this job done. But as technology progressed, automation has its effect in almost all the fields. Classification of plant species through automation is possible through methods like leaf recognition technique. Leaf recognition has its own advantages; apart from the classification plant species, it also helps in recognizing various diseases which can result in their extinction. Much research works happened in the field of leaf recognition. Most of the researchers used extraction techniques involving spatial or frequency domain. Only few researchers used features extracted using both frequency and spatial domain. In this paper we are trying to come up with a reason for less researches being conducted in the field of leaf database in India.

Introduction:-

Earth is a huge planet consisting of number of known and unknown species. It is necessary to stand by each other and co-operate with each other and share the resources available in order to lead a happy and prosperous life. Unnecessarily interfering in each other's space may lead to extinction of particular species. [1]Moreover plant species has some special abilities which benefits both flora and fauna. The oxygen has to be circulated in cycles in order to balance the natural consumption. So the cycle is maintained from time to time for the purpose of balancing the life process. The temperature level of atmosphere is also balanced by plant species by the process of transpiration. Leaf recognition is very useful for biodiversity.

Classification is very important when it comes to taxonomy. As a result, such categories of species can be grouped or clustered to a particular location like a park or a dedicated garden. [2]Identification of plant species has many advantages including identification of medicinal plants used in Ayurveda as well as poisonous plants. This process is tedious because human senses are limited and hence identification becomes very difficult especially when there are

Corresponding Author:- Aby Joseph.
Address:- Dept. of Computer Science, Christ University, Bengaluru, India.

large number of species. So it is necessary to train the computer system using algorithms and develop a system in such a way that it can identify the species by capturing the images of different parts of the plants.

Leaf recognition technique works by identifying the patterns present on the leaf surface and based on the patterns the system can identify the species of the plants. The categorization of both medicinal and poisonous species of plants will help to conserve the medicinal species under extinction and keep away from poisonous plant species and thereby saving lives.

Leaf recognition is based on different characters present on the leaf surface. This characters will determine the category to which a particular plant species belong to. The leaf recognition technique has been followed by botanists to categorize particular plant species. Leaf recognition is limited to factors like climate and plant health because the external surface of the leaf can have change in color due to diseases or climate changes. This makes it even more difficult to proceed with the process.

The climate of a particular region is a key factor which determines the leaf characteristics. The plant has to adopt to different climatic conditions. Thus the plant structure has to be in tune with the particular region in which it is growing. In tropical regions there is high availability of rains and moist winds. Thus the barks and other parts will be thin and the leaves should be capable of enduring strong winds and heavy rains. Desert plants are thornier in nature because loss water should be prevented. As compared to tropical and desert region of the world, databases on Indian leaf species and South Indian in particular are limited to very small numbers. So my intention is to bring this crisis under attention of the researchers. So that, in future more databases will be generated in South Indian regions. This will in turn help the future researchers to do more study on South Indian leaf classification. Unlike, developed countries, India has limited research facility and resources for students and researches to work upon. As young researchers require more references, more databases should be made available for further studies. Thus more and more researchers will come forward to invest time and energy in this field of research.

Literature Review:-
There is lot of surveys conducted on leaf recognition from a very long time. Most of the works are concentrated on multiple features of same plant. As time progressed the size of the database started growing and the number of plant species known to mankind also increased considerably. As the number of surveys increases the size of plant database will also increase. The increase in information will result in increase of knowledge regarding plant species. Discovery of more medicinal herbs can cure day to day changing human conditions. Discovery of more poisonous plant species can save lives as human beings and animals can keep away from them. The gap between man and nature can be bridged and reduced day by day.

[3] As part of his contribution, has concentrated on fruit/leaf analysis. He used the technique of wavelet based analysis for classifying the obtained images. This classification helps in identifying the damage caused by pests on fruit/leaf without the involvement of human beings by manually inspecting the fruit/leaf. The method suggested by him has more than 90% accuracy. The method involves neural network for classifying the diseases or damages caused by pests. His method was successful using neural network up to a great extent. Since time consumption is also reduced due to limited human interference in the process.

[4] Proposed a method which is less complicated yet effective in terms of computation using leaf recognition. This method was dedicated for plants with flat leaves which were two dimensional in nature. The base point on the leaf and reference point were decided by the user and with reference to these base points the leaf shape is generated followed by a binary image. Set of features are collected by segmenting the leaves across major axis at the same time segmented across parallel minor axis. The next step is to normalize the feature points. These inputs are taken to neural network.

[5] Here twelve features are used for the purpose of classification. In order to have a faster processing the dimensions are reduced i.e., twelve dimensions are reduced into five dimensions. Probabilistic Neural Network (PNN) technique was used for the above mentioned purpose. Features including leaf area, psychological width, smooth processing, psychological length, perimeter ratio factor, aspect ratio, narrow factor, rectangularity, form factor, perimeter ratio diameter, perimeter ratio of psychological length and psychological width, vein features. Even though there are different sizes of leaves of same plant a standard size should be used, otherwise features of the leaf may vary.

[6] Developed a method for identifying 60 types of leaves. The identification was based on color patterns on the leaves. Even if the plant leaves have the same shape the color on the leaf surface will be different for the leaves. Gray-level co-occurrence matrix (GLCM) and geometric features, Zernike moments were used for the identification of the leaves feature. There were two approaches used for identification. The first being distance measure and the second being Probabilistic Neural Network (PNN).

[7] The proposed technique is based on hybrid image segmentation. A histogram is developed based on the obtained image segmentation. The result of the above operation contains PSNR (peak to signal ratio) and calculation of energy mean square. Based on these two parameters the quality of segmentation is obtained.

[8] Technique of feature extraction and shape analysis is used for leaf recognition. They are given as input for the neural network. The neural network is based on the algorithm with multilayer propagation architecture. The result is ranging between 90 and 94 in its feature category. The error range is between 0-0.7. As a conclusion it is very clear that the accuracy rate is about 94%.

[9] Used 12 features of morphology in which 5 principal variables where is used. In order to classify 32 types of plants by using 1800 leaves probabilistic neural network system was used. The input is obtained using scanners or digital cameras. The following step is to select the base points from the obtained leaves.

[10] Proposed leaf classification based on ontology. Vein recognition is extracted using combined neural network approach as well as thresholding. This will result in improved detailing of leaves. Here automation of the system is carried out by machine learning techniques. For better performance, modified Fourier descriptor and curvature scale space techniques are employed. In order to identify the basic shapes, CCD code system is proposed. Since plant leaf shape is considered to be 2D, an efficient method should be devised to obtain proper characterization of plant species.

[11] Came up with radial basis probabilistic neural network which uses orthogonal least square algorithm as well as optimized recursive square algorithm. The partially damaged leaves are identified using plant recognition. The plant recognition technique is optimized using Fourier descriptors. The final result reveals RBPNN is more efficient when compared to BP neural network and radial basis function neural network.

[12] Proposed a feature extraction technique using natural image. The method comes handy for research students. For obtaining information about leaves, a multidirectional Curvelet transform is employed. The orientation of leaves is not considered, which will enhance the rate of accuracy. SVM classifier uses the coefficients for the purpose of categorization of the obtained results.

[13] Introduced move median centers (MMC) hypersphere classifier for the leaf classification. During classification, pre-processing of data occurs. While doing so it was found that not only storage space is saved but also classification time is reduced. This method is more efficient when compared to contour features because it is very difficult to spot the curvature points.

[14] The approach by them utilized PCA based surface order and complex calculation for the shape highlight delivered a consequence of 89.2% in which the researcher focused on the parameters like angle proportion, conservativeness, linearity. The morphological component extraction strategy has been utilized numerous a times.
[15] Actualized a leaf acknowledgment framework utilizing leaf vein and shape. They proposed the calculation utilizing leaf shape extraction. Projection histogram was utilized to recognize the vein and after that the components were extricated. The exploratory result demonstrated an exactness of 97.19%.

Methodology:-
The remaining parts of plants like flowers are three dimensional in nature. So the process involved in image processing will be more complex as compared to two dimensional image processing. Since the leaf surface is flat it is a two dimensional structure. Image captured can be passed through two dimensional image processing algorithms. After image processing the resultant information characteristics are compared with e-botanists database and the traits will determine the leaf category. This will reduce the human effort by automating the entire process of leaf identification using dynamic programming techniques
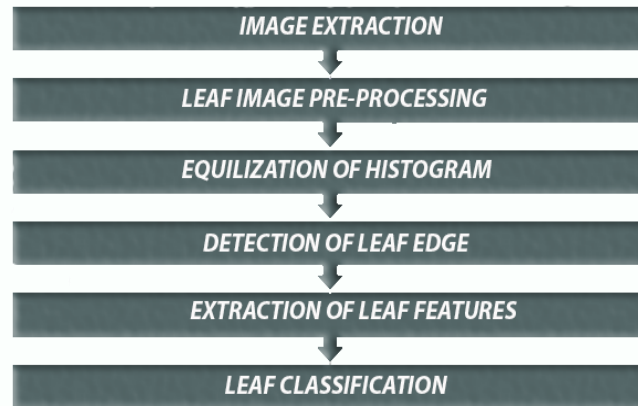
Figure 1:- Image based plant identification system

Image Processing:-

Leaf recognition process involves capturing of images. Image has to be captured using a 2D camera or a scanner. The equipment should be good enough to capture the external surface of the leaves crisp and clear. So that, veins or patterns on the leaves can be properly identified. High quality equipment is used for image capturing. Cameras capable of capturing high definition contents are used for this purpose. Once this process is over the captured image in digital format will be forwarded to the system for further processing. Here 2D cameras are used because the leaf surface is flat and at the same time the amount of details and algorithm complexity required for processing will be very less when compared to 3D image processing. Hence time complexity and space complexity can be reduced up to a great extent.

Image conversion involves the following steps:
- The captured image from the camera in RGB format.
- RGB format will be converted into Grayscale image.
- Grayscale image gets converted to Binary image.
- Next step is to smoothen the image.
- Final pre-processing step involves filtering of the obtained image after smoothening.
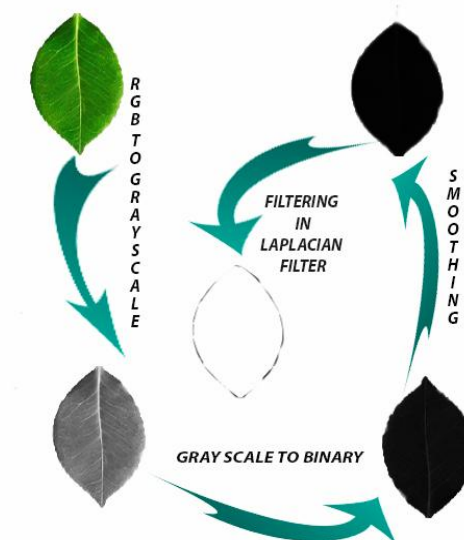

Figure 2:- Image pre-processing steps

The process of leaf recognition involves quite a few steps. The foremost step is to capture the leaf image, followed by two parallel processes. Leaf Contour extraction and Leaf Vein extraction. The following processes will determine contour-centroid distance and leaf direction determined using Projection histogram. The next process is distance sequence rearrange. There are subsequent processes for distance sequence rearrangement including FFT (it can be used for feature extraction), Feature extraction based on geometric and digital morphological features, and finally Feature extraction based on Convex Hull. Using FFT method, analysis is focusing more on frequency domain. Even though same plant can have different shapes for its leaves, the basic structure of the leaves will remain same. This makes the classification easier. After performing these complex techniques, the final result is obtained.

Conclusion:-

A number of methods were discussed in the above surveys. But very few works are available for the future researchers to work on and contribute to their research. As discussed earlier, databases for South Indian leaf species are few in numbers. So my work is to bring the attention of researchers all over the world towards the lack of availability of databases of the endangered leaf species of South India. So that they won't suffer extinction in the near future. There are wide varieties of medicinal herbs available in southern region of India which are facing extinction in the near future. So these species will also be preserved by the creation of South Indian leaf databases. With the creation of more precise methods of extraction of Indian based leaves more plant species can be protected.

Acknowledgements:-

References:-

1. A review on plant recognition and classification techniques using leaf images anantbhardwaj*1, manpreet kaur2
2. Sandeep kumar e. "leaf, colour, area and edge features based approach for identification of indian plants" indian journal of computer science and engineering (ijcse), issn : 0976-5166, vol. 3 no.3 jun-jul 2012.
3. Brendon j. woodford, nikola k. kasabov and c. howardwearing: fruit image analysis using wavelets.
4. J. hossain and m.a.amin, "leaf shape identification based plant biometrics", 13th international conference on computer and information technology (iccit), pp. 458-463, 2010.
5. Stephen gang wu, forrest sheng bao, eric you xu, yu-xuanwang, yi-fan chang and qiao-liangxiang (2007) "a leaf recognition algorithm for plant classification using probabilistic neural network", arxiv:0707.4289vl [cs.ai].
6. Abdul kadir, l.e. nugroho, a.susanto, p.insapsantosa, "experiments of zernike moments for leaf identification", journal of theoretical and applied information technology, 15 july 2012. vol. 41 no.1.
7. N.valliammal,dr.s.n.geethalakshmi .:hybrid image segmentation algorithm for leaf recognition and characterization, ©2011 ieee
8. Jyotismitachaki, ranjanparekh.: plant leaf recognition using shape based features and network classifiers,(ijacsa) international journal of advanced computer science and applications, vol. 2, no. 10, 2011.
9. S. wu, f. bao, e. xu, y. wang, y. chang, and q. xiang, "a leaf recognition algorithm for plant classification using probabilistic neural network," ieee 7th international symposium on signal processing and information technology, december 2007.
10. H. Fu, Z. Chi, D. Feng, and J. Song, "Machine learning techniques for ontology-based leaf classification," IEEE 2004 8th International Conference on Control, Automation, Robotics and Vision, Kunming, China, 2004.
11. J. Du, D. Huang, X. Wang, and X. Gu, "Shape recognition based on radial basis probabilistic neural network and application to plant species identification," in Proceedings of 2005 International Symposium of Neural Networks, ser. LNCS 3497. Springer, 2005.
12. S. Prasad, P. Kumar and R.C. Tripathi, "Plant leaf species identification using Curvelet transform", 2nd International Conference on Computer and Communication Technology (ICCCT), Pp. 646–652, 2011.
13. J.-X. Du, X.-F. Wang and G.-J. Zhang, "Leaf shape based plant species recognition," Applied Mathematics and Computation, vol. 185, 2007.
14. T.H. Reiss, "The Revised Fundamental Theorem of Moment Invariants", IEEE Trans. Pattern Anal. Machine Intell., Vol. PAMI-13, No. 8, August 1991, pp 830-834.
15. SidharthaMaître,"Moment invariant ",IEEE Preceding vol. 67, no. 4.