INTERNATIONAL JOURNAL OF RESEARCH CULTURE SOCIETY ISSN: 2456-6683 Volume - 4, Issue - 4, Apr – 2020
Monthly, Peer-Reviewed, Refereed, Indexed Journal Scientific Journal Impact Factor: 5.245
Received on : 04/04/2020 Accepted on : 17/04/2020 Publication Date: 30/04/2020

STUDY AN IMAGE PROCESSING BASED APPROACH TO IDENTIFICATION OF PLANT DISEASE USING DATA PROCESSING TECHNIQUES

Dr. R. Balu
Faculty
Department of Computer Applications, School of Computer Science and Engineering,
Bharathiar University, Tamil Nadu, India
Email – rvkbalu@yahoo.co.in

Abstract: *In this research paper is an image processing based approach is proposed and used for plant disease detection. Author test our results on five diseases which effect on the plants; they are: Early scorch, Cottony mold, ashen mold, late scorch, tiny whiteness. The proposed approach is image processing based and is much supported K-Means clustering technique and Artificial Neural Network (ANN). The approach consists of 4 main phases; after the preprocessing phase, the pictures at hand are segmented using the K-means technique, then some texture features are extracted during which they're skilled a pre-trained neural network.*

Key Words: *Image Processing, Data Mining, Leaf Disease, Artificial Neural Network.*

1. INTRODUCTION:

India is an agricultural country wherein most of the population depends on agriculture. Research in agriculture is aimed towards increase of productivity and food quality at reduced expenditure, with increased profit. So as to get more valuable products, a product internal control is essentially mandatory. Diseases are impairment to the traditional state of the plant that modifies or interrupts its vital functions like photosynthesis, transpiration, pollination, fertilization, germination etc. Many studies show that quality of agricultural products could also be reduced thanks to plant diseases. These diseases are caused by pathogens viz., fungi, bacteria and viruses, and thanks to adverse environmental conditions. Therefore, the first stage diagnosis of disease is a crucial task. This paper provides an introductory part includes importance of plant disease detection; plant leaves analysis, various sorts of leaf diseases [1].

1.1. DATA MINING TECHNIQUES:

Data mining refers to extracting or mining the knowledge from great deal of knowledge. Data processing is employed to extract implicit and previously unknown information from data. Data collection and storage technology has made it possible for organizations to accumulate huge amounts of knowledge at lower cost. Data processing is that the process of exploration and analysis, by automatic or semiautomatic means, of huge quantities of knowledge so as to get meaningful patterns and rules. Data processing is about solving problems by analyzing data already present in databases. Data processing is that the process which provides an idea to draw in attention of users thanks to high availability of giant amount of knowledge and wish to convert such data into useful information. Data processing is additionally stated as essential process where intelligent methods are applied so as to extract the info patterns [2]. 1.2.

1.2. BACKGROUND STUDY:

Plant leaf diseases identification features a broad application prospective in agriculture and medicine, and is particularly significant to the biology diversity research. Plant leaf classification finds application in botany and in tea, cotton and other industries. the primary step is to show a computer the way to classify plants. Leaf recognition plays a crucial role in plant classification. Plants are basically identified supported flowers and fruits. However these are three dimensional objects and increases complexity. Plant identification supported flowers and fruits require morphological features like number of stamens in flower and number of ovaries in fruits. Identifying plants using such keys may be a very time consuming task and has been administered only by trained botanists. However leaves also play a crucial role in plant identification. Moreover, leaves are often easily found and picked up everywhere in the least seasons, while flowers can only be obtained at blooming season. Shape of plant leaves is one among the foremost important features for characterizing various plants visually. Plant leaves have two-dimensional nature and thus they're best suited for machine processing. Our paper presents survey of various leaf diseases identification [3].

1.3. Types of Diseases in Plant Leaf [1]:

Generally the diseases of plant in two forms the diseases can be:

- Infectious diseases caused by Fungi, Bacteria, Viruses, etc.
- Non-infectious diseases or disorders caused by mineral toxicities, soil acidity, nutrient deficiencies, or environmental factors.

1.4. Bacterial disease symptoms: The disease is characterized by tiny pale green spots which soon inherit deem water- soaked. Then it appear as dry dead spots as e.g. bacterial leaf spot have brown or black water-soaked spots on the foliage, sometimes with a yellow halo, generally identical in size. Under dry conditions the spots have a speckled appearance.

(a)Bacteria leaf spot (b) Virus affected plant

1.5. Viral disease symptoms: Among all plant leaf diseases, those caused by viruses are the foremost difficult to diagnose. Viruses produce no telltale signs which will be readily observed and sometimes easily confused with nutrient deficiencies and herbicide injury. Aphids, leafhoppers, whiteflies and cucumber beetles insects are common carriers of this disease, e.g. Mosaic Virus, search for yellow or green stripes or spots on foliage. Leaves could be wrinkled, curled and growth could also be stunted. (a) Bacterial leaf spot (b) mosaic virus Bacterial and Viral disease on leaves.

1.6. Fungal disease symptoms: Among all plant leaf diseases, those caused by fungus a number of them are discussed below and e.g. blight caused by the fungus Phytophthora infesters. It first appears on lower, older leaves like water-soaked, graygreen spots. When fungal disease matures, these spots darken then white fungal growth forms on the undersides. Early blight is caused by the fungus Alternariasolani. It first appears on the lower, older leaves like small brown spots with concentric rings that form a bull‟s eye pattern. When disease matures, it spreads outward on the leaf surface causing it to show yellow. In false mildew yellow to white patches on the upper surfaces of older leaves occurs. These areas are covered with white to greyish on the undersides. Figure (a) blight (b) early blight (c) false mildew.

(a) Late blight (b) early blight (c) downy mildew

Fungal disease on leaves

2. METHODOLOGY:

An order to detect and diagnosis the leaf infection/disease various research works are administered and various methods or algorithms are proposed. for instance candied citrus peel diseases was analyzed by color texture features analysis.

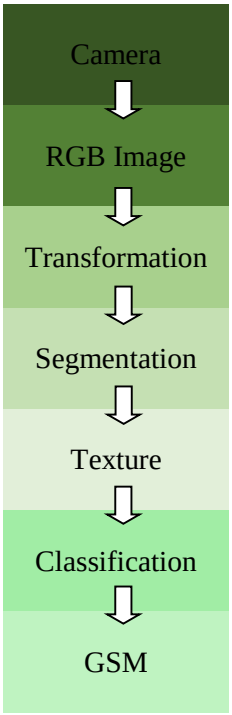

Figure.1 Flow Diagram for Proposed System

The feel feature analysis is intern categorized into structural, statistical, model based and transform method. Similarly in another method of study Hue Saturation Intensity [HSI] transformation is applied to the input image, then it's segmented using Fuzzy C-mean algorithm.This algorithm is hottest method utilized in image segmentation because it's robust characteristics for ambiguity and may retain far more information than hard segmentation methods. Although the traditional FCM algorithm works well on most noise free images, feature extraction stage deals with the colour, size and shape of the spot and eventually classification is completed using neural networks. In proposed project leaf infection detection and diagnosis is formed through image processing technique because Images form important data and knowledge in biological sciences. Digital image processing and image analysis technology supported the advances in microelectronics and computers has many applications in biology and it circumvents the issues that are related to traditional photography. This new tool helps to enhance the pictures from microscopic to telescopic range and also offers a scope for his or her analysis. Therefore it's many applications in biology. the tactic for detection and classification of leaf diseases is predicated on masking and removing of green pixels, applying a selected threshold to extract the infected region and computing the feel statistics to guage the diseases using JAVA and a few of the part using MATLAB. The implementation phase begins with leaf's sample being captured using regular camera with white background with the assistance of a stand. The image is loaded into MATLAB for processing. The features like texture and color features are extracted for identifying and classifying like healthy or diseased sample image.

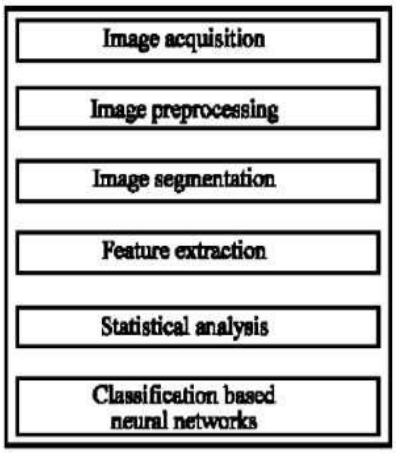

Figure. 2. Basic Procedure of Leaf Disease Detection System

2.1 Feature Extraction:

Once the feature extraction was complete, two files were obtained. They were: (1) Training texture feature data and (2) Test texture feature data. The files had 192 rows each, representing 32 samples from each of the six classes of leaves. Each row had 10 columns representing the ten texture features extracted for a specific sample image. Each row had a singular number (1, 2, 3, 4, 5 or 6) which represented the category of the actual row of knowledge. 1 represented Early scorch disease infected leaf. 2 represented Cottony mold disease infected leaf. 3 represented ashen mold disease infected leaf. 4 represented late scorch disease infected leaf. 5 represented tiny whiteness disease infected leaf and 6 represented normal leaf.

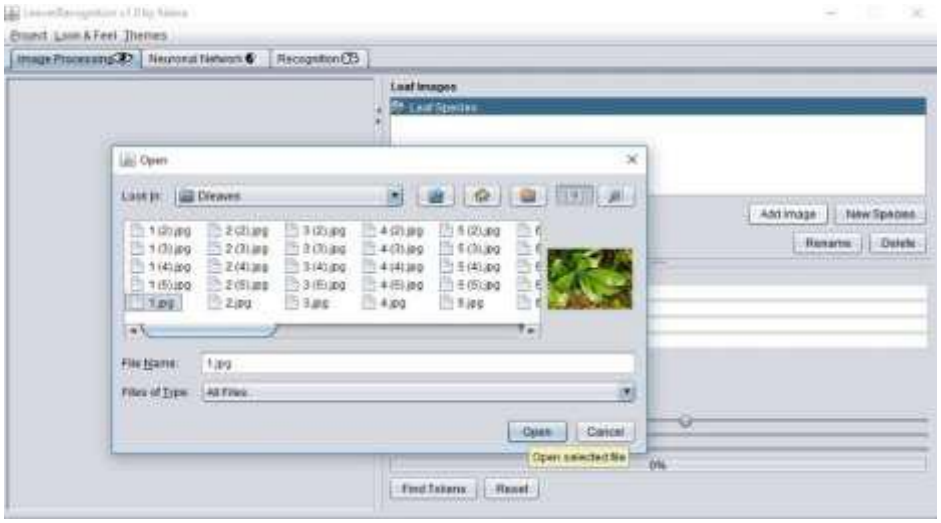

Figure 3. Training texture feature data

Figure 4. Test texture feature data

2.2 Classification using neural network:

A software routine was written in MATLAB that might absorb .mat files representing the training and test data, train the classifier using the train files then use the test file to perform the classification task on the test data. Consequently, a MATLAB routine would load all the info files (training and test data files) and make modifications to the info consistent with the proposed model chosen. The architecture of the network utilized in this study was as follows. the amount of hidden layers within the neural network was 10. The number of inputs to the neural network is adequate to the amount of texture features listed above. the amount of output is 6 which is that the number of classes representing the 6 diseases (Early scorch, Cottony mold. Ashen mold, late scorch, tiny whiteness) and therefore the normal leafs. The neural network used is that the feed forward back propagation. The performance function was the Mean Square Error (MSE) and therefore the refore the number of iterations was 10000 and the maximum allowed error was 10^{-5}.

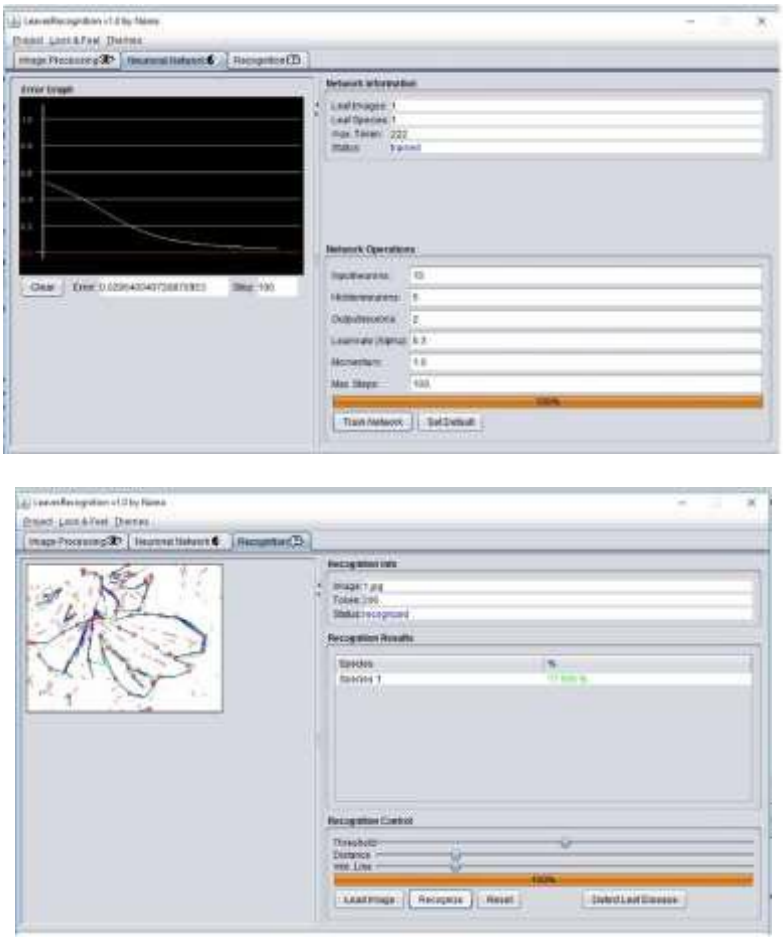

Figure.5. Classifying leaves using neural network

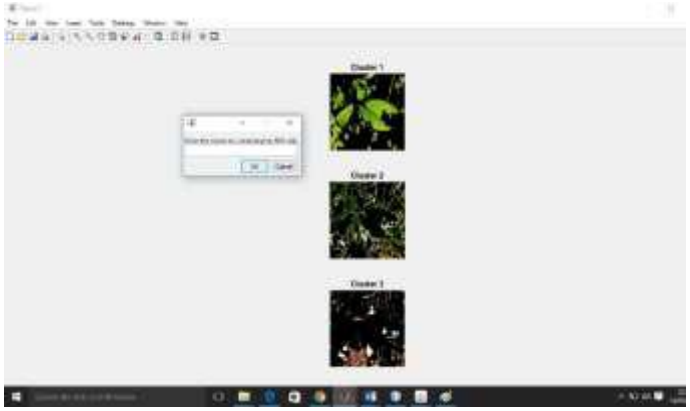

Figure. 6. Classifying leaves using ROI

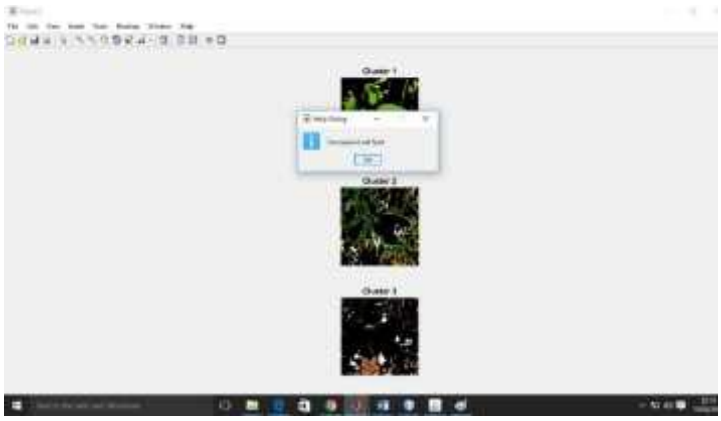

Figure.7. Cercospora Leaf spot

Figure 8. Cercospora Leaf – Affected Area and accuracy

3. RESULT AND OBSERVATIONS:

With these parameters, the network was trained. Once the training was complete, the test data for every class of leaves was tested. The results for NN classification strategy that were used are given in Table 1. The results shown in T able 1 were obtained employing a NN classifier for five different diseases.The results reported better classification accuracies for all the info models. especially , model Ml achieved the very best overall classification ac curacy. Model Ml achieved an overall accuracy of 89.5%, model M2 achieved an overall accuracy of 84.0% and model M3 achieved an overall accuracy of 83.66%. However, it should be noted that models M4 and M5 involve calculation of intensity texture features, which is disadvantageous in terms of computational complexity. Therefore, it's deciphered that model Ml is that the overall best model during this classifier. another advantage of using Ml is that the decrease in computational time for training and classification. this is often due to the elimination of intensity features and since of the less number of features present within the model. The recognition rate for NN classification strategy of HS and HSI models for early scorch, cottony mold and normal leaf image were shown in Table 2. The results shown in Table 1 were obtained using Neural Network supported Back Propagation principle for 10 testing images of the three testing types. especially , model Ml achieved better overall classification rates than model M5. Model Ml achieved an overall accuracy of 99.66% and model M5 an accuracy of 96.66%. However, it should be noted

that model M5 involves calculation of intensity texture features as already explained within the above paragraphs. Therefore, model Ml is that the overall best model during this classifier. In general, Table 1 and a couple of prove that the results for NN classifier supported statistical classification perform well in both cases, it's not useful in world applications, since choosing only intensity could also be detrimental to the classification task thanks to inherent intensity variations in an outside lighting environment. Hence, model Ml emerges because the best model in classifiers supported statistical classification. it's evident from Table 1 and a couple of that, for models M4 and M5, the classification accuracies for a few classes of leaves were inconsistent with the superb results that were obtained for other models. Similarly, from Table 1 and a couple of , the results for neural network classifiers also show some discrepancies in terms of accuracies for a few models. within the case of neural network with back propagation also like radial basis functions, model M4 (with only intensity features) performs poorly. this will be attributed to the very fact that the network didn't have enough information (in other words, there was overlapping of knowledge clusters) to form an ideal classification hyperplane to separate the info clusters belonging to varied classes of leaves..

The intensity of the leaves when considered as an entire might not have incorporated enough information, for the network to form correct classification decisions. This proves the very fact that for neural network classifiers using only intensity texture features won't yield good classification. One significant point to be noted in neural network classifiers is that the results might not be consistent across several trials using an equivalent input and parameters. this is often because the load initialization within the network is random. Hence, the outputs vary. The results for neural network classifiers that were shown during this research were the typical of outputs (classification accuracies) for 3 successive trials. Model Ml emerged because the best model among various models. it had been noted earlier, that this was partially due to the elimination of the intensity texture features. Elimination of intensity is advantageous during this study because it nullifies the effect of intensity variations. Moreover, it reduces the computational complexity by foregoing the necessity to calculate the CCM matrices and texture statistics for the intensity pixel map. However, in an outside application, elimination of intensity altogether may have an impact on the classification, since the ambient variability in outdoor lighting isn't taken into consideration.

Table 1: Percentage classification accuracy results of the test data from various diseases

Model	Color features	Early scorch	Cottonv mold	Ashen mold	Late scorch	Tiny Whiteness	Normal	Overall Average
M1	HS	95	93	80	83	87	100	89.50
M2	H	89	86	69	79	83	98	84.00
M3	S	88	87	72	80	77	98	83.66
M4	I	89	88	79	81	83	97	86.16

Table 2: Percentage classification accuracy results for neural network using back propagation

Model	Color features	Early scorch	Cottonv mold	Normal	Overall Average
M1	HS	99	100	100	99.6
M5	HS1	93	97	100	96.6

Table 3: Classification results per class for neural network with back propagation

From Species	Earlv scorch	Cottonv mold	Ashen mold	Late scorch	Tiny Whiteness	Normal	Overall Average
Earlv scorch	25	0	0	0	0	1	100.0
Cottonv mold	0	24	0	1	0	0	96.00
Ashen mold	0	0	25	1	1	1	100.00
Late scorch	0	1	0	20	1	0	80.00
Tiny Whiteness	0	0	0	1	22	0	88.00
Normal	0	0	0	2	1	23	92.00
Average							97.00

Table 3 shows the amount of leaf samples classified into each category for the neural network classifier with back propagation algorithm using model Ml. The results show that a couple of samples from late scorch and tiny whiteness leaves were misclassified. For the case lately scorch infected leaves, five test images were misclassified. Two leaf samples were misclassified as belonging to normal leaf class and therefore the others as a Cottony mold, Ashen mold and Tiny whiteness infected leaves. Similarly, within the case of small whiteness images, three test images from the category were misclassified as belonging to ashen mold, late scorch and normal classes. generally , it

had been observed in various trials that mis classifications mainly occurred in four classes; cottony mold, late scorch, tiny whiteness and normal..

4. CONCLUSION AND FUTURE WORK

During this study an image-processing-based approach is proposed and used for plant disease detection. We test our program on five diseases which effect on the plants; they are: Early scorch, Cottony mold, ashen mold, late scorch, tiny whiteness. The proposed approach is image processing based and is much supported K-Means clustering technique and Artificial Neural Network (ANN). The approach consists of 4 main phases; after the preprocessing phase, the pictures at hand are segmented using the K-means technique, then some texture features are extracted during which they're skilled a pre-trained neural network. Present experimental results indicate that the proposed approach may be a valuable approach and may significantly support accurate and automatic detection of leaf diseases. Supported our experiments, the developed Neural Network classifier that's supported statistical classification perform well and may successfully detect and classify the tested diseases with a precision of around 93%.

For future research, they need been some directions, such as, developing better segmentation technique; selecting better feature extraction and culling classification algorithms. This project has been considered just for four diseases and further it are often extended for various diseases. In future it are often extended to seek out the share of the infection in accordance with the world affected.

REFERENCE:

1. N. Sasirekha and N. Swetha, “An Identification of Variety of Leaf Diseases Using Various Data Mining Techniques”, International Journal of Advanced Research in Computer and Communication Engineering, Vol. 4, Issue 10, 2015
2. S. Sumathi and S.N. Sivanandam, “Introduction to Data Mining Principles”, Studies in Computational Intelligence (SCI) 29, 1–20, 2006
3. Rupa M, Pathur Nisha S, Geetha A and Shanthi P, “An Efficient Technique for Predicting Plant Leaf Diseases in Digital Image Processing”, Imperial Journal of Interdisciplinary Research (IJIR), Vol 2, Issue 12, 2016, ISSN: 2454-1362
4. Aldrich, B. and M. Desai, 1994. Application of spatial grey level dependence methods to digitized mammograms, Image Analysis and Interpretation, 1994. Proceedings of the IEEE Southwest Symposium on, April 21 -24, UK., pp: 100-105. Ah,
5. S.A., N. Sulaiman, A. Mustapha and N. Mustapha, 2009. K-means clustering to improve the accuracy of decision tree response classification. Inform. Technol. J., 8: 1256-1262.
6. Babu, M.S.P. and B. Srinivasa Rao, 2010. Leaves recognition using back-propagation neural network-advice for pest and disease control on crops. T echnical Report, Department of Computer Science and Systems Engineering andhra University, India.
7. Bauer, S.D., F. Korc and W. Forstner, 2009. Investigation into the Classification of Diseases of Sugar Beet Leaves Using Multispectral Images. In: Precision Agriculture, van Henten, E.J., D. Goense and C. Lokhorst, Wageningen Academic Publishers, WageningenDC., pp: 229-238.
8. Camargo, A. and J.S. Smith, 2009. An image-processing based algorithm to automatically identify plant disease visual symptoms. Biosyst. Eng., 102: 9-21.
9. Brown, Charles, Frank Chaloupka, and Michael Grossman. "The Demand for Cocaine by Young Adults: A Rational Addiction Approach." NBER Working PaperSeries.Working Paper 5713. August 1996.
10. Lili N. A / International Journal on Computer Science and Engineering,Classification Of Herbs Plant Diseases Via Hierarchical Dynamic ArtificialNeural Network After Image Removal Using Kernel RegressionFramework, ISSN: 0975-3397, Vol. 3 No. 1 Jan 2011, PN: 15-20 .
11. Lili N. A / International Journal on Computer Science and Engineering,Classification Of Herbs Plant Diseases Via Hierarchical Dynamic ArtificialNeural Network After Image Removal Using Kernel RegressionFramework, ISSN: 0975-3397, Vol. 3 No. 1 Jan 2011, PN: 15-20.
12. Devraj and Jain R (2011) PulsExpert: An expert system for the diagnosis and control of diseases in pulse crops. AgriSci J 38(9).
13. Subbaiah, K. M., In Proceedings of the First International Workshop on Biological Control of Chromolaenaodorata (ed)
14. R Balu, T Devi, Design and Development of Automatic Appendicitis Detection System using Sonographic Image Mining, International Journal of Engineering and Innovative Technology (IJEIT), Volume 1, Issue 3, Pages 67 – 74
15. Balu, Writing A Good Ph.D Research Synopsis, International Journal of Research in Science and Technology, Volume 5 Issue 5, Pages 38-48.