iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 8 Issue: V Month of publication: May 2020

DOI: http://doi.org/10.22214/ijraset.2020.5450

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue V May 2020- Available at www.ijraset.com

Plant Disease Detection using Image Processing

Shubham Malvankar[1], Bilwa Prabhu Moye[2], Riya Shirodkar[3], Vipul Soman[4], Nora Naik[5], Saurabh Naik[6]
[1, 2, 3, 4] *Department of Computer Engineering, Agnel Institute of Technology & Design, Goa-India*
[5, 6] *Assistant Professor, Department of Computer Engineering, Agnel Institute of Technology & Design Goa-India*

Abstract: Plant disease detection system has been a necessity for cultivators as detection of a correct disease might lead to a proper yield and effective growth of the plant. The proposed plant disease detection methodology aims at identifying tomato plant diseases from digital images of affected plant leaves. This proposed technique will help farmers or any tomato cultivators to detect plant diseases digitally as the same process manually might lead to incorrect detection. In the proposed project, various image processing steps are performed on the leaf images along with classification methods such as K-Nearest Neighbor Algorithm and Naïve Bayes algorithm to distinguish between different plant diseases.
Keywords: Image acquisition, Image segmentation, Image post-processing, Feature Extraction, Classification.

I. INTRODUCTION

The quantity of food consumed by the United Kingdom is almost equal to the quantity of food wasted in India. Up to 40% of food produced in India is wasted as reported by the United Nations [1]. Food shortage and wastage has been the greatest issue in recent times.

Following are the primary types of Food Losses:-

1) *Production Stage:* Due to Diseases, Spillage, mechanical error.
2) *Handling and Storage:* Losses due to storage and transportation during farms and distribution networks.
3) *Processing:* Losses during domestic and Industrial Processing like juice production, canning, sorting for unsuitability.
4) *Distribution:* The waste generated by market distribution systems, at wet markets, supermarkets, wholesale markets and retailers.
5) *Consumption:* Wastes produced by consuming food at a domestic level.

This project covers a potential solution to reduce the wastes caused during the production stage due to plant diseases. One of the existing methods to detect a plant disease is by simple observation performed by the experienced farmers or experts. In case of a large plantation, a considerable team of experts is required to examine the plants which increases the overall cost and time required. The proposed system facilitates in predicting the diseases using image processing and other techniques which makes the detection process faster and easier, it also assists the new farmers in detecting the diseases.

Controlling plant diseases is very essential in the production of food since it involves the use of various resources like land and water for cultivation of food crops, fuel and other energy sources used for machinery and equipment. All plants carry intrinsic disease resistance, yet there are various examples of plant diseases which have caused devastating impacts for instance the Great Famine of Ireland, soybean cyst nematode, rice blast and citrus canker. Farmers worldwide face major losses due to plant diseases.

II. EXISTING SYSTEM

Plant leaf diseases are very common in plantations and farms, this system will help the farmers in detecting plant disease with the help of just an image of the affected leaf. This is not possible with traditional methods as they are slow, time-consuming and might also lead to incorrect detection. So it is important to implement a modern method that would help in such problems with accuracy and speed.

Leaves differ from each other based on different physical features such as shape, color etc. Some plant diseases cannot be predicted by looking at the external features and guessing the disease. Also, laboratory predictions of certain diseases might be 100% accurate but it is a lengthy process. Plants which are very sensitive require faster detection and faster measures to cure them therefore the farmers cannot opt for laboratory detection technique since this takes time. It may also result in wrong recognition, so they need an efficient and foolproof system to aid in their work. The aim of our system is to help farmers detect plant diseases without any delay and inaccuracies.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue V May 2020- Available at www.ijraset.com

III. PROPOSED SYSTEM

The objective of this project is to develop a plant disease detection system using the domains of image processing. It aims at detecting plant diseases up to a certain extent. In order to implement the system to detect plant diseases, we extract the data from the plant leaf image using Image Processing. Then using a classifier the image is classified as either a healthy leaf or an infectious leaf.

Architecture of the Proposed System

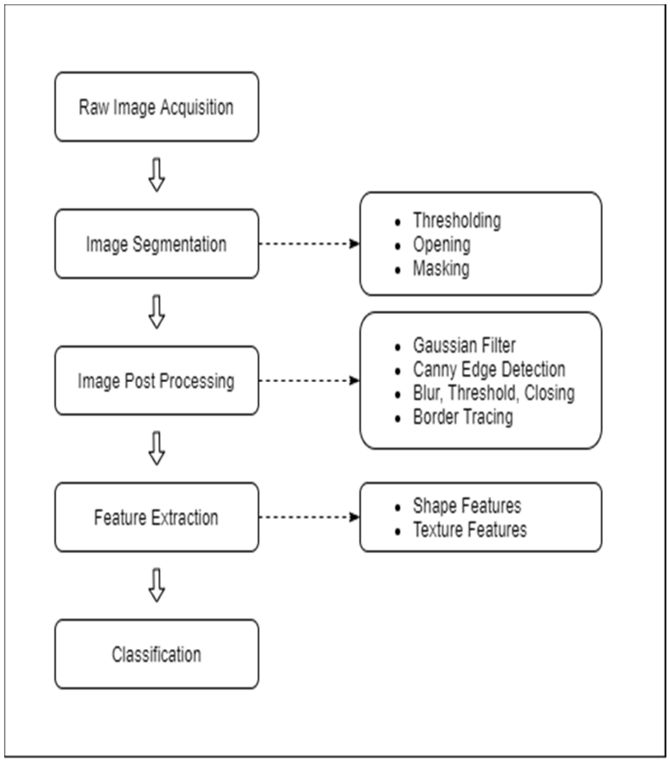


Fig. 1 Block Diagram

A. Image Acquisition

Image Acquisition is the first step in disease recognition. It is a process of acquiring an image of a target leaf using a digital camera. The image is stored for further processing. All possible optical details of the leaf are captured. The images used in this project are of tomato leaves. The dataset is categorized into 6 different categories. The first dataset consists of healthy tomato leaves and the rest are of leaf images with diseases namely bacterial spots, late blight, leaf curl, mosaic virus and septoria leaf spot. This categorization will later be used as an input for training and testing the classifiers in the classification phase.

Fig.2 Original image acquired

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue V May 2020- Available at www.ijraset.com

B. Image Segmentation

Image segmentation refers to the method of fragmenting a digital image into multiple parts. We use this method to separate the background of a leaf from the foreground on which further analysis is done. In order to extract numerical features, we only require the leaf part which is shades of green and yellow colors in majority and shades of brown and black for the diseases. Rest of the pixel data is dumped to avoid outliers. In order to accomplish this task, we use color image thresholding and other procedures. To perform thresholding, the RGB image is transformed into HSV which is done due to the following reasons:

1) RGB is a composite color model
2) HSV is good for visualization
3) Using only the Hue component will make the algorithm less prone to lighting variations.
4) Channel-wise arithmetical operation performed in RGB space may give very inaccurate or even wrong results

The next step is masking in which a mask is created of the original RGB image with the help of the HSV image obtained In the previous step. After which a bitwise AND operation is performed between the actual image and the masked image in order to obtain the segmented image.

Fig.3 Output after Image segmentation

C. Image Post Processing

Pre-processing among other things include filtering techniques. However, this application requires the use of leaf edges for its functioning which may be softened on such smoothing. For this reason, we first segment the image to accurately detect the outer edge and then use post-processing to smoothen the image which removes various noises.

1) *Gaussian Filter:* Gaussian Filtering operator which is a 2-D convolution operator is used to remove off the details and noise by blurring off images much like mean filtering but with a different kernel which has a shape of a Gaussian hump.

1-D form of Gaussian distribution

$$G(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{\frac{-x^2}{2\sigma^2}}$$

2-D form of Gaussian distribution

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{\frac{-(x^2+y^2)}{2\sigma^2}}$$

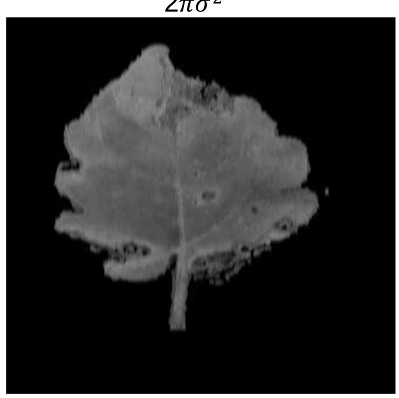


Fig.4 Output after Gaussian Filter

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue V May 2020- Available at www.ijraset.com

2) *Canny Edge Detection:* The Canny edge detection algorithm detects edges in images and is a multi-stage algorithm. To compute the intensity of the gradients it uses a filter based on the Gaussian derivative. The Gaussian reduces the effect of noise present in the image, it then uses Non-Maximal Suppression which thins down the potential edges to near 1-pixel curves. After which Hysteresis Thresholding is carried out which decides whether the edge pixels are kept or removed. The output of Canny can be adjusted using three parameters which are the width or the kernel of the Gaussian, the lower threshold and the higher threshold of the hysteresis.

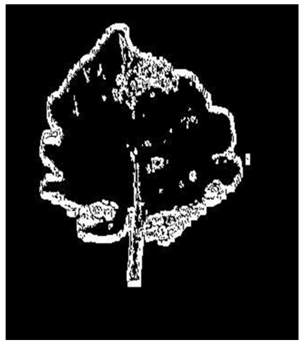


Fig.5 Output after Canny Edge Detection

3) *Blur-Threshold-Close:* This process involves three steps- The first step involves blurring the image with a Gaussian Filter. The Gaussian Filter spreads the intensity of the pixel blobs in its neighboring pixels. The second step performs thresholding on these pixel spots and removes the outer pixel belt using a certain threshold value, following which a closing operation is performed which removes the reduced pixel spots in the image. This process is performed several times which eliminates the pixel spots in the background of the leaf.

Fig.6 Output after Blur-Threshold-Close

4) *Border Tracing:* The result of Canny Edge Detection is an image with a thin multi-pixel border. This method of Border Tracing yields a border with 1-pixel thickness. It starts with an initial border point of the object in the binary image which is followed by an iterative search in its 8-neighborhood periphery. The algorithm is called Radial Sweep because it hypothetically draws a line segment in all of its neighborhood to check for a border pixel.

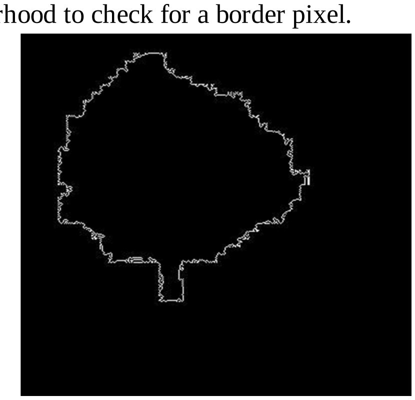


Fig.7 Output after Border tracing

D. Feature Extraction

A Feature in an image may be parameters such as a piece of spatial information, properties such as shape, curves or edges, borders, they may also be features such as blobs and ridges. On successful extraction of a feature, the result; called the feature descriptor or feature vector is then stored for further processing. This complete set of features will then be given to the classifier for training and testing.

There are two ways in which features are extracted from a leaf: shape features, texture features

1) *Shape Features:* Length- length of the two terminals of the central vein.
   a) *Width*- The highest distance between the points of the orthogonal intersection of the length and the border is the physiological width.
   b) *Area* - number of set pixels on the segmented image.
   c) *Leaf Perimeter*- Refers to the number of border pixels in the leaf margin.
   d) *Aspect Ratio*- The aspect ratio refers to the ratio of physiological length to physiological width of a leaf.
   e) *Form Factor*- Is used to specify the difference between a leaf and a circle. It is defined as $4\pi A/P^2$, where P denotes the perimeter of the leaf margin and A represents the leaf area.
   f) *Rectangularity*- Rectangularity refers to the likeness between a rectangle and a leaf. The formula is defined as (L*W)/A, where L is the length, W is the width and A is the leaf area.

2) *Texture Features:* Angular Second Moment- ASM measures orderliness, how regular or orderly the pixel values are in the window.
   a) *Energy*- Obtains the sum of squares of every element in the GLCM. (Square root of ASM)
   b) *Entropy*- Measure uncertainty of the image (variations)
   c) *Contrast*- Calculates the local variations of the gray-level co-occurrence matrix.
   d) *Homogeneity*- Refers to the adjacency of the elemental distribution in the GLCM and GLCM diagonal.

Feature	Value
Area	26376
Form Factor	0.19402943676497836
Entropy Mean	1.6719866350768626
Entropy Range	4.204318021092474
Homogeneity Mean	0.682312377753786
Homogeneity Range	0.08485436691438598
Perimeter	1307.0
Aspect Ratio	0.9004080511633751
Contrast Mean	267.73086219520883
Contrast Range	1268.4025336131135
Sum of Squares Mean	3202.456126228885
Sum of Squares Range	9838.438988702219
Rectangularity	1.9748485653549155
Angular Second Moment Mean	0.4357475156795096
Angular Second Moment Range	6.465582331463026E-4

Fig.8 Output after Feature Extraction

E. Classification

The primary goal of performing classification is to predict the target class for each case in the data accurately. In the proposed system, classification algorithms like K-Nearest Neighbor and Naive Bayes algorithms are used to classify diseases based on the numerical dataset derived from feature extraction.

1) *K-Nearest Neighbor Algorithm:* KNN is used to solve classification as well as regression problems. In the proposed system, KNN is used for classification purposes. It is a supervised machine learning algorithm which relies on input data whose characteristics or parameters are known or also called as labelled data and gives as output a result based on its learning patterns. In the proposed project, KNN takes as input the image of the leaf and all its characteristics and features that are extracted in previous phases and it creates a data model. This data model is then given as an input a test data that is a leaf image whose characteristics are not known and the algorithm predicts the characteristics, in our case the disease of the leaf.
2) *Naive Bayes Algorithm:* Naive Bayes Algorithm is also used to differentiate between different data objects based on some characteristics. It follows a Probabilistic machine learning model which is utilized for learning tasks. The core of this classifier is based on Bayes theorem.

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Naive Bayes classifiers presume that a feature value is independent of other feature values of a class variable.
In simple terms Naive Bayes model follow conditional probability for classification.

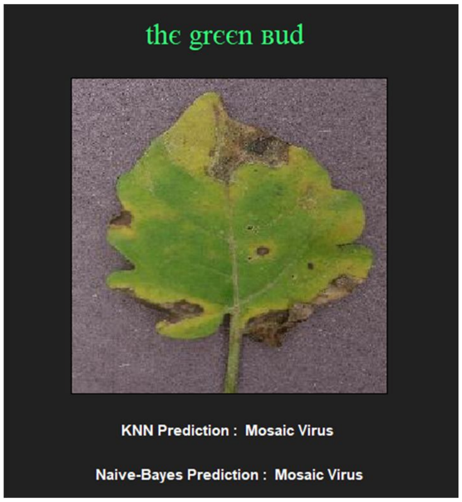


Fig.9 Result after performing the K-Nearest Neighbor Algorithm and Naive Bayes Algorithm

IV. CONCLUSION

The primary motive of this project was to recognize diseases which are still difficult for the farmers to detect without any assistance. During the study of this project, we researched the diseases of tomato leaves and its characteristics, various image processing techniques and classification algorithms. The first classification algorithm we have used is Naive Bayes and second is K Nearest Neighbours algorithm. Naïve Bayes Classifier has an accuracy of 65% and K Nearest Neighbour has an average accuracy of 80 %. KNN was found to be better as it has higher accuracy. The training dataset consists of a minimum of 500 images with each disease having at least 100 images for training. This paper infers an efficient and accurate approach for the detection of plant diseases.

REFERENCES

[1] Food Wastage in India," [Online]. Available: https://thecsrjournal.in/food-wastage-india/

[2] S. G. Wu, F. S. Bao, E. Y. Xu, Y. Wang, Y. Chang and Q. Xiang, "A Leaf Recognition Algorithm for Plant Classification Using Probabilistic Neural Network," 2007 IEEE International Symposium on Signal Processing and Information Technology, Giza, 2007, pp. 11-16, doi: *10.1109/ISSPIT.2007.4458016*.

[3] Singh and A. K. Misra, "Detection of plant leaf. diseases using image segmentation and soft computing techniques," Information Processing in Agriculture, vol. 4, no. 1, pp. *41-49, 2017*.

[4] P. K. Sethy, B. Negi, S. K. Behera, N. K. Barpanda and A. K. Rath, "An Image Processing Approach for Detection, Quantification, and Identification of Plant Leaf Diseases -A Review," International Journal of Engineering and Technology, vol. 9, no. 2, pp. *635-648*, April 2017.

[5] S. Radha, "Leaf Disease Detection using Image Processing," Journal of Chemical and Pharmaceutical Sciences, March 2017.

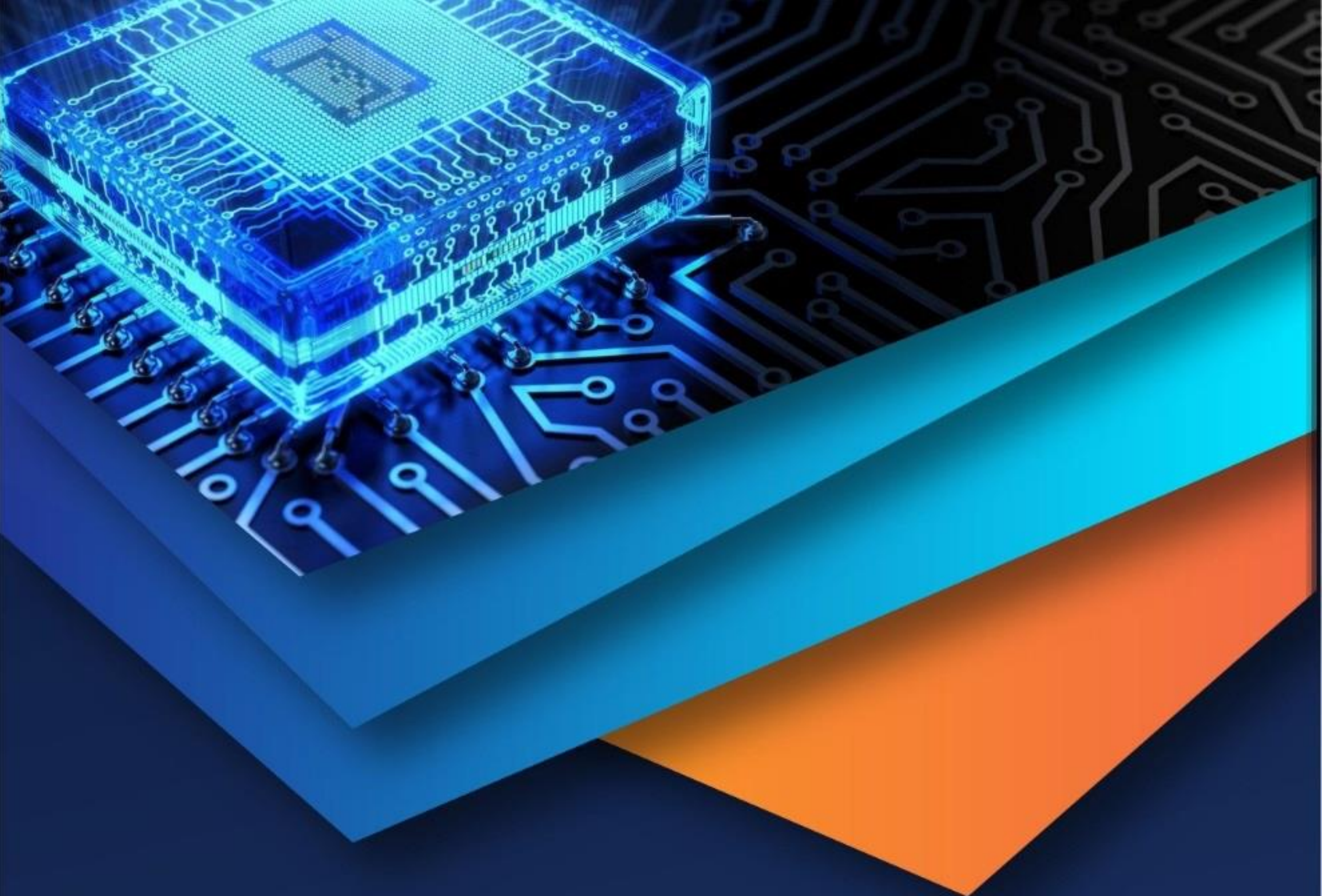
[6] R. Patil, S. Udgave, S. More, D. Nemishte and M. Kasture, "Grape Leaf Disease Detection Using K-means Clustering Algorithm," International Research Journal of Engineering and Technology, vol. 3, no. 4, 2016.

[7] P. M. Mainkar, S. Ghorpade and M. Adawadkar, "Plant Leaf Disease Detection and Classification Using Image Processing Techniques," International Journal of Innovative and Emerging Research in Engineering, vol. 2, no. 4, 2015.

[8] S. R. Kamlapurkar, "Detection of Plant Leaf Disease Using Image Processing Approach," International Journal of Scientific and Research Publications, vol. 6, no. 2, pp. 73, February 2016.

[9] M. A. Hussein and A. H. Abbas, "Comparison of Features Extraction Algorithms Used in the Diagnosis of Plant Diseases," Ibn AL- Haitham Journal For Pure and Applied Science, pp. 512, doi: *10.30526/2017*.IHSCICONF.1785

[10] Priya P., D. A. D'souza, "Study of Feature Extraction Techniques for the Detection of Diseases of Agricultural Products," International Journal Of Innovative Research In Electrical, vol. 3, no. 1, pp. 2321-5526, April 2015.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY