International Journal of Scientific Research & Engineering Trends
Volume 5, Issue 2, Mar-Apr-2019, ISSN (Online): 2395-566X

Disease and Cure Prediction in Crops Using Neural Networks

Prathyusha Mukkarla **Prof. Asha** **Phani Deepthi Yarlagadda**
nagireddy.mukkarla2015@vit.ac.in yarlagaddaphani.deepthi2015@gmail.com
Dept. of Computer Science Engg.
Vellore Institute of Technology
Vellore, TN, Chennai

Abstract- **Detection and classification of plant diseases at appropriate times and accurately are very important factors in production of plants and also in the reduction of losses in crop yield. Detection of plant diseases using traditional methods like naked eye detection is ineffective as they are microscopic in nature. The aim is to predict the leaf type, symptoms, cure for the disease. The algorithm we used follows three steps:-image preprocessing and analysis, feature extraction, and recognition of plant disease. Our aim is to detect the symptoms of the occurred disease in an accurate manner. Once the image is captured it is preprocessed using the masking technique, various properties like color, shape and size are extracted and fed to the neural networks for classification.**

Keywords- **Disease, Neural network, preprocessing, analysis.**

I. INTRODUCTION

One of our driving motivations is to provide solutions for tomorrow's problems using tomorrow's technology. One of the vital sectors of Indian Economy is Agriculture. Employment to virtually five hundredth of the countries personnel is provided by Indian agriculture sector. India is thought to be the world's largest producer of pulses, rice, wheat, spices and spice merchandise. Farmer's economic process depends on the standard of the merchandise that they manufacture, that depends on the plant's growth and also the yield they get.

Therefore, in field of agriculture, detection of diseases in plants plays ANN semantic role. In this project we are implementing technique to detect and classify the diseases based on neural networks. Firstly the image will be preprocessed. Preprocessing includes converting the image into grayscale image and masking it. In masking the green pixels will be masked and the percentage of RGB value of the affected portion 13 will be found out using image processing.

Once the preprocessing is done the features like shape, size and color of the leaf will be extracted. After comparing the leaf details with the training dataset leaf's name, its disease and cure will be predicted. Artificial intelligence techniques offer great solutions to detect diseases of sick plants by visual inspection of damaged leaves. Our solution of automated early disease detection is based on an artificial neural network. Main advantages of our solution are processing speeds and accuracy. It can also detect scorching and mold abnormalities on plants apart from diseases. Our solution mainly focuses on the following crops:

- Brinjal
- Mango Manual detection of diseases is very time consuming and also it cannot be trusted upon as we may overlook things which are not visible to our naked eyes. This system is designed such that it captures even the tiniest details which naked eye may miss.

II. MOTIVATION

Although weeds are the major cause of crop loss on a global scale, major losses are suffered by agricultural crops due to insect damage and plant diseases. The worldwide annual production tonnage % age lost to various pests at the start of the 21st century have been estimated as follows: losses due to animal pests, 18%;

- Microbial diseases, 16% (and 70-80% of these
- Losses were caused by fungi);
- weeds, 34%;
- Making a grand total of 68% average annual loss of crop production tonnage (data from Oerke, 2006). We all are aware that every year there is a humongous loss not only because of variety of diseases but also due to the incapability of timely disease detection.

In such cases farmers have to go through both economic as well as mental pressure. One of the main motivations of this system is the disadvantages of manual leaf disease systems. The speed of the manual system is very slow when compared to automated systems. Moreover there are chances that manual systems may give inaccurate results. Humans cannot be programmed to do efficient multitasking and therefore it is not economical.

© 2019 IJSRET

III. LITERATURE SURVEY

There is the existence of the list of methods used by researchers for the detection of plant leaf diseases. Introduces an image processing-based approach that is used to detect leaves, stem's disease. The work tested this program on five diseases in which the plants had an impact; they are: initial shocks, Cotcinsy mold, Aisha mold, and late sweat, small white, proposed approach is based on image processing. In this approach, images on hand are broken by using K-means technique. In the next stage, the images containing the segment are passed throughout the Pre- skilled neural network. The result show that proposed approach leaves a lot to help in the automatic and accurate 20 detection of disease. Based on our experiment, the developed neural network classifiers, which is based on statistical classifications, and it performs well and effectively identifies and classifies the exact test cases of about 93%. [3] In the first paper [4], there are two steps to recognize the affected area of the disease.

At first, border identification based image is split, as well as lastly image analysis and illnesses are classified using our proposed HPCDDD algorithm. The goal of the network was to expand an advanced computing method that could classify the affected area of cotton leaves by using image analysis techniques. The second paper [5] develops processing network mainly include of four steps, the initial of which, used for the input RGB images, a color change construction is formed, for the reason that it is for RGB color creation and the image of converted RGB is used for that, HSI is used for color descriptor in the next phase, by the threshold value, green pixels are covered and removed.

In the third, using a precompiled threshold level, green is used to remove the pixels and covering for useful segments, which are extract earlier in this phase, while the image is the segment. And division is done in the last phase. The third paper [6] includes a few stages in the ID procedure of the disease, the four noteworthy advances be as per the following: at first, utilized for the info RGB image, a color change development is being used, and after that with an exact limit value, green pixels Are computed and evacuated, which are utilized to figure literary information keeping in mind the end goal to acquire facilitate helpful regions after the division procedure. All things considered, group patients are utilized for ordering offices. The proposed calculation demonstrates its adequacy with 94% debasement in the fruitful location and arrangement of the tested illness. 21 The quality of the proposed calculation is demonstrated by utilizing test aftereffects of roughly 500 plants in a database.

In the fourth paper [7], an application for printed observation has been clarified in the recognizable proof and association of plant leaf problem. The strategy was tested on huge numbers of the plants. Particular maladies were taken for the treatment of those plants. Trial comes about demonstrate that the proposed approach have the capacity to recognize and arrange leaf disease with modest computational exertion. To enhance the character rate of the problem in various stages, the preparation tests can be enhanced and the highlights alongside the ideal offices can be given as the info status of the disease acknowledgment.

IV. ARCHITECTURE DIAGRAM

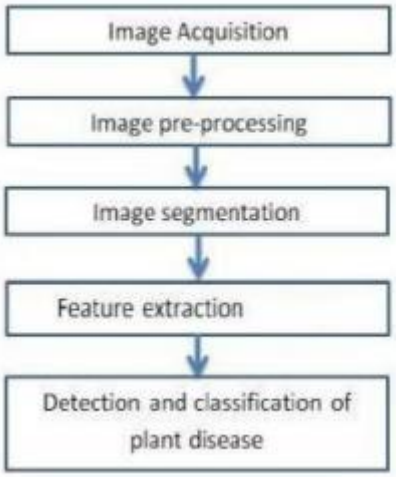

Fig. 1 Architecture Diagram.

V. RESEARCH METHODOLOGY

1. Image Acquisition- In this process the image input as input is given. The sampled leaf images are to be collected from field as well as internet. Some images are needed to train the system and some images are necessary for testing the leaf. First the image has to be resized to some pixels to make the task easier. The collected images should be resized to one particular pixel size to be used as an input because it is difficult if we use different sizes. Here, we are using 256*256 as pixel size.

2. Image Preprocessing- The image always might not be perfect; there may be additional noise in the image. It is difficult to identify the disease accurately if the image has noise present in it. In pre-processing the Median Filter is used to removing the noise and for producing the enhanced images. Median filter is performed with salt and pepper noise by choosing appropriate threshold. An image has salt and pepper noise will have dark pixels (black dots or pepper) in bright region and bright pixels (white dots or salt) in dark region. Method used to remove such type of noise is median filter. Median filter is non-linear filter.

© 2019 IJSRET

The median filter is to find the median value by, Ewing, across the window, replacing each entry in the window with the median value of the pixel. It's best in removing salt and pepper noise and impulse noise. Median filter erases black dots called the pepper and fills in the white holes in the image, called salt. It simply replaces each pixel values by the median of the intensity level in the neighborhood of the pixel.

3. Image segmentation- This consists of three modules:

3.1 Color Features- Firstly, the image has to be converted into grey scale to make the work easier. Four features are required to identify the color feature of the image.

1) Mean
2) Standard Deviation
3) Kurtosis
4) Skewness
5) Texture Feature:
6) After the color feature is extracted, texture feature should be extracted.

3.2 Texture feature- Here we are using HOG technique. The first step is to divide the input image into blocks. Each block is further divided into small regions, called cells. Usually, it should be divided so that the blocks overlap each other, the same cell may be in several blocks. For each pixel within the cell the vertical and horizontal gradients are obtained. The method to do that is to use 1-D Sobel vertical and horizontal operators: Gx(y,x) = Y(y,x+1) – Y(y,x-1); Gy(y,x) = Y(y+1,x) – Y(y-1,x). Y(y,x) is the pixel intensity at coordinates x and y. Gx(y,x) is the horizontal gradient, and Gy(y,x) is the vertical gradient.

The magnitude and phase of the gradient are determined as: Next, the HOG is created for each cell. For the histogram, Q bins for the angle are chosen (for example Q=9). Usually unsigned orientation is used, so angles below 0o are increased by 180o . Since different images may have different contrast, contrast normalization can be very useful. Normalization is done on the histogram vector v within a block. One of the following norms could be used: A descriptor is assigned to each detector window. This descriptor consists of all the cell histograms for each block in the detector window. The detector window descriptor is used as information for object recognition. Training and testing happens using this descriptor.

3.3 Shape Feature- Here different cluster techniques are used for the classification or segmentation of image. A K-means clustering technique is used here. The technique is used to segment the diseased leaf into four clusters. Segmentation is a technique which classifies the image into different segments. We cannot get all the meaning full information from original image so, segmentation should be done. In this we are going to mask green pixels of image and remains with the infected portion of image. The principle used is to characterize k centroids, one for each cluster. These centroids ought to be set as a result of various area causes distinctive outcome. Along these lines, the better decision is to put them however much as could be expected far from each other. The following stage is to take 35 each guide having a place toward a given informational index and partner it to the closest centroid.

At the point when no point is pending, the initial step is finished and an early gathering age is finished. Now we have to re-figure k new centroids as ban focuses of the clusters coming about because of the past advance. After we have these k new centroids, another coupling must be done between similar informational collection focuses and the closest new centroid. A circle has been created. Because of this circle we may see that the k centroids change their area well-ordered until the point that no more changes are finished. At the end of the day centroids don't move any more. 6. Feature Extraction: The Features Extraction is used to reduce the data sets and to create the new subsets of features by a combination of the existing features.

The texture feature, shape feature and pixel coefficient value is extracted. The texture, shape and pixel value are extracted using the GLCM. The features extracted from GLCM are-

- Mean
- Standard Deviation
- Energy
- Entropy
- Skewness
- Variance
- Kurtosis
- Smoothness
- Contract .
- Correlation

3.4 Contrast- Returns a measure of the intensity contrast between a pixel and its neighbour over the whole image. Range = [0 (size (SGDM, 1)-1) ^2] Contrast is 0 for a constant image. C(i,j) is pixel value of location (i,j). $Contrast = \sum (i,j)(i,j)2\ N-1\ i,j=0\ 2)$.

3.5 Energy= Returns the sum of squared elements in the SGDM. Range = [0 1]. Energy is 1 for a constant image. $Energy = \sum (i,)2\ N-1\ i,j=0$

3.6 Homogeneity- = Returns a value that measures the closeness of the distribution of elements in the SGDM to the SGDM diagonal. Range = [0 1] Homogeneity is 1 for a diagonal SGDM. $Homogeneity = \sum$ C(i,j)/(1 + $(i-j)2\ N-1\ i,j=0$).

3.7 Correlation- = Returns a measure of how correlated a pixel is to its neighbor over the whole image. Range = [-1 1] Correlation is 1 or -1 for a perfectly positively or

© 2019 IJSRET

negatively correlated image. $Correlation$ = (∑{i × j} × C(i,j) − {μx − μy})/ σx × σy.

3.8 Classification- Here the classifier is used is NN (neural network). It is a feed forward network which is widely used in pattern recognition problems and classification. PNN is mainly consist of four layers first is input layer second is pattern layer third one is summation layer and the latter is yield layer. Once trained, the network weights are frozen and can be used to compute output values for new input samples Along with the prediction of disease the cure that is to taken is displayed.

Features	Formulas
Mean	$\mu = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N} p(i,j)$
Standard Deviation	$\sigma = \sqrt{\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}(p(i,j)-\mu)^2}$
Energy	$E = \sum_{i,j} p(i,j)^2$
Entropy	$h = -\sum_{k=0}^{i-1} prk(log_2 prk)$
Skewness	$S = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(\frac{(p(i,j)-\mu)^3}{\sigma}\right)$
Variance	$Var = (S.D)^2$
Homogeneity	$H = \sum_{i,j}\frac{p(i,j)}{1+\lvert i-j\rvert}$
Kurtosis	$K = \left\{\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(\frac{(p(i,j)-\mu)^4}{\sigma}\right)\right\} - 3$
Smoothness	$R = 1 - \frac{1}{1+\sigma^2}$
Contrast	$C = \sum_{i=1}^{M}\sum_{j=1}^{N}(i,j)^2 p(i,j)$
Correlation	$cor = \sum_{i=1}^{M}\sum_{j=1}^{N}\frac{(i-\mu)(j-\mu)p(i,j)}{\sigma_1\sigma_2}$

Fig. 2 Formulae of GLCM Matrix.

1. Paper title- This information should be placed at the top of the first page in 24 point, Times New Roman in Title case, and centered. This style is defined under the style menu of this document as "TITLE."

2. Insert a 11 point blank line between the Title and the

3. Author listing and between the Author listing. Author listing: 11 point, Times New Roman, centered. This style is defined under the style menu of this document as "Author." For Special and Panel Session, list all panelists as authors. Include only the author names in the author listing. Use the full first name for the authors. If an author is submitting more than one paper please provide the same name on both papers. Be consistent as the name you provide on the paper will be what is used to create the author index for the proceedings.

4. Author Affiliations- 10 point, Times New Roman, centered under the author name. This style is defined under the style menu of this document as "Affiliations." Information for each author such as email, department or college, university, city, state and zip can be listed. If you list the email address make sure the Hyperlink is not attached. If the email address turns a color that means the hyperlink is active. Right click on the address and select Remove Hyperlink.

There are 3 basic types of setups that will be encounter: One author, multiple authors from the same institution and multiple authors from multiple institutions. If the paper has one author use the affiliation layout as shown at the top of this document where all the author information is listed under the paper title. For multiple authors, from the same institution, list the common university, then list the email addresses at the end in the same order as the authors appear in the Author Listing. For multiple authors from multiple institutions list the author names under the title, then list the author emails under their names. Additional information for multiple authors can be added to the Author Information section at the end of the paper.

VI. PAPER BODY FORMAT

1. Column Format Instructions- The title and author data is in one-column format, while the rest of the paper is in two-column format. To accomplish this, most word processors have a section break that is installed to separate the one and two-column format. For example, in Word, under the Insert menu select Breaks – Continuous. After you enter the title and author information enter a few blank lines and then insert a Continuous section break. Now you must define this section to be in two-column format. To do this in Word, under the "Format" menu select "Columns." This option will open the Columns window.

It has an input box for the number of columns. Enter 2 and then set the spacing to 0.2" and select equal column widths. If you have the margin widths set correctly, the width of the column should display as 3.40." If it does not, go to the "File" menu and select "Page Setup - Margins." This will open the "Page Setup" window that will allow you to set the top and bottom margins to 1" and the right and left margins to 0.75." Every word processor will have its own method of accomplishing the above setting, however, most follow the same format. For "Work in Process" or "Panel Session" submissions, if the paper does not fill the entire page and covers more than 4 inches of the first column, please balance the paper within the two-columns. The easiest way to do this is to insert a section break at the end of the document.

2. Font and Spacing Instructions- Use the full justify option for your columns, and use two columns in all pages. The two columns must always exhibit equal lengths and you should try to fill your last page as much as possible. To obtain such results, you are free to adjust the figure sizes provided this does not compromise their clarity. Use one line of space between text and section headings. Use one line of space between text and

© 2019 IJSRET

captions, equations, and tables. Use the spelling and grammar checkers. Do not use the “hyphenation” feature in Word. Please use the following font and alignment instructions:

2.1 Body text- 10 point, Times New Roman, full justified, single space, no blank lines between the paragraphs. Indents - first paragraph of section - none (this style is defined under the style menu of this document as "First Paragraph") Indent - all other paragraphs - .25" (this tyle is defined under the style menu of this document as "Body Text"). Follow the examples shown in this document.

2.2 Section Headings- 10 point, Times New Roman, bold, centered, use SMALL CAPS, leave one blank line above and below. For example, "Page Layout" on page one of this document is a Section Heading (this style is defined under the style menu of this document as "Section Headings"). To convert the text to small caps, click on the “Home” tab, and select the pull down arrow in the “Font” submenu. This will open the “Font Window” where you can select the Small Cap option. **2.3 Section Sub-headings-** 10 point, Times New Roman, italic, left justified, leave one blank line above and below. For example, "Font and Spacing Instructions" on this page is a Section Sub Heading (this style is defined under the style menu of this document as "Subheading").

2.4 Bullets- 10 point, Times New Roman, left justify and indent the text 0.25.” Insert a blank line after the bullet list but not before, follow the examples in this document (this style is defined under the style menu of this document as "Bullets").

3. Page Numbering The page number on your manuscript is centered at the bottom of the footer. The final page number will be modified by the conference before the file is saved into the proceedings

VII. FIGURES, TABLES AND EQUATIONS

All figures and tables must fit either one or two-column width, 3.4" or 7" wide respectively. It is suggested that you use one-column whenever possible. If your table or figure will not fit into one-column, then insert a continuous section break before and after the table or figure, as described above and define it as one-column. To make the paper read easier you may want to position any table or figure that requires one-column either at the bottom of the page or the top of a new page. Do not abbreviate “Table”; use Roman numerals to number tables. Use the following format guidelines for figures and tables:

1. Figure and table headings- 8 point, Times New Roman UPPERCASE, centered. Place below the figure and above the Table, (this style is defined under the style menu of this document as "Figure Heading")

2. Leave one blank line above and below each Table or Figure.

3. Figure and table captions- 8 point, Times New Roman, SMALL CAPS, centered. Place below the figure or table headings (this style is defined under the style menu of this document as "Figure Caption")

4. Table text- 8 point, Times New Roman, (this style is defined under the style menu of this document as "Table text"). Table I and Figure 1 below illustrates proper Table and Figure formatting. Avoid placing figures and tables before their first mention in the text. IEEE has the following rules for inserting graphics as figures: The manuscript’s graphics should have resolutions of 600 dpi for monochrome, 300 dpi for grayscale, and 300 dpi for color.

Graphics should be inserted into the manuscript file using the “Insert – Picture – From File” option in the “Insert” tab. This means you must save every graphics as a separate file. Do not use cut and paste to insert graphics. Do not link to a graphic. When inserting figures or tables be sure you insert the figure and not just a link to the figure. The best way to make sure you are doing this correctly is to save your paper, then open the file on a different machine and make sure all your figures are correct. If you insert the link instead of the figure or table, a box with a big red x will appear in the location where the table or figure is supposed to be located.

TABLE I
POINT SIZES AND TYPE STYLES

Points	Place of Text	Type Styles
8	Table number	ROMAN NUMERALS
8	Table text	
8	Figure and Table Headings	UPPERCASE
8	Figure and Table Captions	SMALL CAPS
8	Footnote	
8	Reference list	
10	Footer	**Bold**
10	Abstract	**Bold**
10	Index Terms	
10	Section Titles	**SMALL CAPS, BOLD**
10	Main Text and Equations	
10	Subheadings	*Italic*, Left justified
11	Authors' names	
24	Title	Title Case

Fig.3 Logo of The Institute For Electrical And Electronics Engineers.

Number equations consecutively with equation numbers in parenthesis flush with the right margin, as in (1) ∂z = ∂u/∂2jk x∂2u/ 2 + k2 (n2 β- 2) u . (1) Refer to “(1),” not “Eq. (1)” or “Equation (1),” except at the beginning of a sentence: “Equation (1) is….” Make sure you only use the “Symbol Font” for all your symbols, or embed all your different symbol fonts within the file when you save the document.

© 2019 IJSRET

5. Headers And Footers - Select Page Layout tab, and then select Margins – More Margins. This set of actions will open the Page Setup Window; there, select the Layout tab. Set the Header and Footer to 0.5 inches. Please use the following format guidelines for the Header and Footer.

5.1 Header text- should say Session n as shown on the top of this document. Once your paper is assigned to its final session, the General Chair will modify this to the correct session. Text should be 14 point Times New Roman, right justified, bold (this style is defined under the style menu of this document as "Header")

5.2 Footer text- should be 10 point Times New Roman, bold (this style is defined under the style menu of this document as "Footer"). The text of the footer should say the same as shown on the bottom of this document. Please copy and paste this information into your document exactly as shown on this page. Once your paper is assigned to its final session, the General Chair will modify the page number to the correct session.

Acknowledgment - The preferred spelling of the word "acknowledgment" in American English is without an "e" after the "g." Use the singular heading even if you have many acknowledgments. Please put the sponsor acknowledgments in this section; do not use a footnote on the first page.

Copyright Form - An IEEE copyright form should accompany your final submission.

In Text References All material from any research resource must be accompanied by a bracketed in-text reference. This reference must correspond to its end-text full bibliographic information in the References section. Failure to properly reference all resource material used in a paper leaves the paper's author open to charges of plagiarism. Follow these specifications for in-text references: Bracket all in-text references, for example [1]. In text references must be numbered sequentially in the text, beginning with [1] for the first reference. In other words, the first source from which you quote, paraphrase or use information must be referenced in your paper as [1].

The next source from which your quote, paraphrase or use information must be [2]. If, later in your paper, you use information from the same source and same page as [1], then your in-text reference number will again be [1]. Do not use "Ref. [3]" or "reference [3]." Simply use the bracketed number. For material summarized from several sources, use the appropriate bracketed numbers, for example [3]-[5]. Bracketed reference numbers should appear after the quotation marks on an in-text quote, but before the final punctuation of the quote. For example, "Here's the quote" [3]. Bracketed references for paraphrases or summaries should appear after the paraphrase or summary, but before the final punctuation of the sentence or passage. For example, Here's the paraphrased material [4].

REFERENCES

Place references in a separate References section; do not footnote references. Number the references sequentially by order of appearance, not alphabetically. Refer simply to the reference number, as [3] or [5]-[8]. Do not use "Ref. [3]" or "reference [3]" except at the beginning of sentence: "Reference [3] shows….." Provide up to three authors' names; replace the others by "et al." Do not put figures or anything else except for author information after the references.

- **Reference text-** 8 point, Times New Roman, full justified, no space between the references (this style is defined under the style menu of this document as "References")
- Use box numbers with square brackets [ ] within text. Do not use superscripts or subscripts. Do not use ( ) for references, since these are used to refer to equations.

Use the following as the guide for references: [1] Author's Last name, First initial, Middle initial, "Title," Journal or book (italics), Vol, No #., date, pp. [2] Author's Last name, First initial, Middle initial, "Title," Journal or book (italics), Vol, No #., date, pp.

Author Information Use this section to list author information. For example: Allen Katz, Professor, Department of Electrical and Computer Engineering, The College of New Jersey Media Microblogs using Open Source Tools: An Empirical Study Eman M.G. Younis International Journal of Computer Applications (0975 – 8887) Volume 112 – No. 5, February 2015.

© 2019 IJSRET