Digital Image Based Potato Leaf Disease Detection

Tasneem Tazeen Rashid[1,12] and Sonia Corraya[2,12]
[1] Universität Bonn, Germany
s6tarash@uni-bonn.de
[2] Jagannath University, Bangladesh
sonia.corraya@gmail.com

Abstract

This paper emphasizes on digitally recognizing two most vulnerable diseases of potato: 'Early blight' and 'Late blight'. The image of the affected leaf is processed through a number of steps: blob detection, shape analysis of the detected blobs, measuring diameter of the blobs, color feature extraction and detection of the concentric ring in the affected region. Initial analysis and findings show promising results with 89.82% accuracy and this approach could contribute directly in the overall agricultural economy.

1 Introduction

Economic development and food security in most developing countries like Bangladesh significantly depend on agriculture production. Despite progresses made in agricultural techniques, farmers still face difficulties detecting different diseases during cultivation period. In Bangladesh potato alone covers 54% of total annual vegetable production [16].

In general, 30% to 50% of the food production is lost due to different diseases [17]. Early detection of diseases can ensure having a good growth of crops. 'Early blight' and 'Late blight' are two of the most common yet vulnerable diseases that cause high amount of reduction in potato production all over the world[2][3]. Early detection of these two diseases will ensure better and secure profit as well as will reduce loss in overall agriculture economy.

On the other hand, digital image processing is nowadays a very significant sector of Computer Science. Algorithm and method like thresholding, regional labeling, color feature extraction, shape detection etc. are being used in crop image analysis [8].

As per literature review of [18] , up to year 2009 there was no digital image processing based work specifically for potato leaf diseases ('Early blight' and 'Late blight'). After that, most of the relevant research works are based on plant leaf image analysis [13]. General plant diseases can be identified by advance image processing [9][10].. However, this system requires well trained data set which costs result accuracy and time [14]. Again decision is made based on three general features color, texture and edge. A more specific system for potato leaf detection is presented in [15], where output is generated based on only three components and three separate color histogram analysis is required,

[1] Designed the word-document
[2] Reviewed the documents

though area was overlooked as a strong component. Recently a different disease: zebra chip disease of crop is detected in [1]. Objective of our research is to develop a state-of the art technique for automatic potato leaf diseases identification from digital image using five specific, distinct and independent features. This paper intended to collaborate in technique of image processing to help the farmers identifying these two vital diseases in potato leaf. If this process can be anticipated in any platform of mobile, following a server system, then it can be an effective approach and can help the farmers considerably.

In the proposed method, firstly, five distinct and independent feature components are identified which are mentioned in Table 1. A decision table (Table 2) is created from the symptoms of the disease associated with these five components and then the proposed algorithm is developed based on the truth or decision table (Table 2). Finally this system is investigated for its performance evaluation. Initial performance of the proposed approach is highly encouraging with fifty five carefully selected sample images of potato leaf. Basically, this paper highlights on determining if the potato leaf is affected with 'Early blight' or 'Late blight'. And the overall accuracy is around 89.82%.

Rest of the paper is organized as follows: section 2 describes about the methodologies (basic terms, algorithm and equations) used in this research. Section 3 is about data acquisition process of this work. After that, section 4 explains the implementation of the proposed method. Results are discussed in section 5 and finally conclusion is drawn in section 6.

2 Methodology

2.1 Symptoms of 'Early Blight'

Different symptoms indicate whether the crop is being affected by 'Early blight' or not as shown in Figure 1 (a). Symptoms listed are [4]:

- Small, irregular to circular dark brown spots on the lower leaves appears as foliar symptoms. These spots may be sized from a pinpoint to 1/8 (one eight) inch in diameter.
- As the spots enlarge, they become restricted by leaf veins and take on an angular shape.
- Early in the growing season, those lesions may become larger – up to 1/2 (half) inch in diameter and may seem like a series of dark concentric rings alternating with bands of light tan tissue
- A band narrow in shape of chlorotic tissue sometimes surrounding each lesion, and extensive chlorosis of infected foliage develops over time.
- At the closing stage of the growing season, the upper leaves of infected potato plants may be sprinkled with numerous small early blight lesions and gradually, lesions may join together to cover a large area of the leaf.
- Heavily infected leaves eventually wither and die but usually remain attached to the plant.

2.2 Symptoms of 'Late Blight'

Similarly, different symptoms indicate whether the crop is being affected by 'Late blight' or not as shown in Figure 1 (b). Symptoms listed [5]:

- Small, circular to irregular in shape lesions, which show up three to five days after infection in the leaves with purple ranged color.

- The lower leaves are primarily affected, where the micro climate is more humid.
- Nevertheless, if weather conditions are favorable and the pathogen has been carried into the field by air currents they may occur on upper leave.
- Generally development of the lesions begins to on the compound leaf near the point of attachment to the petiole (which is often cupped) or at the leaf edges, where dew is retained longest.
- In time of cool, moist weather, lesions expand rapidly into large, brown or black dark spots, often surrounded by a pale green to yellow border.
- Lesions are not bounded by leaf veins and they can cause young expanding leaves can be misshaped if formed at leaf tips or edges.
- Entire leaves may become blighted and killed within a few days as new infections occur and existing lesions coalesce.
- The lesions continue developing along the length of the stem and can remain active even in hot, dry weather

2.3 Algorithm

At first, the affected regions are considered as the blobs in the leaf. Steps for detecting the diseases are given in Figure 2. After taking the image of the affected leaf, these steps are followed to achieve the desired output. MATLAB 2010a is used for the system implementation. Execution and analysis is done in Intel corei5 64 bit processor, Windows 8.

(a) (b)

Figure 1 (a) Early blight affected potato leaf [4]; (b) Late blight affected potato leaf [5] .

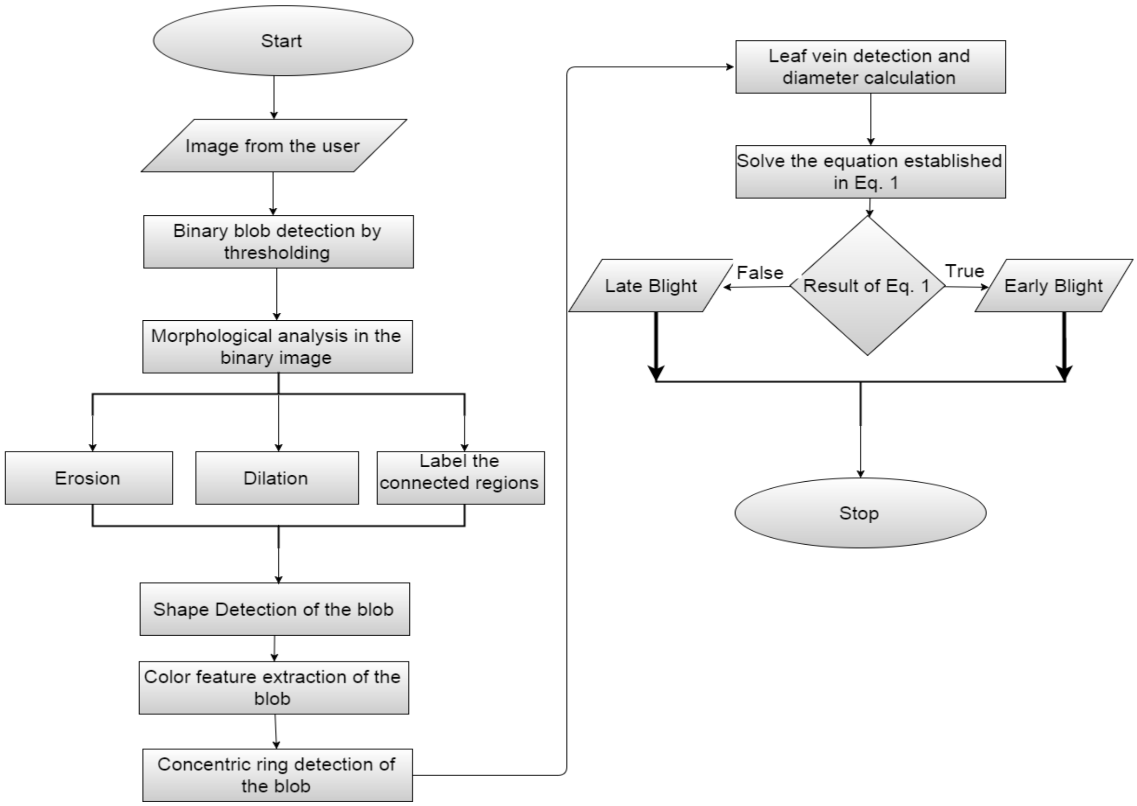


Figure. 2 Steps of digital image based potato leaf disease ('Early blight' and 'Late blight') detection

2.4 Equation

Based on symptoms of these two diseases, five different features are extracted. Through logical reasoning, an equation is generated to determine whether the leaf is affected by 'Early blight' or 'Late blight'. The logical components are set at 'True' or 'False' conditions. The relation between the features and logical components are given in Table 1. By using all 5 components presented in Table 1, a truth table is constructed as presented in Table 2.

1. Roundness, R =	{	True, if blob is round in shape False, if blob is not round
2. Color, Cl =	{	True, if blob is purple in color False, if blob is brown in color
3. Concentric ring, Cr =	{	True, if concentric ring is found in the blob False, if concentric ring is not found
4. Leaf Vein, V =	{	True, if blob is limited by vein False, if blob is not limited by vein
5. Diameter, D =	{	True, if diameter is more than 1/8 inches False, if diameter of the blob is likely 1/8 inches

Table 1: Logical condition of components

R	Cr	Cl	V	D	Output
0	0	0	0	0	EARLY BLIGHT
0	0	0	0	1	LATE BLIGHT
0	0	0	1	X	EARLY BLIGHT
0	0	1	0	X	LATE BLIGHT
0	0	1	1	X	EARLY BLIGHT
0	1	0	0	X	EARLY BLIGHT
0	1	0	1	X	EARLY BLIGHT
0	1	1	0	X	EARLY BLIGHT
0	1	1	1	X	EARLY BLIGHT
1	0	0	0	0	EARLY BLIGHT
1	0	0	0	1	LATE BLIGHT
1	0	0	1	X	EARLY BLIGHT
1	0	1	0	X	LATE BLIGHT
1	0	1	1	X	EARLY BLIGHT
1	1	0	0	X	EARLY BLIGHT
1	1	0	1	X	EARLY BLIGHT
1	1	1	0	X	EARLY BLIGHT
1	1	1	1	X	EARLY BLIGHT

Table 2: Truth table based on all 5 components of Table 1

With each five of the components from Table 1 and all the symptoms from both of the diseases, the truth table, Table 2 is constructed manually for all the combination. This truth table represents final experimental observation and investigation data. From this Table 2, using K-Map [11], an equation is generated, which is given bellow:

$$Y = Cr + R'Cr'V + RCr'Cl'V + Cr'Cl'V'(R \oplus D)' \qquad (1)$$

The output is obtained from the Boolean value of Y. If Y is true then the disease is ‘Early blight’ otherwise it is a case of ‘Late blight’. Equation (1) is the key originality and contribution of our proposed method.

3 Data Acquisition

3.1 Camera

Digital camera with minimum five megapixel resolution is used to take images of the affected potato leaf.

3.2 Data Collection

Images are collected from the department of agriculture extension, Ministry of Bangladesh [4][5]. Images are taken in a close range, focusing on the leaf disease part only [13]. Cropping is used for better focusing of the subject area in wide background [12].

4 Implementation of the Proposed Methodology

4.1 Blob Detection

Blob detection process is applied on the image as shown in Figure 3(a) using MATLAB by thresholding with respect to green pixels. There are many approaches on selecting threshold like OSTU [6]. However, this paper proposed an equation which did not need any method to calculate threshold. And this is a striong feature of this system. Pixel masking need only the below simple equation:

$$\begin{matrix}(\text{image}(i,j,1) < \text{image}(i,j,2)) \\ \&\& \\ (\text{image}(i,j,2) > \text{image}(i,j,3))\end{matrix} = \begin{cases} \text{True (1); black} \\ \text{False (0); white} \end{cases} \quad (2)$$

In this equation, "image" is a built-in method in MATLAB which takes three parameters. Here 'i' and 'j' are the pixel values of row and column respectively of the 2D image. The numerical values:

- '1' indicates the value of red.
- '2' indicates the value of green.
- '3' indicates the value of blue in RGB color sector.

Generally for green pixels, value of green component has been always higher than red and blue components. And all the pixel values of the image matching (2) are turned into black and all the others are turned into white. After this procedure the image is converted to a binary image as shown in Figure. 3(b).

4.2 Morphological Analysis

On the binary image, filling method is applied to fill up any wholes inside the white affected region. Morphological analysis for blob detection is followed by erosion and dilation first mentioned in [7]. Lastly, connected regions are labelled.

- **Erosion:** It is used to remove all the noises from the binary image shown in Figure 3(b). In this paper, default erosion method in MATLAB 2010a is used. It is observed that, more or less affected regions are oval to disk shaped. So the image is probed by structuring element of 'disk' valued with ten times. This value of ten is determined by testing all the images in the storage and ten results the minimal error compared to other trial numbers. At this point all unwanted regions are demolished, as shown in Figure 4(a).

- **Dilation:** Dilation meant expanding the shapes and probing with the structuring element. The same structuring element used in erosion stage is used for dilation. Basically, to get back the actual shapes of the non-noise part, dilation is applied. After this procedure, affected regions are extracted more clearly as shown in Figure 4(b).

- **Labeling Connected Regions**: Labeling the connected regions is used to find the connected components of a binary image shown in Figure 4(b). All the pixels in connected components are given a label. The searching of connected components is executed in the column-wise fashion. To continue the process, eight is used as the parameter. This method segmented the entire affected region in the binary image and returned each region separately as shown in Figure 5.

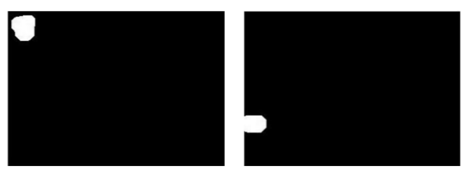
(a) (b)

Figure. 3 (a) Original 2D color (RGB) image, (b) Converted binary image

(a) (b)

Figure. 4 (a) Binary image after erosion, (b) Binary image after dilation

(a) (b)



(c) (d)

Figure. 5 Separated blobs after dilation

4.3 Shape Detection

At this stage the region properties are extracted from a method called ‘regionprop’ in MATLAB. Shapes of the blobs are identified here. 2 conditions are checked to resolve this factor.

- **Count ratio:** Ratio of maximum and minimum length of the blob’s axis is found. With the images in the storage, it is observed that roundness is true if value of the ratio is less than 1.7

$$\text{Ratio} = \frac{\text{maxAxixLength}}{\text{minAxixLength}} < 1.7$$

- **Calculate circularity:** ‘Area’ and ‘Perimeter’ of the region are spotted and circularity is measured with the mathematic formula of :

$$\text{Circularity} = \frac{\text{Perimeter}^2}{4*\pi*\text{Area}} < 1.65$$

It is found with the existing storage that if the value is less than 1.65, then the region is considered round of the affected region.

Both the conditions should be true to declare the roundness of the blobs to be true. This is because; blobs are not perfectly round in case of early blight shown in Figure. 6. To ensure that semi round shapes are figured correctly to match the equation, both the conditions are checked.

4.4 Color Feature Extraction

Color feature extraction is done by getting the value from an image histogram [7]. The separated images found in the labelling stage are mapped into the RGB image one by one to extract RGB blobs. Each RGB blob image is cropped by a fixed size from the centroid. Then the cropped image is converted into HSV color scale. Following that, the HSV values are plotted in a histogram. From the comparative analysis, it is determined that 'Brown' regions are valued at 0.2 and 'Purple' regions are valued at 0.8. Average of these two values is considered as the threshold to determine the color scale of the blobs.

$$\text{Threshold} = \frac{0.2+0.8}{2} = 0.5$$

If the value is less than 0.5, then the region color is brown otherwise the color is purple.

4.5 Concentric Ring Detection

Appearance of concentric ring helps to differentiate between 'Early blight' and 'Late blight'. Only 'Early blight' has concentric rings. To determine whether the affected blob has any concentric ring, a very easy yet accurate algorithm is used. For this procedure two sets of images are taken.

Image with RGB color blob is shown in Figure 7(a) and Figure 8(a) .Image with the binary blob is shown in Figure 7(b) and Figure 8(b)

The RGB color blob image is then converted to binary as shown in Figure 7(c) and Figure 8(c) with the default method in MATLAB. Area is calculated for all the binary images. For 'Early blight' the binary converted image of blob shown in Figure. 7(c), has lesser white pixels and area then the binary image of blob shown in Figure. 7(b). On the other hand, for 'Late blight' the white pixel and area difference between Figure 8(c) and Figure 8(b) is comparatively very small. Thus, if the difference between these 2 binary images is more than 60% (more specifically exact 0.58) then the blobs have concentric rings.

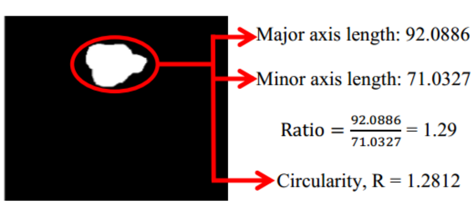
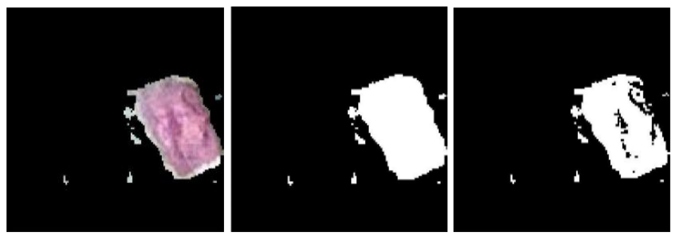


Figure. 6 Shape detection of a blob

(a) (b) (c)

Figure. 7 (a) RGB color blob for ‘Early blight’, (b) Binary blob for ‘Early blight’, (c) Binary conversion of RGB color blob



(a) (b) (c)

Figure. 8 (a) RGB color blob for ‘Late blight’, (b) Binary blob for ‘Late blight’, (c) Binary conversion of RGB color blob

4.6 Leaf Vein Extraction and Diameter Calculation

Currently, the system implemented in this paper is basically emphasized on the first three features. In this initial stage, whether the blobs are bounded by the veins and the calculation of the diameter of the blobs are provided by user.

5 Result and Discussion

For performance evaluation of our proposed method, system outcome is compared and verified with agriculture experts’ feedback. Each image is categorized both by manual investigation by expert and by the new proposed system. Finally both results are compared for each of the 55 sample images and the investigation findings are summarized in Table 3, and in Table 4 covering both qualitative and quantitative investigations.

The tables above basically emphasize on the comparison of expected output regarding the disease of the leaf according to an expert’s opinion and the output from the system implemented. Percentage of matched and mismatched outcomes of the both side is recorded. Expert’s opinion is taken as to the best of our knowledge there is no existing similar system for comparison.

We also have explored and identified specific reasons for those special cases resulting 18.18% unsuccessful rate.

Agriculture expert’s opinion and proposed system’s output	Percentage
Match	89.82%
Mismatch	11.18%

Table 3: System output validation result

5.1 System Performance Limiting Reason Analysis

Identified reasons for result variation between agriculture expert's opinion and system output and their explanation as well as possible solutions are discussed in here. Reasons are:

1. Blob position inside leaf vein: Some images of 'Late blight' in initial stage are found to be similar to symptoms of 'Early blight'. As per expert agricultural personnel, Figure 9 is an image of late blight in early stage. However, as per Table 2 (row number 10) the individual component is matching with the condition of early blight.
2. False detection of concentric ring: The system implemented in this paper results that the affected blob of Figure 10 has a concentric ring. So, as per Table 2 (row number 15), this blob would be detected as an early blight whereas this is neither 'Late blight' nor 'Early blight'. The reason for this is, the detection process of concentric ring of this work only compare the value of white pixel. For which this image falls inside the formulated range and declares the blob as having ring. To solve this problem in next version of our work, we will consider the continuity and adjacency of neighbor pixels for ring detection.
3. Overlooking the presence of concentric ring: Sometimes for high brightness value of input images, concentric rings are not clearly visible. To handle such scenarios, picture resolution and background light should be carefully adjusted while capturing the image or specific brightness limit can be set particularly for this purpose.
4. Recognition of background as blob: Green pixel masking formula given in Equation 2 might detect the background as blob though it is very rare. For this, images should either be taken of a single leaf or cropped with the disease affected parts only.

Table 4 represents the individual percentage of images which produce incorrect results for specific reason. It is need to be noted that a single image might produce incorrect result for more than one reason mentioned in section 5.1. This analysis is important for probability and tolerance level set.

Reason for limiting output	Percentage
Blob position inside leaf vein	44.44%
False detection of concentric ring	55.56%
Overlooking the presence of concentric ring	22.22%
Recognition of background as blob	11.11%

Table 4: Percentage of Image Which Resulted Incorrect Output

Figure. 9 Early stage of 'Late blight' affected leaf

Figure. 10 Potato leaf affected by deforming rust disease

Of course as a human work, our system has some limitations which are the challenges and future working scopes. As mentioned in section 4, diameter calculation and testing of blobs bounding by the leaf vein or not are not fully automatic. Again other edge detectors (like Sobel, Canny etc.) could be tested for shape detection to identify the most appropriate one among different edge detector and MATLAB function 'regionprop'. With more training and investigation of image data, Ground truth image database could be generated for 'Early blight', 'Late blight' and potato leaf shape boundary which would increase accuracy in shape analysis outcomes. Only affected images need to be input in the system as the proposed system is for detecting specific disease, not for identifying affected or unaffected leaves. For including this, in our next version of this system, we plan to differentiate only the 'Early blight' and 'Late blight' affected images at the very early step of inputting. This could be done by filtering images by checking the presence of purple (0.8) and brown (0.2) color values. Adding to that, the observations from section 5.1 will also be a great aspect for future exploration.

6 Conclusion

A new approach to detect potato leaf diseases: 'Early blight' and 'Late blight' is presented in this paper. With digital image processing this research is aimed to determine only two leaf diseases of potato: 'Early blight' and 'Late blight' initially. Other crop diseases will be analyzed in extended versions of this work. At first five distinct and independent feature components are identified. A decision table is created from analysis result of these five components (Roundness, Color, Concentric ring, Leaf vein and Diameter) and then the proposed algorithm is developed based on the truth or decision table. Performance evaluation result of the proposed approach with the experimental data (fifty five carefully selected sample images of potato leaf) is highly promising with an overall accuracy of around 89.82%. Reasons affecting output results are also identified and analyzed thoroughly. Initial findings show promising results however, with an intense storage, this approach can be improved by further investigation with larger number of sample image to get more accuracy even if the images have other environmental noises. Early detection of these two diseases with this new method will ensure better and secure profit as well as will reduce loss in overall agriculture economy.

References

[1] Pei-Shih Liang, Ronald P.Haff, Sui-Sheng T.Hua, Joseph E.Munyaneza, TariqMustafa, Siov Bouy L.Sarreal. Nondestructive detection of zebra chip disease in potatoes using near-infrared spectroscopy. Published by Elsevier -Biosystems Engineering,Volume 166, February 2018, Pages 161-169

[2] L. Liu and G. Zhou; Extraction of the rice leaf disease image based on bp neural network. International Conference on Computational Intelligence and Software Engineering, 2009, pp. 1-3.

[3] P. Wharton and W. Kirk. Early blight, Department of Plant Pathology, Michigan State University, Extension Bulletin E-2991, New, May 2004 Available online: http://www.potatodiseases.org/earlyblight.html [Retrieved on 26 January, 2016].

[4] W. Kirk, P. Wharton, R. Hammerschimidt, F. A. Salman and D. Douches. Late blight, Department of Plant Pathology, Michigan State University, Extension Bulletin E-2945, New,

December 2004. Available online: http://www.potatodiseases.org/lateblight.html [Retrieved on 26 January, 2016].

[5] N. Otsu; A threshold selection method from gray-level histograms. IEEE Transactions on System, Man and Cybernetics, vol.smc-9, no.1, January 1979.

[6] S. Naikwadi and N. Amoda; Advances in image processing for detection of plant disease, vol. 2, Issue 11, November 2013, ISSN 2319 – 4847, 2013.

[7] R. C. Gonzalez and R. E.Woods; Digital image processing. Englewood Cliffs, NJ: Prentice Hall, 1992.

[8] P. Salembier and J. Serra. Flat zones filtering, connected operators and filters by reconstruction. IEEE Transaction on image processing, vol. 4, no. 8, August 1995.

[9] R. Chakravarti and X. Meng. A study of colour histogram based image retrieval, Sixth International Conference on Information Technology: New Generations, 2009.

[10] A. Vibhute and S. K. Bodhe. Applications of image processing in agriculture: a survey, International Journal of Computer Applications (0975 – 8887), vol. 52 - no.2, August 2012.

[11] M. M. Mano. Digital logic and computer design, 4th edition.

[12] J. W. Orillo, J. Dela Cruz, L. Agapito, P. J. Satimbre and I. Valenzuela; Identification of diseases in rice plant (oryza sativa) using back propagation. Artificial Neural Network. International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management, 2014, Palawan, 2014, pp. 1-6, 2014.

[13] S. Weizheng, W. Yachum, C. Zahliang and W. Hongda. Grading method of leaf spot disease based on image processing. International Conference on Computer Science and Software Engineering, 2008, , Wuhan, Hubei.

[14] G. M. Choudhury and V. Gulati. Advance in image processing for detection of pant diseases. International Journal of Advanced Research in Computer Science and Software Engineering, vol. 5, Issue 7, July 2015, ISSN: 2277 128X, 2015.

[15] G. Athanikar and P. Badar; Potato leaf diseases detection and classification system. International Journal of Computer Science and Mobile Computing, Vol. 5, Issue. 2, February 2016, pg.76 – 88, ISSN 2320–088X, 2016.

[16] M. M. I. Masum, S. M. M. Islam, M. S. Islam and M. H. Kabir. Estimation of loss due to post harvest diseases of potato in markets of different districts in Bangladesh, African Journal of Biotechnology, Vol. 10(56), pp. 11892-11902, 26 September, 2011, ISSN 1684–5315 © 2011 Academic Journals, 2011.

[17] M. T. Hossain, S. M. M. Hossain, M. A. Bakr, A. K. M. M. Rahman and S. N. Uddin; Survey on major diseases of vegetable and fruit crops in Chittagong region. Bangladesh J. Agril. Res. 35(3): 423-429, September 2010, ISSN 0258-7122, 2010.

[18] J. K. Patil and R. Kumar; Advances in image processing for detection of plant diseases. Journal of Advanced Bioinformatics Applications and Research ISSN 0976-2604 ,Vol 2, Issue 2, June-2011, pp 135-141, 2011.