The Development of Diseases Identification System in Paddy Plant Using Image Processing Technique

Alex Wenda[1*], Nanda Putri Miefthawati[1], Mas'ud Zein[2]

[1]Electrical Engineering Department, Faculty of Science and Technology, UIN SUSKA Riau
[2]Educational Islamic Religion S2 Programme Department, Tarbiyah and Teacher Training Faculty, UIN SUSKA Riau
Jl. HR. Soebrantas No. 155 Simpang Baru, Panam, Pekanbaru, 2829
Email: alexwenda@uin-suska.ac.id[1*], nandamiefthawati@uin-suska.ac.id[1], mas'ud.zein@uin-suska.ac.id[2]

Abstract. There are three types of paddy leaf disease that have similar symptoms, making it difficult for farmers to identify them, namely Blast Disease, Brown-Spot Disease, and Narrow Brown-Spot Disease. This paper aims to develop an application to identify paddy leaf disease automatically. Several important aspects of the development of software engineering such as usability, interactivity, and simplicity have been considered. Image processing techniques, namely Blobs analysis and color segmentation are used to get the characteristics of diseased leaf; these characteristics are then used to identify the type of diseases using a rule-based expert system. The results obtained indicate that the developed system recognition capability is considered satisfactory with an accuracy of 94.7%.

Keywords: Blast Disease, Blobs, Brown-spot, Interactivity, Narrow brown-spot, Simplicity, Usability

INTRODUCTION

A quality paddy plant with a high level of productivity is produced by paying attention to many things including land processing, seed selection, seeding, maintenance and control of pests and diseases. If the selection of seeds and land processing has been done properly, the most important thing in increasing the amount of paddy production is to control the pests and diseases of paddy plants (Bbpadi, 2016).

There are various types of pests and diseases in paddy plants, including blast, brown spot, narrow brown spot, Sheath Blight, Bacterial Leaf Blight, Tungro, white pest, stem borer, rat pest, green ladybug and stink bug (Scardaci et al. 1997). In this study will focus on 3 types of paddy disease, blast disease, brown spot disease and narrow brown disease. Three types of this disease have characteristics that are almost the same as each other so that this makes it difficult for farmers to identify the disease and require special expertise to be able to distinguish the three types of disease (Liu Zhan 2007, Kobayasi et al. 2001). Apart from that, the three types of disease are very widespread and can be found in more than 80 countries (Robert W, 2001). The most important thing of the three types of disease is the result if the paddy plant has been infected with this disease, namely a marked decrease in yield, because it causes the panicle to rot or break, this will inhibit the process of filling the grain, causing the grain to become empty. This will certainly lead to a decrease in paddy production and if production decreases, it will certainly have an impact on losses. The disease caused by this fungus is estimated to have caused a production loss of US $ 55 million every year in South Asia and Southeast Asia. Losses are even higher in the East Asia region and some areas that have higher climate temperatures (Liu Zhan, 2007).

Image processing is a processing process whose input is the image and the output can be an image or a set of characteristics or parameters related to the image. Image processing has several functions, including being used as a process to improve image quality so that it is easily interpreted by humans or computers. Several studies that have been carried out in the field of paddy farming that use image processing techniques are the development of diagnoses of paddy plant disease systems (Rossilawati et al. 2003), diagnosis of paddy plant disease system (Nurul Ashikin, 2002), expert systems for paddy management (Edrees et al. 2001, Phadikar 2008, Kurniawati et al. 2009).

Based on the background above, this study aims to develop applications that can identify leaf diseases in paddy plants, namely blast disease, brown spot disease and narrow brown spot disease using image processing techniques by analysing the texture of diseased paddy leaf images using Blobs analysis, then analysing the color texture paddy leaf use color segmentation techniques to get the characteristics of diseased leaf, this characteristic is then used to identify the type of disease using a rule-based expert system.

MATERIALS AND METHODS

This methodology is divided into two important parts, the first is about algorithms for detecting paddy diseases and the second is issues related to application design. The application development framework for identifying paddy plant diseases is shown in Figure 1.
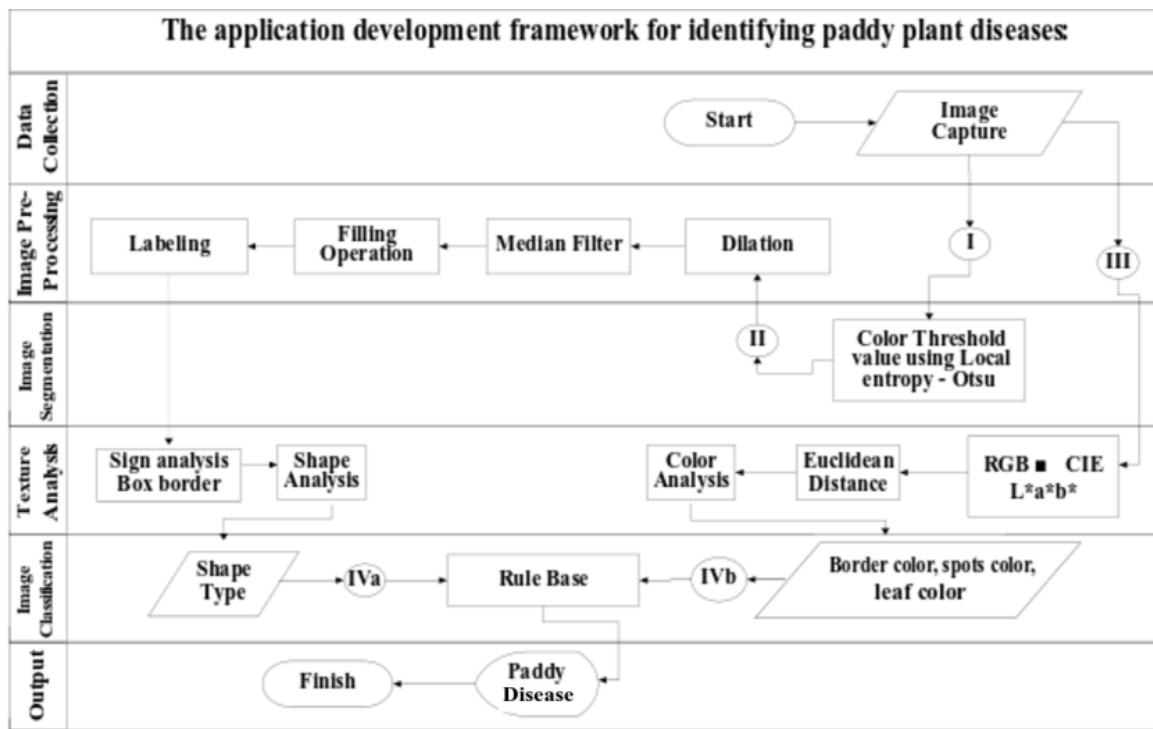
Figure 1. Development frame work.

Data Collection

Capturing images of diseased leaf paddy is the first part of this study. To get an image that has a uniform brightness level, the environment must be controlled, some important parameters are, flash camera and distance of the object from the camera. The distance between the object and the camera is between 15-30cm and using the flashlight on the camera. The image was taken during the day and the process of taking the paddy leaf image must be carried out quickly after the paddy leaf are cut from the tree. Otherwise, the leaf will immediately roll. Image of paddy leaf is resized into images with a resolution of 109 × 310 and stored in bitmap format (* .BMP). Examples of sample images and characteristics of the lesion to diseased paddy leaf as shown by Figure 2.
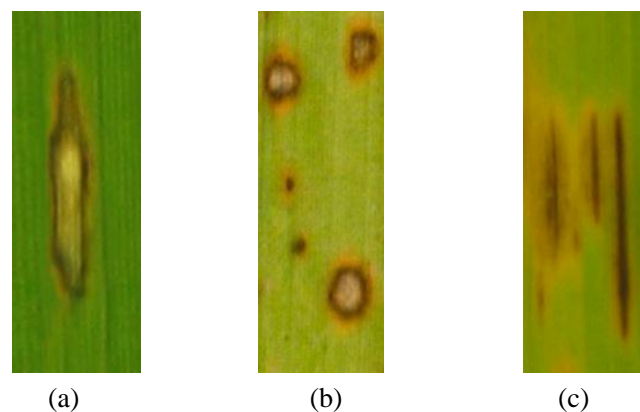

Figure 2. (a) Blast disease, (b) Brown spot disease (c) Narrow brown spot disease.

Image Pre-Processing and Segmentasion

The pre-processing task involves some procedures to prepare the images enhancement. Median filter and morphological operators are applied to remove unnecessary spots by using a region filling technique. As a result, a noise free binary image is produced.

The image segmentation method in this study is by combining the local entropy color threshold with the Otsu threshold method (Kurniawati et al. 2009). The combined algorithm of local entropy and Otsu color threshold values as shown in algorithm 2.1.

Table 1. Algoritma 2.1 Image segmentation algorithm; Combines the local threshold value method of entropy and Otsu.

Algoritma 2.1 Image segmentation algorithm; Combines the local threshold value method of entropy and Otsu		
Input: Image of paddy leaf diseased (image RGB-MHB), $img0$		
Output: Image binary, img_P		
1:	Get all the pixel values of the original image based on the amount of red, green and blue. Calculate the value of a two-dimensional matrix of events histogram for each color pixel in the image size N × N, where N is the maximum value of the grayscale.	
2:	Calculate the local entropy value based on the two-dimensional event matrix. If M is chosen as the threshold value, then the matrix for the four parts (Quadrant A, B, C, and D). Calculate the local entropy value in quadrants A and C. The threshold value will maximize the amount of entropy for each quadrant that represents the selected object and background. Suppose that P_{ij} is a matrix, it is normalized so that the number of values for each quadrant is 1. Determine a bright object is an element in quadrant C divided by P_o and elements in quadrant A divided by P_B. $$P_{ij}^O \equiv \frac{P_{ij}}{P_O}; \quad P_{ij}^B \equiv \frac{P_{ij}}{P_B}$$ with $$P_O \equiv \sum_{i=m+1}^{N-1} \sum_{j=m+1}^{N-1} P_{ij}; \quad P_B \equiv \sum_{i=0}^{M} \sum_{j=0}^{M} P_{ij}$$ Thus, the entropy values of $E_O(M)$ and $E_P(M)$ for the object and background is $$E_O(M) \equiv \frac{1}{2} \sum_{i=0}^{M} \sum_{j=0}^{M} P_{ij}^O \log_2 P_{ij}^O$$ The threshold value will maximize the amount of entropy $$E_T(M) \equiv E_O(M) + E_B(M)$$	
3:	Each value entropy color Red, green and blue will be set as follows: $$\rho(x,y) \equiv \begin{cases} 1 \ if & \bar{h}(x,y) \geq E_{M,H,B} \\ 0 \ else & \bar{h}(x,y) < E_{M,H,B} \end{cases}$$ with, $\bar{h}$ is the input image; ρ is image output,; (x, y) is the pixel point, E_{MHB} is threshold value entropy.	
4:	Change the entropy image to a binary image using the Otsu method and do the process to produce a binary image (img_P) with a white object and a black background.	

Image Texture Analysis

Image texture analysis is divided into two parts, namely shape texture analysis to get the type of lesion and color texture analysis to get the boundary color, lesion color, and leaf color lesion.

Texture analysis of object shape in this study was used to get symptoms of types of the lesion from diseased paddy leaf. There are five types of lesion, namely spindle, round, oval, spot, and taper. Blobs analysis is used to calculate the height and width values of the object boundary box. To determine the height and width of the object as 100 images have been analyzed, consisting of 20 images for each type of lesion. Each image is analyzed to get the height and width values than obtained a reference to determine the type of lesion for each form of lesion.

Color texture analysis in this study was used to obtain color characteristics of diseased leaf boundaries, spots, and color of the diseased leaf. The object boundary color (diseased part) is obtained by specifying 8 coordinates in different locations of the object boundary, namely left-top, top-left, top-right, right-top, right-down, down-right, down-left coordinates, left-down, and top left. According to

experts, for spindle shapes, oval shapes and round shapes, the color of the diseased leaf must be determined, while for the tapered shape and the shape of the boundary color spots lesion is not needed. The color of the lesion spots is obtained by determining the center coordinates of the object. While the leaf color lesion is obtained by determining the coordinates of each pixel in the leaf image except the pixels on the object.

Color texture analysis is used to determine the color of the diseased object boundary, spots, and leaf color by changing the RGB color space into CIEL * a * b * color space and calculating the Euclidean distance to get an identical color with the color determined by the expert. Experts have set visually related colors in determining the color of the boundary, the color of the spots and the color of diseased leaf. These colors were obtained from previous researchers [10]. There are three categories for boundary color lesion, namely brown, orange and yellow. In lesion spots, there are two color categories, namely gray and brown. While in leaf color, there are four color categories, namely brown, orange, yellow and green. The color texture algorithm in this study is shown in algorithm 2.2.

Table 2. Algoritma 2.2. Color texture analysis algorithm.

Algorithm 2.2 Color texture analysis algorithm	
Input: Image of diseased paddy leaf (RGB image), $img0$ and binary image, img_P	
Output: Boundary color, spot color, and color of the leaf, ws, wb, wd	
1:	Determine the number of objects, bil from binary image, img_P using labeling operations.
2:	Determine the 8 coordinate boundary of each object, namely top-left, top-right, right-top, right-down, down-right, bottom-left, left-down, and top left. Spot coordinates and leaf coordinates.
3:	Get RGB values for each boundary coordinates, spots, and leaf. Change the RGB color space to CIEL * a * b * color space using, $$R\,G\,B\ \left(R_{ij}, G_{ij}, B_{ij}\right) \xrightarrow{\text{color space CIEL}*a*b*} (L^*_{ij}, a^*_{ij}, b^*_{ij})$$
4:	Determine the value $L^*a^*b^*$ boundary color, spot color, and leaf color.
5:	Calculate Euclidean distance (ΔE), to calculate the colors most similar to color rather than experts.

Image Classification

This study uses only three classes according to the type of disease, namely blast disease, brown spot disease, and narrow brown spot disease, so the use of knowledge-based systems as a method in image classification is very suitable to be used.

Rules are generated based on the characteristics of each class. In this study, there are four characteristics that paddy disease symptoms are the type of lesion, boundary color, spot color and leaf color. If the type of lesion that is produced is a spindle, oval shape and round shape then the rules that will be produced have 4 conditions. This is because, according to [10], these three types of shapes require the use of boundary color. In contrast to the tapered shape and spot shape, the resulting rule has a total of three terms, which do not require the use of boundary color. Then the number of rules generated if he has 4 conditions are:
Number of Rule =Type of shape $\times$ spot Color $\times$ boundary Color $\times$ leaf Color
Number of Rule =$3 \times 2 \times 3 \times 4 = 72$.

Whereas for those who have 3 conditions, the number of rules is:
Number of Rule = Type of shape $\times$ spot Color $\times$ leaf Color
Number of Rule =$2 \times 2 \times 4= 16$.

Then the total rule generated is as many as 84 rules. To facilitate the classification process, all of these rules will be stored in the database as a knowledge base.

Interface Design Issue

There are three important elements should be considered in order to produce a good user interface design, which is, the guidelines of documents and processes, the user interface aided-software and evaluation of usability. In order to get a good interface, in this study user-based evaluations were used which included three factors: Usability Factor, Simplicity and Interactivity.

Usability means ease of use, including such measurable attributes as learnability, the speed of user task performance, user error rates, and subjective user satisfaction.

Simplicity does not mean lack of functionality. The goal for a good user interface design should be easy to learn but have a built-in path to higher proficiency. The golden rule of simplicity is to design such a system with components which will actually be used by the user, without any unnecessary extra elements.

Interactivity factor is another important thing that should be considered. An interactive and attractive appearance of the user interface will give users a comfortable feeling to further using the system. It should react to the users in a way they expect the system should react. The style of interaction offered by a user interface is actually the dialog used to communicate human and computer. The most common styles of interaction are menu, instructions language, WIMP (window, icon, menu, pointing device), form-filling, questions and answers dialog, 3D interface and direct manipulation. Each of these dialog types has their own structures and standard parameters.

RESULTS AND DISCUSSION

This research has successfully developed an application to identify the paddy diseases using image processing techniques. Figure 3 shows the main user interface, the upload button is used to insert images of diseased paddy leaf, the "original Image" is a place to display

the original image of paddy leaf, the "Binary Image" is a place to display binary images. The Button Process is used to perform the analysis and identification process. Text box "Penyakit" is a place where the type of disease is identified. The "Diagnose" button is used to diagnose in detail.
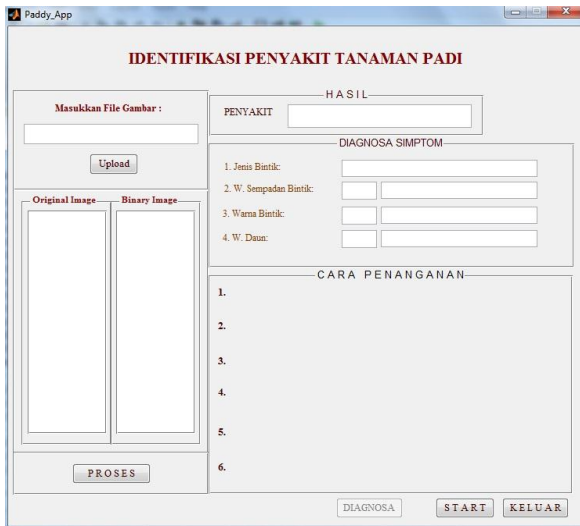

Figure 3. Main user interface.

Figure 4 shows the process that occurs after insert the diseased paddy leaf into the system and press the process button. From this picture it can be seen that this type of disease is Blast, the type of shape is spindle, the color of the bordering is yellow, the color of spots is brown and the color of the leaf is orange.
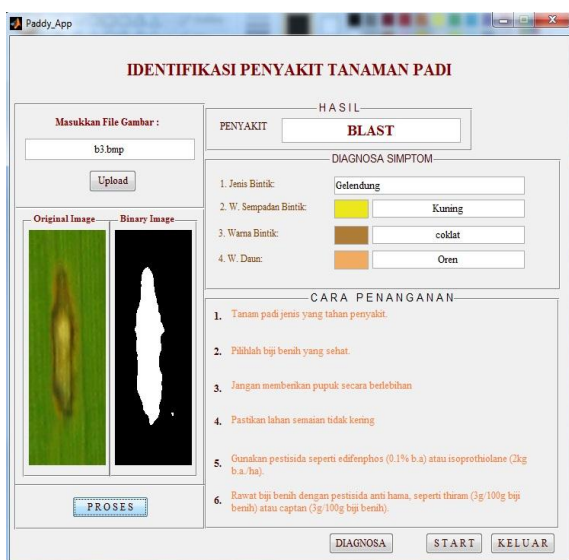

Figure 4. Result.

If we want to diagnose in more detail, there is a "Diagnosa" button. A detailed diagnosis of each process will appear as shown in Figure 5 where the image segmentation process is displayed in detail, such

as the shape of the lesion, object height, object width, object shape, color spots, and bordering color spots. More detail for the bordering color also displays the color around spots and leaf color.

Figure 5. Detail analysis.

To validate the application that has been developed, it has been tested on 94 samples of images of diseased paddy leaf. As a result, 89 paddy leaf images can be analyzed and identified the type of disease correctly, while the remaining 5 paddy leaf are not identified correctly. So that the level of accuracy of the applications that have been developed is 94.7%.

This study has also conducted user-based evaluations where the application developed was tested by 35 respondents. Each of them has been asked to answer a questionnaire consisting of ten questions to get their feedback from experience using the developed application. Some important aspects of good software design such as usability, simplicity, and interactivity were evaluated during the survey. Likert scale, starting from 1 (not good) to 5 (good) used in this survey. The average score for each question is around 3.6 to 4.6, which indicates that all the factors that have been evaluated during this questionnaire can be considered to function well. Generally, users are satisfied and find that this software is good, consistent and achieves its objectives in identifying and analyzing paddy plants.

CONCLUSIONS

Overall this research has successfully achieved the goal to be achieved is to develop an application to identify the disease of paddy plants using image processing techniques. The results of application testing in identifying paddy plant diseases have an accuracy rate of 94.7%.

While the results obtained from the perspective of user evaluation are considered satisfactory where most of the questions asked in the form of questionnaires obtained an average score of 4, which indicates good performance. The software developed can be used to identify paddy plants.

REFERENCES

BBpadi, http://bbpadi.litbang.pertanian.go.id/index.php/berita/info-teknologi/content/240-penyakit-blas-pada-tanaman-padi-dan-cara-pengendaliannya

Edrees SA, El-Sayed El-Azhary, Rafea, A. 2001. Expert System for Paddy Production Management. 9th International Conference of Artificial Intelligence :155-162.

Kurniawati, N. N., Abdullah, S. N. H. S., Abdullah, S., & Abdullah, S. (2009). Investigation on image processing techniques for diagnosing paddy diseases. In SoCPaR 2009 - Soft Computing and Pattern Recognition (pp. 272-277)

Kobayashi, T., Kanda, E., Kitada, K., Ishiguro, K., Torigoe, Y., 2001. Detection of rice panicle blast with multispectral radiometer and the potential of using airborne multispectral scanners. Phytopathology, 91(3):316-323.

Liu Zhan-Yu et al. 2007, Characterizing and estimating rice brown spot disease severity using stepwise regression, principal component regression and partial least-square regression*, Journal of Zhejiang University SCIENCE B., pp. 738-744.

.

Nurul Ashikin, M.N. 2002. Sistem Pakar Diagnosis Penyakit Padi (SPDPP). Tesis Sarjana Muda. Universiti Kebangsaan Malaysia.

Noorashikin Binti Mustafa. 2006. Sistem Pakar Kabur Berasaskan Imej bagi Diagnosis Penyakit Padi. Tesis Sarjana. Universiti Kebangsaan Malaysia.

Phadikar, Santanu & Sil, Jaya. 2008. Rice Disease Identification using Pattern Recognition Techniques. Proceeding of 11th International Conference on Computer and Information Technology ICCIT'08 :420-423.

Robert W. Herdt, 2001. Research Priorities for Rice Biotechnology, in Rice Biotechnology, G.S. Khush and G.H. Toenniessen (eds.), Alden Press Ltd., London, pp. 35-37.

Rossilawati, S., Siti Norul Huda, S.A., Mohammed, Y., Azuraliza, A.B., Sharizan, R., Noorashikin, M., Saad, A. & Nik Mohd Noor, N.S. 2003. The development of diseases diagnosing system in paddy plant (E-Paddy). Prosiding Antarabangsa Teknologi Maklumat (ITSIM'03), hlm. 424-432.

Scardaci SC, Webster RK, Greer CA, Hill JE, William JF, Mutters RG, Brandon DM, McKenzie KS, Oster JJ, 1997, Rice blast: A new disease in California. Agronomy Fact Sheet Series. 1997-2. Departement of Agronomy and Range Science, University of California, Davis.