1

Implementation of Gray Level Co-occurrence Matrix on the Leaves of Rice Crops

Lilis Indrayani[1], Raden Wirawan[2]

[1]STMIK Kreatindo Manokwari, Jl. Kali Bambu Reremi Puncak, Manokwari, Papua Barat, 98314, Indonesia
[2]STMIK Bina Adinata, Jl. Serikaya No.08, Bulukumba, Sulawesi Selatan, 92513, Indonesia

ARTICLE INFO

ABSTRACT

Rice is one of the cultivation plants that are very important for human survival, the success of rice harvesting affects the level of farmers' income. But farmers often suffer losses, as a result of illness in rice. Rice plants infected with the disease will show symptoms in the form of patches that have certain patterns and colors on some parts of the body of rice such as stems, leaves, and roots. Disease symptoms that arise on the leaves are most easily identified because the leaves have a wider cross-section than other body parts of rice. Therefore, in this study leaf is used as an initial step parameter for disease detection in rice. This research aims to identify diseases that exist in rice plants using the method of Gray Level Co-occurrence Matrix (GLCM). The GLCM method is a feature extraction method. The disease detection process on the leaves of the rice plants is done by retrieving the original image for the initial step, then the original image is segmented before converting to greyscale imagery. Once it is done extraction features by using GLCM features namely Entropy, Eccentricity, Contrast, Energy, Correlation, Homogeneity. The results showed 90% accuracy results using GLCM extraction. The introduction of the resulting disease can help to recognize the type of rice leaves infected with the disease.

Corresponding Author:

Raden Wirawan,
STMIK Bina Adinata, Jl. Serikaya No.08, Bulukumba, Sulawesi Selatan, 92513, Indonesia
Email: liliraden12790@gmail.com

1. INTRODUCTION

For farmers with rice, planting is their livelihood, Rice harvest success affects the level of farmers' income because most farmers rely on their life from the rice harvest. Paddy has a scientific name of Oryza sativa and belongs to the tribe of paddy-paddy or Poaceae. Paddy produces rice which is the staple food of most of our nation, so rice is one of the fields of agriculture that affects daily life. However, farmers often suffer losses [1][2], but one factor is the disease of rice. Although farmers have already gained some training and knowledge of how to care for and know the disease of paddy plants but mistakes remain occasionally occur in determining the disease [3][4]. This error occurs when human abilities are limited in knowing the disease visually [5]. Also, the characteristics of paddy disease are almost identical between one disease and the other. Referring to previous research by Dadi Rosadi and Asril Hamid under the title of research on rice crop diagnosis system using Forward Chaining method, which was designed using the Borland Delphi 7 programming language [6] And Sri Wulandari et al with the title of research system for diagnosis of pests and diseases of rice crops with Bayes method [7].

One factor of declining rice production is the disease that affects rice crops, especially in the leaves. Types of diseases in the leaves of rice crops are Blast, leaf blight, Tungro, Leaf burned [5]. The symptoms of diseases that arise in rice leaves are most easily identifiable because rice leaves have a wider cross-section than other parts of the rice body So that discoloration and spot-shape can be visible [4]. Therefore, the rice leaves can be used as the first step of disease detection in rice. With the problem, the researcher made a system for the introduction of rice crop disease using leaf parameters. For this study applied the GLCM (Gray-Level Co-occurrence Matrix) method. GLCM is used for the extraction of rice leaf characteristics by using Matlab.

This research develops the image implementation using the GLCM method with feature extraction of 6 features, contrast, eccentricity, energy, homogeneity, entropy, and correlation with angles of 0 °, 45 °, 90 °, 135 ° and detect four types of rice disease are Blast, Leaf Blight, Burning Leaves, and Tungro.

2. RESEARCH METHOD

2.1. System Design

The first built system user will open the application created using Matlab application [8][9]. The steps to be used in this study are 1) image Capture (Image Acquisition), 2) image Improvement (preprocessing), 3) feature extraction, and 4) object identification [10]. It can be shown in Figure 1. The object will be taken in the form of a digital image in the "JPG" format [11]. Improved imagery by doing changes to the pixel's intensity with grayscale function and resize. Then the result of image improvement will be extracted feature with GLCM, and finally for the introduction of the object with the closest distance calculation (Euclidean Distance) [12]. The value of each object's characteristics will be calculated, by contrast, Eccentricity, Correlation, Homogeneity, and Energy [13][14][15]. Euclidean Distance is an imagery premeasurement for measuring the most commonly used resemblance, usually calculated from raw data, not from standardized data.

To configure GLCM, the co-occurrence matrix for gray levels often calculates the intensity of the pixel (gray level) of n in a spatial relation to a pixel of m, primarily the spatial relationship defined as a pixel of interest and its adjacent pixel horizontally on the right directly, the element (n, m) produced in the co- occurrence matrix is simply the sum of the number of pixels that have a value of n in the spatial relationship specified for the pixel of m in the input image [16][17].
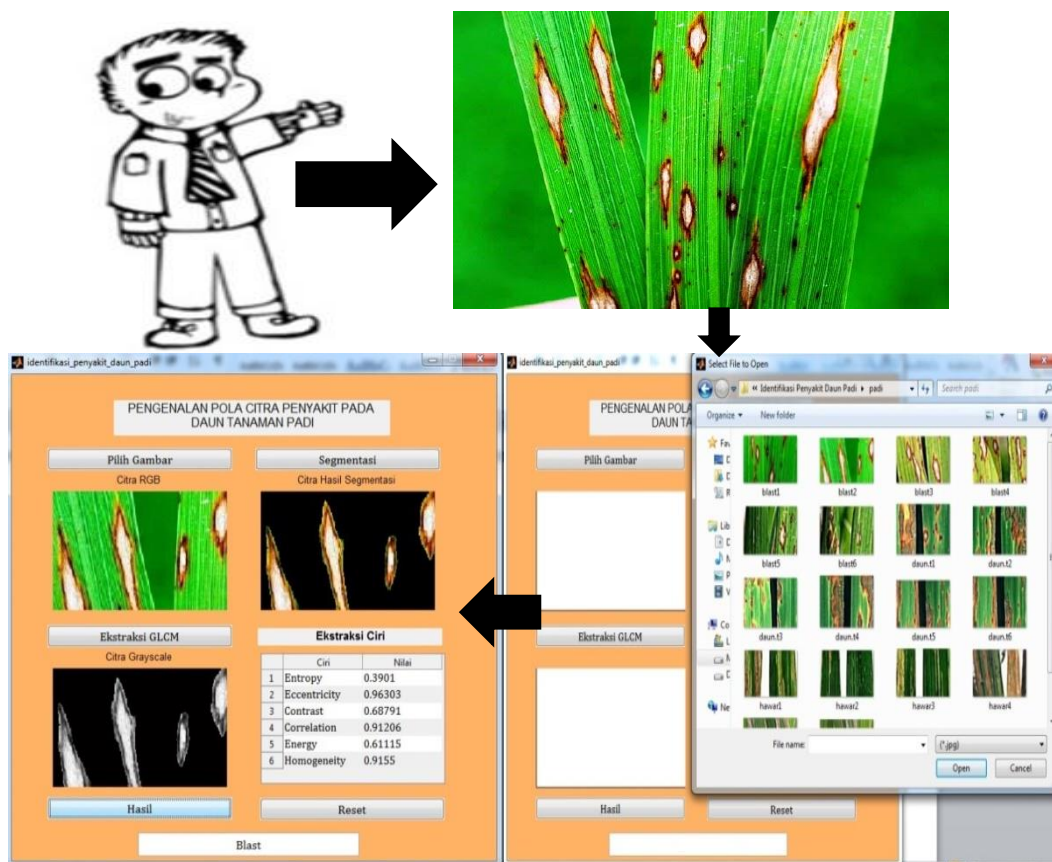

Fig. 1. System Design

To configure GLCM, the co-occurrence matrix (Figure 2) for gray levels often calculates the intensity of the pixel (gray level) of in spatial relation to a pixel of m, primarily the spatial relationship defined as a pixel of interest and its adjacent pixel horizontally on the right direction. The element (n, m) produced in the co-occurrence matrix is simply the sum of the number of pixels that have a value of n in the spatial relationship specified for the pixel of m in the input image. Because processing requires a calculation of the co-occurrence matrix for the full variable range in the image, this is not desirable, so using the measures to reduce the number of density values in gray images from 256 to 8, the number of gray levels determines the size of the co-occurrence matrix [17][18]. Texture applications are divided into two categories, first is for segmentation purposes, where texture is used to perform separation between one object with another object, Second is for texture classification, which uses texture features for object classification [19].
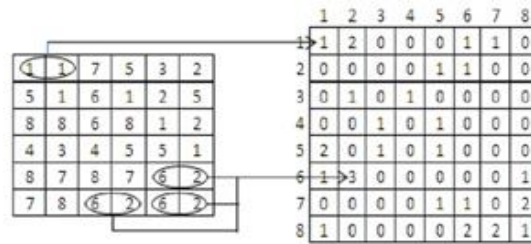
Fig. 2. Co-Occurrence Matrix

It can be seen that the number of values from column 1 and column 2 and so on are inserted into the coo-currency matrix according to rows and columns. Some GLCM features are described as follows. Energy, used to measure texture uniformity, energy will be of high value when the value of a pixel is similar to each other, Otherwise, a small value signifies the value of GLCM normalization is heterogeneous. The maximum energy value is 1 means that the pixel distribution is in a constant condition or the periodic (not random) shape. The equation of energy is

$$E = \sum_{i=1} \sum_{j=1} \{p(i,j)\}^2 \tag{1}$$

Where, i is row, j is column, and $p(i,j)$ is declare the value that belongs to line I and column J on the Co-Occurrence Matrix.

Entropy shows the amount of information of the image that is needed for image compression. Entropy measures the loss of information or message in a transmitted signal and also measures the image information. The entropy can be seen as

$$C_e = \sum_{i=1} \sum_{j=1} p(i,j) \, log_2 p(i,j) \tag{2}$$

Contrast, special frequencies of the imagery and differences of the resulting GLCM moment. The difference is the difference between the high and the low pixels. Contrast is 0 if the pixel's thickness value is equal. The contrast is

$$C = \sum_{i=1} \sum_{j=1} (i - j)^2 \, p(i,j) \tag{3}$$

Homogeneity, also known as Inverse Difference Moment. Homogeneity is used to measure the level of image homogenization. This value is used because it is very sensitive to the value generated by the same pixel or uniform it will be of high value. The opposite of contrast, a big value if the value of energy is still the same pixel value.

$$C_h = \sum_{i=1} \sum_{j=1} \frac{p(i,j)}{1 + |i - j|} \tag{4}$$

Correlation measures the linear dependency of grey levels of neighboring pixels. Digital Image Correlation is an optical method that employs tracking & image registration techniques for accurate 2D and 3D measurements of changes in images. This is often used to measure deformation, displacement, strain, and optical flow, but it is widely applied in many areas of science and engineering. One very common application is for measuring the motion of an optical mouse [20][15].

$$C_{cor} = \sum_{i=1} \sum_{j=1} \frac{(i - m_r)(j - m_c)p(i,j)}{\sigma_r \sigma_c} \tag{5}$$

The formulation and extraction of the Enamt image feature rendered extracted using Matlab to calculate GLCM as images cannot be directly provided as inputs to be implemented using FPGAs. The extraction method of the image feature used in this paper is given in Figure 3 [19].
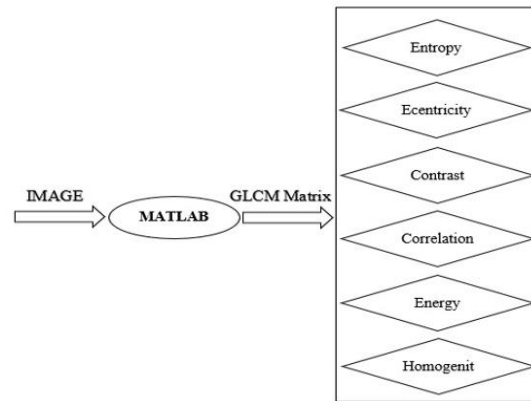
Fig. 3. Extraction of Image Features

2.2   Training Process

Digital Image Acquisition, digital image acquisition aims to determine the data needed and choose digital image recording methods. In this stage, the researchers used a digital image recording method by searching for herbal plant image data through the Google search engine.

Preprocessing, the preprocessing phase aims to simplify the process of image identification. This stage consists of changing the pixel size of the original image to 688x800 pixels, changing the background color of the image in the segmentation process and changing the color of the RGB image to grayscale, LAB, and binary image to get the extraction of the shape value from the image[12][8].

Labeling, at this stage labeling of each image in the training, data is done. Labeling aims to separate data based on labels that will be used in segmentation and classification.

GLCM, at this stage, texture analysis is performed using the GLCM feature. This process is related to the quantization of image characteristics into a group of corresponding characteristic values. Texture analysis is generally used as an intermediary process for image classification and interpretation. The extracted features are Entropy, Energy, Eccentricity, contrast, correlation, and homogeneity [15].

Training, at this stage, the training process is carried out using a set of training data that contains parameter features or features that are used to differentiate between one object and another object. The characteristics used are texture analysis with GLCM and leaf shape recognition. The training process maps training data towards the training target through an algorithm formulation used. Image identification using GLCM features and image classification using k-means clustering segmentation.

The formula used for testing the results of this system is demonstrated by

$$Accuracy = \frac{TP + TN}{TOTAL\ DATA} \tag{6}$$

2.3   Disease in Rice Plants

Rice leaves infected with the disease have different texture patterns, color, and shape of leaves between one disease to another, the types of disease discussed in this study are:

a.   Blast (Figure 4), on the leaves, brown spots with a gray-white center, over time the panicle's neck began to rot or break so that the process of filling the panicle was interrupted and many were hollow in rice.


Fig. 4 Blast Disease

b.   Leaf blight, attacked leaves will be green-gray folding and rolling, in a severe state capable of causing leaves to curl, wither and die. it can be shown in Figure 5.

Fig. 5. Leaf Blight Disease

c.   Burning Leaves, symptoms usually happen on the tips of old leaves, round spots, infected leaves usually dry out and turn brown, the edges are fading. it can be shown in Figure 6.
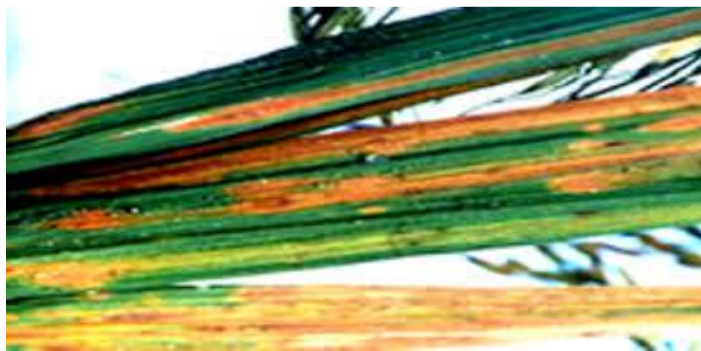
Fig. 6. Burning Leaves Disease

d.   Tungro, symptoms can be seen with changes in small plants tillers are reduced, the leaves turn yellow from the shoot to rice straw. it can be shown in Figure 7.

Fig. 7. Tungro Disease

3.   RESULT AND DISCUSSION

Feature Extraction Flowcharts can be shown in Figure 8. Feature extraction is the process of obtaining the main characteristics contained in the imagery, which has been transformed into the form of contrast, eccentricity, energy, homogeneity, entropy, and correlation with angles of 0 °, 45 °, 90 °, 135 °. Once all these values are obtained it will be averaged. Here's the extraction Flow feature: Grayscale will produce a grayscale matrix that has been quantized, that matrix that is used at this stage, this stage will calculate 5 statistical values of the coo-currency namely.

Feature extraction values to be searched are the value of Contrast, Eccentricity, Homogeneity, Energy, Entropy, Correlation. With a 3x3 pixel size that has 16 degrees of grey with a range of 0-15. The Training process by using the features of the GLCM method to calculate each extraction feature of each disease, thereby generating the extraction value of the features of Entropy, Contrast, Eccentricity, Correlation, Energy, homogeneity. To detect 4 types of diseases using each of 20 training data for each type of disease, by producing the following training data as shown in Table 1.
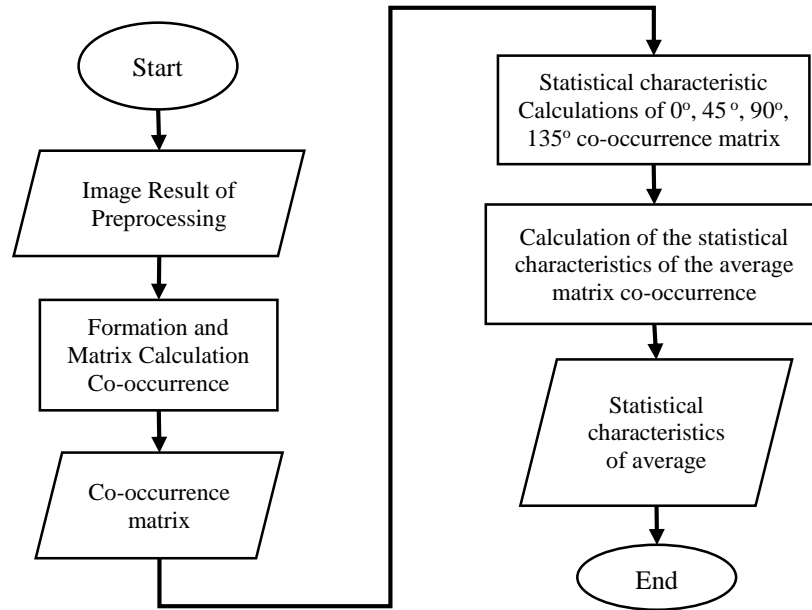
Fig. 8. Feature Extraction Flowcharts

Table 1 Extraction Features

Nama File	Entropy	Eccentricity	Contrast	Correlation	Energy	Homogeneity	Information
Blast1	0.3125	0.9863	0.2612	0.8568	0.7907	0.9550	Correct
Blast2	0.5731	0.9146	0.2851	0.8586	0.7320	0.9447	Correct
Blast3	0.1802	0.9835	1.4112	0.6935	0.5299	0.8675	Correct
Blast4	0.3214	0.9589	0.8607	0.7939	0.4247	0.8739	Correct
Blast5	0.1663	0.9989	0.8453	0.9289	0.5701	0.9278	Correct
Lb1	0.3091	0.9667	0.2930	0.8487	0.8061	0.9543	Correct
Lb2	0.3901	0.9630	0.6879	0.9121	0.6111	0.9155	Correct
Lb3	0.4605	0.7213	0.3266	0.8595	0.7929	0.9537	Correct
Lb4	0.3504	0.9792	1.2013	0.8564	0.5746	0.8957	Correct
Lb5	0.5845	0.9163	1.1547	0.8658	0.4694	0.8776	incorrect
Bl1	0.1127	0.9990	0.3674	0.7987	0.7854	0.9527	Correct
Bl2	0.4545	0.6951	1.2355	0.8606	0.3477	0.8462	Correct
B13	0.4226	0.9153	0.4788	0.7606	0.8098	0.9557	Correct
B14	0.0601	0.9997	0.3893	0.9062	0.6465	0.9415	Correct
B15	0.0944	0.9408	1.2467	0.8041	0.6220	0.9125	Correct
Tungro1	0.8058	0.8576	0.4816	0.9013	0.5915	0.9291	Correct
Tungro2	0.3063	0.8862	0.9925	0.8319	0.2512	0.8457	Correct
Tungro3	0.5719	0.9030	0.7769	0.8955	0.2284	0.8924	Correct
Tungro4	0.1345	0.9966	1.6127	0.7844	0.5996	0.8970	Correct
Tungro5	0.5127	0.9513	0.7572	0.8886	0.3062	0.8983	Correct

Here are the stages of training GLCM implementation. This Matlab Program has 3 Button processes after Matlab Run which is where there is a Button select image, Process segmentation, and feature extraction. The first process is "select image" the leaf that will be extracted, the image is taken from the existing folder in the computer, for a faster process in searching parameters for input data image, the initial input image is resizing form the original image size 32x32 pixels in the RGB color method. The application MATLAB used to detect diseases of the rice leaves. The tools used are Select Image (*Pilih Gambar*), GLCM Extraction (*Ekstraksi GLCM*),

Segmentation (*Segmentasi*), Feature Extraction (*Ekstraksi Fitur*), result (*Hasil*), and Reset, it can be shown in Figure 9.

Button "Segmentation" is used to see the image segmentation results from the original image. Of this stage, the conversion to greyscale imagery is carried out. Image segmentation is a binary image in which the desired object is white (1), while the background that you want to eliminate is black (0). The first step of the segmentation process is to take the original image, then do the image filtering with λ = 4 and θ = 45, After that perform thresholding operations against the magnitude image with a threshold value of 1000. Once a segmented binary image is obtained, then visualizing the segmentation results against the original image. In Matlab using the command dst=rgb2gray(img), "dst" is output variables from image conversion to gray, and "src" is input image variable RGB, it can be shown in Figure 10.

After processing the image input from the folder in the computer, and through the segmentation process and resizing stage, the next step is to change the RGB image to greyscale. This process aims to simplify the pixel value in an image that initially on each pixel has three values, namely RGB to one gray value, the equation used for the greyscale process is:

$$X = 0.12 * R + 0.72 * G + 0.07 * B \tag{7}$$

Next is pressing the Button "extraction GLCM", To display the extraction result of the features of each feature, it can be shown in Figure 11.
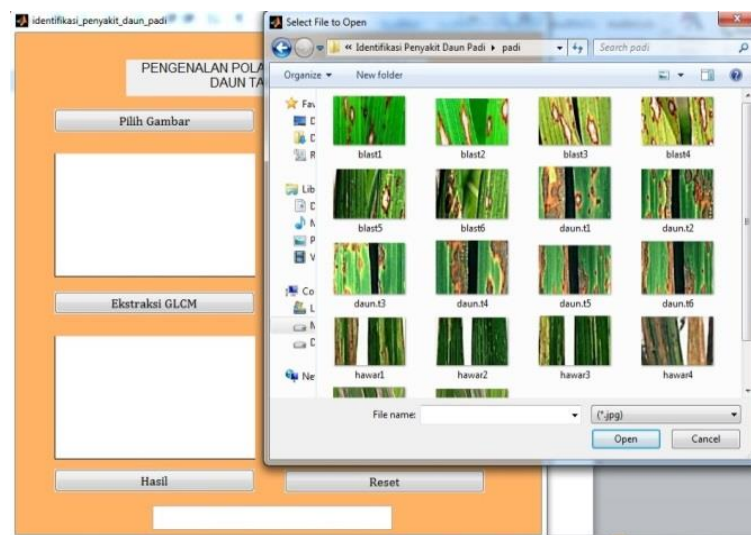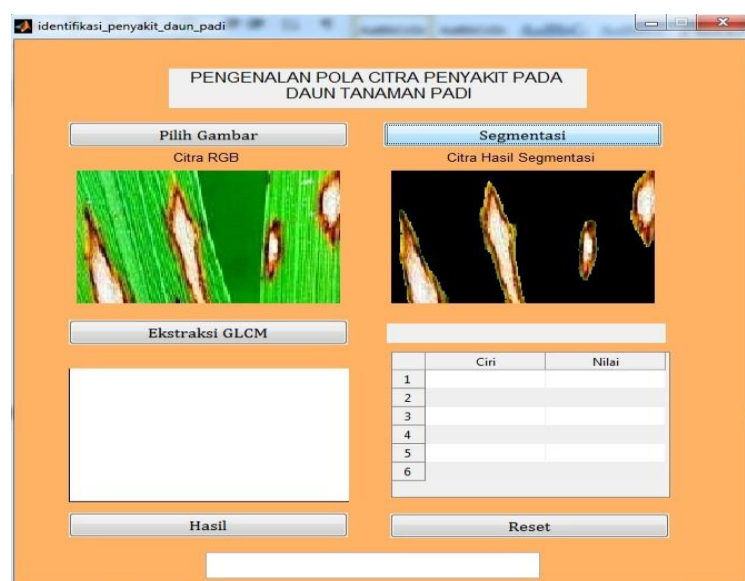

Fig. 9. Select Original image
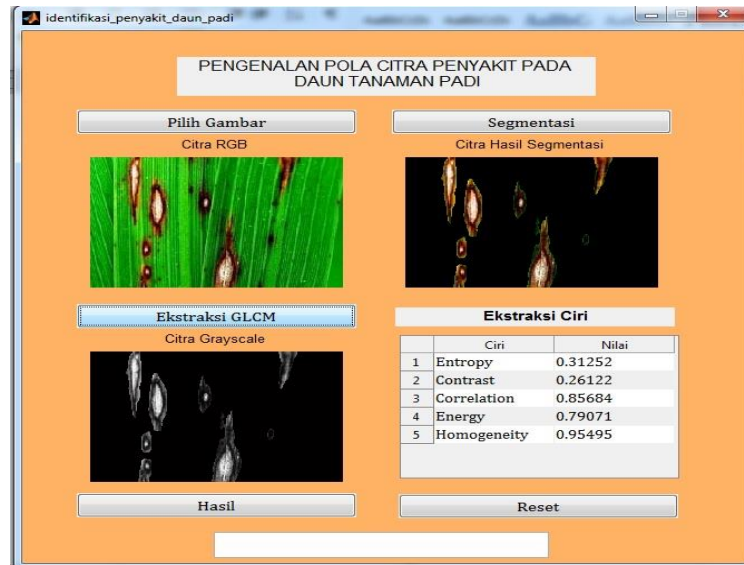

Fig. 10. Segmentation process

Fig. 11. Feature extraction Display

Once you find the extraction result of the sixth feature, then press the "Result" button to find out what kind of leaf disease has been extracted. With the value of Entropy = 0.3901, Eccentricity = 0.96303, Contrast = 0.68791, Correlation = 0.91206, Energy = 0.61115, Homogeneity = 0.9155, thereby detecting the results of Blast disease. it can be shown in Figure 12.
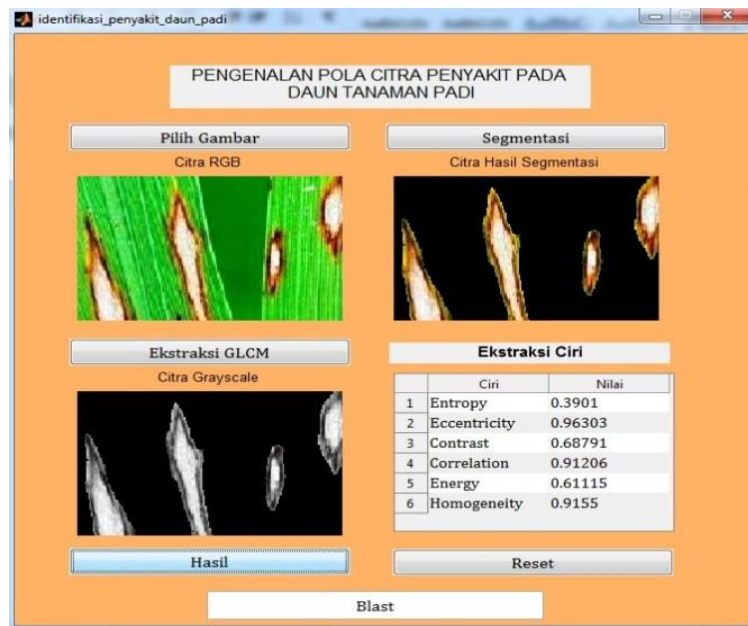

Fig. 12. Display results

4.   CONCLUSION

Based on the results of research and system testing conducted from the feature extraction process using GLCM method, Then the author concludes that is the result of disease detection on the leaves of paddy plants with a total sample data of 40 image data which is where the training data as much as 40 image and test data as much as 20 images based on some features GLCM used, indicating an accuracy level of 90%. The result is not enough 100% due to several factors that affect the extraction result of which is the error in the detection process caused by a poor image color detection factor. The advantage of this research with previous research is using the GLCM method, which is one of the accurate methods to perform the feature extraction process in a texture. The suggestion that the authors convey from the results of this research is necessary to do further research using other methods to perform the extraction of the characteristics on the leaves of the rice crop.

REFERENCES

[1] Y. Sari, A. R. Baskara, and F. Arya, "Klasifikasi grade daun padi sebagai penentu pemupukan urea dengan metode ekstraksi fitur," Prosiding Seminar Nasional Lingkungan Lahan Basah, vol. 4, no. 1, pp. 148–151, April 2019. Google Scholar

[2] A. Candra, "Prototype Sistem Kontrol Air Sawah Otomatis Berdasarkan Level Air Berbasis Mikrokontroler Atmega8535 Pada Desa Bontoraja Kabupaten Bulukumba," *JEECOM*, vol. 2, no. 1, pp. 22–33, 2020. Google Scholar

[3] S. Sudewi, A. Ala, Baharuddin, and M. Farid BDR., "Keragaman Organisme Pengganggu Tanaman (OPT) pada Tanaman Padi Varietas Unggul Baru (VUB) dan Varietas Lokal pada Percobaan Semi Lapangan," *J. Agrik.*, vol. 31, no. 1, pp. 15–24, 2020. DOI: 10.24198/agrikultura.v31i1.25046

[4] N. Trisna, Y. Elva, and A. I. Jmhur, "Implementasi Sistem Pakar Diagnosa Penyakit Tanaman Padi Dengan Menggunakan Metode Forward Chaining," *Jurnal Informasi Komputer Logika*, vol. 1, no. 3, pp. 1–10, 2019. Google Scholar

[5] J. Kusanti and N. A. Haris, "Klasifikasi Penyakit Daun Padi Berdasarkan Hasil Ekstraksi Fitur GLCM Interval 4 Sudut," *J. Pengemb. IT*, vol. 03, no. 01, pp. 1–6, 2018. DOI: 10.30591/jpit.v3i1.669

[6] D. Rosadi and A. Hamid, "Sistem Pakar Diagnosa Penyakit Tanaman Padi Menggunakan Metode Forward Chaining," *J. Comput. Bisnis*, vol. 8, no. 1, pp. 43–48, 2014. Google Scholar

[7] S. Wulandari, M. F. Noor, A. K. Wardhana, and Kusrini, "Sistem Pakar Diagnosa Hama Dan Penyakit Tanaman Padi Dengan Metode Bayes," *J. Inf.*, vol. 5, no. 2, pp. 59–64, 2019. DOI: 10.46808/informa.v5i2.83

[8] A. Priatmoko and E. Harahap, "Implementasi Algoritma DES Menggunakan MATLAB," *J. Mat.*, vol. 16, no. 1, pp. 11–19, 2017. DOI: 10.29313/jmtm.v16i1.3360

[9] I. D. Kurniawati and I. A. Kusumawardhani, "Implementasi Algoritma Canny dalam Pengenalan Wajah menggunakan Antarmuka GUI Matlab," *Institution of Engineering and Technology*, 2017. Google Scholar

[10] A. P. Manullang, A. Prahutana, and R. Santoso, "Penerapan Metode Simple Additive Weighting (Saw) Dan Weighted Product (Wp) Dalam Sistem Penunjang Pemilihan Laptop Terfavorit Menggunakan Gui Matlab," *J. Gaussian*, vol. 7, no. 2018, pp. 11–22, 2018. Google Scholar

[11] E. Nurraharjo, "Implementasi Pemrograman Interfacing MATLAB-Arduino," *J. Teknol. Inf. Din.*, vol. 20, no. 2, pp. 100–105, 2015. Google Scholar

[12] E. Maria, Yulianto, Y. P. Arinda, Jumiaty, P. Nobel, "Segmentasi Citra Digital Bentuk Daun Pada Tanaman Di Politani Samarinda Menggunakan Metode Thresholding," *Jurnal Rekayasa Teknologi Informasi (JURTI)*, vol. 2, no. 1, 2018. DOI: 10.30872/jurti.v2i1.1377

[13] I. Amalia, Indrawati, and Y. M. Amin, "Ekstraksi Fitur Citra Songket Berdasarkan Tekstur Menggunakan Metode Gray Level Co-occurrence Matrix (GLCM)," *J. Infomedia*, vol. 3, no. 2, pp. 64–68, 2018. DOI: 10.30811/jim.v3i2.715

[14] M. Latief and R. Yusuf, "Gorontalo Medicinal Plants Image Identification System Using Artificial Neural Network with Back Propagation," *IJITEE (International Journal of Information Technology and Electrical Engineering)*, vo. 2, no. 2, 2018. DOI: 10.22146/ijitee.42154

[15] Priyanka and D. Kumar, "Feature Extraction Using and Selection Kidney Ultrasound Images Using GLCM and PCA," *Procedia Comput. Sci.*, vol. 167, no. 2019, pp. 1722–1731, 2020, DOI: 10.1016/j.procs.2020.03.382

[16] C. Liu and X. Zheng, "Comparative Investigation on Objective Evaluation Methods for Fabric Smoothness," *FIBRES Text. East. Eur.*, vol. 2, no. 140, pp. 43–49, 2020. Google Scholar DOI: 10.5604/01.3001.0013.7313

[17] K. H. Thanoon, "Proposed Algorithm for Using GLCM Properties to Distinguishing Geometric Shapes," *Raf. J. Comp. Math's*, vol. 13, no. 1, pp. 32–47, 2019. DOI: 10.33899/csmj.2020.163501

[18] R. Andrian, D. Maharani, M. Ardhi, and A. Junaidi, "Butterfly identification using gray level co-occurrence matrix (glcm) extraction feature and k-nearest neighbor (knn) classification," *Sci. J. Inf. Syst. Technol. Available*, vol. 6, no. 1, pp. 11–21, 2020, DOI: 10.26594/register.v6i1.1602

[19] P. Mohanaiah, P. Sathyanarayana, and L. Gurukumar, "Image Texture Feature Extraction Using GLCM Approach," *Int. J. Sci. Res. Publ.*, vol. 3, no. 5, pp. 1–5, 2013. Google Scholar

[20] H. A. Al-beiruti and H. A. Jeiad, "OPG Images Automatic Segmentation and Feature Extraction for Dental Lesion Diagnosis Purposes," *Int. J. Sci. Res.*, vol. 7, no. 11, pp. 717–724, 2018, DOI: 10.21275/ART20192396. Semantic Scholar

BIOGRAPHY OF AUTHORS

Lilis Indrayani, have completed Bachelor Computer-Information System in 2016, then Master Computer-Computer system in 2018. Currently, she is working as Lecturer and Chairman of information System Study Program in STMIK Kreatindo Manokwari, Papua Barat, Indonesia. Email : lilisindrayani8@gmail.com

Raden Wirawan, has finished Computer Master-Computer System in 2015. Currently, she is working as a lecturer of the Computer Systems Study Program in STMIK Bina Adinata Bulukumba, Sulawesi Selatan, Indonesia. Email : liliraden12790@gmail.com