ISSN:0975-9646

Gasim et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 4 (1) , 2013, 188 - 193

Merging Feature Method on RGB Image and Edge Detection Image for Wood Identification

Gasim[#], Agus Harjoko[*], Kudang Boro Seminar[**], Sri Hartati[*]

[#] *Faculty of Information Technology,*
Multi Data Palembang Bachelor Program, Palembang, Indonesia

[*] *Dept. of Computer Science and Electronics,*
Gadjah Mada University, Yogyakarta, Indonesia

[**] *Dept. of Mechanical and Biosystem Engineering,*
Bogor Agricultural University (IPB), Bogor, Indonesia

***Abstract*— The identification of timber types has become very important especially for identifying the name because this name of a timber, determining the quality and price of timber, determining the tax for particular timbers, determining the price of a timber based on the quality. Until now, these activities are still done manually in two ways: (1) using a magnifying glass to identify the cross-section of the timber, and (2) using a microscope, which is done in the laboratory basis. Some of the shortcomings of these manual methods include slowness, subjectivity, less accuracy and strongly dependent upon the human expert availability. This research attempts to resolve these problems by proposing a intelligent computational method with ANN. The proposed method utilizes merging feature on RGB image and edge detection image. Based on the experimental results, the proposed method gives the results of 100% and 95% accuracy of the training data and testing of data respectively. This research using 20 names of wood, who all come from Indonesia. And then the photo shoot of cross section is used for training and testing. For training using 2000 image, and use 500 images for testing. It is important to notice that proposed method only requires seven features of RGB images, and six features of edge detection image.**

***Keywords*— wood identification, ANN, RGB, edge detection, image processing, pattern recognition.**

I. Introduction

Knowing the type of wood (timber name) is very important when it is related to illegal logging, prices, taxes, travel documents, compliance the timber with product, and others. By knowing the type of wood through the identification, then it can be solved quickly.

Wood of a particular species can be identified by its unique features. These features include strength, density, hardness, odor, texture and color. Reliable wood identification usually requires the ability to recognize basic differences in cellular structure and wood anatomy. Each species has unique cellular structure that creates differences in wood properties and ultimately determines the suitability for a particular use. Cellular characteristics provide a blueprint for accurate wood identification [1].

Wood is composed of many small cells and its structure is determined by the type, size, shape and arrangement of these cells. The structure and characteristics of wood can vary between species and within the same species. With practice, a small hand lens (10x) can be used to distinguish the different cell types and their arrangements [1].

The diversity of types of wood to make wood identification activities can only be performed by experienced experts, even a highly experienced can identify the type of wood only with a magnifying glass (loop) that has a magnification of at least 10 times, because that analysis are the cross-sectional of the timber which form pore texture and other elements around the pore.

Besides identification using a loop, to improve the accuracy of the identification, analysis is also usually done under a microscope that has been in the form of slice. There are 163 [10] variable to be filled from these observations.

How to identify the type of wood that has been described above, showing the complexity of the activity. Besides it takes a long time, due to several factors such as the location of wood and experts are far apart, very few number of experts, etc.. However, with the development of computer science, specifically image processing and pattern recognition. Author believes it can solve this problem, although it is still in the research stage.

In this paper, the author developed a method of feature extraction from RGB image and edge detection image, with a magnification 60 times, and used neural networks to recognition. This research used 20 types of wood.

This paper is divided into six sections. Section 2 describes the theory relevant with this research, and also presented some research beforehand. Section 3 describes the research method. Section 4 presented the results of research. Section 5 presented some conclusions and suggestions for future research.

II. Teori

A. Previous Research

There have been some research about this object, the method, the number types of wood, and mixed results. Author also has produced some research results of this object, among other things: 1) Using texture-analyst features and RGB component with magnification 24 times, using five different types of wood [3], 2) Using the texture-analysts features, the method used is ANFIS, and uses five types of wood [4], 3) Then the comparison of recognition rate based on features the input with enlarged 24 times, using five different types of wood [7], 4) The next research using 15 types of wood, the texture-analyst and RGB, using ANNBP, and gives the recognition rate 95% [8], this value is quite high, because the number types that used only 15 types, and the data test that are used most of the images are sourced from the same sample with the training image.

And in this paper, the author uses 20 types of wood, using RGB images (converted to gray scale) and edge detection. With hope this method can reduce the number of features that are used, in this case a features duplication still counted one feature. That is expected to increase the number of types of wood that can be identified to be even more.

B. *Wood Anatomy*

Wood of a particular species can be identified by its unique features. These features include strength, density, hardness, odor, texture and color. Reliable wood identification usually requires the ability to recognize basic differences in cellular structure and wood anatomy. Each species has unique cellular structure that creates differences in wood properties and ultimately determines the suitability for a particular use. Cellular characteristics provide a blueprint for accurate wood identification [1].

Wood is composed of many small cells and its structure is determined by the type, size, shape and arrangement of these cells. The structure and characteristics of wood can vary between species and within the same species. With practice, a small hand lens (10x) can be used to distinguish the different cell types and their arrangements. By using the proper techniques, you can become efficient and accurate at wood identification [1].

Classifying wood surfaces into three categories of geometric field of reference, namely cross section, radial section, and tangential section as shown in Fig. 1[1].

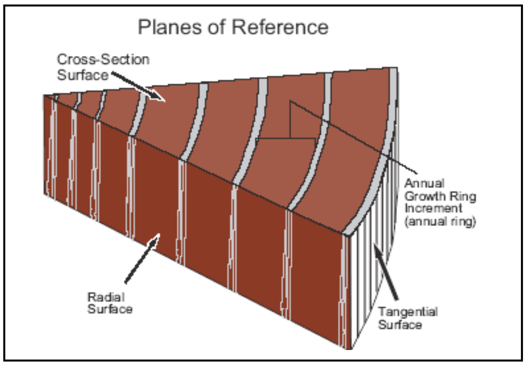

Fig. 1 Reference geometric surface area of wood[1]

C. *Artificial Neural Network*

An *artificial neural network* is an information-processing system that has certain performance characteristics in common with biological neural networks. Artificial neural networks have been developed as generalizations of mathematical models of human cognition or neural biology, based on the assumptions that [2] :

- Information processing occurs at many simple elements called neurons,
- Signals are passed between neurons over connection links,
- Each connection link has an associated weight, which, in a typical neural net, multiplies the signal transmitted,
- Each neuron applies an activation function (usually nonlinear) to its net input (sum of weighted input signals) to determine its output signal.

A neural network is characterized by (1) its pattern of connections between the neurons (called its *architecture*), (2) its method of determining the weights on the connections (called its *training*, or *learning, algorithm*), and (3) its *activation function*[2].

D. *Representing Digital Images*

An image may be defined as a two-dimensional function, *f*(*x*, *y*), where x and y are *spatial* (palne) coordinates, and the amplitude of *f* at any pair of coordinates (x, *y*) is called the *intensity* or *gray level* of the image at that point. When x, y, and the intensity value of f are all finite, discrete quantities, we call the image a *digital image*. The field of *digital image processing* refers to processing digital images by means of a digital computer. Note that a digital image is composed of a finite number of elements, each of which has a particular location and value. These elements are called *picture elements, image elements, pels,* and *pixels*. Pixel is the term used most widely to denote the elements of a digital image[5].

Numerical arrays are used for processing and algorithm development. In equation form, we write the representation of an M x N numerical array as [5],

$$f(x,y)=\begin{bmatrix} f(0,0) & f(0,1) & \cdots & f(0,N-1) \\ f(1,0) & f(1,1) & \cdots & f(1,N-1) \\ \vdots & \vdots & & \vdots \\ f(M-1,0) & f(M-1,1) & \cdots & f(M-1,N-1) \end{bmatrix} \quad (1)$$

Both sides of this equation are equivalent ways of expressing a digital image quantitatively. The right side is a matrix of real numbers. Each element of this matrix is called an *image element, picture element, pixel*, or *pel*. The terms *image* and *pixel* are used throughout the book to denote a digital image and its elements[5].

E. *Canny Edge Detection*

Edge detection is the approach used most frequently for segmenting images based on abrupt (local) changes in intensity. There are there fundamental steps performed in edge detection[5] :

- Image smoothing for noise reduction,
- Detection of edge points,
- Edge localization.

The remainder of this section deals with techniques for achieving these objectives.

Canny's approach is based on three basic objectives [5]:

1) *Low error rate* : All edges should be found, and there should be no spurious responses. That is, the edges detected must be as close as possible to the true edges.

2) *Edge points should be well localized* : The edges located must be as close as possible to the true edges. That is, the distance between a point marked as an edge by the detector and the center of the true edge should be minimum.

3) *Single edge point response* : The detector should return only one point for each true edge point. That is, the number of local maxima around the true edge should be minimum. This means that the detector should not identify multiple edge pixels where only a single edge point exists.

F. Texture Analysis

Texture analysis [6] refers to the characterization of regions in an image by their texture content. Texture analysis attempts to quantify intuitive qualities described by terms such as rough, smooth, silky, or bumpy as a function of the spatial variation in pixel intensities. In this sense, the roughness or bumpiness refers to variations in the intensity values, or gray levels [9].

Some of the most commonly used texture measures are derived from the Grey Level Co-occurrence Matrix (GLCM). The GLCM is a tabulation of how often different combinations of pixel brightness values (gray levels) occur in a pixel pair in an image [9].

G. Gray-Level Co-Occurrence Matrix (GLCM) [6]

A statistical method of examining texture that considers the spatial relationship of pixels is the gray-level co-occurrence matrix (GLCM), also known as the gray-level spatial dependence matrix. The GLCM functions characterize the texture of an image by calculating how often pairs of pixel with specific values and in a specified spatial relationship occur in an image, creating a GLCM, and then extracting statistical measures from this matrix [9].

To illustrate, the Fig. 2 shows how *graycomatrix* calculates the first three values in a GLCM. In the output GLCM, element (1, 1) contains the value *1* because there is only one instance in the input image where two horizontally adjacent pixels have the values *1* and *1*, respectively. *glcm(1, 2)* contains the value *2* because there are two instances where two horizontally adjacent pixels have the values *1* and *2*. Element (1, 3) in the GLCM has the value 0 because there are no instances of two horizontally adjacent pixels with the values 1 and 3. *graycomatrix* continues processing the input image, scanning the image for other pixel pairs (i, j) and recording the sums in the corresponding elements of the GLCM [9].

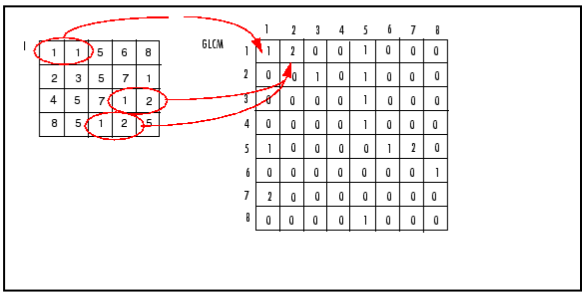

Fig. 2 Process used to creat the GLCM

From this matrix is used to calculate some statistical variables. These statistics provide information about the texture of an image. Lists the statistics are [9] :

- Contrast : Returns a measure of the intensity contrast between a pixel and its neighbor over the whole image, the formula for contrast is in (2).

  Range = [0 (size(GLCM,1)-1)^2]

  Contrast is 0 for a constant image

$$\sum_{i,j}|i-j|^2 p(i,j) \qquad (2)$$

- 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 negatively correlated image. Correlation is NaN for a constant image. The formula for contrast is in (3).

$$\sum_{i,j}\frac{(i-\mu i)(j-\mu j)p(i,j)}{\sigma_i\sigma_j} \qquad (3)$$

- Energy : Returns the sum of squared elements in the GLCM. The formula for contrast is in (4).

  Range = [0 1]

  Energy is 1 for a constant image

$$\sum_{i,j} p(i,j)^2 \qquad (4)$$

- Homogeneity : Returns a value that measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal. The formula for contrast is in (5).

  Range = [0 1]

  Homogeneity is 1 for a diagonal GLCM

$$\sum_{i,j}\frac{p(i,j)}{1+|i-j|} \qquad (5)$$

H. Entropy

Entropy is a statistical measure of randomness that can be used to charaterize the texture of the input image. Entropy is defined as [9] :

*-sum(p.*log2(p))* *(6)*

Where p contains the histogram counts returned from `imhist` [9].

III. Research Method

A complete research activities is shown in Fig. 3,

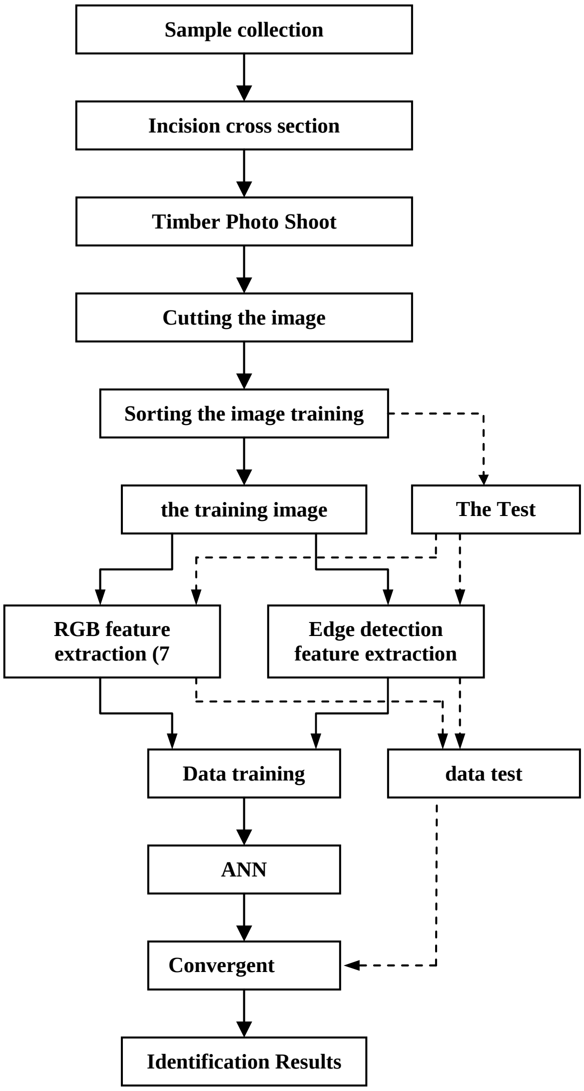

Fig. 3 Research Method

Explanations of each step in Fig. 3 is:

1) *Sample collection*: The collection of wood samples consisting of 20 types of wood (Table 1). Each type of wood taken from several regions in Indonesia.

TABLE I
TYPE OF WOOD USED

No.	*Trade Name (Scientific Name)*
1.	Bakau (Rhizophora apiculata Bl.)
2.	Cenge (Mastixia trichotoma Bl.)
3.	Jabon (anthocepalus cadamba)
4.	Jabon merah (Anthocephalus macrophyllus)
5.	Kembang semangkok (Scaphium macropodum J.B.)
6.	Kruing (Dipterocarpus gracilis Bl.)
7.	Kruing (Dipterocarpus kunstleri King)
8.	Kulim (Scorodocarpus borneensis Becc.)
9.	Mempisang (Mezzetia parviflora)
10.	Meranti Kuning (Shorea acuminatissima sym)
11.	Meranti Merah (Shorea acuminata)
12.	Meranti Merah (Shorea ovalis Bl.)
13.	Meranti Putih (Shorea Javanica k.ot. val)
14.	Merawan (Hopea spp.)
15.	Merbau (Intsia bijuga O.K.)
16.	Merbau (Intsia palembanica)
17.	Mersawa (Anisoptera)
18.	Penjalin (Celtis Philippinensis)
19.	Perupuk (Lophopetalum javanicum)
20.	Rasamala (Hamamelidaceae)

2) *Incision cross section*: Section to be photographed is a cross-sectional of the timber. This incision using a cutter knife, and carried on some parts of a cross section.

3) *Timber Photo Shoot*: Photo Shoot done in cross section has been sliced. This shoot using the camera microscope, with 60 times optical zoom. The results of this shoot, is the RGB image with size 1280x1024, Fig.4.

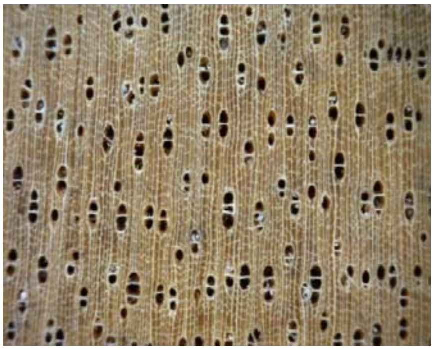

Fig. 4 Example the image captured (minimized)

4) *Cutting the image*: At this stage the image is cut to the size of 500 x 500 pixels. Performed on all the images that have been photographed Fig. 5.

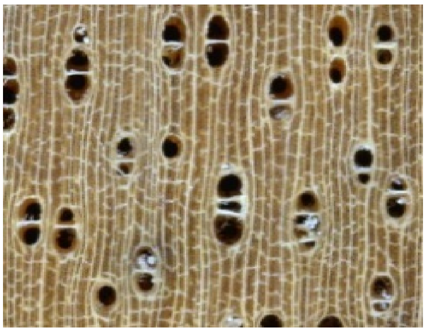

Fig. 5 Example image of the cutting (minimized)

5) *Sorting the image training*: This stage is done because not all of the image of the cuts is a good image. This activity will produce a collection of the image training and a collection of test images.

6) *The training image*: A collection of training the image result from training the image sorting process.

7) *The test images*: A collection of test images that result from the process of training image sorting.

8) *RGB feature extraction*: this extraction is done to the training image and test images. In this process the image in RGB then converted to gray scale and then extracted. Features are taken Contrast, Energy, Homogeneity, Correlation, Entropy, Standard Deviation, Gray Level.

9) *Edge detection feature extraction*: this extraction is done to the training images and test images. In this process, the edge detection use canny algorithm (Fig.6). Feature taken are Contrast, Energy, Homogeneity, Correlation, Entropy, Standard Deviation.

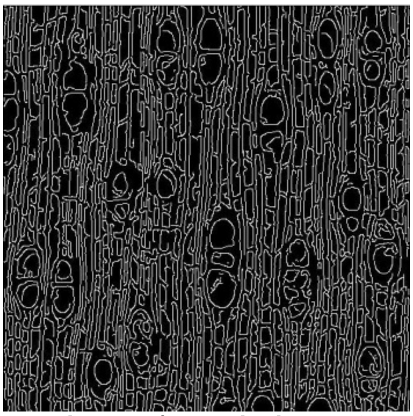

Fig. 6 Example image of canny edge detection (minimized)

10) *The training data* : results of two previous activities, formed a table. The first seven rows of data containing the extracted RGB and six next line contains data extracted edge detection. This stage is also performed on the data test.

11) *ANN*: An architecture for the ANN training process.

12) *Convergent ANN*: An ANN who already have a balance on the weights that are changing during the training to recognize 20 types of wood. Convergent ANN is used for testing. ANN technical data can be seen in Table 2.

13) *Identification Results*: This is the output value of the ANN Convergent form values paired with the targets set when training.

TABLE II
SPECIFICATION OF NEURAL NET WORK

Characteristic.	***specification***
Architecture, algorithm	3 hidden layers, back propagation
Neuron input	Extraction from Gray Level of image 500x500 : Contrast, Energy, Homogeneity, Correlation, Entropy, Standard Deviation, Gray Level Extraction from edge detection of image 500x500 : Contrast, Energy, Homogeneity, Correlation, Entropy, Standard Deviation.
Neuron of hidden layer	73, 73, 73
Neuron of output layer	Number of wood
activation function	Sigmoid binary
MSE	1e-32
learning rate	0,1
epoch	2000
Number of image each wood for data training	100
Number of image each wood for data testing	25

IV. RESULTS

There are three forms of testing in this research, and each gave different results, because different sources of test data. The tests are:

- Tests on the training image, meaning image that are used for testing is the same image used for training. Test results on the image gives the recognition rate of 100%. This means that all image can be recognized accurately. The amount of data is 100 images for each type of wood, so the overall number of images is 2000 image
- test on images of the same sample, meaning that the image is used for testing, came from the same timber samples with the sample for training, but with a different image to the training images. The recognition rate with this image is 90%, depending on the image.
- Tests on the outside image of the sample, meaning that the image used for the test is not derived from the same training samples. Testing of this image gives the recognition rate 85%, of course, depending on the condition of the image.

V. CONCLUSIONS AND SUGGESTIONS

There are several conclusions from this research, that is:

- With feature extraction texture analysts on RGB images were converted to gray scale, and the canny edge detection, this method can provide a fairly good recognition rate.
- Magnification level that is used (60 times) also provides fairly detailed texture of microscopic cross section.
- spacious areas observed to be very influential on the level of recognition. However it is quite difficult to do, because it is associated with an incision, photo shoot, and the level of magnification.

Some suggestions from the author for further research are:

- There needs to be additional features of the method, so that the recognition rate can be even higher.
- It needs a experiment of this method by combining image size. It aims to form a range of features values wider between each type of wood.
- Author hopes this research can be developed further to get a more accurate recognition rate.

ACKNOWLEDGMENT

The authors would like to thank (1) STMIK Multi Data Palembang (www.mdp.ac.id), (2) the Department of Computer Science and Electronics, Gadjah Mada University (http://mkom.ugm.ac.id) Yogyakarta Indonesia, and Dept. of Mechanical and Biosystem Engineering, Bogor Agricultural University (IPB), Bogor, Indonesia (http://ipb.ac.id) that provides technical support for the research, and (3) Department of Forestry Laboratory of Wood Anatomy Bogor Indonesia. With these three elements this research can be run well.

REFERENCES

[1] Bond B. and Hamner P. "*Wood Identification for Hardwood and Soft wood Species Native to Tennese,*" http://www.utextension.utk.edu/, 2002.

[2] Fausett, L., "Fundamentals Of Neural Network Architectures : Algorithm and Applications," Prectice-Hall, Inc., 1994.

[3] Gasim, "*The Design and Implementation of an Image-Based Wood Variety Recognition System Using ANN*," Proceeding The 9th INTERNATIONAL CONFERENCE on QUALITY in RESEARCH

(QiR).Information and Computation Engineering, ISSN : 114-1284, 2006.

[4] Gasim, Hartati, S., "*Arsitektur ANFIS untuk Pengenalan Kayu Berbasis Citra Cross-Section*," The International Conterence on Computer and Mathematical Sciences 2010, 29 June 2010 UiTM and UGM Collaboration, Jogjakarta, 2010.

[5] Gonzales, R. C. & R. E. Woods. "*Digital Image Processing*." Addison Wesley, Massachusetts, 1992.

[6] Haralick, RM., K. Shanmugam and Itshak Dinstein. "*Textural Features For Image Classification*," IEEE Transaction On System, Man and Cybernetics. Vol 3, No. 6. 1973.

[7] Harjoko, A., Gasim, "*Comparison of Some Features Extraction of Wood*," Proceeding The 2nd International Conference on Distributed Frameworks and Applications, Jogjakarta, 2010.

[8] Harjoko, A., Gasim, Rulliaty, S.S., Damayanti, R., "*Identification Method for 15 Names of Commercial Wood With Image of Texture Pore as an Input*," Proceeding International Conference on Informatics for Development, Jogjakarta, 2011.

[9] Mathwork Inc., "*Neural Network Toolbox for Use With Matlab*" The Mathwork Inc. Natick, USA, 2012.

[10] Wheeler, E.A., P. Baas, and P.E. Gasson. 1989. IAWA List Of Microscopic Features For Hardwood Identification, IAWA Bulletin N.S. 10(3); 221-332

AUTHORS PROFILE

Gasim is a lecturer at the Faculty of Information Technology, Multi Data Palembang (STMIK MDP), Bachelor Program, Palembang, Indonesia. He is graduated as Bachelor of Computer in STMIK Bandung, Indonesia. He is received his Master of Computer at Bogor Agricultural University (IPB), Bogor, Indonesia. He is currently taking his Doctoral Program at the Departement of Computer Science and Electronics, Gadjah Mada University in Yogyakarta, Indonesia.

Agus Harjoko is a Associate Professor and Chair of Electronics and Instrumentation Lab. He is graduated from the Electronics and Instrumentation study program, Faculty of Mathematics and Natural Sciences, UGM Yogyakarta, Indonesia. He got M.Sc. and Ph.D in Computer Science from the University of New Brunswick, Canada.

Kudang Boro Seminar is Professor in Computer Technology. He is head of laboratory bioinformatics engineering, director of communications and information systems, IPB, and Honorary Member of AFITA (Asian Federation for Information Technology in Agriculture) .He is graduated from Bogor Agricultural University (IPB) Bogor, Indonesia. He got M.Sc. and Ph.D in Computer Science at University of New Brunswick, Canada.

Sri Hartati is a Associate Professor and Chair of Computer Science Graduate Program, Gadjah Mada University, Yogyakarta, Indonesia. She got M.Sc. and Ph.D in Computer Science Dept, at University of New Brunswick, Canada.