Detection of Heart Abnormalities Based On ECG Signal Characteristics using Multilayer Perceptron with Firefly Algorithm-Simulated Annealing

Sofiah Ishlakhul Abda[1], Auli Damayanti[1,*] & Edi Winarko[1]

[1]Mathematics Department, Faculty of Science and Technology, Universitas Airlangga

[*]Corresponding author: auli-d@fst.unair.ac.id

Abstract. Heart disease is one of the causes of death worldwide. Therefore, detecting heart disease is very important to reduce the increased mortality rate. One of the methods used to detect the abnormalities or disorders of the heart is to use computer assistance to determine the characteristics of an electrocardiogram. Electrocardiogram (ECG) is a test that detects and records the activity of the heart through small metal electrodes attached to the skin of one's chest, arms and legs. This test shows how fast the heart beats and whether the rhythm is stable or not. The purpose of this thesis is to apply a multi-layer perceptron model with firefly algorithm and simulated annealing in detecting cardiac abnormalities based on the ECG signal characteristics. The initial step of this research is image processing. The stages of ECG image processing are grayscale, thresholding, edge detection, segmentation and normalization processes. The results of this image processing are used as input matrices in the perceptron multilayer network training using firefly algorithm and simulated annealing. In the training process, we will get optimal weights and biases for validation tests on test data. The training data in this thesis uses 20 ECG images and in the validation test process uses 10 ECG images. The validation results in the validation test show that the accuracy in detecting heart abnormalities based on the characteristics of ECG signals using multi- layer perceptron with firefly algorithm and simulated annealing is 100%.

Keywords: *electrocardiogram, firefly algorithm, heart disease, multi-layer perceptron, simulated annealing.*

1   Introduction

Heart disease is one of the causes of death worldwide. There are three million causes of death from cardiovascular disease each year at 41 percent of the total. One of the methods used to detect the heart abnormalities or disorders is to use computer assistance to determine the characteristics of an electrocardiogram. Electrocardiogram (ECG) is a test that detects and records the activity of the heart through small metal electrodes attached to the skin of one's chest, arms and legs. This test shows how fast the heart beats and whether the rhythm is stable or not [1]. Although knowing how the ECG works is relatively easy, knowing the information contained in the electrocardiogram (ECG) data is very difficult. Experience and knowledge about heart disease and its symptoms are needed in order to be able to read ECG records. The manual pattern recognition of

information on the ECG signals is very inefficient because of the large amount of data that must be observed [2].

Artificial Neural Network (ANN) is one of the artificial representations of the human brain that always tries to simulate the learning process in the human brain [3]. There are several learning methods of ANN, one of which is the multi-layer perceptron. Multi-layer Perceptron (MLP) is a learning method by updating the weight and comparing it to the target to produce output that is close to the target value [4]. To improve the performance of artificial neural network training, optimization algorithms are needed to determine the optimal weight in the learning process of ANN. Firefly algorithm is a multimodal optimization, inspired by nature based on the flashing behavior of fireflies, which can be effectively used for the extraction of features present in input data [5]. In addition, simulated annealing is also one of the algorithms for optimization. This algorithm can be used to find approaches to the local optimum solution of a problem [6]. The advantage of the simulated annealing algorithm compared to other methods is its ability to avoid optimal local traps [7].

Based on the explanation, this research applies multi-layer perceptron with firefly algorithm and simulated annealing to detect heart abnormalities based on ECG signal characteristics.

2   Image Processing

Image processing is the process or attempts to transform an image / image into another image using certain techniques. The image processing aims to improve the quality of the image, which the image result can provide information about the characteristics of the image that has been in the form of numerical clearly [9]. Some image processing phases that will be carried out are grayscale, thresholding, edge detection and segmentation processes.

a.   Grayscale

Grayscale are pixel colors within the black and white gradation range. Grayscale intensity is usually stored as 8 bit image data or 256 gray color intensities from 0 (black) to 255 (white), while the gray image is in between. Gray color images can be generated from RGB color images by taking the average value of the three main color components, namely red, blue and green, or can be written like equation (1) [8].

$$Gray = \frac{R+G+B}{3} \qquad (1)$$

where
R is the value on red layer
G is the value on green layer

B is the value on blue layer

b. Thresholding

Thresholding is used to adjust the number of gray degrees of the image. By using thresholding, the degree of gray can be changed as desired. This thresholding process is basically the process of changing the quantization of the image, so that the formula can be used to do thresholding with a degree of gray is as follows[9] :

$$x = b.int\left(\frac{w}{b}\right) \tag{2}$$

$$b = int\left(\frac{256}{a}\right) \tag{3}$$

where :

w is the gray degree value before thresholding

x is the value of the degree of gray after thresholding

a is the threshold value used (16)

c. Edge Detection

Edge detection is a process that produces the edges of an image. One of the methods used for edge detection is to use the sobel operator. The advantage of this method is the ability to reduce noise before performing edge detection calculations. The equation used for the sobel edge detection process is as follows [10]:

$$G(x,y) = \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix} * A \tag{4}$$

where A is the value of each pixel of the image.

d. Segmentation

In general, image segmentation can be interpreted as dividing the image into its constituent parts or smaller segments and it is further expected to process data faster. The stages in the segmentation process are each segment of an image looking for the average of segment. To calculate the average of segment, the intensity or grayscale level of all pixels in one segment is added up, then divided by the number of pixels of that segment. So that each segment consists of pixels that have the same grayscale level [9].

3   Multi Layer Perceptron

Multi-layer Perceptron (MLP) is a feedforward Artificial Neural Network (ANN) consisting of a number of neurons connected by connecting weights. The neurons are arranged in layers consisting of one input layer, one or more hidden layers, and one output layer. The input layer receives the signal from the outside, then passes it to the first hidden layer, which will continue until it finally reaches the output layer [11]. Backpropagation is a supervised training algorithm. The training with backpropagation method consists of three steps, namely: feedforward, backpropagation of error, weight renewal and bias [3]. Here are the steps for backpropagation training:

Step 0.   Initialization of weights and biases.
                 Weights and biases can be initialized with random real numbers at intervals (0,1) or (-1,1).
Step 1.   If the stop condition is still not met, run steps 2 through 9. Step 2.
                 For each training data, do steps 3 to 8.

Feedforward
Step 3.   Each input unit $(X_i, i = 1, ... , n)$ receive the x_i input signal and spread the signal across all hidden units.
Step 4.   Each hidden unit $(Z_j, j = 1, ... , p)$ will add up the weighted and biased input signal :

$$z_{in_j} = v_{0j} + \sum_{i=1}^{n} x_i v_{ij} \qquad (5)$$

and use the pretedetermined activation function to calculate the corresponding output signal and hidden unit :
$$z_j = f(z_{in_j}) \qquad (6)$$
Then send the output signal to all units in the output unit.
Step 5.   Each output unit $(Y_k, k = 1, ... , m)$ will add up the weighted and biased input signal :

$$y_{in_k} = w_{0k} + \sum_{j=1}^{p} z_j w_{jk}$$

And use the predetermined activation function to calculate the output signal and output unit concerned :
$$y_k = f(y_{in_k}) \qquad (8)$$

Backpropagation of error
Step 6.   Each output unit $(Y_k, k = 1, ... , m)$ receive a target that matches the training data input to calculate the error between the target and the output produced by the network :

$$\delta_k = (t_k - y_k)f'(y_{in_k}) \tag{9}$$

Like the training data input, the training data output t_k has also been scaled according to the activation function used.

The δ_k factor is used to calculate error correction (Δw_{jk}) which will later be used to update w_{jk}, with :

$$\Delta = \alpha \delta_k z_j \tag{10}$$

In addition, the bias correction Δw_{0k} is also calculated which will be used to update w_{0k}, with :

$$\Delta w_0 = \alpha \delta_k \tag{11}$$

The δ_k factor is then sent to the layer located.

Step 7.　Each hidden unit $(Z_j, j = 1, \dots, p)$ add up the input delta

$$\delta_{in_j} = \sum_{k=1}^{m} \delta_j^k \tag{12}$$

Then multiplied by the derivation of its activation function to calculate the error information :

$$\delta_j = \delta_{in_j} f'(z_{in_j}) \tag{13}$$

Then calculate the weight correction :

$$\Delta v_{ij} = \alpha \delta_j x_i \tag{14}$$

In addition, the bias correction Δv_{0j} is also calculated which will be used to update v_{0j}, with :

$$\Delta v_{0j} = \alpha \delta_j \tag{15}$$

Update of Weight and Bias

Step 8.　Each output unit ($k = 1, \dots, m$) updating weights and biases ($j = 0, \dots, p$):

$$(new) = w_{jk}(old) + \Delta w_{jk} \tag{16}$$

Each hidden unit ($Z_j, j = 1, \dots, p$) updating weights and biases ($i = 0, \dots, n$):

$$(new) = (old) + \Delta v_{ij} \tag{17}$$

Step 9.　Check stop condition. If the stop condition has been met, network training can be stopped.

The notations used in the training algorithm are as follows:

$x = (x_1, \dots, x_i, \dots, x_n)$　　　: training data for input

$t = (t_1, \dots, t_i, \dots, t_n)$　　　: training data for output

α　　: learning rate is a parameter that controls changes in weights during training. If the learning rate is large, the network learns faster, but the results are less accurate. Learning rate is valued at intervals (0, 1).

X_i　: i-th input unit. For input units, the incoming and outgoing signal in a unit is represented by the same variable x_i.

Z_j : j-th hidden unit. The unit signal at Z_j is denoted by . The output (activation) signal for Z_j is denoted by z_j.

v_{0j} : The bias for the j-th hidden unit.

v_{ij} : Weight between i-th input unit and j-th hidden unit.

Y_k : k-th output unit. The unit signal at Y_k is denoted by . The output (activation) signal for Y_k is denoted by y_k.

t_k : k-th output target.

w_{0k}: Bias for k-th output unit.

w_{jk} : Weight between j-hidden unit and k-output unit.

δ_k : Error correction factor for w_{jk} weights.

δ_j : Error correction factor for v_{ij} weights.

4   Firefly Algorithm

The firefly algorithm is a metaheuristic algorithm that is inspired by the flickering behavior of fireflies. There are several important things that are related in the firefly algorithm which are the intensity of light, attractiveness, distance and movement. [5].

A. Light intensity
   Light intensity is obtained from the objective function value. The value of light intensity is obtained from the following formula:

$$I(x) = \frac{1}{1+f(x)} \tag{18}$$

   where:
   (x) : level of light intensity in firefly x
   $f(x)$: the value of the objective function of the firefly x

B. Attractiveness
   Attractiveness is the second most important thing in the firefly algorithm. Where the value of attractiveness is obtained from the following formula:

$$(r) = \beta e_0^{-\gamma r^2} \tag{19}$$

   where :
   β_0 : the attractive value of early firefly
   γ : coefficient of light absorption
   r : the distance between the firefly

C. Distance

The distance between firefly i and j at location x are x_i and x_j which can be determined when laying the point where the firefly is randomly distributed in the Cartesian diagram with the formula:

$$r_{ij} = \|x_i - x_j\| = \sqrt{\sum_{k=1}^{d}(x_{i,k} - x_{j,k})^2} \qquad (20)$$

where:
$x_{i,k}$: k- component from firefly i
d : number of components

D. Movement
The movement of firefly i that moves towards the best level of light intensity can be seen from the following equation:

$$x_i^{new} = x_i + \beta_0 e^{-\gamma r_{ij}^2}(x_j - x_i) + \alpha(rand - \tfrac{1}{2}) \qquad (21)$$

where :
α: random parameter coefficient.
$rand$: random real numbers at intervals [0,1].

5   Simulated Annealing

Simmulated Annealing was first applied in 1983. Simulated annealing is an algorithm for optimization. This algorithm can be used to find approaches to the local optimum solution of a problem [6]. The simulated annealing procedure can be outlined as follows :

a. Initial temperature initialization T_{now}.
b. Initial solution initialization v_i.
c. Evaluate the initial solution as a temporary solution (v_i).
d. Modify the initial solution v_m and re-evaluation as a new solution (v_m).
e. If $(v_m) \leq (v)$ then the temporary solution is the same as the new solution. If not, then real numbers r are generated randomly at intervals [0,1]. Calculate the probability $P = \exp(-\frac{\Delta I}{T})$, if $P > r$ then the new solution is still accepted as a temporary solution. If it reaches the maximum iteration, the algorithm stops. If not, calculate the temperature change $T_i = aT_{now}$ with $0 < a < 1$ and return to step d.

6 Heart Abnormalities detection based on ECG signal characteristics using multi layer perceptron with firefly algorithm and simulated annealing

In the training process data, applying multi layer perceptron with firefly algorithm and simulated annealing (MLP-FA-SA) is used to obtain optimal weight and bias. The steps are:

a. Input MLP-FA-SA parameters. FA parameters are the number of fireflies population, initial attractiveness, air absoption coefficient, random coefficient. For the SA parameters are the value of the initial temperature, final temperature and temperature reduction, while the MLP parameters are the error limit, the maximum iteration and learning rate.

b. Generate initial populations at FA randomly at intervals [0,1]. The population raised is the weight and bias that will be used in MLP.

c. Covertion of individuals in FA into weights and biases in MLP

d. Calculate the new weights and biases with the MLP process in each pattern found in each individual. Furthermore, for each individual MSE value is calculated using the following formula:

$$MSE = \frac{1}{m} \times \{(t_{k1} - y_{k1})^2 + (t_{k2} - y_{k2})^2 + \cdots + (t_{km} - y_{km})^2\}$$

e. Check whether the MSE value obtained is less than the error limit or the maximum iteration. If it meets the error limit or maximum iteration, the solution is optimal. If not, then the FA and SA proceed processes.

f. Change weights and biases in MLP into individuals in the FA.

g. Calculate the value of light intensity of each individual fireflies based on MSE obtained from the MLP process using the following formula:

$$f(x) = \frac{1}{1 + MSE}$$

h. Compare the light intensity of each firefly with the other firefly.

i. Calculate distance, attractiveness, and movement when finding firefly with greater light intensity

j. Calculate the MSE value and the light intensity value of each firefly after the movement.

k. Simulated annealing process by :
   - Determine the worst firefly as the initial solution
   - Modify the worst firefly x' and recalculate the value of the light intensity as a new solution $I(x')$.
   - If $(x) \leq (x')$, the temporary solution is a new solution. If not, then real numbers are randomly generated r at intervals [0,1]. Calculate the probability $P = \exp(-\frac{\Delta I}{T})$, if $P \geq r$ then the new solution is accepted as a temporary solution.
   - Make a decrease in temperature
   - If $T_{now} \leq T_{end}$, then the process stops. If not, then return to step ii

l. Combine all solutions found.

m. Determine the best firefly on the iteration. The best firefly is firefly with the highest light intensity.

n. Update G*, by comparing the existing G* with the best firefly on the iteration.

o. Do the best firefly movement.

p. After the FA and SA processes will get a new population that will become a weight and bias in NLP, then repeat the process (f).

7 Result and Discussion

The data used in this paper is ECG data in the form of electrocardiogram (ECG) soft file output with the extension .jpg. The data were taken as many as 30 ECG images consisting of 15 ECG soft file images from heart abnormalities patients and 15 ECG soft file images from normal patients, respectively.

In this study there are 2 stages that must be done, namely the image processing and the heart abnormalities detection using MLP-FA-SA. The first stage is image processing. All images used will go through several process, which are : grayscale, thresholding, edge detection and segmentation. In the segmentation process, the image size from 180 x 90 pixels will be 20 x 10 pixels. Figure 1 shows the results of the image processing.
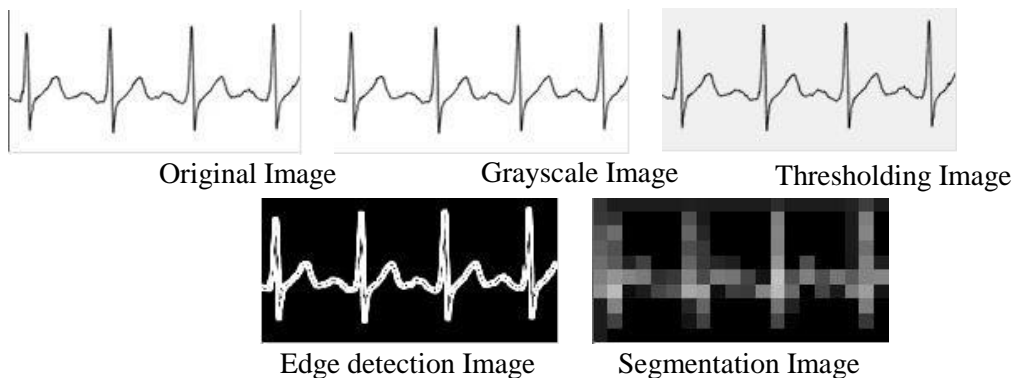

Original Image   Grayscale Image   Thresholding Image

Edge detection Image   Segmentation Image

Figure 1. Image Processing Process Results

The second stage is process of detecting heart abnormalities from ECG images using MLP-FA-SA which consists of a training process and a validation test process. MLP network architecture used consists of 3 layers, namely 200 input layer neurons, 200 hidden layer neurons and 1 output layer neuron. In the training process, MLP networks with firefly algorithm and simulated annealing are used to get optimal weights and biases. This MLP network training uses several parameters consisting of parameters in firefly algorithm training, namely the number of firefly individuals, air absoption coefficient, random value coefficient and initial attractiveness. As for the simulated annealing the

parameters are initial temperature, decreasing temperature, final temperature, meanwhile the multi layer perceptron uses the error limit, maximum iteration and learning rate as its parameter. Variations of parameters used in the MLP training process can be seen in Table 1.

Table 1. Variation of Parameters Used in the MLP Training Process

Individual	Air Coefficient	Random Coefficient	Initial Attractiveness	Learning Rate	Initial Temperature	Final Temperature	Reduction in Temperature
3	0,1	0,2	0,1	0,1	1000	1	500
5	0,2		0,5	0,5			
			1	0,9			

In the testing process the training data uses optimal weights and biases from the training results with some values for parameter variations and the best results obtained can be seen in Table 2.

Table 2. Training Data Testing Results

No	Individual	Initial Attractiveness	Air Coefficient	Learning Rate	MSE	Iteration	Percentage of Test Results
1	3	1	0,2	0,1	0,028031	1	100%
2	5	0,5	0,2	0,5	0,0137485	1	100%

The validation test process is performed on 10 ECG images using optimal weights and biases whose results can be seen in Table 3.

Table 3. Validation Test Results

No	Individual	Initial Attractiveness	Air Coefficient	Learning Rate	MSE	Iteration	Percentage of Test Results
1	3	1	0,2	0,1	0,028031	1	100%
2	5	0,5	0,2	0,5	0,0137485	1	100%

8   Conclusion

To conclude, the implementation of the program for detecting heart abnormalities based on the characteristics of ECG signals was successfully carried out with 100% in the *training* dan *testing* process with many individuals of 5, MSE 0,0137485 with 1 iteration, using the random coefficient = 0,2, the air coefficient = 0,2, initial attractiveness = 0,5, initial temperature = 1000, final temperature = 1, reduction in temperature = 500, learning rate = 0,5 and error limit = 0,1. By using the optimal weights and biases, the program can recognize the validation test data patterns with 100% success.

9   References

[1]   Hammad, M., A. Maher, K. Wang, F. Jiang, M. Amrani, 2018, Detection of Abnormal Heart Conditions Based on Characteristics of ECG Signals, Measurement.

[2]   Scarmoth, L., 1990, *An Introduction to Electrography,* Blackwell Scientific Publication, Oxford.

[3]   Fausett, L., 2003, Fundamentals of Neural Networks: Architectures, Algorithms, and Applications, Printice-Hall Inc., London.

[4]   Siang, J.J., 2005, Jaringan Syaraf Tiruan dan Pemrogamannya menggunakan MATLAB, Penerbit ANDI, Yogyakarta.

[5]   Yang, X. S., 2010, Engineering Optimization: An Introduction with Metaheuris-tic Applications, Wiley & Sons, Inc, New Jersey.

[6]   Kirkpatrick, C. D. Gelatt, M. P. Vecchi, 1983, Optimization by Simulated Annealing, Science, Vol 220, No 4598, Halaman 671-680.

[7]   Baharuddin, 2008, *Perancangan Alokasi Kanal Dinamik pada GSM*, Teknik A, Vol.1, N0.29, April 2008, ISSN: 0854-8471.

[8]   Putra, darma, 2008, Sistem Biometrika : Konsep Dasar Teknik Analisis Citra, penerbit Andi, Yogyakarta.

[9]   Basuki, A., Palandi, J. F., dan Fatchurrochman, 2005, *Pengolahan Citra Digital Menggunakan Visual Basic*, Penerbit Graha Ilmu, Yogyakarta.

[10]   Ahmad, U., 2005, Pengolahan Citra Digital dan Teknik Pemrogamannya, Graha Ilmu, Yogyakarta..

[11]   Riedmiller, M., 1994, Advanced Supervised Learning in Multi-Layer Perceptrons from Backpropagation to Adaptive Learning Algorithms, International Journal on Computer Standards and Interfaces.