Classification of tumor based on magnetic resonance (MR) brain images using wavelet energy feature and neuro-fuzzy model

A Damayanti¹, and I Werdiningsih²

¹,²Department of Mathematics, Faculty of Science and Technology, Universitas Airlangga, Surabaya, Indonesia

Email : ¹aulid@fst.unair.ac.id

Abstract. The brain is the organ that coordinates all the activities that occur in our bodies. Small abnormalities in the brain will affect body activity. Tumor of the brain is a mass formed as a result of cell growth not normal and unbridled in the brain. MRI is a non-invasive medical test that is useful for doctors in diagnosing and treating medical conditions. The process of classification of brain tumor can provide the right decision and correct treatment and right on the process of treatment of brain tumor. In this study, the classification process performed to determine the type of brain tumor disease, namely Alzheimer's, Glioma, Carcinoma and normal, using energy coefficient and ANFIS. Process stages in the classification of images of MR brain are the extraction of a feature, reduction of a feature, and process of classification. The result of feature extraction is a vector approximation of each wavelet decomposition level. The feature reduction is a process of reducing the feature by using the energy coefficients of the vector approximation. The feature reduction result for energy coefficient of 100 per feature is 1 x 52 pixels. This vector will be the input on the classification using ANFIS with Fuzzy C-Means and FLVQ clustering process and LM back-propagation. Percentage of success rate of MR brain images recognition using ANFIS-FLVQ, ANFIS, and LM back-propagation was obtained at 100%.

1. Introduction

The brain is the organ that coordinates all the activities that occur in our bodies. Small abnormalities in the brain will affect body activity. The human brain weighs 2% of the adult's body weight which receives 20% of cardiac output and requires 20% of the body's oxygen consumption and about 400 kcal of energy per day. Therefore, brain tissue is very susceptible to changes in oxygen and blood glucose, stop blood flow 10 seconds alone can eliminate human consciousness and can damage the brain permanently if blood flow stops within minutes [1]. One of the brain disorders is a brain tumor. Tumor of the brain is one of the deadly illness at the world, namely the growth of abnormal cells in the brain structure [2]. Tumor of the brain is a lump that results from the growth of morbid and unbridled cells in the brain. Tumor of the brain can become malignant or benign. The benign brain tumor is a non-cancerous type and is not harmful to surrounding tissue, known as the non-metastatic tumor. While brain cancer is a cell that develops very quickly and destroys the surrounding tissue [3]. In Indonesia, within a decade of brain tumors can occur in children, are adults at the age of 30-70 with shoulders aged 40-65 years. occurred brain tumor incidents [4]. In America, there are 35,000 new
cases of brain tumors each year, and central nervous system primary tumors are found in 10% of all neurological diseases found in General Hospital [5]. Decision-making becomes very complicated with many types of tumor of brain.

Therefore the classification of MRI images of the brain is very important to classify the types of brain tumors, which can affect people's lives. Magnetic resonance imaging (MRI) is a device of medicine on radiological diagnostic examination which results in high-quality image recording of the structure of anatomies body of human, particularly the brain and can be used for soft tissue imaging [6]. MRI is a noninvasive medical test that is beneficial to physicians in diagnosing and treating medical conditions. The process of classification of brain tumors can provide the right decision and correct treatment and appropriate in the process of treatment of tumor of brain. Classification of brain tumors can be done by using supervised classification and unsupervised classification, one of which is support vector machine (SVM) [7] and knearest neighbors (kNN) [8], self-organization map (SOM) [7] and fuzzy c-means [9]. All the classification methods that have been done, achieved good results, and the supervised classification of the success rate (accuracy) of the classification is better than the unsupervised classification. The method of classification used in this research is the Neuro-Fuzzy systems method, is a combination of inference fuzzy system mechanism described in the neural network architecture and a special approach to the development of Neuro-Fuzzy, Adaptive Neuro-Fuzzy Inference Systems (ANFIS). ANFIS has denoted prominent results in modeling functions of nonlinear. Anfis learning process through features of many data and based on the value of parameters and errors given [10]. In biomedical engineering, the most successful implementation of ANFIS is the process of classification [11].

The transform of the wavelet is an effective technique for extraction of a feature because wavelet transforms can perform various image analysis with various resolutions and can perform MRI data extraction process [6]. Since the results of wavelet transform require large memory storage, a method is needed to reduce the dimensions of the data. Energy coefficient of wavelet decomposition can be used to reduce dimensions and improve classification performance [12].

In this study, will use the coefficient of the energy of transformation of wavelet for dimension reduction in MR Brain Images data and be using ANFIS for MR Brain Images data classification process according to the type of diseases, such as Glioma, Carcinoma, and Alzheimer's. In previous studies, Damayanti (2014) have done classifications MR brain images using gradient descent backpropagation with the level of success of recognition pattern MR brain images of 95%. In this study classifications MR brain images using anfis , expected percentage this classifications success rate better than neural-networks.

2. Methodology

Stages in this study there are three, namely: feature extraction, feature reduction, and classification process. Feature extraction stage is the stage to get the feature of MR brain image that is approximation of vector using wavelet decomposition. The feature reduction stage is a process step to reduce the feature by using energy coefficient of the approximation vector. The last stage is the MRI brain image classification stage using ANFIS. Classification results are divided into four kinds, namely normal, Alzheimer, Glioma, and Carcinoma. The process stages in this study can be seen in figure 1.
Figure 1. The system block diagram

2.1. Feature Extraction

Feature extraction is the process of capturing features of an image. At this study, feature extraction of MR brain image using wavelet decomposition, is decomposition of image process making use wavelet transform. The type of wavelet used is Haar. Wavelet haar is one of the simplest types of discrete wavelet transforms because the wavelet haar uses a square wave period [13]. The result of the feature extraction process is a vector feature. This feature vector is the output of the wavelet decomposition of the decomposition vector C and the matrix S [C, S]. Wavelet decomposition used up to level 3.

Feature Extraction is a process that takes the features of an image. In this study, feature extraction of MR brain image using wavelet decomposition is the process of image decomposition using wavelet transformation. The type of wavelet used is Haar. Wavelet Haar is one of the simplest types of discrete wavelet transforms because the wavelet Haar uses a square wave period [13]. The result of the feature extraction process is a vector feature. This feature vector is the output of the wavelet decomposition of the decomposition vector C and the matrix S [C, S].

Wavelet transform is a windowing technique using variable sizes to adjust the information and time of the signal. The advantage of wavelet transformation is taking a scale that is a time-scale rather than a time-frequency of a signal. [14]. Transformation of the Discrete wavelet is the implementation of wavelet which uses a scale and position. The basic of the wavelet of discrete is, suppose x(t) is a square-integrated function, then the continuous x(t) relative to w(t) is formulated as Eq.(1)-(2)

$$W_{\psi} (a,b) = \int_{-\infty}^{\infty} x(t) \psi_{a,b}(t) dt$$

where

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \varphi \left( \frac{t-a}{b} \right)$$

The wavelet calculation is derived from the primary wavelet value with the translation and dilation process, parameter of a is the factor of the dilation and the parameter of b is the parameter of the translation which both must be real positive. on the development of wavelet analysis, The most commonly used wavelet is the Harr wavelet, which is the simplest and most preferred wavelet in many applications [15].

The discrete wavelet transform, which can be expressed as follows.

$$ca_{j,k} (n) = DS \left[ \sum_{n} x(n) g_{j}^{*} (n - 2^{j} k) \right]$$

$$cd_{j,k} (n) = DS \left[ \sum_{n} x(n) h_{j}^{*} (n - 2^{j} k) \right]$$
ca_{j,k} and cd_{j,k} is approximation components coefficients. g(n) and h(n) are the low-pass filter and high-pass filter. Index of j and k is scale of the wavelet and factors of translation. Operator of DS represent down sampling [16]. Schematic diagram of 2D DWT can be seen at figure 2 with images of four sub-bands (LL, LH, HH, HL) at every scale.

2.2. Feature Reduction
The feature reduction is a process for reducing the feature by using the energy coefficient of the vector approximation. In this study, the feature reduction was performed by summing 100 energy per feature of the coefficient energy value obtained by squaring each vector approximation. The vector approximation is the result of the extraction of the MR brain images coefficients.

2.3. Adaptive Neuro-Fuzzy Inference Systems (ANFIS)
The ANFIS method was first introduced by Jang with embedded Fuzzy Inference System (FIS) into an adaptive networking framework. The network of adaptive is a network structure with a number of interconnected nodes. The output of this adaptive node depends on the modifications associated with this node. The learning rule determines how this parameter should be updated to minimize errors. Architecture of ANFIS can be seen in figure 3, it appears that the neuro-fuzzy system consists of five layers with different functions for each layer. Each layer consists of several vertices denoted by a box or circle. The square symbol represents an adaptive node meaning that its parameter value may change with learning and the circle symbol represents a non-adaptive node whose value is fixed.

The layers in the ANFIS structure can be explained as follows:

**Layer 1.** In this layer, all nodes are adaptive nodes (parameters may change) with function of node:

\[ O_{i,j} = \mu_{A_i}(x) \]  
\[ O_{i,j} = \mu_{B_{i-2}}(y) \]

With x and y being inputs at the i node, Ai (or Bi) is the membership function of each node. Oj node I serves to declare the membership degree of each input to the fuzzy set A and B. The membership
function used is the generalized bell type. The parameters $a$, $b$, $c$, in the generalized bell membership function are called adaptive premise parameters. 

*Layer 2.* In this layer all nodes are non-adaptive (fixed parameters). The function of this node is to multiply every incoming input signal. Node function:

$$O_{2,j} = w_j = \mu_{A_i}(x)\mu_{B_i}(y), \quad \text{for} \quad i = 1,2$$

Each node output is the firing strength of each fuzzy rule. This function can be expanded if the premise section has more than two fuzzy sets. The number of vertices in this layer shows the number of rules that are formed. The multiplication function used is the interpretation of the hyphen and by using the t-norm operator.

*Layer 3.* Each node in this layer is a non-adaptive node that displays function of firing strength has normalized that is the output ratio of the $i$ node of the previous layer to the entire output of the previous layer, with the shape of the node function:

$$O_{3,j} = \frac{w_j}{w_1 + w_2}, \quad i = 1,2$$

When more than two rules are established, the function can be extended by dividing $w_i$ by the total number $w$ for all rules.

*Layer 4.* Each node in this layer is an adaptive node with a node function:

$$O_{4,j} = \bar{w}_j f_i = \bar{w}_j (p_i x + q_i y + r_i)$$

by is the normalized degree of activation of layer 3 and the parameters $p$, $q$, $r$ denotes the adaptive consequent parameters.

*Layer 5.* In this layer, there is only one fixed node whose function is to sum all the entries. Node function:

$$O_{5,j} = \sum_i \bar{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i}$$

The adaptive network with these five layers is equivalent to the TSK fuzzy inference system [10].

In the ANFIS structure of Fig.3, the adaptive nodes are in the first and fourth layers. The node in the first layer contains the nonlinear premise parameter while in the fourth layer contains the linear consequent parameter. Updates of ANFIS parameters are obtained from the learning process of the ANFIS neural network. Learning on ANFIS uses hybrid learning algorithm, that is forward pass using Recursive Least Square Estimator method and backward pass using gradient descent method. [10]

In the forward direction, the premise parameter is fixed. Using the Recursive Least Square Estimator (RLSE) method, the consequent parameters are corrected based on the input-output data pair. The RLSE method can be applied because the consequent parameters that are fixed are linear parameters. The RLSE method will speed up the hybrid learning process. Then after the consequent parameters are obtained, the input data is passed through the adaptive network again and the output of this adaptive network is compared with the actual output. [10]

In the backward direction, the consequent parameters are fixed. The error that occurs between the adaptive network output and the actual output is propagated back by using a gradient descent to correct the premise parameter [10]. This learning is known as back-propagation-error Algorithm. Stages of the learning process and ANFIS testing process can be seen in figure 4.

2.4. **Fuzzy Learning Vector Quantization (FLVQ)**

Fuzzy learning vector quantization (FLVQ) is a method that combines learning method of Learning Vector Quantization (LVQ) neural network and data grouping method of fuzzy C-Means (FCM). This method was first introduced by Tsao (1994). Stages of Fuzzy learning vector quantization training are as follows:
1. Initialization parameters include the total of clusters \((C)\), values \(m_i\) and \(m_f\) , maximum iterations and error limits.

2. Determination of the initial value of each cluster \((v_0)\)

3. Calculation of the weight value of each iteration

\[
m = m_i + k \left( \frac{m_f - m_i}{N} \right)
\]

4. Calculation of the partition matrix of each cluster, using the FCM equation as follows

\[
a_{ij,k} = \left\{ \sum_{r=1}^{M} \left[ \frac{1}{d(v_i - v_r)^2} \right] \right\}^{-m}
\]

dengan \(1 \leq i \leq M \) dan \(1 \leq j \leq C\)

5. Calculation of cluster learning rate \((\eta)\).

\[
\eta_{j,k} = \left( \sum_{i=1}^{M} a_{i,j,k} \right) \quad 1 \leq j \leq C
\]

6. Determination of new cluster center

\[
v_{j,k} = v_{j,k-1} + \eta_{j,k} \left( \sum_{i=1}^{M} a_{i,j,k} \right) (v_i - v_{j,k-1}) \quad 1 \leq j \leq C
\]

7. Calculation of the error value.

\[
E_k = \sum_{j=1}^{C} \left\| v_{j,k} - v_{j,k-1} \right\|^2
\]

The iteration process continues until the target error is reached or the maximum iteration is reached.

**FIGURE 4.** (a) Flowchart of ANFIS training process and (b) Flowchart of ANFIS testing process
2.5. Algorithm of Learning Back-propagation Neural Network

2.5.1. Gradient Descent Back-propagation
In the learning process using gradient descent backpropagation consists of 3 stages, namely the forward bait, propagation error, and renewal of weight. In the forward feed process, each input unit will spread the signal to each hidden unit, then each hidden unit will calculate its activation. From the hidden unit will send a signal to each unit of output. Then, each unit of output will be counted its activation function which is the result of a response to the input given the network. The process of propagation error is the process of comparing the activation function of the output with the target value. If the results obtained are different then it will result in an error value. Based on the error, the error correction factor used to distribute the error and output back to the previous layer. Then we perform the process of updating the weights on all layers [17].

2.5.2. Levenberg-Marquadt Algorithm
According to [18], Levenberg Marquardt (LM) algorithm is the development of backpropagation algorithm with the determination of changes of weight and bias using hessian matrix approach (H) which can be calculated with equation 16. and the gradient can be calculated by equation 17. The changes of weight on the network are calculated by equation 18 and weight fixing is determined by equation 19.

\[
H = J^T e \\
g = J^T J \\
\Delta X = (J^T J + \mu I)^{-1} J^T e \\
X = X + \Delta X
\]

where \( X \) = function of network weight and bias
\( e \) = vector that states all errors in the network output
\( \mu \) = learning rate
\( I \) = identity matrix

3. The Dataset and Experimental Result
The MR Brain image dataset used for this research were downloaded from the Harvard Medical School website (http://www.med.harvard.edu/AANLIB/home.html). The MR Brain dataset image consists of 12 normal and 48 abnormal images. In the abnormal images consist of glioma, carcinoma, and Alzheimer's. Out of a total of 60, MR Brain image datasets are divided for 40 training data and 20 data testing. Brain data images used in this study are .gif type with pixel size 256 x 256. Image of MR brain can be shown in figure 5.

![Figure 5. MR Brain images (a). Alzheimer (b). Glioma (c). Carcinoma (d). Normal](image-url)
Stages in this study are feature extraction, reduction of feature, and process of classification. In feature extraction process consists of four stages, namely: wavelet decomposition, coefficient extraction, normalization, and computational energy. The result of feature extraction is vector approximation of each level and ignore level 1. This is because at the decomposition of wavelet level 1 has many frequencies and noise detail. The size of the vector approximation at level 2 is 64 x 64 pixels and at level 3 is 32 x 32 pixels. Since the size of the approximation vector is too large, a feature reduction is performed. The result of feature reduction for energy coefficient of 100 per feature is 1 x 52 pixels. This is the vector that will be entered into the classification process. The feature reduction uses the energy coefficients of the vector approximation. Research stage after feature reduction is classification. And the classification method used in the next process is ANFIS.

The classification step using ANFIS is the first step is the determination of input on ANFIS. The input data used in the training process are 40 images consisting of 32 abnormal images of 12 Alzheimer, 8 Carcinoma, 12 Glioma and 8 normal images while the data testing as much as 20 images consisting of 16 abnormal images of 6 Alzheimer, 4 Carcinoma, 8 Glioma and 4 normal. The second stage is the determination of membership function. Membership functions are used generalized-bell because generalized-bell has the smoothness and simplicity of notation with much information and degrees of freedom to improve the steepness at the crossover point is the point where membership value is 0.5 [10]. The third stage is the determination of membership function number. The fourth stage is the initialization of the premise parameter. The premise parameter to be initialized with Bell's membership function is \{a, b, c\} for each membership function. The value of parameter b is one, but for parameter a and c are initialized randomly. The fifth stage is the ANFIS learning process. In the process of learning ANFIS used 2 methods for the clustering process of Fuzzy C-Means (FCM) and fuzzy learning vector quantification (FLVQ). The last stage is the process of testing the MR brain images.

In this study, ANFIS classification results will be compared with the results of the classification process LM learning backpropagation and the result of Damayanti (2014). In ANFIS process using variation of learning rate parameter value (lr) and momentum (mc), that is 0.1, 0.4, and 0.9. In ANFIS classification using FCM for clustering process, 100% accuracy percentage is obtained for parameter variation lr = 0.1, mc = 0.1; lr = 0.1, mc = 0.4; lr = 0.4, mc = 0.9; and lr = 0.9, mc = 0.4, the results can be seen in figure 6. But the variation of the parameters resulting in the smallest error and with minimal iteration is at lr = 0.1 and mc = 0.4 with error of 0.00000664 with iteration as much 4. While the ANFIS classification with FLVQ for clustering process, obtained 100% accuracy percentage for all variations of learning rate parameter and momentum. The results of classification of MR brain images using ANFIS can be seen in table 1.

In classification LM learning back-propagation using a variation of the number of hidden layer node (hl) and learning rate (\(\mu\)). The number of nodes in the hidden layers used are 26, 52 and 78. While the variation of learning rate value used is 0.1, 0.4 and 0.9. The result of the best accuracy percentage of 100% is at the value of hl = 26 and \(\mu = 0.1, \mu = 0.9\), the result can be seen in figure 8. The results of the comparison of the process of classification of MR brain images using ANFIS, LM learning and result of Damayanti (2014) can be seen in table 2.
Figure 6. The results of the ANFIS classification process in MR brain images using FCM where (a) lr = 0.1, mc = 0.1, (b) lr = 0.1, mc = 0.4, (c) lr = 0.4, mc = 0.9, and (d) lr = 0.9, mc = 0.4

Table 1. The results of classification of MR brain images using ANFIS

| Learning Rate | Momentum | Percentage Accuracy ANFIS | Percentage Accuracy ANFIS-FLVQ |
|---------------|----------|---------------------------|-------------------------------|
| 0.1           | 0.1      | 100                       | 100                           |
| 0.1           | 0.4      | 100                       | 100                           |
| 0.1           | 0.9      | 85                        | 100                           |
| 0.4           | 0.1      | 85                        | 100                           |
| 0.4           | 0.4      | 90                        | 100                           |
| 0.4           | 0.9      | 100                       | 100                           |
| 0.9           | 0.1      | 95                        | 100                           |
| 0.9           | 0.4      | 100                       | 100                           |
| 0.9           | 0.9      | 95                        | 100                           |
Table 1 shows that the classification using ANFIS with FLVQ is the best classification result compared to ANFIS with FCM.

![Testing Using Data Testing: Recognition Rate = 100.00 %](image1)
![Testing Using Data Testing: Recognition Rate = 100.00 %](image2)

(a) (b)

**Figure 7** The result of classification LM learning where (a) \( h_l = 26 \) and \( \mu = 0.1 \), (b) \( h_l = 26 \) and \( \mu = 0.9 \)

| Methods                  | Percentage Accuracy |
|--------------------------|---------------------|
| GD Back-propagation      | 95%                 |
| (Damayanti,2014)         |                     |
| LM Back-propagation      | 100%                |
| ANFIS-FLVQ               | 100%                |
| ANFIS                    | 100%                |

**Table 2. The Result of Classification MR brain images**

4. Conclusion

The classification of MR brain images by type of disease, among others Glioma, carcinoma, and Alzheimer's using ANFIS is the aim of this study. Stages of the classification of MR brain images are feature extraction, reduction of feature, and process of classification. The resulting feature of the feature reduction process for 100 energy coefficient is 52 features, which is the input of the classification process. The result of classification MR brain image using ANFIS-FLVQ, ANFIS, and LM back-propagation by 100%.

**References**

[1] Wilson P 2006 Anatomi dan Fisiologi Sistem Syaraf *EGC Publisher* Jakarta
[2] Nagori M, Mutkule S and Sonarkar P 2013 Detection of Brain tumor by mining MRI Images *Volume 2 Issue 4 and ISSN: 2278-1021*
[3] John P 2012 Brain Tumor Classification Using Wavelet and Texture Based Neural Network *International Journal of Scientific & Engineering Research Vol 3 Issue 10 pp 1-7*
[4] Jardi I Penyakit Alzheimer 24 Februari 2012 [http://www.repository.usu.ac.id/bitstream/](http://www.repository.usu.ac.id/bitstream/)
[5] Berttolone 1982 Tumor of The Central Nervous System Concepts in Cancer Medicine pp 649-659
[6] El-Dahshan E A, Hosny T and Salem A M 2010 Hybrid intelligent techniques for MRI Brain Images classification Cairo pp 433–441
[7] Chaplot S, Patnaik L M and Jagannathan N R 2006 Classification of magnetic resonance brain images using wavelets as input to support vector machine and neural network. Biomedical Signal Processing and Control 1 pp 86–92
[8] Fletcher-Heath L M, Hall L O, Goldgof D B et al 2001 Automatic segmentation of non-enhancing brain tumors in magnetic resonance images Artificial Intelligence in Medicine 21 pp 43–63
[9] Maitra M and Chatterjee A 2008 Hybrid multi-resolution slantlet transform and fuzzy c-means clustering approach for normal pathological brain MR image segregation Medical Engineering and Physics 30(5) pp 615–623
[10] Jang J-SR 1997 ANFIS : Adaptive Neuro-Fuzzy based Inference Systems IEEI Trans On System Man and Cybernetics
[11] Belal S Y, Taktak A F G, Nevill A J, Spencer S A, Roden D and Bevan S 2002 Automatic detection of distorted plethysmogram pulses in neonates and paediatric patients using an adaptivenetwork-based fuzzy inference system Artificial Intelligence in Medicine 24 pp 149–165
[12] Mousa R, Munib Q and Mousa A 2005 Breast Cancer Diagnosis System Based on Wavelet Analysis and Fuzzy Neural Jordan pp 713–723
[13] Roy K and Bhattacharya P 2008 Optimal features subset selection and classification for Iris recognition. Journal Image Video Process doi:10.1155/2008/743103
[14] Schmeelk J 2002 Wavelet transforms on two-dimensional images Mathematical and Computer Modelling 36(7) pp 939–948
[15] Kaur L, Gupta S, Chauhan R C and Saxena S C 2007 Medical ultrasound image compression using joint optimization of thresholding quantization and bestbasis selection of wavelet packets Digital Signal Processing 17(1) pp 189–198
[16] Nanni L and Lumini A 2008 Wavelet decomposition tree selection for palm and face authentication Pattern Recognition Letters 29(3) pp 343–353
[17] Fausett L 2003 Fundamental of Neural Network Architecture, Algorithm, and Application Printice-Hall Inc London
[18] Hagan M and Menhaj M B 1994 Training Feedforward Networks With The Marquard Algorithm IEEE Transactions On Neural Networks 5(6) pp 989-993