

Internal content classification of ultrasound thyroid nodules based on textural features

Anan Nugroho^{1*}, Hanung Adi Nugroho², Noor Akhmad Setiawan², Lina Choridah³

¹Department of Nursing, Surya Global Institute of Health Science

²Department of Electrical Engineering and Information Technology, Faculty of Engineering, Universitas Gadjah Mada

³Department of Radiology Sardjito Hospital, Faculty of Medicine, Universitas Gadjah Mada

Article history:

Received: 30 October 2016 / Received in revised form: 25 November 2016 / Accepted: 29 November 2016

Abstract

Ultrasound (US) is one of the best imaging modalities on thyroid identification. The suspicious thyroid is indicated in the existence of palpable nodules whose solid or cystic composition. Solid nodules have high possibility to be malignant than cystic. An effort to detect and classify the internal content of thyroid nodule has become challenge problem in radiology area. Operator dependence of ultrasound imaging makes it complicated due to missing interpretation among radiologists. Objective Computer Aided Diagnosis (CAD) was designed to solve it which works on texture analysis of histogram statistic, gray level co-occurrence matrices (GLCM) and gray level run length matrices (GLRLM). The fine-needle aspiration cytology (FNAC) is not needed because the textural pattern is significantly different between solid and cystic nodules. Multi-layer perceptron (MLP) was adopted to do classification process for 72 US thyroid images yield an accuracy of 90.28%, the sensitivity of 87.80%, specificity of 93.55% and precision of 94.74%.

Keywords: Thyroid; Nodule; Ultrasound; CAD

1. Introduction

Thyroid nodules are solid or liquid lumps in the thyroid gland. The incidences of nodule are the common problem in society and increasing with age [1]. About 4-7% of adults possess palpable nodules, biopsy, and ultrasound imaging detects these well [2]. Thyroid nodules are more common in elderly patients, women and people with iodine deficiency [3]. More 90% of thyroid nodules are benign and do not require special treatment. However, 5% nodules are malignant, require early detection and comprehensive treatment [4].

Several studies have explained identification of ultrasound thyroid nodules whether benign or malignant [5,6,7]. An important parameter to identify thyroid malignancies is classifying the internal content of nodule [8,9]. Most malignant nodules are solid on ultrasound imaging with sensitivity 69%-75% [5]. However, ultrasounds are dependent operator modalities, leads to high subjective interpretation among radiologists whose vary on experience and expertise [1]. Fine-needle aspiration cytology (FNAC) is a further procedure to ensure nodules composition. Noted, not all of the nodules should undergo this procedure, FNAC is probably unnecessary if the nodule dominantly cystic [5]. An unwelcome thing if the surgery is undergoing on suspected malignant nodule, but it turns out the results of surgery shown a benign nodule. Therefore, we proposed an objective CAD to overcome these matter. Textural features embedded on CAD to recognize a solid and cystic pattern.

Main parameter malignancy of thyroid nodules can be measured from the uniformity of dark texture in ultrasound images. Textural features produced by statistical histogram [10,11], Gray Level Co-occurrence Matrices (GLCM) [12] and Gray Level Run Length matrices (GLRLM) [13,10]. Harralick [14,15] and Galloway [16] respectively initiators of GLCM and GLRLM texture pattern analysis in 2D images. Both were formulated 21 GLCM's and 11 GLRLM's textural features.

We have prepared 72 ultrasound image taken from department radiology of Sardjito hospital. All of the textural values are classified by multi-layer perceptron. The performance of the system is validated by confusion table consist of accuracy, sensitivity, specificity, precision and Negative Predictive Value (NPV). This study aims to classify thyroid nodules into solid and cystic as core parameters on thyroid cancer. Hopefully, the results can be used as objective consideration in making an accurate diagnosis.

2. Materials and Methods

The steps of this study are illustrated on the flowchart as shown in Fig.1. The theoretical foundations of textural algorithms and perceptron and a briefly review of ultrasound imaging and pre-processing were presented here. Features of extraction were carried out to obtain measurable textural values on varies images of thyroid nodules based on the statistical histogram, GLCM, and GLRLM. MLP was adopted to classify all of the features and validated by confusion matrix.

* Corresponding author.

Email: anan.sie13@mail.ugm.ac.id

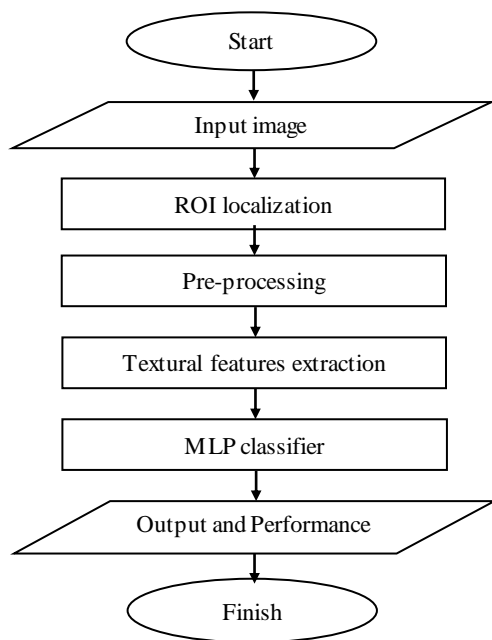


Fig. 1. Steps of work

2.1. Ultrasound Images

Digital ultrasound images were acquired from Doppler ultrasound machine of General Electric P3 Voluson series within 7 Mhz probe transducer as shown in Fig.2. The amount of 72 images were taken during 2011 to 2014 periods and divided into 41 cysts and 31 solids. All of the images had been consulted by the expert radiologist. See Fig.3 for the sample. A pair of images in Fig.3 representing for right lobe and left lobe. Two masks in the bottom of Fig.3 illustrate location of nodule.



Fig. 2. General Electric P3 Voluson Doppler ultrasound machine

Region of interest (ROI) localization was applied to get nodule area. Localized nodul area provide low computational cost and giving accurate features. ROI localization is shown

in Fig. 4. Dominant dark or gray local area is indicated cystic or solid nodule respectively.

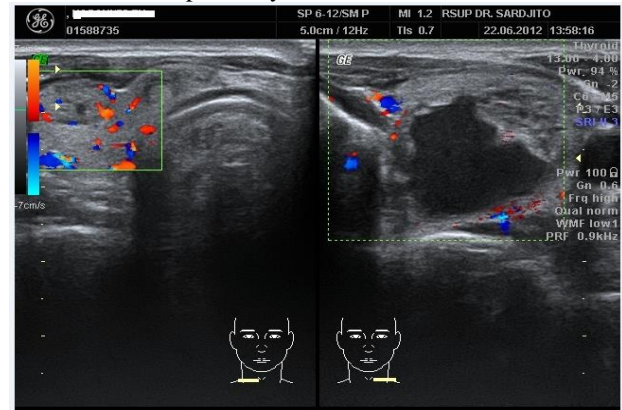


Fig. 3. US thyroid nodule

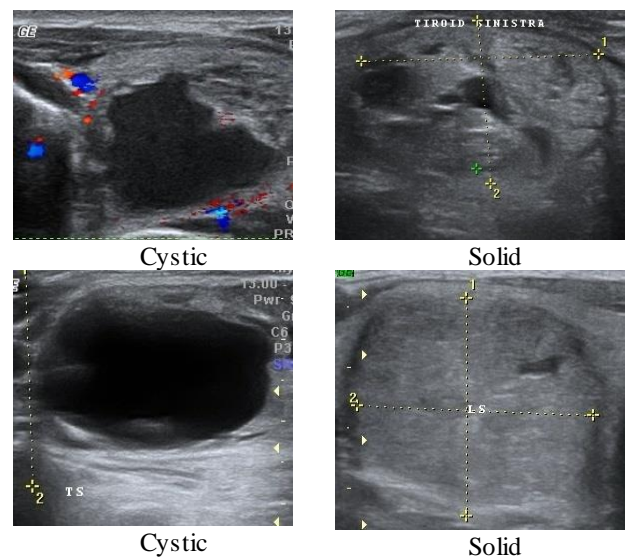


Fig. 4. ROI of cystic and solid thyroid nodule

2.2. Pre-processing

Pre-processing is mean an effort to reduce marking area in which nodule exist by the median filter. Marking area lead to improper extracted textural features. The main idea of the median filter is to run through each pixel along an image and replace each intensity with the median of neighboring values. The mask pattern of neighbors is called window which slides pixel by pixel over an image. The median is defined by middle value after all the entries in the window are sorted computationally.

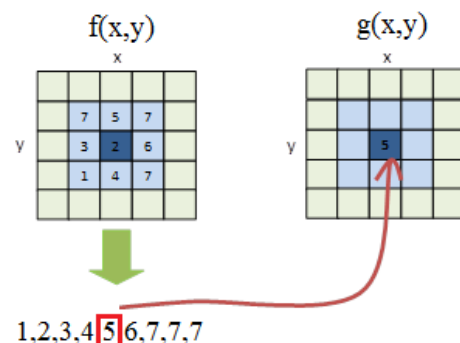


Fig. 5. Windowing of median filter

Fig.5 illustrates eight neighborhood pixels on 3×3 window of median filtering.

2.3. Textural Features

Jain [17] defines texture as a relationship neighbor pixels intensity which is repeated forming a basic pattern. The basic pattern is mean texture elements or texels. But sometimes there is a complicated repetition pattern and poorly understood by human vision. Hence, the sum of methods is required to describe related pattern into measurable values as an objective information namely texture features.

The simplest way, texture features are extracted by the statistical histogram. Here, histogram means as the discrete bar chart illustrating probabilities of each pixel intensity appears in the whole image. A great value states that the pixels on an intensity that often arise. In the gray-scale image as used in this study, the number of gray levels (L) as much as 256 level from 0 to 255. A value of 0 for dark black imagery and a maximum of 255 to a maximum white image. The dominance of pixels with levels approaching 255 will tend to make the image appear bright white. Histogram of various textures can be seen in Fig. 6.

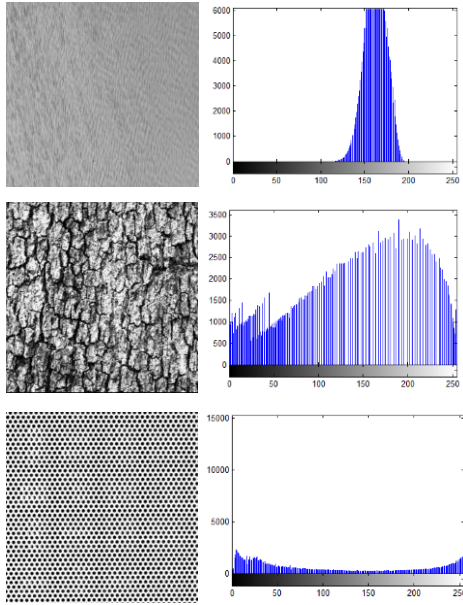


Fig. 6. Histogram various gray-scale images [18]

A histogram is useful to observe the spread of the intensity value that can be used as a basis to represent the image of the texture pattern. The values which are derived from histogram also named as first order statistic features. The following values are explained below.

1) Mean [17]

$$m = \sum_{i=0}^{L-1} i \cdot p(i) \quad (1)$$

In this case, i is the level of gray in the image of f and $p(i)$ represents the probability of occurrence of i and L declare the highest value of gray levels. The above formula will produce an average brightness of objects.

2) Mode [17]

Mode values stated gray level that appears most of all pixels.

3) Variance

$$\sigma^2 = \sum_{i=1}^{L-1} (i - m)^2 p(i) \quad (2)$$

σ^2 can be called variance or second order normalized moment because $p(i)$ is a function of probability. M is the Mean value in the Eq. (1).

4) Standard Deviation

$$\sigma = \sqrt{\sum_{i=1}^{L-1} (i - m)^2 p(i)} \quad (3)$$

5) Skewness [17]

$$\text{Skewness} = \sum_{i=1}^{L-1} (i - m)^3 p(i) \quad (4)$$

6) Kurtosis [17]

$$\text{Kurtosis} = \sum_{i=1}^{L-1} (i - m)^4 p(i) - 3 \quad (5)$$

7) Entropy [17]

$$\text{Entropy} = - \sum_{i=0}^{L-1} p(i) \log_2(p(i)) \quad (6)$$

8) Energy [18]

$$\text{Energy} = \sum_{i=0}^{L-1} [p(i)]^2 \quad (7)$$

9) Smoothness [18]

$$\text{Smoothness} = 1 - \frac{1}{1 + \sigma^2} \quad (8)$$

Gray Level Co-occurrence Matrices (GLCM) was published by Haralick et al. [19] with 14 characteristic value that represents the texture pattern. GLCM was known as the 2nd order of texture pattern calculation. Texture measurements on first order statistical of the histogram are not paid attention to the relationship of neighborhood pixels, while the 2nd order statistic, the relationship between the couple two original image pixel is calculated [18].

For example, $f(x,y)$ is an image with a size of N_x and N_y which has a pixel with a possibility of up to L level and \vec{r} is a vector spatial direction. $GLCM_{\vec{r}}(i,j)$ is defined as the number of pixels with $j \in 1, \dots, L$ effect at the offset \vec{r} to pixels with values $i \in 1, \dots, L$, which can be expressed by the following formula [20] :

$$GLCM_{\vec{r}}(i,j) = \# \left\{ (x_1, y_1), (x_2, y_2) \in (N_x, N_y) \times (N_x, N_y) \mid f(x_1, y_1) = i \wedge f(x_2, y_2) = j \wedge \vec{r} = (x_2 - x_1, y_2 - y_1) \right\} \quad (9)$$

where the offset \vec{r} an angle and pixel distance. Fig.7 [18] shows the four directions of GLCM.

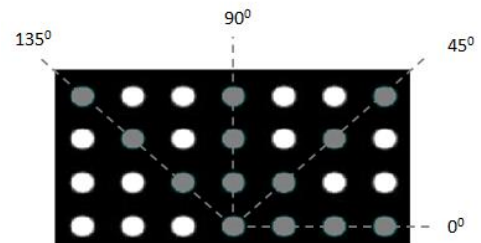


Fig. 7. GLCM angle 0°, 45°, 90°, and 135° [16]

Fig. 8 [16] gives an example GLCM matrix for angle 0° with 1-pixel distance. Notation (0,0) states the relationship of two horizontal rows of pixels with pixel value 0° followed by pixel value 0 on the right. Arranged matrix from this calculation was called GLCM framework matrix.

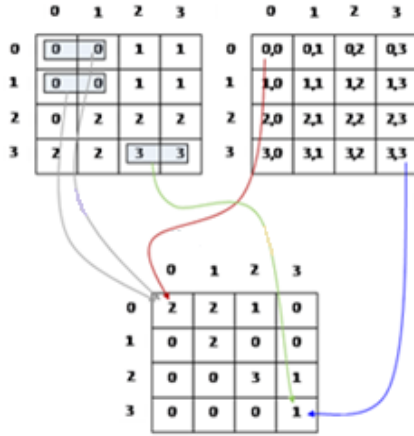


Fig. 8. GLCM matrix angle 0° distance 1 pixel [16]

Haralick et al. [19] formulated 14 texture features that can be extracted from GLCM matrix. The notations that will be used are:

- $p(i, j)$ is a $GLCM_{\bar{r}}(i, j)$ matrix which has been normalized by the size of the row i and column j .
- $p_x(i)$ a new matrix is the sum of all elements in the matrix row $p(i, j)$ or can be written as

$$p_x(i) = \sum_{j=1}^{N_g} p(i, j) \quad (10)$$

where N_g equal with L that is the number of quantization gray levels, that is 256.

$$p_y(j) = \sum_{i=1}^{N_g} p(i, j) \quad (11)$$

$$p_{x+y}(k) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) \quad (12)$$

where $k = i + j$; $k = 2, 3, 4, \dots, 2N_g$

$$p_{x-y}(k) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) \quad (13)$$

where $k = |i - j|$; $k = 0, 1, 2, \dots, N_g - 1$

Textural features of Haralick et al. [19] are defined as follows.

1) Angular Second Moment (ASM)

$$f_1 = \sum_{i=0}^{L-1} i \cdot p(i) \quad (14)$$

ASM is often also referred to as 2nd order energy or uniformity.

2) Contrast

$$f_2 = \sum_{n=0}^{N_g-1} n^2 \left\{ \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) \right\} \quad (15)$$

3) Correlation

$$f_3 = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (16)$$

where μ_x , μ_y , σ_x and σ_y area mean and deviation standard from p_x and p_y matrix respectively.

4) Sum of Squares : Variance

$$f_4 = \sum_i \sum_j (i - \mu)^2 p(i, j) \quad (17)$$

where μ is means of normalized GLCM $p(i, j)$ matrices.

5) Invers Different Moment (IDM)

$$f_5 = \sum_i \sum_j \frac{1}{1 + (i - j)^2} p(i, j) \quad (18)$$

IDM also referred to as *Homogeneity*.

6) Sum Average

$$f_6 = \sum_{i=2}^{2N_g} i p_{x+y}(i) \quad (19)$$

7) Sum Entropy

$$f_7 = - \sum_{i=2}^{2N_g} p_{x+y}(i) \log_2 \{p_{x+y}(i)\} \quad (20)$$

Just in case probability may be zero, and the log (0) is not defined, then the value of “log” was added with the arbitrary constant small value ($p + e$) where e constant is small positive in the calculation of entropy.

8) Sum Variance

$$f_8 = \sum_{i=2}^{2N_g} (i - f_7)^2 p_{x+y}(i) \quad (21)$$

9) Entropy

$$f_9 = - \sum_i \sum_j p(i, j) \log_2 p(i, j) \quad (22)$$

10) Difference Variance

$$f_{10} = \text{Variance of } p_{x-y} \quad (23)$$

11) Difference Entropy

$$f_{11} = - \sum_{i=0}^{N_g-1} p_{x-y}(i) \log_2 \{p_{x-y}(i)\} \quad (24)$$

12) Information Measures of Correlation 1

$$f_{12} = \frac{HXY - HXY1}{\max\{HX, HY\}} \quad (25)$$

where

$$HXY = f_9 \quad (26)$$

$$HXY1 = - \sum_i \sum_j p(i, j) \log_2 \{p_x(i) p_y(j)\} \quad (27)$$

$$HX = p_x \quad (28)$$

$$HY = p_y \quad (29)$$

13) Information Measures of Correlation 2

$$f_{13} = (1 - \exp[-2(HXY2 - HXY)])^{1/2} \quad (30)$$

where

$$HXY2 = - \sum_i \sum_j p_x(i) p_y(j) \log_2 \{p_x(i) p_y(j)\} \quad (31)$$

14) Maximal Correlation Coefficient

$$f_{14} = (\text{second largest eigenvalue of } Q)^{1/2} \quad (32)$$

where

$$Q = \sum_k \frac{p(i,k)p(j,k)}{p_x(i)p_y(k)} \quad (33)$$

Soh and Tsatsoulis [22] used ten features based on GLCM matrix to analyze radar image texture. Of the ten values that are used, there are five new features that have not been employed by Haralick et al. [19] that are:

1) Autocorrelation

$$f_{15} = \sum_i \sum_j (ij) \cdot p(i, j) \quad (34)$$

2) Dissimilarity

$$f_{16} = \sum_i \sum_j |i - j| \cdot p(i, j) \quad (35)$$

3) Cluster Shade

$$f_{17} = \sum_i \sum_j (i + j - \mu_x - \mu_y)^3 p(i, j) \quad (36)$$

where

$$\mu_x = \sum_i \sum_j i \cdot p(i, j) \quad (37)$$

$$\mu_y = \sum_i \sum_j j \cdot p(i, j) \quad (38)$$

4) Cluster Prominance

$$f_{18} = \sum_i \sum_j (i + j - \mu_x - \mu_y)^4 p(i, j) \quad (39)$$

5) Maximum Probability

$$f_{19} = \text{MAX } p(i, j) \quad (40)$$

f_{19} is the greatest value entry element of $p(i, j)$ matrix.

Additional features of GLCM are also derived from Clausi [12].

1) Inverse Difference

$$f_{20} = \sum_i \sum_j \frac{p(i, j)}{1 + |i - j|} \quad (41)$$

2) Inverse Difference Normalized

$$f_{21} = \sum_i \sum_j \frac{p(i, j)}{1 + |i - j| / N_g} \quad (42)$$

3) Inverse Difference Moment Normalized

$$f_{22} = \sum_i \sum_j \frac{p(i, j)}{1 + |i - j|^2 / (N_g)^2} \quad (43)$$

Overall features based on GLCM matrix are 22 values.

Gray Level Run Length Matrices (GLRLM) is a method to represent the texture published by Galloway [16]. Gray Level is a level value of image intensity at successive either vertically, horizontally or diagonally. The Run Length is the number of pixels occupied by the value of the intensity level.

		0	1	2	3
		0	2	3	3
Picture:		2	1	1	1
		3	0	3	0

Run length:					
G	0°	1	2	3	4
r					
a	0	4	0	0	0
y					
l	1	1	0	1	0
e					
v	2	3	0	0	0
e					
l	3	3	1	0	0

45°	1	2	3	4
0	4	0	0	0
1	4	0	0	0
2	0	0	1	0
3	3	1	0	0

Fig. 9. GLRLM matrix [16]

GLRLM features are represented by the notation $p(i, j)$ refer to matrix $M \times N$ consist of row i intensity gray levels and column j of run length in succession in the direction of 0° (horizontal), $45^\circ/135^\circ$ (diagonal) or 90° (vertical). Fig. 9 gives an example of the formation of GLRLM matrix of 4x4 image with gray levels 0 to 3 (4 bit) at an angle 0° and 45° .

Xiaoou Tang [23] defines four new matrix of GLRLM values to simplify the texture which are proposed by Galloway [16].

1) Gray Level Run Length Pixel Number Matrix

$$p_p(i, j) = p(i, j) \cdot j \quad (44)$$

2) Gray Level Run Number Vector

$$p_g(i) = \sum_{j=1}^N p(i, j) \quad (45)$$

3) Run Length Run Number Vector

$$p_r(j) = \sum_{i=1}^M p(i, j) \quad (46)$$

4) Gray Level Run Length One Vector

$$p_0(i) = p(i, 1) \quad (47)$$

Based on GLRLM matrix, many textures that can be extracted. There are five texture features which is proposed by Galloway [16].

1) Short Run Emphasis (SRE)

$$SRE = \frac{1}{n_r} \sum_{i=1}^M \sum_{j=1}^N \frac{p(i, j)}{j^2} \quad (48)$$

Based on matrix which is proposed by Xiaoou Tang [23] above, the value of SRE in Eq. (48) can be simplified into :

$$SRE = \frac{1}{n_r} \sum_{j=1}^N \frac{p_r(j)}{j^2} \quad (49)$$

where n_r is the sum total of all value elements GLRLM or it can be written as

$$n_r = \sum_{i=1}^M \sum_{j=1}^N p(i, j) = \sum_{i=1}^M p_g(i) = \sum_{j=1}^N p_r(j) \quad (50)$$

2) Long Run Emphasis (LRE)

$$LRE = \frac{1}{n_r} \sum_{i=1}^M \sum_{j=1}^N p(i, j) \cdot j^2 = \frac{1}{n_r} \sum_{j=1}^N p_r(j) \cdot j^2 \quad (51)$$

3) Gray Level Nonuniformity (GLN)

$$GLN = \frac{1}{n_r} \sum_{i=1}^M \left(\sum_{j=1}^N p(i, j) \right)^2 = \frac{1}{n_r} \sum_{i=1}^M p_g(i)^2 \quad (52)$$

4) Run Length Nonuniformity (RLN)

$$RLN = \frac{1}{n_r} \sum_{j=1}^N \left(\sum_{i=1}^M p(i, j) \right)^2 = \frac{1}{n_r} \sum_{j=1}^N p_r(j)^2 \quad (53)$$

5) Run Percentage (RP)

$$RP = \frac{n_r}{n_p} \quad (54)$$

where n_p is the size of the value matrix GLRLM resolution or number of pixels contained in the matrix GLRLM. The value of n_p can be written as

$$n_p = M \times N \quad (55)$$

Chu *et al* [24] added 2 texture features that can be derived from GLRLM matrix as a follow.

1) Low Gray-level Run Emphasis (LGRE)

$$LGRE = \frac{1}{n_r} \sum_{i=1}^M \sum_{j=1}^N \frac{p(i, j)}{i^2} = \frac{1}{n_r} \sum_{i=1}^M \frac{p_g(i)}{i^2} \quad (56)$$

2) High Gray-level Run Emphasis (HGRE)

$$HGRE = \frac{1}{n_r} \sum_{i=1}^M p_g(i) \cdot i^2 \quad (57)$$

Dasarathy and Holder [25] also added four other texture features that can be derived from the value of GLRLM namely:

1) Short Run Low Gray-level Emphasis (SRLGE)

$$SRLGE = \frac{1}{n_r} \sum_{i=1}^M \sum_{j=1}^N \frac{p(i, j)}{i^2 j^2} \quad (58)$$

2) Short Run High Gray-level Emphasis (SRHGE)

$$SRHGE = \frac{1}{n_r} \sum_{i=1}^M \sum_{j=1}^N \frac{p(i, j) \cdot i^2}{j^2} \quad (59)$$

3) Long Run Low Gray-level Emphasis (LRLGE)

$$LRLGE = \frac{1}{n_r} \sum_{i=1}^M \sum_{j=1}^N \frac{p(i, j) \cdot j^2}{i^2} \quad (60)$$

4) Long Run High Gray-level Emphasis (LRHGE)

$$LRHGE = \frac{1}{n_r} \sum_{i=1}^M \sum_{j=1}^N p(i, j) \cdot i^2 \cdot j^2 \quad (61)$$

All of GLRM textural features are 11 values.

2.4. Multi Layer Perceptron

Multilayer Perceptron (MLP) is an arrangement of a neural network with programming paradigm that is inspired by the performance of the microstructure of the human brain [26]. Artificial neural networks are used extensively in the disciplines of artificial intelligence for pattern recognition and classification [27]. MLP architecture can be seen in Fig. 10. This network consists of a collection of neurons (nodes) in a hidden layer and on any ramifications contained weight value

will always change as the process of learning. Input layer contains all textural features of a statistical histogram, GLCM, and GLRLM.

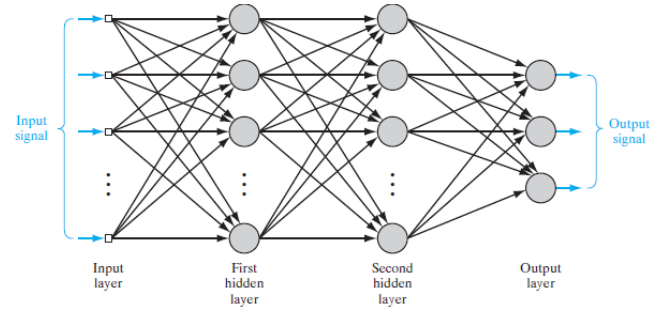


Fig. 10. Architecture of MLP [28]

The amount of hidden layer and neuron can be adapted depend on the complexity of textural classification problems. Each value entered in neurons will produce output values by a sigmoid activation function as shown in Eq. (62).

$$y_j = \frac{1}{1 + e^{-ay_{jnet}}} \quad \text{where} \quad 0 < a < 1 \quad (62)$$

Good knowledge is generated by new weights at each branch network at the end of the learning process. Back Error Propagation (BEP) is a supervised learning algorithm used in architecture MLP [26]. BEP allows the adjustment of the price weights hidden neurons by means of creeping and turning error output. There are two stages in the learning process of the BEP is the direction forward and backward direction.

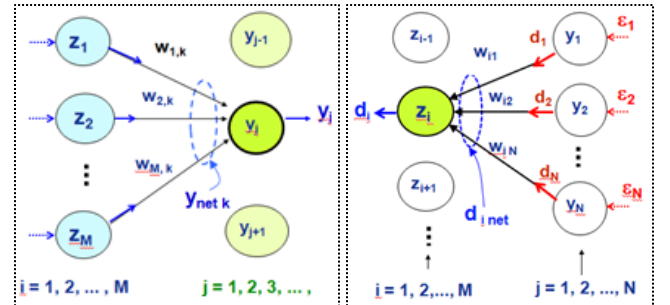


Fig. 11. Forward and backward BEP

In the forward direction, the signal propagated from the input layer to the hidden layer 1, the hidden layer 2, to the output layer. Then calculated the output of each neuron in each layer with activation function. In the backward direction, the error factor propagated from the output layer to the hidden layer two hidden layers 1, to the input layer. Then the weight of the branching connection in each layer is calculated and updated. The process is carried out continuously every copies and epoch. A copy is one training process involving of textural features, while one epoch is one training process involving all data. Illustration of the learning process in the forward and backward direction can be seen in Fig.11.

For each neuron in a layer (eg, y_j) is calculated sum of the input value with the following formula:

$$y_{jnet} = \sum_{i=1}^M w_{ij} z_i \quad (63)$$

As for the output neurons in a layer is calculated activation function in Eq. (62), if $a = 1$, the first derivative of this is:

$$y_j' = y_j(1 - y_j) \quad (64)$$

The value of y_j and y_j' are stored for calculation in backward direction. This process is done for the entire layer in the forward direction. The weight w_i at every branch is updated on the backward direction following this equation:

$$\Delta w_i = \eta \sum_{j=1}^M (t_i - y_{jk})(y_{jk}(1 - y_{jk}))z_i \quad (65)$$

or

$$\Delta w_i = \eta \sum_{j=1}^M (t_i - y_{jk})y_j' z_i \quad (66)$$

$$w_{i \text{ new}} = w_{i \text{ previous}} + \Delta w_i \quad (67)$$

where η is learning rate with the value between 0 to 1. And t_i is training target which is used for learning reference of the system.

Each neuron (eg z_i) in backward direction layer are applied in calculation of error factor d . Then neurons in the output layer are applied:

$$d_j = (t_j - y_j)y_j' \quad (68)$$

$$d_j = \varepsilon_j y_j' \quad (69)$$

to neurons in other layers then :

$$d_{i \text{ net}} = \sum_{j=1}^N w_{ij} d_j \quad (70)$$

and

$$d_i = d_{i \text{ net}} y_j' \quad (71)$$

Adjustment of weight factor w_{ij} is done by following Eq. (68) to Eq. (71). According to Fig. 10, every weight that is interconnected with neuron z_i on backward direction is calculated using the following formula:

$$\Delta w_{ij} = \eta d_j z_i \quad (72)$$

where z_i is the result of the previous calculation of the forward direction. The value d_i is then propagated return for the next layer to the input layer of the MLP network. The new weights were adjusted by Eq. (67). Its process continues to next input copies until all epoch is completed. The new weights will appear at every training then at the end of the process resulting small error. Thus the MLP network is ready for use in testing process with thyroid nodule dataset.

2.5. Performance Indices

Validation methods are required to perform quality of features and classification. Performance indexes are determined by the following formulas:

1) Accuracy

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \times 100\% \quad (73)$$

where TP (True Positive) is the number of positive data on the target that is classified positive on the system. TN (True Negative) is the number of negative data on the target that is classified negatively on the system. FP (False Positive) representation of the amount of negative data on the target that is classified positive systems and FN (False

Negative) is the amount of positive data on the target system are classified negatively.

2) Sensitivity

$$\text{Sensitivity} = \frac{TP}{TP + FN} \times 100\% \quad (74)$$

3) Specificity

$$\text{Specificity} = \frac{TN}{TN + FP} \times 100\% \quad (75)$$

4) Positive Predictive Value (PPV) atau Precision

$$\text{Precision} = \frac{TP}{TP + FP} \times 100\% \quad (76)$$

5) Negative Predictive Value (NPV)

$$\text{NPV} = \frac{TN}{TN + FN} \times 100\% \quad (77)$$

The measurement index above is grouped in confusion matrix table. It shows the relationship between the actual class or target with predicted class. Actual class is mean radiologist diagnosis and clinical pathology report from the hospital. Predicted class is output classification from CAD. By looking at the value of Accuracy, Sensitivity, Specificity, Precision and Negative Predictive Value (NPV), the performance of the classification on the classification of thyroid nodules can be obtained. The arrangement of confusion matrix can be seen in Table. 1.

Table 1. Performance index of confusion matrix

		Target Class		
		Kistik	Solid	
Predicted Class	Kistik	True Positives	False Positives	Sensitivity
	Solid	False Negatives	True Negatives	Specificity
		Precision	NPV	Accuracy

3. Results and Discussion

3.1. CAD System

All of the steps in Fig.1 are compiled in the executable graphical user interface (CAD) developed on MATLAB. Its application supported radiologist to identify thyroid nodules composition objectively. CAD was built to decrease dependency of the ultrasound machine and miss-interpretation diagnosis among radiologist as shown in Fig. 12.

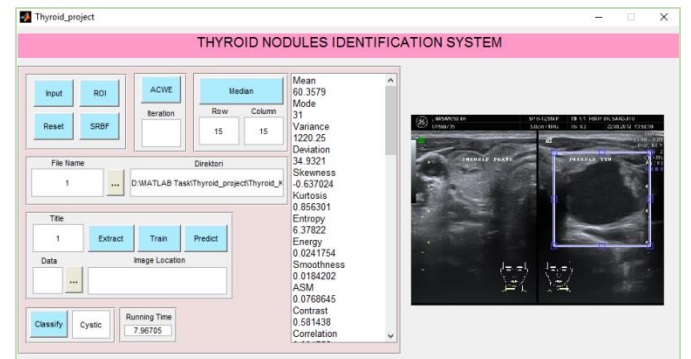


Fig. 12. CAD of thyroid nodule classification based on textural features

3.2. Median Filtering

Pre-processing stage in this study is a screening process to remove the marking area on the ultrasound image with the median filter algorithm. Marking area makes the results of feature extraction less precise. Sample pre-processing results from the ROI image is displayed in Fig. 13.

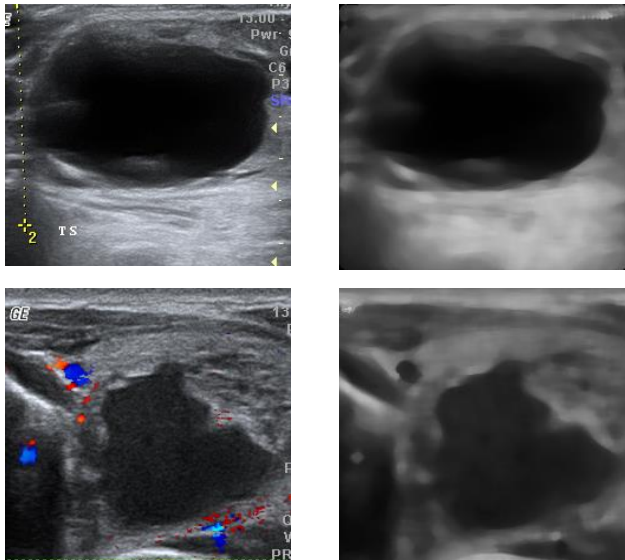


Fig. 13. Original (left) & filtered images (right)

3.3. Classification and Evaluation

The classification process begins with the preparation of MLP classifier architecture for later conducted training and testing process on the system of identification of thyroid nodules. Determination of the number of hidden layer neurons and MLP architecture that will be applied to the system should also be tested with Weka machine learning. Thus thyroid nodule identification system used will be more accurate and reliable in a variety of different data and keep fit at a time whenever used. From the test of neurons number and layers, it can be obtained MLP architecture that consists of one hidden layer which composes of 100 neurons. The correlation of hidden layer and neurons number to accuracy of the classification results is shown in Fig. 14 and Fig. 15.

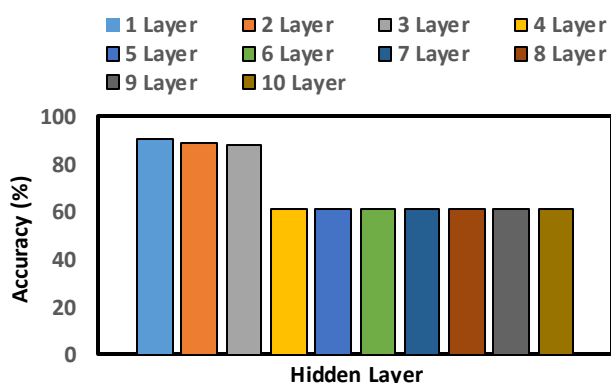


Fig. 14. Correlation number of hidden layer versus accuracy

As in bar chart, Fig. 14 shows that more and more layers are added to the MLP network system, the accuracy of the

test results has narrowed. Accuracy stood at 60% and did not change again begin to layer 4 and so on. Single layer produce the highest accuracy of 90% compared with the number of hidden layer to another. Fig. 15 shows that accuracy is not changed even though the number of neurons increased. Start neurons 40 to 100 unchanged at all. So that the MLP architecture applied in this classification system is a single layer with 100 number of neurons.

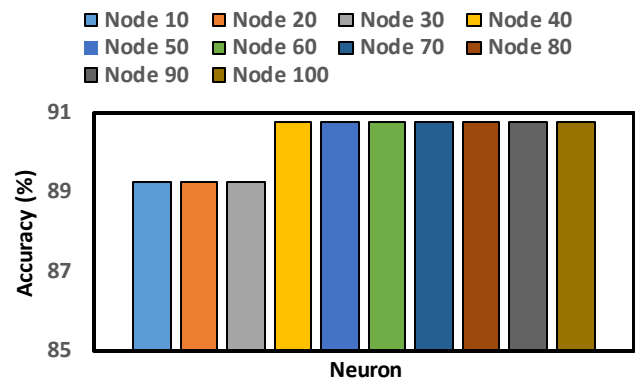


Fig. 15. Correlation number of neuron versus accuracy

From the conducted experiment results, the full combination features of the statistical histogram, GLCM and GLRLM has given best result than using each groups separately. Fig. 16 shows the results of each performance.

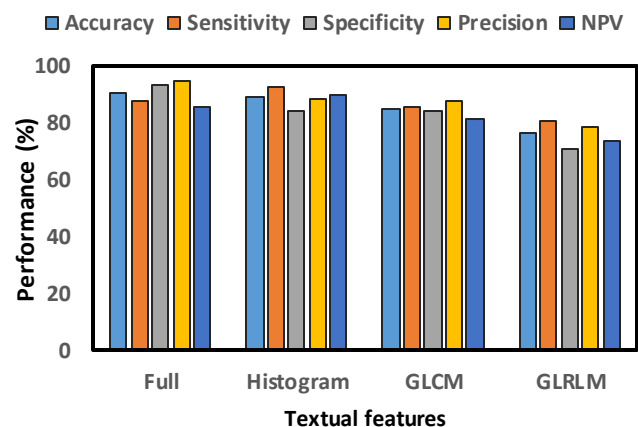


Fig. 16. Comparison of textural features performance

Two confusions matrix are presented to evaluate textural features extraction and classification. The first confusion matrix is used to validate MLP learning outcomes with testing images. All 72 thyroid images were divided into 36 training set and 36 testing set. Its depends on the number of the training set, as much of training set then the CAD become smarter. Table 2 illustrates the result which there are five incorrect classification consists of 2 FP and 3 FN. It means that two cysts are predicted as solids and three solids as cystic.

Table 2. The peformance of testing set

		Target Class		
		Cystic	Solid	
Predicted Class	Cystic	17	2	89.5 %
	Solid	3	14	82.4 %
		85.0 %	87.5 %	86.1 %

The second confusion describes the performance of 10-fold cross validation. Each of data has possibility to become the training and testing set. The number of fold determine how much sample data and how many time training. As shown in Table 3, cross-validation presented better accuracy than training-testing method. The number of misclassifying data in cross-validation was higher than training-testing method, i.e., 2 FP, and 5 FN. Sensitivity also decreased from 89.5% to 87.80%. The better performance could be obtained by some thyroid ultrasound images addition.

Table 3. 10-fold cross validation

		Target Class		
		Cystic	Solid	
Predicted Class	Cystic	36	2	87.80 %
	Solid	5	29	93.55 %
		94.74 %	85.29 %	90.28 %

4. Conclusion

The goal of CAD thyroid nodule based on textural features to classify cystic and solid category has been going well. Textural patterns of thyroid nodules on ultrasound image are the main visual characteristic for the radiologist in identification. The calculation of statistical histogram and measurable values of GLCM also GLRLM can describe the textural pattern of cystic or solid nodules accurately. It increases the objectivity of doctors in diagnosing thyroid nodules and reduces dependence on the ultrasound operator. In future work, it is suggested to use segmentation technique for localizing nodule area specifically. The work of MLP is influenced by some the dataset, as much of it then we get a robust network. Utilities another classifier are also needed to compare the quality of features.

References

- 1 I. Legakis, M. A. Savelonas, D. Maroulis, and D. K. Iakovidis, *Computer-Based Nodule Malignancy Risk Assessment in Thyroid Ultrasound Images*, Int. J. Comput. Appl. (2011).
- 2 G. Braunstein, *Thyroid Cancer*, (2012).
- 3 H. J. Moon, J. Y. Kwak, E.-K. Kim, M. J. Kim, C. S. Park, W. Y. Chung, and E. J. Son, *The Combined Role of Ultrasound and Frozen Section in Surgical Management of Thyroid Nodules Read as Suspicious for Papillary Thyroid Carcinoma on Fine Needle Aspiration Biopsy: A Retrospective Study*, World J. Surg. 33 (2009) 950–957.
- 4 H. G. Moon, E. J. Jung, S. T. Park, W. S. Ha, S. K. Choi, S. C. Hong, Y. J. Lee, Y. T. Joo, C. Y. Jeong, D. S. Choi, and J. W. Ryoo, *Role of ultrasonography in predicting malignancy in patients with thyroid nodules*, World J. Surg. 31 (2007) 1410–1416.
- 5 M. C. Frates, C. B. Benson, J. W. Charboneau, E. S. Cibas, O. H. Clark, B. G. Coleman, J. J. Cronan, P. M. Doubilet, D. B. Evans, J. R. Goellner, I. D. Hay, B. S. Hertzberg, C. M. Intenzo, R. B. Jeffrey, J. E. Langer, P. R. Larsen, S. J. Mandel, W. D. Middleton, C. C. Reading, S. I. Sherman, and F. N. Tessler, *Management of thyroid nodules detected at US: Society of Radiologists in Ultrasound consensus conference statement*, Ultrasound Q. 22 (2006) 231–240.
- 6 T.-C. Chang, *The Role of Computer-aided Detection and Diagnosis System in the Differential Diagnosis of Thyroid Lesions in Ultrasonography*, J. Med. Ultrasound. 23 (2015) 177–184.
- 7 M. Savelonas, D. Maroulis, and M. Sangriotis, *A computer-aided system for malignancy risk assessment of nodules in thyroid US images based on boundary features*, Comput. Methods Programs Biomed. 96 (2009) 25–32.
- 8 W. J. Moon, S. L. Jung, J. H. Lee, D. G. Na, J.-H. Baek, Y. H. Lee, J. Kim, H. S. Kim, J. S. Byun, and D. H. Lee, *Benign and malignant thyroid nodules: US differentiation—multicenter retrospective study*, Radiology. 247 (2008) 762–770.
- 9 A. Gürsoy and M. F. Erdoğan, *Ultrasonographic Approach to Thyroid Nodules: State of Art*, Thyroid Int. (2012).
- 10 C. Chang, M. Tsai, and S. Chen, *Classification of the Thyroid Nodules Using Support Vector Machine*, (2008) 3093–3098.
- 11 M. P. Wachowiak, a. S. Elmaghraby, R. Smolikova, and J. M. Zurada, *Classification and estimation of ultrasound speckle noise with neural networks*, Proc. IEEE Int. Symp. Bio-Informatics Biomed. Eng. (2000).
- 12 D. a. Clausi, *An analysis of co-occurrence texture statistics as a function of grey level quantization*, Can. J. Remote Sens. 28 (2002) 45–62.
- 13 H. Wibawanto and A. Susanto, *Discriminating Cystic and Non Cystic Mass using GLCM and GLRLM-based Texture Features*, Int. J. Electronic Engineering Research. 2 (2010) 569–580.
- 14 R. M. Haralick, K. Shanmugam, and I. Dinstein, *Textural Features for Image Classification*, IEEE Trans. Syst. Man. Cybern. 3 (1973).
- 15 R. M. Haralick, *Statistical and structural approaches to texture*, Proc. IEEE. 67 (1979) 786–804.
- 16 M. M. Galloway, *Texture analysis using gray level run lengths*, Comput. Graph. Image Process. 4 (1975) 172–179.
- 17 A. K. Jain, *Fundamentals of digital image processing*, 46 (1989).
- 18 A. Kadir and A. Susanto, *Image Processing Theory and Application 1 (in Indonesian)*, (2012).
- 19 R. M. Haralick, *Statistical and structural approach to texture*, Proceeding IEEE. 67 (1979) 786–804.
- 20 S. D. Newsam and C. Kamath, *Comparing Shape and Texture Features for Pattern Recognition in Simulation Data*, (2005) 106–117.
- 21 A. Kadir and A. Susanto, *Image Processing Theory and Application 2 (in Indonesian)*, (2012).
- 22 U. Gray, L. C. Matrices, L. Soh, C. Tsatsoulis, and S. Member, *Texture Analysis of SAR Sea Ice Imagery*, 37 (1999) 780–795.
- 23 X. Tang, *Texture Information in Run-Length Matrices*, 7 (1998) 1602–1609.
- 24 A. Chu, C. M. Sehgal, and J. F. Greenleaf, *Use of gray value distribution of run lengths for texture analysis*, Pattern Recognit. Lett. 11 (1990) 415–419.
- 25 B. V. Dasarathy and E. B. Holder, *Image characterizations based on joint gray level—run length distributions*, Pattern Recognit. Lett. 12 (1991) 497–502.
- 26 H. Demuth, *Neural Network Toolbox*, (2000).
- 27 R. Duda, P. Hart, and D. Stork, *Pattern Classification*, (1973).
- 28 S. Haykin, *Neural Networks and Learning Machines*, 3 (2008).