Medical Texture Recognition Based on Intelligent Technique

Dr. Hasanen S. Abduallah
Computer Science Department, University of Technology/ Baghdad
E-mail: qhasanen@yahoo.com
Hajer Mazin Salih
Computer Science Department, University of Technology/ Baghdad

Received on: 22/3/2015 & Accepted on:12/11/2015

ABSTRACT

The aim of this paper is to improve a disease diagnosing system through classifying the skin diseases depending on the skin images of many patients by using the multiwavelet transformation. First, extract features from multiwavelet coefficients. Second, the skin images are classified as Warts, Vitiligo, Hemangioma, and Normal depending on the decision rules generated by the decision tree using the ID3 learning algorithm.
Keywords: Texture Recognition, Multiwavelet Transformation, Decision Tree, Skin Diseases Recognition, ID3.

تمييز الانسجة الطبية بالاعتماد على تقنيات ذكية

الخلاصة

الهدف من هذا البحث هو تحسين نظام تشخيص المرض من خلال تصنيف الأمراض الجلدية اعتمادا على صور للجلد تابع للعديد من المرضى باستخدام التحويل متعدد المويجات. أولا، يتم استخلاص ميزات الصورة من خلال معاملات التحويل متعدد المويجات. ثانيا، يتم تصنيف صور الجلد أما ان تكون مصابة بالثؤلول، البهاق، الوحمة الدموية (الورم الوعائي)، أو ان تكون بشرة طبيعية اعتمادا على قواعد اتخاذ القرارات التي تولدها شجرة القرار باستخدام خوارزمية التعلم (Iterative Dichotomiser3 "ID3").

INTRODUCTION

In the past twenty years, a huge improvement has been noticed in imaging techniques targeted for biomedical applications. Significant progress was also recorded in dermatology domain [1]. One of the significant features of images that has been exploited in medical image analysis and classification is texture. One of the main stages in collecting these features through texture analysis process is extraction of texture features [2]. Image analysis technique has played a significant role in many medical applications. Generally, the applications include the automatic extraction of features from an image that are used then for a diversity of classification tasks, such as distinguishing abnormal tissues from normal ones. Based on specific classification task, the extracted features hold certain textural properties, or color properties of an image. The textural features extracted are associated with the application domain to be used [3]. Wavelet theory proposed by Mallat has become a mathematical framework. Basically, the wavelet transformation job is mapping an image into a low resolution image and a series of detail images. The entropy or energy of the detail images are the commonly used features for texture segmentation and classification problems.

Multiwavelet, which is an extension from scalar wavelet, has received considerable attention from the wavelet research communities both in theory as well as in applications such as denoising, signal compression, and image processing applications [4].

https://doi.org/10.30684/etj.34.1B.14
2412-0758/University of Technology-Iraq, Baghdad, Iraq
This is an open access article under the CC BY 4.0 license http://creativecommons.org/licenses/by/4.0

Theoretical Background

Image classification is one of classical problems concerning the image processing. The aim of image classification is to predict the types of the entered image using its extracted features [5].

A decision tree is one of the main tools in solving the task of classification. These tools have been successfully used in many fields like remote sensing, medical diagnosing, radar signal classification, character recognition, speech recognition, expert systems and others [6].

One of the most significant steps in classification is transforming. The main limit in transforming the two-dimension wavelet is not getting information from different directions. As known, the wavelet transform can only extract information in vertical, horizontal and diametrical direction. To get rid of this problem, researchers use Multi-wavelet which is a modern concept added to wavelet transform, offers symmetry, short support, and simultaneous orthogonality, which are not possible with scalar wavelet systems [3].

Transformations take important part in various signal processing applications like filtering, pattern recognition, restoration, spectrum estimation, signal enhancement, localization and compression [7].

Wavelet has an associated scaling function φ(t) and wavelet function ψ(t), Multiwavelets, have two or more wavelet and scaling functions. The set of scaling functions can be written by using the vector notation $\varphi(t) \equiv [\varphi_1(t), \varphi_2(t) \ldots \varphi_r(t)]^T$, where φ(t) is called the multiscaling function. Likewise, the Multiwavelet functions are defined from the set of wavelet functions as $\psi(t) \equiv [\psi_1(t), \psi_2(t) \ldots \psi_r(t)]^T$. When (r=1), ψ(t) is called a scalar wavelet, or just wavelet, while theoretically (r) can be arbitrary large number.

The Multiwavelets have two scale (r=2) equations similar to those for scalar wavelets as [8]

$$\psi(t) = \sqrt{2}\sum G_k \psi(2t-k) \quad \ldots (1)$$

$$\emptyset(t) = \sqrt{2}\sum H_k \emptyset(2t-k) \quad \ldots (2)$$

Where

$[G_k]$ and $[H_k]$ are matrix filters, i.e., G_k and H_k are (r × r) matrices, $\psi(t), \emptyset(t)$ are wavelet function and scaling function respectively.

The matrix elements in these filters give more degree of freedom than the scalar wavelet.

The Proposed Intelligent Skin Diagnosis System

ISD system is used to classify different skin diseases depending on dataset of skin images. The ISDS is first trained by using the training dataset in order to generate production rules, which give the class of the disease. Once the system is trained, a new image from testing dataset is entered to the system and can be classified directly by the production rules generated by the training phase.

The training phase is significant to construct the classifier model that produces the production rules. Training phase has many stages each stage has a different function.

The testing phase or prediction phase uses images that are different from the images used in training phase. The testing phase has the same stages in training phase except the classifier model stage. The new entered images are classified directly based on the production rules generated by the training phase, as shown in figure (1).

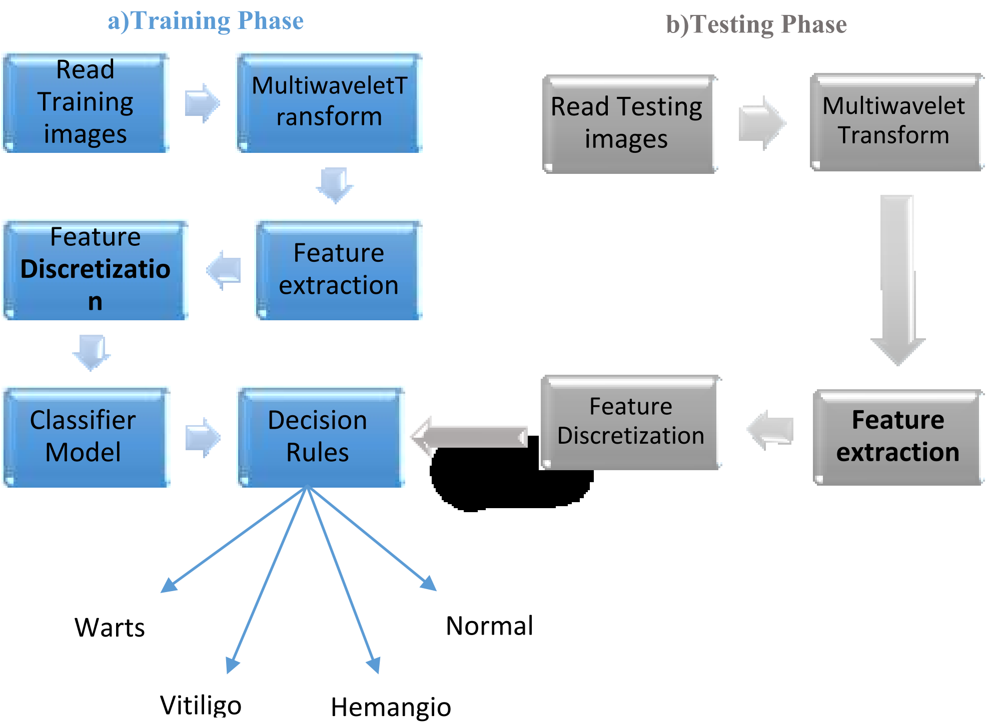


Figure (1): The Block diagram of ISDS. a) Illustrates the block diagram of the training phase of the ISD system. b) Illustrates the block diagram of the testing phase of ISD system.

The Dataset

The dataset used consists of images for different skin diseases such as (Vitiligo, Wart, Hemangioma, and Healthy skin). These images are collected form random medical websites. The size of these images is (256*256) and their format is (JPEG). These images are colored images but first they are converted to gray-scale images which have only intensity information. In gray-scale images, all color bands (R, G, B) have equal values in order to produce Gray-Level Co-occurrence Matrix (GLCM) which used to extract texture feature from the skin images. Sample images of skin images are shown in the Figure (2).

(a) (b) (c) (d)

Figure (2): Samples of skin images a) Warts. b) Hemangioma. c) Vitiligo. d) Normal skin.

The Multiwavelet Transformation

Image transformation regarded as a preprocessing step before extracting features. It is an important step in this work because it reduces the noise in images. The Multiwavelet transforms the image from spatial domain to frequency domain, the resultant image after one level of multiwavelet decomposition the image is divided into two low pass sub-bands and two high pass sub-bands. The high pass sub-bands contain the high frequencies which are mostly noise and the low pass sub-bands contain the detail frequencies which are important in many applications such as object segmentation and texture recognition. The resultant image is consists of 16 sub-images.

Figure (3) shows the steps of the Multiwavelet transformation:

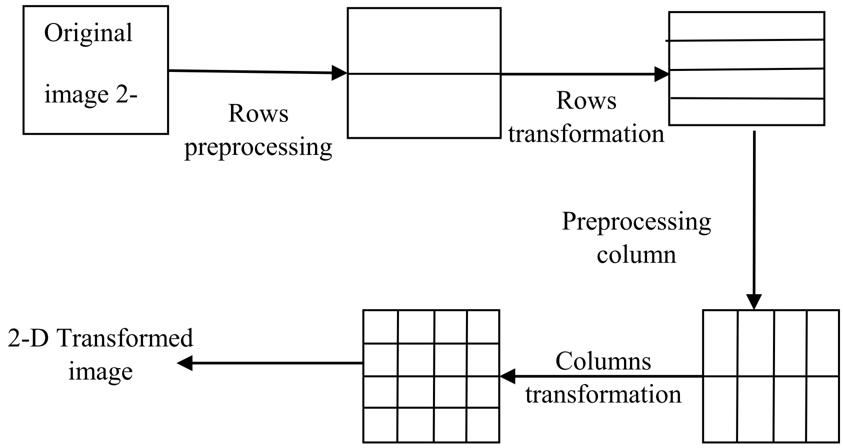
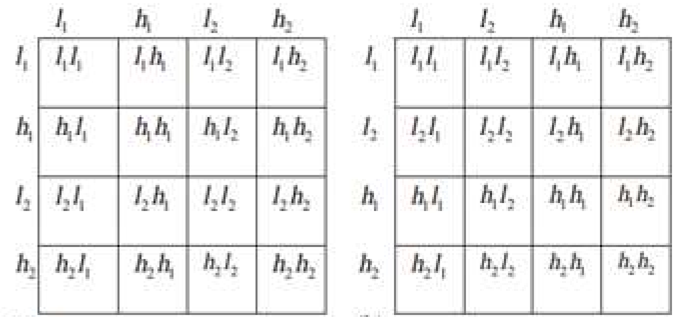


Figure (3): One level of 2-D Multiwavelet decomposition of an image.

Figure (4-a) illustrates the final result of one level Multiwavelet decomposition. Rearrangement of the sub-images will lead to Figure (4-b).



Feature Extraction

Feature Extraction is a significant step for image classification, it is used to discriminates the images and improve the efficiency of image classification. Texture is an important feature used in image processing to recognize the surface of a given object. To extract the texture from the skin image first the symmetrical Gray-Level Co-occurrence Matrix (GLCM) will be computed for each entered skin image. The extracted features from the skin image are Energy, Contrast, Homogeneity, Mean, and Variance.

Feature extracted from symmetric normalized GLCM can be clarified below:

Energy: is also known as Angular Second Moment (ASM), the equation to calculate the Energy is [9]:

$$\text{Energy}=-\sum_{i=0}^{n-1}\sum_{j=0}^{n-1}\left(\frac{P\,(i,j)}{R}\right)^2 \quad \ldots(3)$$

Where *P(i,j)*: the elements of GLCM matrix, R: the total number of pixels.

Contrast: is a measure of local gray level variation of an image. The equation used to find the contrast is [9]:

$$\text{Contrast}=-\sum_{i=0}^{n-1}\sum_{j=0}^{n-1}(i-j)\left(\frac{P(i,j)}{R}\right) \quad \ldots(٤)$$

Where *P(i,j)* and *R* are defined previously.

Homogeneity: also known as Inverse difference moment IDM, the equation [10]:

$$\text{Homogeneity}=\sum_{i,j=0}^{n-1}\frac{P(i,j)}{1+|i-j|} \quad ..(٥)$$

Where *P(i,j)* is defined previously.

Entropy: is a measure, which is inversely related to energy. It measures the randomness of the image [9]. It reaches its largest value when all elements in P matrix are the same.

$$\text{Entropy}=-\sum_{i=0}^{n-1}\sum_{j=0}^{n-1}\left(\frac{P(i,j)}{R}\right)\log\left(\frac{P(i,j)}{R}\right) \quad \ldots(6)$$

Where *P(i,j)* and *R* are defined previously.

Mean: is measured by the following equation [11]:

$$\text{Mean}=\frac{1}{x*y}\sum_{\mathrm{i}=0}^{\mathrm{x}-1}\sum_{j=0}^{y-1}X(i,j) \quad \ldots(7)$$

Where *x,y* are the dimensions of an image and *X(i,j)*: gray level value.

Variance: This measure puts quite high weights on the elements that vary from the average value of the P (i, j) as in equation [9]:

$$\text{Variance}=\sum_{i=0}^{x-1}\sum_{j=0}^{y-1}(X(i,j)-Mean) \quad \ldots(8)$$

Where *x,y* , *Mean,* and *X(i,j)* are previously defined.

Standard Deviation: it is the square root of variance. It tells something about the contrast. It describes the diffusion in the data, so a low contrast image will have a low variance and a high contrast image will have a high variance, as in equation [9]:

$$\text{Standard Deviation}=\sqrt{\text{Variance}} \quad \ldots(9)$$

The LL sub-band contains the detail information, which is used to extract texture features. Each sub-band has four sub-images with different brightness. The best sub-image l_2l_1 is used in feature extraction stage because it has a moderate brightness value and showed better results from the other sub-images. The l_2l_1 is considered as an input for the feature extraction stage, and then the GLCM is built for each entered sub-image to extract texture features from it.

Feature Discretization

Feature discretization is an essential step for the decision tree-building stage. It transforms the continuous values of features extracted from the previous step to discrete values in order to prepare them for decision tree, because decision tree does not deal with the continuous values. Feature discretization step aims to give symbols for the features which were extracted previously. These symbols are used by the decision tree. This process requires the choosing of an integer number. Dividing the intervals of the continuous values of the feature is done depending on the chosen number. In this work, number 3 has proved to be the best number to achieve the desired results depending on the trail-and-error method.

Image Classification

After the skin images have been transformed by multiwavelet transformation and the texture features have been extracted from these transformed images as described previously, now the classifier model will be constructed. Classifier model construction is the most significant step in classification process; this model is used to predict the disease of the new entered skin image.

Classifier Model

The skin images dataset (Training Set) contains attribute values represented with five nominal attributes (they are Energy, Entropy, Contrast, Homogeneity, and Standard Deviation), and disease class for many cases. These attributes are entered to the classifier model for learning phase. The prediction of new cases depends on the classifier built in learning phase, all the skin images used in prediction phase are new skin images that are not exist in the learning phase. The prediction phase used to test the skin images (Testing Set). The classifier model is developed in this work based on decision tree depending on the training skin images. The classifier model in the prediction phase gives the decision for the skin images which is mean that the classifier model classified the skin images as Vitiligo, Warts, Hemangioma, or Healthy.

Decision Tree

A decision tree is an illustration of a decision process for defining the class of a particular instance. Every node of the tree identifies either a class name or a specific test that divides the space of instances at the node according to the possible results of the test.

A decision tree is a powerful method for data classification. It is used to make a good decision. The decision tree is used in various disciplines including medical diagnosis, artificial intelligence, game theory, and data mining.

Decision tree can be implemented by using the Iterative Dichotomiser 3 (ID3) learning algorithm which is supervised learning. ID3 algorithm can produce the production rules based on the generated decision tree that depends on the attributes. The decision tree is constructed depending on entropy and information gain.

ID3 Decision Tree Induction Algorithm

ID3 algorithm, which is a supervised learning method, has the ability of producing rules through a decision tree. Information gain is computed for every attribute to build the decision tree then choose the attribute that has the maximum information-gain in order to be considered as a root node. The root node labels attribute, the arcs represents the possible values of the attribute. To see possible instances whether they are lying under the same class or not they must be tested. If all the instances are lying under the same class, the node is denoted with a class name, otherwise choose the splitting attribute in order to classify the instances [6].

Entropy: Entropy measures the goodness of a split. It estimates the amount of information in the attribute [7].

$$\text{Entropy}(S) = \sum_{I=1}(-P(I)\log_2 P(I)) \quad \ldots (10)$$

Where P(I) is the probability of the attribute (S).

Information Gain: The entropy is measures of the randomness in a group of training dataset; information gain measures the effectiveness of the attribute in classifying the training dataset. It is a predictable reduction in the entropy measure generated by splitting the dataset based on this attribute. More precisely, the information gain, Gain(S, A) of an attribute A, relative to a collection of examples S, is illustrated as [7]:

$$\text{Gain}(S, A) = \text{Entropy}(S) - \sum\left(\left(\frac{|SV|}{|S|}\right) \times Entropy(SV)\right) \quad \ldots (11)$$

Where: $\sum$ is summation of all the possible values of the attribute A,

SV= subset of S for which attribute A has value V,

|SV|=number of element in SV,

|S|= number of element in S.

Production Rules

After the decision tree has been built by using the ID3 algorithm depending on the training set of skin images, now the decision rules can be generated from this tree using if-then rules. The decision rules are easier to understand by the computer environment. In the testing phase, the classifier classifies the testing set of the skin images depending on the decision rules produced by the ID3 algorithm. The testing phase has the same steps of the training phase except that in the testing phase the skin images are tested depending on the decision rules generated by the decision tree.

Performance Metrics of ISDS

The performance of ISDS is estimated by using running time and classification accuracy. The running time and classification accuracy of ISDS performance can be shown in Table (1).

The running time is measured in seconds, that is computed by finding the difference between the time of the beginning of the implementation and the time of the end of the implementation of the ISDS .The running time is calculated five times to make sure that result is close enough to the real time it takes, because the CPU might be busy with another process when ISDS is running.

Table (1): Performance Metrics of ISDS

Class Type	Computational-Time	Classification Accuracy
Hemangioma	60 Sec	85%
Wart	70 Sec	90%
Vitiligo	70 Sec	90%
Healthy	80 Sec	100%
Total	75 Sec	91.25%

The equation (12) was used for calculating the accuracy of the classification process. This equation gives the accuracy ratio of the image classification process [8]:

$$\text{Accuracy} = \frac{\text{Number of Correctly Classified Images}}{\text{Total Number of Images}} \times 100\% \qquad \dots (12)$$

The results obtained previously are due to the dataset used in this work, and checked the system according to the images of the testing dataset.

Decision Maker

After the testing phase is over, the imprecisely classified images are entered to this stage to improve the system classification accuracy. The decision maker is an expert system. Expert system is an intelligent system that performs as a human expert in the application field. Expert system is usually composed of two parts: (1) an inference engine and (2) a knowledge base. Knowledge base can be represented in various forms. The most common ones are production (if-then) rules. The knowledge base consists of the information about a specific problem field.

The use of an expert system make the ISDS more reliable and conformance for more instances, and gives the ISDS a powerful point in updating the existent knowledge base.

The performance of the ISDS after using the decision maker technique has been improved and the classification accuracy has been obviously increased as illustrated in Table (2):

Table (2): Performance Metrics of ISDS after DM

Class Type	Classification Accuracy	Improved Classification Accuracy
Hemangioma	85%	95%
Wart	90%	100%
Vitiligo	90%	95%
Healthy	100%	100%
Total	91.25%	97.5%

CONCLUSIONS

In the earlier sections, some remarks associated with the behavior and performance of the suggested skin diagnosis system. A summary of some important conclusions could be the following:

The ISDS helps the dermatologists to diagnose the skin diseases. ISDS is not substitutable for the dermatologists but assists them. Multiwavelet is a frequency transformation that is useful in image compression and reduces the noise because it isolates the high frequencies from the important detail frequencies. The using of ID3 learning algorithm has a significant role in decision tree building process in addition to its efficiency in classifying cases because it deals with large amount of data efficiently. The ID3 learning algorithm is used for classifying the infected skin images, some cases are imprecisely classified to overcome this issue an expert system is used. The expert system supports the ISDS results and increases the classification accuracy up to 97.5%.

REFERENCES

[1] Ogorzałek, L., G. and A., "Modern Techniques for Computer-Aided Melanoma Diagnosis", InTech, 2011.

[2] Anh, "Optimization of Texture Feature Extraction Algorithm" M.Sc. thesis, Faculty of Electrical Engineering, Mathematics and Computer Science, 2010.

[3] Waleed F., “Design of Hierarchical Architecture of Multilevel Discrete Wavelet Transform Using VHDL Language”, Eng. & Tech. Journal, Vol.28, No.7, 2010.

[4] LI, W., L., "Texture Classification Using Discrete Multiwavelet Transform", International Journal of Science and Research.

[5] Shi, Y., T., G., Y., and J., "Geochemical Analyses", 2011.

[6] E. Jabbar, and M. Jabbar, "Classification of Images Using Decision Tree", Eng. & Tech. Journal .Vol31,Part (B), No. 6 , 2013

[7] A. Neymark, "ID3 Algorithm", Spirng, 2007.

[8] B. Patel, K. Rana, "A survey on Decision Tree Algorithm for Classification", International Journal of Engineering Development and Research, 2001.

[9] Sumesh, C., R., and V.," Level Based Normal-Abnormal Classification of MRI Brain Images", International Journal of Computer Engineering and Technology (IJCET), Volume 4, Issue 2, March – April (2013).

[10] Selvarajah, and R., "Analysis and Comparison of Texture Features for Content Based Image Retrieval", International Journal of Latest Trends in Computing 108 Volume 2, Issue 1, March 2011.

[11] Vinoth, R., D., M., D., "Probabilistic Based Rock Texture Classification", International Journal of Advances in Engineering & Technology, Jan. 2014.