An Improved Fuzzy C-means Algorithm learned wavelet network for segmentation of Dermoscopic image

LishaRajan T L,Jemila Rose R

M.E. Student, Dept. of Computer &Communication Engineering,
St.Xavier's Catholic College of Engineering, Nagercoil, Tamilnadu, India.
Assistant Professor, Dept. of Information Technology,
St. Xavier's Catholic College of Engineering, Nagercoil, Tamilnadu, India.

Abstract-**Dermoscopy is one of the major imaging aspects used in the skin lesions. This paper presentsa new tacticfor the segmentation of skin abrasionsin dermoscopic images based on fuzzy C-means algorithm learned wavelet network (WN). The WN offeredhere is a member of fixed-grid WNs that is designedwith no requirementof training. Fuzzy C-means techniqueis used to enhancethe network structure. In addition, this methodhas the capabilityto allocateone data point into more than one cluster. Since average shift can rapidlyand consistentlyobtain cluster centers, the absoluteschemeis capable of successfullyidentifyingareaswithin an image. The clusters are combinedinto lesion and surrounding skin region, yielding the segmented dermoscopy image. Then, the image is segmented and the skin lesions preciseedgeis decidedappropriately.Thesegmentation algorithm wererelated to 30 dermoscopic images and calculatedby means ofthe segmentation outcomeachievedby a skilled pathologist as the ground truth.**

Keywords-Dermoscopy,image segmentation, melanoma diagnosis, Fuzzy C-Means, wavelet network (WN).

I. INTRODUCTION

Malignant melanoma (MM), the furthermostlethalform of skin cancer, is one of the most quicklythrivingcancers in the world, with an predictableyearlyoccurrenceof 70,230 and 8790 mortalities in the United States in 2011 [7]. The priorthe diagnosis, the lesserthe metastatichazard. The surveyshaverevealed that the healrate is nearly 100% if the skin cancer is recognized early enough and treated surgically [2]. Duringthe formerfewyears,the clinical analysesof MM were generallybased on the ABCD rule, in which factorsof symmetry, border irregularity, color, and dimension were examined,or on the seven-point checklist where various features such as color, shape, and texture wereisolated.Following on, the more specificand progressiveimagingsystem, namely dermoscopy, offered a noninvasive method for *in vivo* observation of pigmented skin lesions used in dermatology.

The typicalmethodin automatic dermoscopic image evaluationhas usually three steps: 1) image segmentation 2) feature extraction and feature selection 3) lesion classification. Owing to diverseshapes and colors of skin, segmentation is the most key stage of all. Thisapproachpresentsa new methodfor the segmentation of skin lesions in dermoscopic images that uses wavelet network (WN). The WN used is a member of fixed-grid WNs that is formedwith no requirementof training.

II. RELATED WORK

In [4] an automated dermatological tool for the parameterization of melanomas is presented. The system depends on the typicalABCD Rule and dermatological Pattern Recognition protocols.In[3], anovelclass of wavelet networks (WNs) for nonlinear system recognition isoffered. In the new networks, the classicconfiguration for a high dimensional system is selectedto be a superimposition of a number of functions with fewer variables. In[8]a wide-rangingmodel using supervised learning and MAP evaluationthat is capable of achievingmany generictasks in automated skin lesion diagnosis issuggested. The model is directedto segment skin lesions, findblockinghair, and identify the dermoscopic structure pigment network. The probabilistic descriptionof the model to createreceiver operating characteristic curves, show powerfulvisualizations of pigment networks, and provide classified informationintervals on segmentations.

III. PROPOSED MODEL

1. PREPROCESSING

Impulse noise is commonlyencountered in acquisition, transmission, and storage and processing of images. The existenceof impulse noise in an image may be eithercomparatively high or low. Thus, it could brutallydestroythe image quality and cause some loss of image information details. Filtering a digital image to eliminatenoise while preservingthe image featuresis annecessarypart of this task. The filtering approach used for removing impulse noise is median filter. This also includes the extraction of R, G, B values which is given as the input to the Wavelet Network [1].

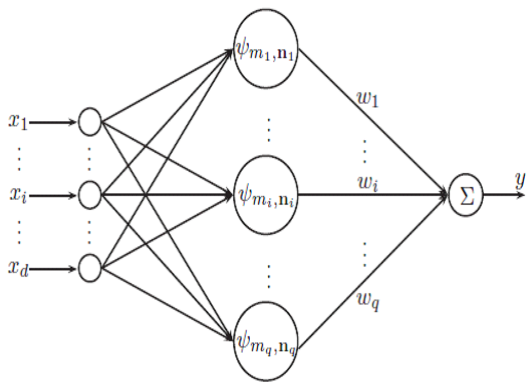


Fig. 2.1.1. Three-layered WN structure with one hidden layer.

2.ALGORITHM FOR BUILDING AN FGWN

A major benefitof WNs over other neural designsis the availability of effective construction algorithms for developing the network structure[5]. In FGWN, after forming the structure, the weights *wi*can be obtained through fuzzy estimation techniques. In this study, a constructive method is employed to build an FGWN. It can be described as follows

1)Consider M input–output data in the form $\{(\mathbf{x}(k), y(k)), k = 1, 2, \ldots, M\}$, where $\mathbf{x}(k) =[x(k)1 , \ldots , x(k)d]^T$is the input d-dimensional vector and inputs matrix is the form$X = [\mathbf{x}(1) \ldots \mathbf{x}(k) \ldots \mathbf{x}(M)]$. The output vector is considered as $\mathbf{y} = [y(1) \ldots y(k) \ldots y(M)]^T$. The FGWN structure is developed by a ten-stage algorithm. In many cases, input data of WN vary within a wide range and this variability reduces the efficiency of WN. Thus, this first stage is considered as the data preprocessing stage in which the input data are normalized to a certain range in order to avoid data scattering. If for kth inputT_k= max$_q$=1,..., $dx_q^{(k)}$,t_k= min$q$=1,...,$dx_q^{(k)}$then for mapping the input data to range $[a, b]$, the following equation is used

$$x_q^{(k)} = \frac{b-a}{T_k - t_k} x_{q,old}^{(k)} + \frac{aT_{k-b}t_k}{T_k - t_k} \qquad (1)$$

Where$x_{q,old}^{(k)}$ the jth input of kth sample is, $x_{q,new}^{(k)}$ is its value after the normalization process is carriedout. In the same way, all of the vector values $x_{new}^{(k)}$ =[$x_{1,new}^{(k)}, \ldots . x_{q,new}^{(k)}$,….$x_{d,new}^{(k)}$]fall within the range$[a, b]$

2) Selecting the mother wavelet: Due to better regularities and also the ease of frame generation in comparison with wavelet basis (orthonormal or biorthonormal), in multidimensional single scaling wavelet

frame is employed. In this study, the *d*-dimensional Mexican hat radial wavelet is used to implement WN. It is expressed as

$$\psi(\mathbf{x}) = \|\mathbf{x}\| = (d - \|\mathbf{x}\|^2)$$
$$\exp(-\|\mathbf{x}\|^2 /2) \qquad (2)$$

3) (Choose the scale and shift parameters) In this stage, minimum and maximum scale levels in the form$[m_{min}, m_{max}]$ and shift parameter in the form nj =[n1, . . . , nt, . . . , nd]T , where nt ∈ $[n_{t,min}, n_{t,max}]$, t =1, . . . , d and j = 1, . . . , $(n_{t,max} - n_{t,min} + 1)$, are to be employed.

4) (Formation of wavelet lattice) In this step regarding a hypershape on wavelet parameters space that was selected in the previous stage, the wavelet function is calculated for all the input vectors according to following equation:

$$\psi_{mi,\mathbf{n}j}(\mathbf{x}) = 2^{-mi\,d/2}\psi(2^{mi}\mathbf{x} - \mathbf{n}_j) \quad (3)$$

Where, $i = 1, \ldots, m_{max} - m_{min} + 1$,$d$ is input dimension. The number of nodes in a wavelet lattice is too many; therefore, it is obligatory that the number of these nodes be lowered and the shift and scale parameters of effective wavelets be selected. It is done through two stages of screening as follows.

5) *(Primary screening)* In this stage, for every scale level selected in stage 4, I_k set is formed for each input vector according to

$$I_k = \{(m, n) : |\psi_{mi,nj}(x)| \geq \max_i|\psi_{mi,nj}(x)|\} \qquad (4)$$

Where∈is a chosen small positive number (typically ∈=0.5). Also for the reason of simplicity in writing, the index of shift and scale parameters is eliminated. In fact, in this stage, the effective support of wavelets is selected.

6) (Secondary screening) In this stage, the shift and scale parameters of wavelets that are selected in at least two set of the sets in stage 5 are determined. In this way, set *I* is formed as follows:

$$I = \{(m, \mathbf{n}) : \text{if}[(m, \mathbf{n}) \in Ikr, (m, \mathbf{n}) \in Ikl] \Rightarrow r \neq l\} \qquad (5)$$

7) *(Formation of wavelet matrix)* Suppose that the number of selected wavelets in the last stage as *L*. In addition, to make the writing simpler, the couple index of $(m, \mathbf{n})$ is replaced with single index of $\{l = 1, \ldots, L\}$. In this stage, $W_{M\times L} = [\boldsymbol{\psi}1, \ldots, \boldsymbol{\psi}l, \ldots, \boldsymbol{\psi}L]$ matrix is calculated for all the input vectors and for all the selected shift and scale parameters that are in set *I*. In this matrix, $\boldsymbol{\psi}l$ vectors is considered as regressors . This

matrix is called wavelet matrix and calculated as follows:

$$W = \begin{bmatrix} \Psi 1(\mathrm{x}(1)) & \dots & \Psi \mathrm{L}(\mathrm{x}(1)) \\ \vdots & \dots & \vdots \\ \Psi 1(\mathrm{x}(\mathrm{M})) & \dots & \Psi \mathrm{L}(\mathrm{x}(\mathrm{M})) \end{bmatrix} \qquad (6)$$

The output vector is then constructed as

$$\mathrm{y}=\sum_{\mathrm{i}=1}^{\mathrm{L}} \mathrm{w_i} \Psi_\mathrm{i} = \mathrm{W}\theta \qquad (7)$$

Where weight vector $\boldsymbol{\theta}_{L\times 1} = [w_1. \ . \ . \ w_L]^T$is included of the weights between the wavelons of hidden layer and output layer.

8) *(Performing Fuzzy c-means algorithm)* In the previous stages, the output of the WN was expressed in terms of expanding the wavelet matrix members.In fuzzy clustering data elements can belong to more than one cluster, and relatedwith each element is a group of membership levels. These specifythe powerof the relationshipbetween that data element and a specificcluster. Fuzzy clustering is a practiceof allocatingthese membership levels, and then operatingthem to assign data elements to one or more clusters.It is based on minimization of the following objective function:

$$J_m = \sum_{i=1}^{N} \sum_{j=1}^{C} u_{ij}^m \left\| x_i - c_j \right\|^2 , \quad 1 \le m < \infty \qquad (8)$$

where m is any real number greater than 1, u_{ij} is the degree of membership of $x_i$in the cluster j, x_i is the ith of d-dimensional measured data, c_j is the d-dimension center of the cluster, and ||*|| is any norm expressing the similarity between any measured data and the center. According to the Fuzzy C-means[6] algorithm, to select the best subset of W, assuming that the size of this subset is known and denoted as s, the following steps should be taken. At first, the most significant wavelets in W is selected. After employing this stage, WN is constructed as

$$\mathrm{y}=\sum_{i=1}^{N} w_i \Psi_\mathrm{i}(\mathrm{x}) \qquad (9)$$

Where$s$ is the number of wavelons in the hidden layer and $wi$is the weight of wavelons. After performing the fuzzy algorithm, W is composed of orthonormal matrix Q and upper triangular matrix A. The above equation can be written as

$$\mathrm{y} = \mathrm{QA}\theta. \qquad (10)$$

9) *(Selecting the number of wavelons)* Wavelons are the nodes creating the hidden layer of the WN. By choosing the ideal number of wavelons, the system performance index is calculated. Then, the number of wavelons will change until the desired error measure is achieved

$$\mathrm{MSE}=\frac{1}{M}\sum_{k=1}^{M}(\hat{y}^{(\mathrm{k})}-\hat{y}^{(\mathrm{k})})^2 \qquad (11)$$

Here, *y*is the function approximation and *s* is the number of wavelons in hidden layer. The index of model performance in this method is mean squared error (MSE).Another method for selecting the number of wavelons is calculated by the generalized cross validation method.

10)*(Calculating wavelons weight coefficient):* This stage is the last stage of the algorithm.

Then, an FGWN is formed with three inputs, a hidden layer, and an output. In order to form the FGWN, the values of three color matrices are measuredas network inputs. In this way, the FGWN is developed. After that, the three matrices R, G, and B for each pixel values are considered as FGWN inputs, and the output of FGWN is a binary image that shows the segmented of original image.

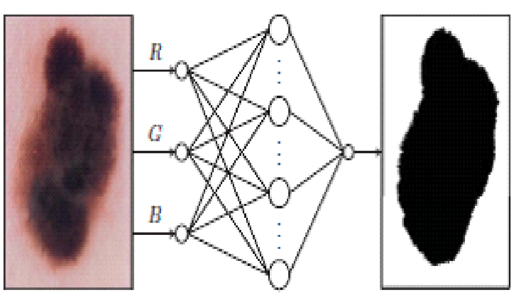


Fig.2.2.1 FGWN for segmentation of dermoscopic images.

The R, G, and B values of a dermoscopy image are regarded as FGWN inputs and the Fuzzy C-means algorithm is used to calculate the network weights and to optimize the network structure. The FGWN algorithm detect the lesion boundary carefully therefore, the lesion boundary that is the most important feature in detecting melanoma is extracted by FGWN with an acceptable accuracy.

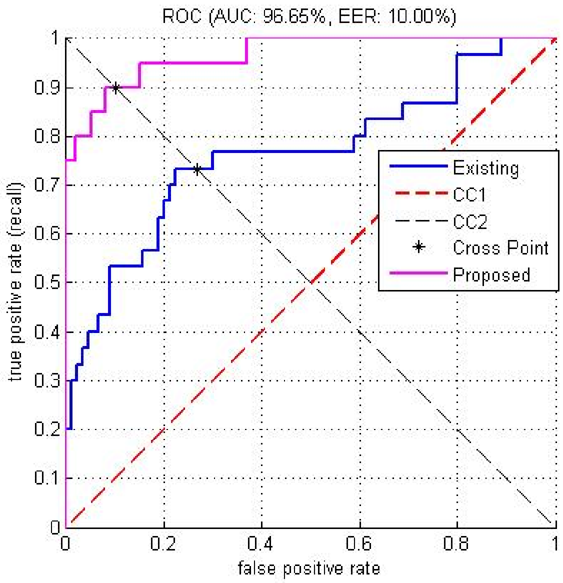
IV.EXPERIMENTAL RESULT

For our tentativecalculationwe used a PC with Intel(R) Core(TM)2 Due CPU T9550 (2.66 GHz) and 4 GB RAM. All the approacheswere recognizedby MATLAB 7.12.As statedbefore, segmentation is the most significantand acutestage of the three stages of instinctiveanalysisof melanoma which has a very substantialrole in the finalresult.Because of this reason, the performance of this state should beexamined by means of appropriate criteria.In a similar vein, it is evident from this table that the FGWN algorithm has an appropriate level of specificity. This means that the FGWN algorithm diagnoses the lesion boundary properly.

The Fuzzy C-means algorithm offers high accuracy in the detection of lesion boundary. As the true positive rate is high, there exist smooth border detection. Here the network is being trained using Fuzzy

ISSN: 2278 – 1323

algorithm that gives better result for overlapped datasets which is better than K-means algorithm.



V. CONCLUSION

In this work, a new approach is proposed for segmentation of dermoscopic images based on FGWN. An effective segmentation of dermoscopic images can be done by using wavelet network. The existence of two stage screening increases globality of the wavelet lattice and provide better estimation of the function especially for larger scales. The lesion boundary that is the most significant featurein detecting melanoma is extracted by FGWN with an acceptableaccuracy.The wavelet network presented here is a member of fixed-grid WNs that is formed with no need of training. Here the network weightcan be done by using Fuzzy C-means algorithm.This segmentation method has high accuracy and less error when compared to various algorithms.

REFERENECES

[1] Amir Reza Sadri, Maryam Zekri, Saeed Sadri, Niloofar Gheissari, Mojgan Mokhtari,and Farzaneh Kolahdouzan,"Segmentation of Dermoscopy Images Using Wavelet Networks," *IEEE Trans. On Biomed.* vol. 60, no. 4, April 2013.

[2] Di Leo G, C. Liguori, A. Paolillo, P. Sommella, An Improved Procedure for the Automatic Detection of Dermoscopy Structures in Digital ELM Images of Skin Lesions, in: Proceedings of the IEEE International Conference on Virtual Environment, Human-Computer Interfaces, and Measurement Systems, Istanbul, Turkey, 2008.

[3] Galvao R, V. M. Becerra, and M. F.Calado,"Linear waveletnetworks," *Int. J. Appl. Math. Comput. Sci.*, vol. 14, no. 2, pp. 221–232, Aug. 2004.

[4] Isasi A.G, B. G. Zapirain, and A. M.Zorrilla, "Melanomas non-invasive diagnosis application based on the ABCD rule and pattern recognition image processing algorithms," *J. Comput. Biol. Med.*, vol. 41, no. 9, pp. 742–755, Sep. 2011.

[5] Oussar Y and G. Dreyfus, "Initialization by selection for wavelet network training," *Neurocomputing*, vol. 34, no. 1, pp. 131–143, Sep. 2000.

[6] Schmid P, "Segmentation of digitized dermatoscopic images by twodimensional color clustering,"*IEEE Trans. Med. Imag.*, vol. 18, no. 2,pp. 164–171, Feb. 1999.

[7] Siegel R, E. Ward, O. Brawley, A. Jemal, Cancer statistics, 2011, CA: A Cancer Journal for Clinicians 61 (4) (2011) 212–236.

[8] Wighton P, T. K. Lee, H. Lui, D. I. McLean, and M. S. Atkins, "Generalizing common tasks in automated skin lesion diagnosis," *IEEE Trans. Inf.Tech. Biomed.*, vol. 15, no. 4, pp. 622–629, Jul. 2011.

