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 presents a new tactic for the segmentation of skin abrasions in dermoscopic images based on fuzzy C-means algorithm learned wavelet network (WN). The WN offered here is a member of fixed-grid WNs that is designed with no requirement of training. Fuzzy C-means technique is used to enhance the network structure. In addition, this method has the capability to allocate one data point into more than one cluster. Since average shift can rapidly and consistently obtain cluster centers, the absolute scheme is capable of successfully identifying areas within an image. The clusters are combined into lesion and surrounding skin region, yielding the segmented dermoscopy image. Then, the image is segmented and the skin lesions precise edge is decided appropriately. These segmentation algorithms were related to 30 dermoscopic images and calculated by means of the segmentation outcome achieved by 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 furthermost lethal form of skin cancer, is one of the most quickly thriving cancers in the world, with a predictable yearly occurrence of 70,230 and 8790 mortalities in the United States in 2011 [7]. The prior the diagnosis, the lesser the metastatic hazard. The surveys have revealed that the heal rate is nearly 100% if the skin cancer is recognized early enough and treated surgically [2]. During the former few years, the clinical analyses of MM
were generally based on the ABCD rule, in which factors of symmetry, border irregularity, color, and dimension were examined, or on the seven-point checklist where various features such as color, shape, and texture were isolated. Following on, the more specific and progressive imaging system, namely dermoscopy, offered a noninvasive method for in vivo observation of pigmented skin lesions used in dermatology.

The typical method in automatic dermoscopic image evaluation has usually three steps: 1) image segmentation 2) feature extraction and feature selection 3) lesion classification. Owing to diverse shapes and colors of skin, segmentation is the most key stage of all. This approach presents a new method for 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 formed with no requirement of training.

II. RELATED WORK

In [4] an automated dermatological tool for the parameterization of melanomas is presented. The system depends on the typical ABCD Rule and dermatological Pattern Recognition protocols. In [3], a novel class of wavelet networks (WNs) for nonlinear system recognition is offered. In the new networks, the classic configuration for a high dimensional system is selected to be a superimposition of a number of functions with fewer variables. In [8], a wide-ranging model using supervised learning and MAP evaluation that is capable of achieving many generic tasks in automated skin lesion diagnosis is suggested. The model is directed to segment skin lesions, find blocking hair, and identify the dermoscopic structure pigment network. The probabilistic description of the model to create a receiver operating characteristic curves, show powerful visualizations of pigment networks, and provide classified information intervals on segmentations.

III. PROPOSED MODEL

1. PREPROCESSING

Impulse noise is commonly encountered in acquisition, transmission, and storage and processing of images. The existence of impulse noise in an image may be either comparatively high or low. Thus, it could brutally destroy the image quality and cause some loss of image information details. Filtering a digital image to eliminate noise while preserving the image
features is an inevitable part 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].

![Three-layered WN structure with one hidden layer.](image)

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

2. ALGORITHM FOR BUILDING AN FGWN

A major benefit of WNs over other neural designs is the availability of effective construction algorithms for developing the network structure[5]. In FGWN, after forming the structure, the weights \( w_i \) 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 \( \{(x(k), y(k)), k = 1, 2, \ldots , M\} \), where \( x(k) = [x(k)1, \ldots , x(k)d]^T \) is the input \( d \)-dimensional vector and inputs matrix is the form \( X = [x(1) \ldots x(k) \ldots x(M)] \). The output vector is considered as \( y = [y(1) \ldots y(k) \ldots y(M)]^T \). 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 \( k \)th input \( T_k = \max_q=1,\ldots ,d x_q^{(k)},T_k = \min_q=1,\ldots, d x_q^{(k)} \) then for mapping the input data to range \([a, b] \), the following equation is used

\[
x_q^{(k)} = \frac{b-a}{T_k-t_k} x_q^{(k)} + \frac{aT_k-n_t}{T_k-t_k} (1)
\]

Where \( x_{q,old}^{(k)} \) the \( j \)th input of \( k \)th sample is, \( x_{q,new}^{(k)} \) is its value after the normalization process is carried out. In the same way, all of the vector values \( x_{new}^{(k)} = [x_{1,new}^{(k)}, \ldots, x_{q,new}^{(k)}, \ldots, x_{d,new}^{(k)}] \) 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

\[
\psi(x) = \|x\|^2 - \|x\|^2 \exp(-\|x\|^2/2)
\]  

(2)

3) (Choose the scale and shift parameters) In this stage, minimum and maximum scale levels in the form \([m_{\text{min}}, m_{\text{max}}]\) and shift parameter in the form \(n_j = [n_1, \ldots, n_t, \ldots, n_d]\) are 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:

\[
\psi_{mi, nj}(x) = 2^{-mi d/2} \psi(2^{mi} x - nj)
\]  

(3)

Where, \(i = 1, \ldots, m_{\text{max}} - m_{\text{min}} + 1, d\) 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

\[
I_k = \{(m, n) : |\psi_{mi, nj}(x)| \geq \max_i |\psi_{mi, nj}(x)|\}
\]  

(4)

Where \(\epsilon\) is a chosen small positive number (typically \(\epsilon = 0.5\)). 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:

\[
I = \{(m, n) : \text{if}(m, n) \in I_k \land (m, n) \in I_l \Rightarrow r \neq 1\}
\]  

(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 \((m, n)\) is replaced with single index of \(\{l = 1, \ldots, L\}\). In this stage, \(W_{M \times L} = [\psi l, \ldots, \psi l, \ldots, \psi L]\) 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, \(\psi l\) vectors is considered as regressors. This
The output vector is then constructed as
\[ y = \sum_{i=1}^{L} w_i \Psi_i = W\theta \quad (7) \]

Where weight vector \( \theta_{L \times 1} = [w_1, \ldots, w_L]^T \) 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 related with each element is a group of membership levels. These specify the power of the relationship between that data element and a specific cluster. Fuzzy clustering is a practice of allocating these membership levels, and then operating them to assign data elements to one or more clusters. It is based on minimization of the following objective function:

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

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 \( i \)th of \( d \)-dimensional measured data, \( c_j \) is the \( d \)-dimension center of the cluster, and \( \| \cdot \| \) 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 wavelons in \( W \) is selected. After employing this stage, WN is constructed as

\[ y = \sum_{i=1}^{N} w_i \Psi_i(x) \quad (9) \]

Where \( s \) is the number of wavelons in the hidden layer and \( w_i \) 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

\[ y = QA\theta. \quad (10) \]

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

\[ \text{MSE} = \frac{1}{M} \sum_{k=1}^{M} (\hat{y}^{(k)} - \hat{y}^{(k)})^2 \quad (11) \]
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 measured as 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.

![FGWN for segmentation of dermoscopic images.](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 tentative calculation, we used a PC with Intel(R) Core(TM)2 Duo CPU T9550 (2.66 GHz) and 4 GB RAM. All the approaches were recognized by MATLAB 7.12. As stated before, segmentation is the most significant and acute stage of the three stages of instinctive analysis of melanoma which has a very substantial role in the final result. Because of this reason, the performance of this state should be examined 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
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 feature in detecting melanoma is extracted by FGWN with an acceptable accuracy. The wavelet network presented here is a member of fixed-grid WNs that is formed with no need of training. Here the network weight can be done by using Fuzzy C-means algorithm. This segmentation method has high accuracy and less error when compared to various algorithms.

REFERENCES


