International Journal of Modern Trends in Engineering and Research
www.ijmter.com

Computer aided classification of Bascal cell carcinoma using adaptive Neuro-fuzzy Inference System

Rajvi Parikh[1]
[1]Asst. Prof., I.T. Department, G H Patel College of Engg & Tech, rajviparikh@gcet.ac.in

Abstract— For skin lesion detection pathologists examine biopsies to make diagnostic assessment largely based on cell anatomy and tissue distribution. However in many instances it is subjective and often leads to considerable variability. Whereas computer diagnostic tools enable objective judgments by making use of quantitative measures. Paper presents a diagnosis system based on an adaptive Neuro-fuzzy inference system for effective classification of Bascal cell carcinoma images from the given set of all types of skin lesions. System divide in three parts. Image Processing, Feature Extraction, and classification. First part deals with the noise reduction and artifacts removing from the set of images. Second part deals with extracting variety of features of Bascal Cell Carcinoma using the Greedy feature flip algorithm (G-flip), and classification method using ANFIS algorithm and finally Part three deals with the results that is classification of BCC images from the variety of pre-cancerous stage images that is Actinic Keratosis and also other images called psoriasis which looks as cancer images at a first look . The results confirmed that the proposed ANFIS model has potential in classifying the skin cancer diagnosis.

I. INTRODUCTION

Bascal cell carcinoma is the most common and largely affected skin cancer around the globe. Paper describes the way that will detect the Bascal cell carcinoma from the variety of images such as images of mole, HSS (healthy surrounding skin and, lesion images. Basically all the images having noise or unwanted artifacts in it so for initial stage image preprocessing is required so it will remove all the noisy data from an image. The paper introduce ANFIS (Adaptive-Network-Based Fuzzy Inference Systems) as a diagnosis system. ANFIS has proven to be an excellent function approximation tool, where it implements a first order Sugeno-style fuzzy system. The basic idea behind these neuro-adaptive learning techniques is very simple. These techniques provide a method for the fuzzy modelling procedure to acquire information about a data set, in order to compute the membership function parameters that best allow the associated fuzzy inference system track the given input/output data. The data inputs are features selected by the G-flip algorithm; from a number of features that were extracted from images of three different skin lesions (Fig. 1), are described in the next section These lesions can be classified into three groups: (1) Actinic Keratosis or malignant melanoma, a type of skin cancer known also as a solar keratosis, can be considered as the first step of the development of skin cancer); (2) Basal Cell Carcinoma is a cancer that begins in the deepest basal cell layer of the epidermis (the outer layer of the skin); and (3) Psoriasis is a chronic skin condition which tends to run in families [1]. The necessary features have been selected from the whole large set of features which are sufficient enough to predict the targeted class well. A small good set of features could achieve high performance level so feature selection is crucial for efficient learning. Feature extraction is the process which identifies the features that discriminates key features from the set of features.

(a) BCC (b) Psoriasis (c) Actinic Keratosis

Figure 1: Different Organisms

II. Extraction and selection of features

Feature extraction involves simplifying the amount of resources required to describe a large set of data accurately. The features that were extracted in total 75 characteristics or parameters. In order to avoid highly redundant features we have extracted the correlation coefficient matrix and discarded those features with a correlation coefficient above 0.98 with respect to other features. In this way only 39 out of the initial 75 features have been selected. These features are extracted by using different algorithm of image processing: (1) Edge detection, with two main categories methods Gradient and Laplacian. The Gradient method, mainly represented by three types (Sobel, Prewitt, and d Canny), Refs. [2-4] detects the edges by looking for the maxi mum and minimum values that satisfy the first derivative of the image. The Laplacian method searches for the zero crossing in the second order derivative of the images to find edges. These parameters have been extracted from database of 150 images. Refs [5-7] the feature selection is the task of choosing a small set out of a given set of features that capture the relevant properties of the data. The need for feature selection arises to avoid the presence of large number of weakly relevant and redundant features in the data set [8]. A good choice of features is a key for building compact and accurate classifiers. In this paper the Greedy feature flip (G-flip) [10] algorithm is used to select a small set of features, which can be used as inputs to our diagnosis system. G-flip is a greedy search algorithm for maximizing the evaluation function e(F), where F is a set of features. The algorithm repeatedly iterates over the feature set and updates the set of chosen features. In each iteration it is decided to remove or add the current feature to the selected set by evaluating the margin term in the Eq. (1) with and without this feature. The following equation shows the evaluating function for a training set S and a weight vector w according to the second definition of the margin in ref[9].

$$e(w) = \sum \Theta s w / x \, (x) \qquad (1)$$

where, $x \in s$

This algorithm converges to a local maximum of the evolution function as each step increases its value and the number of possible feature sets is finite. The computational complexity of one pass over all features of G flip is

$$\Theta(N^2 m^2) \qquad (2)$$

Where n is the number of features and m is the number of instances.

III. Classification

ANFIS uses the hybrid learning method since it combines gradient descent and the least-squares methods [10] In a fuzzy inference system, there are three types of input space partitioning: grid, tree, and scattering partitioning. The "curse of dimensionality" refers to a situation where the number of fuzzy rules increases exponentially with the number of input variables [12,13]. Therefore, six features were used in the diagnosis system. These features have to be very accurate, so the features selection algorithm, Greedy feature flip (G-flip) algorithm was used to determine our best 6 features out of the dataset of features that were extracted previously. Paper describes ANFIS algorithm in order to classify the given trial images in to images that falls in category of Bascal Cell Carcinoma type of skin lesion. ANFIS network structure shows multi-layered neural network as shown in fig 2

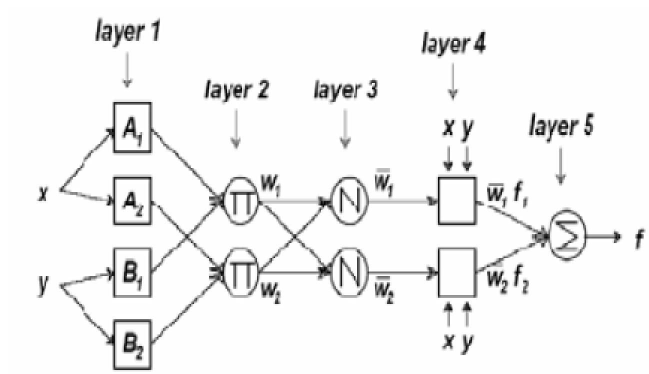

Figure 2: type -3 ANFIS

First layer executes a fuzzification process, the second layer executes the fuzzy AND of the antecedent part of the fuzzy rules, the third layer normalizes the membership functions (MFs), the fourth layer executes the consequent part of the fuzzy rules, and finally the last layer computes the output of fuzzy system by summing up the outputs of fourth layer. Here for ANFIS structure (Fig. 2) two inputs and two labels for each input are considered. The feed forward equations of ANFIS are as follows: [10-12]

$$w_i = \mu_{A_i}(x) \times \mu_{B_i}(y), i = 1,2 \tag{3}$$

$$\overline{w}_i = \frac{w_i}{w_1 + w_2}, i = 1,2 \tag{4}$$

$$f = \frac{w_1 f_1 + w_2 f_2}{w_1 + w_2} = \overline{w}_1 f_1 + \overline{w}_2 f_2 \tag{5}$$

Where $f_1 = p_1 x + q_1 y + r_1$; and $f_2 = p_2 x + q_2 y + r_2$; In order to model complex nonlinear systems, the ANFIS model carries out input space partitioning that splits the input space into many local regions from which simple local models (linear functions or even adjustable coefficients) are employed. The ANFIS uses fuzzy MFs for splitting each input dimension. The input space is covered by overlapping MFs, which means that several local regions can be activated simultaneously by a single input. As simple local models are adopted in ANFIS model, the ANFIS approximation ability will depend on the resolution of the input space partitioning, which is determined by the number of MFs in ANFIS and the number of layers. Usually MFs are used as bell-shaped with maximum equal to 1 and minimum equal to 0 such as

$$\mu_{A_i}(x)=\frac{1}{1+\left[\left(\frac{x-c_i}{a_i}\right)^2\right]^{b_i}} \quad \mu_{A_i}(x)=\exp\left\{\left[\left(\frac{x-c_i}{a_i}\right)^2\right]^{b_i}\right\} \tag{6}$$

$$\mu_{A_i}(x)=\exp\left\{\left[\left(\frac{x-c_i}{a_i}\right)^2\right]^{b_i}\right\} \tag{7}$$

Where {ai, bi, ci} are the parameters of MFs which are affected in shape of MFs. The ANFIS uses member function for each input. The training was run for 10 iterations. The network performance was evaluated on the checking set, after each iteration, by calculating the root-mean-square errors (RMSE)

$$RMSE=\sqrt{\frac{\sum_{K=1}^{K}\left(Y_K-\hat{Y}_K\right)}{K}} \tag{8}$$

Where the k is the pattern number, k=1, 2...K, *K Y* is the correct value, and *k Yˆ* is the output value from the ANFIS. The RMSE was also calculated on training data set for every iteration. The optimal number of iterations obtained was 9 epochs by the time RMSE reached its minimum value. We then convert the error from RMSE to percentage error.

IV. Result and Discussion

Image dataset has been divided in to variety of groups including all three types of images and hence two classes are being made i.e. cancerous image is class 1 and non-cancerous image is class 2 The training data set was made up from 80% of the overall data and the other 20% of the data is considered as testing data set. Three versions of these data sets are used where each version was randomly disordered in order to cross–validate the results. After applying the methodology and running the classification algorithm for 10 iterations, it reached the minimum RMSE value at the ninth epoch. The ANFIS structure information is shown in table below.

Number of nodes	1503
Number of nonlinear parameters	50
Number of training data pairs	80
Number of fuzzy rules	700
Number of linear parameters	5103
Total number of parameters	5157
Number of checking data pairs	20

Table 1: ANFIS structure Information

This work was implemented by using MATLAB under Windows XP with Intel Centrino processor running at 1.87 GHz. The time spent to get the result of this classification was an average 120 to 150 minutes. The ANFIS classification algorithms KNN classifier optimized with GA represents a valid tool to study the significance of different features for a given diagnosis problem [1,2]. We have tested other classifiers such as KNN classifier optimized with GA, and artificial neural networks (ANN) (Multilayer Perceptron) yielding lower level of accuracies and requiring much longer computing times. The low performance level of the ANN classifiers is mainly due to the limited number of samples (200 images) of the database.

V. Conclusion

This paper introduces the adaptive neuro-fuzzy inference system (ANFIS) as a diagnosis system for the diagnosis of skin lesions type BCC. This system showed good performance accuracy, It also validates the optimization technique of the different features to a high level of classification accuracy, where these features were extracted by image processing and then selected by using G-flip algorithm. The result of this classification method showed that by using ANFIS, produces better result than with other algorithms for diagnosis systems for skin lesion. The system can also be tested for other types of skin lesion as future work.

References

[1] E. Ros, M.M. Rodríguez, S. Mota, J.L. Bernier, I. Rojas,C.G. Puntonet, E.W. Lang, Skin lesions diagnosis basedon fluorescence image processing, simple parametersscanning, LNCS 2562 (2003) 858-867.

[2] S. Odeh, E. Ros, I. Rojas, J. Palomares, Skin lesion diagnosis using fluorescence images, Proceeding of ICIAR06, International Conference on Image Analysis and Recognition, 2006, pp. 648-659.

[3] J. Canny, A computational approach to edge detection, IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI, 1986, pp. 679-698.

[4] J.R. Parker, Algorithms for Image Processing andComputer Vision, John Wiley & Sons, 1997, pp. 23-29.

[5] E. Ros, M.M. Rodríguez, S. Mota, J.L. Bernier, I. Rojas, C.G. Puntonet, E.W. Lang, Skin lesions diagnosis based on fluorescence image processing, simple parameters scanning, LNCS 2562 (2003) 858-867.

[6] J.S. Lim, Two-Dimensional Signal and Image Processing, Englewood Cliffs, NJ, Prentice Hall, 1990, pp. 478-488.

[7] R. Gonzalez, R. Woods, Digital Image Processing, Addison Wesley, 1992, pp.414-428.

[8] P.M. Narendra, K. Fukunaga, A branch and bound algorithm for feature subset selection, IEEE Transactions on Computers 26 (1977) 917-922.

[9] R.G. Bachrach, A. Navot, N. Tishby, Margin based feature selection: theory and algorithms, ICML, 2004.

[10] S. Odeh, J. Hodali, M. Sleibi, I. Salsa', Cursor movement control development by using ANFIS algorithm, The International Arab Journal of Information Technology (2009) 448-453.

[11] F.A. Alturki, A.B. Abdennour, Neuro-fuzzy control of a steam boiler turbine unit, Proceeding of the 1999 IEEE, International Conference on Control Applications, 1999,pp. 1050-1055

[12] V.S. Ghomsheh, M.A. Shoorehdeli, M. Teshnehlab, Training ANFIS structure with modified PSO algorithm, Proceeding of the 15th Midetaerian on Control and Automation, 2007.

[13] V.S. Ghomsheh, M.A. Shoorehdeli, M. Teshnehlab, Training ANFIS structure with modified PSO algorithm, Proceeding of the 15th Midetaerian on Control and Automation, 2007

