IJIREEICE

ISSN (Online) 2321-2004
ISSN (Print) 2321- 5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

NFTPCOS-18

National Conference on Future Technologies in Power Control and Communication Systems

College of Engineering Perumon, Kollam, Kerala

Vol. 1, Special Issue 2, March 2018

Back Propagation Neural Network (BPNN) based Melanoma Classification on Dermscopy Images

Aseena A[1], Anjali A[2], Arsha A S[3], Athulya A[4]

Assistant Professor, Electronics and Communication Engineering, College of Engineering Perumon, Kollam, India[1]

Students, Electronics and Communication Engineering, College of Engineering Perumon, Kollam, India[2,3,4]

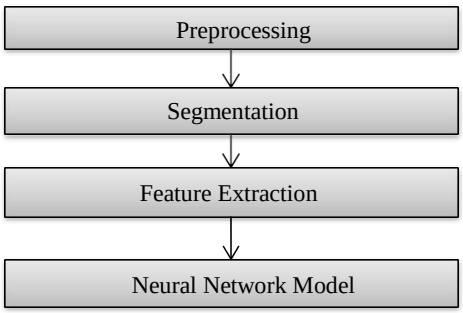
Abstract: Melanocytic tumors as Benign or Malignant can be diagnosed using different methods. One such method is a Digital Dermoscopic Image. Many conventional methods employ Support Vector Machines (SVMs), K- Nearest Neighbor (KNN), Adaboost, etc have been widely used for lesion classification. An Artificial Neural Network (ANN) is a computational model based on the structure and functions of biological neural networks. ANN can be used widely for medical diagnosis and tumor detection. There are various ANN that can be employed for this purpose. Our proposed method employs BPNN for melanoma classification on Dermoscopy images. We develop a novel method for classifying melanocytic tumors as Benign or Malignant by the analysis of digital dermoscopic images.

Keywords: Dermoscopy image, Benign, Malignant, BPNN (Back Propagation Neural Network)

I. INTRODUCTION

Both benign and malignant are the type of tumors. The incidence of malignant melanoma, the most dangerous skin cancer, has increased rapidly during the last decade, and the figures are still rising. As per the census in 2016, an estimated total of 19,640 deaths were recorded due to skin cancer. Dermoscopy is an imaging technique that allows magnified visualization of skin surface and thereby increasing the accuracy. It has been reported that the well-trained and experienced dermatologists are able to reach the diagnostic accuracy of about 75% in visual preoperative classification, the discriminating ability of digital image analysis was evaluated in more than 350 malignant melanoma and benign melanocytic lesions had been successful. The existing method include Computer Aided Diagnosis (CAD), KNN and ABCD rule that analyses asymmetry, border irregularity, color variation and different structures of a lesion. Both are with an accuracy of 92.2%, sensitivity of 97.9% and specificity of 90.3% but, images may not contain entire lesion [1] [3]. This can be overcome by using our technique in which tumors descriptive feature are utilized which are fed to neural network ensemble model that is trained to differentiate malignant lesion from benign lesion.

II. BLOCKDIAGRAM



Most of the existing literature regarding a machine learning approach to classify melanocytic lesions malignant or benign, using dermoscopic images The lesion features used in the classification framework are inspired on border, texture, color and structures used in popular dermoscopy algorithms performed by clinicians by visual inspection. Selection of a set of weights and threshold are the main issues of dermoscopy algorithms which seems not be robust or independent of population. To overcome this issue, we prefer machine learning technique. Pre-processing stage includes hair removal filtering, each image is automatically segmented using well known image segmentation algorithms. Each lesion is processed on the basis of a feature vector that includes shape, color and texture information

IJIREEICE

ISSN (Online) 2321-2004
ISSN (Print) 2321- 5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

NFTPCOS-18

National Conference on Future Technologies in Power Control and Communication Systems

College of Engineering Perumon, Kollam, Kerala

Vol. 1, Special Issue 2, March 2018

as well as local and global parameters [7]. Adaboost with c4.5 decision tree can be used for learning and classification process. On an automatically segmented database, this model delivered a specificity of 77% with sensitivity of 90% and a specificity of 85% with sensitivity of 90% on a manually segmented database. The reminder of the paper is organized as follows. Section A describes preprocessing and segmentation. In section B, feature extraction is presented section C introduces the neural network model.

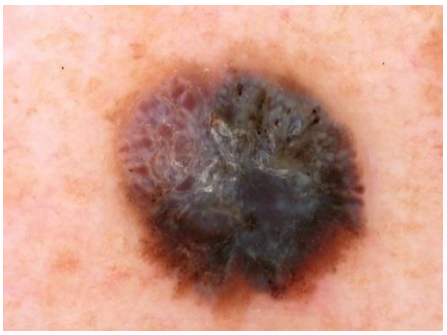


Fig 1: input image

A. Preprocessing and Segmentation: The key problems for the precise segmentation and analysis of the skin malignant melanoma image with hair are the repair of hair occluded information. In this paper an unsupervised repair algorithm for the hair-occluded information is proposed. This includes 3 steps: first, image with hairs are enhanced by morphologic closing based top-hat operator and then segmented through static threshold. Second, the hairs are extracted based on the elongated of connected region; third, the hair- occluded information is repaired by the PDE – based image inpainting [5]. The goal of image segmentation is to divide an image into constituent parts that correlate to objects with in the image. In our work, K-means clustering aims to partition 'n' observations into K clusters in which each observation belongs to the cluster with the nearest mean, serving as a proto type of the cluster.

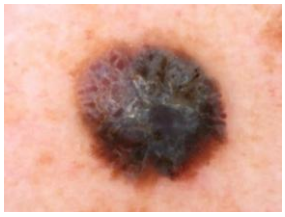


Fig 2: preprocessed image

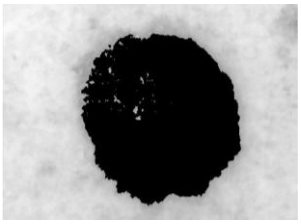


Fig 3: K means cluster

B. Feature Extraction: In this section the features that were used to characterize the skin lesion images are described. We describe a set of widely used color and texture features as well as asset of lesion border features on incomplete lesions.

Region division on dermoscopy images: We divide the lesion object into two regions; diffusion region and an inner lesion region. Each lesion object is separated from the background skin using SGNN method and manual interaction. From this automatic border detection is performed. Many clinical features such as asymmetry and border irregularity are calculated from the border. The extraction of other important color or texture related clinical features critically depends on the accuracy of border detection.

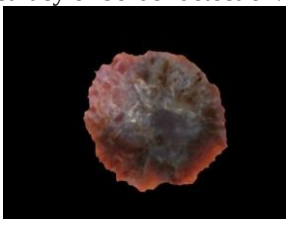


Fig 4 : Inner lesion

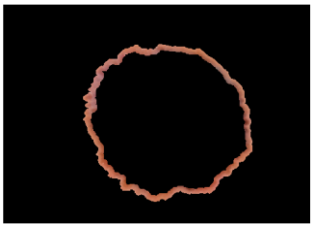


Fig 5 : Diffusion lesion

Description of Color and Texture Features: In order to quantify the colors present in a lesion, two statistics were (mean and standard deviation) over the channels of six different color spaces and several color asymmetries, histogram

IJIREEICE
ISSN (Online) 2321-2004
ISSN (Print) 2321- 5526
International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering
NFTPCOS-18
National Conference on Future Technologies in Power Control and Communication Systems
College of Engineering Perumon, Kollam, Kerala
Vol. 1, Special Issue 2, March 2018

distance and centroidal distance features were calculated. Mean and standard deviation: The mean and standard deviation values calculated over a particular channel quantify the average color. Color features were calculated as follows :(6 color spaces) (3 channels in each color space) (2 statistics: mean and standard deviation) (3 regions {lesion, inner periphery, outer periphery}). The ratios and differences of two statistics over the three regions were calculated (outer/inner), (outer/lesion), (inner/lesion), (outer-inner), (outer-lesion) and (inner-lesion). LUV histogram distance: In order to determine the color similarity of two regions, the histogram distance in the L*U*V color space was used. For histogram computation, the color space was coarsely quantized into 4x8x8 bins. The color similarity between the two regions was quantified by the L1- and L2- norm histogram distance. The use of these norms is justified because the color space is coarsely quantized and there is negligible correlation between adjacent histogram bins. The histogram distance between pairs of three regions that is, lesion, inner periphery, outer periphery using the two distance measures were calculated. The total number of color features in this category was 6.

Centroidal distances: The centroidal distance for a channel is defined as the distance between the geometric centroid (of the binary object) and the brightness centroid of that channel. The brightness centroid was calculated similarly to the geometric centroid except that the moment calculations were weighted by the pixel values. If the pigmentation in a particular channel is homogeneous, the brightness centroid will be close to the geometric centroid and thus the centroidal distance for that channel will be small. In order to achieve invariance in scaling, the distance values were divided by the lesion diameter. The centroidal distance values were calculated for all 3 channels of the six color spaces. The total number of color features in this category was 18 [2].

Description of texture features: In order to quantify the texture present in a lesion, a set of statistical texture descriptor based on the gray level co-occurrence matrix (GLCM) were employed. GLCM based texture is one of the most well-known and widely used methods in literature [4]. Although many statistics can be derived from the GLCM, eight gray level shift invariant statistics were used in this study in order to obtain a texture characterization that is robust to linear shift in the illuminate intensity. These statistics were maximum probability, energy, entropy, dissimilarity, contrast, inverse difference, inverse difference moment and correlation.

Description of Border Features:

A lesion's border is represented by pixels comprising the lesion's boundary, obtained as a result of the lesion segmentation process. The border criterion is basically measured by the irregularity of its shape, which may be regular (usually related to in malignant lesion) or irregular (usually related to benign lesion) in clinical images, or even an abrupt cut- off of network at the border of lesion in dermoscopy images. Several methods have been proposed for assessing the border irregularity. In order to identify the sharp transition between inside and outside region of a 100 cm concerning its border. For each region, the ratio of color intensity inside and outside the lesion and the gradient of the color intensity were computed in particular color channels.

Lesion concavity features: Benign lesion can have roughly elliptical with convex border that exhibit few concavities whereas malignant lesions tend to have highly irregular border with more frequent concavities. For identifying border concavities, convex hull is used. Fig shows the convex hull of both malignant and benign segmented lesion. Lesion convex hull can be used to define 6 feature parameter that are descriptive of the degree of concavities occurring on the lesion border Let Nr be the number of distinct concave segments along the border of the lesion object, and let *li*, *di* and *RAi* be the span, depth and areas of the concavity. Then three pairs of sample statistics are defined on the lesion object as follows:

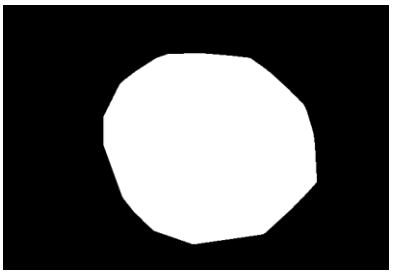


Fig 6: Convex hull

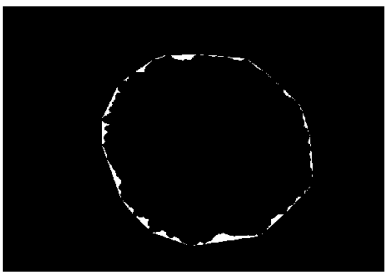


Fig 7: Concavity

IJIREEICE
ISSN (Online) 2321-2004
ISSN (Print) 2321- 5526
International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering
NFTPCOS-18
National Conference on Future Technologies in Power Control and Communication Systems
College of Engineering Perumon, Kollam, Kerala
Vol. 1, Special Issue 2, March 2018

(a) Mean and standard deviation of span:

$$\mu_s = \frac{1}{N_r}\sum_{i=1}^{N_r} l_i \quad (1)$$

$$\delta_s = \sqrt{\frac{1}{N_r}\sum_{i=1}^{N_r}(l_i - \mu_s)^2} \quad (2)$$

(b) Mean and standard deviation of depth:

$$\mu_d = \frac{1}{N_r}\sum_{i=1}^{N_r} d_i \quad (3)$$

$$\delta_d = \sqrt{\frac{1}{N_r}\sum_{i=1}^{N_r}(d_i - \mu_d)^2} \quad (4)$$

(c) Mean and standard deviation of average thickness

$$\mu_t = \frac{1}{N_r}\sum_{i=1}^{Nr}\frac{RA_i}{l_i} \quad (5)$$

$$\delta_t = \sqrt{\frac{1}{N_r}\sum_{i=1}^{N_r}\left(\frac{R\,Ai}{li} - \mu_t\right)^2} \quad (6)$$

Benign lesions can be characterized as having similar concavities, shallower, and more uniform in both span and depth. Malignant lesions tend to be highly variable spans, depth and thickness. Therefore, the values of these 6 features parameters computed on malignant lesions are larger than on the benign lesions. This is true even for incomplete objects; note that the artificial boundaries at form of image are not included in the computation.

Separation between inner and outer cavity lesion: Benign lesions tend to have well-defined outer lesion border (yellow contour) and inner lesion border (blue contour). The diffusion region between inner and outer borders is generally of uniform width, whereas malignant melanoma exhibit diffusion region that is less regular, with uniform width. The separation between inner and outer is highly variable on malignant lesion as compared to benign lesion. Let outer and inner denote the outer and inner borders respectively, and d (pi; pj) denote the Euclidean distance between any two pixels pi and p_j. Then define the distance Di from the i^{th} pixel on the outer border to the inner border as:

$$Di = \min_j(d(pi; pj)); \; pi \in \Gamma outer; \; pj \in \Gamma inner \quad (7)$$

The degree of variability of the separation between the inner and outer borders can then be represented as the variance of the distance from the outer border pixels to the inner border: As compared with the border characteristics of benign lesions, the inner and outer borders of malignant lesions tend to follow different paths causing highly variable separation between the inner and outer borders.

IJIREEICE

ISSN (Online) 2321-2004
ISSN (Print) 2321- 5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

NFTPCOS-18

National Conference on Future Technologies in Power Control and Communication Systems

College of Engineering Perumon, Kollam, Kerala

Vol. 1, Special Issue 2, March 2018

Feature normalization: In classification takes the features that characterizes the sample quite often have different ranges. Many classifiers such as k-nearest neighbors and neural network require that the features be normalized so that their values fall within a specified region. One of the most common normalization methods is the z-score transformation given by,

$$z_{ij} = \frac{(x_{ij} - \mu_j)}{(3\delta_j)+} \quad (8)$$

where x_{ij} is the value of the j^{th} feature of the i^{th} sample, and μj and σj are the mean and standard deviation of the j^{th} feature, respectively.

Assuming that each feature is normally distributed, the transformation guarantees 99% of z_{ij} be in the [0,1] range. The out- of- range values are truncated to either 0 or 1.

Feature dimensionality reduction: Principle Component Analysis (PCA) is a very popular technique for dimensionality reduction [6]. Given a set of data on n dimensions, PCA aim to find a linear subspace of dimension d lower than n such that the data points lie mainly on the linear subspace Such a reduced subspace attempts to maintain most of variability of data. The linear space can be specified by the orthogonal vectors that form a new coordinate system, called the 'principle components'. The principal component is orthogonal, linear transformation of the original points so there can be no more than n of them. However, the hope is that only d<n principal components are needed to approximate the space spanned by the n original axis.

C. NEURAL NETWORK MODEL

For training neural networks involves hidden neurons have caused a resurgence of interest in non- algorithmic supervised learning. A supervised learning scheme is implemented using a database which consists of a set of input pattern together with corresponding target. The objective of the training is to let the trainee extract relevant information from the database in order to classify future input patterns. Each network individual is initially trained using the same training sample on a neural network model. Then, the individual network outputs must be weighted to form the overall output.

$$y(x, a) = \sum_{i=1}^{p} a_i y_i(x) \quad (9)$$

Where x is the input feature vector and a_i is the weight applied to the output of the i^{th} individual network.

A neural network model of independently trained network can make a collective classification in several ways. The most powerful voting rule appears to be plurality in which the collective decision is the classification reached by mix networks than any other. Simpler to analyze is the majority voting rule which choose the classification made by more than half the networks. When there is no agreement among more than half the networks, the result is considered an error. Different types of Neural Networks are Feedforward Neural Network, Radial basis function Neural Network, Kohonen self-organizing Neural Network, Recurrent Neural Network etc.

Proposed method using BPNN: In general, neural network model formation consists of two steps. For the use in data analysis systems it is desirable to automate both of these steps. The first step is to form the structure and to train a number of component networks which will be included in a model as preliminary part. The second step is to select the network, the solution of which will produce a final solution and to define a way and parameter of a common solution formation. Back propagation is method used in artificial neural network to calculate a gradient. This gradient is used to calculate the weights to be used in the network. Back propagation algorithm is the most common method of training multi-layer networks. For BPNN classification, the analysis is divided into five aspects: (1) input data preprocessing, structure and encoding (2) output encoding and extraction of classes (3) network architecture (4) training algorithms (5) comparisons to conventional classifiers Final prediction depicts the identification of benign and malignant tumor.

IJIREEICE
ISSN (Online) 2321-2004
ISSN (Print) 2321- 5526
International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering
NFTPCOS-18
National Conference on Future Technologies in Power Control and Communication Systems
College of Engineering Perumon, Kollam, Kerala
Vol. 1, Special Issue 2, March 2018

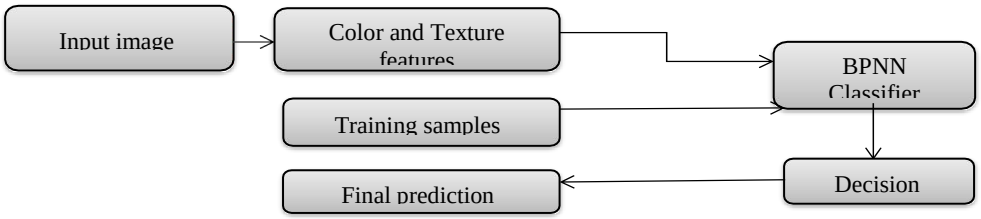


Fig 8: Block diagram of classification of melanoma using BPNN

III. EXPERIMENTAL RESULTS

The output is obtained by using different images such as skin cancer cells, normal skin, moles etc. The 8 output samples are as shown in fig. MATLAB is used to simulate and confirm if it is benign or malignant. If the image is affected by cancer the final image obtained highlights the border of the cancer affected area.

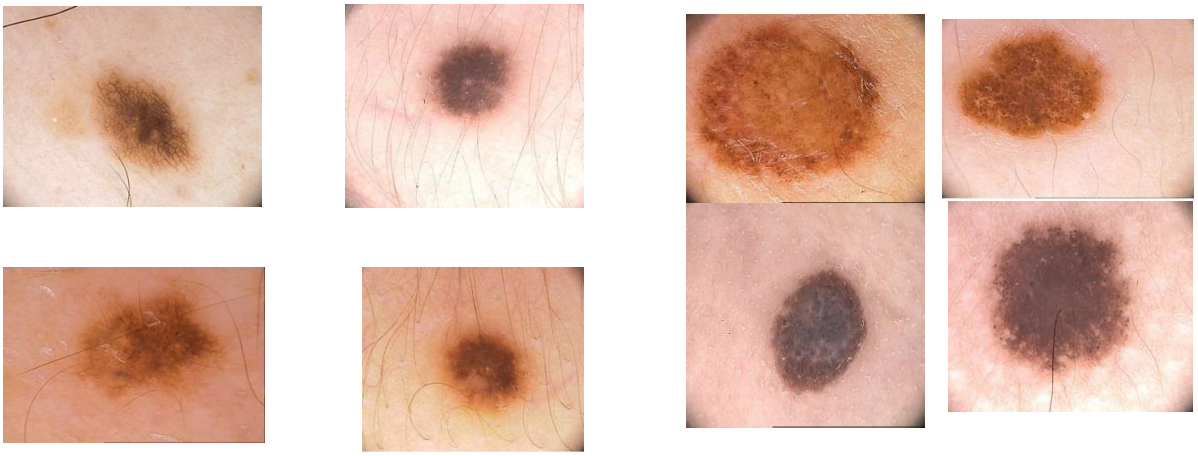


Fig 9: Dermoscopy image samples for BPNN training

IV. CONCLUSION AND FUTURE SCOPE

In this paper, a simple method for classifying melanocytic tumors as malignant or benign cancers has been described. The proposed system includes lesion extraction, feature description and finally classification based on the features extracted. Also, for better results these are practiced on a dermoscopy image. Feature dimensionality reduction was used to eliminate noisy features. For classification purpose, a neural network model is used. This can overcome the problems such as complexity, time consumption etc. often faced in fuzzy systems. The classification results by the designed model were shown to be more accurate than other existing methods. This project can be modified to predict the depth of the cancerous lesion. In future works, new tools like ANFIS can be used to analyzing 3D images of lesion.

REFERENCES

[1] T. Schindewolf, W. Stolz, R. Albert, et al., "Classification of melanocytic lesions with color and texture analysis using digital image processing" American Society of Cytology, vol. 15, no. 1, pp. 1-11, 1993.
[2] B.Kusumoputro,A. Ariyanto,"Neural network diagnosis of malignant skin cancers using principal component analysis as a preprocessor" Neural Networks Proceedings, 1998. IEEE World Congress on Computational Intelligence, pp. 310–315, 1998.
[3] M.E. Celebi, H.A. Kingravi, B. Uddin, et al., "A methodological approach to the classification of dermoscopy images" Computerized Medical Imaging and Graphics, vol. 31, no. 6, pp. 362-373, 2007.
[4] D.A. Clausi, "An analysis of co-occurrence texture statistics as a function of grey level quantization" Canadian Journal of remote sensing, vol. 28, no. 1, pp. 45-62, 2002.
[5] F.Y. Xie, S.Y. Qin, Z.G. Jiang, R.S. Meng, "PDE-based unsupervised repair of hair-occluded information in dermoscopy images of melanoma" Computerized Medical Imaging and Graphics, vol. 33, no. 4, pp. 275-282,2009.
[6] A. Ghodsi, "Dimensionality reduction a short tutorial" Department of Statistics and Actuarial Science, Univ. of Waterloo, Ontario, Canada, 2006.
[7] C. Barata, M. Ruela, M. Francisco, et al. "Two systems for the detection of melanomas in dermoscopy images using texture and color features" Systems Journal, IEEE, vol. 8, no. 3, pp. 965-979, 2014.