Detection and Classification of Skin Lesions in Dermoscopic Images

J.Junitha Persi Rajam[1], A.A.Haseena Thasneem[2]
[1]P.G Student, Electronics and Communication Engineering, National College Of Engineering, Tirunelveli,TamilNadu,India
[2]Asisstant Professor, Electronics and Communication Engineering, National College Of Engineering, Tirunelveli,TamilNadu,India
[1]junithapersirajam@gmail.com
[2]haseenajmal@gmail.com

Abstract: -Among all the types of skin cancer Malignant Melanoma (MM) is the most dangerous skin cancer. Skin cancer is commonly called as melanoma. There are two types in melanoma namely, Benign Melanoma and Malignant Melanoma. Both benign and malignant melanoma appears similar at the initial stages. So that it is difficult to differentiate both the melanomas, which is the main problem in the detection of skin cancer. Only an expert dermatologist will be able to provide an accurate classification as to which is benign and which is malignant. The standard approach in automatic dermoscopic image analysis consists of three stages: 1) image segmentation 2) feature extraction 3) lesion classification. Main advantage of this Computer Aided Diagnosis (CAD) is that only the patient confirmed with malignant melanoma need to undergo various painful diagnoses like Biopsy and others with benign melanoma need not. In this paper three segmentation techniques were applied to segment the lesion boundary. Accurate segmented output can be taken out by comparing three performance metrics namely sensitivity, accuracy and border error. Two classifiers are used to classify the types of melanoma, namely Neural Network(NN) and Support Vector Machine (SVM).

Keywords— Malignant Melanoma, Dermoscopy, Neural Network (NN), Support Vector Machine (SVM)

I. INTRODUCTION

Skin lesion is a part of the skin that has an abnormal growth or appearance compared to the skin around it. Some lesions have irregular boundaries and in some cases there is a smooth transition between the lesion and the skin. Malignant Melanoma is a type of skin cancer, if identified early it can be cured. Derma to scopy is the examination of skin lesions with a derma to scope. Dermoscopy images of pigmented lesions are taken at x10 magnification under lighting at a low angle of incidence. Accurate skin lesion segmentation from the background skin is important for diagnosis. This paper focuses on segmenting dermoscopic images for skin lesions. Comparison of the segmented outputs by calculating accuracy, sensitivity and border error is also performed. Then the accurate segmented output can be used for further classification.

II. SEGMENTATION METHODS

Adaptive Thresholding (AT) [1,2], K-Means clustering [3-9] and Neural Networks(NN)[10] techniques are used to segment the skin lesions in dermoscopic images.The segmented output of each technique is analysed using three metrics namely sensitivity, accuracy and border error. The following figure clearly illustrates the parameters namely TP,TN,FP and FN as well as the method of measurement.GT(Ground Truth) is obtained from dermatologist.

Fig (2.1) Measurement method based on four parameters

- TP (True Positive) - TP is the number of pixels that were classified both by GT and SA as lesion pixels.
- TN (True Negative) - TN is the number of pixels that were classified both by GT and SA as non-lesion pixels.
- FP (False Positive) - FP is the number of pixels where a non-lesion pixel was falsely classified as part of a lesion by SA.

- FN (False Negative) - FN is the number of pixels where a lesion Pixels was falsely classified as non-lesion by SA.

Table I segmentation performances on 10images

Techniques	Sensitivity	Accuracy	Border error
Adaptive Thresholding	82.89%	94.20%	36.38%
K-Means Clustering	88.82%	95.73%	26.26%
Neural Networks	91.80%	96.93%	15.3%

For each segmentation algorithm the median value are given.

III. CLASSIFICATION METHODS

As depicted in the above table, the NN algorithmhas an appropriate level of accuracy. This means that the NN segmentation algorithm segment the lesion boundary properly.

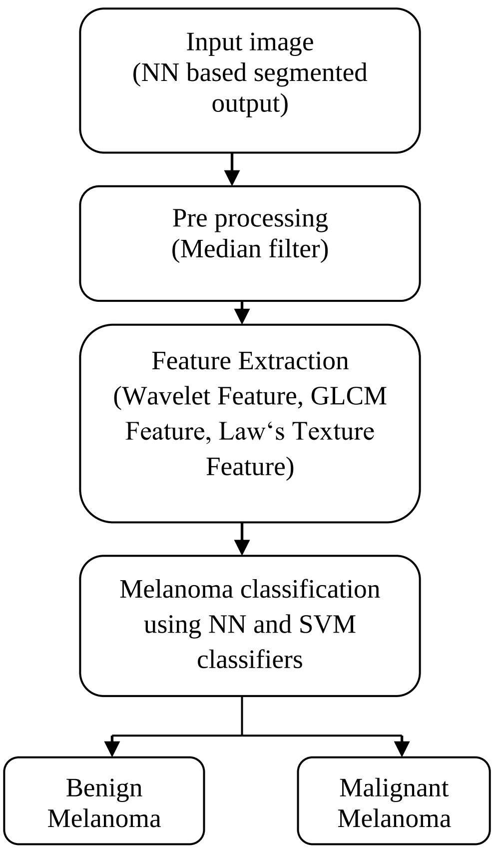

Fig 3.1 Block diagram representation

A. Pre-Processing

In this paper we have used Median filtering to perform noise removal. In median filtering, the neighbouring pixels are ranked according to brightness (intensity) and the median value becomes the new value for the central pixel. Median filters can do an admirable job of rejecting certain types of noise, in particular, hair removal in dermoscopic images.

B. Feature Extraction

Wavelet feature, GLCM (Gray Level Co-occurrence Matrix) feature and Law's Texture features are extracted from the NN based segmented image.

1) Wavelet feature

Wavelet feature selection algorithm is based on statistical dependence[11]. This algorithm is improved by combining the dependence between wavelet feature and the evaluation of individual feature component.

Wavelet Transforms enable the decomposition of the image into different frequency sub bands, similar to the way the human visual system operates. This property makes it especially suitable for the segmentation and classification of texture images.

For texture classification, absolute features need to be extracted to obtain a representation that is as discriminative as possible in the transform domain. A widely used wavelet feature is the energy of each wavelet sub band.

The 2D discrete wavelet transform is applied to the image. 2D-DWT performs single level two-dimensional wavelet decomposition. Then Low Low(LL), Low High (LH), High Low(HL) and High High (HH) values from the wavelet decomposed image is stored.

2) GLCM feature

The Gray Level Co-ocurrence Matrix (GLCM) method is a way of extracting second order statistical texture features from the image. GLCM method is widely used in many texture analyses.

A GLCM is a matrix where the number of rows and columns is equal to the number of gray levels, G, in the image[12].

The matrix element P $(i, j \mid \Delta x, \Delta y)$ is the relative frequency with which two pixels, separated by a pixel distance $(\Delta x, \Delta y)$, occur within a given neighborhood, one with intensity 'i' and the other with intensity 'j'.

The matrix element P $(i, j \mid d, \theta)$ contains the second order statistical probability values for changes between gray levels 'i' and 'j' at a particular displacement distance d and at a particular angle (θ).

Using a large number of intensity levels G implies storing a lot of temporary data, i.e. a G × G matrix for each combination of (Δx, Δy) or (d, ϴ). Due to the large dimensionality of the image, the GLCM's are sensitive to the size of the texture samples on which they are estimated. Thus, the number of gray levels can be reduced.

Entropy

The entropy of a dermoscopic image can be defined as a measure of the uncertainty associated with a random variable.

Entropy quantifies, in the sense of an expected value, the information contained in an image. Entropy shows the amount of information of the image that is needed for the image compression.

Entropy measures the loss of information or message in a transmitted signal and also measures the image information.

$$\text{Entropy} = \sum_i \sum_j C_{ij} \log C_{ij}$$

C - GLCM matrix

Energy

Angular second moment is also known as Uniformity or Energy. It is the sum of squares of entries in the GLCM. Angular second moment measures the image homogeneity.

Angular second moment is high when image has very good homogeneity or when pixels are very similar.

$$\text{Energy} = \sum_i \sum_j C^2_{ij}$$

Homogenity

Homogenity can measure the closeness of the distribution of elements in GLCM to the GLCM diagonal.

$$\text{Homogenity} = \sum_i \sum_j \frac{C_{ij}}{1+|i-j|}$$

These are the Gray Level Co-occurence Matrix(GLCM) features extracted from the input image during the training stage.

3) Law's Texture feature

Laws observed that certain gradient operators such as Laplacian and Sobel operators accentuated the underlying microstructure of texture within the dermoscopic image[13]. This is the basis for a feature extraction scheme based on a series of pixel impulse response arrays obtained from combinations of 1-D vectors shown in Figure (3.2).

Each 1-D array is associated with an underlying microstructure and labeled using an acronym accordingly. The arrays are convolved with other arrays in a combinatorial manner to generate a total of 25 masks, typically labeled as L5L5 for the mask resulting from the convolution of the two L5 arrays.

Level L5 = [1 4 6 4 1]
Edge E5 = [-1 -2 0 2 1]
Spot S5 = [-1 0 2 0 -1]
Wave W5 = [-1 2 0 -2 1]
Ripple R5 = [1 -4 6 -4 1]

Figure(3.2). Five 1-D arrays identified by Laws

These masks are subsequently convolved with a texture field to accentuate its microstructure giving an image from which the energy of the microstructure arrays is measured together with other statistics.

C.NN classifier

The NN based classification consists of two stages

- Training stage
- Testing stage

Various back propagation algorithms are used to train the NN. In this paper the Levenberg-Marquardt back propagation algorithm is used to train the feed forward neural network. The Levenberg-Marquardt algorithm is very sensitive to the initial network weighs.

A classifier classifies the given datasets into benign melanoma and malignant melanoma. Here a computer based classifier implemented in MATLAB software is used for classify the melanoma[14].

The above extracted features are given as input to the NN. The activation function used is tan sigmoid function. The output of the network will be 0 or 1. Zero indicates benign melanoma and one indicates malignant melanoma condition.

ANN classifier is designed in MATLAB software. During each epoch, the weights are updated so that error between desired output and actual output is minimum. Database consists of fiftyimages for classification.

Both benign and malignant melanoma dermoscopic images are used to train the network. At the testing stage, 16 benign images and 34 malignant dermoscopic images are used to the test the network.

During the training stage of NN, both benign and malignant features are saved as mat file. In the testing stage the dermoscopy image is given as the input to the NN. The network extracts the features from the input image and comparing with the mat file which is saved in training stage. If the features are similar with the benign melanoma means the network produce 0 as the output. That is benign melanoma. If the extracted features are similar with malignant means the network produce 1 as the output. That is malignant melanoma.

Figure (3.3) Benign image

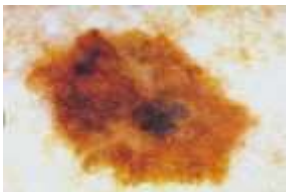

Figure (3.4) Malignant image

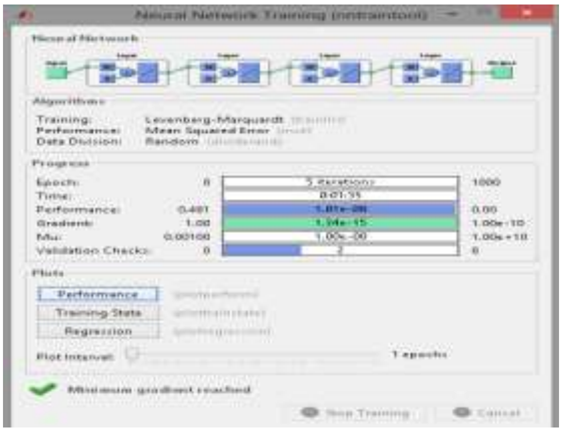

Figure (3.5) ANN training in MATLAB

D. SVM classifier

Support Vector Machines (SVMs) are supervised learning models with associated learning algorithms that analyze data and also recognize patterns and also used for classification and regression analysis.

SVM classifier takes a set of input data and predicts, for each given input, which of two possible classes forms the output, building it a non-probabilisticbinarylinear classifier [15].

Given a set of training examples to the classifier, each marked as belonging to one of two categories, an SVM training algorithm builds a representation that assigns new examples into one category or the other.

SVM model is a representation of the examples as points in space, mapped. Theexamples of separate categories are divided by a clear gap that is as wide as possible.

New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall on.

To perform linear classification, SVMs can efficiently perform non-linear classification using what is called the kernel trick, implicitly mapping their inputs into high-dimensional feature spaces.

An *n*-dimensional pattern (object) x has *n* coordinates, $x=(x_1, x_2, \ldots, x_n)$, where each x_i is a real number, $x_i \in R$ for $i = 1, 2, \ldots, n$.

Each pattern x_j belongs to a class $y_j \in \{-1, +1\}$. Consider a training set T of m patterns together with their classes, $T=\{(x_1, y_1), (x_2, y_2), \ldots, (x_m, y_m)\}$. Consider a dot product space S, in which the patterns x are embedded, $x_1, x_2, \ldots, x_m \in S$.

Any hyperplane in the space S can be written as

$$\{\mathbf{x} \in S \mid \mathbf{w} \cdot \mathbf{x} + b = 0\}, \quad \mathbf{w} \in S, b \in R$$

The dot product w•x is defined by:

$$\mathbf{w} \cdot \mathbf{x} = \sum_{i=1}^{n} w_i x_i$$

A training set of patterns is linearly separable if it has at least one linear classifier defined by the pair (w, *b*) which correctly classifies all training patterns (Figure 3.7).

The linear classifier is represented by the hyperplane H (w•x+*b*=0) and defines a region for class 0 patterns (w•x+*b*>0) and another region for class 1 patterns (w•x+*b*<0).

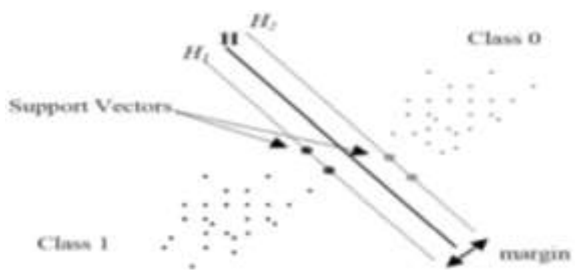

Figure 3.7 Linear classifier defined by the hyper plane

After training, the linear classifier is ready to predict the class membership for new patterns, different from those used in training.

Therefore, the classification of new patterns depends on the sign of the expression **w•x**+*b*.

Both benign and malignant melanoma dermoscopic images are used to train the network. At the testing stage, 16 benign images and 34 malignant dermoscopic images are used to the test the network.

E. Performance evaluation

Each classifiers performance is analysed using three metrics namely Senitivity, Specificity, Accuracy.

$$\text{Sensitivity} = \frac{TP}{TP+FN}$$

$$\text{Specificity} = \frac{TN}{TN+FP}$$

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}$$

Where,

TN (True Negative) – Benign images are correctly classified as Benign.

TP (True Positive) – Malignant images are correctly classified as Malignant.

FP (False Positive) – Benign images are wrongly classified as Malignant.

FN (False Negative) – Malignant images are wrongly classified as Benign.

The database consists of 16 benign images and 34 malignant images.

Classifiers	TN	TP	FP	FN	SV %	AC %	SP %
NN	12	31	4	3	91	86	75
SVM	13	34	3	0	100	94	81

Comparison of two classifiers

SV – Sensitivity

AC – Accuracy

SP – Specificity

IV. CONCLUSION

Early detection of malignant melanoma using computer based techniques is more efficient than the conventional Biopsy methods. Three segmentation techniques were applied to the dermoscopic images to segment the skin lesions and evaluated with 3 different metrics, namely sensitivity, accuracy and border error. Segmentation performance shows that Neural Network based lesion segmentation has high sensitivity, accuracy and less border error. Hence it can be said that the NN algorithm segments the lesion boundary properly. NN based segmented output is further classified using NN and SVM. Performance evaluation shows that the SVM based classification has 100% sensitivity.

REFERENCES

Images are downloaded from the following website
www.skincancer.org
http://www.mmmp.org/MMMP/import.mmmp?page=dermoscopy.mmmp.
http://www.dermnetnz.org/doctors/dermoscopy-course/malignant-melanoma.html.
http://www.eastmed.co.nz/melanoma-mole-photos-images-pictures.php.

[1] M. Silveria, J. C. Nascimento, J. S. Marques, A. R. S. Marcal, T. Mendonca, S. Yamauchi, J. Maeda, and J. Rozeira, "*Comparison of segmentation methods for melanoma diagnosis in dermoscopy images," IEEE Trans. Signal Process.,* vol. 3, no. 1, pp. 35–45, Mar. 2009.

[2] Maglogiannis, S. Pavlopoulos, D. Koutsouris, "*An Integrated Computer Supported Acquisition,*

Handling, and Characterization System for Pigmented Skin Lesions in Dermatological Images," IEEE transactions on information technology in biomedicine, vol. 9, no. 1, march 2005.

[3] Wei Sun and Yaonan Wang, "*Segmentation Method of MRI Using Fuzzy Gaussian Basis Neural Network Neural Information Processing", Letters and Reviews*, Vol.8, No.2, August 2005, pp.19-24.

[4] Ibrahiem M. M. El Emary, "*On the Application of Artificial Neural Networks in Analyzing and Classifying the Human Chromosomes", Journal of Computer Science*, vol.2(1), 2006, pp.72-75.

[5] Bouchet A, Pastore J and Ballarin V, "*Segmentation of Medical Images using Fuzzy Mathematical Morphology", JCS and T,* Vol.7, No.3, October 2007, pp.256-262.

[6] Wilson, E., D.W. Sutter, and R.W. *MahThe Application of Watershed in Segmentation of Overlapping Cells.*. 2004. Honolulu, HI, United States: Acta Press, Anaheim, CA, United States.

[7] A.M Uso,F.Pla,P.GSevila," *Unsupervised Image Segmentation using a Heirarchical Clustering Selection Process".Structural Syntactic and Statistical Pattern Recognition* Vol 4109,pp.799-807,2006.

[8] A.Z Arifin,A.Asano,*"Image Segmentation by histogram thresholding using hierarchical cluster analysis"Pattern Recognition Letters*,Vol.27,no.13,pp. 1515-1521,2006.

[9] H.S.Prasanthaet. *(IJCSE) International Journal on Computer Science and Engineering* Vol. 02, No. 04, 2010, 1209-1218 Medical Image Segmentation.

[10] Dr. J. Abdul Jaleel, SibiSalim, Aswin.R.B Dept. of Electrical & Electronics Engineering, TKM College of Engineering, Kollam, Kerala, India. *Artificial Neural Network Based Detection of Skin Cancer International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering* Vol. 1, Issue 3, September 2012 .

[11] Ke Huang and SelinAviyente, Member IEEE,*Wavelet Feature Selection for Image Classification*. IEEE Transactions On Image Processing, Vol. 17, No. 9, September 2008.

[12] P. Mohanaiah, P. Sathyanarayana,L. Guru*Kumar Image Texture Feature Extraction Using GLCM approach. International Journal of Scientific and Research Publications,* Volume 3, Issue 5, May 2013.

[13] P. Shantha Kumar and P. Ganesh Kumar *"Performance Analysis Of Brain Tumor Diagnosis Based On Soft Computing Techniques*" American Journal of Applied Sciences 11 (2): 329-336, 2014.

[14] Aswin.R.B, J. Abdul Jaleel, SibiSalim*," Implementation of ANN Classifier using MATLAB for Skin Cancer Detection" International Conference on Mobility in Computing-ICMiC13, Organized by Mar Baselios College of Engineering and Technology during December 17-18, 2013 at Trivandrum, Kerala, India, pg.87 – 94.*

[15] Daljit Singh, KamaljeetKaur,*"Classification of Abnormalities in Brain MRI Images Using GLCM, PCA and SVM"* International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-1, Issue-6, August 2012.