

IMPLEMENTATION OF SUPERVISED LEARNING FOR MELANOMA DETECTION USING IMAGE PROCESSING

Siddiq Iqbal¹, Sophia.M², Divyashree.J.A³, Mallikarjun Mundas⁴, Vidya.R⁵

¹Dept. of Telecommunication, B.M.S. Institute of Technology & Management, Bangalore

²Dept. of Telecommunication, B.M.S. Institute of Technology & Management, Bangalore

³Dept. of Telecommunication, B.M.S. Institute of Technology & Management, Bangalore

⁴Dept. of Telecommunication, B.M.S. Institute of Technology & Management, Bangalore

⁵Dept. of Telecommunication, B.M.S. Institute of Technology & Management, Bangalore

Abstract

Among the different types of skin cancers, Melanoma is one of the most threatening type of cancer. This cancer is most often caused due to over exposure to ultraviolet radiation from the sun which causes unrepaired DNA damage to skin cells which further develops into cancerous tumours. This unrepaired damage to the skin usually affects the melanocytes, which are skin cells containing a pigment called melanin which is responsible for the colour of the skin, hence the name melanoma. If melanoma is recognised in the early stages it is proven to be curable. If not, the cancer advances and spreads to all other parts of the body and becomes incurable leading to death. One of the traditional methods of analysing melanoma is biopsy, which is a painful and time consuming process. To overcome this we have implemented a computer aided method for automatic melanoma detection and classification of Dermoscopic skin images with the help of Digital Image Processing and Artificial Intelligence. This paper proposes that using artificial intelligence for Melanoma detection increases the accuracy of classification.

Keywords- Melanoma, ultraviolet, tumor, biopsy, asymmetry, computer aided, image processing, artificial intelligence.

1. INTRODUCTION

Skin is the largest organ and is the outermost layer of the human body. It consists of has several layers of ectodermal tissues and guards the underlying muscles, organs etc. It forms a defensive outer covering against external bodies and it also plays a role in regulating the body mechanisms. The skin consists of melanin which is provided by the melanocytes. These cells are responsible for the absorption of the dangerous Ultraviolet Radiation entering the earth's atmosphere. This Radiation in excessive amounts can damage the DNA repair enzymes present in the skin that reverses the effect of UV damage, which may result in skin Cancer.

Cancer can be of two types, benign or malignant. Benign tumors are the non-cancerous tumors. Cells in benign tumors do not spread to other parts of the body and can be easily curable. Malignant tumors are cancerous and are made up of cells that grow out of control. Cells in these tumors can invade nearby tissues and spread to other parts of the body. In most cases Malignant Melanoma results in the death of a person.

The most dangerous and common type of Skin cancer is Melanoma. It is a cancer that begins in the skin cells called melanocytes. Melanoma can develop anywhere on the skin, but they are more likely to start on the trunk (chest and back) in men and on the legs in women. The neck and face are other common sites for the growth of melanoma.

Melanoma accounts for about 77% of skin cancer related deaths. However, it is very important that melanoma is recognised in the early stages. It has been proven that if it is diagnosed in the early stages it is curable. Adding to the importance of melanoma being detected in the early stages is that the treatment is simple and cost effective. If it is not detected in the early stages, the cancer advances and spreads to all the other parts of the body and becomes incurable leading to death. There is a great need for an automatic detection system for skin cancer in which are non-invasive diagnosis techniques are used which provides quantitative information about the lesion. It is more advantageous than the traditional biopsy method where a layer of skin is scrapped off which is then subjected to many laboratory tests for diagnosis which are both painful and time consuming. In this paper we propose a Computer Aided Automatic Diagnosis for classifying and identifying benign and malignant melanoma. The methodology proposed for automatic diagnosis of melanoma makes use of Digital Image Processing techniques and Artificial Intelligence. The procedure includes- obtaining the epiluminescence image of the skin lesion, filtering the image to remove hair and noise, followed by segmentation to obtain the Region Of Interest. From the obtained Region Of Interest unique features of the skin lesion are extracted using various feature extraction techniques. These features are used in the Neural Network for the purpose of classification. The Neural Network used is a Feed Forward Back Propagation system consisting of three layers- an input layer, a hidden layer and an output layer.

2. SYSTEM DESIGN

The design of a computer aided system for melanoma detection and classification based on Image Processing and Artificial Intelligence include the following stages:

2.1 Stage 1: Image Acquisition

Images of the affected skin area is captured using an imaging technique with the help of an Epiluminescence microscope. This device is also known as Dermoscopy, Dermatoscopy or Epiluminescence Light Microscopy (ELM). The image acquired with this device is a high resolution digital image known as a Dermoscopic image (Fig 1). An oil immersion is placed between the skin and the optics; and the lens of a microscope is placed in direct contact with the skin. The device with a lighting at an angle of 450 brings out the pigmented structure of the epidermis which is not visible to the naked eye. It also enhances the visibility of deeper layers of the skin enabling the depth of infection to be examined.

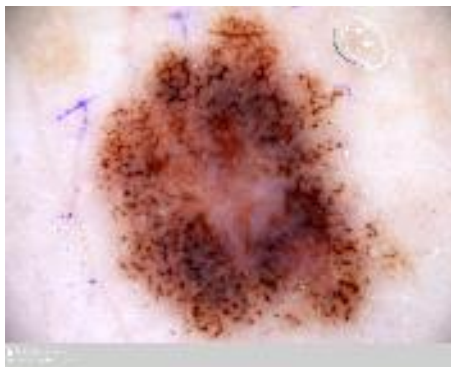


Fig 1. Original Image

2.2 Stage 2: Image Processing

Human skin surface has various in skin tones and is accumulated with hair, scars, bubbles etc. Due to this, feature from the captured image may be of less accuracy which in turn may reduce the accuracy of classification. Image processing is required to remove the hair and other noise in the captured image. It includes two procedures: Pre-processing and Post-processing.

Preprocessing includes removal of hair, air bubbles etc present on the skin and removal of noise by filtering. The filtering is performed by an Averaging filter. This filtering operation as shown in Fig 2 is based on correlation. Pre-processing also includes resizing the image so as to access the affected skin lesion accurately. Post processing includes Image Enhancement and Contrast Adjustment. This plays a major role in increasing the accuracy of classification.

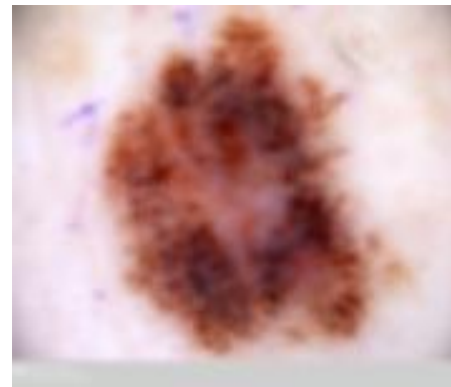


Fig 2. Filtered Image

2.3 Stage 3: Image Segmentation

Image segmentation is a crucial stage in obtaining the accurate results in the subsequent stages. It is carried out in order to obtain the Region Of Interest from the processed image. [3]. The Image after being subjected to processing consists of cancerous region and healthy skin only. If the entire image is taken into consideration for feature extraction there would be a compromise in the overall efficiency of the system. Hence features only from the cancerous region is required. Thereby, the Region Of Interest which is the cancerous region has to be separated from the background skin.

In this paper, the segmentation technique used is Otsu's Segmentation. It is a simple and powerful technique based on the thresholding principle in which a gray level image is converted to a binary image.

The Otsu algorithm assumes the image that is subjected to threshold to have two classes of pixels as shown in Fig 3 or bi-modal histogram. The two classes of pixels are foreground and background pixels. The algorithm calculates the optimum threshold that separates the two classes in such a way that it has minimum variance.

In Otsu's method the threshold that minimizes the intra-class variance (the variance within the class) is selected by trial and error method. Within class variance is defined as a weighted sum of variances of the two classes:

$$\sigma_w^2(t) = \omega_1(t) \sigma_1^2(t) + \omega_2(t) \sigma_2^2(t) \dots \dots \dots (1)$$

Weights ω_i are the probabilities of the two classes separated by a threshold t and σ_i^2 are variances of these classes.

Otsu shows that minimizing the intra-class variance is the same as maximizing inter-class variance (between class variance).

$$\begin{aligned} \sigma_b^2(t) &= \sigma^2 - \sigma_w^2(t) \dots \dots \dots (2) \\ &= \omega_1(t) \omega_2(t) [\mu_1(t) - \mu_2(t)] \end{aligned}$$

The class probability $\omega_1(t)$ is computed from the histogram as t:

$$\omega_1(t) = \sum_0^t p(i) \dots \dots \dots (3)$$

While the class mean $\mu_1(t)$ is:

$$\mu_1(t) = [\sum_0^t p(i) \times i] / \omega_1 \dots \dots \dots (4)$$

Where $x(i)$ is the value at the center of the i th histogram bin. Similarly, you can compute $\omega_2(t)$.

And μ_2 on the right-hand side of the histogram for bins greater than t.



Fig 3. Black and White

2.4 Stage4: Feature Extraction

Feature extraction is a method where unique features of the skin lesion are extracted from the Region Of Interest (ROI) to differentiate malignant melanoma from benign melanoma. By feature extraction a large data is reduced down to a relevant narrow set which makes the classification easier. Melanoma can be classified into malignant and benign based on features such as color, brightness or luminance, texture and shape. Malignant melanoma has a mix of red, green, and dark coloration whereas benign lesion has uniform color pattern; the other features that differentiates malignant melanoma from benign is asymmetry the lesion, irregular border and size.

Asymmetry

Asymmetry is calculated, firstly by dividing the image over its closest line of symmetry (i.e. centroid as shown in Fig 4). Secondly by finding the area of the non overlapping sections to obtain the difference between these areas. The obtained result is then divided by the total area.

The mathematical expression used to calculate the percentage of asymmetry is:-

$$\text{Asymmetry} = (\Delta P / P) * 100 \dots \dots \dots (5)$$

Where, ΔP = Pixel difference
 P = Total Pixel count of lesion

The asymmetric value is then given as an input to the neural network.



Fig 4. Asymmetry calculation

Border

In case of melanoma the borders of the lesion are irregular, ragged, notched, or blurred. The edge or the border of the lesion is first recognized and then fetched from the image excluding the inner and outer parts of the mole (Fig 5).

The border irregularity is calculated using

$$CI = [(\text{perimeter})^2 / 4\pi A] \dots \dots \dots (6)$$

Where, A= Area of the lesion, $\pi = 22/7$

Compactness index (CI), estimates the roundness of a 2D object.

This value is also given as an input to the neural network.



Fig 5. Border detection

Colour

The colour of the lesion varies over different shades of brown, black, red, white or blue. The sum of the amount of colour is given as an input to the neural network.

Major Axis and Minor Axis

The length of the major and minor axis of the skin lesion are extracted and are fed as an input to the neural network. These two features help in determining the size of the lesion.

2.5 Stage 5: Neural Network Classification

The Neural Network is a massive network made up of artificial neurons. This network consists of three layers namely, the input layer, the hidden layer and the output layer. The neural network is trained to perform useful

computation through a process of learning. In order to do so the network employs a massive interconnection between the neurons, which are also known as processing units.

Classification of melanoma into malignant and benign is done with the help of a classifier. The classifier used here is an Artificial Neural Network which is based on Feed-Forward Back Propagation algorithm to train the network. The features that have been extracted are fed as inputs to the input layer neurons. In this paper the neural network consists of 5 input neurons and 10 hidden neurons (Fig 6)

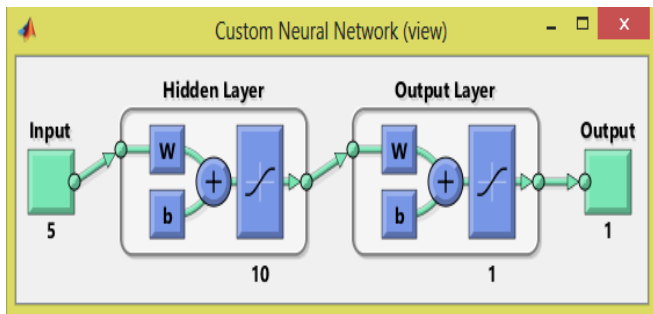


Fig 6. Structure of the Neural System

Initially the weights of the hidden layer are chosen randomly which are later adjusted in the training phase based on the error output. In this algorithm, the signal flows in the forward direction and an output is generated based on the value of the weights in the hidden layer. If the output generated does not match the desired output for which the network was trained then an error signal is generated.

$$\text{Error} = \text{Desired Output} - \text{Actual Output}$$

This error signal propagates in the backward direction to readjust the weights till the error is minimized and accuracy is improved. The re-adjustment of the weights is based on the gradient of the error curve, in the direction of local minima. Thus this method is reliable in classification of skin lesions.

3. RESULTS

25 images were taken into consideration whose results were known previously. All the images considered were high resolution Dermoscopic images. These 25 images were used for the purpose of training. 25 images were taken again for the purpose of testing. Out of these 25 images, 20 images were classified as benign, 5 were classified as malignant and there were 2 miscalculations, resulting in a 8% error. Hence this paper proposes a computer aided method for skin cancer detection with an accuracy of 92%.

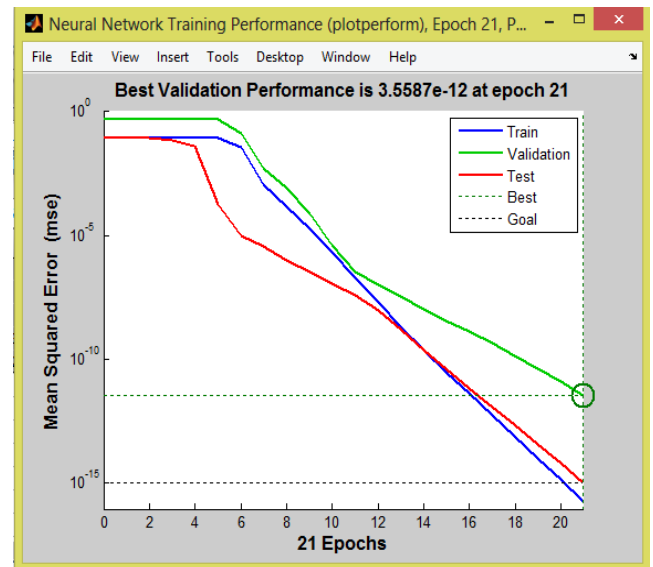


Fig 7. Performance Plot

The performance plot (Fig 7) hold informations regarding the training of the network and it determines if the training process need to be changed. In addition to this it also specifies whether the data set has to be changed. From the obtained performance plot we can conclude that the training need not be changed.

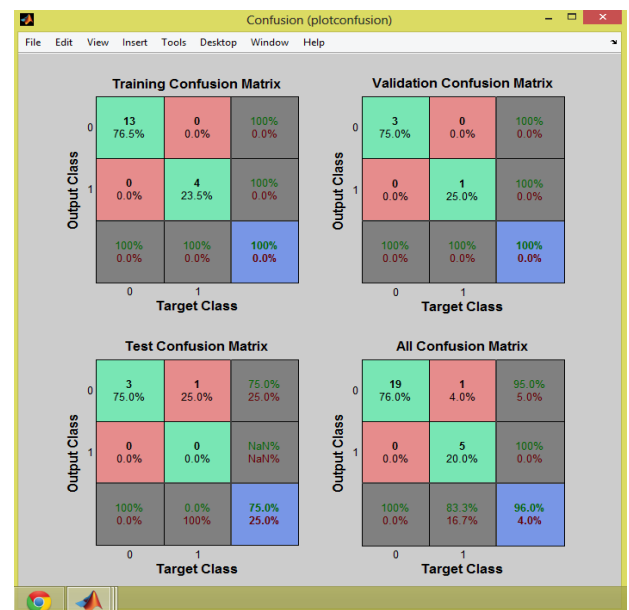


Fig 8. Confusion Matrix

The ANN Classifier classified the given data into benign and malignant melanoma. The confusion matrix (Fig 8) of the test images has an accuracy of 96%.

4. CONCLUSION

In this paper, we have implemented Image Processing and Artificial Intelligence for cancer Detection. The Neural Networks were trained based on the Asymmetry, Border, Colour, Major Axis and Minor Axis values with the help of MATLAB. We have chosen this because of its effectiveness in automatic detection of melanoma. This Methodology has

got an accuracy of 92%. Hence this paper shows that by varying the Image processing techniques and training algorithms of Artificial Neural Network, the accuracy of the system has been improved and the images are classified as cancerous and non-cancerous.

REFERENCES

- [1] Aswin R.B, J. Abdul Jaleel, Sibi Salim, "Hybrid Genetic Algorithm Artificial Neural Classifier for Skin Cancer Detection" IEEE 2014
- [2] Ms. H. R. Mhaske, Mrs. D. A. Phalke, "Melanoma Skin Cancer Detection and Classification Based On Supervised and Unsupervised Learning", IEEE 2013.
- [3] J Abdul Jaleel, Sibi Salim, Aswin.R.B "Computer Aided Detection of Skin Cancer",IEEE 2013
- [4] Arushi Bhardwaj, Dr. J.S Bhatia. "An Image Segmentation Method for Early Detection and Analysis of Melanoma", IOSR Journal of Dental and Medical Sciences (IOSR-JDMS) e-ISSN: 2279-0853, p-ISSN: 2279-0861. Volume 13, Issue 10 Ver. V (Oct. 2014).
- [5] "Analysis of Skin Cancer Using Fuzzy and Wavelet Technique – Review & Proposed New Algorithm", International Journal of Engineering Trends and Technology (IJETT) – Volume 4 Issue 6- June 2013
- [6] T Y Satheesha, D Satyanarayana, M N Giriprasad. "A Pixel Interpolation Technique for curved hair removal in skin images to support melanoma detection". Journal of Theoretical and Applied Information Technology 31st December 2014.
- [7] J. Breckling, Ed., The Analysis of Directional Time Series: Applications to Wind Speed and Direction, ser. Lecture Notes in Statistics. Berlin, Germany: Springer, 1989, vol. 61.
- [8] Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification, IEEE Std. 802.11, 1997.