|| Volume 6 || Issue 1 || January 2021 || ISSN (Online) 2456-0774
INTERNATIONAL JOURNAL OF ADVANCE SCIENTIFIC RESEARCH
AND ENGINEERING TRENDS

SKIN CANCER DETECTION USING IMAGE PROCESSING

Prajakta Sanjay Gaikwad[1], Amit Sanjay Shete[2], Manasi Hanumantrao Patil[3], Aniket Sanjay Rane[4]

Computer Department, NBN Sinhgad School Of Engineering Pune, India [1,2,3,4]

gaikwadprajakta244@gmail.com, amit.shete98@gmail.com, patilmanasih.pune@gmail.com, aniketrane2000@gmail.com

---

Abstract:- Technology is playing important role in day-to-day life. Image processing plays a crucial role in medical field. Melanoma skin cancer is dangerous and harmful for human beings. Detection of Melanoma skin cancer in early stages is very much essential for the patient because this disease directly lead to the death of a person. Melanoma skin cancer is completely curable if it is detected at early stage. Due to irregularities in shapes and edges clinical diagnosis is difficult. This paper proposes the idea.0 for early detection of melanoma using Multiclass support vector machine (MSVM).

Keyword -*Support Vector Machine*

I INTRODUCTION

Skin cancer is increasing day by day due to large amount of UV radiation. There are basic three types of skin can- cer Melanoma ,Basal Cell Carcinoma (BCC) and Squamous Cell Carcinoma (SCC). Melanoma skin cancer is developed in melanocytes skin cells which are directly responsible to produce melanin. When the pigment producing cells becomes cancerous Melanoma occurs. If melonoma is detected at early stage then it is curable. Melanomas mostly occurs in mouth, eye and intestines, where for women's it mostly occurs in legs. The melanoma can be easily identified by the change in colour, itchiness of skin, size of moles, irregularity in edges. To distinguish and flawlessly identify lesion area is the most important thing from a doctor's point of view. If it is not correctly identified in early stages it may lead to advanced stages of cancer. Therefore, it is an important task for dermatologist to dectect it in early stages. Eventhough research has been done on detection of Melanoma skin cancer but still issue exists for higher accuracy for the detection and classification of Melanoma skin cancer. In the process of skin cancer screening, clinicians usually detect the suspected lesion region by visual checkup that is highly dependent on observer skills and is likely to have human error. The main aim of the proposed research work is to develop a detection and classification framework for detecting cancer in early stages. The proposed system is a automated based Classification where the image is classified by matching. The proposed system contains image acquisition, color space conversion, feature extraction, classifier algorithm and segmentation.

II. SYSTEM MODEL

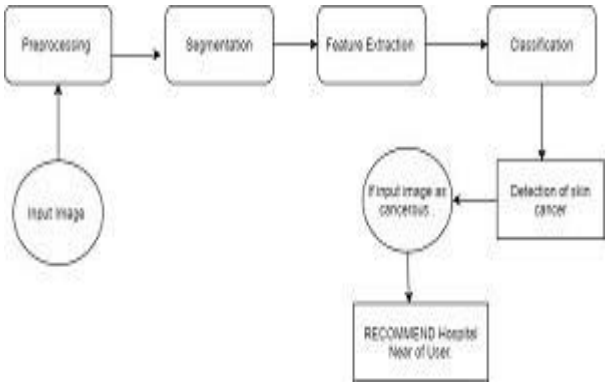

Fig. 1. Block diagram of proposed method.

A. Input Image

Input of skin cancerous image is an important procedure in automated skin lesion analysis system. Because the input image will undergo the rest of the process in the system such as feature extraction, MCSVM training, Segmentation and Area calculation which may not be achievable or the obtained result will not be a reasonable one. Hence the input image should have healthy in resolution and clarity.

2..Pre-Processing

The dataset used in this system is obtained from different sources. The sizes of these images are non-standard. This stage consists of 2 steps. The first step is to resize the images to have fixed width but variable height[1]. The second step is to eliminate the background noise and the undesirable objects from the images such as fine hair and air bubbles [9, 13]. Such objects are not a part of the cancerous cell where they might reduce the accuracy of border detection or segmentation, and increase the computation time. In this system two filters were used; Winner and Median filters. Winner filter was used to eliminate the undesirable objects and reduce noise. While, Median filter which is standard nonlinear signal processing technique developed by Turkey [9,14] was used for smoothing images, and suppressing the noise by removing the outliers that are the extreme pixel values. Before the system processes the color digital images to be segmented it is converted to the grey level.

C. Segmentation

Segmentation is nothing but a partitioning of image so that the region of interest is obtained[1]. In Thresholding technique peak value for skin and peak value for lesion [3] is determined and then threshold is selected in between these two peak points[3]. The pixel intensity values those are greater than threshold value are set as 0 while intensity values those are less than threshold value are set as 1.

Fig. 2. Segmented image.

For segmentation of skin cancer images region growing [4] and merging techniques is used but the technique called statistical region merging is not giving good results for the Back Propagation Neural network[5]. In statistical region merging regions are formed having similar characteristics 0 neighboring pixel and then it get merged[1]. Melanoma skin cancer is also classified by Multi Layer Perceptron where accuracy is 78% It has been observed from literature survey that for dermatoscopic images efficiency and accuracy of results is higher [6] and it is greater by about 1% to 3%. The segmented image is shown in figure 2

D. Feature Extraction

This part of algorithm computes the features related to shape, texture and color from the segment mole. Using these three features categories, this approach collects ABCD data of lesion mole where A defines the asymmetry, B corresponds to border Evolving, C is related to color variation and D is the diameter of the mole These features are represented with v, u,n, symbols respectively. The shape features includes image diameter, image compactness, ulnar variance and the image asymmetry. Compactness refers to the ratio of the object perimeter to its area. Ulnar variance is the measure of relative length of articular surfaces of some particular radius and the image asymmetry is the measure of asymmetry of the cancerous mole. The texture features include coarseness and Gray Level Co-occurrence Matrix (GLCM). Coarseness is the measure of different angle texture representation. GLCM is a histogram of co-occurring grayscale values for given offset over the image and provides the feature discriminatory attributes. It consists of different parameter which includes mean, correlation, homogeneity, contrast, energy, dissimilarity and kurtosis. The color related features include variance, skewness, and entropy. Variance is the measure of dispersion in the image. Entropy is the proportion of randomness and skewness is the measure of distributed asymmetry.

E. Classification

Classification of images into cancer type and skin type or non cancer type is done using supervised and unsupervised learning. In supervised learning the trained data is provided in order to predict the result or class of data but in case of unsupervised learning known data is not provided the classifiers does the work itself to find out in which class or category the data belongs. [n classification based on rules the data is get categorized for different classes. [n this paper classification results for different classifiers are compared.

• Neural Network:- Back propagation Neural Network is used for classification in which 20 neurons are present in middle layer. Feature vector is applied to the Neural Network and after applying feature vector activation function calculates the category for the input. In Neural Network classification, class labeling is done for input data. Class labeling denotes that in which category input data exist. Results are obtained by weight and bias values. Initially the weight values are randomly selected and the neural network adjusts the weight value for getting min- imum errors. Training is performed in Back Propagation Neural Network until the minimum gradient. is reached. The gradient is calculated accordingly and results are obtained. In training the delta rule is applied on a training set and then coefficients are found. In testing the images are checked for different type of patterns or categories. For testing purpose the back propagation neural network is simulated and after simulation of network output value is calculated. Based on output value the category that is whether the image is belonging into cancer type or skin type is determined [6]. After training of network the class labeling is done for input images. [n Neural Network the feature vector is applied accordingly and activation function is calculated. [n detection of Melanoma skin cancer two categories are formed one is for cancer type and another is for skin type. For cancer images labeling is done with 0 and for skin images labeling is done with 1. After training of neural network it is simulated and value is generated as result. If the value is less than or equal to 0.5 then the identified category is of cancer type and if the value after simulation is greater than 0.5 then the identified category is of skin type. After categorization the overall accuracy in terms of percentage is calculated.

• **K-Means Clustering Algorithm**:- K-means classification is based on unsupervised learning. In unsupervised learn- ing no one is there to instruct and monitor the results. In k-means classification the numbers of clusters are formed. These clusters are nothing but a collection of data points. Each cluster classifies the data point for the different categories. After getting the separate clusters the accuracy is calculated. The features are obtained using parameters like as mean, median, standard deviation, minimum, variance and maximum. These features are given as an input to the k-means classification Here two clusters are formed one cluster represent maximum possibility of cancer images and other cluster represent maximum possibility for skin images. K-means algorithm is based on the measures in term of Euclidian distance. The data points are divided into two clusters. Initially random center points for the clusters are considered. The distance between data points and centroid is calculated. Data points those are having minimum distance towards no more movement for data points. The aim of k-means algorithm is to reduce the error function as

$$E(U) = \sum_{m=1}^{k} \sum_{n=1}^{c} \left(x_m - y_m \right)^2$$

where

kil = Number of data points in cluster m. k = number of centers for clusters.

K means algorithm is as follows:

• Let D denotes set of data points and P denotes set of cluster centers. Initially randomly select the centers.

• Calculate Distance between data points and centroid using formula

$$P_j = (1 / k) \sum_{n=1}^{k} D_j$$

• Data points having minimum distance to that centroid will belong to that cluster.

• Recalculate the new centroid for the clusters.

• Recalculate distance between data points and new centroid.

• Repeat step 3 until each data point fitted correctly.

• Support Vector Machine:- Support Vector Machine con- structs the hyperplane which creates the different classes for given training points. The main aim to construct a hyperplane having maximum distance between nearest training data points. Support Vector Machine predicts the classes for input data which is feature vector. It classifies training data in terms of two classes with cancer type it is 0 and for skin type it is [. The hyperplane equation for given x input is w.x - b or w.x + b according to category. Where w is normal

vector to hyperplane. [t classifies the data using formula.

Where Si is the support vector, ai is the weight, b is the bias, and k is a kernel function. In the case of a linear kernel, k is the dot product. If c 2: 0, then x is classified as a member of the first group, otherwise it is classified as a member of the second group.

III. RESULTS

For Unsupervised learning using k-means algorithm the classification result obtained is 52.63%. In k-means algorithm n data points are divided into k clusters. In case of melanoma skin cancer detection two clusters are formed one cluster is for cancer and another one is for non cancer detection. From the Table 2 it has been observed that the accuracy for Back Propagation Neural Network is 60% to 75% and by using Support Vector Machine it is 80% to 90% so Support Vector Machine is giving better results that the K-means clustering technique and Back Propagation Neural Network classification. K-means algorithm is worked on basis of unsupervised learning.

Support vector machine classifier accuracy is improved than the Back Propagation Neural network and K-Means clustering algorithm. In support vector machine hyperplane is formed which exactly separates data points for different categories and produces high accuracy results. Figure 3 shows skin non cancer type image while figure 4 represent cancer type image. Figure 5 represents all stages in detection of melanoma skin cancer.

Fig. 3. Skin non cancer type image.

Fig. 4. Skin cancer type image.

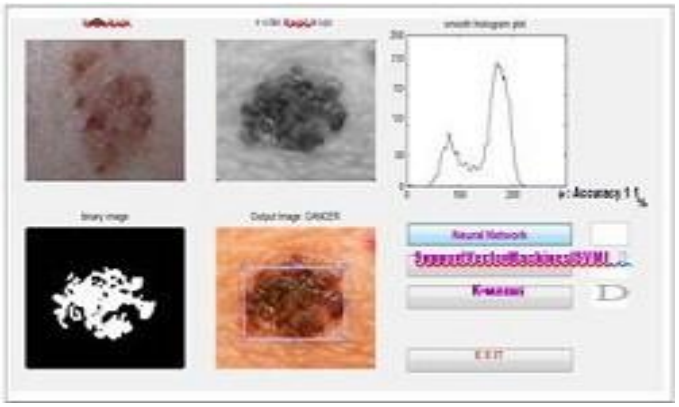

Fig. 5. All stages in detection of melanoma skin cancer.

REFERENCES

[1]Swati D. Shirke, C.Rajabhushanam "Iris Recognition from an Image at Lengthy Distance by using Deep Belief Neural Network (DBN)" Published in International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249-8958, Volume-8, Issue-5S3, July 2019

[2]Vijayalakshmi MM" Melanoma Skin Cancer Detection using Image Processing and Machine Learning" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-3 —Issue-4, June 2019

[3]Swati D. Shirke, C. Rajabhushanam, "Personal Iris Recognition from an Image at Long Distance using Back Propagation Neural Network" Published in

International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-8, Issue-9S4, July 2019

[4]Ms. Swati D. Shirke, Dr. C. Rajabhushnam," Iris Recognition Using Visible Wavelength Light Source and Near Infrared Light Source Image Database: A Short Survey" published in Proceedings of the Third International Conferenceon Trends in Electronics and Informatics (ICOEI 2019) IEEE Xplore Part Number: CFP19J32-ART; ISBN: 978-1-5386- 9439-8

[5]Swati D. Shirke, C. Rajabhushnam, "Correct Personal Iris Recognition at Long Distance by using Dougman's Rubber Sheet Model" published in International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8, Issue-1S5, June 2019

[6]W. Zhou, M.Yang, H.Li,X.Wang, Y.Lin,and Q.Tian, "Towards codebook-free: Scalable cascaded hashing for mobile image search,"IEEE Transactions on Multimedia, pp. 601–611, April 2014.

[7]Swati D. Shirke, C. Rajabhushanam" Biometric Iris Recognition of Person from an Image at Long Distance using Chronological Monarch Butterfly Optimization based Deep Belief Network" Published in Inter- national Journal of Recent Technology and Engineering (IJRTE) ISSN:

2277-3878, Volume-8, Issue-2S8, August 2019

[8]IOSR Journal of Engineering (IOSRJEN) e-ISSN: 2250-3021, p-ISSN:

2278-8719 Vol. 3, Issue 2 (Feb. 2013), ——V1—— PP 52-59

[9]Swati D. Shirke, C. Rajabhushnam" Correct Personal Iris Recognition at Long Distance by using Dougman's Rubber Sheet Model" published in International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8, Issue-1S5, June 2019

[10]International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056 Volume: 04 Issue: 04 — Apr -2017 www.irjet.net

[11]R.Suganya, Anautomated computer aided diagnosis of skin lesions detection and classification for dermoscopy images. April 2016.

[12]Cancer Facts and Statistics. http://www.skincancer.org/skin-cancer information/skin-cancer-facts.

[13]M.A. Farooq, M.A.M.Azhar ,and R.H.Raza, Automatic Lesion Detection System(ALDS)for Skin Cancer Classification Using SVM and Neural Classifiers. Oct 2016

[14]M. W. Rashad and M.Takruri, Automaticn on invasivere cognition of melanoma using Sup- port Vector Machines. Dec 2016.

[15]Automated Classification of Common Skin Lesions Using Bioinspired Features D. Csabai, 1 K. Szalai, 1,2 M. Gyngy.

[16]H.T.Lauand A.Al-Jumail, Automatically Early Detection of Skin Cancer: Study Based on Neural Network Classification. Dec 2009