Classification of Melanoma using Deep learning

[1]Dr.C.Mahiba, [2]Indhumathi E, [3]Kaleeswari P, [4]Kanagavalli A
[1]*Assistant Professor, Computer Science and Engineering, PSN College of Engineering and Technology, Tirunelveli, India*
[2,3,4] *Final year Student, Computer Science and Engineering, PSN College of Engineering and Technology, Tirunelveli, India*

***Abstract*— Dermatological diseases are one of the most pressing medical concerns in the twenty-first century. Owing to their high complexity and cost of diagnosis as well as the difficulty and subjectivity of human interpretation in the case of lethal disorders such as melanoma early detection is critical in evaluating the likelihood of cure. We believe those using automated methods will aid in early diagnosis particularly when dealing with a large number of photos with a variety of diagnoses as a result we describe In this article a fully automated approach for recognising dermatological disease through lesion photographs a machine intervention in contrast to traditional medical personnel-based identification. We proposed an approach to detect the melanoma ask in cancer and feature extraction through various image processing techniques. Our model is divided into three stages data gathering and augmentation model creation and prediction. We applied a variety of AI techniques including convolution neural network and support vector machine and amalgamated it with image processing tools to form a better structure leading to higher accuracy of 85%.**

***Index Terms:* Data gathering, Augmentation model, prediction, AI Techniques.**

1.INTRODUCTION

Skin is the outer most region of our body and it is likely to be exposed to the environment which may get in contact with dust, Pollution, micro-organisms and also to UV radiations. These may be the reasons for any kind of Skin diseases and also Skin related diseases are caused by instability in the genes this makes the skin diseases more complex. The human skin is composed of two major layers called epidermis and dermis. The top or the outer layer of the skin which is called the epidermis composed of three types of cells flat and scaly cells on the surface called SQUAMOUS cells, round cells called BASAL cells and MELANOCYTES, cells that provide skin its color and protect against skin damage. As the diagnostic classification currently do not represent the diversity of the disease, these are not sufficient enough to make a correct prediction and also treatment to be provided for that disease. Adding to this cancer cells are often diagnosed late and treated late, it is diagnosed when the cancer cells have mutated and spreads to the other internal parts of the body. At this stage therapies or treatments are not very effective. The other reasons for which the disease might have taken over to a very serious state can be because of people's ignorance and also that people try using home remedies without knowing the severity of the problem and also sometimes these may lead to another kind of skin rashes or even increasing these verity of the problem. The general stages of these diseases are as: STAGE1-diseasesinsitu, survival99.9%, STAGE 2- diseases in high risk level, survival 45-79%, STAGE 3-regional metastasis, survival 24-30%, STAGE 4- distant metastasis-survival7-19%. Existing segmentation algorithm (Otsu-RGB) for standard digital camera images proposed by Cavalcanti used three channels for thresholding. Shadings are removed using morphological operations To segment the skin lesion local textural variability information, principal component analysis are used. The results of have been proven that Otsu-RGB algorithm has reduced segmentation errors, but this algorithm has low segmentation accuracy. The segmented image is given to extract the centre of the mass by calculating the mean value of the image repeatedly to acquire the centre, in order to measure the geometrical features like Area, Perimeter, MinorAxis Length, Major Axis Length appropriately. The existing system provided only whether affected by melanoma or not. The important steps in a diagnosis of melanoma skin cancer are: image acquision of lesion image, Segmentation of the lesion area from the outer area, Extraction of the feature from the input image. Feature extraction is the intent of extracting the features from the lesion image in order to characterize the melanoma. The Stolz algorithm is used for feature classification stage, the extracted features that are proceeded in order to assort the image as mole, benign, suspicious, highly suspicious skin lesions.

2. SYSTEM ARCHITECTURE

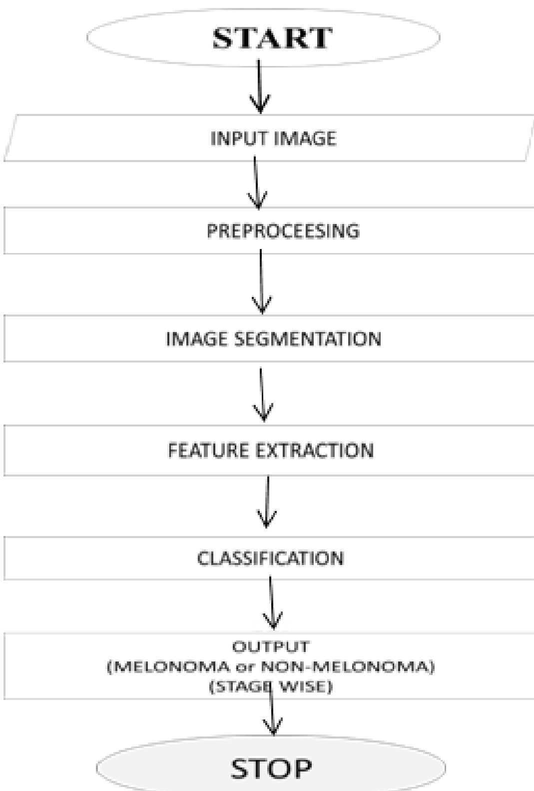

Fig. System Architecture

3. MODULE DESCRIPTION

3.1 DATASETCOLLECTION

The images were collected from the ISIC dataset; the ISIC dataset provide the collection of images for melanoma skin cancer. ISIC melanoma project was undertaken to reduce the increasing deaths related to melanoma and efficiency of melanoma early detection. This ISIC dataset contains approximately 23,000 images of which we have collected 1000-1500 images and trained and tested over these images.

Hair Removal: for the above collected images hair removal method was applied this method was performed using Hough transform, Hough transform is basically used to identify lines or elliptical or circular shapes. Performing hair removal for the images that has hair within the tumor provides us an clear image of tumor which also helpsusto make further more enhancements.

Shading removal: The images that is taken from the dataset contains shade around the region of the tumor this shade for few images is dark and for few is light, removal of the shade in the region of tumor also provides us an clear vision of the tumor which is also helpful in the further enhancements. We have used the MATLAB filters to remove the shade for images in the dataset. Glare Removal: sometime the images are captured from camera the images will contain glare this glare is not visible to the naked eyes, we remove this glare using the MATLAB filter, this minute noise sometimes may affect the accuracy at the end.

3.2 PRE-PROCESSING

It is very important step in segmentation of skin lesions from digital images. In this proposed work standard digital camera is used to capture skin lesion images. Illumination variation occurs due to shadows, this feature allows misclassification of the shadows as skin lesion .In this paper Multistage Illumination Modeling algorithm (MSIM) is implemented to remove illumination variation.

3.3SEGMENTATION

It is the second important step. In this paper Modified Texture Distinctiveness lesion segmentation algorithm (M-TDLS) is proposed to segment the skin lesion. M-TDLS algorithm involves two steps. TD metric Calculation and feature extraction to Region Classification

3.3.1TD METRIC CALCULATION

Original RGB image is converted to XYZ colour space, which gives efficient skin lesion detection. Texture vectors are extracted for each pixel to find representative texture distributions [4], probability of distinctiveness between two texture distributions is calculated. TD metric is calculated to find dissimilarity between two texture distributions. Skin lesion distributions have high TD metric due to having high pixel intensity variation, where normal skin distributions are same.TD metric is based on only lesion texture distributions

3.3.2 REGION CLASSIFICATION

The second step is to over segment the input image to classify the regions as normal or lesion. Otsuthreshold value is used to divide the set of texture distributions into two classes, which classifies texture distributions belongs to normal or lesion. Region Distinctiveness metric is used for region classification which is based on TD metric. After classification morphological dilation operator is used to refine the lesion border which provides accurate estimate of lesion order.

3.4 FEATURE EXTRACTION

After segmenting the skin lesion, to classify it as melanoma or non-melanoma some unique feature is extracted, these features are given as input to the classifier. Unique features of melanoma are asymmetry, border irregularity, colour variation and diameter. To find asymmetry of lesion solidity, equivalent diameter features are calculated, for Border irregularity mean and variance are calculated. To determine colour variation maximum and minimum pixel intensities of RGB channels are calculated. As per ABCD rule the features which we need to extract include Asymmetry Index Border Color Index Diameter.

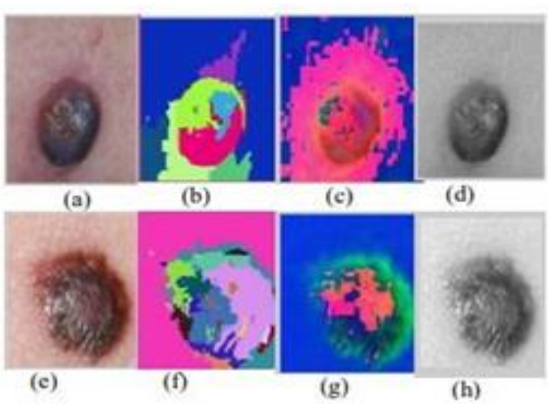

Fig.2 Feature Extraction

3.4.1 ASYMMETRYINDEX

Asymmetry Index is computed with the following equation: AI=(A1+A2)/2Ar Where, A1= Area of non-overlapped region along minor axis of the lesionA2= Area of non-overlapped region along major axis of the lesion Ar= Area of lession Implementation. The Area of lesion (Ar) can be calculated using bw area over the binary image of the segmented region. For calculating non overlapped area over axis. The segmented region is divided along the lines passing through centroid of the region Two Separate areas are generated which are then adjusted so that the areas will be overlapped by flipping one area. Using XOR over the area will generate the non-overlapped region whose area is calculated using bw area function To generate area along x axis the bisection will be generated using first Gx pixels and the next Gx pixels along x axis and bisecting line on y axis. To generate area along y axis the bisection will be generated using first Gy pixels and the next Gy pixels along x axis and bisecting line on y axis. After calculating area of the regions Asymmetry index is calculated Border

3.4.2 IRREGULARITY

In order to calculate border irregularity, there are different measures such as: compactness index, fractal index, edge abruptness. 1) Compact Index: Compact Index can be determined by using the following equation: CI=(P2L)=(4AL)Where, PL = Perimeter of the Lesion.AL=Area of the Lesion. Fractal Dimension: Fractal set is provided by the" box counting" method. It returns two variables whose differential log ratio provides the fractal dimension as the mean value along 4-8 index. Edge Variation: Edge variation is calculated using the following equation EI= ((Max Min)%6+2)/100; Where, Max and min are length of major and minor axis. Axis lengths are calculated using region props function.

3.4.3 COLOR INDEX

Color index is calculated by converting the input image to HSV image value by checking the presence of the following colors. Length of all the available pixels with given values is divided by total number of pixels. The presence of color is dependent on the value of resultant not equal to zero. For each color present theColorIndexis+1.

3.4.4 DIAMETER

The diameter value is said to be 5 if the diameter of lesion is greater than6mm. For other values the diameter is one less than its actual rounded value. To calculate Diameter the region props function is used to get the minor axis length of the lesion region. Resultant value is converted into mm value and the value is assigned to diameter [2].

4. CLASSIFICATION

It is the final step. There are many classification algorithms to classify segmented lesion image like ANN classifier, hybrid classifier, SVM classifier. In this paper Support Vector Machine (SVM) classifier algorithm is used to classify the segmented lesion as melanoma or non-melanoma.SVM classifier [2] provides good classification results in image processing. SVM constructs hyper planes to classify a set of data. Unique features of melanoma are asymmetry, border irregularity, colour variation and diameter. To find asymmetry of lesion solidity, equivalent diameter features are calculated, for Border irregularity mean and variance are calculated. To determine colour variation maximum and minimum pixel intensities of RGB channels are calculated.

4.4.1METHODLOGY

4.4.1.1 MSIM ALGORITHM

The main advantage of MSIM algorithm is: It has the following special features than existing algorithms. It can be used for complex texture images. It maintains consistent skin lesion color after removing shadows i.e. skin lesion color does not vary.

It provides good illumination correction. It avoids misclassification of skin lesion. MSIM algorithm involves three steps: Segmentation map, Illumination map and Reflectance map

4.4.1.1 SEGMENTATION MAP

The steps of the algorithm is explained in the following section.

Step1: Start the process.

Step2: Get the input skin lesion image.

Step3: Construct four connected graph using horizontal and vertical pixels.

Step4: Calculate image gradient to find edge pixels. Image gradient provides the information about pixel intensity variation in horizontal and vertical direction.

Step5: Generate pairs using four connectivity graphs.

Step6: Group the pairs as regions based on pixel intensity similarity.

Step7:Combine regions based on merging predicate using the equation1),(2)&(3).P (R, R′) = – ≤ -- (1) b (R, R′) = -- (2) b(R)= --(3) Where R′ a and Ra are observed average pixel intensities in the regions R' and R respectively, b (R,R′) is merging threshold, P(R,R′) is merging predicate which combines two regions(R,R′) of the image, g is maximum pixel intensity, Q is defined as segmentation parameter which decides the number of regions to be segmented in the image I.

If K is set of regions with l pixels, is defined as1/6|.

Step8: Stop the process.

4.1.1.2 ILLUMINATION MAP

After classifying the pixels as skin lesion or normal skin the original RGB image is converted into HSV colour space to get illumination map. Skin lesion photograph is illuminated through white light. Due to this only V channel is down scaled. Hue and saturation channels are not used.

4.1.1.3 REFLECTANCE MAP

To get final illumination corrected image reflectance map should be estimated, it is obtained from V channel pixel intensity and illumination map. Finally hue and saturation channels are added to value channel to corrected image. v(s)=i(s).r(s).Where s is pixel location, v(s) is V channel pixel intensity, i(s) is illumination component, r(s)is reflectance component

4.1.1.4 SVM

SVM is a supervised learning algorithm. This means that SVM (Support Vector Machine) trains on a set of labelled data. SVM studies the labelled training data and then classifies any new input data depending on what it learned in the training phase. A main advantage of SVM is that it can be used for both classification and regression problems. Though SVM is mainly known for classification, the SVR (Support Vector Regressor) is used for regression problems.SVM can be used for classifying non-linear data by using the kernel trick. The kernel trick means transforming data into another dimension that has a clear dividing margin between classes of data. After which you can easily draw a hyper plane between the various classes of data. What is support vectors in SVM? we start of by drawing a random hyper plane and then we check the distance between the hyper plane and the closest data points from each class.

These closest data points to the hyper plane are known as support vectors. And that's where the name comes from, support vector machine. In this paper we have used SVM to classify the malignant and benign skin cancer images, this done by passing the segmented and feature extracted images into SVM where SVM write the hyper plane and groups all the nearby similar features into different classes. The performance of the SVM classifier was very accurate for even a small data set.

5. CONCLUSION

This paper aims to determine the accurate prediction of skin cancer and also to classify the skin cancer as malignant or non-malignant melanoma. To do so, some pre-processing steps were carried out which followed Hair removal, shadow removal, glare removal and also segmentation. SVM and Deep Neural networks will be used to classify. The classifier will be trained to learn the features and finally used to classify. The novelty of the present methodology is that it should do the detection in very quick time hence aiding the technician stopper fect their diagnostic skills. The dataset used is from the available ISIC (International Skin Image Collaboration) data set, hence any dataset can be used to find the efficiency Incident rates of melanoma skin cancer have been rising since last two decades. So, early, fast and effective detection of skin cancer is paramount importance. If detected at an early stage, skin has one of the highest cure rates, and the most cases, the treatment is quite simple and involves excision of the lesion. Moreover, at an early stage, skin cancer is very economical to treat, while at a late stage, cancerous lesions usually result in near fatal consequences and extremely high costs associated with the necessary treatments. After all, the best way to lower the risk of melanoma is to limit the exposure to strong sunlight and other source of Ultraviolet light. Finally take care of all the necessary measures such as: protecting skin with clothing, wearing hat, using sunscreen, staying in the shade (etc.). Moreover, always stay alert about skin and do monthly skin-self exams to reduce the chance of getting any skin cancer which is a risk to human life.

REFERENCES

[1] Jeffrey Glaister and David A. Clausi, Senior Member, "Segmentation of Skin Lesions From Digital Images Using Joint Statistical Texture Distinctiveness" IEEE Transactions on Biomedical Engineering, Vol., 61, No.4, pp.1220-230,April.2017.

[2] Pedram Ghamisi, Micael S.Couceiro, Fernando M.L.Martins, JonAtli Benediktsson" Multilevel Image Segmentation Based on Fractional-Order Darwinian Particle Swarm Optimization", IEEE Transactions on Geo science and Remote sensing,vol,52,no.5,pp.2382-2394,May2014.

[3] A.Clausi, "MSIM: Multi-stage illumination modeling of dermatological photographs for illumination corrected skin lesion analysis," IEEE Transactions on Biomedical Engineering, vol. 60, no.7,pp.1873-1883, Jul. 2017.

[4] C.Scharfenberger, A.Wong, K.Fergani, J.S.Zelek, and D.A.Clausi, "Statistical textural distinctiveness for salient region detection in natural images," In IEEE Conference on Computer Vision and Pattern Recognition. Jun. 2013, pp.979–986.

[5] M.Celebi, H.Kingravi, and P.A.Vela, "A comparative study of efficient initialization methods for the k-means clustering algorithm," Expert Systems with Applications, vol.40, no.1, pp.200–210, Sep.2016.

[6] P.G.CavalcantiandJ.Scharcanski,"Automatedprescreeningofpigmentedskinlesionsusingstandardcameras",Elsevier,ComputerizedMedicalImagingGraphics , vol.35, no.6, pp. 481–491,Sep. 2011.

[7] M. Silveira, J. Nascimento, J. Marques, A. R. S. Marcal, T. Mendonca, S. Yamauchi, J. Maeda, andJ. Rozeira, "Comparisonof segmentation methods for melanoma diagnosis in dermoscopy images," IEEE Journals of Selected Topics in Signal Processing.vol.3, no. 1,pp. 35–45, 2009.

[8] R. Nock and F. Nielsen, "Statistical region merging," IEEE Transactions on Pattern Analysis and Machine Intelligence , vol.26,no.11, pp. 1452–1458, Nov2004.