MELANOMA DETECTION FROM DIGITAL IMAGES USING TEXTURE DISTINCTIVENESS BASED SEGMENTATION

Sinu S V[1] ,Dr.R.A Jaikumar[2], Abhilash S Vasu[3]
[1]PG Scholar, Dept. of ECE, SHM Engineering College, Kollam, Kerala, India
[2]Professor, Ponjesly College of Engineering, Nagercoil, Tamilnadu
[3]Asst.Professor, Dept. of ECE, SHM Engineering College, Kollam, Kerala, India

ABSTRACT

Melanoma is the deadliest form of skin cancer. It should be diagnosed early because of its aggressiveness. Usually Melanoma detection is done by clinical analysis and biopsy tests. These kinds of methods are expensive andtime consuming also have many side effects. Thus, an automated melanoma prescreening system is needed to diagnose melanoma using images acquired in digital cameras. Accuracy of segmentation and melanoma detection can be improved by examining the textural features of skin lesion. Computer aided melanoma detection system using image processing techniques is proposed for accurate and early detection of melanoma. This system has different stages which include preprocessing for image enhancement, segmentation of skin lesion using textural features to improve accuracy, feature extraction and classification. The input image is preprocessed using contrast stretching for image enhancement. The enhanced image is segmented using Texture Distinctiveness Lesion Segmentation (TDLS) algorithm to extract the lesion area from the background skin. In this work, different types of melanoma are classified as Superficial Spreading Melanoma(SSM), Nodular Melanoma(NM), Lentigo Maligna Melanoma(LMM) using multi-SVM (Support Vector Machine). Support Vector Machines (SVM) are supervisedlearning models with associated algorithms that analyze images for classification analysis.

Index Terms:- Melanoma Detection, TDLS Segmentation, Classification

I. INTRODUCTION

Skin is the largest organ present in the human body. Its functions include regulation of body temperature as well as indicating any malfunction within the body with the change in color or pigmentation. The structure of skin comprises of an outermost layer called Epidermis. The top layer of epidermis contains a tough, fibrous protein called Keratin while the bottom layer contains Melanin which is a dark pigment which protects the body from the harmful rays of sun. [3].

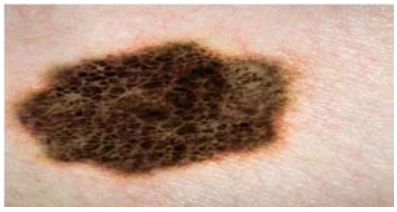

Fig. 1 Melanoma skin lesion image

Malignant cells or cancer cells are usually found in the outer layers of skin. It is divided into two major forms: Melanomas and Non-Melanomas. Melanoma is a form of skin cancer that affects the melanocytes in the epidermis whereas, Non-Melanoma further is of two primary forms: Basal cell skin cancer which is common and typically appear on the head, neck, arms and other body parts which are frequently exposed to the sun. Squamous cell carcinomas originate in the outer layers of the epidermis. It progresses rapidly to involve deeper layers of dermal skin tissue but is unlikely to spread to other parts of body. The common types of melanoma areSuperficial

spreading melanoma (SSM*),* Nodular melanoma (NM),Lentigo melanoma (LM).

SSM is the most frequent form of MM in the Caucasian population and is diagnosed in about 65% of all MM cases. At the beginning, a lesion of SSM is flat and grows horizontally, subsequently its surface becomes irregular as circumscribed infiltrated papules or nodules develop,signaling vertical growth. The prognosis is relatively favorable in early phases (horizontal growth). The risk of metastasis significantly increases when vertical growth and dermal invasion occur. Nodular melanoma*,* this form of MM shows an early vertical growth with rapid invasion of the dermis, makes prognosis unfavorable even in the early phases. Lentigo melanoma (LM),This clinical form accounts for about 10% of the MM cases and can growth for years or even decades until it develops malignant features. The prognosis of this type is more favorable, as the vertical growth occurs only late. Incident rates of melanoma have been increasing among non-Hispanic white males and females, but survival rates can be increased if detected early [2].

By extracting the cancerous tissue, melanoma can be cured completely in the initial stage itself. Otherwise it might go to the secondary stage in which recovery becomes impossible. Proper melanoma detection mechanism is needed to detect and diagnose melanoma in the initial stage. Fig.1 shows the example of melanoma skin lesion image. Clinical analysis and biopsy tests are commonly used for melanoma detection and diagnosis. Clinical analysis is done using a dermatoscope by trained dermatologists. A dermatoscope is an optical device used by the dermatologists to get a magnified and enhanced view of skin structure using skin surface reflection. Here melanoma is detected based on a visual examination. Visual examination by a dermatologist is based on ABCDE [3] rule where A denotes asymmetry, B denotes border, C denotes color and E denotes evolving structures. If any trace of melanoma is found in the initial clinical screening using dermatoscope, the patient is referred for further biopsy tests. Biopsy test is usually done in an outpatient department by a trained pathologist where all or part of the skin lesion will be extracted for further analysis. The clinical analysis and biopsy tests are very expensive and time consuming. Clinical analysis sometimes gives false positive or false negative results which may lead the patient to the secondary stage and further biopsy tests. The biopsy tests also have many side effects like scarring and other health issue.To overcome the disadvantages of Clinical analysis and biopsy tests, an automated system is needed to detect and diagnose melanoma in the initial stage itself so that clinical trials and biopsy tests can be eliminated. More accurate results will be obtained in the initial stage itself without going for expensive dermatological screening and biopsy tests and the treatment can be immediately started if detected early and thereby reducing the death rates. The objective of this project is to develop a novel, less complex, accurate and automatic screening software on MATLAB platform to diagnose a patient's risk of melanoma by using their skin lesion images which can be developed by applying image processing techniques.

II. LITERATURE REVIEW

The standard method to evaluate skin growth in order to rule out melanoma is by biopsy followed by histopathological examination. The challenge lies in identifying the lesions that have the highest probability for being melanoma. Mostly dermoscopy is used for the diagnosis of skin cancer. Which is a non-invasive skin imaging technique which uses a hand-held lighted magnifier to analyse skin lesions by observing newly defined and descriptively named subsurface structures (e.g., dots, streaks, veils, networks). While due to thedifficulty of human interpretation, analysis of dermoscopy images has become an important research area. The main approach for CAD is to determine an estimate of the probability of disease.Diagnosis of skin lesion and also to find the location of lesion. it is required to understand the relevance of the different dermoscopic features, which are lesion specific features , color , symmetry and pattern analysis. The general architecture of CAD system includes selection of training samples, image pre-processing, segmentation, feature extraction and classification. The aim of the pre-

processing step is to eliminate the background noise and improve the image quality for the purpose of determining the focal areas in the image. Image segmentation is an important step in image analysis, pattern recognition, and computer vision. A better segmentation of skin images can help the diagnosis to define well the region of cancer. In 1985, recognizing the need to educate physician and the public to recognize melanoma in its 7 early clinical presentation, group from New York University coined the ABCD acronym (Asymmetry, Border irregularity, Color variegation, Diameter > 6mm) [5]. For melanoma skin lesion detection, ABCD features are most widely used for feature extraction which is based on morphological analysis of dermatoscopic image of skin lesion. When features are determined, the next step is to distinguish the malignant structures from their counterparts. In this step, a region of interest of lesion image is assigned to one of the classes of cancerous, benign, or healthy. We can also possible to classify the malignancy level of the tissues.

Existing Melanoma detection systems make use ofIterative Stochastic Region Merging [6], Statistical Region Merging (SRM) [5], Multilevel Thresholding [7] and Color Enhancement and Iterative Segmentation [8]. SRM can be used only for dermoscopic images. It does not consider any texture information and it is noise sensitive. Iterative Stochastic Region Merging method is suitable only for macroscopic images which are robust to noise and artifacts, structural irregularities, illumination variations but the procedure is more complex. This method of segmentation does not consider texture analysis so this method is less accurate. Multilevel Thresholding is suitable for both dermoscopic and digital images but does not include textural information during segmentation and it is computationally more complex. In Color Enhancement and Iterative Segmentation, texture based segmentation is not used. So, it is less accurate and complexity is high due to iterative segmentation. Existing Melanoma detection systems are comparatively less accurate and do not include texture analysis for segmentation of skin lesion from the background skin. Texture based segmentation enables accuratefeature extraction and segmentation of skin lesion. Classification of the segmented images must be done to accurately classify the images asmelanoma or not after extracting various features of skin lesion. Textural based segmentation improves accuracy of the overall melanoma detection system and classification results.

III. PROPOSED SYSTEM

The proposed melanoma detection system mainly consist of texture based segmentation and classification . The proposed framework is shown in Fig.2.

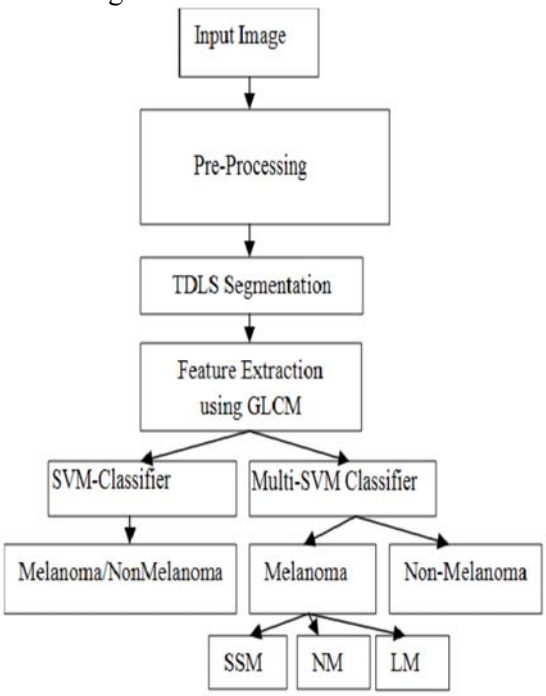

Fig 2: Proposed System

This system consists of different modules which include preprocessing, segmentation, feature extraction and classification. Preprocessing is done using histogram equalization. The preprocessed image is segmented using Texture Distinctiveness Lesion Segmentation [2] (TDLS) method to extract the lesion area from the background skin. The extracted lesion area is used for feature extraction which is done by gray level 12 cooccurrence matrix (GLCM). Using the extracted features, the system is trained using SVM classifier in order to classify the images as malignant or benign melanoma. The melanoma images are again categorized as Superficial Spreading Melanoma, [4] Nodular Melanoma, Lentigo Maligna Melanoma

A. PRE-PROCESSING

Preprocessing 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. These images may contain contrast variations illumination artifacts and unwanted shading. So, these images cannot be used directly for segmentation of the lesion area since it reduces the accuracy of the overall system. Therefore, the image must be enhanced before segmentation. Histogram Equalization is used in the preprocessing stage to enhance the image contrast and to produce better quality images. In Histogram Equalization, contrast of the image is enhanced using the image histogram. Image histogram is the graphical representation of the tonal distribution in a digital image in which number of pixels of each tonal value is plotted. In this contrast adjustment, intensities are distributed on the histogram so the low contrast pixel intensities are enhanced to high contrast intensity values by spreading out the most frequent intensity values. in Fig. 1(b) and (d). In both examples, shadows that appeared on the left side of the uncorrected images have been removed. The corrected images are used as the input to the segmentation algorithm

B. TDLS Segmentation

Most important step in automatic melanoma detection is the segmentation of lesion area from the background skin. InSegmentation a digital image is divided into different segments which aims to produce the representation of an image into something that is easier for further analysis. The input images contain both skin and lesion area. The lesion area is used for feature extraction and further analysis. If the image contains skin surface, accurate feature extraction cannot be done and it affects the classification results. Inorder to improve the accuracy of the system, lesion area must be properly extracted from the background skin and used for further analysis. Texture Distinctiveness Lesion Segmentation (TDLS) method is used for extracting the lesion from the skin surface. In TDLS algorithm, the input image in RGB color space is first converted into XYZ color space. For each pixel in the image, a local texture vector is found out which contains the neighborhood pixels of size *n* centered on a particular pixel of interest *s*. For an image of size N * M, a set of N * M texture vectors are obtained

$T = \{tsj | 1 \leq j \leq N \times M\}$. (1)

Two step clustering process is used after extracting the texture vectors of each pixel. First k-means clustering is used which divides the image into *k* number of clusters where each cluster contains the different texture data. Next a Gaussian Mixture Model (GMM) clustering is used since k-means clustering does not consider any probabilistic information of the texture. For each cluster, a [2] Gaussian distribution is assumed and model parameters like distribution mean μ and distribution covariance Σ are calculated. Next step is to define a metric *lj,k*given in (2) which represents the similarity of two texture distributions and it is asymmetric since $\Sigma i \neq \Sigma j$. It gives the probability that the mean of one texture distribution is a realization of the mean of other texture distribution. Let the mean and covariance of each texture distribution *Tj*be *tj*and Σj.The average of *lj,k*and *lk,j*gives the measure of similarity *Lj,k*which is defined in (3) [2]

$$l_{j,k} = \frac{1}{\sqrt{(2\pi)^{n*n*a}|\Sigma_j|}} exp^{\left(-1/2\left(t_j^r - t_k^r\right)^T \Sigma_J^{-1}\left(t_j^r - t_k^r\right)\right)} (2)$$

$$L_{j,k} = 1/2(l_{j,k} + l_{j\prime k}) \quad (3)$$

The texture distributions will be different form one another. Next step is to find the distinctiveness of the texture distributions by calculating a metric *dj,k*defined in (4) which gives the probability of one texture distribution is different from other texture distributions. Let *P(Tk| I)* be the probability of occurrence of a pixel [2] being associated with a particular texture distribution *Tk*. The measure of dissimilarity of a texture distribution *Tj*from other texture distributions is defined using a Texture Distinctiveness (TD) metric *Dj*which is given in (5). For normal skin, the dissimilarity of skin texture distributions will be very small so the TD metric of normal skin texture distributions will be small. But the lesion textures are dissimilar from one another so the TD metric will be large.

$$d_{j,k} = 1 - L_{j,k} \quad (4)$$

$$D_j = \sum_{k=1}^{K} d_{j,k} P(T_k^r | I) \quad (5)$$

The next main step in TDLS algorithm is to find and classify regions in the input image as being a part of the lesion based on the texture distributions and associated TD metric. Statistical Region Merging (SRM) [5] is used to divide the lesion image into number of regions .The two main steps in SRM are sorting and merging of pixels. In sorting step, the pixels in image are sorted to find the order in which pixels are compared. In the merging step, [2] the pairs of pixels are combined to form regions based on their similarity. For each texture distribution, a TD metric is computed and it is combined with the contents of each region to find out a regional TD metric, *DR* (6) which represents the average TD of a region *R* where probability of a pixel being associated with the *j*th texture distribution of a region *R* is given as *P(Tk| I)*.

$$D_R = \sum_{j=1}^{K} D_j P\left(T_j^r / R\right) \quad (6)$$

After computing the *DR* metric, regions in the image must be classified (7) as skin or lesion using the average TD metric *DR*. The classification of the regions is done using a threshold τ which divides the set of texture distributions into two categories, normal skin and lesion. It acts as a decision boundary between the normal skin and the lesion class and it is computed using Otsu's threshold method [2]

$$y(R) = \begin{cases} 1, & if\ D_R \geq \tau(lesion) \\ 0, & otherwise(normal) \end{cases} \quad (7)$$

A. Feature Extraction

The segmented lesion is used for extracting various features which is given as input for the classification algorithms to classify the lesion as malignant or benign. Textural features such as Angular second moment (ASM), Contrast, Inverse Difference Moment (IDM), Entropy, Variance, Correlation, Cluster shade, [10] Cluster Prominence, Homogeneity, Difference entropy, Dissimilarity and sum average are extracted from the segmented lesion. Gray Level Cooccurrence Matrix (GLCM) is used to extract the textural features of segmented skin lesion. GLCM is a matrix representation of the image A.

Classification

Feature extraction step is followed by the classification of images. The set of extracted features are given as input to the classification stage. Based on the extracted features, the system must classify the images as Melanoma or Non-Melanoma. Melanoma images must be again categorized as SSM, NM, LMM. Two classifiers, Support Vector Machine (SVM) and Multi-(SVM) [11] are used in the proposed system. Using the extracted feature set, SVM classifier is trained to classify the images as Melanoma or not Binary pattern recognition involves constructing a decision rule to classify examples into one of two classes based on a training set of examples whose classification is known a priori. Support Vector Machines(SVMs) construct a decision surface in the feature space that bisects the two categories and maximizes the margin of separation between two classes of points

IV. RESULTS AND DISCUSSIONS

In the proposed system, digital images of melanoma and non-melanoma were collected from internet and hospitals. Preprocessing is done to enhance the quality of all the images. TDLS algorithm is implemented to segment the lesion area from the collected images. Features are extracted from the segmented lesion area and they are given for classification of melanoma images. SVM classifiers were used to classify the images as Melanoma or Non-Melanoma. The system is implemented in MATLAB on a computer with an Intel Core i3-380M CPU (2.5 GHz, 4GB RAM). For performance analysis, FeatureExtraction is done for both raw images and segmented images and then given to train the classifiers. 12 features are extracted and given to train theclassifier. Classification done using features extracted from the segmented image gives better accuracy than classification done using features from raw images without segmentation. A set of 126 images from the Dermquest database [13] were used for training and testing the classifiers and classified as Melanoma and Non-Melanoma.

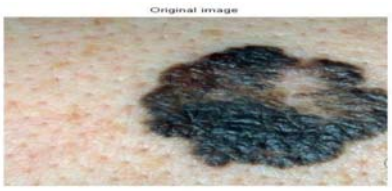

Fig 3 Input image

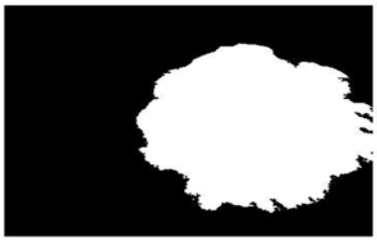

Fig 4 morphology operation

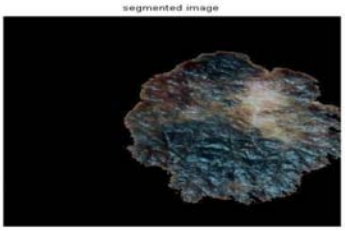

Fig 5 Segmented image

Using Multi-SVMclassifier, melanoma images are again classified into 3 categories namely Superficial Spreading Melanoma, Nodular Melanoma and Lentigo Maligna Melanoma

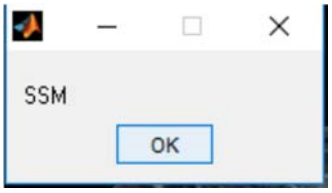

Fig 6 output of multi-SVM

Performance of the segmentation algorithms are compared by calculating accuracy and the formula is given in (8) where TP denotes number of true positive pixels, FP denotes the number of false positive pixels, TF denotes the number of true negative pixels, and FN denotes number of false negative pixels [3].

Accuracy = (TP+TN) / (TP+FN+TN+FP)

Classifier	Accuracy	Sensitivity	Specificity
SVM	92%	93%	91%
Multi-SVM	96%	97%	95%

Table 4.1 Performance analysis Measurements

V. CONCLUSION

Melanoma detection in the initial stage is very important since it can reduce the death rates to a great extent. Early detection can be done only with the help of trained dermatologists and proper diagnosis. Due to the lack of trained dermatologists, expensive and time consuming diagnosis procedures, automated system for melanoma detection is required. An automated melanoma detection system using texture based lesion segmentation, feature extraction using GLCM and classification of images using SVM and multi-SVM classifiers is implemented. Using multi-SVM classifier, melanoma images are again classified into 3 categories namely Superficial Spreading Melanoma, Nodular Melanoma and Lentigo Maligna Melanoma. This proposed system gives better diagnosis and accuracy than conventional clinical screening and biopsy tests since it makes use of texture based analysis and classification. Overall accuracy of this study ranges from 92% to 96%. This automated detection system can be used for both dermatological images and digital images taken by a standard camera. So the dermatological screening can be eliminated. By extracting more features from the segmented image, the accuracy of the overall system can be further improved and these feature set can be used to classify the images to find out the stages of melanoma skin cancer.

REFERENCES

[1] N. Howlader, A. M. Noone, M. Krapcho, J. Garshell, N. Neyman, S.F. Altekruse, C. L. Kosary, M. Yu, J. Ruhl, Z. Tatalovich, H. Cho, A.Mariotto, D. R. Lewis, H. S. Chen, E. J. Feuer,

and K. A. Cronin, "SEER cancer statistics review, 1975-2010," Nat. Cancer Inst., Bethesda, MD,USA, Tech. Rep., 2013
2. F. Jerants, J. T. Johnson, C. D. Sheridan, and T. J. Caffrey, "Early detection and treatment of skin cancer," Amer. Family Phys., vol. 62, no. 2,pp. 1–6, Jul. 2000.
3. Public Health Agency of Canada. (2013). Melanoma skin cancer.[Online]. Available:http://www.phac-aspc.gc.ca/cd-mc/cancer/melanoma skin cancer-cancer peaumelanome-eng.php
4. Jemal, M. Saraiya, P. Patel, S. S. Cherala, J. Barnholtz-Sloan, J. Kim,C. L. Wiggins, and P. A. Wingo, "Recent trends in cutaneous melanoma incidence and death rates in the united states, 1992-2006," J. Amer. Acad. Dermatol., vol. 65, no. 5, pp. S17.e1–S17.e11, Nov. 2011.
5. K. A. Freedberg, A. C. Geller, D. R. Miller, R. A. Lew, and H. K. Koh,"Screening for malignant melanoma: A cost-effectiveness analysis," J.Amer. Acad. Dermatol., vol. 41, no. 5, pt. 1, pp. 738–745, Nov. 1999.
6. "FUNDAMENTALS OF DIGITAL IMAGE PROCESSING" by Anil K. Jain
7. "Digital Image Processing" by Rafael C. Gonzalez
8. "DIGITAL IMAGE PROCESSING & ANALYSIS" by Bhabatosh Chanda
9. "Digital Image Processing Using MATLAB" by Rafael c. Gonzalez
10. "Digital Image Processing" by Sanjay M. Shah Munesh Chandra
11. Trivedihttps://www.mathworks.com/discovery/digital-image-processing.html
12. http://proceedings.spiedigitallibrary.org/volume.aspx?volumeid=5040
13. https://www.tutorialspoint.com/dip/
14. https://stanford.edu/class/ee368/
15. https://www.amazon.com/Digital-Image-Processing-Rafael-Gonzalez/dp/013168728X