Available online at www.ijarbest.com

International Journal of Advanced Research in Biology, Ecology, Science and Technology (IJARBEST)
Vol. 1, Issue 2, May 2015

Segmentation Of Skin Lesions Using Texture Distinctiveness Lesion Segmentation Algorithm

Bindhu S[1], Manu Mohan C M[2]

Post-Graduate Scholar, Department of Electronics and Communication, FISAT, Ernakulam, India [1]
Assistant Professor, Department of Electronics and Communication, FISAT, Ernakulam, India [2]

***Abstract*— The most deadliest form of skin cancer is known as Melanoma. If it is detected in the earlier stage, the survival rates can be improved. Due to the costs for dermatologists to screen every patient, there is a need for an automated system to assess a patient's risk of melanoma using images of their skin lesions captured using a standard digital camera. The main challenge in implementing such a system is locating the skin lesion in the digital image. A texture based skin lesion segmentation algorithm called TDLS(Texture Distinctiveness Lesion Segmentation) is used to locate skin lesion in the captured image. In this method, first a set of sparse texture distributions that represent skin and lesion textures are learned. Second, the TD metric is calculated and it is used to classify regions in the image as part of skin class or lesion class. The proposed algorithm is implemented using MATLAB.**

***Index Terms*— Melanoma, segmentation, skin cancer, texture.**

I. INTRODUCTION

Melanoma is a cancer of the melanocytes, the cell found in the skin's epidermis that produces melanin. These cancerous growth develop when unrepaired DNA damage to skin cells most often caused by ultraviolet radiation from sunshine or tanning beds. This triggers mutation that lead the skin cells to multiply rapidly and form malignant tumours.

Melanoma accounts for approximately 75% of deaths associated with skin cancer . In 2013, it is estimated that 76,690 people diagnosed with melanoma and 9,480 people died of melanoma in the United States. In Canada, 1 in 74 men and 1 in 90 women will develop melanoma in their lifetime . For young adults ages 15-30, melanoma is one of the most commonly diagnosed forms of cancer .If melanoma is detected at Stage I, the survival rate is 96% ; but detected in Stage IV , survival rate decreases to 5% [1] . With the rising incidence rates in certain subsets of the general population, early melanoma screening is beneficial.

A dermatoscope is a device used by dermatologist to detect skin cancer. It consists of a magnifier (typically x10), a non-polarised light source, a transparent plate and a liquid medium between the instrument and the skin. This act as a filter and magnifier and helps the inspection of skin lesions unobstructed by skin surface reflections . The main reasons against using the the dermatoscope include a lack of training or interest and its cost for melanoma screening [2].

The limitations with visual melanoma screening can be overcome through the use of computer-aided diagnosis of melanoma. These computer algorithms take an image of the skin lesion as an input and extract a set of useful features. The features are used to identify the skin lesion as malignant melanoma with an accurate estimate of the lesion border.

The existing segmentation algorithms like simple thresholding, active contours, and region merging only use features derived from pixel color to drive the segmentation. Another approach to find skin lesions is to incorporate textural information, because normal skin and lesion areas have different textures. Textures describes a pixel characteristics include smoothness, roughness, or the presence of ridges, bumps or other deformations and are visible by variation in pixel intensities in an area.

A new segmentation algorithm based on texture distinctiveness (TD) can be proposed to locate the skin lesions in photographs. This algorithm is called the texture distinctiveness lesion segmentation (TDLS) algorithm. The TDLS algorithm consists of two main steps. First, a set of sparse texture distributions that represent skin and lesion textures are learned. A TD metric is calculated to measure the dissimilarity of a texture distribution. Second, the TD metric is used to classify regions in the image as part of the skin class or lesion class.

In the case of normal skin texture distributions, the dissimilarity of one skin texture distribution from other skin texture distributions is very small. The TD metric for skin

International Journal of Advanced Research in Biology, Ecology, Science and Technology (IJARBEST)
Vol. 1, Issue 2, May 2015

texture distributions is small overall. But lesion texture distributions are dissimilar from other normal skin texture distributions, so the textural distinctiveness metric is large. Based on this we can classify the skin region is cancerous or non-cancerous.

This paper has the following structure: section II explains about the related works, the proposed method is given in section III. Section IV and V gives implementation details and results and section VI concluded the paper.

II. RELATED WORKS

Image segmentation is the division of an image into regions or categories, which correspond to different objects or parts of objects. Every pixel in an image is allocated to one of these categories. Segmentation of melanocytic lesions can be an extremely hard problem. Besides the presence of hair, many lesions present diffuse borders, difficult to determine even for dermatologists.

Skin lesions are typically visually screened for melanoma. Visual algorithms that doctors used as a guide to assess skin lesions based on the ABCD [3] scale. The ABCD scale is an acronym for Asymmetry, Border irregularity, Colour variegation, and Diameter ,and has been proposed to be expanded to include Evolving. These guidelines are somewhat helpful, but the problem is that many normal moles are not completely symmetrical in their shape or color.

Many segmentation algorithms have been proposed to locate skin lesion in images automatically. The majority of proposed segmentation algorithms are only applicable to dermoscopy images, which has better contrast between the lesion and surrounding skin area for certain types of lesions. Algorithms include using simple thresholding[4], active contours, and region merging[5] .The majority of algorithms only use features derived from pixel color to drive the segmentation. This includes the blue channel from the RGB color space, the luminance channel from the CIELUV etc .But it is difficult to accurately segment lesions with fuzzy edges when relying only on color features.

Thresholding is based on the notion that regions corresponding to different regions can be classified by using a range function applied to the intensity values of image pixels. Lesion segmentation can be obtained by comparing the color of each pixel with a threshold .The pixel is classified as active (lesion) if it is darker than the threshold .The output of this step is a binary image. But it is useful in discriminating foreground from background. The major drawback to threshold-based approaches is that they often lack the sensitivity and specificity needed for accurate classification. The problem gets severe in case of multi-modal histograms with no sharp or well-defined boundaries and only applicable to dermoscopy images.

Finally, region-based segmentation algorithms operate iteratively by grouping together pixels which are neighbors and have similar values and splitting groups of pixels which are dissimilar in value. Consists of initially subdividing the image into a set of arbitrary, disjoint regions then to merge and/or to split the regions in an attempt to satisfy the conditions .The two key steps of the algorithm are as follows: First ascertain a sort function, by which the adjacent region are sorted according to the size of the function. Secondly ascertain a merging predicate, which confirms whether the adjacent regions are merged or not. It is obvious that sort function and merging predicate are basis of the algorithm and they are interactive with each other. It is difficult to identify the point to split a region and not provide a unique solution, when starting with a different initial partition, we get a different result. This algorithm does not guarantee that each subset is merged with the best of its supersets .A small variation in the segmentation process induces large differences in the result.

III.PROPOSED METHOD

The segmentation algorithm based on texture distinctiveness (TD), used to locate skin lesions in photographs is referred to as the Texture Distinctiveness Lesion Segmentation (TDLS) algorithm. The TDLS algorithm can be implemented in two steps. In first step, a set of sparse texture distributions that represent skin and lesion textures are learned and a TD metric is calculated to measure the dissimilarity of a texture distribution. In Second step, the TD metric is used to classify regions in the image as part of the skin class or lesion class. The Figure3.1 shows the overall process to learn the representative texture distributions and calculate the TD metric.

A. Texture Distinctiveness

Dictionary-learning algorithm is a common method to learn a sparse texture model . In this case a set of texture patches that can best match details in the original image is learned. Here we use probabilistic information to learn sparse texture distributions, rather than texture models. To learn whether each texture distribution belongs to the skin or lesion class, a TD metric is formulated. Texture gives us information about the spatial arrangement of the colors or intensities in an image.

International Journal of Advanced Research in Biology, Ecology, Science and Technology (IJARBEST)
Vol. 1, Issue 2, May 2015

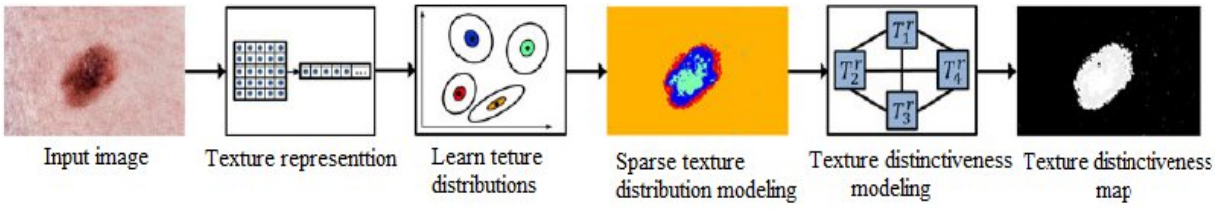

Fig .1: Algorithm flowchart displaying the steps to learn the representative texture distribution and calculate TD metric

B. Representative Texture Distribution

The main advantage of using a joint probabilistic sparse model is that the .sparse texture distributions can model both local and global texture characteristics. To learn the sparse texture model, a local texture vector is obtained for each pixel in the image. The input image contains N ×M pixels and each pixel has a channels. The texture vector contains pixels in a neighborhood of size n centered on the pixel of interest. Let s be a pixel location (x, y) in the photograph. Then, the vector **t**s represents the n × n × a texture patch centered at pixel s.

In the case of multiple channels, $\mathbf{t}_{A,s}$ is the texture patch centered at pixel s and corresponding to channel A. The texture vector is constructed by concatenating each $\mathbf{t}_{A,s}$ corresponding to the same pixel across all channels. For example, if the color image contains three channels {R,G,B} for each pixel, three texture vectors, $\mathbf{t}_{R,s,}$ $\mathbf{t}_{G,s}$, and $\mathbf{t}_{B,s}$, are extracted and concatenated such that $\mathbf{t}_s = [\mathbf{t}_{R,s,}\ \mathbf{t}_{G,s}\ \mathbf{t}_{B,s}]$. After extracting the set of texture vectors for an image, we get a set of N ×M texture vectors , with each vector of size n × n ×a:

$$\mathcal{T} = \{\boldsymbol{t}_{s_j} | 1 \le j \le N \times M\}. \qquad (1)$$

To find the representative texture distributions and the sets of texture vectors corresponding to each representative distribution, an unsupervised clustering algorithm is used. The k[th] representative texture distribution is defined as τ_k^r . A small set T^r comprised of K representative texture distributions can be used instead of using all the local texture vectors, to reduce the computational complexity and memory requirements ,

$$\mathcal{T}^r = \{T_k^r | 1 \le k \le K\}. \qquad (2)$$

For learning texture distribution we assume a Gaussian probability distribution , so θ_k contains the two required parameters to define a multivariate Gaussian distribution. The mean and covariance of the k[th] texture distribution are represented by $\mathbf{t}_{rk}$ and Σk , respectively and $P(\mathbf{t}_{sj} | T^r{}_k)$ is the probability of the j[th] texture vector given the parameters of the k[th] texture distribution. Texture distributions are chosen to maximize the log-likelihood of the mixture model,

$$\hat{T}^r = \arg\max_{T^r} \sum_{k=1}^{K} \sum_{\boldsymbol{t}_{s_j} \in C_k} \log\left(\mathrm{P}(\boldsymbol{t}_{s_j} | \mathrm{T}_k^r)\right). \qquad (3)$$

C. TD Metric

The next step is to formulate a TD metric using the learned sparse texture model. Since we are only interested in two classes, normal skin and lesion, but have learned many texture distributions. Multiple texture distributions must represent the same class. To measure similarity of two texture distributions, we first measure the probability that the mean of one texture distribution is a realization of the mean of the other texture distribution, which is defined in Eqn (4). Because we assume that the texture distributions are Gaussian, $\mathbf{t}_{rj}$ and Σj are the mean and covariance of distribution Tr. After Lj,k has been calculated for each pair of texture distributions, they are normalized to be between 0 and 1.

$$l_{j,k} = \frac{1}{\sqrt{(2\pi)^{n\times n\times a}|\Sigma_j|}} \exp\left(-\frac{1}{2}(t_j^r - t_k^r)^T \Sigma_j^{-1}(t_j^r - t_k^r)\right) \qquad (4)$$

$$L_{j,k} = \frac{1}{2}(l_{j,k} + l_{j,k}). \qquad (5)$$

For example, lesion texture distributions are both dissimilar from the normal skin texture distributions and also from other texture distributions, due to color variegation and textural patterns found in skin lesions. The probability that a texture distribution is distinct from another texture distribution is given by $d_{j,k}$:

$$d_{j,k} = 1 - L_{j,k}. \qquad (6)$$

A TD metric D_j is used to capture the dissimilarity of texture distribution from other texture distributions. The metric is defined in (3.6) measures the expected distinctiveness of $T^r{}_j$ given the photograph I, where $P(T^r{}_k | I)$ is

International Journal of Advanced Research in Biology, Ecology, Science and Technology (IJARBEST)
Vol. 1, Issue 2, May 2015

the probability of occurrence of a pixel being associated with a texture distribution T^r_k ,

$$D_j = \sum_{k=1}^{K} d_{j,k} P(T_k^r | I). \quad (7)$$

D. Region Classification

The last step in the TDLS algorithm is to find and classify regions in the input image as being part of the lesion based on the sparse texture distributions and their associated TD metric. First each region is independently classified as representing normal skin or lesion based on the textural contents of that region. Then, post processing steps refine the lesion segmentation.

$$y(R) = \begin{cases} 1, & \text{if } \mathcal{D}_R \geq \tau \text{ (lesion)} \\ 0, & \text{otherwise (normal skin).} \end{cases} \quad (8)$$

Where texture distribution metric across the region R,

$$\mathcal{D}_R = \sum_{j=1}^{K} D_j P(T_j^r | R). \quad (9)$$

In the TDLS algorithm, the threshold is found that divides the set of texture distributions into two classes such that the total intraclass variance of the TD metric for each class is minimized as

$$\tau = \arg\min_{\tau} \left(\sigma^2_{C_1(\tau)} P\left(T^r | C_1(\tau)\right) + \sigma^2_{C_2(\tau)} P\left(T^r | C_2(\tau)\right) \right). \quad (10)$$

The threshold τ is used to divide the set of texture distributions into two classes C1(τ) and C2 (τ).The probability that a texture distribution is in the class C for a given τ is P(Tr |C(τ)) and the variance of the TD based on the elements in the class is $\sigma_{C(\tau)}$. This threshold is known as the Otsu's threshold.

After the regions are classified as being normal skin or lesion, the following post processing steps are applied to refine the lesion border: morphological dilation and region selection. First, the morphological dilation operator is applied to fill holes and smooth the border. Morphological dilation is a process that expands binary masks to fill small holes .The shape and amount that the binary mask is expanded is controlled by a structuring element, which is a disc with a radius of 5 pixels in the TDLS algorithm.

IV. IMPLEMENTATION DETAILS

A. Color Space

The input image is converted to the XYZ color space .It was found that the XYZ color space be an efficient color space for the segmention of the skin region of human faces[6] .This colour space is designed to better model colour perception and reduce correlation between the XYZ channels, compared to the standard RGB colour space.

B. Learning Representative Texture Distributions

For learning representative texture distribution, a two-step clustering algorithm is used. First, a *k*-means clustering algorithm is run, which is followed by expectation-maximization algorithm for learning a finite mixture model. *K*-means clustering[7] is used as an initial step to increase the robustness and to speed up the number of iterations required for the finite mixture model to converge. *K*-means clustering finds *K* clusters of texture data points that minimizes C in Eqn(11)the sum of squared error between cluster members and the cluster mean and is given by

$$\hat{C} = \arg\min_{C} \sum_{k=1}^{K} \sum_{t_{s_j} \in C_k} \|t_{s_j} - \mu_k\|^2. \quad (11)$$

The main limitation with *k*-means clustering is that it does not take into account any probabilistic information .So in the second step we apply a finite mixture model clustering which maximize the log-likelihood function.

Expectation-maximization is an iterative algorithm[8]. The initial parameters for the Gaussian mixture model are obtained from the results of the *k*-means clustering. That is, the initial Gaussian means are equal to the *k*-means cluster means μ_k then calculate covariance k and mixing coefficient α_k in eqn (12)and update the value of TD metric .

$$\begin{aligned} \mu_k &\leftarrow \mu_k \\ \Sigma_k &\leftarrow \text{cov}(cluster(K)) \\ \alpha_k &\leftarrow \frac{\text{Number of points in k}}{\text{Total number of points}} \end{aligned} \quad (12)$$

V. EXPERIMENTAL RESULTS

The TDLS algorithm is implemented using MATLAB software. The images from the Dermquest database [9] are used to test the segmentation algorithms. Fig 2 represent the segmented image of cancerous skin and border of the lesion is detected here. The Fig 3 shows non cancerous skin and we cannot see the border because of the similarity of texture pattern as normal skin.

International Journal of Advanced Research in Biology, Ecology, Science and Technology (IJARBEST)
Vol. 1, Issue 2, May 2015

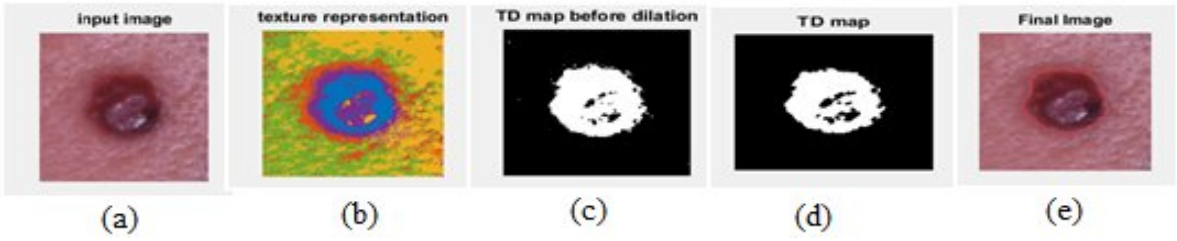

Fig.2.(a)input image (b)texture representation (c)TD map before (d)after dilation and (e)final segmented image of a cancerous skin

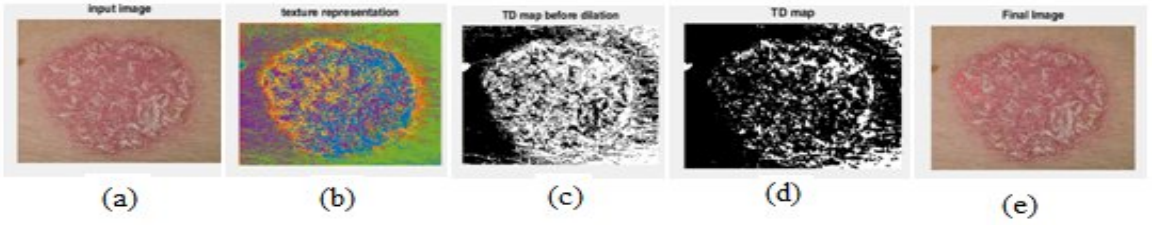

Fig.3 (a)input image (b)texture representation (c)TD map before (d)after dilation and (e)final segmented image of a non-cancerous skin

VI. Conclusion

A new lesion segmentation algorithm using the concept of texture distinctiveness is proposed. A probabilistic TD metric is introduced based on a learned model of normal skin and lesion textures. Representative texture distributions are learned from the image itself and the TD metric captures the dissimilarity between pairs of texture distributions. Then, the image is classified as lesion or skin based on the TD map. The experimental results show that the proposed method is able to segment the lesion in images of different scales and levels of quality.

Acknowledgment

This work is supported and guided by my research guide. I am very thankful to my research guide Mr. Manu Mohan C M, Assistant Professor, Electronics and Communication Engineering, Federal Institute of Science and Technology (FISAT), India, for her guidance and support.

References

1 A. 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

2 H. C. Engasser and E. M.Warshaw, "Dermatoscopy use by US dermatologists: A cross-sectional survey,"J. Amer. Acad. Dermatol., vol. 63, no. 3, pp. 412–419, 2010

3 R. J. Friedman, D. S. Rigel, andA. W.Kopf, "Early diagnosis of cutaneous melanoma: Revisiting the ABCD criteria," *CA: A Cancer J. Clinicians*, vol. 35, no. 3, pp. 130–151, May 1985.

4 B. Erkol, R. H. Moss, R. Joe Stanley, W. V. Stoecker, and E. Hvatum,"Automatic lesion boundary detection in dermoscopy images using gradient vector flow snakes," Skin Res. Technol., vol. no. 1, pp. 17–26, 2005.

5 M. E. Celebi, H. A. Kingravi, H. Iyatomi, W. V. Stoecker, R. H. Moss, J. M. Malters, J. M. Grichnik,A,"Border detection in dermoscopy images using statistical region merging,"Skin Res.Technol., vol. 14, no. 3, pp. 347–353, 2008.

6 .J.-C. Terrillon, M. Shirazi, H. Fukamachi, and S.Akamatsu, "Comparative performance of different skin chrominance models and chrominance spaces for the automatic detection of human faces in color images," in Proc. 4th IEEE Int. Conf. Autom. Face Gesture Recog., 2000, pp. 54–61.

7 M. Celebi, H. Kingravi, and P. A. Vela, "A comparative study of efficient initialization methods for the k-means clustering algorithm," Expert Syst. Appl., vol. 40, no. 1, pp. 200–210, Sep. 2012.

8 M. A. T. Figueiredo and A. K. Jain, "Unsupervised learning of finite mixture models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 3,pp. 381–396, Mar. 2002.

9 DermQuest,(2012).[Online].vailable:http://www.dermque st.com

Bindhu S received the BTech degree in Electronics and Engineering from Cusat University, Kerala, India, in 2008. Currently, she is post graduate student with the Department of Communication Engineering, Federal Institute of Science and Technology (FISAT), Kerala, India. Her current research area includes antenna design and image processing.

Manu Mohan C M received the Bachelor's degree in Electronics and Communication Engineering in 2010, from MG University, Kerala,India and M.Tech in Communication Engineering from Federal Institute of Science and Technology (FISAT), Kerala, India in 2012. Since February 2014, he has been working as a Assistant Professor in Federal Institute of Science and Technology (FISAT), Kerala, India