ISSN 2349-7815

International Journal of Recent Research in Electrical and Electronics Engineering (IJRREEE)
Vol. 2, Issue 4, pp: (79-85), Month: October - December 2015, Available at: www.paperpublications.org

Skin Lesion Segmentation Using TDLS Algorithm and Pattern Identification

Anu Thankaraj

M tech Student, Electronics and Communication Department, MG University, India

***Abstract*: The rapid growing rate of Melanoma makes researchers to study in depth about it. It is the most deadliest form of skin cancer. Due to the higher costs for dermatologists, images of skin lesion of the patient are taken and texture-based skin lesion segmentation algorithm is proposed. Different regions belonging to different textures are studied. Finally the image is classified as normal skin or lesion. Different model based pattern classification of the lesion is proposed. We mainly classify the whole pigmented lesion into three possible patterns: globular, reticular, and homogeneous.**

Keywords: **Melanoma, segmentation, texture, classification.**

I. INTRODUCTION

Melanoma, also known as malignant melanoma, is a type of cancer that develops from the pigment-containing cells known as melanocytes. Melanomas typically occur in the skin but may rarely occur in the mouth, intestines, or eye. In women they most commonly occur on the legs, while in men they are most common on the back .Sometimes they develop from a mole with concerning changes including an increase in size, irregular edges, change in color, itchiness, or skin break down. Melanoma is the most dangerous type of skin cancer. Globally, in 2012, it occurred in 232,000 people and resulted in 55,000 deaths. Australia and New Zealand have the highest rates of melanoma in the world. There are also high rates in Europe and North America while it is less common in Asia, Africa, and Latin America. They are more common in men than women. Melanoma has become more common since the 1960s in areas that are mostly Caucasian. With the rising incidence rates in certain subsets of the general population, it is beneficial to screen for melanoma in order to detect it early. To reduce costs of screening melanoma in the general population, development of automated melanoma screening algorithms have been proposed. . Images acquired through a digital dermatoscope are referred to as dermoscopy images and have relatively low levels of noise and consistent background illumination. Pattern analysis seeks to identify specific patterns, which may be local or global. The melanocytic lesions are identified by their general dermoscopic features, defining their global pattern, or by specific dermoscopic criteria that determine their local patterns.

Thus, a lesion is categorized by a global pattern, although it can present more than one local pattern. Global features permit a broad classification of pigmented skin lesions, while a description of the local features provides more detailed information about a given lesion [5]. The local features represent individual or grouped characteristics that appear in the lesion. The global features are presented as arrangements of textured patterns covering most of the lesion. The main global patterns are: Reticular pattern, Globular pattern, Cobblestone pattern, Homogeneous pattern, Parallel pattern, Starburst pattern, and Multicomponent pattern. They are associated with the predominant local pattern: Reticular pattern with pigment network, Globular pattern with globules, Cobblestone pattern with globules, Homogeneous pattern with pigmentation, Parallel pattern with furrows and ridges, Starburst pattern with streak, and Multicomponent pattern with a combination of three or more above patterns.

The main aim of this paper is the classification of a entire pigmented lesion into Reticular pattern, Globular pattern, or Homogeneous pattern by texture analysis. Likewise, in a further evaluation the Multicomponent pattern is analyzed.

There are different reasons behind this decision instead to address the classification of the seven patterns mentioned above. Globules are also predominant in the Cobblestone pattern, however they are larger and more closely aggregated than in Globular pattern, for what can be considered a special case of Globular pattern. Consequently, in our database, images belonging to Cobblestone pattern have been included in the Globular class. Regarding Parallel pattern, its automatic detection does not have a significant interest for the clinical community because lesions with this pattern are only located in palm or sole. Starburst pattern is characterized by the presence of pigmented streaks at the edge of a given lesion. As our objective is the texture analysis of an entire lesion, this type of lesion escapes from our study.

Pattern analysis allows to dermatologist not only the distinction between benign and malignant growth features but it also determines the type of a lesion. 1) Reticular pattern represents the dermoscopic hallmark of benign acquired melanocytic nevi in general and of thin melanomas in particular; 2) Globular pattern and the Cobblestone pattern are commonly seen in congenital nevus, superficial type; 3) Homogeneous pattern represents morphologic hallmark of blue nevus.

Segmenting digital photographs of skin lesions is a more difficult problem due to illumination variation. Special segmentation algorithms are required to take into account illumination variation, which causes shadows and bright areas to appear throughout the photograph. Hance *et al.* [6] explored different algorithms, including thresholding, active contours and split-and merge, and modified them to be usable on lesion photographs. For example, the thresholding algorithm has to be modified to account for bright areas where there is reflection of the camera's flash. In this paper, we propose a segmentation algorithm based on texture distinctiveness (TD) to locate skin lesions in photographs. This algorithm is referred to as the TD lesion segmentation (TDLS) algorithm. The main contributions are the introduction of a joint statistical TD metric and a texture-based region classification algorithm.TD captures the dissimilarity between learned representative texture distributions

II. SEGMENTATION

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 from all other texture distributions.Second, the TD metric is used to classify regions in the image as part of the skin class or lesion class. In this section, the first step is described in detail and Fig. 2 illustrates the overall process to learn the representative texture distributions and calculate the TD metric

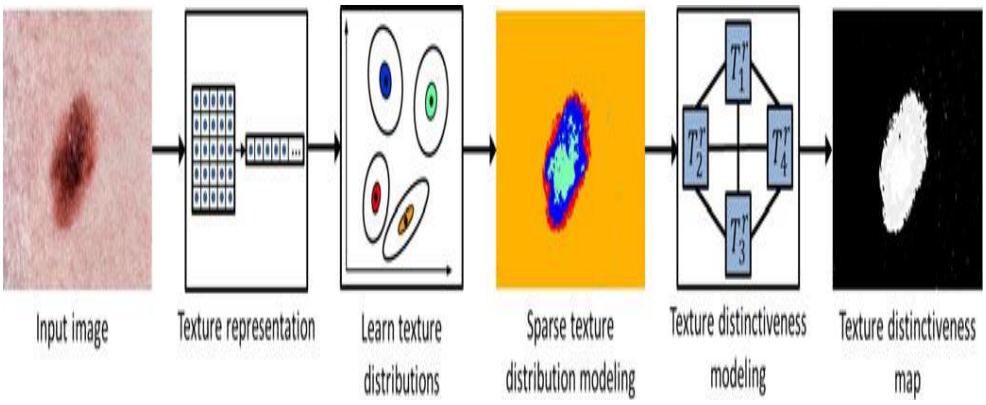

Fig.1.TDLS Algorithm

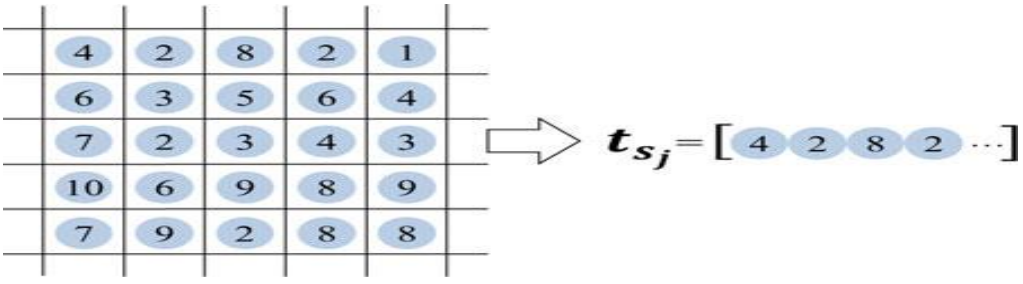

Fig.2.Texture vector extraction

ISSN 2349-7815

International Journal of Recent Research in Electrical and Electronics Engineering (IJRREEE)
Vol. 2, Issue 4, pp: (79-85), Month: October - December 2015, Available at: www.paperpublications.org

An existing sparse texture model algorithm is modified to find representative sparse texture distributions from the input photograph. Our proposed sparse texture model algorithm incorporates statistical information. The advantage of using a joint probabilistic sparse model is that the sparse texture distributions can model both local and global texture characteristics. After extracting the set of texture vectors for an image, we have a set of $N \times M$ texture vectors is extracted, with each vector of size $n \times n \times a$.[1].

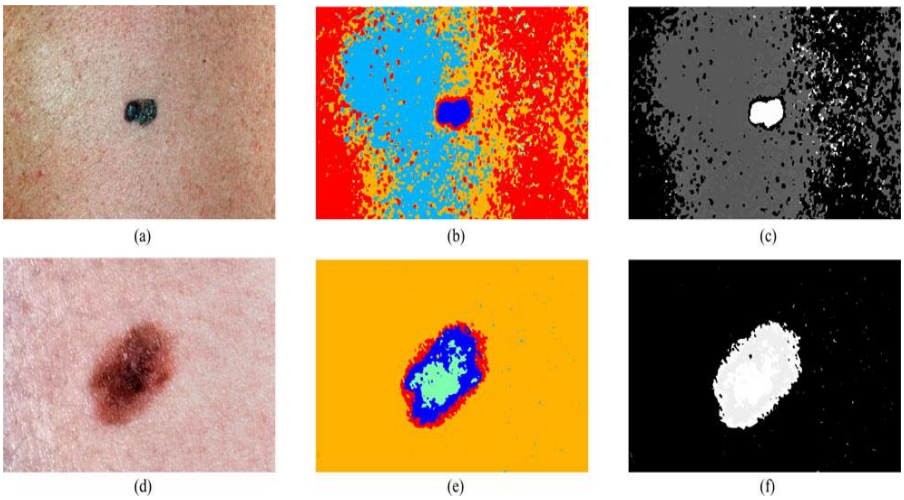

Fig.3.TD Map

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 texture distributions is small overall. Lesion texture distributions are dissimilar from other skin and lesion texture distributions, so the textural distinctiveness metric is large. Fig. 4(c) and (f) give illustrative examples of the TD metric corresponding to each pixel in the images. A brighter pixel corresponds to a higher TD metric. In both figures, the lesion is predominately white, meaning that the lesion texture distributions have higher TD metrics, as expected. In Fig. 3(f), there are two texture distributions that correspond to the lesion class and have high TD. However, in Fig. 3(c), some normal skin pixels to the right of the lesion also have high TD. This can occur when there are unique texture patterns in normal skin areas. This commonly occurs, motivating the region classification step of the TDLS algorithm. The region classification step allows the algorithm to be more robust and minimize misclassification of pixels.

A. Region Classification:

The second main 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, the image is oversegmented, which results in the image being divided into a large number of regions. Next, each region is independently classified as representing normal skin.The corrected lesion image is divided into a large number of regions. This initial oversegmentation step is incorporated to increase the TDLS algorithm's robustness to noise. SRM [1] contains two main steps: a sorting step and a merging step. SRM sorts pixels in an image to determine the order in which pixels are compared, and then merges pairs of pixels into regions based on their similarity. A four-connected graph is constructed so that each pixel in the photograph is connected with its neighbors. The pixels are sorted based on their similarity with their neighboring pixel. Both horizontal and vertical neighboring pixels are considered when sorting the pixels [1].

B. Summary of the TDLS Segmentation Algorithm:

1) Convert the corrected image to the XYZ color space.

2) For each pixel *s* in image *I*, extract the texture vector to obtain the set of texture vectors [1]

3) Cluster the texture vectors to obtain the representative texture distributions.

4) Calculate probability that two texture distributions are distinct for all possible pairs of texture distributions.

5) Calculate the textural distinctiveness metric for each texture distribution.

6) Apply the SRM algorithm to find the initial regions.

7) Calculate the region distinctiveness metric for each initial region .

8) Calculate the threshold τ between the normal skin and lesion classes .

9) Classify each region as normal skin or lesion based on the results of steps 7 and 8 .

10) Apply a morphological dilation operator to the initial lesion classification.

11) For each contiguous region in the initial segmentation, count the number of pixels in the region[1]

IV. PROPOSED MODEL

In this section, the proposed model-based classification methods are detailed. The aim is the classification of a whole lesion, not only of a sample or patch of it. It is important to note that, in this paper, two different training sets of images are used, depending on the method implemented. Complete lesions compose the first dataset, whereas the second set is constituted by individual patches, each patch extracted from a different lesion of the first dataset. The extraction of these patches was performed randomly. The test set is constituted by complete lesions. None of the lesions included in the test dataset are included in the training dataset. In order to analyze a whole lesion, the lesion is divided into overlapping patches. Taking into account that our images have a spatial resolution of 768 ×512 pixels, different patch sizes were tested: 40× 40 , 50× 50 ,81 ×81 and 100 ×100. Finally, patch size was fixed to 81× 81 pixels achieving a trade-off between computational cost and size that should be large enough to distinguish and detect different textures. A displacement equal to nine rows or/and nine columns on the lesion is applied to obtain the next patch. A displacement of 27 rows or/and 27 columns instead of nine is shown in Fig. 4(e) and (f) in order to be appreciated. In Fig. 4 individual patches of the three global patterns under study as well as an example of a lesion divided into overlapping patches can be seen. Only the patches without background or with a background area of up to 10% the patch area are taken into account.

A. Gaussian Model-Based Method:

This approach is based on the assumption that the MRF features of the patches or samples constituting a test lesion follow a multivariate Gaussian distribution model[2]

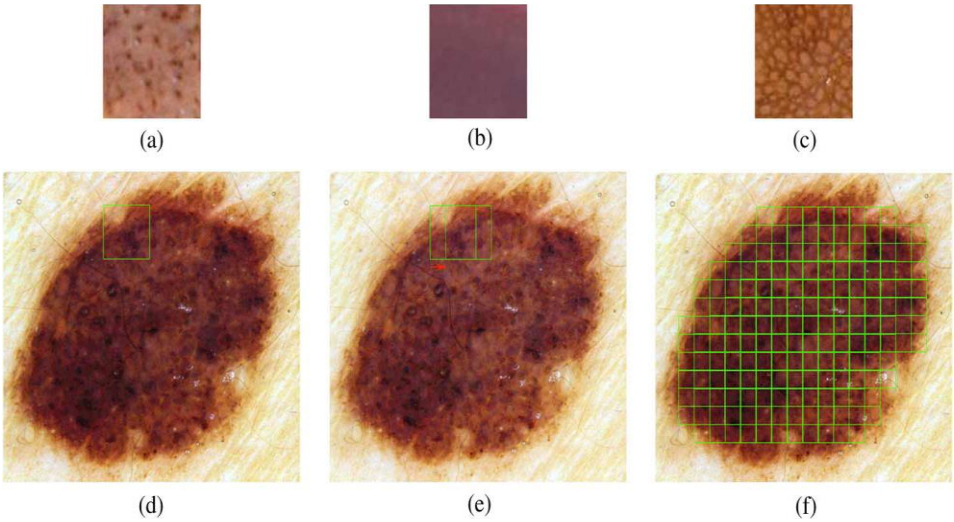

Fig.4.Examples of individual patches and complete lesions

Apart from this assumption, two different scenarios regarding to the training set have been considered.

1) GM1: the training set is constituted by individual patches. MRF features of each class in this training set are supposed to follow a multivariate Gaussian distribution

2) GM2: full lesions constitute the training set.MRF features of the patches within each training lesion are supposed to follow a multivariate Gaussian distribution.

Different distance metrics are used in order to compare the multivariate Gaussian distributions of the test lesion and those from the training sets. Symmetric Kullback–Leibler distance [9], Bhattacharyya distance [10] and Frechet distance [11], which is the closed form solution of the earth movers distance (EMD) in the case of two Gaussian distributions, are analysed

B .Gaussian Mixture Model-Based Methods:

In this approach MRF features extracted from patches constituting a test lesion are supposed to follow a Gaussian mixture model. Based on this assumption, other two scenarios regarding to the training set are considered[2].

1) GMM1: individual patches constitute the training set. The MRF features of the individual training patches belonging to each class follow a Gaussian mixture distribution

2) GMM2: the training set consists of full lesions that are supposed to follow a Gaussian mixture distribution different distance metrics between Gaussian mixture models are used: the symmetric Kullback–Leiblerdivergence[12], the Bhattacharyya-based distance metric [12], EMD

C. Bag of Features:

The last approach is based on the representation of an image as a bag of features (BoF). This approach finds its origin, on the one hand, in the texture recognition by textons (basic elements of texture) and, on the other hand, in the bag of words scheme used for text categorization and text retrieval [4]. The idea is to model an image as a frequency histogram of visual words (bag of features). These visual words are built from the quantification of descriptors (in our case the descriptors are MRF features) of local patches sampled from the training set. This quantification is usually carried out by a clustering algorithm such as k-means. The centroid of each cluster represents a visual word. The set of visual words forms a codebook[2].

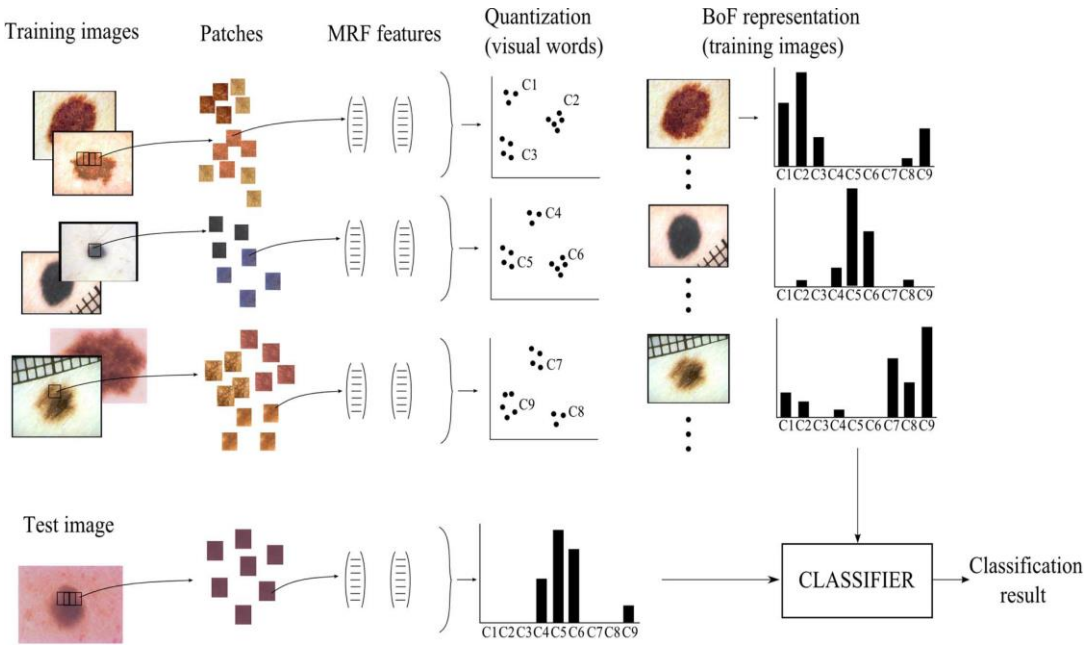

Fig.5.BoF approach

V. IMAGE DATABASE

The image database used in this work is formed by 30 images of each type of pattern, a total of 90 images. These 30 images from each global pattern were randomly chosen. However it should be emphasized that some low quality images (blurry or low-contrast images) had to be replaced. This is due to the fact that they have been acquired in different hospitals without following an acquisition protocol. As it has already been mentioned, globules are predominant in Globular and Cobblestone pattern, however, for the second case, they are larger and more closely aggregated than in Globular pattern. Thus, Cobblestone pattern can be considered a special case of Globular pattern. Eight images of the 30

categorized as globular pattern, belong to Cobblestone pattern. All images were extracted from the Interactive Atlas of Dermoscopy, published by Edra Medical Publishing New Media [3], which is a multimedia project for medical education with images of pigmented skin lesions from different centers and hospitals. The selected database include both images with a clear diagnosis and images difficult to classify depending on the type of the lesion. Each image presents an unique global pattern.

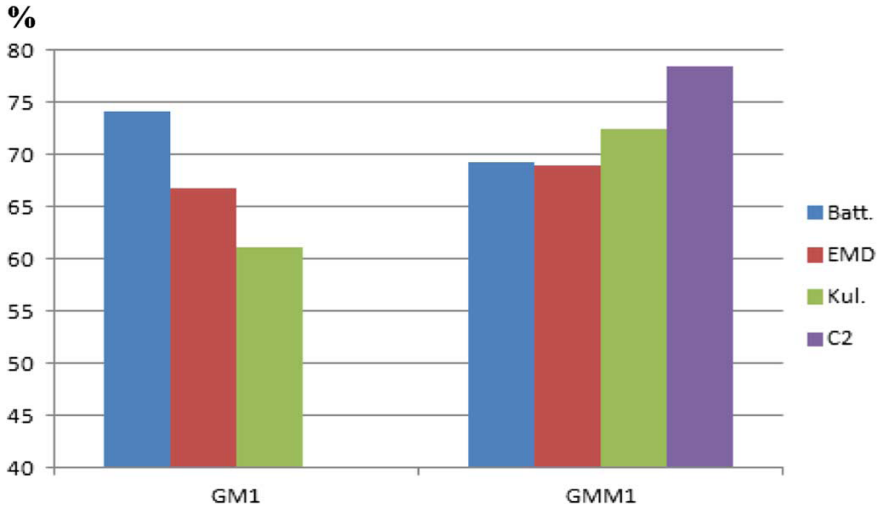

Fig.6: Performance of individual patches

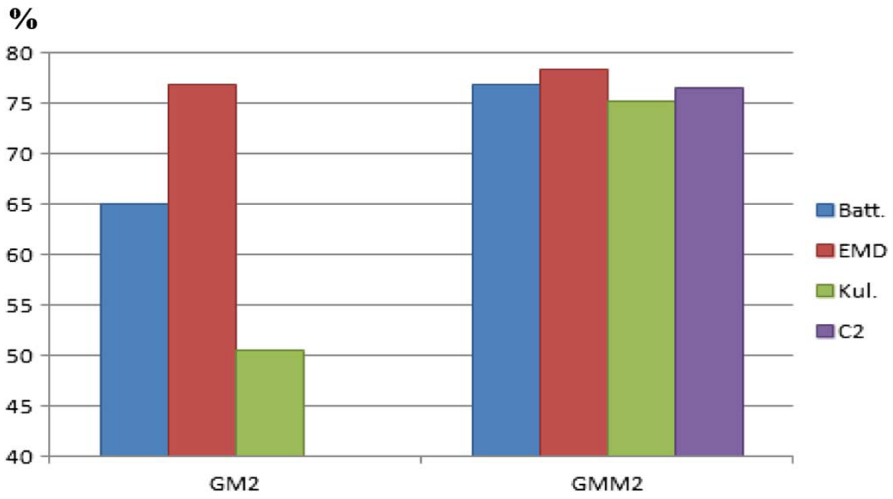

Fig.7: Performance of complete lesions.

VI. CONCLUSION

In summary, a novel lesion segmentation algorithm using the concept of TD 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 divided into numerous smaller regions and each of those regions are classified as lesion or skin based on the TD map. The entire proposed framework is tested by using the illumination corrected images as the input to the texture-based segmentation algorithm. It is compared to state-of-art lesion segmentation algorithms, including three algorithms designed for lesion images. The proposed framework produces the highest segmentation accuracy In this paper, different classification methods for global dermoscopic patterns have been proposed. The aim is to classify each lesion as a particular global pattern. This unique-label classification is motivated by the fact that a lesion is characterized by a global pattern and by one or more local patterns. The majority of the classification approaches in the literature are based on a feature extraction step followed by a classifier whose inputs are the features extracted. On the contrary, this paper proposes techniques based on modeling in different senses. First, an image is modeled by a MRF on the color space. The

estimated parameters of this model are treated as features. And then, these features within a lesion are supposed to follow three different models. In the first one, it is supposed that a lesion follows a multivariate Gaussian distribution. The idea is to measure distances between Gaussian models (GM) and then to apply a KNN algorithm. The same idea remains in the second approach proposed although a GMM assumption substitutes to GM. As in the previous case different distance metrics between GMMs are analyzed. The third model-based classification technique is a Bag of Features approach, where a image is modeled by a frequency histogram of visual words. In this case, different distances between histograms have been studied.

REFERENCES

[1] Jeffrey Glaister,Alexander Wong and David A. Clausi, ,"segmentation of skin lesions from digital images using joint statistical texture distinctiveness" Biomedical Engineering, Vol. 61, No. 4, April 2014.

[2] Aurora Sáez*, Carmen Serran and Begoña Acha,"Model based classification methodsof global patterns in dermoscopic images" Medical Imaging, Vol. 33, No. 5, May 2014.

[3] G. Argenziano and H. Soyer et al., Interactive Atlas of Dermoscopy.Milan, Italy: EDRA-Medical Publishing New Media, 2000.

[4] D. D. Lewis, "Naive (Bayes) At Forty: The Independence Assumption In Information Retrieval," In Proc. 10th Eur. Conf. Mach. Learn., 1998, Pp. 4–15.

[5] H. Soyer, G. Argenziano, V. Ruocco, and S. Chimenti, "Dermoscopy of pigmented skin lesions (Part II)," *Eur. J. Dermatol.*, vol. 11, no. 5, 2001.

[6] [6] G. Hance, S. Umbaugh, R. Moss, and W. Stoecker, "Unsupervised color image segmentation: with application to skin tumor borders," IEEE Eng. Med. Biology Mag., vol. 15, no. 1, pp. 104–111, Jan./Feb. 1996

[7] C. Scharfenberger, A. Wong, K. Fergani, J. S. Zelek, and D. A. Clausi,"Statistical textural distinctiveness for salient region detection in natural images," in Proc. IEEE Conf. Comput. Vis. Pattern Recog., Jun. 2013,pp. 979–986.

[8] 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.

[9] S. Kullback, *Information Theory and Statistics*. Mineola, NY: Dover, 1997.

[10] A. Bhattacharyya, "On a measure of divergence between two statistical populations defined by their probability distributions," Bull. Calcutta Math. Soc., vol. 35, pp. 99–109, 1943.

[11] D. Dowson and B. Landau, "The Fréchet distance between multivariate normal distributions," J. Multivariate Anal., vol. 12, no. 3, pp. 450–455, 1982.

[12] G. Sfikas, C. Constantinopoulos, A. Likas, and N. Galatsanos, An Analytic Distance Metric for Gaussian Mixture Models With Application in Image Retrieval. New York: Springer, 2005, vol. 3697, LNCS, pp. 835–840.