Implementation of the 7-Point Checklist for Melanoma Detection on Smart Handheld Devices

Tarun Wadhawan, Ning Situ, Hu Rui, Keith Lancaster,
Xiaojing Yuan, *Member IEEE*, and George Zouridakis, *Senior Member IEEE*

Abstract— In this paper we implement the 7-point checklist, a set of dermoscopic criteria widely used by clinicians for melanoma detection, on smart handheld devices, such as the Apple iPhone and iPad. The application developed is using sophisticated image processing and pattern recognition algorithms, yet it is light enough to run on a handheld device with limited memory and computational speed. When combined with a commercially available handheld dermoscope that provides proper lesion illumination, this application provides a truly self-contained handheld system for melanoma detection. Such a device can be used in a clinical setting for routine skin screening, or as an assistive diagnostic device in underserved areas and in developing countries with limited healthcare infrastructure.

Fig. 1. *The SkinScan application running on the Apple iPhone.*

I. INTRODUCTION

The new generation of smart handheld devices with sophisticated hardware and operating systems has provided a portable platform for running medical diagnostic software, such as the heart rate monitoring [1], diabetes monitoring [2], and experience sampling [3] applications, which combine the usefulness of medical diagnosis with the convenience of a handheld device. Their light operating systems, such as the Apple iOS and Google Android, the support for user friendly touch gestures, the availability of an SDK for fast application development, the rapid and regular improvements in hardware, and the availability of fast wireless networking over Wi-Fi and 3G make these devices ideal for medical applications. In this paper we present an innovative application for melanoma detection running on Apple iOS based smart devices, like the iPhone and iPad. Our long term objective is to develop an automated system for melanoma detection deployed on handheld devices, based on well known diagnostic criteria used by dermatologists to analyze dermoscopic images. Dermatologist use pattern analysis to detect the presence of certain features or patterns of texture and color in a skin lesion. Some widely used criteria are the ABCD rule [4], Menzies rule [5], and the 7-point checklist [6]. The latter includes seven dermoscopic features that can

be detected with high sensitivity and decent specificity by even less experienced clinicians [6]. The seven points of the list are subdivided into three major and four minor criteria, reflecting their importance in defining a melanoma. To score a lesion, the presence of a major criterion is given two points and that of a minor one point. If the total score is greater than or equal to 3, the lesion is classified as melanoma.

We have recently developed a library of image processing and texture analysis algorithms for low level features of a lesion that runs on embedded devices [7]. In this paper we present an application for automatic detection of melanoma, called SkinScan. It uses high level features, such as those on the 7-point checklist, and is implemented on Apple iOS4, as shown in Figure 1. The major components of this application can be easily configured for the diagnosis of other skin conditions, such as wounds and ulcers.

II. METHODS

The 7-point checklist includes the following criteria:
1) Major criteria
    a) Atypical pigment network: Black, brown, or gray network with irregular meshes and thick lines.
    b) Blue-whitish veil: Confluent gray-blue to whitish-blue diffuse pigmentation associated with pigment network alterations, dots/globules and/or streaks.
    c) Atypical vascular pattern: Linear irregular or dotted vessels not clearly combined with regression structures and associated with pigment network alterations, dots/globules and/or streaks.

T. Wadhawan and N. Situ are with the Department of Computer Science and R. Hu and K. Lancaster are with the Department of Electrical and Computer Engineering.

X. Yuan (xyuan@uh.edu) is with the Department of Engineering Technology.

G. Zouridakis (zouridakis@uh.edu; phone: +1-713-743-8656; fax: +1-713-743-1250) is with the Departments of Engineering Technology, Computer Science, and Electrical and Computer Engineering, University of Houston, Houston, TX 77204-3058 USA.
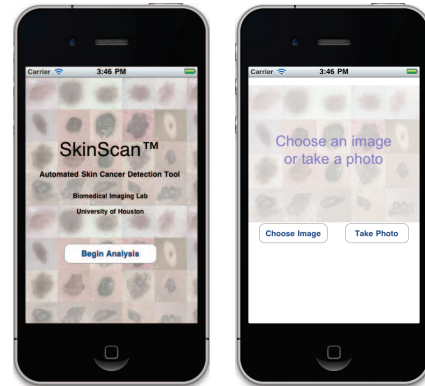
2) Minor criteria

    a) Irregular streaks: Irregular more or less confluent, linear structures not clearly combined with pigment network lines.

    b) Irregular pigmentation: Black, brown, and/or gray pigment areas with irregular shape and/or distribution.

    c) Irregular dots/globules: Black, brown, and/or gray round to oval, variously sized structures irregularly distributed.

    d) Regression structures: Associated white scar-like and gray blue, peppering, multiple blue gray dots.

A. Dataset

In this study, we used images from a large commercial library of skin cancer images annotated by expert dermatologists [8] that were uploaded to the phone. Intra-observer and inter-observer agreement could be low for certain criteria [9]. To demonstrate the feasibility of our automated system, we only chose images considered as low difficulty by the experts [8]. There were 385 low difficulty images in the database [8] and our segmentation methods could provide a satisfactory boundary for 347 (90.13%) of them. All images were segmented manually by one of the authors to provide a ground truth against which the automated techniques employed by our application were compared. In the selected set of 347 images: 110 were classified by the 7-point list as melanoma and the remaining 237 were classified as benign

B. Feature Extraction

A skin lesion image can be acquired using the smartphone camera (with or without an external attachment which can provide illumination and magnification) or can be loaded from the photo library to provide the diagnosis in real time. To identify a region of interest (ROI), an image is first converted to greyscale, and then fast median filtering [10] for noise removal is performed, and followed by ISODATA segmentation [11], and several morphological operations. From the ROI, we extract color and texture features relating to each criterion on the 7-point checklist, as follows.

1) Texture Features: They provide information on the various structural patterns [8] of 7-point checklist, such as pigmentation networks, vascular structures, and dots and globules present in a skin lesion. Elbaum et al. [12] used wavelet coefficients as texture descriptors in their skin cancer screening system MelaFind®. Our previous work [13] has demonstrated the effectiveness of Haar wavelet coefficients and local binary patterns [14] for melanoma detection.

Haar Wavelet: From the ROI we select non-overlapping $K \times K$ blocks of pixels, where K is a user defined variable. Computation time for feature extraction in directly proportional to the block size K. The block of pixels is decomposed using a three-level Haar wavelet transform [15] to get 10 sub-band images. We extract texture features by computing statistical measures, like the mean and standard deviation, on each sub-band image, which are then combined to form

a vector $W_i = \{m_1, sd_1, m_2, sd_2, ...\}$. Haar wavelet extraction for texture feature is as follows:

1) Convert the color image to greyscale; select a set of points in the ROI using a rectangular grid of size M pixels.
2) Select patches of size $K \times K$ pixels centered on the selected points.
3) Apply a 3-level Haar wavelet transform on the patches.
4) For each sub-band image compute statistical measures, namely mean and standard deviation, to form a feature vector $W_i = \{m_1, sd_1, m_2, sd_2, ...\}$.
5) For all feature vectors W_i extracted, normalize each dimension to zero-mean and unit-variance.
6) Apply K-means clustering [16] to all feature vectors W_i from all training images to obtain L clusters with centers $C = \{C_1, C_2, ..., C_L\}$.
7) For each image build an L-bin histogram H_i. For feature vector W_i, increment the jth bin of the histogram such that $\min_j \|C_j - W_i\|$.

The value of parameter M is a trade-off between accuracy and computation speed. When the algorithm runs on a handheld device we want to reduce computation time, so we choose $M = 10$ for the grid size, $K = 24$ for patch size, and $L = 200$ as the number of clusters in the feature space. By exhaustive parameter exploration in our previous studies [13], we found that these parameters are reasonable settings.

Local Binary Pattern (LBP): LBP is a robust texture operator [14] defined on a greyscale input image. It is invariant to monotonic transformation of intensity and invariant to rotation. It is derived using a circularly symmetric neighbor set of P members on a circle of radius R denoted by LBP_{PR}^{riu} [14]. The parameter P represents the quantization of angular space in the circular neighborhood, and R represents the spatial resolution. A limited number of transitions or discontinuities (0/1 changes in LBP) are allowed to reduce the noise and for better discrimination of features. We restrict the number of transitions in LBP to P, and transitions greater than that are considered equal. An occurrence histogram of LBP with useful statistical and structural information is computed as follows:

1) Convert the color image to greyscale.
2) Select pixels belonging to ROI and compute local binary pattern LBP_{PR}^{riu} [14].
3) Build an occurrence histogram, where we increment the jth bin of the histogram if number of transitions in LBP is j.
4) Repeat steps 2 and 3 for all pixels in ROI.

We build the occurrence histograms for $LBP_{16,2}$ and $LBP_{24,3}$ and concatenate them to form a feature vector L_i.

2) Color Features: Detection of the 7-point checklist criteria, such as blue-whitish veil and regression, which consist of mixtures of certain colors, can be achieved by analyzing the color intensity of pixels in the lesion. Stoeker et al. [17] investigated the discriminatory capability of color histograms in classification of malignant lesions from benign lesions. To reduce the variance due to the lighting conditions

in which dermoscopic images were taken, we considered also the HSV and LAB color spaces, which are invariant to illumination changes [18].

Color Histograms: To extract the color information of a lesion, we compute a color histogram from the intensity values of pixels belonging to the ROI. Additional images in the HSV and LAB color spaces are obtained from the original RGB image. The intensity range of each channel is divided into P fixed-length intervals. For each channel we build a histogram to keep count of the number of pixels belonging to each interval, resulting in a total of nine histograms from three color spaces. Statistical features, such as standard deviation and entropy (Eq. 1), of the nine histograms are also extracted as features for classification. More specifically, entropy is defined as

$$entropy = \sum_{i=1}^{P} f(histogram[i]) \qquad (1)$$

where $histogram[i]$ is the normalized pixel count of ith bin and

$$f(n) = \begin{cases} n \times log2(n) & \text{if } n > 0 \\ 0 & \text{if } n = 0 \end{cases}$$

The color histogram feature extraction steps are as follows:

1) Obtain skin lesion image in HSV and LAB color space from input RGB image
2) For each channel in all three color spaces build a separate P bin histogram
3) For all pixels belonging to ROI, increment the jth bin of histogram where $j = \frac{I_c}{P} \times M_c$. I_c and M_c are pixel intensity and maximum intensity in the specified color space.
4) Compute the standard deviation and entropy of the histogram.
5) Repeat steps 3 and 4 for all the channels in RGB, HSV, and LAB color space.

Color histogram and statistical features are combined to form a feature vector C_i.

C. Classification

The features from color histogram C_i, Haar wavelet H_i, and LBP L_i are combined to form $F_i = \{C_i, H_i, L_i\}$. For each criterion in the 7-point checklist we perform a filtered feature selection to obtain a subset of F_i with the highest classification accuracy. Correlation coefficient values (between F_i and each criterion) are used as the ranking criterion in the filters. The size of the subset and the parameters of the linear support vector machine (SVM) are obtained by grid search. Each criterion requires both training and testing.

Training: The training algorithm is as follows:

1) Segment the input image to obtain region of interest.
2) Extract Color histogram, Haar wavelet, and Local binary pattern. Concatenate them to form feature vector F_i.
3) Repeat steps 1 and 2 for all training images.
4) Perform filter feature selection to choose a subset of features S_i from F_i.
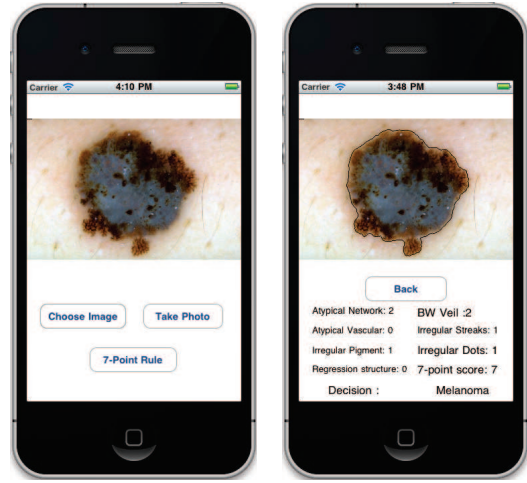

Fig. 2. *SkinScan application on Apple iPhone device*

5) Input S_i to linear SVM to obtain maximum margin hyperplane.

Testing: Image classification is performed as follows:

1) Read the input image and perform steps 1 and 2 from the training algorithm.
2) For each criterion use the SVM coefficients obtained from training to make a prediction.
3) Scores from all major and minor criteria are summed up. If the score is greater than or equal to 3 then lesion is classified as melanoma, otherwise as benign.

III. RESULTS

In this section we present an instantiation of the 7-point checklist on the Apple iPhone. We compare the classification accuracy of each criterion separately against the overall final classification by expert physicians [8]. We have developed a menu based application that implements the automated procedure for classification, as shown in Figure 2. The end user can take an image of the skin lesion using the 5 megapixel built in camera with LED flash, or load the image from media library. The image is analyzed in quasi real time and the final result displayed on the screen.

A. Classification Results

We performed 10-fold cross validation on the set of 347 images. We divided the dataset into 10 folds, nine folds with 11 melanoma and 23 benign lesions and the remaining fold with 11 melanoma and 30 benign lesions. Of the 10 folds, we used nine for training and one for test. We performed 10 rounds of validation where we chose each fold for testing and the rest for training, to get 10 experiments. In the following sub-sections we compare the classification accuracy of each criterion and the overall decision of the 7-point checklist with dermatology and histology.

Classification Results for each Criterion: In Table I we present the sensitivity and specificity of our algorithm in classification of each of the 7-point checklist criterion. We have lower accuracy for the regression structures, because

TABLE I

Classification Results for all Criteria.

	Sensitivity	Specificity
Atypical Pigment Network	72.86%	70.40%
Blue-Whitish Veil	79.49%	79.18%
Atypical Vascular Pattern	75.00%	69.66%
Irregular Streaks	76.74%	79.31%
Irregular Pigmentation	69.47%	74.21%
Irregular Dots and Globules	74.05%	74.54%
Regression Structures	64.18%	67.86%

TABLE II

Confusion Matrix of the automated Decision.

Confusion Matrix		Predicted	
		Melanoma	benign
Dermoscopy	Melanoma	96	14
	Benign	68	169

TABLE III

Classification accuracy

	Sensitivity	Specificity
7-Point Checklist	87.27%	71.31%
Ignoring 7-Point Checklist	74.78%	70.69%

they are usually indistinguishable from the blue-whitish veil via dermoscopy [8]. However, this is not an issue, as is it only necessary to obtain a minimum score of 3 to correctly detect a melanoma.

Classification and Overall Decision: In Table III we compare the sensitivity and specificity of our algorithms with the decision made by expert clinicians via dermoscopy. In Table II we present the confusion matrix computed using the sum of ten confusion matrices from the ten test sets, given from the 10-fold cross validation.

We also performed another classification experiment using SVM, where we ignore the 7-point checklist and directly classify each skin lesion as melanoma or benign. The feature vectors, feature selection scheme, and final ground truth (melanoma/benign) are the same as the classification using the automated 7-point checklist. Table III shows that classification accuracy is much lower when we ignore the 7-point checklist.

Execution Time: We compare the time needed for classification using the ISODATA segmentation algorithm [11] on the Apple iPhone 3G with a typical desktop computer (2.26 GHz Intel Core 2 Duo with 2GB RAM). The average image size in the dataset is 552×825 pixels. The classification time includes time taken for feature extraction. Table IV shows computation time in seconds for both platforms. It can be seen that the whole procedure takes under 10 sec to complete. This proves that the application is light enough to run on a smartphone which has limited computation power.

IV. DISCUSSION AND CONCLUSIONS

In this paper, we showed an implementation of 7-Point checklist on the Apple iPhone. The application, written in C/C++, is light enough to run on devices with limited memory and computation speed. We implemented sophisticated image processing and pattern recognition algorithms that run

TABLE IV

Mean computation time on the iPhone 4 and desktop computer

Mean time (sec)	Apple iPhone 3G	Desktop computer
Segmentation	2.3910	0.1028
Classification	7.4710	0.2415

with limited resources and provide classification results in quasi real time. These finding demonstrate the feasibility of developing medical diagnostic applications on off-the-shelf smart handheld devices, which have the advantages of low cost and portability.

REFERENCES

[1] M. Morris and F. Guilak. Mobile heart health: Project highlight. *Pervasive Computing, IEEE*, 8(2):57–61, April-june 2009.
[2] A. Logan, W. McIsaac, A. Tisler, and et al. Mobile phone-based remote patient monitoring system for management of hypertension in diabetic patients. *American journal of hypertension*, 20(9):942–8, 2007.
[3] J. Hicks, N. Ramanathan, D. Kim, M. Monibi, and J. Selsky. Andwellness: an open mobile system for activity and experience sampling. In *Wireless health 2010*, WH '10, pages 34–43, New York, NY, USA, 2010. ACM.
[4] W. Stolz, A. Riemann, AB. Cognetta, and L. Pillet. Abcd rule of dermoscopy: a new practical method for early recognition of malignant melanoma. *Eur J Dermatol*, 4:521–527, 1994.
[5] S. W. Menzies, C. Ingvar, K. A. Crotty, and W. H. McCarthy. Frequency and morphologic characteristics of invasive melanomas lacking specific surface microscopic features. *Archives of Dermatology*, 132(10):1178–1182, 1996.
[6] G. Argenziano, G. Fabbrocini, P. Carli, V. De Giorgi, E. Sammarco, and M. Delfino. Epiluminescence microscopy for the diagnosis of doubtful melanocytic skin lesions. *Arch Dermatol*, 134:1563–70, 1998.
[7] T. Wadhawan, N. Situ, K. Lancaster, X. Yuan, and G. Zouridakis. Skinscan: A portable library for melanoma detection on handheld devices. In *Proceeding*. ISBI, 2011.
[8] G. Argenziano, H. P. Soyer, V. D. Giorgi, D. Piccolo, P. Carli, M. Delfino, A. Ferrari, R. Wellenhof, D. Massi, G. Mazzocchetti, M. Scalvenzi, and I. Wolf. *Dermoscopy: a tutorial*. EDRA: medical publishing and new media, 2000 edition, Feburary 2000.
[9] G. Argenziano, H. P. Soyer, S. Chimenti, and et al. Dermoscopy of pigmented skin lesions: Results of a consensus meeting via the internet. *J Am Acad Dermatol*, 48(5):679–93, 2003.
[10] T. Huang, G. Yang, and G. Tang. A fast two-dimensional median filtering algorithm. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 27:13– 18, Feb 1979.
[11] J.C. Dunn. A fuzzy relative of the isodata process and its use in detecting compact well-separated clusters. *Journal of Cybernetics*, 3:32–57, 1973.
[12] M. Elbaum, A.W. Kopf, H.S. Rabinovitz, R.G.B. Langley, H. Kamino, M.C. Mihm, et al. Automatic differentiation of melanoma from melanocytic nevi with multispectral digital dermoscopy: A feasibility study. *Journal of the American Academy of Dermatology*, 44(2):207–218, 2001.
[13] N. Situ, X. Yuan, J. Chen, and G. Zouridakis. Malignant melanoma detection by bag-of-features classification. In *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, pages 3110–3113, 2008.
[14] T. Ojala, M. Pietikinen, and M. Topi. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24:971–987, 2002.
[15] E. J. Stollnitz, T. D. DeRose, and D. H. Salesin. Wavelets for computer graphics: A primer. *IEEE Computer Graphics and Applications*, 15:76–84, 1995.
[16] J. A. Hartigan and M. A. Wong. A k-means clustering algorithm. *Applied Statistics*, 28:100–108, 1979.
[17] R. J. Stanley, W. V. Stoecker, and R. H. Moss. A relative color approach to color discrimination for malignant melanoma detection in dermoscopy images. *Skin Research and Technology*, 13(1):62–72, 2007.
[18] K. van de Sande, T. Gevers, and C. Snoek. Evaluating color descriptors for object and scene recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(9):1582–96, 2010.