A Cascade Classifier for Diagnosis of Melanoma in Clinical Images

P. Sabouri, *Student Member*, *IEEE*, H. GholamHosseini, *Senior Member, IEEE*, T. Larsson, and J. Collins

***Abstract*— Computer aided diagnosis of medical images can help physicians in better detecting and early diagnosis of many symptoms and therefore reducing the mortality rate. Realization of an efficient mobile device for semi-automatic diagnosis of melanoma would greatly enhance the applicability of medical image classification scheme and make it useful in clinical contexts. In this paper, interactive object recognition methodology is adopted for border segmentation of clinical skin lesion images. In addition, performance of five classifiers, KNN, Naïve Bayes, multi-layer perceptron, random forest and SVM are compared based on color and texture features for discriminating melanoma from benign nevus. The results show that a sensitivity of 82.6% and specificity of 83% can be achieved using a single SVM classifier. However, a better classification performance was achieved using a proposed cascade classifier with the sensitivity of 83.06% and specificity of 90.05% when performing ten-fold cross validation.**

I. Introduction

Skin cancer has been one of the most common form of cancers and melanoma is one of the leading cause of death particularly in USA, Australia and New Zealand [1, 2]. Melanoma has also risen faster than any of the most common cancers among white population [3]. Early diagnosis of malignant melanoma significantly reduces the morbidity, mortality and cost of the medication [4]. Computer aided diagnosis (CAD) systems and computer vision have been used to help skin cancer specialists for better detection of melanoma lesion mainly based on ABCD rule (asymmetry, border irregularity, color variation and regions with diameter greater than 6mm) and seven-point checklist. There are several non-invasive methods that have been developed on pigmented skin lesion (PSL) images. Multi-spectral and hyper-spectral imaging systems have led to renewed interest in diagnosis of melanoma [5, 6]. Dermatological photographs (clinical images) and dermoscopy are also non-invasive methods for diagnosis of PSLs which have been widely studied in the dermatological imaging realm using CAD [7].

P. Sabouri is with Auckland University of Technology, Department of Electrical and Electronic Engineering, Auckland, New Zealand (e-mail: psabouri@aut.ac.nz).

H. GholamHosseini is with Auckland University of Technology, Department of Electrical and Electronics Engineering, Auckland, New Zealand (corresponding author phone: +64-9-9219999; fax: +64-9-9219973; e-mail: hgholamh@aut.ac.nz).

T. Larsson is with Mälardalen University, Västerås, Sweden (e-mail: thomas.larsson@mdh.se).

J. Collins is with Auckland University of Technology, Department of Electrical and Electronics Engineering, Auckland, New Zealand (e-mail: jcollins@aut.ac.nz).

In this work a dataset of clinical images that have been collected online is formed with the aim of developing a handheld embedded vision application. This paper has been structured into eight sections: following the introduction the image dataset is described, preprocessing and proposed method for lesion segmentation are given in section three, feature extraction and classification methodologies are presented in sections four and five, system development in section six and finally, sections seven and eight deal with experimental results and conclusion.

II. Dataset

A dataset of 370 images collected from web resources (Table 1) has been used for developing the algorithms. The dataset is divided into two groups: with 175 images of Malignant Melanoma (MM) and 195 images of non-melanoma (benign nevus). Non-melanoma images include: Atypical (compound, junctional, dermal, and combined), dysplastic, seborrheic keratosis, blue nevus, congenital, spitz, halo and neurofibromatosis.

TABLE.1 LIST OF ONLINE RESOURCES

Web resource	URL
DermQuest	http://www.dermquest.com
DanDerm	http://www.danderm-pdv.is.kkh.dk/atlas/index.html
DermAtlas	http://www.dermatlas.net/
DermIs	http://www.dermis.net
DermNetNz	www.dermnetnz.org

III. Preprocessing and Segmentation

A. Preprocessing

Presence of artifacts such as hairs, ruler marks and light illumination in dermatological images may affect the segmentation and therefore imperfect feature extraction results [8]. In this work, DullRazor tool was employed for hair artifacts removal from lesion images [9]. This tool performs three main tasks: 1) Detecting the location of dark hairs in the image using morphological filters, 2) Replacing the identified hairs by neighboring pixels 3) Applying adaptive median filter to smooth the final image. After hair artifact removal, all lesion images are cropped manually and resized to construct a dataset of unified 512x512 pixel images.

B. Lesion segmentation

After preprocessing of skin images, lesion segmentation (border detection) is performed in order to remove the undesirable background skin part from the foreground

978-1-4244-7929-0/14/$26.00 ©2014 IEEE

lesion part. A number of algorithms have been proposed for border segmentation of clinical and dermoscopic images based on color histogram thresholding (e.g. Otsu and thresholding) [7]. A discussion of the accuracy of current methods and the performance of these border segmentation techniques can be found in the literature [10, 11]. The main drawback of applying threshold based segmentation/clustering to our dataset was the variation in lighting condition (e.g. brightness, contrast and reflection of flash light on the skin's surface) of the images.

In this study, an interactive object recognition method was adopted for lesion segmentation [12]. The proposed method consists of the following steps: extracting color features from the region of interest by user, applying a nearest neighbor search to the specified points on the lesion by using color features, applying morphological filters: opening, closing and median for smoothing and removing the artifacts, labeling connected component on the lesion and then preserving the largest part, and finally filling holes after creating the binary mask.

Besides, the result of above mentioned method was compared with the histogram segmentation approach using the following algorithm: 1) converting the preprocessed image to HSV color space, 2) applying median filter to smooth the image, 3) applying Otsu's thresholding method on each color channel, 4) labeling connected components on the binary mask and finally 5) removing small objects (remained artifacts).

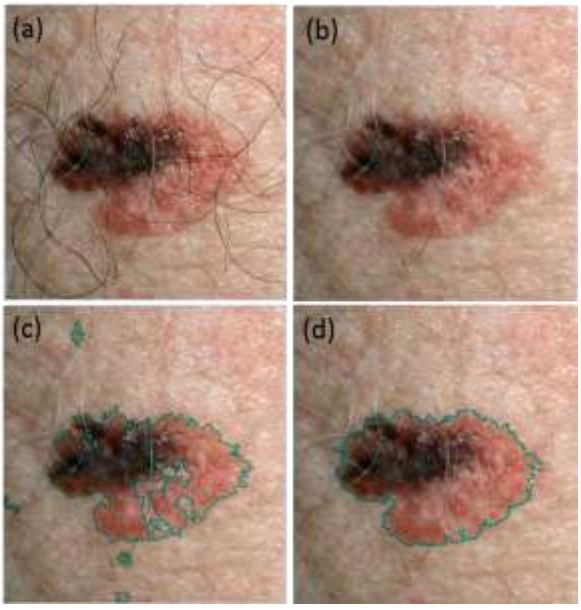

Fig. 1. Results of border segmentation: (a) original image, (b) after preprocessing and hair removal, (c) border segmentation using histogram thresholding on HSV color space, (d) border segmentation using adopted interactive object recognition.

Fig. 1 shows the results of preprocessing and border segmentation phases. It indicates that using the proposed method border of the lesion is well detected (Fig. 1 d) compared with threshold based segmentation on HSV color space (Fig. 1 c).

IV. FEATURE EXTRACTION

This section describes how the feature attributes were extracted from the segmented lesion of clinical images. While shape analysis can improve the performance of melanoma detection significantly, we only performed general color and feature analysis for this study and left the shape analysis for future development.

A. Color Features

Color analysis is one of the most important methods for analyzing medical images. In the skin image analysis realm, typically original RGB image is transformed to different color domains in order to measure corresponding color information from color channels. Although RGB images can be transformed to various color spaces, other image formats might be superior for a specific application [13]. In this paper, color analysis is performed on RGB and HSV color spaces by measuring the average, standard deviation, skew and entropy of each color channel.

B. Texture Features

Texture analysis is a technique for extracting shape attributes and spatial structure of images. There are various commonly used texture analysis methods. In this paper, texture attributes were extracted from the gray level co-occurrence matrix (GLCM) [14]. The GLCM is calculated from the RGB segmented lesions and created by averaging different orientations matrices (angles of 0, 45, 90 and 135 degree). Then the following features were calculated from the average matrix: energy, entropy, correlation, inverse different moment, and inertia.

V. CLASSIFICATION

In this study, five classification algorithms were used in order to assess their performance using the extracted features from the dataset by Weka tool [15]. KNN (k=10), Multi-Layer Perceptron (MLP), Naive Bayes (NB), Random Forest (RF) and Support Vector Machine (SVM) using LibSVM (c=14, Radial Base Function, γ=0.08) were used and the accuracy of each classifier was compared [16]. In bioinformatics applications, overfitting is one of the fundamental issues for supervised learning classification. This phenomena occurs when training set is adopted by a fixed set of data rather than learning from the trend; therefore, the accuracy would be too optimistic [17]. In order to overcome this issue, ten-fold cross validation was applied for training and testing of all input images. After performing classification, specificity and sensitivity of each classifier was calculated for performance analysis. Furthermore, the following five feature sets were extracted in order to find the best features and a proper classifier:

1- Combination of RGB color and texture features (25 feature columns),
2- Combination of HSV color and texture features (25 feature columns),
3- RGB color attributes (15 feature columns),

4- HSV color attributes (15 feature columns),
5- Texture only features (10 feature columns).

VI. System development

By using massively parallel hardware features, an effective and responsive stand-alone system can be created even on commodity hardware. Interestingly, there is ample evidence of how compute-intensive tasks in medical imaging can be solved successfully by taking advantage of GPU architectures, which leads to generous speed-ups [18].

The presented cascade classifier relies mainly on operations such as image denoising and segmentation, which are realized by applying a sequence of highly data-parallel image operations. Thus, it is expected that an efficient GPU implementation can perform both the preprocessing and lesion detection phases in merely a fraction of a second. Similarly, the feature extraction procedures involve RGB to HSV color transforms as well as the computation of a number of statistical measures over different color channels, which again give good opportunities to exploit the embedded fine-grained data-parallelism. The SVM training phase is the most time-consuming part, but fortunately this is an offline computation; it can be run once, and then the results are simply used to classify the extracted feature vectors, which is a very efficient operation.

Therefore, to enable more extensive testing and further progress, a natural next step would be to implement an efficient parallel version of the proposed cascade classifier on a modern handheld device with a powerful GPU, a high quality camera, and a suitable display for user interaction. Currently, two different types of hardware setups are considered. First, smart phones, and in particular iPhone or iPad running iOS, as they are widely available and have the necessary features. Unfortunately, general-purpose GPU computations through a standardized Application Programming Interface (API) such as OpenCL is not yet publicly available, but access to the GPU is still possible through, e.g., OpenGL ES shaders and CoreImage. Second, a heterogeneous high-performance research platform, called GIMME3+, is currently under development. It comes with an AMD quad core CPU with up to 16 GB DDR3 RAM, an integrated GPU sharing the CPU memory, and 12 Mgate FPGA with >256 MB, DDR2 memory, and a high quality stereo camera. The system fits in 70 x 70 mm^2 board and it can deliver up to 1 TFLOPS at 30 W maximum. This system offers much more flexibility in terms of reconfigurability and programmability [19].

VII. Experimental results

Table II illustrates comparison of five classifiers in terms of specificity and sensitivity based on different feature selection. The first set of features is a combination of RGB color and texture attributes (RGB + TEXTURE). The results show that the sensitivity of NB is the highest (89.2%) but it has the lowest specificity of 51% among the selected group. Therefore, the high true positive rate (melanoma images were correctly classified) of NB is compromised by its relatively low false positive rate (half of the benign nevus images were classified as melanoma). In the selected group, SVM with the sensitivity of 82.6% and specificity of 83% shows a better performance.

The next set of results compares the same classifiers using HSV color attributes and texture features (HSV + TEXTURE). Table II, shows that the sensitivity of all classifiers is slightly improved specially for KNN (from 79% to 86.3%) and SVM from (82.6% to 84.5%) but the specificity of KNN is reduced from 74% to 69.1% and 83% to 76.6% for SVM. Therefore, no significant improvement can be observed for these sets of features. The third and fourth sets demonstrate the comparison of the selected classifiers using solely RGB and HSV color features (separate analysis for each set). The NB classifier ranked the highest sensitivity but low specificity for both features. In the selected group, RF and SVM achieved the highest specificity of 82.4% and 81.9% for HSV attributes, respectively. Finally, the accuracy of classifiers was compared based on texture only features as shown in Table II. In this category, accuracy of SVM ranked the first with sensitivity of 82.8% and specificity of 67.9%; while NB ranked the last with sensitivity of 51.6% and specificity of 58%.

To further improve the performance of classification and diagnosis, we propose a cascade classifier (Fig. 2) consisting of two SVM stages with the following details:

1) The input images are divided into two dataset1 and

TABLE II. ACCURACY COMPARISON OF KNN, MLP, NB, RF AND SVM (SE is Sensitivity and SP is Specificity).

	RGB + TEXTURE		HSV + TEXTURE		RGB		HSV		TEXTURE	
Classifier	**SE**	**SP**	**SE**	**SP**	**SE**	**SP**	**SE**	**SP**	**SE**	**SP**
KNN	79	74.5	86.3	69.1	76	77.1	83.3	66	80.4	53.7
MLP	76.6	77.7	78.6	78.9	77.2	81.4	73.8	79.8	88.7	51.6
NB	89.2	51.6	89.9	59.6	89.2	47.9	85.7	58	51.6	58
RF	73.1	80.6	78.8	83.7	77.8	76.1	78.9	82.4	68.5	70.7
SVM	82.6	83	84.5	76.6	75	79.9	82.2	81.9	82.8	67.9

dataset2 using a threshold filter; the filter controls only two HSV color feature values; i.e. if either the entropy of S and V channels is less than zero. Therefore, dataset1 would have more non-melanoma images (86 benign and 15 melanoma images) and dataset2 contains more melanoma images (109 benign and 169 melanoma images).

2) A SVM classifier (SVM #1) using normalized HSV color features (-1, 1) is applied to dataset1 to classify images into melanoma and non-melanoma. (HSV features were ranked the most effective features among other features for SVM #1).
3) A second SVM classifier (SVM #2) using a combination of color and texture features is applied to dataset2. At this stage, correlation-based feature ranking method (CFS) is used in order to find the most effective features and nine features were selected out of 40 features (2 RGB, 2 texture and 5 HSV features).

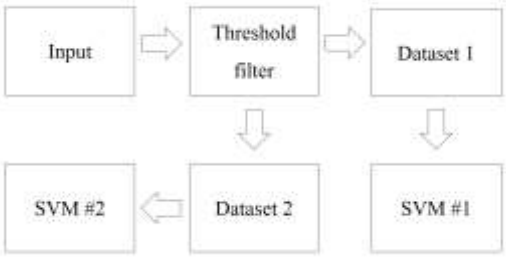

Fig. 2. Classification using proposed cascade classifier

Using the proposed cascade classifier, we achieved sensitivity of 83.06% and specificity of 90.05%. It was found that by employing the proposed cascade classifier, the accuracy of the diagnosis system increased. Furthermore, the cascade classifier was tested using a new set of data (26 benign and 16 melanoma images) and achieved the sensitivity of 89.28% and specificity of 100%. Moreover, the execution time was remarkably improved after applying normalization on the data set.

VIII. Conclusion

This paper reports on preprocessing, feature extraction, design and evaluation of classifiers in order to discriminate melanoma from non-melanoma lesions. Five different classifiers were selected to determine their diagnosis accuracy based on color and texture features of clinical images. In addition, an interactive object recognition method was adopted for optimal lesion border segmentation. Experimental results show that although the combination of color and texture may result in more optimal classification performance than separate color and texture features analysis, the overall performance of the classifier is still insufficient for an accurate discrimination. Therefore, we proposed a cascade classifier using SVM which achieved sensitivity of 83.06% and sensitivity of 90.05% for more accurate diagnosis. To implement an efficient parallel version of the proposed cascade classifier, two options of using smart phones and GIMME3+ platform are being considered.

References

[1] T. Diepgen and V. Mahler, "The epidemiology of skin cancer," *British Journal of Dermatology,* vol. 146, pp. 1-6, 2002.

[2] A. W. Kopf, T. G. Salopek, J. Slade, A. A. Marghoob, and R. S. Bart, "Techniques of cutaneous examination for the detection of skin cancer," *Cancer,* vol. 75, pp. 684-690, 1995.

[3] A. Jemal, S. S. Devesa, P. Hartge, and M. A. Tucker, "Recent trends in cutaneous melanoma incidence among whites in the United States," *Journal of the National Cancer Institute,* vol. 93, pp. 678-683, 2001.

[4] L. A. Goldsmith, F. B. Askin, A. E. Chang, C. Cohen, J. P. Dutcher, R. S. Gilgor*, et al.*, "Diagnosis and treatment of early melanoma," *JAMA: The Journal of the American Medical Association,* vol. 268, pp. 1314-1319, 1992.

[5] A. Sahu, C. McGoverin, N. Pleshko, K. Sorenmo, and C.-H. Won, "Hyperspectral imaging system to discern malignant and benign canine mammary tumors," in *SPIE Defense, Security, and Sensing*, 2013, pp. 87190W-87190W-8.

[6] S. Kiyotoki, J. Nishikawa, T. Okamoto, K. Hamabe, M. Saito, A. Goto*, et al.*, "New method for detection of gastric cancer by hyperspectral imaging: a pilot study," *Journal of biomedical optics,* vol. 18, pp. 026010-026010, 2013.

[7] K. Korotkov and R. Garcia, "Computerized analysis of pigmented skin lesions: A review," *Artificial intelligence in medicine,* 2012.

[8] Q. Abbas, M. E. Celebi, I. Fondón García, and M. Rashid, "Lesion border detection in dermoscopy images using dynamic programming," *Skin Research and Technology,* vol. 17, pp. 91-100, 2011.

[9] T. Lee, V. Ng, R. Gallagher, A. Coldman, and D. McLean, "Dullrazor®: A software approach to hair removal from images," *Computers in Biology and Medicine,* vol. 27, pp. 533-543, 1997.

[10] M. E. Celebi, G. Schaefer, H. Iyatomi, and W. V. Stoecker, "Lesion border detection in dermoscopy images," *Computerized medical imaging and graphics: the official journal of the Computerized Medical Imaging Society,* vol. 33, p. 148, 2009.

[11] P. Wighton, T. K. Lee, H. Lui, D. I. McLean, and M. S. Atkins, "Generalizing Common Tasks in Automated Skin Lesion Diagnosis," *Information Technology in Biomedicine, IEEE Transactions on,* vol. 15, pp. 622-629, 2011.

[12] G. Friedland, K. Jantz, and R. Rojas, "Siox: Simple interactive object extraction in still images," in *Multimedia, Seventh IEEE International Symposium on*, 2005, p. 7 pp.

[13] M. Faal, M. Baygi, M. Hossein, and E. Kabir, "Improving the diagnostic accuracy of dysplastic and melanoma lesions using the decision template combination method," *Skin Research and Technology,* vol. 19, pp. e113-e122, 2013.

[14] M. A. Sheha, M. S. Mabrouk, and A. Sharawy, "Automatic Detection of Melanoma Skin Cancer using Texture Analysis," *International Journal of Computer Applications,* vol. 42, 2012.

[15] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The WEKA data mining software: an update," *ACM SIGKDD explorations newsletter,* vol. 11, pp. 10-18, 2009.

[16] C.-C. Chang and C.-J. Lin, "LIBSVM: a library for support vector machines," *ACM Transactions on Intelligent Systems and Technology (TIST),* vol. 2, p. 27, 2011.

[17] D. de Ridder, J. de Ridder, and M. J. Reinders, "Pattern recognition in bioinformatics," *Briefings in bioinformatics,* vol. 14, pp. 633-647, 2013.

[18] A. Eklund, P. Dufort, D. Forsberg, and S. M. LaConte, "Medical image processing on the GPU–Past, present and future," *Medical image analysis,* vol. 17, pp. 1073-1094, 2013.

[19] C. Ahlberg, J. Lidholm, F. Ekstrand, G. Spampinato, M. Ekstrom, and L. Asplund, "Gimme-a general image multiview manipulation engine," in *International Conference on Reconfigurable Computing and FPGAs (ReConFig)*, 2011, pp. 129-134.