Pattern Recognition 46 (2013) 86–97

Contents lists available at SciVerse ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

Pattern classification of dermoscopy images: A perceptually uniform model

Qaisar Abbas [a,b,c,*], M.E. Celebi [d], Carmen Serrano [e], Irene Fondón García [e], Guangzhi Ma [a,b]

[a] Department of Computer Science and Technology, Huazhong University of Science and Technology, 1037 Luoyu Road, Wuhan 430074, China
[b] Center for Biomedical Imaging and Bioinformatics, Key Laboratory of Image Processing and Intelligent Control of Ministry of Education, Wuhan, China
[c] Department of Computer Science, National Textile University, Faisalabad 37610, Pakistan
[d] Department of Computer Science, Louisiana State University, Shreveport, Louisiana, LA, USA
[e] Escuela Superior de Ingenieros, Universidad de Sevilla, Camino de los Descubrimientos, s/n, 41092 Sevilla, Spain

ARTICLE INFO

Article history:
Received 15 February 2011
Received in revised form
20 March 2012
Accepted 31 July 2012
Available online 14 August 2012

Keywords:
Dermoscopy
Pattern classification
Steerable pyramid transform
Human visual system
AdaBoost
Multi-label learning

ABSTRACT

Pattern classification of dermoscopy images is a challenging task of differentiating between benign melanocytic lesions and melanomas. In this paper, a novel pattern classification method based on color symmetry and multiscale texture analysis is developed to assist dermatologists' diagnosis. Our method aims to classify various tumor patterns using color–texture properties extracted in a perceptually uniform color space. In order to design an optimal classifier and to address the problem of multi-component patterns, an adaptive boosting multi-label learning algorithm (AdaBoost.MC) is developed. Finally, the class label set of the test pattern is determined by fusing the results produced by boosting based on the maximum a posteriori (MAP) and robust ranking principles. The proposed discrimination model for multi-label learning algorithm is fully automatic and obtains higher accuracy compared to existing multi-label classification methods. Our classification model obtains a sensitivity (*SE*) of 89.28%, specificity (*SP*) of 93.75% and an area under the curve (*AUC*) of 0.986. The results demonstrate that our pattern classifier based on color–texture features agrees with dermatologists' perception.

1. Introduction

Non-invasive malignant melanoma (MM) is widely diagnosed using digital dermoscopy. In fact, dermoscopy is one of the most cost-effective methods for the detection and analysis of pigmented (PSLs) and non-pigmented skin lesions (Non-PSLs). Many dermatologists [1,2] use clinical ABCD (*A*: Asymmetry, *B*: Border, *C*: Color, *D*: Differential structures); Menzies' method; 7-point checklist and pattern analysis methods to diagnose and classify the lesions. In particular, it is very difficult to distinguish among lesions and even experienced dermatologists [3] have a diagnostic accuracy below 85%. Therefore, recently computer-assisted diagnosis systems (CADs) have been developed. For CADs to decide if a lesion is benign, melanoma or suspect [4], it would be desirable to have automated systems that can provide assistance to less experienced dermatologists. Automated systems for dermoscopy images [5–7] usually have four stages: (1) artifact removal, (2) lesion segmentation, (3) ABCD and texture related feature extraction and optimization, and finally (4) classification. In practice, the classification algorithms, which utilizes low-level feature extraction such as color, texture and shape, are quite sophisticated and have achieved considerable success. The extraction of low-level image features that correlate with high-level image semantics, however, remains a challenging task.

A few discrimination approaches have been proposed in the literature that attempt to bridge this semantic gap between low-level features and high-level semantics. However, many of them use non-uniform color spaces e.g., RGB, HSV and some of them use approximately uniform color space such as CIEL*a*b* [8]. These techniques do not correlate well with the physician's perception due to the use of a non-uniform color space or limited gray scale image properties. Moreover, color appearance models are not utilized in these classification techniques. These appearance models correlate to well with human perception and are capable to predict a wide variety of visual phenomena, which is not possible in color spaces.

Various lesion classification systems have been proposed in the literature. In [9], Ganster et al. proposed an automated melanoma recognition by the nonparametric "KNN classifier". A machine learning algorithm was developed in [10] to characterize melanoma by a feature vector that contained shape, color and texture information, as well as local and global parameters. Burroni et al. [11] used the K-NN classifier for classification of melanoma. Automated melanoma classification systems [12,13] were developed too. Tanaka et al. [14] developed a method for pattern classification of nevus with texture analysis. Their

* Corresponding author at: Department of Computer Science, National Textile University, Faisalabad 37610, Pakistan.
Tel.: +92 41 9230081x140; fax: +92 41 9230082.
E-mail addresses: drqaisar@ntu.edu.pk, qaisarabbasphd@gmail.com (Q. Abbas), emre.celebi@lsus.edu (M.E. Celebi), cserrano@us.es (C. Serrano), irenef@us.es (I. Fondón García), maguangzhi@mail.hust.edu.cn (G. Ma).

0031-3203/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.patcog.2012.07.027

research was devoted to categorize three texture patterns: globular, reticular, and homogeneous patterns with 94% accuracy. Iyatomi et al. [15] developed a classification system based on acral volar skin with three detectors such as parallel ridge, parallel furrow, and fibrillar patterns. They used more than 46 texture features with maximal value of between-class variance, and Mahalanobis distance. In [16], a pattern analysis based on clinical color, architectural order, symmetry of pattern and homogeneity (CASH) technique was modeled by Markov random field (MRF). In that study, mean and variance of each plane of CIEL*a*b* color space was used to extract the color related features. For seven pattern classes, the authors reported 86% classification accuracy. A different approach was developed in [17] to detect and visualize only pigment network structures based on cyclic graphs. Recently in [18], a pattern classification system (PCS) based on the CASH rule was presented to recognize among six classes instead of multicomponent pattern.

1.1. Aim and approach

The primarily aim of this paper is to measure the color, architectural order, symmetry of pattern and homogeneity (CASH) of lesions instead of clinical ABCD rule [7,15]. By using CASH method, physicians can classify between benign and malignant lesions. Benign melanocytic lesions tend to have few colors, an architectural order, symmetry of pattern and are homogeneous. Malignant melanomas often have many colors, architectural disorder, asymmetry of pattern and are heterogeneous. To automate CASH model, some effective methods [14–18] were proposed. Particularly in [16], CIEL*a*b* color space [16] model was utilized to identify seven patterns using maximum likelihood (ML) criteria. Melanoma or Clark nevus lesions often contain multicomponent patterns meaning that a lesion contain 2, 3 or more pattern classes. As a result, if a classification decision rule is used based on ML [16] then it cannot provide multiple decisions to match with multiple pattern classes simultaneously because it was based on the single state maximum probability concept. However, if single input pattern is provided to match a single class output then it may provide better classification results. Therefore, this study is focused on providing multi-patterns as an input to match with multi-class outputs, concurrently as shown in Fig. 1. In this example, an input lesion is classified into three patterns.

Physicians are more capable of interpreting color–texture information than any automated method. This is because; the human visual system (HVS) plays an important role in the categorization and recognition of objects. Similarly, it has been proven that HVS used model based approach to take decisions based on the fuzzy logical model of perception and signal detection theory. Humans are using these two models to discriminate among patterns and to make decisions. By combining these two models, the conclusion is that we should focus on developing a pattern analysis model based on human perception. As a result, the first aim of this study is to develop an effective pattern classification model based on CASH, which is closer to physician perception. Secondly, to develop an efficient and optimized pattern classification method in a CIECAM02 (*JCh*) perceptually uniform color space. Thirdly, an adaptive boosting (AdaBoost.MH) multi-label input machine learning algorithm is used to develop (AdaBoost.MC) multi-label output method for effective patterns detector and to solve the multicomponent pattern problem in dermoscopy images. The AdaBoost.MC algorithm is developed by integrating maximum a posterior probability (MAP) along with ranking concept. The MAP probability technique is utilized since it provides most popular statistical criteria to get optimality. In this multi-label boosting algorithm, the classes are ranked according to their level of similarity. The class is ranked first with the highest probability; second best probability is ranked second and so on based on label-weighted score.

In this study, the classifications of different pattern classes in the lesion diagnosis process are summarized as follows:

(a) *Reticular pattern or pigmented network:* It is the most common global feature present in a junctional nevus, compound nevus, lentigo or melanosis.
(b) *Globular pattern:* It presents itself as small aggregated globules and may have different colors, which has high specificity for diagnosis of compound and intradermal nevi.
(c) *Cobblestone pattern:* It is similar to Globular pattern but they are large, closely aggregated and angulated.
(d) *Homogeneous pattern:* diffuse and homogeneous blue-grayish pigmentation is present and absence of pigmented network, which characterizes the blue nevi.
(e) *Parallel ridge pattern* (*PRP*)**:** The specific type of pattern found in palm or sole, which may be benign melanocytic nevi and acral melanomas if it has parallel ridge pattern.
(f) *Starburst pattern:* It is characterized by the presence of pigmented streaks in a radial arrangement. It is commonly seen in Red nevi or pigmented Spitz nevi.
(g) *Multicomponent pattern***:** This pattern has high specificity for diagnosis of melanoma and consists of presence of three or more dermoscopic features in a single lesion.

The multicomponent pattern is shown in Fig. 1, while the rest of the abovementioned dermoscopic patterns are shown in Fig. 2.

2. Outline of the proposed pattern classification model

To analyze the above mentioned patterns, a computerized CASH model is proposed as illustrated in Fig. 3. From each dermoscopy image, Region-of-interest (ROI) is selected first, which is then transformed it into the CIECAM02 (*JCh)* uniform color space. Afterwards color attributes such as the number of colors, percentage of occurrence and their similarities are measured. For texture feature

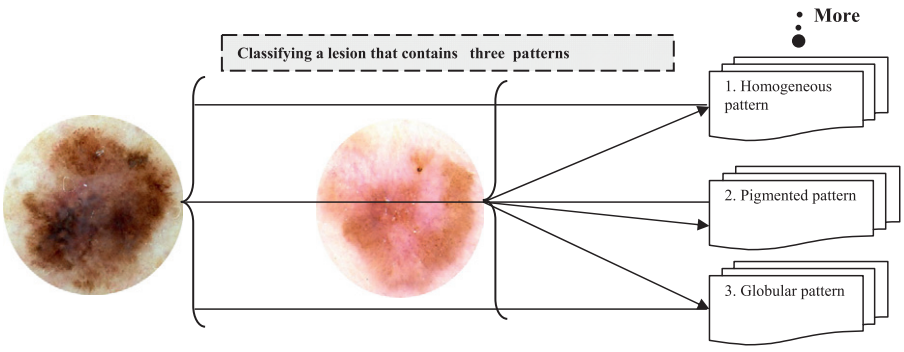

Fig. 1. Automatic classification of multicomponent patterns to match the class label sets for diagnosis the melanoma tumors.

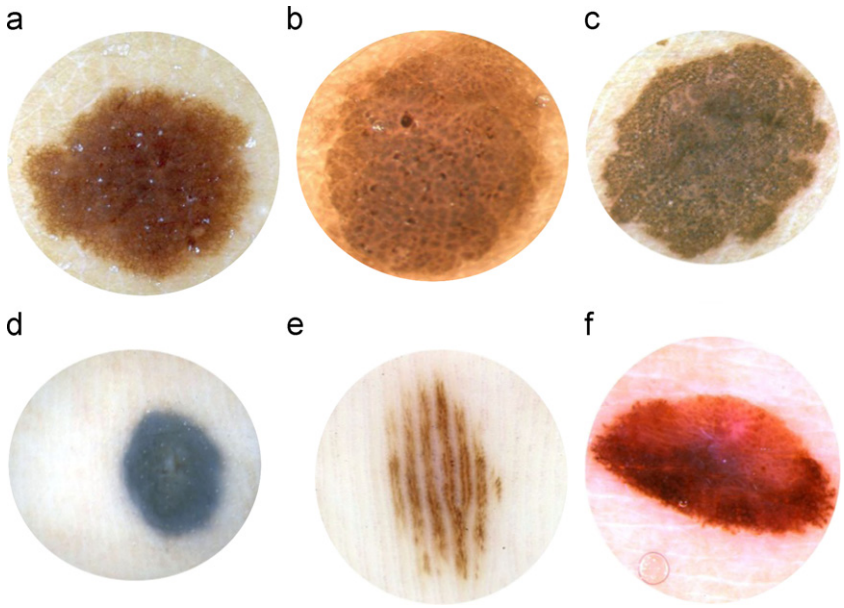

Fig. 2. Example of the different patterns present in dermoscopy images. (a) Clark nevus shows pigmented network, (b) melanocytic nevus shows globular pattern. (c) dermal nevus shows Cobblestone pattern, (d) blue nevus shows Homogeneous pattern, (e) acral melanomas shows parallel pattern, and (f) reed Spitz nevus shows starburst pattern.

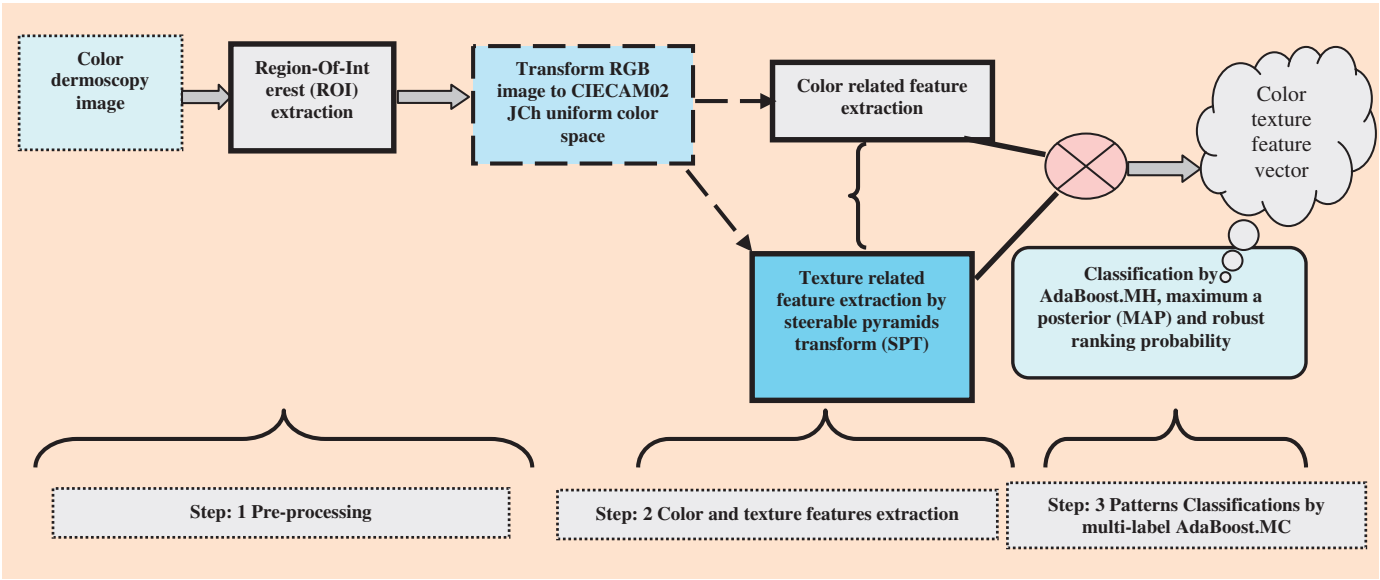

Fig. 3. Flow chart of the proposed pattern classification model for dermoscopy images.

analysis, the local and global statistical properties are extracted using the multiscale steerable pyramid transform (SPT) technique.

By combining these color–texture related features, a feature vector is constructed, which is finally learned by AdaBoost.MC multi-label classifier to distinguish among pattern classes. All steps are performed in a uniform color space (*JCh*) of CIECAM02 color appearance model. Detailed information about each step is presented in the subsequent sections.

3. Region-of-interest extraction and color space transformation

3.1. Region-of-interest (ROI) extraction

In order to obtain effective pattern extraction and classification, a region of size (450×450) is automatically selected from the center of each dermoscopy image having 768×512 pixels. This step is called ROI extraction. Seven pattern groups of total 350 are selected from a data set of 1039 dermoscopic images. A detailed description of this selected data set is presented in Section 6.1.

3.2. Color space transform

The proposed perceptually adapted pattern classification is intended to make the early diagnosis of skin cancer easier to the physician. Accordingly, the algorithm has to emulate the perception of the medical expert in order to categorize tumors. Therefore, the algorithm must be related to dermatologist's perception.

Since dermoscopy images are color images, any algorithm designed to process them must take into account color information or otherwise, the technique would waste a valuable source of

information. Consequently, the image has to be represented in one of the color spaces presented in the literature.

Combining both arguments, the conclusion is that the chosen color space must be related to human perception. That is, it must be a uniform color space. When dealing with almost uniform color spaces, CIEL*a*b* [16] is the most widely used, especially when combined with the advanced distance metrics CIE94 and CIEDE2000. This color system, accounts for chromatic adaptation when the illuminant of the scene is near day-light, the background is medium gray and the surround levels of luminance are moderate. However, it does not account for changes in background or luminance, cannot predict brightness and colorfulness and give erroneous results when the illuminant of the scene is largely different from day-light.

Due to the bright illumination of the image acquisition device in a dermoscopy context, certain details of the lesion can become indistinguishable. The illumination changes the contrast within the scene, making some patterns disappear, especially in dark areas of the image [19]. Therefore, it is necessary to predict the perceived appearance of the dark lesion in a bright surrounding. Moreover, as stated in the literature, among the existing color spaces with their corresponding color difference formula, it is preferable the selection of a color appearance based uniform color space which is capable of considering viewing conditions [20].

A color appearance model, CAM, provides us the scene [21] as we would actually see it. CIECAM02 is the recently adopted CIE color appearance model [22] and it is the most advanced among all existing CAMs due to its successful combination of the best features from existing color models. The CIECAM02 simplifies its previous version, CIECAM97s, which was adapted to the color appearance transformation of CAT2000. CIECAM02 defines six dimensions of color appearance: brightness *Q*, lightness *J*, colorfulness *M*, chroma *C*, saturation *s*, and hue *h*.

In CIECAM02 color space, the input data [23] are the adapting field luminance (normally taken to be 20% of the luminance of white in the adapting field), *La*, the relative tristimulus values of the stimulus, *XYZ*, the relative tristimulus values of white in the same viewing conditions, *XwYwZw*, and the relative luminance of the background, Y_b. Also, the parameters *c*, for the impact of surround, *Nc*, a chromatic induction factor, and *F*, a factor for degree of adaptation, must be selected. Afterwards, an initial chromatic adaptation transform is used to change from the stimulus viewing conditions to corresponding colors under equal-energy-illuminant reference viewing conditions. First, tristimulus values for both the sample and white are transformed to spectrally-sharpened cone responses, using the transformation.

CIECAM02 does not explicitly construct a color space [23]. However, a cylindrical color space can be built with CIECAM02 lightness, chroma, and hue correlates (*J*,*C*,*h*) and *J*, *C*cos(*h*), *and C*sin(*h*) that could be used if a rectangular color space is preferable. The JCh color space is not totally uniform (the same as CIEL*a*b*). Although, this is one of the most uniform color spaces in existence that accounts for all the perceptual phenomena as mentioned above. It has also an advanced metric very similar to CIEDE2000 in CIEL*a*b* to correct the remaining non-uniformity [24] by Eq. (1).

$$\Delta E_{02-OPT} = [(\Delta J/k_J S_J)^2 + (\Delta C/k_C S_C)^2 + (\Delta H/k_H S_H)^2]^{1/2}$$

$$S_J = 0.5 + (\bar{J}/100)^2; \quad S_C = 1 + 0.02\bar{C}; \quad S_H = 1 + 0.01\bar{C}$$

$$k_J = k_C = k_H = 1$$

$$\Delta H = 2\sqrt{C_s C_r}\sin\left(\frac{\Delta h}{2}\right)$$

$$\bar{C} = \frac{C_s + C_r}{2}; \quad \bar{J} = \frac{J_s + J_r}{2} \tag{1}$$

where, $\bar{C}$is the mean of the reference color chroma C_r and the sample color chroma C_s and $\bar{J}$ is the corresponding value for lightness. S_J, S_C, and S_H are parametric functions that allow the adaptation of the formula to the experimental data set for whom the equation is derived. k_J, k_C and k_H are weighting factors that are set to one by default that can be changed regarding to viewing conditions. Accordingly, the *JCh* uniform color space is the color system adopted by the proposed method because of its better uniformity and adaptation to human perception.

4. Pattern analysis and feature vector construction

Pattern or texture analysis plays an important role in many image processing tasks such as in remote sensing, medical, natural scenes and content based image retrieval (CBIR) systems. The main benefit of extracting the effective texture features is that they provide better classification results. A number of researchers have proposed algorithms for texture analysis, but they are limited to gray scale or have used non-uniform color space. In this paper, optimized color–texture features are extracted in a uniform color space. The proposed pattern analysis method is based on color and texture features and consists of the following steps: (1) extraction of color features such as color occurrence and color similarity, and (2) extraction of texture features by using the multiscale *SPT* decomposition approach. Afterwards, a normalized feature vector is constructed. These steps are further explained in the following subsections.

4.1. Color-related features extraction

Color feature extraction from dermoscopy images [7] plays an important role for early diagnosis of melanoma and benign skin lesions. To examine skin lesions, dermatologists usually have identified six colors. These colors are light-brown, dark-brown, white, red, blue, and black. In fact, different individuals perceive skin lesion colors in a distinct manner. Especially, there are problems with separating light-brown from dark-brown, but problems also occur with red and dark-brown due to a rather reddish glow of the dark-brown color in skin lesions. To perform better differentiation among objects, the human visual system (HVS) [25] utilizes color features separate from texture ones. By following these assumptions, color attributes are calculated by using *JCh* (lightness, chroma and hue) uniform color space of the CIECAM02 color appearance model to improve its adaptation to dermatologist's perception.

For color feature extraction, the spatially adaptive six shades of dominant colors are calculated. The presented approach is followed from [26] but adjusted to uniform color space. To calculate color features, a number of locally adapted dominant colors and their corresponding percentage of occurrence of each color within a certain neighborhood, are calculated as

$$\alpha_{JCh}(x,y,N_{x,y}) = \left\{(c_i^{JCh}, p_i),\ i = 1,2,3,..,M = 6, p_i \in [0,1]\right\} \tag{2}$$

where each of the dominant colors, c_i^{JCh} is a 3-D vector in *JCh* color space, and p_i are the corresponding percentages. $N_{x,y}$ denotes the neighborhood of the pixel at location (*x*,*y*) and *M* is the total number of colors in the neighborhood. To measure the spatial adaptive dominant colors, the follow technique is presented in [27], which is generalized form of iterative *k*-means clustering. This developed technique is adaptive and includes spatial constraints to segment the image into fixed *k* number of clusters. A typical value is $k=6$. In this clustering technique, every pixel of the image is represented by a color that is equal to the average color of the pixels in its neighborhood that belong to that class. As in the immediate neighborhood of a pixel, we can assume that the dominant colors

are slowly varying and are approximately constant. Accordingly, the color composition feature illustration of the Eq. (2) at each point in the image consists of the K characteristics colors of each class and their relative percentages. After calculating colors their percentage of occurrence, color symmetry of tumor regions are needed to calculate because it is an important measure of the pigment distribution in a certain neighborhood.

To find out color symmetry, a technique based on the perceived color differences of the set of six dominant colors is computed from Eq. (2). The difference measure is adopted by the proposed method that is corresponding to Eq. (1). This measure corrects the remainder non-uniformity, which is present in *JCh* color space. Afterwards, symmetric distance differences between two locally adapted dominant colors α^1_{JCh} and α^2_{JCh}, $D_{JCh}(\alpha^1_{JCh}, \alpha^2_{JCh})$, are calculated as

$$D_{JCh}(\alpha^1_{JCh},\alpha^2_{JCh}) = \sum_{i=1}^{M} \Delta E_{02-OPT}(c_i^{JCh}(1),c_i^{JCh}(2)) * p_i \tag{3}$$

where $c_i^{JCh}(1)$, $c_i^{JCh}(2)$ and p_i are the matched colors and their respective percentages. ΔE_{02-OPT}represents color difference formula in *JCh* uniform color space, which is calculated from Eq. (1). Consequently, the color percentage and symmetric distance differences are utilized for the definition of color features. By using Eqs. (2) and (3), f^{JCh}_{color} color feature vector is constructed.

$$f^{JCh}_{color} = D_{JCh}(\alpha^1_{JCh},\alpha^2_{JCh}),\alpha_{JCh}(x,y,N_{x,y}) \tag{4}$$

4.2. Texture-related features extraction

The aforementioned color features rely only on the color characteristics of the dermoscopy image. However, color features do not give important characteristics to quantify the texture characteristics of the lesions.

Several studies [12,13,15–18] proposed to extract texture features in skin lesion images. However, these studies focused more on extracting statistical properties and hence did not consider both local and global spatially correlated relationships among pixels. In case of multicomponent, PRP and homogeneous patterns, it is very difficult to analyze texture by just considering global texture [16] properties. As a result, the differentiation of Melanoma from other pigmented skin lesions becomes a difficult task. To solve this problem, a solution is developed based on the multiscale frequency decomposition algorithm in a perceptually uniform color space. Steerable pyramids transformation (SPT) as a multiscale frequency decomposition algorithm is applied to the J plane of *JCh* color space to get an advantage of the perceptual uniformity.

In general, multiscale feature extraction provides an effective solution for pattern recognition as compared to co-occurrence [28] matrix. Among multiscale feature extraction algorithms, the most popular are discrete wavelet transform (DWT) [29,30], Gabor wavelets (GWs) [31], local binary pattern (LBP) [32], and the Steerable pyramid transform (SPT) [33,34]. These approaches have been shown to be very useful in capturing texture characteristics with high discriminatory power.

The SPT decomposition algorithm is a multiscale and multidirectional representation of frequency transform similar to DWT or GWs, but with interesting translation and rotation invariance properties. In addition to this, SPT combines the benefits of both GWs and wavelet transform, making a multiscale form of the image in a pyramid hierarchy. SPT transform was first introduced by Freeman and Adelson [33,34]. Recently, several studies have investigated the discriminating power of steerable pyramid-based features in various applications including: image denoising, textures classification, digital watermarking and image processing. Particularly in [26], SPT decomposition algorithm was effectively adapted for segmentation of objects to achieve its approximate adaptation to human visual cortex (V1 area) that can be used to produce any number of orientations subbands. The physiological studies of visual cortex [35] also proved that a feature vector constructed in the frequency domain achieves an effective texture discrimination power. Besides that SPT decomposition provides local texture characteristics, while keeping global properties. Accordingly, this SPT decomposition method is adapted for pattern analysis in dermoscopy images.

Multiscale texture feature extraction by one-level and 4-oriented SPT decomposition is shown in Fig. 4. One-level decomposition is empirically determined to be sufficient for dermoscopy texture analysis. In order to extract multiscale texture features, first input J plane image of size (450×450) pixels is divided into eight rectangular blocks of size (32×32) pixels. As determined by experimentation, the total eight blocks divisions are performed for measuring the statistical properties of different tumor's region. However, in order to apply multiscale decomposition step, the whole area of J plane dermoscopy images is considered.

The J plane input is first decomposed into highpass and lowpass band images by using steerable [34] high and lowpass filters. Next, a lowpass image is decomposed into sub-band images by using four orientation band-pass filters. The mathematical representation of the decomposition step is given as follows:

Let $I_J(x,y)$ be 2-D J image in *JCh* color space, which is first decomposed into highpass residual band and a lowpass band by using steerable filters H_0 and L_0, respectively. This lowpass band is then split into a set of oriented bandpass subbands $B_0,B_1,...,B_k$ and a lower lowpass subband L_1 where $k=4$. For a given $I_J(x,y)$ image, its steerable pyramid decomposition $I^{mn}_{SPT}(x,y)$ is defined as:

$$I^{mn}_{SPT}(x,y) = \sum_{x'}\sum_{y'} I_J(x',y')B_{mn}(x-x',y-y'), \quad n=0,1,...,k-1 \tag{5}$$

where B_{mn} denotes the directional bandpass filters at level $m=1$, and orientation $n=0,1,...,k-1$. After an image is decomposed into set of subimages ($I^n_{SPT}(x,y)$), statistical properties are calculated for measuring the texture characteristics and these properties are extracted on each block of size (32×32) pixel of lowpass, highpass band and every subband images. Mean (μ), Standard deviation (σ) and Skewness (ρ) texture characteristics are extracted from lowpass, highpass band and every subband $I^{mn}_{SPT}(x,y)$ oriented images. In order to minimize the prediction error of the classification model, we applied a feature selection step to reduce the dimensionality of feature vectors by selecting the most informative features. Since, the subband images have shared some of the common statistical texture features compared to lowpass and highpass band images that must be minimized. Therefore, principal-component analysis (PCA) method is applied to get more informative features. 24-diagonal eigenvalues ($D_{\lambda 24}$) are selected by PCA method to each statistical attributes extracted from four subband $I^{mn}_{SPT}(x,y)$ decomposed images. We have empirically determined that these 24 eigenvalues describe the best discrimination among the textures of each subband images compared to other values. In total $(8 \times 3 \times 2)=48$ μ, σ and ρ statistical attributes are also extracted from each block of lowpass and highpass band images denoted by $L_{8\mu}$, $L_{8\sigma}$, $L_{8\rho}$, $H_{8\mu}$, $H_{8\sigma}$, $H_{8\rho}$. Consequently, each input image is transformed into a $(48+24)=72$-dimenional multiscale texture features vector.

Next a texture feature vector based on statistical properties of low, high and lowpass subband images is created by combining these values as

$$f^J_{texture} = L_{8\mu},L_{8\sigma},L_{8\rho},H_{8\mu},H_{8\sigma},H_{8\rho},D_{\lambda 24} \tag{6}$$

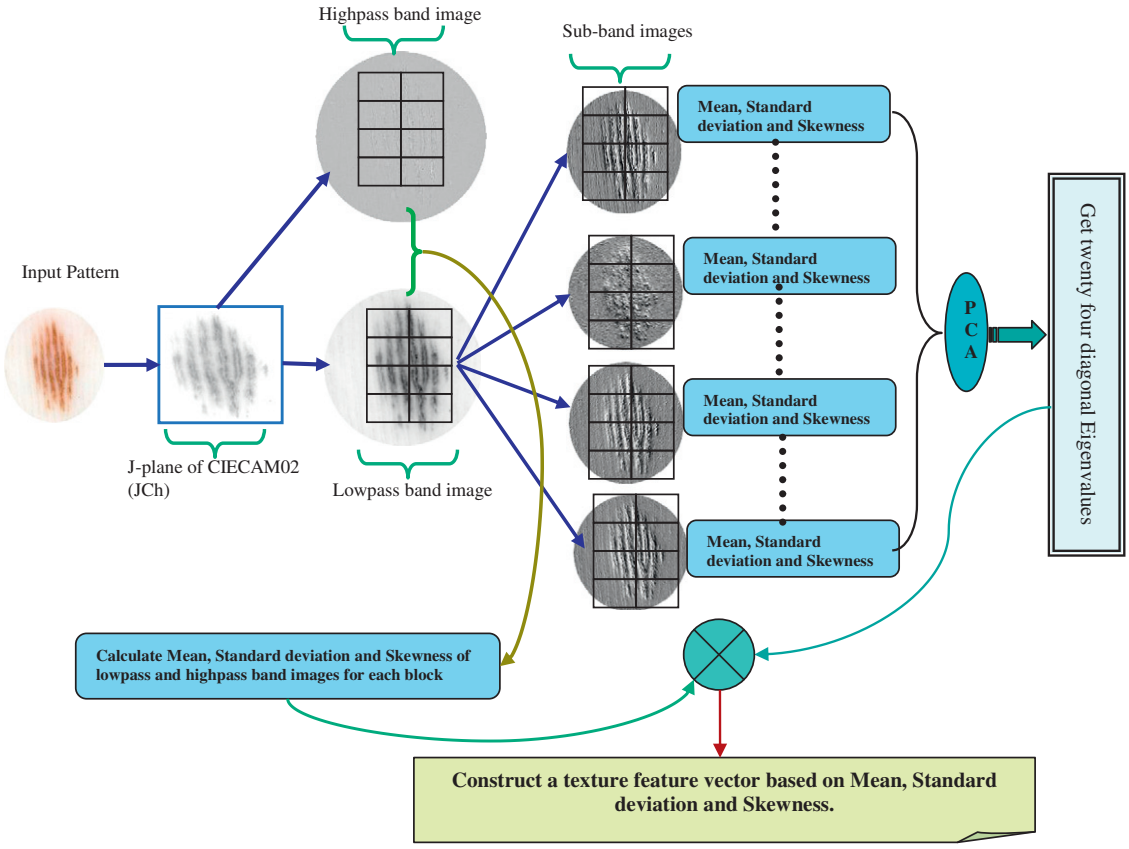

Fig. 4. Multiscale texture features extracted by one-level steerable pyramids decomposition method.

An optimal feature vector is generated based on the combination of color and texture properties that is further normalized using normal-probability density function (PDF). The PDF function is utilized to transform the feature vector into zero mean and unit variance. This feature vector is shown in Eq. (7), where m denotes the number of color features and n represents the number of texture features.

$$f_i = [f^{JCh}_{color}(m), f^{J}_{texture}(n)], \quad \forall i = 1,2,3,\ldots,m+n \tag{7}$$

5. Pattern classification model by multi-label learning

After extracting the set of features (f_i), the next step is to devise a suitable machine learning algorithm to assess the features' discriminative power. Melanomas often exhibit multicomponent patterns so any learning algorithm must take multi-label input and provide multi-label output at the same time. Multi-label learning [36] refers to the classification problem where each input pattern can be assigned to multiple class labels, simultaneously. It has found applications in many domains, such as computer vision, human computer interaction, bioinformatics, and physiology. Among multi-classes and multi-label learning algorithms, the most popular algorithms are support vector machine (ML-SVM) [37], ML-KNN [38] and multi-label ranking [39].

Boosting algorithms [40] are a set of nonparametric meta-learning algorithms, which can provide optimal classification results. In general, the advantages of using adaptive boosting (AdaBoost) algorithm over other machine learning algorithm are its computational efficiency, better robustness and no regressions. In medical image analysis studies, SVM classifiers have been widely used, but as AdaBoost can choose good informative features from potentially very large feature data sets [41]. This can significantly reduce the need for experts for selection of useful features based on their knowledge. However the major problem of boosting is that, it does not provide multiple class-label outputs with max-margin without class-correlation. Therefore nowadays, the boosting algorithms [37,42] are extended to multi-class outputs, which are based on AdaBoost.MH [43].

For dermoscopy pattern discrimination as discussed in Section 1.1, a multi-class input boosting algorithm (AdaBoost.MH) is adopted, which is extended to multi-class output by using maximum a posterior (MAP) and robust ranking principles named as AdaBoost.MC classifier. Compare to existing multi-label ranking algorithms, the AdaBoost.MC classifier is robust in the sense that its rank classes are based on optimized scores, which are calculated during evaluation of MAP and boosting. As a result, this AdaBoost.MC multi-label learning algorithm is designed to provide maxim-margin without the class-correlation problem.

5.1. AdaBoost.MC classification algorithm

Assume that $x_i \in f_i$, $i = 1,2,3,\ldots,n$ are the collection of training examples, which belong to a set of class labels y_i denoted as $\{(x_1,y_1),(x_2,y_2),\ldots,(x_n,y_n)\}$ where $x_i \in X$ and $y_i \subset Y = \{y_1,y_2,\ldots,y_n\}$. For solving this multi-class problem, AdaBoost.MH [43] algorithm is adopted, which is further extended to multi-class label outputs by MAP and robust ranking principles called AdaBoost.MC. The detailed pseudocode of the algorithm AdaBoost.MC is given in Fig. 5. In this classification algorithm, decision stumps are used as weak learners during boosting process. A decision stump is a machine learning model consisting of a one-level decision tree. Decisions stumps components are often used as a weak learner for ensemble machine learning algorithms such as Boosting. By using 1-vs-All technique in the Boosting algorithm, a binary

Input Let training features data and corresponding class labels $\{(x_1,y_1),(x_2,y_2),...,(x_n,y_n)\}$ are denoted where $x_i \in f_i, \quad i=1,2,3,...,n$ and $y_i \subset Y=\{y_1,y_2,...,y_n\}$, respectively.

Training Function multi_**AdaBoost** (x,y,n,m,T)

Initialize $W(x_i,y_j)=1/mn, \quad \forall\, i=1,2,...,n$ and $j=1,2,...,m$, $\alpha=1$, T=250

For $t=1,2,3\ \ldots,\ T$ and while $x_t \neq 0$ do

A. Train the **Weak learner** algorithm for distribution $W(x_i,y_j)$ using decision stumps.

B. Get the Weak hypothesis h_t from the **Weak learner.**

C. Calculate the error of

$$h_t:\varepsilon_t=\sum_{h_t(x)\neq y_j} W(x_i,y_j)$$

D. Set $\alpha_t=1/2(\ln(1-\varepsilon_t)/\varepsilon_t)$

E. Update the weights and hypothesis

$$h_t(x_i,y_j)=h_t(x_i,y_j)+\alpha_t h(x_i,y_j)$$

$$W_{t+1}(x_i,y_j)=W_t(x_i,y_j)\exp(-h(x_i,y_j).h_t(x_i,y_j))/z_t$$

Where $z_t=\sum_i^n\sum_j^m W_t(x_i,y_j)\exp(-h(x_i,y_j).h_t(x_i,y_j))$ is the normalization factor such that $\sum_i^n\sum_j^m W_{t+1}(x_i,y_j)=1$.

End.

Function End.

Output Repeatedly called multi_**AdaBoost** function for all class labels (l) to generate generalized classification model and then get the Output of the final hypothesis:

$$H(l)=\sum_{t=1}^{T} h_t(x,y)$$

Testing **Maximum a posteriori (MAP):** Given an unlabeled feature vector $x_i \in f_i, \quad i=1,2,3,...,n$, to match with class-label sets.

1. Obtained prior and maximum a posterior probabilities of each class labels with maximum weighted score by weak learner from the training data.

2. Finally, use the Bayes rule and ranking probability to select the maximum probabilities of $x_i \in f_i, \quad i=1,2,3,...,n$ belongs to a set of class labels with max-margin.

Fig. 5. Multi-label pattern classification (AdaBoost.MC) by boosting (AdaBoost.MH), Maximum a posteriori (MAP) and robust ranking principles.

classifier is obtained for each label class without learning the complex decision boundaries.

To generate a sequence $h_1,h_2,...,h_t$ of weak hypotheses, AdaBoost.MH works by iteratively (t) calling a weak learner. At the end of this iteration step, the final hypothesis H is obtained as a sum $H(l)=\sum_{t=1}^{T} h_t$ of these weak hypotheses for class labels (l). A weak hypothesis is a function i.e., $h_t: X\times Y\rightarrow R$. We interpret the sign of $h_t(x_i,y_j)$ as the prediction of h_t on whether x_i belongs toy_j, i.e. $h_t(x_i,y_j)>0$ means that x_i is believed to belong to y_j while $h_t(x_i,y_j)<0$ means it is believed not to belong toy_j.

After that principle of maximum a posteriori (MAP) is utilized to determine the class label set for the dermoscopy test pattern. Bayes rule was used to describe a combination of feature values and a class value. Traditionally, Bayes rule is mostly used to select features for multi-instance learning. Therefore, this rule is integrated in this Boosting algorithm to ensemble the results produce by MAP method. Finally, to get multi-label class decision in multiple states is performed by a robust ranking principle. The implementation details of MAP rule and robust ranking methods are described in the subsequent paragraphs.

The corresponding rule for determining the class label set for input feature vector $(f_1,f_2,f_3,...,f_n)$ by using MAP principle can be defined as follows:

$$\vec{y}(l)=\underset{y_j\in Y}{\operatorname{argmax}}\, P(f_1,f_2,f_3,...,f_n|y_j) \tag{8}$$

where $f_1,f_2,f_3,...,f_n$ is the set of feature values that describe the new instance, and $\vec{y}(l)$ is the most probable hypothesis to find

out class label set for a test pattern. Using Bayes rule, Eq. (9) can be rewritten as follows:

$$\overrightarrow{y}(l) = \underset{y_j \in Y}{\operatorname{argmax}} ((P(y_j)P(f_1,f_2,f_3,...,f_n|y_j))/P(f_1,f_2,f_3,...,f_n)) \quad (9)$$

Next, the formulation of multi-class labels problem is given by:

$$\overrightarrow{y}(l) = \underset{y_j \in Y}{\operatorname{argmax}} P(y_j)P(f_1,f_2,f_3,...,f_n|y_j) \quad (10)$$

Using training data the two terms $P(y_j)$ and $P(f_1,f_2,f_3,...,f_n|y_j)$ have to be calculated. The class prior probability $P(y_j)$ can be easily estimated by a maximum-weighted score of occurrence of the class value, during a training-phase of weak learner (decisions stump). Let w_i^l be the cumulative weighted probability for multi-classifier decision stumps of each attribute (l), and s_i^l is the maximum score of distinguish points find out from right and left child of decision stump by w_i^l. This prior probability step can be represented as

$$P(y_j) = \underset{y_j \in Y}{\operatorname{argmax}} \sum_{i=1}^{n} |w_i^l|, s_i^l \quad (11)$$

Alternatively, estimating the posterior probability $P(f_1,f_2,f_3,...,f_n|y_j)$ terms is calculated by evaluating the multi-AdaBoost model hypothesis in terms of decision i.e. $H(l)$. In this evaluation step of boosting, the classes with highest posterior probability are returned. The class label set is selected for input pattern based on these prior and maximum posterior probabilities.

The objective of multi-label learning algorithms is to learn a set of labels with max-margin without class-correlations problem, which must have these characteristics among classes without imposing the strict boundary conditions. Moreover, MAP based on Bayes rule provides initial probabilities of each class to match with the set of input class, but cannot rank that output class labels. As a result, we need to reformulate Eq. (10) in order to introduce efficient optimization, max-margin and provide ranking for multi-label learning problem. Under these assumptions, the problem can be expressed in terms of ranking probability as

$$\underset{w_k}{\operatorname{argmin}} P_{ij}|\phi_c(w_k(x_i,y_j)) - \phi_c(w_{k-1}(x_i,y_j))| \le \delta, \quad \forall k = 1,2,...,n \quad (12)$$

where P_{ij} denotes the maximum a posterior (MAP) probability of the test pattern, which was calculated from Eq. (10). Also, the multi-label classifier is ranked according to δ parameter, which is the maximum significance level of every class (K) belongs to particular category. Accordingly, the value of this objective function ($\operatorname{argmin}_{w_k} P_{ij}|\bullet|$) must be less than or equal to δ parameter. The value of δ parameter is experimentally defined as 0.99. The $\phi_c(w_k(x_i,y_j))$ function represents the weighted probability $w_k(x_i,y_j)$ of every class label $\phi_c(\bullet)$ calculated during training step. Since, the aim is to minimize this objective function by following the ranking probability scheme. Let θ_k be the weighted probability of each label class $\phi_c(w_k(x_i,y_j))$ then Eq. (12) can be rewritten as

$$\underset{w_k}{\operatorname{argmin}} P_{ij}|\phi_c(\theta_k) - \phi_c(\theta_{k-1})| \le \delta, \quad \forall k = 1,2,,...,n \quad (13)$$

Next, we select first, second, third and so on classes (q_{ij}) based on the weighed-ranking score to the test pattern features during MAP calculation, which is denoted by S_c. Finally, the probability of each class (P_k) is maximized with respect to other classes by the weighted score S_c, which is calculated as

$$\theta = S_c : \max_{P_k} \arg(P_k/(1-P_k))/(P_{k-1}/(1-P_{k-1})) \quad (14)$$

where, we have to choose the class label set for test pattern based on the dynamic selection criteria for every value of $(i,j) \in n$ classes as

$$q_{ij} = \begin{cases} 1, & \text{if } i=j=a,-a \\ \theta/(1+\theta), & \text{if } j>i, i<a \\ 1/(1+\theta), & \text{if } j<i, i>-a \\ 0, & \text{otherwise} \end{cases} \quad (15)$$

Such that

$$a = (\ln(1-\Psi)/\ln(\theta)), \quad \Psi = 0.05 \quad (16)$$

where, Ψ denotes the threshold level of significance, which equals 0.05. The value of Ψ is determined by experimental analysis. Also, the condition on θ_k is applied such that $\theta_k \approx \theta_{k-1} + \delta$. This formulation for learning a max-margin multi-label outputs classifier is projected by using robust ranking scheme. To achieve a perfect ranking of all labels, the optimized weights are derived based on prior and posterior probabilities. Consequently to develop AdaBoost.MC classifier, multi-label classification algorithm is acquired that maximizes the margins among class labels.

6. Experimental setup

6.1. Dermoscopy data set

Skin lesions of dermoscopy images are used in the experiment, which has been collected as a CD resource from the two European university hospitals as part of the EDRA-CDROM, 2002 [44]. This data set contained 1039 color images in total with spatial resolution of 768×512 pixels. All these images were captured during routine clinical assessments to imitate the a priori probabilities of the clinical diagnosis. In total 350 dermoscopic images are selected from this data set as Reticular (50), Globular (50), Cobblestone (50), Homogeneous (50), Parallel ridge (50), Starburst (50) and Multicomponent (50). After selecting the 350 dermoscopy images, a region of size (450×450) is automatically selected from the center of those images. Next, each image is preprocessed by transforming into CIECAM02 color appearance model of (*JCh*) uniform color space. After that, the color and texture features are extracted and then construct a normalize feature vector to classify pattern classes. In the classification step, 20% of dermoscopy images are used of each class for testing and 80% for training the AdaBoost.MC classifier. Finally, the input test pattern is matched with the set of class label based on MAP and robust ranking techniques, which gives best discrimination power among patterns allowing the property of max-margins.

6.2. Statistical analysis

The performance of proposed model based on concept of multi-class instances and multi-label outputs is evaluated using statistical methods. To evaluate the diagnostic performance of this model, the area under curve (*AUC*) [45] of the receiver operating characteristic (ROC) analysis is used. Specifically, *AUC* is used to investigate the sensitivity: *SE* and specificity: *SP*. The area under the curve (*AUC*) is a commonly used index to assess the overall discrimination. The *AUC* ranges from 0.50 to 1.0 and the greater its value the higher is the pattern classification accuracy. To calculate *SE* and *SP*, the true positive rate (*TPR*) and false positive rate (*FPR*) are measured, respectively as

$$SE = TPR = TP/(TP+FN) \quad (17)$$

$$FPR = FP/(FP+TN) \quad (18)$$

and

$$SP = 1 - FPR \tag{19}$$

Table 1
Pattern classification performance of different dermoscopic structures.

Pattern detectors classes	*SE* (%)	*SP* (%)	*E*	*AUC*
Reticular	87.11	97.96	0.459	0.981
Globular	86.25	97.21	0.477	0.997
Cobblestone	87.76	93.23	0.555	0.990
Homogeneous	90.47	95.10	0.697	0.996
Parallel ridge	85.25	89.50	0.524	0.989
Starburst	89.62	90.14	0.634	0.966
Multicomponent	98.50	93.11	0.344	0.989

Number of classifiers in case of melanomas with four pattern detectors as; *SE*=Sensitivity, *SP*=Specificity, *E*=average standard deviation training error during learning by AdaBoost.MC, and *AUC*=Area under the Receiver operating characteristic (ROC) curve.

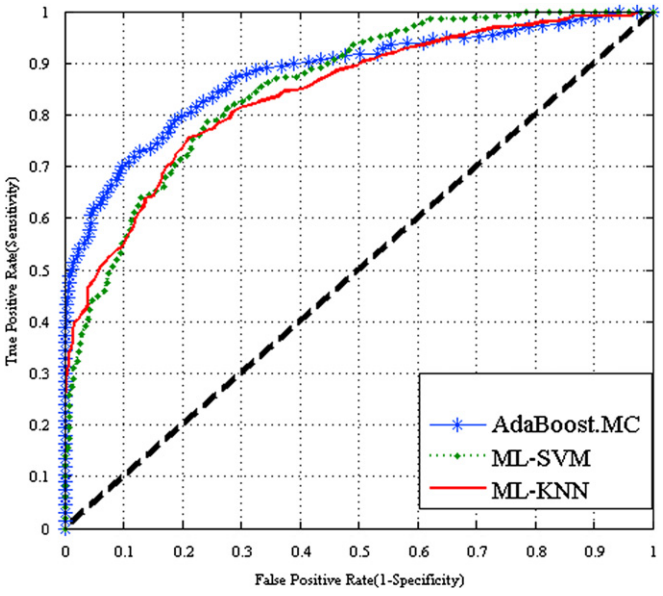

Fig. 6. Receiver operating characteristics (ROC) curves for the multi-label learning algorithms.

where *TP* and *FP* represent the number of true positives and false positives, and *TN* and *FN* denote the number of true negatives and false negatives, respectively. The *TPR* value is determined by a classifier performance on classifying positive instances correctly among positive data set and the *FPR* value is defined how many incorrect positive results occur among negative samples. For experimental analysis, the data set is divided into training (80%) and testing (20%). During training the proposed boosting classifier of each round, we also calculated the training error in terms of standard deviation of mean error (*E*). The significance level of AdaBoost.MC was compared with state-of-the-art classification algorithms such as (ML-SVM [37] and ML-KNN [38]) using 10-fold cross-validation test. The parameters used for ML-SVM and ML-KNN are type=RBF, para=1, cost=1 and rounds=250 along with $k=7$(fixed cluster) values.

6.3. Computational time

The proposed system is currently implemented in MATLAB 7.6.0.324® (The Mathworks, Natick, MA) on a 2.0 Core to Duo 32-bit Intel processor system with 1 GB DDR2 RAM, running Windows 7. On average, it takes 0.346 s to transform 230 RGB dermoscopy image to *JCh* color space. In order to extract color and texture features, on average 2.23 s and 1.78 s are spent on each dermoscopy images for constructing the training and testing the data set, respectively. Since, for training and testing the AdaBoost.MC classifier on this data set, 7.56 and 3.65 s are consumed on average, respectively. This time is calculated on fixed 250 iterations. In case of testing an input pattern of image size (768×512) pixels, on average 3.65 s is taken for output class matching. This time can be further reduced by using optimized C/C++ implementation.

7. Results and discussion

The average results of the proposed model on this data set are shown in Table 1. Table 1 demonstrates each pattern classifier by using boosting, maximum a posterior (MAP) and robust ranking principles in terms of sensitivity: *SE*, specificity: *SP* and average error(*E*) along with area under the *ROC* curve (*AUC*) analysis. Fig. 6 shows the corresponding receiving operating characteristic curve

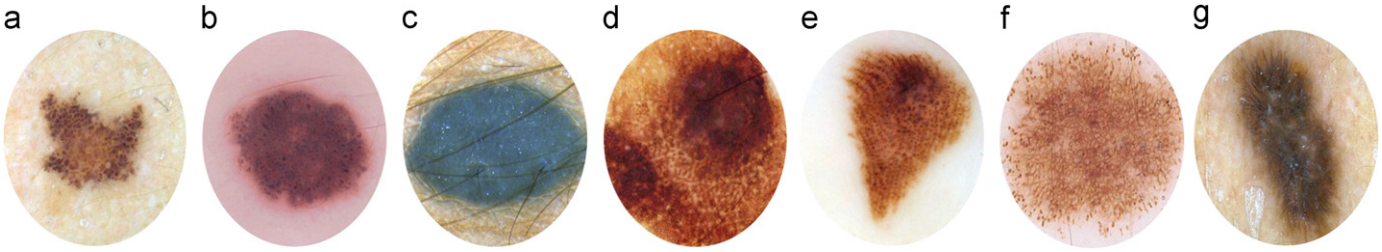

Fig. 7. Seven different dermoscopy images match with pattern classes, where (a) cobblestone pattern, (b) globular pattern, (c) homogeneous pattern, (d) multicomponent pattern includes (cobblestone, globular, homogeneous, pigmented network), (e) parallel pattern, (f) Reticular or pigmented network, and (g) starburst pattern.

Table 2
Comparisons of classification performance by different methods using 10-fold cross validation.

No.	[a]Classifiers	Patterns detection errors						
		Reticular (%)	Globular (%)	Cobblestone (%)	Homogeneous (%)	Parallel-ridge (%)	Starburst (%)	Multicomponent (%)
1	ML-SVM	14	16	12	9	10	13	8
2	ML-KNN	16	20	16	13	14	18	19
3	AdaBoost.MC	3	6	4	2	4	4	3

[a] The classifiers are used in this comparisons were; ML-KNN=K-Nearest Neighbor, ML-SVM=Support vector machine and AdaBoost-MC=proposed AdaBoost classification algorithm.

(*ROC*) of each multi-label machine learning algorithm. In this figure, each classifier performance is shown based on the average *TPR* and *FPR* values over the seven pattern classes. It shows that significantly better performance is obtained by the proposed multi-label AdaBoost.MC learning algorithm as compared to others. As displayed also in Table 1, it can be noticed that in case of Homogeneous; Globular; Cobblestone and Multicomponent patterns, the best performance has been measured i.e., *AUC*: 0.997. The proposed pattern classifier significantly improves the performance with the average value of *SE*: 89.29, *SP*: 93.25 and *AUC*: 0.986. Moreover, the sample results that are matched with pattern classes are also shown in Fig. 7.

On the other hand, the classification accuracy of AdaBoost.MC is compared against the sate-of-the-art classification algorithms such as ML-KNN and ML-SVM by 10-fold cross validation test using 1-vs-all approach. The results are displayed in Table 2. Hence, the performance of the proposed patterns detector by AdaBoost.MC has been shown to be better than ML-KNN and ML-SVM algorithms. In fact, ML-KNN is based on the ranking scheme and to predict a new input data for ranking can be problematic. But the classification accuracy of new input multi-component pattern for ML-SVM is somewhat lower than AdaBoost.MC, which is overall better than ML-KNN. AdaBoost.MC algorithm performs well due to the fact that it selected the most discriminating feature, which is getting characteristics to maximize the margins based on optimized Bayes rule and robust ranking probabilities for multi-label output. However, ML-SVM does not select most discriminating features but provide maximum margins. By empirical analysis, it was observed that ML-KNN and MK-SVM are not designed to select appropriate feature sets, and effectively rank the final multi-label output classification decision, which are robustly implemented in AdaBoost.MC algorithm. Time computation is an important factor to evaluate the classification algorithms for showing the performance in real-time applications. As shown in Table 3, the AdaBoost.MC is faster when compared to ML-KNN and ML-SVM. On average, AdaBoost.MC (training: 7.56 and testing: 3.65) seconds (s) are taken, while ML-KNN (training: 19.89 s and testing: 13.22 s) and ML-SVM (training: 14.27 s and testing: 9.24 s).

In contrast of these comparisons with the state-of-the-art algorithms, a separate evaluation of the proposed AdaBoost.MC model is also performed in different color spaces such as RGB, CIE L∗a∗b and *JCh*. Table 4 summarizes the mean classification errors of the proposed method in RGB, CIE L*a*b* and *JCh* color spaces. As shown in Table 3, by using uniform color space (*JCh*), the pattern classification results are greatly enhanced. The reason is that by using uniform color space (*JCh*), the color and texture features are effectively extracted in comparison to other color spaces.

The proposed pattern classification algorithm based on CASH is highly accurate, when tested on this dermoscopy data set. However, there are some dermoscopy images in which this system provides false positive output. This is often due to lesion covered by artifacts such as heavy hair (Fig. 8(a)), effect of dermoscopy-gel (Fig. 8(b)) or some diffused-regions (Fig. 8(c)). For instance, the lesion in Fig. 8(a) is considered as a parallel pattern-class, while it belongs to cobblestone pattern. Similarly, the other two lesions are regarded as cobblestone, and (homogeneous, parallel) patterns instead of globular and homogeneous patterns, which are shown in Fig. 8(b) and Fig. 8(c), respectively. Due to these problems, the algorithm is sensitive to the artifacts. Accordingly, the classification accuracy can be increased by removing these artifacts from lesions.

Table 3
Comparisons of classification performance in terms of average computational time.

No.	[a]Classifiers	Patterns detection errors	
		Training(s)	Testing (s)
1	ML-KNN	19.89	13.22
2	ML-SVM	14.27	9.24
3	AdaBoost.MC	7.56	3.65

[a] The classifiers are used in this comparisons were; ML-KNN=multi-label K-Nearest Neighbor, Support vector machine (ML-SVM) and Ada-Boost-MC=proposed AdaBoost multi-label classification algorithm.

8. Conclusion

In this paper, a novel pattern classification model related to dermatologist's perception is proposed. The purpose of this study

Table 4
Comparisons of proposed Adaboost.Mc pattern classifier by using mean errors in terms of different color spaces.

No.	[a]Color spaces	Pattern detection errors						
		Reticular (%)	Globular (%)	Cobblestone (%)	Homogeneous (%)	Parallel-ridge (%)	Starburst (%)	Multicomponent (%)
1	HSV	5	7	12	11	17	14	12
2	CIEL*a*b*	3	4	8	9	13	11	7
3	*JCh*	2	1	3	5	4	2	5

[a] The proposed classifier are compared with different color spaces such as in HSV, CIEL*a*b* and *JCh*.

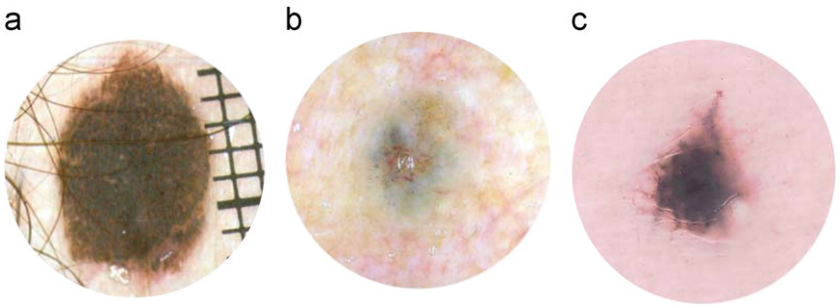

Fig. 8. Example of wrongly classified deromscopy images, where (a) cobblestone pattern, (b) globular pattern and (c) homogeneous pattern.

was to develop an effective pattern classification (AdaBoost.MC) model based on extraction of color symmetry and multiscale-texture feature in uniform CIECAM02 (*JCh*) color space. To the best of our knowledge, many studies for lesions classification of dermoscopy have been devoted towards clinical ABCD rule but few studies on pattern analysis have been developed in the literature. Since, none of them consider the problem of multi-component patterns, which are mostly observed in melanoma and Clark nevi lesions. In fact, previous pattern detection methods attempted to classify multi-pattern lesions by just considering single pattern class. In contrast, in our study, multiple pattern input technique is presented to develop classifier model for providing multiple class-label outputs. By developing AdaBoost.MC multi-label algorithm, the groups such as melanoma and nevi lesions are easily classified. The overall performance of the proposed algorithm is measured in terms of effectiveness and time. On average, sensitivity (*SE*): 89.28%, specificity (*SP*): 93.75% and area under the curve (*AUC*): 0.986 are obtained. The AdaBoost.MC algorithm achieved significantly higher classification rate and faster than state-of-the-art classification methods due to use of robust ranking with max-marginal characteristics, which avoids class correlation problem. Furthermore, an entirely automated system is developed, which has more sophisticated pattern classifier for dermoscopy images that focused more on color and textural properties of the lesion regions in a uniform color space derived from a color appearance model and making use of an advanced color distance measure. The proposed system appears to be sufficiently accurate, robust, and computationally fast for discrimination of lesions, which helps the dermatologists in a "screening support system" or in a CAD tool. In the future work, more local pattern classes will be added with the combination of ABCD technique to classify among lesions and pre-process them before extraction of color–texture features.

Conflict of interest

All authors in this paper have no potential conflict of interest.

Acknowledgments

This research was supported by the National 863 High Technology Research and Development Program of China (Grant no. 2006AA02Z347) and the Chinese Scholarship Council (CSC) (Grant no. 2008GXZ143). We would like to thank anonymous reviewers for their valuable comments and suggestions.

References

[1] R.H. Johr, Dermoscopy: alternative melanocytic algorithms—the ABCD rule of dermatoscopy, menzies scoring method, and 7-point checklist, Clinics in Dermatology 20 (3) (2002) 240–247.
[2] H. Blum, Luedtke, U. Ellwanger, et al., Digital image analysis for diagnosis of cutaneous melanoma. Development of a highly effective computer algorithm based on analysis of 837 melanocytic lesions, British Journal of Dermatology 151 (5) (2004) 1029–1038.
[3] G. Argenziano, H.P. Soyer, S. Chimenti, R. Talamini, R. Corona, F. Sera, et al., Dermoscopy of pigmented skin lesions: results of a consensus meeting via the Internet, Journal of the American Academy of Dermatology 48 (5) (2003) 679–693.
[4] P. Rubegni, M. Burroni, G. Cevenini, et al., Digital dermoscopy Analysis and artificial neural network for the differentiation of clinically atypical pigmented skin lesions: A retrospective study, The Journal of Investigative Dermatology 119 (2) (2002) 471–474.
[5] G. Capdehourat, A. Corez, A. Bazzano, R. Alonso, P. Musé, Toward a combined tool to assist dermatologists in melanoma detection from dermoscopic images of pigmented skin lesions, Pattern Recognition Letters 32 (16) (2011) 2187–2196.
[6] A.G. Isasi, B.G. Zapirain, A.M. Zorrilla, Melanomas non-invasive diagnosis application based on the ABCD rule and pattern recognition image processing algorithms, Computers in Biology and Medicine 41 (9) (2011) 742–755.
[7] M.E. Celebi, H.A. Kingravi, B. Uddin, et al., A methodological approach to the classification of dermoscopy images, Computerized Medical Imaging and Graphics 31 (6) (2007) 362–373.
[8] A. Mojsilovic, J. Kovacevic, J. Hu, R.J. Safranek, S.K. Ganapathy, Matching and retrieval based on the vocabulary and grammar of color patterns, IEEE Transactions on Image Processing 1 (1) (2000) 38–54.
[9] H. Ganster, A. Pinz, R. Rohrer, E. Wilding, M. Binder, H. Kittler, Automated melanoma recognition, IEEE Transactions on Medical Imaging 20 (3) (2001) 223–239.
[10] A. Safi, M. Baust, O. Pauly, V. Castaneda, T. Lasser, D. Mateus, et al., Computer-aided diagnosis of pigmented skin dermoscopic images, medical content-based retrieval for clinical decision support, Lecture Notes in Computer Science 7075 (2012) 105–115.
[11] M. Burroni, R. Corona, G. Dell'Eva, F. Sera, et al., Melanoma computer-aided diagnosis reliability and feasibility study, Clinical Cancer Research 10 (6) (2004) 1881–1886.
[12] I. Maglogiannis, D.I. Kosmopoulos, Computational vision systems for the detection of malignant melanoma, Oncology Reports 15 (2006) 1027–1032.
[13] D. Ruiz, V. Berenguer, A. Soriano, B. Sánchez, A decision support system for the diagnosis of melanoma: a comparative approach, Expert Systems with Applications 38 (12) (2011) 15217–15223.
[14] T. Tanaka, S. Torii, I. Kabuta, K. Shimizu, M. Tanaka, Pattern classification of nevus with texture analysis, IEEJ Transactions on Electrical and Electronic Engineering 3 (1) (2008) 143–150.
[15] H. Iyatomi, H. Oka, M.E. Celebi, K. Ogawa, G. Argenziano, H.P. Soyer, H. Koga, T. Saida, et al., Computer-based classification of dermoscopy images of melanocytic lesions on acral volar skin, Journal of Investigative Dermatology 128 (8) (2008) 2049–2054.
[16] C. Serrano, B. Acha, Pattern analysis of dermoscopic images based on markov random fields, Pattern Recognition 42 (2009) 1052–1057.
[17] M. Sadeghi, M. Razmara, T.K. Lee, M.S. Atkins, A novel method for detection of pigment network in dermoscopic images using graphs, Computerized Medical Imaging and Graphics 35 (2011) 137–143.
[18] Q. Abbas, M.E. Celebi, I.F. Garcia, Computer-aided pattern classification system for dermoscopy images, Skin Research and Technology, http://dx.doi.org/10.1111/j.1600-0846.2011.00562.x, in press.
[19] E.F. Kelley, M. Lindfors, J. Penczek, Display daylight ambient contrast measurement methods and daylight readability, Journal of the Society for Information Display 14 (11) (2006) 1019–1030.
[20] D.H. Alman, R.S. Berns, G.D. Snyder, W.A. Larsen, Performance testing of color-difference metrics using a color tolerance dataset, Color Research & Application 14 (3) (1989) 139–151.
[21] N. Moroney, M. Fairchild, R. Hunt, C. Li, M. Luo, T. Newman, The CIECAM02 color appearance model, in: Proceedings of the IS&T/SID 10th Color Imaging, Scottsdale, Arizona, November 2002, pp. 23–27.
[22] C.J. Li, M.R. Luo, Testing the robustness of CIECAM02, Color Research & Application 30 (2) (2005) 99–106.
[23] M.D. Fairchild, A revision of CIECAM97s for practical applications, Color Research & Applications 26 (6) (2001) 418–427.
[24] P. Urban, R.S. Berns, M.R. Rosen, Constructing Euclidean color spaces based on color difference formulas, in: Proceedings of the IS&T/SID 15th Color Imaging conference: Color Science and Engineering Systems. Technologies, and Applications, Albuquerque, New Mexico, November 2007, pp.77–82.
[25] T. Mäenpää, M. Pietikäinen, Classification with color and texture: jointly or separately? Pattern Recognition 37 (8) (2004) 1629–1640.
[26] J. Chen, T.N. Pappas, A. Mojsilovie, B.E. Rogowitz, Adaptive perceptual color-texture image segmentation, IEEE Transactions on Image Processing 14 (10) (2005) 1524–1536.
[27] T.N. Pappas, An adaptive clustering algorithm for image segmentation, IEEE Transactions on Signal Processing 40 (4) (1992) 901–914.
[28] A. Samal, J.R. Brandle, D.S. Zhang, Texture as the basis for individual tree identification, Information Sciences 176(5), (2006) 565–576.
[29] S. Li, J.T. Kwok, H. Zhu, Y. Wang, Texture classification using the support vector machines, Pattern Recognition 36 (12) (2003) 2883–2893.
[30] A. Sengur, Wavelets transform and adaptive neuro-fuzzy inference system for color texture classification, Expert Systems with Applications 34 (3) (2008) 2120–2128.
[31] G.M. Haley, B.S. Manjunath, Rotation-invariant texture classification using a complete space-frequency model, IEEE Transactions on Image Processing 8 (2) (1999) 225–269.
[32] H. Zhou, R. Wang, C. Wang, A novel extended local binary pattern operator for texture analysis, Information Sciences 178 (22) (2008) 4314–4325.
[33] W.T. Freeman, E.H. Adelsonk, The design and use of steerable filters, IEEE Transactions on Pattern Analysis and Machine Intelligence 13 (9) (1991) 891–906.
[34] E.P. Simoncelli, W.T. Freeman, E.H. Adeslon, D.J. Heeger, Shiftable multi-scale transforms, IEEE Transactions on Information Theory 38 (2) (1992) 587–607.
[35] John G. Daugman, Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters, Journal of the Optical Society of America 2 (7) (1985) 1160–1169.
[36] M.R. Boutell, J. Luo, X. Shen, C.M. Brown, Learning multi-label scene classification, Pattern Recognition 37 (9) (2004) 1757–1771.

[37] Z.-H. Zhou, M.-L. Zhang, Multi-instance multi-label learning with application to scene classification, in: Proceedings of the Advances in Neural Information Processing Systems (NIPS), Vancouver, Canada, 2006, pp. 1609–1616.
[38] M.-L. Zhang, Z.-H. Zhou, ML-kNN: a lazy learning approach to multi-label learning, Pattern Recognition 40 (7) (2007) 2038–2048.
[39] R. Jin, S. Wang, Z.-H. Zhou, Learning a distance metric from multi-instance multi-label data, in: Proceedings of the IEEE Computer Society on Computer Vision and Pattern Recognition, Miami, FL, USA, June 2009.
[40] J.H. Morra, Z. Tu, L.G. Apostolova, A.E. Green, A.W. Toga, et al., Comparison of AdaBoost and support vector machines for detecting Alzheimer's disease through automated Hippocampal segmentation, IEEE Transactions on Medical Imaging 29 (1) (2010) 30–43.
[41] E. Song, D. Huang, G. Maa, C.-C. Hung, Semi-supervised multi-class Adaboost by exploiting unlabeled data, Expert Systems with Applications 38 (6) (2011) 6720–6726.
[42] M.-L. Zhang, J.M. Peña, V. Robles, Feature selection for multi-label naive bayes classification, Information Sciences 179 (9) (2009) 3218–3229.
[43] R.E. Schapire, Y. Singer, BoosTexter: a boosting-based system for text categorization, Machine Learning 39 (2) (2000) 135–168.
[44] G. Argeniano, P.H. Soyer, V.G. De, P. Carli, M. Delfino, Interactive Atlas of Dermoscopy CD, EDRA medical publishing and New media, 2002.
[45] A.P. Bradley, The use of the area under the ROC curve in the evaluation of machine learning algorithms, Pattern Recognition 30 (7) (1997) 1145–1159.

Qaisar Abbas received his BSc degree in Computer Science from the Bahauddin Zakaryia University (BZU) (Multan, Pakistan) in 2001. He received his MSc in Computer Science from BZU and Doctor of engineering degree (PhD) from the University of HUST at (Wuhan, China) in 2004 and 2011, respectively. He was working as a software manager under Business Craft company (Spanish), which provided IT solutions from 2004–2009 and also as a lecture from 2006–2008 in the same University, where he did masters. He is currently working as an Assistant Professor and the founding director of the Image Processing and Analysis Laboratory in the Department of Computer Science at the National Textile University in Faisalabad. His research interests include: image processing, medical image analysis, Genetic programming and pattern classification.

M. Emre Celebi received his BSc degree in computer engineering from Middle East Technical University (Ankara, Turkey) in 2002. He received his MSc and PhD degrees in computer science and engineering from the University of Texas at Arlington (Arlington, TX, USA) in 2003 and 2006, respectively. He is currently an associate professor in the Department of Computer Science at Louisiana State University, Shreveport, LA, USA). His research interests include medical image analysis, color image processing, content-based image retrieval, and open-source software development.

Irene Fondón received the M.S. degree in Telecommunication Engineering from the University of Seville, Spain, in 2004. She has completed her PhD degree in the year of 2011. She has been working as an associated professor since 2005 in the Signal Processing and Communications Department of the University of Seville. Her current research activities include works in the field of image processing and its medical applications. She is author of papers both in journals and international conferences.

Carmen Serrano received the MS degree in Telecommunication Engineering from the University of Seville, Spain, in 1996 and the PhD degree in January 2002. In 1996, she joined the Signal Processing and Communication Department at the same university, where she is currently an Associate Professor. Her research interests concern image processing and, in particular, color image segmentation, classification and compression, mainly with biomedical applications. She is author of numerous papers both in journals and international conferences. Dr. Guangzhi Ma is an Associate Professor at Huazhong University of Science and Technology. His research interests concern about Data mining and Knowledge Discovery. He is author of numerous papers both in journals and international conferences.

Guangzhi Ma is an Associate Professor at Huazhong University of Science and Technology. His research interests concern about Data mining and Knowledge Discovery. He is author of numerous papers both in journals and international conferences.