

# Melanoma Decision Support Using Lighting-Corrected Intuitive Feature Models

Robert Amelard, Jeffrey Glaister, Alexander Wong and David A. Clausi

**Abstract** Skin lesion analysis using standard camera images has received limited attention from the scientific community due to its technical complexity and scarcity of data. The images are privy to lighting variations caused by uneven source lighting, and unconstrained differences in resolution, scale, and equipment. In this chapter, we propose a framework that performs illumination correction and feature extraction on photographs of skin lesions acquired using standard consumer-grade cameras. We apply a multi-stage illumination correction algorithm and define a set of high-level intuitive features (HLIF) that quantifies the level of asymmetry and border irregularity about a lesion. This lighting-corrected intuitive feature model framework can be used to classify skin lesion diagnoses with high accuracy. The framework accurately corrects the illumination variations and achieves high and precise sensitivity (95 % confidence interval (CI), 73.1–73.5 %) and specificity (95 % CI, 72.0–72.4 %) using a linear support vector machine classifier with cross-validation trials. It exhibits higher test-retest reliability than the much larger state-of-the-art low-level feature set (95 % CI, 78.1–79.7 % sensitivity, 75.3–76.3 % specificity). Combining our framework with these low-level features attains sensitivity (95 % CI, 83.3–84.8 %) and specificity (95 % CI, 79.7–80.1 %), which is more accurate and reliable than classification using the low-level feature set.

**Keywords** Melanoma · Decision support system · Pigmented skin lesion · Feature extraction · Illumination correction · Standard camera

---

R. Amelard (✉) · J. Glaister · A. Wong · D. A. Clausi  
Department of Systems Design Engineering, University of Waterloo, Waterloo, Canada  
e-mail: ramelard@uwaterloo.ca

J. Glaister  
e-mail: jlglaist@uwaterloo.ca

A. Wong  
e-mail: a28wong@uwaterloo.ca

D. A. Clausi  
e-mail: dclausi@uwaterloo.ca

## Introduction

Melanoma is the most deadly form of skin cancer and is caused by the development of a malignant tumour of the melanocytes [42]. It was estimated that 76,250 men and women were diagnosed with melanoma in 2012 and 9,180 men and women died of it in the US alone [31]. One in five Americans are expected to be diagnosed with melanoma in their lifetime [54]. One of the most alarming facts about melanoma is that it is the most common form of cancer for young adults [30] and it is one of the few cancers where the incidence rate is increasing for men [34]. Fortunately, the five-year survival rate is 98 % if the lesion is extracted while the cancer is still confined to its primary location. However, the five-year survival rate decreases to 15 % if the cancer has spread to remote parts of the body [31].

Some clinical tools exist that can assist dermatologists diagnose skin lesions. For example, the Asymmetry, Border irregularity, Colour variation, and Diameter (ABCD) rubric serves as a guide for dermatologists to check skin lesions in a systematic manner [41, 50]. However, expert dermatologists using the ABCD rubric with a dermatoscope (a specialized tool that optically magnifies and enhances skin structures) reported a sensitivity of 76.0–87.7 % and a specificity of 61.0–77.8 % [5]. Furthermore, only 48 % of US fellows of the American Academy of Dermatology reported using a dermatoscope [25]. Recent developments include melanoma detection using standard camera images [1, 8, 9], but there is still much room for technological advances.

This chapter presents a systematic framework to analyse and assess the risk of melanoma using dermatological photographs taken with a standard consumer-grade camera. The framework consists of illumination preprocessing and feature extraction, and is validated using a simple malignancy classification scheme. The preprocessing step consists of a multi-stage illumination modeling algorithm. The proposed features that are extracted are high-level intuitive features (HLIF) describing lesion asymmetry and border irregularity. The segmentation is obtained using a manually-drawn ground-truth border and an existing classification algorithm is used. Automatic segmentation is not discussed here.

The first step in the proposed framework is a preprocessing step, where the image is corrected for illumination variation. This preprocessing step serves to improve performance of the subsequent steps. Illumination correction tries to remove illumination variation, such as shadows, so that healthy skin is a more consistent colour throughout the photograph. Since the lesion is typically a darker colour than the surrounding healthy skin, a segmentation algorithm may misclassify shadows as lesion. The illumination correction can improve the classification algorithm by standardizing features reliant on underlying pixel values. For example, features that rely on lesion colour are affected by shadows.

Once an image has been preprocessed, features are extracted and used to classify the image as “malignant” or “benign”. The quality of this feature set is very important in differentiating the two classes. Furthermore, it is generally preferable to project the image into a low-dimensional feature space, since high-dimensional

feature spaces are usually associated with larger computational cost and possibly various classification problems such as overfitting and the curse of dimensionality. In the proposed framework, we extract HLIFs that were designed specifically to model a human-observable characteristic. These features may take more effort to design than low-level features, but we show in the experimental results that classification accuracy improves dramatically by integrating HLIFs with low-level features.

The rest of the chapter is organized as follows. A literature review of existing methods and algorithms is presented in “Background”. The multi-stage illumination modeling algorithm is described in “Illumination Correction Preprocessing”. A set of high-level intuitive features for describing skin lesions is presented in “Feature Extraction”. Experimental results of the proposed unified framework are shown in “Results” and conclusions are drawn in “Conclusion”.

## Background

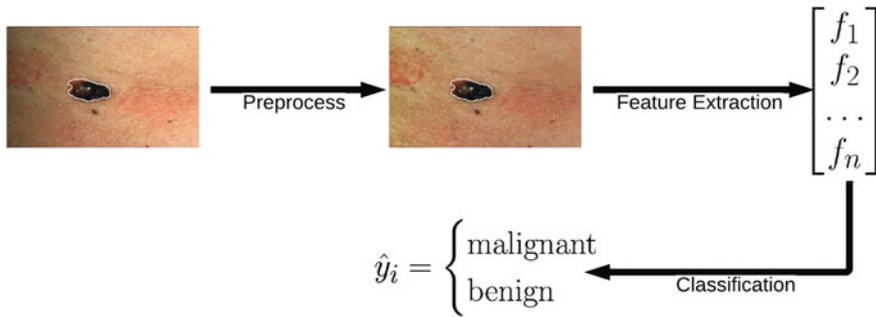
The problem of diagnosing cutaneous cancer has received more attention from the technical community in recent years. Unlike classical medical imaging modalities that produce an image using dedicated equipment (e.g., magnetic resonance imaging and X-ray), skin diagnosis is performed using visual information. This workflow poses more difficult computer vision problems as there is no standard modality with which data is captured.

The majority of the melanoma detection methods in the literature analyse images acquired using dermoscopy [21], also known as epiluminescence microscopy (ELM) or the unfortunate similarly-named “dermatology”. Dermoscopy is a non-invasive technique whereby the dermatologist uses a handheld dermatoscope to visually analyse the skin. Dermatoscopes optically magnify the area of interest, and most can elucidate sub-surface structures by applying a dermoscopic oil on the skin or, more recently, employing light cross-polarization. These produce images with standardized lighting conditions, and show sub-surface microstructures of the epidermis. However, it has been reported that only 48 % of American Academy of Dermatology fellows use dermatoscopes [25]. There is, therefore, a large demand for methods that analyse images taken without a dermatoscope.

Our framework’s workflow is summarized in Fig. 1. We present a literature review of preprocessing, feature extraction, and classification methods in the following sections, as it pertains to our framework.

### *Illumination Correction Preprocessing*

Most illumination correction algorithms are not specific to skin lesion photographs and can be applied to any scene. Histogram equalization adjusts the distribution of pixel intensities, minimizing illumination variation globally [47]. Other algo-



**Fig. 1** Workflow of the proposed melanoma decision support framework

gorithms correct for local illumination variation. These algorithms typically assume a multiplicative relationship between illumination and reflectance components. The estimated illumination component is estimated and used to find the reflectance component. The illumination component is assumed to be low-frequency, while the high-frequency detail is in the reflectance component. Using this assumption, there are many different algorithms that estimate illumination. One of the earliest is the Retinex algorithm, which uses a set of Gaussian filters of different sizes to remove detail and to estimate illumination [27, 36]. Morphological operators [49], bilateral filters [24], Monte Carlo sampling [53] and total variation [16] approaches have also been used to estimate illumination.

Other methods involve correction algorithms that are specific to images of skin lesions. Earlier algorithms enhance images taken with a dermatoscope to better separate lesion pixels from healthy skin. These algorithms include colour calibration [29] and normalization [33] to improve lesion classification or contrast enhancement [13, 46] to improve segmentation.

Recent work focuses on correcting photographs of skin lesions acquired using standard digital cameras to improve segmentation and classification. Work by Cavalcanti et al. [10] apply morphological operators to estimate the illumination component. The initial estimate of illumination is used to fit a parametric surface using the illumination intensities in the four corners of the photograph. The reflectance component is estimated using the parametric surface. Initial work on the correction algorithm outlined in this chapter was initially presented by Glaister et al. [28].

### *Feature Extraction and Classification*

Most existing feature sets have been designed to model the ABCD criteria using dermoscopic images. Lee and Claridge propose irregularity indices to quantify the amount of border irregularity [37]. Aribisala and Claridge propose another border irregularity metric based on conditional entropy [6]. Celebi et al. propose shape,

colour, and texture features with rationale, and using a filter feature selection method [12]. Colour features are primarily taken either in the RGB space (usually mean and standard deviation of the three channels), or a perceptually-uniform CIE colour space. Good overviews of existing features can be found in [35, 39].

Features designed to analyse dermoscopic images may not necessarily be suitable for the noisy unconstrained environment of standard camera images. Some work has been done to identify suitable features for standard camera images [1, 8, 9], however the focus of these methods has primarily been in the preprocessing and segmentation stages, resulting in large sets of low-level features. For example, Cavalcanti and Scharcanski [9] propose the same low-level feature set as Alcon et al. [1] with a few minor adjustments. Amelard et al. proposed the first set of high-level asymmetry and border irregularity features that were modeled assuming standard camera images [3, 4], which are used in this chapter.

Most of the methods use existing classification schemes, such as support vector machines (SVM), artificial neural networks (ANN), decision trees, and k-nearest neighbour (K-NN) [35]. Ballerini et al. designed a hierarchical classification system based on K-NN using texture and colour features to classify different types of non-melanoma skin lesions with 93 % malignant-versus-benign accuracy and 74 % inter-class accuracy [8]. Piatkowska et al. achieved 96 % classification accuracy using a multi-elitist particle swarm optimization method [44]. Thorough reviews of existing classification schemes can be found in [35, 39].

Some emphasis has been placed on constructing content-based image retrieval (CBIR) frameworks for recalling similar lesions. These methods rely on constructing a representative feature set that can be used to determine the similarity of two images. Ballerini et al. extracted basic colour and texture features such as colour mean, covariance, and texture co-occurrence matrix calculations, and used a weighted sum of Bhattacharyya distance and Euclidean distance to find visually similar lesions [7]. Celebi and Aslandogan incorporated a human response matrix based on psychovisual similarity experiments along with shape features to denote similarity [11]. Aldridge et al. designed CBIR software and experiments which showed with high statistical significance that diagnostic accuracy among laypeople and first-year dermatology students was drastically improved when using a CBIR system [2].

There has been some work done to extract melanin and hemoglobin information from skin images. The melanin and hemoglobin information can be very useful in trying to identify the stage and type of the lesion. All of the proposed methods rely on various physics-based models of the skin to characterize the reflectance under some assumptions about the absorption, reflectance, and transmission of the skin layers. Work primarily led by Claridge explores using multispectral images using spectrophotometric intracutaneous analysis to analyse melanin, hemoglobin, and collagen densities [19, 40]. Claridge used a physics-based forward parameter grid-search to determine the most feasible skin model parameters assuming standardized images [17]. Tsumura et al. used an independent component analysis (ICA) scheme to decompose the image into two independent channels, which they assumed are the melanin and hemoglobin channels [51]. D'Alessandro et al. used multispectral images obtained using a nevoscope and used a genetic algorithm to estimate melanin

and blood volume [20]. Madooei et al. used blind-source separation techniques using a proposed corrected log-chromaticity 2-D colour space to obtain melanin and hemoglobin information [38].

## Illumination Correction Preprocessing

The proposed framework first corrects for illumination variation using the multi-stage algorithm outlined in this section. The illumination correction algorithm uses three stages to estimate and correct for illumination variation. First, an initial non-parametric illumination model is estimated using a Monte Carlo sampling algorithm. Second, the final parametric illumination model is acquired using the initial model as a prior. Finally, the parametric model is applied to the reflectance map to correct for illumination variation. The three stages are outlined in this section.

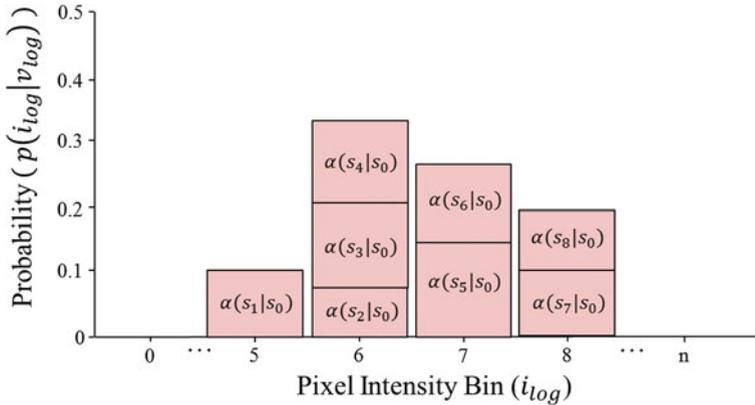
### *Initial Non-parametric Illumination Modeling*

The first stage involves estimating the initial non-parametric illumination model. This stage is required to estimate illumination robustly in the presence of artefacts, such as hair or prominent skin texture. Certain assumptions are made about the illumination in the dermatological images. The images are assumed to have been taken inside a doctor's office, in a controlled environment and beneath overhead lights. This means that the illumination model does not need to account for sudden changes in lighting conditions. Instead, the illumination will change gradually throughout an image. Illumination variation is assumed to be produced by white lights, so the correction algorithm does not need to correct colour variation. Finally, a multiplicative illumination-reflectance model is assumed [36]. In this model, the V (value) channel in the HSV (hue-saturation-value) colour space [48] is modeled as the entry-wise product of illumination  $i$  and reflectance  $r$  components. After applying the logarithmic operator, this relationship becomes additive (1).

$$\begin{aligned} v(x, y) &= i(x, y) \times r(x, y) \\ v_{log}(x, y) &= i_{log}(x, y) + r_{log}(x, y) \end{aligned} \quad (1)$$

To estimate the illumination map  $i$ , the problem can be formulated as Bayesian least squares (2), where  $p(i_{log}|v_{log})$  is the posterior distribution.

$$\begin{aligned} \hat{i}_{log} &= \underset{i_{log}}{\operatorname{argmin}} \left\{ E((i_{log} - \hat{i}_{log})^2) | v_{log} \right\} \\ &= \underset{i_{log}}{\operatorname{argmin}} \left\{ \int (i_{log} - \hat{i}_{log})^2 p(i_{log}|v_{log}) di_{log} \right\} \end{aligned} \quad (2)$$

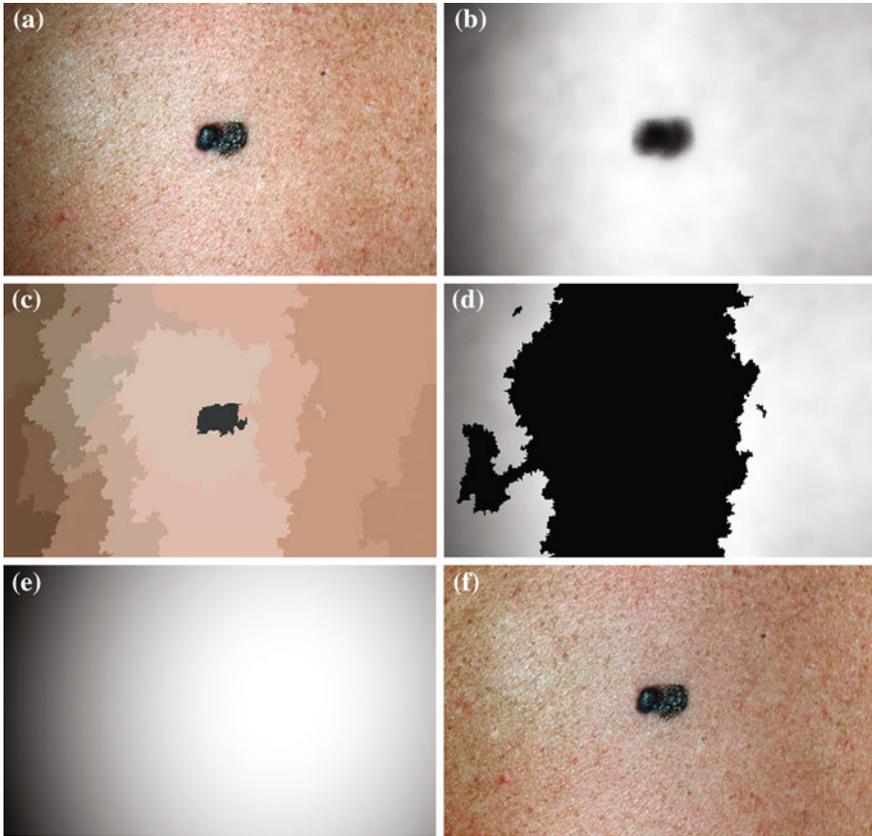


**Fig. 2** Sample posterior distribution  $\hat{p}(i_{log}|v_{log})$ , built from pixels accepted in the set  $\Omega$ . Each stacked element corresponds to a pixel  $s_k$  in  $\Omega$ , where the height is  $\alpha(s_k|s_0)$  and bin location is  $s_k$

To estimate the posterior distribution, a Monte Carlo posterior estimation algorithm is used [15]. A Monte Carlo estimation algorithm is used to avoid assuming a parametric model for the posterior distribution. In this Monte Carlo estimation strategy, candidate samples are drawn from a search space surrounding the pixel of interest  $s_0$  using a uniform instrumental distribution. An acceptance probability  $\alpha$  is computed based on the neighbourhoods around the candidate sample  $s_k$  and pixel of interest  $s_0$ . The Gaussian error statistic used in this implementation is shown in (3). The parameter  $\sigma$  controls the shape of the Gaussian function and is based on local variance and  $h_k$  and  $h_0$  represent the neighbourhoods around  $s_k$  and  $s_0$  respectively. The term  $\lambda$  in the denominator normalizes the acceptance probability, such that  $\alpha(s_k|s_0) = 1$  if the neighbourhoods around  $s_k$  and  $s_0$  are identical. The elements in the neighbourhoods are assumed to be independent, so the acceptance probability is the product of the probabilities from each site  $j$ .

$$\alpha(s_k|s_0) = \prod_j \frac{\frac{1}{2\pi\sigma} \exp\left[-\frac{(h_k[j]-h_0[j])^2}{2\sigma^2}\right]}{\lambda} \tag{3}$$

The candidate sample is accepted with a probability of  $\alpha$  into the set  $\Omega$  for estimating  $p(i_{log}|v_{log})$ . The selection and acceptance process is repeated until a desired number of samples were found in the search space. The posterior distribution is estimated as a weighted histogram, using  $\alpha$  as the weights associated with each element. A sample histogram is shown in Fig. 2. The estimate of the log-transformed illumination map  $\hat{i}_{log}$  is calculated using (2), as outlined in [26]. The initial illumination estimate  $\hat{i}$  is acquired by taking the exponential of  $\hat{i}_{log}$ . An example of an image with visible illumination variance is shown in Fig. 3a and the associated non-parametric illumination model is shown in Fig. 3b.



**Fig. 3** Methodology to estimate illumination map: **a** original image of a skin lesion, where the *top* edge is noticeably darker than the *bottom* edge; **b** illumination map determined via non-parametric modeling using Monte Carlo sampling; **c** segmentation map found using Statistical Region Merging; **d** regions included in the subset of skin pixels, where pixels *black* in colour are not classified as normal skin; **e** new illumination map determined by using (**d**) as a prior to the quadratic surface model; **f** resulting corrected image using the multi-stage illumination correction algorithm

### *Final Parametric Illumination Modeling*

The initial non-parametric illumination model results in an estimate of the illumination variation in healthy skin, but does not properly model illumination near the lesion. Instead, the initial model identifies the lesion as a shadow. Using the initial model to correct the image would result in a significant bright spot around the lesion, which is obviously undesirable. To better model the illumination, a second stage is added, which results in a parametric model of illumination that uses the initial illumination pixel values. The parametric model can adequately estimate the illumination variation because illumination is assumed to change slowly throughout the image.

The subset of pixels that are used to fit the parametric surface correspond to healthy skin in the original image.

To find the subset of healthy skin pixels, the original image is segmented into many regions. The segmentation algorithm used in this implementation is Statistical Region Merging [43]. The resulting segmented image is shown in Fig. 3c, where each region is represented as a single colour. Any regions that touched  $20 \times 20$  pixel regions in the four corners of the image are considered part of the “healthy skin” class. While this method does not yield a perfect segmentation of the healthy skin and lesion classes, only an estimate of healthy skin pixels is required for fitting the parametric model. The regions that are considered “healthy skin” are shown in Fig. 3d.

The final illumination model  $\hat{i}'$  is estimated as a parametric surface (4) with coefficients  $c_1$  to  $c_6$ , which is fit to the initial illumination values  $\hat{i}$  corresponding to pixels in the “healthy skin” subset  $S$  using maximum likelihood estimation (5). The final parametric illumination model is shown in Fig. 3e.

$$i'(x, y) = c_1x^2 + c_2xy + c_3y^2 + c_4x + c_5y + c_6 \quad (4)$$

$$\hat{i}' = \underset{\hat{i}'}{\operatorname{argmax}} \prod_{(x,y) \in S} P(\hat{i}(x, y) | \hat{i}'(x, y)) \quad (5)$$

$$\text{where } P(\hat{i}(x, y) | \hat{i}'(x, y)) \text{ i.i.d. } \mathcal{N}(\hat{i}'(x, y), \sigma^2)$$

## Reflectance Map Estimation

The reflectance map is calculated by dividing the V channel  $v$  from the original image in the HSV colour space by  $\hat{i}'$ . The reflectance map  $\hat{r}$  replaces the original V channel and is combined with the original hue (H) and saturation (S) channels. The resulting image is corrected for illumination. An example of a corrected image is shown in Fig. 3f.

## Feature Extraction

Once the image has been preprocessed, descriptive features are extracted to describe the lesion as a vector of real numbers. One of the most prominent clinical methods for diagnosing a skin lesion is using the ABCD rubric [41, 50], where the dermatologist looks for signs of asymmetry, border irregularity, colour variations, and diameter. However, this is done in a very subjective manner, and results in discrete categorical values. For example, the score assigned to a lesion’s asymmetry is determined by identifying whether the lesion is asymmetric across two orthogonal axes chosen by

the dermatologist, resulting in a score  $A \in \{0, 1, 2\}$  [50]. This type of subjective visual analysis leads to large inter-observer bias as well as some intra-observer bias [5]. We aim to create continuous high-level intuitive features (HLIFs) that represent objective calculations modeled on a human's interpretation of the characteristic.

## *High-Level Intuitive Features*

A “High-Level Intuitive Feature” (HLIF) is defined as a feature calculation that has been designed to model a human-observable phenomenon (e.g., amount of asymmetry about a shape), and whose score can be qualitatively intuited. As discussed in “Background”, most skin lesion features are low-level features. That is, they are recycled mathematical calculations that were not designed for the specific purpose of analysing a characteristic of the lesion shape.

Although designing HLIFs is more time-consuming than amalgamating a set of low-level features, we show in “Results” that the discriminative ability of a small set of HLIFs is comparable to a large set of low-level features. Since the HLIF set is small, the amount of required computation for classification decreases, and the risk of overfitting a classifier in a highly dimensional space is reduced, especially with small data sets.

In the next two sections, we describe nine HLIFs to evaluate the asymmetry and border irregularity of a segmented skin lesion. These HLIFs are designed to model characteristics that dermatologists identify. This work is described in detail in [3, 4], so we limit our analysis here to a general overview of the features.

## *Asymmetry Features*

Dermatologists try to identify asymmetry with respect to shape or colour to indicate malignancy. These visual cues result due to the unconstrained metastasis of melanocytes in the skin.

### **HLIF for Colour Asymmetry**

Asymmetry with respect to colour can be quantified by separating the lesion along an axis passing through the centre of mass (centroid) such that it represents the maximal amount of colour asymmetry. This “maximal” axis was found iteratively. First, we calculate the major axis of the lesion. The major axis is that which passes through the centroid and describes the maximal variance of the shape. The lesion image was then converted to the Hue-Saturation-Value (HSV) space so we can use the illumination-invariant hue measure to analyse the colours. The normalized hue histograms of both sides of the axis were smoothed using a Gaussian filter for robustness and were then

compared to generate the following HLIF value:

$$f_1^A = \max_{\theta} \left\{ \frac{1}{2} \sum_{i=1}^{nbins} |H_1^\theta(i) - H_2^\theta(i)| \right\} \quad (6)$$

where  $H_1^\theta$  and  $H_2^\theta$  are the normalized smoothed hue histograms according to the separation axis defined by rotating the major axis by  $\theta$ , and  $nbins$  is the number of discretized histogram bins used for binning hue values (we used 256 bins). Noticing that  $f_1^A \in [0, 1]$ ,  $f_1^A$  can be intuited as an asymmetry score ranging from 0 (completely symmetric) to 1 (completely asymmetric).

Figure 4 depicts an example of this calculation. The lesion has a dark blotch on one side of it, rendering it asymmetric with respect to colour, which is reflected in the calculated value of  $f_1^A = 0.3866$ .

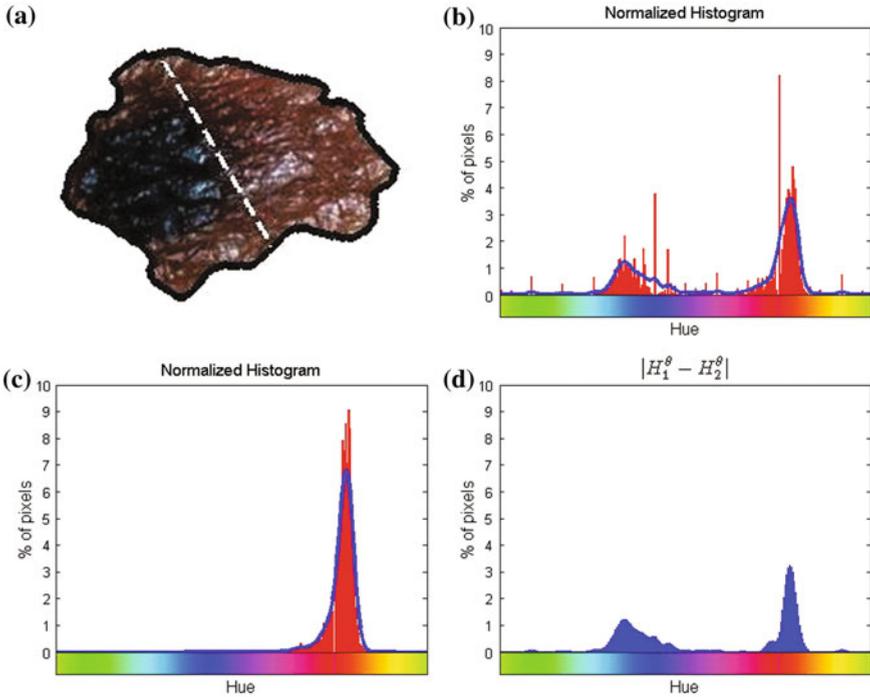
### HLIF for Structural Asymmetry

Making the observation that symmetry can usually not be found in highly irregular shapes, we express the segmented lesion as a structure and analyse its simplicity. Fourier descriptors apply the Fourier series decomposition theory to decomposing some arbitrary shape into low-frequency and high-frequency components. In particular, the points on the shape border are mapped to 1D complex number via  $F : (x, y) \mapsto x + iy$ , where  $i$  is the complex number. The Fourier transform is performed on this set of complex numbers. We can then compare a low-frequency (“simple”) reconstruction with the original shape to determine the amount of asymmetry due to shape irregularity.

First, since we must use the discretized Fourier transform, the lesion border was uniformly sampled at a fixed rate to ensure the same decomposition in the frequency domain across all images. The Fast Fourier Transform (FFT) was then applied, and only the two lowest frequency components were preserved. These two frequencies represent the zero-frequency mean and the minimum amount of information needed to reconstruct a representative border. The inverse FFT was applied on these two frequencies to reconstruct a low-frequency representation of the lesion structure. Comparing this shape with the original is ill-advised, as this would yield a metric that is more suitable for border irregularity. Instead, we applied the same procedure to reconstruct a structure using  $k$  frequencies, where  $k$  is some small number that reconstructs the general shape of the lesion. This shape is compared to the original shape to generate the following HLIF calculation:

$$f_2^A = \frac{area(S_k \oplus S_2)}{area(S_k \cup S_2)} \quad (7)$$

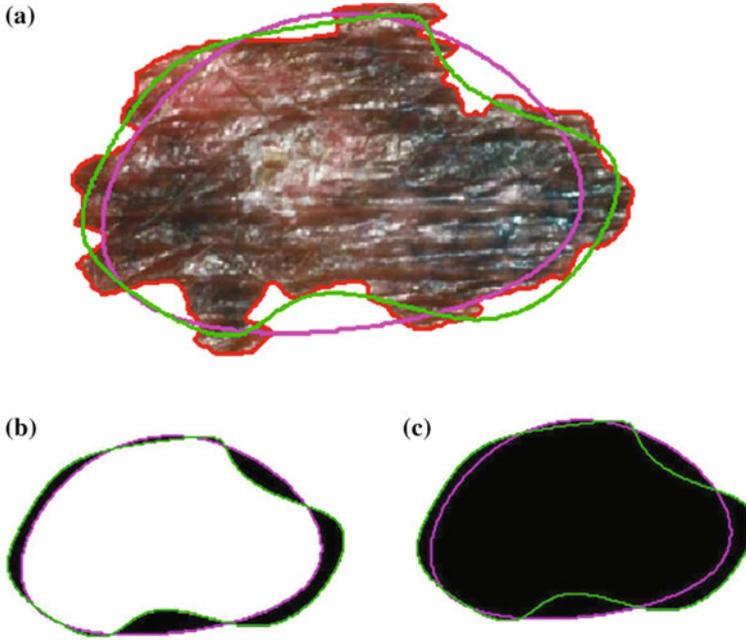
where  $S_k$  and  $S_2$  are the  $k$ -frequency and 2-frequency reconstructions of the original lesion shape. This feature value can be intuited as a score representing the general



**Fig. 4** Example of the design of  $f_1^A$  by comparing the normalized hue histograms of both sides of the separation axis. The *red bars* represent the original binned histogram of hue values, and the *blue line* represents these histograms smoothed by a Gaussian function ( $\sigma = 2$  bins), which allows us to compare hue histograms robustly. In this example,  $f_1^A = 0.3866$ , representing a lesion with asymmetric colour distributions. **a** Image separated by the axis that produces maximal hue difference. **b** Normalized hue histogram of the *left* side of the lesion. Note the prominence of *blue* pixels along with the *red* ones in the histogram due to the *dark blotch* in the image. **c** Normalized hue histogram of the *right* side of the lesion. Note the prominence of *red* pixels in the histogram in correspondence with the image. **d** Absolute difference of the two hue histograms. The amount of *blue* and lack of *red* pixels in the first histogram are reflected by the two “humps” in the difference histogram

structural variations. We found empirically that linearly sampling the border with 1000 points and setting  $k = 5$  yielded good results.

Figure 5 depicts an example of this calculation. Notice how the lesion has a very abnormal shape that does not seem to contain any symmetry. The logical XOR (Fig. 5b) captures this structural variation, and the calculation represents the dark area with respect to the union in Fig. 5c.



**Fig. 5** Example of the design of  $f_2^A$  by comparing a baseline reconstruction of the lesion (*magenta*) with a low-frequency reconstruction (*green*) that incorporates the structural variability, if any exists. In this example,  $f_2^A = 0.1609$ , representing a structurally asymmetric lesion. **a** Original lesion, and reconstruction using 2 (*pink*) and 5 (*green*) frequency components. **b** Logical XOR of the two reconstructions (*dark area*). **c** Union of the two reconstructions (*dark area*)

### HLIF for Asymmetric Areas

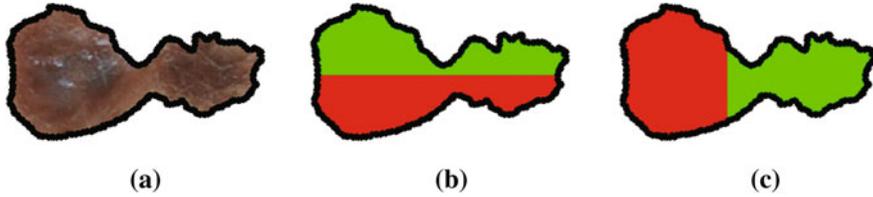
Cavalcanti and Scharcanski propose the following four HLIFs as a measure of area asymmetry [9]:

- $f_8 : (B_1 - B_2)/A$  with respect to  $L_1$
- $f_9 : (B_1 - B_2)/A$  with respect to  $L_2$
- $f_{10} : B_1/B_2$  with respect to  $L_1$
- $f_{11} : B_1/B_2$  with respect to  $L_2$

where  $L_1, L_2$  are the major and minor axes of the lesion, and  $B_1, B_2$  are the areas of each side of  $L_1$  or  $L_2$ . No explicit constraint exists on the relative sizes of  $B_1$  vs  $B_2$ . Thus, for very asymmetric shapes,  $\{f_8, f_9\}$  can be either positive or negative, and  $\{f_{10}, f_{11}\}$  can be large or small. For clarity, we add the following constraint to the features:

$$f_3^A = (A_1 - A_2)/A \text{ with respect to } L_1, \quad (8)$$

$$f_4^A = (A_1 - A_2)/A \text{ with respect to } L_2, \quad (9)$$



**Fig. 6** Example of the design of  $\{f_3^A, f_4^A, f_5^A, f_6^A\}$  by comparing the areas to each side of the lesion's major and minor axes. Notice that the red area appears much larger than the *green* area when separated by the minor axis. In this example,  $f_3^A = 0.0052$ ,  $f_4^A = 0.1560$ ,  $f_5^A = 0.0105$ ,  $f_6^A = 0.3698$ , representing a structurally asymmetric lesion about the minor axis. **a** Original lesion. **b** Major axis ( $L_1$ ). **c** Minor axis ( $L_2$ )

$$f_5^A = (A_1 - A_2)/A_2 \text{ with respect to } L_1, \quad (10)$$

$$f_6^A = (A_1 - A_2)/A_2 \text{ with respect to } L_2 \quad (11)$$

such that:

$$A_1 = \max \{B_1, B_2\},$$

$$A_2 = \min \{B_1, B_2\}$$

where  $B_1, B_2$  are as before. This way,  $\{f_3^A, f_4^A\}$  represent the positive difference in areas with respect to the total area, and  $\{f_5^A, f_6^A\}$  represent the positive difference in areas with respect to the smaller area.

Figure 6 depicts an example of this calculation. When this lesion is separated by the minor axis which passes through the centroid, the red area is much larger than the green area, indicating a large amount of structural asymmetry, which is reflected in the scores  $f_4^A = 0.1560$  and  $f_6^A = 0.3698$ .

## ***Border Irregularity Features***

Clinically, border irregularities are usually defined by spiky non-uniform pigmentation. Benign lesions usually have relatively smooth borders that form an oval-like shape. We therefore want to design features to capture this “spiky” nature.

### **HLIF for Fine Irregularities**

One indicator of the malignancy of a lesion is the amount of small spiky deviations from a theoretical smooth border. We therefore wish to quantify the degree to which a lesion border contains these fine irregularities. To do so, it seems natural to compare the border to a “smoothed” version of itself, thus preserving the overall structure and

capturing the spikes. This can be accomplished using morphological operations. In particular, given an arbitrary shape and a structuring element, morphological closing tends to smooth over abrupt *exterior* peaks. Conversely, morphological opening tends to smooth over abrupt *internal* peaks. If the morphological operation yields any changes to the original shape due to abrupt structural elements, the modified shapes will have a different area than the original. These areas are compared to generate the following HLIF calculation:

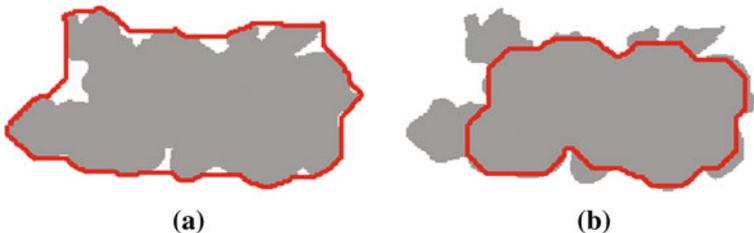
$$f_1^B = \frac{A_{closed} - A_{lesion}}{A_{lesion}} + \frac{A_{lesion} - A_{opened}}{A_{lesion}} \quad (12)$$

where  $A_{lesion}$  is the area of the original lesion shape, and  $A_{closed}$  and  $A_{opened}$  are the areas of the modified shape under the specific morphological operations. The sum of the two normalized areas indicates the level of fine irregularities in a lesion's border.

Figure 7 depicts an example of this calculation. The red border denotes the resulting area from the morphological operation. notice how morphological closing produces a larger area that fills in the gaps from extreme irregularities of the border, and morphological opening produces a smaller area that crops these extreme irregularities.

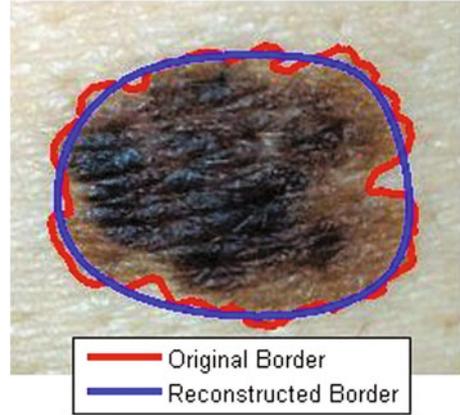
### HLIF for Coarse Irregularities

Another indicator of the malignancy of the lesion is the amount of structural deviation from a standard circular shape. This shape is influenced by the non-linear spatial reproduction of melanocytes in the skin. We can use Fourier descriptors again to characterize these coarse irregularities. In particular, we can compare a low-frequency reconstruction of the lesion shape to the original lesion shape. These two shapes will differ significantly if the lesion has a varying border. We capture this information by comparing the perimeters of the two shapes to generate the following HLIF



**Fig. 7** Example of the design of  $f_1^B$  by comparing the areas resulting from morphological closing and opening. The *red* borders denote the resulting area from the respective morphological operation. Notice how the morphological closing fills in the *white* areas that are present due to the external spikes of the border, and morphological opening crops those spikes. In this example,  $f_1^B = 0.3063$  representing a border with abrupt spikes. **a** Morphological closing. **b** Morphological opening

**Fig. 8** Example of the design of  $f_2^B$  by comparing the perimeters of the original (red) and low-frequency reconstructed (blue) borders. Notice how the reconstructed border follows the general shape of the lesion, but does not take into account the large amounts of irregularity. In this example,  $f_2^B = 0.24951$  representing a border with coarse irregularities



calculation:

$$f_2^B = \frac{|P_{lesion} - P_{low}|}{P_{lesion}} \quad (13)$$

where  $P_{lesion}$  and  $P_{low}$  are the perimeters of the original and low-frequency reconstruction of the lesion.

Figure 8 depicts an example of this calculation. The reconstructed (blue) border follows the general shape of the original lesion border (red), however it does not account for the coarse irregularities present in the actual border.

### HLIF for Comparing Against Average Malignant Lesion

Over time a doctor will start to recognize malignant lesions based on the visual similarity to previously diagnosed cases. We can model this “learning” procedure by comparing a new case against the average malignant lesion border found across the training data. To perform this comparison, we must be able to compare the Fourier descriptors in a translation-, scale-, rotation-invariant manner. Given a set of Fourier coefficients  $C = \{C_0, C_1, \dots, C_{N-1}\}$ , Fourier descriptor normalization is performed using the following three steps [52]:

1. *Translation Invariance*: set the first Fourier component (i.e., the DC component) to 0 ( $C_0^* = 0$ ).
2. *Scale Invariance*: divide each  $k - 1$  Fourier coefficient by the complex magnitude of the second Fourier coefficient ( $C_k^* = \frac{C_k}{|C_1|}$ ).
3. *Rotation/Point-Order Invariance*: consider only the real-valued complex magnitude of each Fourier component ( $C_k^* = |C_k^*|$ ).

For each image in the training set, we sampled the border at a fixed rate (1,000 points on the border using linear interpolation produced good results), applied this normalization process, and computed the average normalized Fourier coefficients

across these images to obtain  $\bar{C}^* = \{\bar{C}_0^*, \bar{C}_1^*, \dots, \bar{C}_{N-1}^*\}$ , where  $C_k^* = |C_k|/|C_1|$ . This ‘‘average malignant’’ frequency representation is compared to a new case using the sum of squared differences to generate the following HLIF calculation:

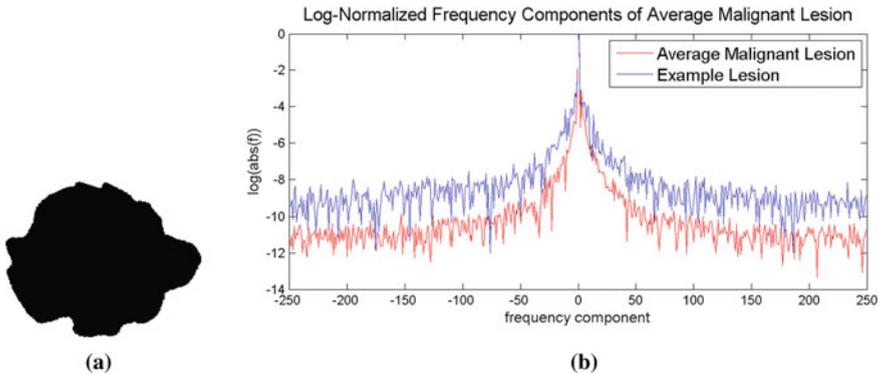
$$f_3^B = \sum_{u=0}^{N-1} (|C_u^*| - |\bar{C}_u^*|)^2 \tag{14}$$

where  $C^* = \{C_0^*, C_1^*, \dots, C_{N-1}^*\}$  is the normalized set of Fourier coefficients for a new lesion image. Intuitively, the feature value corresponds to the amount of deviation from the ‘‘typical’’ malignant lesion.

Figure 9 depicts an example of this calculation. Note that the frequency components of the average malignant lesion (computed from the training data) and the extracted frequency components of the new lesion’s border are quite similar, even in log space. This represents a ‘‘typical’’ malignant lesion in terms of structure.

## Results

We validated the proposed framework in two phases using publicly accessible skin lesion images. First, we validated the resulting photographs after applying the initial illumination correction preprocessing algorithm. The photographs before and after correction were compared visually. We then tested the complete framework by extracting the HLIFs and low-level features from the corrected photographs and



**Fig. 9** Example of the design of  $f_3^B$  by comparing the frequency components of a lesion to the computed average malignant frequency components via sum of squared differences. This will give a low value for cases that look like a standard malignant lesion, based on training data. In this example,  $f_3^B = 0.0087$ , representing a typical malignant case. **a** Example melanoma lesion shape. **b** Log-frequency components of the average malignant lesion in the training data (red) and the example lesion on the left (blue). Log-space was used for visualization only

performing classification using a linear SVM model. This allowed us to validate the separability of the proposed framework's feature space.

## ***Data and Setup***

Our data set comprises 206 standard camera images which were taken in completely unconstrained and varying environments. These images were obtained from the Dermatology Information System [22] (43 malignant melanomas, 26 nevi) and DermQuest [23] (76 malignant melanomas, 61 nevi). Each image contains a single lesion of interest, which was manually segmented by the authors. In order to ensure rotation- and scale-invariance, each image was automatically rotated and scaled such that the major axis of the lesion lies parallel to the horizontal axis, and the lesion bounding box fit within a  $200 \times 200$  rectangle prior to feature extraction. The test infrastructure was implemented in MATLAB.

## ***Illumination Correction Results***

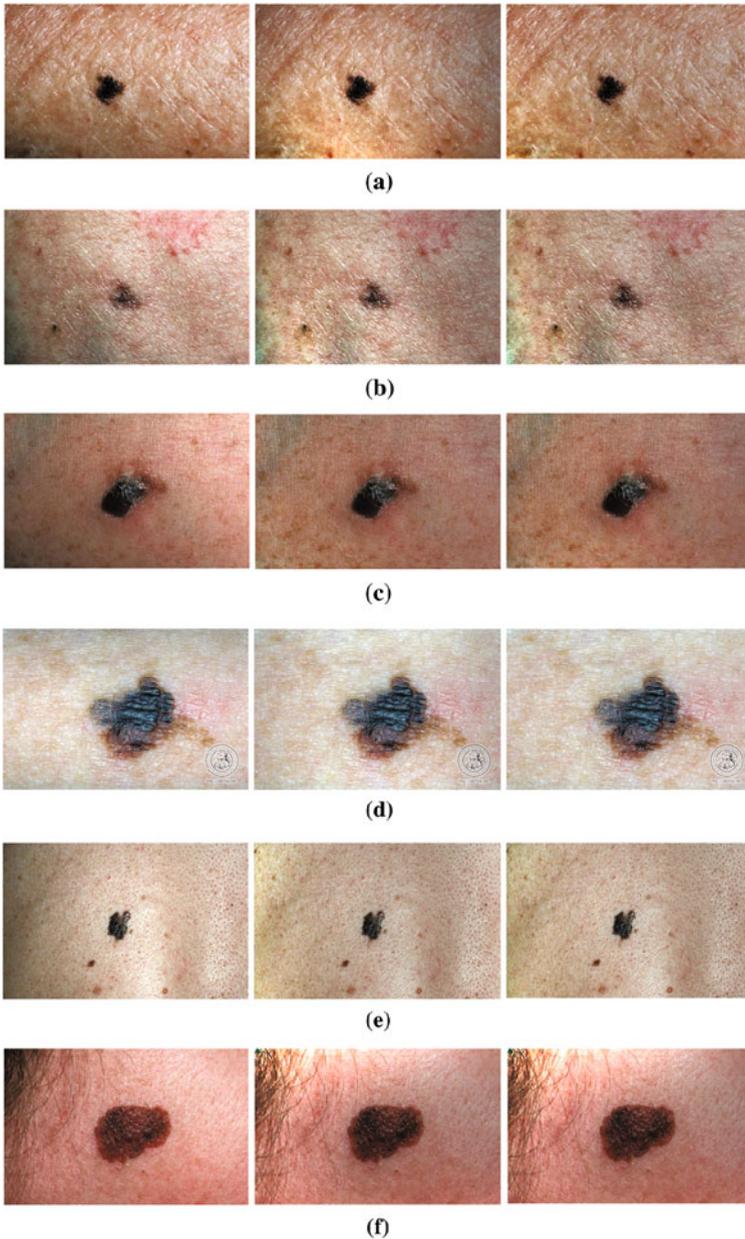
After applying the illumination correction algorithm, the images were compared visually to the original images. The images were also compared visually to the Cavalcanti et al. illumination correction results. Figure 10 shows a set of images for comparison. To allow for a fair comparison between corrected and uncorrected images, the dynamic range of the V channel intensities were normalized.

In Fig. 10a–e, there is a visual improvement between the corrected and uncorrected images. Furthermore, in Fig. 10a, the framework's correction algorithm performs better than the Cavalcanti et al. algorithm for correcting illumination variation. This is because the correction algorithm uses a much larger subset of pixels in the parametric illumination model. The Cavalcanti et al. algorithm only uses pixels in the four corners, whereas the framework's correction algorithm uses any regions that touch the corners.

Figure 10f is an example of poor correction of illumination variation by both algorithms. This occurs when the illumination is complicated and cannot be modelled using a quadratic surface. For example, in Fig. 10f, the quadratic surface model is a false assumption due to the large patch of dark hair in the top of the photograph. As a result, the top left corner of the photograph becomes too bright.

## ***Feature Extraction and Framework Classification Results***

Upon applying the illumination correction algorithm, we extracted the asymmetry and border irregularity HLIFs (see "Feature Extraction") as well as Cavalcanti and



**Fig. 10** Examples of dermatological photographs corrected for illumination variation. First column is the original photograph. Second column is the illumination correction results using the algorithm outlined in [10]. Third column is the illumination correction result using the proposed framework’s multi-stage algorithm outlined in “Illumination Correction Preprocessing”. In (a–e), the illumination variation is removed or reduced in the corrected image, while in (f), the illumination variation is not reduced. Examples of dermatological photographs corrected for illumination variation, continued

Scharcanski’s low-level standard camera feature set for comparison [9]. For convenience, we use the following naming convention in the discussion:

- $F_C$ : Cavalcanti and Scharcanski’s low-level feature set.
- $F_{HLIF}$ : set of HLIFs from “Feature Extraction”.
- $F_T$ : Total combined feature set containing both of the above feature sets ( $F_T = F_C \cup F_{HLIF}$ ). Note that four features from  $F_C$  are replaced by the HLIFs in “Feature Extraction”.

Due to the small data set, leave-one-out cross-validation (LOO CV) trials were used to calculate the success metrics.

## Classification

We used a linear soft-margin SVM model [18] to calculate sensitivity, specificity, and accuracy. We used the LIBSVM implementation of SVM [14]. Linear SVM was chosen to emphasize the separability of the proposed framework’s feature space due to illumination correction and descriptive features, rather than emphasizing the performance of an advanced classifier.

## Parameter Optimization

For each feature set we found the optimal soft-margin constraint ( $c$ ) and benign class weight ( $w$ ) using a grid-search algorithm. In particular, for each set of parameters ( $c$ ,  $w$ ), we calculated the F-score [45] over 100 CV trials using randomized 80%/20% splits of the data for training and testing, respectively. Recall the standard F-score formula:

$$F_\beta = (1 + \beta^2) \frac{\textit{precision} \times \textit{recall}}{\beta^2 \times \textit{precision} + \textit{recall}} \quad (15)$$

where

$$\textit{precision} = \frac{\textit{TP}}{\textit{TP} + \textit{FP}} \quad (16)$$

$$\textit{recall} = \frac{\textit{TP}}{\textit{TP} + \textit{FN}} \quad (17)$$

$$\textit{accuracy} = \frac{\textit{TP} + \textit{TN}}{\textit{TP} + \textit{FP} + \textit{TN} + \textit{FN}} \quad (18)$$

where TP, TN, FP, FN are the number of true positive (i.e., malignant), true negative (i.e., benign), false positive, and false negative cases from classification. The F-score measure is the weighted harmonic mean of the precision and recall. For stability of the results due to the relatively restricted data set, we use  $\beta = 1$  to calculate the harmonic mean of precision and recall.

The training and testing data points were chosen at random from the entire data set for each iteration. In accordance with the authors’ suggestions [32], we varied the value of  $c \in \{2^i | i = -5, \dots, 11\}$  and  $w \in \{2^i | i = -3, \dots, 3\}$ . Let  $(c^*, w^*)$  be the parameter values that exhibit the highest  $F_1$  score. Upon determining  $(c^*, w^*)$ , we sub-sampled the parameter space  $\{(\frac{1}{2}c^*, \frac{1}{2}w^*), (2c^*, 2w^*)\}$  at  $c_i = 2^{0.15}c_{i-1}$  and  $w_i = 2^{0.5}w_{i-1}$ . The optimal parameter values found according to  $F_1$  in this sub-sampling over 100 independent cross-validation trials with 80%/20% train/test split were used in the classification.

**Discussion**

Table 1 shows that classification using the nine HLIFs exhibits higher test-retest reliability as compared to classification using the 52 low-level features. In particular, although mean sensitivity (95% confidence interval (CI), 73.1–73.5%) and specificity (95% CI, 72.0–72.4%) of the small HLIF set are slightly lower than the sensitivity (95% CI, 78.1–79.7%) and specificity (95% CI, 75.3–76.3%) of the large low-level set, classification using the HLIFs is much more reproducible as shown through the more narrow confidence interval. This is a powerful observation, since the HLIF set is only one-fifth the size of the low-level feature set. Since the features are in a lower dimensional space, the cost of computation and curse of dimensionality are not as pervasive as with the large low-level feature set.

Moreover, combining the HLIFs with the low-level features yields by far the best results (95% CI, 83.3–84.8% sensitivity, 79.7–80.1% specificity). Adding only the nine HLIFs to the low-level feature set increases the number of features to 59, but yields non-trivial sensitivity and specificity improvements while also increasing reproducibility (i.e., decreasing standard deviation). This can be attributed to the HLIFs’ ability to replicate human-observable phenomena in the data, whereas using many low-level features to model a high-level characteristic introduces a lot of variability in the measure, since the features were not designed specifically for the intended purpose of diagnosing skin cancer.

**Statistical Significance**

We wish to investigate each feature’s ability to uniquely separate a malignant lesion from a benign lesion. The  $t$ -test indicates whether there is enough statistical signifi-

**Table 1** Comparing classification results of different feature sets over 100 cross-validation trials see “Results” for feature set descriptions

Feature set	# features	Sensitivity (%)		Specificity (%)		Accuracy (%)	
		$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$
$F_C$	52	78.89	4.21	75.80	2.40	76.51	1.08
$F_{HLIF}$	9	73.32	<b>0.92</b>	72.21	0.99	72.52	<b>0.49</b>
$F_T$	59	<b>84.04</b>	3.67	<b>79.91</b>	<b>0.98</b>	<b>81.26</b>	1.31

$\mu$  mean,  $\sigma$  standard deviation

cance to reject a null hypothesis about a population using data from a sample of that population. In a two-sample  $t$ -test, the null hypothesis states that the means of two normally-distributed populations are equal ( $H_0 : \Delta\mu = \mu_1 - \mu_2 = 0$ ). A  $p$ -value can be calculated using  $\Delta\mu$ , which indicates probability that we can observe a test statistic at least as extreme as the one observed assuming  $H_0$  is indeed valid.

Representing the populations as a particular feature’s scores across malignant (population 1) or benign (population 2) cases, we seek to reject this null hypothesis with a low  $p$ -value, thus showing that a particular feature’s scores are significantly different between malignant and benign classes (i.e., the feature separates malignant and benign cases well). We therefore assume that the population responses of the malignant and benign cases follow a normal distribution. We use Welch’s  $t$ -test, which assumes populations with different variances.

Table 2 summarizes the  $p$ -value scores for each HLIF using Welch’s  $t$ -test. That is, for each feature, we set the null hypothesis  $H_0 : \Delta\mu = \mu_m - \mu_b = 0$ , where  $\mu_m$  and  $\mu_b$  are the mean response values for malignant and benign feature scores. The  $t$ -statistic is calculated as follows:

$$t = \frac{\mu_m - \mu_b}{\sqrt{\frac{s_m^2}{N_m} + \frac{s_b^2}{N_b}}}, \tag{19}$$

where  $s_m, s_b$  are the sample standard deviations of malignant and benign feature scores, and  $N_m, N_b$  are the number of malignant and benign cases in the data set. The associated  $p$ -value is the area under the normal curve to the right of the calculated  $t$ -score.

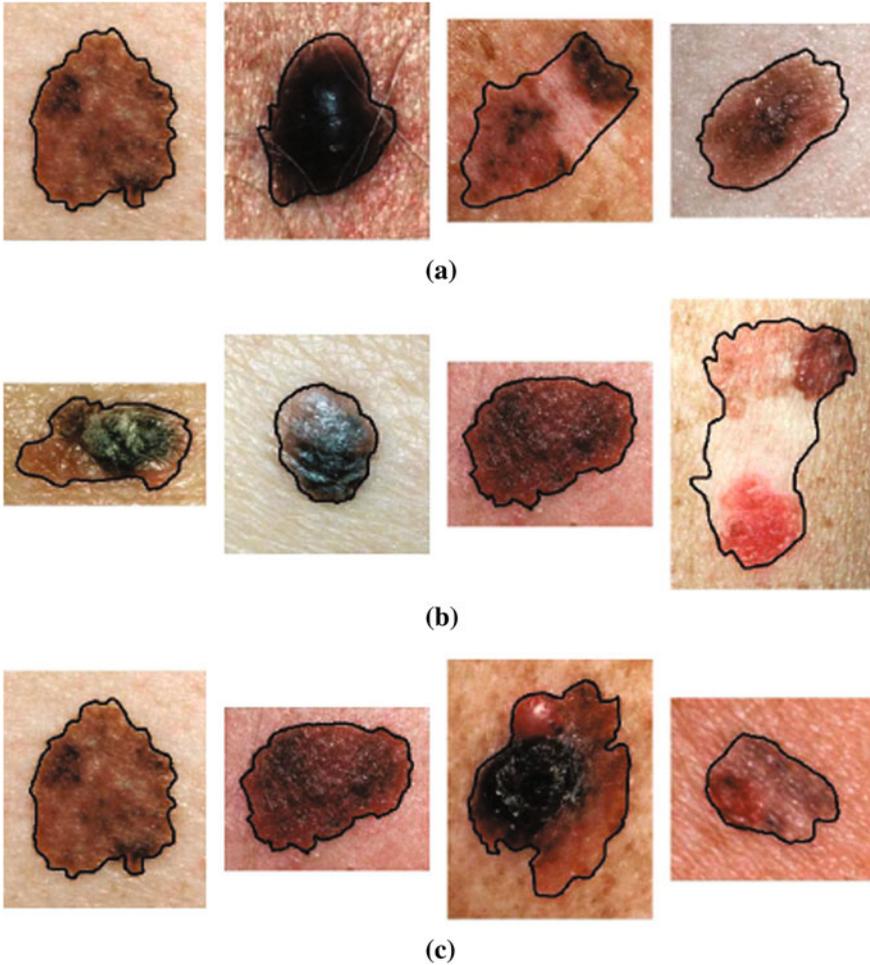
Most of the  $p$ -values are relatively low, indicating a high ability to separate malignant from benign. In particular,  $\{f_1^A, f_2^A, f_1^B, f_2^B\}$  seem to be very good predictors of malignancy as indicated by their very low  $p$ -values. Recall that  $f_3^A$  and  $f_5^A$  are the relative area differences with respect to the major axis. So, although these  $p$ -values are high, it is probably worth keeping them in the feature set because they complement the minor axis features  $f_4^A$  and  $f_6^A$ .

### Sources of Error

Figure 11 provides some examples of false negative cases (i.e., misidentified malignant cases) using the framework classification results. Using the  $F_{HLIF}$  feature set (Fig. 11a), we see lesions that have a fairly smooth and regular border, and most are fairly symmetric as well. It appears as though the colour patterns would be the

**Table 2** Performing Welch’s two-sample  $t$ -test on the set of HLIFs

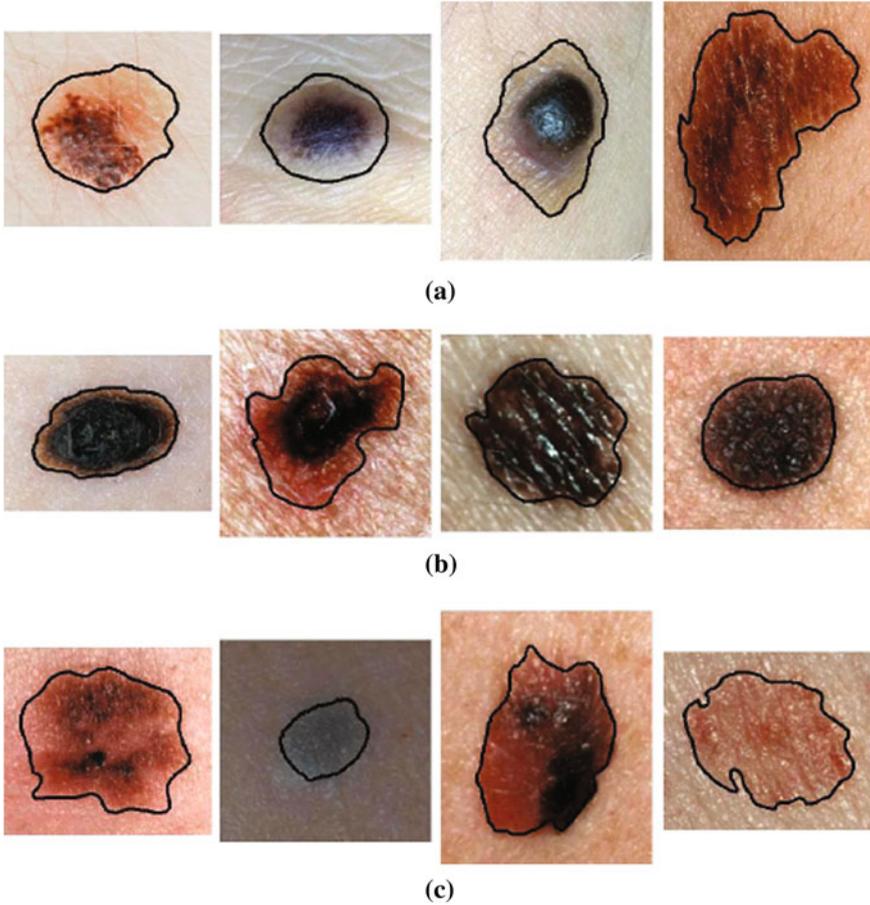
HLIF	$f_1^A$	$f_2^A$	$f_3^A$	$f_4^A$	$f_5^A$	$f_6^A$	$f_1^B$	$f_2^B$	$f_3^B$
<b><math>p</math>-value</b>	< 0.0001	0.0340	0.7208	0.1215	0.6822	0.1372	0.0006	< 0.0001	0.3308



**Fig. 11** Classification false negative examples using each feature set. **a** False negatives using  $F_{HLIF}$ . **b** False negatives using  $F_C$ . **c** False negatives using  $F_T$

primary characteristic in determining the malignancy, of which there are no features in  $F_{HLIF}$ . Using the  $F_C$  feature set (Fig. 11b), we see lesions with varying colour distributions and symmetries. It is therefore no surprise that using the  $F_T$  feature set (Fig. 11c) also misclassifies lesions with prominent colour patterns.

Figure 12 provides some examples of false positive cases (i.e., misidentified benign cases) in the above classification results. Using the  $F_{HLIF}$  feature set (Fig. 12a), we see lesions that have asymmetric colour distributions and irregular borders, thus being classified as malignant. However, similar to the false negative cases above, their colour is fairly uniform, and could be an indicator of its benign nature. Using the  $F_C$  feature set (Fig. 12b), we see lesions with varying borders and some colour



**Fig. 12** Classification false positive examples using each feature set. **a** False positives using  $F_{HLIF}$ . **b** False positives using  $F_C$ . **c** False positives using  $F_T$

asymmetry, although some appear symmetric, smooth border, and uniform colour, making it peculiar that some were classified as malignant. However, by adding  $F_{HLIF}$  to this set, we see that the false positive cases of  $F_T$  (Fig. 12c) are very suspicious lesions, with irregular borders and strange symmetries.

### Conclusion

In this chapter we have proposed a novel framework for aiding in the diagnosis of skin lesions that uses lighting-corrected intuitive feature models. The multi-stage preprocessing step correctly adjusted the illumination across the standard camera image to ensure consistent analysis of that lesion. High-level intuitive features (HLIF) that characterize asymmetry and border irregularity were extracted and combined

with state-of-the-art low-level features. Designing HLIFs ensures that the feature scores can be conveyed to the doctor with intuitive rationale, as they are modeled in an intuitive manner.

Experimental results indicate that the illumination correction algorithm produces photographs taken with standard cameras that have been better corrected for illumination variation compared to a state-of-the-art algorithm [10]. Furthermore, linear classification using the small set of extracted HLIFs produces accurate and reliable results compared to the large state-of-the-art low-level feature set, and when combined with this large set we obtained improvements in sensitivity, specificity, and test-retest reliability. This unified framework can be used along with segmentation and advanced classification methods to provide a robust automatic diagnostic system for analysing skin lesions.

**Acknowledgments** This research was sponsored by Agfa Healthcare Inc., Ontario Centres of Excellence (OCE), and the Natural Sciences and Engineering Research Council (NSERC) of Canada.

## References

1. Alcon, J.F., Ciuhu, C., Ten Kate, W., Heinrich, A., Uzunbajakava, N., Krekels, G., Siem, D., de Haan, G.: Automatic imaging system with decision support for inspection of pigmented skin lesions and melanoma diagnosis. *IEEE J. Sel. Top. Signal Process.* **3**(1), 14–25 (2009)
2. Aldridge, R.B., Glodzik, D., Ballerini, L., Fisher, R.B.: Utility of non-rule-based visual matching as a strategy to allow novices to achieve skin lesion diagnosis. *Acta Derm. Venereol.* **91**(3), 279–283 (2011)
3. Amelard, R., Wong, A., Clausi, D.A.: Extracting high-level intuitive features (HLIF) for classifying skin lesions using standard camera images. In: *CRV'12: Ninth Conference on Computer and Robot Vision*, Toronto, pp. 396–403 (2012a)
4. Amelard, R., Wong, A., Clausi, D.A.: Extracting morphological high-level intuitive features (HLIF) for enhancing skin lesion classification. In: *EMBC'12: 34th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, San Diego, pp. 4458–4461 (2012b)
5. Argenziano, G., Soyer, H.P., Chimenti, S., Talamini, R., Corona, R., Sera, F., Binder, M., Cerroni, L., De Rosa, G., Ferrara, G.: Dermoscopy of pigmented skin lesions: results of a consensus meeting via the internet. *J. Am. Acad. Dermatol.* **48**(5), 679–693 (2003)
6. Aribisala, B.S., Claridge, E.: A border irregularity measure using a modified conditional entropy method as a malignant melanoma predictor. In: Kamel, M., Campilho, A. (eds.) *Image Analysis and Recognition*, Lecture Notes in Computer Science, vol. 3656. Springer, Heidelberg, pp. 914–921 (2005)
7. Ballerini, L., Li, X., Fisher, R.B., Rees, J.: A query-by-example content-based image retrieval system of non-melanoma skin lesions. In: Caputo, B., Müller, H., Syeda-Mahmood, T., Duncan, J., Wang, F., Kalpathy-Cramer, J. (eds.) *Medical Content-Based Retrieval for Clinical Decision Support*, Lecture Notes in Computer Science, vol. 5853, pp. 31–38. Springer, Heidelberg (2010)
8. Ballerini, L., Fisher, R.B., Aldridge, B., Rees, J.: A color and texture based hierarchical K-NN approach to the classification of non-melanoma skin lesions. In: Celebi, M.E., Schaefer, G. (eds.) *Color Medical Image Analysis*, Lecture Notes in Computational Vision and Biomechanics, vol. 6, pp 63–86. Springer, Netherlands (2013)
9. Cavalcanti, P.G., Scharcanski, J.: Automated prescreening of pigmented skin lesions using standard cameras. *Comput. Med. Imaging Graph.* **35**(6), 481–491 (2011)

10. Cavalcanti, P.G., Scharcanski, J., Lopes, C.B.O.: Shading attenuation in human skin color images. In: Bebis, G., Boyle, R., Parvin, B., Koracin, D., Chung, R., Hammoud, R., Hussain, M., Kar-Han, T., Crawfis, R., Thalmann, D., Kao, D., Avila, L. (eds.) *Advances in Visual Computing*, Lecture Notes in Computer Science, vol. 6453, pp. 190–198. Springer, Heidelberg (2010)
11. Celebi, M.E., Aslandogan, Y.A.: Content-based image retrieval incorporating models of human perception. In: *ITCC'04: International Conference on Information Technology: Coding and Computing*, vol. 2, pp. 241–245 Las Vegas (2004)
12. Celebi, M.E., Kingravi, H.A., Uddin, B., Iyatomi, H., Aslandogan, Y.A., Stoecker, W.V., Moss, R.H.: A methodological approach to the classification of dermoscopy images. *Comput. Med. Imaging Graph.* **31**(6), 362–373 (2007)
13. Celebi, M.E., Iyatomi, H., Schaefer, G.: Contrast enhancement in dermoscopy images by maximizing a histogram bimodality measure. In: *ICIP'09: 16th IEEE International Conference on Image Processing*, Cairo, pp 2601–2604 (2009)
14. Chang, C.C., Lin, C.J.: Libsvm: A library for support vector machines. *ACM Trans. Intel. Syst. Technol.* **2**(3), 27:1–27:27 (2011). <http://www.csie.ntu.edu.tw/~cjlin/libsvm>
15. Chen, M.H.: Importance-weighted marginal bayesian posterior density estimation. *J. Am. Stat. Assoc.* **89**(427), 818–824 (1994)
16. Chen, T., Yin, W., Zhou, X.S., Comaniciu, D., Huang, T.S.: *Pattern Anal. Mach. Intel. IEEE Trans.* **28**(9), 1519–1524 (2006)
17. Claridge, E., Cotton, S., Hall, P., Moncrieff, M.: From colour to tissue histology: physics-based interpretation of images of pigmented skin lesions. *Med. Image Anal.* **7**(4), 489–502 (2003)
18. Cortes, C., Vapnik, V.: Support-vector networks. *Mach. Learn.* **20**, 273–297 (1995)
19. Cotton, S.D.: A non-invasive imaging system for assisting in the diagnosis of malignant melanoma. PhD thesis, University of Birmingham, UK (1998)
20. D'Alessandro, B., Dhawan, A.P.: 3-d volume reconstruction of skin lesions for melanin and blood volume estimation and lesion severity analysis. *IEEE Trans. Med. Imaging* **31**(11), 2083–2092 (2012)
21. Day, G.R., Barbour, R.H.: Automated melanoma diagnosis: where are we at? *Skin Res. Technol.* **6**(1), 1–5 (2000)
22. *Dermatology Information System*: (2012). <http://www.dermis.net>. Accessed 08 Nov 2012
23. *DermQuest*: (2012). <http://www.dermquest.com>. Accessed 08 Nov 2012
24. Elad, M.: Retinex by two bilateral filters. In: Kimmel, R., Sochen, N.A., Weickert, J. (eds.) *Scale Space and PDE Methods in Computer Vision*, Lecture Notes in Computer Science, vol. 3459, pp. 217–229. Springer, Heidelberg (2005)
25. Engasser, H.C., Warshaw, E.M.: Dermatoscopy use by US dermatologists: a cross-sectional survey. *J. Am. Acad. Dermatol.* **63**(3), 412–419 (2010)
26. Fieguth, P.: *Statistical image processing and multidimensional modeling*, vol. 155, p. 65. Springer, New York (2010)
27. Frankle, J.A., McCann, J.J.: Method and apparatus for lightness imaging. US Patent 4,384,336 (1983)
28. Glaister, J., Wong, A., Clausi, D.A.: Illumination correction in dermatological photographs using multi-stage illumination modeling for skin lesion analysis. In: *EMBC'12: 34th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, San Diego, pp 102–105 (2012)
29. Haeghen, Y.V., Naeyaert, J.M.A.D., Lemahieu, I., Philips, W.: An imaging system with calibrated color image acquisition for use in dermatology. *IEEE Trans. Med. Imaging* **19**(7), 722–730 (2000)
30. Herzog, C., Pappo, A., Bondy, M., Bleyer, A., Kirkwood, J.: *Cancer Epidemiology in Older Adolescents and Young Adults 15 to 29 Years of Age*, National Cancer Institute, Bethesda, MD, chap Malignant Melanoma, pp 53–64. NIH Pub. No. 06–5767 (2006)
31. Howlader, N., Noone, A.M., Krapcho, M., Neyman, N., Aminou, R., Altekruse, S.F., Kosary, C.L., Ruhl, J., Tatalovich, Z., Cho, H., Mariotto, A., Eisner, M.P., Lewish, D.R., Chen, H.S., Feuer, E.J.: *Seer cancer statistics review, 1975–2009 (vintage 2009 populations)*. Technical report, Bethesda, MD (2012)

32. Hsu, C.W., Chang, C.C., Lin, C.J.: A practical guide to support vector classification (2010). <http://www.cs.sfu.ca/people/Faculty/teaching/726/spring11/svmguide.pdf>. Accessed 22 Nov 2012
33. Iyatomi, H., Celebi, M.E., Schaefer, G., Tanaka, M.: Automated color calibration method for dermoscopy images. *Comput. Med. Imaging Graph.* **35**(2), 89–98 (2011)
34. Jemal, A., Siegel, R., Xu, J., Ward, E.: Cancer statistics, 2010. *CA Cancer J. Clin.* **60**(5), 277–300 (2010)
35. Korotkov, K., Garcia, R.: Computerized analysis of pigmented skin lesions: A review. *Artif. Intell. Med.* **56**(2), 69–90 (2012)
36. Land, E.H., McCann, J.J.: Lightness and retinex theory. *J. Opt. Soc. Am.* **61**(1), 1–11 (1971)
37. Lee, T.K., Atkins, M.S., Gallagher, R.P., MacAulay, C.E., Coldman, A., McLean, D.I.: Describing the structural shape of melanocytic lesions. In: *SPIE Medical Imaging*, pp. 1170–1179 (1999)
38. Madooei, A., Drew, M.S., Sadeghi, M., Atkins, M.S.: Intrinsic melanin and hemoglobin colour components for skin lesion malignancy detection. In: Ayache, N., Delingette, H., Golland, P., Mori, K. (eds.) *Medical Image Computing and Computer-Assisted Intervention MICCAI 2012*, Lecture Notes in Computer Science, vol. 7510, pp. 315–322. Springer, Heidelberg (2012)
39. Maglogiannis, I., Doukas, C.N.: Overview of advanced computer vision systems for skin lesions characterization. *IEEE Trans. Inf. Technol. Biomed.* **13**(5), 721–733 (2009)
40. Moncrieff, M., Cotton, S., Claridge, E., Hall, P.: Spectrophotometric intracutaneous analysis: a new technique for imaging pigmented skin lesions. *Br. J. Dermatol.* **146**(3), 448–457 (2002)
41. Nachbar, F., Stolz, W., Merkle, T., Cagnetta, A.B., Vogt, T., Landthaler, M., Bilek, P., Braun-Falco, O., Plewig, G.: The ABCD rule of dermatoscopy: high prospective value in the diagnosis of doubtful melanocytic skin lesions. *J. Am. Acad. Dermatol.* **30**(4), 551–559 (1994)
42. National Center for Biotechnology Information: Melanoma—PubMed Health (2012). <http://www.ncbi.nlm.nih.gov/pubmedhealth/PMH0001853>. Accessed 08 Nov 2012
43. Nock, R., Nielsen, F.: Statistical region merging. *IEEE Trans. Pattern Anal. Mach. Intel.* **26**(11), 1452–1458 (2004)
44. Piatkowska, W., Martyna, J., Nowak, L., Przystalski, K.: A decision support system based on the semantic analysis of melanoma images using multi-elitist PSO and SVM. In: Perner, P. (ed.) *Machine Learning and Data Mining in Pattern Recognition*, Lecture Notes in Computer Science, vol. 6871, pp. 362–374. Springer, Heidelberg (2011)
45. van Rijsbergen, C.: *Information Retrieval*, 2nd edn. Butterworth-Heinemann, Newton (1979)
46. Schaefer, G., Rajab, M.I., Iyatomi, H.: Colour and contrast enhancement for improved skin lesion segmentation. *Comput. Med. Imaging Graph.* **35**(2), 99–104 (2011)
47. Shan, S., Gao, W., Cao, B., Zhao, D.: Illumination normalization for robust face recognition against varying lighting conditions. In: *AMFG'03: IEEE International Workshop on Analysis and Modeling of Faces and Gestures*, Nice, pp. 157–164 (2003)
48. Smith, A.R.: Color gamut transform pairs. *SIGGRAPH Comput. Graph.* **12**(3), 12–19 (1978)
49. Soille, P.: Morphological operators. *Handb. Comput. Vis. Appl.* **2**, 627–682 (1999)
50. Stolz, W., Riemann, A., Cagnetta, A., Pillet, L., Abmayr, W., Holzel, D., Bilek, P., Nachbar, F., Landthaler, M., Braun-Falco, O.: ABCD rule of dermatoscopy: a new practical method for early recognition of malignant melanoma. *Eur. J. Dermatol.* **4**(7), 521–527 (1994)
51. Tsumura, N., Ojima, N., Sato, K., Shiraishi, M., Shimizu, H., Nabeshima, H., Akazaki, S., Hori, K., Miyake, Y.: Image-based skin color and texture analysis/synthesis by extracting hemoglobin and melanin information in the skin. *ACM Trans. Graph.* **22**(3), 770–779 (2003)
52. Wallace, T.P., Wintz, P.A.: An efficient three-dimensional aircraft recognition algorithm using normalized fourier descriptors. *Comput. Graph. Image Proces.* **13**(2), 99–126 (1980)
53. Wong, A., Clausi, D.A., Fieguth, P.: Adaptive monte carlo retinex method for illumination and reflectance separation and color image enhancement. In: *CRV'09: Canadian Conference on Computer and Robot Vision*, Kelowna, pp. 108–115 (2009)
54. World Health Organization: WHO |Skin cancers (2012). <http://www.who.int/uv/faq/skincancer/en/index1.html>. Accessed 08 Nov 2012