

Learning Features for Streak Detection in Dermoscopic Color Images using Localized Radial Flux of Principal Intensity Curvature

Hengameh Mirzaalian¹, Tim K Lee^{1,2}, and Ghassan Hamarneh¹

¹Medical Image Analysis Lab, Simon Fraser University, BC, Canada.

²Photomedicine Institute, Department of Dermatology and Skin Science, University of British Columbia, Vancouver Coastal Health Research Institute, and Cancer Control Research Program, BC Cancer Agency, BC, Canada.

{hma36,hamarneh}@sfu.ca, tlee@bccrc.ca

Abstract

Malignant melanoma (MM) is one of the most frequent types of cancers among the world's white population. Dermoscopy is a noninvasive method for early recognition of MM by which physicians assess the skin lesion according to the skin subsurface features. The presence or absence of "streaks" is one of the most important dermoscopic criteria for the diagnosis of MM. We develop a machine-learning approach for identifying streaks in dermoscopic images using a novel melanoma feature, which captures the quaternion tubularness in the color dermoscopic images, is sensitive to the radial features of streaks, and is localized to different lesion bands (e.g. the most periphery band where streaks commonly appear). We validate the classification accuracy of SVM using our novel features on 99 dermoscopic images (including images in the absence, presence of regular, and presence of irregular streaks). Compared to state-of-the-art, we obtain improved classification results by up to 9% in terms of area under ROC curves.

1. Introduction

Malignant melanoma (MM) is one of the most frequent type of cancers among the world's white population [1, 2]. Early diagnosis of MM is an important factor for prognosis and treatment of melanoma. Dermoscopy (also called epiluminescence microscopy) is a noninvasive method for early recognition of MM, allowing a better visualization of the skin structures. Using the dermoscopic images, physicians assess the skin lesion based on the presence or absence of the different global (e.g. homogeneous, starburst, parallel patterns) or local (pigment network, dots, streaks, blue-whitish veil, regression structures, hypopigmentation, blotches, vascular structures) dermoscopic features.

Recently, a considerable amount of research has focused on automating the feature extraction and classification of

dermoscopic images as a key step toward performing machine learning for computer aided diagnosis. An overview of the existing feature extraction methods for skin lesion characterization has been recently reported in [12]. These techniques investigate global appearance descriptors like border asymmetry and irregularity [5], color variation [8], and texture patterns, e.g. using the Fourier power spectrum [17], statistics of the wavelet transform coefficients [16], Gaussian derivative kernels [19], or Laws's kernels [3]. The aforementioned methods are developed to learn features to differentiate melanoma from melanocytic nevi, or detect the pigment network, homogeneous pattern, globular pattern, reticular pattern, vascular pattern, and blue-white veil. Due to the clinical importance of the absence or presence of the regular or irregular streaks in dermoscopic images [4], in this study, we focus on extracting features for a machine learning system for streak detection. To the best of our knowledge, only a few studies have focused on streak detection, in which global appearance descriptors, such as color variation and border irregularities of the lesion, are extracted [5, 9].

Streaks, also referred to as radial streamings, appear as linear structures located at the periphery of a lesion and are classified as either regular or irregular depending on the appearance of their intensity, texture, and color distribution [4]. Examples of dermoscopic images in the absence and presence of the streaks are shown in the first column of Figure 1.

Since streaks have a ridge like appearance, their analysis stands to benefit from state-of-the-art research in image analysis of curvilinear structures, most notably the plethora of works on vasculature (cf. recent survey [11]). But unfortunately, the small body of work on streak detection fades in comparison. In this work, we utilize, for the first time, Hessian based tubularness filters to enhance streak structures in dermoscopic images (Section 2.1). We are the first to make use of orientation information for streak classification

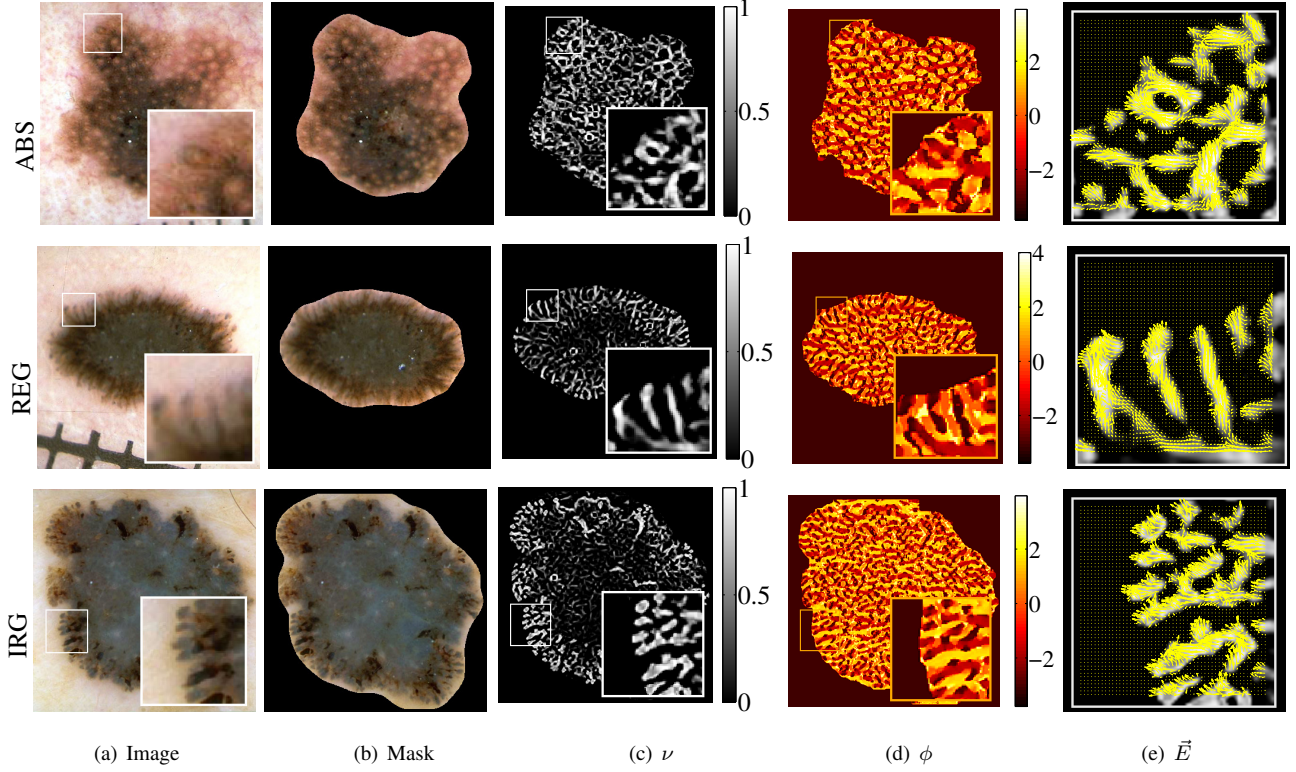


Figure 1. Examples of dermoscopic images from [4] and the enhanced images using the tubularness filter responses [10]. The results are shown for dermoscopic images in the absence (first row), presence of regular (second row), and irregular streaks (third row). (a) Dermoscopic image [4]. (b) The segmented lesion using graph-cuts. (c-d) Streak-enhancement and estimated streak-direction resulting from applying Frangi et al.’s filter (1). (e) Vector field of the streaks according to (2).

through the use of eigenvalue decomposition of the Hessian matrix (Section 2.2). Given the estimated tubularness and direction of the streaks, we define a vector field in order to quantify radial streaming pattern of the streaks. In particular, we compute the amount of flux of the field passing through iso-distance contours of the lesion. We construct our appearance descriptor based on the mean and variance of the flux through the different concentric bands of the lesion, which in turn allows for more localized features without the prerequisite of explicitly calculating a point-to-point correspondence between the lesion shapes (Section 2.3). We validate the classification accuracy of a SVM classifier based on our extracted features (Section 3). Our reported results on 99 dermoscopic images show that we obtain improved classification (by up to 9% in term of area under ROC curves), compared to state-of-the-art (Section 4).

2. Methods

2.1. Tubularness Filter for Streak Enhancement

Frangi et al. [10] proposed to measure the tubularness $\nu(\mathbf{x}, s)$ at pixel $\mathbf{x} = (x, y)$ for scale s using:

$$\nu(\mathbf{x}, s) = e^{-\frac{R^2(\mathbf{x}, s)}{2\beta^2}} \left(1 - e^{-\frac{S^2(\mathbf{x}, s)}{2c^2}} \right), \quad (1)$$

$$R(\mathbf{x}, s) = \frac{\lambda_1(\mathbf{x}, s)}{\lambda_2(\mathbf{x}, s)}, \quad S(\mathbf{x}, s) = \sqrt{\sum_{i \leq 2} \lambda_i^2(\mathbf{x}, s)}$$

where $\lambda_i(\mathbf{x}, s)$, $i = 1, 2$ ($|\lambda_1| \leq |\lambda_2|$) are the eigenvalues, resulting from singular value decomposition (SVD), of the Hessian matrix of image I computed at scale s . R and S are measures of blobness and “second order structureness”, respectively. β and c are parameters that control the sensitivity of the filter to the measures R and S .

Figure 1 shows the computed tubularness based on (1) for dermoscopic images of the different types: in the absence, presence of regular, and presence of irregular streaks, denoted by ABS, REG, and IRG, respectively.

2.2. Flux Analysis of the Streaks’s Principle Curvature Vectors

While computing tubularness of the streaks using (1), we make an estimation of the streak direction $\phi(\mathbf{x}, s)$. It is computed as the angle between the x-axis and the eigenvector corresponding to $\lambda_1(\mathbf{x}, s)$, which points along the direction of the minimum intensity curvature. Given ϕ and ν , we define a “streak vector field” as:

$$\vec{E} = (\nu \cos(\phi), \nu \sin(\phi)) \quad (2)$$

Examples of the computed ϕ and \vec{E} are shown in the last two columns of Figure 1. To quantify the radial streaming pattern of the vector field with respect to a lesion contour C , we measure the amount of the flow of \vec{E} parallel and perpendicular to C , denoted by ψ_{\parallel} and ψ_{\perp} , respectively, using:

$$\begin{aligned}\psi_{\parallel}(\vec{E}, C) &= \oint_C \|\vec{E} \times \vec{n}\| \, dc \\ \psi_{\perp}(\vec{E}, C) &= \oint_C |\langle \vec{E}, \vec{n} \rangle| \, dc\end{aligned}\quad (3)$$

where \vec{n} is the normal vector to C , \times and $\langle \cdot \rangle$ denote cross and dot products between the vectors, and $\|\cdot\|$ and $|\cdot|$ measure the L_2 norm of the vector and the absolute value of the scalar, respectively. By computing (3), we state our hypothesis as: in the presence of streaks on the contour C , ψ_{\parallel} and ψ_{\perp} would take low and high values, respectively, capturing the known radial characteristic of the streaks. In Section 2.3, we discuss how to utilize the measured flux to construct a feature vector for streak detection.

Note that in our implementation, we make an initial estimation of C by applying a binary graph cut segmentation [6], where the data term and regularization terms are set using the distribution of the pixel intensities and the Pott's model, respectively [7]. The intensity distributions of the foreground and background are estimated by clustering, in color space, the image pixels into two distinct clusters.

2.3. Streak Detection Features

We measure ψ_{\parallel} and ψ_{\perp} according to (3) over iso-distance contours of the lesion, where each contour is the loci of the pixels which have equal distance from the outer lesion contour C_o . We calculate the distance transform (DT) of the lesion mask to extract the iso-distance contours, denoted by C_d , where d represents the distance between C_d and C_o . Figure 2 shows an example of the computed DT of a lesion mask and the iso-distance contours C_d . We compute the mean and variance of the flux of the different bands of the lesion, where the K^{th} band of thickness Δ , $B_{K,\Delta}$, is defined as the region limited between the contours $C_{K\Delta}$ and $C_{(K-1)\Delta}$ and is given by:

$$B_{K,\Delta}(\mathbf{x}) = \chi(C_{K\Delta}(\mathbf{x})) \cap (1 - \chi(C_{(K-1)\Delta}(\mathbf{x}))) \quad (4)$$

where $\chi(C)$ is the region inside contour C . Therefore, the mean and variance of the flux over band $B_{K,\Delta}$ are given by:

$$\mu_{K,\parallel} = \sum_{d=(K-1)\Delta}^{K\Delta} \psi_{\parallel}(C_d) / \int_{\mathbf{x} \in \Omega} B_{K,\Delta}(\mathbf{x}) \, d\mathbf{x}, \quad (5)$$

$$\sigma_{K,\parallel} = \sqrt{\sum_{d=(K-1)\Delta}^{K\Delta} (\psi_{\parallel}(C_d) - \mu_{K,\parallel})^2 / \int_{\mathbf{x} \in \Omega} B_{K,\Delta}(\mathbf{x}) \, d\mathbf{x}},$$

$$\begin{aligned}\mu_{K,\perp} &= \sum_{d=(K-1)\Delta}^{K\Delta} \psi_{\perp}(C_d) / \int_{\mathbf{x} \in \Omega} B_{K,\Delta}(\mathbf{x}) \, d\mathbf{x}, \\ \sigma_{K,\perp} &= \sqrt{\sum_{d=(K-1)\Delta}^{K\Delta} (\psi_{\perp}(C_d) - \mu_{K,\perp})^2 / \int_{\mathbf{x} \in \Omega} B_{K,\Delta}(\mathbf{x}) \, d\mathbf{x}}\end{aligned}$$

where Ω is the image domain. Note that the denominator in (5) corresponds to the area of the K^{th} band, which is used to normalize the extracted features. After computing μ and σ of the flux of the N different bands ($K = \{1, 2, \dots, N\}$), our SVD-flux based feature vector, denoted by SVD-FLX, is constructed by concatenating the measurements of the different bands and is given by:

$$\text{SVD-FLX} = \{\mu_{1,\parallel}, \sigma_{1,\parallel}, \mu_{1,\perp}, \sigma_{1,\perp}, \dots, \mu_{N,\parallel}, \sigma_{N,\parallel}, \mu_{N,\perp}, \sigma_{N,\perp}\}. \quad (6)$$

Note that to make use of color information in the computed tubularness in (1), the tubularness is measured using the eigenvalues of the quaternion Hessian matrix of the color image [14]. We denote the feature vector utilizing quaternion Hessian matrix by QSVD-FLX and provide a comparison between the classification accuracies of SVD-FLX and QSVD-FLX in Section 4.

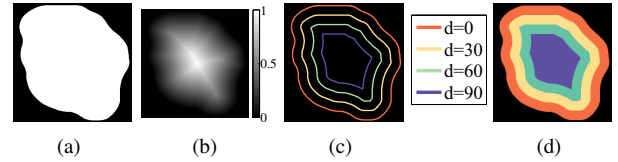


Figure 2. The iso-distance contours and subbands of a lesion. (a) Lesion mask. (b) Distance transform of (a). (c) Iso-distance contours C_d of the lesion, where d represents the distance between C_d and the lesion border in (a). (e) Bands of the lesion defined according to (4) between the contours in (d).

3. Machine Learning for Streak Classification

The final step in our approach is to learn how the extracted descriptors can best distinguish the three different classes: the absence (ABS), presence of regular (REG), or presence of irregular (IRG) streaks in the dermoscopic images. The 3-class classification task is realized using an efficient pairwise classification. The pairwise classification is based on a non-linear SVM, trained and then validated according to a leave-one-out scheme [13].

The SVM classifier requires the setting of two parameters: ξ , which assigns a penalty to errors, and γ , which defines the width of a radial basis function [18]. We compute the false positive (FP) and true positive (TP) rates of the classifier for different values of ξ and γ in a logarithmic grid search (from 2^{-8} to 2^8) to create a receiver operating characteristic (ROC) curve. Therefore, each pair of the parameters (ξ_i, γ_j) would generate a point (FP_{ij}, TP_{ij}) in the graph. The ROC curve is constructed by selecting the

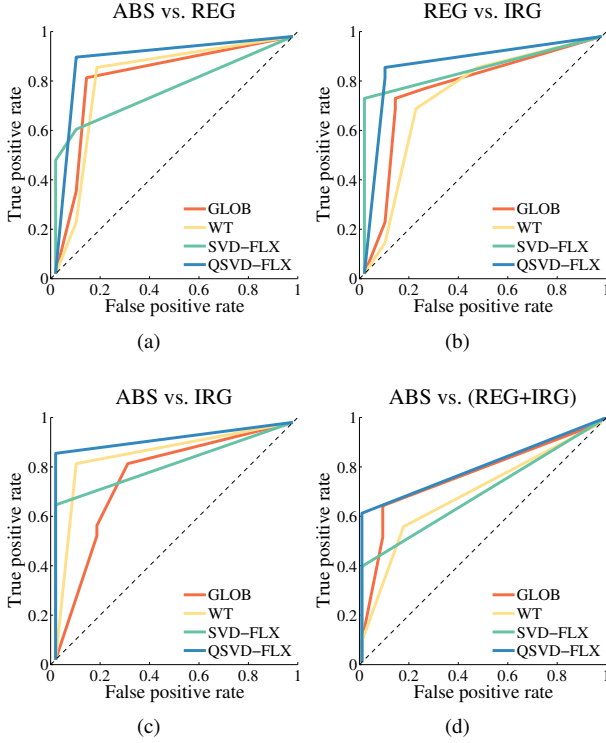


Figure 3. ROC curves of the pairwise classifiers resulting from using the different descriptors. Areas under the ROC curves are reported in Table 1.

set of optimal operating points. Point (FP_{ij}, TP_{ij}) is optimal if there is no other point (FP_{mn}, TP_{mn}) such that $FP_{mn} \leq FP_{ij}$ and $TP_{mn} \geq TP_{ij}$. We use the area under the generated ROC curves obtained from classification involving different descriptors to compare their discriminatory power.

4. Results

The proposed algorithm has been tested on $99\,768 \times 512$ -pixel dermoscopic images of Argenziano et al.’s atlas of dermoscopy [4], including 33 images in which the streaks are absent (ABS), 33 images in which regular streaks are present (REG), and 33 images with irregular streaks (IRG). Note that the whole dataset in [4] consists of 527 images of different resolutions, ranging from 0.033 to 0.5 mm/pixel. The 99 out of 527 images are selected such that a complete lesion occupying more than 10% of the image can be seen, since only then the lesion texture is reasonably visible and suitable for analysis.

Figure 3 and Table 1 show the classification accuracies of the different descriptors in terms of the ROC curves and the areas under them, where GLOB, WT, SVD-FLX, and QSVD-FLX denote the global descriptors used in [9]¹, WT-

¹GLOB is constructed using the mean and variance of pixel intensities in different color spaces (RGB, HSI, and Luv) and the border irregularities, where the latter is measured via the change in the lesion contour pixels’ coordinates relative to the lesion’s centroid and the ratio between the lesion

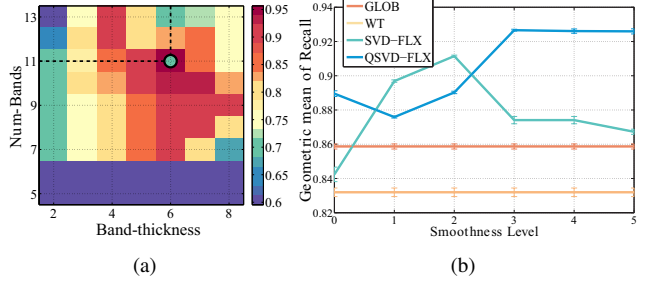


Figure 4. Classification accuracy of our flux-based descriptor, QSVD-FLX, vs. different parameters: (a) Accuracies (pixel intensities) vs. different number of bands (Y-axis) and band thicknesses (X-axis). The green dot indicates the maximum accuracy. Note that the accuracies are reported in terms of the geometric-mean of the AUCs of the pairwise classifiers. (b) Classification accuracy (Y-axis) vs. different smoothness levels of the lesion border (X-axis).

based descriptors used in [16]², and our flux-based descriptors using the eigenvalues of the luminance and RGB images in (1), respectively. Note that we measure the classification accuracies of the flux-based descriptors for different numbers of bands and thicknesses in (5), $K \in [5, 13]$ and $\Delta \in [2, 8]$ pixels, and report the optimum accuracies (Figure 4(a)). In the last row of Table 1, we provide multi-class classification accuracies as the geometric mean of the pairwise classifiers, as suggested in [15]. Furthermore, since it is expected that the classification accuracies of the flux-based descriptors be sensitive to the lesion border C , in Figure 4(b), we report the change in classification accuracies as C is varied, which in turn is done by applying different levels of regularization (smoothness) to C . It can be seen that our proposed descriptors produce higher classification accuracy for a wide range of smoothness levels.

The results indicate that, averaged over all the groups (the last row in Table 1), we obtain 91% accuracy; up to 10% increase in terms of area under ROC curves compared with GLOB and WT. Furthermore, it can be noticed that QSVD-FLX, which is obtained by considering the color information, results in an average of 4% improvement in the classification accuracy compared with SVD-FLX. It should be mentioned that many of the existing feature extraction methods for skin lesions have been developed to detect the presence or absence of streak structures. Therefore, we report the classification accuracies for ABS vs. (REG+IRG) in the last row of Table 1. The results indicate the superiority of QSVD-FLX compared with SVD-FLX.

contour length and the maximum axis of the contour’s convex hull (details in [9]).

²The WT-based descriptors are constructed by concatenating the mean and variance of the WT-coefficients of the different sub-bands of the WT using Haar wavelets and three decomposition levels (details in [16]).

| Group1 vs. Group2 | Area under the ROC curves | | | | Selected Descriptor(s) |
|-------------------|---------------------------|---------|---------|-------------|------------------------|
| | GLOB [9] | WT [16] | SVD-FLX | QSVD-FLX | |
| ABS vs. REG | 0.83 | 0.83 | 0.80 | 0.91 | QSVD-FLX |
| REG vs. IRG | 0.78 | 0.75 | 0.87 | 0.89 | QSVD-FLX |
| ABS vs. IRG | 0.77 | 0.87 | 0.88 | 0.93 | QSVD-FLX |
| Geometric Mean | 0.79 | 0.81 | 0.85 | 0.91 | QSVD-FLX |
| ABS vs. (REG+IRG) | 0.78 | 0.70 | 0.70 | 0.80 | QSVD-FLX |

Table 1. Area under the ROC curves in Figure 3. The last column shows the descriptor(s) that resulted in the highest AUC. Note that we report multi-class classification accuracies in terms of geometric mean (GM) of the pairwise classifiers (as done in [15]). The last row compares the ability of shape descriptors to discriminate between the absence and presence of the streaks (ABS vs. (IRG+REG)).

5. Discussion and Conclusion

Automating feature extraction and classification of dermoscopic images is of utmost importance for early detection of potential malignant melanoma (MM). Presence or absence of the characteristic streaks is one of the most important dermoscopic criteria for MM diagnosis. We proposed a novel appearance descriptor that captures the tubularness in the color dermoscopic images, is sensitive to the radial features of streaks, and is localized to different lesion bands (e.g. the most periphery band where streaks commonly appear). The experimental results show that we achieve improved classification results compared to the state-of-the-art global and wavelet transform based descriptors. We plan to extend our method to detect and classify the presence of other dermoscopic features (e.g. pigment network, dots, vascular structures), moving us an important step forward towards a machine-learning based computer aided diagnosis system for early detection of MM.

Acknowledgements

We would like to thank Dr. Giuseppe Argenziano at the University of Naples for sharing the dermoscopy data [4].

References

- [1] American cancer society. Cancer facts and figures. 2009. Atlanta, USA. 57
- [2] Canadian cancer society’s steering committee. Canadian cancer statistics. 2009. Toronto, Canada. 57
- [3] M. Anantha, R. H. Moss, and W. V. Stoecker. Detection of pigment network in dermatoscopy images using texture analysis. *Computerized Medical Imaging and Graphics*, 28(5):225–234, 2004. 57
- [4] G. Argenziano, H. Soyer, V. Giorgio, D. Piccolo, P. Carli, M. D. A. Ferrari, R. Hofmann, D. Massi, G. Mazzocchi, M. Scalvenzi, and H. Wolf. *Interactive Atlas of Dermoscopy*. Edra Medical Publishing and New Media, 2000. 57, 58, 60, 61
- [5] G. Betta, G. D. Leo, G. Fabbrocini, A. Paolillo, and M. Scalvenzi. Automated application of the 7-point checklist diagnosis method for skin lesions: Estimation of chromatic and shape parameters. *Instrumentation and Measurement Technology Conference*, 3:1818–1822, 2005. 57
- [6] Y. Boykov and G. Funka-Lea. Graph cuts and efficient N-D image segmentation. *Int. J. Comput. Vision*, 70(2):109–131, Nov. 2006. 59
- [7] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *IEEE TPAMI*, 23:2001, 1999. 59
- [8] M. E. Celebi, H. Iyatomi, W. V. Stoecker, R. H. Moss, H. S. Rabinovitz, G. Argenziano, and H. P. Soyer. Automatic detection of blue-white veil and related structures in dermoscopy images. *Computerized Medical Imaging and Graphics*, 32(8):670 – 677, 2008. 57
- [9] G. Fabbrocini, G. Betta, G. Leo, C. Liguori, A. Paolillo, A. Pietrosanto, P. Sommella, O. Rescigno, S. Cacciapuoti, F. Pastore, V. Vita, I. Mordente, and F. Ayala. Epiluminescence image processing for melanocytic skin lesion diagnosis based on 7-point check-list: A preliminary discussion on three parameters. *The Open Dermatology Journal*, 4:110–115, 2010. 57, 60, 61
- [10] A. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever. Multiscale vessel enhancement filtering. *MIC-CAI*, pages 130–137, 1998. 58
- [11] C. Kirbas and F. K. H. Quek. A review of vessel extraction techniques and algorithms. *ACM Computing Surveys*, 36:81–121, 2004. 57
- [12] I. Maglogiannis and C. Doukas. Overview of advanced computer vision systems for skin lesions characterization. *IEEE TITB*, 13(5):721 –733, 2009. 57
- [13] S. Park and J. Frnkranz. Efficient pairwise classification. In *ECML*, volume 4701, pages 658–665. Springer, 2007. 59
- [14] L. Shi, B. Funt, and G. Hamarneh. Quaternion color curvature. In *Color Imaging*, pages 338–341, 2008. 59
- [15] Y. Sun, M. Kamel, and Y. Wang. Boosting for learning multiple classes with imbalanced class distribution. In *IEEE ICDM*, page 592602, 2006. 60, 61
- [16] G. Surowka and K. Grzesiak-Kopec. Different learning paradigms for the classification of melanoid skin lesions using wavelets. In *IEEE EMBS*, pages 3136 –3139, 2007. 57, 60, 61
- [17] T. Tanaka, S. Torii, I. Kabuta, K. Shimizu, M. Tanaka, and H. Oka. Pattern classification of nevus with texture analysis. *IEEE EMBS*, pages 1459–1462, 2004. 57
- [18] V. Vapnik. *Statistical Learning Theory*. Wiley, 1998. 59
- [19] H. Zhou, M. Chen, and J. M. Rehg. Dermoscopic interest point detector and descriptor. 1:1318–1321, 2009. 57