

Modeling the Dermoscopic Structure Pigment Network Using a Clinically Inspired Feature Set

Maryam Sadeghi^{1,2,3}, Majid Razmara¹, Paul Wighton^{1,2,3},
Tim K. Lee^{1,2,3}, and M. Stella Atkins¹

¹ School of Computing Science, Simon Fraser University, Canada
msa68@sfu.ca

² Department of Dermatology and Skin Science, University of British Columbia and
Vancouver Coastal Health Research Institute, Canada

³ Cancer Control Research Program and Cancer Imaging Department, BC Cancer
Research Centre, Canada

Abstract. We present a method to detect and classify the dermoscopic structure pigment network which may indicate early melanoma in skin lesions. We locate the network as darker areas constituting a mesh, as well as lighter areas representing the ‘holes’ which the mesh surrounds. After identifying the lines and holes, 69 features inspired by the clinical definition are derived and used to classify the network into one of two classes: *Typical* or *Atypical*. We validate our method over a large, inclusive, real-world dataset consisting of 436 images and achieve an accuracy of 82% discriminating between three classes (*Absent*, *Typical* or *Atypical*) and an accuracy of 93% discriminating between two classes (*Absent* or *Present*).

1 Introduction

Melanoma, a cancerous lesion in the pigment-bearing basal layers of the epidermis, is the most deadly form of skin cancer, yet treatable via excision if detected early. The cure rate for early-stage melanoma is nearly 100%. A recent study [1] has concluded that dermoscopy increases the early detection of melanoma, only if the practitioner is sufficiently trained. In fact, dermoscopy *decreases* accuracy if training is insufficient. There is, therefore, a demand to develop computer-aided diagnostic systems to facilitate the early detection of melanoma. This paper follows a relatively new trend in clinical dermatology: to identify specific ‘dermoscopic structures’ in the lesions such as the pigment network which is then used to arrive at a diagnosis [2]. A pigment network can be classified as either *Typical* or *Atypical*, where a working definition of a typical pigment network (TPN) is “a light-to-dark-brown network with small, uniformly spaced network holes and thin network lines distributed more or less regularly throughout the lesion and usually thinning out at the periphery” [3]. For an atypical pigment network (APN) we use the working definition “a black, brown or gray network with irregular holes and thick lines” [3]. The goal is to automatically classify a given image to one of three classes: *Absent*, *Typical*, or *Atypical*. Figure 1 exemplifies these 3 classes.

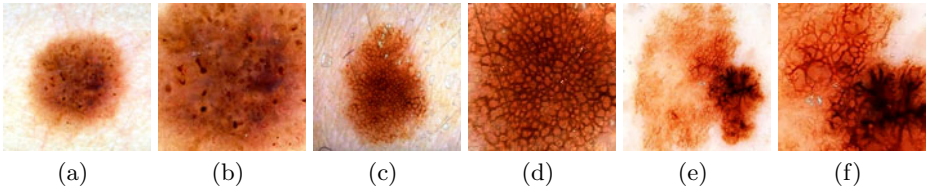


Fig. 1. The 3 classes of the dermoscopic structure pigment network: a-b) *Absent*; c-d) *Typical*; and e-f) *Atypical*. b),d),f) are magnifications of a),c),e) respectively.

We use these definitions to subdivide the structure into the darker mesh of the pigment network (which we refer to as the ‘net’) and the lighter colored areas the net surrounds (which we refer to as the ‘holes’). After identifying these substructures we use the definitions above to derive a several structural, geometric, chromatic and textual features suitable for classification. The result is a robust, reliable, automated method for identifying and classifying the structure pigment network.

2 Previous Work

The automated detection of pigment network has received some recent attention [4,5,6,7,8,9,10]. Fleming et al. [4] report techniques for extracting and visualizing pigment network via morphological operators. Fischer et al. [5] use local histogram equalization and gray level morphological operations to enhance the pigment network. Anantha et al. [6] propose two algorithms for detecting pigment networks in skin lesions. They are most successful when employing a weighted average of two Laws’ energy masks whose weights are determined empirically and report a classification accuracy of approximately 80%. Betta et al. [7] begin by taking the difference of an image and its response to a median filter. This difference image is thresholded to create a binary mask which undergoes a morphological closing operation to remove any local discontinuities. This mask is then combined with a mask created from a high-pass filter applied in the Fourier domain to exclude any slowly modulating frequencies. Results are reported graphically, but appear to achieve a sensitivity of 50% with a specificity of 100%. Di Leo et. al. [8] extend this method and compute features over the ‘holes’ of the pigment network. A decision tree is learnt in order to classify future images and an accuracy of 71.9% is achieved. Shrestha et. al. [9] begin with a set of 106 images where the location of the APN has been manually segmented. If no APN is present, then the location of the most ‘irregular texture’ is manually selected. They then compute several texture metrics over these areas (energy, entropy, etc.) and employ various classifiers to label unseen images. They report accuracies of approximately 95%. Recently, Sadeghi et al. [10] presented an approach whereby ‘holes’ of the network are detected using a graph-based loop search on the result of an edge detection algorithm. They then create a graph

based on the distance of these ‘holes’ which they use to propose a new spatial feature called ‘density ratio’ for detecting a pigment network in a given image.

Although these studies have certainly made significant contributions, there has yet to be a comprehensive analysis of pigment network detection on a large number of dermoscopic images under ‘real-world’ conditions. All work to date has either: 1) not reported quantitative validation [4,5]; 2) validated against a small ($n < 100$) number of images[7]; 3) only considered or reported results for the 2-class problem (e.g. *Absent/Present* rather than *Absent/Typical/Atypical*) [6,7,8,9,10]; 4) not explicitly identified the location of the network [6]; or 5) has made use of unrealistic exclusion criteria and other manual interventions [9].

This paper presents an effective method for pigment network segmentation and classification, which is validated on a large ($n = 436$) ‘real-world’ dataset.

3 Method

An overview of our method for the identification and classification of pigment network is given in Fig. 2. After pre-processing, Sadeghi et al.’s [10] ‘hole detector’ is employed to generate a ‘hole mask’ indicating the pixels belonging to the holes of the pigment network. Next, a ‘net mask’ is created, indicating the pixels belonging to the net of the pigment network. We then use these masks to compute a variety of features including structural (which characterizes shape), geometric (which characterizes distribution and uniformity), chromatic and textural features. These features are fed into a classifier to classify unseen images.

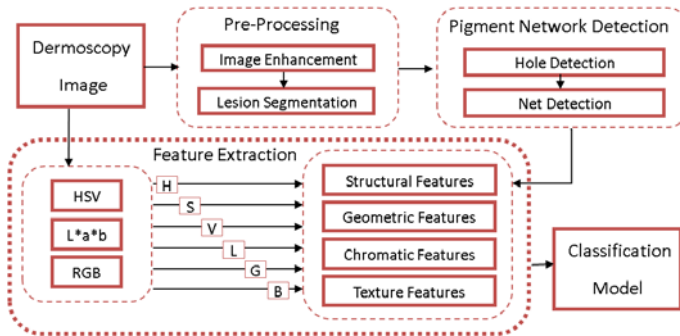


Fig. 2. Overview of construction of our classification model

3.1 Pre-Processing

Image Enhancement: The image is first enhanced so that the pigment network is more visible using manner similar to [5]. A contrast-limited adaptive histogram equalization is used [11]. Images are then further sharpened by subtracting a blurred version of the image from itself.

Lesion Segmentation: Next, in order to prevent unnecessary analysis of the pixels belonging to the skin, lesions are segmented using Wighton et. al.'s method [12] which employs supervised learning and the random walker algorithm. The output of the segmentation is a 'lesion mask' which indicates which pixels belong to the lesion (as opposed to the surrounding healthy skin).

3.2 Pigment Network Detection

Hole Detection: To find the holes of the pigment network, we used Sadeghi et al.'s method [10] which employs a Laplacian of Gaussian edge detector to create a graph-based structure which is used to identify holes of the pigment network. They demonstrate that *globules* (another dermoscopic structure which is difficult to discriminate from *pigment network*) are separable from the holes of the pigment network according to the difference between the average intensity of inner pixels and the average intensity of the border [10]. The result of this process is a 'hole mask' which indicates which pixels belong to the holes of the pigment network.

Net Detection: In order to identify the net of a pigment network, we apply the Laplacian of Gaussian (LoG) filter to the green channel of the image. The LoG filter identifies high frequency components of an image and therefore makes an ideal net detector. The major issue with applying this operator is that its response is strongly dependent on the relationship between the frequency of the structures and the size of the Gaussian kernel used. We used $\sigma = 0.15$, which is an appropriate value for images of the two atlases used in our experiment [13,14], however it can be tuned for a given imageset according to scale and magnification. In our experiment, we observed that the average thickness of the pigment network is proportional to the average size of holes of the network. We therefore set the size of the LoG window size to half of the average hole size in the image. The average window size over all images of our data set is 11 pixels. We then threshold the filter response, resulting in a 'net mask' which indicates which pixels belong to the net of the the pigment network. Furthermore, we skeletonize this mask, resulting in a 'skeleton mask'. Figure 3 illustrates the net extraction process.

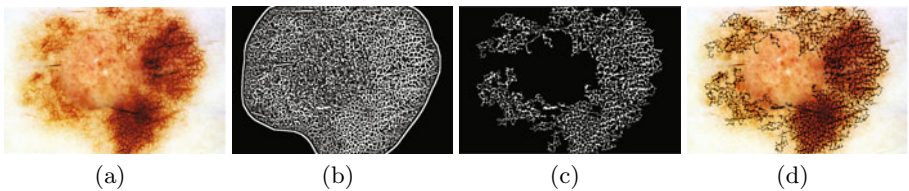


Fig. 3. Net detection. a) A dermoscopic image, b) response of the LoG filter, c) the resulting 'net mask', and d) the extracted net of the pigment network overlaid on the original image

3.3 Feature Extraction

Based on our working definitions of TPN and APN, we use the resulting masks of sections 3.2 and 3.1 (namely the lesion, hole, net and skeleton masks) to propose a set of features capable of discriminating between the 3 classes (*Absent*, *Typical* and *Atypical*). We propose a set of structural (shape), geometric (spatial) chromatic and textural features.

Structural Features (20): Diagnostically important characteristics of an *Atypical* network include the thickness and variation in thickness of net as well as the size and variation in size of network holes.

For each spatially disjoint section of the net mask, we compute its size (number of pixels in the net mask) and length (number of pixels in the skeleton mask). Our features are then the mean, standard deviation and ratio (*mean/std*) of the sizes and lengths of net sections. Thickness is also computed by measuring the distance from each pixel in the net mask to the closest pixel in the skeleton mask. The mean, standard deviation and ratio of thickness as well as a 6-bin thickness histogram are also included as features.

For each spatially disjoint section of the hole mask, we compute the size (number of pixels) and include as features the mean, standard deviation and ratio (*mean/std*) of hole size as well as the total number of holes.

We also include the ratio of the network size (number of pixels in the net and hole masks) to the lesion size (number of pixels in the segmentation mask).

Geometric Features (2): Clinically, there is an emphasis on the ‘uniformity’ of the network in order to differentiate between TPN and APN. Sadeghi et al. [10] have proposed the feature ‘density ratio’ of holes which, while useful in discriminating between the absence and presence of a pigment network, does not reliably discriminate between the TPN and APN. We include ‘density ratio’ of holes as a feature, as well as a new feature, ‘hole irregularity’. This feature is computed by constructing a graph as in [10] where 2 holes are connected if they are less than 3 times the average diameter of the holes. ‘Hole irregularity’ is then the number of edges in this graph.

Chromatic Features (37): Color also plays a crucial role in clinical diagnosis. We therefore convert the image to HSV colourspace and compute features over each channel as well as the original green channel of the image. In each channel, for the hole, net and lesion masks respectively we compute the mean, standard deviation and ratio (*mean/std*) of the intensity values. Additionally, we also propose a new chromatic feature called the ‘atypicality measure’ which is the sum of the intensity values over the green channel of the pixels in the net mask.

Textural Features (10): We use five of the classical statistical texture measures of Haralick et al. [15]: entropy, energy, contrast, correlation and homogeneity which are derived from a grey level co-occurrence matrix (GLCM). The GLCM is a tabulation of how often different combinations of pixel brightness values (gray levels) occur in a specific pixel pairing of an image. We construct

2 GLCMs and extract the 5 texture metrics from each. The first GLCM is constructed over the entire lesion (using the pixels in the lesion mask) and the second is constructed over the pigment network (using the pixels in the net and hole masks).

3.4 Classification

Finally, these 69 features are fed into a classifier so that new images can be classified. We employ the WEKA's [16] implementation SimpleLogistic which uses a powerful boosting algorithm LogitBoost. Boosting is a method for combining the performance of many features to produce a powerful classifier. SimpleLogistic fits logistic models by applying LogitBoost with simple regression functions as base learners.

4 Evaluation and Results

We applied the method described above to a set of dermoscopic images taken from two atlases of dermoscopy [13,14]. In [13] each image is labeled as *Absent*, *Typical* or *Atypical*, representing the presence and regularity of the dermoscopic structure pigment network. However the images in [14] have been labeled by 40 experts, each one assigning a label of either *Absent*, *Typical* or *Atypical* to each image. Overall labels for these images are generated by majority voting. In total, our dataset consists of 436 images (161 *Absent*, 154 *Typical*, 121 *Atypical*). We compute results for both the 3-class (*Absent*, *Typical* or *Atypical*) and 2-class problems (*Absent*, *Present*). Ten-fold cross validation was used to generate all results. For comparison, the feature set described in [8] was also implemented and results over our imagesets computed. Table 1 summarizes these results.

Table 1. Comparing accuracy, precision, recall and f-measure of our proposed features with Di Leo et al.'s features using the same set of 436 images. The last three rows summarize the results from previous work on different image sets.

Absent-Typical-Atypical Classification					
	Precision	Recall	F-measure	Accuracy	N
Absent	0.905	0.950	0.927	-	161
Typical	0.787	0.792	0.790	-	154
Atypical	0.750	0.694	0.721	-	121
Weighted Avg	0.820	0.823	0.821	0.823	436
Di Leo et al. [8]	0.709	0.711	0.709	0.719	436
Absent-Present Classification					
Absent	0.893	0.932	0.912	-	161
Present	0.959	0.935	0.947	-	275
Weighted Avg	0.935	0.933	0.934	0.933	436
Di Leo et al. [8]	0.875	0.876	0.875	0.876	436
Absent-Present Classification on Different Image Sets					
Anantha et al. [6]	-	-	-	0.80	155
Betta et al. [7]	-	-	-	0.66	30
Sadeghi et al. [10]	-	-	-	0.926	500

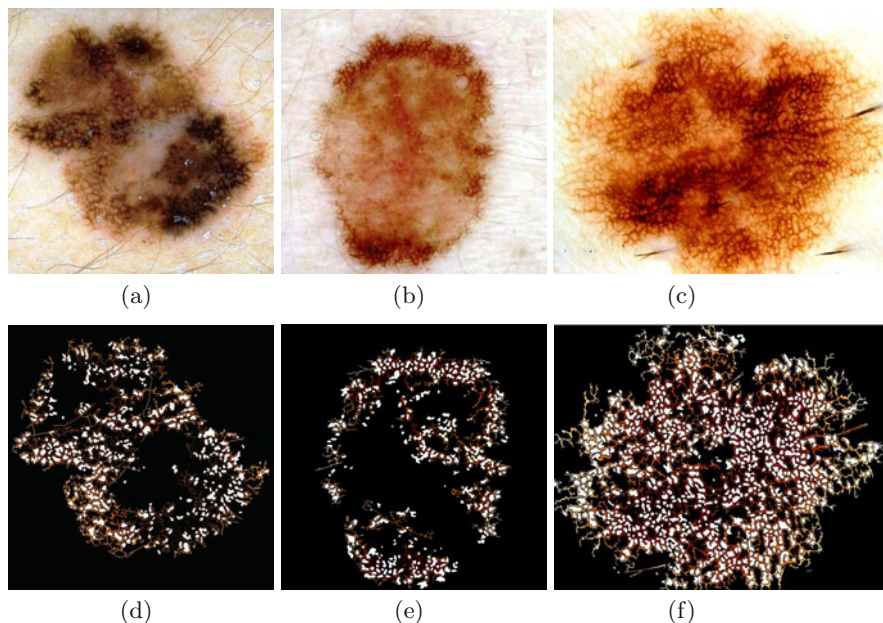


Fig. 4. Three images of the image set: the top row shows the original images of a APN, APN, and TPN. The bottom row shows their corresponding pigment networks.

Comparing these results with the results generated by the others using different datasets may be tenuous, nevertheless these reported values are also summarized in Table 1. As can be seen, this work outperforms other previous work on the 2-class problem and is the only one to report quantitative results of the 3-class problem. Additionally, qualitative results of detecting pigment network ‘net’ and ‘holes’ is illustrated in Figure 4.

5 Conclusion and Future Work

We have described techniques to identify the sub-structures of the dermoscopic structure pigment network. Furthermore, we have proposed and validated a set of clinically motivated features over these sub-structures suitable for classification. Our feature set has proven to be extremely robust, outperforming previous work on a more inclusive ‘real-world’ dataset consisting of 436 images, which is the largest validation to date on the 3-class problem.

This feature set can be used to aid in the automation of clinical dermoscopic algorithms [2]. Future work will be motivated by the notion of a fully automatic, robust automation of clinical dermoscopic algorithms and will therefore focus on reliably identifying and classifying other dermoscopic structures in a similar fashion.

Acknowledgements

This work was funded by the Canadian Natural Sciences and Engineering Council (NSERC), CIHR-Skin Research Training Centre (SRTC) and a grant from the Canadian Health Research Project (CHRP).

References

1. Kittler, H., Pehamberger, H., Wolff, K., Binder, M.: Diagnostic accuracy of dermoscopy. *The Lancet Oncology* 3(3), 159–165 (2002)
2. Johr, R.: Dermoscopy: alternative melanocytic algorithm—the ABCD rule of dermatoscopy, Menzies scoring method, and 7-point checklist. *Clinics in Dermatology* 20(3), 240–247 (2002)
3. Argenziano, G., Soyer, H., et al.: Dermoscopy of pigmented skin lesions: results of a consensus meeting via the Internet. *Journal of the American Academy of Dermatology* 48(5), 679–693 (2003)
4. Fleming, M., Steger, C., et al.: Techniques for a structural analysis of dermatoscopic imagery. *Computerized Medical Imaging and Graphics* 22(5), 375–389 (1998)
5. Fischer, S., Guilloid, J., et al.: Analysis of skin lesions with pigmented networks. In: *Proc. Int. Conf. Image Processing, Citeseer*, pp. 323–326 (1996)
6. Anantha, M., Moss, R., Stoecker, W.: Detection of pigment network in dermatoscopy images using texture analysis. *Computerized Medical Imaging and Graphics* 28(5), 225–234 (2004)
7. Betta, G., Di Leo, G., et al.: Dermoscopic image-analysis system: estimation of atypical pigment network and atypical vascular pattern. In: *MeMea 2006*, pp. 63–67. IEEE Computer Society, Los Alamitos (2006)
8. Di Leo, G., Liguori, C., Paolillo, A., Sommella, P.: An improved procedure for the automatic detection of dermoscopic structures in digital ELM images of skin lesions. In: *IEEE VECIMS*, pp. 190–194 (2008)
9. Shrestha, B., Bishop, J., et al.: Detection of atypical texture features in early malignant melanoma. *Skin Research and Technology* 16(1), 60–65 (2010)
10. Sadeghi, M., Razmara, M., Ester, M., Lee, T., Atkins, M.: Graph-based Pigment Network Detection in Skin Images. In: *Proc. of SPIE*, vol. 7623 (2010)
11. Zuiderveld, K.: Contrast limited adaptive histogram equalization. In: *Graphics gems IV*, pp. 474–485. Academic Press Professional, Inc., London (1994)
12. Wighton, P., Sadeghi, M., Lee, T., Atkins, M.: A fully automatic random walker segmentation for skin lesions in a supervised setting. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) *MICCAI 2009*. LNCS, vol. 5762, pp. 1108–1115. Springer, Heidelberg (2009)
13. Argenziano, G., Soyer, H., et al.: *Interactive atlas of dermoscopy*. EDRA-Medical Publishing and New Media, Milan (2000)
14. Soyer, H., Argenziano, G., et al.: Dermoscopy of pigmented skin lesions. In: *An Atlas Based on the Consensus Net Meeting on Dermoscopy 2000*, Edra, Milan (2001)
15. Haralick, R., Dinstein, I., Shanmugam, K.: Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics* 3(6), 610–621 (1973)
16. Goebel, M.: A survey of data mining and knowledge discovery software tools. *ACM SIGKDD Explorations Newsletter* 1(1), 20–33 (1999)