Automatic detection of basal cell carcinoma using vascular-extracted features from dermoscopy images

Hagar Maher

11th October 2018

1 Abstract

- Timely diagnosis of BCC is an important factor in the prognosis of the disease.
- A key to BCC diagnosis are vascular structures of the lesion. Detection and recognition of cutaneous vasculature provide critical information on diagnosis accuracy and assessment precision.
- In this paper, they present an effective method to extract vascular information towards lesion diagnosis. Given a dermoscopy image, we first segment vascular structures of the lesion by decomposing the image using independent component analysis into melanin and hemoglobin components and further applying shape filters at different scales. A vessel mask is generated as a result of global thresholding. A set of vascular features are then extracted from the final vessel image of the lesion and fed into a Random Forest classifier.
- The method demonstrates performance of 90.3% in terms of AUC in differentiating BCC from benign lesions

2 Introduction

- The presence of vascular structures in lesions, their morphology and architectural arrangement are specific signs for malignancies and certain skin abnormal conditions. Moreover, formation of new vessels could be an indicator of tumor development.
- The presence of telangiectasia (small dilated blood vessels near the skin surface) has been considered as a diagnostic indicator of Basal Cell Carcinoma (the most common type of skin cancer).
- Detection and quantification of cutaneous blood vessels provide critical information for diagnosis and assessment of skin conditions.
- Dermoscopy was primarily used to study pigmented skin lesions and there is very few studies on skin vasculature in dermoscopy.
- Visual inspection, as the only current technique for assessing vasculature, suffers from subjectivity and lack of precision. Moreover, these features are small and normally occluded by other structures which make the detection a challenging task.
- There are very few studies on the automated analysis of cutaneous vasculature. Most of these studies focus on detecting erythema by color processing.
- Blood vessel segmentation has also been addressed in retinal image analysis using texture features and supervised or semi supervised classification.
- There are major differences between the vessel segmentation problem in dermoscopic and retinal images.
- Retinal vessels are usually larger and hence more detectable than cutaneous vessels. Moreover, in retinal vessel analysis, the anatomy of the retina could be used as a priori information towards the segmentation problem, whereas in skin, the shape of the lesion varies from image to image.
- The presence of skin pigmentation and hair occludes vessel visibility and adds to the problem challenges.
- To address the specific challenges of vessel detection in human skin, in this paper they present a novel segmentation technique for accurate feature extraction of cutaneous blood vessels in dermoscopy images for BCC classification. The presented method is the first in the field that accounts for both the underlying color components of the skin and vascular shape. This generalizes the applicability of our method from erythema detection to vessel segmentation, compared to the existing methodologies. It results in extracting more accurate vascular features.

3 Method

3.1 Blood Vessel Segmentation

- A major challenge in skin vessel segmentation is that skin pigmentation occludes the visibility of vessels. As a solution, they propose an approach based on skin decomposition.
- Human skin has a layered structure, with different pigments being responsible for different color components.
- Among all skin pigments, melanin and hemoglobin are the most dominant with the latter being responsible for blood color.
- They used Independent Component Analysis (ICA) to decompose the skin color image into melanin and hemoglobin channels.

  They define a color vector as $I_{x,y} = [r_{x,y}, g_{x,y}, b_{x,y}]$ where $r_{x,y}, g_{x,y}, b_{x,y}$ are normalized color values of pixel (x, y) in RGB color space.
- Skin color density can then be modeled as

$$I_{x,y} = c^m q^m_{x,y} + c^h q^h_{x,y} + \Delta \qquad (1)$$

  where c^m and c^h are pure color per unit density of melanin and hemoglobin respectively; q^m and q^h are relative quantities of melanin and hemoglobin in each pixel (x,y); Δ is the stationary column vector caused by other skin pigments and structures.
- FastICA was applied to the input and relative quantities of melanin and hemoglobin and their pure color vectors were extracted as

$$\left[q^m_{x,y}, q^h_{x,y}\right] = \tilde{C}^{-1} I_{x,y} - \tilde{C}^{-1}\Delta$$
$$\acute{I}_{x,y} = \tilde{C}\left(K\left[q^m_{x,y}, q^h_{x,y}\right]^T + jE\right) + j\Delta \qquad (2)$$

  where $\widetilde{C}$ is the estimated pure color density vector; $I_{x,y}$ is the synthetized skin color.
- Setting j = 0, k = diag[1, 0], and k = diag[0, 1] respectively, extracts each color component of the skin.
- To learn reference color values for normal skin, pigmented skin and vessels, reference regions of the three clusters were outlined by an expert among 100 selected images and over 500000 pixels.
- For each cluster, mean and standard deviation of R, G and B channels of hemoglobin component were derived as the reference color vector.

- Using the reference values, a clustering framework is designed using the Mahalanobis distance on the hemoglobin component.
- Going through every pixel, the Mahalanobis distance from the reference color values of the three clusters (normal skin, pigmented skin and blood vessels) in hemoglobin component is calculated and the pixel is classified into the group with the closest distance.
- This step would cluster the erythematous areas from the normal skin and the rest of the lesion.

$$E^{i}_{x,y} = \sqrt{\frac{(r_{x,y}-r_i)^2}{\sigma_{ri}} + \frac{(g_{x,y}-g_i)^2}{\sigma_{gi}} + \frac{(b_{x,y}-b_i)^2}{\sigma_{bi}}} \qquad (3)$$

where $E^{n}_{x,y}, E^{p}_{x,y}, E^{v}_{x,y}$ are Mahalanobis distances of pixel (x, y) from normal skin, pigmented skin and blood vessels, respectively.

- If $E^{v}_{x,y} = min(E^{n}_{x,y}, E^{p}_{x,y}, E^{v}_{x,y})$ then pixel is classified as belonging to the erythematous cluster. The next phase of the method is shape filtering. Cutaneous vasculature seen in BCC are normally of tubular shape.
- They used frangi filter, in order to measure the tubularness of a pixel X = (x,y) at different scales $s^1, ..., s^k$

$$V(X,s) = \begin{cases} 0 & \lambda_2(X,s) < 0 \\ e^{\frac{-R^2(X,s)}{2\beta^2}}\left(1 - e^{\frac{-S^2}{c^2}}\right) & otherwise \end{cases} \qquad (4)$$

$$R = \frac{\lambda_1(X,s)}{\lambda_2(X,s)}\ ,\ S = \sqrt{\sum_{i=1}^{2} \lambda_i^2(X,s)} \qquad |\lambda_1| \le |\lambda_2|$$

where R, S are measures of blobness and second order structureness; $\lambda^i(X,s), i = 1, 2(|\ \lambda^1\ | \le |\ \lambda^2\ |)$ are the eigenvalues of the Hessian matrix of image calculated at scale s; β, c are control parameters.

- For each pixel, this function calculates the probability that the pixel belongs to a tubular structure.
- The Frangi filter is applied to the extracted red areas resulting from the previous step.
- Otsu's global thresholding is finally applied to segment the vascular structures and produce a binary vessel mask.

3.2 Feature Extraction Classification

- The endpoint of the analysis is automated disease classification to differentiate skin cancer (BCC) from benign lesions.
- For this purpose, using the vessel mask resulted from the segmentation step, they define and extract a set of 12 vascular features.
- Table 1 demonstrates the extracted features and their description.

TABLE I. VASCULAR FEATURES

Feature	Description
Max_Length	Maximum length of all vessel segments in the lesion
Mean_Length	Average length of all vessel segments in the lesion
S.D._Length	Standard deviation of length of all vessel segments in the lesion
Max_Width	Maximum width of all vessel segments in the lesion
Mean_Width	Average width of all vessel segments in the lesion
S.D._Width	Standard deviation of width of all vessel segments in the lesion
Max_Area	Maximum area of all vessel segments in the lesion
Mean_Area	Average area of all vessel segments in the lesion
Area_Ratio	Ratio of vessel area to lesion area
Num_Branch	Number of vascular branches
Branch_Ratio	Ratio of number of branches to lesion area

The above mentioned 12 features were fed as the input to a Random forest classifier of 100 trees, each constructed while considering 4 random features to perform a two-class classification of cancerous vs. benign lesions.

4 Result and Discussion

- Three different sources of data were used in this paper from Atlas of dermoscopy by Argenziano, University of Missouri and Vancouver Skin Care Centre.
- The proposed method was implemented on a dataset of 659 images among which 299 were BCCs and 360 were non-BCC.
- Matlab was used to implement the method.
- Segmentation of the lesions were provided by an expert.
- Figure 1 demonstrates the decomposition of two BCC lesions into melanin and hemoglobin channel.

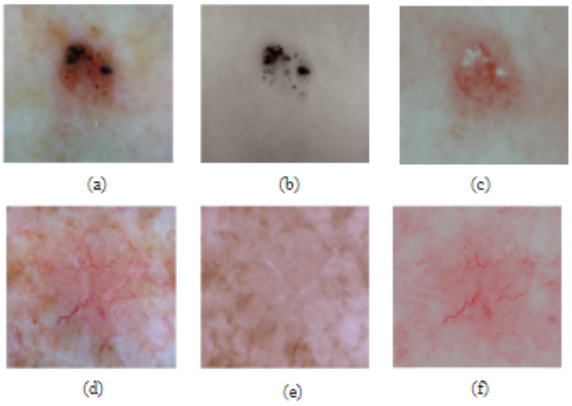

Fig. 1. (a) and (d) are original dermoscopy images of two BCC lesions.(b) and (e) are their corresponding decomposition Melanin channel and (c) and (f) hemoglobin channel

- Figure 2 shows the results of segmentation of the red areas of the lesions using the Mahalanobis distance clustering of the hemoglobin channel.

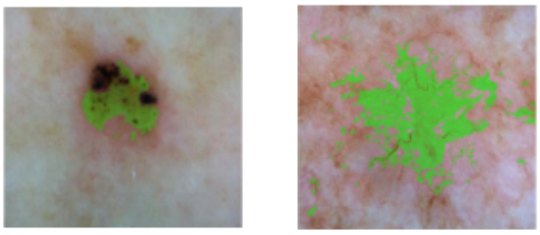

Fig. 2. Segmented red areas of the two previous lesions based on clustering of the hemoglobin channel

- Figure 3 demonstrates two shape probability maps applied on the segmented red areas of the two basal cell carcinomas.

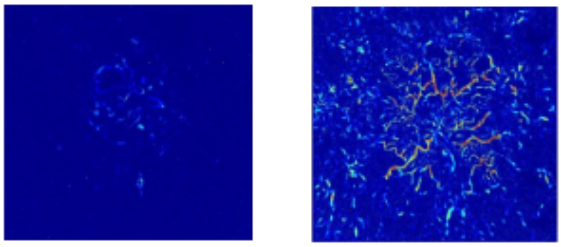

Fig. 3. Shape probability maps

- Figure 4 demonstrates the final vessel mask of the two previous BCC lesions, after Otsu's thresholding, overlaid on the original image.

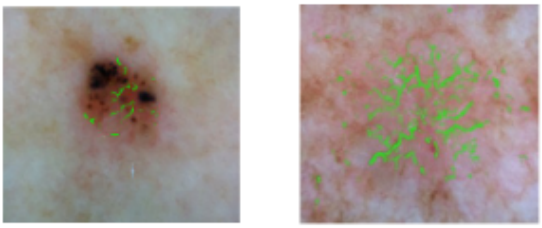

Fig. 4. Vessel mask overlay on the lesions

- Ground truth for the location of vessels was obtained by an expert manual segmentation of cutaneous vasculature.
- Segmentation performance results were tested on 500000 pixels was 89% and 86% in terms of sensitivity and specificity respectively.
- 12 Vascular features were extracted from the vessel mask and fed into a random forest classifier.
- The classification task was performed using Weka. Total time for the classification was 0.44 seconds.

- Table 2 demonstrates the statistics of the classification task for the two classes.

TABLE II. PERFORMANCE MEASURES

	TP Rate	FP Rate	Precision	Recall	ROC Area
BCC	0.736	0.103	0.856	0.736	0.903
Non-BCC	0.897	0.264	0.803	0.897	0.903
Weighted Average	0.824	0.191	0.827	0.824	0.903

- The performance of classification is 90.3% in terms of the area under the ROC curve as demonstrated by figure 5.

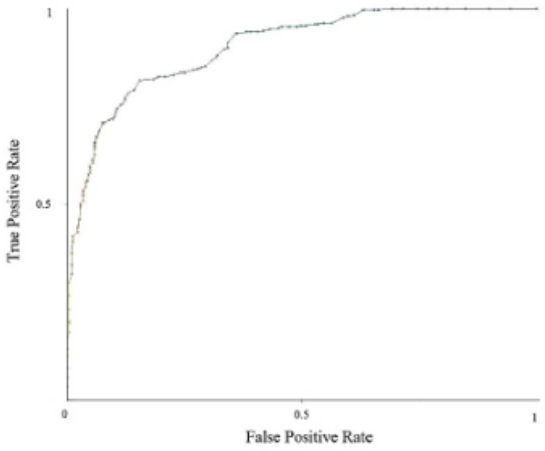

Fig. 5. ROC curve of BCC classification

5 Conclusion

- Compared to previous studies, this study accounts for both shape and color information.
- The segmentation method has been used as the base for vascular feature extraction towards BCC classification.
- Experimental results show an area under the ROC curve of 90.3% for differentiating BCC vs. benign lesions in a set of 659 dermoscopy images.
- This technique may be applied to quantification of skin vasculature in a variety of applications.

6 Acknowledgment

- This work was partly supported by Canadian Institute of Health Research and Skin Research and Training Centre. The authors would like to thank Dr. Mohammed AlJasser, Dr. William V. Stoecker and Dr. Emre Celebi for providing parts of the data.

- 978-1-4673-8721-7/16/$31.00 ©2016 IEEE

- 2016 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)

7 Summary

7.1 Motivation

Visual inspection, as the only current technique for assessing vasculature, suffers from subjectivity and lack of precision. Moreover, these features are small and normally occluded by other structures which make the detection a challenging task. Also retinal vessels are usually larger and hence more detectable than cutaneous vessels. Moreover, in retinal vessel analysis, the anatomy of the retina could be used as a priori information towards the segmentation problem, whereas in skin, the shape of the lesion varies from image to image.

7.2 Dataset

Three different sources of data were used in this paper from Atlas of dermoscopy by Argenziano, University of Missouri and Vancouver Skin Care Centre. The proposed method was implemented on a dataset of 659 images among which 299 were BCCs and 360 were non-BCC.

7.3 Framework

They present an effective method to extract vascular information towards lesion diagnosis. Given a dermoscopy image, they first segment vascular structures of the lesion by decomposing the image using independent component analysis into melanin and hemoglobin components and further applying shape filters at different scales. A vessel mask is generated as a result of global thresholding. A set of vascular features are then extracted from the final vessel image of the lesion and fed into a Random Forest classifier.

7.4 Challenge

A major challenge in skin vessel segmentation is that skin pigmentation occludes the visibility of vessels. As a solution, they propose an approach based on skin decomposition.

7.5 Results

The method demonstrates performance of 90.3% in terms of AUC in differentiating BCC from benign lesions.