# CAPILLARY DETECTION FOR CLINICAL IMAGES OF BASAL CELL CARCINOMA

*Adam Huang[1,2,3], Wen-Yu Chang[4,5], Hsin-Yi Liu[1,2], and Gwo-Shing Chen[6]*

[1]Research Center for Adaptive Data Analysis/[2]Center for Dynamical Biomarkers and Translational Medicine/[3]Graduate Institute of Biomedical Engineering, National Central University, Jhongli, Taiwan
[4]E-Da Hospital, I-Shou University, Kaohsiung, Taiwan
[5]Graduate Institute of Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan
[6]Department of Dermatology, Kaohsiung Medical University Hospital, Kaohsiung, Taiwan

## ABSTRACT

Dilated capillaries are an important characteristic of basal cell carcinoma (BCC). Detecting capillaries in images can improve a computer-aided skin cancer diagnosis system. In this study, we investigate the feasibility to extract capillaries from clinical images of skin lesions recorded by a regular digital camera. First, we used a compact set of 1 curvilinear and 2 color parameters to train a support vector machine (SVM) classifier to identify capillary pixels. Second, the identified pixels were grouped by a region-growing algorithm to form capillary candidates. Last, the likelihood to be a true capillary was estimated based on the distance to the red color in the "CIE Lab" color space. The method was tested on a dataset of 21 BCC images with visible capillaries and 28 benign pigmented lesions without visible capillaries. The accuracy, sensitivity, and specificity of the proposed method were 89.8% (44/49), 90.5% (19/21), and 89.3% (25/28) respectively. We found capillaries recorded by a regular digital camera can be detected successfully.

***Index Terms***— Computer-aided detection, basal cell carcinoma, vascular pattern detection, color segmentation, support vector machines

## 1. INTRODUCTION

Basal cell carcinoma (BCC) is a common malignant skin cancer with locally invasive properties. BCC characteristic features including arborizing vessels, ulceration, large blue-gray ovoid nests, leaflike areas, multiple blue-gray globules and spoke-wheel areas [1]. The presence of certain capillary patterns is crucial for clinical diagnosis of BCC. Arborizing vessels (Fig. 1), which are characterized by focused vessels with different diameters and tree-like branches, have been reported with a presence frequency of 60.7% [1] of BCCs and have positive predictive value of 94.1% [2], [3]. Automatic detection of arborizing capillaries can potentially improve a computer-aided diagnosis system for skin cancer.

Although the vascular patterns in BCC are visible to the naked eye and can be captured by regular digital photography, the majority of recent studies in vascular patterns of BCCs have been conducted by using dermoscopy [1]-[3]. Dermoscopy is a new technique to improve diagnostic accuracy of pigmented skin lesions because it allows in vivo evaluation of colors and microstructures not visible to the naked eye. As dermoscopy has become a popular modality in diagnosing pigmented skin lesions, fewer new studies have been devoted to improve the diagnostic performance of a computer-aided diagnosis (CAD) system that processes more easily accessible clinical images taken by regular digital cameras.

In our hospital, clinical images are constantly taken for observation and recording purposes. Therefore, combining information extracted from both dermoscopic and clinical images could possibly improve a CAD system's accuracy. Further, a specialized CAD system devoted to analyze clinical images has the potential to improve health care quality for rural hospitals that lack of dermatological specialists and dermoscopic equipments.

In this study, we investigate the feasibility of detecting capillaries in skin lesions by using regular digital photographic images that are more accessible in most hospitals.

## 2. MATERIALS AND METHODS

### 2.1. Data Acquisition

The digital pictures of skin lesions used in this study were selected from a skin lesion database from the Kaohsiung Medical University Hospital in Taiwan. In this database, a suspicious region was photographed by a dermatologist with a Nikon D70 6.1 megapixel digital single-lens reflex (SLR) camera and sigma 18-50mm F2.8 macro lens for recording purposes before biopsy. Biopsy was performed and histopathology result was available for each lesion.

From the database, a training dataset of 25 BCC images with visible, various capillary and pigment patterns were selected to train the proposed capillary pixel identification algorithm. A separate dataset of 21 BCC images with visible capillaries and 28 (15 nevi and 13 seborrheic

keratosis) benign pigmented lesions without visible capillaries were selected to test the proposed capillary detection algorithm.

### 2.2. Capillary Pixel Identification

Capillary has two main visual characteristics, reddish color and curvilinear, tree-like branching shapes. However, regular photographic images may contain reflections, shadows, and artifacts caused by various photo shooting conditions. Fully automatic capillary detection remains a difficult task. In addition, different shades of skin color and various capillary patterns have made a decision-tree type of algorithm very difficult to construct and optimize. In order to tackle this complex problem, we took a machine learning approach in the very beginning stage of the proposed method. We first identified and extracted the most notable features of capillary from a diverse training dataset of various capillary images, and then we trained a support vector machine (SVM) classifier to identify capillary pixels.

*2.2.1. CIELab color features*

We first convert the image from the RGB color space to the CIELab [4] color space. The reason for this is to reduce the memory requirement and improve computation time. As a skin lesion generally has tens or hundreds of thousands of pixels in a digital image, it may take several minutes to compute the SVM score of all pixels in the image. Instead, using a precomputed table of SVM scores will cut down the computation time significantly. However, using RGB and curvilinear characteristics requires a matrix of 4 dimensions stored in a large memory space. To resolve this dilemma, the CIELab system was chosen because the color components are represented in two dimensions (a and b) only. Another advantage of CIELab is that the Lab color is designed to approximate human vision. Its L component closely matches human perception of lightness. The L component in Fig 1 is shown in Fig 2, which is the common form of grayscale image for a color image. Further, the perceptual deference between colors is proportional to the Cartesian distance in the CIELab color space.

*2.2.2. Curvilinear shape features*

Observed from Fig 2, capillaries appear like dark lines and can be extracted by using common linear structure extraction methods for grayscale images. Ridge/trench-shaped linear structures are generally processed by using the Hessian matrix [5]-[7] or matched filters [8], [9].

From our earlier unpublished study on skin clinical images, we found that a multiscale Hessian matrix approach [7] generally created noisier results than a multiscale matched filters approach. However, the original matched filters method proposed in [8] requires 12 templates in all directions. In order to improve computation efficiency, we use only one matched filter template in the ridge direction, which can derived from the eigenvectors of 2D Hessian:

$$\mathbf{H} = \begin{bmatrix} F_{xx} & F_{xy} \\ F_{xy} & F_{yy} \end{bmatrix} \quad (1)$$

where $F$ is a grayscale intensity function $F(x, y)$. In the framework of multiscale approaches, differentiation is defined as a normalized convolution with derivatives of Gaussians:

$$F_{x^i y^j}(x, y; \sigma_f) = \left\{ \sigma_f^{i+j} \frac{\partial^{i+j}}{\partial x^i \partial y^j} G(x, y; \sigma_f) \right\} * F(x, y) \quad (2)$$

where G is an isotropic 2D Gaussian function with standard deviation $\sigma_f$ and $\sigma_f^{i+j}$ is the normalization factor.

The ridge-shaped matched filter is defined as

$$F_M(x, y; \sigma_f) = \exp\left( \frac{[(P - P_0) \cdot \mathbf{v}]^2}{2\sigma_f^2} \right) \quad (3)$$

where $P$ is any point $(x,y)$, $P_0$ is the center of the filter, and $\mathbf{v}$ is the eigenvector along the ridge direction derived from (1). The detection of a ridge-shaped structure from a grayscale image $I$ at a pixel location $(x, y)$ by using the filter defined in (3) has an upper limited value using Schwartz's inequality [9]

$$\left( \sum_{i=-w}^{w} \sum_{j=-w}^{w} I(x+i, y+j) \times F_M(x+i, y+j) \right)^2 \le \left( \sum_{i=-w}^{w} \sum_{j=-w}^{w} (I(x+i, y+j))^2 \right) \times \left( \sum_{i=-w}^{w} \sum_{j=-w}^{w} (F_M(x+i, y+j))^2 \right)$$

where $w = 3\sigma_f$ for the filter window size. Rearrange the above formula and let Z denote the filter response at $(x, y)$ by a matched filter

$$Z(x, y; \sigma_f) = \frac{\left( \sum\sum I \times F_M \right)^2}{\left( \sum\sum I^2 \right)\left( \sum\sum F_M^2 \right)} \times 255 \quad (4)$$

so that $0 \le Z \le 255$, which is suitable to be discretized into 8-bit precision number for precomputed table indexing.

Our multiscale curvilinear shape filtering approach repeats the computation (1)-(4) using a set of 4 discrete standard deviation values $\sigma_f \in \{1, \sqrt{2}, 2, 2\sqrt{2}\}$ for each pixel and the maximum response $Z(x, y; \sigma_f)$ is collected. Fig 3 illustrates the resultant curvilinear filtering response image.

*2.2.3. SVM classifier*

Capillaries in the training dataset of 25 BCC images were identified by a dermatologist. A sample of 6015 pixels from the identified capillaries and 54048 pixels from the non-capillary areas were selected manually by a technician. Color components (a and b in the CIELab color space) and maximum curvilinear matched filtering responses for the sample pixels were computed and used to train a SVM classifier, LIBSVM, which is available freely on the web [10]. Once the classifier was trained, the SVM scores of all possible combinations of 8-bit precision color components a,

b values, and matched filter responses were computed by the trained classifier to build the precomputed SVM score table, a 256x256x256 matrix.

### 2.3. Region Growing

Once the curvilinear responses and color a, b components in an image were computed, the SVM score for each pixel was obtained from the precomputed table. (The SVM scores for Fig 1 is shown in Fig 4) The top 2% of pixels with highest SVM scores were selected to form capillary clusters. Clusters with a population lower than 100 pixels were removed. The operating parameters of 2% and 100 pixels were chosen empirically. Last, the survived clusters were dilated by one pixel to form the final capillary candidates (shown as green areas in Fig 5).

### 2.4. Likelihood Estimation

In this step, the likelihood of a candidate to be a true capillary was estimated based on the mean distance to the red color for all the color pixels within the capillary candidate. (The red color is (255, 0, 0) in the RGB color space, and (138, 209, 198) in the CIELab color space.) Finally, the likelihood of an image with visible capillaries was estimated by averaging the mean distances to the red color from the top three capillary candidates within the image.

Three distance computations were tested in this study: 1) Cartesian distance in the 3D CIELab space, 2) Cartesian distance in the 2D space of color components 'a' and 'b', and 3) difference in color component 'a' only. We used the "area under the ROC curve" (short noted as AUC) to compare the performance of three different distance computations.

## 3. RESULTS

The test results for three distance computations were illustrated in Fig 6. The better ROC curve (blue curve in Fig 6), with a higher "true positive rate" and a lower "false positive rate", is located on the upper left side. The AUC values of 3D Lab space (red curve), 2D a-b space of (green curve), and 1D along the 'a' direction were 0.863, 0.877, and 0.923 respectively. The distance computed on the color axis 'a' in the Lab space performed significantly better than the other two approaches. The accuracy, sensitivity, and specificity of the blue curve at the upper left corner were 89.8% (44/49), 90.5% (19/21), and 89.3% (25/28) respectively.

## 4. CONCLUSIONS

We found capillaries in clinical images recorded by a regular digital camera can be accurately detected by a SVM classifier using a compact set of 3 parameters (2 Lab color components and one matched filtering response). This compact set allows us to build a precomputed SVM scoring table to shorten computation time for processing a high resolution digital image.

## ACKNOWLEDGMENTS

This research was supported by the National Science Council of Republic of China (Taiwan) under Grants NSC 100-2221-E-008-009 and NSC 99-2911-I-008-100.

## REFERENCES

[1] T. Micantonio, A. Gulia, E. Altobelli, A. Di Cesare, R. Fidanza, A. Riitano, M.C. Fargnoli, and K. Peris, "Vascular Patterns in Basal Cell Carcinoma," Journal of the European Academy of Dermatology and Venereology, vol.25, pp. 358-361, 2011.

[2] S.W. Menzies, K. Westerhoff, H. Rabinovitz et al, "Surface Microscopy of Pigmented Basal Cell Carcinoma," Arch Dermatol, vol. 136, pp. 1012-1016, 2000.

[3] G. Argenziano, I. Zalaudek, R. Corona et al, "Vascular Structures in Skin Tumors: A Dermoscopy Study," Arch Dermatol, vol. 140, pp. 1485-1489, 2004.

[4] Schwiegerling, J., *Field Guide to Visual and Ophthalmic Optics*, SPIE Press, Bellingham, WA, 2004.

[5] S. Aylward, S.M. Pizer, E. Bullitt, and D. Eberly, "Intensity Ridge and Widths for 3D Object Segmentation and Description," Proc. IEEE Workshop Math. Methods in Biomedical Image Analysis, pp. 131-138, 1996.

[6] J. Staal, M.D. Abramoff, M. Niemeijier, M.A. Viergever, and B. van Ginneken, "Ridge-Based Vessel Segmentation in Color Images of the Retina," IEEE Transactions on Medical Imaging, vol. 23, pp. 501-509, 2004.

[7] A. Huang, G.M. Nielson, A. Razdan, G.E. Farin, D.P. Baluch, and D.G. Capco, "Thin Structure Segmentation and Visualization in Three-Dimensional Biomedical Images: A Shape-Based Approach," IEEE Transactions on Visualization and Computer Graphics, vol. 12, pp. 93-102, 2006.

[8] S. Chaudhuri, S. Chatterjee, N. Katz, M. Nelson, and M. Goldbaum, "Detection of blood vessels in retinal images using two-dimensional matched filters," IEEE Transactions on Medical Imaging, vol. 8, pp. 263-269, 1989.

[9] A. Huang, G. Nielson, A. Razdan, G. Farin, D. Capco, and P. Baluch, "Line and Net Pattern Segmentation Using Shape Modeling," Proc. SPIE Electronic Imaging Symp. Visualization and Data Analysis, 2003.

[10] C.C. Chang and C.J. Lin, "LIBSVM: A Library for Support Vector Machines," ACM Transactions on Intelligent Systems and Technology, vol. 2, pp. 27:1-27:27, 2011. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm



Fig 1. Arborizing capillaries in a clinical image of BCC. They appear like curvilinear, branching structures (blue arrows).

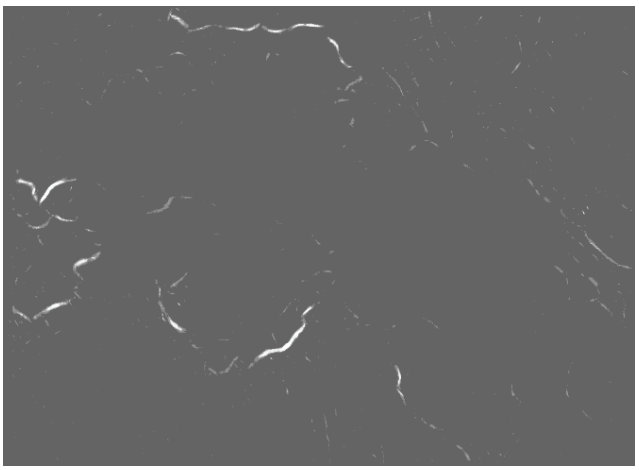


Fig 4. The SVM scores (higher in white) to be true capillary pixels estimated by a SVM classifier using matched filter response and CIE Lab color components a, b.

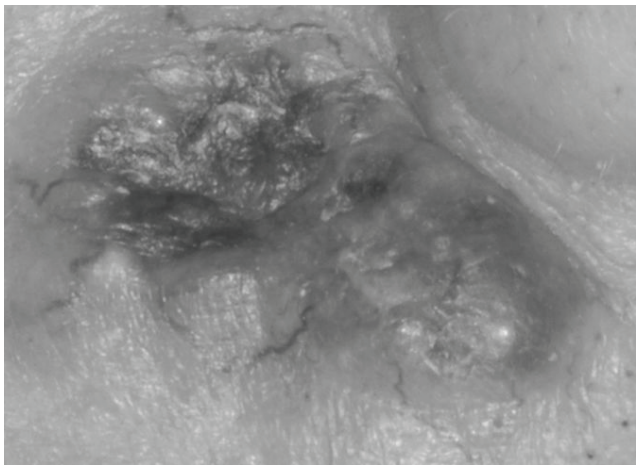


Fig 2. Capillaries appear like dark lines in a grayscale image.

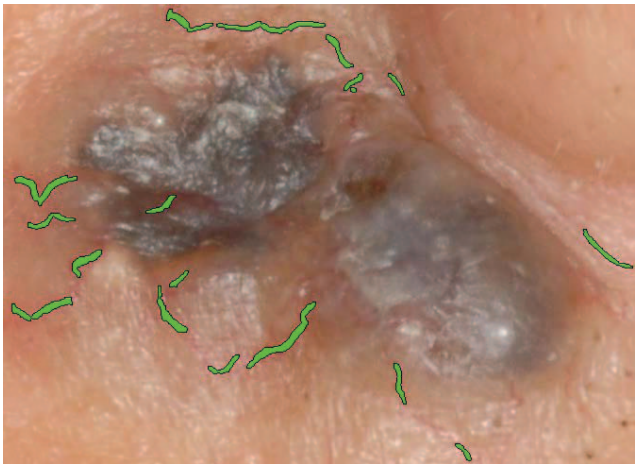


Fig 5. Extracted capillaries colored in green.

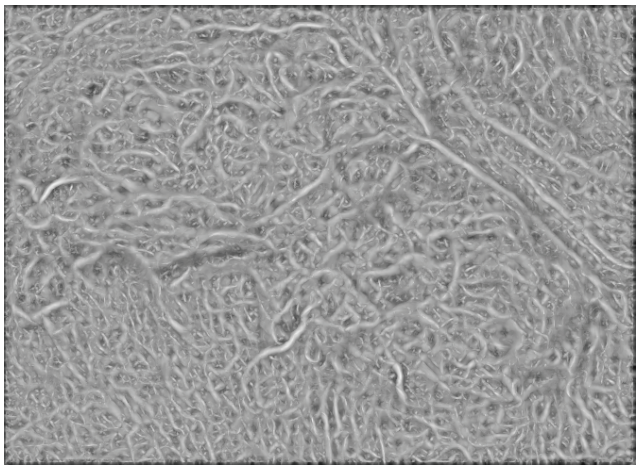


Fig 3. Matched filter responses of the grayscale image shown in Fig 2.

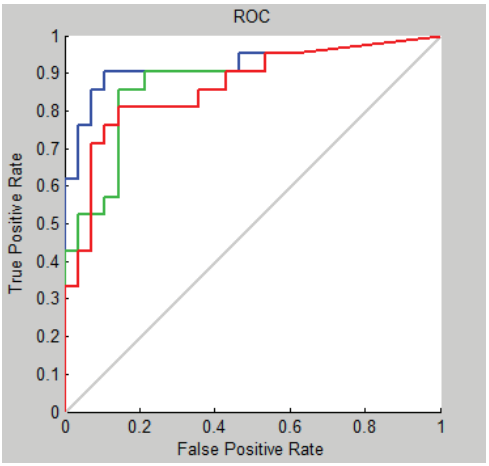



Fig 6. ROC curves: distance computed in the Lab space (red), in the 'a' and 'b' space (green), and in the 'a' axis (blue).