Proceedings of the 28th IEEE
EMBS Annual International Conference
New York City, USA, Aug 30-Sept 3, 2006

SVM-based Texture Classification and Application to Early Melanoma Detection

Xiaojing Yuan, Member IEEE, Zhenyu Yang, George Zouridakis*, Senior Member IEEE, and Nizar Mullani

Abstract — We have recently developed a decision support system for early skin cancer detection that relies on analysis of the pigmentation characteristics of a skin lesion, detected using crosspolarization imaging, and the increased vasculature associated with malignant lesions that is detected using transillumination imaging. Current system uses size difference based on lesion physiology and achieves great overall accuracy (86.9%). In this paper, we explore texture information, one of the criteria dermatologists use in the diagnosis of skin cancer, but has been found very difficult to utilize in an automatic manner. The overarching goal is to improve the overall decision support capability of the DSS. The objective is to use texture information ONLY to classify the benign and malignancy of the skin lesion. A three-layer mechanism that inherent to the support vector machine (SVM) methodology is employed to improve the generalization error rate and the computational efficiency. The performance of the algorithm is validated with a series of benchmark texture images and then tested on 22 pairs of real clinical skin lesion images. Our experimental results show that a 4th-order polynomial kernel can reach an average accuracy of 70% in determining the malignancy of any pixel within any given skin lesion image. Further study will look at whether multi-channel filtering based feature extraction algorithm will improve the accuracy rate, and the performance comparison between SVM-based texture classification and decision tree-based texture classification in both the spatial and frequency domain.

I. INTRODUCTION

A. Early Melanoma Detection

Skin cancer is the most common one and represents 50% of all new cancers detected each year. In about 1.2 million new skin cancer cases detected in the U.S. every year, melanoma is the deadliest one, accounting for over 7300 deaths per year in the U.S. However, if detected at an early stage, skin cancer has a very high cure rates, and requires rather simple and economical treatments. On the contrary, when detected at a late stage, cancerous lesions have very high morbidity and mortality rates, and an extremely high cost associated with accurate diagnosis and the necessary treatment.

We have recently developed an integrated early skin-cancer detection system[1] which was based on automated image analysis and on principles of cancer physiology. It showed an 86.9% overall accuracy in correctly classifying a skin lesion as benign or malignant. More specifically, we analyzed statistical characteristics of the pigmentation area, which was obtained using surface crosspolarization epiluminescence (XLM) imaging, and we also estimated the increased vasculature typically associated with a malignant lesion, using side transillumination (TLM) [2] imaging of the same lesion.

Pigmented skin lesions are typically evaluated by dermatologists using the "ABCD" rule; which analyze the Asymmetry, Border irregularity, Color variation and Diameter of a lesion. Practically, rule A, B, and C analyze texture information, and this confirms the importance of texture and its value in skin cancer diagnosis, as has been recently summarized in two review papers [3,4]. With regard to texture features, skin lesions have been classified by physicians manually into three classes as shown in Fig.1, namely, reticular, globular, and homogeneous.

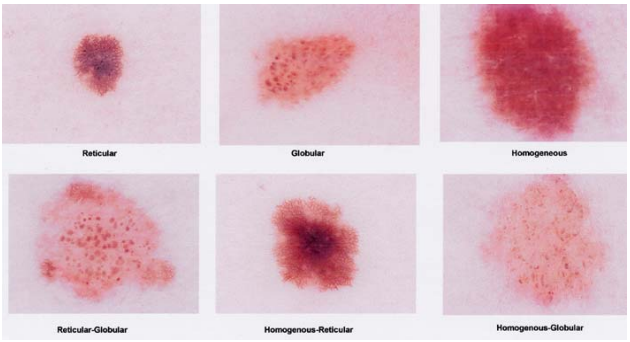


Fig.1 Texture characteristics of skin lesions [3,4].

Lesions having only one or two of the above features are found to be benign with a probability of 90% and 70%, respectively, while lesions possessing all three features have a 90% probability of being malignant and should definitely be excised. In other cases, whereby only partial features are detected, lesions have a 70% probability of being benign. Despite the high diagnostic value of texture information, it is very difficult to utilize it in automated skin cancer detection due to complexity of the texture and high noise and interference of the obtained lesion images. In this paper, we describe a recently developed SVM-based texture identification and classification module and evaluate the performance of it when applying to 22 pairs of real clinical skin lesion images. The performance is very satisfying considering the very simple texture feature used and simple kernel functions selected for SVM. Further comprehensive study is under way to identify the best combination of feature extraction and classification system.

B. Texture Analysis

Texture information has always been an important and efficient measure to estimate the structure, orientation, roughness, smoothness, or regularity of various regions in a set of images that enables us to distinguish between different

Manuscript received April 3rd, 2006.

This work was supported in part by the Higher Education Assistant Fund and Environmental Institute of Houston and by TransLite, LLC. Z. Yang (zhenyu@cs.uh.edu), X. Yuan (xyuan@uh.edu), and G. Zouridakis (zouridakis@uh.edu) are with University of Houston, Houston, Tx, 77204. N. Mullani is with Translite LLC, Sugar Land, TX.

objects. Even though the human visual system can easily classify different textures in a real world view, it is difficult to quantitatively characterize it. Hence, texture analysis is regarded as one of the most difficult challenges and one of the least utilized tools in computer vision and image analysis. Image texture can be broadly defined as a certain pattern repeated in a local area of an image [5]. It is generally agreed that texture exhibits periodicity of some basic patterns. The texture identification and classification have been proved to be critical in image segmentation, content-based image retrieval and MRI-, or X-ray- based diagnosis and prognosis.

Texture analysis is usually performed in two steps, as Fig. 2 shows, namely, (i) texture feature extraction and (ii) texture identification and classification. Common texture extraction methods include statistical, model-based, and filtering-based methods, among which multi-channel filtering is the most efficient and accurate one. Some of the typically used filter banks are Laws masks, the dyadic Gabor filter bank, and wavelet transforms [6], while common classification methods used for texture identification and classification include neural networks, decision trees, and support vector machines.

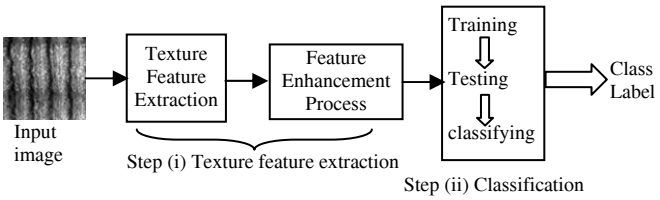


Fig. 2 Steps for traditional texture classification

In the texture feature extraction step, the input image is fed into a feature extraction algorithm (e.g., local Auto-Regression, or Gabor Filter Banks) and a set of texture feature vectors are generated. Then, a proper nonlinearity transformation, smoothing, and energy computation are applied on the filter responses for feature enhancement. The resulting images are the sought-after feature images. The smoothing function averages the texture features in a local area, usually a Gaussian window. It is obvious that a feature of a pixel is the energy on a specific frequency channel *f* and orientation together with local pixels. Unlike the pixels in the original image, the feature vector generated for each pixel tells the complex connection with neighboring pixels.

In the texture classification step, a classifier is trained and tested, and then used to classify a new filter response of a textured image. The supervised classification step includes three stages, namely, training, testing, and classification. At the training stage, a classifier is trained on a set of training feature vectors with an assigned label. At the testing stage, the classifier is applied to a set of test feature vectors that have been reserved for testing purposes (and are different from training feature vectors), and performance of the classifier is evaluated. At the classification stage, after a satisfactory classifier is built, it is used to classify any new instances coming in.

C. *Support Vector Machine*

Recently, significant progress has been made in statistical learning theory and machine learning. The problem of learning is much better understood thanks for the work of several researchers such as Vapnik [7,8], and algorithms with high classification accuracy such as SVMs have been developed and successfully applied to many problems [9,10]. Traditional approaches from statistics failed to address the problem of "the curse of dimensionality". However, SVM tries to minimize the empirical error while controlling the complexity of the mapping function. This enables it to achieve much better generalization performance on new data. We review the support vector machines briefly in this section.

Given a two class data set $(x_i, y_i)_{i=1}^{l}$, a SVM tries to find a hyper-plane which can separate the two classes of data well. Figure 3 shows two class data in a 2-D space. SVM achieves maximization of the margin ($\frac{1}{\|w\|^2}$) by solving

$$\max \frac{1}{\|w\|^2}$$

The

$$subject \quad to \quad y_i(w^T \cdot x_i + b) \geq 1 \quad for \quad all \quad i = 1,...,l$$

decision rule $f: x - y$ is then $f(x) = \text{sgn}(\sum_{i=1}^{l} y_i w_i \cdot x + b)$

The problem is usually solved in a dual form

$$\max W(\alpha) = \sum_{i=1}^{l} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{l} \alpha_i \alpha_j y_i y_j (x_i \cdot x_j)$$

$$subject \quad to: \sum_{i=1}^{l} \alpha_i y_i = 0, \quad 0 \leq \alpha_i \leq \frac{C}{l}, for \quad i = 1,...,l$$

and the decision rule is then

$$f(x) = \text{sgn}(\sum_{i=1}^{l} y_i \alpha_i x_i \cdot x + b)$$

The data x_i with nonzero $_i$ is called a support vector.

Often the data in input space is not linearly separable, then the data are first mapped to a high dimensional feature space by mapping function $\phi: x - \phi(x)$, and then the SVM solves

$$\max W(\alpha) = \sum_{i=1}^{l} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{l} \alpha_i \alpha_j y_i y_j (\phi(x_i) \cdot \phi(x_j))$$

$$subject \quad to: \sum_{i=1}^{l} \alpha_i y_i = 0, \quad 0 \leq \alpha_i \leq \frac{C}{l}, for \quad i = 1,...,l$$

and the decision rule is $f(x) = \text{sgn}(\sum_{i=1}^{l} y_i \alpha_i \phi(x_i) \cdot \phi(x_j) + b)$

The feature space induced by is usually high dimension and the direct computation is usually impossible. However, since in SVM everything is dot product, a kernel function k(x,x) is used. If the symmetric kernel satisfies Mercer's theorem, then there is a mapping such that

$$k(x, \tilde{x}) = \phi(x) \cdot \phi(\tilde{x}).$$

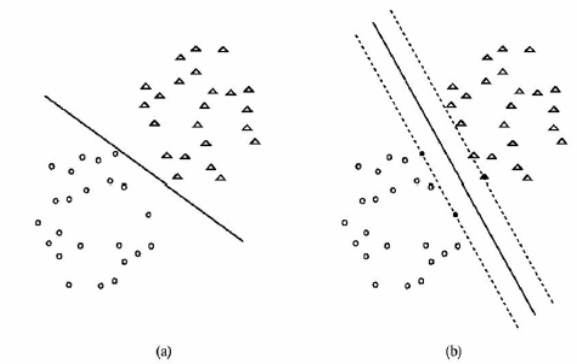
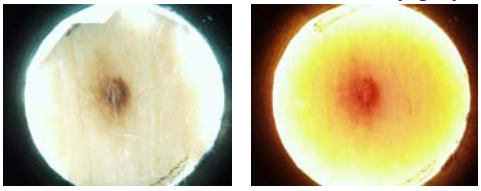


Fig.3 SVM tries to find hyper-plane to separate data and have maximal margin. (a) the hyper-plane separates the data but with small margin; (b)the hyper-plane found by SVM achieves maximum margin.

D. Subjects and Procedure

The skin lesion images used in this study (samples are shown in Fig. 4) were taken using a Nevoscope device capable of obtaining both TLM and XLM images [13, 14]. The device uses an optical lens (Nikon, Japan) to achieve a standard 5X magnification and an Olympus C2500 digital camera for capturing the images. Original images have a spatial resolution of 1712 x 1368 pixels. For each lesion two distinct images are acquired in the TLM and the XLM modality. Clinical imaging is carried out at the University of Texas in Houston under the direction of a board certified physician. Full IRB approvals are obtained for all studies. A total of 60 lesions were imaged from consecutively and prospectively enrolled clinic patients undergoing routine skin exams, who had clinically suspicious skin lesions of less than 1*cm* in size. To avoid clinician bias, all dermatology patients were considered for participation in the study regardless of risk of melanoma or past history.

In our experiment, 22 pairs of images that are pathology-proven cases with non-disputable class label (including 13 benign lesions and 9 malignant ones) are used. Each image, whether TLM or XLM, undergoes some pre-processing steps before the lesion images were fed into the SVM-based texture identification system. The detailed pre-processing steps have been presented recently [14]. It includes image resizing, masking, cropping, hair removal (or attenuation), and conversion from RGB color to intensity grey image.



(a) XLM (b) TLM

Fig. 4 Sample skin lesion image pair

II. SVM-BASED TEXTURE IDENTIFICATION AND CLASSIFICATION FOR EARLY MELANOMA DETECTION

Recently, SVM has been applied to texture classification [11]. There are three layers in the SVM-based texture classification algorithm (Fig. 5). The input layer accepts gray-level values of the M x M window defined in the image. Instead of using all the pixels in the image AR features [5] is used. This reduces the size of the feature vector from M^2 to 4M-3, resulting in an improved generalization performance and classification speed. The hidden layer applies a nonlinear transformation from the input spatial imaging space to the feature space F, where the inner products are computed and becomes linear separable. The size of the hidden layer is determined as the number of SVs identified by the SVM. Since there are usually much fewer SVs than the size of the input feature vectors, the size of the hidden layer is moderate. The output layer is a hyper plane classifier, and the output of the SVM represents the class of the central pixel in the input window.

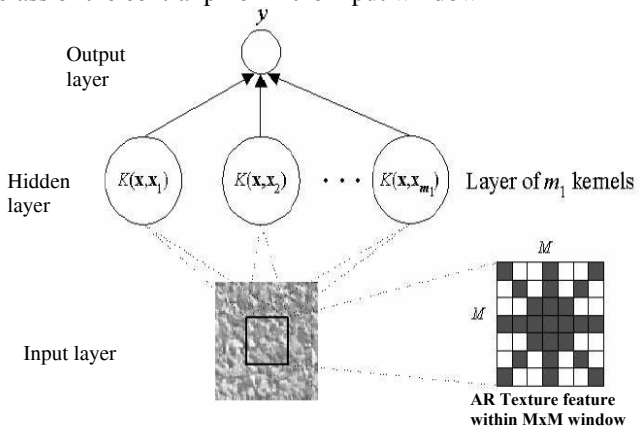


Fig. 5 SVM-based Texture Classification

III. EXERPIEMNT RESULTS AND ANALYSIS

A binary-class benchmark classification problem is used as an example to validate our algorithm (Fig.6). SVMs with a varying polynomial order *p* were trained with 2,000 training examples obtained by randomly selecting 1000 samples from both classes. The input window size was progressively to 9x9, 11x11, 13x13, 15x15, 17x17, and 19x19. The testing errors for different window sizes and degrees are shown in Fig.6. The error rate increase after 4^{th} degree may due to over-fitting. From these results, the optimal degree for the polynomial kernel is selected as 4, while the optimal window size is 17x17. The set of experiments validate our implementation of support vector machine algorithm.

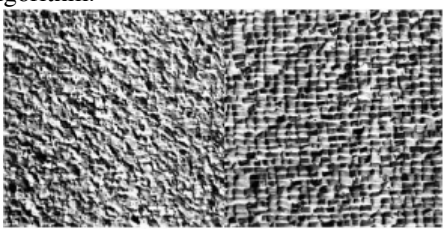


Fig. 6 Benchmark binary texture classification problem. Two textures are D4 and D84 from [12].

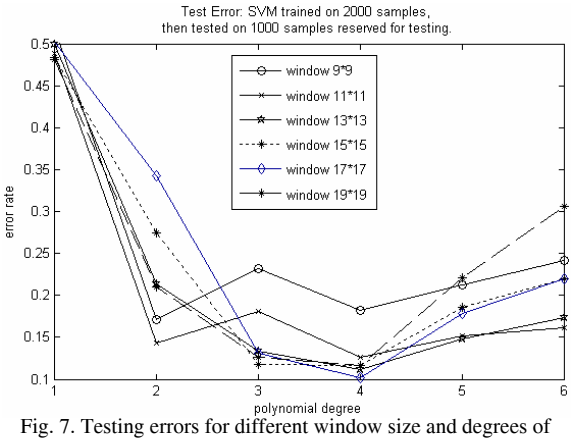


Fig. 7. Testing errors for different window size and degrees of polynomial kernels

Several tests have been performed for the 22 pairs of skin lesion images. One and two hundred feature vectors are extracted from each of them, within which 20% are reserved as testing samples, while the remaining data set are used for training. The average accuracy from 10 experiment (to reduce the variance from different experiments) for 100 feature vectors is about 60%, while the accuracy for 200 feature vectors is about 70% for determining the benign and malignancy of any given images, using 35x35 window and 4^{th}-degree polynomial kernel. The 35x35 window size is chosen after performance evaluation with window size progressively fixed at 21x21, 25x25, 29x29, 35x35, and 39x39. The texture information in the skin lesion images has much lower frequency than those in the benchmark textured images. This results in larger window size.

Analysis of the skin lesion images tells us that the real lesion may occupy from 20% to 40% of the whole image. However, for our experiments, all feature vectors were assigned labels based on whether the whole lesion image is classified as benign or malignant. Experiments results show better classification accuracy when using skin lesion images from TLM modality, which confirms our observation. New experiments are designed to solve this problem by discard the background feature vectors, which will improve the classification performance.

IV. CONCLUSION AND FUTURE WORKS

In this paper we described a SVM-based texture classification algorithm for early melanoma detection. All the experiments were done using a simple texture feature – AR features. We tested the algorithm on a benchmark binary texture classification problem, and select the optimal window size 17x17 and 4^{th}-degree polynomial kernel for SVM based on performance analysis. After following the same performance analysis routine, we identified 35x35 as optimal window size, and 4^{th} degree polynomial kernel as optimal degree. Using this set of parameters we did experiments using 22 pairs of skin lesion images with non-disputable class label. The average accuracy for the binary classification is about 70% when 200 feature vectors are selected from each image. Future study look at whether more advanced texture feature extraction algorithms, such as the multi-channel filtering based texture feature extraction [5] will improve the classification accuracy for early melanoma detection. Performance comparison will be studied between SVM-based texture classification and the most popular data mining algorithm, decision-tree based texture classification, in both spatial and frequency domain.

REFERENCES

[1] Dhawan AP, Gordon R, Rangayyan RM: Noevoscopy: Three-dimensional computed tomography for nevy and melanoma by transillumination. IEEE Trans of Med Imaging. MI-3:54-61; 1984

[2] G. Zouridakis, M. Doshi, M. Duvic, and N.A. Mullani, Transillumination Imaging for Early Skin Cancer Detection, Technical Report, UH-CS-05-05, March, 2005

[3] Hofmann-Wellenhof R, Blum A, Wolf IH, Piccolo D, Kerl H, Garbe C, Soyer HP. Dermoscopic classification of atypical melanocytic nevi (Clark nevi). Arch Dermatol.,137(12):1575-1580, 2001.

[4] Blum A, Soyer HP, Garbe C, Kerl H, Rassner G, Hofmann-Wellenhof R. The dermoscopic classification of atypical melanocytic naevi (Clark naevi) is useful to discriminate benign from malignant melanocytic lesions. Br J Dermatol. 149(6): 1159-1164, 2003.

[5] T. Randen, J. Husoy, "Filtering for Texture Classification: A Comparative Study", IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol.21 No. 4, pp. 291-310, 1999.

[6] X. Yuan**,** J. Zhang, and B. Buckles, "Multi-scale Feature Identification using Evolution Strategies," Image and Vision Computing Journal, vol. 23/6, pp. 555-563, 2005.

[7] V.N. Vapnik, Statistical Learning Theory, John Wiley and Sons, New York (1998).

[8] C. Cortes and V.N. Vapnik, Support vector networks, Machine Learning 20 (1995), pp. 273–297.

[9] M. Kanevski, A. Pozdnukhov, S. Canu, M. Maignan. Advanced Spatial Data Analysis and Modelling with Support Vector Machines. International Journal on Fuzzy Systems 2002. p. 606-615.

[10] Y.B.Dibike, S. Velickov, D.P.Solomatine, and M.B.Abbot, Model induction with support vector machine: Introduction and applications, J. Comput. Civil Eng. 15(2001), pp. 208-216.

[11] Kwang In Kim, Keechul Jung, Se Hyun Park, and Hang Joon Kim, "Support Vector Machines for Texture Classification," IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 24, No. 11, Nov. 2002.

[12] P. Brodatz, Textures: A Photographic Album for Artists and Designers. New York: Dover, 1966.

[13] http://www.tlite.com/

[14] Doshi M. Automatic Segmentation of Skin Cancer Images, MS Thesis, University of Houston, 2004.