Neural Network Diagnosis of Malignant Skin Cancers Using Principal Component Analysis as a Preprocessor

Benjamin Kusumoputro and Aripin Ariyanto
Faculty of Computer Science, University of Indonesia
Jl. Salemba Raya no 4, Jakarta Indonesia 10002
E-mail : kusumo@cs.ui.ac.id

Abstract

This paper presents an artificial neural network which is used to separate the malignant melanoma from benign categories of skin cancers based on cancer shapes and their relative color. To reduce the computational complexities, while increasing the possibility of not being trapped in local minima of the Back-Propagation neural network, we applied PCA (principal component analysis) to the originally training patterns, and utilized a cross entropy error function between the output and the target patterns. By using this method, more built-in features of the cancer image through its color and the cancer shapes could be used as the input of the system, leading to higher accuracy of finding the differences between malignant cancer from the benign one. Using this approach, for reasonably balance of training/testing sets, above 91,8% of correct classification of malignant and benign cancers could be obtained.

Key words : Skin cancers, Malignant melanoma, Neural network

1. Introduction

Malignant melanoma is the deadliest form of all skin cancers. Melanoma cells produce a dark protective pigment called melanin, and has a tendency to spread rapidly. Dermatologist believe that early diagnosis of this cancer through its certain physical features is very important for successful treatment. When diagnosing skin cancers, dermatologists usually made their decisions based on clinical experiences, as well as on complex inferences and extensive pathological knowledge through a biopsy procedure. However, since this procedure involved some expense and morbidity, other alternative techniques which have rapid and convenient procedure are being sought. One of the solutions is by using an artificial neural network technique. In this approach, we used a color images of skin cancers instead of living tissue, and utilized an image processing technique combined with neural network algorithm. Supervised learning back-propagation is chosen in this experiment, since this algorithm has high probability in finding the best correlated-response for the input that is similar but not identical, to that already used in the training stage. To memorize the clinical experience of the dermatologist, the weights of the interconnected neurons are updated via the training/testing paradigm. Two statistically independent disjoint sets were designed and utilized as a training set and a test set, to obtain unbiased results from the test set.

Friedman *et.al* described that early malignant melanoma can be classified through mnemonic 'ABCD', namely : Asymmetry, Border irregularity, Color and Diameter [1]. This mnemonic is usually used as the main extracted-features from the cancer images. A benign cancer, namely dysplastic nevi, intradermal nevi and seborcheic keratoses, as outlines by The National Cancer Institute have been characterized also based on its color, shape, surface and size. Since malignant and benign cancers are just differ slightly on their physical characteristics and their color, especially in the early phase, it is difficult to classify the cancers by taking just from the above descriptions.

Friket *et al.* [2] tried to diagnose the malignant melanoma through its color images using standard back-propagation neural network. The used 14 features derived from the color images and obtained about 80% correct classification. The importance of shape analysis for distinguishing malignant from benign cancer is reported by Feig *et al.* [3] and Sickles [4] in mammograms cancer images. Liang Shen *et al* [5] developed a procedure of using a set of shape features to make a mammographic classification more accurate. This analysis indicates that break down of the built-in features of the images into a sets of shape features (irregularity, size, asymmetry, etc.) and a set of color features (RGB, hue, etc.) are necessary, and a collection of these features, rather than features from shape or color alone, is needed to obtain a satisfactory classification of the cancers.

0-7803-4859-1/98 $10.00©1998 IEEE

2. Built-in Features of the Cancer Images

The features of the shape which are used in this analysis include 4 features of moments, one feature of Fourier descriptor and one feature of irregularity index. Other features are derived from color images, including 3 features of color variance, 3 features of spherical color coordinates, 3 features of relative chromatocity, and 3 features of the Intensity-Hue-Saturation (IHS) transformation.

Moments : Moments have been utilized as pattern features in many applications. The definition of moments in this study is usually used for a sequence of contour pixels. One dimensional statistical moments are extracted from a sequential representation of the boundary of a region; where the boundary is characterized by an ordered sequence that represents the Euclidean distances *(z(i))*, between the centroid of the region and all contour pixels of the digitized region (m_1). Shen *et al.* [5] proposed 4 new features of moments that are used in this paper, which overcomes the shortcomings of the original feature set derived by Gupta *et.al* [6], as can be written :

$$F_1 = \frac{(M_2)^{\frac{1}{2}}}{m_1} = \frac{\left[\frac{1}{N}\sum_{i=1}^{N}\left[z(i) - m_1\right]^2\right]^{1/2}}{\frac{1}{N}\sum_{i=1}^{N} z(i)}$$

$$F_2 = \frac{(M_3)^{\frac{1}{3}}}{m_1} = \frac{\left[\frac{1}{N}\sum_{i=1}^{N}\left[z(i) - m_1\right]^3\right]^{1/3}}{\frac{1}{N}\sum_{i=1}^{N} z(i)}$$

$$F_3 = \frac{(M_4)^{\frac{1}{4}}}{m_1} = \frac{\left[\frac{1}{N}\sum_{i=1}^{N}\left[z(i) - m_1\right]^4\right]^{1/4}}{\frac{1}{N}\sum_{i=1}^{N} z(i)}$$

$$F_4 = F_3 - F_1$$

The values of these moments are dimensionless, and invariant to transition, rotation and scaling. In general, the smoother the shape contour, the larger the value of moments.

Fourier Feature: Fourier Feature (FF) is extracted by, firstly, transforming the region of image into a complex plane with imaginary and real axis of $Z_i = x_i + j\ y_i$. Secondly, we derive Fourier Descriptor by using FFT algorithm of $Z_i = x_i + j\ y_i$:

$$A(n) = \frac{1}{N}\sum_{i=0}^{N-1} Z_i \exp(-j2\Pi\, ni / N)$$

$n = 0, 1, ..., N-1$

To eliminate the dependence of the Fourier Descriptor to position, size, orientation and starting point of the region, normalization procedure is performed by :

$$NFD(k) = \begin{cases} 0 & k = 0 \\ A(k)/A(1) & k = 1, 2, ..., N/2 \\ A(k+N)/A(1) & k = -1, -2, ..., -N/2 + 1 \end{cases}$$

Then the Fourier Feature (FF) is defined as [5]:

$$FF = \frac{\sum_{k=-N/2+1}^{N/2} \|NFD(k)\| \ / \ |k|}{\sum_{k=-N/2+1}^{N/2} \|NFD(k)\|}$$

If the shape of the boundary is more complex or rougher, the value of FF will be smaller.

Irregularity Index : The malignant cancer can also be characterized partially by irregularity in its border image which provides a simple measure of counter complexity versus area enclosed. This index is also independent of translation, rotation and scale. The irregularity index is defined as:

$$I = \frac{p^2}{4\Pi A}$$

where p : perimeter of the cancer (in pixels) and A : area of the tumor (in pixels). For a circle contour, irregularity index is obviously the minimum and equal to one. Most melanoma have high irregularity index.

Color variance : Since most of the malignant has higher variance in the red (R), green (G) and blue (B) color components, whilst benign cancer usually has lower color variance, RGB color variance can also be utilized as the extracted-feature of the cancer images. Example of color variance for red is :

$$V_R = \frac{\sum_{i=1}^{N}(R_i - R_{avg})^2}{N}$$

Spherical Color Coordinates: Spherical color spaces are extracted using a coordinate transformation from RGB color spaces, through [7]:

$$L = \sqrt{R^2 + G^2 + B^2}$$

$$\alpha = \cos^{-1}\left[\frac{B}{L}\right]$$

$$\beta = \cos^{-1}\left[\frac{R}{L\sin(\alpha)}\right]$$

This transformation splits the color space into 2-D color space (angle α and angle β) and 1-D intensity space (L).

Relative Chromaticity : Relative chromaticity is defined as the normalized value of the color in the cancer area that is subtracted from the normalized value of that color in the background. Relative chromaticity of the blue, for example, is

$$B_R = \frac{B_t}{R_t + G_t + B_t} - \frac{B_b}{R_b + G_b + B_b}$$

where Bt, Gt and Rt denote the RGB components of the cancer, and Bb, Gb and Rb the RGB components of the background. These features are believed to reduce the small variation of lighting, printing and digitization process of the image and also to equalize the variations of individual human color [2].

Intensity-Hue-Saturation : Although the RGB coordinate is a good representation of all colors, this space is not enough to represent the perception of color. HIS spaces coordinates is believed to have color perception better for human eye [8]. The relation between RGB and IHS can be derived through :

$$I = R + G + B$$

$$S = 1 - \frac{3 \min(R,G,B)}{I}$$

$$H = \cos^{-1}\left\{\frac{\frac{1}{2}[(R-G)+(R-B)]}{\sqrt{(R-G)^2+(R-B)(G-B)}}\right\}$$

3. Principal Component Analysis and Neural Network

Schematic diagram of the developed system for diagnosing the skin cancer is illustrated in Fig.1. Since automated boundary segmentation is not a successful tool at present [2], boundary segmentation is done manually with help of some software tool. The representation and description processes of the images are in mutual accord with the definition of the built-in features described previously. Extracted built-in features of the image are then processed by PCA. The purpose of PCA process is to reduce the number of features of the images into orthogonal features that still have the whole information. Since longer computing time of the training process is one of the drawback in the back-propagation neural network, the reduced-number of the features by PCA will increase the computing speed without sacrificing the information. Mathematically, the PCA is solved through eigen-decomposition of a matrix.

In this research, multilayer perceptron (MLP) neural network with backpropagation algorithm is used. At each iteration, the error between the actual output and the desired output will be reduced, by changing the value of the connection weight. However, this algorithm has two drawbacks that are often pointed out, i.e. : the very slow computing speed and the possibility of being trapped in local minima. To that purpose, the PCA as the preprocessor of the neural network, are used to reduce the complexity and computing time, while for increasing the probability of not being trapped in local minima, a cross entropy error function is used, instead of usual quadratic error function [9][10].

4. System and Experimental Design

The experiments are conducted with a personal computer Pentium 100 MHz. The software is developed by our group and written in Borland C++ language in Windows 95 environment. All images are 360 X 360 pixel color images with 24 bits per pixel, i.e. 8 bits for each of the R, G, B planes. For the system implementation reported here, 63 digital images, 29 malignant melanoma and 34 benign, of skin cancers are observed. All the color photographic slides are obtained from Medical Faculty of the University of Indonesia. Images are converted to digital data using scanner and recorded as bit map pattern (BMP) files. The experiments are conducted with 3 sets of training/testing paradigm; the percentage of malignant and benign cancers can be seen in Table 1.

The developed MLP consists of three layers, with input neurons in the input layer has the same dimensions as the extracted-features from the PCA. For different percentage of principal component used in the experiments, different input neurons are utilized. Output neuron is only one, which determined whether the cancer is malignant or not. Hidden layer has one layer with 9 neurons, and weights are initialized with Nguyen-Widrow initialization rule.

5. Experimental Results

Table 2 shows the relation between extracted-feature with percentage of principal component for shape features, or color features, or shape and color features involved. It is clearly seen that reduced-feature for color and shape analysis could be dropped into 14 features for 100% principal component, 7 features for 95% principal component, and 4 features for 85% principal component. These percentage of principal component will be used to test the recognition probability of the system.

Figure 2 presents the relation between recognition probability for different percentage of principal component for each of the experiment cases, while Table 3 shows the results of the overall experiments, including the experiment without PCA, for comparition. It is clearly seen that by increasing the percentage of principal component, the recognition probability increases significantly.

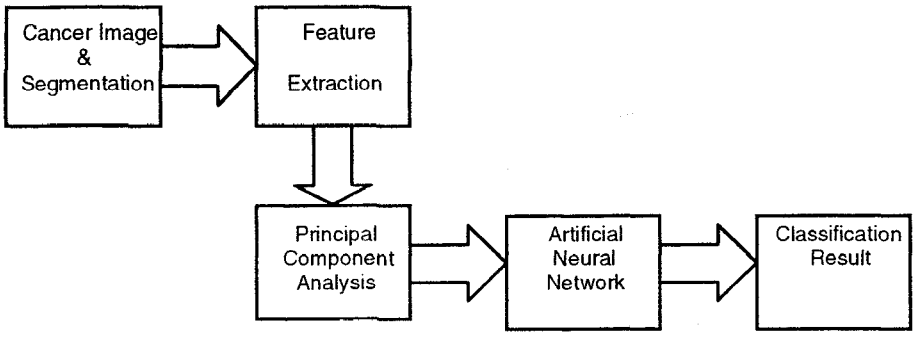

Figure 1. Schematic diagram of the developed diagnosis system

Table 1. Percentage of training/testing set of the benign and malignant images

	Case 1		Case 2		Case 3	
	Number	%	Number	%	Number	%
Training Set	**35**	**55.56**	**37**	**58.73**	**32**	**50.79**
Benign Tumor	18	28.57	20	31.75	17	26.98
Malignant Tumor	17	26.98	17	26.98	15	23.81
Test Set	**28**	**44.44**	**26**	**41.27**	**31**	**49.21**
Benign Tumor	16	25.40	14	41.28	17	26.98
Malignant Tumor	12	19.05	12	41.29	14	22.22

Table 2. Number of principal component and the percentage of PC

	Color & Shape Feature Total Feature : 18			Color Feature Total Feature : 12			Shape Feature Total Feature : 6		
No	Case1	Case2	Case3	Case1	Case2	Case3	Case1	Case2	Case3
1	0.403	0.376	0.390	0.616	0.557	0.612	0.729	0.646	0.601
2	0.666	0.606	0.635	0.750	0.752	0.781	0.888	0.867	0.814
3	0.764	0.742	0.754	0.846	0.847	0.861	0.981	0.969	0.947
4	0.828	0.823	0.822	0.918	0.915	0.925	0.995	0.993	0.995
5	0.885	0.878	0.883	0.951	0.957	0.966	1,000	1,000	1,000
6	0.928	0.927	0.926	0.973	0.975	0.985			
7	0.951	0.950	0.948	0.990	0.991	0.995			
8	0.968	0.966	0.966	0.999	0.999	0.999			
9	0.979	0.997	0.981	1,000	1,000	1,000			
10	0.989	0.985	0.991						
11	0.995	0.993	0.996						
12	0.998	0.998	0.998						
13	0.999	0.999	0.999						
14	1,000	1,000	1,000						

For 100% principal component, the neural can recognize within average probability as high as 91.83%. This high recognition probability is comparable with the experiment of neural network without PCA. However, as can be seen in Fig.3, the computing time of the neural network without PCA is longer, with its number of epoch is about twice as that with the neural using 100% principal component as a preprocessor.

6. Conclusion

Analysis of malignant skin cancers through its shape and color is performed using a neural network based recognition system. Principal component analysis and cross entropy error function are utilized to reduce the disadvantages of the multilayer perceptron neural network. Results showed that the system can distinguished the malignant cancers from the benign one with recognition probability as high as 91.83%. This analysis can be done in an inexpensive personal computer enables dermatologists to perform optical signal processing directly.

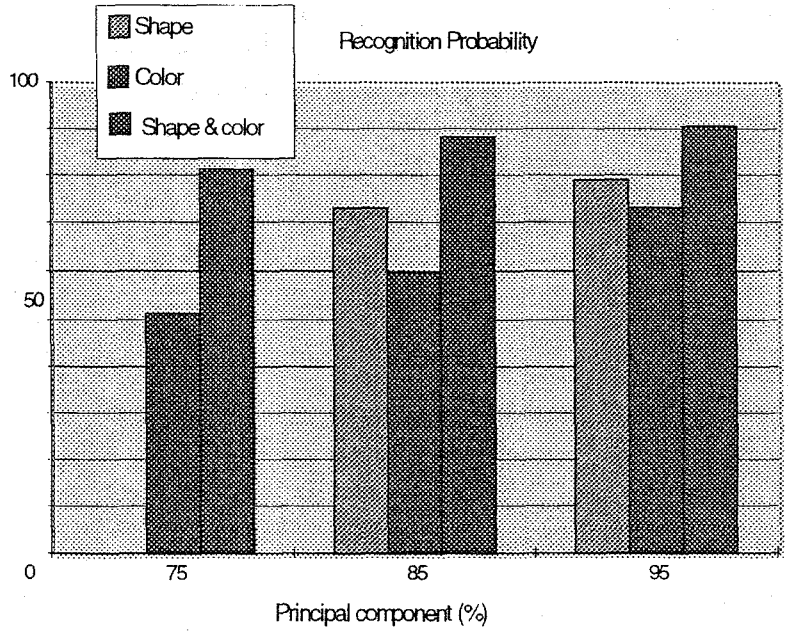

Figure 2. Recognition Probability for different %PC of each cases.

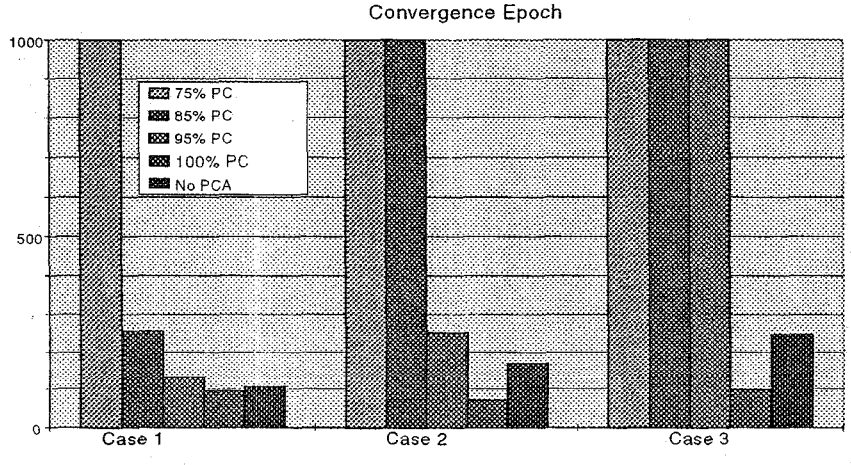

Figure 3. Number of computational epoch for the overall experiments

Table 3. Recognition probability of the %PC and without PCA for the overall cases.

		Case 1			Case 2			Case 3			Mean
%PC	Features	Correct	FALSE	%	Correct	FALSE	%	Correct	FALSE	%	%
75	3	23	5	82.14	22	4	84.62	24	7	77.42	81.39
85	4	24	4	89.28	23	3	88.46	27	4	87.10	88.28
95	7	26	2	92.86	24	2	92.31	27	4	87.10	90.76
100	14	26	2	92.86	24	2	92.31	28	3	90.32	91.83
No-PCA	19	26	2	92.86	24	2	92.31	28	3	90.32	91.83

Acknowledgments

The authors would like to express their gratitude to Profs. Dr. T. Nakamoto and Dr. T. Moriizumi, both from the Tokyo Institute of Technology, for their valuable discussion and collaborations. This work is supported by Ministry of Research and Technology of Indonesia, through RUT IV Program with contract number 271/SP/RUT/BPPT/IV/96.

References

[1] Friedman, D.S. Rigel and A.W. Kopf, "Early detection of malignant melanoma: The role of physician examination and self-examination of the skin", *Ca-A Cancer J. Clinicians*, vol. 35(3), 1985, pp. 130-151.

[2] Friket, C.Anurag, W.V. Stoecker, H.H. Lee and R.H. Moss, "Neural network diagnosis of malignant melanoma from color images", *IEEE Trans. Biomed.Eng.* vol. 41(9), 1994, pp. 837-844.

[3] S.A.Feig, B.M.Galkin and H.D. Muir, "Evaluation of breast microcalcifications by means of optically magnified tissue specimen radiographs", *Recent Results in Cancer Res.*,vol 105, 1987, pp.111-123.

[4] E.A.Sickles, "Breastcalcifications: Mammographic evaluation", *Radiology*, vol 160, 1986, pp. 289-293.

[5] Shen, R.M. Rangayyan and J.E.L. Desautels, "Application of shape analysis to mammographic calcifications", *IEEE Trans. Med. Imaging*, vol 13(2), 1994, pp 263-274.

[6] L.Gupta and M.D. Srinath, "Contour sequence moments for classifications of closed planar shapes", *Pattern Recognition*, vol 20(3), 1987, pp. 267-272.

[7] M.Celenk, "A color clustering technique for image segmentation", *Comput.Vision Graphics and Image Processing*, vol 52, 1990, pp 145-170.

[8] Gonzales and R.E. Woods, *Digital Image Processing*, Addison-Wesley Inc., Massachusets, 1992

[9] Humpert, "*Improving backpropagation with a new error function*", *Neural Networks*, vol 7(8), 1994

[10] Matsuoka and J.Q. Yi,' Backpropagation based on the logaritmic error function and elimination of local minima, *Int. Joint Conf. on Neural Networks*, vol 2, 1991