Texture Analysis Based Automated Decision Support System for Classification of Skin Cancer Using SA-SVM

Ammara Masood, Adel Al-Jumaily, and Khairul Anam

School of Electrical, Mechanical and Mechatronic Engineering,
University of Technology, Sydney, Australia
{ammara.masood,khairul.Anam}@student.uts.edu.au,
Adel.Ali-Jumaily@uts.edu.au,

Abstract. Early diagnosis of skin cancer is one of the greatest challenges due to lack of experience of general practitioners (GPs). This paper presents a clinical decision support system aimed to save lives, time and resources in the early diagnostic process. Segmentation, feature extraction, and lesion classification are the important steps in the proposed system. The system analyses the images to extract the affected area using a novel proposed segmentation method H-FCM-LS. A set of 45 texture based features is used. These underlying features which indicate the difference between melanoma and benign images are obtained through specialized texture analysis methods. For classification purpose, self-advising SVM is adapted which showed improved classification rate as compared to standard SVM. The diagnostic accuracy obtained through the proposed system is around 90% with sensitivity 91% and specificity 89%.

Keywords: Skin Cancer, diagnosis, feature extraction, classification, self-advising support vector machine.

1 Introduction

Malignant melanoma is one of the deadliest forms of skin cancer. A rapid increase in melanoma cases is observed in Europe, North America, and Australia over the last decade. Over 76,250 new cases of invasive melanoma were diagnosed in the US in 2012 [1]. An estimated 1,890 Australians die from skin cancer each year [2]. From treatment point of view, skin cancer is one of the most expensive forms of cancer, but early diagnosis can make the situation better as melanoma has near 95% cure rate if diagnosed and treated in early stages [1].

A Computer Aided Diagnostic (CAD) system for diagnosis of skin cancer is aimed to find the exact boundaries of a lesion automatically and also to provide an estimate of the probability of a disease. There are various diagnostic systems proposed in literature [3-5] but as we discussed in [6] more research is required to make the best choice and for setting the benchmarks for diagnostic system development and validation. This paper presents a part of our research being carried out to come up with the best combination of segmentation, feature extraction and classification algorithms which can consequently form the basis of a more generalized and efficient skin cancer diagnostic system. The diagnostic model proposed here is shown in Fig. 1.

C.K. Loo et al. (Eds.): ICONIP 2014, Part II, LNCS 8835, pp. 101–109, 2014.

The proposed decision support system uses adaptive median filter for pre-processing of image to reduce the ill effects and various artifacts like hair that may be present in the images. It is followed by the detection of the lesion by our Histogram based fuzzy C means thresholding algorithm presented in [7]. This algorithm provided efficient segmentation results as compared to other segmentation methods used in literature; the comparative analysis is presented in [8]. Once the lesion is localized, texture based features are quantified. Finally, Self-advisable Support Vector machine (SA-SVM) is used for classification of cancerous and non-cancerous skin lesions.

This paper is organized as follows: Section 2 describes the computer-aided diagnosing (CAD) system which consists of pre-processing, segmentation, features extraction and classification stages. Section3 presents the experimental results, comparative analysis and discussion, and final section for conclusions and future work.

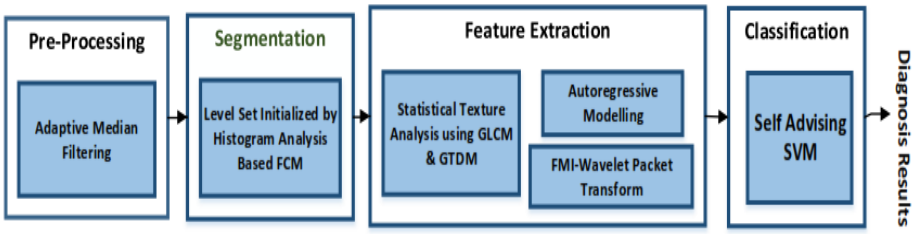

Fig. 1. Computer Aided Diagnostic Support System

2 Proposed CAD System

2.1 Pre-processing

Skin images have certain extraneous artifacts such as skin texture, dermoscopic gel and hair that make border detection a bit difficult. It is necessary to pre-process the images with a smoothing filter like adaptive median filter. The median filter also performs well as long as the spatial density of the impulse noise is not too large. However the adaptive median filtering has a better capability to handle impulse noise with even larger probabilities. An additional benefit of the adaptive median filter is that it seeks to preserve details while smoothing the non-impulse noise [9]. Considering the high level of noise that may be present in skin lesion images and the need of preserving structural details, the adaptive median algorithm performed quite well.

2.2 Segmentation

In one of our previous work [10], We proposed a segmentation algorithm, histogram analysis based fuzzy C mean algorithm for Level Set initialization (H-FCM-LS) as presented in figure 2. In the proposed method, histogram analysis of image was done to see average intensity distribution in the images and then the hard threshold was selected between classes with dominant intensity values. This method was further

used as an initializing step for complex segmentation method like Level set having spatial information. Segmentation results for some of the skin lesion images are shown in figure 2. For details of algorithm refer to [10].

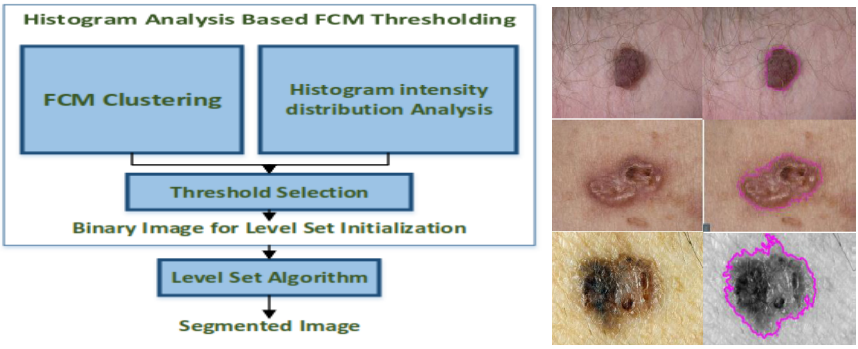

Fig. 2. Block Diagram & Results of Segmentation Algorithm (H-FCM-LS)

2.3 Feature Extraction

Texture analysis can potentially expand the visual skills of the expert eye by extracting features that are relevant to diagnostic problem and not necessarily visually extractable. Three types of methods for texture based feature extraction are used here 1) model-based 2) statistical and 3) transform-based.

Gray Level Co-occurrence Matrix (GLCM)
GLCM was introduced by Haralick [11] provides one of the most popular statistical methods in analysis of grey tones in an image. The GLCM functions characterize the texture of image by calculating how often pairs of pixel with specific values and in a specified spatial relationship occur in an image, creating a GLCM, and then extracting statistical measures from this matrix. Details for GLCM based features that are used for classification stage can be found in [12].

Grey-Tone Difference Matrix (GTDM)
GTDM was suggested by Amadasun [13] in an attempt to define texture measures correlated with human perception of textures. A GTDM matrix is a column vector containing G elements. Its entries are calculated based on the difference between intensity level of a pixel and average intensity computed over a square, while sliding window centered at the pixel. Suppose the image intensity level $I_L(x,y)$ at location (x,y) is I, i=0,1,… G-1. The average intensity over a window centered at (x,y) will be $I_L(i) = I(x,y) = \frac{1}{W-1}\sum_{m=-K}^{K}\sum_{n=-K}^{K} f(x+m, y+n)$, where K specifies the window size and W= (2K+1)2. The i^{th} entry of GTDM x is $s(i)=\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}|i - f_i|$ for all pixel having intensity level I, otherwise, s(i)=0. Mathematical formulae for GTDM based features used here are given in Table 1.

Table 1. GTDM based Features

Feature	Mathematical Equation
Coarseness	$\left(\varepsilon + \sum_{i=0}^{G-1} p_i\, s(i)\right)^{-1}$
Contrast	$\left[\frac{1}{N_t(N_t-1)}\sum_{i=0}^{G-1}\sum_{j=0}^{G-1} p_i p_j (i-j)^2\right]\left[\frac{1}{n}\sum_{i=0}^{G} s(i)\right]$
Business	$\frac{\sum_{i=0}^{G-1} p_i\, s(i)}{\sum_{i=0}^{G-1}\sum_{j=0}^{G-1} \lvert i p_i - j p_j\rvert}\ p_i \neq 0, p_j \neq 0$
Complexity	$\sum_{i=0}^{G-1}\sum_{j=0}^{G-1} \frac{\lvert i-j\rvert}{n(p_i+p_j)}\left[p_i s(i) + p_j\, s(j)\right]\ p_i \neq 0, p_j \neq 0$
Texture Strength	$\frac{\sum_{i=0}^{G-1}\sum_{j=0}^{G-1}(p_i+p_j)(i-j)^2}{\varepsilon + \sum_{i=0}^{G-1} s(i)}\ p_i \neq 0, p_j \neq 0$

Fuzzy-Mutual Information Based Wavelet Packet Transform

The wavelet packet method is a generalization of wavelet decomposition and offers a richer signal analysis. Different extensions of wavelet packet transform are present in literature for different applications [14,15]. It is observed that features extracted using wavelet transforms provide significant increase in the classification accuracy. After converting skin images to corresponding vectors, following Fuzzy mutual-information based wavelet packet transform (FMI-WPT) is used:

1) For each original image vector, perform a full WPT decomposition to the maximum level J (taken as 3 here). For all j = 0, 1,.., J and k = 0, 1, . . , 2j − 1, construct features according to relation $E_{\Omega j.k} = \log\left(\frac{\sum_n (w^T{}_{j,k,n}\, x)^2}{N/2^j}\right)$. where Ωj,k is the decomposition subspace with j denoting scale and k denoting sub-band index within the scale [15].

2) Construct associated fuzzy sets and compute fuzzy entropies and mutual information. Then evaluate classification ability of n number of features using fuzzy-set based criterion F_i where $F_i = I(C; f_i)/H(f_i)$ for i = 1, 2, . . . , n.

Note: I (f;C) = H(f) – H(f |C) where H(f) is marginal entropy of f and H(f |C) is conditional entropy of f and C [16].

3) Determine the optimal WPT decomposition X, being the one that corresponds to the maximum value of F.

4) The set X is the final FMIWPT-based decomposition.

Autoregressive Modeling Based Features

Autoregressive modeling is an all pole modeling which is widely used to get a robust spectral estimation of one dimensional signal. Yule-walker method is the most widely used method to estimate autoregressive coefficients. In order to determine the coefficients, it uses Levinson-Durbin algorithm to minimize error. This estimation method minimizes square of forward prediction error and finds out autoregressive parameters by solving autocorrelation function (1) as expressed in [16].

$$\sigma^2 = \frac{1}{N}\sum_{n=-\infty}^{\infty}\left|x(n) - \sum_{k=1}^{p} a(k)x(n-k)\right|^2 \quad (1)$$

Where x(n) is input and a(k) demonstrate autoregressive parameters. As images are 2-dimensional signals, so for doing autoregressive modeling, skin lesion images were converted into corresponding vectors. Estimated autoregressive parameters which are poles of 1-dimensional signals are used as features vectors extracted from images.

2.4 Classification

Classification of the lesion as cancer or non-cancer is the final step. For classification of skin lesions, an improved version of support vector machine (SVM), named Self Advising SVM is adapted here. SVM is a well-known machine learning method proposed by Vapnik [17] and a lot of literature is available for applications of SVM. The idea of SVM is to construct a maximized separating hyper plane that can separate data in the feature space. The classic SVM ignores the training data that has not been separated linearly by the kernels during the training phase. Thus, if data that is similar or identical to this misclassified data appears in the test set, it will be classified wrongly. This misclassification is not reasonable and it can be handled if the available data and information in the training phase has not been ignored by SVM algorithm.

In this study a non-iterative self-advising approach [18] for SVM is adapted that extracts subsequent knowledge from training phase. The misclassified data can come from two potential sources 1) outliers 2) data that have not been linearly separated by using any of the types of kernels. SA-SVM deals with the ignorance of SVM from the knowledge that can be acquired from misclassified data by generating advice weights based on use of misclassified training data, and through use of these weights together with decision values of SVM in the test phase. These weights also help the algorithm to eliminate the outlier data. The details of SA-SVM algorithm are as follows:

Training Phase

1. Finding hyperplane by using decision function$f(x) = sign(\sum_{\alpha_{i>0}} y_i \alpha_i \, k(x, x_i)$ +b) i.e. the normal SVM training. Note here that the kernel function we used is radial Basis Function so $K(x_i x_j) = e^{-\gamma |x_i - x_j|^2}$.

2. To benefit from the misclassified data of the training phase, the misclassified data sets (MD) in the training phase is determined as

$$MD = \cup_{i=1}^{N} x_i \; | y_i \neq sign(\sum_{\alpha_j > 0} y_i \, \alpha_j \, k(x_i, x_j) + b) \tag{2}$$

The MD set can be null, but experimental results have revealed that the occurrence of misclassified data in training phase is a common occurrence. Note that xi is the input vector corresponding to the ith sample and labeled by yi depending on its class and αi is the nonnegative Lagrange multiplier as used in standard SVM [19].

3. If the MD is null, go to the testing phase else compute neighborhood length (NL) for each member of MD. NL is given as

$$NL(x_i) = minimum_{x_j} (\|x_i - x_j\| \mid y_i \neq y_j) \tag{3}$$

Where xj , j=1, …., N are the training data that do not belong to the MD set. Here the training data is mapped to a higher dimension, the distance between xi and xj is be computed according to the following equation with reference to the related kernel k (RBF).

$$\|\theta(x_i) - \theta(x_j)\| = (k(x_i, x_i) + k(x_j, x_j) - 2k((x_i, x_j))^{0.5} \quad (4)$$

Testing Phase

1. Compute the advised weight AW(x_k) for each x_k, from the test set. Where AW is computed as (5), These AWs represent how close the test data to the misclassified data is.

$$\begin{cases} 0 & \forall x_i \in MD, \|x_k - x_i\| > NL(x_i) or\ MD = NUL, \\ \sum 1 - \frac{\sum_{x_i} \|x_k - x_i\|}{\sum_{x_i} NL(x_i)} & x_i \in MD, \|x_k - x_i\| \le NL(x_i) \end{cases} \quad (5)$$

2. Compute the absolute value of the SVM decision values for each xk from the test set and scale it to [0, 1].

3. For each xk from the test set, If (AW (x_k) < decision value (x_k) then y_k = sign ($\sum_{\alpha_j>0} y_j \alpha_j k(x_k, x_j) + b$) which is normal SVM labeling, otherwise $y_k = y_i \mid (\|x_k - x_i\| \le NL(x_i) and\ x_i \in MD)$

3 Experimental Results

A clinical database of dermoscopic and clinical view lesion images were obtained from different sources but most of the images came from Sydney Melanoma Diagnostic Centre, Royal Prince Alfred Hospital. A total of 168 images (56 benign and 112 melanoma) were included in experimental data set. All the images were rescaled to a resolution of 720x 472 with bit depth 24 and size around 526KB. After pre-processing and segmentation, a total of 45 features (15 GLCM, 5 GTDM, 15 FMI_WPT, and 10 Autoregressive) were extracted for each image. For training the SA_SVM, 84 images are used and 84 images were used for testing. The whole process is implemented using MATLAB software R2013 and simulated by a system with corei5 3.10 GHz processor and 4 GB memory under Windows7 operating system.

The constructed feature sets are used separately as well as in different combinations for feeding the classifier. The contribution of each feature extraction method used as well as of the proposed set of 45 texture based features for classification can be seen from figure 3. It can be seen clearly that use of GLCM and GTDM based features resulted in better sensitivity ($\frac{True\ Positive}{True\ Positive + False\ Negative}$) but lower specificity ($\frac{True\ Negative}{True\ Negative + False\ Positive}$). On the other hand use of autoregressive and FMI-WPT based features lead to better specificity but relatively lower sensitivity. However, the experimental analysis clearly indicated that a proposed combination of (45 features) using these four feature extraction methods resulted in a feature set that formed a good basis for classification using SA_SVM. The proposed diagnostic system achieved an overall accuracy of around 90%, with sensitivity 91% and specificity 89%.

For cross validating the results, hold-out validation a specific type of k-fold cross validation is used. For each fold, the skin images are randomly divided into two equal sized sets S1 and S2. Then SA-SVM is trained using S1 and tested on S2. This is followed by training using S2 and testing using S1. This has the advantage that our training and test sets are both large, and each data point is used for both training and validation on each fold. In order to ensure better validation of classifier performance the hold-out validation was repeated 5 times and each time the experimental data set was shuffled and split it into two parts.

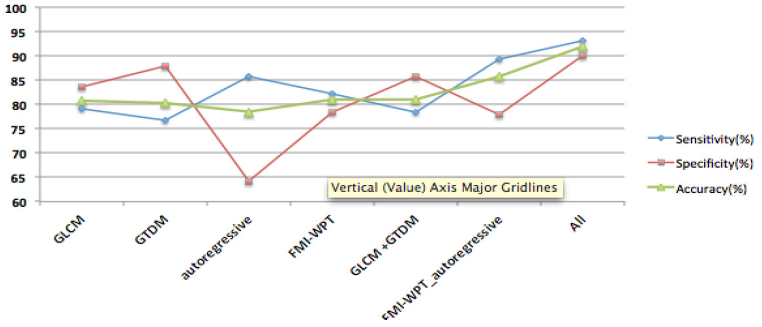

Fig. 3. Comparative Diagnosis results of SA-SVM using proposed feature set

The performance of diagnostic system is analyzed using statistical parameters such as sensitivity, specificity and accuracy. The classification results of SA-SVM are compared with standard SVM (both linear and kernel based). Higher value of both sensitivity and specificity shows better performance of the system. The experimental analysis also shows that the results obtained by the self-advising SVM are significantly better than the results of traditional SVM.

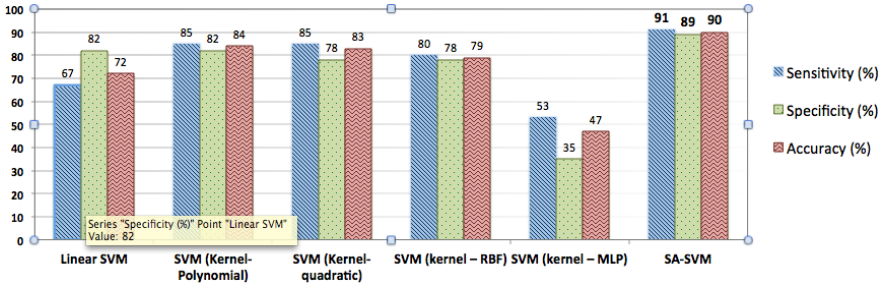

Fig. 4. Relative Performance Measure of SA-SVM

4 Conclusion and Future Work

In this paper, an automated skin cancer diagnostic system is proposed based on SA-SVM. SA-SVM uses information generated from misclassified data in the training

phase and thus, improves performance by transferring more information from training phase to the test phase. Features used for differentiating melanoma and benign images are extracted using four texture analysis methods. A set of features is proposed that worked best for SA-SVM. The diagnostic results obtained are quite satisfactory with sensitivity of 91% and specificity of 89%.

Despite the high accuracy that can be achieved by the proposed system, for developing a more reliable diagnostic system we intend to test multiple classifier based systems as well to undo the chances of any misclassification due to classifier limitations. We intend to do experiments combining different classification algorithms like neural networks, support vector machine and extreme learning machine. Such tools may serve as diagnostic adjuncts for medical professionals. And it will provide the opportunity of implementing more accurate, faster and reliable diagnostic systems.

References

1. Siegel, R., Naishadham, D., Jemal, A.: Cancer statistics. CA: A Cancer Journal for Clinicians 62(1), 10–29 (2012)
2. Causes of death, C.W.o. Australia, Editor, Australian Bureau of Statistics: Canberra (2010)
3. Ganster, H., et al.: Automated melanoma recognition. IEEE Transactions on Medical Imaging 20(3), 233–239 (2001)
4. Rubegni, P., et al.: Automated diagnosis of pigmented skin lesions. International Journal of Cancer 101(6), 576–580 (2002)
5. Ruiz, D., et al.: A decision support system for the diagnosis of melanoma: A comparative approach. Expert Systems with Applications 38(12), 15217–15223 (2011)
6. Masood, A., Jumaily, A.: Computer Aided Diagnostic Support System for Skin Cancer: A Review of Techniques & Algorithms. International Journal of Biomedical Imaging, 22 (2013)
7. Masood, A., Jumaily, A.: FCM Thresholding based Level Set for Automated Segmentation of Skin Lesions. Journal of Signal & Information Processing 4(3B), 66–71 (2013)
8. Masood, A., Al-Jumaily, A.: Automated segmentation of skin lesions: Modified FCM thresholding based LS method. In: 16th International Multi Topic Conference, pp. 201–206 (2013)
9. Sun, T., Neuvo, Y.: Detail-preserving median based filters in image processing. Pattern Recognition Letters 15(4), 341–347 (1994)
10. Masood, A., Al-Jumaily, A.A., Maali, Y.: Level Set Initialization Based on Modified Fuzzy C Means Thresholding for Automated Segmentation of Skin Lesions. In: Lee, M., Hirose, A., Hou, Z.-G., Kil, R.M. (eds.) ICONIP 2013, Part III. LNCS, vol. 8228, pp. 341–351. Springer, Heidelberg (2013)
11. Haralick, R.M., Shanmugan, K.: Textural features for image classification. IEEE Transactions on Systems, Man, and Cybernetics 3(6), 610–621 (1973)
12. Nitish Zulpe, V.P.: GLCM Textural Features for Brain Tumor Classification. International Journal of Computer Science Issues 9(3), 354–359 (2012)
13. Amadasun, M., King, R.: Textural Features Corresponding to Textural Properties. IEEE Transactions on System, Man Cybernetics 19(5), 1264–1274 (1989)
14. Khushaba, R.N., Al-Jumaily, A., Al-Ani, A.: Novel feature extraction method based on fuzzy entropy and wavelet packet transform for myoelectric Control. In: International Symposium on Communications and Information Technologies, pp. 352–357 (2007)

15. Khushaba, R.N., Kodagoa, S., Lal, S., Dissanayake, G.: Driver Drowsiness Classification Using Fuzzy Wavelet-Packet-Based Feature-Extraction Algorithm. IEEE Transactions on Biomedical Engineering 58(1), 121–131 (2011)
16. Proakis, J.G., Manolakis, D.G.: Digital Signal Processing Principles, Algorithms, and Applications. Prentice-Hall, New Jersey (1996)
17. Vapnik, V.N.: The nature of statistical learning theory, 2nd edn. Springer, New York (2000)
18. Maali, Y., Al-Jumaily, A.: Self-advising support vector machine. Knowledge-Based Systems 52, 214–222 (2013)