Automatic Detection of Melanoma Skin Cancer using Texture Analysis

Mariam A.Sheha
Cairo University

Mai S.Mabrouk
MUST University

Amr Sharawy
Cairo University

ABSTRACT

Melanoma is considered the most dangerous type of skin cancer. Early and accurate diagnosis depends mainly on important issues, accuracy of feature extracted and efficiency of classifier method. This paper presents an automated method for melanoma diagnosis applied on a set of dermoscopy images. Features extracted are based on gray level Co-occurrence matrix (GLCM) and Using Multilayer perceptron classifier (MLP) to classify between Melanocytic Nevi and Malignant melanoma. MLP classifier was proposed with two different techniques in training and testing process: Automatic MLP and Traditional MLP. Results indicated that texture analysis is a useful method for discrimination of melanocytic skin tumors with high accuracy. The first technique, Automatic iteration counter is faster but the second one, Default iteration counter gives a better accuracy, which is 100 % for the training set and 92 % for the test set.

Keywords
Texture analysis, GLCM, CAD, Melanocytic Nevi, Melanoma, ANN, MLP.

1. INTRODUCTION
Melanoma is a serious and sometimes life-threatening cancer. It can occur in any part of the body that contains melanocytes. The main cause of melanoma is excessive exposure to ultraviolet (UV) radiation reaching the skin. UV rays from the sun and other sources (such as tanning booths) can damage skin cells, causing the cells to grow abnormally. Fair-skinned people have a higher risk for melanoma and other types of skin cancer if they have too much exposure to sunlight. In 2003 it was estimated that 105,000 people would receive a diagnosis of melanoma and a further 33,000 would die from the decease that year worldwide [1]. In 2011, an estimated 70,230 adults (40,010 men and 30,220 women) in the United States were diagnosed with melanoma. It is estimated that 8,790 deaths (5,750 men and 3,040 women) from melanoma will occurred this year [2].

Melanoma is a particularly deadly form of skin cancer and although it accounts for only 4% of all skin cancers it is responsible for 75% of all skin cancer deaths [1]. If melanoma is found and treated in its early stages, the chances of recovery are very good but if the diagnosis becomes late, melanoma can grow deeper into the skin and spread to other parts of the body. Once melanoma has spread to other parts of the body beyond the skin, it is difficult to treat.

Dermoscopy is a non-invasive examination technique based on the use of incident light and oil immersion to make possible the visual examination of sub surface structures of the skin. The rate of detection of melanoma using dermoscopy is higher than detection only with unaided observation [3].

Since differential diagnosis of melanoma from melanocytic nevi is often not straight forward especially in the early stage, the diagnostic accuracy of dermoscopy is also depending on the training of the dermatologist. So that automatic diagnosis is essential tool for less experienced physicians. It considered to be a "double reading'' system were physicians take into consideration the information provided by computer before making decision.

Computer is not more intelligent than human but it may be able to extract some information, such as texture features, that may not be readily perceived by human eyes. Several scoring systems and algorithms such as the ABCD rule for epiluminescence, the seven-point checklist, and the Menzies method [4], [5] have been proposed to improve the diagnostic performance of less experienced clinicians. Most of the proposed techniques require segmentation process that considers being a fatal problem due to the irregularity of the tumor, where dermoscopy views of histological tissues show structures mostly arranged in a variety of patterns. So that, automatic segmentation of different structures, like nuclei, cytoplasm, vessels etc., is difficult, and cannot be done in general approach [6]-[10].

This study Proposes, an automated system for discrimination between melanocytic nevi and malignant melanoma avoiding segmentation process using texture analysis. Texture analysis refers to the characterization of regions in an image by their texture content. Texture analysis is used in a variety of applications, including remote sensing, automated inspection, and medical image processing. Some of the most commonly used texture measures are derived from the Grey Level Co-occurrence Matrix (GLCM) [11]. Recently, computer-aided dermoscopy using artificial neural networks (ANNs) has been reported to be an accurate tool for the evaluation of pigmented skin lesions (PSLs) [12]-[14]. ANN is considered as important way for classification, it is computational paradigms based on mathematical models that unlike traditional computing have a structure and operation that resembles that of the mammal brain [15]. The ANN classification procedure used is MLP; represent the classes of benign common nevi and malignant melanoma. To check the applicability of the method for prognosis, the classification results are applied to the individual cases.

2. RELATED WORK
Various kinds of techniques have been proposed to improve the accuracy of skin cancer diagnosis. The dermoscope or epiluminescence microscope (ELM) was first described on 1987 [16]; it facilitates non-invasive diagnosis process based on using of incident light, oil immersion and a magnifier. But its accuracy is still mainly depends on physician experience. The research in automatic detection for skin cancer has been conducted in last few decades with several methods and

techniques.

In 2003,M. Wiltgen, A. Gerger and J. Smolle used a method of tissue counter analysis (TCA), which is based on the partition of the image into square elements of equal size where the features, are calculated out of each square element. The features, based on grey level histogram and co-occurrence matrix, allow the differentiation of homogeneous and high contrast tissue areas. The highest classification accuracy obtained by this approach was 92.7% of nevi elements and 92.1% of melanoma elements in the learning set. In the test set, discriminate analysis based on the percentage of "malignant elements" showed a correct classification of all cases [17]. In 2011,Daniel Ruiz, Vicente Berengue, Antonio Sorianoand Belén Sánchez proposed types of ANN classifiers, which area multilayered perceptron, a Bayesian classifier and the algorithm of the K nearest neighbors. These methods work independently and also in combination making a collaborative decision support system. The classification rates obtained are around 87%[18].

3. METHODOLOGY

The proposed methodology of discrimination between malignant and nevi tumors is shown in Figure 1.The method uses the steps of Pre-processing, feature extraction, feature selection, classification and then evaluation. System enables texture analysis without being limited by selection and detection of structure of interest. In this way a prior definition and segmentation process, was avoided. The work at hands is implemented using Matlab ver.10.
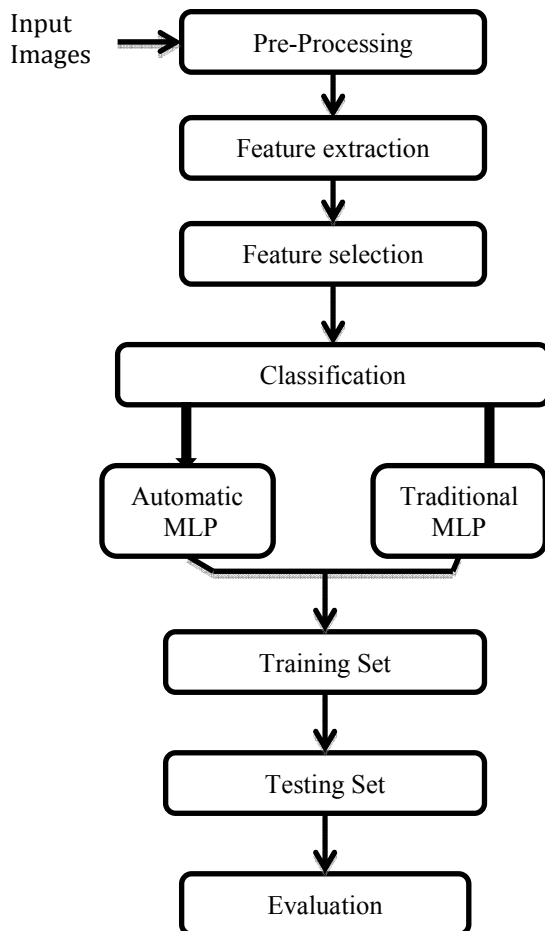

Figure 1: Automated diagnosis block diagram.

3.1 Experimental setup

Development of non-invasive tools to improve early diagnosis results in 2 approaches, dermoscopy and digital image analysis [19]. For the study set, 102 dermoscopyimages for different cases were randomly sampled from the skin diseases Atlases and doctors' clinics. Image atlases of reference pictures are widely available [20], [21]. This set consists of 51dermoscopy images of melanocytic nevi and 51dermoscopy images of malignant melanoma. Regarding that all images are divided to two classes 75% for learning set and 25% for test set. All images were resized to be [512*512] pixel.

3.2 Pre-Processing

This study based on texture analysis, In contrast to other methods in medical image analysis, segmentation process is avoided. And so,preprocessingshows simplicity compared to other methods of detection. Although it's simple process, it shows two main steps.First,image resizing to adjust imagesto fixed scale (512*512) so that it supports the classification with clear and accurate features. After that, convert images from RGB to grey level where the features are based on grey level co-occurrence matrix.

3.3 Feature extraction

The purpose of feature extraction is to reduce the original data set by measuring certain properties, or features, that distinguish one input pattern from another. All feature extracted in this study are based on texture analysis using GLCM. The GLCM is a powerful tool for image feature extraction by mapping the grey level co-occurrence probabilities based on spatial relations of pixels in different angular directions. The feature extracted based on GLCM are: Autocorrelation, Contrast, Correlation, Cluster Prominence, Cluster Shade, Dissimilarity, Energy, Entropy, Homogeneity, Maximum probability, Sum of squares Variance, Sum average, Sum variance, Sum entropy, Difference variance, Difference entropy, Information measure of correlation 1,Information measure of correlation 2,Inverse difference homogenous (INV), Inverse difference normalized (INN) and Inverse difference moment normalized [22].

3.4 Feature selection

In this step, only the most significant features are selected to increase the detection accuracy using fisher score ranking that calculated according to equation (1). The Fisher score ranking technique calculates the difference, described in terms of mean and standard deviation, between the positive and negative examples relative to a certain feature. This technique is described in detail in [23].

$$R_i = \frac{|\mu_{i,p} - \mu_{i,n}|}{|\sigma_{i,p} + \sigma_{i,n}|} \qquad (1).$$

Where, **R_i** is the rank of feature i, the bigger the R_i the bigger the difference between the values of positive and negative examples relative to feature i.By fisher score ranking equation and a co-occurrence matrix C_d one can draw out some important features for texture classification. According to the input dermoscopy images and fisher score ranking, 12 features are selected where the highest score features seems to be the most effective, these features are Contrast, Correlation, Cluster Prominence, Dissimilarity, Homogeneity, Difference variance, Difference entropy, Information measure of correlation 1,Information measure of correlation 2,Inverse difference homogenous (INV), Inverse difference normalized (INN) and Inverse difference moment normalized.

3.5 Classification

The various measurements based on co-occurrence matrix textural features, representing the classes of melanocytic nevi and malignant melanoma,are given as input to neural network classifier which consists of two processes: for training and test set. A common example of such a network used in that study, the Multilayer Perceptron (MLP).As shown in Figure 2,it contains three layers: input layer, hidden layer and output layer.
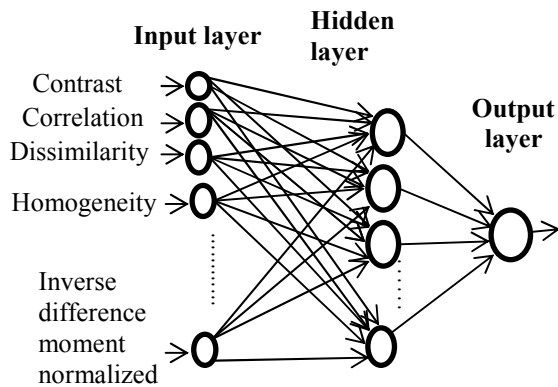

Figure 2: A multilayered feed forward network

The MLP is a feed forward network, capable of generating nonlinear boundaries. It has been successfully applied to solve some difficult and diverse problems [24]-[26]. Two techniques for that classifier were used. The First, Automatic MLP divides the given data to three subsets 60% of data is train, 20 % test, and 20 % validation. Training subset is to train classifier, Validation subset is used to stop training early if the network performance on the validation subset fails to improve or remains the same for six epochs in a row and the weights and biases at the minimum of the validation error are returned. Test subset is used as a further check that the network is generalizing well, but do not have any effect on training [11]. The Second, Traditional MLP use all data as it is and completes the 1000 iteration (default) or as alternative for training step.

3.6 Performance

When classification is done results could have an error rate, either fail to identify an abnormality, or identify an abnormality that is not present. Some kind of performance measure to the diagnosis done by the terms of true and false positive, true and false negative [27],[28].

According to these terms the clinical performance of a classification is described where sensitivity, specificity, and accuracy are calculated for ANN classification through the following equations (2,3,4).

$$\text{Sensitivity} = \frac{TP}{TP+FN} * 100\% \qquad (2)$$

$$\text{Specificity} = \frac{TN}{TN+FP} * 100\% \qquad (3)$$

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} * 100\% \qquad (4)$$

4 RESULTSANDDISCUSSION

This section details the results of automatic classification on images that acquired by means of dermoscopy technique. Database consists of 102 dermoscopy images, previously diagnosed, 51 of them are melanocytic nevi and 51 are melanomas. GLCM features were used for feature extraction and neural network for classification.12 features are selected according to fisher score method from the total 23 of GLCM and these input fed to neural input layer. Table 1 shows GLCM features for malignant and nevi class. To train and to test the system we have used a MLP classifier that was exhibited in two techniques, Automatic MLP and Traditional MLP.

The proposed method trained with 75% and tested with 25% of the total number of images. At the end of the training process updated weight values are stored. Then, performance measures are done by computed output of TP, TN, FP and FN as discussed in the upper section. Table 2 and 3 show the computed sensitivity, specificity and accuracy for training and testing data respectively of MLP techniques. According to the results showed, it's clear that the second technique, Traditional MLP is more effective for detection of melanoma from melanocytic nevi.

5 CONCLUSION

Melanoma considered the most dangerous type of skin cancer. As melanoma diagnosis requires experience, where early stages may look identical to harmless moles. Automatic diagnosis is essential tool for less experience physicians. In this work, an automated system of melanoma classification was applied on dermoscopy images to be an assisting tool in the early diagnosis of malignant melanoma and melanocytic nevi lesions. With main advantage that it is in contrast to other methods in medical image analysis segmentation process is avoided using texture analysis.

At first, Pre-processing adjust all images to a fixed scale [512*512] to support extraction of accurate features; so can obtain clear cut off difference between two types of lesions. It has investigated a classification of dermoscopy images using GLCM features. The texture features obtained from co-occurrence matrix contain 23 sufficient features. The most significant features were selected using fisher score method. Despite fisher's score simplicity, appears to be a good feature selection method. According to fisher score method 12 features were selected that represent the most significant features.

Afterwards, classification process was implemented using MLP classifier that was proposed in two techniques. The performances of the classifier techniques presented different classification accuracy. The first technique: Automatic MLP proposed 93.4% and 76% for training and testing accuracy respectively. The second technique: Traditional MLP, proposed 100% and 92% for training and testing accuracy respectively. The results indicated that the Traditional MLP yielded the better performance when compared to the first one.

In conclusion, this study shows that combination between co-occurrence matrix and ANN is a promising technique for discrimination between malignant melanoma and melanocytic nevi dermoscopy images.

Table 1: GLCM Features Value for Malignant and Nevi Class

Features	Malignant 1	Nevi1	Malignant 2	Nevi2	Malignant 3	Nevi3
Contrast	0.0956152	0.1592465	0.0875007	0.1920330	0.0985926	0.109134
Correlation	0.9785536	0.8042529	0.9861547	0.8904056	0.9826944	0.944586
Cluster Prominence	121.79231	23.69187	276.10653	72.473045	293.16681	33.63818
Dissimilarity	0.0942774	0.1107356	0.0794054	0.1447147	0.0831740	0.098493
Homogeneity	0.9529950	0.9492526	0.9610813	0.9320725	0.9599293	0.951783
Difference Variance	0.0956152	0.1592465	0.0875007	0.1920330	0.0985926	0.109134
Difference entropy	0.3146659	0.3507624	0.2806824	0.4269775	0.2929147	0.327785
Information measure of correlation 1	-0.7864709	-0.5797063	-0.8229712	-0.595457	-0.8117864	-0.739550
Information measure of correlation 2	0.9567472	0.7898776	0.9695077	0.8532012	0.9641905	0.929534
Inverse difference homogenous (INV)	0.9530765	0.9504054	0.9613402	0.9332840	0.9606344	0.952166
Inverse difference normalized (INN)	0.9895394	0.9881743	0.9912591	0.9843856	0.9909207	0.989165
Inverse difference moment normalized	0.9985310	0.9976929	0.9986747	0.9971896	0.9985152	0.998347

Table 2: Training Performance Measure

Classifier (MLP)	Sensitivity	Specificity	Accuracy
Automatic MLP	94.5%	92.3%	93.4%
Traditional MLP	100%	100%	100%

Table 3: Testing Performance Measure

Classifier (MLP)	Sensitivity	Specificity	Accuracy
Automatic MLP	70.5%	87.5%	76%
Traditional MLP	92.3%	91.6%	92%

REFERENCES

[1] Fredrik Georgsson and Tor-BjörnHolmström,Master Thesis inComputing Science "*A Survey and Evaluation of Features for Diagnosis of Malignant Melanoma*" August 2005.

[2] www.cancer.net/patient/Cancer+Types/Melanoma?sectionTitle=Statistics

[3] Kittler et al.H. Kittler, K. Pehamberger, K. Wolf and M. Binder, "*Diagnostic accuracy of dermoscopy, The Lancet Oncology3*"2002, pp. 159–165.

[4] Pehamberger H, Binder M, Steiner A, Wolff K. "*In vivo epiluminescence microscopy: improvement of early diagnosis of melanoma.*"J Invest Dermatol 1993,100:356S–62S.

[5] Bafounta ML, Beauchet A, Aegerter P, Saiag P. "*Is dermoscopy (epiluminescence microscopy) useful for the diagnosis of melanoma? Results of a meta-analysis using techniques adapted to the evaluation of diagnostic tests*". Arch Dermatol 2001,137:1343–50.

[6] McGovern TW, Litaker MS. Clinical predictors of malignant pigmented lesions"*a comparison of the Glasgow seven-point checklist and the American Cancer Society's ABCDs of pigmented lesions*" J DermatolSurgOncol 1992,18:22–6.

[7] H. Harms, H.M. Aus, M. Haucke, U. Gunzer, "*Segmentation of stained blood cell images measured at high scanning density with high magnification and high numerical aperture optics, Cytometry*" 1992; 7; 522-531

[8] W. Stolz, T. Vogth, M. Landthaler, S. Hempfer, P. Bingler, W. Abmayr, "*Differentiation between maligant melanomas and benign melanocytic nevi by computerized DNA cytometry of imprint specimens*", J CutanPathol 1994, 21; 7-15

[9] A.R. Brown, "*Combined immunecytochemical staining and image analysis for the study of lymphocyte specificity and function in situ*" J Immunol Methods 1990; 130: 410-414

[10] P.W. Hamilton, P.H. Bartels, R. Montironi, et al., "*Automated histometry in quantitative prostate pathology, Analytical and Quantitative Cytology and Histology*". 1998 , 20; 443-460

[11] http://www.mathworks.com/

[12] Binder M, Steiner A, Schwarz M, Knollmayer S, Wolff K, PehambergerH: "*Application of an artificial neural network in epiluminescencemicroscopy pattern analysis of pigmented skin lesions:a pilot study.*" Br J Dermatol1994, 130:460-465.

[13] Piccolo D, Ferrari A, Peris K, Diadone R, Ruggeri B, Chimenti S "Dermoscopic diagnosis by a trained clinician vs. a clinician with minimal dermoscopy training vs. computer-aided diagnosis of 341 pigmented skin lesions: a comparative study". Br JDermatol2002, 147:481-486

[14] Blum A, Luedtke H, Ellwanger U, Schwabe R, Rassner G, Garbe C: "Digital image analysis for diagnosis of cutaneous melanoma".

[15] Margarita Sordo"*Introduction to Neural NetworksinHealthcar*e"October 2002.

[16] McGovern TW, Litaker MS. Clinical predictors of malignant pigmented lesions"*a comparison of the Glasgow seven-point checklist and the American Cancer Society's ABCDs of pigmented lesions*" J DermatolSurgOncol 1992,18:22–6.

[17] M. Wiltgen, A. Gerger and J. Smolle "Computer aided diagnosis of melanocytic tumors"Volume 69, Issue 1, January 2003, Pages 17-28

[18] Daniel Ruiz, Vicente Berenguer, Antonio Soriano and Belén Sánchez "A decision support system for the diagnosis of melanoma: A comparative approach" 13 June 2011.

[19] Br J Dermatol"*Development of a highly effective computer algorithm based on analysis of 837 melanocytic lesions*" 2004, 151:1029-1038.

[20] Dermoscopy, Eric EHRSAM, www.dermoscopic.blogspot.com

[21] www.dermoscopyatlas.com

[22] Markus Gipp, Guillermo Marcus, Nathalie Harder, ApichatSuratanee, Karl Rohr,RainerKönig, ReinhardMänner "*Haralick's Texture Features Computedby GPUs for Biological Applications*" IJCS_36_1_09,36:1,

[23] T.Golub, D.Slonim, P.Tamaya, C.Huard, M.Gaasenbeek, J.Mesirov, H.Coller, M. Loh, J. Downing, M. Caligiuri, C.Bloomfield, and E.Lander. "*Molecular Classification of cancer:Class discovery and class prediction by gene expression monitoring,*" scince,1999، 286:531-537

[24] J. A. Freeman and D. M. Skapura, "*Neural Networks, Algorithms, Applications and Programming Techniques*", Addison-Wesley Publishing Company, 2002.

[25] C.M Bishop, "*Neural networks for Pattern Recognition*", Oxford U.K: Clarendon,1995.

[26] S. Haykin, "*Neural networks: A comprehensive Foundation*", 2nd ed. Englewood Cliffs, NJ: Prentice Hall, 1999.

[27] R.M. Nishikawa, M.L. Giger, K. Doi, C.J. Vyborny, and R.A. Schmidt, "*Computer aided detection of clustered microcalcifications in digital mammograms,*" Med. Biol. Eng. Comp. 1995 , Vol. 33, pp. 174-178.

[28] S. Theodoridis and K. Koutroumbas, "*Pattern Recognition*", Academic Press, San Diego, 1999.