Analysis on Skin Lesion Classification Systems and Dermoscopic Feature Analysis for Melanoma

S Mohan Kumar¹, Dr. T. Kumanan²

¹Research Scholar, MAHER, Meenakshi University, Chennai, & Associate Professor, New Horizon College of Engineering, Bangalore, India.
²Principal, Faculty of Engineering and Technology, MAHER, Meenakshi University, Chennai, India

Abstract: Cancer in the recent years is one of the significantly complex diseases for diagnostics. The early detection of a deadly form a cancer ie., skin cancer can be done by using image based computer aided diagnostic systems. These systems have a significant potential to screen and detect malignant melanoma. This paper discusses efficient methods using various algorithms for classification of skin cancer. This paper also reviews the different processing steps involved, the various features extracted and classifiers used to classify as melanoma or benign. We also review and examine the current practices, problems and suggest appropriate modeling methodologies.

Keywords: Cancer, Melanoma, Thresholding, SVM, GLCM, Skin Cancer, Classifier, Segmentation, Image Processing, Texture analysis, Edge detection, ABCD Feature Selection

I. INTRODUCTION

Cancer is a complex disease which is caused by genetic instability and accumulation of multiple molecular alternations. It involves abnormal increase in the number of cells which has the potential to spread to other parts of the body. Cancer is generally described by the body part they originate from. Some body parts contain multiple types of tissue so to be more precise, cancers are classified by the type of cells the tumor cells originate from. Some common types of cancer are breast cancer, lung cancer, blood cancer and skin cancer. Skin cancer is one of the deadliest forms of cancer. Skin protects the human body against heat, sunlight, injury and infection. Skin also plays a vital role in controlling the body temperature and in storage of water and fat. Skin cancer is one of the most frequent and most malignant type of cancer which is caused by excessive exposure to uv rays. The incidence of skin cancer is increasing worldwide especially in countries where the ozone layer is thinning. Now at present, the rate of incidence has increased from 3% to 7% in fair skinned population. It is the fifth most common cancer in men and seventh most common cancer in women. According to a statistics by the skin cancer foundation, 1 in 5 cancers diagnosed is skin cancer. Skin cancer is basically classified into two major types, ie., melanoma and non-melanoma. Non-melanoma occurs at the basal and squamous cells and can be cured if it is detected early and appropriate treatment is administered. Malignant melanoma is the deadliest form of cancer and so studies prove that morality rate can be reduced if detected at an early stage. Malignant melanoma arises from cancerous growth in pigmented skin lesion. The pigments that give color to the skin are called melanocytes and they start in a small region and spread to other areas in the skin through the lymphatic system or blood. Melanin is produced in the outer most layer of the human skin which is the dermis.
This melanin protects the human skin from harmful ultraviolet radiations. Melanoma is capable of deep invasion and it spreads widely via lymphatic vessels and blood vessels. This is very dangerous as it may spread over the whole body very quickly. This emphasizes the need for early detection. The most common diagnosis techniques are visual inspection by a dermatologist. So dermatologist usually relies on experience of visual assessment to distinguish between benign and malignant skin lesions. The differentiation of melanoma from other skin lesions is not an easy task even for a well experienced dermatologist. Some primary care physicians underestimate melanoma at early stages. Here the accuracy is disappointing, this requires high level of experience and expertise to differentiate skin lesions. Since advanced cutaneous melanoma is non-curable, the patient may succumb to this deadly disease. Some of the characteristics of skin cancer are

A. Blue white veil – these are irregular structure less area of conflict having blue pigments with a ground glass haze
B. Multiple brown spots – they may be sun spots or liver spots caused by aging, but it is better to consult a dermatologist as it may even be a sign of skin cancer
C. Pseudo pods – these are temporary cytoplasmic projections of cell membrane
D. Radical streaming – these are irregularly deep pigmented lesions in chronic phase of regression
E. Globules – this is a small globular body or mass
F. Multiple colors – Variety of colors like shades of brown / tan or black can appear. These can also become red / white or blue.
G. Multiple blue / grey dots – those are melanin loaded macrophages
H. Pigmented network – it is an important structure in dermoscopy. The pigmented network is disrupted when there is malignancy

Dermoscopy is a noninvasive diagnostic technique which links dermatology and dermopathology by enabling visualization of morphological features, non-distinguishable with naked eye. Even with accomplished dermatologist use dermoscopy for diagnosis, the percentage for accuracy is only 75 - 80%. The use of medical image processing [36,37] to assist doctors in diagnostics of diseases is referred to as computer aided diagnosis. Advancements in the field of image processing have helped automated diagnosis of various diseases. Computers are very smart but cannot obtain information like color variation, asymmetry and texture features. But recent advancements in the field of image processing have enabled computers to adapt to the advanced requirements of image quantification. This paper focuses on selecting suitable algorithms required during various phases of automatic diagnosis procedure for ensuring the timely diagnosis of skin cancer. We would also discuss the different issues that affects the success of classification.

II. METHODOLOGY

The general methodology used in computer aided diagnosis of melanoma detection using image processing is as follows. The input for the system is the image of the suspicious skin lesion. Preprocessing of the image is done by which the quality of the image may be enhanced by the removal of unwanted noise and structures like hair. This is a main step as it differentiates lesion from healthy skin. The transition between the lesion to healthy skin is smooth, which makes the detection of lesion tough. The dermoscopic images even contain uneven illumination, black frames, ink markings, dermoscopic gel, rulers, air bubbles and intrinsic cutaneous features which may affect border detection. Segmentation refers to partitioning of an image into disjoint regions that are
homogeneous with respect to chosen property such as luminescence color and texture thereby changing representation of images to more meaningful ones. The image segmentation is done by using automatic thresholding process, masking operation and border detection is done. The next step is feature extraction by which the original dataset is reduced by measuring certain properties. The geometrical, color based, wavelet coefficients, dermoscopic features and texture features are the main features considered. The next step is classification through which the lesion is classified as melanoma or benign by various classification algorithms like K-nearest algorithm, decision tree algorithm, logistic regression, artificial neural network and support vector machines.

III. LITERATURE REVIEW
Extensive research has been done on computer aided diagnosis of skin cancer. The input image which is a dermoscopic image is segmented using existing systems which are either manual, semi-automatic or fully automatic border detection methods. Various literature shows different border detection methods which include histogram thresholding, global thresholding on optimized color channels followed by morphological operations, hybrid thresholding. Various features based on shape, color, texture and luminescence are extracted and used in different papers. The ABCD rule of dermoscopy suggests that asymmetry is the most important among the four features of asymmetry, border irregularity, color and diameter. In some techniques, symmetry feature is calculated based on the geometrical measurements on the lesion as a whole. Example, symmetric distance and circularity. According to other studies, circularity index is proposed as a measure of irregularity of borders in dermoscopy images. This paper compares the performance of various classifiers, on the skin lesion diagnostic problem. Di Leo[1] et al, proposes a system for the automated diagnosis of early melanoma using the ELM 7-point checklist. ELM is the epi luminescence microscopy non-evasive technique that uses different light invasive techniques with an oil immersion technique. The 7 point checklist refers to the a typical network pigment network, blue whitish veil, a typical vascular pattern, irregular streaks, irregular pigmentation, irregular dots and regression structures. The input of the Computer Aided Diagnostic (CAD) system will be digital images obtained by ELM, which are processed through different algorithms. Then the images are processed in three main stages in which first the boundary detection is done followed by feature extraction where different morphological and chromatic features are considered, followed by classification. Here the decision tree classifiers belonging to supervised machine learning techniques are used. The decision tree classifier is a predictive model and is preferred as it is fast to train and apply and the rules are easy to understand. V. De Vita[2] et al., proposes a automatic detection system for melanoma which uses statistical techniques and approaches to improve the performance of different algorithms for automatic detection of dermoscopic criteria provided by 7-point checklist method. Here, the boundary detection in done by a technique based on adaptive thresholding and also on an unsupervised approach based on statistical region merging. Feature extraction is done by taking into account the first order low level features. These features are measured by techniques like color segmentation which is a statistical region merging technique belonging to the region growing and merging group and texture extraction which is a combination of two different techniques namely structural and spectral methods. Structural technique is intended to search for primitive structures such as lines or points which can constitute a texture. Classification is done by using decision tree classifiers. Spectral technique is based on Fourier analysis of grey level image. Mabrouk[3] et al., in his paper proposes a automated method for melanoma diagnosis. The input images are a set of dermoscopic images. The features are extracted based on grey level co-occurrence matrix. Here the classifier used is a multilayer perceptron classifier which uses 2 different techniques in training and testing process which is the automatic multilayer perceptron classifier and traditional multilayerperceptron classifier. This comes under the neural network classifiers. Results obtained from this method indicate that the texture analysis is a useful method in diagnosis of melanocytic skin tumors with a high level of accuracy. Karol[4] et al., proposes a computer aided diagnosis – a CAD[Refer 31,32,33] system, which is a decision support system based on semantic analysis of melanoma images. The input isdermoscopic images from Jagiellonian university skin lesions database. The images are then segmented and the objects are extracted which leads to border extraction. The binary border mask is generated and objects are extracted by running simple region growing algorithms. When every object is separated from this image, feature extraction is done as the next step. The color-based features are extracted and classification is done using classification algorithms likes support vector machines and neural networks. Support vector machines is used with four kernels namely linear, polynomial, radial and sigmoid. For Neural networks, radial basis function is used. The classification is done for six object groups which are skin regions, red regions, black regions, light and dark brown regions and grey blue regions. Best results of 98.3% is obtained for objects corresponding to dark brown region of images. The second best is the standard skin region which is 97.5%. The classification accuracy for black regions is 93.89% and for pink red regions it is 94.3%. The least accuracy rate of 80.07% is achieved in the blue grey veil color, because areas covered by them could easily belong to other region types as blue grey veil appear can appear over any other region class. This paper summarizes that the support vector machines with linear kernel prove to perform best in
classifying melanoma. Omar Abuzaghleh[11] et al, proposes a completely utility based system that makes use of a mobile phone application to help in the early detection of melanoma. The whole process has two segments where in the first segment, constant alarms are offered to the clients to prevent the skin smolder’s that are caused by excessive exposure to sunlight. The second segment is that which is a computerized picture investigation module for pictures securing, hair discovery, sore division and highlight extraction. A total of around 200 samples were considered in which an computational accuracy of 96.3% was obtained. Sarika Choudhani[10] et al, proposes a Neural Network based method for the detection of cancer. The dermoscopic images are collected and these are the input image to the CAD[Refer 40,41] system. These images are filtered for removal of noise using various high definition filters and then the segmentation is done using the maximum entropy threshold by which the Region Of interest is calculated. The feature extraction in this particular model is done by GLCM and classification is done using Artificial Neural Networks. The accuracy achieved in this system is around 86.66%

Ezzeddine[12] et al., proposes a system in which the input images are color images of skin lesions. As a first step in the dataset analysis, a preprocessing sequence is done using medial filters to remove noise and undesired elements. The last preprocessing step is the application of Karhunen-Loeve transform that enhances edge detection. Secondly, an automated segmentation approach separates suspicious lesions by region growing after a preliminary step based on fuzzy sets. The lesion ROI and its associated boundary is obtained by gathering pixels into a homogeneous region according to similarity criterion upon the obtained fuzzy set image. The pneumonic device, ABCD is used to describe several features that help to distinguish melanoma from non-cancerous growth. The last step, the classification process is done using artificial neural networks which is a supervised classification. This method possesses universal approximation capability by which it can be approximate any given function with arbitrary precision as long as the number of hidden units is large enough. After defining its architecture, the network is trained with supervision using the back-propagation algorithm. The dataset consisted of 200 images and an accuracy of 79.1% was obtained. N. C. F. Codella[13] et al., proposes a system that combines recent developments in deep learning with established machine learning approaches creating novel methods capable of skin lesion segmentation, analysis of detected area and surrounding tissues for melanoma detection. Hand coded feature extractors, sparse coding methods and SVM’s with more machine learning techniques which include deep residual networks and convolutional neural networks are combined into ensembles whose primary goal is melanoma recognition and segmentation for the given dermoscopy images. The dataset used here was released by the international skin imaging collaboration for 2016.
Once the feature vectors have been extracted, a non-linear support vector machine was used as a classifier over feature vector to discriminate for melanoma. A classification accuracy of 76% when tested on a dataset containing 900 training and 379 testing images is obtained.

![SVM Classifier](image)

SujayaSaha[14] et al., proposes a system in which input image is acquired with a digital camera under consistent lighting. After image is acquired, artifacts such as hair and air bubbles are removed by median filtering and edges are preserved by using another filter called the Gaussian filter. When the Gaussian filter is implemented, it creates a blur which makes it difficult to analyze the lesion. Thresholding is carried out and the image is transformed to a binary image in the form of 0 & 1. Various classifiers are used for melanoma classification. The lesion is segmented and discriminative features like entropy, color compactness, coarseness, radial variance and border are extracted. LequanYu[15] et al., proposes a novel melanoma diagnostic method by leveraging very deep convolutional neural networks. Substantially deeper networks acquire richer and more discriminative features that enable more accurate recognition. Convolutional neural networks with hierarchical feature learning capability have lead to breakthroughs in the field of medical imaging. This novel approach is made up of a two stage approach based on deep CNN. For accurate skin lesion segmentation, there FCRN is used which is the Fully Convolutional Residual Network based on residual blocks. This takes an arbitrary size image as input and outputs an equal sized prediction score mask. The main advantage of FCRN is that it can make pixel-wise prediction which is very vital in skin cancer detection. Next comes the classification stage in which integration of two stages is done. Several networks with different scales are fused to predict the results. A deep residual network is constructed to classify skin lesions based on segmentation results. The softmax classifier and the SVM[Refer 30,42,43,44,45] are employed and averaged to get the final result. An accuracy of 85% had been obtained when tested on a dataset comprising of 900 images as training set and 350 images as testing set. E.Zagrouba[25] at al., proposes an accelerated system for melanoma diagnosis based on subset feature selection. The accuracy obtained was 77.7%. Aaron Par[20] et al., proposes a system that classifies skin cancer methods based on Confocal Raman Spectroscopy using maximum a posterior probability, fuzzy algorithm and support vector machine. The data is then preprocessed to remove various disturbances using different steps which consist of data normalization with minmax method and dimension reduction with principal component analysis. The Raman spectra is the divided into three significant protein bands which correspond to Amid I Mode composed of two bands, NOR and BCC tissue and the Amid III mode. In this particular method, all the features are extracted independently. The test data comprises of around 216 species with a classification accuracy of 97.2%. This is evident proof to show that the three band approach is an effective method in skin cancer detection. Martin Kreutz[16] et al., presents a sophisticated artificial neural network approach for the classification of skin lesions. In this method highly sophisticated image processing, feature extraction, pattern recognition and various other fields from statistics and artificial neural networks are combined to achieve reliable diagnosis. The feature extraction is done by employing a hybrid method that combines statistical clustering of the color space and hierarchical region-growing algorithm. Features based on specialized Gabor wavelets which are scale, translation and rotation invariant are employed. Classification is done using neural networks where mixture of experts architecture and multilayer perceptron is used and an accuracy of 90% is achieved considering a dataset of around 423 cases of different skin lesions. Taouil[19] et al., proposes a CAD system for detection of melanoma. This paper focuses more on the segmentation step in which different methods used in various researches are combined to produce a hybrid approach that could be integrated into the system. The author has used three different methods of automatic segmentation based on...
thresholding, morphology functions and active contours. Snakes are integrated into the entire process of melanoma detection to replace the supervised methods. (Snakes). Nidhal K. EL Abbadi[30] et al proposes a CAD system for melanoma where initially the input images are filtered using Markov and Laplace Filter to remove undesired elements and finally detect the edge of the lesion. The color image is converted to YUV color space and the U channel is selected for processing. Otsu’s thresholding converts image to binary form. The most important step is the classification based on ABCD rule where asymmetry is based on rotation of the lesion and divided into two parts horizontally and vertically followed by which the mismatched pixels are counted. The dataset consists of 220 samples and accuracy obtained was 95.4%. UsmaBanoAnzari[17] proposes a method to detect the cancerous cells by using Gray level co-occurrence matrix and SVM classifier. The input images are taken from a derma scope which is a hand held magnifier used to take pictures of skin lesions. The next step being the preprocessing step in which three steps are involved which are Gray scale conversion, Noise removal and image enhancement. The main goal of the preprocessing stage is to remove unwanted distortions and enhance the image features. The gray scale images have brightness information and each pixel corresponds to the amount of quality of light. The main advantage of gray scale images is that it is more easy and faster to process. Noise removal is done by using the median filter. The median filter is a non-linear filter that leaves edges invariant. Next step involves enhancement in which the images are processed to improve visibility of the feature of interest. The feature extraction uses GLCM for texture image analysis. This is used to capture the spatial dependency between the image pixels. The common features that are considered are contrast, mean, energy and homegenity. The Support Vector Machine Classifier is used for melanoma and benign classification and regression analysis. The accuracy obtained in this system is about 95%.

IV. OBSERVATIONS

A. Various different Features are selected from skin images
B. Classification done using Neural network and SVM classifier
C. Classifier is trained using features to classify skin cancer and normal
D. Performance Measures like sensitivity, accuracy and specificity are computed on testing set to evaluate outcome of proposed method
E. Training and testing data sets utilized to conduct the research should be increased.

V. CONCLUSION

The last two decades has seen a very steady rise in the occurrence of melanoma. So, early detection of skin cancer is essential. If detected at an early stage, skin cancer can be cured if appropriate treatment is administered. Moreover, at an early stage, skin cancer is very economical to treat, while at a late stage, melanoma skin is mostly fatal. Different types of tests are done and the results are helpful in melanoma classification. The main preventive measure to be followed is avoiding excessive exposure to sunlight which is a main source of UV rays which in turn is found to be the main cause of melanoma.

REFERENCES


[40] Dr. S. Mohan Kumar, Statistical Features Based Classification of Micro calcification in Digital Mammogram using Stochastic Neighbour Embedding-International Journal of Advanced Information Science and Technology - ISSN No:2319-2682 Volume 07, Issue 07 , November 2012, Page Numbers: 20-26.(IF:0.621)


