Diagnosis of Skin Cancer Using Image Processing

Ms. P.Gayathri[1], S. Poonkodi[2]
[1]Assistant Professor, Dept of computer science and engineering
[2]Associate professor, Dept of computer science and engineering
[1, 2] Karpagavinayaga college of engineering and technology

Abstract- *The people were affected by different types of cancer, skin cancer becomes apparent in that. In recent years, there is no screening tools for early detection of skin cancer. It can detected by doctor or through certain technique which only capture image or through invasive method like biopsy in abnormal region. Melanoma is considered a fatal variety of skin cancer. However, it is extremely hard to differentiate it from nevus because of their identical look and symptoms. The proposed system is to detect and distinguish melanoma from nevus based on the image processing technique. The Gaussin filter are used to remove noise from the lession skin of the image which is followed by K-Mean clump to the session out of lession. Probabilistic Neural Network are employed for the classification of skin cancer into melanoma and nevus.*

I. INTRODUCTION

Melanoma may be very serious form of skin cancer. According to the American Academy of Dermatology, one person dies from malignant melanoma every hour. The overall incidence rate for the disease is increasing faster than that of any other cancer and an American's lifetime risk of developing melanoma will be one in 75 melanoma is the malignant neoplasm deriving from melanocytes through malignant transformation. One among the foremost prominent functional features of melanocytes is to produce the melanin pigment as a response to UV radiation in order to protect the skin structures from sunburn damage. In about 2% of occurrences, the disease is present even though no skin discoloration occurs, this is often called amelanotic melanoma. When melanoma spreads, it affects other places on skin, lymph nodes, lungs, liver, brain or bones. Such secondary spread is referred to as Metastatic Melanoma Clinical reports indicate same treatments produced very different outcomes on different cases. Early melanoma classification and stage recognition would be critical predicting effectiveness or uneffectiveness of potential treatment. As determination of melanoma type and stage is often crucial to the assignment of appropriate treatment. Genetics researchers have proposed that discrete and previously unrecognizable melanoma taxonomy can be identified by viewing the systematized data from gene expression experiments.

Skin Cancer is that the cancer affecting the skin. It may appear as malignant or benign form. Benign Melanoma is simply appearance of moles on skin. Malignant melanoma is appearance of sores that cause bleeding. Malignant Melanoma is the deadliest sort of all skin cancers. It arises from cancerous growth in pigmented skin lesion. Malignant melanoma is named after the cell from which it presumably arises, the melanocyte. If diagnosed at the right time, this disease is curable.

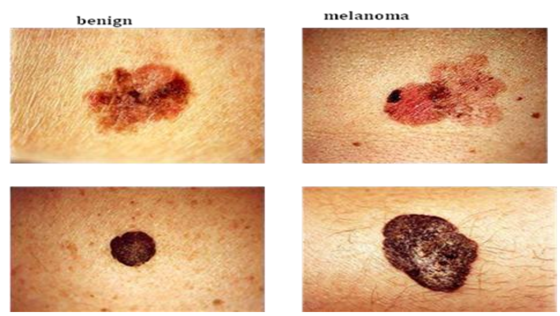


Figure.1 Sample Benign and Melanoma

II. PROJECT METHODOLOGY

In severe cases where the melanoma may have spread to remote tissues, more intense treatment may be administered, including chemotherapy and lymphadenectomy (removal of lymph nodes). Although several visual metrics exist for identifying melanoma, one of the most widely used ones is the ABCD metric. In this analysis, dermatologists attempt to identify:

- **A**symmetry (of color and structure)
- **B**order irregularity
- **C**olor

Diameter **Asymmetry**

It is determined with respect to both color and structure patterns. Two orthogonal axes are determined by the dermatologist who produces the smallest amount of visual asymmetry.

A score of 1 is assigned to each axis if it produced asymmetry and 0 otherwise. The final score is calculated as the sum of the two axis scores. Thus, the final asymmetry score yields A∈ {0, 1, 2}.

Border irregularity

Irregularity with respect to shape and pigmentation is identified along the border. The lesion is visually split into eight radial segments, each of an equivalent size, consistent with eight axes at 45◦ intervals. A score of 1 is assigned to each segment if it portrays border irregularity and 0 otherwise. The final score is calculated as the sum of the axis scores. Thus, the final border irregularity score yields B ∈ {0, 1, 2,...,8}.

Color

Particular colors have been historically observed in melanomas. This color distribution is easily identified using a dermatoscope. The number of unique colors in the lesion is identified and counted. Possible colors include white, red, light/dark brown, blue-gray, and black pigments. Thus, the final color score yields C ∈ {1,
2,...,6}.

Diameter

If the lesion is more than six millimeters in diameter, it is highly likely to be melanoma. This large radius may be a by-product of the radial growth phase.

III. IMAGE PROCESSING TOOLS

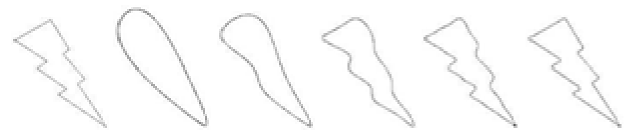
COLOR SPACES AND PERCEPTUAL UNIFORMITY

Color spaces, such as the popular RGB (redgreen-blue) space, are usually three- or four-dimensional spaces that map a physical color to a coordinate in the space, where each axis represents a certain characteristic for that color space. Most images are expressed in the RGB color space, where a color is represented by "adding" certain amounts of red, green, and blue $_{\text{light}}$ to obtain a pixel's color. Color displays are generally made up of individual pixels, each of which is in turn comprised of three RGB light sources. It is therefore a natural fit to use the RGB color space for image color representation, as it is a direct mapping to the output's display.

FOURIER DESCRIPTORS

Fourier descriptors are an application of Fourier theory onto shape analysis. The Fourier transform is widely used to transform a signal into the frequency domain, where the signal is represented as a weighted combination of sinusoidal frequencies.

It is an invertible process, meaning that the frequencies can be uniquely mapped back to the original signal. Given a set of N complex numbers $\{x_i : x_i \in C\}_i$, the discrete Fourier transform (DFT).



(a) Original (b) n_f= 2 (c) n_f= 4 (d) n_f= 8 (e) n_f= 16 (f) n_f= 32

Figure.2 Example of shape reconstruction using Fourier Descriptor

The original border was sampled with 1000 points, and reconstructed using different numbers of frequencies (n_f). Each reconstruction uses the lowest $n_f$frequencies.

$$F_k = \sum_{n=0}^{N-1} f_n \exp(-2\pi ikn/N$$

$$f_n = \frac{1}{N}\sum_{k=0}^{N-1} F_k \exp(2\pi ikn/N$$

Where, F is the set of contributions from the uniformlyspaced frequencies. Thus, F_0 is the DC component (i.e., offset), F_1 is the amplitude of the first low-frequency sinusoidal contributing to the original signal, and so on.

MORPHOLOGICAL OPERATIONS

Morphological operations are techniques for analyzing geometric shapes using set theory. Each operation relies on the definition of a structuring element, which is usually some simple geometric shape, such as a disk or a line. $M_o(A,B)=B_x$

B ⊆A $M_c(A,B) = M_o(A^*,B^*)$

Where, X^* is the complement of X. An intuitive description of morphological opening and closing using a disk structuring element could be easier to grasp. Given a shape A, if we "roll" the disk over the form, the disk will roll over some areas without falling into them. The amount of area that it "rolls over" is dependent on the size of the shape and structuring element. If we fill in that area, we obtain the product of morphological closing. Similarly, if we roll the disk on the interior of the shape, and fill in any gaps in which the disk does not fall, we get the product of morphological opening.

Morphological operations, like Fourier descriptors, are very useful for analyzing shapes in a robust and consistent manner.

IV. PROPOSED SYSTEM

The proposed system has two major components. The first component may be real-time aware of help users to prevent from skin burn caused by sunlight, a equation to compute the time for skin to burn is there by introduced. The second component is an automatic image analysis which consists of image acquisition, hair detection and exclusion, skin segmentation, feature extraction, and classification, where the user getting to be ready to capture the pictures of skin moles and our image processing module will classify under which category the moles fall into benign, a typical, or melanoma. An alert will be provided to the user for medical help if the mole belongs to the atypical or melanoma category. In this system uses PH2 Dermoscopy image database from Pedro Hispano Hospital for the development and testing purposes. The image database contains a complete of 200 dermoscopy images of lesions, including benign, atypical, and melanoma cases.

To help the users avoid skin burn caused by sun exposure, and hence, to prevent skin cancer, our system would calculate the time for skin to burn and the system will deliver a true time aware of the user to avoid the daylight and seek shade to stop developing skin cancer. The system created a model by deriving an equation to calculate the time for skin to burn namely, Time to Skin Burn (TTSB). This model is derived based on the knowledge of burn frequency level and UV index level.

Melanoma is that the deadliest variety of carcinoma. Incidence rates of malignant melanoma are increasing, particularly among non-Hispanic white males and females, however survival rates are high if detected early.

The projected framework has higher segmentation accuracy compared to all or any alternative tested algorithms. GLCM algorithm that is employed to feature extracts the image.

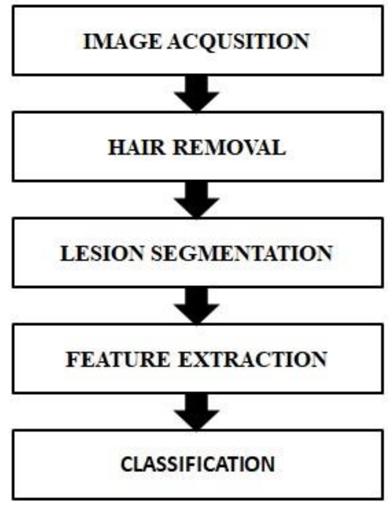


Figure.3 Process of proposed system

The techniques used in the proposed system are,

- Pre-processing - median filter
- Image Segmentation - Texture filters.
- Feature Extraction - Gray Level Co-occurrence Matrix
- Classification - Neural Network (PNN).

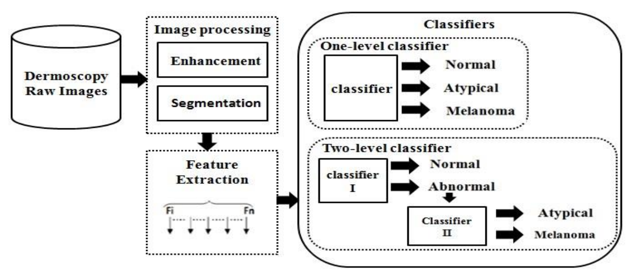


Figure.4 Filter image 3.18

PREPROCESSING

Preprocessing is the process of getting an input image for the process and that image has been converted into a gray scale image. From the Gray scale image, the thresholding can be used to create binary images (digital images that has two possible values for each pixel). It carries the intensity information of black and white. In the preprocessing process, it increases the pixel size from 256 pixel as 512 pixels to show the damaged part clearly in large size. It may also filter the noise within the input image by using the median filter. To reduce the noise like salt and pepper, Gaussian filter etc.

IMAGE SEGMENTATION

Texture is used for segmentation process. Entropy filter is used to create the texture image. Texture is that innate property of all surfaces that describes visual patterns, each having properties of homogeneity. It contains important information about the structural arrangement of the surface, such as; clouds, leaves, bricks, fabric, etc. It also describes the relationship of the surface to the surrounding environment. In short, it is a feature that describes the distinctive physical composition of a surface. Texture features are extracted from Generalized Co-occurrence Matrices (GCM) that is the extension of the co-occurrence matrix to multispectral images. Texture properties include Coarseness, Contrast, Directionality, Line-likeness, Regularity and Roughness.
Texture is one among the foremost important defining features of a image. It is characterized by the spatial distribution of gray levels during a neighborhood. In order to capture the spatial dependence of gray-level values, which contribute to the perception of texture, a two-dimensional dependence texture analysis matrix is taken into consideration. This two-dimensional matrix is obtained by decoding the image file; jpeg, bmp, etc. The most popular statistical representations of texture are Cooccurrence Matrix, Tamura Texture, Wavelet Transform.

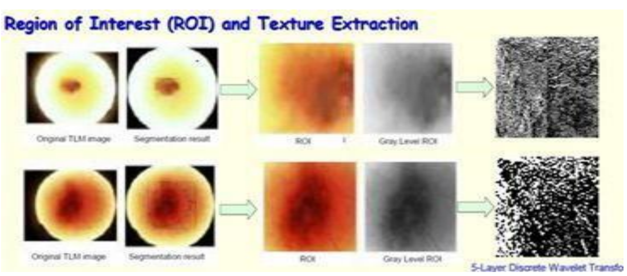


Figure.5 Texture Extraction

FEATURE EXTRACTION

Feature extraction involves performing specific pre-defined calculations on the preprocessed and segmented image. The goal is to generate a feature vector (i.e., a vector of real numbers) from the image that aptly describes important characteristics of the image. Thus, each image is represented by some extend in some ndimensional space (called the feature space).

The goal of feature extraction is to construct a feature space such that the inherent image classes (e.g., malignant and benign) are easily separable during this space. Mathematically, given image I, feature extraction is that the mapping f: I →7 F ∈R^n, where F is computed feature vector. Many carcinoma feature extraction methods propose features that model the ABCD criteria employed by doctors in clinical settings.

Feature Extraction is completed by the Grey Level Co-occurrence Matrix. It's also called GLCM. The cooccurrence matrix originally proposed by R.M. Heraldic, the co-occurrence matrix representation of texture features explores the grey level spatial dependence of texture. A mathematical definition of cooccurrence matrix is as follows,

- Given a position operator **P (i, j)**,
- let **A** be an n x n matrix
- whose element **A[i][j]** is the number of times that points with grey level (intensity) **g[i]** occur, in the position specified by **P**, relative to points with grey level **g[j]**.
- Let **C** be the **n x n** matrix that is produced by dividing **A** with the total number of point pairs that satisfy **P**. **C[i][j]** is a measure of the joint probability that a pair of points satisfying **P** will have values **g[i]**, **g[j]**.
- **C** is called a co-occurrence matrix defined by **P**.
- Examples for the operator **P** are: "**i** above **j**", or "**i** one position to the right and two below **j**", etc. This can also be illustrated as follows… Let t be a translation, then a co-occurrence matrix C_t of a region is defined for every greylevel (a, b).

CLASSIFICATION

Neural networks are predictive models loosely supported the action of biological neurons. The extracted part of the skin lesion is classified by the Probabilistic Neural Network (PNN). It has many classifiers for identifying the correct type of the skin cancer. It compares with the feature set images.

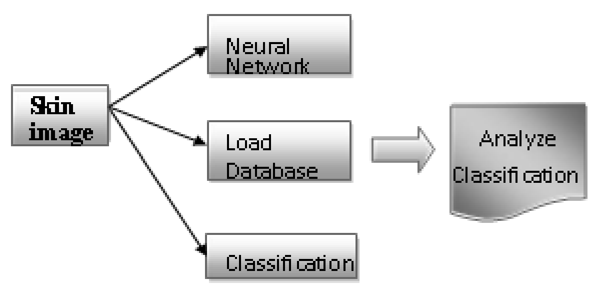


Figure.6 Classification

EVALUATION METRICS

The Evaluation metrics for both segmentation and classification performance in our method we used, include sensitivity (SE), specificity (SP), accuracy (AC), Jacquard index (JA) and Dice coefficient (DI). The organizer first calculated these criteria for each test dermoscopy image and then averaged each criterion on the whole testing dataset to get the

$$AC = \frac{N_{tp} + N_{tn}}{N_{tp} + N_{fp} + N_{fn} + N_{tn}},$$

$$SE = \frac{N_{tp}}{N_{tp} + N_{fn}},$$

$$SP = \frac{N_{tn}}{N_{tn} + N_{fp}},$$

$$JA = \frac{N_{tp}}{N_{tp} + N_{fn} + N_{fp}},$$

$$DI = \frac{2 \cdot N_{tp}}{2 \cdot N_{tp} + N_{fn} + N_{fp}}$$

Where N_{tp}, N_{tn}, $N_{fp}$and N_{fn}denote the number of true positive, true negative, false positive and false negative, respectively, and they are all defined on the pixel level. A lesion pixel is considered as a true positive if its prediction is lesion; otherwise it is regarded as a false negative. A non-lesion pixel is considered as a true negative if its prediction is non-lesion; otherwise it is regarded as a false positive. Participants are ranked based on the results of JA, as it is generally considered as the most important criterion for segmentation. In this case, the false positive rate should be relatively small and the true negative rate should be relatively large, resulting in most points fall in the left part of the receiver operating characteristic (ROC) curve.

V. RESULTS AND DISCUSSSION

The roundness, skewness, diameter and the kurotsis of the melanoma can be determined by the image processing techniques. The accuracy obtained by the proposed system is 97% and the time consumption is very less compared to the existing system.

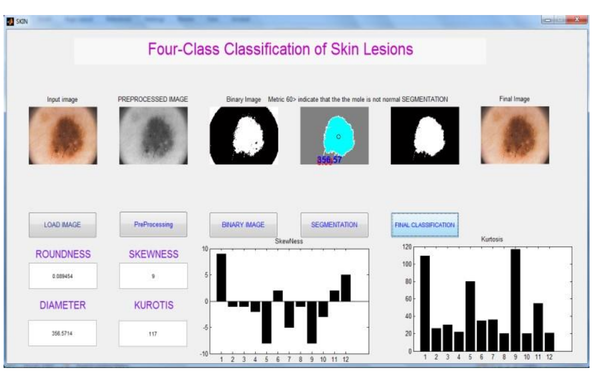


Figure.7 Final Classified Images KURTOSIS

Measure of the degree of peachiness of a distribution. In some cases, a distribution may have its values concentrated near the mean so the distribution has large peak. In other cases, the distribution may be relatively flat. It gives about the central peak is high & sharp or short & broad.

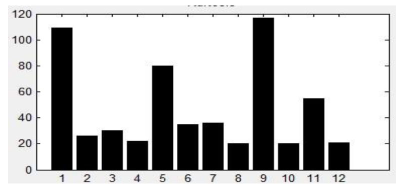


Figure.8 graph of kurtosis

SKEWNESS

Skewness is a measure of the asymmetry of a histogram. A distribution is said to be symmetric if the left and right of the center point is same. If longer tails occur to right the distribution is said to be skewed to right, while if the tails occur to the left it is said to be skewed to the left. Skewness can be defined as the ratio of the third cumulate K3 and the third power of the square root of the second cumulate.

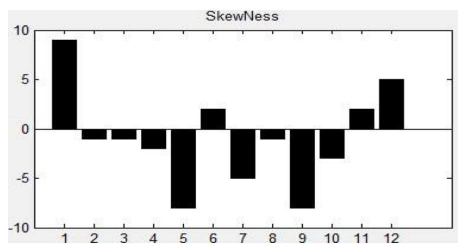


Figure.9 graph for skewness

VI. CONCLUSION

A Computer based early detection of skin cancer system is proposed. It proves to be a far diagnosis method than the traditional Bioscopy method. The diagnosing methodology uses Digital Image Processing Techniques and Artificial Neural Networks for the classification of melanoma from other skin diseases. Dermoscopic images were collected and that they are processed by various Image processing techniques. The cancerous region is separated from healthy skin by the tactic of segmentation. The unique features of the segmented images are extracted using 2-D Wavelet Transform. Based on the features, the images are classified as Cancerous and Non-cancerous.

Orientation estimation and correction is applied to detect low contrast and fuzzy streak lines and the detected line segments are used to extract clinically inspired feature sets for orientation analysis of the structure.

VII. FUTURE ENHANCEMENT

The future scope of the skin cancer detection is more accurate and efficient. The system may provide the encryption of data and authentication for the users so that there is no unauthorized access of the data of the patient, because if there is unauthorized access is performed on the data then the data integrity may be lost. In future it's more interactive and use friendly for checking the lesion that if it is cancerous or not.

REFERENCES

[1] HoTak Lau and Adel Al-Jumpily, Automatically Early Detection of SkinCancer: Study Based on Neural Network Classification, InternationalConference of Soft Computing and Pattern Recognition, IEEE, pp 375-380, 2009.

[2] FikretErcal, AnuragChawla, William V. Stoecker, Hsi-Chieh Lee, and Randy H.Moss, Neural Network Diagnosis of Malignant Melanoma FromColor Images, IEEE Transactions on Biomedical Engineering. vol. 41, No. 9, 1994

[3] Ali AI-Haj, M.H.KabirWavelets Pre-Processing of Articial Neural Networks Classifiers, IEEE Transactions on Consumer Electronics, Vol. 53 Issue2, pp. 593-600, 2008.

[4] PankajAgrawal, S.K.Shriwastava and S.S.Limaye, MATLAB Implementation of Image Segmentation Algorithms, IEEE Pacic Rim Conference.M. E. Celebi, et al., "A methodological approach to the classification of dermoscopy images," Comput. Med. Imag. Grap., 2007.

[5] E. Ahn, et al., "Automated Saliency-based Lesion Segmentation inDermoscopic Images," EMBC, 2015.

[6] M. E. Celebi, et al., "Automatic detection of bluewhite veil and related structures in dermoscopy images," Compute. Med. Image. Grip. 2008.

[7] N. Situ, et al., "Evaluating sampling strategies of dermoscopicinterestPoints," ISBI, 2011.

[8] C. Barata, et al., "Improving dermoscopy image classification usingColor constancy," IEEE JBHI, 2014.

[9] C. Barata, et al., "Two systems for the detection of melanomas inDermoscopy images using texture and color features," IEEE SystemsJournal, 2014.

[10] C. Barata, et al., "Evaluation of color based key points and features for the classification of melanomas using the bag-of-features model," in Advances in Visual Computing, 2013.

[11] Suez, et al., "Model-based classification methods of global patterns in dermoscopic images," IEEE TMI, 2014.

[12] M. Hayat, et al., "Reverse training: An efficient approach for image set classification," ECCV, 2014.

[13] T. Lee, et al., "Dull razor? A software approach to hair removal from images," Comput. Biol. Med., 1997.

[14] F. Riaz, et al, "Detecting melanoma in dermoscopy images using scale adaptive local binary patterns," EMBC, 2014.

[15] D. G. Lowe, "Distinctive image features from scaleinvariant key points," Int. J. Comput. Vision, 2004.

[16] F. Perronnin, et al., "Fisher kernels on visual vocabularies for image categorization," CVPR, 2007.

[17] C.-C. Chang, et al., "LIBSVM: A library for support vector machines,"ACM TIST, 2011.

[18] T. Mendonca, et al., "PH 2-A dermoscopic image database for researchAnd benchmarking, "EMBC, 2013.

[19] J. S. Marques, et al., "On the role of texture and color in the classification of dermoscopy images, "EMBC, 2012.