Implementation of Feature Selection Method to Diagnosis of Skin Disease By using Classification Techniques

T.D.Srividya, V. Arulmozhi

Abstract: ***Skin cancer detection develops into a thought-provoking concern in identifying the correct location in the skin. Image processing occupies a significant part in diagnosis. For initial diagnosis of skin cancer, this image analysis feature is applied. Skin cancer appears to be the most mutual disease in recent years. Initial diagnosis decreases the death rate. This is accomplished by combining feature extraction and segmentation methods. The best appropriate feature subclasses of skin lesion like color, texture are removed and accepted for classification. A method is developed to choose features from huge data sets which contain more inappropriate or repeated features. In this paper, a three-step procedure aimed at selecting features is proposed. The system in the Genetic Algorithm is a classical machine learning technique originates its routine from a representation of evolution. Improving the quality of the image for perception is the main task of GA in order to improve the Feature selection, when applied to classification for skin cancer images, signifies the importance of feature selection while decreasing the number of features required to design classifier. For classification to improve, the images of the affected skin are removed from the normal skin. An automatic skin growth classification scheme is proposed with the associations of skin lesions through the ANN training network.***

Index terms: ***classification, feature extraction, genetic algorithm, segmentation.***

I. INTRODUCTION

Abnormal growth on skin with unpredictable marks of malignancy is skin cancer. If skin cancer is not diagnosed early results in death. The death rate of skin cancer is alarming every year. One method of reducing the death rate is to detect at an early stage. Traditional techniques are inefficient.

Malignancy of skin looks like tumors as benign or malignant. Malignant is dangerous among other cancer types. All types of tumors seem to be matching but malignant differs in terms of irregularity, border anomaly, color dissimilarity and diameter [1]. Automatic skin cancer is a classification method that differentiates malignant skin cancer from other cancer. AI is used for the classification perspective.

The images are transformed into digital format. The noises in the image are removed by pre-processing techniques[2]. Feature extraction is done to reserve the edges [3]. GA has the capability to govern the optimum segments after segmentation or select certain features apt to the scope of heuristic edges.

Experimental outcomes designate that feature selection eradicates large sum of redundant features.

II. LITERATURE SURVEY

Sanjay Jaiswar et al [1] specifies a methodology based threshold-based segmentation, clustering and edge detection based are employed after pre-processing. The ABCD and TDS are estimated aimed at feature abstraction. The suggested work afford encryption of data and validation for users. A useful communicating and accessible system is projected at some point.

Saudamini S. Jivtode et al [2] advocated a process of removing hairs and air bubbles by Dull Razor Filtering, changing to a grayscale image, enhancing contrast, filtering noise and Max Entropy segmentation for segmentation. The specificity and sensitivity of NN by Raman scales remained familiar for the exertion.

Saranya et al[3] propose a division method in which ABCD rule used for preprocessing. Clustering is grounded on the unsupervised approach. Segmentation based on ANN and K- Mean besides Fuzzy C-Mean clustering discussed. An assessment of the entire accessible subdivision methods is presented here.

Lucia Ballerini et al [4] recommend a system that employs classifier based on K-NN and the color and texture skins removed commencing a lesion. Related to this, 5 various skin lesion modules, the patterns intended for numerous classifications are assembled as 3 core groupings of color, texture equations are given. Individually three experimentation is carried out and mean and standard deviation are given.

Ammara Masood et al [5] reports statistics and results from most implementations. He suggests that one factor that complicates the detection of the border is insufficient contrast. The sensitivity ratio is given for various dermoscopic techniques. Various shape features are mentioned. Several classification algorithms are proposed.

Ron Kohavi [6] considers the comprehensibility and accuracy of concepts for irrelevant and weekly relevant features. The practical algorithms in machine learning, the degree of performance when unnecessary features are predicted for output are dealt with in this work. the method of selection of a good subset of features for machine learning, statistics and pattern recognition are dealt with. Several algorithms are used. The evaluation function is calculated. He suggests that the search stops when five node

Revised Manuscript Received on July 07, 2019.

T.D.Srividya, (Ph.D) Computer Science, TiruppurKumaran College Post box no. 18S.R.Nagar, Mangalam Road, Tiruppur

Dr.V. Arulmozhi, Associate Professor, TiruppurKumaran College Post box no.18S.R.Nagar, Mangalam Road Tiruppur

expansions do not yield improved performance of more than 0.1%. Finally, the algorithms discussed attempts to improve prediction accuracy.

Sri Krishna et al [7] introduce the feature subset for optimal clusters with Automatic Feature Subset Selection(AFSGA) using GA. Clustering is classified as hierarchical and partitional. This algorithm contracts with choosing the ideal initial sources and deals with the development of feature subset selection while clustering by selecting CS measure as the fitness function. A new Genetic Algorithm based technique of selecting features by wrapper method in addition to the classification of hyperspectral appearance data using Support Vector Machine is proposed. The AFSGA methods and procedures are given. In the result, the AFSGA is compared with the classical clustering algorithm. The results are demonstrated to improve the efficiency of AFSGA.

M Chaitanya Krishna et al [8] uses CAD to differentiate normal and melanoma skin cancer lesion. Pre- processing of the image is done based on image brightness equalization, normalization of color series, and correction of image scale or image resolution regularization. The segmentation techniques used are threshold based, edge detection based and clustering techniques. The step and change detection are employed here and TDS formula is given.

Munya A Arasi et al [9] projected a review paper for Proximal Support Vector Machine(PSVM). The anticipated system here is the Neuro- Fuzzy system & PNN. The K-Nearest Neighbor, Artificial Neural Network and Support Vector Machines are adapted.

Nabin K.Mishra et al [10] deliberates the numerous characteristics of subdivision of lesion. Artificial Neural Network, Support Vector Machines, and Several cataloging techniques are employed here.

Asmaa Aljawawdeh et al [11] proposes a classification using NN backpropagation and threshold for selected values. K-Means algorithm is used in this work. A classification system is built which distinguishes malignant from benign. A pseudo- code methodology is given. Formulas are given and three segmentation models are proposed.

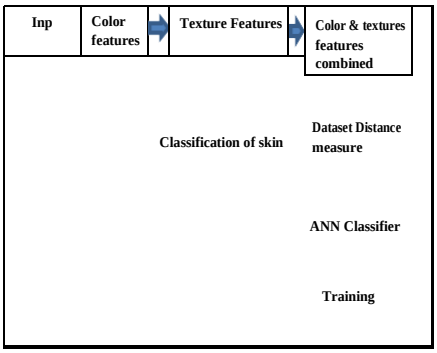
Alessia Amelio et al [12] suggests the ability to differentiate injuries the of skin for separating the color image centered on genetic algorithm. It is proved from the investigative findings for isolating cancer of the skin, the methodology for subdivision is capable of finding lesion boundaries exactly joined with a method of integration of adjacent region. CAD methods are used. Results presented on C-GeNcut. He concludes the experiments to various imageries of dermatology for the noble presentation of the methodology.

III. EXISTING SYSTEM

An effective way to recognize skin cancer in the initial phase, digital descriptions of skin cancer remains explored. For achieving this objective, the process of feature extraction is well-thought of as important for image analysis. For analyzing image accurately segmentation is crucial ever since it disturbs the accurateness of successive phases. Due to unlimited variations of lesions in terms of outlines, dimensions, and shades for various kinds of texture and different skin, segmentation will be difficult[4].

IV. PROPOSED METHODOLOGY & RESULTS

Proposed work includes the implementation of three main steps on the skin lesion. The first step is preprocessing to remove inappropriate features such as hairs, air bubbles, and noise. In the subsequent step, the features are clustered using K-Means to eliminate redundancy, and the next phase is a classic algorithm for selecting the features[5]. An illustration of the projected scheme is in fig(1).



Fig(1) Block diagram of the proposed system

A Pre-processing

Preprocessing of skin cancer involves:

- Get the image as input
- Transform it into a grayscale image.
- Filter the image to remove hairs, noise, and air bubbles.
- The filtered image is resized.
- Obtain image contrast for enhancement.

B Image Segmentation Using A Genetic Algorithm

In a segmented image, all regions fulfill the properties of connectivity and homogeneity. A region is assumed to be homogeneous if all pels in the region fulfill the condition of homogeneity like color and texture[6]. As skin cancer imaging has a very big solution space, a genetic algorithm is used here[7]. The edge detection method is employed in this work[8].

C Feature Extraction

The features extracted are color, texture and GLCM features. Few significant properties related are height, width, irregularity, and location. In this work, we have mined the skin lesion from the image and analyzed its area of irregularity, area, and length. The elementary process of feature extraction is depicted.

D Classification

Depending on the lesion, the features are extracted. Centered on the shape, the texture of the skin, the image is classified as benign or not.

The projected work employs the abstraction of color features in skin images[9]. The RGB color space is employed here because RGB gives good accuracy in case of

skin images. The texture features are mined using Gray Level Co-occurrence Matrix(GLCM)[10]. Other features related to texture as energy, entropy, contrast are also retrieved [11]. To enhance contrast gamma correction method is used.

In combining color and texture features by distance metrics method, the estimated distance of input and database image are analyzed [12]. ANN classifier is utilized for classification purpose to analyze input and training set. The testing stage will be the input image and all images are training images. Ultimately, the image is differentiated as different types.

V. NUMERICAL ILLUSTRATION

A sample of 5 benign and 5 malignant images are considered as shown in the figure. The pre-processing of the images are carried out and the grayscale of these images are shown. The color feature is calculated for benign and malignant images and provided with statistical ratios. A sample of benign and malignant images are taken for texture feature classification. The texture features considered are the mean and standard deviation. A graph of texture features is plotted for these sample images along with values.

Fig (2a) Benign Skin cancer images

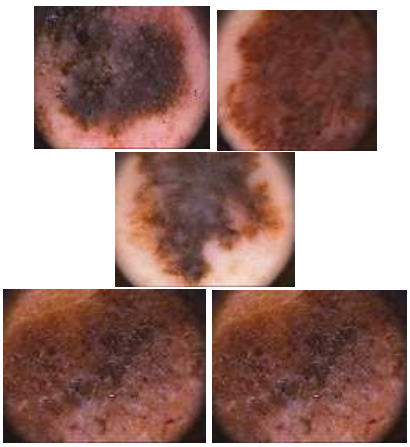


Fig (2b) Malignant Skin cancer images

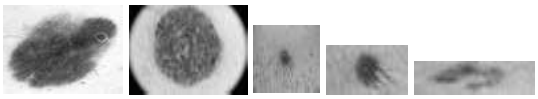


Fig (2c)Grayscale images for Benign cancer after cropping

Fig (2d) Grayscale images for Malignant cancer after cropping

Fig (3) Applying gamma correction to Benign skin cancer image

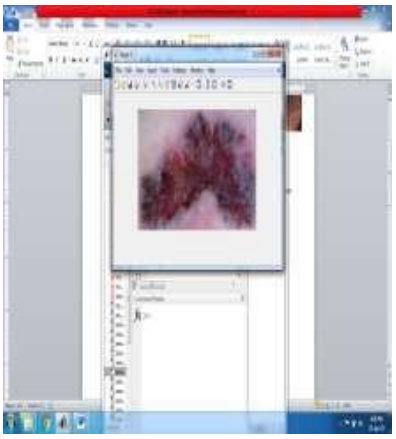


Fig (4) Applying gamma correction to the malignant image to enhance contrast

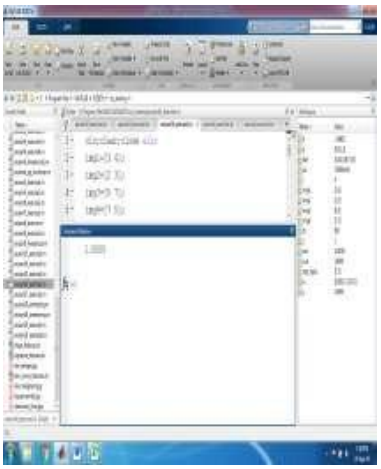


Fig (5) Sample data set for classification

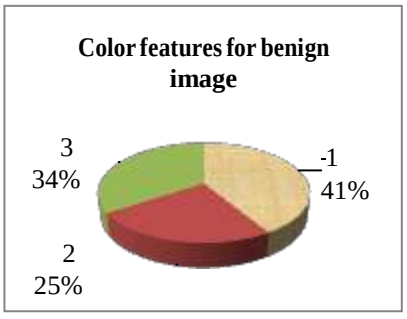


Fig (6) Chart depicting color features for benign

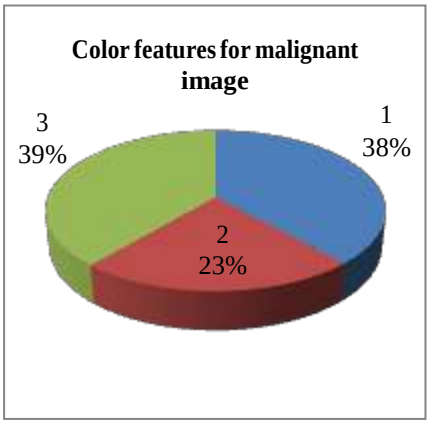


Fig (7) Chart depicting color features for malignant

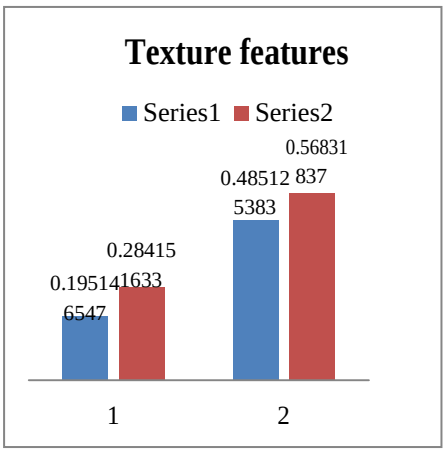


Fig (8) chart illustrating the texture features for benign and malignant images

VI. CONCLUSION

Genetic algorithm has various benefits in procuring the optimized solution. In a vast solution space, GA proves to be the most dominant optimization procedure. Several tasks starting from image contrast, elaborate enhancement, to intricate filters are used in this pattern. Improvement of the image is employed to increase the excellence of digital image, irrespective of understanding of the degradation source. The future enhancement stretches to results of training and testing.

REFERENCES

1. Sanjay Jaiswar, Mehran Kadri, Vaishali Gatty "Skin cancer detection using digital image processing" IJSER ISSN (Online): 2347-3878
2. Saudamini S. Jivtode1, Amit Ukalkar2 " Neural network based detection of melanoma skin cancer" International Journal of Science and Research (IJSR) ISSN (Online): 2319-7064
3. Saranya D, M Malini " A review of segmenation techniques on melanoma detrection" ISSN:2277 128X, Volume 5, Issue 4,2015
4. Lucia Ballerini, Robert B. Fisher , Ben Aldridge,Jonathan Rees "Non-melanoma skin lesion classification using colour image data in a hierarchical k-nn classifier
5. Ammara Masood and Adel Ali Al- Jumaily,"Computer aided diagnostic support system for skin cancer: a review of techniques and algorithms", International Journal of Biomedical Imaging Volume 2013, Article ID 323268, 22 pages http://dx.doi.org/10.1155/2013/323268
6. Ron kohavi, "Feature subset selection as search with probabilistic estimates",AAAI Fall Symposium-1994
7. A. Srikrishna1, B. Eswara Reddy2, V. Sesha Srinivas1 "Automatic feature subset selection using genetic algorithm for clustering" Int. J. on Recent Trends in Engineering and Technology, Vol. 9, No. 1, July 2013
8. M.Chaithanya Krishna 1, S.Ranganayakulu 2, DR.P.Venkatesan3"Skin cancer detection and feature extraction through clustering technique" ISSN(Online): 2320-9801
9. Munya A. Arasi, 2El-Sayed A. El-Dahshan, 3El-Sayed M. El-Horbaty, 4Abdel-Badeeh M. Salem "Malignant melanoma detection based on machine learning techniques: a survey"Egyptian Computer Science Journal (ISSN: 1110 – 2586) Volume 40 – Issue 03, September 2016
10. Nabin K. Mishraa and M. Emre Celebib a Stoecker & Associates, Rolla, MO, USA- "An overview of melanoma detection in dermoscopy images using image processing and machine learning"
11. Asmaa Aljawawdeh, Esraa Imraiziq, Ayat Aljawawdeh "Enhanced k-mean using evolutionary algorithmsfor melanoma detection and segmentation in skin images", (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 8, No. 12, 2017
12. Alessia Amelio, Clara Pizzuti "Skin lesion image segmentation using a color genetic algorithm", Inst. for High Perf. Comp. and Networking(ICAR), Via P. Bucci 41C, 87036 Rende(CS), Italy

AUTHOR'S BIOGRAPHY

Mrs.T.D.Srividya completed her M.Phil(Computer Science) from Mother Teresa Women's University in 2015. She is currently a Ph.D Student under the guidance of Dr.V.Arulmozhi,. Her research is centered on Skin Cancer detection- A Genetic algorithm approach.

Dr.V.Arulmozhi, Associate Professor & Research Head, Department of Research – Ph.D Computer Science, TiruppurKumaran College for women, Tiruppur, has received best paper awards and published more journal papers, conference proceedings and guiding M.phil and Ph.D research scholars. Her specialization is Chemoinformatics, Machine learning techniques, Soft Computing and Evolutionary Computational informatics.