International Journal of Engineering & Technology, 7 (2.4) (2018) 131-135

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research Paper

Detection of skin cancer- A genetic algorithm approach

T. D. Srividya [1] *, Dr. V. Arulmozhi [2]

[1] Ph. D, Associate professor, Research Head & Supervisor, Department of Research-Ph. D Computer Science
[2] Tiruppur kumaran College for women, Tiruppur
**Corresponding author E-mail:*

Abstract

In the present scenario skin cancer is found highly risk in human beings. Many forms of skin cancer are affecting the human life. Among the form of skin cancer the unpredictable diseases is Melanoma cancer. Skin cancer the fatal form is primarily diagnosed visually leads to death, if not diagnosed in its early stage. It can be identified by tedious lab testing with more time and cost. There are vast numbers of computational techniques helpful to predict diseases. A challenging task in skin lesion classification is due to the smooth variation, in the appearance of skin lesions. Image processing techniques like segmentation is used in medical science to identify the region of significance.. This paper focuses Genetic algorithms by means of adaptive parameters (adaptive genetic algorithms, AGAs), an important and promising alternative to genetic algorithms. The extent for accurate solution and convergence speed is significantly measured by employing of crossover along with mutation from which genetic algorithms appear.

Keywords Segmentation; Lesion; Genetic Algorithm; Mutation; Crossover; Feature Extraction; Feature Selection.

1. Introduction

Skin is the essential part of our human body which protects the internal parts from the outside world. The skin gets affected by various factors such as exposure to sun (Ultraviolet radiation), sunlamps, medicines which increase the risk of skin cancer [1]. Dermatology is a deviation of medicine that is concerned with diagnosis of skin cancer. Cancer can be defined as a category of diseases or disorders characterized by uncontrolled partition of cells and the capability of these cells to attack other tissues [2]. According to cancer institute one out of five Americans develop skin cancer in their life span [3].skin cancer is one among the rapid growing cancer types and most dangerous in the world, the important is malignant melanoma. Skin cancer occurs from non-pigmented cells. Therefore the premature diagnosis of melanoma is a critical problem to dermatologists. The death rate of melanoma is three times than other forms [4]. The foremost step towards a detailed study of skin lesion is to distinguish the lesion from the healthy skin. It has been observed that dermoscopy images often include artifacts such as uneven illumination, dermoscopic gel, black frames, ink markings, rulers, air bubbles, and built-in cutaneous features that can affect texture[2]. Thus, it requires several preprocessing steps to make segmentation process possible by the removal of unwanted objects. Segmentation partition an image into its vital images. The paper is structured as follows: Image preprocessing is explained in Section 3, Segmentation in section 4, the Feature extraction is presented in section 5, feature selection in section 6 and section 7 focuses on Adaptive Genetic Algorithm (AGA) with adaptive parameters.

2. Literature Review

Nisha et al [1] proposed a method for histogram equalization which will share out all the colors of the image in a way that the brightest spot of the processed image will be color and the darkest point will be black. Saudamini S. Jivtode et al [2]suggested a method for which filtering is done by Dull Razor filtering for removing hairs and air bubbles in the image, converting to gray scale image, contrast enhancement , noise filtering , segmentation using Max entropy threshold. The sensitivity and specificity by NN of Raman spectrum were identified in this work.

Nishima sachdeva et al [3] analyses the ABCD feature extraction and used Otsu segmentation and PCA method. The proposed scheme uses wavelet transformation for improvement in image, denoising and histogram analysis. In this paper GA are used to select the concurrent important features as input to ANN and determine the optimal number of hidden node automatically. Here PCA method is used for classification.

Ebtihal Almansour et al [4] devised a methodology that is compared with the state of the art methodology. The four color feature formulas are given in this method. The GLCM method is used. The classical statistical texture textures like entropy, energy, contrast, correlation & homogeneity are discussed here. Analysing the results of GLCM feature set, LBP, TF, CF, TCF it is proved that color is an important factor to distinguish melanoma & non-melanoma.

Lucia Ballerini et al [5] proposed a methodology which uses K-NN classifier and the color with texture feature extracted from skin lesions. For the study he took 5 different classes of skin lesions the schemes for multiple classifiers are grouped into 3 main categories. Three experiments are conducted independently and performance reported as mean and standard deviation over 3 experiments.

Munya a Arasi et al [6] introduces a survey paper in which the classification methods such as K-NN, ANN & SVM are analysed.Proximal Support Vector Machine (PSVM) was adopted in this work.

Savita ET a l [7] analyses the problem of finding Maximum Clique-Problem to find the solution largest subgraphs connected to one another. She developed a new technique to find the maxi-

mum clique problem. The operators of crossover and mutation of GA are employed and given by formula. In this paper the GA and graph theory concepts are merged to optimize the problem.

Anil Jain et al [8] uses SFFS for feature subset selection. It is based on SAR satellite images using 4 different texture models. The geometric pattern recognition technique is used here. The crossover and mutation of GA methods are used. The Branch and Bound algorithm is employed. For training the network Squared-error cost function are used. The SFS and SBS algorithms are compared. It is proved that SFFS shows good performance on very high dimensional problems. The goal is to decide whether the classification error reduced by applying feature selection.
Saudamini S. Jivtode et al [9]suggested a method for which filtering is done by Dull Razor filtering for removing hairs and air bubbles in the image, converting to gray scale image, contrast enhancement , noise filtering , segmentation using Max entropy threshold. The sensitivity and specificity by NN of Raman spectrum were identified in this work.
A S Deshpande et al [10] put-forth a methodology in which the image is pre-processed using median filter fro noise removal. For segmentation Fuzzy C-Means is used. GLCM is used and classification is done by SVM. He concludes that SVM always correct.
Ruchika Sharma et al [11] proposed segmentation methods such as edge detection, thresholding, region based , based on clustering. The unsupervised learning algorithms such as clustering by K-Means & Fuzzy C-means are used.ANN segmentation also used. This work shows the comparision between all available segmentation techniques.
Mohammed H.Sadaaghi et al [12] introduces an adaptive genetic algorithm in which Bayes classifier is used. The feature selection techniques employed are SFS, SBS, SFFS He concludes that SFFS is dominant in classification error. The GA approach is focused mainly in this work which includes three reasons for subset feature selection. Another coding method is employed in which integer values are used. The population diversity equation is given and the selection policy is cross generational. A simple multi-point crossover operator is applied. He concludes that AGA is applied to optimize the results of feature subset selection algorithm. And it is proved from this work AGA yield an improvement in total error rate.
Damogoj Jakbovic et al [13]employs a Self-Contained GA with steady-state selection. This alternative to genetic algorithm utilizes empirically based methods for calculating its control parameters. The strength of GA has two major goals. He analyses two characteristics are optimizing multimodal functions. He concluded that the balance between these characters affect the way the genetic operators are performed.
Eric Pellerin et al [14] uses parameter setting prior to running a GA and remains unaffected during execution. He suggested the GA's basic mechanism depend on choice of several key parameters. He analyzed a high crossover rate and a low mutation rate might be very good in exploration of new solutions. A hypothesis was also given in this work. The Travelling Salesman Problem is used for training.
Sanjay Jaiswar et al [15] given a method in which after image pre-processing the segmentation techniques of threshold based, clustering techniques, edge detection based are used. And the feature extraction features such as ABCD and TDS is calculated. The proposed work amy provide encryption of data & authentication for user's A more interactive and user-friendly system is proposed in the near future.
Ruchika Sha rma et al [16] proposed segmentation methods such as edge detection, thresholding, region based, based on clustering. The unsupervised learning algorithms such as K-Means clustering & Fuzzy C-means are used.ANN segmentation also used. This work shows the comparision between all available segmentation techniques.

3. Image pre-processing

The most important step in pre-processing is a detailed examination of pigmented skin lesion. Detection of the lesion is a major problem in dermoscopic images, since the transition between the lesion and the surrounding skin is smooth and challenging even for trained dermatologist to diagnose accurately. The dermoscopy images often contain artefacts such as dermoscopic gel, uneven illumination, gel, ink markings, black frames, air bubbles as well as intrinsic cutaneous features[2][4]. These features can affect border detections of hairs, skin lines and so on. The border detection is made complicate by the irrelevant features which leads to loss of accuracy.

3.1 The menzies method

A popular approach of Pattern analysis has become significant because of its use nowadays. A simplified dermoscopy method was adopted by menzies[1] for inexperienced clinicians to detect and diagnose melanoma. The above method gives a sensitivity and specificity of 85% and 99% accordingly[2]. This has enabled the dermatologists to increase their sensitivity for the identification of melanoma. In Table 1, the Menzies method is shown. For identifying melanoma it should include one or more than one positive features and two negative morphological features.

Table 1: The Menzies Method

Negative Features(neither can be found in melanoma)	Positive Feature(Atleast one feature found)
Symmetry of Pigmentation pattern Presence of merely a single color	Blue-White veil Multiple brown dots Pseudopods Radial steaming Scar like depigmentation Peripheral black dots Multiple 5-6 colors Multiple Blue-Grey dots Broadened network

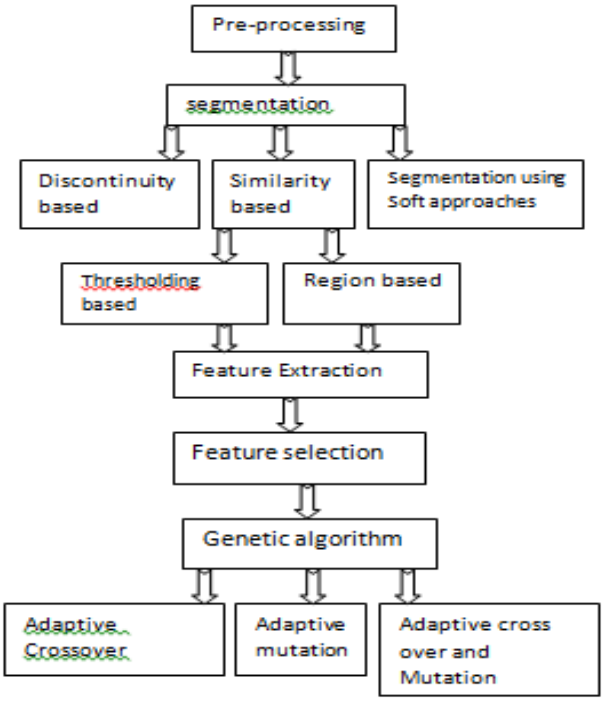

Fig. 1: An Outline of the Proposed System.

4. Image segmentation

Image segmentation is the method of partitioning an image into several sets of pixels. The purpose of segmentation is to transform the manifestation of an image to make it more significantl than the

original and easy to analyze. Thus segmentation partition an image pixels into constituent image regions [2]. It assigns a label to every pixel in an image such that pixels with the same label share positive image characteristics. Each pixel in a particular region differ with respect to color, brightness, appearance while the nearest pixels share identical characteristics [3]. One of the main important and complicated task in image processing is Image Segmentation. The first step in analysis of skin lesion is detection of lesion borders [3]. The border detection analysis is two fold. In the first step , the border structure provides essential information for correct diagnosis. Many of the clinical features such as asymmetry, border irregularity, and abrupt border cut-off present in the lesion can be determined from the border [5]. The second step critically relies on other clinical features like atypical pigment networks, globules and blue-white areas [11].Segmentation becomes difficult since there is a variation in lesion shape, sizes and colors with different textures in human skin[15]. On a broader perspective image segmentation algorithms can be classified into two main categories: discontinuity based and similarity based segmentation.

4.1. Discontinuity based

For discontinuity based segmentation the partition of an image is agreed upon abrupt changes in the variation of intensity levels in an image, or by rapid change in gray level of an image. The main aim of these algorithms is to identify isolated points, lines or edges in an image.

4.2. Similarity based

In similarity based segmentation algorithms those pixels in the image are grouped together which are similar in some pattern. The similarity based segmentation of thresholding and region based are given here.

4.2.1. Thresholding based

One of the simplest approach to segmentation is thresholding. The pixels are classified into groups and based on that criterion the threshold value is decided. It is based on the information that pixel values that have lesion change from the background[11] . Therefore by selecting an upper and lower value it is likely to segregate those pixels whose values are within this range. This is represented in the image histogram, in which the different objects are denoted as peaks. These peak bounds are estimated as good limits. This simple thresholding cannot be used in all cases since image histograms of skin lesions are not always multi-modal.
Thresholding operation is testing a pixel against a function T, where
T= T [(x, y), p(x, y), f(x, y)]
Thus, T is a function of any combination of three terms, that is;
(X, y) pixel location
F(x, y) pixel intensity at (x, y)
P(x, y) local property in a neighbourhood centered at (x, y). E.g. average intensity value
Within neighbourhood.
Depending upon this combination T can be local threshold, Global threshold or
Adaptive/dynamic threshold.
T [f(x, y)] Global threshold
T [p(x, y), f(x, y)] Local threshold
T [(x, y), p(x, y), f(x, y)] Adaptive/ dynamic threshold

4.2.2 Region based

In region based segmentation approach, the image is usually seperated into regions that satisfy a homogeneity criterion[5]. Both of the segmentation techniques discussed above has its own advantages and disadvantages:

Table 2: Comparison of Segmentation Approaches

Segmentation methods	Color	image histogram	Results
Thresholding	when good contrast Between lesion and normal skin	Bimodal	good
Region based	textured skin or different colors	Squared	moderate

4.3. Segmentation using soft computing approaches

Soft Computing is an upcoming field that consists of important applications of self-computing which is image segmentation [2]. One of the soft computing advance to image segmentation is Genetic Algorithm based approach [7].Some researchers discussed that physical border detection is better when compared to computer-detected borders in order to detach the problems of feature extraction from the problems of automated lesion border detection. Due to the advancement of automated diagnostic system for detecting skin lesion, it is essential to develop automatic segmentation algorithms. A critical step in the investigation of analyzing lesion images is segmentation, evolved into an important areas of research.

5. Feature extraction

The aim of feature extraction is to diminish the original data set by measuring certain property that distinguish one input pattern from another [3]. The feature extraction is performed by dimensions on the pixels that represent a segmented object allowing various features to be computed. Unfortunately, the feature extraction step is often subject to error.

6. Feature selection

For mny machine learning tasks, Feature selection is an important pre-processing step. Its purpose is to decrease the dimensionality of the feature space by eliminating redundant, irrelevant or noisy features. Feature selection can be defined as a process to select the best optimal subset of m features from the original set of n features [8]. The original feature set can be formed by combining the features formed by different feature extraction methods. It includes selection of , the subset of features in the input features that has maximum prediction power for the output. For the efficient performance of the classifier, the amount of data can be reduced by eliminating the dependent variables. The importance of feature selection phase has led to the expansion of a variety of techniques for selecting an best possible subset of features from a larger set of possible features. Feature selection does not create new features since it uses the input features itself to lessen their number. Once a feature selection criterion is selected, a procedure must be developed to find the subset of useful features [12]. In other words, the search for the optimal feature subset requires an evaluation measure to estimate the goodness of subsets and a search strategy to generate candidate feature subsets.
Feature selection aims mainly two goals: (1) reducing the cost of extracting features and (2) enhancement of the classification accuracy of a practical classifier. Especially, the second goal has received a great deal of attention in recent years according to the increase of the problem size.

7. Genetic algorithm

Genetic Algorithm is stochastic search algorithms mainly inspired by the genetic process of biological organisms. It searches for the finest solution in search space sharply in an iterative order to attain a new generation from the old one. Genetic algorithms are used to find the subset of features where the chromosome bits correspond the feature included or not. Genetic algorithms are dominant search algorithms that can be useful to a wide range of problems [3]. For organizing a genetic algorithm parameter setting

is employed, which remains unaffected during execution. The interesting problem here is the self-adaptive parameter adjustment of genetic algorithm. This paper proposes an approach for controlling the parameters of genetic algorithm that can be determined within the chromosome of each individual. Depending on the problem situation, the parameter values are totally dependent on the evolution mechanism. Our initial analysis demonstrate that GA has the ability to learn and assess the value of self-set parameter depending on the degree of contribution to the decree of the problem [8]. These results indicate the possible methodolgy for the development of GAs with self-adaptive parameter settings which do not require the user to modify parameter at the beginning. The basic mechanism of GA depends on various key parameters such as crossover operators, mutation operators, crossover probability, mutation probability, mechanism of selection and the size of population [13]. All these parameter have a great impact on GA's performance [12].One can thus describe the parameters quickly and later modify the parameters in a exact manner depending on the problem. A high rate of crossover and low rate of mutation might be very good in the study of new solutions for the first generations formed by the algorithm [12] [13]. Similarly while the algorithm is close to the optimal solution, it becomes unfavourable. One of the probable solution to decide the best set of parameters e in the exist in use of learning, involving an enrichment of the GA's performance. It is suggested that this improvement relies on self-adaptive parameters in GA. In this study a hypothesis of self-adaptive parameters in GA is proposed for further work. A self-adaptive approach to GA parameter based on the problem is considered in this work.The individual is based on two learning levels. (1) a genetic algorithm is applied to the knowledge of new sets of parameters which results in an improvement of the individual's adaptation to the problem to be solved. (2) reinforcement learning is used to learn the finest parameter settings. This learning occurs in communication with the problem situation. Lastly, the individual evaluate the quality of learned parameter settings [14]. An algorithm of the self-adaptive parameter approach is proposed which is a customized version of the basic genetic algorithm structure.

Step 1: generate random population
Step 2: estimate the individual population fitness value
Step 3: a new population is created- repeat steps until stopping condition is reached
Step 3a: to discover the best parameter settings for each individual use reinforcement learning depending on the problem.
Step 3b: depending on the fitness value, chooset two parents from a population
Step 3c: perform search in genetic crossover space search
Step 3d: to form best fitness new offspring cross the two individuals
Step 3e: to identify the best parameter for individual based on fitness and the problem, search in genetic mutation
Step 3f: in each position of the chromosome- mutate the new offspring
Step 3g: place new offspring in the population
Step 3h: develop a new set of parameters
Step 3i: based on the reinforcement location select two parameter settings
Step 4: continue by generating newly developed population
Step 5: terminate if end condition satisfies-return the best solution
Step 6: return to step [2]

The proposed method is provoked in part by the theory of independent individual. Various authors have recommended that the control of GAs' parameters could be encoded in the chromosome of every individual of the population [15]. This suggest the inclusion of a mechanism, which inserts the parameters in the individual's chromosomes, the changes of parameter values are thus completely reliant on the evolution mechanism. There are no suggested best settings for parameters to any problem. The operators used are crossover operators, mutation operators, crossover probability and mutation probability.

7.1. Adaptive crossover

The crossover evolution mechanism perform in two steps:

- Search for best parameter with respect to fitness value and problem domain in genetic crossover
- Crossing the two individuals

The only position of parameters, which survives, is measured to be the best individual. Test the fastest convergence of the algorithm to coincide with good parameter setting. If all individuals have same genetic code, then the algorithm converges quickl

7.2. Adaptive mutation

Mutation operation is carried out as:

- Search for best parameter with respect to fitness value and problem domain in genetic crossover
- Using the best parameter at each point, mutate the new offspring.

The parameters that last a long time result exist from many successful repeated mutations.

7.3. Adaptive crossover and mutation probabilities

The above cases are based on the adaptive crossover and mutation probabilities among the genetic diversity of population. To improve the study of genetic algorithm [14], a dynamic adaptaion of probability based on genetic diversity measure (gdm) is used. The gdm is defined as the percentage of the difference between the means and maximum values of the fitness value in each generation. The value of gdm is in the range of [0, 1]. Thus, the learning velocity avoids convergence since crossover probability is reduced and mutation rate increased.

8. Conclusions and future work

A unique algorithm for adaptive genetic algorithm is proposed. Segmentation of skin lesion by considering border irregularities are employed. The different segmentation methods are compared. In the future this AGA can be developed using soft computing approaches Genetic algorithm with adaptive parameters is considered in this paper. A learning based self-adjustment of GA can be adopted for further research.

References

[1] Nisha Oommachen, Vismi V, Soumya S, Jeena C D. Melanoma Skin Cancer Detection Based on Skin Lesions Characterization, IOSR Journal of Engineering (IOSRJEN) e-ISSN: 2250-3021, p-ISSN: 2278-8719 Vol. 3, Issue 2 (Feb. 2013), V1, PP 52-59.

[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] Nishima Sachdeva(Mtech scholar), Professor Rohan Gupta "Hybrid Approach To Investigate The Probability Of Skin Cancer By ABCD And PCA Method" ISSN: 2278 – 909X International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE) Volume 6, Issue 4, April 2017

[4] Ebthihal Almansour and M.Arfan Jaffar "Classification Of Dermoscopic Skin Cancer Images Using Color And Hybrid Texture Features" IJCSNS International Journal of Computer Science and Network Security, VOL.16 No.4, April 2016

[5] Lucia Ballerini, Robert B. Fisher , Ben Aldridge, Jonathan Rees "Non-Melanoma Skin Lesion Classification Using Colour Image Data In A Hierarchical K-NN Classifier

[6] 1Munya 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

[7] Savita, Sudha "New Technique Of Genetic Algorithm For Finding Maximum Clique Problem International Journal of Science, Engi-

neering and Technology Research (IJSETR), Volume 3, Issue 8, August 2014
[8] Ani1 K. Jain, Fellow, IEEE, Robert P.W. Duin, and Jianchang Mao "Statistical Pattern Recognition: A Review" IEEE Transactions On Pattern Analysis And Machine Intelligence, Vol. 22, No. 1, January 2000
[9] Hintz-Madsen, Mads; Hansen, Lars Kai; Larsen, Jan; Olesen, Eric; Drzewiecki, K.T Design and evaluation of neural classifiers application to skin lesion classification Published in: Proceedings of the 1995 IEEE Workshop on Neural Networks for Signal Processing
[10] Anil Jain, Douglas Zongker "Feature selection: Evaluation, Application and Small sample performance"
[11] Miss. Komal R. Hole1, Prof. Vijay S. Gulhane2, Prof. Nitin D. Shellokar3 "Application of Genetic Algorithm for Image Enhancement and Segmentation" ISSN: 2278 – 1323 International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 2, Issue 4, April 2013.
[12] Douglas Zongker, Anil Jain "Algorithms for Feature Selection: An Evaluation"
[13] Mineichi Kudo1 And Jack Sklansky "A Comparative Evaluation Of Medium and Large-Scale Feature Selectors For Pattern Classifiers" Kybernetika — Volume 34 (1998), Number 4, PAGES 429-434
[14] Ose Bins, Bruce A.Draper "Feature Selection from Huge Feature Sets"
[15] Alireza Fasih, Jean Chamberlain Chedjou " Cellular Neural Networks-Based Genetic Algorithm for Optimizing the Behavior of an Unstructured Robot " Transportation Informatics Group, Institute of Smart Systems Technologies, University of Klagenfurt Klagenfurt
[16] International Journal of Computational Intelligence Systems, Vol.2, No. 2 (June, 2009), 124-131
[17] Cesar Guerra Salcedo, Stephen Chen, Darrell Whitley, Stephen Smith "Fast and Accurate Feature Selection Using Hybrid Genetic Strategies".
[18] Tao Wang1, Yong Rui2, Shi-min Hu1, Jia-guang Sun1 "Adaptive tree similarity learning for image retrieval"-Multimedia Systems 9: 131–143 (2003) Multimedia Systems
[19] Ron Kohavi "Feature Subset selection as Search with Probabilistic estimates" AAAI Fall Symposium in 1994.
[20] Mohammad H. Sedaaghi, Constantine Kotropoulos, and Dimitrios Ververidis "Using Adaptive Genetic Algorithms to Improve Speech Emotion Recognition".
[21] Daamagoj Jakbovic and Marin Golub "Adaptive Genetic Algorithm" Journal of Computing and Information Technology", CIT-7, 1993, 3, 229-235.
[22] Eric Pellerin, Luc Pigeon, Sylvain Delisle "Self-Adaptive Parameters in Genetic Algorithms".