ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 5, May 2015

Classification of Dermoscopic Images for Studying Cancer and Non-Cancer

T.Selvakumar[1], Dr. L.S. Jayashree[2]

P.G. Scholar, Department of CSE, RVS College of Engineering and Technology, Coimbatore, India[1]

Professor and Dean, Department of CSE, RVS College of Engineering and Technology, Coimbatore, India[2]

ABSTRACT: Main advantage of this computer based classification is that patient need not to go hospitals or undergo various painful diagnosing techniques like biopsy. In this computer aided classification, dermoscopic image of skin cancer is taken and is subjected to various preprocessing procedures for image enhancement. The cancer affected region is segmented from the healthy skin using thresholding technique. Level set techniques are used to mark the edges exactly. In order to reduce the complexity of classification, some important features of malignant and benign melanoma are extracted. Features are extracted using first order derivatives of Gray Level Co-occurrence Matrix (GLCM) method. Ant Colony Optimization (ACO) algorithm is used to select the best features among the large feature set. Feature selection improves the classification accuracy. Based on selected features dermoscopic images have been classified into cancer or non-cancer. There are two types of classification technique available supervised learning and unsupervised learning.To classify the dermoscopic image using classification technique like Back propagation, Support vector machine, Artificial immune system. Ensample method is a combination of classifiers used to improve the classification accuracy.

KEYWORDS: GLCM, Feature selection, Ant colony optimization, Support vector machine, Artificial immune system.

I. INTRODUCTION

Classification is a form of data analysis that extracts models describing important data classes. Such models, called classifiers, predict categorical (discrete, unordered) class labels. For example, we can build a classification model to categorize dermoscopic images as cancer or non-cancer. Such analysis can help provide us with a better understanding of the data at large. Many classification methods have been proposed by researchers in machine learning, pattern recognition, and statistics. Most algorithms are memory resident, typically assuming a small data size. Classification has numerous applications, including fraud detection, target marketing, performance prediction, manufacturing, and medical diagnosis.

Reduction of pattern dimensionality via feature extraction is one of the most important tasks for pattern recognition and classification. Feature selection has considerable importance in areas such as bioinformatics, signal processing, image processing, text categorization, data mining, pattern recognition, medical diagnosis and remote sensor image recognition. The goal of feature selection is to choose a subset of available features by eliminating unnecessary features. To extract as much information as possible from a given image set while using the smallest number of features, we should eliminate the features with little or no predictive information, and ignore the redundant features that are strongly correlated. As a result, a large amount of computation time can be saved. The selected subset of features used to represent such classification function influences several aspects of image classification, including the time required to learn a classification function, the accuracy of the learned classification algorithm, the time-space cost associated with the features, and the number of samples required for training. ACO is a multi-agent system where communications among artificial ants result in a positive feedback behavior to guide the ant colony to converge to the optimal solution. GLCM model used for extracting the features from the given input image. Before extract the features from the image preprocessing and segmentation have to be done.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 5, May 2015

II. DATA FLOW DIAGRAM

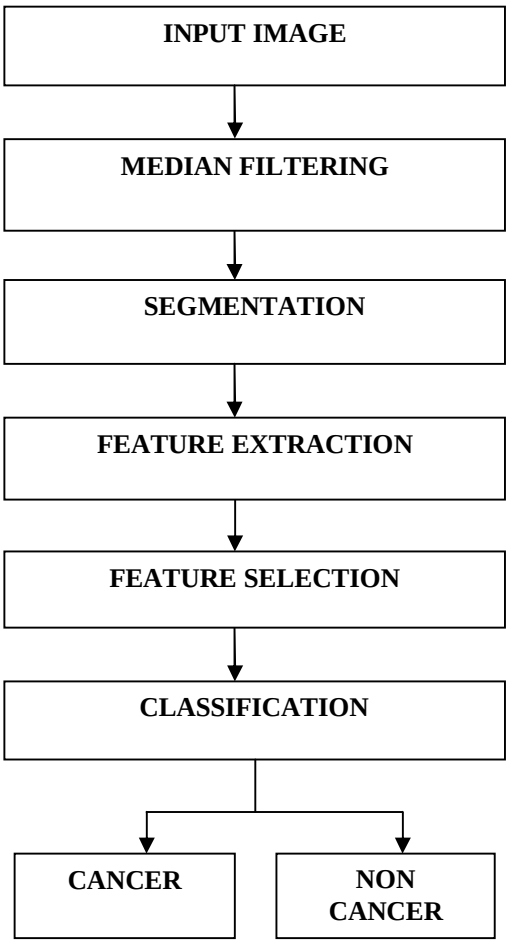

Fig 1 Steps for preprocessing, image segmentation, feature extraction, feature selection and classification.

III. DETAILS OF METHODOLOGY

1) PREPROCESSING:

Clahe operates on small regions in the image, called tiles, rather than the entire image. Each tile's contrast is enhanced, so that the histogram of the output region approximately matches the histogram specified by the 'Distribution' parameter. The neighboring tiles are then combined using bilinear interpolation to eliminate artificially induced boundaries. The contrast, especially in homogeneous areas, can be limited to avoid amplifying any noise that might be present in the image.

2) THRESHOLDING:

The simplest method of image segmentation is called the thresholding method. This method is based on a threshold value to turn a gray-scale image into a binary image. There is also a balanced histogram thresholding.

The key of this method is to select the threshold value (or values when multiple-levels are selected). Several popular methods are used in industry including the maximum entropy method, Otsu's method, and k-meansclustering. Recently, methods have been developed for thresholding computed tomography (CT) images. The key idea is that, unlike Otsu's method, the thresholds are derived from the radiographs instead of the (reconstructed) image.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 5, May 2015

3) IMAGESEGMENTATION:

Image segmentation[3] is the process of partitioning a digital image into multiple segments (sets of pixels, also known as super pixels). The goal of segmentationis to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze. Image segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images. More precisely, image segmentation is the process of assigning a label to every pixel in an image such that pixels with the same label share certain characteristics.

The result of image segmentation is a set of segments that collectively cover the entire image, or a set of contours extracted from the image. Each of the pixels in a region are similar with respect to some characteristic or computed property, such as color, intensity, or texture. Adjacent regions are significantly different with respect to the same characteristic(s).When applied to a stack of images, typical in medical imaging, the resulting contours after image segmentation can be used to create 3D reconstructions with the help of interpolation algorithms like Marching cubes.

4) LEVEL SET METHODS:

The level set method tracks the motion of an interface by embedding the interface as the zero level set of the signed distance function. The motion of the interface is matched with the zero level set of the level set function, and the resulting initial value partial differential equation for the evolution of the level set function resembles a Hamilton-Jacobi equation. In this setting, curvatures and normal may be easily evaluated, topological changes occur in a natural manner, and the technique extends trivially to three dimensions. This equation is solved using entropy-satisfying schemes borrowed from the numerical solution of hyperbolic conservation laws which produce the correct viscosity solution.

5) FEATURE EXTRACTION:

A gray level co-occurrence matrix (GLCM) to extract second order statistical texture features for motion estimation of images. The four features MEAN, ENTROPHY, VARIENCE, and STANDARD DEVIATIONS are to be calculated. Gray Level Co-Occurrence Matrix (GLCM) has proved to be a popular statistical method of extracting textural feature from images. According to co-occurrence matrix, Haralick defines fourteen textural features measured from the probability matrix to extract the characteristics of texture statistics of remote sensing images.

In statistical texture analysis, texture features are computed from the statistical distribution of observed combinations of intensities at specified positions relative to each other in the image. According to the number of intensity points (pixels) in each combination, statistics are classified into first-order, second-order and higher-order statistics. The Gray Level Co-occurrence Matrix (GLCM) method is a way of extracting second order statistical texture features. The approach has been used in a number of applications, Third and higher order textures consider the relationships among three or more pixels. These are theoretically possible but not commonly implemented due to calculation time and interpretation difficulty. A GLCM [5] is a matrix where the number of rows and columns is equal to the number of gray levels, G, in the image. The matrix element P (i, j | Δx, Δy) is the relative frequency with which two pixels, separated by a pixel distance (Δx, Δy), occur within a given neighborhood, one with intensity 'i' and the other with intensity 'j'. The matrix element P (i, j | d, ɵ) contains the second order statistical probability values for changes between gray levels 'i' and 'j' at a particular displacement distance d and at a particular angle (ɵ). Using a large number of intensity levels G implies storing a lot of temporary data, i.e. a G × G matrix for each combination of (Δx, Δy) or (d, ɵ). Due to their large dimensionality, the GLCM's are very sensitive to the size of the texture samples on which they are estimated. Thus, the number of gray levels is often reduced. Here one pixel offset is used (a reference pixel and its immediate neighbor). If the window is large enough, using a larger offset is possible. The top left cell will be filled with the number of times the combination 0,0 occurs, i.e. how many time within the image area a pixel with grey level 0 (neighbour pixel) falls to the right of another pixel with grey level 0(reference pixel). Tab no.1 shows the sample feature extraction work of both cancer and non-cancer.

6) FEATURE SELECTION:

The goal of feature selection is to choose a subset of available features by eliminating unnecessary features. To extract as much information as possible from a given image set while using the smallest number of features, we should eliminate the features with little or no predictive information, and ignore the redundant features that are strongly correlated. As a result, a large amount of computation time can be saved.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 5, May 2015

Pheromone information, which simulates the chemical substance the real ants lay on the route they passed, is assigned to the edges of the graph. Artificial ants are used in ACO to travel in the graph to search for optimal paths according to the pheromone and problem specific local heuristics information. The pheromone on each edge is evaporated at a certain rate at each iteration. It is also updated according to the quality of the paths containing this edge. Artificial ants are usually associated with a list that records their previous actions, and they may apply some additional operations such as local search, crossover and mutation, to improve the quality of the results obtained. Compared with GA, ACO has some advantages such as allowing positive feedback, distributed computing, and constructive greedy heuristic search.

The framework of the ACO based feature selection algorithm[9] ACOFS is shown as follows.

6.1) ACO Algorithm:

Algorithm: ACOFS (ACO based feature selection)
Input: DG: The directed graph;
t: The initial pheromone matrix;
Output: S_{best}: The solution of the feature selection;
Begin
1. Set the initial values of parameters;
2. While not termination condition do
3. Starting fromv_0, the m ants traverse on the directed graph according to the probability formula (1) on each node. After all the m ants reach the node v_n, m subsets of features are formed;
4. Evaluate the fitness of them feature subsets by classifying the training image sets;
5. Update the pheromone and heuristic information on each arc;
6. Select the solution with the highest fitness value found so far as
S_{best};
7. End While;

For n features, most ACO based feature selection methods use a complete graph with $O(n^2)$ edges. This means that $O(n^2)$ pheromone and heuristic information will be computed and stored.

7) CLASSIFICATION:

Classification is a form of data analysis that extracts models describing important data classes. Such models, called classifiers, predict categorical (discrete, unordered) class labels. Many classification methods have been proposed by researchers in machine learning, pattern recognition, and statistics. Before classification we have to do preprocessing, segmentation and optimization steps refer data flow diagram fig no.1.

Data classification is a two-step process, consisting of a learning step (where a classification model is constructed) and a classification step (where the model is used to predict class labels for given data).

7.1) BACK PROPAGATION:

Back propagation learns by iteratively processing a data set of training tuples, comparing the network's prediction for each tuple with the actual known target value. The target value may be the known class label of the trainingtuple (for classification problems) or a continuous value (for numeric prediction). For each training tuple, the weights are modified so as to minimize the mean-squared error between the network's prediction and the actual target value. These modifications are made in the "backwards" direction (i.e., from the output layer) through each hidden layer down to the first hidden layer (hence the name backpropagation). Although it is not guaranteed, in general the weights will eventually converge, and the learning process steps.

7.2) ARTIFICIAL IMMUNE SYSTEM:

In artificial intelligence, artificial immune systems (AIS) are a class of computationally intelligent systems inspired by the principles and processes of the vertebrate immune system[6]. The algorithms typically exploit the immune system's characteristics of learning and memory to solve a problem.

The field of Artificial Immune Systems (AIS) is concerned with abstracting the structure and function of the immune system to computational systems, and investigating the application of these systems towards solving computational problems from mathematics, engineering, and information technology[5]. AIS is a sub-field of

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 5, May 2015

Biologically-inspired computing, and Natural computation, with interests in Machine Learning and belonging to the broader field of Artificial Intelligence.Artificial Immune Systems (AIS) are adaptive systems, inspired by theoretical immunology and observed immune functions, principles and models, which are applied to problem solving.

Tab no.1 Feature Extraction for Sample Cancer and Non-Cancer Images

Features/Images	Cancer1	Cancer2	Cancer3	Cancer4	Cancer5	Non-Cancer1	Non-Cancer2	Non-Cancer3	Non-Cancer4	Non-Cancer5
Autocorrelation	10268.94	9787.12	9924.24	10718.59	9582.45	9455.47	10082.03	10812.33	10225.89	9406.23
Contrast	7469.96	6683.57	6673.23	8000.40	7252.10	6822.51	6763.73	6839.49	8718.68	7721.03
Correlation: matlab	0.04	0.02	0.02	0.03	-0.02	0.00	0.01	-0.02	-0.06	-0.02
Correlation	0.04	0.02	0.02	0.03	-0.02	0.00	0.01	-0.02	-0.06	-0.02
Cluster Shade	8386.70	25838.98	32229.14	-47208.59	28648.62	35740.39	-3765.71	-28812.23	-32962.7	50834.47
Dissimilarity	69.97	66.61	66.07	73.39	70.14	67.43	67.00	67.72	78.55	72.35
Energy: matlab	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Entropy	10.54	10.59	10.58	10.49	10.58	10.59	10.59	10.59	10.49	10.54
Homogeneity: matlab	10.54	0.04	0.05	0.04	0.04	0.04	0.04	0.04	0.04	0.04
Homogeneity	0.02	0.02	0.02	0.01	0.01	0.02	0.02	0.02	0.01	0.02
Maximum probability	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sum of squares: Variance	13393.37	12850.46	13289.35	14474.40	13202.37	12329.87	13476.99	14250.13	14124.16	12951.87
Sum average	201.12	197.17	198.54	205.69	196.63	194.44	200.64	208.56	204.58	194.92
Sum variance	46222.02	43573.04	44095.91	48491.80	43330.01	42419.86	44795.21	47706.09	47261.71	43107.97
Sum entropy	5.86	5.81	5.81	5.88	5.81	5.81	5.81	5.79	5.85	5.83
Difference variance	7469.96	6683.57	6673.23	8000.40	7252.10	6822.51	6763.73	6839.49	8718.68	7721.03
Difference entropy	5.15	5.11	5.11	5.17	5.14	5.12	5.11	5.12	5.20	5.16
Information measure of correlation1	0.00	0.00	0.00	-0.01	0.00	0.00	0.00	0.00	-0.01	0.00
Information measure of correlation2	0.20	0.08	0.13	0.26	0.11	0.10	0.10	0.07	0.25	0.21
Inverse difference normalized (INN)	0.77	0.77	0.78	0.76	0.76	0.77	0.77	0.77	0.74	0.76
Inverse difference moment normalized	0.87	0.88	0.88	0.86	0.87	0.88	0.88	0.87	0.85	0.86

7.3) SUPPORT VECTOR MACHINE:

In machine learning, support vector machines are supervised learning models with associated learning algorithms that analyze data and recognize patterns, used for classification and regression analysis[4]. Given a set of training examples, each marked as belonging to one of two categories, a SVM training algorithm builds a model that assigns new examples into one category or the other, making it a non-probabilisticbinarylinear classifier. A SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall on.In addition to performing linear classification, SVMs can efficiently perform a non-linear classification using what is called the kernel trick, implicitly mapping their inputs into high-dimensional feature spaces.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 5, May 2015

Comparison between the classifiers:

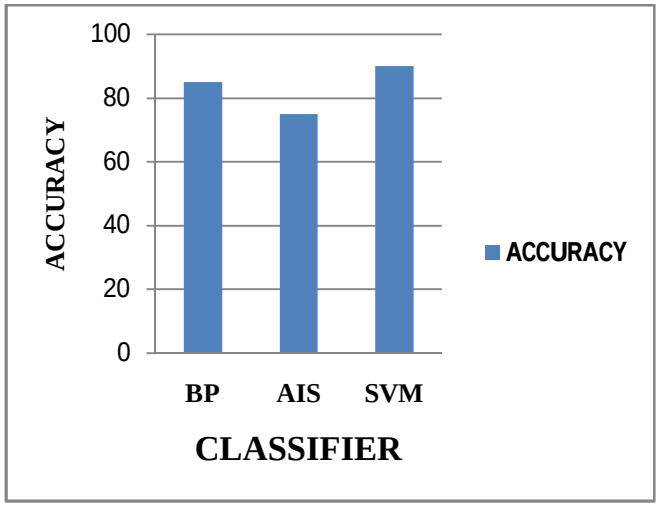

Fig no. 2 Comparison of Back propagation, AIS and SVM Classifier

IV CONCLUSION AND FUTURE WORK

In this work, features are extracted from dermoscopic images. CLAHE is a Contrast-limited adaptive histogram equalization technique used to remove noise. Edge detection techniques like threshold and level set methods were used to detect the edges. Texture analysis is also considered for segmentation. Gray level co-occurrence matrix (GLCM) using this approach 21 features and above were extracted from the segmented region. These extracted features are optimized by using Ant Colony Optimization (ACO). ACO is used to select the best features among the all features. Classify the image using selected features by Back propagation, Artificial immune system and Support vector machine.Accuracy also calculated using confusion matrix. Fig no.2 Comparing with all these classifier Support Vector Machine has greater accuracy.From the comparisonSVM performing well while classifying dermoscopic images into cancer or non-cancer.Enhancement of this work is to classify the data's with combination of classifiers. Ensemble method is a combination of classifiers to improve the classification accuracy.

REFERENCES

[1]. G. Sfikas, C. Constantinopoulos, A. Likas, and N. Galatsanos(2005),"An Analytic Distance Metric for Gaussian Mixture Models With Application in Image Retrieval"., New York: Springer, vol. 3697, LNCS, pp.835–840
[2]. H. Pehamberger, A. Steiner, and K. Wolff(1987), "In vivo epiluminescence microscopy of pigmented skin lesions. I. Pattern analysis of pigmented skinlesions,"J. Am. Acad. Dermatol., vol. 17, no. 4, pp. 571–583.
[3]. M.Tuceryan andA.Jain,,C.H.Chen,L.F.Pau,andP.S.P.Wang, Eds(1998), "Texture analysis," inHandbook of Pattern Recognition and Vi-sion, 2nd ed. River Edge, NJ: World Scientific.
[4]. C.W. Hsu and C. J. Lin (2002)," A comparison of methods for multiclass support vector machines", IEEE Transactions on Neural Networks, 13(2):415-425.
[5]. U.OjhaandMo-Yuen Chow(2010)," An analysis of Artificial Immune System and Genetic Algorithm in urban path planning",IECON 2010 - 36th Annual Conference on IEEE Industrial Electronics Society ,1553-572X , 1064 – 1069.
[6]. D.Simic, and S.Simic (2011)," Artificial immune system in risk of falling classification", Nature and Biologically Inspired Computing (NaBIC), IEEE, Date of Conference, 19-21 Oct., 627 – 632.
[7]. Yun-Chia Liang and Alice E. Smith (2004), "An Ant Colony Optimization Algorithm for the Redundancy Allocation Problem (RAP)", IEEE transactions on reliability, vol. 53, no. 3.
[8]. A. Haritha, L. Pavan Krishna, Y. Suresh, K. Pavan Kumar, P.V.S. Lakshmi (2013), " In Time Access of Biomedical Data through Ant Colony Optimization ",Advances in Intelligent Systems and Computing Volume 199, 2013, pp 303-310.
[9]. M.E.Basiri, N.Ghasem-Aghaee, M.H.Aghdam (2008),"Using ant colony optimization-based selected features for predicting post-synaptic activity in proteins", EvoBIO 2008, Lecture Notes in Computer Science, vol.4973,Springer-Verlag,Berlin,Heidelberg,Italy, pp. 12–23.