ISSN: 2278 – 909X
International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE)
Volume 4, Issue 9, September 2015

Classification of Lung Diseases using Particle Swarm Optimization

Tejinder Kaur
Department of Electronics and Communication
Amritsar College of Engineering & Technology

Er. Neelakshi Gupta
Department of Electronics and Technology
Amritsar College of Engineering & Technology

***Abstract*- - Lung diseases are the disorders that affect the lungs, the organs that allow us to breathe and it is the most common medical conditions worldwide especially in India. In this paper, new feature selection technique is applied to lung CT images so as to classify various lung diseases The lung CT image is engaged as the input. The guided image filter is applied to remove noises and the pre processed images are given as input for feature extraction where the useful features of the images are extracted. New feature selection technique that is hybridization of genetic and PSO is used to select features after extracting features using MAD technique. Quantitative analysis was performed on the basis of geometry Accuracy, Bit Classification Rate and Bit Error Rate.**

KEYWORD- CT image, Genetic, Particle Swarm Optimization, MLP-NN, feature selection, classification

1. Introduction

Lung diseases are some of the most common medical conditions in the world. Tens of millions of people suffer from lung disease in the U.S. Smoking, infections, and genetics are responsible for most lung diseases. Lung disease refers to many disorders affecting the lungs such as asthma, chronic obstructive pulmonary (COPD) disease, infections such as tuberculosis, influenza, lung cancer, pneumonia and other breathing problems. Feature selection is used to identify powerfully predictive subsets of field within a database and reduce the number of fields presented to the computation method. Feature selection affects several pattern classification aspects, including the accuracy of the learned classification algorithm [10]. Better performance can be achieved by discarding some features. Thus, we imply eliminating noisy, irrelevant and redundant data, while maintaining the discriminating power of data by Feature selection.

2. Literature survey

In September 2012 [2], Ms.Swati P. Tidke, Prof. Vrishali A. Chakkarwar provided a Computer Aided Diagnosis System (CAD) for early detection of lung cancer nodules from the Chest Computer Tomography (CT) images. There were five main phases involved in the proposed CAD system. They were image pre-processing, extraction of lung region from chest computer tomography images, segmentation of lung region, feature extraction from the segmented region, classification of lung cancer as benign or malignant. Initially total variation based de noising was used for image de noising, and then segmentation has been performed using optimal thresholding and morphological operations. Textural features extracted from the lung nodules using gray level co-occurrence matrix (GLCM). For classification, SVM classifier was used. The main aim of the method was to develop a CAD (Computer Aided Diagnosis) system for finding the lung tumor using the lung CT images and classify the tumour as Benign or Malignant. In 2014 [4], C. Gunavathi , K. Premalatha used Genetic Algorithm (GA) for effective feature selection. Informative genes were identified based on the T-Statistics, Signal-to-Noise Ratio (SNR) and F-Test values. The initial candidate solutions of GA were obtained from top-m informative genes. The classification accuracy of k-Nearest Neighbor (kNN) method was used as the fitness function for GA. In this work, kNN and Support Vector Machine (SVM) was used as the classifiers. The experimental results show that the proposed work was suitable for effective feature selection. With the help of the selected genes, GA-kNN method achieves 100% accuracy in 4 datasets and GA-SVM method achieves in 5 out of 10 datasets. The GA with kNN and SVM methods are demonstrated to be an accurate method for microarray based tumor classification.

3. Methodology

An experiment using MATLAB has been performed using a CT image. The procedure took for the implementation of the proposed technique is explained below:

Image Acquisition

For experimental purpose, abdominal CT scans of various patients were collected from different sources. The medical images have a unique format i.e. DICOM (Digital Imaging and Communication in Medicine). The DICOM files have a header which contains information of patient, type of scan and image dimensions. For research purpose, we have used files displayed as a gray scale image of 256*256. The entries of a gray scale image range from 0-255, where 0 shows total black colour and 255 shows total white colour. Entries between these ranges vary in intensity from black to white.

Pre-processing

The pre-processing performs background noise removal, intensity inhomogenity correction and removal of irrelevant organs such as kidney, spleen, gallbladder automatically from abdominal images to suppress their interference in the classification process. In classification process it is very important that that image should be clear and edges should be smooth. Median filter is used to eliminate salt and pepper noise and guided image filter is used as an edge preserving smoothing operators.

Feature Extraction

Creating the subset of new features from already existing features is called Feature Extraction. Purpose of feature extraction technique is to reduce original dataset by measuring features that distinguish one region of interest from other. Sub-image is taken from top left corner of the original image and texture features and pixel coefficients values are extracted. Features are extracted using Gabor filter and Walsh- Hadamard transform. The extracted features are fused using Median Absolute Deviation (MAD) technique.

Feature Selection

Feature selection is a dimensionality reduction technique widely used for data mining and knowledge discovery and it allows elimination of (irrelevant/redundant) features, whilst retaining the underlying discriminatory information, feature selection implies less data transmission and efficient data mining. It also brings potential communication advantages in terms of packet collisions, data rate, and storage. Feature selection method is of three types- Filter method, Wrapper method, Hybrid method. Filter techniques rank the feature by the intrinsic properties of the data, independent of the choice of classifier. Wrapper methods are so called because they wrap a classifier up in a feature selection algorithm. Typically: a set of features is chosen; the efficiency of this set is determined; some perturbation is made to change the original set and the efficiency of the new set is evaluated .Hybrid is similar to wrapper method, while multiple algorithms can be combined in embedded method to perform feature subset selection.

Classification

Image classification analyzes the numerical properties of various image features and organizes the data into categories. Classification algorithms mainly employs into two phases of processing such as training and testing. In initial training phase, the characteristic properties of typical image features are isolated based on the training class, is created. In the subsequent testing phase, these feature-space partitions are used for classification of image features. There are various classification methods such as decision tree induction, Bayesian networks, KNN , MLP-NN . in my work I am using MLP-NN classification method.

4. Results and Discussions

In this paper, feature selection using hybrid genetic and particle swarm optimization and classification of lung CT images using MLP-NN is analysed. The lung CT images are collected from various medical sites. The implementation of the framework is done in MATLAB. The performance metrics like Geometry Accuracy, Bit Classification Rate and Bit Error Rate are computed for quantitative comparison.

For qualitative analysis, output images of the implemented framework are shown in Fig 1(a)-(i). The images show the original image, then original image is converted into Grey scale image. Then filters are applied on image in order to remove noise and enhance the quality of image. Further the features are extracted using MAD technique. Extracted features are selected using GAPSO algorithm. And at last we classify the features using MLP-NN. Resulted image is obtained using GAPSO-MLPNN.

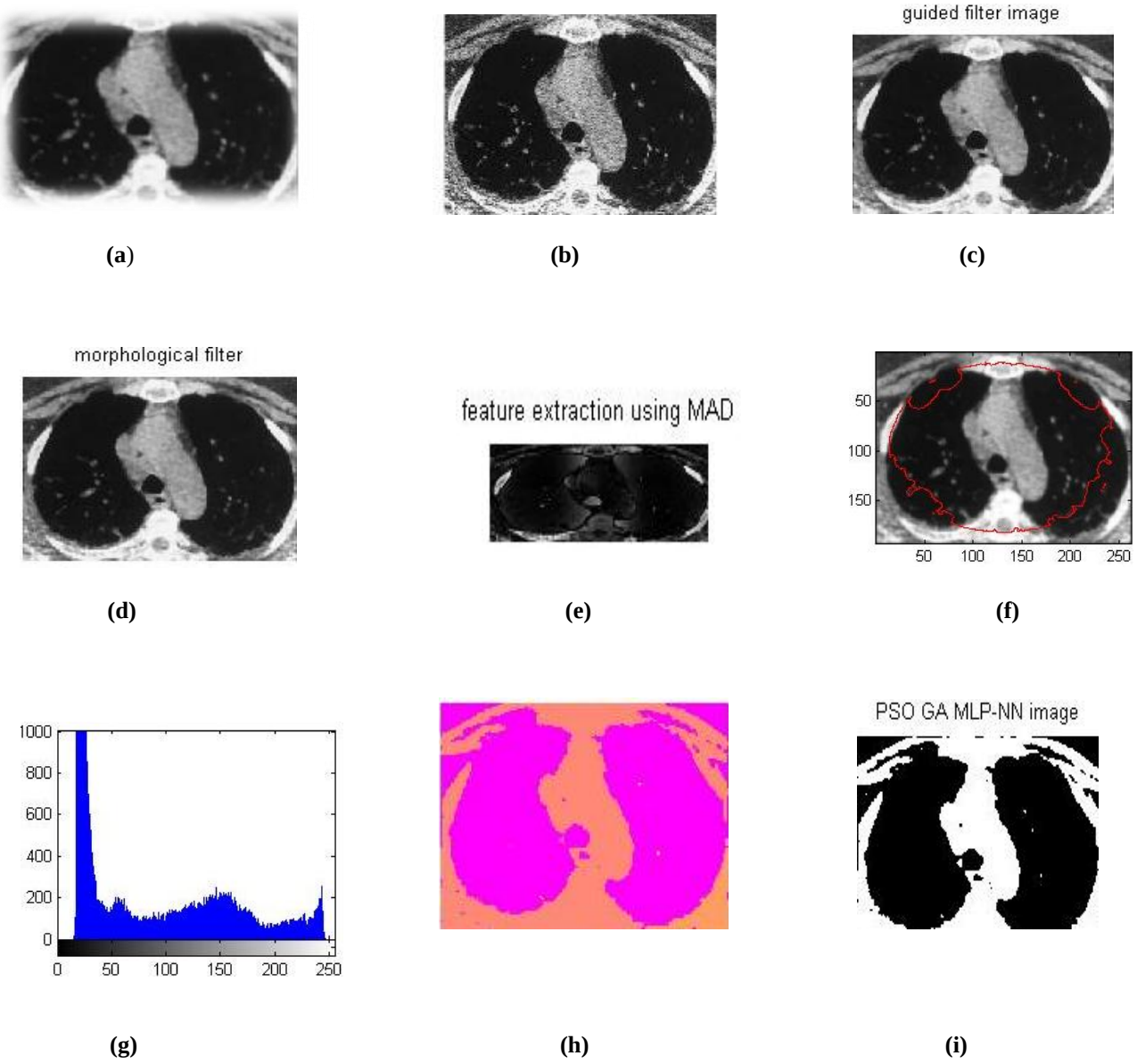

Figure 1: (a) Original CT image of abnormal lung (b) Conversion of image to grey scale image (c) Guided filter applied on image (d) Obtained image with morphological filter (e) Features are extracted using MAD technique (f) Image showing the maximum energy (g) Energy graph (h) Extracted features are selected using GAPSO algorithm (i) Final image after applying MLP-NN classifier

Quantitative Analysis

The quantitative results are presented in tables which gives the comparison between the existing and proposed technique on basis of Geometry Accuracy, Bit Classification Rate and Bit Error Rate . The results were taken on 15 CT scan images.

Geometry Accuracy Analysis:

Table below shows the comparative output of the existing and proposed technique on basis of Geometry Accuracy. The results were taken on fifteen CT images.

Image .No.	Existing	Proposed
1	91.2	98.8
2	83	99
3	90	98.5
4	88.5	100
5	91.5	98.5
6	89	98
7	92.5	100
8	88	98.5
9	90	99.5
10	80	97
11	89	98
12	96	100
13	84	99
14	93	100
15	85	97

The figure below shows the graphical implementation of the above table. The graph clearly depicts that the results of the proposed technique are much better than those obtained from the existing technique.

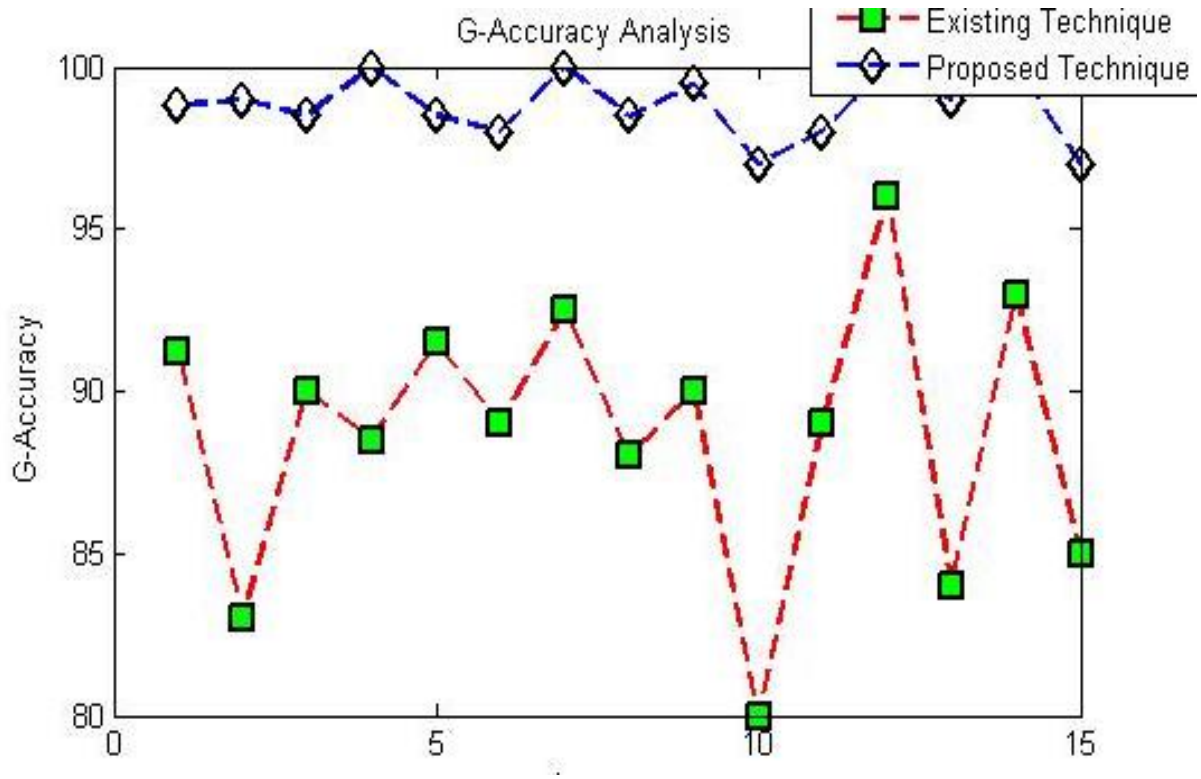

Fig 2: Graphical representation of existing and proposed technique in terms of G-Accuracy.

Bit Classification Rate Analysis:

Table below shows the comparative output of the existing and proposed technique on basis of bit classification rate. The results were taken on fifteen CT images.

Image No.	Existing	Proposed
1	0.90	0.95
2	0.87	0.91
3	0.92	0.95
4	0.85	0.92
5	0.86	0.95
6	0.94	0.96
7	0.89	0.95
8	0.89	0.94
9	0.90	0.94
10	0.85	0.92
11	0.92	0.97
12	0.90	0.92
13	0.88	0.94
14	0.95	0.98
15	0.89	0.90

The figure below shows the graphical implementation of the above table. The graph clearly depicts that the results of the proposed technique are much better than those obtained from the existing technique.

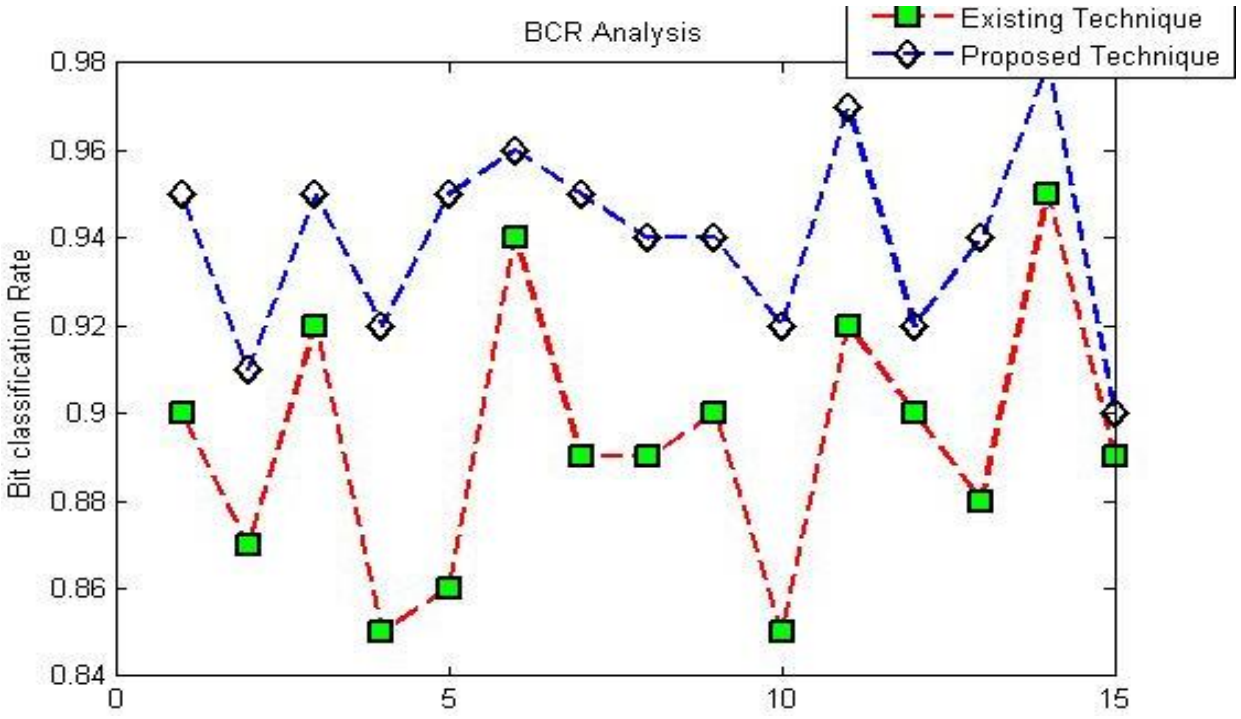

Fig 3: Graphical representation of existing and proposed technique in terms of Bi Classification Rate.

Bit Error Rate Analysis:

Table below shows the comparative output of the existing and proposed technique on basis of bit classification rate. The results were taken on fifteen CT images.

Image No.	Existing	Proposed
1	10	5
2	13	9
3	8	5
4	15	10
5	15	6
6	9	6
7	12	6
8	11	6
9	10	6
10	15	8
11	8	3
12	10	8
13	12	6
14	5	2
15	11	10

The figure below shows the graphical implementation of the above table. The graph clearly depicts that the results of the proposed technique are much better than those obtained from the existing technique.

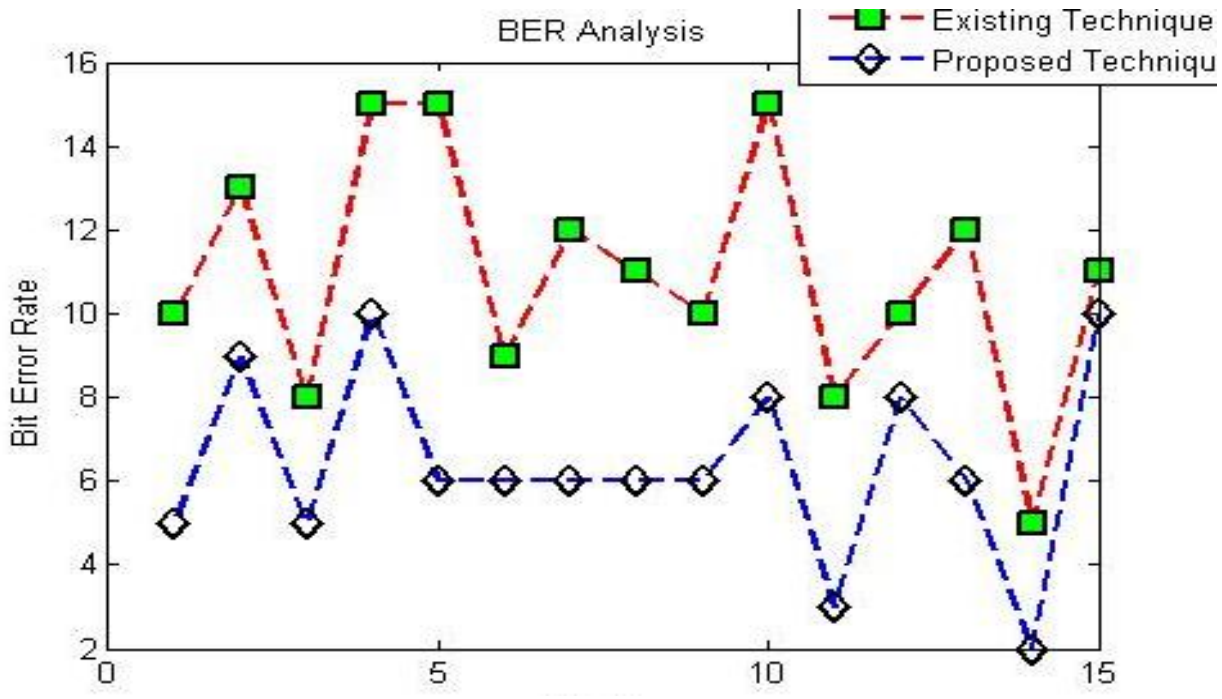

Fig 4: Graphical representation of existing and proposed technique in terms of Bit Error Rate

5. Conclusion and Future Scope

In this paper new Feature Selection Technique that is Hybridization of Genetic/Particle Swarm Optimization ia Proposed. Extracted and selected features are then classified applying MLP-NN classifier. . The experimental result shows high Geometry Accuracy, high Bit Classification Rate and low Bit Error Rate in various testing data. This method effectively works well for the detection of lung diseases. In future, work can be extended by adding more feature extraction and selection techniques so as to classify more disease.

References

1. Lu Meng, Hong Zhao,"A New Lung Segmentation Algorithm for Pathological CT images", IEEE 2009.
2. Ms. Swati P. Tidke, Prof. Vrishali A. Chakkarwar," Classification of Lung Tumor using SVM", Volume 2 Issue. 5, IJECR Sep 2012.
3. C.Bhuvaneswari1, P.Aruna, D.Loganathan," Classification of Lung Diseases by Image Processing Techniques Using Computed Tomography Images", International Journal of Advanced Computer Research Volume-4 Number-1, 14 March-2014.
4. C. Gunavathi, K. Premalatha," Performance Analysis of Genetic Algorithm with kNN and SVM for Feature Selection in Tumor Classification", International Journal of Computer, Control, Quantum and Information Engineering , 2014.
5. Lauge Sørensen, Marco Loog, Pechin Lo, Haseem Ashraf, Asger Dirksen, Robert P.W. Duin, and Marleen de Bruijne," Image Dissimilarity-Based Quantification of Lung Disease from CT", MICCAI 2010.
6. Ms. Swati P. Tidke, Prof. Vrishali A. Chakkarwar," Classification of Lung Tumor using SVM", Volume 2 Issue. 5, IJECR Sep 2012.
7. Shengjun Zhou, Yuanzhi Cheng, Shinichi Tamura," Automated lung segmentation and smoothing techniques for inclusion of juxtapleural nodules and pulmonary vessels on chest CT images", Volume 13, Elseiver 2014.
8. Choa-Lung Yang, R.J. Kuo, Chia-Hsuan Chien, Nguven Thi Phuong Oueyen," Non-dominated sorting Genetic Algorithm using Fuzzy membership chromosome for categorical data clustering.", Volume 30, Elseiver2015.
9. Gurpreet Kaur, Harpreet Singh," An Intelligent System for Lung Cancer Diagnosis Using Fusion of Support Vector Machines and Back Propagation Neural Network", International Journal of Science and Research (IJSR), 2013.
10. Mingyuan Zhao, Chong Fu, Luping Ji, Ke Tang, Mingtian Zhou," Feature Selection and Parameter optimization for support vector machines: A new approach based on genetic algorithm with feature chromosomes", Elseiver 2010.
11. I. Guyon and A. Elisseeff, "An Introduction to Variable and Feature Selection," Journal of Machine Learning Research, vol. 3, pp:1157-1182, 2003
12. Bahriye Akay," A Study on Partile Swarm Optimization and Artifical Bee Colony algorithms for Multilevel Thresholding", Elseiver June2013.
13. N. Suguna1, and Dr. K. Thanushkodi, "An Improved k-Nearest Neighbor Classification Using Genetic Algorithm", IJCSI International Journal of Computer Scien ce Issues, , Vol. 7, 2, July 2010.
14. Fardin Ahmadizar, Khabat Soltanian, Fardin Akhlaghian, Loannis Tsoulos," Artificial Neural Network development by means of a novel combination of Grammatical Evolution and Genetic Algorithm", Volume 39, Elseiver March 2015.
15. Rina D.Rudyanto, Sjoerd Kerkstra,Eva Mvan Rik Xoart," Comparing algorithms for automated vessel segmentation in Computed Tomography scans of lung", Volume 18, Elseiver Oct 2014.