

Application of Machine Learning Techniques for the Diagnosis of Lung Cancer with ANT Colony Optimization

Rashmee Kohad
Department of Computer Science
Jawaharlal Nehru Engineering College
Aurangabad, India.

Vijaya Ahire
Department of Computer Science
Jawaharlal Nehru Engineering College
Aurangabad, India.

ABSTRACT

Lung cancer is the leading cause of cancer-related death in the world-wide. The prognosis is poor, with less than 15% of patients surviving 5 years after diagnosis. The poor prognosis is attributable to lack of efficient diagnostic methods for early detection and lack of successful treatment metastatic disease. However, persons with early lung cancer have lower lung cancer-related mortality than those with extensive disease, suggesting early detection and treatment of lung cancer might be beneficial. Computer Tomography is a clinical recommended imaging tool for the detection, diagnosis and follows up of many cancers. The most essential and challenging task for physicians is detection of lesions. Detection of lesions is found to be most difficult because of low contrast of an image or sometimes their residual. This paper is aimed to design computer aided diagnosis (CAD) system which has four different phases for detection of cancerous nodules from lung CT (Computer Tomography) images and they are preprocessing, feature extraction, feature selection and classification. Ant colony optimization as feature selection technique is being used to get more accurate result. Two kinds of machine learning techniques, viz., SVM (Support Vector Machine) and ANN (Artificial Neural Network) have been presented which classify the abnormal or normal lung image. ANN has proved that it gives the best accuracy as compared to SVM. The system accuracy using SVM is 93.2% and ANN is 98.40%. In this study, our methods are validated via a series of experimentation conducted with a data set of 250 lung CT images and the procedure is implemented using MATLAB tool.

Keywords

Computer aided diagnosis (CAD) system, lung nodule, thresholding, ant colony optimization (ACO), support vector machine (SVM), artificial neural network (ANN).

1. INTRODUCTION

We Lung cancer is the leading cancer killer in both men and women worldwide. According to the World Health Organization (WHO) [1], approximately 14 million new cases and 8.2 million cancer related death in 2012. As per American Lung Association [2] - [3], deaths due to Lung Cancer is more than the next most common cancers combined (colon, breast and pancreatic). An estimated 158,040 Americans are expected to die from lung cancer in 2015, accounting for approximately 27 percent of all cancer deaths. The number of deaths due to lung cancer has increased approximately 3.5% between 1999 and 2012 from 152,156 to 157,499. In 2012,

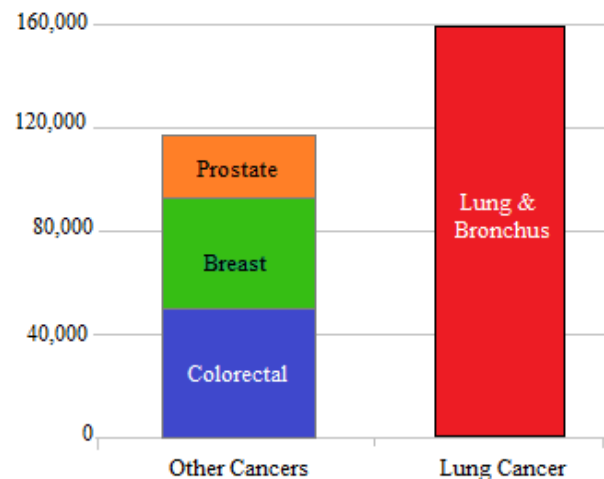


Fig. 1 Estimated Cancer Deaths by Site, 2015

there were 86,740 deaths due to lung cancer in men and 70,759 in women. Figure 1 shows the estimated cancer deaths by site, 2015.

1.1 Lung Cancer

Cancer [4], that forms in tissues of the lung, usually in the cells lining air passages. There are three main types of lung cancer.

1.1.1 Small cell lung cancer

In other word, it is also called as oat cancer. About 10%-15% of lung cancers are small cell lung cancers. This type of lung cancer tends to spread quickly.

1.1.2 Non-Small cell lung cancer

This is the most common type of lung cancer. About 85% of lung cancers are non-small cell lung cancers. Squamous cell carcinoma, adenocarcinoma, and large cell carcinoma are all subtypes of non-small cell lung cancer.

1.1.3 Lung carcinoid tumor

Fewer than 5% of lung cancers are lung carcinoid tumors. They are also sometimes called lung neuroendocrine tumors. Most of these tumors grow slowly and rarely spread. Cancer mortality can be reduced if cases are detected and treated early. There are two components of early detection efforts: Early diagnosis and Screening. Healthy lung tissues form darker regions in CT images compared to other parts of the chest. Lung nodules are small masses of tissue in the lung which are quite common. They appear as round, white shadows on the chest computerized tomography scan. They are usually about 0.2 inch (5 millimeters) to 1 inch (25 millimeters) in size. A larger lung nodule such as nodule with 25 mm or larger size is more likely to be cancerous nodule. A pulmonary nodule is a small round or oval-shaped growth in

the lung. It is sometimes also called a spot on the lung or a coin lesion. Pulmonary nodules are generally smaller than 3 centimeters in diameter. If the growth is larger than that, it is known as a pulmonary mass. Lung cancer may be found as a mass or tumor on a chest computerized tomography of a patient. Tumors can be benign or malignant. Benign tumors usually can be removed and do not spread to other parts of the body. Malignant tumors, a term used to refer to cancerous cells or tumors [5]. For the radiologist, detection of such lesions from chest radiographs is one of the challenging tasks. Such lesions may not be detected due to the fact that they may be camouflaged by underlying anatomical structure or low-quality of images or subjective and variable decision criteria used by radiologist. In order to detect lung cancer in its early stages, regular screening is very important to survive patient's life because symptoms appear in the advanced stages where the chances of survival are very low. Recently some medical researchers have proven that analysis of lung nodule can assist for successful diagnosis of lung cancer. Hence the aim of this paper is to present the CAD system model that would be the great support for pathologist and radiologist to handle large amounts of data. CAD system in general consists of four stages; pre-processing, feature extraction, feature selection and classification. Figure 2 shows a block diagram of the main processing component of CAD system.

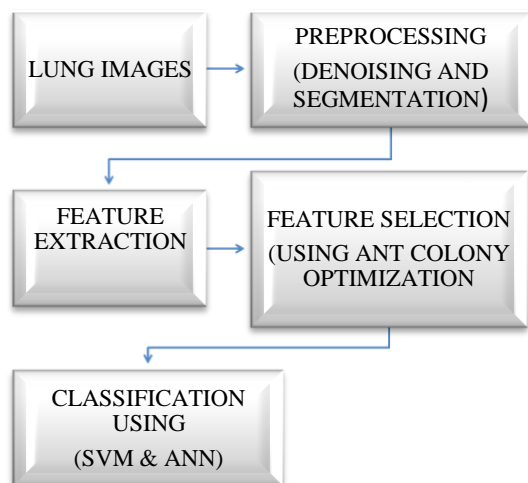


Fig. 2 Proposed CAD system

2. RELATED WORK

Different types of machine learning techniques have been developed to improve the diagnosis accuracy rate of lung cancer. For that, accordingly CAD systems are presented. S. Ashwin, et al. [6] and Ada, et al. [7] presented CAD system based on Artificial Neural Network for the lung nodule detection. Their proposed systems are given accuracy 96.7% and 96.03% respectively. Anam Tariq, et al. [8] presented a proposed method based on Neuro-Fuzzy classifier for diagnosis of lung cancer which gives system accuracy 95%. Using local binary method, for cancer diagnosis, Yeni Hardi, et al. [9] presented probabilistic neural network which has accuracy 78% for 3D and 43% for 2D. Fatma Taher [10]-[11] presented two proposed method, one is based on Bayesian classification and other one is based on Artificial Neural Network and Fuzzy Clustering method for diagnosis of lung cancer having accuracy of system is 88.6%. Hamada R.H., et al. [12] proposed CAD system based on K-nearest neighbor classifier which has accuracy 96.5%. In this paper novel contribution of proposed system lies in the following: A. Using Support vector machine B. Artificial neural network.

The author [13] has implemented CAD system using SVM, which gives the accuracy 96.6 % on database of 150 samples for detection of lung cancer. The main objective of our proposed work is to extract optimum distinguishing features from lung nodule, which will help the CAD system to examine the accuracy of classifier ANN with SVM for lung cancer diagnosis. Since the existing system [13] has worked on 150 samples. In the proposed system, 250 samples are used and accordingly accuracy rate using SVM and ANN has been examined.

3. PROPOSED SYSTEM

The proposed system is composed of following different stages; Database, Pre-processing, Feature extraction, Feature selection and Classification.

3.1 Data Base

The proposed CAD system has worked on total 250 lung CT images, out of which 125 samples are cancerous cases with single and multiple nodules and other 125 samples are normal lung images. Nearly 190 samples are collected from web site cancerimagingarchives.net and the rest of samples are from Mahatma Gandhi Mission and Tapadia Diagnostic Center situated at Aurangabad. All images are in JPEG form with 512*512 pixels resolution. The size of nodules in database is ranging from 3.7 mm to 17.2 mm in diameter. The criteria for deciding the size of lung nodule/mass is based on [14].

3.2 Pre-Processing

The objective of image pre-processing is to smoothness, enhancement and segmentation. Because of pre-processing technique, the quality of CT lung image will be improved and emphasize certain features that makes features extraction/selection and classification more effective. In proposed system, pre-processing involves following steps.

3.2.1 De-noising

Improvement in the quality of degraded images due to the effect of CT scan can be achieved by using application of different enhancement techniques. There are different types of noises that corrupt the image ex. Additive noise, Poisson noise, Gaussian noise, etc. In order to identify the tumors or cancer, the edges must be preserved. In proposed system, we have used un-sharp masking in order to highlight fine details within an image. Figure 3(b) shows the de-noised lung image.

3.2.2 Segmentation

Image segmentation is an essential process for image analysis. Image description and recognition depend on the segmentation results. Segmentation subdivides an image into its constituents regions or objects. The level of detail to which the subdivision is carried out depends upon problem being solved. That is, segmentation should stop when objects or regions of interest have been detected. In this paper, work has been carried out on global thresholding technique [15]. For extracting lung lobes from rest of chest parenchyma, segmentation plays an important role so that the next steps will be ease to work. Thresholding is one of the most powerful tools for image segmentation. The segmented image obtained from thresholding has the advantages of smaller storage space, fast processing speed and ease in manipulation, compared with gray level image which usually contains 256 levels. Therefore, thresholding techniques have drawn lot of attention during the past 20 years. Global thresholding has good performance in case of separation between lung lobes from rest of lung parenchyma. The following steps are involved in segmentation based on global thresholding. The resultant segmented lung image is as shown in figure 3(b).

Step 1: Select the computed tomography image and de-noised it.

Step 2: To extract the object of interest from background is to select an initial estimate for global threshold, T.

Step 3: Segment the image $g(x, y)$ is given by,

$$g(x, y) = \begin{cases} 1 & \text{if } f(x, y) > T \\ 0 & \text{if } f(x, y) \leq T \end{cases} \quad (1)$$

Which results two group of pixels: G1 consisting of all pixels with intensity values $> T$, and G2 consisting of all pixels with intensity values $\leq T$.

Step 4: Compute average mean intensity values m_1 and m_2 for pixels in G1 and G2 respectively.

Step 5: Compute a new threshold value:

$$T = 1/2(m_1 + m_2).$$

Step 6: Repeat steps b to 4 until the difference between values of T in successive iterations is smaller than a predefined parameter ΔT .

3.2.3 Post processing Enhancement

Mathematical morphing is used as final step for smoothing the region of interest and describing the shapes with certain chosen structuring element. Erosion with structuring removes small-scale details from an image but simultaneously reduces size of regions of interest too. Dilation has the opposite effect to erosion; it adds the pixels to both inner and outer boundaries of regions of an image. Opening and closing operations are derived from erosion and dilation of morphing. Opening is used to smoothes the contour of an object; breaks narrow isthmus and eliminate thin protrusions. Closing is also used to smooth the contour but removes the small holes; fuses narrow breaks, long thin gulfs and fill the gaps in contour. The resulting lung image after pre-processing is as shown in figure 3.

$$\text{Erosion } A \ominus B = \{e | (B)e \subseteq A\} \quad (2)$$

$$\text{Dilation } A \oplus B = \{e | (B)e \cap A \neq \emptyset\} \quad (3)$$

$$\text{Opening } A \ominus B = (A \ominus B) \oplus B \quad (4)$$

$$\text{Closing } A \oplus B = (A \oplus B) \ominus B \quad (5)$$

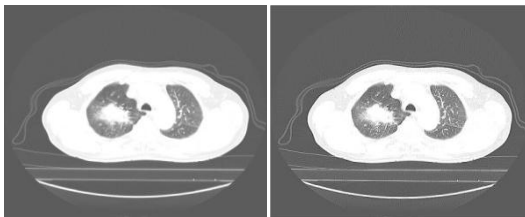


Fig. (a)

Fig. (b)

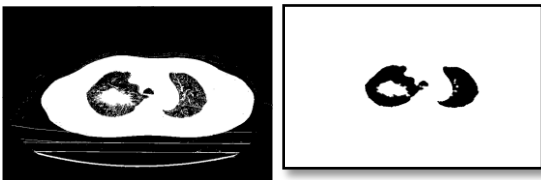


Fig. (c)

Fig. (d)

Fig. 3. Lung image after pre-processing; (a) Original lung CT image Fig. (b) Gray level image Fig. (c) Thresholded image Fig (d) Ssegmented lung image after post-processing

3.3 Feature Extraction

Feature extraction is the process of acquiring higher level information of an image, such as color, shape, and texture. Features contain the relevant information of an image and will be used in image processing (e.g. searching, retrieval, storing). Feature extraction involves reducing the amount of resources required to describe a large set of data. In the proposed system, for extracting the feature vector, GLCM (Gray level co-occurrence matrix) as feature extraction technique has been used. This method was proposed by Haralick in 1973 and still one of the most popular means of texture analysis [16]. The GLCM functions characterize the texture of an image by calculating how often pairs of pixel with specific values and in a specified spatial relationship occur in an image. Working with GLCM, main steps involved are creating GLCM, specifying offsets and then extracting statistical measures from this matrix. Using GLCM, total extracted features are 22, out of which 19 texture features and 3 shape based features. Texture features are most essential for identifying object or ROI (region of interest) and shape based features describe the structure of nodule in lung CT Image. These features are further used to input for feature selection algorithm.

3.4 Feature Selection

The feature selection technique is used in order to reduce the dimension of dataset. It significantly reduces the spurious information those are irrelevant, redundant and noisy and retain a feature subset which is most relevant. As a result, number of good outcomes can be expected from applications, such as speeding up data mining algorithms, improving mining performance and comprehensibility of result. In this paper proposed method presents a feature selection algorithm by utilizing the strategy of Ant Colony Optimization [17]. ACO is proposed by Marco Dorigo [18] for image feature selection. ACO algorithm is inspired by the natural behavior of ant species which deposit pheromone on their path to guide their foraging. The flowchart of proposed ACO algorithm is as shown in figure 4. In proposed method, each feature is represented as node and all nodes are independent of each other. Nodes (i.e. features n) are selected according to their selection probability.

3.4.1 Steps for following ACO algorithm

Based on [19], ACO algorithm follows two main steps and they are:

3.4.1.1 Route Construction

Initially, the moving ants construct a route randomly on their way to food. However, the subsequent ants follow a probability-based route construction scheme.

3.4.1.2 Pheromone Update

This stage involves two important stages. Firstly, a special chemical “pheromone” is deposited on the path traversed by the individual ants. Secondly, this deposited pheromone is subjected to evaporate. The quantity of pheromone updated on an individual path is cumulative effect of these two stages.

3.4.2 ACO-Based feature selection algorithm

3.4.2.1 Initialization Process

Parameters α , β and heuristic information are to be set. The number of ants is $K = \sqrt{M}$ where M is feature vector length. Each ant is propagated to one feature node. Initial value of pheromone matrix is set to a constant value.

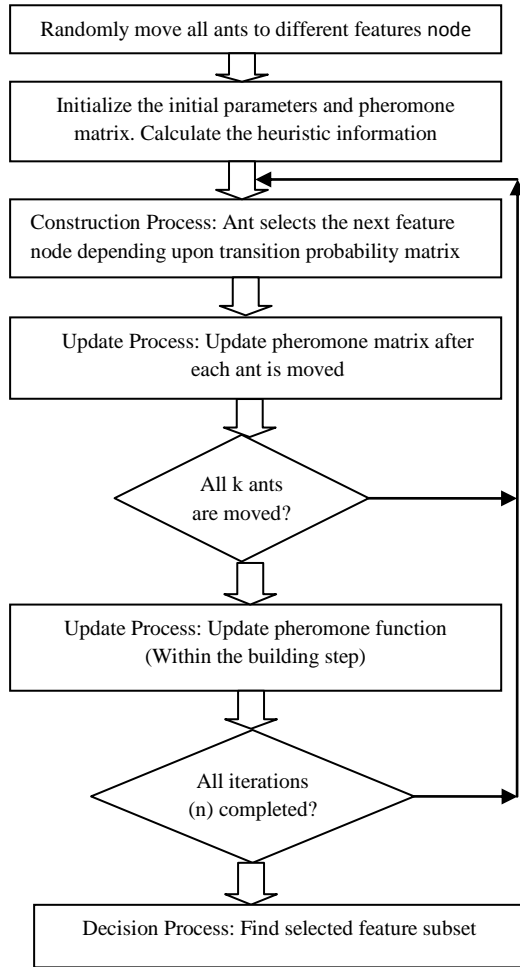


Fig. 4 Flowchart of Ant Colony Optimization Algorithm

3.4.2.2 Construction Process

To select the next feature, ant moves from current feature node I to the neighboring feature node l . For that probability transition matrix has to be calculated for which contains the probability of transition values for each feature node. Probability transition matrix is given by,

$$p_k(i) = \frac{(\tau(i))^\alpha \cdot (\eta(i))^\beta}{\sum_{d \in N_i^k} (\tau(i))^\alpha \cdot (\eta(i))^\beta} \quad (6)$$

Where,

$\tau(i)$: Pheromone value of node(feature) i .

N_i^k : “feasible” neighborhood of ant k , i.e., all features as yet unvisited by ant k .

d : the neighboring feature node of i .

η_i : the heuristic information of the feature node i .

For all $i=1, 2, \dots, n$.

3.4.2.3 Update Process

In this process, after each ant is moved, the pheromone matrix would be updated.

$$\tau_i^n = \begin{cases} ((1 - \rho) \cdot \tau_i^n + \rho \cdot \Delta_i^{(k)}) & \text{if } i \text{ belongs to best result} \\ \tau_i^n & \end{cases} \quad (7)$$

Where ρ is evaporation rate

Δ_i^k is determined by heuristic matrix i.e. $\Delta_i^k = \eta_i$

For all $i=1, 2, \dots, n$.

The heuristic information is added into each ant’s memory and used for further steps. The second update is made at the end of each building step i.e. all ants K within the step have been moved, the pheromone matrix will be updated and is represented by,

$$\tau^n = (1 - \psi) \cdot \tau^n + \psi \quad (8)$$

Where

Where, ψ is pheromone decay coefficient.

3.4.2.4 Decision Process

Being a very important process as it incorporates the result from the previous steps to determine the best features; pheromone matrix has to be visualized. In the proposed system, threshold value for pheromone has been used for selecting the feature. Pheromone value which is greater than the threshold, select the corresponding feature and generate feature subset s . this feature subset s is then used for classification.

3.5 Classification

All classification algorithms are based on the assumption that the image in question depicts one or more features and that each of these features belongs to one of several distinct and exclusive classes. The classes may be specified *a priori* by an analyst (as in supervised classification) or automatically clustered (i.e. as in *unsupervised* classification) into sets of prototype classes, where the analyst merely specifies the number of desired categories. Hence, the task of classifier component of system is to use the feature vector provided by the feature extractor to assign the object to a category. In the proposed system two classifiers have been presented, support vector machine and artificial neural network. While working with these classifiers, it has found that the system accuracy with artificial neural network is quite higher than support vector machine even though the samples which are used to train SVM are same as that for ANN. The working of these two classifiers and management of dataset for training and testing are explained below.

3.5.1 Using Support Vector Machine

Support vector machine was initially developed for pattern classification or recognition task [20]. Pattern classification or recognition is to label the object into one of the specified category called class. It is also called as binary classifier when the data samples are having exactly two classes, like our data has abnormal and normal lung image. SVM classifiers [21] are based on the class of hyper planes. The optimal hyper plane, as shown in figure 5 defined as one with the maximal margin of separation between the two classes (normal and abnormal).

$$(w \cdot x_i + b) \geq +1, \quad \text{for } y_i = +1 \quad (9)$$

or

$$(w \cdot x_i + b) \leq -1, \quad \text{for } y_i = -1, i=1,2,\dots,k. \quad (10)$$

These constraints can be re-written as

$$y_i(w^T \cdot x_i + b) \leq 1 \quad \text{Where } i=1, 2, 3, \dots, k. \quad (11)$$

This specific problem formulation may not be useful in practice because in many cases that the model does not classify correctly. In such situation, SVM [22] finds hyper plane that maximizes the margin and minimizes the

misclassification, for that purpose, slack variables, denoted by the data may not be completely separable by a hyper plane. In this case, slack variables, denoted by ξ can be used to relax the separability constraints in (11) as follows:

$$y_i(w^T \cdot x_i + b) \leq 1 - \xi_i, \quad \xi_i \geq 0; i=1, 2, \dots, k.$$

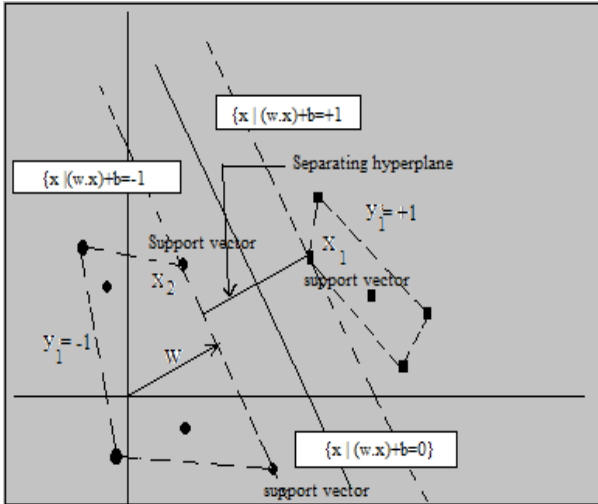


Fig. 5 SVM with linear separating hyper plane

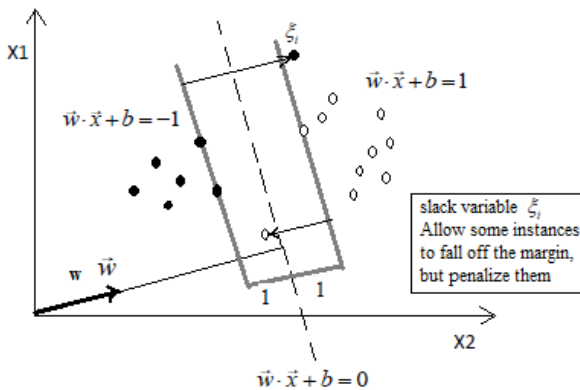


Fig. 6 SVM with separating hyper plane using slack variable

An ideal SVM analysis should produce a hyper plane that completely separates the cases into two non-overlapping classes. The simplest way to separate two groups of data is with a straight line (1 dimension), flat plane (2 dimensions) or an N-dimensional hyper plane. However, there are situations where a nonlinear region can separate the groups more efficiently. SVM handles this by using a kernel function (nonlinear) to map the data into a different space where a hyper plane (linear) cannot be used to do the separation. It means a non-linear function is learned by a linear learning machine in a high-dimensional feature space while the capacity of the system is controlled by a parameter that does not depend upon the dimensionality of the space. This is called kernel trick that means kernel function transform the data into a higher dimensional feature space to make it possible to perform linear separation. There are following four basic kernels.

- Linear: $K(x_i, x_j) = x_i^T x_j$,
- Polynomial: $K(x_i, x_j) = (\gamma x_i^T x_j + r)^d, \gamma > 0$

- Radial basis function (RBF):

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2), \gamma > 0.$$

- Sigmoid: $K(x_i, x_j) = \tanh(\gamma x_i^T x_j + r)$.

If the number of features is too large, then it may not require mapping the data into higher dimension space. That is, non-linear mapping does not improve the performance then the linear kernel will be the good choice of mapping. In the proposed method, SVM works with RBF kernel that will map the data into higher dimension space and the result has been analyzed.

3.5.2 Using Artificial Neural Network

The proposed neural network [5]-[6] consists of three layers of nodes: an input layer, hidden layer and an output layer. Input units depend upon the outcomes of feature selection algorithm (ACO). As the feature subset contains the most optimum features (in the proposed method 3 features are selected) only that much number of input unit of ANN will be used. Hidden units are 10 and one output unit that will give two outputs whether the case is normal or abnormal. The architecture of neural network is as shown in figure 7.

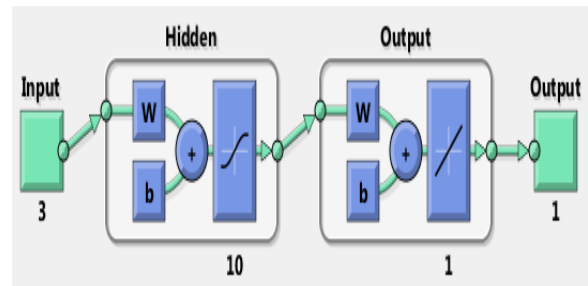


Fig.7 Neural Network Architecture

Every neural element in a layer is interconnected with every element in adjacent layers. The strength of every interconnection is characterized by its weight. Information prorogates from the input layer to the output layer. Each neural element weights the inputs it receives from the elements in the previous layer using the appropriate interconnection weights. Subsequently, the sum of the weighted inputs is filtered through a logistic function to produce an output from the element. The outputs from the output layer represent the final prediction of the neural network. The parameters of feed-forward network were tan-sigmoid transfer function in the hidden layer and linear transfer function in the output layer with train function "trainlm". The output of the ANN represented the likelihood of abnormality: 0 was normal lung and 1 for abnormal lung. And this difference was based on that the output value > 0.5 would be lung cancer.

3.5.3 Training and Testing networks

Like any machine learning technique, involves two basic steps namely training and testing. Training classifier involves feeding known data to classifier along with previously known decision values, thus forming a finite training set. From the training set, the classifier gets its intelligence to classify unknown data. In testing phase, unknown data are given and classification is performed using trained classifier. For abnormal case is and for normal case is 0.

4. PERFORMANCE MEASURES AND RESULTS

4.1 Using Support Vector Machine

For SVM classifier, 200 samples are used for training and 50 samples are used for testing. Results of abnormal and normal lungs are shown in fig 8(a) and fig 8(b) respectively.

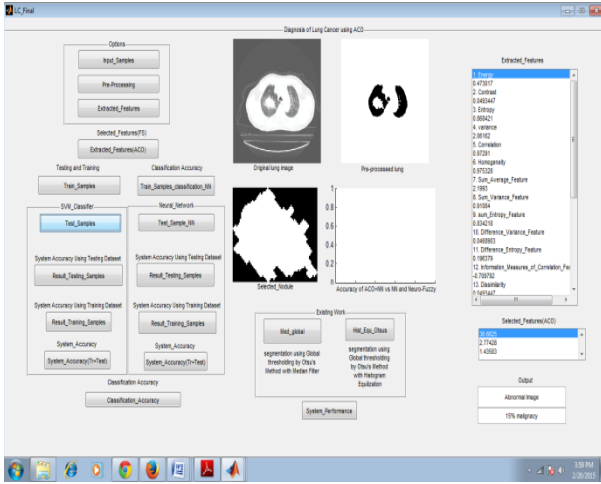


Fig. 8(a) Result of Abnormal Lungs

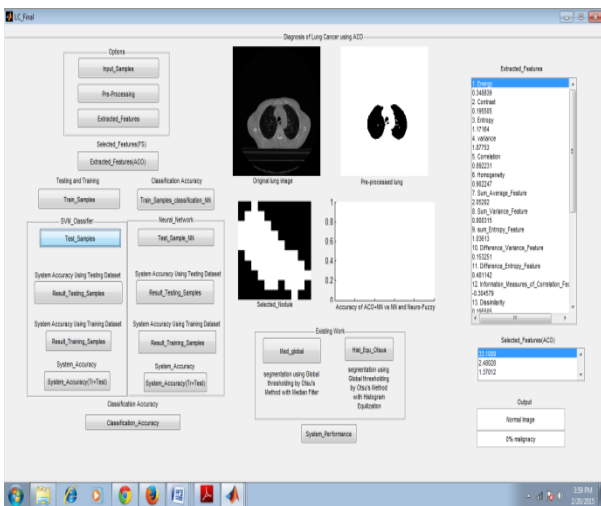


Fig. 8(b) Result of Normal Lungs

4.2 Using Artificial Neural Network

For ANN classifier, the entire data set (250 samples) was randomly divided, where 70% of data is used for training, 10% data is used for validation, which is used to validate if the network is generalizing and thus stop training before over fitting and 20% data is used for testing which have no effect on training, but provide completely independent test of network generalization.

4.2.1 Mean Squared Error

The mean squared error is the difference between the output and the target. Validation vectors prompts to stop training early if the network performance on the validation vectors fails to improve or remains the same, as indicated by an increase in the mean square error of the validation samples. We further perform a check based on test vectors to see if the network is generalizing well, but do not have any effect on training. The best validation performance is 0.0032944 at epoch 70 as shown in figure 9(a).

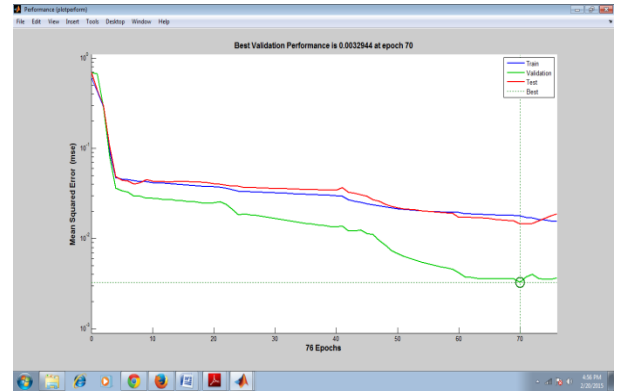


Fig. 9(a) Performance Curve

4.2.2 Regression Analysis

Figure 9(b) shows the regression plot of network outputs with respect to targets for training, testing, validation and test sets. Solid line indicates the best fit linear regression line between outputs and targets. R indicates the relationship between the outputs and targets, if R=1, then there is exact relationship between outputs and targets. If R=0 then there is no relationship between outputs and targets.

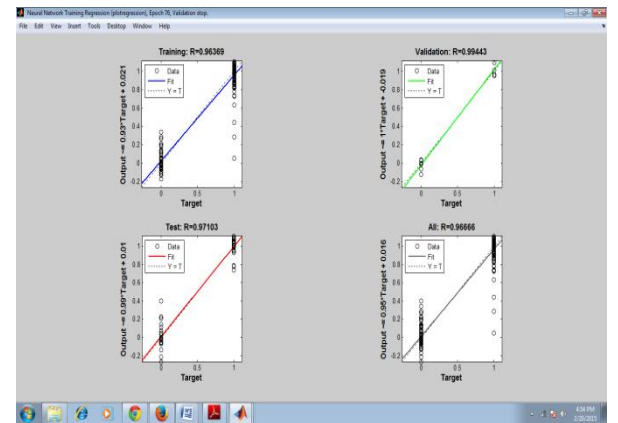


Fig. 9(b) Regression Plot



Fig. 9(c) Confusion Matrix

4.2.3 Confusion Matrix

A confusion plot can also be used to visualize the performance of neural network as shown in figure 9(c). Here the confusion matrix is plotted across all samples. The confusion matrix shows the percentage of correct and incorrect classification. Correct classifications are the green squares on the matrices diagonal and incorrect classifications form red squares. The result shows the best recognition by proposed neural network.

4.2.4 Error Histogram

Figure 9(d) shows the error histogram of proposed network which indicates how the error sizes are disseminated. The Three performance measure terms Accuracy (AC), Sensitivity (SE) and Specificity (SP) are used to evaluate the performance of the classifier. Sensitivity relates to the test's ability to identify a condition correctly. Specificity relates to the test's ability to exclude a condition correctly. Classification accuracy is depends on the number of samples correctly classified. These terms are defines follow.

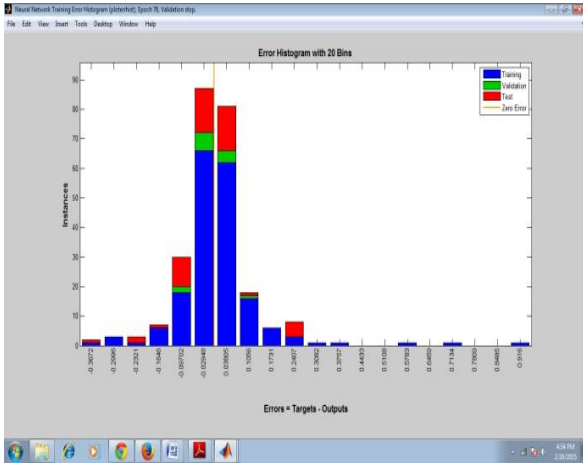


Fig 9(d) Error Histogram

$$AC = \frac{TP + TN}{TP + FP + TN + FN} \quad (15)$$

$$SP = \frac{TN}{TN + FP} \quad (16)$$

$$SE = \frac{TP}{TP + FN} \quad (17)$$

Where,

True positive = correctly identified abnormal lung.

False positive = incorrectly identified abnormal lung.

True negative = correctly rejected abnormal lung.

False negative = incorrectly rejected abnormal lung.

In order to check the efficiency of classifiers of proposed method, ACO_SVM is compared with ACO_ANN. Table 1 is represented as confusion matrix. Table 2 represents performance of system.

Table 1. Confusion matrix of proposed algorithm

Screen test outcome	ACO_SVM		ACO_ANN	
	Predicated		Predicted	
	Abnormal (Positive)	Normal (Negative)	Abnormal (Positive)	Normal (Negative)
Abnormal (Positive)	TP=112	FP=04	TP=121	FP=0
Normal (Negative)	FN=13	TN=121	FN=04	TN=125

Table 2. Performance of proposed algorithm

Performance	ACO_SVM	ACO_ANN
Features Extracted	22	22
Features selected	3	3
Accuracy	93.2%	98.40%
Sensitivity	89.60%	96.8%
Specificity	96.8%	100%

5. CONCLUSION

In this paper, the classifiers SVM and ANN were applied to the medical image diagnosis of lung cancer and the result of feedback ACO_SVM was compared with ACO_ANN. ACO algorithm has higher processing speed. With ACO algorithm, smaller feature subset has been obtained that ultimately increases the classification speed. The accuracy of proposed system using ANN classifier is higher than SVM classifier. The reason of using ANN's comes from the fact that they have several advantageous properties, such as higher generalization capability than SVM and the capability of learning from training data, even where the rules are not known a priori. Using ACO_ANN, the system accuracy reaches to 98.40 % which higher than the accuracy is gained by SVM.

6. REFERENCES

- [1] World health organization, <http://www.who.int/mediacentre/factsheets/fs297/en/>
- [2] American Cancer Society <http://www.cancer.org/research/cancerfactsstatistics/cancerfactsfigures2014/>
- [3] American Lung Association <http://www.lung.org/lung-disease/lung-cancer/resources/facts-figures/lung-cancer-fact-sheet.html>
- [4] American Cancer society <http://www.cancer.org/cancer/lungcancer/>
- [5] Mayo clinic, "Lung Cancer" <http://www.mayoclinic.com/health/lung-nodules/AN01082>.
- [6] S. Ashwin, J. Ramesh , "Efficient and reliable lung nodule detection using NN based CAD system". IEEE, ICETEEEM, PP 135-142, 2012.
- [7] Ada, Rajneet Kaur , " Early Detection and Prediction of Lung Cancer Survival using Neural Network Classifier", IJAIEEM, Volume 2, Issue 6,PP 375-383,June 2013
- [8] Anam Tariq,M. Usman , " Lung Nodule Detection in CT images using neuro fuzzy classifier". IEEE , CIMI, PP 49-53, 2013.
- [9] Yeni Hardi yeni, "Diagnosis of lung cancer using 2D and 3D local binary pattern". IJACSA, Vol 3, No. 4, PP 89-95, 2012.
- [10] Fatma Taher , "Bayesian classification and ANN for diagnosis of lung cancer", IEEE, PP 773-776, 2012 .
- [11] Fatma Taher , "Lung Cancer Detection by Using Artificial Neural Network and Fuzzy Clustering

- Methods”, American Journal of Biomedical Engineering, PP 136-142, 2012.
- [12] Hamada R. H., A- Absi ,” CAD System Based on M/C Learning Techniques for Lung Cancer”, IEEE , ICCIS, PP 295-300, 2012.
- [13] Rashmee Kohad, Vijaya Ahire, “Diagnosis of Lung Cancer Using Support Vector Machine with Ant Colony Optimization Technique”, IJACST, Vol.3, No.11, Pages:19-25 (2014)
- [14] Radiology assistant, <http://www.radiologyassistant.nl/en/p460f9fcd50637/solitarypulmonary-nodule-benign-versus-malignant.html>
- [15] Anjali Gautam, H.S. Bhadauria,” White Blood Nucleus Segmentation Using an Automated Thresholding and Mathematical Morphing”, ICAET-2014.
- [16] Robert M. Haralick,” Texture Features for Image Classification”, IEEE Transaction on systems, MAN And Cybernetics, PP 610-621, November 1973.
- [17] Ling Chen, Bolun Chen, Yixin Chen, Image Feature Selection Based on Ant Colony Optimization
- [18] The United States of America. Library of Congress Cataloging-in-Publication Data . Dorigo, Marco. Ant colony optimization / Marco Dorigo, Thomas Stützle. p. cm.
- [19] P.Thukaram,” Image Edge Detection Using Improved Ant Colony Optimization Algorithm”, International Journal of Research in Computer and Communication Technology, PP 1256-1260, Vol 2, Issue 11, November-2013
- [20] Bottou, L., and Chih-Jen Lin. Support Vector Machine Solvers. Available at <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.64.4200&rep=rep1&type=pdf>
- [21] Support Vector Machine, <http://pages.cs.wisc.edu/~jerryzhu/cs540/handouts/heardst98-VMtutorial.pdf>
- [22] Chih-Wei Hsu, Chih-Chung Chang, and Chih-Jen Lin, “A Practical Guide to Support Vector Classification”