ISSN 2320 -2602

**International Journal of Advances in Computer Science and Technology (IJACST)**, Vol.3 , No.11, Pages : 19-25 (2014)
*Special Issue of ICCAAC 2014 - Held on December 08,2014,Hyderbad, India*

# Diagnosis of Lung Cancer Using Support Vector Machine with Ant Colony Optimization Technique

**Rashmee Kohad[1], Vijaya Ahire[2]**

Ms. Rashmi Kohad
*Department Of Computer Science, ME-Student,*
*Jawaharlal Nehru College of Engineering,*
*Aurangabad, India.*
Email-id: - rashmee_kohad@rediffmail.com

Ms. Vijaya Ahire
*Department Of Computer Science, Faculty,*
*Jawaharlal Nehru College of Engineering,*
*Aurangabad, India.*
Email-id: - vijayahire19@gmail.com

**Abstract:** Cancer is a leading cause of death worldwide. About 60% cancerous cases occur in Africa, Asia and south and Central America, from which 30% of cancers could be prevented. Lung cancer is a type of cancer mortality contributing about 1.3 million deaths/year globally. To reduce the mortality of lung cancer, early detection of cancer greatly increases the chances of successful treatment.

In this paper an automated diagnosis system has been developed for diagnosis of lung nodule. The system consists of four phases, pre-processing, features extraction, features selection and classification. We are aiming to get the more accurate result by using Ant Colony Optimization as feature selection technique using which most relevant features are obtained for building robust learning model. In this study, MATLAB have been used through every procedures made. The standard database has been used to test the automated system and accuracy level of over 96% can be achieved.

**Keywords:** Lung nodule, Computer aided diagnosis (CAD) system, thresholding, ant colony optimization (ACO), support vector machine (SVM).

## INTRODUCTION

Cancer is currently most common cause of death in both men and women every year. Radon gas is a natural radioactive gas that is a natural decay product of uranium. Uranium decays to form products, including radon, that emit a type of ionizing radiation. Radon gas is a known cause of lung cancer, with an estimated 12% of lung-cancer deaths attributable to radon gas, or about 20,000 lung-cancer-related deaths annually in the U.S., making radon the second leading cause of lung cancer in the U.S. [1]. In the UK, female lung cancer deaths will reach 95,000 annually in 2040, from 26,000 in 2010 – a rise of more than 350%. Male annual lung cancer deaths will increase by 8% over the same period, to 42,000 in 2040 from 39,000 in 2010 [2]. Lung cancer can be broadly classified into two main types based on the cancer's appearance under a microscope: non-small cell lung cancer and small cell lung cancer. Non-small cell lung cancer (NSCLC) accounts for 80% of lung cancers, while small cell lung cancer accounts for the remaining 20%. [3]. Fig.1 shows the cancer survival rates for a group of 1,015 metastatic non-small cell lung cancer patients who were diagnosed between 2000 and 2009. Each patient in the group was first diagnosed at CTCA and/or received at least part of their initial course of treatment at CTCA.

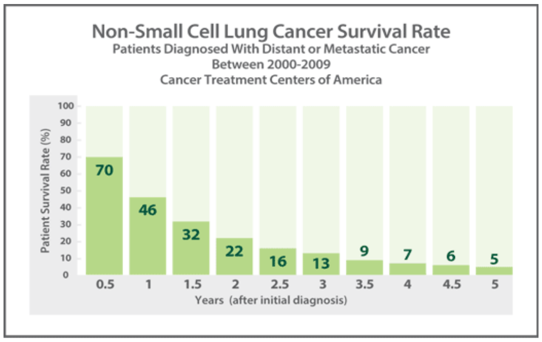

Fig. 1 Non-Small Cell Lung Cancer Survival Rate

The use of low-dose spiral computed tomography in the screening of a high-risk population has demonstrated the possibility of diagnosing small peripheral tumors that are not seen on conventional X-ray.

Healthy lung tissues form darker regions in CT images compared to other parts of the chest such as the heart and the liver. Lung nodules are small masses of tissue in the lung which are quite common. They appear as round, white shadows on a chest computerized tomography (CT) scan. They're usually about 0.2 inch (5 millimeters) to 1 inch (25 mm) in size. A larger lung nodule, such as one that's 25 mm or larger, is more likely to be cancerous nodule [4]. 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 (CT) 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.

One of the most important and difficult task the radiologist has to carry out, consists of the detection and diagnosis of cancerous lung nodules from chest radiographs. Some of these lesions may not be detected due to the fact that they may be camouflaged by the underlying anatomical structure or low-quality of the images or the subjective and variable decision criteria used by radiologist.

In order to detect lung cancer in its early stage regular screening is very important to survive the patient's life because symptoms appear in an advanced stage where the chances of survival are very low.

The goal of automatic screening will offer several advantages, such as improving the sensitivity of the test. Recently, some medical researchers have proven that the analysis of lung nodule can assist for a successful diagnosis of lung cancer, for this reason we attempt to come with a CAD system for detecting the lung cancer in its early stages based on the analysis of the computer tomography (CT) images, therefore the CAD system would be a great support for pathologists and radiologist to handle larger amounts of data.

In the literature there are number of proposals which deal with the problem of detecting cancerous nodules on digitized images from chest radiographs.

Neural network based approach [5], [6], Neuro-fuzzy based approach [7], Local binary method [8], Bayesian classifier and FCM and HNN based approach [9]-[10], CAD system based on M/C learning [11]. These methods correctly detect the nodules in images but small nodules can be detected by few and if detected then blood vessels are also detected along with nodules.

Nodules are difficult to detect in digital images because of low contrast, large variations in density, varying size and location of lung nodule within area of complicated anatomy (such as helium and ribs).

The main objective of proposed work is to extract optimum distinguishing features from lung nodule, which will help the CAD system in classifying the lung tumor.

In this paper, Computer-aided-diagnosis (CAD) system for automated detection of pulmonary nodules in computed-tomography (CT) images is proposed. The automated system is organized into five different phases, pre-processing, segmentation, generating features vector, selection of features using ACO and support vector machine based classification

## PROPOSED SYSTEM

Proposed system as shown in Fig.2 involves the following stages: A. data acquisition, B. pre-processing, C. features extraction, D. features selection, E. classification using SVM. Figure of proposed system is shown below.

A. Data Base

The database is composed of total 150 images out of 120 lung computed tomography (CT) images which have been collected from routine cases on web site cancerimagingarchives.net and rest of 30 images from Mahatma Gandhi Mission (MGM) hospital and Tapadia diagnostic center at Aurangabad. In our database, near about 73 cancerous cases with single and multiple nodules and all are confirmed by computed tomography (CT). All images are in JPEG form and have dimension 512*512.

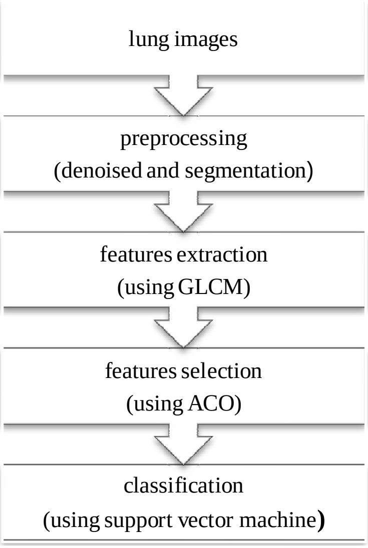

Fig.2 Proposed System

The nodule size in database ranging from 3.7 mm to 17.2 mm in diameter. The criterion used for detecting the size of nodule is based on [12].

B. Preprocessing

The main objective of image pre-processing is to enhance, smoothness, remove noise that caused by defects of CT scanner and improve the quality and emphasizes certain features of an image so that it makes segmentation or classification easier and more effective. In our paper preprocessing involves following three steps.

1. De-noising

Convert the original RGB lung computed tomography (CT) image as shown in Fig.3 into gray level image. The gray-scale image contains all the details of information; it is easy for understanding and has not ambiguities typical of black-and-white images. The medical image quality parameter is mainly noise. In order to identify tumours or cancer, the edges must be preserved.

In proposed system we have used unsharpened masking in order to highlight fine details within an image.

The resulting image after pre-processing is shown in Fig.4.

2. Segmentation

Segmentation is the process of partitioning an image into distinct non overlapping regions. The purpose of segmentation is to separate objects and background and only remain the object of interest. Nowadays numerous segmentation algorithms are developed [9], [13]. In the proposed system, global thresholding technique has been implemented.

Threshlding is the simplest and most important technique for segmentation of lung computed tomography image, where pixels are partitioned into foreground and the background depending on the distribution of gray levels or texture of an object. Global thresholding [13] has good performance in case of separation between white tumor and background gray levels. The following steps are involved in segmentation based on global thresholding.

a. Select the de-noised lung computed tomography image.
b. Select an initial estimate for global threshold, T.

c. Segment the image using T,

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

This will produce two groups of pixels, G1 consisting of all pixels with intensity values > T and G2 consisting of pixels with values <= T.

d. Compute the average(mean) intensity values m1 and m2 for the pixels in G1 and G2 respectively.

e. Compute a new threshold value :

$$T=1/2(m1 + m2) \quad (2)$$

f. Repeat steps 2 through 4 until the difference between values of T in successive iterations is smaller than a predefined parameter ∆T.

As shown in Fig.5. thresholded lung image has been obtained.

3. Post processing Enhancement

Mathematical morphing is used as the final step for smoothing the region of interest. The operations which are used as basic morphological operations are erosion and dilation. Suppose we have image A and structuring element B, then the operations are denoted as,

$$\mathit{Erosion} \quad A \ominus B = \{e|(B)e \subseteq A\} \quad (3)$$

$$\mathit{Dilation} \quad A \oplus B = \{e|(B)e \cap A \neq \emptyset\} \quad (4)$$

The opening and the closing operations are derived from the erosion and the dilation of morphing. Opening used to smoothes the contour of an object; breaks narrow isthmuses, and eliminate thin protrusions. Closing is also used to smooth contours but remove the small holes; fuses narrow breaks, long thin gulfs, and fill the gaps in contour.

$$\mathit{Opening} \; A \ominus B = (A \ominus B) \oplus B \quad (5)$$

$$\mathit{Closing} \; A \cdot B = (A \oplus B) \ominus B \quad (6)$$

Lung image after post-processing is shown in Fig.6.

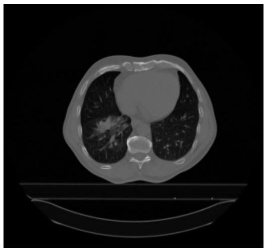

Fig. 3 Original lung CT image

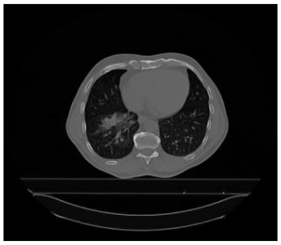

Fig. 4 De-noised lung image

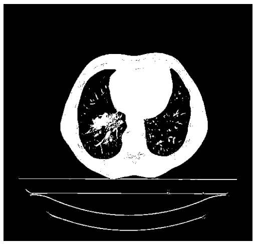

Fig. 5 Threshold image

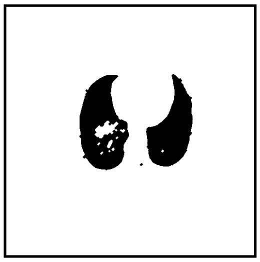

Fig. 6 Lung image after post-processing.

C. Feature Extraction

Feature extraction involves reducing the amount of resources required to describe a large set of data. When performing analysis of complex data one of the major problem stems from the number of variables involved. Analysis with a large number of variables generally request large amount of memory and computation power or a classification algorithm which over fits the training sample and generalizes poorly to new samples.

Feature extraction is general term for methods of constructing combinations of variables to get around these problems while still describing the data with sufficient accuracy. For extracting feature vector, GLCM technique has been used.

Gray-Level-Co-occurrence matrix

This method was first proposed by Haralic in 1973 and still one of the most popular means of texture analysis [14]. Haralick suggested the use of co-occurrence matrix or gray level co-occurrence matrix. It considers the relationship between two neighbouring pixels, the first pixel is known as a reference and the second is known as a neighbour pixel. In the following, we will use {I(x, y), $0 \le x \le Nx-1$, $0 \le y \le Ny-1$} to denote an image with G gray levels. The G * G gray level co-occurrence matrix $P_d^\theta$ for a displacement vector d=(dx,dy) and direction theta is defined as follows. The element (i,j) of gray level co-occurrence matrix is $P_d^\theta$the number of occurrence of the pair of gray levels I and j which the distance between i and j following direction theta is d, and Equation is,

$$P_d^\theta = \#\{((r,s),(t,v)):I(r,s) = i, I(t,v) = j\} \quad (7)$$

Where, $(r,s),(t,v) \in Nx * Ny$; $(t,v) = (r + dx, s + dy)$.

In this paper, we have extracted 19 texture features which are most essential for identifying objects or regions of interest in computer tomography (CT) image of lung and 4 shape based features, shape is an important visual feature and used to describe structure of nodule in computer tomography image of lung. These extracted texture and shape based features are further used for selection technique, ant colony optimization (ACO)

D. Feature Selection

In this paper we present a feature selection algorithm by utilizing the strategy of Ant Colony Optimization [15]. ACO is proposed Marco Dorigo [16] for image feature selection using ant colony optimization. Ant colony optimization (ACO) is an optimization algorithm inspired by the natural behavior of ant species, which deposit pheromone on the ground to guide their foraging.

After the feature extraction step, the optimum subset of features is selected with ACO based feature selection algorithm. The flowchart of our proposed ACO algorithm is presented in below Fig.7.

In proposed method, each feature represents a node, and all nodes are independent of each other. Nodes (i.e. features n) are selected according to their selection probability. The feature selection problem is to find a minimal feature subset of size s *(s< n)* while maintaining a fairly high classification accuracy in representing the original features. Using the approach [17], there are two main steps in ACO algorithms. These are:-

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.

Pheromone update**:** This step involves two important stages. Firstly, a special chemical "pheromone" is deposited on the path traversed by the individual ants. Secondly, this deposited pheromone is subject to evaporation. The quantity of pheromone updated on an individual path is a cumulative effect of these two stages.

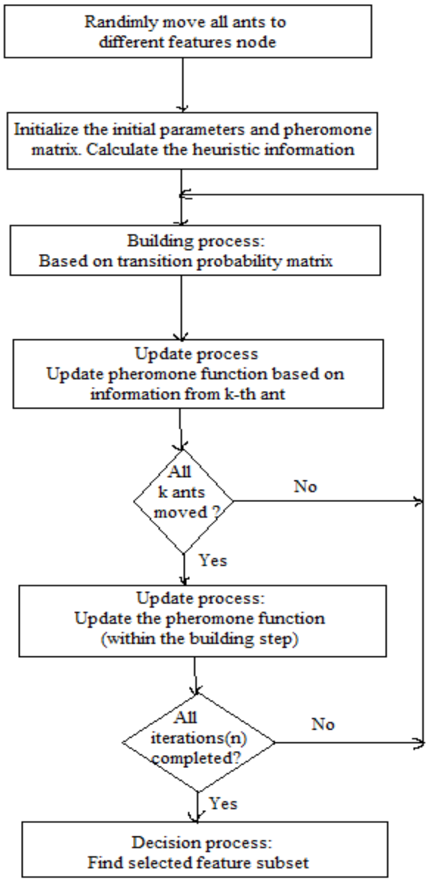
Randimly move all ants to different features node
Initialize the initial parameters and pheromone matrix. Calculate the heuristic information
Building process: Based on transition probability matrix
Update process Update pheromone function based on information from k-th ant
All k ants moved ?
No
Yes
Update process: Update the pheromone function (within the building step)
All iterations(n) completed?
No
Yes
Decision process: Find selected feature subset

Fig.7 Flowchart of proposed Ant Colony Optimization Algorithm

ACO-Based Feature selection approach

1. Initialization Process

The parameters α and β are initialized. The heuristic information is set. The number of ants is K=√M where M is the length of feature vector. All the K ants are propagated on feature vector 1 D such that at most one ant is on each feature. Every feature in the feature vector is a node and the initial value of the pheromone matrix is set to a constant value.

2. Construction Process

At each building step, an ant, which is chosen from the K ants, moves L steps on the image I. The ant k moves from feature node i to its neighbouring node l. According to the probabilistic transition matrix defined as,

$$P_k(i) = \frac{(\tau(i))^{\alpha}.(\eta(i))^{\beta}}{\sum_{l\epsilon N_i} k\ (\tau(i))^{\alpha}.(\eta(i))^{\beta}} \qquad (8)$$

$\tau(i)$ is the pheromone value of node(feature) *i*
$N_i^k$ is the "feasible" neighbourhood of ant *k*, that is, all features as yet unvisited by ant *k*.
l Is the neighbouring feature node of i.
$\eta_i$ Is the heuristic information of the feature node i.
For all i=1, 2… n.

3. Update Process

The update process, which updates the pheromone matrix after each ant is moved, is

$$\tau_i^n = \begin{cases} ((1-\rho), \tau_i^n + \rho \cdot \Delta_i^{(k)}) & \text{if i belongs to best result} \\ \tau_i^n \end{cases} \qquad (9)$$

Where $\rho$ Is evaporation rate
$\Delta_i^k$ Is determined by heuristic matrix i.e. $\Delta_i^k = \eta_i$
For all i=1, 2… n.

The heuristic information is added into the ant's memory and used for further steps. The second update is made at the end of each building step i.e. all the ants K within the step have moved. Since all the ants have moved at the end of the building step, the equation is

$$\tau^n = (1-\psi).\tau^n + \psi \qquad (10)$$

Where, $\psi$ is pheromone decay coefficient

4. Decision Process

Being a very important process as it incorporates the results from the previous steps to determine the best features; pheromone matrix has to be visualized. In this paper we have used threshold value for pheromone is 1, for selecting features. Hence, pheromone value which is greater than 1, select the corresponding feature and generate feature subset s. This feature subset is used for classification.

Using Ant Colony Optimization optimum 3 features subset is selected.

1. Cluster shade: Cluster shade is the measure of skewness of matrix, in other words the lack of symmetry. When cluster shade is high, the image is not symmetry with respect to its texture values.

2. Autocorrelation: When the series containing non-random patterns of behaviour, it is likely that any particular item in the series is related in some way to other item in the same series.

3. Standard deviation: standard deviation tells how closely the sample is located to mean.

This algorithm can obtain higher processing speed and smaller feature set than other existing methods. Higher quality classification results are obtained using such smaller feature set.

E. Classification

The reason for choosing support vector machines come from the fact that it has several advantageous properties such as they infrequently give high generalization capability and to deal with the problems with low samples and high input features. Because of its effectiveness in pattern classification and recognition, we have used support vector machine as a classifier for diagnosis of lung cancer.

Support vector machine was initially developed for pattern classification task [18]. Pattern recognition or classification is to label some object into one of the specified categories called classes. Support vector machine relies on the structural risk minimization (SRM) principle founded on the statistical learning theory, which enhances generalization capabilities.

We can use support vector machine when the data has exactly two classes. Hence, it is also called as binary classifier which gives rise to a linear hyper-plane which distinguishes a set of positive examples from a set of negative examples with extreme margin. The margin is outlined by the distance of hyper-plane to the bordering.

Support vector machine with linear separating hyper plane is shown in Fig.8.

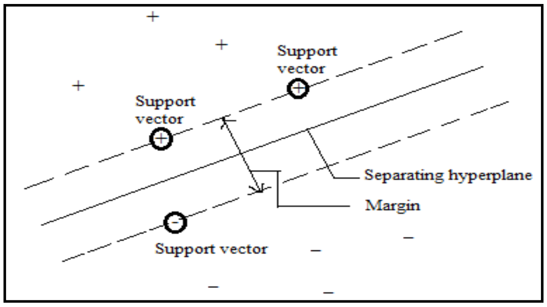

Fig. 8 Support vector machine with linear separating hyper plane

Consider the training sample set G= {(Xi,Yi), i=1….M}, where M is the number of samples and each sample Xi ∊ $R^d$ fits into a class by Yi ∊{+1,-1}. In the case of linearly separable data set, it is possible to separate the given data into two classes using hyper-plane:

$$W^T.X + b=0 \qquad (11)$$

Where, W=M dimensional vector
B=scalar bias term.

The vector W and scalar term b are employed to describe the position of separating hyper-plane.

The following decision function , D(Xi), can be employed to categorize input data into either positive class or negative class. So for a known input data Xi :

$$D(X1) = W^T Xi + b \begin{cases} > 0 \; for \; Yi = \; +1 \\ < 0 \; for \; Yi = -1 \end{cases} \qquad (12)$$

If the input data are not linearly separable, the original input space is mapped into a high-dimensional space called the feature space. It is then possible to determine a hyper-plane that allows linear separation in the feature space. SVM employs kernels to transform the data into higher dimensions. The most common kernel function is Radial Basis Functions (RBF) which is defined as:

$$K(x, y) = \exp\left(-\gamma \parallel x - y \parallel^2\right) \qquad (13)$$

In the proposed algorithm, the use of SVM, like any other machine learning technique, involves two basic steps namely training and testing. Training SVM involves feeding known data to the SVM along with previously known decision values, thus forming a finite training set. It is from the training set that an SVM gets its intelligence to classify unknown data. In testing phase, unknown data are given and the classification is performed using trained classifier. 1 for abnormal cases and 0 for normal cases.

Result of abnormal and normal lungs are shown in Fig.9 (a) and Fig.9 (b).

## PERFORMANCE MEASURES AND RESULTS

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.

$$AC = \frac{TP + TN}{TP + FP + TN + FN}$$

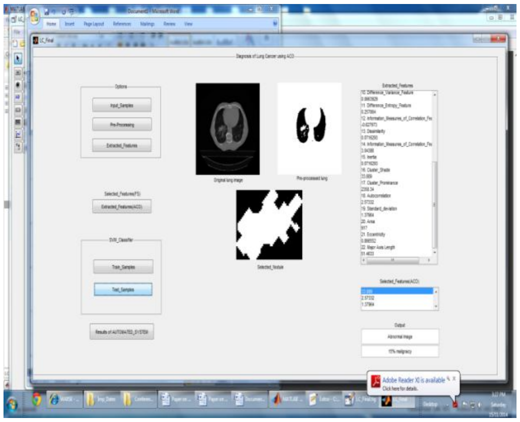

Fig. 9(a) Result of Abnormal Lungs

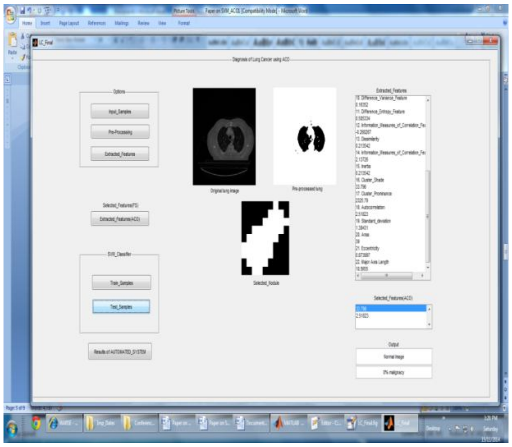

Fig. 9(b) Result of Normal Lungs

$$SP = \frac{TN}{TN + FP}$$

$$SE = \frac{TP}{TP + FN}$$

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.

A total of 60 normal cases and 60 abnormal cases are used for training and 30 cases (15 normal, 15 abnormal) for testing. The training set used for training the SVM network and test set used for estimating the accuracy of the model. In order to check the efficiency of the proposed method, ACO_SVM method is compared with NN and Neuro-fuzzy. Table 1 is represented as confusion matrix; Table 2 represents performance of system and Fig.10 shows the graphical representation of result analysis of proposed algorithm.

**Table 1:** Confusion matrix of proposed algorithm

| **Screen test outcome** | **Predicated** | |
|---|---|---|
| | **Abnormal (Positive)** | **Normal (Negative)** |
| Abnormal (Positive) | TP=16 | FP=0 |
| Normal (Negative) | FN=1 | TN=13 |

**Table 2:** Performance of proposed algorithm

| **Performance** | **ACO_SVM** |
|---|---|
| Features Extracted | 22 |
| Features selected | 3 |
| Accuracy | 96.6% |
| Sensitivity | 94.1176 |
| Specificity | 100% |

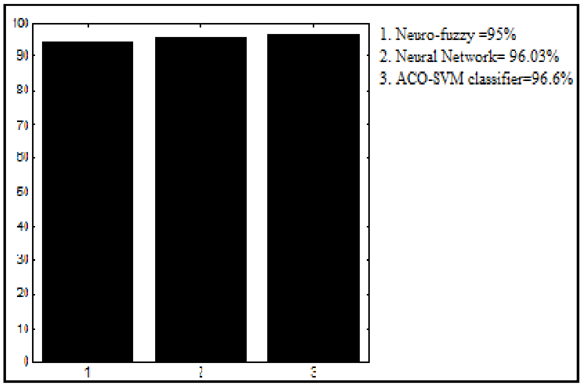

Fig. 10 Result analysis of classifiers

**Table 3:** Comparisons of proposed method with existing systems

| **Sr. No.** | **Author** | **Techniques** | **Accuracy** |
|---|---|---|---|
| 1. | S. Ashwin, J. Ramesh[5] | Pre-processing: Median filter, Adaptive histogram equalization.<br>Classification: NN using BFGS quasi-newton BP algorithm | 92% |
| 2. | Rajneet Kaur[6] | Pre-processing : Histogram Equalization and Morphological Operations.<br>Feature Extraction: GLCM and Binarization<br>Classification: PCA with Neural Network | 90.04% |
| 3. | Anam Tariq, M. Usman[7] | Pre-processing: Median filter,Thresholding<br>Feature Extraction: Area, Eccentricity, Energy, Mean, Standard deviation.<br>Classification: Neuro fuzzy classifiers | 95% |
| 4. | Yeni Hardi yeni[8] | Pre-processing: Morphological operations and thresholding based on statistical criteria.<br>Feature Extraction: Appling 2D LBP and 3D LBP.<br>Classification: Probabilistic neural network (RBF) | 3D LBP: 78%<br>2D LBP: 43% |
| 5. | Proposed Method | Pre-processing: Global thresholding with Morphological operations.<br>Feature Extraction: GLCM<br>Feature Selection: ACO<br>Classification: SVM | 96.6% |

## CONCLUSION

In this paper, the ACO-SVM algorithm was applied to the medical image diagnosis of lung cancer and the results of the feedback ACO-SVM was compared with those of the neural network trained using the back propagation algorithm and neuro-fuzzy classifier. ACO algorithm can obtain higher processing speed and smaller feature set than other existing methods. Higher quality classification results are obtained using such smaller feature set. It shows the best result with highest true positive (16) and lowest false positive rate (0). ACO_SVM gives 96.6% result as compared to other classifiers as indicated in Table 3.

## REFERENCES

[1] http://www.medicinenet.com/lung_cancer/page5.htm#how_is_lung_cancer_diagnosed.

[2] http://www.medicalnewstoday.com/info/lung-cancer/.

[3] http://www.cancercenter.com/lung-cancer/statistics/.

[4] http://www.mayoclinic.com/health/lung-nodules/AN01082.

[5] S. Ashwin, J. Ramesh , “Efficient and reliable lung nodule detection using NN based CAD system”. IEEE, ICETEEEM, PP 135-142, 2012.

[6] Ada, Rajneet Kaur ,” Early Detection and Prediction of Lung Cancer Survival using Neural Network Classifier”, IJAIEM, Volume 2, Issue 6, PP 375-383, June 2013.

[7] Anam Tariq,M. Usman ,” Lung Nodule Detection in CT images using neuro fuzzy classifier”. IEEE , CIMI, PP 49-53, 2013.

[8] Yeni Hardi yeni,”Diagnosis of lung cancer using 2D and 3D local binary pattern”. IJACSA, Vol 3, No. 4, PP 89-95, 2012.

[9] Fatma Taher , “Bayesian classification and ANN for diagnosis of lung cancer”, IEEE, PP 773-776, 2012 .

[10] Fatma Taher , “Lung Cancer Detection by Using Artificial Neural Network and Fuzzy Clustering Methods”, American Journal of Biomedical Engineering, PP 136-142, 2012.

[11] Hamada R. H., A- Absi ,” CAD System Based on M/C Learning Techniques for Lung Cancer”, IEEE , ICCIS, PP 295-300, 2012.

[12] http://www.radiologyassistant.nl/en/p460f9fcd50637/solitary-pulmonary-nodule-benign-versus-malignant.html

[13] Anjali Gautam, H.S. Bhadauria,” White Blood Nucleus Segmentation Using an Automated Thresholding and Mathematical Morphing”,ICAET-2014

[14] Robert M. Haralick,” Texture Features for Image Classification”, IEEE Transaction on systems, MAN And Cybernetics, PP 610-621,November 1973.

[15] Ling Chen, Bolun Chen, Yixin Chen, Image Feature Selection Based on Ant Colony Optimization

[16] The United States of America. Library of Congress Cataloging-in-Publication Data . Dorigo, Marco. Ant colony optimization / Marco Dorigo, Thomas Stützle. p. cm.

[17] 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

[18] Bottou, L., and Chih-Jen Lin. Support Vector Machine Solvers. Available at http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.64.4209 &rep=rep1&type=pdf.