

Enhanced Lung Cancer Detection Using Firefly Optimization and SVM on CT Images

Clara Shanthi D¹, Preethi D², M.S.Nidhya³, T.S Lakshmi⁴, Saahira Banu Ahamed⁵

¹Associate Professor, School of Computer Science and IT, Jain (Deemed to be University), Jayanagar, Bangalore, India.

ORCID: 0009-0009-1644-6967

²Assistant Professor, School of Computer Science and IT, Jain (Deemed to be University), Jayanagar, Bangalore, India.

ORCID: 0009-0007-4135-2140

³Associate Professor, School of Computer Applications, Dayananda Sagar University, Bangalore, India.

ORCID: 0000-0002-1290-4520

⁴Assistant Professor, School of Computing Sciences, Vels Institute of Science Technology and Advanced Studies, Chennai, India.

ORCID: 0009-0008-8765-9364

⁵Department of Computer Science, College of Engineering and Computer Science, Engineering and Technology research center, Jazan University, Saudia Arabia.

ORCID: 0000-0002-7019-3085

ABSTRACT

Lung cancer is a critical health concern globally, and early detection is essential for improving patient survival. This study proposes the use of Firefly Optimization algorithms for accurate segmentation of lung nodules from CT images. Features are extracted using a hybrid soft computing technique, and classification is performed using Support Vector Machines (SVM). The primary objective is to improve the accuracy of classifying tumor and non-tumor images, thereby supporting early and reliable lung cancer diagnosis.

This paper proposes the use of optimization-based algorithms for the segmentation of lung nodules and machine learning techniques to enhance the accuracy of lung cancer diagnosis. Computed Tomography (CT) scans are utilized to detect and segment lung nodules. Relevant features are then extracted from the segmented images, and an optimal subset is selected to reduce the computational time during classification. The primary objective of this study is to improve the accuracy of classifying tumor and non-tumor images, thereby supporting early and reliable lung cancer diagnosis. Lung cancer is a dangerous disease diagnosed worldwide. Early detection is crucial for patient survival, and there are many possible methods to identify lung cancer in its initial periods. Lung cancer is divided into two main categories: NSCLC and SCLC. NSCLC is further classified into carcinoma, adenocarcinoma, squamous cell carcinoma. In this paper, our aim is to utilize optimization-based algorithms for the segmented of lung nodules and soft computing approaches to increase the accuracy of diagnosing cancerous cells. This will aid in the optimal classification of tumor and non-tumor images. A CT scan is used to identify and segment lung nodules from images and to select the optimal set of features removed from the segmented lung nodules, thereby decreasing the processing time of classification. The main objective of this work is to increase the accurateness of diagnosing cancerous cells for optimal classification.

Keywords : *Lung cancer , Optimization , Tumor, Nodules, Classifier*

1. INTRODUCTION

Lung cancer is a dangerous disease diagnosed worldwide, with initial detection being critical for patient survival. Image processing and soft computing procedures are broadly used in many medical areas for image enhancement and better diagnostic accuracy. Lung cancer is the utmost commonly found type of cancer in both women and men and is also known as a lung tumor. [1] Lung cancer is commonly categorized as : NSCLC and SCLC.

[2] Lung cancer may be identified through computerized tomography (CT) scans of a patient's chest. Various methods to detect lung cancer include image processing, soft computing, and deep learning to improve computer-aided diagnosis systems. In this paper, MRI scan images, ultrasound images, and computed tomography scan images are used. [3][4]The algorithms employed are effective in terms of accurateness. The model proposed contains of the followed phases: preprocessing, nodule detection, feature extraction , collection of lung images, nodule segmentation and classification for Computed Tomography and Magnetic Resonance images.

[5] Based on the type of tissue involved cancers can be categorized into six main classes: myeloma, lymphoma, carcinoma, , sarcoma, leukemia. The initial place of origin are based on cancers can be specified as prostate, breast, brain, oral, kidney, lung or liver cancers.[6]

An abnormal growth of new tissue, known as a tumor, can form in any organ of the human body. The utmost common type of lung tumor is NSCLC, one of the leading causes of death for human beings.[7][8] survival prediction of NSCLC and Computer-aided diagnosis are crucial for the treatment and diagnosis of suffering people from lung cancer. This study focuses on developing an accurate segmentation of lung tumors using image processing algorithms. Tumors can be malignant or benign.[9] Malignant tumors spread and grow aggressively into other parts of the body, while benign tumors can be stopped from spreading and removed.

[10]Nodule detection and treatment of lung cancer performance a very vital role in medical imaging. It is an efficient and accurate technique for diagnosis. Differentiating between non-cancerous and cancerous nodules using chest Computed Tomography images has achieved a high level of accuracy. The developed approach is purely based on image processing and soft computing, utilizing different images for lung nodule segmentation and feature extraction.[11]

The segmented system can be mostly separated into three main phases. [12][13]During the first phase, firefly-based optimization algorithms are used for the segmentation of lung nodule images. Secondly, the image is separated so that individual part is divided using hybridization soft computing approaches to select the optimum set of features extracted from the segmented lung nodules, decreasing the processing time of classification. The third stage involves using a SVM to categorize lung cancer images correctly into tumor or non-tumor categories.

2. LITERATURE REVIEW

Buse Ugur et al. [14] proposed image processing methods for segmented lung tumors based on different lung tumor segmentation systems. The lung image database consists of a Computed Tomography image dataset presented in the field of DIP. Noise is removed from images using median filters and erosion , which are individually applied. In Computed Tomography images for applying filters, segmentation is one of the furthest common methods used. Through the segmented technique a exact value of grayscale image turns into a binary image. Then, the binary areas are opened to remove small objects from nodule images. Different filtered images and images with small objects removed are taken to create tumor cells.

Chang-Chiun Huang et al. [15] using an adaptive Wiener filter a proposed preprocessing filter is to optimize images. Lung segmented uses the employs histogram shifting to create images and Fast Otsu technique for rapid segmentation and edge search. To enhance gray level, nodule enhanced adopts a collective method. The Fast Otsu technique is used for candidate detection to obtain primary candidates. FP are reduced using quadratic SVM to attain the nodules.

Elchouemi et al. [16] presented a new model that uses a median filter instead of Gabor and Gaussian filters in the preprocessing stage. Subsequently, watershed segmentation is processed by the preprocessing method. Feature extraction and accuracy detection in cancer nodules are performed. However, classification as malignant or benign has not been executed. Therefore, using a SVM the classification of cancer nodules is completed.

Ke Yan et al. [17] in lung cytological images an automatic method for classifying malignant and benign cells is proposed. The structural design of the DCNN technique is described, a cytological image dataset is collected, and accuracy is grouped.

Mohamed Shakeel et al. [18] to predict lung cancer a new ML and optimized IP method is proposed. Lung cancer dataset images are collected from CT scans for diagnosing NSCLC. The quality of the lung image is increased and the grouped images are observed by multilevel intensity for each pixel. Affecting the segmented region in terms of network layers and extracted features. Noise is removed from Computed Tomography lung images using an improved DNN.

Md Rashidul Hasan et al. [19] proposed a lung tumor cancer detection method within the lung cancer. Using a marker-controlled watershed segmented image the Computed Tomography image is pre-processed, which is used for feature extraction and detection of lung tumors. Both unsupervised and supervised classifiers are used for identifying lung cancer. The accurateness rate of the proposed system is 73.3% using a SVM.

Ozge Gunaydin et al. [20] observed and compared ML techniques for identifying lung cancer nodules. Methods such as Principal Component Analysis, K-NN, ANN, SVM, Naïve Bayes, Decision Trees are applied to detect anomalies.

Ayshath Thabsheera et al. [21] different techniques for lung diagnosis using Computed Tomography images. It was detected that the enhancement of the Gabor filter provides improved performance. Area and eccentricity features are calculated and the classifier to decide with the given as input if lung nodule in the Computed Tomography image is non-cancerous or cancerous.

Hongtao Xie et al. [22] a novel automatic pulmonary nodule detection outline to assist with CT images. Firstly, two region proposal networks are adjusted based on the structure of Faster RCNN, and nodule candidates are detected by a deconvolutional layer. Secondly, an enhancing design based on 2D Faster RCNN is considered for TP reduction, to separate true nodules from the person which acts as a classifier.

Ayshathil Bushra et al. [23] a hybrid combination of RBFNN and BNMB algorithms to detect lung cancer, which was compared to using the BNMB algorithm alone for better results. The system classifies lung CT images as abnormal or normal, identifying PBR regions. The CAD system at an early stage helps radiologists to classify lung cancer due to increasing accuracy.

Gustavo Ramirez et al. [24] proposed an ODNN with feature extraction, verifying better classification compared to other techniques for lung CT images. An lung cancer classification method automatically decreases the manual classification time. The method aims to achieve better accuracy in recognizing abnormal and normal lung cancer images through machine learning techniques.

Junjie Zhang et al. [25] introduced the detection of malignant nodules, which is important for diagnosing early primary lung tumors. Computer aided development schemes are implemented to identify PN and assist radiologists in better diagnosis. Deep learning-based and feature engineering CAD schemes include the types of PN.

Tjeng Wawan Cenggoro et al. [26] proposed a weighted clustering, non-maximum suppression (NMS), loss function, and modified RetinaNet neural network backbone for the detection method. Computed Tomography scan images is based on the LIDC public dataset are most commonly used for lung cancer. Lung nodular malignancies can increase the map rate by 7.3%. Textures is poorly studied in computer aided development by the method proposed for classification and detection of lung nodular. From natural image datasets increases detection performance and decreases training time this method also proves that transferring weight learning.

Syed Omer Gilani et al. [27] analyze nodules at initial stages in CT images and automated detection techniques to identify. Several techniques displays possible improvement but still needs development to overcome many tasks, such as classifying different nodules based on their high sensitivity with a low (FP) rate, PACS, size, shape, and locations for different databases.

Pengyi Hao et al. [28] using a one-stage structure relying on a FPN a diagnosis technique for LA is proposed. The network has two benefits: generate high-resolution mapping features and it can classify and localize LAs simultaneously. In common, feature construction at multiple levels and reuses feature maps for multi-scale objects it can classify and localize LAs simultaneously. It can create high resolution, where it can exactly classify small objects and securely cover.

3. PROPOSED WORK

The proposed model for lung cancer diagnosis implements nodule segmentation through an optimization-based approach. The diagnostic process involves two key tasks: differentiating between cancerous and non-cancerous nodules. The lung nodules are then segmented and their features extracted. NSCLC to help diagnose lung cancer(LC) patients. Two common types of primary lung cancer are NSCLC and SCLC.

During preprocessing, the model optimizes CT images for further analysis. FA is used to perform the Lung nodule(LC) segmentation. This study presents a complete model for lung tumor(LT) segmented and feature extraction, utilizing a combination of image processing and soft computing techniques. Lung nodules are segmented and the results are placed on the original image using optimization.

The segmented lung nodules is achieved, and the outcomes are superimposed on the original image using optimization. The enhancement of nodules is done using a cumulative technique to improve the RGB levels. Initial candidate nodules are identified using the firefly algorithm. The tumor and non-tumor images final classification of including the identification of benign or malignant tumors, is conducted to increase the accuracy using SVM. Figure 1 illustrates the different phases involved in the image processing approach.

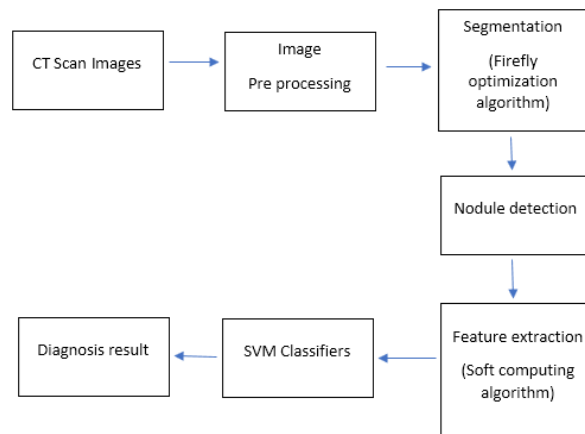


Fig 1: Proposed Model

3.1 IMAGE PRE PROCESSING

In lung cancer detection is the initial and important technique involved in image pre-processing. To remove some regions to identify of CT image such as background and surrounding tissues or vessels and pre-processing technique is needed to improve the detection accuracy and it enhances the reduces the distortion in image and relevant features.

3.2 SEGMENTATION

In image segmentation the FF algorithm which is the process of partitioning a lung nodule into grouping of region on certain similarity criterion. For detecting nodule the comparison can be determined using specific properties for representinh objects in the image. Segmented in other word the FF algorithm allows the finding of similarities in the image.

3.2.1 FIREFLY OPTIMIZATION (FF- ALGORITHM)

A nodules method for firefly optimization algorithm is proposed and contains of 3rd stages. In the first phase, the firefly algorithm examines the brighter region of the given data set to determine the lung nodules images. In the 2nd stage, Attractiveness is related to brightness, the lesser bright nodule region cannot be detected whereas brighter FF can. In the third phase brightness of an FF is influenced or segmented nodule for an objective function. Identifying nodule by using firefly algorithm by data set are encoded by each and every solution of the nodules. The solution will be given in equation,

$$S = A_1\{s_1, s_2, \dots, s_d\}, A_2\{s_1, s_2, \dots, s_d\}, A_3\{s_1, s_2, \dots, s_d\} \quad (1)$$

s_i is a segmentation characteristic that explains nodule centre and $s_i, \in S$, where S is the collection of the data set nodule images for every pixels attribute. Accordingly, each nodules at centre a_i is defined by 'n' nodule image segmentation of feature s_1, s_2, \dots, s_d . As a result, every single solution has an actual size of $(c_c * n_c)$ where ' c_c ' represents cancerous nodule and ' n_c ' indicates the non-cancerous lung images can be identified by RGB channel in firefly algorithm. From the above given equation (1)

$$C = S_i\{s_1, s_2, \dots, s_d\} + \begin{cases} c_c = 1(true) \\ n_c = 0(false) \end{cases} \quad (2)$$

From the above equation (1), the nodule lung image dataset are observed based on the size. In equation (2), for the nodules images. c represents cancer cell in lungs, S_i is the process for segmentation in lung nodule images. $\{s_1, s_2, \dots, s_d\}$ size of the image to be detected in cancer cells. c_c represents cancerous where 1 denotes whether patients have tumor which spread around the cell and n_c denotes non-cancerous where 0 represents patients have cancer which does not spread. In cancerous patients we identify the tumor and non tumor size of the cancer cells by firefly algorithm using RGB channel for better identification. The following stages in equation (3)

$$c_c = \begin{cases} x = x_f \cup x_b \\ x_f(i, j) = R(i, j), G(i, j), B(i, j) > \mu \\ x_b(i, j) = R(i, j), G(i, j), B(i, j) > \mu \end{cases} \quad (3)$$

From the above equation (2), cancerous cell images are detected. In equation (3), As the input lung Computed Tomography image X, according to the image brightness value μ is composed into two different sub-images, the images that are composed named as the foreground image and background image that are represented as, x_b and x_f . (i, j) presented technique efficiently observes each nodules image and its size for identifying the tumor is present in image. RGB are known as color channels. Color image has three channels RGB each channel has 256 different values of intensity ranging 0-255. If RGB have zero value then is called non-cancerous and while RGB channel values is one then called cancerous by firefly algorithm in following equation. (4)

$$FF(x_f \cup x_b) = \begin{cases} F(e) = F_0 f^{-\alpha e^2} \\ \gamma(e) = \gamma_0 f^{-\alpha e^2} \end{cases} \quad (4)$$

From the above equation (3) the two attraction are obtained. Equation (4) where (e) = distance of the inter-firefly, F_0 = light intensity of initial stage and α = light absorption of coefficient which is the value of the firefly fixed in the light intensity variation. The brightness level of a F and γ is usually fixed in FA is directly related to the nodule images. where γ_0 is the brightness at $e = 0$ Then, the attraction of the i and j firefly can be formulated by equation (5)

$$i_x = j_x + \gamma(i_x - j_x) + \beta(rand(i_x - j_x)) \quad (5)$$

From equation (4), intensity of i and j firefly detected by nodule images. From above equation (5), β and γ is the attractiveness of the firefly. $rand$ is the random number generated for choosing lung cancer nodules range [0,1]. i_x and j_x are the firefly determination for identifying the cells. From the following equation (6).

$$x_{i,j} = \sqrt{\sum_{s=1}^S (x_{i,j} - x_{i,s})^2} \quad (6)$$

From the above equation (5) segmentation is addressed as the detection method. From equation (6) the detection of $x_{i,j}$ is the segmentation between i and j can be formulated. S denotes the size of an

optimization problem for an segmentation process by analysing set of features to be used in the diagonalising rules. $x_{i,s}$ is applied for an extraction of CT scan image to detect the lung caner region. By using equation (2) to detect the cancer nodules from the lung images segmentation algorithm is applied . For lung nodule detection equation (7)

$$N_d(C_c, N_c) = \begin{cases} c_c = 1(true) \\ n_c = 0(false) \end{cases} \sum_0^1 C_c + N_c - \sum_0^1 C_c + N_c \quad (7)$$

From the equation (6), the nodules for lung cancer images for an optimization algorithm. From the above equation (7), N_d denotes th nodule detection. C_c and N_c are cancerous and non cancerous Along with their locations in patients cells to detect changes lung nodules pattern. To separate true nodules as one $c_c = 1(true)$ from false nodules as zero $n_c = 0(false)$, in order to have a high accuracy but the effort is to detect as many true lung nodules. true cancerous nodules that are perfectly identified and false non cancerous nodules that are incorrectly rejected stand for where $C_c + N_c$. They are different nodules identified in equation (8)

$$N = \begin{cases} S_n(C_c) + (S_n + 1) = S_n(s) \cup P(s), \text{ for } \alpha(S_n) + \mu_0, \mu_1 \\ S_t(C_c) + (S_t + 1) = S_t(s) \cup P(s), \text{ for } \alpha(S_t) + \mu_0, \mu_1 \\ N_v(C_c) + (N_v + 1) = N_v(s) \cup P(s), \text{ for } \alpha(s) + \mu_0, \mu_1 \end{cases} \quad (8)$$

From the equation (7), for applying optimization algorithm different segmented nodules of lung images are present. From above equation (8) , N is the nodules present. μ_0, μ_1 are the color mean of class. s is a neighbour size of the lung image and contains $P(s)$ more than one at a time. α is the color difference. Different lung nodule detection is to analyse different kinds of CT images and segmented of solitary nodules S_n , while identifying the dark region has an advantage over the others on segmented lung nodules follow to N_v vessels. semi-transparent nodules S_t for segmentation easier to analyse.

3.3 FEATURE EXTRACTION

3.3.1 HYBRIDIZATION SOFT COMPUTING

HSC is known for its effective feature extraction which syndicates the advantage of the Firefly algorithm. The FA-based feature selection can divide the features into an automatic in terms of attractiveness through the light intensity of variation. To select the optimum set of features extracted from the segmented lung nodules. In HSC, the features of lung cancer dataset are encoded in the form of classification in order to determine the cancerous nodules . After the segmentation lung nodule is used for feature extraction.

$$H_{sc}(F') = \beta \cdot \frac{F'}{A} + (1 - \beta) \cdot P \quad (9)$$

From the above equation (9), H_{sc} is the objective function. F' denotes the feature set for an lung images. A denotes the feature boundary classifier and P represents classification of the processing time due to accuracy. β dimension of the features. From equation (8) we classify the extracted nodules for the lung images. By identifying the segmented features in the search space for brighter images and for the entire RGB image. For equation (10)

$$F'(n) = \beta \cdot \frac{F'}{A} + \begin{cases} S_n(C_c) \\ S_t(C_c) + [S_{n1,i,d}, S_{t2,i,d}, \dots, S_{tg,i,d}] \\ N_v(C_c) \end{cases} \quad (10)$$

Form the equation (9), the following features to be optimised. From the above equation (10), $F'(n)$ addressed as the selecting features of the nodules. Each nodules generation $i= 1,2$ is number of

dimensional features. d and g denotes the number of fetures calculated. The features like statistical based intensity and geometric features are extracted. Shape of cancerus cell are characterized by the appearance of an CT image. Some of the features were considered to be extracted by the following equation (11)

$$A * P = n[1] + \left[x_n x_1 + \sum_{i=1}^{n-1} [x_i x_{i+1}] \right] \quad (11)$$

From the equation (10), different nodules is introduced for a feature set of lung images. From above equation (11), $A * P$ denotes the area and perimeter of the feature boundary of the lung nodule images. $n[1]$ represents the couting feature set. $x = [x_1, \dots, x_n]$ is the set of consecutive features identified for a cancerous images. For an texture and the intensity of the CT image color constant image are calculated by equation (12).

$$P(x, y) = \sum \sum p(x, y) \log p(x, y) \sum \sum (x - y)^2 p(x, y) \quad (12)$$

From the equation (11), the intensity and texture images are obtained. From the above equation (12), $P(x, y)$ is the feature location The numerical measure of chance can be used to characterize the image texture. Where p is the number of RGB matrices in contrast. Measures the local dissimilarities in the color channel. It calculates the contrast between a feature and its segment for the whole image. Contrast is 1 for a RGB correlation image. For equation (13)

$$\sigma_x * \sigma_y = \sum_{x=0}^{c-1} \sum_{y=0}^{c-1} \frac{(x - \omega_x)(y - \omega_y)p(x, y)}{\sigma_x \sigma_y} \quad (13)$$

From the equation (12), lung images are characterized for CT scan for texture image nodules. From the above equation (13), σ_x and σ_y specifies the color gradient of image nodules. c represents the co-occurrence of the specified image. ω_x and ω_y are the weighted function for an nodule images. p is defined as the probability of the image brightnes. To compare the RGB channel for maximum contrast to identify the lung nodule images. The following equation in (14)

$$\alpha(\beta) = \log_2 \sum_{N=0}^{n-1} (F(T)) + (1 - \alpha + F(T)) \left\{ \begin{array}{l} \sum_{x,y} p(x, y)^2 \\ \sum_{x,y} \frac{p(x, y)}{1 + [x - y]} \end{array} \right. \quad (14)$$

From the equation (13), processing time for the segmented classification method for the above satisfied equation (14), $\alpha(\beta)$ is the regression process for segemented time. Log is calculated for the feature is 1 for the segmented cancerous nodules will reduce the processing time of classification T . $p(x, y)^2$ provides the elements to be squared for the color gradient and α distribution of the closest measure is diagonized. From the segmented image features are extracted and is used as the input to classify the lung cancer nodule images for feature extraction.

3.4 CLASSIFICATION

3.4.1 SVM CLASSIFIER

Support vector machine (SVM) is used as a classification method. In lung cancer images in order to correctly classify tumour or non-tumour images. After the nodules are segmented and detected, the next method is to classify the nodule as malignant or benign. This method is called supervised machine learning. SVM data are classified into two classes :

$$T(x) = \begin{cases} B^T i_x + b + 1 \text{ for } j_y = 0, \text{ normal} \\ M^T i_x + m + 1 \text{ for } j_y = 0, \text{ severe} \end{cases} \quad (15)$$

From the above equation (15), $B^T i_x + b$, i_x is the training inputs. j_y is for condition to be normal or severe. $T(x)$ for tumor classification. B^T denotes as benign tumor, which is removed and stopped spreading to other parts of the body. M^T denotes as malignant grows and spread to other parts of the body. The extracted features from a CT scan, lung cancer images and using the SVM classifier is used to reduce dimension of the feature to recognize lung nodules cancer images as malignant or benign for the following equation (16).

$$M_T + B_T = \sum_{y=1}^a \sum_{x=1}^{C_s} (m_x^y - \beta_y)(m_x^y - \beta_y)^T + (b_x^y - \beta_y)(b_x^y - \beta_y)^T \quad (16)$$

From equation (15), tumor-based features were extracted and SVM classification of malignant and benign nodules. From above equation (16), $M_T + B_T$ represents the malignant and benign method. a denotes the class of tumor images and m_x^y, b_x^y, C, β_y are testing class for nodules images. By observing this equation it seems that malignant is having the severe growth and spread to other parts of the cells. So, in this proposed work malignant is addressed as tumor cell shown in equation (17).

$$M^T = \left[\frac{\prod_{x=1}^N (m_x - m) \prod_{x=1}^N ((m_x - m)^T)}{\rho} \right] + \text{for } M^T > 0.6 \quad (17)$$

From the equation (16), similarity to malignant nodules is high to be obtained. From the above equation (17), m denotes the classes of the malignant. $[x.....n]$ is the classifiers of the malignant to be identified for the lung cancer images. SVM classifier is applied for an direct differentiate hypothesis. We have identified the nodule with a value greater than 0.6 as a malignant nodule, otherwise it is observed as a benign nodule. Proposed model classifies as malignant and benign with better accuracy and used to detect lung CT scan image for the cancerous nodule using firefly algorithm segmented for detection and Support Vector Machine for classified of nodule as benign or Malignant. Classifying the features include shape, appearance and size of the nodule for the equation (18).

$$C(m, b) = \frac{\left(\sum_{x=1}^x \sum_{y=1}^y C_{xy} m_x b_y + \sum_{x=1}^X \beta_x m_x + \sum_{y=1}^Y \alpha_y b_y > 1 \right)}{\left(\sum_{x=1}^x \sum_{y=1}^y C_{xy} m_x b_y + \sum_{x=1}^X \beta_x m_x + \sum_{y=1}^Y \alpha_y b_y < 1 \right)} \quad (18)$$

For equation (17), we classify the tumor which is better for differentiate. From the above equation (18), C_{xy} represents the tumor interaction for symmetric appearance for visible and non visible. $m_x b_y$ is the hidden size of the cell to detect the tumor. β_x and α_y are the partiality term. x and y number of visible cell. $C(m, b)$ are the classification for SVM method. Features include shape, appearance and size and growth of the nodule. The size of the tumor is greater than 1 are malignant and less than 1 are expected to be benign. The benign tumor shape is circular, however for malignant tumor is halo. Benign tumor would be even but malignant have uneven size. Respectively the growth rate of the malignant tumor is faster rather than benign tumor. Firefly algorithms for the segmented of lung nodules images is detected by using equation (8) and (7) which results in the improvement of color image by extracted feature is shown in equation (13). Thus, this work classifies the deep learning for detected the SVM classifiers from the lung cancer nodules by using equation (18) and (17). Using this equation the lung cancer image is obtained from 8x8 size image. Classification for tumor detection is evaluated for finding classification of SVM. The tumor of the lung nodules on shape, appearance and size are classified as malignant tumor by the SVM classifier. These classified methods are identified as tumor or non tumor for nodules lung cancer images.

3.5 ACCURACY ASSESSMENT

The accuracy of the classification was calculated using accuracy, specificity and sensitivity. The accuracy of the segmented tumors was evaluated based on the size of the lung nodule images. TP (true positive) signifies lung images segmented correctly and have a tumor (C_c). FN (false negative) signifies lung images segmented but do not have a tumor (N_c). TN (true negative) represents lung images that do not have a tumor and were segmented correctly (T_s). FP (false positive) represents lung images were not segmented correctly but have tumors (S_T).

$$Accuracy = \frac{C_c + N_c}{C_c + T_s + S_T + N_c} \tag{19}$$

$$Sensitivity = \left[\frac{C_c}{C_c} + N_c \right] * 1 \tag{20}$$

$$Specificity = \left[\frac{N_c}{N_c} + S_T \right] * 1 \tag{21}$$

4. RESULTS & DISCUSSIONS

Our proposed work is to utilize image processing and soft computing approaches in order to rise the accuracy of diagnosing cancerous cells for the optimal classification of tumour and non-tumour images. Segmentation of lung nodules images by Firefly optimization-based algorithms SVM method for lung cancer images in order to correctly classify tumour or non-tumour images. The evaluated performance method by on-line CT scan lung images from the database[14]. SVM is evaluated using the sensitivity, specificity of the classification.

Table 1: Proposed firefly optimization lung image for nodules images

Images (CT Scan)	Accuracy			Sensitivity			Specificity		
	FF	CS	PSO	FF	CS	PSO	FF	CS	PSO
ID_0005_AGE_0 048	0.1054 1	0.0270 7	0.0988 6	97.231	90.129 4	89.484 6	2.522 4	0.660 5	2.370 1
ID_0004_AGE_0 056	0.2022 1	0.0319 6	0.1999	96.835 1	80.024	79.782 8	4.858	0.765	4.791 5
ID_0001_AGE_0 069	0.2068 4	0.0381 8	0.1799 9	96.111 9	82.003 1	79.353 2	1.655 3	0.318	1.444 1
ID_0006_AGE_0 075	0.2447	0.0448 9	0.2297 4	95.500 1	77.038 6	75.597 2	1.956 4	0.363	1.839 8
ID_0058_AGE_0 082	0.2765 4	0.0498	0.2577 2	95.095 2	74.231 9	72.355 4	2.219 3	0.390 3	2.060 3

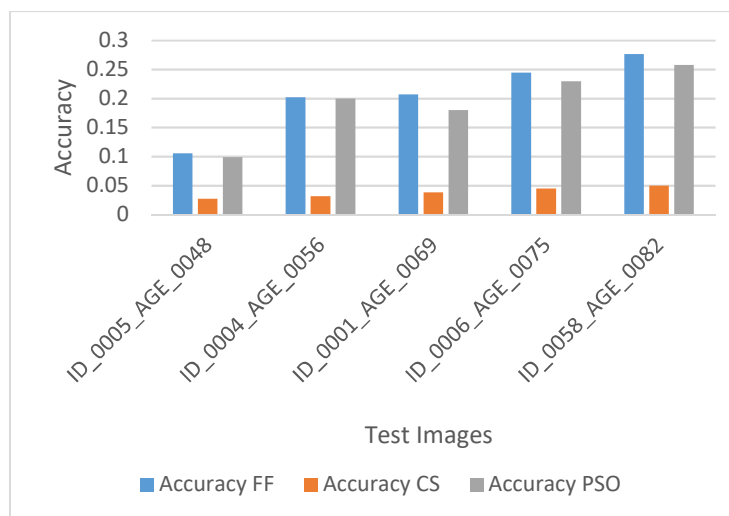


Fig 1: Performance measurements of Accuracy

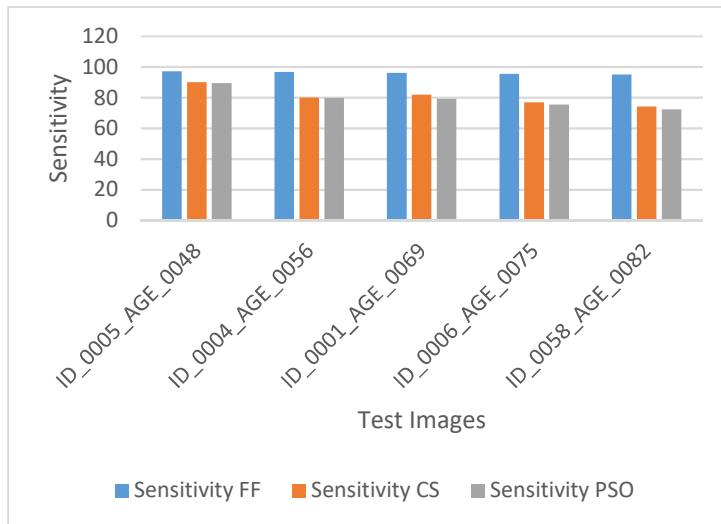


Fig 2: Performance measurements of Sensitivity

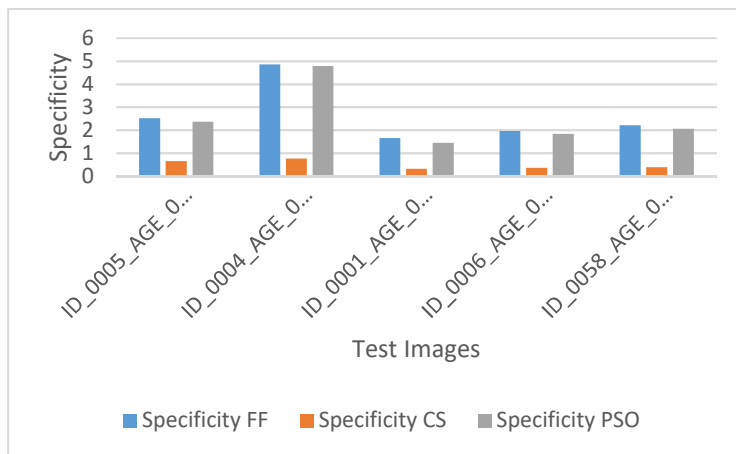


Fig 3: Performance measurements of Specificity

Table 2: Comparison of processing time of codebook size of 16X16

Approaches	Processing Time	Type	Accuracy
Soft Computing	30 to 50 sec	Existing	86.4
hybridization soft computing	04 to 15 sec	Proposed	88.7

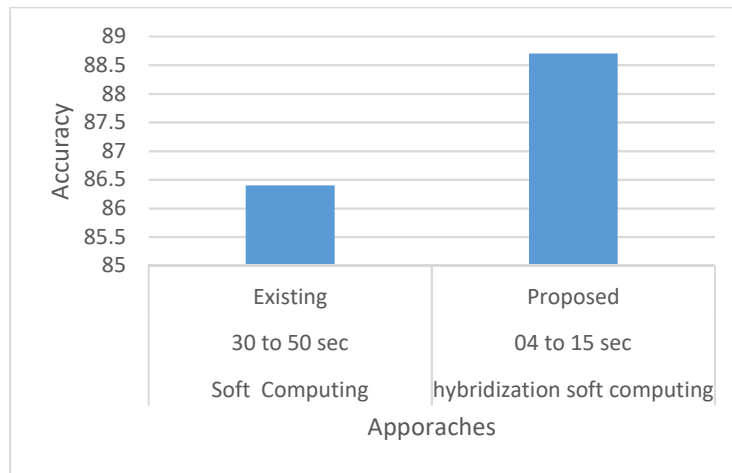


Fig 4: Performance measurements of HSC

Table 3: Proposed CT lung image for SVM classification

Type	Accuracy	Sensitivity	Specificity
Malignant	94.22%	93.52%	91.3%
Benign	85.22%	86.52%	82.5%

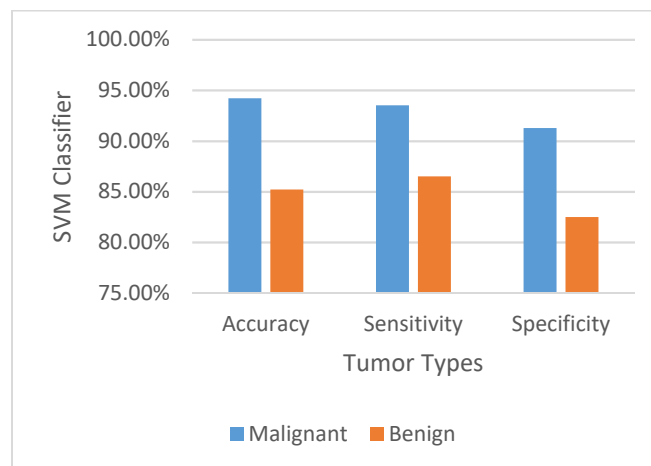


Fig 4: Analysis of SVM Classifier

Based on Table 1: Our proposed result shows that the accuracy of algorithms which is optimization-based for the segmentation of lung nodule images is higher for CT test images. Therefore, this algorithm provides more accuracy in identifying tumor cells.

From Table 2: The approach selects from the segmented lung nodules of the the optimal set of extracted features, reduces the processing time of classification and improves the accuracy of nodule detection.

From Table 3: Our proposed work aims to build a model that classifies lung cancer images using DL method to appropriately identify non-tumor and tumor images. The roposed approach is performed by evaluating using lung cancer images. The proposed optimization and classification methods for lung cancer nodules in computed tomography images helps to detect tumor and non-tumor cells. The false positive rate and true positive rate are measured as accuracy, sensitivity, and specificity.

Nodules for lung cancer detection are denoted as 0 for non-cancerous and 1 for cancerous detection using the FF algorithm, where $N_d(C_c, N_c)$. Comparative analysis for the SVM classifier, where $M_T + B_T$ represents tumor detection, shows the accuracy of the SVM classifier. The $H_{sc}(F')$ approach is used for classification to achieve maximum accuracy from the segmented lung nodules for HSC by

extracting the optimal set of features, thereby decreasing the processing time of classification. The color gradient for RGB, denoted as $\sigma_x * \sigma_y$, enhances the true positive and false positive rates.

The proposed method efficiency is measured using the performance metrics such as accuracy, sensitivity, and specificity across different sets of CT images. A analysis has been compared its showed between the existing optimization, proposed model and classification techniques. The analysis shows that the FireFly algorithm achieves optimal results for the segmented lung nodule images. As illustrated in the figure 4, the accuracy of the SVM classifier has been compared with various tumor types for lung cancer detection. By integrating a hybrid soft computing approach, the accuracy of cancerous cell diagnosis is improved, enabling more optimal classification between non- tumor and tumor images.

5. CONCLUSION

The main aim of proposed methodology is to determine the presence of tumors or non-tumorous regions in Computed Tomography images, focusing on diagnosis of lung cancer nodules. In this proposed system, images are assessed to detect the lung nodules with an support vector machine classifier optimized for tumor cell identification. Additionally, to enhance the lung image detection firefly optimization is used. This will integrate feature detection for analysing nodule images. The challenge lies in identifying the relevant regions in Computed Tomography scans accurately using segmentation techniques. The three main phases are used in lung cancer detection are segmentation, feature extraction and classification. The proposed system improved the accuracy in classifying lung nodules into Small Cell Lung Cancer(SCLC) and Non-Small Cell Lung Cancer(NSCLC) and also distinguishes the cancerous and non-cancerous cases. The system's performance enhanced the precision of detecting the lung cancer nodules by analyzing lung cancer presence in CT images.

REFERENCES

1. Disha Sharma, Gagandeep Jindal" Identifying Lung Cancer Using Image Processing Techniques" International Conference on Computational Techniques and Artificial Intelligence (ICCTAI'2011).
2. Saadya Fahad Jabbar1" A classification model on tumor cancer disease based mutual information and firefly algorithm" Periodicals of Engineering and Natural Sciences ISSN 2303-4521 Vol. 7, No. 3, September 2019, pp.1152-1162
3. Amudha K *, Balu S 2, Sakthivel K 2" Performance Analysis Of Firefly Search Fuzzy C-Means (Fsfcm) For Detecting Lung Cancer Nodules" International Research Journal Of Pharmacy www.irjponline.com ISSN 2230 – 8407
4. Jaspinder Kaur, Nidhi Garg, Daljeet Kaur" A survey of Lung Cancer Detection Techniques on CT scan Images" International Journal of Scientific & Engineering Research, Volume 5, Issue 6, June-2014 ISSN 2229-5518
5. Sakshi Sharma, Maninder Pal Singh, Baljeet Kaur Nagra" CDCT: CT Scan Images based on Mechanism for Lung Cancer Detection" International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8, Issue-2S6, July 2019
6. Retz Mahima Devarapalli1, Hemantha Kumar Kalluri, Venkatesulu Dondeti" Lung Cancer Detection of CT Lung Images" International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-7, Issue-5S4, February 2019
7. Khin Mya Mya Tun, Aung Soe Khaing" Feature Extraction and Classification of Lung Cancer Nodule using Image Processing" International Journal of Engineering Research & Technology (IJERT) ISSN: 2278-0181 Vol. 3 Issue 3, March – 2014
8. VR. Nagarajan, Dr. D. Vimal Kumar" Hybridized Firefly And Differential Evolution Optimization Algorithm Based Feature Selection For Disease Prediction Techniques" International Journal Of Scientific & Technology Research Volume 8, Issue 10, October 2019 Issn 2277-8616
9. 1Ayushi Shukla, 2Chinmay Parab, 3Pratik Patil, 4Prof. Savita Sangam" Lung Cancer Detection using Image Processing Techniques"

Volume: 05 Issue: 04 | Apr-2018 www.irjet.net.

10. A. Asuntha¹, A.Brindha¹, S.Indirani¹, Andy Srinivasan²” Lung cancer detection using SVM algorithm and optimization techniques” www.jchps.com, ISSN: 0974-2115
11. Li Zheng¹ and Yiran Lei¹” A Review of Image Segmentation Methods for Lung Nodule Detection Based on Computed Tomography Images” MATEC Web of Conferences232,02001(2018)<https://doi.org/10.1051/mateconf/201823202001EITCE> 2018
12. G.Niranjana, Dr.M.Ponnaivaikko” A Review on Image Processing Methods in Detecting Lung Cancer using CT Images” 978-1-5090-4797-0/17 \$31.00 © 2017 IEEE DOI 10.1109/ICTACC.2017.16
13. A. Asuntha*, Neelima Singh** and Dr.Andy Srinivasan” PSO, Genetic Optimization and SVM Algorithm used for Lung Cancer Detection” Journal of Chemical and Pharmaceutical Research, 2016, 8(6):351-359 ISSN : 0975-7384
14. Selin Uzelaltinbulata, Buse Ugurb* “ Lung tumor segmentation algorithm” 1877-0509 © 2018 The Authors. Published by Elsevier B.V. Peer-review under responsibility of the scientific committee of the 9th International Conference on Theory and application of Soft Computing, Computing with Words and Perception. 10.1016/j.procs.2017.11.221.
15. Chung-Feng Jeffrey Kuoa, Chang-Chiun Huang^a” Automatic lung nodule detection system using image processingtechniques in computed tomography” <https://doi.org/10.1016/j.bspc.2019.101659> 1746-8094/© 2019 Elsevier
16. Abeer Alsadoona , A. Elchouemic” Lung Cancer Detection using CT Scan Images” Procedia Computer Science 125 (2018) 107–114 2.
17. Ke Yanc, Ling Zhangd, Kazuyoshi Imaizumib” Automated classification of benign and malignant cells from lung cytological images using deep convolutional neural network” Informatics in Medicine Unlocked 16 (2019) 100205.
18. P. Mohamed Shakeel¹, M. A. Burhanuddin², Mohammad Ishak Desa¹” Automatic lung cancer detection from CT image using improved deep neural network and ensembleclassifier”NeuralComputingandApplications <https://doi.org/10.1007/s00521-020-04842-6>, springer(0123456789()),-volV)
19. Md Rashidul Hasan, Muntasir Al Kabir “Lung Cancer Detection and Classification basedonImageProcessingandStatisticalLearning”<https://www.researchgate.net/publication/328642003>
20. Özge GÜNAYDIN, Melike GÜNAY” Comparison of Lung Cancer Detection Algorithms” 978-1-7281-1013-4/19/\$31.00 ©2019 IEEE.
21. A. P. Ayshath Thabsheera , T. M. Thasleema and R. Rajesh” Lung Cancer Detection Using CT Scan Images: A Review on Various ImageProcessing Techniques” © Springer Nature Singapore Pte Ltd. 2019 https://doi.org/10.1007/978-981-13-2514-4_34
22. Hongtao Xie a , Dongbao Yang a , b , c , *,” Automated pulmonary nodule detection in CT images using deep convolutional neural networks” Pattern Recognition 85 (2019) 109–119.
23. K. AyshathilBushra^a, SunithaLasradob, K.S. Prasada” Detection of Lung Cancer by Modified Irregular Tree Structure Bayesian Network Model Based Image Segmentation” Proceedings 11 (2019) 1130–1138
24. Lakshmanaprabu S.K., Sachi Nandan Mohanty, Shankar K.,Arunkumar N., Gustavo Ramirez” Optimal deep learning model for classification of lung cancer on CT Images”FutureGenerationComputerSystems(2018),<https://doi.org/10.1016/j.future.2018.10.009>
25. Junjie Zhanga, Yong Xiaa,b,*” Pulmonary nodule detection in medical images: A survey” Biomedical Signal Processing and Control 43 (2018) 138–147.
26. William Harsono, I., Liawatimena, S., Wawan Cenggoro, T., Lung Nodule Detectionand Classification from Thorax CT-Scan Using RetinaNet with Transfer Learning, Journal of King Saud University - Computer and Information Sciences (2020), doi: <https://doi.org/10.1016/j.jksuci.2020.03.013>

27. Muhammad Zia ur Rehmana,b,*, Muzzamil Javaida, Syed Irtiza Ali Shaha,Syed Omer Gilania, Mohsin Jamila “An appraisal of nodules detection techniques for lung cancer in CTimages” Biomedical Signal Processing and Control 41 (2018) 140–151
28. Pengyi Hao a , Kun You” Lung adenocarcinoma diagnosis in one stage” <https://doi.org/10.1016/j.neucom.2018.11.110> Neurocomputing