Future Generation Computer Systems 92 (2019) 374–382

Contents lists available at ScienceDirect

Future Generation Computer Systems

journal homepage: www.elsevier.com/locate/fgcs

Optimal deep learning model for classification of lung cancer on CT images

Lakshmanaprabu S.K. [a,*], Sachi Nandan Mohanty [b], Shankar K. [c], Arunkumar N. [d], Gustavo Ramirez [e]

[a] Department of Electronics and Instrumentation Engineering, B.S.Abdur Rahman Crescent Institute of science and Technology, Chennai, India
[b] Department of Computer Science & Engineering, Gandhi Institute for Technology, Bhubaneswar, India
[c] School of Computing, Kalasalingam Academy of Research and Education, Krishnankoil, India
[d] Department of Electronics and Instrumentation Engineering, SASTRA University, Tanjavur, India
[e] Department of Telematics, University of Cauca, Colombia

HIGHLIGHTS

- An innovative approach is proposed for the automated diagnosis of lung cancer in CT images.
- Modified Gravitational Search Algorithm is applied to train the Optimal Deep Neural Network.
- The proposed classifier provides the sensitivity of 95.26%, specificity of 96.2% and accuracy of 96.2%.

ARTICLE INFO

Article history:
Received 25 July 2018
Received in revised form 16 September 2018
Accepted 4 October 2018
Available online 16 October 2018

Keywords:
Image processing
Computed tomography
Lung cancer
LDA
Optimization
Classification

ABSTRACT

Lung cancer is one of the dangerous diseases that cause huge cancer death worldwide. Early detection of lung cancer is the only possible way to improve a patient's chance for survival. A Computed Tomography (CT) scan used to find the position of tumor and identify the level of cancer in the body. The current study presents an innovative automated diagnosis classification method for Computed Tomography (CT) images of lungs. In this paper, the CT scan of lung images was analyzed with the assistance of Optimal Deep Neural Network (ODNN) and Linear Discriminate Analysis (LDA). The deep features extracted from a CT lung images and then dimensionality of feature is reduced using LDR to classify lung nodules as either malignant or benign. The ODNN is applied to CT images and then, optimized using Modified Gravitational Search Algorithm (MGSA) for identify the lung cancer classification The comparative results show that the proposed classifier gives the sensitivity of 96.2%, specificity of 94.2% and accuracy of 94.56%.

1. Introduction

Medical image analysis has extraordinary supremacy in the field of health sector, particularly in noninvasive treatment and clinical examination [1]. The acquired restorative images such as X-rays, CT, MRI, and ultrasound imaging are used for specific diagnosis [2]. In medical imaging, CT is one of the filtering mechanism which use attractive fields to capture images in films [3]. Lung cancer is one-of-its-kind of cancer that leads to 1.61 million deaths per year. In Indonesia, lung cancer is ranked in the third position among the prevalent cancers, for the most part, found in the MIoT centers [4]. The survival rate is higher if the cancer is diagnosed at the beginning stages. The early discovery of lung cancer is not a simple assignment. Around 80% of the patients are diagnosed effectively only at the center or propelled phase of cancer [5]. Lung cancer is positioned second among males and tenth among females [2] globally. The information given in these studies is a general portrayal of lung cancer location framework that contains four basic stages. The lung cancer is the third most frequent cancer in women, after breast and colorectal cancers [6,7]. Feature extraction process is one of the simplest and efficient dimensionality reduction techniques in image processing [8,9]. One of the striking features of CT imaging is its non-obtrusive character. The rise of

* Corresponding author.
E-mail addresses: prabusk.leo@gmail.com (Lakshmanaprabu S.K.), Sachinandan09@gmail.com (S.N. Mohanty), shankarcrypto@gmail.com (Shankar K.), arun.nura@gmail.com (Arunkumar N.), gramirez@unicauca.edu.co (G. Ramirez).

https://doi.org/10.1016/j.future.2018.10.009
0167-739X/

Nomenclature

CT	Computed Tomography
LDA	Linear Discriminate Analysis
ODNN	Optimal Deep Neural Network
GSA	Gravitational Search Algorithm
MGSA	Modified Gravitational Search Algorithm
KNN	K-Nearest Neighbor
ANN	Artificial Neural Network
SVM	Support Vector Machine
CAD	Computer-Aided Diagnosis
DNN	Deep Neural Network
DCNN	Deep Convolutional Neural Network
CNN	Convolutional Neural Network
GLCM	Gray Level Co-occurrence Matrix
ELDA	LDA based on Euclidean Distance
ROIs	Regions of Interest
ELDA	Regularized Linear Discriminate Analysis
PCA	Principal Components Analysis
DWT	Discrete Wavelet Transform
EC	Evolutionary Computation
DBN	Deep Belief Network
RBM	Boltzmann Machine
NN	Neural Network
PPV	Positive Predictive Value
NPV	Negative Predictive values

angles, which might be viewed, is odd when compared to parallel imaging modalities [10].

The selected or extracted features set will extract the relevant information from the input data to the reduction process [11]. The reduced features are assigned to a support vector machine for the purpose of training and testing. The models used for lung cancer image classification are neural network models with binarization image pre-processing [12]. The existing research work for lung cancer classification was performed using a neural network model which provided 80% accuracy [13]. Various investigations have been conducted regarding lung cancer classification and Classifiers, for example, 'SVM, KNN and ANN' [14]. The SVM is a universal useful learning method based on statistical learning hypothesis [15]. However, these techniques are expensive and detect lung cancer at its advanced stages due to which the chance for survival is very low. The early detection of cancer can be helpful in curing the disease completely. So, the requirement of developing a technique to detect the occurrence of cancerous nodule in the early stage is increasing [16].

The contribution of the current work considers two important phases: First phase is the CT lung cancer classification processes where the selected features are extracted to LDA reduction process and in the second phase, optimal deep learning classifier with MGSA optimization algorithm is used to classify the CT lung cancer images. The proposed method outperformed over other methods and also it is shown that the performance improvement is statistically significant. In the rest of this paper, Section 2 discuss about literature study where Section 3 depicted the current issues of the classifier. Further, the 4thsection extravagantly contemplated the proposed philosophy. At that point, Section 5 contains the execution and investigation of this work followed by the conclusion with recommendations for future work.

1.1. Causes and detection of lung cancer

- The general visualization of lung cancer is poor since specialists are unable to discover the infected region until the point when it reaches propelled stage. Five-year survival is around 54% for beginning time lung cancer which is restricted to the lungs, yet just around 4% in the advanced stage of inoperable lung cancer.
- The danger of lung cancer increments with the number of cigarettes smoked after some time; specialists allude to this hazard as far as pack-long periods of smoking history. A little segment of lung cancers occurs in individuals with no known hazard factors for the illness. A portion of these may very well be arbitrary occasions that there may not be an outside reason.
- To examine lung cancer, patients normally undergo X-ray or CT or MRI scans to distinguish anomalous developments in lungs. In any case, exceptionally sensitive CT can identify little knobs that could conceivably be cancerous.
- Early detection of lung cancer: The earlier identification of lung cancer can have greater treatment alternatives and a far more possibility of survival. Be that as it may, just 16% of the individuals are diagnosed in the beginning stage when the sickness is generally treatable.

2. Literature review

In 2018, Yutong Xie et al. [17] recommended an algorithm for lung nodule classification that circuits the Texture, Shape and Deep model-learned data (Fuse-TSD) at the choice level. This algorithm utilizes a GLCM-based surface descriptor, a Fourier-shape descriptor to portray the heterogeneity of nodules and a DCNN to train the features of nodes.

Hiba Chougrad et al. [18] investigated a CAD framework based on CNN to classify the breast cancer. Deep learning generally requires expansive datasets to prepare systems while transfer learning method uses a little datasets of medical images. The CNNs optimally trained with the help of transfer learning method. The CNN accomplished the best outcomes in terms of accuracy i.e., 98.94%. Heba Mohsen et al. [19] demonstrated the DNN classifier for brain tumor classification where the DNN is combined with wavelet transform and principal component analysis.

In 2015, Alok Sharma et al. [20] proposed a method of regularized linear discriminant analysis, in which the regularization parameter computed traditional cross-validation algorithm. In order to investigate the medical data for prediction of disease needs a proper set of features. There have been many evolutionary algorithm has been applied to obtain the optimal selection of features. Recently, gravitational search algorithm and Elephant Herd optimizations are utilized for the selection of optimal features [21,22].

Kuruvilla, J. and Gunavathi, K (2014) developed a ANN based cancer classification for CT images. The statistical used for the classification model developed. The paper claimed that feed forward back propagation network provide better accuracy compared to feed forward networks. Also, the skewness feature has more significance in enhancement of classifier accuracy [23].

In the study conducted by Hao Wang et al. during 2016, proposed a summed up LDA method based on Euclidean norm called ELDA method to defeat the existing disadvantages in the conventional LDA procedure. Multi-class SVM is connected to execute step classification. The trial results exhibited that this algorithm accomplishes better results similarly with high accuracy and viability than any other gait recognition procedures in the model [24].

3. Existing problems of classifiers

- In existing techniques, the lung images were captured and subjected to segmentation specifically after which the SVM classifier was applied and then the accuracies were measured [23].
- The current framework had a limitation since it could not predict the sort and shape or size of the tumor and it dealt with a number of pixels which is not valuable for the earlier detection of cancer [17]. At the point, when ANN produces a testing solution, it does not provide some insights as to why and how. This reduces confide in the network.
- Neural networks are 'black box' and have been restricted in their ability to expressively recognize the conceivable causal relationships. NN particularly has profound networks with many hidden layers and are capable of modeling complex structures. However, the training algorithm is again more complex and dynamically sensitive, which can cause a few issues [17,18].
- It is estimated that by utilizing this model, various existing data mining and image processing strategies could be made to work on together in multiple ways. The main disadvantage of the LDA technique is that it only distinguishes the images containing anomalies [25].
- The GSA drawback is the metropolis standard for contrasting the places of moving particles along with controlling the molecule to move and conquer their arbitrariness.

4. Methodology

The proposed approach used to classify the CT images of the human lung which has a few stages such as preprocessing, feature extraction, reduction and finally the classification. Initially, the CT images were considered to improve the quality of images followed by the feature extraction procedure to extract the features (histogram, Texture, and wavelet) of the images based on strategies. After the feature extraction, dimensionality reduction technique considered reduce the feature for the classification process, the purpose of dimension reduction is reducing the computational time and cost of our classifying Method, feature reduction is utilized that is LDA. The LDA based feature reduction technique is applied in the proposed classifying method for reducing the computational time and cost. The maximum features utilized for classification increase the computation time as well as the storage memory. During classification phase, the CT lung images are classified as normal, benign and malignant based on the extracted features. Generally, the classification issue has two phases such as training and testing phases; the classifier is trained with the chosen features of the training data. On the other hand, during the testing phase, the outcomes of the classification procedure signify whether the images contain the lung cancer regions or the non-cancer regions. The current study utilized ODNN classifier and MGSA optimization is used for the optimized structure. This approach is illustrated in Fig. 1 with outstanding straightforwardness and minimal effort in both trainings as well as in testing process in the classification of CT lung image.

4.1. Filtering and contrast enhancement phase

The collected medical database images were adulterated with some kinds of noise. This filter, if the image is noisy and target pixels neighboring pixel worth is somewhere around 0's and 255's by then supplanting pixel respect with the middle esteem. After removing noise of databases, it is considered to contrast enhancement process as adaptive histogram equalization.

$$Contrast(i,j) = rank * \max\_intensity\,(i,j) \;\therefore\; Initially\ rank\ = 0$$

$$rank = 0 + 1 \tag{1}$$

The histogram in the primary position of each line is acquired utilizing the principal position of the last row, by subtracting the trailing column including the new leading row. The complexity of the CT images is to be increased and set with the limit, so that it consequently recognizes the gray level of the image and adaptively modify the dispersing two neighboring gray levels in the new histogram.

4.2. Feature extraction

The reason of feature extraction technique is to represent the image in its compact and unique form of single values or matrix vector. Feature extraction computes dimensionality reduction in image processing based on which image can be used for classification. It includes diminishing the input data into a reduced representative set of features. The features are utilized as contributions to classifiers that assign them to the class what they represented. The aim of feature extraction is to reduce the data by estimating positive properties. In the current study, histogram features, texture feature and wavelet features whereas all these features are extracted from different bands of CT images.

4.2.1. Histogram features

In histogram features, the image is represented in terms of pixels. The histogram demonstrates the number of pixels in an image at each power value. Transforming the power values of the histogram of the images approximately matches a predetermined histogram. From the input image, a total range of gray levels is evaluated by the histogram method. Here, there are 256 gray levels which ranges from 0 to 255. It has some common features, for example, Variance, mean, skewness and kurtosis, standard.

Variance: The variance gives the number of gray level fluctuations from the mean gray level value. The statistical distributions, like the variance in the length of lines of a particular limit, it could be utilized to distinguish low profile contrasts in the texture.

Mean: The mean gives the average gray level of each region and it is helpful only as a harsh idea of power not by any stretch of the image texture.

Standard Deviation: The definition of Standard Deviation is the square root of the variance denoting image contrast. The image contrast level is evaluated by high and low variance values. This denotes that a high contrast image has high variance whereas a low contrast image has low variance.

Skewness: The image skewness is calculated based on the tail of the histogram. The tail of the histogram value is categorized into two sets, positive and negative

Kurtosis: It is a measure of the possible distribution of a real-valued random variable and it depicts how anomaly the image is. Kurtosis and Skewness are used in the statistical analysis to get an insight into shape of distribution.

4.2.2. Texture features

Texture features are extracted from the input image only next to histogram features. Since the abnormality is generally spread in the image, the textural orientation of each class is extraordinary, which helps to attain better classification accuracy. The gray level co-occurrence matrix symbolizes a statistical method of reviewing the surface that takes into account the spatial relationship of pixels. The GLCM functions spell out the texture of an image, by estimating the recurrence of occurrences of pairs of the pixel with the same values. Generally, these features are calculated by utilizing GLCM probability values and it has somewhere in the range of 22 features

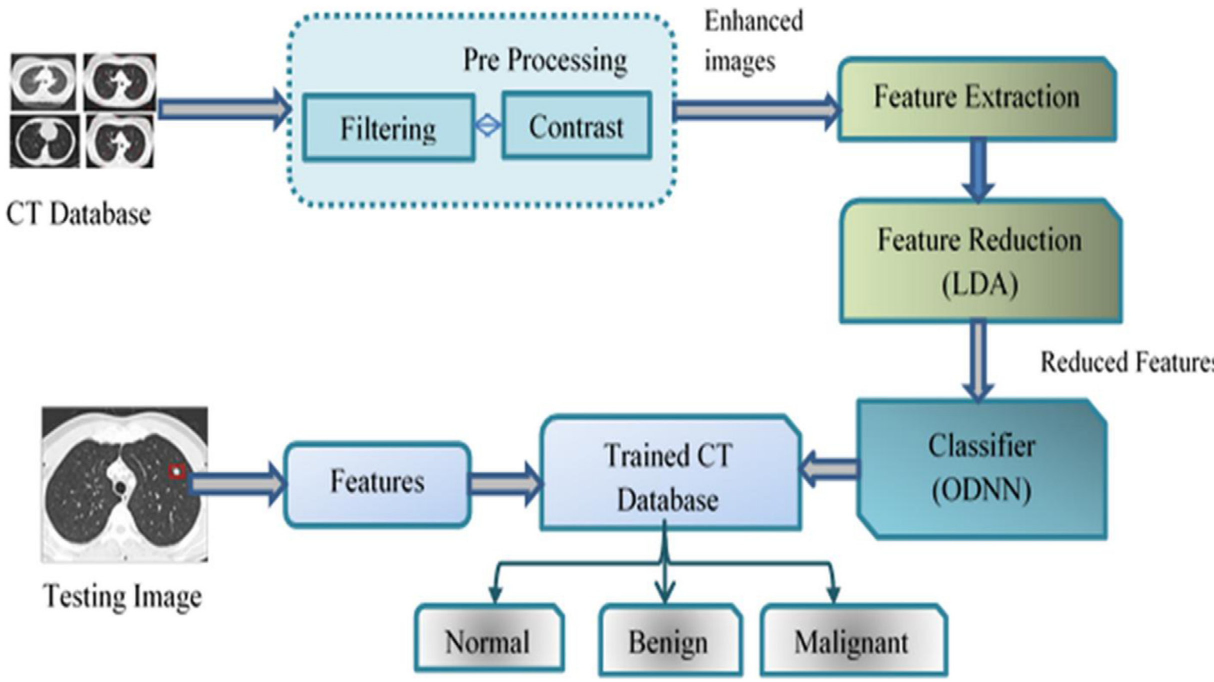


Fig. 1. Block diagram for proposed CT image classification.

among which a few features is considered for the current study regarding CT lung image classification process.

$$G_{P_{ij}} = {}^{F_{ij}}/\textstyle\sum_{i,j=0}^{L-1} {}_{F_{ij}} \tag{2}$$

In the above equation, F_{ij} denotes the 'frequency of occurrences between two gray levels', L is the Number of quantized gray levels, 'i' and 'j' for a given displacement vector for the specified window size.

Energy: This guarantees that the maximum constant values or intermittent consistency in gray level distribution will shape the maximum vitality of surface.

Entropy: It refers to the quantity of data in the image which is required for the compression process. The image with low entropy exhibits tiny contrast and large runs of image pixels in the assigned values.

Homogeneity: The homogeneity constraint is generally called the contrast minute which evaluates the image homogeneity assuming that the prevalent values for minor gray tone changes in pair components. Along these lines, the homogeneity is an evaluation which characterizes prevalent values for minor contrast images.

Contrast: This is the one that calculates the spatial recurrence of an image and the varying moments of GLCM. It symbolizes the variance between the maximum and the base values of a neighboring arrangement of pixels.

Correlations: Correlation evaluates the linear dependence of gray levels of adjoining pixels. The tracking of the digital image correlation stands for an optical procedure, which misuses tracking and the images registration is approached for the measurements of variations in images.

4.2.3. Wavelet-based features

The wavelet transform gives an image handling information because of its beneficial features. The DWT speaks to a linear transformation, which is the function on the data vector whose length is related energy. In the wavelet transform, the feature extractions are carried out by means of two stages as follows. First, the subband of the natural image is developed and these subband are evaluated with the help of various resolutions. Wavelet is an extraordinary numerical method to include extraction and has been used to separate the wavelet coefficients from images. The mean prediction of DWT coefficients is figure out by taking the normal coarse coefficient.

$$Coeff\,[a_t] = \delta_{a_t} \tag{3}$$

where δ_{a_t} is the mean value for approximation coefficient since initially, the images are outfitted to the low pass channel which screens the low recurrence image within the cut off recurrence. Thereafter, the image signals are outfitted to a high pass filter which screens the high-frequency beat signals surpassing the cut-off recurrence.

4.3. Feature reduction: Linear discriminant analysis

The objective is to decrease the first informational index by estimating specific characteristics or features that differentiate one data design from another. All the bad features of CT is to be combined so as to reduce the features. LDA is a dimensionality reduction process, where the original input space is transformed into an autonomous feature space with a dimension that is free of alternate dimensions. The LDA model is illustrated in Fig. 2. It is used as a dimensionality reduction factor for feature vectors before the classification process without any loss of data.

The feature reduction matrix is given as,

$$M_w = \sum_{j=1}^{c} \sum_{i=1}^{N_s} \left(m_i^j - \alpha_j\right) \left(m_i^j - \alpha_j\right)^T \tag{4}$$

where 'c' denotes the number of classes and m_j, N_s and α_j are a test of a class, Number of tests in class and meaning of class. The reduction matrix is calculated using blow equations such as follows.

$$R_s = \sum_{j=1}^{k} (m_j - m) \quad (m_j - m)^T \tag{5}$$

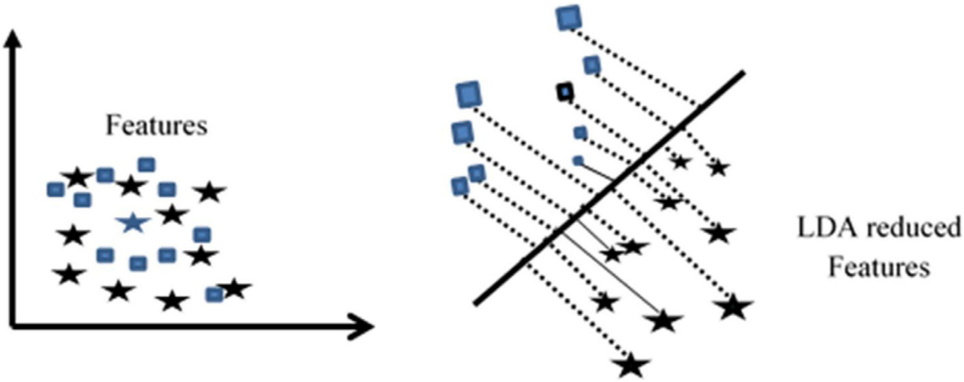


Fig. 2. LDA.

where 'm' is 'mean of all classes' in which LDA strategies are applied with the direct discriminant hypothesis. This standard attempts to expand the proportion of determinant of 'between-class disperses grid of the anticipated examples' to the determinant of the 'inside-class disseminate network of the anticipated examples'. The multi-class LDA is considered where the connection between a set of classes is not same as another set. From the minimal features available for classification procedure, the CT images are classified.

4.4. Classification of lung CT images

In the CT image classification model, the current study proposed DNN in view of profound learning approach. DL structure broadens the customary NN by adding more hidden layers to the system design between the input and output layers so as to demonstrate more unpredictable and nonlinear connections. After the features selection, the grouping step is performed with the help of DNN on the resultant component vector. This classifier works with the help of two capacities such as profound DBN and RBM. In order to improve the classification performance of the proposed model, MGSA optimization is considered which involved steps of optimal deep learning model described in the section below along with an illustration of optimal DNN as Fig. 3.

4.4.1. Deep belief network

During the training stage, a DBN is utilized which is a deep design and feed-forward neural network, i.e. with various hidden layers. The DBN model awards the system to deliver evident-starts based on its hidden units' states which depicts the system conviction. The parameters of a DBN are the weights among the units of layers in addition to the bias of layer. It is a principal challenging task to set up the parameters to train DNN help of a confined restricted RBM [18].

4.4.2. Restricted Boltzmann machine (RBM)

RBM is a two-layer rehashed neural framework in which the stochastic twofold sources of information are associated with stochastic paired yields by symmetrically-weighed affiliations. A preparation case is demonstrated in which the class check is ignored and it is expanded stochastically through RBM in condition (6). This vector is also coursed the other way through RBM which impacts in a confabulation (re-trying) of the remarkable information data.

$$F\left(w,h\right)=-\sum_{i=1}^{i}\sum_{j=1}^{j}I_{ij}w_ih_j-\sum_{i=1}^{I}\alpha_iw_i-\sum_{j=1}^{J}\beta_jh_j \tag{6}$$

where I_{ij} represents the symmetric interaction term between the visible unit w_i and the hidden unit h_j, α, β are the bias terms, i, j are the number of visible and hidden units.

4.4.3. Modified gravitational search algorithm for weight optimization

The novel population-based heuristic algorithm is based on the law of gravity and mass interactions. So, the majority of interactions cooperate for an immediate type of correspondence through the gravitational force. The GSA approach provides a solution to the issue by its mass position and gravitation; also the fitness function of the algorithm is determined by its inertial masses. Subsequently, each mass over an answer and the strategy is directed [22] by properly adjusting the gravitational and inertial masses. The new position is updated for the probability function which is utilized as a part of random value selection following the technique considered for the optimization process.

(i) Weight initialization

Initially, 'w' sets of agents are considered, their positions specified and represented as follows: In Eq. (7) shows w_i^1 the position of agent and $.w_i^s$a search space of agent to choose weights.

$$w=\{w_i^1\ldots w_i^s\} \tag{7}$$

(ii) Fitness evaluation

In this CT lung image classification, the maximum specificity ratio based on the trained and tested structure of DNN is considered as the fitness function and it is shown in the equation below (8).

$$Fit=MAX(spec) \tag{8}$$

(iii) Mass and force updates for generation of new solution

The force between two particles is proportionally relative to their masses and conversely corresponding to their distance, each one of the particles moves towards those particles which are heavy in their mass. This is derived under the Eqs. (9) and (10).

$$Mass_{i(t)}={}^{D_{i(t)}}/{\textstyle\sum}\, D_{i(t)} \tag{9}$$

$$D_{i(t)}=\frac{Fit_{i(t)}-worst_{(t)}}{Best_{(t)}-worst_{(t)}} \tag{10}$$

where 'Fit' represents the fitness value of the particle 'i' at a time 't'. For the maximization problem and for estimating the acceleration of an agent, a set of total force applied from heavier masses has to be taken into account.

(iv) Force evaluation

To give a stochastic trademark to GSA, the total force that follows up on the molecule 'i' in the dth measurement is set to be a randomly weighted total of search segments of the forces.

$$Force_{i(t)}=\sum_{j\in k\ best}Rn_jgr_{(t)}\frac{Mass_{j(t)}*Mass_{i(t)}}{Ed_{ij}+\varepsilon}(w_{j(t)}-w_{i(t)}) \tag{11}$$

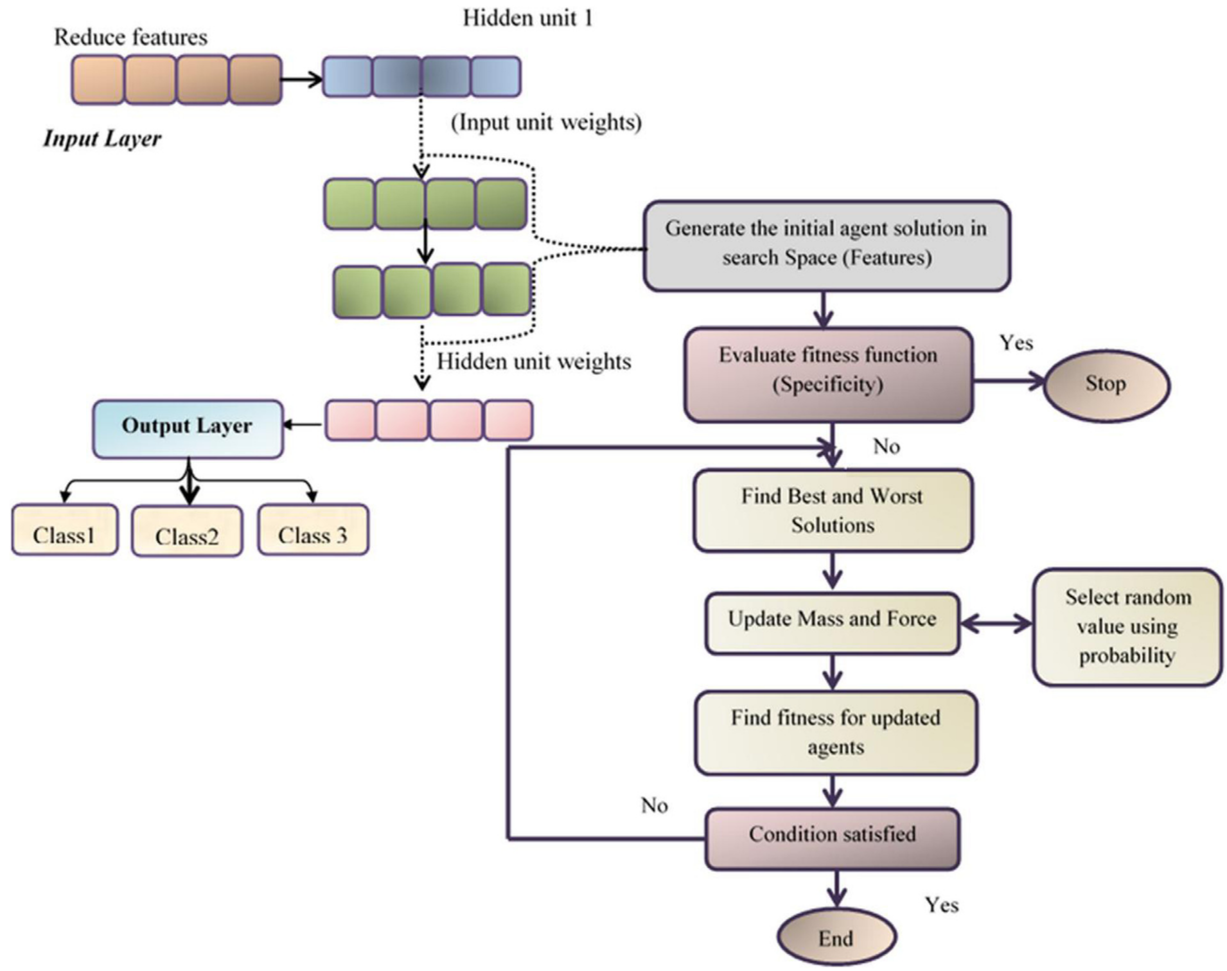


Fig. 3. Optimal DNN structure.

Here $w_{j(t)}$ represents the position of the *i_th* particle in the dimension; Mass *i* and Mass *j* denote the gravitational mass related to the particles '*i*' and '*j*'; Then $gr_{(t)}$ is the gravitational constant; Ed_{ij} denotes the Euclidian distance between the particles '*i*' and '*j*' in the generation '*t*'; 'ε' is a small constant which is bigger than 0.

(v) Modification process to select a random value

In Eq. (11), the random values are selected by calculating the probability function of mass and force updating process and the equation is as follows.

$$prob = 0.3\,(1 - {}^{l}/_{l_{\max}}) \tag{12}$$

This algorithm is iteratively associated for various iterations to join at a sufficiently adequate solution. It justifies saying that the random walk of the insect is compelled by the investigation rate at the present cycle.

(vi) The optimal solution with the termination process

The best solutions which fulfill the objective function are discovered and the algorithm is prepared to give exact solutions in light of maximizing the accuracy of CT lung images in the classification process. If unable to get optimal results in iteration 1, then move for iteration _New = iteration + 1, until get the optimal weights of DNN process, the steps will be repeated.

4.4.4. Fine tuning phase

The working rule of this stage depends on the normal back-propagation algorithm. To detect and classify the abnormal, an output layer is proposed as the highest point of the DNN. Additionally, there is'N' number of input neurons (based on the features), and three hidden layers are used in the current study DNN. The optimized weight is planned through the training stage with the assistance of a training data set, where back propagation begins with the weights that were achieved in the pre-training stage. From the optimal weights, the layer work is refreshed and is shown as follows.

$$\begin{aligned} T(m_i = 1/n) &= \sigma(m_i + \sum opt\_w_{ji} n_j) \\ T(n_i = 1/m) &= \sigma(n_i + \sum opt\_w_{ji} m_j) \end{aligned} \tag{13}$$

where *m* and *n* denote the bias vectors for visible and hidden layers and ff is a logistic function with the range of (0, 1). Further, the training dataset is skilled until the optimized weight is grasped, or maximum accuracy is attained with the help of Eq. (13). Finally, on the basis of the optimal weight (w), the lung images are classified in the testing stage by testing the data set.

5. Result and discussion

The proposed CT lung image classification models were implemented in the working platform of MATLAB 2016 with system configurations such as an i5 processor with 4 GB RAM. In this cancer image classification process, standard CT database was used and the proposed model was compared with existing classifiers like NN, SVM, KNN, DNN and so on, based on different measures of the classification model.

Fig. 4. Sample database images.

Table 1
Performance metrics.

Metrics	Formula
	True positive-TP True negative-TP False positive-FP False negative-FN
Sensitivity	$Sen = \frac{TP}{TP+FN}$
Specificity	$Spc = \frac{TN}{TN+FP}$
Accuracy	$Acc = \frac{TP+TN}{TP+TN+FP+FN}$
PPV	$PPV = \frac{TP}{TP+FP}$
NPV	$NPV = \frac{TN}{TN+FN}$

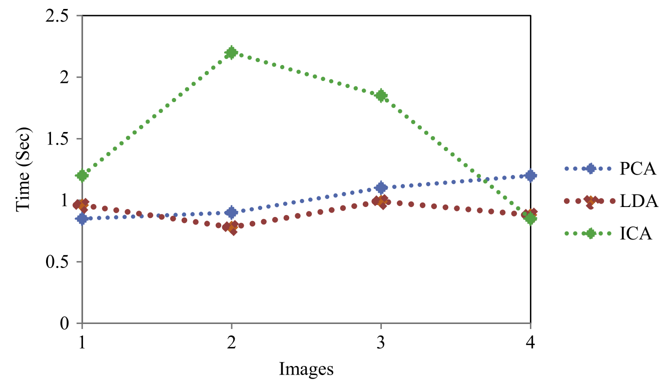


Fig. 5. Feature reduction time comparison.

Table 2
Database images for training and testing analysis.

Phase	Target images	ODNN classifier			
		Normal	Malignant	Benign	Total images
Training	Normal	22	1	4	27
	Malignant	2	18	2	22
	Benign	1	20	0	21
	Total Images	25	39	6	70
Testing	Normal	6	0	2	8
	Malignant	1	9	1	11
	Benign	0	11	0	11
	Total Images	7	20	3	30

Table 3
Optimal weights based hidden layer vs error rate.

Hidden layer with weights	Number of features	Error rate	
		Training	Testing
1 (0.8)	8	0.22	0.28
2 (0.65)	4	0.26	0.33
3 (0.25)	6	0.29	0.35
4 (0.33)	3	0.35	0.31

5.1. Database description

In the proposed work, the database comprised of 50 low-dosage and recorded lung cancer CT image dataset are used for the detection purpose [26]. The CT scan images with 1.25 mm slice thickness were attained by a single breath. The location of nodules was recognized by the radiologist and also provided in the dataset. The test images considered for the proposed work are shown in Fig. 4.

5.2. Performance metrics

The most commonly used evaluation methods for a classification model include in the Table 1. The data based image for training and testing images of ODNN-based classification process is tabulated in Table 2. For lung image investigation, 70 images were considered for training and the remaining 30 images were considered for the testing process.

The results attained by the ODNN model were illustrated with point operations. From the results, the performance of the proposed model is determined by the ability to detect cancer or non-cancerous lung image. Based on the testing data, the model is able to predict the medical conditions of the lung in a new patient's record.

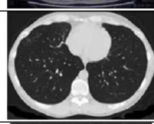
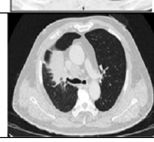
The LDA was utilized for lessening the difficulty of the framework in feature reduction time as illustrated in Fig. 5. The dimensional value of the feature vector got reduced from the images. The comparison graph clearly shows that the proposed technique achieves less computational time coupled with high classification accuracy (because of LDA-based feature reduction). The ideal opportunity for network training is not considered since the weights/biases of the LDA should keep unchanged unless the properties of images change a great deal. Confining the element vectors to the part picked by the LDA, prompts an expansion in precision rates.

Table 3 demonstrates the hidden layers with weights training and testing error values where the diminished features are considered for the training-testing process. In the hidden layer 4, three features extraction is assigned as input factors to give minimum Mean Square Error (MSE) to training data and minimum MSE for testing information.

Table 4 demonstrates the accuracy level of lung cancer image classification rates for the proposed approach. In this test, two classifiers based on supervised machine learning are exhibited, for CT image, as normal or benign or malignant. The proposed ODNN is compared with existing classifiers and it demonstrated that proposed algorithm provides better classification results. It is inferred from the tabulated results that the proposed work expels sensitivity to initial values of clustering because of the evolutionary classification algorithm. Secondly, the texture and color features are considered for grouping CT lung cancer datasets to enhance the classification proficiency. From the validation analysis, maximum accuracy and the significant variation in the accuracy can be observed between the kernel function.

Fig. 5 provides the comparative analysis of classifiers with various measurements like PPV, NPV, Accuracy, Sensitivity, specificity, and accuracy. In this investigation, two classifiers based on supervised machine learning are displayed for CT normal/abnormal human lung classification. It was inferred that the proposed techniques produce classification accuracy of 99% in proposed classifier, which is demonstrated during the testing phase. The classification, it is accuracy of 82.29% in NN 90.54% in SVM, and 74.55%

Table 4
Proposed CT lung image classification with pre-processed results.

CT image	Contrast Enhanced	Type	Accuracy	Sensitivity	Specificity
		Normal	95.21	92.2	86.5
		Benign	85.48	88.52	82.1
		Malignant	92.22	93.2	91.2
		Normal	96.45	91.58	89.2
		Malignant	94.58	90.52	93.2
		Benign	92.22	86.2	84.5

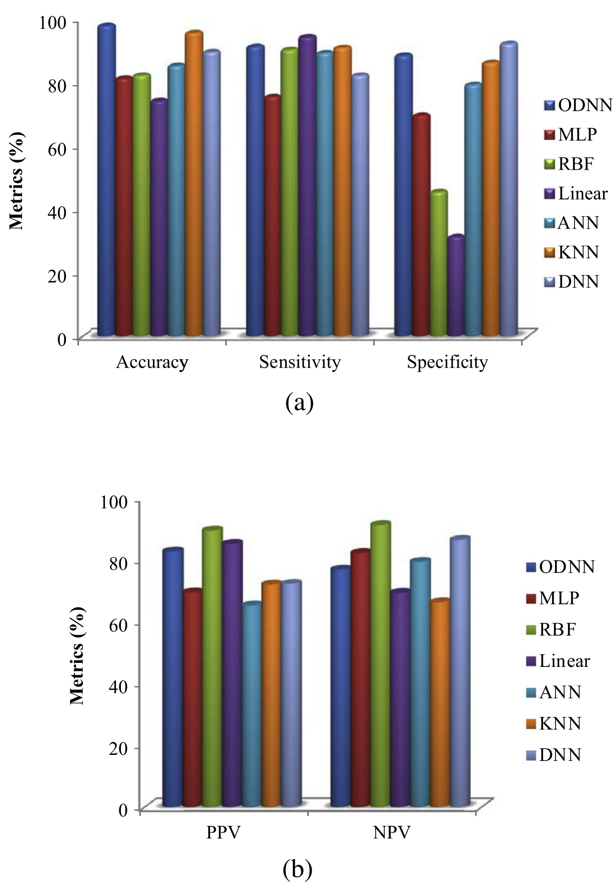


Fig. 6. Classifier comparative analysis.

Table 5
K folds' validation results.

Number of folds	Accuracy	Sensitivity	Specificity	PPV	NPV
10	92.12	88.56	88.54	77.45	90.1
15	93.11	89.45	71.2	72.2	82.2
20	96.2	90.2	83.5	88.65	56.2
25	88.54	75.5	93.2	90.1	67.1
30	89.52	90.2	88.41	79.5	90.1
35	96.52	86.45	92.1	82.45	93.3
40	95.45	92.85	89.4	88.5	59.2

in DNN. The PPV and NPV values indicate better performance as nearly 98% in proposed model. After completing the analysis, the classification specificity was 95% which is not considered as a decent performance, similarly sensitivity parameter. This may be because of the commotion that was exhibited in the phase data due to which the image was misclassified (see Fig. 6).

Table 5 demonstrates the classification of K fold validation and consequences of CT lung cancer classification, accuracy getting approximately 100% in proposed classifier. To be sure, the median operator creates the most noticeable bad classification performance even not as much as single feature extraction. Every time, an overlap is utilized for training and the rest is utilized for the test. The results variance is reduced with a larger k. All the observations are utilized for both training and validation and each observation are utilized for validation for only once.

6. Conclusion

The proposed ODNN with feature reduction demonstrated the better classification in case of lung CT Images compared with others classification techniques. An automatic lung cancer classification approach reduces the manual labeling time and avoids a human mistake. Through machine learning techniques, the researchers planned to achieve better precision and accuracy in recognizing a normal and abnormal lung image. According to the experimental outcomes, the proposed technique is effective for the classification of the human lung images in terms of accuracy, sensitivity, and specificity with its values 94.56%, 96.2%, and 94.2% respectively. The accuracy level has clearly evident that the proposed algorithm is deeply proficient in recognizing cancer-affected parts in CT images. The classification performances of this investigation demonstrate the advantages of this strategy: it is speedy, simple to operate, non-invasive and cheap. In future work, we will use high dosage CT lung images and optimal feature selection with multi-classifier consisted to cancer detection process.

References

[1] S. Rattan, S. Kaur, N. Kansal, J. Kaur, An optimized lung cancer classification system for computed tomography images, in: Image Information Processing (ICIIP), 2017 Fourth International Conference on, IEEE, 2017, pp. 1–6.
[2] P. Naresh, R. Shettar, Image processing and classification techniques for early detection of lung cancer for preventive health care: a survey, Int. J. Recent Trends Eng. Tech. 11 (1) (2014) 595.
[3] F.C. Detterbeck, Lung Cancer Stage Classification: What Does It Mean on Main Street, eighth ed., 2017, pp. 1–26.
[4] J. Li, Y. Wang, X. Song, H. Xiao, Adaptive multinomial regression with overlapping groups for multi-class classification of lung cancer, Comput. Biol. Med. (2018) 1–22.
[5] D.U. Wutsqa, H.L.R. Mandadara, Lung cancer classification using radial basis function neural network model with point operation, in: Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), 2017 10th International Congress on, IEEE, 2017, pp. 1–6.
[6] D. Sharma, G. Jindal, Computer aided diagnosis system for detection of lungcancer in CT scan images, Int. J. Comput. Electr. Eng. 3 (5) (2011) 714.
[7] D. Bhatnagar, A.K. Tiwari, V. Vijayarajan, A. Krishnamoorthy, Classification of Normal and Abnormal Images of Lung Cancer, in: IOP Conference Series: Materials Science and Engineering, vol. 263 (4), IOP Publishing, 2017, p. 042100.
[8] X. Sui, W. Jiang, H. Chen, F. Yang, J. Wang, Q. Wang, Validation of the stage groupings in the eighth edition of the tnm classification for lung cancer, J. Thoracic Oncol. 12 (11) (2017) 1679–1686.
[9] B. El-Sherbiny, N. Nabil, S.H. El-Naby, Y. Emad, N. Ayman, T. Mohiy, A. AbdelRaouf, BLB (brain/lung cancer detection and segmentation and breast dense calculation), in: Deep and Representation Learning (IWDRL), 2018 First International Workshop on, IEEE, 2018, pp. 41–47.
[10] F. Chen, D. Zhang, J. Wu, B. Zhang, Computerized analysis of tongue sublingual veins to detect lung and breast cancers, in: Computer and Communications (ICCC), 2017 3rd IEEE International Conference on, IEEE, 2017, pp. 2708–2712.
[11] M.S. Al-Tarawneh, Lung cancer detection using image processing techniques, Leonardo Electr. J. Practices Tech. 11 (21) (2012) 147–158.
[12] Akay Mehmet Fatih, Support vector machines combined with feature selection for breast cancer diagnosis, Expert Syst. Appl. 36 (2) (2009) 3240–3247.
[13] Y. Xie, J. Zhang, Y. Xia, M. Fulham, Y. Zhang, Fusing texture, shape and deep model-learned information at decision level for automated classification of lung nodules on chest ct, Inf. Fusion 42 (2018) 102–110.
[14] D. Sharma, G. Jindal, Computer aided diagnosis system for detection of lungcancer in CT scan images, Int. J. Comput. Electr. Eng. 3 (5) (2011) 714.
[15] Shankar, S.K. Lakshmanaprabu, Deepak Gupta, Andino Maseleno, Victor Hugo C. de Albuquerque, Optimal feature-based multi-kernel SVM approach for thyroid disease classification, J. Supercomput. (2018) 1–16.
[16] P. Sarker, M.M.H. Shuvo, Z. Hossain, S. Hasan, Segmentation and classification of lung tumor from 3D CT image using K-means clustering algorithm, in: Advances in Electrical Engineering (ICAEE), 2017 4th International Conference on, IEEE, 2017, pp. 731–736.
[17] Y. Xie, J. Zhang, Y. Xia, M. Fulham, Y. Zhang, Fusing texture, shape and deep model-learned information at decision level for automated classification of lung nodules on chest ct, Information Fusion 42 (2018) 102–110.
[18] H. Chougrad, H. Zouaki, O. Alheyane, Deep convolutional neural networks for breast cancer screening, Comput. Methods Programs Biomed. 157 (2018) 19–30.
[19] H. Mohsen, E.S.A. El-Dahshan, E.S.M. El-Horbaty, A.B.M. Salem, Classification using deep learning neural networks for brain tumors, Future Comput. Inf. J. (2017) 1–4.
[20] A. Sharma, K.K. Paliwal, A deterministic approach to regularized linear discriminant analysis, Neurocomputing 151 (2015) 207–214.
[21] S. Nagpal, S. Arora, S. Dey, Feature selection using gravitational search algorithm for biomedical data, Proc. Comput. Sci. 115 (2017) 258–265.
[22] S.K. Lakshmanaprabu, K. Shankar, A. Khanna, D. Gupta, J.J. Rodrigues, P.R. Pinheiro, V.H.C. De Albuquerque, Effective features to classify big data using social internet of things, IEEE Access 6 (2018) 24196–24204.
[23] J. Kuruvilla, K. Gunavathi, Lung cancer classification using neural networks for CT images, Comput. Methods Programs Biomed. 113 (1) (2014) 202–209.
[24] Z. Wang, J. Tao, A fast implementation of adaptive histogram equalization, in: Signal Processing, 2006 8th International Conference on, vol. 2, IEEE, 2006, pp. 1–4.
[25] H. Wang, Y. Fan, B. Fang, S. Dai, Generalized linear discriminant analysis based on euclidean norm for gait recognition, Int. J. Machine Learn. Cybern. 9 (4) (2018) 569–576.
[26] http://www.via.cornell.edu/lungdb.html.

Lakshmanaprabu S.K. completed his Bachelor of Engineering (B.E.) degree in Electronics and Instrumentation Engineering from the R.M.K. Engineering College, Chennai in the year 2009. He did his Masters in M.E. Industrial Engineering from Sudharsan Engineering College, Pudukkottai, Tamil Nadu in the year 2011. He is a senior research fellow and he is currently pursuing his Ph.D. degree in multivariable process control in the Department of Electronics and Instrumentation Engineering of B.S. Abdur Rahman Crescent Institute of Science and Technology, Chennai, India. He received award of National Fellowship from University Grants Commission, Govt. of India, Delhi, for doing his Ph.D. degree for the Year of 2013-2018. His area of interests includes Multivariable Control, Evolutionary Algorithm, Fuzzy Logic Control, Image Processing, Artificial Intelligence, Internet of Things, Model Based Development and Hardware in the loop Testing.

Prof. Dr. Sachi Nandan Mohanty, received his Ph.D. from IIT Kharagpur, India in the year 2014, with MHRD scholarship from Govt of India. He has recently joined as Associate Professor in the Deaprtment of Computer Science & Engineering at Gandhi Institute for Technology Bhubanewar. His research areas include Data mining, Big Data Analysis, Cognitive Science, Fuzzy Decision Making, Brain-Computer Interface, Cognition, and Computational Intelligence. Prof. S N Mohantyhas received 2 Best Paper Awards during his Ph.D at IIT Kharagpur from International Conference at Benjing, China, and the other at International Conference on Soft Computing Applications organized by IIT Rookee in the year 2013. He has awarded Best thesis award first prize by Computer Society of India in the year 2015. He has published 15 International Journals of International repute and has been elected as Member of Institute of Engineers and IEEE Computer Society. He also the reviewer of IJAP, IJDM International Journals.

K. Shankar is an assistant professor in the Department of Computer Science and Information Technology at the Kalasalingam University, Krishnankoil, Tamilnadu, India. He received his master degrees of Master of Computer Applications, Master of Philosophy in Computer Science and Ph.D. degree in computer science from Alagappa University, Karaikudi, India. He has several years of experience working in the research, academia and teaching. His current research interests include Cryptography and Network Security, Cloud security, Image Processing and Soft Computing Techniques.

Arunkumar N. has completed in his BE, ME and PhD in Electronics and Communication Engineering with specialization in Biomedical Engineering. He has a strong academic teaching and research experience of more than 10 years in SASTRA University, India. He is appreciated for his innovative research oriented teaching related practical life experiences to the principles of engineering. He is active in research and has been giving directions to active researchers across the globe.

Gustavo Ramírez González is a professor in department of telematics engineering in Universidad del Cauca, Colombia. He has published several research papers and worked on many research projects.