Hybrid Bio-inspired Algorithm and Convolutional Neural Network for Automatic Lung Tumor Detection

Surbhi Vijh
Department of Information technology, KIET Group of Institution, Ghaziabad
surbhivijh428@gmail.com

Prashant Gaurav
Department of Computer Science and Engineering, Amity University, Sector 125, Noida, Uttar Pradesh
prashantgaurav36@gmail.com

Hari Mohan Pandey* (Corresponding Author)
Department of Computer Science, Edge Hill University, Ormskirk, Lancashire, England
Pandeyh@edgehill.ac.uk

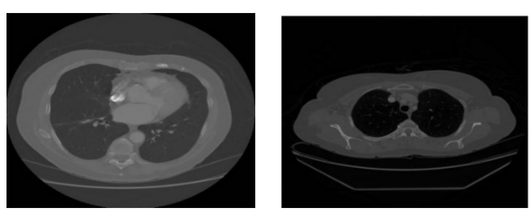
Abstract: In this paper, we have proposed a hybrid bio-inspired algorithm which takes the merits of whale optimization algorithm (WOA) and adaptive particle swarm optimization (APSO). The proposed algorithm is referred as Hybrid WOA_APSO algorithm. We utilize convolutional neural network (CNN) for classification purpose. Extensive experiments are preformed to evaluate the performance of the proposed model. Here, pre-processing and segmentation are performed on 120 Lung CT images for obtaining the segmented tumored and non-tumored region nodule. The statistical, texture, geometrical and structural features are extracted from the processed image using different techniques. The optimized feature selection plays a crucial role in determining the accuracy of the classification algorithm. The novel variant of whale optimization algorithm and adaptive particle swarm optimization, hybrid bio-inspired WOA_APSO is proposed for selecting optimized features. The feature selection grouping is applied by embedding Linear Discriminant analysis which helps in determining the reduced dimensions of subsets. Two fold performance comparisons are done. First, we compare the performance against the different classification techniques such as support vector machine (SVM), artificial neural network (ANN) and CNN. Second, the computational cost of the Hybrid WOA_APSO is compared with the standard WOA and APSO algorithms. The experimental result reveals that the proposed algorithm is capable of automatic lung tumor detection and it outperforms the other state-of-the-art methods on standard quality measures such as accuracy (97.18%), sensitivity (97%) and specificity (98.66%). The results reported in this paper are encouraging, hence, these results will motivate other researcher to explore more in this direction.

Keywords: Medical Imaging, Artificial Intelligence**,** Feature extraction, Hybrid WOA_APSO, Convolutional Neural Network

1. Introduction

According to WHO statistics, the estimated worldwide death of people from cancer is 9.6 million in 2018. However, it is observed that 30-50% of people can be prevented from cancer by providing preventive measures and treatment at an early stage of cancer. Lung cancer is also referred to as lung carcinoma characterized by uncontrollable cell growth in tissues which generally have been categorized as small cell and non-small cell carcinoma on the basis of cellular structure [1]. On the basis of tumor lymph node location and tumor size, there are four stages of lung cancer from I to IV [2, 3].

A Computed tomography (CT) imaging is considered the finest way for analyzing and visualizing the abnormalities present due to less distortion [4]. Image preprocessing consisting of normalization and enhancement are usually performed to improve the quality of image and reduce the distortion. There are different types of filters applied for image enhancement techniques in spatial and frequency domain [5]. The wiener filter minimizes the noise using low pass filter and performs deconvolution using high pass filter (Inverse filtering) on lung CT image obtained from Lung Image Database consortium [6]. The segmentation is the most challenging task in medical imaging for appropriately extracting the features from the segmented tumor nodule region. There are various segmentation technique [7] such as Watershed transform edge-based segmentation, Region-based segmentation [8], Thresholding [9]. The global thresholding and morphological post-processing operation are performed for detecting the tumored region from surrounding and segmenting the lung region nodule. The different statistical features, textural feature, shape, geometrical based features are extracted from the segmented region [10]. The number of features is prominently more in neuro-imaging, so the technique used for feature extraction is Grey Level co-occurrence Matrix (GLCM), Grey Level run-length Matrix (GLRLM), Histogram features, Grey Level Dependence Matrix (GLDM) and Local Binary pattern (LBP) [11, 12]. These set of extracted features provide significant information from medical imaging which helps in evaluating the pattern and decision making process. To reduce the dimensional features space of different modalities of image, the hybrid WOA_APSO feature selection algorithm is proposed, which removes the redundant features [13, 14]. The feature selection technique optimizes the selection of extracted features and provides more dominant information. The extracted feature selection are grouped using a Linear Discriminant algorithm [15-18] for selecting the more dependent and relevant features on class values to improve the performance of classification. The subset of the best feature is selected to reduce the dimensionality of problem space and to maximize the performance of the learning algorithm.



(i) Tumored (ii) Non-tumored

Fig 1. (i) Tumored Lung CT Image (ii) Non-Tumored Lung CT image

Classification is the most critical aspect in determining the performance parameters, i.e. accuracy, sensitivity and specificity of the model that helps in the process of prognosis of lung cancer done by the experts. The comparative classification analysis is performed by considering techniques such as artificial neural network, support vector machine and convolutional neural network [19 ,20]. Support vector machine uses the Multiclass model learning technique to determine the prediction of medical imaging. SVM consider a linear function hypothetical space in higher dimension features which are instructed using a learning bias originated from statistical learning theory [21]. The artificial neural network represents a multilayer network consisting of three different layers with the backpropagation algorithm [22, 23]. Deep learning [24] is a promising field and provides enhanced performance in various medical imaging application [24, 25]. The fully connected convolutional neural network consisting of mesh connection of neurons comprises of activation function and backpropagation for adjustment of weights [26, 27]. The presented state of art outperforms for the fully connected convolutional neural network in computer-aided diagnosis of lung cancer imaging by showing better performance.

The key contribution of this paper is highlighted as follows:

- Firstly, we propose an intelligent lung tumor segmentation algorithm for the detection of tumor and treatment of cancer patients.
- Secondly, we show the hybridization of two metaheuristic algorithms, namely, whale optimization algorithm (WOA) and adaptive particle swarm optimization (APSO). The proposed algorithm is referred as Hybrid WOA_APSO algorithm. Hybrid WOA_APSO is implemented for the selection of the optimized features subset. Here, feature selection grouping is performed by embedding linear discriminant analysis. Our proposed model utilizes a CNN for classification purpose.
- Thirdly, extensive computer simulations are performed to determine the effectiveness of the proposed model. We reported two-fold performance comparison. First, the performance of the proposed model is tested against different state-of-the-art classification techniques and evaluated accuracy, sensitivity and specificity. Secondly, the performance of the proposed algorithm is compared with the standard WOA and APSO algorithms based on the computational cost of convergence to the optimal results.

The rest of the paper is organized as follows: **Section 2** presents the related work; The proposed model is comprehensively discussed in **Section 3**; **Section 4** represents the experimental results and computational performance measures; conclusion and future avenue of the proposed work is given in **Section 5**.

2. Related work

Cancer is the deadly disease worldwide whose detection at an early stage provides preventive measures and treatment to increase the survival rate of patients. The research in medical imaging is growing rapidly to recognize the pattern of disease showing the development of computationally intelligent systems. Data visualization for medical image analysis is one the most promising field in the development of a robust expert system using artificial intelligence [28 ,29], computer vision [30] and pattern recognition application [31-33]. The computer-aided medical diagnosis is the most challenging task in the interpretation of decision making process by a radiologist for extracting the abnormalities in the image [34-35]. The related research work and findings are presented in this section.

Uzelaltinbulat et al. [36] presented lung tumor segmentation technique using Otsu thresholding and morphological operations. Kumar et al. [37] proposed a hybrid of 2D Otsu method and modified artificial bee colony method for segmentation of lung CT image. The performance is computed by evaluating the correlation values. Joon et al. [38] showed the segmentation of cancerous and non-cancerous lung region nodule using k-mean clustering and fuzzy c-mean technique. The structural and texture extracted features are used in classification performed by the support vector machine. Prabukumar et al. [39] proposed a hybrid segmentation technique comprising of Fuzzy C-means (FCM) and region growing algorithm to segment the nodule. The statistical, texture and geometrical features are extracted from the segmented nodule and optimized features are selected using a cuckoo search algorithm. The optimal features are considered for evaluating the classification using support vector machine while obtaining the overall accuracy of 98.5%. Mittal et al. [40] proposed a methodology for multilevel image thresholding for segmentation and introduced exponential kbest gravitational search technique.

Shankar et al. [41] stated a methodology for Alzheimer detection in which various features are extracted using grey level run length matrix, histogram features, grey level occurrence matrix, the local binary pattern features, scale-invariant transform. The feature selection is performed using the grey wolf optimization algorithm to attain the optimized features for performing the classification. The convolutional neural network classification technique is used to achieve the accuracy, sensitivity and specificity as 96.23%, 94% and 96%. Vijh et al. [42] proposed an approach for developing the computer-aided lung tumor segmentation system. The whale optimization algorithm is used for feature selection and support vector machine for classification. The performance of methodology is compared using different SVM kernels. However, the RBF support vector kernel provided the accuracy, sensitivity, specificity of 95%, 100%, 92%.

Reddy et al. [43] presented algorithm consists of parallel thresholding, feature extraction and fuzzy neural network for identifying the lung tumor on CT imaging evaluating the accuracy of 96.5%. Zhang et al. [44] stated a novel approach for computer aided diagnosis for lung tumor detection by implementing multiscale mask region-based convolutional neural network on PET imaging. The performance computed by estimating recall, precision and F-value as 1, 0.90 and 0.95. Uçar et al. [45] presented the methodology for automatic detection of lung nodule using deep learning convolutional neural network architecture with Laplacian of gaussian filter model obtaining the accuracy of 72.97%. Naqi et al. [46] proposed a novel hybrid approach consisting of 3D neighbourhood connectivity, Active contour model (ACM) and geometric properties for 3D nodule candidate detection. The comparative analysis of classification is performed using Naïve Bayes, KNN, SVM and AdaBoost to evaluate the model effectiveness. The Table 1 presents a summary of the existing works.

Table 1. Comparative study of the existing techniques. The comparision is done based on the dataset, segmentation methods, classification method and performance of the techniques. For performance standard measures such as specificity, accuracy, precision, recall etc have been reported.

Author	Dataset	Segmentation Methods	Classification	Performance
Manikandan et al. [47]	Lung CT images	Fuzzy auto seed cluster	SVM kernel's	Specificity-93% Accuracy -94%
Kavitha P et al. [48]	Lung CT images	Otsu's Thresholding and Morphological segmentation method	Artificial Neural network	Accuracy-92.68% Precision-87.50% Recall -100%
Kumar et al. [49]	Brain MRI T1 weighted	Gradient vector flow model	Principal component analysis and Artificial neural network (PCA-ANN)	Accuracy -91.17%
El Abbadi et al. [50]	MRI Brain Images	Morphological segmentation	Probabilistic Neural Network	Accuracy -98%
Nibali et al.[51]	LIDC/IDRI dataset -Lung CT Images	----	Residual network (ResNet) - Deep residual network	Accuracy -89.90% Sensitivity-91.07% Specificity-88.64%

Ali et al. [52]	LIDC/IDRI dataset -Lung CT images	----	Reinforcement learning-convolutional neural network	-----
Duarte et al. [53]	MIAS	Fisher discriminant analysis & texture features	--------	Accuracy- 0.94± 0 .019
Esteva et al. [54]	Clinical Images		Deep Learning–Convolutional neural network architecture	Accuracy -72%
Vijh et al. [55]	Brain MRI	Otsu Thresholding +APSO, Morphological operation	Convolutional neural network (CNN)	Accuracy -98% Loss function - 4%

3. Proposed Methodology

This section shade light on the proposed methodology. The simulated phases implemented in the proposed methodology of computer aided automatic diagnosis system for detection of lung tumor as shown in Fig 2 are (i) Image Acquisition and Normalization (ii) Image pre-processing (iii) Image segmentation (iv) Mathematical Morphological operations (v) Feature extraction (vi) Feature selection with grouping (vii) Classification.

1. 120 samples of Lung CT images is obtained from the NCI Lung cancer database consortium [56]. After the acquisition of lung CT images, the normalization is performed on tumored and non-tumored CT images which are available in cancer imaging archive.

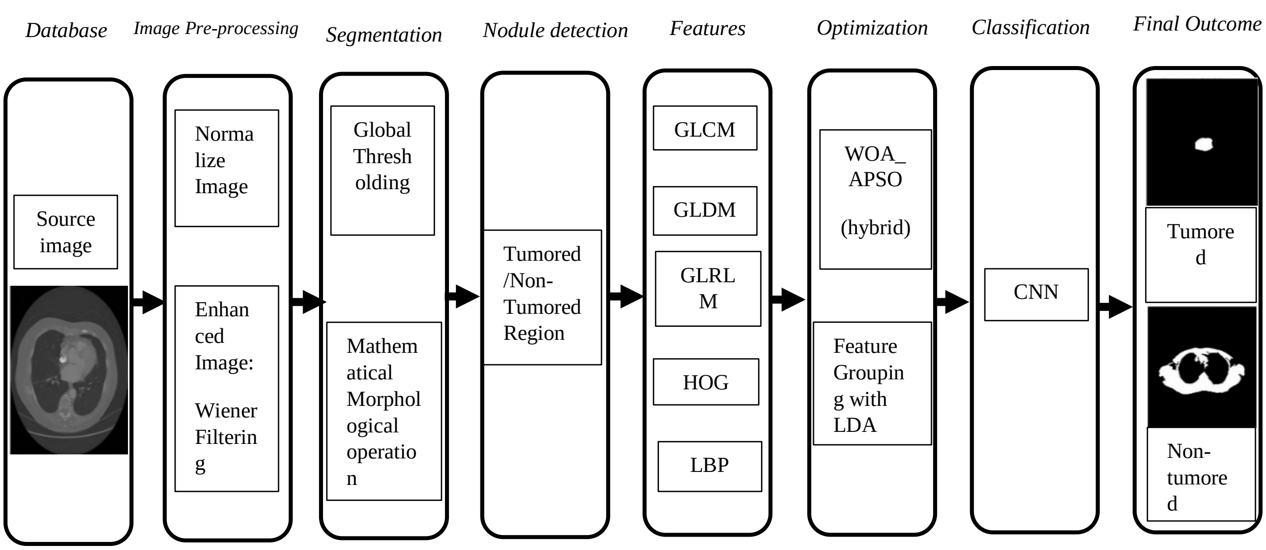


Fig 2. Flow process of proposed methodology.

2. Image preprocessing is considered an important phase in clinical research application of neuroimaging for improving the quality. Image enhancement is achieved by applying a Wiener filter for denoising the Image with a minimum mean square error. The Wiener filtering is the statistical approach for reducing the blurring and smoothing effect in the Image. The F[n] represents the outcome of wiener filter as depicted in Eq.(1).

$$F[n] = \sum_{k=0}^{n} a_k w[n - i] \quad (1)$$

3. Image segmentation is the process of partitioning the Image in multiple regions consisting of a different set of pixel. Global thresholding technique is projected to partition the Image depending upon the intensity of the gray level pixels for threshold T. The segmented image acquired from global thresholding can be represented as H(x, y) and using Eq. (2). Here , $t(x,y)$ is the pixel value of image.

$$\text{H(x, y)} = \begin{cases} 1 & if\ t(x,y) > T \\ 0 & if\ \ t(x,y) \leq T \end{cases} \tag{2}$$

4. Mathematical Morphological operations are estimated by applying the certain structuring element at all possible location for smoothing the region of interest. The mathematical operation are performed where D is binary image and F as structuring element as depicted in Eq.(3), (4),(5) and (6) respectively.

$$\text{Erosion: D} \ominus \text{F} = \{A \mid (F)_A \subseteq \text{D}\} \tag{3}$$

$$\text{Dilation: D} \oplus \text{F} = \{A \mid (F)_A \cap D \neq \emptyset\} \tag{4}$$

$$\text{Opening: D} \ominus \text{F} = \text{D} \ominus \text{F} \oplus \text{DF} \tag{5}$$

$$\text{Closing: D} \ominus \text{F} = \text{D} \oplus \text{F} \oplus \text{F} \tag{6}$$

5. Feature extraction is the most important phase in obtaining the pattern information of the segmented nodule. In the proposed methodology, the total 60 different geometrical, statistical, texture and structure features are extracted from each segmented nodule. The techniques used for extraction are Grey level co-occurrence matrix (GLCM) , Grey level run length matrix (GLRLM), Histogram Oriented Gradient features (HOG), Grey Level Dependence Matrix (GLDM) and Local Binary pattern (LBP). GLCM is referred to as a second-order statistics method which considers the spatial relationship between a couple of pixels. GLRLM helps in obtaining higher-order statistical features consisting of a set of continuous pixels having similar gray level [57]. GLDM extracts the features by computing gray level absolute difference method between two pixels separated by specific displacement [58].Histogram oriented gradient extracts feature by focusing on the structure of Image and uses the feature descriptor for counting the occurrence of gradient orientation in localized portion [59]. LBP considers the shaping based LBP operator for Lung CT image which threshold the neighbouring pixels based on the value of current pixels [60]. Table 2 shows the name of features extracted from segmented tumored and non- tumored lung images for the analysis.

Table 2. Extracted Features

Feature	**Name of Feature**	**Formula**	**Feature**	**Name of Feature**	**Formula**
GLCM	Dissimilarity	$\sum_{i,j}^{g} \lvert i - j \rvert H(i,j)$	**GLCM**	Cluster Prominence	$\sum_{i,j=0}^{g-1} H_{ij}\,(i - F_i + j - F_j)^3$
	Energy	$\sum_{i,j=0}^{g-1} -\log(H\,ij)^2$		Inverse difference moment normalized (IDM)	$\sum_{i,j=0}^{g-1} \frac{H_{ij}}{1+(i-j)^2}$

	Inverse difference normalized (IDN)	$\sum_{i,j=0}^{g-1} \frac{H_{ij}}{1+\mid i-j\mid}$		Sum variance	$-\sum_{i,j=1}^{2g}(1-\mu)^2 IH(i,j)$
	Autocorrelation	$\sum_{i,j=0}^{g}(ij)\, H(i,j)$	**LBP**	Local binary Pattern	$\sum_{h=0}^{H-1} 2^g y(G_h - G(Z_{a1}Z_{a2}))$
	Contrast	$\sum_{i,j=0}^{g-1} -H(i,j)\,(i-j)^2$	**GLDM**	GLDM	$prob[K\alpha(e,r)]$
	Maximum probability	$\mathrm{Max}(GP_{ij})$		Skewness	$\sum_{h=0}^{g-1}(o_i - mean)^3 H(o_i)$
	Sum of squares	$\sum_{i=0}^{g-1}\sum_{j=0}^{g-1}(i-\mu)^2\, H(i,j)$		Kurtosis	$\sum_{h=0}^{g-1}(o_i - mean)^4 H(o_i)$
	Sum average	$\sum_{i,j=2}^{2g-2} i\, H_{i+j}(i)$		Mean	$\sum_{h=0}^{g-1} o_i H(o_i)$
	IM of correlation 1	$\frac{X_{ij}-X_{ij}\,1}{\max\{X_i,X_j\}}$	**HOG**	Variance	$\sum_{h=0}^{g-1}(o_i - mean)^2 H(o_i)$
	Sum entropy	$\sum_{i,j=0}^{2g-2} H_{i+j}(i)\log(H_{i+j}(i))$		Stand Deviation	$\sqrt{\sum_{h=0}^{g-1}(o_i - mean)^2 H(o_i)}$
	IM of correlation 2	$\sqrt{1-\exp(Xij2 - Hij)}$		High grey level Run Emphasis	$\frac{1}{t}\sum_{i,j} i^2 H(i,j)$
	Cluster Shade	$\sum_{i,j=0}^{g-1} H_{ij}(i-F_i+j-F_j)^4$		Run Length Non Uniformity II.	$\frac{1}{t}\sum_{i}\left(\sum_{w} H(i,j)\right)^2$
	Homogeneity	$\sum_{i,j=0}^{g-1}\frac{H_{ij}}{1+(i-j)^2}$		Run Percentage	$\sum_{i,j}\frac{w}{H(i,j)w}$
	Correlation	$\sum_{i,j=0}^{g-1} H_{ij}\frac{(i-\mu)(j-\mu)}{1+(i-j)2}$		Inertia	$\sum_{i=0}^{g-1}\sum_{j=0}^{g-1}(i-j)*H(i,j)$
	Difference variance	$\sum_{i=1}^{g-1}(1-\mu)^2\, H(i,j)$	**GLRLM**	Long Run Emphasis I.	$\frac{1}{t}\sum_{i,j} w^2 H(i,j)$
	Difference entropy	$-\sum_{i=0}^{g-1} H_{i+j}(i)\log(H_{i+j}(i))$		Low grey level Run Emphasis	$\frac{1}{t}\sum_{i,j}\frac{X(i,w)}{i^2}$
	Angular second Moment	$\sum_{i=0}^{g-1}\sum_{j=0}^{g-1} H(i,j)^2$		Gray-Level Non Uniformity	$\frac{1}{t}\sum_{i}\left(\sum_{w} H(i,j)\right)^2$
	Energy	$\sum_{i,j=0}^{g-1} -\log(H\,ij)^2$		Short Run Emphasis	$\frac{1}{t}\sum_{i,j}\frac{H(i,j)}{i^2}$

6. Nature Inspired Meta-heuristic optimization algorithms mimic a physical or biological phenomenon for solving the real world optimization problems. The novel hybrid algorithm comprising of whale optimization algorithm and adaptive particle swarm optimization (WOA_APSO) is proposed for selection of the optimized dimension of features and for providing the effective results.

6.1 Mathematical Formulation of the Hybrid WOA_APSO optimization algorithm

Initially, the hybrid of WOA_APSO algorithm begins with a random solution. However, search agents modify their position with respect to specific agent behaviour. The objective function Fit_{ob} used for each iteration is outlined in Eq.(7) and also 2D Matlab plot is depicted in Fig 3.

$$Fit_{ob}=E*(1+\beta)/ RF \quad (7)$$

$$RF = m / S \quad (8)$$

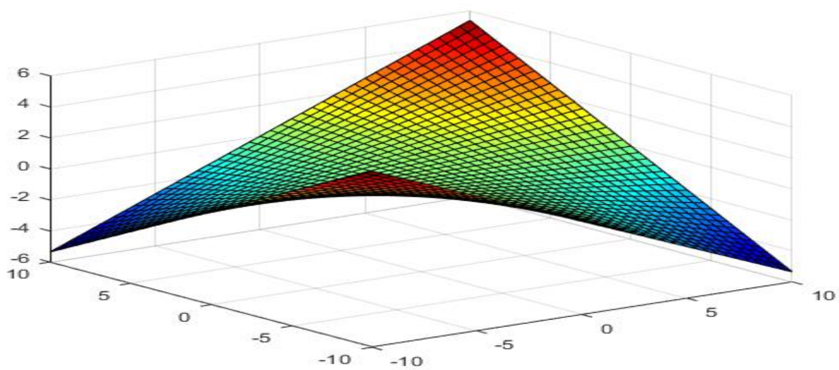


Fig 3. 2D Matlab Plot for F(x, y) = E*(1+β)/ RF

Where E calculates the overall error, β is constant having value 0.5, m represents no of selected features, S are the no swarms. The parameters and the values used in hybrid bio-insppired algorithm are represented in Table 3.

Table 3. Usage of Hybrid Bio-inspired algorithm parameters and values.

Algorithm	**Parameter**	**Values**
WOA_APSO	Lb	1
	Ub	160
	Dim	20
	Max_iteration	35
	Gamma	0.5
	SearchAgents_no	25

The hunting functionality is performed with the help of the best search agent chasing the position of prey to encircle. This behaviour can be mathematically represented using the Eq.(9) and (10) respectively.

$$\vec{E} = |\ \vec{L}.\overrightarrow{Y^*}(i) - \vec{Y}(i)\ | \quad (9)$$

$$\vec{Y}(i+1) = \overrightarrow{Y^*}(i) - P.\vec{E} \quad (10)$$

L and P are the coefficient vector, i is the latest ongoing iteration, $\vec{Y}$ shows the position vector and Y^* represents the position vector of best solution acquired. However the coefficient vector are shown in Eq (11) and 12.

$$P=2\ b.r-b \quad (11)$$

$$L= 2.r \quad (12)$$

Where b is linearly decreasing from 2 to 0 and r is random vector [0, 1].

The bubble net phase formulation for spiral updating position to mimic the helix-shaped movement of humpback whales and prey is shown using the Eq.(13).

$$Y(i+1) = E'.e^{al}.\cos(2\pi l) + Y^*(i) \qquad (13)$$

The humpback whales revolve around the prey within spiral-shaped and shrinking circle simultaneously. The prey updates their location by using Eq. (14) ,(15) and (16).

$$\vec{Y}(i+1) = (C_{tj}^{i} * v^{i+1}) + [f_1 * v_{1j}^{i} * (R_{Best,t}^{t} - x_{tj}^{i})] + [(f_2 * v_{2j}^{i} * (R_{Best,t}^{t} - x_{tj}^{i})]. \qquad (14)$$

Here R_{best} represents the local best search

$$y_{tj}(i+1) = y_{tj}(i) + fa * C_{tj}(i+1) \qquad (15)$$

$$G_{best} = G_{best} + G_a = G_{best} + (\max(y_j) - \min(y_j)) \times rand \qquad (16)$$

Here fa represents Adaptive factor and G_{best} global best location

In the exploration phase, the coefficient vector P is used for searching the prey and it can be shown through Eq. (17).

$$\vec{Y}(i+1) = \overrightarrow{Y_{random}} - \vec{P}.\vec{E} \qquad (17)$$

Algorithm 1: Hybrid WOA_APSO

1. Initialize the humpback whales population as Yi = (1, 2…n), initially i=1.
2. Evaluate the fitness function Fit_{ob} for Yi.
3. *while* (i < maximum number of iterations)
4. *for* each search agent
5. Update b, P, L and m
6. if_1 (m<0.5)
7. if_1 (|P| < 1)
8. Modify the position of currents search agent by using Eq. (9).
9. $elseif_2$ (|P|>1)
10. Select a random search agent ($\overrightarrow{Y_{random}}$)
11. Modify the position of the current search agent by using the Eq. (17)
12. Search for the minimum individual best position to determine the global best position.
13 $endif_2$
14 $elseif_1$ (m>0.5)
15. Upgrade the position of the current search by using the Eq. (13).
16. By using Eq. (11) & (12) replace with the new as the global best position.
17. $endif_1$

18. end *for*

19. Upgrade Y^* if better solution founds.

20. i=i+1

21. end *while*

22. Return Y* = best search agent.

Time Complexity: The proposed hybrid bio-inspired algorithm has two inner loops for population 'n' and one outer loop for iteration 't' where n = 20 & t= 35. The extreme case complexity of hybrid WOA_APSO can be O (n2). However, when population size is large, the time complexity of the algorithm can be represented as follows: T (n) = 2n + s + nlog (n).t =O (nlog (n).t).

6.2 Feature selection grouping is performed using Linear Discriminant Analysis for reducing the dimensions and selecting the best-optimized subsets which enhances the classification performance effectively.

7. Convolutional neural network is deep learning classification technique for training and testing the learning network [61, 62]. The neural network comprises of three densely connected layer consisting of activation function connecting one neuron to another neuron as depicted in Fig 4. The backpropagation algorithm is used for updating the weights and deltas with a learning rate of 0.001. The testing is performed on different parameters to identify the best combination for determining the robustness of experiment. The various parameter are as follows: Layer neurons: [5,10],[10,15],[15,20]; Activation function : relu, softmax; Validation split: 0.1,0.2,0.3; Batch size: 1,2,3; Learning rate: 0.1,0.01,0.001; and Epochs : 10,20,40,60,80,100,200. Table 4 represents the parameters and values used for the CNN.

Table 4. Parameter and values used in Convolutional neural network

Algorithm	Parameter	Value
CNN	Layer	3
	activation function	relu,softmax
	Optimizer	Adam
	LOSS	categorical_crossentropy
	Epochs	200
	validation_split	0.2
	BATCH SIZE	2

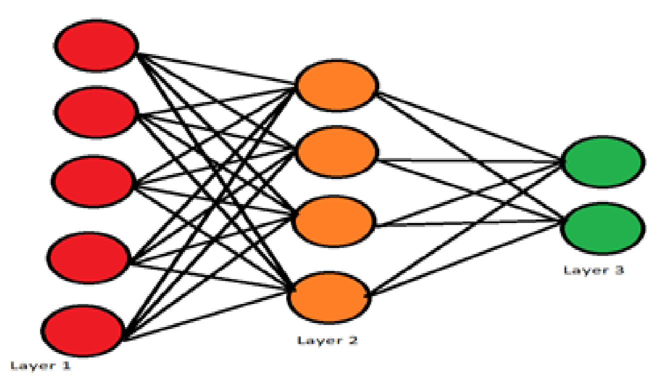
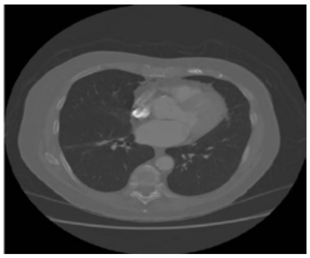


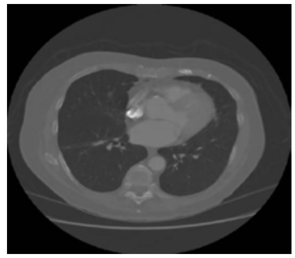
Fig 4. Fully connected convolutional neural network

4. Experiment Result and Analysis

The 120 lung CT tumored and non tumored images acquired from lung cancer database consortium for evaluation of experimental results. The optimized features are taken as input for classification are segregated in 7:3 ratio for training and testing purpose.

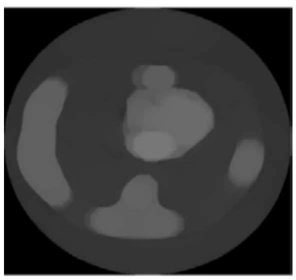


(i) Normalized Image

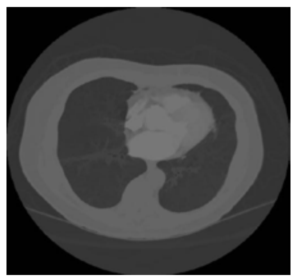


(ii) Filtered Image

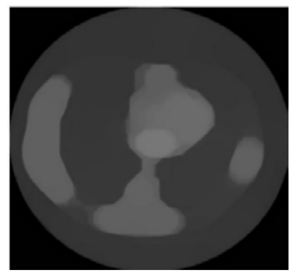
(iii) Global Thresholding



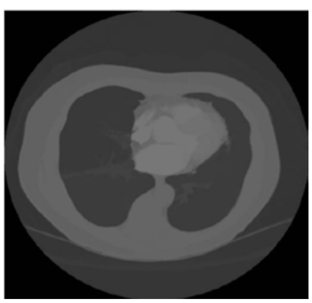
(iv) Opening



(v) Opening by reconstruction



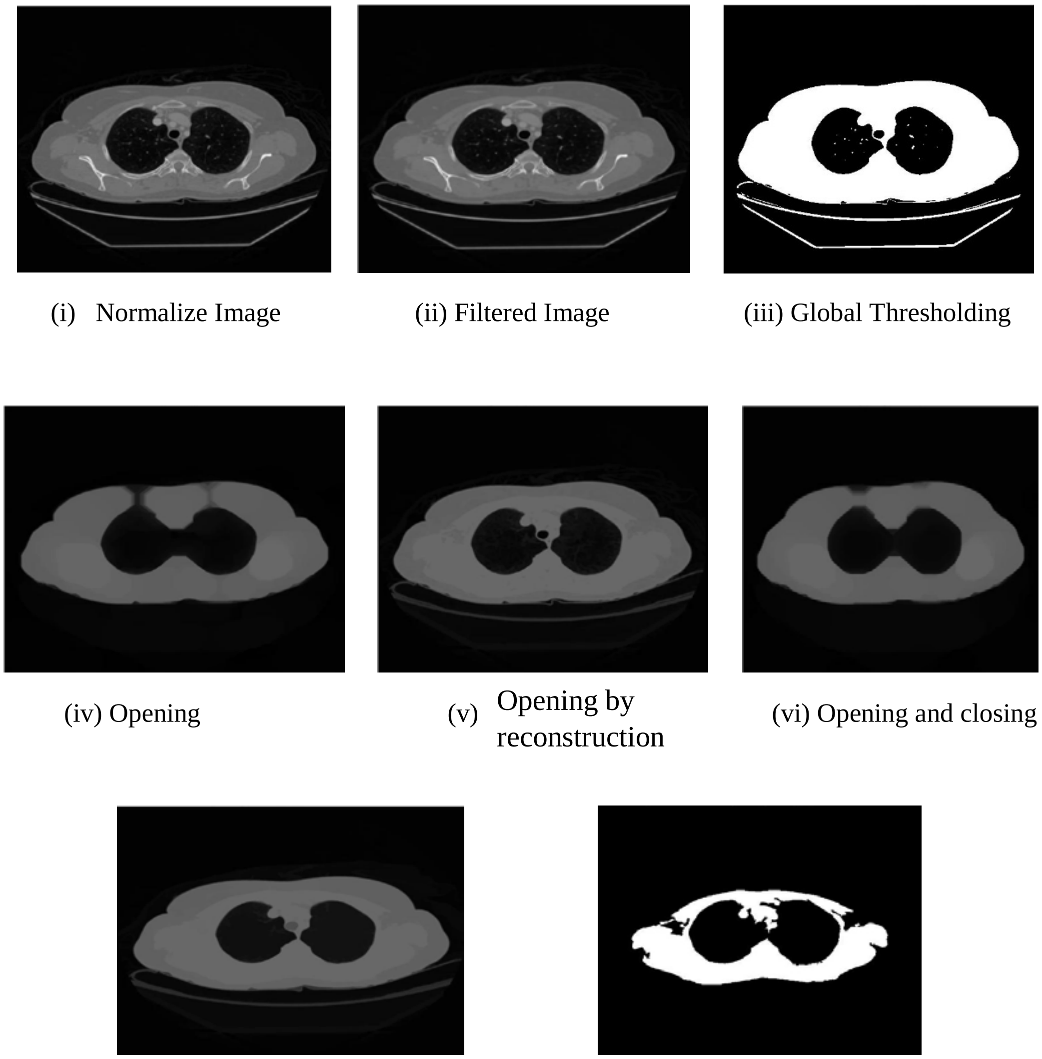
(vi) Opening and closing



(vii) Opening-Closing by Reconstruction

(viii) Regional Maxima of Opening-Closing by Reconstruction

Fig 5. The sample result analysis on Lung tumored Image for detection of Nodule



(i) Normalize Image (ii) Filtered Image (iii) Global Thresholding

(iv) Opening (v) Opening by reconstruction (vi) Opening and closing

(vii) Opening-Closing by Reconstruction (viii) Regional Maxima of Opening-Closing by Reconstruction

Fig 6. The sample result analysis on Lung Non- tumored Image for detection of Nodule

The optimized threshold value achieved from WOA_APSO algorithm is compared with whale optimization algorithm (WOA) and adaptive particle swarm optimization (APSO). The achieved threshold value of WOA_APSO, WOA, APSO are 1.16, 2.09, and 1.9. So, the proposed bio-inspired hybrid WOA_APSO algorithm is giving precise informational optimized features subsets.

Table 5. Optimized Selected Features Subset

Sum Average	**Long Run Emphasis**	**Run Length Non-Uniformity**	**Low Gray level Run Emphasis**	**High Gray Level run Emphasis**	**Mean**	**Skewness**	**Kurtosis**	**Run Length Non-Uniformity**	**Class value**
2.29	19218.84	96.4479	100.4414	5893.224	1.1447	2.01957	5.0786	96.44796	0
2.1947	28150.36	163.6982	97.12592	5331.39	1.0971	2.7202	8.3994	163.6982	0
2.2103	25320.4	80.3332	95.31727	5074.938	1.1049	2.57784	7.6452	80.33326	0
2.1522	47876.4	78.6185	83.83347	3964.353	1.0759	3.20094	11.2460	78.61852	0

2.5389	12683.62	128.558	106.8538	6759.883	1.2689	1.04237	2.0865	128.5588	0
2.5003	13692.13	118.406	106.0758	6529.926	1.2496	1.15684	2.3382	118.4067	0
2.2186	26600.17	85.7822	95.54607	5105.718	1.1090	2.50816	7.2908	85.78228	0
2.2494	30015.29	79.6508	92.85763	4770.801	1.1244	2.27514	6.1762	79.65085	0
2.2017	27257.69	92.2708	95.26133	5067.48	1.1006	2.6546	8.0468	92.27081	0
2.0691	101773.3	232.107	49.68066	2765.702	1.0345	5.10078	27.01794	2.0691	1
2.0328	121067.6	330.5319	39.3437	2633.991	1.01636	7.62568	59.15096	2.0328	1
2.0493	104101.7	285.0864	50.65326	2781.266	1.02461	6.13689	38.66141	2.0493	1
2.0441	107260.8	293.685	47.32784	2730.529	1.02203	6.51332	43.42328	2.0441	1
2.0268	141059.5	439.4875	27.46381	2541.249	1.01336	8.4763	72.84769	2.0268	1

Run Length Non- Uniformity	**LBP 1**	**LBP 2**	**LBP 4**	**LBP 6**	**LBP 8**	**LBP 18**	**LBP 20**	**LBP 22**	**Class Value**
96.4479	0.6241	3.9027	5.6613	5.5266	9.4334	166.0049	178.3931	209.1352	0
163.6982	1.38394	14.4269	6.4487	14.8731	14.9557	187.4467	158.0956	183.8711	0
80.3332	0	2.6145	3.5119	2.0653	1.9645	120.6988	126.473	160.5806	0
78.6185	0	1.2176	0.5654	3.2268	2.6067	71.84998	70.3722	101.3662	0
128.5588	3.75799	8.1659	10.0845	8.1010	11.4551	203.7716	186.0544	235.3808	0
118.406	2.97823	9.1889	12.3220	8.2400	10.1207	201.6179	162.3591	225.4596	0
85.7822	0.68514	1.3914	5.8027	6.5657	4.9089	136.2018	135.387	166.0467	0
79.6508	0.97669	1.5798	4.3095	3.8455	1.6442	143.2166	137.4865	156.7091	0
92.2708	0	6.0355	3.6341	6.9010	4.4248	146.1222	153.3348	176.1286	0
2.0691	0.88508	0.9473	0.8688	2.6195	1.6197	30.97928	55.0529	32.1778	1
2.0328	0	0.8622	0	2.6543	0	52.63489	20.3920	39.4177	1
2.0493	0	1.3256	0.6096	0.6547	2.0256	14.30879	47.1313	16.4106	1
2.0441	0	0.884	1.6993	0.9474	1.4987	45.7859	38.8361	28.7902	1
2.0268	0	0.9572	0.9010	0	0	11.9131	15.9448	17.2959	1

LBP 24	**LBP 34**	**LBP 36**	**LBP 38**	**LBP 40**	**LBP 50**	**LBP 52**	**LBP 54**	**LBP 56**	**Class value**
195.0578	351.533	301.095	312.6975	319.8348	176.9504	203.6467	190.9865	161.6447	0
148.8005	302.4375	208.926	291.9901	216.2246	158.332	182.0684	151.3116	185.1942	0
155.4493	292.4331	207.4571	249.5617	208.6418	123.4223	158.4634	151.4662	119.5931	0
102.5716	181.7596	147.5484	147.1823	136.703	66.70948	101.6018	98.88786	72.21412	0
223.4218	395.9502	286.0657	329.443	339.4295	187.1072	235.5035	225.1967	199.3057	0
217.505	367.9648	290.6264	302.8536	318.8569	165.3334	224.8615	220.4193	198.9715	0
191.1737	366.9759	243.2482	293.0461	228.3787	132.0325	168.4901	186.2506	136.3982	0
161.79	332.7053	170.2044	315.1821	155.3123	136.9516	157.8566	162.9077	145.1957	0
169.3257	332.2018	213.7075	301.8197	231.957	151.542	176.2257	167.812	146.2633	0
55.4414	79.2554	54.8933	66.2689	56.7798	55.1357	31.49076	54.7443	30.1848	1
14.1459	61.3162	22.4648	82.7475	17.5012	18.6852	38.5865	12.4575	51.8421	1
58.1105	41.1567	50.6474	33.6050	52.2777	44.1412	15.8069	55.8905	12.3614	1
36.9702	52.3596	57.4833	62.0686	47.2138	38.1989	26.5194	36.2923	43.05667	1
22.5867	49.9930	18.9753	41.1125	18.3220	15.1034	16.4373	21.7003	11.0442	1

The performance measures of intelligent lung tumor diagnosis system are acquired by comparing the different classification algorithm. The evaluation analysis parameter used for determining the effectiveness of model is accuracy, sensitivity, specificity shown in Eq. (18) ,(19) and (20) respectively.

Accuracy : [TP+TN/ Total] *100 (18)

Sensitivity : [TP/ TP+FN] *100 (19)

Specificity : [1-FPR] *100 (20)

Where TP depicts total number of correctly segmented Image of true positives which are classified properly, FN depicts a total number of correctly segmented image of true negatives which are not classified properly, FPR is the number of incorrect segmented images which are classified properly.

Table 6. Comparative Performance Analysis of Classification

Parameters	SVM	ANN	CNN
Accuracy	80	95.79	**97.18**
Sensitivity	85	89.33	**97**
Specificity	95.12	95.45	**98.66**

The time complexity of the proposed algorithm (Algorithm 1) is presented in Section 3. The computational time of the proposed algorithm against the standard WOA and APSO is reported in Table 7. The presented Table 7 reveals that the proposed algorithm takes less time to converge as compared to other two existing algorithms.

Table 7. Comparison of efficiency in terms of time taken by different meta-heuristic technique

Algorithm	WOA	APSO	WOA_APSO
Time taken in seconds	189.9583	193.55621	**184.48222**

4.1 Result and Discussion

Fig 1 shows the sample of considered CT tumored and non-tumored lung image taken from cancer imaging archive. The segmented image result analysis of proposed algorithm at each successive for efficient detection of the nodule is shown in Fig 5 and 6 respectively. Table 5 contains the optimized and discriminant set of grouped features obtained by applying proposed hybrid whale optimization algorithm and adaptive particle swarm optimization (WOA_APSO) grouped by applying Linear Discriminant analysis (LDA). The comparative performance analysis of different classification technique is shown in Table 6. Fig 7 depicts that convolutional neural network provides better accuracy, sensitivity and specificity in comparison to support vector machine and artificial neural network. The achieved accuracy, sensitivity and specificity are 97.18, 97 and 98.66.

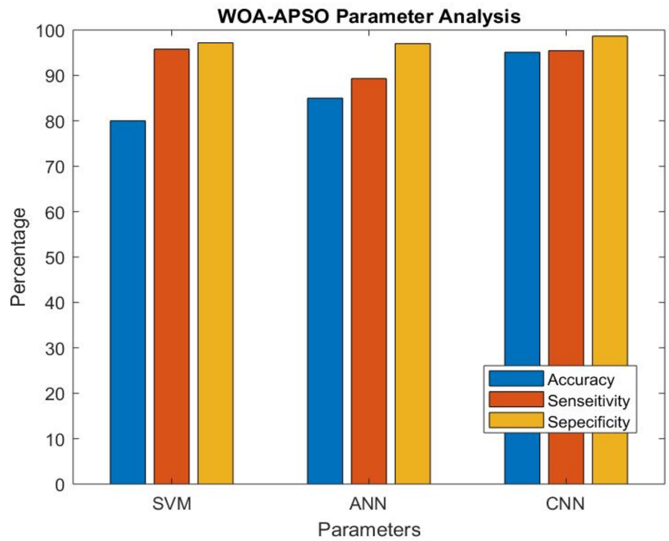


Fig 7. Comparative performance analysis

5. Conclusion and Future Work

In this paper, we have presented a novel approach for early detection, diagnosis and prediction to improve the treatment of patients and take the preventative measures. Here, we implemented a hybrid WOA_APSO algorithm (see Algorithm 1). Image preprocessing and segmentation technique is applied for partitioning and segmenting the tumor region. The different features are extracted to gather the statistics information analysis which assist in decision making process. The proposed state-of-art method provides better consolidated optimized dimension of features selection grouping approach by implementing hybrid WOA_APSO algorithm embedding LDA. The convolutional neural network classification technique outperforms by providing accuracy of 97.5% in comparison to support vector machine and artificial neural network. The methodology demonstrates the effectiveness and promising results for clinical application as compared to exiting algorithms. The proposed method limits the working for 3-dimensional medical imaging .In the future, the work can be extended by using different modalities of medical imaging and other metaheuristic techniques can be incorporated for enhancing the system performance.

Conflict of interest

The authors declare that they have no conflict of interest

References

1. Sluimer, I., Schilham, A., Prokop, M. and Van Ginneken, B., 2006. Computer analysis of computed tomography scans of the lung: a survey. *IEEE transactions on medical imaging*, *25*(4), pp.385-405.
2. Sivakumar, S. and Chandrasekar, C., 2013. Lung nodule detection using fuzzy clustering and support vector machines. *International Journal of Engineering and Technology*, *5*(1), pp.179-185.

3. Dimililer, K., Ever, Y.K. and Ugur, B., 2016, September. ILTDS: Intelligent lung tumor detection system on ct images. In *The International Symposium on Intelligent Systems Technologies and Applications* (pp. 225-235). Springer, Cham.
4. Makaju, S., Prasad, P.W.C., Alsadoon, A., Singh, A.K. and Elchouemi, A., 2018. Lung cancer detection using CT scan images. *Procedia Computer Science*, *125*, pp.107-114.
5. Bari, M., Ahmed, A. and Naveed, S., 2019. Lungs Cancer Detection Using Digital Image Processing Techniques: A Review. *Mehran University Research Journal of Engineering and Technology*, *38*(2), pp.351-360.
6. Perumal, S. and Velmurugan, T., 2018. Preprocessing by contrast enhancement techniques for medical images. *International Journal of Pure and Applied Mathematics*, *118*(18), pp.3681-3688.
7. Srimani, P.K. and Mahesh, S., 2013. A Comparative study of different segmentation techniques for brain tumour detection. *International journal of emerging technologies in computational and applied sciences*, *2*(4), pp.192-197.
8. Mesanovic, N., Grgic, M., Huseinagic, H., Males, M., Skejic, E. and Smajlovic, M., 2011, February. Automatic CT image segmentation of the lungs with region growing algorithm. In *18th International Conference on Systems, Signals and Image Processing-IWSSIP* (pp. 395-400).
9. John, J. and Mini, M.G., 2016. Multilevel thresholding based segmentation and feature extraction for pulmonary nodule detection. *Procedia Technology*, *24*, pp.957-963.
10. Novitasari, D.C.R., Lubab, A., Sawiji, A. and Asyhar, A.H., 2019. Application of Feature Extraction for Breast Cancer using One Order Statistic, GLCM, GLRLM, and GLDM. *Advances in Science, Technology and Engineering Systems Journal*, *4*(4), pp.115-120.
11. Bala, R., 2017. Survey on texture feature extraction methods. *International Journal of Engineering Science*, *10375*.
12. Zu, C., Wang, Y., Zhou, L., Wang, L. and Zhang, D., 2018, April. Multi-modality feature selection with adaptive similarity learning for classification of Alzheimer's disease. In *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)* (pp. 1542-1545). IEEE.
13. Mirjalili, S. and Lewis, A., 2016. The whale optimization algorithm. *Advances in engineering software*, *95*, pp.51-67.
14. Zhan, Z.H., Zhang, J., Li, Y. and Chung, H.S.H., 2009. Adaptive particle swarm optimization. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, *39*(6), pp.1362-1381.
15. Sharma, A., Paliwal, K.K., Imoto, S. and Miyano, S., 2014. A feature selection method using improved regularized linear discriminant analysis. *Machine vision and applications*, *25*(3), pp.775-786.
16. Balakrishnama, S. and Ganapathiraju, A., 1998. Linear discriminant analysis-a brief tutorial. *Institute for Signal and information Processing*, *18*, pp.1-8.
17. Anter, A.M. and Hassenian, A.E., 2018. Normalized multiple features fusion based on PCA and multiple classifiers voting in CT liver tumor recognition. In *Advances in Soft Computing and Machine Learning in Image Processing* (pp. 113-129). Springer, Cham.
18. Vijh, S. and Gaur, D., 2017, December. Secured Biometric Template Matching by Using Linear Discriminant Analysis. In International Conference on Health Information Science (pp. 194-203). Springer, Cham.

19. De Bruijne, M., 2016. Machine learning approaches in medical image analysis: From detection to diagnosis.
20. SeDaTiVe: SDN-enabled deep learning architecture for network traffic control in vehicular cyber-physical systems, A Jindal, GS Aujla, N Kumar, R Chaudhary, MS Obaidat, I You, IEEE Network 32 (6), 66-73
21. Nalepa, J. and Kawulok, M., 2019. Selecting training sets for support vector machines: a review. *Artificial Intelligence Review*, *52*(2), pp.857-900.
22. Bertolaccini, L., Solli, P., Pardolesi, A. and Pasini, A., 2017. An overview of the use of artificial neural networks in lung cancer research. *Journal of thoracic disease*, *9*(4), p.924.
23. Ganesan, N., Venkatesh, K., Rama, M.A. and Palani, A.M., 2010. Application of neural networks in diagnosing cancer disease using demographic data. *International Journal of Computer Applications*, *1*(26), pp.76-85.
24. Singh, A., Aujla, G. S., Garg, S., Kaddoum, G., & Singh, G. (2019). Deep Learning-based SDN Model for Internet of Things: An Incremental Tensor Train Approach. IEEE Internet of Things Journal.
25. Shen, D., Wu, G. and Suk, H.I., 2017. Deep learning in medical image analysis. *Annual review of biomedical engineering*, *19*, pp.221-248.
26. Suzuki, K., 2017. Overview of deep learning in medical imaging. *Radiological physics and technology*, *10*(3), pp.257-273.
27. Anwar, S.M., Majid, M., Qayyum, A., Awais, M., Alnowami, M. and Khan, M.K., 2018. Medical image analysis using convolutional neural networks: a review. *Journal of medical systems*, *42*(11), p.226.
28. Jindal, A., Aujla, G. S., Kumar, N., Prodan, R., & Obaidat, M. S. (2018, December). DRUMS: Demand response management in a smart city using deep learning and SVR. In 2018 IEEE Global Communications Conference (GLOBECOM) (pp. 1-6). IEEE.
29. Gulati, A., Aujla, G. S., Chaudhary, R., Kumar, N., & Obaidat, M. S. (2018, May). Deep learning-based content centric data dissemination scheme for Internet of Vehicles. In 2018 IEEE International Conference on Communications (ICC) (pp. 1-6). IEEE.
30. Aujla, G. S., Jindal, A., Chaudhary, R., Kumar, N., Vashist, S., Sharma, N., & Obaidat, M. S. (2019, May). DLRS: deep learning-based recommender system for smart healthcare ecosystem. In ICC 2019-2019 IEEE International Conference on Communications (ICC) (pp. 1-6). IEEE.
31. Teramoto, A., Fujita, H., Yamamuro, O. and Tamaki, T., 2016. Automated detection of pulmonary nodules in PET/CT images: Ensemble false-positive reduction using a convolutional neural network technique. *Medical physics*, *43*(6Part1), pp.2821-2827.
32. Gore, J.C., 2019. Artificial intelligence in medical imaging.
33. Jiang, J., Trundle, P. and Ren, J., 2010. Medical image analysis with artificial neural networks. *Computerized Medical Imaging and Graphics*, *34*(8), pp.617-631.
34. Chen, C.H., 2015. *Handbook of pattern recognition and computer vision*. World Scientific.
35. Bi, W.L., Hosny, A., Schabath, M.B., Giger, M.L., Birkbak, N.J., Mehrtash, A., Allison, T., Arnaout, O., Abbosh, C., Dunn, I.F. and Mak, R.H., 2019. Artificial intelligence in cancer imaging: clinical challenges and applications. *CA: a cancer journal for clinicians*, *69*(2), pp.127-157.

36. Uzelaltinbulat, S. and Ugur, B., 2017. Lung tumor segmentation algorithm. *Procedia computer science*, *120*, pp.140-147.
37. Kumar, S., Sharma, T.K., Pant, M. and Ray, A.K., 2012. Adaptive artificial bee colony for segmentation of CT lung images. *Int J Comp App iRAFIT*, *5*, pp.1-5.
38. Joon, P., Bajaj, S.B. and Jatain, A., 2019. Segmentation and detection of lung cancer using image processing and clustering techniques. In *Progress in advanced computing and intelligent engineering* (pp. 13-23). Springer, Singapore.
39. Prabukumar, M., Agilandeeswari, L. and Ganesan, K., 2019. An intelligent lung cancer diagnosis system using cuckoo search optimization and support vector machine classifier. *Journal of Ambient Intelligence and Humanized Computing*, *10*(1), pp.267-293.
40. Mittal, H. and Saraswat, M., 2018. An optimum multi-level image thresholding segmentation using non-local means 2D histogram and exponential Kbest gravitational search algorithm. *Engineering Applications of Artificial Intelligence*, *71*, pp.226-235.
41. Shankar, K., Lakshmanaprabu, S.K., Khanna, A., Tanwar, S., Rodrigues, J.J. and Roy, N.R., 2019. Alzheimer detection using Group Grey Wolf Optimization based features with convolutional classifier. *Computers & Electrical Engineering*, *77*, pp.230-243.
42. Vijh, S., Gaur, D. and Kumar, S., 2019. An intelligent lung tumor diagnosis system using whale optimization algorithm and support vector machine. *International Journal of System Assurance Engineering and Management*, pp.1-11.
43. Reddy, U.J., Reddy, B.R.V.R. and Reddy, B.E., 2019. Recognition of Lung Cancer Using Machine Learning Mechanisms with Fuzzy Neural Networks. *Traitement du Signal*, *36*(1), pp.87-91.
44. Zhang, R., Cheng, C., Zhao, X. and Li, X., 2019. Multiscale Mask R-CNN–Based Lung Tumor Detection Using PET Imaging. *Molecular imaging*, *18*, p.1536012119863531.
45. Uçar, M. and Uçar, E., 2019. Computer-aided detection of lung nodules in chest X-rays using deep convolutional neural networks. *Sakarya University Journal of Computer and Information Sciences*, *2*(1), pp.41-52.
46. Naqi, S.M., Sharif, M. and Lali, I.U., 2019. A 3D nodule candidate detection method supported by hybrid features to reduce false positives in lung nodule detection. *Multimedia Tools and Applications*, *78*(18), pp.26287-26311.
47. Manikandan, T. and Bharathi, N., 2016. Lung cancer detection using fuzzy auto-seed cluster means morphological segmentation and SVM classifier. *Journal of medical systems*, *40*(7), p.181.
48. Kavitha P, Ayyappan G (2018) Lung cancer detection at early stage by using SVM classifier techniques. Int J Pure Appl Math 119(12):3171–3180.
49. Kumar, V., Sachdeva, J., Gupta, I., Khandelwal, N. and Ahuja, C.K., 2011, December. Classification of brain tumors using PCA-ANN. In *2011 world congress on information and communication technologies* (pp. 1079-1083). IEEE.
50. El Abbadi, N.K. and Kadhim, N.E., 2017. Brain cancer classification based on features and artificial neural network. *Brain*, *6*(1), pp.123-134.

51. Nibali, A., He, Z. and Wollersheim, D., 2017. Pulmonary nodule classification with deep residual networks. *International journal of computer assisted radiology and surgery*, *12*(10), pp.1799-1808.
52. Ali, I., Hart, G.R., Gunabushanam, G., Liang, Y., Muhammad, W., Nartowt, B., Kane, M., Ma, X. and Deng, J., 2018. Lung nodule detection via deep reinforcement learning. *Frontiers in oncology, 8*, p.108.
53. Duarte, M.A., Pereira, W.C. and Alvarenga, A.V., 2019, September. Calculating Texture Features from Mammograms and Evaluating Their Performance in Classifying Clusters of Microcalcifications. In *Mediterranean Conference on Medical and Biological Engineering and Computing* (pp. 322-332). Springer, Cham.
54. Esteva, A., Kuprel, B., Novoa, R.A., Ko, J., Swetter, S.M., Blau, H.M. and Thrun, S., 2017. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, *542*(7639), pp.115-118.
55. Vijh, S., Sharma, S. and Gaurav, P., 2020. Brain Tumor Segmentation Using OTSU Embedded Adaptive Particle Swarm Optimization Method and Convolutional Neural Network. In *Data Visualization and Knowledge Engineering* (pp. 171-194). Springer, Cham.
56. https://wiki.cancerimagingarchive.net/display/Public/LIDC-IDRI
57. Preetha, K. and Jayanthi, S.K., 2018. GLCM and GLRLM based Feature Extraction Technique in Mammogram Images. *International Journal of Engineering & Technology, 7*(2.21), pp.266-270.
58. Ahmadi, N. and Akbarizadeh, G., 2018. Iris tissue recognition based on GLDM feature extraction and hybrid MLPNN-ICA classifier. *Neural Computing and Applications*, pp.1-15.
59. Dalal, N. and Triggs, B., 2005, June. Histograms of oriented gradients for human detection. In *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)* (Vol. 1, pp. 886-893). IEEE.
60. Dornaika, F., Moujahid, A., El Merabet, Y. and Ruichek, Y., 2017. A comparative study of image segmentation algorithms and descriptors for building detection. In *Handbook of Neural Computation* (pp. 591-606). Academic Press.
61. Anthimopoulos, M., Christodoulidis, S., Ebner, L., Christe, A. and Mougiakakou, S., 2016. Lung pattern classification for interstitial lung diseases using a deep convolutional neural network. *IEEE transactions on medical imaging*, *35*(5), pp.1207-12.
62. Ker, J., Wang, L., Rao, J., & Lim, T. (2018). Deep learning applications in medical image analysis. *Ieee Access*, *6*, 9375-9389.