JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 1, 2022, p. 526-542
https://publishoa.com
ISSN: 1309-3452

LDDC-Net: Deep Learning Convolutional Neural Network-based lung disease detection and classification

[1]N. Sudhir Reddy, [2]Dr.V Khanaa,

[1]Research Scholar, [2]Professor, Dean, Department of Information Technology

[1]Department of Computer Science and Engineering, [2]Department of Information Technology

[12]Bharath Institute of Higher Education and Research (BIHER), Chennai, Tamil Nadu

Abstract: Lung cancer is the crucial disease, which causing to millions of deaths around the globe. Therefore, the early detection and classification of lung cancers can save millions of lives. However, the conventional methods were failed to result the better classification performance. Thus, this is implemented deep learning convolutional neural network (DLCNN) model for lung disease detection and classification operations (LDDC-Net). Initially, preprocessing of Computed Tomography (CT) based lung images were performed using modified non-local trilateral filter (NLTF). Then, segmentation of lung cancer is performed using hybrid fuzzy morphological (HFM) operations, which effectively localizes the region of interest (ROI) of cancer. Further, laplacian pyramid decomposition (LPD) process applied on segmented image to extract the deep seismic features. Further, grasshopper optimization algorithm (GOA) based evolutionary model is used to select the best features. Finally, DLCNN is model performed training, testing operations using extracted features and classifies the benign, malignant lung cancers. The simulation results shows that the proposed LDDC-Net resulted in superior segmentation, classification performance as compared to conventional methods.

Keywords: Computed tomography lung images, lung disease detection and classification, non-local mean filter, hybrid fuzzy morphological segmentation, deep learning convolutional neural network.

1. Introduction

Lung cancer is the one of crucial type of cancer and millions of people are suffering with this lung cancer around the world. Recently, COVID-19 outbreak also resulted in abnormal increment of lung diseases. Further, there is direct relationship between COVID-19 and lung cancer [1], because the patients suffered with lung cancer are high chances to effected by COVID-19. Segmentation and classification of lung cancer is an essential study area, and different investigations have been undertaken [2]. Therefore, the early detection and classification of lung cancers can save millions of people.

However, the hospitals are using the traditional methods such as clinical trials for advanced stage of lung cancer detection [3]. But these traditional approaches are consuming the higher time with nominal accuracy. So, there is huge requirement to deploy the computer-Aided Design (CAD) systems in hospitals, laboratories.

The traditional CAD systems [4] are functioned using basic image processing methods without any intelligence, which resulted in poorer segmentation, classification performance. Recely, CAD systems are developed with the artificial intelligence prototypes such as machine leanring and deep learning models.

However, the conventional machine learning models are suffering with the high computational complexity with reduced classification performance. Thus, it is necessary to use the deep learning models instead of machine learning approaches in CAD systems.

In this work, we employed the DLCNN [5] for lung image classification and the morphological graph cut technique for label generation on the produced images in order to build a dataset for segmentation. It was not necessary to do any manual labelling tasks in order to produce the dataset for pretraining. In order to construct CADs [6] for lung segmentation and classification, computer vision and medical imaging technologies are used in conjunction with each other. Specifically in the field of lung disease, lung parenchyma segmentation is used as a pre-processing step in lung CT image processing, which is very useful.

The pre-processing [7] processes have a direct influence on the ultimate CT scan preparation. Subsequently, speedier and more accurate segmentation procedures for lung CT scans are a fascinating problem deserving of investigation, having urgent realistic necessity and clinical value. Multiple lung division procedures have been investigated, and a fraction of the conventional approaches integrate threshold region-development techniques [8]. Regardless, the results are not very promising, and the procedure is time-consuming and tedious. Therefore, it is still considered uncharted area.

Despite the fact that the division is completed quickly, the quality of the division is poor because the estimated dimensions of the lung borders are the same as those of the windpipe and bronchus area. Deep learning [9] is a fundamental image-segmentation approach that is reliant on the area of the CT scan. It has the ability to quickly and efficiently separate the interstitial lung borders. While this approach is effective, it is time-consuming, and the developing model is sensitive to boundary conditions. The majority of lung segmentation frameworks now in use are hybrid systems that include an edge-respecting technique as well as unanticipated development and other extraction processes. In addition, a variety of exams are performed on the basis of the division of the lung parenchyma in the presence of a lung infection. Thus, the major contributions of this article are illustrated as follows:

- Initially, NLTF is applied to remove the different types of noises from lung CT images and also enhances the cancer region.
- Then, HFM based segmentation is applied to localize the ROI of cancer with morphological opening and closing based fuzzy operations.
- The disease specific features are extracted using LPD process applied on segmented image and GOA is used to extract the deep seismic features.
- Finally, DLCNN is model performed training, testing operations using extracted features and classifies the benign, malignant lung cancers.

Rest of the article is contributed as follows: Section 2 deals with the related work with their problem statement. Section 3 deals with the detailed analysis of proposed LDDS-Net. Section 4 deals with results and discussion and comparison with conventional approaches. Section 5 concludes the article with the possible future enhancements.

2. Literature survey

There are a variety of medical imaging modalities available, each with its own set of unique characteristics. This also contributes to the advancement of processing methods. Multi-scale CT scan classification is widely utilized approach in the scientific community. This procedure is used to combine many medical photographs into a single image. There multiple research surveys [10] are conducted on CT scan-based lung cancer detection with classification. The survey shows that the machine learning models are reduced the classification accuracy and resulted in improper classification.

In [11] authors offered a number of evolutionary methods for lung segmentation with mean filtering-based preprocessing operation. Four algorithms with improved quality were used to the pre-processed CT scans in order to improve their overall quality. In order to ensure realistic findings for 20 sample lung scans, MATLAB was utilized to test the results. The computational complexity [12] of this approach still needs to be reduced. Furthermore, in [13] authors worked on lung cancer diagnosis utilizing improved Median filter based denoising with threshold segmentation, enhanced the accuracy of the preprocessing by Gaussian filter [14] of these photos in this research, which led to the development of an algorithm. This approach has much higher sensitivity, specificity, precision, and accuracy than previous methods, with a reduced percentage of false positives.

In [15] authors focused on implementation of bilateral filter-based preprocessing with random forest classification. Here, image classification is the combination of relevant data from many input CT scans into a single, clarifying segmented image that is then shown. As a result, CT scan preprocessing is performed using trilateral filter [16] based preprocessing in order to achieve noise free image. Lung cancer classification techniques may be classified into two categories based on available features such as spatial-domain based image classification methods and transform-domain based image preprocessing methods such as non-local means (NLM) [17].

Multi-scale image segmentation and classification methods [18] such as pyramid-based classification are now in use, as well as a number of additional multi-scale image classification algorithms. Each of the strategies described in [19] is relevant information from changing parameters. Detailed relevant information is gathered along with the segmented CT scan when the derived details are segmented together as a result of measuring the parameters. In [20] authors used the basic morphological approaches for CT image segmentation, which resulted in considerable performance for benign diseases, but failed to segment the malignant diseases. Image segmentation applications such as adaptive thresholding [21] including as super-resolution processing, denoising, and segmentation are increasingly routinely based on the morphological theory.

In recent years, the SR theory has garnered a great deal of interest in the area of image processing, particularly in the context of image classification, and with good reason [22]. We all know that the dictionary creation of standard SR algorithms may be accomplished using Fuzzy c-means clustering (FCM) [23], and Modified OTSU (MOTSU) [24] for lung image segmentation. However, traditional SR algorithms that use a fixed vocabulary have a number of limitations when it comes to CT scan classification. In [25] proposed SVM model for both image classifications and denoising, which can adaptively generate a compressed dictionary for the classification of CT scans by combining images from different sources. According to researchers [26], the modified spatial frequency with k-means clustering (KMC) used for lung image segmentation, and the fundamental notion of an adaptive selection dictionary can be introduced to SR by this method.

On the basis of artificial intelligence models with two redundant wavelet transformations such as Redundant Wavelet Transform (RWT) based support vector machine (SVM) [27] and Redundant Discrete Wavelet Transform (R-DWT) based convolutional neural network (CNN) [28] used for lung cancer classification approach with multi-view medical imaging. Using their suggested technique, they discovered that the shift-invariance of the R-DWT method may be used to construct high-quality CT scan classifications with little effort. In order to achieve the classification of many views of a CT scan, a method known as pyramid transformation may be performed [29]. When it was initially presented, this approach was quickly embraced and became widely employed in a variety of applications, including computer vision, image compression, and CT scan segmentation [30]. The pyramid transform is now widely used to merge multiple-view clinical CT scans, and it is becoming more popular.

In [31] authors presented the union ResNet51 approach for multi class lung cancer classification, which allowed them to extract a large number of essential characteristics [32] from the segmented CT scans. In [33] authors proposed Region Mosaicking on Laplacian Pyramids for feature extraction with GoogleNet as a method for fusing CT scans collected by a microscope, however it was shown to be susceptible to noise. After that, the laplacian pyramid method with joint averaging was proposed, which resulted in an effective improvement in the output by revealing the image's rich background features.

In [34] authors published their findings on a revolutionary multi-modal medical CT scan classification method. A serial canonical correlation-based classification of numerous texture, point, and geometric features is conducted on the contrast CT scans. In [35], authors implemented the basic mean filter-based preprocessing with U-Net based segmentation. Further, machine learning based SVM is used to classify the benign and malignant classes of lung. But this method consuming the higher time for training.

From the survey, it is observed that the conventional methods are failed to segment the cancer effected region accurately, which resulted in reduced classification. The machine learning models are unable to differentiate the benign and malignant disease specific features,

3. Proposed Method

This section gives the detailed analysis of proposed LDDS-Net and the proposed model performs four major operations such as NLTF based preprocessing, HFM segmentation, LPD based feature extraction, GOA based feature selection, and

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 1, 2022, p. 526-542
https://publishoa.com
ISSN: 1309-3452

DLCNN classification. Figure 1 presents the proposed LDDS-Net framework and Table 1 presents the algorithm of proposed LDDS-Net.

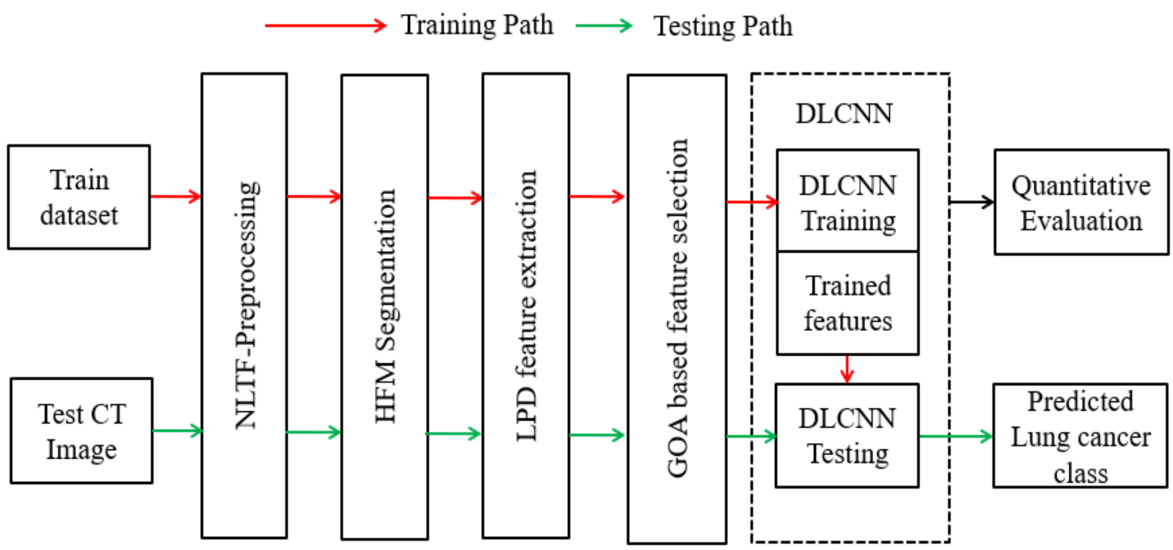


Figure 1. Proposed LDDS-Net model.

Table 1. Proposed LDDS-Net model.

Input: LIDC-IDRI dataset, test CT image **Output:** Segmented image, classified lung cancer class. **Performance measures:** Preprocessing, segmentation, classification metrics.	
Step 1:	Consider the LIDC-IDRI training dataset and apply it to NLTF preprocessing for noise removal, enhancement of CT images.
Step 2:	Apply HFM segmentation for localization of lung cancer effected region.
Step 3:	Appl LPD on HFM segmented outcomes, which extracts the detailed disease specific features.
Step 4:	Apply GOA based bio-optimization approach for selection of optimal features.
Step 5:	Perform DLCNN based training and extracts the trained features and evaluate the segmentation, classification performance.
Step 6:	Consider the test CT image and repeat the step 1 to step5 for extracting the test features.
Step 7:	Apply the DLCNN model for testing, which identifies the lung cancer diseases through prediction.
Step 8:	Evaluate the preprocessing, segmentation, classification metrics and compare with state of art approaches.

3.1 NLTF Preprocessing

In order to maintain the edge of smoothing and visible detail for N-dimensional signaling in computer graphics, image processing and computer vision applications, a high contrast image and mail-in trilateral filter has been introduced as a nonlinearly single-pass filter. In order to identify noisy pixels in random images affected by impulses noise the trilateral filter is used to insert the local image statistics into the BF. A NLTF value quantifies the difference in intensity between a center pixel and its closest comparable neighbors. For reducing Gaussian noise and uniform noisy impulses and mixing, the trilateral filter is suitable. A pipeline architecture that aims to reduce the complexity problem of traditional trilateral filters with trilateral noise filters. Its very high temporal complexity is the major difficulty of the application of such a strong noise filter on real time imaging systems. The NLTF computation is bitwise performed and the exponential function evaluation is partly performed using linear approximations as shown in Figure 2.

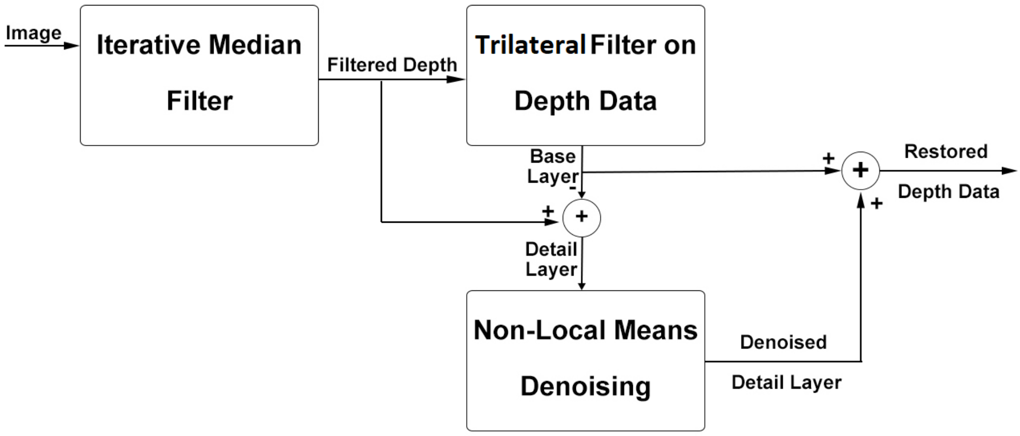


Figure 2. NLTF block diagram.

An adaptive trilateral filter by employing NLTF and Rank Ordered Logarithmic Difference statistic was developed by using trilateral filters. The NLTF is already a decent statistic on pulse noise in imagery; however certain bruising pixels may possess intensity values near their surrounding pixels for randomly valued pulse noise. In such a scenario, the pixel NLTF cannot be large enough to differentiate between it and the free pixels of noise. The logarithmic function is called ROLD function, to enhance the input image.

3.2 HFM segmentation

It is quite beneficial in detecting and assessing surrounding lesions when lung parenchyma segmentation is performed, but it is only effective if particular approaches and frameworks are used. lung parenchyma segmentation is an essential pre-processing step in the design of lung nodules from CT image sequences, which is used in the CAD system. When developing the LDDS-Net, an effective thresholding strategy was employed to lower the complexity of lung segmentation in order to reduce the computing time while simultaneously improving accuracy. The technique was tested on a number of CT scans obtained from the LIDC-IDRI with the aid of experimentation and data analysis. Figure 3 depicts a flowchart of the suggested HFM segmentation approach, which is divided into three stages. The suggested HFM segmentation approach is also detailed in Algorithm 2, which is shown in Table 2.

Let $A(x,y)$ represent the preprocessed CT lung scans that was used as input. The segmentation of the lungs was accomplished using the adaptive global threshold, which was applied to a portion of the lung segment from the CT image that had been intensely threshold. Then, the CT image histogram was used to determine the value of the threshold, which was used to generate the output.

$$A^{\delta}(x,y) = \begin{cases} 1, & if\ A(x,y) \geq \sigma \\ 0, & if\ A(x,y) < \sigma \end{cases} \quad (1)$$

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 1, 2022, p. 526-542
https://publishoa.com
ISSN: 1309-3452

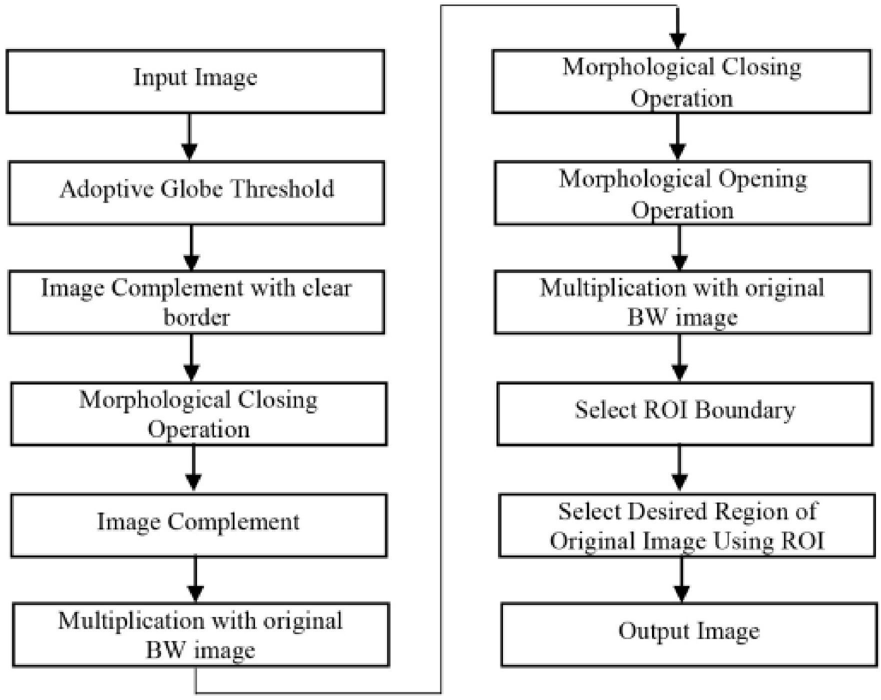


Figure 3. Flowchart of the proposed segmentation method.

The precise global threshold levels generated from $A(x,y)$ is denoted by the symbol σ. Following the application of thresholding to $A(x,y)$, we have the resulting CT scan, which is represented by $A(x,y)$. We now have the CT scan counterpart to the clear border, as demonstrated in the example below:

$$A^{\alpha}(x,y) = C - A^{\delta}(x,y) \quad (2)$$

In this case, C symbolizes a CT scan in which all of the pixel values are equal to one. The intermediate output image is denoted by $A^{\alpha}(x,y)$, which contains different borders of CT images. Now, using the mask B, a morphological closure operation is conducted on $A^{\delta}(x,y)$ in order to acquire the result $A^{\beta}(x,y)$.

$$A^{\beta}(x,y) = A^{\alpha}(x,y) \bullet B \quad (3)$$

Taking the complement of the previous sentence,

$$A^{\gamma}(x,y) = C - A^{\beta}(x,y) \quad (4)$$

Perform the bit wise multiplication operation between $A^{\gamma}(x,y)$ and binary version of input image ($A^{\delta}(x,y)$).

$$A^{\tau}(x,y) = A^{\gamma}(x,y)A^{\delta}(x,y) \quad (5)$$

Table 2. Proposed HFM segmentation algorithm.

Input: A (x, y) **Output:** µ (x, y)
1: Start: 2: $[M,N] \rightarrow$ size $(A\ (x,y))$ 3: C ← ones (M, N) 4: Initialize the structure element (B) 5: if $A\ (x,y)\ \geq\ \sigma$ then

```
A^δ(x,y) ← 1
else
A^δ(x,y) ← 0
end if
Image complement: A^α(x,y) ← C − A^δ(x,y)
Closing operation: A^β(x,y) ← A^α(x,y) • B
Image complement: A^γ(x,y) ← C − A^β(x,y)
Multiplication: A^τ(x,y) ← A^γ(x,y)A^δ(x,y)
Closing operation: A^θ(x,y) ← A^τ(x,y) • B
opening operation: A^ω(x,y) ← A^θ(x,y) ∘ B
μ(x,y) ← A^ω(x,y)A^α (x,y)
end procedure
```

Apply the morphological closing operation on $A^{\tau}(x,y)$ with respect to non-overlapping B values, which generates outcomes as $A^{\theta}(x,y)$.

$$A^{\theta}(x,y) = A^{\tau}(x,y) \bullet B \quad (6)$$

Apply the morphological opening operation on $A^{\theta}(x,y)$ with respect to repeated non-overlapping B values, which generates outcomes as $A^{\omega}(x,y)$.

$$A^{\omega}(x,y) = A^{\theta}(x,y) \circ B \quad (7)$$

Finally, perform the bit wise multiplication operation between $A^{\omega}(x,y)$ and $A^{\alpha}(x,y)$, which generates the segmented outcome as $\mu(x,y)$.

$$\mu(x,y) = A^{\omega}(x,y)A^{\alpha}(x,y) \quad (8)$$

3.3. Feature extraction

The LPD approach was used for feature extraction from the segmented source CT scan in order to acquire the characteristics from $\mu(x,y)$ at varied sizes. The LPD used to decomposed the image into three stages and extracts the low-low (LL) frequency bands as features. To get started, we'll need the Gaussian pyramid of a CT scan with a size of M N pixels. The P_0 layer contains the CT scan that was used as a source. The image (x, y) from layer P_0 was down-sampled with the use of the laplacian kernel in order to generate the P_1 layer (0.5Mx 0.5N) and the P_2 layer (0.5M x 0.5N). The LPD was constructed by repeating the methods outlined above. Figure 4 depicts the three-stage disintegration of an LP in its natural state. The following equation may be used to represent the breakdown of the (l - 1)th layer P_{l-1} into the lth layer P_l:

$$P_l = \Downarrow (P_{l-1}) \quad (9)$$

To create the LP, the first step is to sample each layer of the Gaussian pyramid, which is done in layers. In terms of size, the CT scan to be expanded is deemed to be m n in size. It was necessary to utilize an inverse Gaussian pyramid in order to convert the CT scan into a 2m 2n image, which can be expressed as:

$$P_l^* = \Uparrow (P_l) \quad (10)$$

Now, the feature extraction from $\mu_{fe}(x,y)$ may be accomplished using the Laplacian Pyramid layers, as demonstrated in the following example:

$$\mu_{fe}(x,y) = \begin{cases} P_0, \; for\; l = 0 \\ P_l^* + LP_l, \; for\; 0 < l < L \\ P_l = LP_l, \; for\; l = L \end{cases} \quad (11)$$

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 1, 2022, p. 526-542
https://publishoa.com
ISSN: 1309-3452

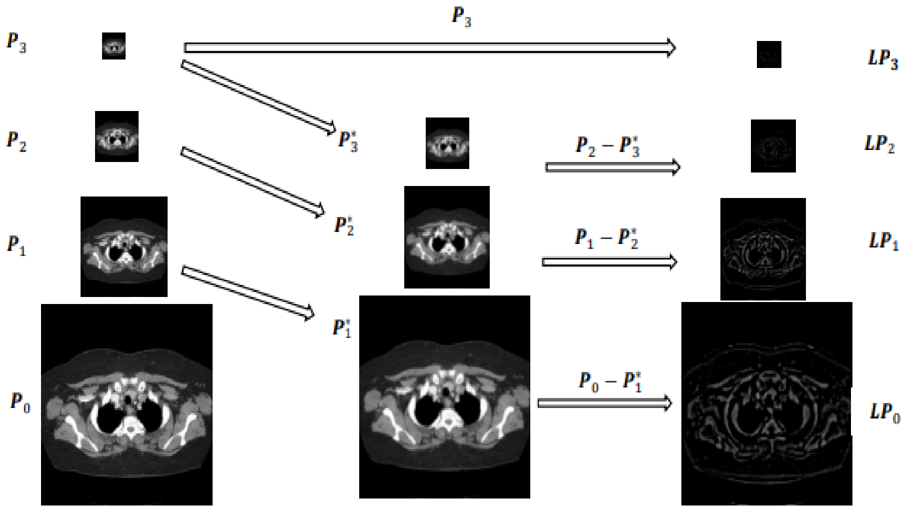


Figure 4. Low level feature extraction using LPD.

3.4 Feature selection

The GOA is a search evolutionary system that mimics grasshopper swarming behavior in the environment. The LDDS-Net is used to select the best features using GOA based bio-optimization features from LPD extracted features as shown in Figure 5. The grasshopper swarm is a characteristic feature that can be noticed in both fairy and real life.

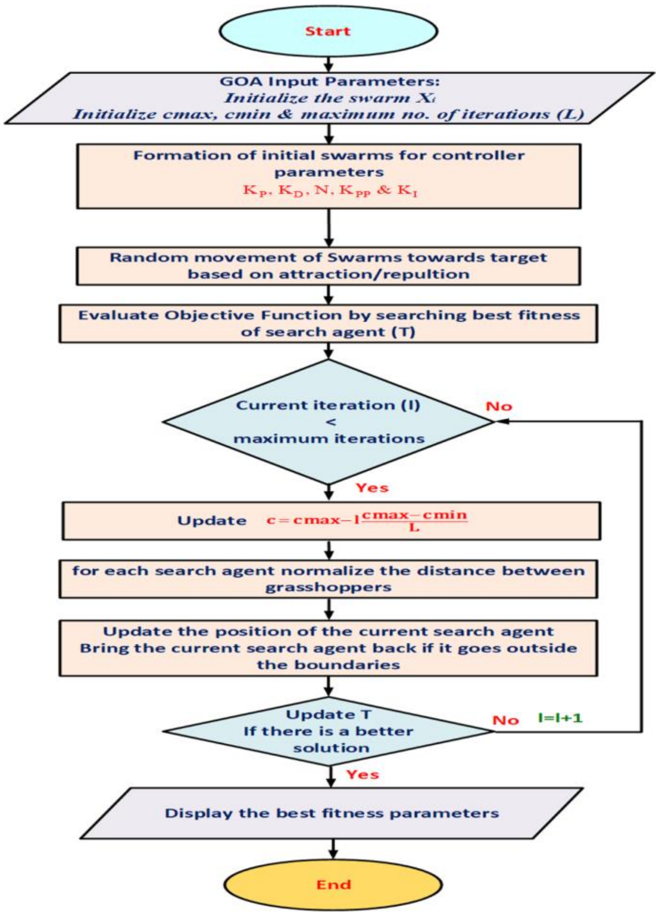


Figure 5. GOA optimization process

In equation 12, the mathematical equation for describing grasshopper swarming behavior is defined.

$$x_i = p_1 Q_i + p_2 G_i + p_3 P_i \qquad (12)$$

Here, i^{th} is the grasshopper Position, social interaction, and gravitational force are represented by chi-square, Si, and Fi, respectively (where p1, p2, & p3 are random numbers). Further, equation 13 is used to calculate the value of Si.

$$S_i = \sum_{k=1,j\neq}^{m_p} t(|\lambda_j - i|)\vec{d}_{jk} \qquad (13)$$

The term $\sum_{k=1,j\neq 0}^{m_p} t(|\lambda_j - i|)\vec{d}_{jk}$ indicates the mean distance between i^{th} & j^{th} grasshoppers; $\vec{d}_{jk}$ is the unit vector between i^{th} and j^{th} grasshoppers. 'ηp' indicates the total number of grasshoppers, and 'S ' represents the strength of the social force of interaction, which is determined using equation 14 as follows:

$$S(r) = \psi e^{-r/l} - e^{-r} \qquad (14)$$

Here, ψ and l specify the intensity of attraction and scale length, respectively. Further, it can be used to determine the remaining two components as follows:

$$\begin{cases} G_i = -g\vec{e}_g \\ W = u\vec{e}_\omega \end{cases} \qquad (15)$$

The gravitational constant is $\vec{e}_g$, and the constant drift is l. The unit vectors to the earth's center and to the direction of the wind are g and $\vec{e}_\omega$, respectively. As a result, equation 15 can be written as:

$$x_i = p_1\left(\sum_{k=1,j\neq i}^{k} s(|\lambda_k - \eta_j|)d_{ij}\right) \qquad (16)$$

Further, Equation 16 is further changed to improve the convergence rate of GOA:

$$x_k = a\left(\sum_{j=1,j\neq i}^{mp} b\,\frac{ta_{dim} - lb_{dim}}{2}\, s(|\mu_k^{dim} - \mu_i^{dim}|)d_{ij}\right) + \lambda_{dim} \qquad (17)$$

Here, d_{im} denotes the dimension of a distinct control variable; $u_{dim} and\ b_{dim}$ are upper and l

ower bounds, respectively, and b is a decreasing coefficient represented as follows:

$$a = 1 - [l(a_{max} - b_{max})/L \qquad (18)$$

According to equation 18, a grasshopper's new updated position is determined using the current position as well as the next position of all other grasshoppers. The first term 'b' is nonlinear, which is analogous to internal weight in PSO, minimizes grasshopper momentums around the target, resulting in a balance between exploration and exploitation. The term $[\lambda_{dim}]$ enables GOA to reduce the search space for exploration and exploitation linearly.

3.5 Classification

The DLCNN models are widely used in many medical image processing applications including multi class lung cancer classification. It is a kind of efficient identification approach that has lately gotten a lot of attention because of its effectiveness. The advantage of DLCNNs is that they are simpler to train and have a much smaller number of parameters than fully linked networks.

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 1, 2022, p. 526-542
https://publishoa.com
ISSN: 1309-3452

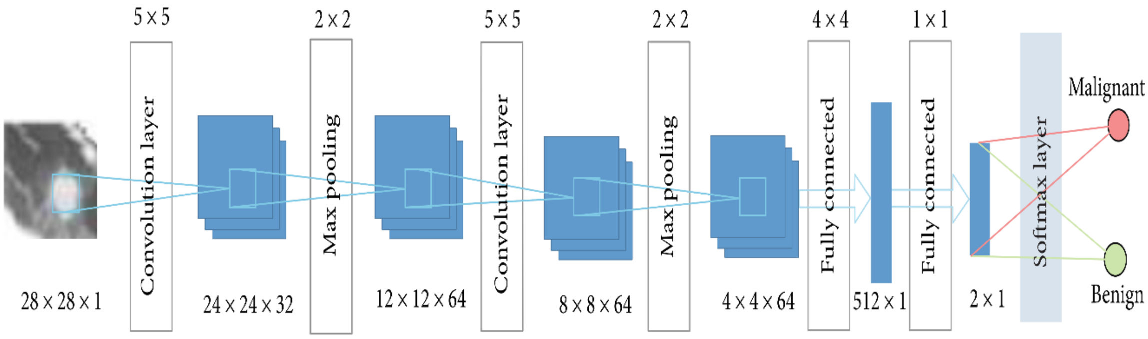


Figure 6. DLCNN classification

Table 3. DLCNN later Information

Layer	Type	Input	Kernel	Output
1	Convolution	$28 \times 28 \times 1$	5×5	$24 \times 24 \times 32$
2	Max pooling	$24 \times 24 \times 32$	2×2	$12 \times 12 \times 64$
3	Convolution	$12 \times 12 \times 64$	5×5	$8 \times 8 \times 64$
4	Max pooling	$8 \times 8 \times 64$	2×2	$4 \times 4 \times 64$
5	Fully connected	$4 \times 4 \times 64$	4×4	512×1
6	Fully connected	512×1	1×1	2×1
7	Softmax	2×1	N/A	Result

Figure 6 presents the detailed layer wise architecture of proposed DLCNN model and Table 6 presents the properties of each layer of DLCNN model. Here, convolution layer is used to extract the deep features using weighted kernel. There are two primary sources of error in feature extraction: the neighborhood size constraint produced and the estimated error in the convolution layer parameter estimation generated by the mean deviation. When using mean pooling, the first mistake may be reduced while still keeping more CT scan background information. With maximum pooling, it is possible to lower the second error while maintaining more texture information.

Then, max pooling layer is used to reduce the number features by selecting the best value in each kernel. The convolution and max-pooling layers are repeated multiple times to generate the best features with low computational complexity. Further, fully connected layer is used to map the input to output features by maintaining all the neuron interconnections. Finally. SoftMax classifier is used to classify the benign and malignants classes from test features.

4. Results and discussions

This section gives the detailed analysis of simulation results with comparisons using state of art approaches. The proposed LDDC-Net and conventional methods utilized the same dataset for implementations. Further, the ablation study of the proposed LDDC-Net also conducted to measure the superiority of each method.

4.1 Datasets

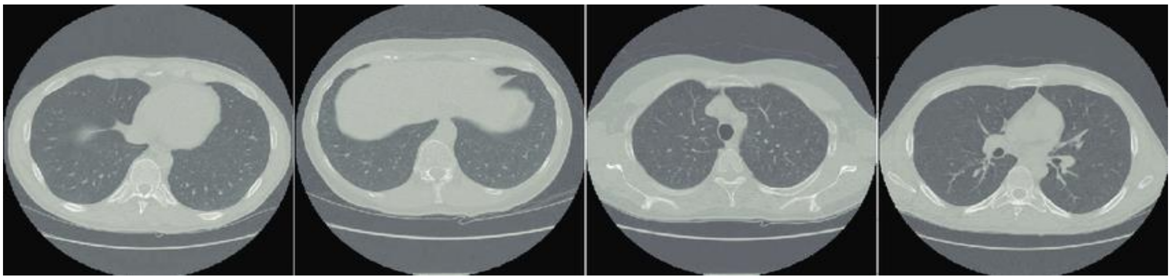
The image collection of the Lung Image Database Consortium (LIDC-IDRI) comprises of diagnostic and screening thoracic CT images with marked-up annotated lesions. It is a web-based worldwide resource for the development, training, and assessment of CAD approaches for the detection and diagnosis of lung cancer. This public-private partnership, initiated by the National Cancer Institute (NCI), advanced by the Foundation for the National Institutes of

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 1, 2022, p. 526-542
https://publishoa.com
ISSN: 1309-3452

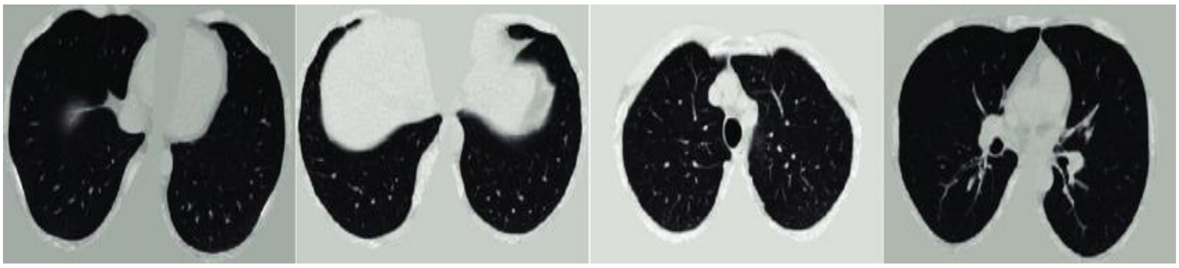
Health (FNIH), and accompanied by the Food and Drug Administration (FDA) through active participation, demonstrates the success of a consortium founded on a consensus-based process. This data collection, which includes 1018 instances, was created in collaboration with seven academic institutions and eight medical imaging firms. Each topic contains pictures from a clinical thoracic CT scan as well as an XML file with the findings of a two-phase image annotation procedure completed by four experienced thoracic radiologists. During the first blinded-read phase, each radiologist independently assessed each CT image and labelled lesions as "nodule > or =3 mm," "nodule 3 mm," or "non-nodule > or =3 mm." During the ensuing unblinded-read phase, each radiologist independently assessed their own markings, as well as the anonymized marks of the three other radiologists, to form a final judgement. The purpose of this procedure was to detect as many lung nodules as feasible on each CT scan without necessitating forced consensus.

4.2 Subjective analysis

This section shows the preprocessing and segmentation based visual subjective analysis. Figure 7 shows the NLTF preprocessing output images, which effectively eliminated the different types of background noises and enhances the cancer presented region. Further, NLTF also improves the pixel spatial, texture properties. Figure 8 shows the segmented output images using HFM approach. Further, first, third columns represent the input images and second, fourth columns represent the segmented images. The proposed HFM method accurately segments the cancer region.



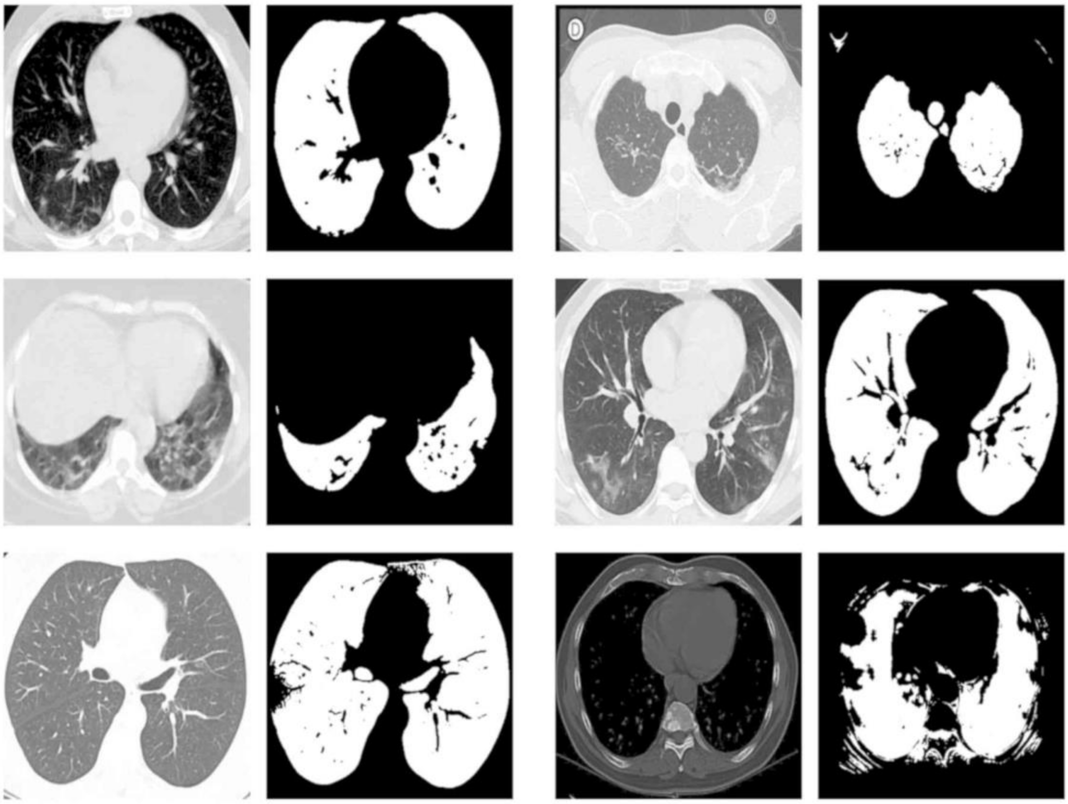
(a) Input CT images



(b) NLTF preprocessed images

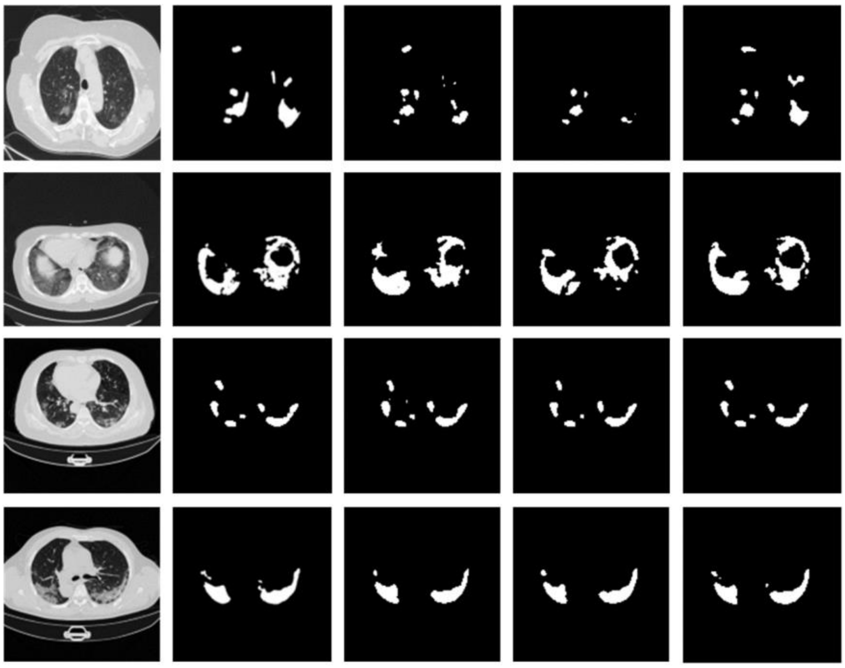
Figure 7. Preprocessing using NLTF

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 1, 2022, p. 526-542
https://publishoa.com
ISSN: 1309-3452



(a) (b) (c) (d)

Figure 8. HFM based segmented output images.



(a) (b) (c) (d) (e)

Figure 9. CT image segmented images using various approaches, (a) input, (b) KMC [26] (c) FCM [23], (d) MOTSU [24], and (e) proposed HMF.

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 1, 2022, p. 526-542
https://publishoa.com
ISSN: 1309-3452

Figure 9 shows the lung CT image segmentation visual performance comparison with conventional approaches. The proposed HFM segmentation method resulted in superior localization of cancer effected region as compared to conventional image processing methods such as KMC, FCM, and MOTSU. The conventional methods resulted in poorer localization of cancer region, which caused to reduce the classification performance.

4.3 Objective performance

This section gives the detailed analysis of subjective analysis of proposed LDDC-Net with conventional models using various metrics such as peak signal to noise ratio (PSNR), structural similarity index metric (SSIM), mean square error (MSE), entropy, standard deviation (STD), and mutual information (MI).

Table 4. Preprocessing methods performance comparison

Method	**PSNR**	**SSIM**	**MSE**	**Entropy**	**STD**	**MI**
Mean filter [11]	48.42	0.236	0.0748	5.33	0.966	3.91
Median filter [13]	51.80	0.259	0.0646	5.70	0.955	4.40
Gaussian filter [14]	52.93	0.358	0.0374	5.74	0.720	4.64
Bilateral filter [15]	53.67	0.412	0.0137	6.70	0.539	6.75
Trilateral filter [16]	54.41	0.604	0.0098	7.26	0.473	6.90
NLM [17]	54.45	0.873	0.0076	8.11	0.351	7.47
Proposed NLTF	55.72	0.989	0.0039	9.71	0.076	8.54

Table 4 compares the performance of proposed NLTF approach with conventional approaches like Mean filter [11], Median filter [13], Gaussian filter [14], Bilateral filter [15], Trilateral filter [16], and NLM [17]. Further, the proposed NLTF resulted in superior performance for all metrics as compared to all existing methods. Figure 10 represents graphical representation of preprocessing methods performance comparison.

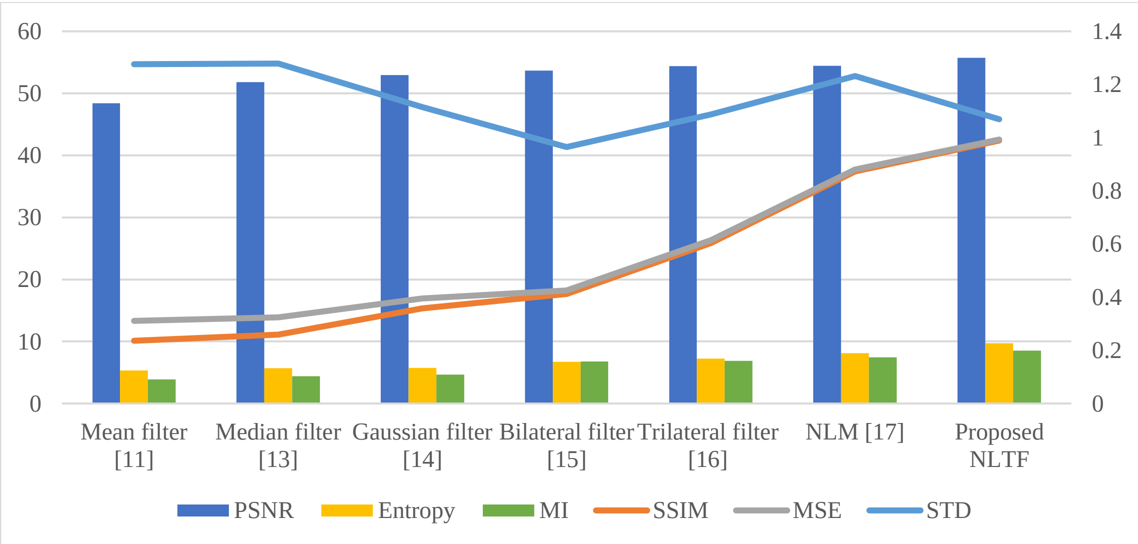


Figure 10. Graphical representation of preprocessing methods performance comparison.

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 1, 2022, p. 526-542
https://publishoa.com
ISSN: 1309-3452

Table 5. Segmentation methods performance comparison.

Method	**SACC**	**SSEN**	**SSPE**	**SF1M**	**SRE**	**SPR**
Morphological [20]	92.41	90.31	90.13	90.10	92.49	90.21
Adaptive Thresholding [21]	92.62	93.22	91.66	92.19	93.13	90.27
FCM [23]	93.25	93.64	92.72	93.18	93.66	94.48
MOTSU [24]	94.44	94.81	94.16	93.36	93.83	96.89
KMC [26]	95.69	95.68	94.99	96.42	94.56	97.24
Proposed HFM	97.53	96.66	97.90	97.93	97.80	97.96

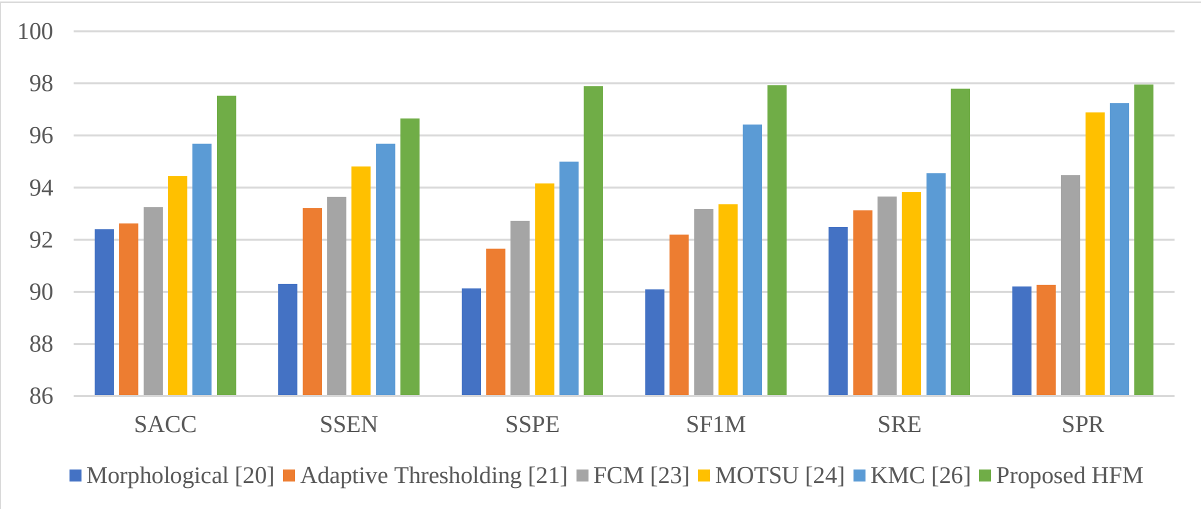


Figure 11. Graphical representation of segmentation methods performance comparison.

Table 5 compares the proposed HFM segmentation performance with conventional approaches like Morphological [20], Adaptive Thresholding [21], FCM [23], MOTSU [24], and KMC [26]. Further, the performance of proposed HFM is superior in terms of all performance metrics such as segmentation accuracy (SACC), segmentation sensitivity (SSEN), segmentation specificity (SSPE), segmentation F1-measure (SF1M), segmentation recall (SRE), and segmentation precision (SPR). Figure 11 represents graphical representation of segmentation methods performance comparison.

Table 6 compares the proposed LDDC-Net performance comparison with conventional approaches like SVM [27], CNN [28], AlexaNet [30], ResNet51 [31], GoogleNet [33], and U-Net with SVM [35]. Further, the performance of proposed LDDC-Net is superior in terms of all performance metrics such as classification accuracy (CACC), classification sensitivity (CSEN), classification specificity (CSPE), classification F1-measure (CF1M), classification recall (CRE), classification precision (CPR), and classification area under curve (CAUC). Figure 12 represents the graphical representation of classification methods performance comparison.

Table 6. Classification methods performance comparison.

Method	**CACC**	**CSEN**	**CSPE**	**CF1M**	**CRE**	**CPR**	**CAUC**
SVM [27]	88.30	89.21	88.02	88.06	89.78	88.35	88.49
CNN [28]	88.33	91.53	92.90	90.91	89.87	89.54	88.83
AlexaNet [30]	90.35	92.44	93.41	91.32	91.13	89.70	91.84
ResNet51 [31]	93.83	93.08	94.58	91.59	92.93	92.12	93.28
GoogleNet [33]	94.79	95.68	94.80	92.97	94.84	95.12	93.30
U-Net with SVM [35]	96.00	95.81	94.88	93.83	96.00	96.03	96.18
Proposed LDDC-Net	97.11	97.11	95.47	97.78	97.77	97.83	97.64

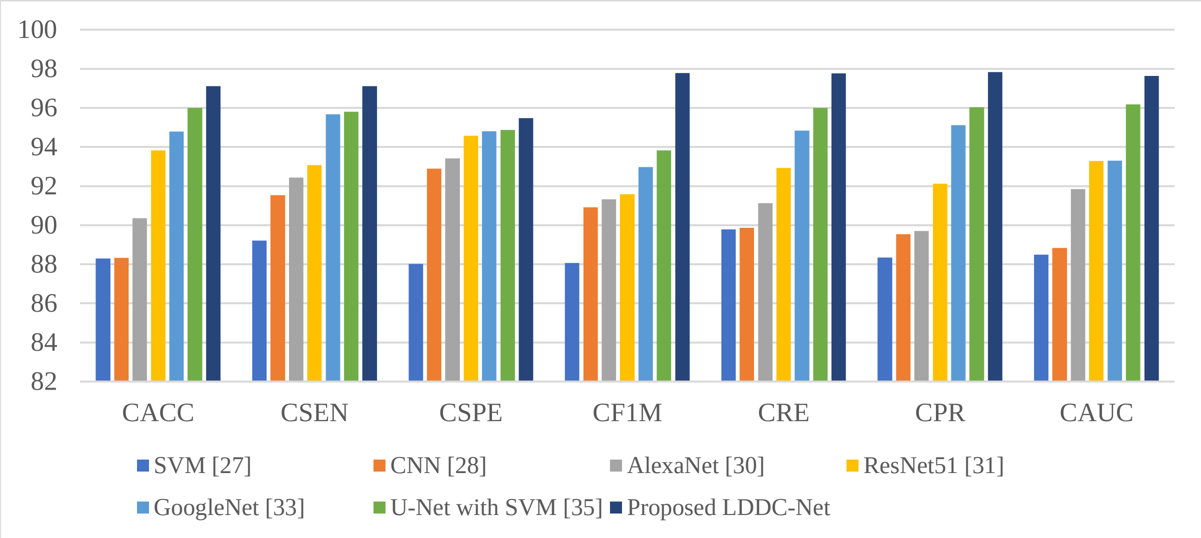


Figure 12. Graphical representation of classification methods performance comparison.

5. Conclusion

This article proposed a deep learning and bio-optimization based LDDC-Net model for segmentation and multi class classification of lung cancers. Initially, NLTF was used to reduce the different types of noises from CT source images and also enhanced the disease effected region. Then, HFM based segmentation was used to localize the disease effected region. Further, LPD was used to extract the deep seismic features from the segmented images. In addition, GOA based bio-optimization approach was used to select the deep interdepend features with disease specific probability dependent features. Finally, DLCNN model was used to classify the benign and malignant classes of CT lung images using optimal features. The simulation results shows that the proposed LDDC-Net resulted in superior preprocessing, segmentation, classification performance as compared to conventional approaches. This work can be extended with other optimization methods and transfer learning models for segmentation, classification.

References

[1]. Sharif, Muhammad Irfan, et al. "A comprehensive review on multi-organs tumor detection based on machine learning." *Pattern Recognition Letters* 131 (2020): 30-37.
[2]. Shakeel, P. Mohamed, M. A. Burhanuddin, and Mohammad Ishak Desa. "Automatic lung cancer detection from CT image using improved deep neural network and ensemble classifier." *Neural Computing and Applications* (2020): 1-14.
[3]. Asuntha, A., and Andy Srinivasan. "Deep learning for lung Cancer detection and classification." *Multimedia Tools and Applications* 79.11 (2020): 7731-7762.
[4]. Yu, Kun-Hsing, et al. "Reproducible machine learning methods for lung cancer detection using computed tomography images: Algorithm development and validation." *Journal of medical Internet research* 22.8 (2020): e16709.
[5]. Bhatia, Siddharth, Yash Sinha, and Lavika Goel. "Lung cancer detection: a deep learning approach." *Soft Computing for Problem Solving*. Springer, Singapore, 2019. 699-705.
[6]. Schwyzer, Moritz, et al. "Automated detection of lung cancer at ultralow dose PET/CT by deep neural networks–initial results." *Lung Cancer* 126 (2018): 170-173.
[7]. Lakshmanaprabu, S. K., et al. "Optimal deep learning model for classification of lung cancer on CT images." *Future Generation Computer Systems* 92 (2019): 374-382.
[8]. Ardila, Diego, et al. "End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography." *Nature medicine* 25.6 (2019): 954-961.
[9]. Riquelme, Diego, and Moulay A. Akhloufi. "Deep learning for lung cancer nodules detection and classification in CT scans." *AI* 1.1 (2020): 28-67.

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 1, 2022, p. 526-542
https://publishoa.com
ISSN: 1309-3452

[10]. Chaunzwa, T. L., Hosny, A., Xu, Y., Shafer, A., Diao, N., Lanuti, M., ... & Aerts, H. J. (2021). Deep learning classification of lung cancer histology using CT images. *Scientific reports*, *11*(1), 1-12.
[11]. Lee, J. H., Sun, H. Y., Park, S., Kim, H., Hwang, E. J., Goo, J. M., & Park, C. M. (2020). Performance of a deep learning algorithm compared with radiologic interpretation for lung cancer detection on chest radiographs in a health screening population. *Radiology*, *297*(3), 687-696.
[12]. Toğaçar, Mesut, Burhan Ergen, and Zafer Cömert. "Detection of lung cancer on chest CT images using minimum redundancy maximum relevance feature selection method with convolutional neural networks." *Biocybernetics and Biomedical Engineering* 40.1 (2020): 23-39.
[13]. Bhandary, Abhir, et al. "Deep-learning framework to detect lung abnormality–A study with chest X-Ray and lung CT scan images." *Pattern Recognition Letters* 129 (2020): 271-278.
[14]. Avanzo, M., Stancanello, J., Pirrone, G., & Sartor, G. (2020). Radiomics and deep learning in lung cancer. *Strahlentherapie und Onkologie*, *196*(10), 879-887.
[15]. Huang, P., Lin, C. T., Li, Y., Tammemagi, M. C., Brock, M. V., Atkar-Khattra, S., ... & Lam, S. (2019). Prediction of lung cancer risk at follow-up screening with low-dose CT: a training and validation study of a deep learning method. *The Lancet Digital Health*, *1*(7), e353-e362.
[16]. Singh, Gur Amrit Pal, and P. K. Gupta. "Performance analysis of various machine learning-based approaches for detection and classification of lung cancer in humans." *Neural Computing and Applications* 31.10 (2019): 6863-6877.
[17]. Nasrullah, N., Sang, J., Alam, M. S., Mateen, M., Cai, B., & Hu, H. (2019). Automated lung nodule detection and classification using deep learning combined with multiple strategies. *Sensors*, *19*(17), 3722.
[18]. Jakimovski, G., & Davcev, D. (2019). Using double convolution neural network for lung cancer stage detection. *Applied Sciences*, *9*(3), 427.
[19]. Masood, Anum, et al. "Computer-assisted decision support system in pulmonary cancer detection and stage classification on CT images." *Journal of biomedical informatics* 79 (2018): 117-128.
[20]. Wang, Shuo, et al. "Predicting EGFR mutation status in lung adenocarcinoma on computed tomography image using deep learning." *European Respiratory Journal* 53.3 (2019).
[21]. Kadir, T., & Gleeson, F. (2018). Lung cancer prediction using machine learning and advanced imaging techniques. *Translational lung cancer research*, *7*(3), 304.
[22]. Ruan, Jingru, et al. "Development of deep learning-based automatic scan range setting model for lung cancer screening low-dose CT imaging." *Academic Radiology* (2022).
[23]. Punithavathy, K., Sumathi Poobal, and M. M. Ramya. "Performance evaluation of machine learning techniques in lung cancer classification from PET/CT images." *FME Transactions* 47.3 (2019): 418-423.
[24]. Polat, Huseyin, and Homay Danaei Mehr. "Classification of pulmonary CT images by using hybrid 3D-deep convolutional neural network architecture." *Applied Sciences* 9.5 (2019): 940.
[25]. Thakur, Shailesh Kumar, Dhirendra Pratap Singh, and Jaytrilok Choudhary. "Lung cancer identification: a review on detection and classification." *Cancer and Metastasis Reviews* 39.3 (2020): 989-998.
[26]. Park, S., Lee, S. M., Do, K. H., Lee, J. G., Bae, W., Park, H., ... & Seo, J. B. (2019). Deep learning algorithm for reducing CT slice thickness: effect on reproducibility of radiomic features in lung cancer. *Korean journal of radiology*, *20*(10), 1431-1440.
[27]. Ozdemir, Onur, Rebecca L. Russell, and Andrew A. Berlin. "A 3D probabilistic deep learning system for detection and diagnosis of lung cancer using low-dose CT scans." *IEEE transactions on medical imaging* 39, no. 5 (2019): 1419-1429.
[28]. Masood, A., Yang, P., Sheng, B., Li, H., Li, P., Qin, J., ... & Feng, D. D. (2019). Cloud-based automated clinical decision support system for detection and diagnosis of lung cancer in chest CT. *IEEE journal of translational engineering in health and medicine*, *8*, 1-13.
[29]. Qin, R., Wang, Z., Jiang, L., Qiao, K., Hai, J., Chen, J., ... & Yan, B. (2020). Fine-grained lung cancer classification from PET and CT images based on multidimensional attention mechanism. *Complexity*, *2020*.
[30]. Chao, H., Shan, H., Homayounieh, F., Singh, R., Khera, R. D., Guo, H., ... & Yan, P. (2021). Deep learning predicts cardiovascular disease risks from lung cancer screening low dose computed tomography. *Nature communications*, *12*(1), 1-10.

[31]. Gerard, S. E., Patton, T. J., Christensen, G. E., Bayouth, J. E., & Reinhardt, J. M. (2018). FissureNet: a deep learning approach for pulmonary fissure detection in CT images. *IEEE transactions on medical imaging*, *38*(1), 156-166.

[32]. Zhou, T., Lu, H., Yang, Z., Qiu, S., Huo, B., & Dong, Y. (2021). The ensemble deep learning model for novel COVID-19 on CT images. *Applied Soft Computing*, *98*, 106885.

[33]. Khan, M. A., Rubab, S., Kashif, A., Sharif, M. I., Muhammad, N., Shah, J. H., ... & Satapathy, S. C. (2020). Lungs cancer classification from CT images: An integrated design of contrast based classical features fusion and selection. *Pattern Recognition Letters*, *129*, 77-85.

[34]. Wang, W., Liu, F., Zhi, X., Zhang, T., & Huang, C. (2020). An integrated deep learning algorithm for detecting lung nodules with low-dose CT and its application in 6G-enabled internet of medical things. *IEEE Internet of Things Journal*, *8*(7), 5274-5284.

[35]. Suresh, S., & Mohan, S. (2020). ROI-based feature learning for efficient true positive prediction using convolutional neural network for lung cancer diagnosis. *Neural Computing and Applications*, *32*(20), 15989-16009.