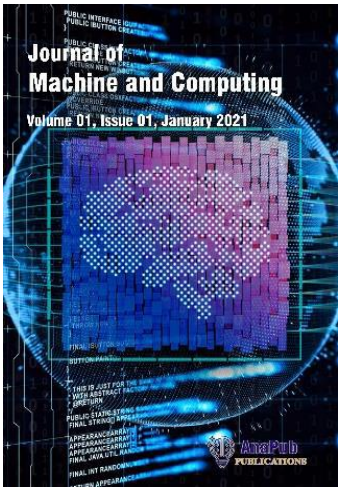
Journal Pre-proof

Enhanced Local Information Weighted Intuitionistic Fuzzy C-Means Clustering for Image Nodule Segmentation



Sandhya L and Marimuthu Karuppiah

DOI: 10.53759/7669/jmc202505184
Reference: JMC202505184
Journal: Journal of Machine and Computing.

Received 20 May 2025
Revised from 31 July 2025
Accepted 03 August 2025

Please cite this article as: Sandhya L and Marimuthu Karuppiah, ‘‘Enhanced Local Information Weighted Intuitionistic Fuzzy C-Means Clustering for Image Nodule Segmentation”, Journal of Machine and Computing. (2025). Doi: https:// doi.org/10.53759/7669/jmc202505184.

Enhanced Local Information Weighted Intuitionistic Fuzzy C-Means Clustering for Image Nodule Segmentation

Sandhya L[1], Marimuthu Karuppiah[2,*]

[1,2]Presidency School of Computer Science and Engineering, Presidency University, Bengaluru, Karnataka, 560119, India.

Email(s): sandhya.l@presidencyuniversity.in, marimuthu.k@presidencyuniversity.in, marimuthume@gmail.com

[*]*Corresponding author: Marimuthu Karuppiah(marimuthu.k@presidencyuniversity.in)*

Abstract

The early stages of the condition are notoriously difficult to diagnose due to the fact that abnormal cells with dimensions less than very small are notoriously difficult to spot by imaging. If identification occurs at an earlier stage, then there is a chance that the probability of extending the lifespan of the individual may increase. Nevertheless, because of the enormous dimensionality of the database space, timely diagnosis is a challenging task. In this paper, nodule segmentation is proposed using Enhanced Local Information Weighted Intuitionistic Fuzzy C-Means (ELWI-FCM) clustering. After pre-processing, segmentation is performed by ELWI-FCM. To optimize the performance of FCM, the improved Golden Eagle Optimization (IGEO) algorithm is used. For classifying the nodules as normal or affected, the pre-trained DenseNet201 model is utilized. Experiments are conducted over the LIDC-IDRI dataset. Experimental results show that the proposed ELWI-FCM-IGEO attains better accuracy, precision, F1-score, sensitivity, and specificity compared to existing models.

Keywords: Enhanced Local Information Weighted Intuitionistic, Fuzzy C-Means clustering, Golden Eagle Optimization, Densenet201

1. Introduction

The emergence of abnormal tissues within the human body leads to the condition known as cancer. Throughout history up to the present day, cancer remains a prevalent and formidable disease that poses a threat to human life. Lung cancer is considered as one of the deadliest disorders [1]. The timely prediction of early stages in cancer can significantly impact and potentially save numerous lives, especially when dealing with tumors in their initial phases [2] [3]. Moreover, it is expected that by the year 2040, this lung cancer may affect the 18 million lives of people. The development of an early cancer detection technique that can be used to lessen the effects of lung cancer is desperately needed [4].

Various diagnostic techniques, including X-rays, CT scans and Magnetic Resonance Imaging (MRI) can be utilized to detect lung cancer. Medical professionals, including doctors and radiologists, utilize CT scans for diagnosing lung cancer [5]. This enables them to characterize disease patterns, assess severity, and directly observe the morphological extent of tumors. Detecting tumor cells becomes challenging because of the misinterpretation of anatomical structures and the intensity variations in CT scan images. In recent times, CAD has emerged as a promising device to aid physicians and radiologists in cancer detection [6].

Despite the development of different models for lung cancer detection, achieving excellent and accurate results remains a substantial challenge. Therefore, effective detection of lung cancer can be achieved through the application of image processing approaches [7].

Recent studies in lung CT image segmentation have primarily concentrated on developing segmentation methods that are precise and efficient [8]. Some of the conventional segmentation approaches are thresholding, active contours, region growing, morphological, deformable, clustering, Markov random field, graph cut, watershed, histogram, fuzzy logic based segmentation, and neural networks. Furthermore, the thresholding is one of the conventional approaches employed in segmentation processes [9].

The utilization of DL as a representation learning model for acquiring hierarchical representation of features has proven to be a significant advancement. The major benefit of employing DL lies in its capacity to generate high level feature representations from image features [10]. This accuracy in data processing has created DL to substantial success, particularly in the field of segmentation of medical images. Some of the limitations in the existing works are: The conventional thresholding encounters challenges when applied to the bronchus and trachea due to their same grey values, resulting in suboptimal performance [11]. The segmentation of lung nodules will be affected by visual features, indistinct shapes, and the surrounding context of the nodules, posing challenges in accurate delineation. The segmentation is impacted by the similarity in visual behavior among lung nodules it influences the overall performance [12]. The variations in the appearance of nodules posed challenges to achieving accurate results. While some of DL models are employed to detect nodules in Images, however, achieving accurate segmentation is still challenging. The difficulties encountered in current lung nodule segmentation methods serve as motivation. A new approach, referred to as ELWI-FCM is introduced for the segmentation of lung nodules.

1.1. Motivation

Detecting nodules robustly presents a formidable challenge, because of the intricate nature of the surrounding environment and the heterogeneous characteristics of lung nodules. The significance of early detection is important to enhance the survival rates of patients affected by lung cancer. The distortion of segmentation processes caused by poor image quality has provided standard strategies ineffective in enhancing accuracy. Numerous scholars have developed CAD systems in the literature works. Despite these efforts, these systems continue to face challenges such as limited visibility and high False Positive Rate (FPR) results when detecting lung lesions. Hence, this work aims to introduce a method for identifying benign and malignant nodules in Images. Influenced by optimizing hyperparameters and visualization techniques, we have innovatively implemented a hybrid approach. This approach incorporates an optimized clustering to fine-tune the model's hyperparameters, aiming to yield optimal results and visually represent both normal and abnormal clusters.

1.2. Objectives

- An automated CAD based segmentation and classification method has been designed to categorize lung cancer diseases.

- To present an efficient pre-processing, segmentation and classification approaches.
- To efficiently segment the lesions using the ELWI-FCM with the IGEO algorithm.
- To overcome convergence issues of standard optimizer, the IGEO is presented.
- To classify the segmented regions as normal and abnormal classes, the DL model DenseNet201 is presented.
- To perform comparative analysis for different DL approaches with respect to the LIDC-IDRI dataset.

The rest sections are: Section 2 is the literature works with respect to the different models are discussed; Section 3 is the proposed lung tumor segmentation; Section 4 discusses results and Section 5 ends the work.

2. Related works

Literature works with respect to the lung tumor prediction using different models and the performance achieved are discussed:

Atiya et al. [16] presented non small cell lung cancers classification model using the TL (transfer learning) based CNN model. Initially, the images were resized and normalized and the augmentation process was carried out. Then, the two stage TL based pre-trained were used for classifying different stages of cancer. Finally, the accuracy value achieved by the ResNet50 was 94%.

Nanglia et al. [17] developed Feed Forward Back Propagation (FFBP) with Support Vector Machine (SVM) for lung cancer disease classification. Speeded Up Robust Features (SURF) was utilized for feature extraction. At last, the FFBP with the Genetic algorithm was utilized for the classification process. Accuracy and sensitivity values achieved were 98% and 96.5% on the ELCAP lung image dataset.

Gopinath et al. [18] presented Deep Fused Features with Cat Optimizer (DFF with CO) for classifying the lung cancer. Then, the saliency maps were utilized for the segmentation and the CNN was utilized for the classification. The DFF with CO was developed for training the features which combined the saliency maps and the DL model.

Siddiqui et al. [19] introduced Enhanced Deep Belief Network (E-DBN) with Gabor filters for lung nodule classification. The Restricted Boltzmann Machine (RBM) based Bernoulli Bernoulli and Gaussian Bernoulli were utilized. Experimental analysis was carried out on the three datasets and better performance was achieved for E-DBN with SVM classifier. .

Ajai et al. [20] introduced Renyi entropy fuzzy based shuffled social sky algorithm for lung nodules segmentation. Then, the texture and statistical features were extracted and RAM (rectified attention model) was utilized for the classification process.

Problem statement

Current lung cancer segmentation techniques encounter several limitations, such as high sensitivity to noise, challenges in differentiating tumors from healthy tissue, and variability in imaging protocols that compromise segmentation accuracy. Traditional methods

often require significant manual intervention and lack consistency, while automated approaches may fail to generalize well across diverse patient data and tumor morphologies. These shortcomings highlight a critical need for more robust solutions. The suggested model addresses these issues by leveraging adaptive clustering to refine segmentation boundaries and enhance tumor detection accuracy. The DenseNet201 architecture, known for its efficient feature reuse and deep representation capabilities, further strengthens the model's ability to distinguish complex tumor structures with greater precision, robustness, and generalizability, even in noisy or varied imaging conditions.

3. Proposed methodology

The advancement of the CAD has played a crucial role in medical analysis, aiding in decision-making regarding diseases of humans. Conventional models for predicting lung cancer faced challenges in accuracy due to poor quality images affecting the segmentation procedure. This paper introduces a novel, optimized approach segmentation model for lung cancer prediction. The collected images undergo pre-processing stage for noise elimination and to enhance the overall lung image quality. Subsequently, the cancerous part is segmented from the noise cleared image using an ELWI-FCM- IGEO. Finally, the features are extracted and classified in the DL model DenseNet201 as shown in Figure 1.

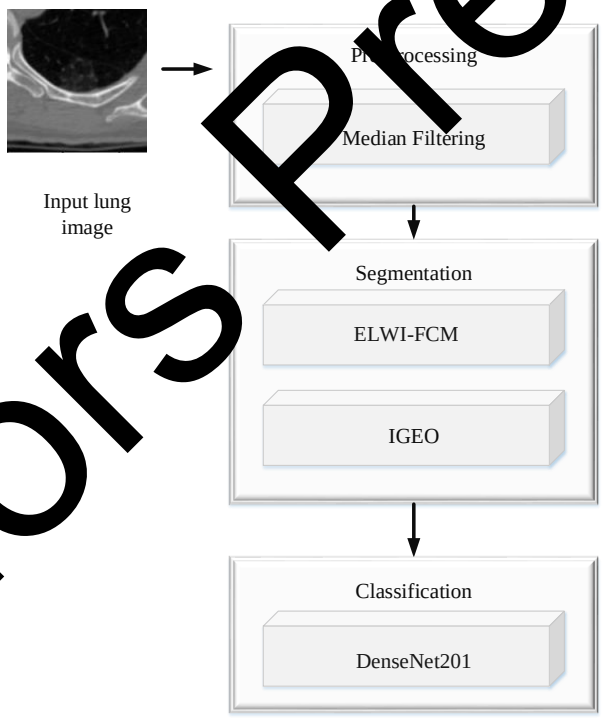


Figure 1: Workflow of the proposed lung tumor model

3.1 Pre-processing

Initially, the MF (median filtering) is utilized for removing noise in the LIDC-IDRI dataset. MF serves as a widely adopted technique to filter noise, operating as the low pass filter that retains information about images while effectively eliminating noise. This method involves filtering a neighborhood L with dimensions $m \times n$ by organizing all neighboring elements in ascending order and selecting the center component from the ordered sequence. Subsequently,

this chosen middle element replaces the pixel in the center. The mathematical representation of the MF is expressed as follows:

$$z_{(m,n)} = Med\{y_{(k.l)}, (k,l) \in L\} \quad (1)$$

3.2 Optimal segmentation

In the segmentation process, the lung images are divided into different categories and this process is performed by the algorithm ELWI-FCM-IGEO. This algorithm improved and simplifies the classification stage. The incorporation of ELWI-FCM allows for precise segmentation by considering local spatial context and intensity variations, which is particularly effective for handling noise and preserving the boundaries of complex tumor structures. This reduces the risk of over-segmentation and ensures more accurate detection of tumor regions. Secondly, the integration with IGEO optimizes the clustering parameters and improves convergence speed, leading to enhanced segmentation performance with reduced computational costs. IGEO, inspired by the hunting behavior of golden eagles, helps in efficiently exploring the solution space and finding optimal solutions, further improving the robustness and accuracy of the segmentation process. Together, these techniques provide a reliable and adaptable framework for lung cancer segmentation, capable of addressing challenges like variability in tumor shape, size, and imaging conditions while maintaining high precision and efficiency.

In the ELWI-FCM, local detail weight k_{lm} is introduced and it influences set of local detail on clustered outcomes. This overcomes local detail relies in less noise regions and avoids the impacts of membership function u_{lm}. In this ELWI-FCM, the objective term $\hat{I}(U,V)$ and k_{lm} is given as:

$$\hat{I}(U,V) = \sum_{l}^{p}\sum_{m} u_{lm}\, d^2(x_m, v_l) + \sum_{l=1}^{c} \pi_l^* \exp(1-\pi_l^*) + k_{lm} G_{lm} \quad (2)$$

where x_m is the object, v_l is the center of cluster, c is the cluster, $d(x_m, v_l)$ is the distance of x_m and v_l. Then, the value of k_{lm} is computed by:

$$k_{lm} = \frac{\sigma_m^2 + \alpha}{\sigma^2 + \alpha} \quad (3)$$

where σ_m^2 is the data's variance in the m^{th} window and α is the small parameter. The values of u_{lm} is computed by:

$$u_{lm} = \frac{1}{\sum_{s=1}^{c}\left(\frac{\| x_m - v_l \|^2 + G_{lm}}{\| x_m - v_s \|^2 + G_{lm}}\right)} \quad (4)$$

Here, the optimized term G_{lm} is given as:

$$G_{lm} = \sum \frac{(1-u_{lk})^n \parallel x_m - v_s \parallel^2}{d_{mk}+1} \quad (5)$$

where n is the fuzzified term and k is the is the neighbour pixels in the m^{th} window.

Finally, the membership function of ELWI-FCM is given as:

$$\hat{u}_{lm} = u_{lm}{}^n + u^*_{lm}{}^n \quad (6)$$

In this algorithm ELWI-FCM, generally the parameters like n and v_l are chosen by the trial and error process. The choice of initial v_l significantly impacts clustering outcomes, directly affecting the lung images segmentation quality. Hence, this work presents a metaheuristic algorithm IGEO is presented for optimizing the clustering performance. the fitness function of the IGEO is given as:

$$Fitness = Max(Dice\ value) \quad (7)$$

The mathematical modelling of the IGEO is presented in this section. The proposed IGEO is a novel approach to addressing global optimization problems, and draws inspiration and mathematical formulation from the intelligent behavior of GE (golden eagles), particularly their adept control over the speed of their spiral flight patterns. In emulating the distinctive characteristics of GE, known for their unique swarm behavior during the initial stages of hunting, IGEO efficiently navigates and explores the solution space. The algorithm's effectiveness lies in its adept manipulation of two key components: cruise propensity and attack propensity. GE employs a strategic hunting approach by systematically exploring various areas in search of superior prey. Central to their hunting methodology is a notable characteristic: their possession of a better memory. This unique attribute enables GE to remember and recall information related to both cruise and attack propensities while in flight. Mathematical modeling of IGEO is explained in this section.

GE's spiral movement: IGEO takes inspiration from the spiral motion exhibited by GEs. Every GE stores in memory the most optimal location it has encountered. The GE is concurrently drawn both towards the pursuit of attacking prey and engaging in a cruising motion to explore and locate superior food sources. In all iterations, every GE j chooses the prey of another GE g in random manner and encircles the better position.

$$g = \{1,2,.....,\ Population\,size\} \quad (8)$$

OBL: Once the population is initialized, opposite solutions are generated and these solutions exploited to enhance diversity within the solution set, thereby expanding the search space. They are derived by considering the opposite position of the solution of candidate $\vec{S}_a$, as determined by the following expression:

$$\vec{S}_a = \vec{lb} + \vec{ub} - \vec{S}_a \quad (9)$$

where $\vec{lb}$ and $\vec{ub}$ are lower and upper bounds.

Selecting prey: During every iteration, all GEs face the task of selecting a prey for executing cruising and attacking strategies in IGEO. The prey, in this context, is conceptualized as the optimum solution identified thus far by the entire flock of GEs. Each individual GE possesses the ability to retain in memory the most optimal solution it has encountered. For all iterations every GE picks a targeting prey from the collective memory of the entire flock. Subsequently, cruising and attacking vectors are computed for every GE eagle in relation to the chosen prey. If the resulting position proves superior to the prior position stored in memory, an update is made to the memory.

Exploitation: The attacking model can be represented by a vector that initiates from the current position of the GE and terminates at the location stored in the GE's memory for the prey. The computation of the GE's attack vector $\vec{A}_j$ is given as:

$$\vec{A}_j = \vec{Y}_g^* - \vec{Y}_j \qquad (10)$$

where $\vec{Y}_g^*$ and $\vec{Y}_j$ are the GE's best position at g and GE's present position at j.

Exploration: The selection of the cruise vector depends on the $\vec{A}_j$. Cruise and attack vectors are designated for the circle and perpendicular. To assess the cruise vector, it's necessary to ascertain the hyper plane's position. The model of the hyper plane is provided by:

$$h_1 y_1 + h_2 y_2 + ... + h_v y_v = d = \sum_{w=1}^{v} h_w t_w = d \qquad (11)$$

where $d = \vec{H}\,\vec{M}$, normal term is $\vec{H} = [h_1, h_2 ....., h_v]$ and parameter term is $Y = [y_1, y_2 ....., y_w]$ and the arbitrary term on the hyper-plane is $\vec{M} = [m_1, m_2 ....., m_w]$.

The value of cruise vector $\vec{C}_j$ is given as:

$$\sum_{w=1}^{v} a_w y_w = \sum_{w=1}^{v} a_w^t y_w^* \qquad (12)$$

where $\vec{A}_j = [a_1, a_2 ....., a_v]$, $Y = [y_1, y_2 ....., y_v]$ is the design parameter and $Y^* = [y_1^*, y_2^* ....., y_v^*]$ is the prey's location.

Movement of new position: The GE's movement undergoes A_j and C_j. The step parameter of GE is represented as:

$$\Delta y_j = \vec{rand}_1 \, m_a \frac{\vec{A}_j}{\left\|\vec{A}_j\right\|} + \vec{rand}_2 \, m_c \frac{\vec{C}_j}{\left\|\vec{C}_j\right\|} \qquad (13)$$

where $\vec{rand}_1$ and $\vec{rand}_2$ are the random vectors; m_a and m_c are the coefficients of attack and cruise; $\left\|\vec{A}_j\right\|$ and $\left\|\vec{C}_j\right\|$ are the Euclidean factors of A_j and C_j. These two values are computed by:

$$\left\|\vec{A}_j\right\| = \sqrt{\sum_{w=1}^{v} a_w^2}, \left\|\vec{C}_k\right\| = \sqrt{\sum_{w=1}^{v} c_w^2} \qquad (14)$$

Changing from exploration to exploitation: IGEO utilizes m_a and m_c for switching from the stage of exploration to exploitation. The values of m_a (initial bias) and m_c (final bias) are computed by:

$$m_a = m_a^0 + \frac{t}{T}\left[m_a^T - m_a^0\right] \qquad (15)$$

$$m_c = m_c^0 + \frac{t}{T}\left[m_c^T - m_c^0\right] \qquad (16)$$

where t and T are current and maximum iterations.

Algorithm 1: Pseudocode of the ELWI-FWM- IGEO
Input: x_m (object), v_l (center of cluster), n (fuzzified term) Population size, Iterations, q_a and q_c
Output: Optimal v_l
Initializing the OBL in the population of GE
Estimate the fitness by expression (7)
Estimate q_a and q_c
for every iteration
Update m_a (initial bias) and m_c (final bias) by expressions (15) and (16)
for all j
Randomly define the prey from population memory
Define $\vec{A}_j$ by expression (9)
If length of $\vec{A}_j$ is $\neq 0$
Define C_j by expression (12)
Define Δy_j by expression (13)

Calculate the fitness value for the newly generated positions
If the fitness value exceeds the memory
Utilize the newly calculated position value instead of the position stored in the GE memory.
end
end
end
end

3.3 Classification

Finally, for classifying normal and abnormal classes, the DL model DenseNet201 is utilized. DenseNet201 [22] has a similar structure to ResNet, but its notable distinction lies in its approach to information flow across layers. In this network, each layer is directly connected to every other layer in a feed-forward manner. This design maximizes the information exchange among network layers, fostering dense connections throughout the model architecture. That is, the DenseNet's layers are interconnected in a way where every layer receives input from all preceding layers and fed its output to all subsequent layers. This design strengthens the robust information flow throughout the network. DenseNet201 offers numerous benefits, including addressing the gradient vanish issue, enhancing the propagation of features, facilitating feature reusability, and significantly reducing parameter count. Consequently, this network is more lightweight and efficient in terms of computation and memory usage. The architecture of DenseNet201 typically comprises 4-layers of convolutional layers, 3-TL (transition layers), and 1-FC (fully connected) layer and softmax layer as shown in Figure 2.

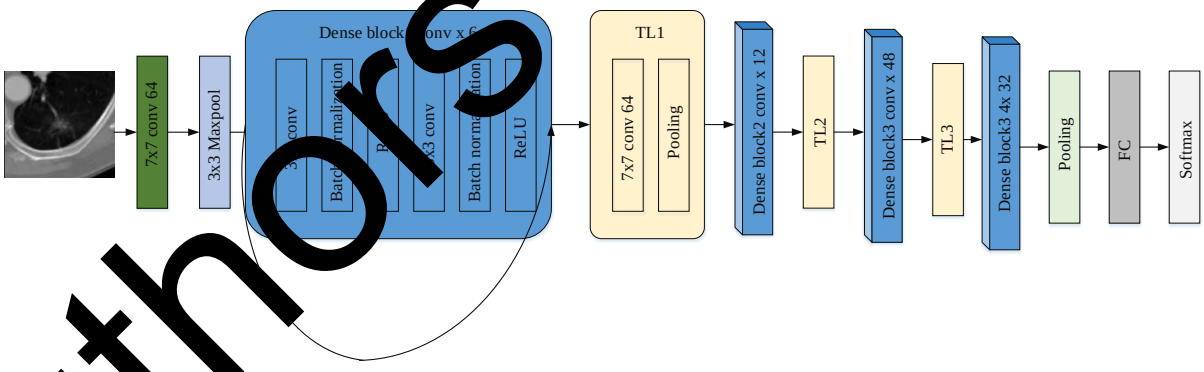


Figure 2: Architecture of DenseNet201

4. Results analysis

The segmentation results of the proposed ELWI-FCM-IGEO are performed in Python software. The experimental parameters like batch size (100), epochs (50), maximum iteration (200) and initial population (50) are considered. The experimental outcomes are validated on 10-cross validation.

4.1 Dataset details

The LIDC-IDRI [21] is a comprehensive repository of diagnostic and lung tumor monitoring thoracic CT modalities, meticulously annotated with tumor. These images are obtained from seven academic institutions and 8 clinical imaging institutions, culminating in a dataset comprising subjects of 1018. Each case comprises images from a medical thoracic Images accompanied by an XML file. The lesions are divided into three distinct classes like nodule ≥3 mm, non-nodule ≥3 mm and nodule <3 mm. Ground truth involves a comprehensive annotation process carried out by expert radiologists.

4.2 Input and Processed Dataset

Figure 3 shows the input, Ground Truth (GT), segmented and contour images of normal and abnormal classes. The segmented images obtained by ELWI-FCM-IGEO model are compared with the GT image and it is noted that the segmented image is matched with the GT.

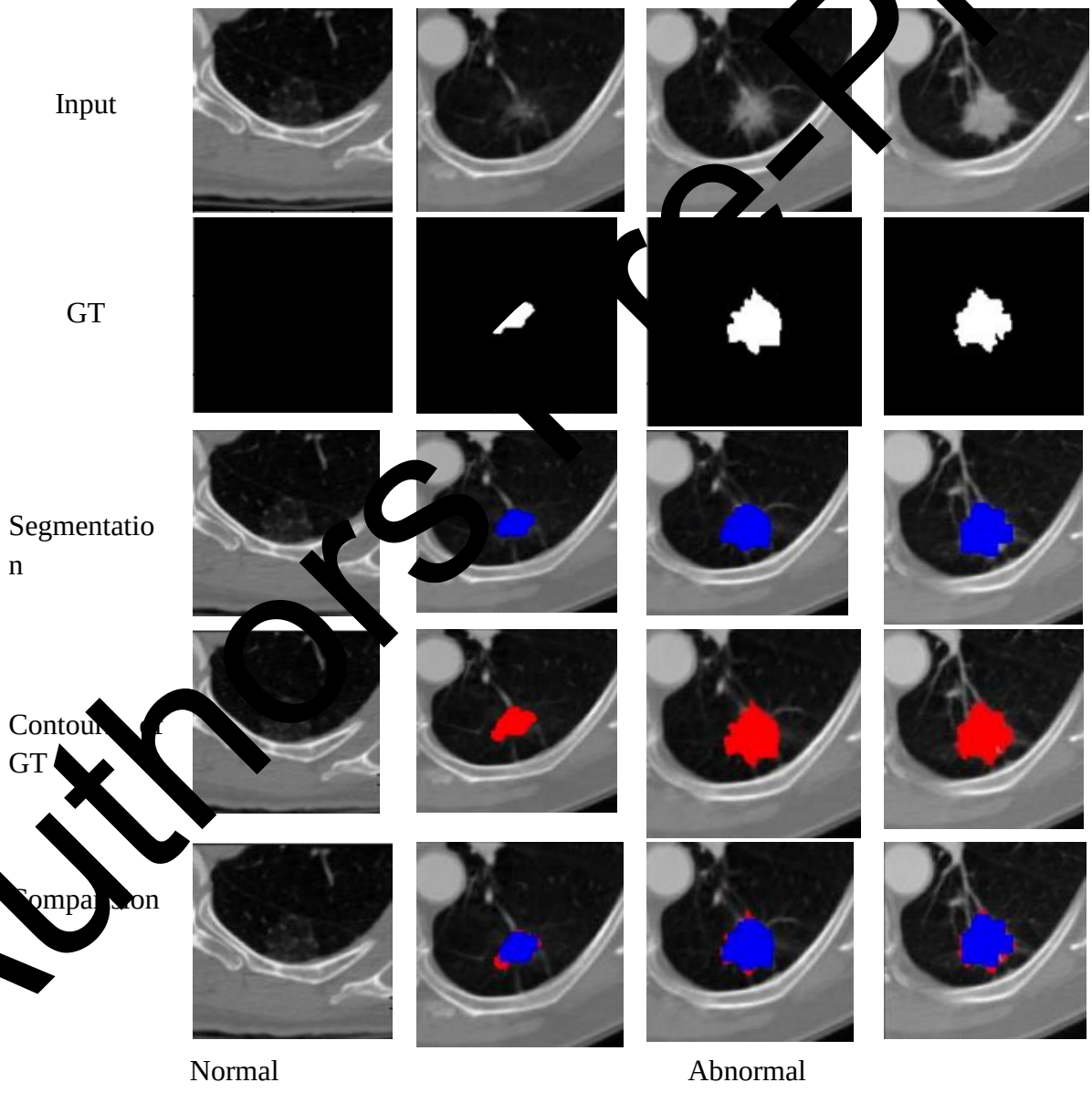


Figure 3: Image analysis of the proposed segmentation

4.3 Performance Metrics

The classification performance is evaluated in terms of the following measures:

$$\textbf{Accuracy} = \frac{TN + TP}{FP + FN + TP + TN} \quad (17)$$

$$\textbf{Specificity} = \frac{TN}{TN + FP} \quad (18)$$

$$\textbf{Sensitivity} = \frac{TP}{TP + FN} \quad (19)$$

$$\textbf{Precision} = \frac{TP}{TP + FP} \quad (20)$$

$$\textbf{Recall} = \frac{TP}{TP + FN} \quad (21)$$

$$\textbf{F1-score} = 2x\frac{precisionXrecall}{precision + recall} \quad (22)$$

In calculating these measures, the four values summarized below are used: **TP:** True Positive, **FP**: False Positive, **FN**: False Negative, **TN**: True Negative .

4.4 Comparison Results

In this section, the performance of the proposed ELWI-FCM-IGEO model has been compared with other clustering models such as traditional FCM, K-means and DBSCAN.

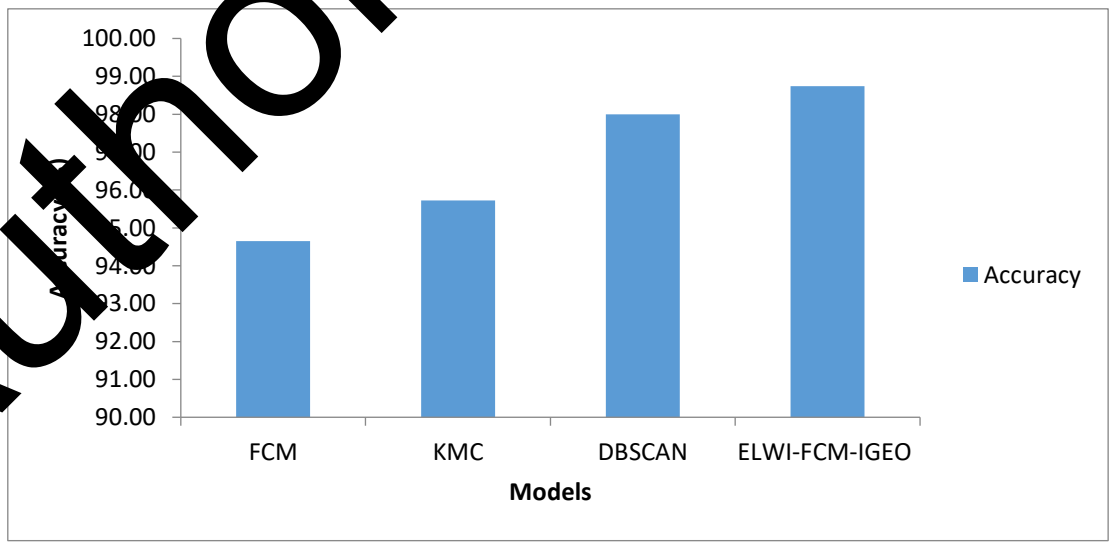


Figure 4(a) Comparison of Accuracy,

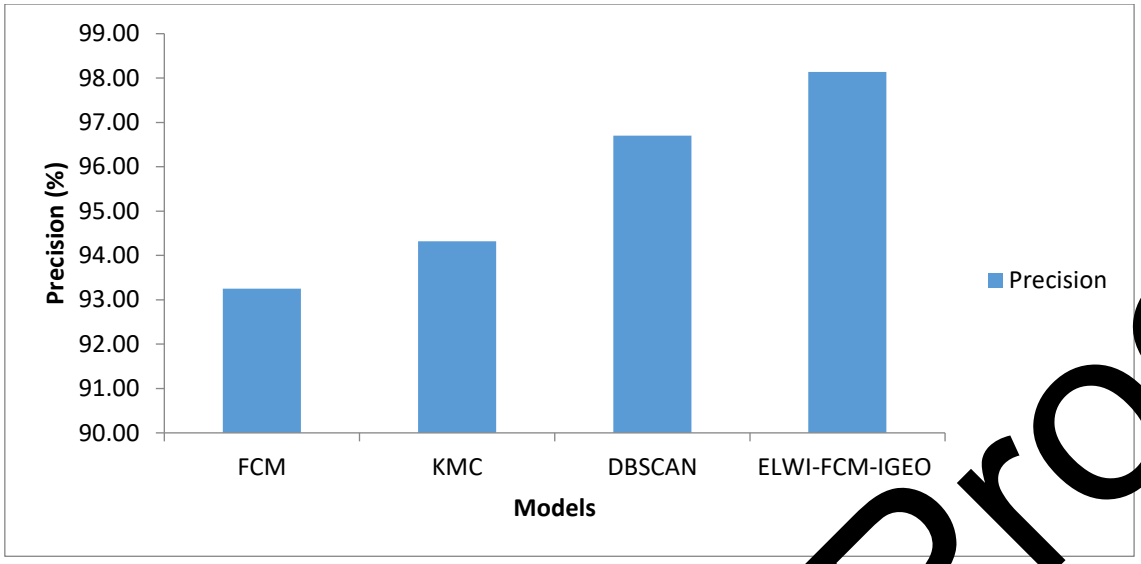


Figure 4(b): Comparison of Precision

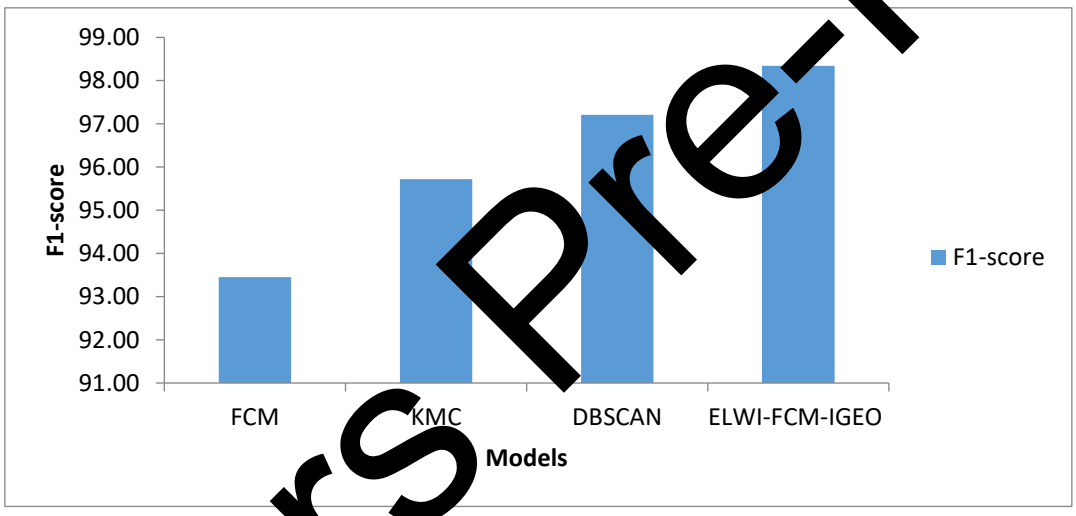


Figure 4(c): Comparison of F1-score

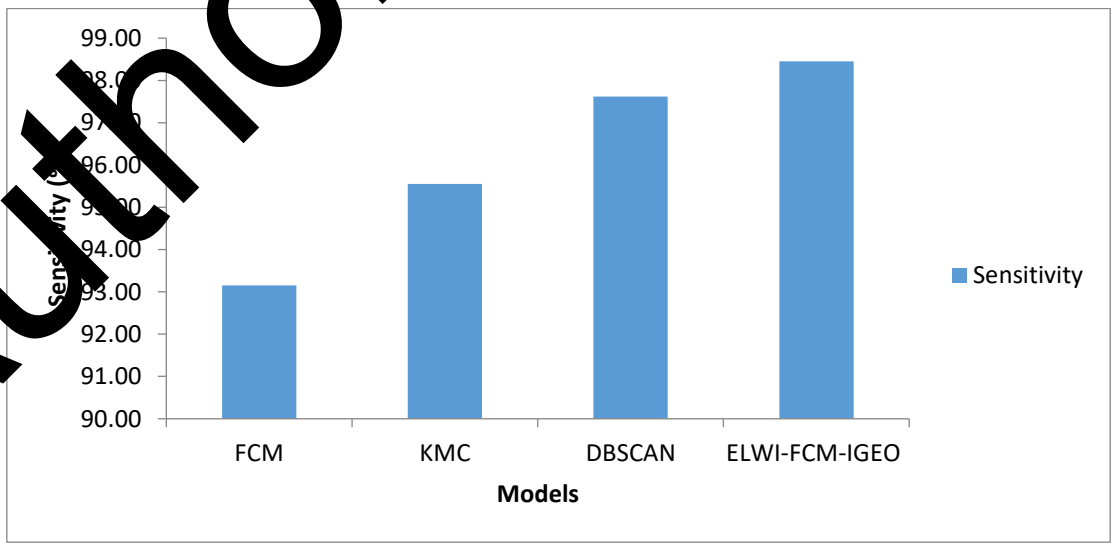


Figure 4(d): Comparison of Sensitivity

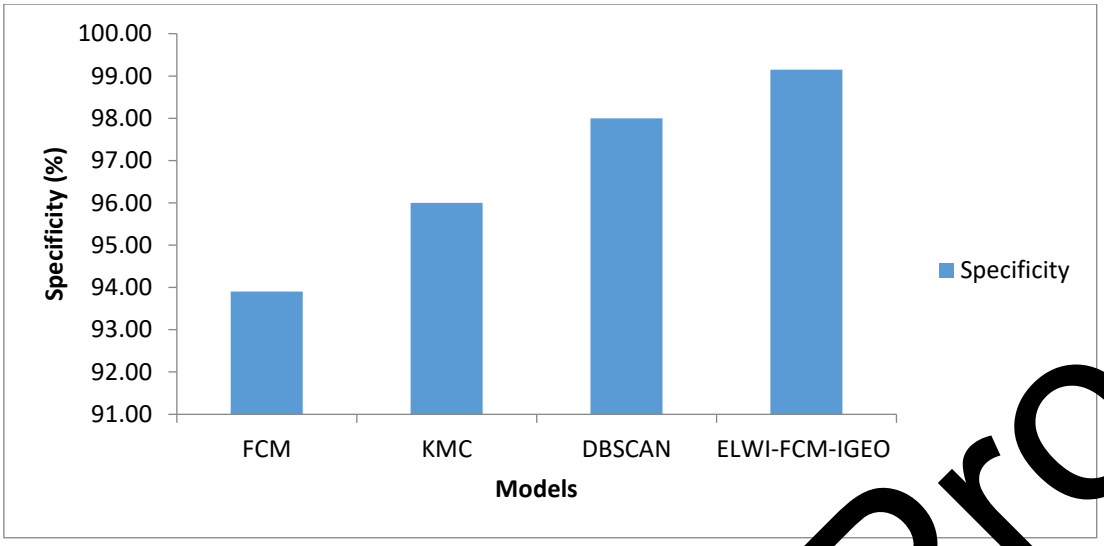


Figure 4(e): Comparison of Specificity

Figure 4(a) shows the comparison results of Accuracy It is observed from the graphical analysis that the proposed ELWI-FCM-IGEO achieved better accuracy of 98.7%, when compared to the other clustering models.

Figure 4(b) shows the comparison results of Precision. It is observed from the graphical analysis that the proposed ELWI-FCM-IGEO achieved better precision of 98.14%, when compared to the other clustering models.

Figure 4(c) shows the comparison results of F1-score. It is observed from the graphical analysis that the proposed ELWI-FCM-IGEO achieved better F1-score of 98.34%, when compared to the other clustering models.

Figure 4(d) shows the comparison results of Sensitivity. It is observed from the graphical analysis that the proposed ELWI-FCM-IGEO achieved better F1-score of 98.45%, when compared to the other clustering models.

Figure 4(e) shows the comparison results of Specificity. It is observed from the graphical analysis that the proposed ELWI-FCM-IGEO achieved better Specificity of 99.15% when compared to the other clustering models.

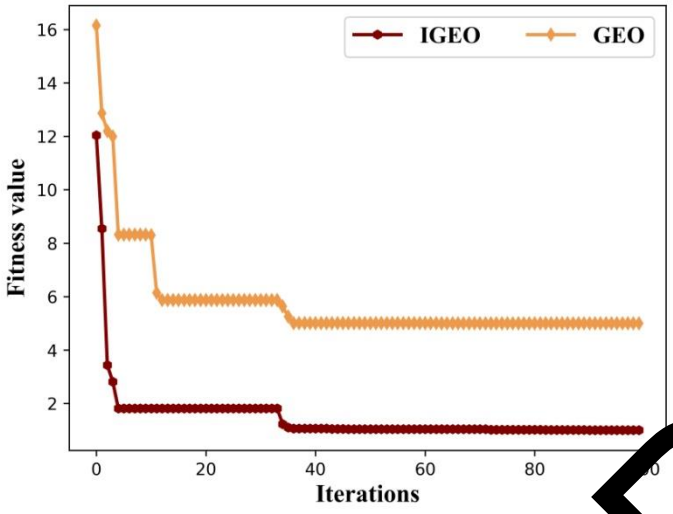


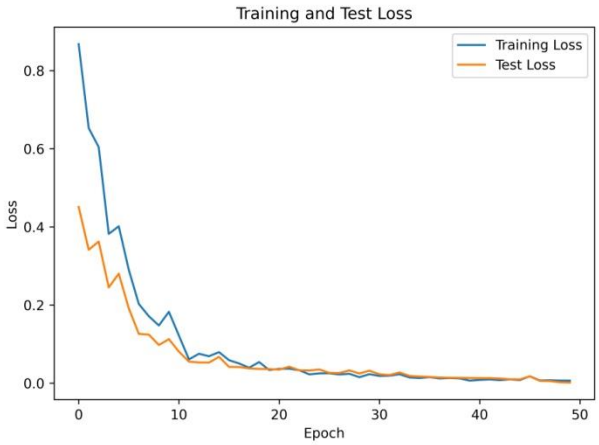
Figure 5: Comparison of optimization approache

Figure 5 illustrates the comparison results of the optimization m els IGEO and GEO. It is evident from the analysis that the proposed IGEO exhibi a n ably higher convergence rate when compared with the GEO. Consequently, the pr d met od emerges as a favorable choice for the segmentation process.

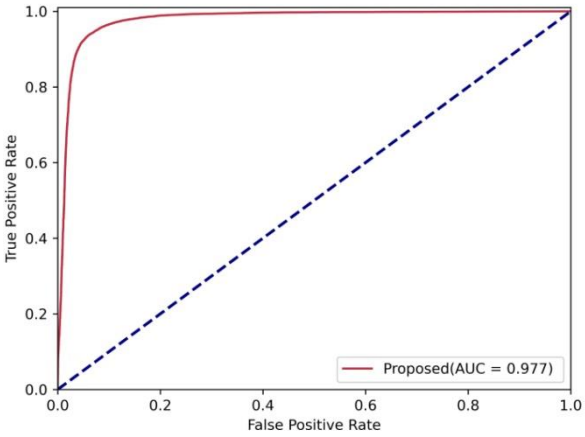
Figure 6(a) and 6(b) depict the tra ng & sting rves of accuracy and loss for ELWI-FCM-IGEO. It is observed that the mode is no nder-fit and over-fit.

Figure 6(c) presents the ROC (region characteristics) of the proposed ELWI-FCM-IGEO and the AUC (area under the curve) value achieved is 0.977.

(a)



(b)



(c)

Figure 6: Analysis of (a) training-testing of accuracy, (b) loss curves and (c) ROC

Table 1: Comparison with recent works

References	Accuracy (%)
Sang et. al [1]	94
Pedrosa et. al [2]	-
Ali et. al [4]	-
Jain et al. [13]	98
Navaneetha krishnan et al. [14]	92
Atiya et al. [16]	94
Proposed	**98.7**

Table 1 presents the comparison with recent works like Sang et. al [1], Pedrosa et. al [2], Ali et. al [4], Jain et al. [13], Navaneetha krishnan et al. [14] and Atiya et al. [16] are compared with proposed segmentation model. It is noted that the proposed segmentation model outperforms other models.

Table 2: Ablation study

Methods	Accuracy (%)
ViT	93.4
UNet	94.2
Attention UNet	96.4
Proposed	98.7

Table 2 suggests the ablation study with respect to the methods like Vision Transformer (ViT), UNet and Attention UNet.

5. Conclusion

In this study, a novel hybrid optimizer combining Golden Eagle Optimization (GEO) and Opposition-Based Learning (OBL) was developed to effectively segment nodules. Initially, the images were pre-processed using a median filter (MF) to eliminate noise. Subsequently, segmentation of the affected regions was performed using Enhanced Local Information Weighted Intuitionistic Fuzzy C-Means (ELWI-FCM). The segmentation performance was further improved with the assistance of the Improved Golden Eagle Optimization (IGEO) algorithm. Finally, classification was carried out using the pre-trained DenseNet201 model. Experiments were conducted on the LIDC-IDRI dataset. The performance of the proposed ELWI-FCM-IGEO model was compared with other clustering techniques, including traditional FCM, K-means, and DBSCAN. The experimental results demonstrate that the proposed model achieved an accuracy of 98.7%, precision of 98.14%, F1-score of 98.34%, sensitivity of 98.45%, and specificity of 99.15%, outperforming the existing models. In future work, various deep learning architectures and optimization strategies will be explored to further enhance both segmentation and classification performance.

CRediT Author Statement

The authors confirm contribution to the paper as follows:

Conceptualization: Sandhya L and Marimuthu Karuppiah; **Methodology**: Sandhya L; **Data Curation**: Sandhya L; **Writing- Original Draft Preparation**: Sandhya L; **Visualization**: Sandhya L; **Supervision**: Marimuthu Karuppiah; **Validation**: Marimuthu Karuppiah; **Writing- Reviewing and Editing**: Sandhya L and Marimuthu Karuppiah; All authors reviewed the results and approved the final version of the manuscript.

Data Availability

No data was used to support this study.

Conflicts of Interest

The author(s) declare(s) that they have no conflicts of interest.

Funding

No funding agency is associated with this research.

References

1. J. Sang, M. S. Alam, and H. Xiang, "Automated detection and classification for early stage lung cancer on Images using deep learning," in *Proc. Pattern Recognition and Tracking XXX*, vol. 10995, pp. 200–207, SPIE, 2019.

2. J. Pedrosa *et al.*, "LNDb challenge on automatic lung cancer patient management," *Med. Image Anal.*, vol. 70, p. 102027, 2021.

3. V. K. Gunjan, N. Singh, F. Shaik, and S. Roy, "Detection of lung cancer in CT scans using grey wolf optimization algorithm and recurrent neural network," *Health Technol.*, vol. 12, no. 6, pp. 1197–1210, 2022.

4. Z. Ali, A. Irtaza, and M. Maqsood, "An efficient U-Net framework for lung nodule detection using densely connected dilated convolutions," *J. Supercomput.*, vol. 78, no. 2, pp. 1602–1623, 2022.

5. T. L. Chaunzwa *et al*., "Deep learning classification of lung cancer histology using CT images," *Sci. Rep.*, vol. 11, no. 1, p. 5471, 2021.

6. A. Bhattacharjee, R. Murugan, and T. Goel, "A hybrid approach for lung cancer diagnosis using optimized random forest classification and K-means visualization algorithm," *Health Technol.*, vol. 12, no. 4, pp. 787–800, 2022.

7. T. Meraj *et al*., "Lung nodules detection using semantic segmentation and classification with optimal features," *Neural Comput. Appl.*, vol. 33, pp. 10737–10750, 2021.

8. S. A. Agnes and J. Anitha, "Efficient multiscale fully convolutional UNet model for segmentation of 3D lung nodule from CT image," *J. Med. Imaging*, vol. 9, no. 5, p. 052402, 2022.

9. S. Kido *et al*., "Segmentation of lung nodules on CT images using a nested three-dimensional fully connected convolutional network," *Front. Artif. Intell.*, vol. 5, p. 782225, 2022.

10. S. Mahesh, "Hybrid optimized MRF based lung lobe segmentation and lung cancer classification using Shufflenet," *Multimed. Tools Appl.*, pp. 1–30, 2023.

11. C. de Margerie-Mellon and G. Chassagnon, "Artificial intelligence: A critical review of applications for lung nodule and lung cancer," *Diagn. Interv. Imaging*, vol. 104, no. 1, pp. 11–17, 2023.

12. M. Tsivgoulis, T. Papastergiou, and V. Megalooikonomou, "An improved SqueezeNet model for the diagnosis of lung cancer in CT scans," *Mach. Learn. Appl.*, vol. 10, p. 100399, 2022.

13. Jain, S. Indora, and D. K. Atal, "Lung nodule segmentation using salp shuffled shepherd optimization algorithm-based generative adversarial network," *Comput. Biol. Med.*, vol. 137, p. 104811, 2021.

14. M. Navaneethakrishnan *et al*., "Deep Fuzzy SegNet-based lung nodule segmentation and optimized deep learning for lung cancer detection," *Pattern Anal. Appl.*, pp. 1–17, 2023.

15. Z. Ren, Y. Zhang, and S. Wang, "LCDAE: data augmented ensemble framework for lung cancer classification," *Technol. Cancer Res. Treat.*, vol. 21, p. 15330338221124372, 2022.

16. S. U. Atiya, N. V. K. Ramesh, and B. N. K. Reddy, "Classification of non-small cell lung cancers using deep convolutional neural networks," *Multimed. Tools Appl.*, pp. 1–30, 2023.

17. P. Nanglia, S. Kumar, A. N. Mahajan, P. Singh, and D. Rathee, "A hybrid algorithm for lung cancer classification using SVM and Neural Networks," *ICT Express*, vol. 7, no. 3, pp. 335–341, 2021.

18. A. Gopinath, P. Gowthaman, M. Venkatachalam, and M. Saroja, "Computer aided model for lung cancer classification using cat optimized convolutional neural networks," *Meas. Sens.*, vol. 30, p. 100932, 2023.

19. E. A. Siddiqui, V. Chaurasia, and M. Shandilya, "Detection and classification of lung cancer computed tomography images using a novel improved deep belief network with Gabor filters," *Chemom. Intell. Lab. Syst.*, vol. 235, p. 104743, 2023.

20. A. K. Ajai and A. Anitha, "Clustering based lung lobe segmentation and optimization based lung cancer classification using CT images," *Biomed. Signal Process. Control*, vol. 78, p. 103986, 2022.

21. C. Jacobs *et al*., "Computer-aided detection of pulmonary nodules: a comparative study using the public LIDC/IDRI database," *Eur. Radiol.*, vol. 26, pp. 2139–2147, 2016.

22. Y. Zhu and S. Newsam, "DenseNet for dense flow," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, pp. 790–794, 2017.