Original Article

Role of Feature Selection in CT Lung Nodule Classification: The modified Particle Swarm Optimization and Topology selection

Mohini Manav, Monika Goyal, Anuj Kumar

Abstract:

Background: In healthcare, machine learning is playing a significant role in computer-aided detection and diagnosis of lung nodules to reduce the load and increase the accuracy of radiologists. In the scenario where the datasize is limited, machine learning poses its advantage. The intrinsic information is discerned utilizing radiomics, which, in turn, serves as the foundational feature set for shaping the algorithmic framework employed in the classification process.

In this study, the Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI) database was used to study the lung nodule classification with different numbers of radiomic features selected.

Materials and Methods: A total of 1018 thoracic CT scans having lung nodules categorized as "nodule ≥3 mm," "nodule <3 mm," and "non-nodule ≥ 3 mm," identified through a two-phase annotation process out of which, 300 CT scans with nodules ≥ 3 mm in size were used in our study. Four machine learning models, namely, Light GBM, Random Forest, XG Boost, and Support Vector Machine, were used with the optimum number of radiomic features selected using binary Particle Swarm Optimization combined with different topologies and time-varying inertia weights.

Results and Conclusion: The XGBoost classifier with ring topology and linear decreasing inertia weight presented the best results, achieving 97.67%, 95.74%, 100%, 95.12%, 97.83%, and 98.75% for accuracy, precision, sensitivity, specificity, f1-score, and AUC, respectively.

This outperformed a wide range of feature selection and machine learning approaches for lung nodule classification on the same data, as documented in numerous published academic papers. The proposed method demonstrated an improved nodule classification performance by utilizing optimal features obtained through the fusion of various inertia weights and different topologies in the feature selection method.

JK-Practitioner2023;28(3-4):74-85

1 Introduction

Lung cancer is the most common reason for cancer-related fatalities worldwide.[1]In 2020, lung cancer leading with nearly 2.21 million deaths. [2] For a good prognosis, detecting lung cancer at an early stage is very important. [3] Calculating the likelihood of malignancy for early malignant lung nodules is a troublesome job. [3,4]Compared to standard chest radiography, computed tomography (CT) screening is more effective at lung cancer screening. [4,5]According to the report by The National Lung Screening Trial, there is a 20% reduction in lung cancer mortality using CT images for diagnosis. [6]

CT has emerged as an imaging technique with greater sensitivity in detecting lung nodules. [7]Pulmonary nodules are lung abnormalities that are crucial indicators of lung cancers visible to lung computed tomography (CT) scans as roughly round opacities. [8] Accurate diagnosis of lung nodules is challenging, laborious, and time-consuming.[9]

Several computer-aided diagnosis and detection techniques have been developed to aid radiologists in detecting and diagnosing lung nodules with greater accuracy and efficiency. [10,11]Lung nodules can be differentiated based on features like shapes, contours, textures, etc.[12]

Author Affiliations
Mohini Manav, Department of Physics, GLA University, Mathura, U.P, India, Department of Radiotherapy, S N Medical College, Agra, U.P., India, **Monika Goyal,** Department of Physics, GLA University, Mathura, U.P, India
Anuj Kumar, Department of Radiotherapy, S N Medical College, Agra, U.P., India

Correspondence
Prof. Anuj Kumar
Department of Radiotherapy,
S N Medical College, Agra, U.P., India
Contact no.: +91-9760047732
Email: toaktyagi@gmail.com

Indexed
EMBASE, SCOPUS, IndMED, ESBCO, Google Scholar besides other national and International Databases

Cite This Article as
Manav M,Goyal M, Kumar A, Role of Feature Selection in CT Lung Nodule Classification: The modified Particle Swarm Optimization and Topology selection, JK ract2023;28(3-4): 74-85

Full length article available at **jkpractitioner.com** one month after publication

Keywords
Machine learning, Particle Swarm Optimization (PSO), inertia weight, LIDC-IDRI, Medical Imaging, Medical Image classification, Computed Tomography

]The role of Machine Learning has shown its remarkable significance in healthcare by assisting in various tasks such as disease diagnosis.[13]

Feature selection is a vital pre-processing step in eliminating the redundant features and lowering the number of features.[14]In classification problems, where the quality of the selected features can significantly impact the accuracy and efficiency of classification models, feature selection is essential.[15] In cancer diagnosis and prognosis, feature selection can also improve models' interpretability and clinical relevance. [16]Deep learning, a subset of machine learning, acts as a black-box solution to many problems, e.g., object recognition, natural language processing, disease detection in healthcare, etc. [17]

In recent years, bio-inspired algorithms have gained popularity due to their capability to resolve complex problems that traditional algorithms struggle with. [18] They are designed to mimic the behaviour of organisms in nature and apply these behaviours to solve optimization problems. [19] Feature selection is one of the most common applications of nature-inspired algorithms. [20] Some popular nature-inspired feature selection algorithms include Genetic Algorithm, Particle Swarm Optimization (PSO), Ant Colony Optimization, Artificial Bee Colony, and Grey Wolf Optimizer. [21-26]PSO is one of the nature-inspired algorithms. PSO is a population-based optimization algorithm that mimics the social behaviour of swarms of birds or fish. The algorithm works by simulating the movement of particles in search of the optimal solution. It has been applied to many optimization problems in recent years. [27,28]

Inthe present paper, we have classified the lung nodule into malignant and benign categories from the CT images. To classify CT lung nodules, we explored multiple topologies, including the ring, star, pyramid, and random topologies, while simultaneously investigating various inertia weight adjustments such as exponential, linear, and nonlinear changes in PSO. [29-34]We utilized the selected features extracted from the CT images to perform classification using several popular supervised machine learning algorithms, namely LightGBM (LGBM), Random Forest(RF), XGBoost (XGB), and Support Vector Machines (SVM). [35-38]

The study results are presented as the performance of the various classification models in terms of accuracy, precision, sensitivity, specificity, f1-score, and AUC score based on the selected features. Through our experimentation, we strive to contribute to the advancement of accurate and efficient lung nodule classification. The subsequent sections will delve into the materials and methods, results, and discussions, and finally, the conclusion provides a comprehensive evaluation of the effectiveness of our integrated approach and the performance of the LGBM, RF, SVM, and XGB algorithms.

2. Materials and methods:

2.1 Dataset used:

The dataset was obtained from the Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI) database, the largest available open-accessible dataset of lung nodules. [39-41]The dataset comprises1018 diagnostic and lung cancer screening thoracic computed tomography (CT) scans with marked-up annotated lesions. These scans were collected from 1010 patients through collaboration between academic centers and medical imaging companies. Four experienced radiologists independently reviewed the scans in a two-phase annotation process and marked the lesions which belong to one of three categories ("nodule > or =3 mm," "nodule <3 mm," and "non-nodule > or =3 mm").[41]

In the present study, we have used an initial 300 scans. The 80% consensus consolidation of the annotation contours was computed, which means the annotations from all annotators by considering regions where at least 80% of the radiologists' annotations agree or overlap were combined. This approach aims to create a consolidated segmentation representing a consensus among the annotators. Also, we limited the scope of our study to the nodules >= 3 mm.

2.2 Pre-processing:

Nodule segmentation: For the nodule segmentation from the segmented lung images, the pylidc library, which is an Object-rational mapping (using SQL Alchemy) for the LIDC dataset, was used. [42]The XML file associated with each scan was analyzed using the scriptsto get the segmentation object.

Feature Extraction: For all the segmented nodules, features such as the First Order Statistics, Shape-based (3D), Gray Level Co-occurrence Matrix, Gray Level Run Length Matrix, Gray Level Size Zone Matrix, Neighbouring Gray Tone Difference Matrix, Gray Level Dependence Matrix were extracted for each nodule. For the feature extraction, the Python library PyRadiomics was used. [43]

Feature selection: For the feature selection process and to investigate the impact of different inertia weight variations with different topologies on the performance of the PSO algorithm, three types of inertia weight functions were employed: exponential, linear, and nonlinear. Each function served to modulate the inertia weight during the optimization process.

2.3Time-Varying Binary Particle Swarm Optimization (TV-BPSO): TV-BPSO is a BPSO algorithm variant that utilizes time-varying inertia weight to improve optimization capabilities. [44]The inertia weight influences the trade-off between exploration (global search) and exploitation (local search) during the optimization process. [45]

By dynamically adjusting the inertia weight over time, TV-BPSO enhances the search capability and helps to avoid premature convergence to suboptimal solutions.

2.4 Time-Varying Inertia Weight (TV-IW) techniques:

In this study, we enhanced the traditional BPSO by incorporating TV-IW techniques. TV-IW modifies the inertia weight during optimization to control the balance between exploration and exploitation. Specifically, we explore three types of time-varying inertia weights: linear, nonlinear, and exponential.[32-34]The following equations were used for the linear, nonlinear, and exponentially decreasing intertia weights respectively.

$$w = (w^{initial} - w^{final} - d_1)\exp\left(\frac{1}{1 + \frac{d_2.iter_t}{iter_{max}}}\right)$$

$$w = w^{final} + (w^{initial} - w^{final})\left(\frac{iter_{max} - iter_t}{iter_{max}}\right)$$

$$w = w^{final} + (w^{initial} - w^{final})\left(\frac{iter_{max} - iter_t}{iter_{max}}\right)^n$$

Where d_1 and d_2 are control factors to control w between $w^{initial}$ and w^{final} and have values 2 and 7, respectively. $w^{initial}$ represents the initial inertia weight at starting inertia weight, and w^{final} is the inertia weight value when the algorithm process runs the max iterations. The value of $w^{initial}$ and $w^{final}$is 1.0 and 0.4, respectively. $Iter_{max}$ denotes the maximum iteration, and $iter_t$ denotes the t^{th}iteration. n is the modulation index and has a value of 1.2 in the present study.

2.5 Different Swarm Topologies:

Topology, which governs the interactionand information sharing among particles, plays a crucial role in the performance of BPSO. The choice of topology can significantly impact search behavior and convergence. The particles' interactions and information exchange in different topologies can impact the convergence speed and exploration capacity of TV-BPSO during the feature selection process. This study aims to comprehensively analyze BPSO using four topologies: Ring, Star, Pyramid, and Random.[46,47]

2.6 Image Classification using supervised machine learning:

Different supervised machine-learning models were used to classify the images using the reduced features from the abovementioned algorithms. These models were the LGBM classifier, RF classifier, XGB classifier, and SVM classifier.[35-38]

2.7Performance evaluation: The metrics used in the evaluation process for the model performance were accuracy, precision, sensitivity, specificity, f1-score, and AUC score.

3. Results:

3.1 Experimental setup: In the present study, the CT scans from the LIDC-IDRI database were classified using different supervised machine learning models with the optimum number of features selected using binary PSO with the combination of different topologies and time-varying inertia weights. The data was divided into train and test data as 80:20. The experiments were performed using the Python programming language, and Google Colab was used as a computational platform to execute the experiments. For different feature extraction, the PyRadiomics library was used. [43]For the feature selection processfrom the training dataset, we utilized the PySwarms library by making specific modifications and customizations tailored to the requirements of our research. [48] For the classification model, the Scikit-learn library was used. [49]Figure 1 demonstrates the proposed method employed in the study for classifying CT scans from the LIDC-IDRI database.

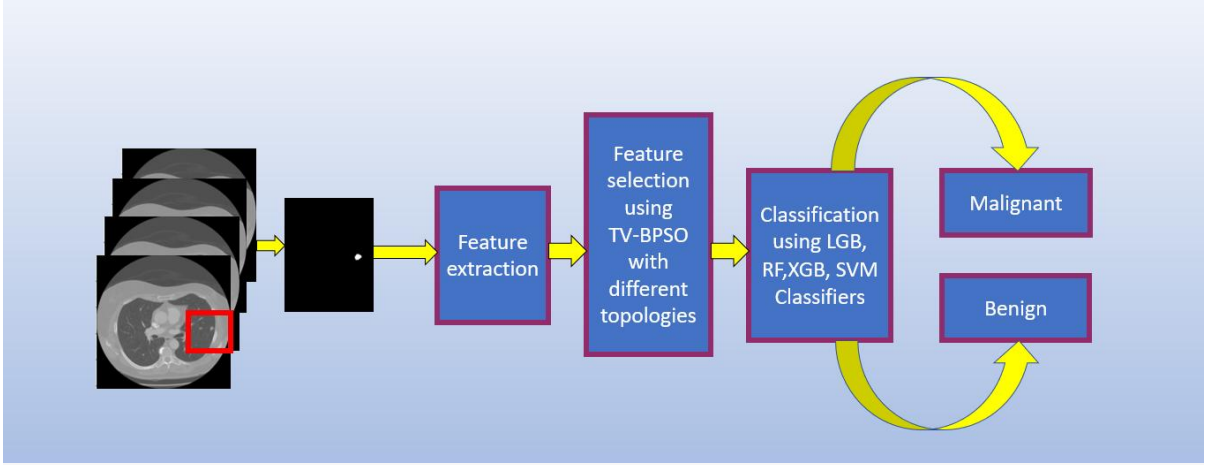


Fig1.Block diagram for proposed method in the present study for classification of CT nodules

3.2 Feature Extraction:

A total of 111 features were extracted from the segmented nodule images.Different features extracted from the CXR images are given in Table 1:

Table 1: Features type and number of features

Sr. No.	Feature Type	Number of features
1.	First Order Statistics	19
2.	Shape-based (3D)	17
3.	Gray Level Co-occurrence Matrix	24
4.	Gray Level Run Length Matrix	16
5.	Gray Level Size Zone Matrix	16
6.	Neighbouring Gray Tone Difference Matrix	5
7.	Gray Level Dependence Matrix	14
	Total	111

Figure 2 shows the graphical representation of selected features with different topologies and inertia weights.

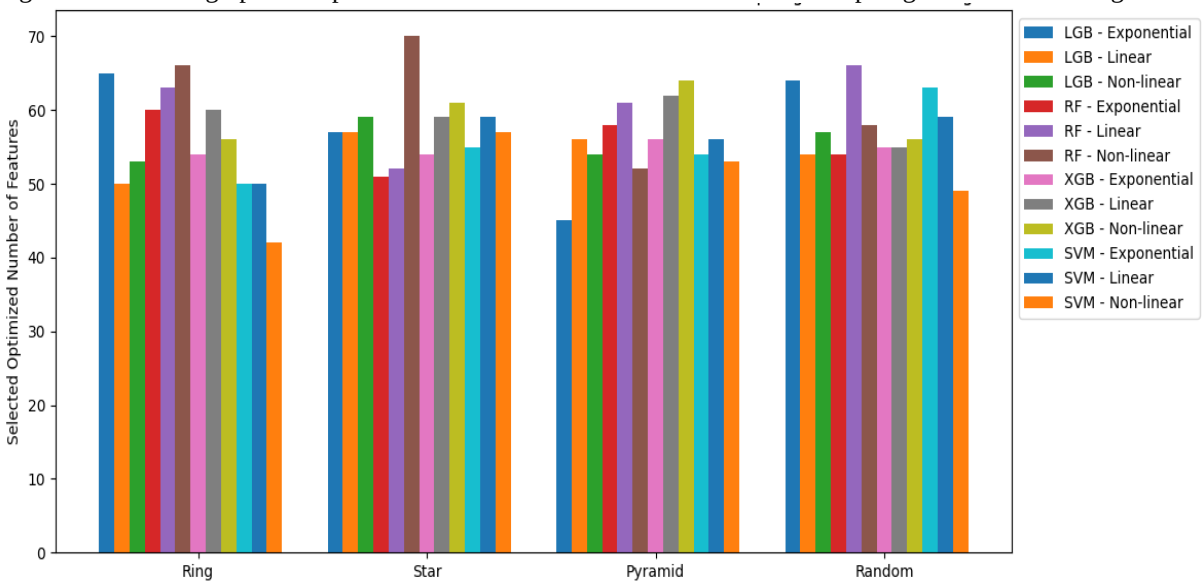


Figure 2 bar plot depicting the selected number of optimal features for the different topology- inertia weight combinations with different classifiers

3.3 Performances of feature selectionalgorithms: Table 2 shows the number of optimal features selected for different combinations of classifiers, topologies, and time-varying inertia weights from the training dataset.

Table 2.Comparative analysis of the selected number of features out of 111 extracted features for different combinations of topologies with time-varying inertia weights for different classifiers

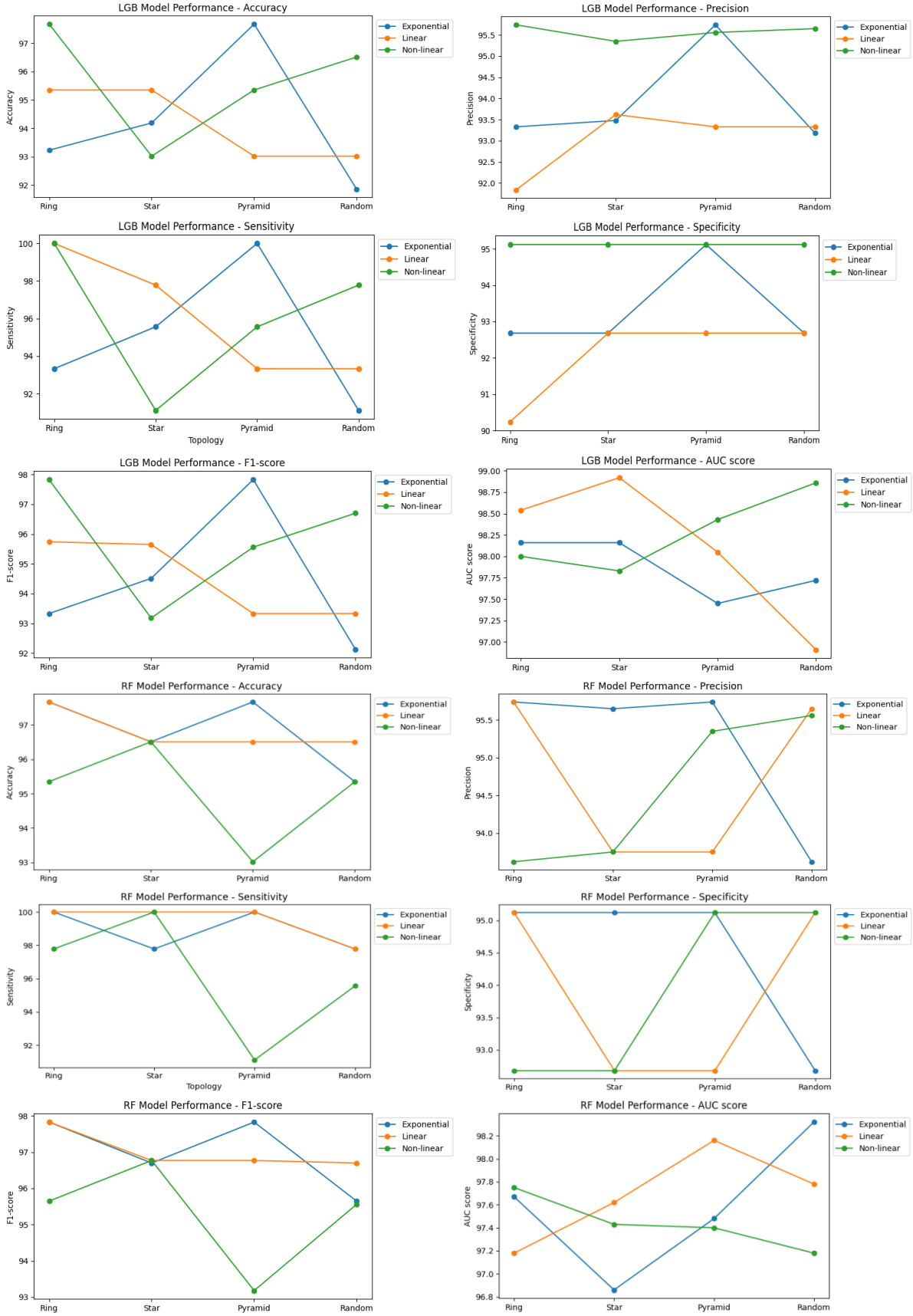
Models used for classification	Topology	Inertia weight	Selected optimized number of features
LGB	Ring	Exponential	65
		Linear	50
		Nonlinear	53
	Star	Exponential	57
		Linear	57
		Nonlinear	59
	Pyramid	Exponential	45
		Linear	56
		Nonlinear	54
	Random	Exponential	64
		Linear	54
		Nonlinear	57
RF	Ring	Exponential	60
		Linear	63
		Nonlinear	66
	Star	Exponential	51
		Linear	52
		Nonlinear	70
	Pyramid	Exponential	58
		Linear	61
		Nonlinear	52
	Random	Exponential	54
		Linear	66
		Nonlinear	58
XGB	Ring	Exponential	54
		Linear	60
		Nonlinear	56
	Star	Exponential	54
		Linear	59
		Nonlinear	61
	Pyramid	Exponential	56
		Linear	62
		Nonlinear	64
	Random	Exponential	55
		Linear	55
		Nonlinear	56
SVM	Ring	Exponential	50
		Linear	50
		Nonlinear	42
	Star	Exponential	55
		Linear	59
		Nonlinear	57
	Pyramid	Exponential	54
		Linear	56
		Nonlinear	53
	Random	Exponential	63
		Linear	59
		Nonlinear	49

3.4 Classification algorithm comparison with different feature selection algorithms: The performance of different classifiers on the test dataset with selected features in terms of accuracy, precision, sensitivity, specificity, F1-score, and AUC is shown in Table 3.

Table3:Performance of classifiers with optimal features selected using different topologies with time-varying inertia weight in BPSO

Models used for classification	Topology	Inertia weight	Accuracy	Precision	Sensitivity	Specificity	F1-score	AUC
LGB	Ring	Exponential	93.23	93.33	93.33	92.68	93.33	98.16
		Linear	95.35	91.84	100	90.24	95.74	98.54
		Nonlinear	97.67	95.74	100	95.12	97.83	98.00
	Star	Exponential	94.19	93.48	95.56	92.68	94.51	98.16
		Linear	95.35	93.62	97.78	92.68	95.65	98.92
		Nonlinear	93.02	95.35	91.11	95.12	93.18	97.83
	Pyramid	Exponential	97.67	95.74	100	95.12	97.83	97.45
		Linear	93.02	93.33	93.33	92.68	93.33	98.05
		Nonlinear	95.35	95.56	95.56	95.12	95.56	98.43
	Random	Exponential	91.86	93.18	91.11	92.68	92.13	97.72
		Linear	93.02	93.33	93.33	92.68	93.33	96.91
		Nonlinear	96.51	95.65	97.78	95.12	96.70	98.86
RF	Ring	Exponential	97.67	95.74	100	95.12	97.83	97.67
		Linear	97.67	95.74	100	95.12	97.83	97.18
		Nonlinear	95.35	93.62	97.78	92.68	95.65	97.75
	Star	Exponential	96.51	95.65	97.78	95.12	96.70	96.86
		Linear	96.51	93.75	100	92.68	96.77	97.62
		Nonlinear	96.51	93.75	100	92.68	96.77	97.43
	Pyramid	Exponential	97.67	95.00	100	95.12	97.83	97.48
		Linear	96.51	93.75	100	92.68	96.77	98.16
		Nonlinear	93.02	95.35	91.11	95.12	93.18	97.40
	Random	Exponential	95.35	95.62	97.78	92.68	95.65	98.32
		Linear	96.51	95.65	97.78	95.12	96.70	97.78
		Nonlinear	95.35	95.56	95.56	95.12	95.56	97.18
XGB	Ring	Exponential	94.19	93.48	95.56	92.68	94.51	97.99
		Linear	97.67	95.74	100	95.12	97.83	98.75
		Nonlinear	95.35	93.62	97.78	92.68	95.65	97.78
	Star	Exponential	96.51	95.65	97.78	95.12	96.70	98.48
		Linear	94.19	95.45	93.33	95.12	94.38	98.05
		Nonlinear	95.35	95.56	95.56	95.12	95.56	97.99
	Pyramid	Exponential	95.35	93.62	97.78	92.68	95.65	98.54
		Linear	93.02	93.33	93.33	92.68	93.33	98.10
		Nonlinear	91.86	91.30	93.33	90.24	92.31	97.45
	Random	Exponential	93.02	95.35	91.11	95.12	93.18	98.37
		Linear	94.19	93.48	95.56	92.68	94.51	97.40
		Nonlinear	95.35	95.62	97.78	92.68	95.65	97.94
SVM	Ring	Exponential	94.19	93.48	95.56	92.68	94.51	97.18
		Linear	96.51	93.75	100	92.68	96.77	96.91
		Nonlinear	95.35	93.62	97.78	92.68	95.65	97.54
	Star	Exponential	91.86	91.30	93.33	90.24	92.31	98.54
		Linear	94.19	93.48	95.56	92.68	94.51	97.53
		Nonlinear	88.37	92.68	84.44	92.68	88.37	96.04
	Pyramid	Exponential	93.02	93.33	93.33	92.68	93.33	96.69
		Linear	93.02	93.33	93.33	92.68	93.33	97.24
		Nonlinear	94.19	95.45	93.33	95.12	94.38	98.05
	Random	Exponential	91.86	93.18	91.11	92.68	92.14	96.91
		Linear	96.51	93.75	100	92.68	96.77	98.00
		Nonlinear	96.51	93.75	100	92.68	96.77	96.26

Figure 3 visually representsthe classification performance of different classifiers on the test dataset used in the present study based on optimal features obtained from BPSO with various combinations of topologies and time-varying inertia weights.



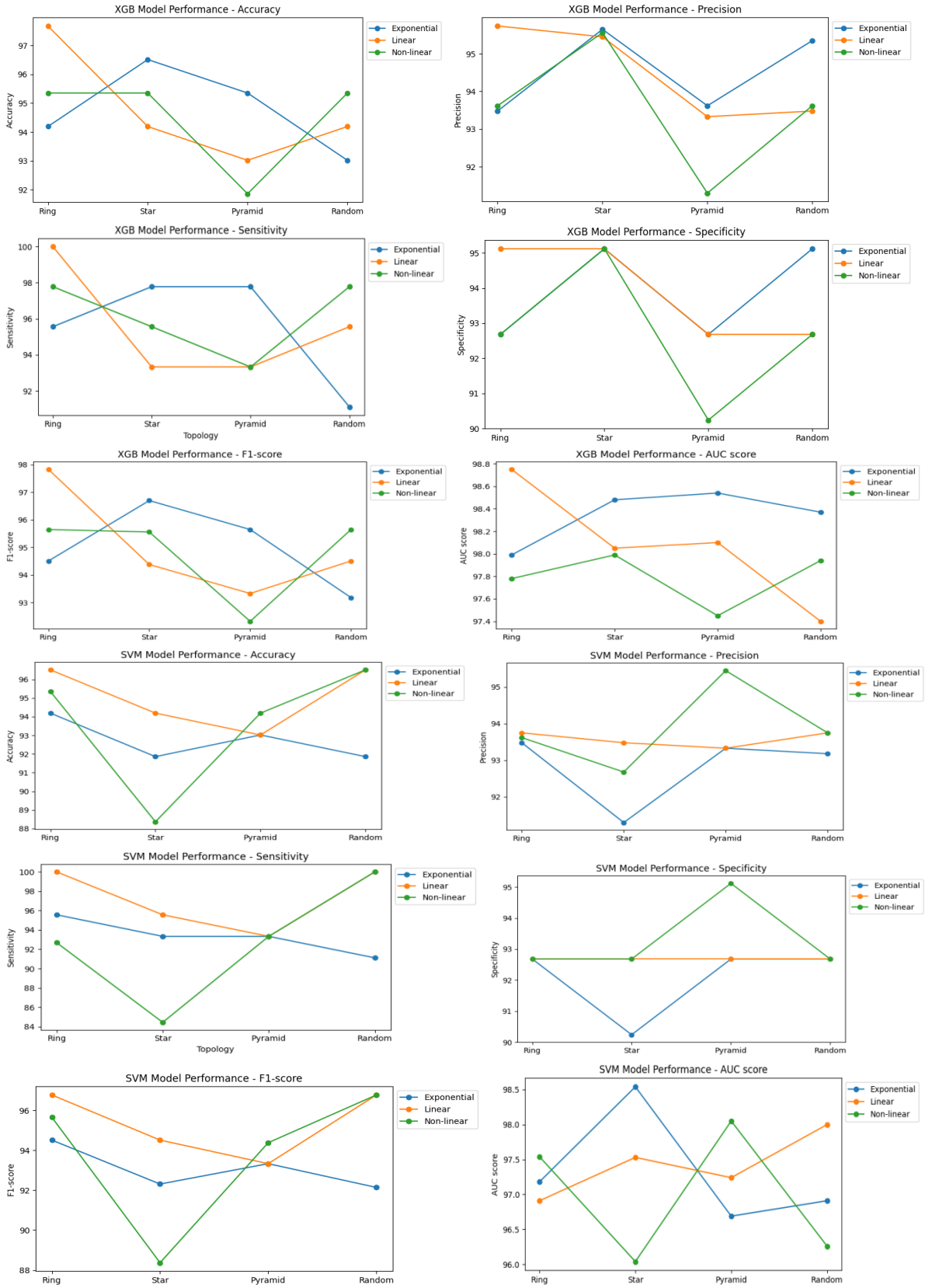


Fig 3: graphical representation of different performance matrices of classifiers on test dataset utilizing distinctive feature sets derived from varied combinations of topologies and time-varying inertia weights in BPSO

4. Discussion

The current study aimed to classify benign and malignant CT nodules from the LIDC-IDRI database using different supervised machine learning models with optimal feature selection through binary PSO. The optimal feature selection was done by exploring

various PSO topologies and different time-varying inertia weights to identify the most discriminative feature subsets. Each topology offers a unique communication mechanism among particles, influencing their exploration and exploitation in the feature space. In addition to exploring various topologies, we also incorporated time-varying inertia weights during the optimal feature selection process, which facilitated adaptability during optimization, allowing the algorithm to balance exploration and exploitation phases over time.

Overall, the best values of accuracy, precision, sensitivity, specificity, f1-score, and AUC on the test dataset were observed for the XGB classifier with ring topology and linear decreasing inertia weight. For this combination, the percentage values of accuracy, precision, sensitivity, specificity, f1-score, and AUC were 97.67, 95.74, 100, 95.12, 97.83, and 98.75, respectively.

On the meticulous evaluation of Table 3, we can divide the results broadly into two parts with an accuracy of more than 97% and more than 96%.For an accuracy of more than 97%, six combinations of different topology and inertia weights were observed for the test dataset.

For all six combinations with accuracy greater than 97%, the F1-score and specificity were also observed, with values of 97.83 and 95.12, respectively.

For ten combinations of different topologies and inertia weight, more than 96% accuracy was observed. Two combinations show AUC greater than or equal to 98%, one with random topology and nonlinear inertia weights for the LGB classifier and the second with star topology with exponential inertia weight for the XGB classifier. In all combinations where the classifiers achieved accuracy greater than 96%, the corresponding F1-score was measured to be 96.70%.

Average accuracy with different combinations of topology and inertia weight for different classifiers LGB, RF, XGB, and SVM is 94.69, 96.22, 93.79, and 94.67, respectively.

Compared to other combinations of classifiers, topologies, and inertia weights, the LGB classifier with ring topology and nonlinear decreasing inertia weight demonstrated the highest values in accuracy, precision, sensitivity, specificity, and AUC score when considering the overall performance.

Machine learning and deep learning have evoked substantial interest among researchers and medical practitioners in detecting, classifying, and predicting the malignancy of lung nodules. Many methodologies are being explored and developed rigorously to augment the precision and dependability of lung nodule analysis and diagnosis. Numerous published academic papers suggested various methodological approaches to estimate the likelihood of malignancy in lung nodules.

Saied etal. explored and developed AI methods for classifying pulmonary nodules from CT scans. They used texture Haralick and local binary pattern features in a machine learning approach, achieving an optimal AUC of 0.885 with random forest and a best accuracy of 0.819 with the support vector machine. [50]

Qiao etal. proposed a Fuse-Long Short-Term Memory-Convolutional Neural Network(F-LSTM-CNN)ensemble learning model to classify benign and malignant nodules by incorporating visual attributes and deep features to categorize benign and malignant nodules from the LIDC-IDRI dataset. They achieved accuracy, sensitivity, and specificity of 0.955, 1, and 0.937 with an AUC of 0.995 for lung nodule classification. [51]

Safta etal. classified nodules from the LIDC-IDRI dataset into malignant and benign categories based on GLCM features. In their study, the classification was performed with SVM for Multiple Instance Learning (MIL-SVM). They achieved AUC, Specificity, Sensitivity, and Accuracy of 0.9767, 0.9524, 0.9111 and 0.9310 respectively. [52]

Jena and George extracted the morphological features from the CT images of the LIDC-IDRI dataset and used a Kernel-based NonGaussian Convolutional Neural Network for classification. They achieved a classification accuracy of 87.3% .[53]

Jiang et al. proposed a novel pixel value space statistics map (PVSSM) for accurately classifying pulmonary nodules in lung cancer diagnosis in the LIDC-IDRI dataset. This study used the singular value decomposition (SVD) method to extract features from the created feature matrixes. In their study, classification accuracies of 77.3%, 80.1%, and 84.2% for KNN, random forest, and SVM classifiers were obtained, respectively. [54]

Chen et al. utilized radiomic features as input, and for the classification algorithm, they used deep attention-based MIL. They achieved a mean accuracy of 0.807 with a standard error of the mean (SEM) of 0.069, a recall of 0.870 (SEM 0.061), a positive predictive value of 0.928 (SEM 0.078), a negative predictive value of 0.591 (SEM 0.155), and AUC of 0.842 (SEM 0.074).[55]

Sahu et al. presented a computer-aided diagnosis system for risk stratification of pulmonary nodules in the CTimages of the LIDC-IDRI dataset by fusing shape and texture-based features in a machine-learning (ML) based paradigm. Using 30 dominant features from the pool of shape and texture-based features, the proposed system achieved classification accuracy, sensitivity, specificity, and AUC of 89%, 88%, 89 %, and 0.92, respectively. [56]

The results of the present study showed that the combination of different topologies combined with time-varying PSO for feature selection, with different supervised machine learning classification algorithms led to better classification performance in the context of accuracy, precision, sensitivity, specificity, and AUC score as compared to the studies as mentioned above [50-56]. Further exploration of the proposed method can help validate and enlighten the

effectiveness and adaptability of the parameters required to achieve the best result.

6. Conclusion

Combining different topologies with time-varying inertia weights has yielded a practical framework for optimal feature selection. This foundational step improved the classification performance of the classifiers by selecting quality discriminative features and discarding the least important ones. Such techniques can assist medical professionals in accurate decision-making, reducing their workloads. Despite being trained on limited data, the classifiers showed promising results with the optimal number of features. Deep learning classifiers can be prone to overfitting if trained with a small dataset, such feature selection techniques combined with different traditional machine learning classifiers can reduce such problems, making more robust systems.

7. ACKNOWLEDGMENT

The authors acknowledge the National Cancer Institute and the Foundation for the National Institutes of Health and their critical role in creating the free publicly available LIDC/IDRI Database used in this study.

Financial support and sponsorship: Nil

REFERENCES:

1. Lung cancer [Internet]. [cited 2023 Jul 30]. Available from: https://www.who.int/news-room/fact-sheets/detail/lung-cancer
2. Cancer [Internet]. [cited 2023 Jul 30]. Available from: https://www.who.int/news-room/fact-sheets/detail/cancer
3. Bjerager M, Palshof T, Dahl R, Vedsted P, Olesen F. Delay in diagnosis of lung cancer in general practice. Br J Gen Pract. 2006 Nov 1;56(532):863–8.
4. Fontana RS. The Mayo Lung Project: a perspective. Cancer. 2000 Dec 1;89(11 Suppl):2352–5.
5. van Beek EJ, Mirsadraee S, Murchison JT. Lung cancer screening: Computed tomography or chest radiographs? World J Radiol. 2015 Aug 28;7(8):189–93.
6. Oken MM, Hocking WG, Kvale PA, Andriole GL, Buys SS, Church TR, et al. Screening by chest radiograph and lung cancer mortality: the Prostate, Lung, Colorectal, and Ovarian (PLCO) randomized trial. JAMA. 2011 Nov 2;306(17):1865–73.
7. National Lung Screening Trial Research Team, Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. N Engl J Med. 2011 Aug 4;365(5):395–409.
8. Gillaspie EA, Allen MS. Computed tomographic screening for lung cancer: the Mayo Clinic experience. ThoracSurg Clin. 2015 May;25(2):121–7.
9. Jassim MM, Jaber MM. Systematic review for lung cancer detection and lung nodule classification: Taxonomy, challenges, and recommendation future works. Journal of Intelligent Systems. 2022 Jan 1;31(1):944–64.
10. Gu Y, Chi J, Liu J, Yang L, Zhang B, Yu D, et al. A survey of computer-aided diagnosis of lung nodules from CT scans using deep learning. Computers in Biology and Medicine. 2021 Oct 1;137:104806.
11. Rubin GD, Lyo JK, Paik DS, Sherbondy AJ, Chow LC, Leung AN, et al. Pulmonary nodules on multi-detector row CT scans: performance comparison of radiologists and computer-aided detection. Radiology. 2005 Jan;234(1):274–83.
12. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. Radiology. 2016 Feb;278(2):563–77.
13. Javaid M, Haleem A, Pratap Singh R, Suman R, Rab S. Significance of machine learning in healthcare: Features, pillars and applications. International Journal of Intelligent Networks. 2022 Jan 1;3:58–73.
14. Cai J, Luo J, Wang S, Yang S. Feature selection in machine learning: A new perspective. Neurocomputing. 2018 Jul 26;300:70–9.
15. Chandrashekar G, Sahin F. A survey on feature selection methods. Computers & Electrical Engineering. 2014 Jan 1;40(1):16–28.
16. Remeseiro B, Bolon-Canedo V. A review of feature selection methods in medical applications. Computers in Biology and Medicine. 2019 Sep 1;112:103375.
17. Taye M. Understanding of Machine Learning with Deep Learning: Architectures, Workflow, Applications and Future Directions. Computers. 2023 Apr 25;12:91.
18. Darwish A. Bio-inspired computing: Algorithms review, deep analysis, and the scope of applications. Future Computing and Informatics Journal. 2018 Dec 1;3(2):231–46.
19. Siddique N, Adeli H. Nature Inspired Computing: An Overview and Some Future Directions. Cognit Comput. 2015;7(6):706–14.
20. Mahalakshmi D, Balamurugan SAA, Chinnadurai M, Vaishnavi D. Nature-Inspired Feature Selection Algorithms: A Study. In: Karrupusamy P, Balas VE, Shi Y, editors. Sustainable Communication Networks and Application. Singapore: Springer Nature; 2022. p. 739–48. (Lecture Notes on Data Engineering and Communications Technologies).
21. Abiodun EO, Alabdulatif A, Abiodun OI, Alawida M, Alabdulatif A, Alkhawaldeh RS. A systematic review of emerging feature selection optimization methods for optimal text classification: the present state and prospective opportunities. Neural Comput Appl. 2021;33(22):15091–118.

22. Sivanandam SN, Deepa SN. Genetic Algorithms. In: Sivanandam SN, Deepa SN, editors. Introduction to Genetic Algorithms [Internet]. Berlin, Heidelberg: Springer; 2008 [cited 2023 Aug 1]. p. 15–37. Available from: https://doi.org/10.1007/978-3-540-73190-0_2
23. Kennedy J, Eberhart R. Particle swarm optimization. In: Proceedings of ICNN'95 - International Conference on Neural Networks. 1995. p. 1942–8 vol.4.
24. Blum C. Ant colony optimization: Introduction and recent trends. Physics of Life Reviews. 2005 Dec 1;2(4):353–73.
25. Karaboga D, Gorkemli B, Ozturk C, Karaboga N. A comprehensive survey: artificial bee colony (ABC) algorithm and applications. ArtifIntell Rev. 2014 Jun 1;42(1):21–57.
26. Mirjalili S, Mirjalili SM, Lewis A. Grey Wolf Optimizer. Advances in Engineering Software. 2014 Mar 1;69:46–61.
27. Gad AG. Particle Swarm Optimization Algorithm and Its Applications: A Systematic Review. Arch Computat Methods Eng. 2022 Aug 1;29(5):2531–61.
28. Tran B, Xue B, Zhang M. Overview of Particle Swarm Optimisation for Feature Selection in Classification. In: Dick G, Browne WN, Whigham P, Zhang M, Bui LT, Ishibuchi H, et al., editors. Simulated Evolution and Learning. Cham: Springer International Publishing; 2014. p. 605–17. (Lecture Notes in Computer Science).
29. Cheng S, Shi Y, Qin Q. Population diversity based study on search information propagation in particle swarm optimization. In: 2012 IEEE Congress on Evolutionary Computation. 2012. p. 1–8.
30. Ni Q, Deng J. A New Logistic Dynamic Particle Swarm Optimization Algorithm Based on Random Topology. The Scientific World Journal. 2013 May 30;2013:e409167.
31. Li T, Shi J, Deng W, Hu Z. Pyramid particle swarm optimization with novel strategies of competition and cooperation. Applied Soft Computing. 2022 May 1;121:108731.
32. Li HR, Gao YL. Particle Swarm Optimization Algorithm with Exponent Decreasing Inertia Weight and Stochastic Mutation. In: 2009 Second International Conference on Information and Computing Science. 2009. p. 66–9.
33. Xin J, Chen G, Hai Y. A Particle Swarm Optimizer with Multi-stage Linearly-Decreasing Inertia Weight. In: 2009 International Joint Conference on Computational Sciences and Optimization. 2009. p. 505–8.
34. Chatterjee A, Siarry P. Nonlinear inertia weight variation for dynamic adaptation in particle swarm optimization. Computers & Operations Research. 2006 Mar 1;33(3):859–71.
35. Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, et al. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. In: Advances in Neural Information Processing Systems [Internet]. Curran Associates, Inc.; 2017 [cited 2023 Aug 1]. Available from: https://proceedings.neurips.cc/paper_files/paper/2017/hash/6449f44a102fde848669bdd9eb6b76fa-Abstract.html
36. Breiman L. Random Forests. Machine Learning. 2001 Oct 1;45(1):5–32.
37. Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining [Internet]. 2016 [cited 2023 Aug 1]. p. 785–94. Available from: http://arxiv.org/abs/1603.02754
38. Cristianini N, Ricci E. Support Vector Machines. In: Kao MY, editor. Encyclopedia of Algorithms [Internet]. Boston, MA: Springer US; 2008 [cited 2023 Aug 1]. p. 928–32. Available from: https://doi.org/10.1007/978-0-387-30162-4_415
39. Clark K, Vendt B, Smith K, Freymann J, Kirby J, Koppel P, et al. The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository. J Digit Imaging. 2013 Dec 1;26(6):1045–57.
40. Armato III SG, McLennan G, Bidaut L, McNitt-Gray MF, Meyer CR, Reeves AP, et al. The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A Completed Reference Database of Lung Nodules on CT Scans. Medical Physics. 2011;38(2):915–31.
41. Armato III SG, McLennan G, Bidaut L, McNitt-Gray MF, Meyer CR, Reeves AP, et al. Data From LIDC-IDRI [Internet]. The Cancer Imaging Archive; 2015 [cited 2023 Oct 13]. Available from: https://wiki.cancerimagingarchive.net/x/rgAe
42. Hancock MC, Magnan JF. Lung nodule malignancy classification using only radiologist-quantified image features as inputs to statistical learning algorithms: probing the Lung Image Database Consortium dataset with two statistical learning methods. J Med Imaging (Bellingham). 2016 Oct;3(4):044504.
43. van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational Radiomics System to Decode the Radiographic Phenotype. Cancer Research. 2017 Oct 31;77(21):e104–7.
44. Kennedy J, Eberhart RC. A discrete binary version of the particle swarm algorithm. In: Computational Cybernetics and Simulation 1997 IEEE International Conference on

Systems, Man, and Cybernetics. 1997. p. 4104–8 vol.5.
45. Mafarja M, Jarrar R, Ahmad S, Abusnaina AA. Feature selection using binary particle swarm optimization with time varying inertia weight strategies. In: Proceedings of the 2nd International Conference on Future Networks and Distributed Systems [Internet]. New York, NY, USA: Association for Computing Machinery; 2018 [cited 2023 Aug 7]. p. 1–9. (ICFNDS '18). Available from: https://doi.org/10.1145/3231053.3231071
46. Kennedy J, Mendes R. Population structure and particle swarm performance. In: Proceedings of the 2002 Congress on Evolutionary Computation CEC'02 (Cat No02TH8600). 2002. p. 1671–6 vol.2.
47. Ni Q, Deng J. A New Logistic Dynamic Particle Swarm Optimization Algorithm Based on Random Topology. The Scientific World Journal. 2013 May 30;2013:e409167.
48. Miranda LJ. PySwarms: a research toolkit for Particle Swarm Optimization in Python. Journal of Open Source Software. 2018 Jan 11;3(21):433.
49. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research. 2011;12(85):2825–30.
50. Saied M, Raafat M, Yehia S, Khalil MM. Efficient pulmonary nodules classification using radiomics and different artificial intelligence strategies. Insights into Imaging. 2023 May 18;14(1):91.
51. Qiao J, Fan Y, Zhang M, Fang K, Li D, Wang Z. Ensemble framework based on attributes and deep features for benign-malignant classification of lung nodule. Biomedical Signal Processing and Control. 2023 Jan 1;79:104217.
52. Safta W, Farhangi MM, Veasey B, Amini A, Frigui H. Multiple Instance Learning for Malignant vs. Benign Classification of Lung Nodules in Thoracic Screening Ct Data. In: 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019). 2019. p. 1220–4.
53. Jena SR, George ST. Morphological feature extraction and KNG-CNN classification of CT images for early lung cancer detection. International Journal of Imaging Systems and Technology. 2020;30(4):1324–36.
54. Jiang H, Ma H, Qian W, Wei G, Zhao X, Gao M. A novel pixel value space statistics map of the pulmonary nodule for classification in computerized tomography images. In: 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). 2017. p. 556–9.
55. Chen J, Zeng H, Zhang C, Shi Z, Dekker A, Wee L, et al. Lung cancer diagnosis using deep attention-based multiple instance learning and radiomics. Medical Physics. 2022;49(5):3134–43.
56. Sahu SP, Londhe ND, Verma S, Singh BK, Banchhor SK. Improved pulmonary lung nodules risk stratification in computed tomography images by fusing shape and texture features in a machine-learning paradigm. International Journal of Imaging Systems and Technology. 2021;31(3):1503–18.