CANCER DIAGNOSIS & PROGNOSIS *6*: 199-213 (2026)
doi: 10.21873/cdp.10519

Genetically Optimized Modular Neural Networks for Precision Lung Cancer Diagnosis: Exploratory Study of Novel Approach

VIJAY L. AGRAWAL[1], TRUSHDEEP AGRAWAL[2,3], ARUNI GHOSE[4,5,6], SOLA ADELEKE[7,8], STERGIOS BOUSSIOS[5,9,10,11,12,13,14] and RAJENDER SINGH ARORA[3]

[1]HVPM's COET, Amravati, India;
[2]Shri Vasantrao Naik Government Medical College, Yavatmal, India;
[3]Sujan Surgical Cancer Hospital & Amravati Cancer Foundation, Amravati, India;
[4]Barts Cancer Centre, St Bartholomew's Hospital, Barts Health NHS Trust, London, U.K.;
[5]Department of Research and Innovation, Medway NHS Foundation Trust, Gillingham, U.K.;
[6]European Interdisciplinary Society for AI in Cancer Research, Milan, Italy;
[7]Cancer Centre at Guy's, Guy's and St Thomas' NHS Foundation Trust, London, U.K.;
[8]Department of Cancer Imaging, School of Biomedical Engineering & Imaging Sciences, Faculty of Life Sciences and Medicine, King's College London, London, U.K.;
[9]Department of Medical Oncology, Ioannina University Hospital, Ioannina, Greece;
[10]Faculty of Medicine, School of Health Sciences, University of Ioannina, Ioannina, Greece;
[11]AELIA Organization, Thessaloniki, Greece;
[12]Kent and Medway Medical School, University of Kent, Canterbury, U.K.;
[13]Faculty of Medicine, Health and Social Care, Canterbury Christ Church University, Canterbury, U.K.;
[14]School of Cancer & Pharmaceutical Sciences, Faculty of Life Sciences and Medicine, King's College London, London, U.K.

Abstract

Background/Aim: Lung cancer is one of the leading causes of cancer deaths. While low-dose computed tomography (CT) screening improves survival, radiological detection is increasingly challenged by a shortage of radiologists. This study aimed to develop and evaluate a novel, precise, and computationally efficient AI-based algorithm for lung cancer diagnosis using chest CT scans.

Patients and Methods: A total of 156 patient chest CT scans were utilized to form Databases I and II. We then conducted extensive feature extraction [statistics, histograms, Fast Fourier Transform (FFT), Discrete Cosine Transform (DCT), Walsh-Hadamard Transform (WHT)] and optimized classifiers [Multi Layer Perceptron (MLP), Generalized Feed

continued

Trushdeep Agrawal, Shri Vasantrao Naik Government Medical College, Waghapur Road, Palswadi Camp, Civil Lines, Yavatmal, 445001, India. E-mail: agrawaltrushdeep@gmail.com; Stergios Boussios, Department of Medical Oncology, Ioannina University Hospital, Stavrou Niarchou Avenue 45500, Ioannina, Greece. E-mail: stergiosboussios@gmail.com

Received December 17, 2025 | Revised December 29, 2025 | Accepted December 30, 2025

Forward Neural Network (GFF-NN), Modular Neural Network (MNN), Support Vector Machine (SVM)] with genetic algorithms. Performance evaluation measures employed were classification accuracy, Mean Squared Error (MSE), Area under the ROC curve (AUC), and computational efficiency.
Results: The MNN (Topology II) classifier employing FFT-based features with momentum learning achieved 100% classification accuracy during cross-validation for both Database I and Database II, consistently yielding perfect average classification accuracy across both datasets.
Conclusion: The genetically optimized MNN (Topology II) classifier shows remarkable performance in lung cancer diagnosis from CT scan images. Its ability to achieve perfect classification accuracy suggests strong potential for clinical application, offering both diagnostic precision, acting as a triage, and workload reduction in healthcare settings.

Keywords: Genetic algorithm, modular neural network, artificial intelligence, lung cancer, diagnosis.

Introduction

The world's leading cause of cancer death is lung cancer, and screening by low dose computed tomography (CT) has proven to improve mortality (1). Radiological detection of lung tumors is a difficult task, and the increasing shortage of radiologists contributes to delayed diagnosis and worse prognosis (2, 3). The use of artificial intelligence (AI) can enhance the overall outcome of the disease (4).

Genetic algorithms are general search algorithms based mostly upon the concept of evolution seen in nature. Even with today's high-performance computers, using an exhaustive search to get the optimal solution for even relatively small problems can be very expensive. Diaz *et al*. (5) utilized genetic algorithm as a method of feature (genes) selection to support vector machine and artificial neural network to classify lung cancer status of a patient. Genetic algorithm (GA) successfully identified genes that classify patient lung cancer status with notable predictive performance.

Daliri *et al*. (6) proposed a hybrid method combining a GA for feature selection with extreme learning machines (ELM) for the classification of lung cancer data. The dimension of the feature space is reduced by the GA in this scheme and the effective features are selected appropriately. The data are then fed to a fuzzy inference system (FIS), which is trained by the fuzzy ELMs approach. The method shows accuracy of 97.5% in the diagnosis of lung cancer.

Dehmeshki *et al.* (7) proposed a shape-based genetic algorithm template matching (GATM) method for the detection of nodules with spherical elements. A spherical-oriented convolution-based filtering scheme is used as a pre-processing step for enhancement. The proposed method resulted in a detection rate of approximately 90%, with the number of false positives at approximately 14.6/scan (0.06/slice). We aimed to develop an optimal classifier based on computational intelligence techniques for the precise diagnosis of lung cancer. To robustly cross-validate the proposed classifiers, two independent lung cancer databases were utilized. Although a previous study by Agrawal *et al*. (8) investigated the same dataset, the methodology adopted in the present work is fundamentally different. In this study, we explore a novel computational intelligence framework with the objective of achieving improved diagnostic performance and more robust classification outcomes.

Patients and Methods

The computational framework for this study utilized MATLAB (with all relevant toolboxes), Neurosolutions version 5.07, and XLSTAT 2011 for data processing, model development, and statistical analysis. The flow chart depicted in Figure 1 shows the process followed to carry

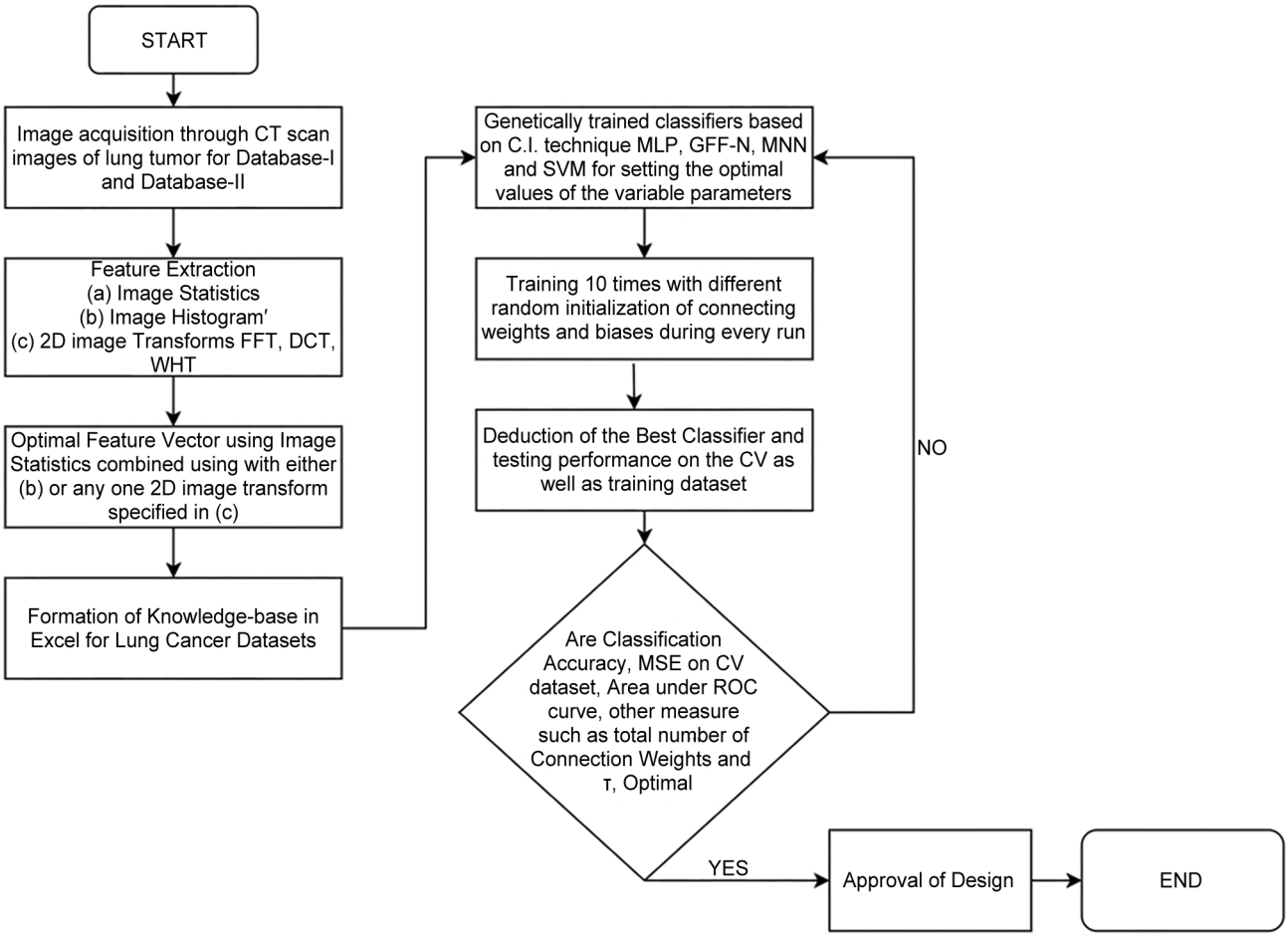

Figure 1. *Process flow diagram for genetic algorithm based classifier using computational intelligence techniques. FFT: Fast Fourier Transform; DCT: Discrete Cosine Transform; WHT: Walsh-Hadamard Transform; MLP: Multi Layer Perceptron; GFF-NN: Generalized Feed Forward Neural Network; MNN: Modular Neural Network; SVM: Support Vector Machine; CV: cross validation; MSE: Mean Squared Error.*

out the research work using genetic algorithms of artificial intelligence techniques.

Image acquisition. Images of chest CT scans were acquired from multiple sources. These images have been segregated into two databases as follows:

Database I: The Database I contains 80 Lung CT scan images of 80 patients. After processing the images, region of interest (ROI) in relation to tumor were identified and located for lung tumors and used from Database I.

Database II: The Database II comprises 76 Lung CT scan images of 76 patients. After processing the images, ROI in relation to tumor were identified and located for lung tumors and used from Database II and the benign images were used from the set of Database I for the development of the system. Annotations of the lung tumor CT scan images in Database II were performed with the assistance of local medical experts, and all tumor samples in Database II were confirmed to be malignant.

Nature of decision boundary of classifier and feature extraction. After acquisition and preprocessing of CT scan images of lung tumors, feature vectors were determined for both Database I and II. In order to form the feature

vector for each tumor image, image statistics parameters, image histogram or transform domain features were used. Feature vector is denoted by feature vector (FV).

FV=[TD1, TD2, ……. TD128, Average, SD, Entropy, Contrast, Correlation, Energy, homogeneity]

Where the Transform Domain (TD) may correspond to the Discrete Cosine Transform (DCT), Fast Fourier Transform (FFT), Walsh–Hadamard Transform (WHT), or histogram coefficients, resulting in a FV of dimension 1×135. In the case of FFT-based features, the FV dimension is reduced to 1×71, as only the first 64 FFT coefficients are retained.

For proper understanding of the input feature space and the nature of the decision boundary separating benign samples from the malignant samples, typical scatter plots were scrupulously examined, where one input feature was plotted against another feature. There were several permutations and combinations of such input features, and this might lead to numerous scatter plots for Database I as well as Database II.

Figure 2 depicts a scatter plot representing the relationship between average and entropy features for database II. Similar plots were obtained for Database I. Meticulous inspection of these scatter plots showed that the decision boundary discriminating between benign and malignant class is not linear. Seemingly, it was highly nonlinear and complex because of partial or complete overlapping of features for both the classes. Such a complex and highly nonlinear decision boundary can't be estimated by any statistical classifier of linear discriminant analysis. In addition, no mathematical equation can be assigned to two classes namely, malignant and benign was intractable and an ill-posed classification problem. Only the classifier based on the neural networks and support vector machines endowed with innovative computational intelligence techniques have proven remarkable ability to solve such complex pattern recognition problems.

The four different types of knowledge bases developed earlier for Database I as well as database II are not only based on:

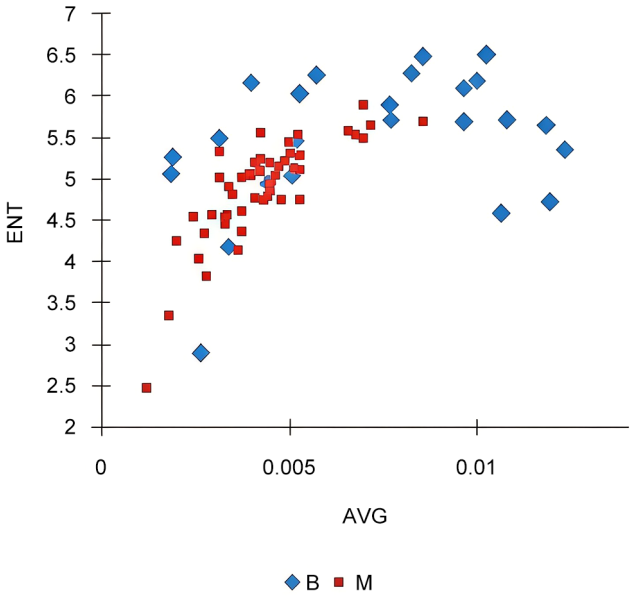

Figure 2. *Scatter plot between average (AVG) vs. entropy (ENT) feature for Database II. B: Benign; M: malignant lung tumor.*

(i) Two Dimensions (2D)-DCT coefficients (128) and Image Statistics parameters (7)

The knowledge base created in Excel is of dimension 80×135 (Database I)

The knowledge base created in Excel is of dimension 76×135 (Database II)

(ii) 2D-FFT coefficients (64) and Image Statistics parameters (7)

The knowledge base created in Excel is of dimension 80×71 (Database I)

The knowledge base created in Excel is of dimension 76×71 (Database II)

(iii) 2D-WHT coefficients (128) and Image Statistics parameters (7)

The knowledge base created in Excel is of dimension 80×135 (Database I)

The knowledge base created in Excel is of dimension 76×135 (Database II)

(iv) Image Histogram coefficients (128) and Image Statistics parameters (7)

The knowledge base created in Excel is of dimension 80×135 (Database I)

Table I. *Experimental results for determining Data Partitioning percentage.*

Tagging in %		LR	PEs	Testing			
Training	CV			CV		Training	
				B	M	B	M
50	50	MOM	23	40	100	5.56	100
		CG	18	40	100	16.67	100
		QP	32	60	91.42	50	95.23
		DBD	23	20	100	5.56	100
70	30	MOM	11	20	100	5.56	100
		CG	25	40	100	100	100
		QP	11	40	94.73	5.56	100
		DBD	1	20	100	5.56	100
80	20	MOM	6	50	90	29.41	97.82
		CG	49	50	90	29.41	97.82
		QP	24	16.67	100	23.52	100
		DBD	8	83.33	90	47.05	71.73
90	10	MOM	28	100	75	68.42	55.76
		CG	36	75	75	42.1	95.23
		QP	5	75	75	36.84	100
		DBD	7	50	75	36.84	96.15

ACA: Average classification accuracy; ANN: artificial neural network; Avg: average; Az: area under ROC curve; B: benign tumor; CA: classification accuracy; CG: conjugate gradient; CI: computational intelligence; Co: covariance; CV: cross validation; DBD: Delta-Bar-Delta; DCT: discrete cosine transform; ENG: energy ENT entropy; FFT: fast fourier transform; FIT: fitness; GFF-NN: generalized feed forward neural network; LL: lower layer; LR: learning rule; M: malignant tumor; MLP: multi layer perceptron; MNN: modular neural network; MOM: momentum; MSE: mean square error; N: connection weights and biases; NMSE: normalized mean square error; NN: neural network; P: Pearson; PCA: Principal Component Analysis; PCs: Principal Components; Pes: processing elements; QP: quick propagation; ROC: receiver operating characteristics; RoI: region of interest; ROU: Roulette; SA: sensitivity analysis; SD: standard deviation; SVM: support vector machine; TF: transfer function; TOPO: topology; UL: upper layer; WHT: Walsh-Hadamard Transform; τ: time elapsed per epoch per exemplar.

Table II. *Combination of advanced genetic options with selection, crossover and mutation used to carry out experimentation on database I and database II.*

Progression	Operators selection	Selection basis	Crossover	Mutation
Generational	Roulette	Rank	Arithmetic Heuristic With no. of tries =3 One Point Two point Uniform With mixed ratio =0.5	Uniform With mutation probability of = 0.01
		Fitness	Arithmetic Heuristic With No. of tries =3 One Point Two point Uniform With mixed ratio =0.5	Uniform With mutation probability of = 0.01

The knowledge base created in Excel is of dimension 76×135 (Database II) but also carefully selected image statistics parameters.

The above knowledge bases were used to simulate Multi-layer Perceptrons (MLP), Generalised Feed Forward Neural Networks (GFF-NN), Modular Neural Networks (MNN) and Support Vector Machines (SVM) using different learning rules. These simulations were performed on knowledge bases constructed from various transform domains such as DCT, FFT and WHT for both database I and database II. In addition, knowledge bases containing Histogram coefficients were also used for simulation of the neural networks.

Selection of activation functions. In order to determine the optimal activation function for experimentation on the DCT, FFT, WHT and HISTOGRAM knowledge bases, experiments were carried out using MLP neural network to achieve maximum average classification accuracy on the Cross Validation (CV) dataset.

The best results obtained for activation function with respect to knowledge bases are listed as below:

i) DCT Knowledge base: TANH; ii) FFT Knowledge base: LINTANH; iii) WHT Knowledge base: TANH; iv) HIST Knowledge base: LINTANH.In this research work, the entire experimentation was carried out using the above mentioned activation function for the classifier with respect to each knowledge base.

Choice of data-partitions. To decide the appropriate data partitioning percentage for the classifiers, experimentation

Table III. *Computer simulation results for genetically optimized and trained neural networks on database I.*

Transform domain	CI based classifier	LR	Topo	PEs		ACA				CV						FITNESS	Total No. of connection weights & biases 'N'	Time elapsed per epoch per exemplar 'τ'
				UL	LL	CV		TR		MSE			NMSE					
						B	M	B	M	B	M	AVG	B	M	AVG			
DCT	DCT-GEN-MNN(80-20)-ROU-RANK-ARITH-UNI-TANH	MOM	I	10	45	100	100	100	100	0.0398	0.0102	0.0250	0.1700	0.0436	0.1068	0.0405	7,594	1.0 ms
DCT	DCT-GEN-MNN(80-20)-ROU-RANK-TWO POINT-UNI-TANH	MOM	III	9	19	100	100	100	100	0.0020	0.0020	0.0020	0.0086	0.0085	0.0085	0.0322	3,868	0.7854 ms
DCT	DCT-GEN-MNN(80-20)-ROU-RANK-TWO POINT-UNI-TANH	MOM	I	28	19	100	100	100	100	0.0308	0.0432	0.0370	0.1315	0.1845	0.1580	0.0244	6,490	0.5754 ms
DCT	DCT-GEN-MNN(80-20)-ROU-FIT-ARITH-UNI-TANH	DBD	I	12	38	100	100	100	100	0.0473	0.0572	0.0523	0.2018	0.2441	0.2230	0.0270	6,904	0.6280 ms
DCT	DCT-GEN-MNN(80-20)-ROU-FIT-ONE POINT-UNI-TANH	MOM	I	3	4	100	100	100	100	0.0335	0.0152	0.0243	0.1428	0.0647	0.1038	0.0424	970	2.02 ms
DCT	DCT-GEN-MNN(80-20)-ROU-FIT-ONE POINT-UNI-TANH	DBD	I	17	3	100	100	100	100	0.0335	0.0152	0.0243	0.1428	0.0647	0.1038	0.0788	2,764	0.9143 ms
DCT	DCT-GEN-GFF(80-20)-ROU-RANK-HEU(TANH)	MOM		4		100	100	100	100	0.0409	0.0225	0.0317	0.1746	0.0959	0.1352	0.0514	554	0.7684 ms
DCT	DCT-GFF(80-20)-ROU-RANK-TWO POINT(TANH)	MOM		45		100	100	100	100	0.0122	0.0033	0.0077	0.0519	0.0139	0.0329	0.0125	6,074	0.4414 ms
DCT	DCT-GFF(80-20)-ROU-RANK-TWO POINT(TANH)	CG		3		100	100	100	100	0.0674	0.0593	0.0593	0.2876	0.2531	0.2703	0.1026	416	0.4261 ms
DCT	DCT-GFF(80-20)-ROU-RANK-TWO POINT(TANH)	DBD		38		100	100	100	100	0.0062	0.0068	0.0068	0.0265	0.0290	0.0278	0.0105	5,246	0.3915 ms
DCT	DCT-GFF(80-20)-ROU-FIT-ARITH(TANH)	QP		26		100	100	100	100	0.0678	0.0892	0.0785	0.2892	0.3807	0.3349	0.0873	3,590	0.4301 ms
DCT	DCT-MLP(80-20)-ROU-FIT-UNIFORM(TANH)	DBD		26		100	100	100	100	0.0348	0.0496	0.0422	0.1483	0.2117	0.1800	0.0683	3,590	0.1359 ms
FFT	FFT-MLP(80-20)-ROU-RANK-HEU(LINTANH)	DBD		46		100	100	100	100	0.0902	0.0703	0.0802	0.3848	0.3000	0.3424	0.0647	3,406	1.16 ms
FFT	FFT-GEN-MNN(80-20)-ROU-RANK-TWO POINT-UNI-LINTANH	MOM	II	23	4	100	100	100	100	0.1151	0.0696	0.0923	0.4911	0.2968	0.3940	0.0841	2,002	0.5729 ms

Table III. *Continued*

Table III. *Continued*

Transform domain	CI based classifier	LR	Topo	PEs		ACA				CV						Total No. of connection weights & biases 'N'	Time elapsed per epoch per exemplar 'τ'
				UL	LL	CV		TR		MSE			NMSE				
						B	M	B	M	B	M	AVG	B	M	AVG	FITNESS	
WHT	WHT-MNN(80-20)-ROU-RANK-TWO POINT	QP	I	6	47	100	100	100	100	0.0613	0.0601	0.0607	0.2617	0.2565	0.2591	0.0984	7,318	0.6682 ms
HIST	HIST-MLP(80-20)-ROU-FIT-UNIFORM(LINTANH)	CG			14	100	100	100	100	0.0280	0.0321	0.0301	0.1195	0.1371	0.1283	0.1402	1,934	0.3075 ms

ACA: Average classification accuracy; ANN: artificial neural network; Avg: average; Az: area under ROC curve; B: benign tumor; CA: classification accuracy; CG: conjugate gradient; CI: computational intelligence; Co: covariance; CV: cross validation; DBD: Delta-Bar-Delta; DCT: discrete cosine transform; ENG: energy ENT entropy; FFT: fast fourier transform; FIT: fitness; GFF-NN: generalized feed forward neural network; LL: lower layer; LR: learning rule; M: malignant tumor; MLP: multi layer perceptron; MNN: modular neural network; MOM: momentum; MSE: mean square error; N: connection weights and biases; NMSE: normalized mean square error; NN: neural network; P: Pearson; PCA: Principal Component Analysis; PCs: Principal Components; Pes: processing elements; QP: quick propagation; ROC: receiver operating characteristics; RoI: region of interest; ROU: Roulette; SA: sensitivity analysis; SD: standard deviation; SVM: support vector machine; TF: transfer function; TOPO: topology; UL: upper layer; WHT: Walsh-Hadamard Transform; τ: time elapsed per epoch per exemplar.

was carried out on MLP neural network with respect to DCT knowledge base and the obtained results are portrayed in Table I.

The work was carried out for database I and database II following Advanced Genetic Options as portrayed in Table II for genetic optimization of neural networks, with a fixed crossover probability of 0.9 across all evaluated combinations.

The range and values of the other parameters during simulation of the genetically optimized neural network are as follows:

No. of Epochs=10,000; Population Size=50; Maximum Generations=100; Maximum Evolution time=60 min; Cross-validation termination=Terminate after 100 epochs without any improvement; Step size optimization=0 to 1; Momentum optimization=0 to 1; The number of processing elements (PEs) was optimized within a range of 1 to 50. Table III presents only those classifiers that achieved 100% average classification accuracy (ACA) on the cross-validation dataset.

Based on the observations of the above computer simulation experiments and the comparison among different classifiers, it was noticed that the knowledge base formed using 2D-FFT coefficients resulted into the best classifier performance, when Modular Neural Network (Topology II) with Momentum learning rule was employed (Figure 3).

This MNN was further re-configured using advanced genetic options Roulette-Rank-Two Point-uniform specifications and the resultant performance measures observed were:

The Average Classification Accuracy on cross-validation: 100%; The number of connection weights and biases, "N": 2,002; The Average Minimum Square Error (MSE) on cross-validation: 0.092; The time elapsed per epoch per exemplar, 'τ': 0.5729 ms; The fitness of chromosomes: 0.0841. The snapshot for the settings of different training parameters for the above-mentioned best classifier is shown in Figure 4.

For database I, the best Neural Network (NN) based classifier was MNN (Topology II). The knowledge base used

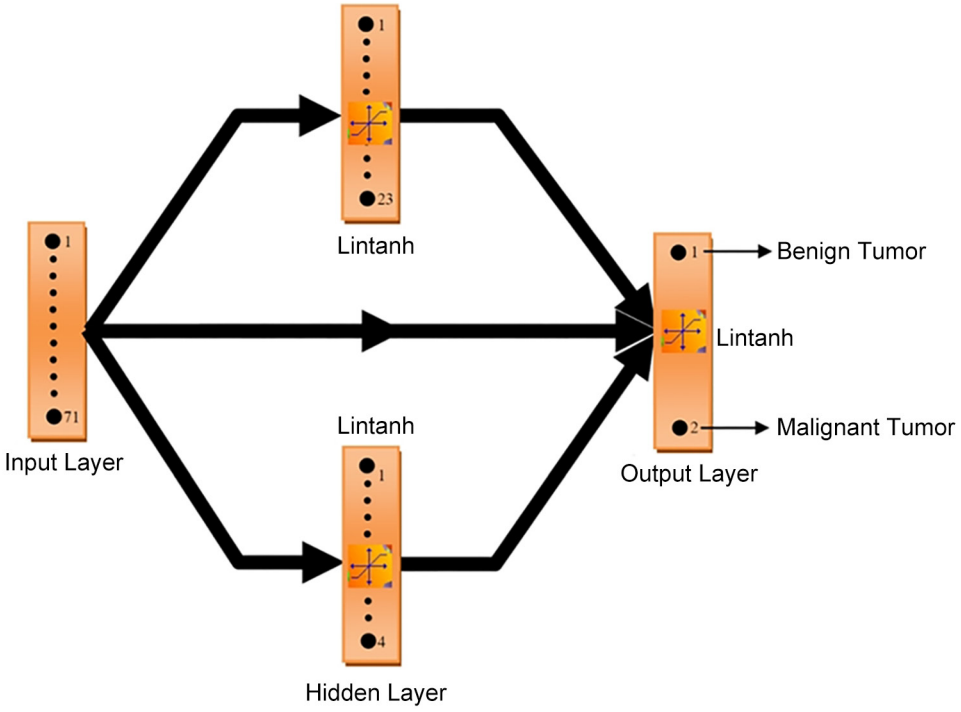

Figure 3. *Modular Neural Network (Topology II-Single Hidden Layer).*

to train and cross-validate this classifier was using 2D-FFT coefficients. The size of the knowledge base was 80×71, which implies that there are 71 different input features.

Similarly, the experimentation was carried out on database II and the results of various classifiers with the use of different learning rules and different knowledge bases of database II were investigated and only the results showing 100% classification accuracy are tabulated as in Table IV.

As MNN (Topology II) was identified as an optimal classifier on FFT knowledge base for database I, the same MNN (Topology II) classifier was applied on database II to obtain optimum results. The best results obtained for database II were:

The Average Classification Accuracy on cross-validation: 100%; The number of connection weights and biases, 'N': 3,556; The average MSE on cross-validation: 0.0028; The time elapsed per epoch per exemplar, 'τ': 0.1268 ms; The area under ROC curve for test on CV Dataset, "AZ": 1.0; The fitness of chromosomes: 0.004548.

Results

From the meticulous observation of Table V, it is obvious that the use of single hidden layer MNN (Topology II) with momentum learning rule on FFT knowledge base for database I and database II resulted in the reasonable and optimal classifier based on Computational Intelligence (C.I.) techniques for the diagnosis of lung cancer as compared to the single hidden layer MLP based classifier.

Thus, MNN (Topology II) based classifier, which is genetically optimized and trained, is recommended as the optimal classifier for the diagnosis of lung cancer from lung tumor CT scan images (Figure 5).

The observation of the above results for database I and database II indicated that two different sets of optimal classifiers with different configurations are required to be designed for best results. To provide a common optimal classifier for different databases, further experimentation was carried out on FFT knowledge base for database I and database II in two stages.

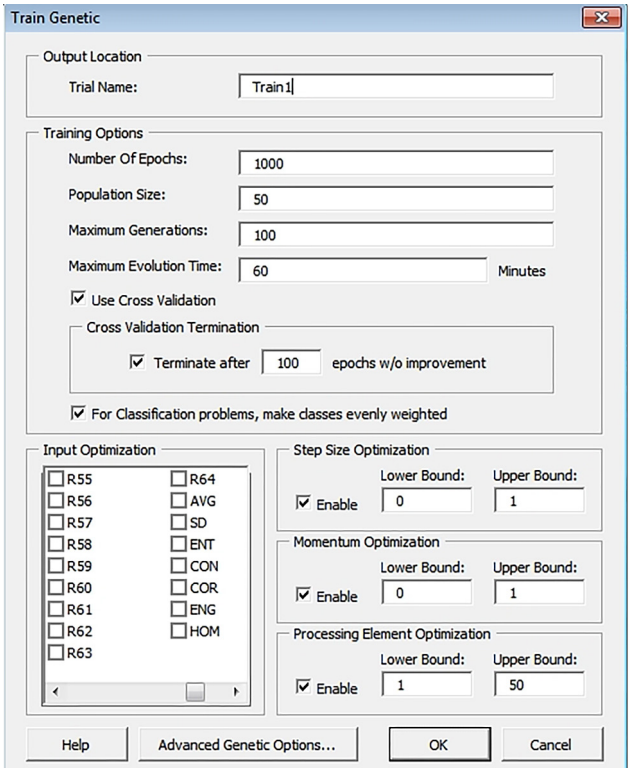

Figure 4. *Parameter settings for Modular Neural Network (MNN) with topology II with Fast Fourier Transform (FFT) transform domain coefficients based features with respect to database I.*

Stage I: As MNN (Topology II) was identified as an optimal classifier on FFT knowledge base for database I, therefore, the same MNN (Topology II) classifier was applied on database II to get optimum results. The best results obtained for database II were:

The Average Classification Accuracy on cross-validation: 100%; The number of connection weights and biases, 'N': 3,556; The average MSE on cross-validation: 0.0028; The time elapsed per epoch per exemplar, 'τ': 0.1268 ms; The area under ROC curve for test on CV Dataset, "AZ": 1.0; The fitness of chromosomes: 0.004548.

Figure 6 shows the plot of area under curve of best Classifier for database II. Area under ROC=1.0; Area under convex hull of an ROC curve=1.0.

Stage II: The MLP was previously identified as an optimal classifier on FFT knowledge base for database II, therefore, the same MLP classifier was applied to database I to obtain optimum results. The best results obtained for database I were:

The Average Classification Accuracy on cross-validation: 100%; The number of connection weights and biases, 'N': 3,406; The Average MSE on cross-validation: 0.03993; The time elapsed per epoch per exemplar, 'τ': 1.12 ms; The area under ROC curve for Test on CV Dataset, "AZ": 0.98; The fitness of chromosomes: 0.064686. Figure 7 depicts the plot of area under curve of best Classifier for database I (Test on CV Dataset).

The area under the ROC curve was 0.98, while the area under the convex hull of the ROC curve was 0.99.

In this study, we propose a novel genetically optimized classifier for lung cancer detection using chest CT images. Based on the meticulous observation of the obtained result, the MNN (Topology II) based classifier, which is genetically optimized and trained, is recommended as the optimal classifier for the diagnosis of lung cancer from the lung tumor CT scan images.

Discussion

AI when integrated with life sciences has ample potential for cancer research and might help in providing better patient outcomes and interventions (9). Diagnosis is a crucial step in providing better results. AI can be a way to help with the increasing shortage of radiologists and AI is doing that without replacing the traditional diagnostic decision steps, and without influencing radiologist's decisions (10). However, successful AI integration is obstructed and needs a lot of effort from the staff as there are major disagreements regarding autonomous AI usage across surveys, pre-, during and post-implementation (11).

Initially developed CNN by Ciompi *et al.,* showed a diagnostic accuracy that was comparable to that of an experienced radiologist (12). Nasrullah *et al*., automated lung nodule detection and classification, and showed reduced misdiagnosis and false positives in early-stage (13). Shen *et al*., produced more interpretable lung cancer

Table IV. *Computer simulation results for genetically optimized and trained neural networks on database II.*

Transfer domain	CI based Classifier	LR	TOPO	PEs		ACA				CV						FITNESS	Total No. of connection weights & biases 'N'	Time elapsed per epoch per exemplar 'τ"
				UL	LL	CV		TR		MSE			NMSE					
						B	M	B	M	B	M	AVG	B	M	AVG			
DCT	DCT-GEN-MNN(80-20)-ROU-RANK-ARITH-UNI-TANH	MOM	1	34	12	100	100	100	97.62	0.00578	0.00431	0.00505	0.02600	0.01941	0.02270	0.81910	6,352	0.1567 ms
DCT	DCT-GEN-MNN(80-20)-ROU-RANK-TWO POINT-UNI-TANH	MOM	III	27	35	100	100	100	100	0.01769	0.00363	0.01066	0.07961	0.01635	0.04798	0.01727	8,560	0.3332 ms
DCT	DCT-GEN-MNN(80-20)-ROU-RANK-TWO POINT-UNI-TANH	MOM	I	2	1	100	100	100	100	0.00322	0.00276	0.00299	0.01450	0.01242	0.01346	0.00484	418	0.1590 ms
DCT	DCT-GEN-MNN(80-20)-ROU-FIT-ARITH-UNI-TANH	DBD	I	15	25	100	100	100	100	0.00179	0.00225	0.00202	0.00803	0.01014	0.00909	0.00327	5,524	0.2385 ms
DCT	DCT-GEN-MNN(80-20)-ROU-FIT-ONE POINT-UNI-TANH	MOM	I	6	40	100	100	94.74	100	0.00426	0.00371	0.00398	0.01915	0.01668	0.01792	0.00645	6,352	0.2345 ms
DCT	DCT-GEN-MNN(80-20)-ROU-FIT-ONE POINT-UNI-TANH	DBD	I	15	23	100	100	100	100	0.00300	0.00346	0.00323	0.01348	0.01558	0.01453	0.00523	5,248	0.1541 ms
DCT	DCT-GEN-GFF(80-20)-ROU-RANK-HEU-(TANH)	MOM		11		100	100	100	100	0.00792	0.01027	0.00910	0.03566	0.04622	0.04094	0.01474	1,520	0.1735 ms
DCT	DCT-GFF(80-20)-ROU-RANK-TWO POINT(TANH)	MOM		37		100	100	100	100	0.00294	0.00302	0.00298	0.01324	0.01358	0.01341	0.00483	5,108	0.2367 ms
DCT	DCT-GFF(80-20)-ROU-RANK-TWO POINT(TANH)	CG		9		100	100	100	97.62	0.00765	0.04977	0.02871	0.03442	0.22397	0.12919	0.04651	1,254	0.2046 ms
DCT	DCT-GFF(80-20)-ROU-RANK-TWO POINT(TANH)	DBD		48		100	100	84.21	100	0.00625	0.00605	0.00615	0.02814	0.02724	0.02769	0.00405	6,626	0.1765 ms
DCT	DCT-GFF(80-20)-ROU-FIT-ARITH(TANH)	QP		37		100	100	100	100	0.00294	0.00302	0.00298	0.01324	0.01358	0.01341	0.03414	5,108	0.1846 ms
DCT	DCT-MLP(80-20)-ROU-FIT-UNIFORM(TANH)	DBD		41		100	100	100	100	0.01578	0.00909	0.01243	0.07100	0.04090	0.05595	0.02014	5,660	0.2459 ms

Table IV. *Continued*

Table IV. *Continued*

Transfer domain	CI based Classifier	LR	TOPO	PEs		ACA				CV						FITNESS	Total No. of connection weights & biases 'N'	Time elapsed per epoch per exemplar 'τ"
				UL	LL	CV		TR		MSE			NMSE					
						B	M	B	M	B	M	AVG	B	M	AVG			
FFT	FFT-MLP(80-20)-ROU-RANK-HEU(LINTANH)	DBD		4		100	100	100	100	0.00055	0.00058	0.00056	0.00247	0.00260	0.00254	0.00091	298	0.0623 ms
FFT	FFT-GEN-MNN(80-20)-ROU-RANK-TWO POINT-UNI-LINTANH	MOM	II	8	40	100	100	89.47	97.62	0.00271	0.00290	0.00281	0.01221	0.01305	0.01263	0.00455	3,556	0.1268 ms
WHT	WHT-MNN(80-20)-ROU-RANK-TWO POINT	QP	I	16	3	100	100	100	88.10	0.00808	0.01481	0.01145	0.03637	0.06666	0.05152	0.01855	2,626	0.2977 ms
HIST	HIST-MLP(80-20)-ROU-FIT-UNIFORM(LINTANH)	CG		16		100	100	100	92.86	0.01045	0.01419	0.01232	0.04703	0.06387	0.05545	0.01996	2,210	0.0718 ms

ACA: Average classification accuracy; ANN: artificial neural network; Avg: average; Az: area under ROC curve; B: benign tumor; CA: classification accuracy; CG: conjugate gradient; CI: computational intelligence; Co: covariance; CV: cross validation; DBD: Delta-Bar-Delta; DCT: discrete cosine transform; ENG: energy ENT entropy; FFT: fast fourier transform; FIT: fitness; GFF-NN: generalized feed forward neural network; LL: lower layer; LR: learning rule; M: malignant tumor; MLP: multi layer perceptron; MNN: modular neural network; MOM: momentum; MSE: mean square error; N: connection weights and biases; NMSE: normalized mean square error; NN: neural network; P: Pearson; PCA: Principal Component Analysis; PCs: Principal Components; Pes: processing elements; QP: quick propagation; ROC: receiver operating characteristics; RoI: region of interest; ROU: Roulette; SA: sensitivity analysis; SD: standard deviation; SVM: support vector machine; TF: transfer function; TOPO: topology; UL: upper layer; WHT: Walsh-Hadamard Transform; τ: time elapsed per epoch per exemplar.

Table V. *Performance measures of optimal classifiers for database I and database II.*

DATA-BASE	Genetically optimized MLP						Genetically optimized MNN (TOPOLOGY II)					
	PE's	N	τ (mS)	ACA %	MSE on CV	Fitness	PE's	N	τ (mS)	ACA %	MSE on CV	Fitness
DATA-BASE-I	46	3,406	1.12	100	0.03993	0.064686	UL : 23 LL : 4	2,002	0.5729	100	0.092	0.0841
DATA-BASE-II	4	298	0.0623	100	0.00056	0.00091	UL : 8	3,556	0.1268	100	0.0028	0.004548

PE's: Processing elements; N: connection weights and biases; τ: time elapsed per epoch per exemplar; ACA: Average Classification Accuracy; MSE: Mean Square Error; CV: Cross Validation; UL: upper layer, LL: lower layer.

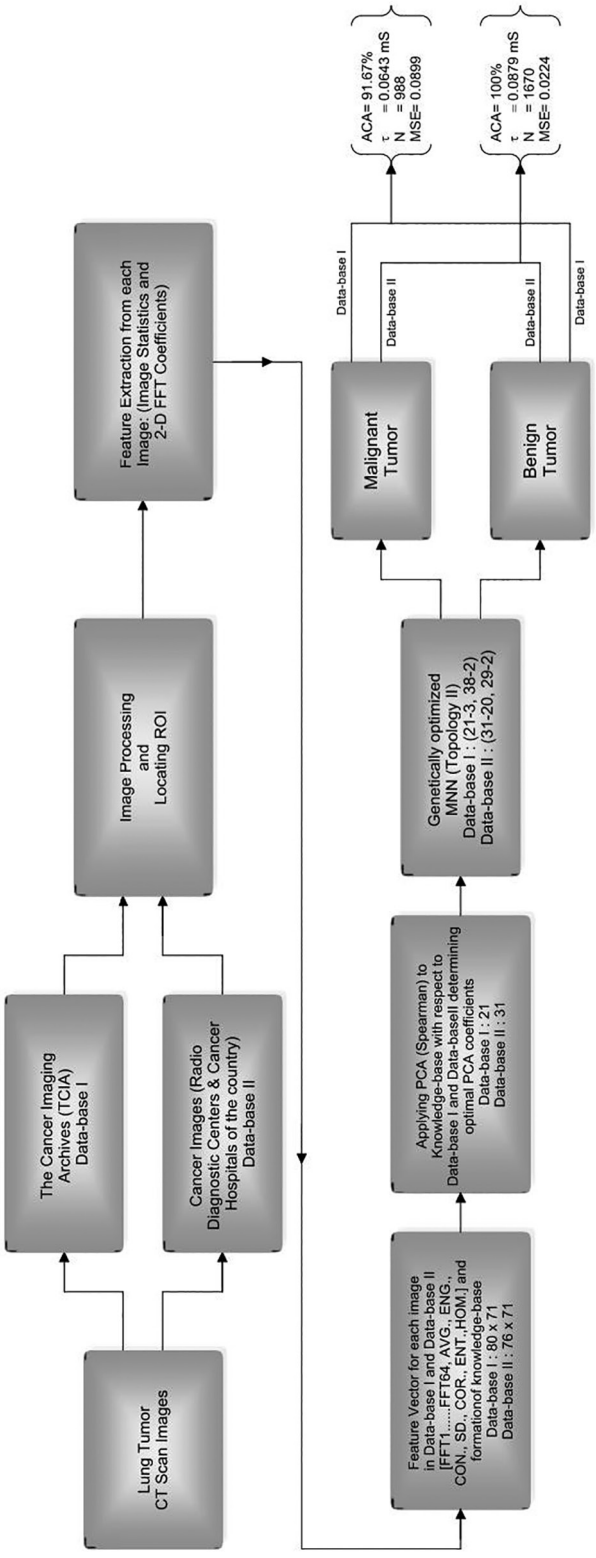

Figure 5. *Genetically optimized Modular Neural Network Classifier using 2D- Fast Fourier Transform (FFT) coefficients and image statistics parameters.*

predictions and achieved significantly better results than a 3D CNN alone (14). Additional contributions that have advanced knowledge in this field include the work of Cui *et al.* (15), whose model demonstrated results highly consistent with expert radiologists in lung nodule classification, and Siddiqui *et al.*, who identified the optimal network configuration as (128–128–20) (16). The majority of these models used the LUNA 16 database, and/or LIDC-IDRI (17, 18). The TCIA lung cancer database has also been used by some of the models (19).

Our novel approach uses genetically optimized modular neural network to achieve better results while trying to be computationally efficient and maintaining the quality of outputs. The results we present are not only aligned with the findings of CNN based models but also add to them (20). Our 100% results though remarkable should be interpreted with caution; this study was conducted with a small sample size compared to its predecessors. Further, the future studies on the model should be conducted as a multi-centered, large sample size, prospective study, similar to the work of Fockens, Kiki *et al*. (21).

While most of these studies utilized the computational strength of the CNN, our model attempted to focus on the GA-MNN. The work of Elayaraja *et al*., has shown excellent results in case of cervical cancer using genetic algorithm, reaching a sensitivity of 99.09%, specificity of 99.39%, and accuracy of 99.36% (22). Similar results were seen in the case of pancreatic cancer by Li *et al*., where genetically optimized back propagated AI was utilized for diagnosis and prognosis (23). In case of lung cancer diagnosis, a recent study showed that AI driven GA-optimized lung cancer segmentation has the potential to provide precise lung cancer diagnosis and at the same time, reduce healthcare disparities in resource-limited settings (24).

Our study aimed to further motivate scientists to apply genetically optimized MNN to their future studies on lung cancer and produce more research that can lead to integration of this technology to a multi-modal approach that helps create patient specific management (25). Efforts to develop multi-modal AI for medical application have been done in case of thyroid carcinoma by Yu *et al.* (26).

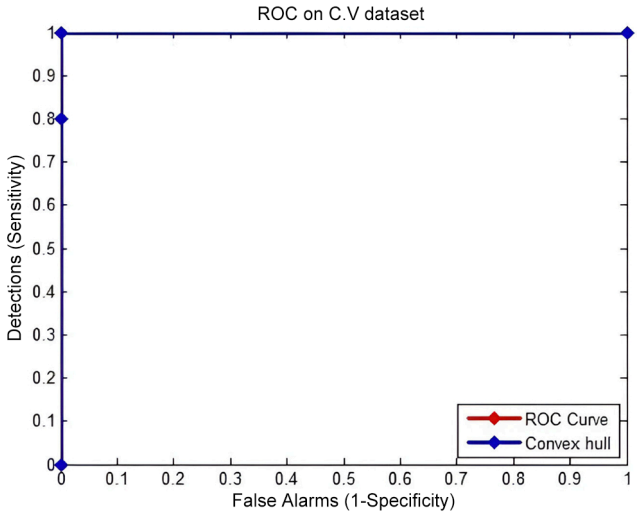

Figure 6. *Receiver operating characteristics curve on Modular Neural Network best classifier for database II [Test on cross validation (C.V) Dataset].*

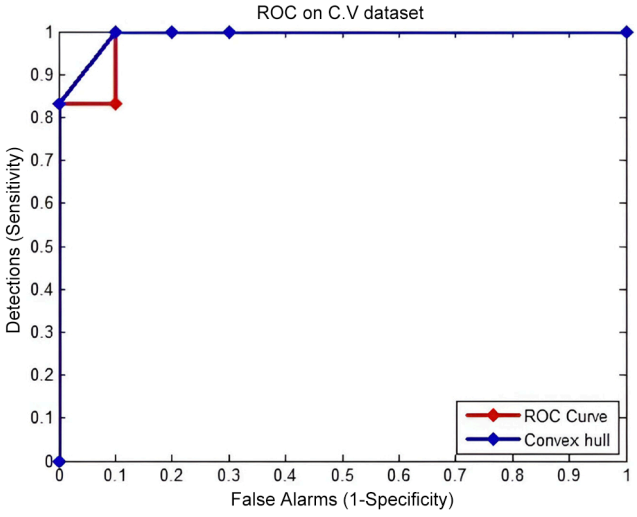

Figure 7. *Receiver operating characteristics curve on Multi Layer Perceptron best classifier for database I [Test on cross validation (C.V) Dataset].*

Another application of multi modal AI was seen in the work of Gao *et al*., where it was utilized for assessing treatment response (27). Both studies show the promise that AI brings in multi modal management, providing better patient care in the future. While developing future AI models for healthcare, it is necessary to keep in mind the regulations on AI use and that many radiologists are still unconvinced about the return on investment that AI in radiology brings (28, 29). We also need to keep in mind as stated by Tjoa *et al*., 'how AI takes decisions is still a black box and we poorly know many machine made decisions' (30).

Conclusion

Our genetically optimized MNN model offers a computationally efficient, highly accurate, and a novel approach for lung cancer diagnosis. This shows a future potential of integration into clinical workflows, which may help address diagnostic delays, serve as a triage in case of patient overloads and radiologist shortage, optimize resource utilization, and ultimately improve patient outcomes. This model might serve as a pathway for further development of future genetically optimized neural network models in medical imaging.

Conflicts of Interest

The Authors have no conflicts of interest to declare in relation to this study.

Authors' Contributions

Dr. Vijay Agrawal: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing-review & editing; Trushdeep Agrawal: Supervision, Visualization, Writing – original draft, Writing – review & editing, Project administration; Dr. Aruni Ghose: Supervision, Writing – review & editing, Project administration, Resources; Dr. Sola Adeleke: Supervision, Writing – review & editing, Project administration, Resources; Dr. Stergios Boussios: Supervision, Writing-review & editing, Project administration, Resources; Dr. Rajender Singh Arora: Conceptualization, Data curation, Validation, Resources, Supervision.

Funding

None.

Artificial Intelligence (AI) Disclosure

No artificial intelligence (AI) tools, including large language models or machine learning software, were used in the preparation, analysis, or presentation of this manuscript.

References

1 Nasim F, Sabath BF, Eapen GA: Lung cancer. Med Clin North Am 103(3): 463-473, 2019. DOI: 10.1016/j.mcna.2018.12.006
2 Legmann P: Imaging and lung disease: uses and interpretation. Tuber Lung Dis 74(3): 147-158, 1993. DOI: 10.1016/0962-8479(93)90003-G
3 Afshari Mirak S, Tirumani SH, Ramaiya N, Mohamed I: The growing nationwide radiologist shortage: Current opportunities and ongoing challenges for international medical graduate radiologists. Radiology 314(3): e232625, 2025. DOI: 10.1148/radiol.232625
4 Pei Q, Luo Y, Chen Y, Li J, Xie D, Ye T: Artificial intelligence in clinical applications for lung cancer: diagnosis, treatment and prognosis. Clin Chem Lab Med 60(12): 1974-1983, 2022. DOI: 10.1515/cclm-2022-0291
5 Diaz JM, Pinon RC, Solano G: Lung cancer classification using genetic algorithm to optimize prediction models. IISA 2014, The 5th International Conference on Information, Intelligence, Systems and Applications: 1-6, 2014. DOI: 10.1109/IISA.2014.6878770
6 Daliri MR: A hybrid automatic system for the diagnosis of lung cancer based on genetic algorithm and fuzzy extreme learning machines. J Med Syst 36(2): 1001-1005, 2012. DOI: 10.1007/s10916-011-9806-y
7 Dehmeshki J, Ye X, Lin X, Valdivieso M, Amin H: Automated detection of lung nodules in CT images using shape-based genetic algorithm. Comput Med Imaging Graph 31(6): 408-417, 2007. DOI: 10.1016/j.compmedimag.2007.03.002
8 Agrawal VL, Dudul SV: Conventional Neural Network approach for the Diagnosis of Lung Tumor. 2020 International Conference on Computational Performance Evaluation (ComPE): 543-547, 2020. DOI: 10.1109/ComPE49325.2020.9200118
9 Sufyan M, Shokat Z, Ashfaq UA: Artificial intelligence in cancer diagnosis and therapy: Current status and future perspective. Comput Biol Med 165: 107356, 2023. DOI: 10.1016/j.compbiomed.2023.107356
10 Martín-Noguerol T, López-Úbeda P, Luna A: Imagine there is no paperwork… it's easy if you try. Br J Radiol 97(1156): 744-746, 2024. DOI: 10.1093/bjr/tqae035
11 Togher D, Dean G, Moon J, Mayola R, Medina A, Repec J, Meheux M, Mather S, Storey M, Rickaby S, Abubacker MZ, Shelmerdine SC: Evolution of radiology staff perspectives during artificial intelligence (AI) implementation for expedited lung cancer triage. Clin Radiol 81: 106704, 2025. DOI: 10.1016/j.crad.2024.09.010
12 Ciompi F, Chung K, van Riel SJ, Setio AAA, Gerke PK, Jacobs C, Scholten ET, Schaefer-Prokop C, Wille MMW, Marchianò A, Pastorino U, Prokop M, van Ginneken B: Towards automatic pulmonary nodule management in lung cancer screening with deep learning. Sci Rep 7: 46479, 2017. DOI: 10.1038/srep46479
13 Nasrullah N, Sang J, Alam MS, Mateen M, Cai B, Hu H: Automated lung nodule detection and classification using deep learning combined with multiple strategies. Sensors (Basel) 19(17): 3722, 2019. DOI: 10.3390/s19173722
14 Shen S, Han SX, Aberle DR, Bui AA, Hsu W: An interpretable deep hierarchical semantic convolutional neural network for lung nodule malignancy classification. Expert Syst Appl 128: 84-95, 2019. DOI: 10.1016/j.eswa.2019.01.048
15 Cui S, Ming S, Lin Y, Chen F, Shen Q, Li H, Chen G, Gong X, Wang H: Development and clinical application of deep learning model for lung nodules screening on CT images. Sci Rep 10(1): 13657, 2020. DOI: 10.1038/s41598-020-70629-3
16 Siddiqui EA, Chaurasia V, Shandilya M: Classification of lung cancer computed tomography images using a 3-dimensional deep convolutional neural network with multi-layer filter. J Cancer Res Clin Oncol 149(13): 11279-11294, 2023. DOI: 10.1007/s00432-023-04992-9
17 Wang N: Luna 16 [Internet]. IEEE Dataport, 2025. Available at: https://dx.doi.org/10.21227/0kjp-g187 [Last accessed on December 30, 2025]
18 Armato SG 3rd, McLennan G, Bidaut L, McNitt-Gray MF, Meyer CR, Reeves AP, Zhao B, Aberle DR, Henschke CI, Hoffman EA, Kazerooni EA, MacMahon H, Van Beeke EJ, Yankelevitz D, Biancardi AM, Bland PH, Brown MS, Engelmann RM, Laderach GE, Max D, Pais RC, Qing DP, Roberts RY, Smith AR, Starkey A, Batrah P, Caligiuri P, Farooqi A, Gladish GW, Jude CM, Munden RF, Petkovska I, Quint LE, Schwartz LH, Sundaram B, Dodd LE, Fenimore C, Gur D, Petrick N, Freymann J, Kirby J, Hughes B, Casteele AV, Gupte S, Sallamm M, Heath MD, Kuhn MH, Dharaiya E, Burns R, Fryd DS, Salganicoff M, Anand V, Shreter U, Vastagh S, Croft BY: The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): a completed reference database of lung nodules on CT scans. Med Phys 38(2): 915-931, 2011. DOI: 10.1118/1.3528204
19 Grove O, Berglund AE, Schabath MB, Aerts HJWL, Dekker A, Wang H, Velazquez ER, Lambin P, Gu Y, Balagurunathan Y, Eikman E, Gatenby RA, Eschrich S, Gillies RJ: Quantitative

computed tomographic descriptors associate tumor shape complexity and intratumor heterogeneity with prognosis in lung adenocarcinoma. PLoS One 10(3): e0118261, 2015. DOI: 10.1371/journal.pone.0118261

20 Quanyang W, Yao H, Sicong W, Linlin Q, Zewei Z, Donghui H, Hongjia L, Shijun Z: Artificial intelligence in lung cancer screening: Detection, classification, prediction, and prognosis. Cancer Med 13(7): e7140, 2024. DOI: 10.1002/cam4.7140

21 Fockens KN, Jukema JB, Boers T, Jong MR, van der Putten JA, Pouw RE, Weusten BLAM, Alvarez Herrero L, Houben MHMG, Nagengast WB, Westerhof J, Alkhalaf A, Mallant R, Ragunath K, Seewald S, Elbe P, Barret M, Ortiz Fernández-Sordo J, Pech O, Beyna T, van der Sommen F, de With PH, de Groof AJ, Bergman JJ: Towards a robust and compact deep learning system for primary detection of early Barrett's neoplasia: Initial image-based results of training on a multi-center retrospectively collected data set. United European Gastroenterol J 11(4): 324-336, 2023. DOI: 10.1002/ueg2.12363

22 Elayaraja P, Kumarganesh S, Sagayam KM, Andrew J: An automated cervical cancer diagnosis using genetic algorithm and CANFIS approaches. Technol Health Care 32(4): 2193-2209, 2024. DOI: 10.3233/THC-230926

23 Li Z, Ma Z, Zhou Q, Wang S, Yan Q, Zhuang H, Zhou Z, Liu C, Wu Z, Zhao J, Huang S, Zhang C, Hou B: Identification by genetic algorithm optimized back propagation artificial neural network and validation of a four-gene signature for diagnosis and prognosis of pancreatic cancer. Heliyon 8(11): e11321, 2022. DOI: 10.1016/j.heliyon.2022.e11321

24 Said Y, Ayachi R, Afif M, Saidani T, Alanezi ST, Saidani O, Algarni AD: AI-driven genetic algorithm-optimized lung segmentation for precision in early lung cancer diagnosis. Sci Rep 15(1): 23058, 2025. DOI: 10.1038/s41598-025-08116-w

25 MacEachern SJ, Forkert ND: Machine learning for precision medicine. Genome 64(4): 416-425, 2021. DOI: 10.1139/gen-2020-0131

26 Yu Y, Ouyang W, Huang Y, Huang H, Wang Z, Jia X, Huang Z, Lin R, Zhu Y, Yalikun Y, Tan L, Li X, Zhao F, Chen Z, Li W, Liao J, Yao H, Long M: Artificial intelligence-based multi-modal multi-tasks analysis reveals tumor molecular heterogeneity, predicts preoperative lymph node metastasis and prognosis in papillary thyroid carcinoma: a retrospective study. Int J Surg 111(1): 839-856, 2025. DOI: 10.1097/JS9.0000000000001875

27 Gao Y, Ventura-Diaz S, Wang X, He M, Xu Z, Weir A, Zhou HY, Zhang T, van Duijnhoven FH, Han L, Li X, D'Angelo A, Longo V, Liu Z, Teuwen J, Kok M, Beets-Tan R, Horlings HM, Tan T, Mann R: An explainable longitudinal multi-modal fusion model for predicting neoadjuvant therapy response in women with breast cancer. Nat Commun 15(1): 9613, 2024. DOI: 10.1038/s41467-024-53450-8

28 Mello-Thoms C, Mello CAB: Clinical applications of artificial intelligence in radiology. Br J Radiol 96(1150): 20221031, 2023. DOI: 10.1259/bjr.20221031

29 Pantanowitz L, Hanna M, Pantanowitz J, Lennerz J, Henricks WH, Shen P, Quinn B, Bennet S, Rashidi HH: Regulatory aspects of artificial intelligence and machine learning. Mod Pathol 37(12): 100609, 2024. DOI: 10.1016/j.modpat.2024.100609

30 Tjoa E, Guan C: A survey on explainable artificial intelligence (XAI): toward medical XAI. IEEE Trans Neural Netw Learn Syst 32(11): 4793-4813, 2021. DOI: 10.1109/TNNLS.2020.3027314