



An intelligent generative adversarial network multistage lung cancer detection and subtypes classification

Mattakoyya Aharonu¹ · Lokesh Kumar Ramasamy¹

Received: 26 August 2024 / Accepted: 27 November 2024 / Published online: 16 December 2024
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2024

Abstract

Cancer remains a leading cause of global mortality, with lung cancer having the highest mortality rate. Early detection through Computed Tomography (CT) imaging is essential for accurately assessing lung cancer stages. However, existing deep-learning classification models for lung CT images often suffer from low efficiency and accuracy in early cancer diagnosis and are typically time-consuming. To address these challenges, this research introduces a novel two-stage network approach, the Deep Convolutional cross max-out kernel graph-based Generative Adversarial Network with Enhanced Prism refraction Search (DCGAN-EPS). This model enhances tumor detection accuracy and efficiency through advanced techniques. In Stage 1, adaptive wavelet denoising reduces noise and eliminates irrelevant pixels in lung CT images, followed by precise image segmentation using an improved attention U-Net. Stage 2 utilizes a weighted local binary pattern to improve feature extraction for tumor detection. The Enhanced Prism refraction Search (EPS) algorithm optimizes the neural network's loss parameters, significantly boosting detection accuracy. Additionally, integrating local interpretable model-agnostic explanations enhances the model's transparency and interpretability. Experimental results on the LIDC-IDRI and Chest CT-Scan datasets demonstrate the proposed method's robust performance, achieving accuracy, recall, precision, F1-Score, and specificity scores of 99.92%, 99.94%, 99.94%, 99.92%, and 99.90% on the LIDC-IDRI dataset, and scores of 99.4%, 99.2%, 99.5%, 99.4%, and 99.2% on the Chest CT-Scan dataset, underscoring its efficiency and suitability for clinical applications.

Keywords Lung cancer detection · Computed tomography · Explainable model · Deep learning · Generative adversarial networks · Image segmentation · Feature extraction

1 Introduction

Lung Cancer (LC) continues to be one of the most widespread and deadly forms of cancer worldwide, posing significant challenges for early detection and powerful dealing. The World Health Organization (WHO) reports that it contributes to around 11.4% of all emerging distortions in India, with an alarming death rate of 84.21%. Early and accurate diagnosis is essential in refining patient consequences and dropping death tariffs [30, 33]. The genesis of LC often stems from uncontrolled tissue growth, characterized by long-lasting changes in DNA sequences, known

as mutations. These mutations are triggered by inherited genetic defects or environmental factors, with tobacco smoke being a primary offender among many carcinogens [25, 36]. While various diagnostic methods like scans, blood tests, chest X-rays, and CT scans aid in detecting lung diseases, distinguishing between malignant and benign cancers remains challenging due to their overlapping features. While classification, preprocessing, and feature extraction are fundamental stages in lung tumor classification, crucial for refining input data quality and extracting pertinent information to enable accurate classification [37]. Ordinary preprocessing stages include typical image resizing, regulating pixels, reducing noise, enhancing contrast, and extracting the Region of Interest (RoI) containing lung tumors. Subsequently, feature extraction methods are applied to capture relevant information from the pre-processed images. Moreover, image segmentation, a critical preprocessing step, involves dividing digital images into segments to simplify analysis. This technique reduces image complexity, making

✉ Lokesh Kumar Ramasamy
lokeshkumar.r@vit.ac.in

Mattakoyya Aharonu
mattakoyya.aharonu2020@vitstudent.ac.in

¹ School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu 632014, India

it easier to handle and interpret, particularly in CT images for various analyses [29].

Recent years have witnessed remarkable advancements in medical imaging technology, coupled with the propagation of Artificial Intelligence (AI) and Deep Learning (DL) approaches, expressively advancing the field of LC recognition and classification. DL, a subset of Machine Learning (ML), facilitates the identification, classification, and quantification of patterns within medical images [12]. The rapid progress in DL, primarily in medical applications, has led to significant improvements in various aspects of healthcare. There are various DL approaches, such as convolutional neural networks (CNN), recurrent neural networks (RNN), attention-based approaches, etc. Among various techniques, Generative Adversarial Networks (GANs) have emerged as a potent tool for medical image analysis [1, 17, 34, 35]. GANs, comprising a generator and a discriminator network, collaborate to yield accurate CT images and differentiate between genuine and synthetic data. By leveraging the adversarial training model, GANs discern complex shapes and structures from medical images, thereby enhancing accuracy in disease detection and classification. Additionally, recent researchers have introduced innovative algorithms that enhance model performance and efficiency like those by Movassagh et al. [21] to enhance performance and efficiency, it still faces limitations, particularly concerning the reliance on high-quality, well-annotated datasets, which restrict its effectiveness in real-world applications. Furthermore, the inherent complexity of medical images poses challenges for accurate interpretation and analysis.

In this framework, a DCGAN-EPS approach for LC recognition and categorization. This approach exploits the capabilities of GANs to analyze medical images, particularly CT scans of the lungs and accurately identify cancerous regions. The main contributions of this research are as follows,

- This research introduces the DCGAN-EPS framework, which improves lung CT classification by converting images into graph structures and utilizing convolutional cells to capture nodal interactions. The integration of a Balance Ratio (BR) metric fine-tunes the generator-discriminator balance, enhancing classification accuracy by effectively distinguishing subtle nodal variations.
- Adaptive Wavelet Denoising Transformer (AWDT), combined with bilateral filtering, enhances noise reduction in lung CT scans by adaptively adjusting thresholds according to local image statistics and preserving edge details. This method effectively reduces Gaussian noise while maintaining the clarity of critical features, ensuring high-quality image preprocessing.
- In DCGAN-EPS, an Improved Attention U-Net (IAU-Net) is utilized to emphasize relevant regions within lung

CT images. By incorporating attention coefficients and skipping connections between encoder and decoder layers, an identity block enhances the preservation of essential information during the segmentation process.

- Moreover, the EPS algorithm effectively enhances the performance of LC recognition by leveraging principles inspired by the refraction of light through prisms. This innovative optimization technique aids in improving the efficiency of the proposed technique by efficiently exploring the solution space and effectively maintaining population diversity.
- Additionally, employing Local, Interpretable, Model-agnostic Explanations (LIME) for CT detection is capable of offering local, interpretable, model-agnostic explanations. By perturbing samples around a target instance and training a localized linear model, LIME clarifies the decision-making process of CT scan detectors.

The organization of this research is as follows: Section 2 provides a summary of existing literature on LC detection. In Sect. 3, the introduced DCGAN-EPS is elaborately presented. Section 4 deliberates the outcomes and discussions. Lastly, Sect. 5 concludes and outlines future works.

2 Related work

This section delivers a complete overview of DL methods for lung tumor detection, including segmentation, non-segmentation, and GAN-based approaches.

2.1 Segmentation based approaches in LC detection

Nowadays Segmentation-based approaches in lung tumor detection involve techniques to recognize and define the margins of tumors within CT scans, typically obtained from lung CT scans. Halder and Dey [1] suggested an approach to segment and classify LC by Atrous Convolution-based CNN (ATCNN), which illustrates lung nodes by portraying different features from the CT scans. Furthermore, this suggested approach with a two-layer atrous pyramid and residual networks confirmed the excellent performance of nodule categorization. Gopinath et al. [8] developed Deep Fused Features-Based Cat-Optimized Networks (DFF-CON), a categorization system that utilized merged attributes and enhanced systems. In this approach, deep CNN was hired to augment classification and overcome the overfitting issues. Balcı et al. [5] developed a Series-based DL(SDL) framework for lung node categorization, among 3D images and series class difficulties. Consequently, self-similarity attributes in CT images were taken by altering the ROI to sequences with a circular image and these sequences were

labeled with U-shaped CNN. Thangavel and Palanichamy [32] designed an Effective DL (EDL) method for segmenting CT images for pulmonary cancer in the thoracic area. It was performed through preprocessing, segmentation, and feature extraction. Next, the suggested NASNet-based classifier categorized the nodules through gathered features. Bishnoi and Goel [6] suggested a Weighted VGG Deep Network (WVDN) for LC Classification. This approach comprised three stages, including preprocessing, tuning, and regularization and system inference in Nvidia TensorRT throughout post-processing for distribution in practical capability. This approach was constructed on complete CT portions rather than the noticeable remarks and helped to develop clinical analysis. Mary and Thanammal [2] suggested a DL-based LC Detection (DL-LCD) model to detect LC images. Initially, those images were pre-processed to eliminate noise biases and enhance CT scan clearness. The lung CT images were segmented using the Honey Badger algorithm. These CT scans were then labeled into cancerous and non-cancerous categories. Balachandran and Ranganathan [4] designed Semantic Context-aware attention UNET (SC-UNET) for LC segmentation and categorization. The goal was to detect and classify nodules in the lung CT scan as malignant or benign. A SC-UNET construction was designed to segment the nodule from the lung CT images.

Bagheri Tofighi et al. [3] developed MobileNet Verison 2 with a Stacked Gated Recurrent Unit (MobileNetV2-SGRU) as a transfer learning-based predictor for LC classification. Relevant features from lung CT images were automatically extracted using MobileNetV2 and subsequently processed through stacked gated recurrent unit layers to capture sequential and temporal information. Additionally, Grad-CAM was incorporated to enhance the model's interpretability and transparency. A limitation of the MobileNetV2-SGRU model is its dependence on large datasets for training, which restricts its effectiveness in situations with limited available data.

2.2 Non-segmentation-based approaches in LC detection

Non-segmentation-based approaches in lung tumor detection state that the methods do not primarily depend on the explicit delineation or segmentation of tumor regions within medical lung CT scans. Bushara et al. [7] suggested a DL called LCD-CapsNet, which combines a CNN and a Capsule Neural Network (CapsNet). This approach aimed to minimize extensive datasets and attain spatial invariance for LC categorization. The main concern of this method was to devise algorithms that categorize images from a dataset to ascertain whether a patient has or is susceptible to developing LC. Siddiqui et al. [31] developed Gabor Filters with an Enhanced Deep Belief Network

(GF-EDBN) to categorize lung CT images. In this model, delay and ruining duration were cut down because the improved filters had smaller amount constraints. Additionally, it utilized a Support Vector Machine (SVM) as a classifier to categorize CT images and improve performance when compared to other approaches. Raza et al. [26] suggested an approach called Lung-EffNet. This approach primarily classified LC by EfficientNet from CT images. It was constructed on the structure of EfficientNet and was further adapted by accumulating top layers in the organization head of the model. Saikia et al. [28] developed an automatic Lung Nodule categorization scheme founded on the Hybrid Transfer Learning (Hybrid TL) model. This approach integrates VGG networks with Support Vector Machine (SVM) and Random Forest (RF) to establish a hybrid model that further helped in diminishing the computation complexity of the model. Nahiduzzaman et al. [23] suggested a Lightweight Parallel Depth-Wise Separable CNN (LPDCNN) with a Ridge regression Extreme Learning Machine (Ridge-ELM) for accurate categorization of three LC using CT images. This approach combined with Contrast-Limited Adaptive Histogram Equalization (CLAHE) for effective preprocessing.

An advanced ML model was developed by Musthafa et al. [22] to enhance the classification of LC stages using CT scan images, aiming to address existing limitations by providing a faster, non-invasive, and reliable diagnostic tool. Extensive preprocessing was performed, including resizing, normalization, and Gaussian blurring. A Double Layered Convolutional Neural Network (DLCNN) was trained on the preprocessed data, and class imbalance was addressed using the Synthetic Minority Over-sampling Technique (SMOTE). However, the model's reliance on extensive preprocessing and data augmentation leads to increased computational requirements and longer training times. A novel CNN architecture, GoogLeNet with Adaptive Layers (GoogLeNet-AL), was suggested by Ma et al. [20] for LC detection. The architecture was enhanced with advanced features such as squeeze-and-excitation blocks, dilated convolutions, depth-wise separable convolutions, group convolutions, non-local blocks, octave convolutions, inverted residuals, and ghost convolutions in the inception layers. However, a potential drawback of GoogLeNet-AL is its increased computational complexity due to the integration of multiple advanced features. A novel framework was developed by Gupta et al. [10] for detecting and classifying LC. The framework was built on a modified U-Net architecture, utilizing multi-scale feature extraction and Differentiable architecture search using CT images (UDCT) for efficient classification. Multilevel Otsu thresholding was incorporated for image preprocessing and lung nodule segmentation. While the framework demonstrated effective and efficient detection, its performance was limited in handling extremely small nodules.

2.3 GAN-based detection approaches

Hong et al. [13] developed the GAN with Long Short-Term Memory (GAN-LSTM-3D) method for the 3D reconstruction of lung tumors. The method improved the reconstruction efficiency by leveraging attention mechanisms to enhance feature extraction and integration, leading to more accurate tumor modeling. Rezaei and Ahmadi [27] developed a GAN-based approach for 3D lung tumor reconstruction that incorporated knowledge transfer techniques. By transferring knowledge from pre-trained models, the method enhanced the reconstruction quality and accuracy of tumor volumes, facilitating improved diagnostic and treatment planning. Kausar et al. [15] developed the Style Distribution Transfer GAN (SD-GAN) approach for detecting Covid-19 from X-ray images. This method employed style transfer techniques to refine the GAN's style distributions, enhancing feature representation and significantly improving the detection of COVID-19-related abnormalities in chest X-rays. Moreover, the limitations of existing techniques are provided in Table 1.

2.4 Problem statement

There are many approaches developed in the field of LC detection. However, existing research exhibits significant flaws, including inaccurate classification, suboptimal segmentation, and detection outcomes. Additionally, many techniques suffer from low precision, high time complexity, and extensive parameter tuning requirements. Moreover, these methods often lack interpretability, making it difficult to understand model predictions and decision-making processes. To address these challenges, this research introduces the efficient DCGAN-EPS-based LC detection framework, providing a robust solution for accurate and efficient LC identification in complex clinical environments.

3 Proposed methodology

The proposed two-stage lung tumor detection model, as depicted in Fig. 1, utilizes DL techniques. In the first stage, the model preprocesses the data using AWDT with bilateral filtering to reduce noise and segments the images with IAU-Net.

Table 1 Limitations of existing approaches in LC detection

References	Methods	Segmentation	Limitations
Halder and Dey [11]	ATCNN	✓	It attained the lowest performance (Accuracy = 95.97%)
Gopinath et al. [8]	DFF-CON	✓	High computational complexity
Balci et al. [5]	SDL	✓	Class imbalance issue occurred
Thangavel and Palanichamy [32]	EDL	✓	It had limited resources
Bishnoi and Goel [6]	WVDN	✓	It was suitable for other modality LC images
Mary and Thanammal [2]	DL-LCD	✓	Preprocessing was not efficient
Balachandran and Ranganathan [4]	SC-UNET	✓	The extracted features were not sufficient for better interpretability
Bushara et al. [7]	LCD-CapsNet	✗	Implementation was more difficult than traditional CNN
Siddiqui et al. [31]	GF-EDBN	✗	A decrease in sensitivity at low False Positive (FP) rates often corresponds with an increase in sensitivity at high FP rates
Raza et al. [26]	Lung-EffNet	✗	The large cell carcinoma class exhibited the lowest performance
Saikia et al. [28]	Hybrid TL	✗	It was computationally expensive and slow when dealing with extensive datasets
Nahiduzzaman et al. [23]	LPDCNN	✗	The transparency and explainability of this approach were not sufficient
Hong et al. [13]	GAN-LSTM-3D	✗	It required extensive computational resources and training time
Rezaei and Ahmadi [27]	GAN	✗	This approach highly depended on the quality of transferred knowledge
Kausar et al. [15]	SD-GAN	✗	Style transfer was not fully capturing all COVID-19 feature
Musthafa et al. [22]	DLCNN	✗	The model's reliance on extensive preprocessing increases computational requirements and training times
Ma et al. [20]	GoogLeNet	✗	Increased computational complexity due to the integration of multiple advanced features
Gupta et al. [10]	UDCT	✗	The model's performance is limited in handling extremely small nodules
Bagheri Tofighi et al. [3]	MobileNetV2-SGRU	✓	The model's effectiveness was limited in scenarios with scarce data due to its reliance on large datasets for training

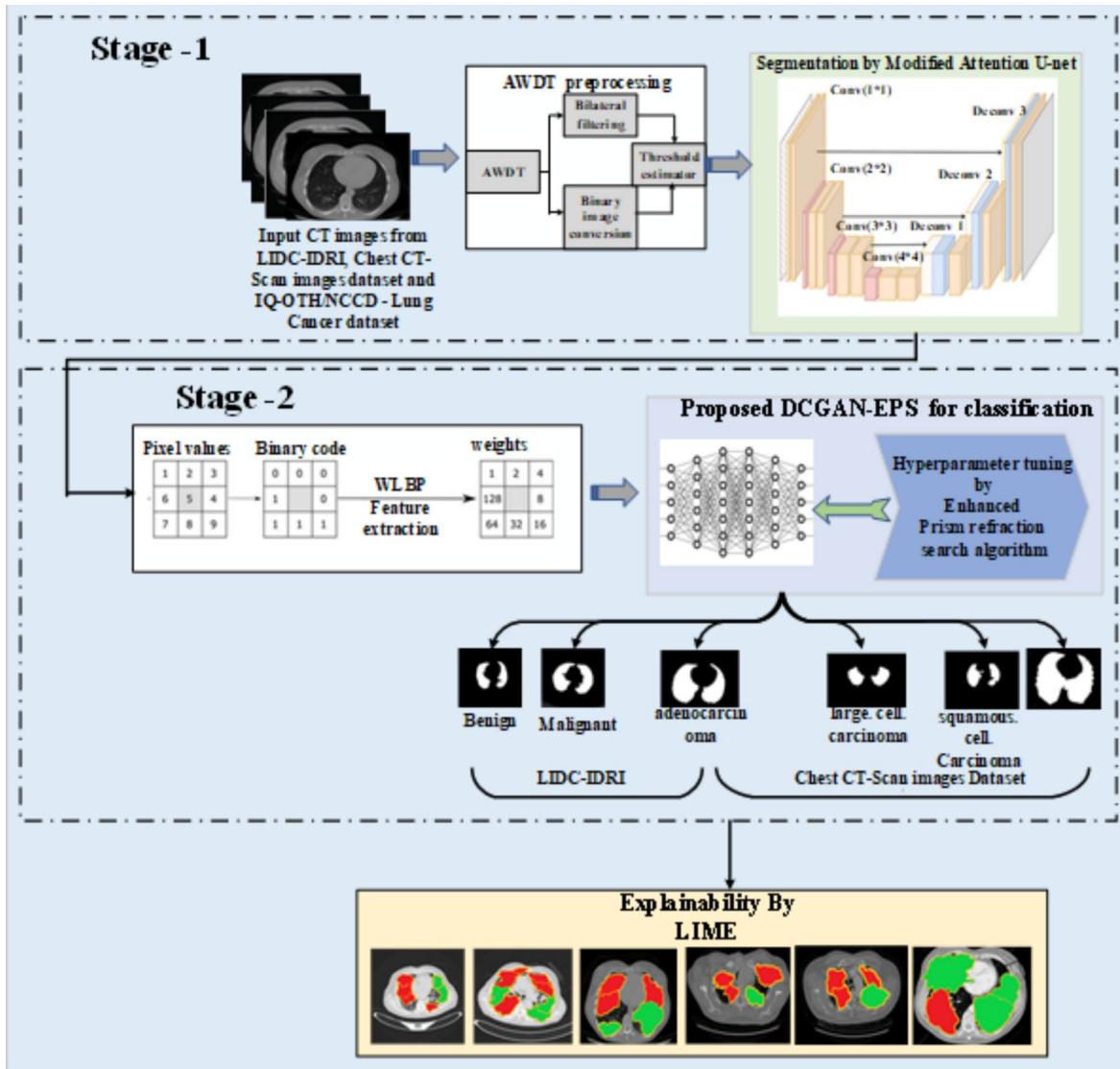


Fig. 1 Block illustration of proposed methodology

In the second stage, feature extraction and classification are performed, with the EPS algorithm further enhancing model performance. During testing, Stage-1 classifies CT images as either healthy or diseased, and Stage-2 employs the DCGAN-EPS for detailed, fine-grained analysis.

3.1 Stage-1

In Stage 1, the lung CT image is first pre-processed, followed by the use of IAU-Net performs segmentation to preserve semantic information.

3.1.1 Image preprocessing

The AWDT technique identifies and removes noisy pixels in lung CT scans by sorting pixels within a window, defining

extreme values, and classifying the central pixel as noisy if it matches these values [18]. Subsequently, the central pixel is examined to determine whether it matches any of these values. If a match is found, the central pixel is classified as noisy. The noisy vector, denoted as N , is affected by the presence of additive white Gaussian noise through Eq. (1).

$$N = V_i + M_i, \text{ where } i = 0, 1, 2, \dots, G-1 \quad (1)$$

where, V_i denotes the noise-free wavelet components of the input and M_i represents independent and identically distributed i th Gaussian noise. Removing noise in the wavelet domain involves the application of an appropriate threshold function along with a specified threshold value. This threshold is recognized for its tendency to effectively eliminate additive Gaussian noise, often leading to an excessively

smooth appearance in the resulting image. This effect arises because the threshold size becomes significant, influenced by its dependence on the number of samples M . To address this, the proposed DCGAN-EPS framework integrates AWDT to effectively reduce noise while preserving important details, ensuring that the model better handles noisy datasets.

AWDT first analyses the local statistical properties of the image, like noise variance and signal strength, to determine a suitable threshold for each region. Then bilateral filtering is applied to adaptively adjust the threshold based on the noise level and image content [14]. By applying a varying threshold, the method distinguishes between noise and actual image features, preventing the over-smoothing of critical details. Additionally, the variable M represents the adjacent pixel count within the adjacent window. The resulting bilateral filtered value is obtained through a specified Eq. (2)

$$P_t(V, M) = V + \frac{1}{2} \left(\sqrt{(V - M)^2 + \eta} - \sqrt{(V + M)^2 + \eta} \right) \quad (2)$$

where, $P_t(V, M)$ represents the bilateral filtered function. η Represents the threshold value. Converting the denoised image into binary involves assigning pixels to the foreground or background based on a η threshold value. This approach addresses the challenge of preserving edges while reducing noise by incorporating bilateral filtering to specifically target noise reduction while maintaining sharpness in edges and fine details. The proposed DCGAN-EPS framework effectively adapts to handle noisy or incomplete datasets by integrating AWDT to identify and eliminate noisy pixels in lung CT scans by analyzing local statistical properties, like noise variance and signal strength, to establish a suitable threshold for each region. By carefully classifying central pixels as noisy based on their matching with extreme values, AWDT ensures that only genuinely corrupted data is targeted for removal. This is crucial for maintaining the integrity of the dataset, as it preserves vital image features necessary for accurate tumor detection.

Furthermore, the incorporation of bilateral filtering allows the framework to dynamically adjust thresholds according to the noise level and image content. This adaptive approach not only mitigates the risk of over-smoothing often a consequence of static thresholding but also ensures that critical edges and fine details remain intact. The careful balance achieved through these methods enhances the quality of the input data, enabling the DCGAN-EPS framework to operate effectively even in the presence of noise or incomplete information. Consequently, this robust preprocessing capability supports more reliable and accurate LC detection

3.1.2 Segmentation by improved Attention U-net

IAU-Net excels in medical image segmentation by combining an encoder and a decoder. The decoder restores image clarity using up-sampling and features from the encoder are fused with up-sampled outputs for accurate localization. It effectively segments medical images with nested and skip connections to preserve semantic information, processing pre-processed lung CT images through its encoding path to extract hierarchical features through Eqn. (3).

$$\gamma = \lambda_2 \left(\psi^t \left(\lambda_1 \left(P_t(V, M) W_f^t + P_t(V, M) W_g^t \right) \right) \right) \quad (3)$$

where γ represents the attention coefficient, t denotes its transpose. λ_1 Corresponds to the ReLU function and λ_2 corresponds to the Sigmoid function, ensuring that the value γ falls within the range of (0, 1). W_f^t and W_g^t assigns transpose weights of matrix f and g to individual pixel features respectively, facilitating interaction among all the pixels in the image features. In the encoding path, attention structures focus on relevant image regions, and identity blocks enhance segmentation accuracy but increase computational complexity. The utilization of identity blocks (κ) is defined through Eq. (4).

$$\kappa = \gamma + P_t(V, M) \quad (4)$$

where, κ represents the identity blocks function. The identity block mirrors the down-sampling block's layers, using uniform strides and an identity skip connection. In decoding, each convolution reduces the feature graph size by half, which is then doubled by up-sampling by Eq. (5).

$$D_o = BR(Conv_{3 \times 3}^1(Up^2(\kappa))) \quad (5)$$

where D_o represents the decoding output, $Conv_{3 \times 3}^1$ denotes the 3×3 conv layer with a step dimension of 1. Up^2 signifies the upsampling layer with a step dimension of 2, and BR signifies the combined layer consisting of the Batch Normalization (BN) layer and the Rectified Linear Unit (ReLU) activation layer [38]. The IAU-Net employs an advanced segmentation approach that utilizes skip connections between encoding and decoding layers, ensuring the preservation of critical details while generating precise segmentation maps. This capability allows the model to classify each pixel in the input image into relevant categories, such as tumors or healthy tissue. By producing a class probability map that indicates the likelihood of each pixel belonging to a specific class, along with feature maps from intermediate layers, the IAU-Net provides valuable insights into the network's learned representations. One of the key advantages of this approach is its integration of attention mechanisms, which enhance segmentation accuracy by focusing on pertinent features and minimizing the influence of irrelevant

areas. While depth features are vital for segmentation, they often fail to capture essential textural details. The IAU-Net addresses this limitation effectively, significantly improving medical image segmentation through superior feature extraction and accuracy. Its encoder-decoder architecture, enhanced by nested and skip connections, not only preserves semantic information but also facilitates the extraction of hierarchical features. This sophisticated design enables a nuanced understanding of complex structures, ultimately allowing for more accurate differentiation between tumors and healthy tissue. Instead, feature extraction in Stage 2 complements depth features, allowing for more detailed texture analysis.

3.2 Stage-2

This stage emphasizes the in-depth classification of disease types by extracting detailed features from the images. It employs DCGAN-EPS to identify and distinguish between different disease varieties. The extracted features facilitate accurate categorization and enhance the overall classification performance.

3.2.1 Feature extraction by a weighted local binary pattern

Following image segmentation, the Weighted Local Binary Pattern (WLBP) is utilized to extract feature vectors from the CT images. It extracts features from a segmented image using a localized binary pattern approach with weighted contributions [19]. WLBP are chosen for textural feature analysis due to their effectiveness in capturing local texture information with computational efficiency. It encodes texture by comparing each pixel with its neighboring pixels, resulting in binary patterns that are robust to minor variations in texture. Additionally, the Local Binary Pattern (LBP) is extended to handle rotation invariance, ensuring consistent texture representation even when textures are rotated. Compared to other traditional feature analysis methods like [13] and Gray-Level Co-occurrence Matrix (GLCM), WLBP offers several advantages. Existing approaches offer detailed texture analysis and rich texture information but come with significant computational costs and complex parameter tuning. In contrast, WLBP provides a simpler and more efficient method for texture analysis. LBP's simple computation and low parameter requirements make it particularly well-suited for scenarios where quick and reliable texture analysis. It operates by associating the intensity of a focal pixel with its neighbors and training the outcome as a binary pattern achieved by Eq. (6).

$$\text{Weighted}_{LBP}(q, s) = \sum_{j=1}^q D_o(M) * 2^{(j-1)}, \text{ where} \\ M = \begin{cases} 1 & \text{if } (v_j - v_d) \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

where s represents the space from adjacent pixels to the central pixel, q signifies the total amount of adjacent pixels. v_d and v_j represents the central pixel and neighboring pixels correspondingly. j represents the index of the neighboring pixels around the center pixel for which the LBP is being calculated. $D_o(M)$ represents the binary value, indicating whether the intensity of an adjacent pixel is higher than or equivalent to that of the central pixel. In considering weighted pixel points, it account for the impact of various pixel positions on the corresponding central pixel within a respective-sized template. Focusing exclusively on pixels located horizontally, vertically, and diagonally relative to the central point, the definite process is designated in Eq. (7).

$$wv_j = \sum_{x=1}^w \frac{2^{w-x}}{10} (v_{z-x} + v_{z+x}) + \left(1 - \sum_{x=1}^w \frac{2^{w-x+1}}{10} \right) v_z \quad (7)$$

where, wv_j represents the central point weight calculation. w represents the weighted number of layers, which calculated the distance from the pixel position, x represents the location left or right away from the j th pixel location to be intended. Subsequently, $v_{j\text{central}}$ denoting the j th element in the template is defined as the half sum of the pixel position values calculated by Eq. (8) for both the inner layer and the outer layer in a specific direction.

$$v_{j\text{central}} = \frac{1}{2} \sum_{w=1}^2 wv_j \quad (8)$$

In the last step, the computed weighted pixel values are associated with the central pixel value, efficiently leveraging the data from the weighted pixels in the binary-layer neighborhood. This effectively increases the ratio of conserving potential features in the target surface.

3.2.2 Proposed deep convolutional cross max out kernel graph-based generative adversarial network-based classification

Following the extraction of features, lung nodules undergo classification as either benign or malignant utilizing the Deep Convolutional cross max out kernel graph-based Generative Adversarial Network (DCGAN) architecture. This architecture is scalable and flexible, employing convolutional cells derived from the data. This is specifically designed to handle a classification task by combining deep convolutional layers, cross-connected max-out activations, kernel mapping technique and

adversarial training with graph-based representations [9, 24, 34, 35]. The network construction of DCGAN is exposed in Fig. 2.

In this approach, lung CT images are denoted as graphs to effectively capture the relationships between different regions within the images. Each node in the graph represents a distinct region or feature of interest within the CT image. The edges between nodes are established based on spatial proximity or similarity in feature characteristics, which allows the model to account for the interdependencies and interactions between different lung regions. Initially, lung CT images into a graph-based representation through generator (G) and discriminator (D) by Eq. (9).

$$I(D, G) = R_{z \approx p_z}[\log D(v_j \text{central})] + R_{z \approx p_G}[\log(1 - D)] \quad (9)$$

where, $I(D, G)$ signifies learned input through discriminator and generator. $R_{z \approx p_z}$ signifies the expectation of real data in the network of the discriminator's outcome when related to real data. $R_{z \approx p_G}$ signifies the expectation over generated data of the network of the discriminator's outcome when related to generated data. $\log(1 - D)$ signifies the discriminator's output when applied to generated data. In that, each element represents different lung nodule regions of ROI. Then, it processes the graph-based data, capturing dependencies and interactions between different lung nodules. After that utilizing Deep Convolutional cross-connected Kernel Mapping (DCKM) and Cross Maxout Convolutional Network (CMCN) as convolutional layers through Eqs. (10) and (11).

$$L_j I(D, G) = DCKM(L_{j-1}), n+1 \leq j \leq n+m \quad (10)$$

$$C_j I(D, G) = CMCN[f(w_1 * (w_2 * z + B_2) + B_1)] \quad (11)$$

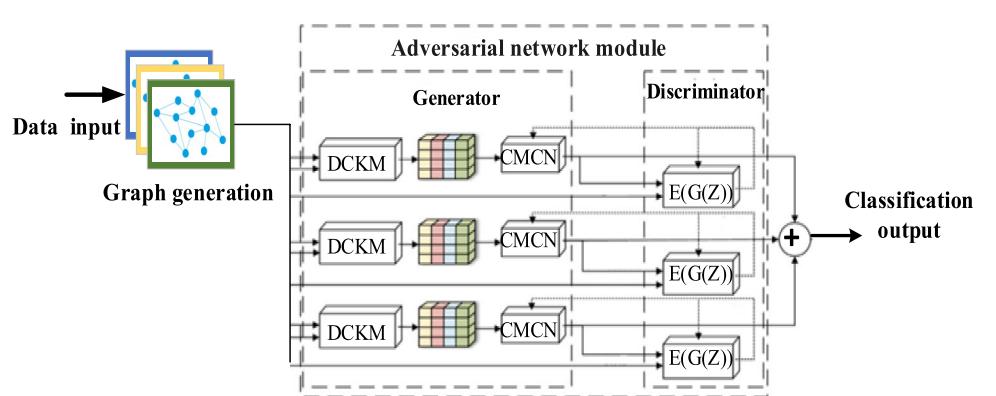
where, $L_j I(D, G)$ represents the deep convolutional cross-connected kernel mapping layer. $DCKM$ represents the kernel function in the convolutional layer. n and m represents the j th mapping of convolutional blocks. $C_j I(D, G)$ represents the cross-max out layer. $CMCN$ signifies cross-maxout function. f represents the ReLU activation function. The

terms w_2 and B_2 represent the trainable weights and bias of the depth-wise convolution, while w_1 and B_1 represent the trainable weights and bias of the point-wise convolution, correspondingly. Ultimately, the layers generate embeddings for nodes in the graph to capture relevant features from the input data. In this network, a successful semantic representation is acquired through training a model to predict a feature, denoted as P_z based on input data z . This graph-based representation enhances the ability to classify and detect lung abnormalities by providing a structured way to analyze complex spatial relationships and feature interactions within lung CT images. In this context, the research illustrates the categorization of lung CT images, wherein the encoder is guided to achieve the inversion of the generator. Specifically, the encoder (E) generates an output, $E(G(z))$ that equals the original feature z through Eq. (12).

$$E(G(z)) = \sum_{j=0}^z L_j I(D, G) \oplus C_j I(D, G) \quad (12)$$

where, $E(G(z))$ represents the final output of the softmax layer. This process is essential for deceiving the discriminator at that particular feature for classification. The generator is designed to simultaneously learn the features of the original data and the patterns within the latent vector. Meanwhile, the discriminator employs the standard network architecture from DCGAN-EPS to differentiate between real and generated data. This setup engages in a minimax plot, where the generator's distribution gradually aligns with the true distribution, causing the discriminator to become increasingly ineffective at distinguishing between them. To measure this configuration, introduce the BR, which evaluates the equilibrium between the sparse data and generated data. During training, random noise is sampled from a multivariate normal distribution, and input features are drawn from the training set. After each gradient descent update, the discriminator is denoted as $D(\theta_D)$, while the generator's states before and after training are represented as $G_{pre}(\theta_G)$

Fig. 2 Network architecture of deep convolutional cross max out kernel graph adversarial network



and $G_{post}(\theta'_G)$, respectively. The BR, as defined in Eq. (13), reflects the effectiveness of the generator and discriminator.

$$BR = \frac{D(G_{post}(\theta'_G)) - D(G_{pre}(\theta_G))}{D(\theta_D) - D(G_{pre}(\theta_G))} \quad (13)$$

A low BR ($< 30\%$) suggests that the generator is insufficiently strong to deceive the discriminator, meaning the generated data are still identifiable as fake. Conversely, a high BR ($> 80\%$) indicates that the discriminator is too weak and is not providing adequate feedback to the generator. The generator's final loss function, incorporating an encoder-decoder-encoder three-sub-network, is computed using Eq. (14).

$$Loss = \gamma_1 L_{CE} + \gamma_2 L_G \quad (14)$$

where γ_1 and γ_2 are weighting parameters, L_{CE} signifies the binary cross-entropy loss, L_G denotes the adversarial loss of the generator network, aimed at misleading the discriminator network. In this stage, the EPS algorithm is utilized to optimize the loss function derived from the classification module, thereby improving the functioning of the network.

In the proposed DCGAN-EPS, magnitude-based weight pruning is strategically integrated to address the computational demands associated with LC detection. The DCGAN-EPS model leverages the power of deep convolutional layers to generate high-quality representations of lung cancer images while employing a generator and discriminator architecture. By incorporating magnitude-based weight pruning, the model systematically identifies and eliminates less significant weights during the training process. This step not only streamlines the network architecture but also significantly reduces the number of parameters, leading to lower memory consumption and faster processing times without compromising the model's ability to accurately detect lung cancer.

After the initial training of the DCGAN-EPS model, magnitude-based weight pruning is applied to filter out weights below a predefined threshold based on their absolute values. This pruning process is followed by a fine-tuning phase, where the remaining weights are optimized to recover any potential performance degradation. The integration of this pruning technique allows the DCGAN-EPS model to maintain high accuracy while achieving a more efficient computational footprint. By evaluating the pruned model's performance, it becomes evident that magnitude-based weight pruning effectively balances computational efficiency with robust LC detection capabilities, making the model suitable for deployment in real-time clinical settings.

3.2.3 Optimize the hyperparameter by using an enhanced prism refraction search algorithm

The EPS algorithm is employed in the DCGAN-EPS model to optimize the loss parameters of the neural network, thereby significantly enhancing tumor detection accuracy. EPS leverages the principles of light refraction and the movement of prisms to explore the search space more effectively, balancing exploration and exploitation. This mechanism allows for more robust parameter optimization, as EPS navigates complex landscapes by avoiding local minima and converging on more optimal solutions. Compared to traditional optimization algorithms, which struggle with high-dimensional spaces or get trapped in suboptimal points, EPS's unique approach facilitates better tuning of the neural network's hyperparameters and loss functions. Consequently, this results in improved model performance and accuracy in tumor detection, making it a superior choice for optimizing neural networks in medical imaging tasks.

In the EPS algorithm, the Stagnation Finding and Replacing (SFR) strategy is used to enhance the functioning of the classification model by describing the inspiration from the refraction of a prism, where the refraction of lights falls on a prism follows specific rules and strategies [16].

- **Initialization**

The initial populations of lights are arbitrarily circulated within the exploration area as follows in Eq. (15).

$$L_{z,v}^o = \text{var}_u^{\min} + \text{random}_{[0,1]}(\text{var}_u^{\max} - \text{var}_u^{\min}), \quad z = 1, 2, 3, \dots, Z; u = 1, 2, 3, \dots, L \quad (15)$$

where, $L_{z,v}^o$ signifies the original value of the u th variable of the z th lights. var_u^{\max} And var_u^{\min} are the highest and lowest acceptable values for the u th variable, correspondingly. $\text{random}_{[0,1]}$ Signifies an arbitrary value with a uniform circulation in the space $[0, 1][0, 1]$. Additionally, particular light refraction within the search space is randomly generated and the quantity of prisms is regulated by the user.

- **Evaluating fitness function**

The fitness of the problem is to optimize both cross-entropy loss and feature-matching loss helps to improve the efficacy of the detection system. The exact formulation of the fitness function is described through Eq. (16).

$$\text{EPS}_{fitness} = \text{Optimize}\{L_G, L_D\} \quad (16)$$

where $\text{EPS}_{fitness}$ represents the fitness function of the specified problem. In the EPS algorithm, the SFR strategy maintains population diversity effectively throughout the optimization

process. This strategy enhances the performance of the search process by incorporating distance-based behavior. Consider the matrix $Z(d) = \{Z_{1D}, Z_{2D}, \dots, Z_{jD}(d), \dots, Z_{nD}(d)\}$ representing the traveling of lights in the current iteration d within a D-dimensional search space. Each vector $Z_{jD}(d)$ signifies the location of the j th lights in the specified area. The matrix $Z(d)$ is adjusted in the first iteration by a uniform arbitrary circulation. In subsequent iterations (when $d \geq 2$), the new positions of the lights are determined using Eq. (17).

$$Z_j(d+1) = \begin{cases} D_j^\beta(d) \times e^{\partial\chi} \times \cos(2\pi d) + H_j(d) & \text{if } j \leq R(d) \\ D_j^\eta(d) \times e^{\partial\chi} \times \cos(2\pi d) + H_R(d) & \text{else} \end{cases} \quad (17)$$

where, $D_j^\beta(d)$ and $D_j^\eta(d)$ represents the $D_j^\eta(d)$ main elements of the SFR strategy. A constant ∂ signifies the shape of the logarithmic spiral and χ signifies the arbitrary number

among the intervals $[-1, 1]$. $H_j(d)$ and $H_R(d)$ signifies the positions of the j th light and the R th light that the constraint R . Utilizing SFR in the EPS algorithm yields improved convergence, accelerating the algorithm's ability to reach satisfactory solutions and reducing the required number of iterations. Also, it effectively balances exploration and exploitation. Finally, the proposed DCGAN-EPS improves feature interpretability by promoting a better understanding of CT image boundaries and feature importance. Importantly, the model's ability to generalize across a variety of lung CT image classifications demonstrates its adaptability and broad utility. The entire process of the EPS algorithm is provided in Algorithm 1.

Algorithm 1. Pseudocode for EPS algorithm

```

1 % Initialize the population through an incident angle, bounds defined by prism angle
2 for iter=1 to MaxIters do
3   for i=1 to OneSolution do
4     for j=1 to Dimensions do
5       % Calculate fitness  $\text{EPS}_{fitness}$  by computing the angle of frequency by Eqn. (16)
6       Calculate fitness
7     end for
8     % Obtain  $BestScore$ 
9     if  $\text{EPS}_{fitness} < BestScore$  then
10       $BestScore = fitness$ 
11    end if
12  end for
13  % Calculate the refractive index
14  for i=1 OneSolution do
15    for j=1 Dimensions do
16      % Update the stagnation finding and Replacing by Eqn. (17)
17      Update the stagnation finding and Replacing
18      % Define an arbitrary number in the range [-1, 1]
19      Define an arbitrary number
20      % Find the incident angle and prism angle
21      Find the incident angle and prism angle
22      % Update the best solution and position
23      Update the best solution and position
24    end for
25  end for
26 end for

```

- Computational complexity

This section details the computational complexity of the proposed DCGAN-EPS framework algorithm, addressing both space and time complexity.

(a) Time complexity

The time complexity of the DCGAN-EPS algorithm is $C(J_{CT}(a \times b))$, where a and b represents the number of rows and columns in each image. Calculating SFR $C(J_{CT}(M_i \times SFR \times a \times b))$ for each pixel involves a specific time cost, while the UNet computations take the maximum number of iterations M_i required for segmentation. The overall complexity for classifying cancer types is

$$C(J_{MSFR}) = C(J_{CT-SFR}(C(J_{CT}(M_i \times SFR \times a \times b))))J_{CT-SFR}(C(J_{CT}(M_i \times SFR \times a \times b))).$$

(b) Space complexity

The space complexity of the DCGAN-EPS algorithm is established during the initial phase and remains consistent throughout. As a result, the overall space complexity is.

The implementation of a complex two-stage process within the DCGAN-EPS model, which incorporates GANs, U-Net architectures, and the EPS algorithm, introduces significant computational overhead. However, this complexity is mitigated during the loss parameter tuning phase, where the EPS algorithm's efficient search capabilities streamline the optimization process. By automatically addressing training issues like vanishing gradients and overfitting, EPS facilitates faster convergence of the neural network. This reduction in training time and computational demands not only enhances the overall efficiency of the model but also leads to a decrease in complexity. Thus, while the model's architecture is sophisticated, the strategic integration of EPS effectively balances the trade-off between performance and computational resources, ensuring a more efficient workflow in tumor detection tasks.

3.3 Explainability for the detection of lung tumors by local interpretable model-agnostic explanations

Medical imaging relies on AI-driven interpretability. This research employs LIME to explain the decision-making process of lung CT scan detectors. LIME disturbs samples around a target instance and utilizes the original classifier to predict class labels for these disturbed samples. It creates a localized weighted sample space based on similarity to the target instance. Within this space, LIME trains an interpretable linear model that approximates the predictions of the original classifier near the target instance. The coefficients of this locally trained model offer perceptions into

categorization, often visualized through masks that highlight crucial image regions. Mathematically, it is represented through Eq. (18).

$$O'_{FC} = \gamma_0 + \gamma_1 F'_1 + \gamma_2 F'_2 + \dots + \gamma_n F'_n + \psi \quad (18)$$

where O'_{FC} represents the predicted class label for the CT sample F' . γ_0 denotes the intercept of the linear model, while F'_1, F'_2 and F'_n signifies the features of the perturbed sample CT sample F' . γ_1, γ_2 and γ_n signifies the coefficients of the linear model and ψ signifies the error term. The trained model coefficients are subsequently utilized to interpret the prediction of the original instance F , which is expressed through Eq. (19).

$$\min q \in Q \sum W(F')(O'_{FC} - q(F'))^2 + \Omega(q) \quad (19)$$

where Q represents the set of potential models, q denoting the model under training. $W(F')$ Represents the weights assigned to each disturbed sample, and $\Omega(q)$ stands for several non-zero coefficients in a linear model. The purpose is to identify the model q that minimizes this unbiased function, which explains the original instance. Incorporating LIME in the LC detection process offers key advantages, particularly in enhancing model transparency and trust in AI-driven decisions. LIME allows clinicians to understand the rationale behind tumor detection by identifying which regions of lung CT scans contributed most to the model's predictions. This is crucial for validating AI decisions in a clinical setting, as it enables the verification of important image features like tumor boundaries. Additionally, LIME's model-agnostic nature allows it to be applied across various classification models, ensuring flexibility and consistency in generating interpretable results without modifying the underlying algorithms.

4 Experimental outcomes and discussions

To assess the performance of the introduced two-stage DCGAN-EPS in contemporary LC detection environments.

Table 2 Experimental parameter descriptions

Hyperparameters	Description
Epochs count	100
Batch Size	32
Learning rate	0.003
Activation function	ReLU
Iterations count	200
Search agents count	30
Optimizer	EPS

This research conducted experiments using Python 3, Keras 2.2.4, and TensorFlow on the considered datasets. The experiments are carried out on a computer running Ubuntu 19.04, equipped with an Intel i9-9900KF CPU, an NVIDIA RTX 2080 Ti GPU, and 64 GB of memory. Moreover, the hyperparameters of the proposed DCGAN-EPS technique are listed in Table 2.

The hyperparameters within the fully connected layers of this DCGAN-EPS model are fine-tuned using EPS optimization. These parameters determine the overall performance of the network and their choice of the specific application of the particular layer employed in the proposed architecture.

4.1 Dataset description

The proposed DCGAN-EPS underwent evaluation through two benchmark datasets, which are described below.

LIDC-IDRI dataset (Dataset A): This dataset comprises an entire of 4335 CT images for assessment. Among these, 2160 images are categorized as benign, while 2175 are labeled as malignant. A comprehensive investigation of the LIDC-IDRI dataset is conducted by four radiologists, focusing on the depiction of nodule sizes in lung CT scans.

Chest CT-Scan images Dataset (Dataset B) [23]: This extensive dataset includes four distinct classes:

adenocarcinoma, and large. Cell. Carcinoma, squamous. Cell. Carcinoma and normal, with a total of 1000 CT images. These datasets are divided into training, testing, and validation sets. Within two datasets, 20% of the images serve as test data, while 20% of the remaining 80% are allocated for validation, aiding in the identification of the best-performing model.

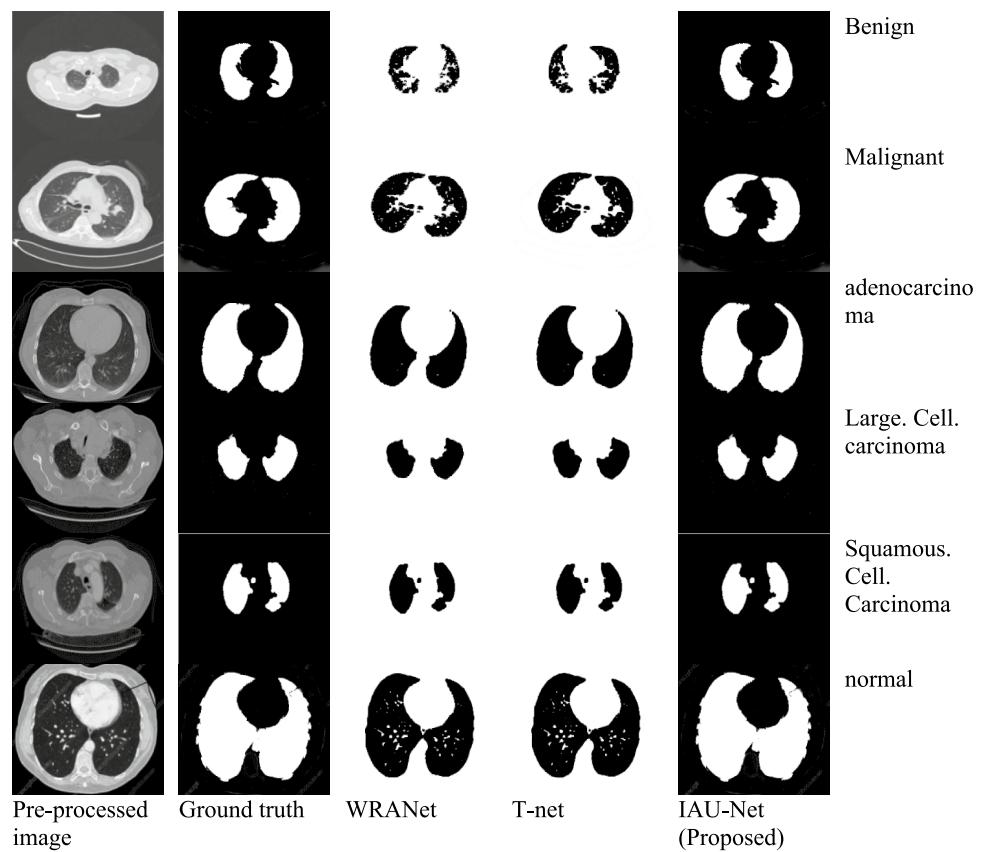
IQ-OTH/NCCD - Lung Cancer Dataset (Dataset C): The IQ-OTH/NCCD LC Dataset (Dataset C) was collected from the Iraq-Oncology Teaching Hospital (IOTH) and National Center for Cancer Diseases (NCCD) over a 3-month period in 2019. It includes 1190 CT scan images of 110 patients, categorized into three classes: 40 malignant cases, 15 benign cases, and 55 normal cases. The CT scans, originally in DICOM format, are captured using Siemens SOMATOM scanners with a slice thickness of 1 mm and specific window parameters.

4.2 Qualitative analysis

This section elaborates on the effectiveness and efficacy of the DCGAN-EPS in identifying lung tumors through visual inspection and interpretation.

Figure 3 presents a comparison of segmentation results across different networks, including the pre-processed image, ground truth, WRANet [38], T-net (Thangavel and

Fig. 3 Visualization of segmentation results from different networks



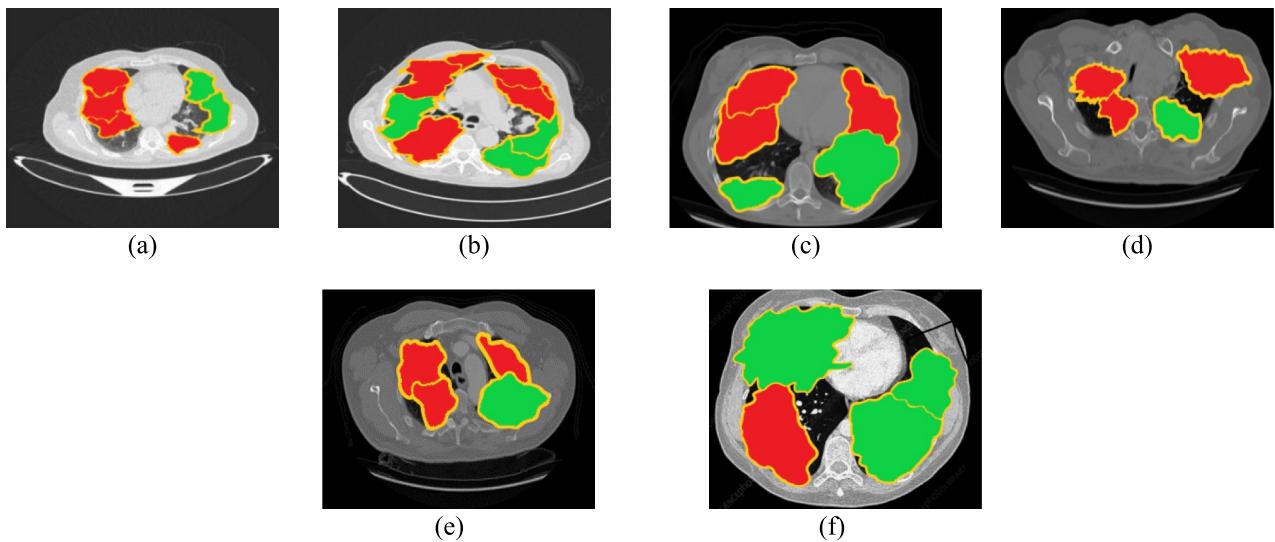


Fig. 4 Explainability visualization for **a** Benign **b** Malignant **c** adenocarcinoma **d** large. Cell. Carcinoma **(e)** squamous. Cell. Carcinoma **(f)** normal classes

Palanichamy 2024a), and the IAU-Net. The IAU-Net demonstrates notable improvements, as it provides more accurate and detailed segmentation of the six classes compared to WRANet and T-net. The proposed IAU-Net model effectively leverages advanced attention mechanisms to focus on relevant features, resulting in improved precision and better alignment with the ground truth. Additionally, IAU-Net handles variability in the image data more effectively, ensuring superior class separation and reduced misclassifications. Figure 4 illustrates the explainability visualization using the LIME model across various classes. The LIME model enhances interpretability by highlighting the regions of the input images most influential to the classification outcomes for each class. The different colors are used to visualize the impact of different features on the model's predictions. Red typically represents features that have a positive impact on the predicted class, while green indicates features that have a negative impact. This approach provides clear visual explanations of the model that differentiate between different types of lung conditions, offering insights into the decision-making process.

Figure 5a–c presents a confusion matrix illustrating the DCGAN-EPS performance in classifying LC images. This visual representation enables an evaluation of the introduced model's classification accuracy by depicting True Positive (TP), True Negative (TN), FP, and False Negative (FN) predictions. The introduced DCGAN-EPS attains an accuracy of 99.92%, 99.4%, and 99.92% on Datasets A, B, and C respectively.

Training loss and accuracy curves for different networks are illustrated in Fig. 6a–f, correspondingly. In that, compare DCGAN-EPS with several conventional classification

models like ATCNN, DFF-CON, WVDN, DL-LCD, VGG-CapsNet, LungEffNet, VGG19+SVM, VGG16+SVM, VGG19+RF and VGG16+RF. All models are trained and tested in the same environment, with their loss and accuracy.

The comprehensive processing time includes 10 min for data preprocessing, 15 min for segmentation, 8 min for feature extraction, and 12 min for post-processing, totaling 45 min. Model training with DCGAN-EPS takes approximately 4 h and 10 min, considering 5 min per epoch over 100 epochs. Therefore, the overall processing time per dataset is approximately 4 h and 55 min.

4.3 Performance analysis

Figure 7 depicts the relationship between the kernel parameter size and the performance metrics of this approach. The kernel parameter determines the shape and flexibility of the boundary used for CT image classification. By tuning the kernel parameter size, adjusting the model's capacity to capture complex patterns in the data. Finding the optimal kernel parameter size is essential for achieving the best balance between model flexibility and generalization ability. In that, the maximum performance parameters are achieved at a kernel size of 3.6.

Table 3 showcases the performance evaluation of DCGAN-EPS and its various configurations on Dataset A and Dataset B. Notably, DCGAN-EPS exhibits the highest accuracy among the different DCGAN-EPS variants tested on both LC datasets. It attained 99.91% accuracy on benign and 99.95% accuracy on malignant in Dataset A. Similarly, in Dataset B achieved 99.94%, 99.86%, 99.96% and 99.92%

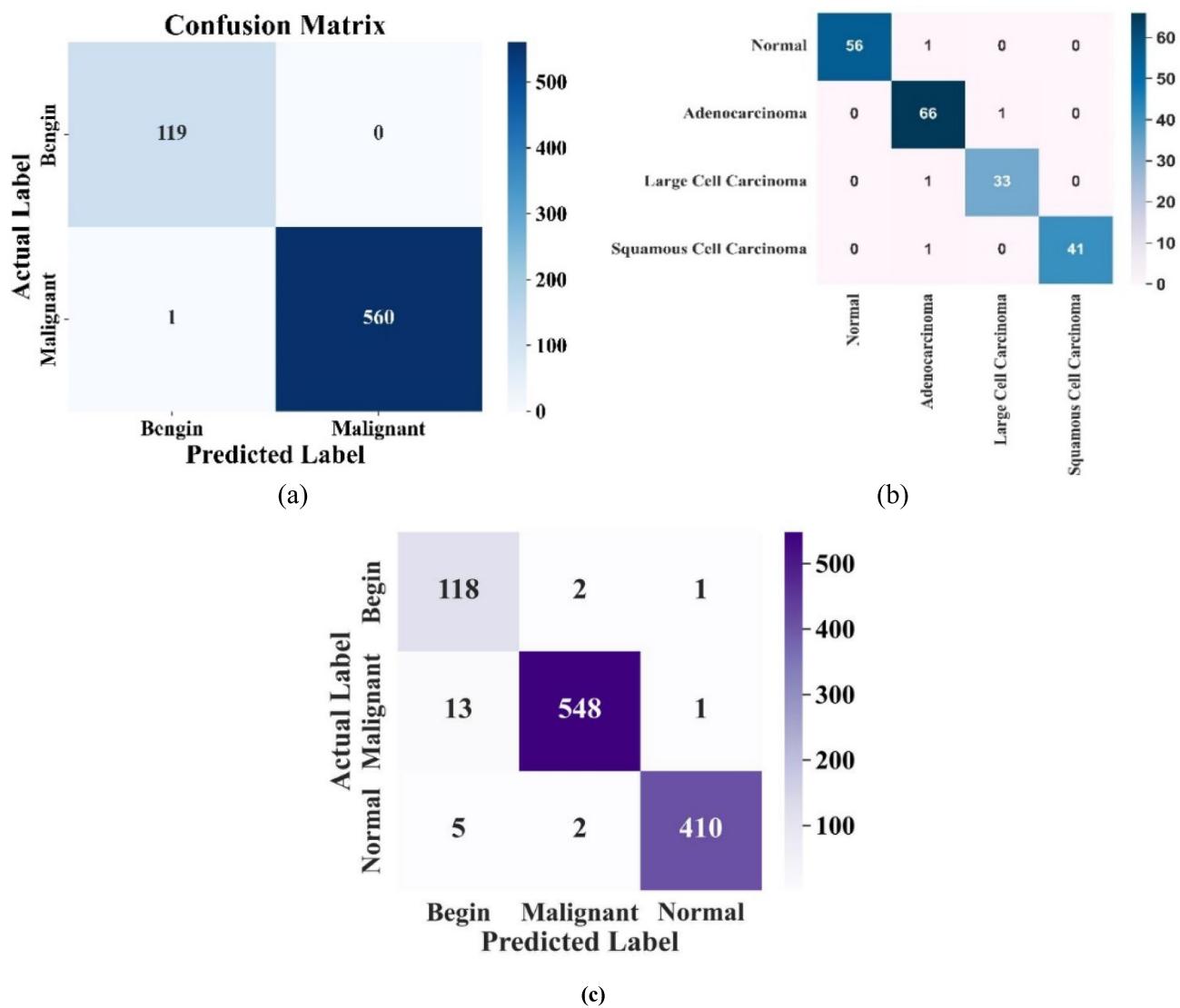


Fig. 5 Confusion matrix for DCGAN-EPS **a** Dataset A **b** Dataset B and **c** Dataset C

for adenocarcinoma, large. Cell. Carcinoma, squamous. Cell. Carcinoma and normal classes respectively.

Figure 8a–c illustrate the Area Under the Curve Receiver Operating Characteristic (AUC-ROC) performance of the introduced approach. This curve is a visual illustration that assesses the performance of a binary categorization at various threshold levels. The performance comparison of AUC-ROC curves across various methods including SDL, VGG-CapsNet GF-EDBN, LCD-CapsNet, Lung-EffNet, LPDCNN, and proposed DCGAN-EPS. The proposed achieved Area Under the Curve (AUC) values of 0.998 on Dataset A, 0.944 on Dataset B, and 1 on Dataset C, which represent the superior performance of the proposed approach.

The purpose of the K-fold test is to assess the model's performance by dividing the dataset into multiple subsets

(folds) and ensuring that each subset is used as both training and validation data across different iterations. In each iteration, the model is trained on $(K-1)$ folds and validated on the remaining fold, allowing the evaluation of its performance across all data points. This technique helps mitigate the bias from random partitioning, provides a more comprehensive understanding of the model's generalization, and highlights any variability in performance. It ultimately ensures that the model is not overfitting and is capable of accurately detecting and classifying LC across different subsets of the data. In this research, K was selected as 5, 10, 15, and 20 to observe the variation in the model's performance with different levels of data partitioning. Lower values of k , such as 5, provide faster validation but have higher variance due to fewer training data in each fold, while higher values like 20 give a more comprehensive evaluation by ensuring that

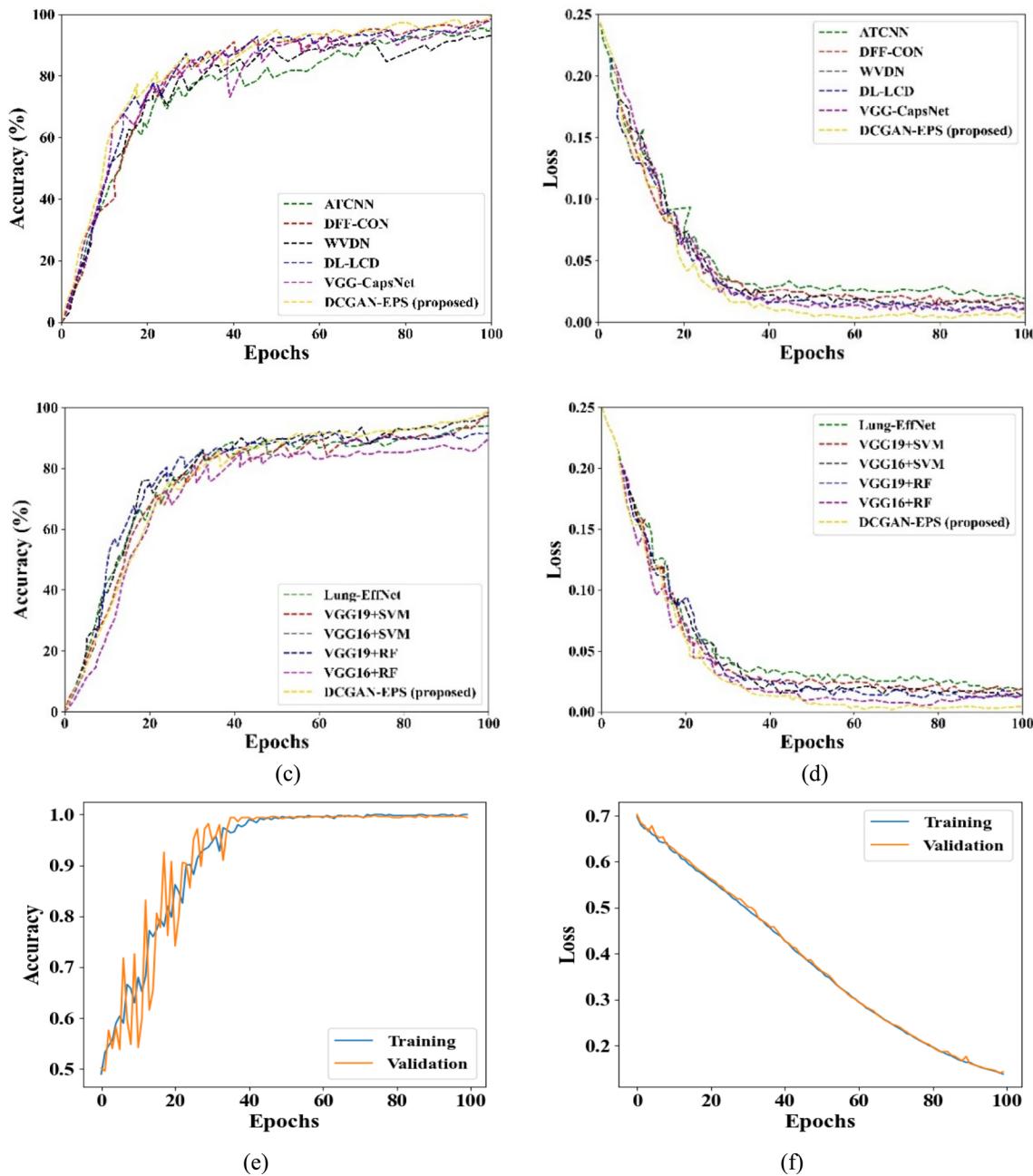


Fig. 6 Training loss and accuracy curves for different networks (a) Accuracy of Dataset A (b) Loss of Dataset A (c) Accuracy of Dataset B (d) Loss of Dataset B (e) Accuracy of Dataset C (f) Loss of Dataset C

most of the data is used for training, leading to more stable and reliable performance metrics. This range of k-values allows for a detailed assessment of the model's generalization ability, aiding in determining the optimal k for accurate LC detection and classification while balancing computational efficiency.

Table 4 presents a comparison of the k-fold validation performance achieved using both datasets. During the training process with DCGAN-EPS, the dataset is partitioned

into 20-fold subsets of around equivalent size. The model undergoes training 20 times, where each iteration utilizes $(k-1=20-1)$ subsets as training data and the remaining subset as validation data. Through this assessment, various performance metrics have been calculated for each fold of the validation process. Notably, in the 20th fold, the model achieved optimal performance than other folds. By comparing the performance metrics across different folds, assess the variability in model performance and determine the overall

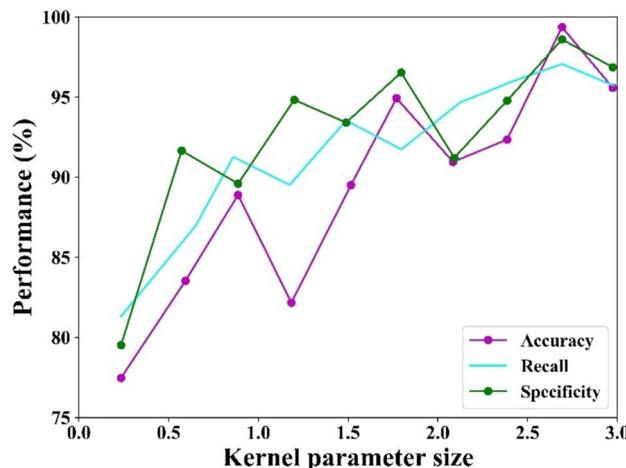


Fig. 7 Graph visualization between overall performance and kernel parameter size

effectiveness of the model in accurately detecting and classifying LC.

In Table 5, when applying 10% pruning, the accuracy across the datasets remains high at 99.95% for Dataset A, 99.7% for Dataset B, and 99.95% for Dataset C, while inference speeds improve from 150.4 to 120.3 ms for Dataset A, indicating a notable efficiency gain. Furthermore, memory usage decreases from 250 to 230 MB, and the model size reduces from 2,350,000 parameters to 1,984,000. As pruning increases to 20% and 30%, accuracy slightly declines (99.94% for Dataset A with 20% pruning) while inference speeds and memory usage continue to improve. For instance, with 30% pruning, the inference speed drops to 85.7 ms and memory usage to 190 MB, indicating that computational demands are effectively minimized without compromising accuracy significantly. Overall, these results illustrate that magnitude-based weight pruning not only streamlines the DCGAN-EPS model but also maintains high accuracy levels while enhancing inference efficiency.

4.4 Generalizability experiment

To assess DCGAN-EPS generalization ability, the model is re-trained and tested on two distinct datasets encompassing various LC types.

Figure 9a shows the performance of a cancer classification model on three different types of cancer cells of Dataset A. In that, the model performs best at classifying benign and malignant cells, but still with an accuracy of over 95%. Figure 9b shows the performance of a DCGAN-EPS model on four different types of cancer cells of Dataset B. In that, the model performs best at classifying adenocarcinoma cells, with an average performance of 99.5%. Also, it performs less performance on squamous cell carcinoma and large cell carcinoma, but still with an accuracy of over 90%.

For LC detection and subtype classification using both datasets, the experimental setups are as follows, GAN-LSTM-3D employs a batch size of 32, a learning rate of 0.0002, Leaky ReLU for the GAN and Tanh for LSTM outputs, with 64 units in LSTM and 3D convolutional layers with BN. GAN uses a batch size of 64, a learning rate of 0.0001, ReLU activation for the generator, and Sigmoid for the discriminator output, featuring dense layers with BN and transposed convolutions.

SD-GAN operates with a batch size of 16, a learning rate of 0.0005, ELU for the generator, and Leaky ReLU for the discriminator, incorporating spatial attention mechanisms in the generator and spatial discriminative layers in the discriminator. GAN-CNN has a batch size of 64, a learning rate of 0.0003, ReLU activation for CNN layers, and Tanh for the generator output, with convolutional layers followed by up-sampling in the generator and BN in the discriminator. Moreover, the Comparison of GAN models under different epochs is shown in Table 6.

At 25 epochs, the DCGAN-EPS model outperforms others with accuracy and recall values of 95.51% and 95.11% for Dataset A, and 96.16% and 96.75% for Dataset B. At 50 epochs, DCGAN-EPS achieves 97.13% accuracy and 97.15% recall for Dataset A, and 98.15% accuracy and 97.98% recall for Dataset B. At 75 epochs, DCGAN-EPS continues to lead with 98.43% accuracy and 98.65% recall for Dataset A, and

Table 3 Different components of Proposed DCGAN-EPS results on different classes of LC

Model under experiment	Dataset A		Dataset B			
	Benign	Malignant	adenocarcinoma	large. Cell. carcinoma	squamous. Cell. Carcinoma	Normal
Graph representation + DCKM	99.84%	99.84%	99.97%	99.75%	99.89%	99.85%
Graph representation + CMCN	99.85%	99.85%	99.92%	99.84%	99.90%	99.87%
DCKM + CMCN	99.89%	99.91%	99.91%	99.85%	99.94%	99.91%
Proposed DCGAN-EPS	99.91%	99.95%	99.94%	99.86%	99.96%	99.92%

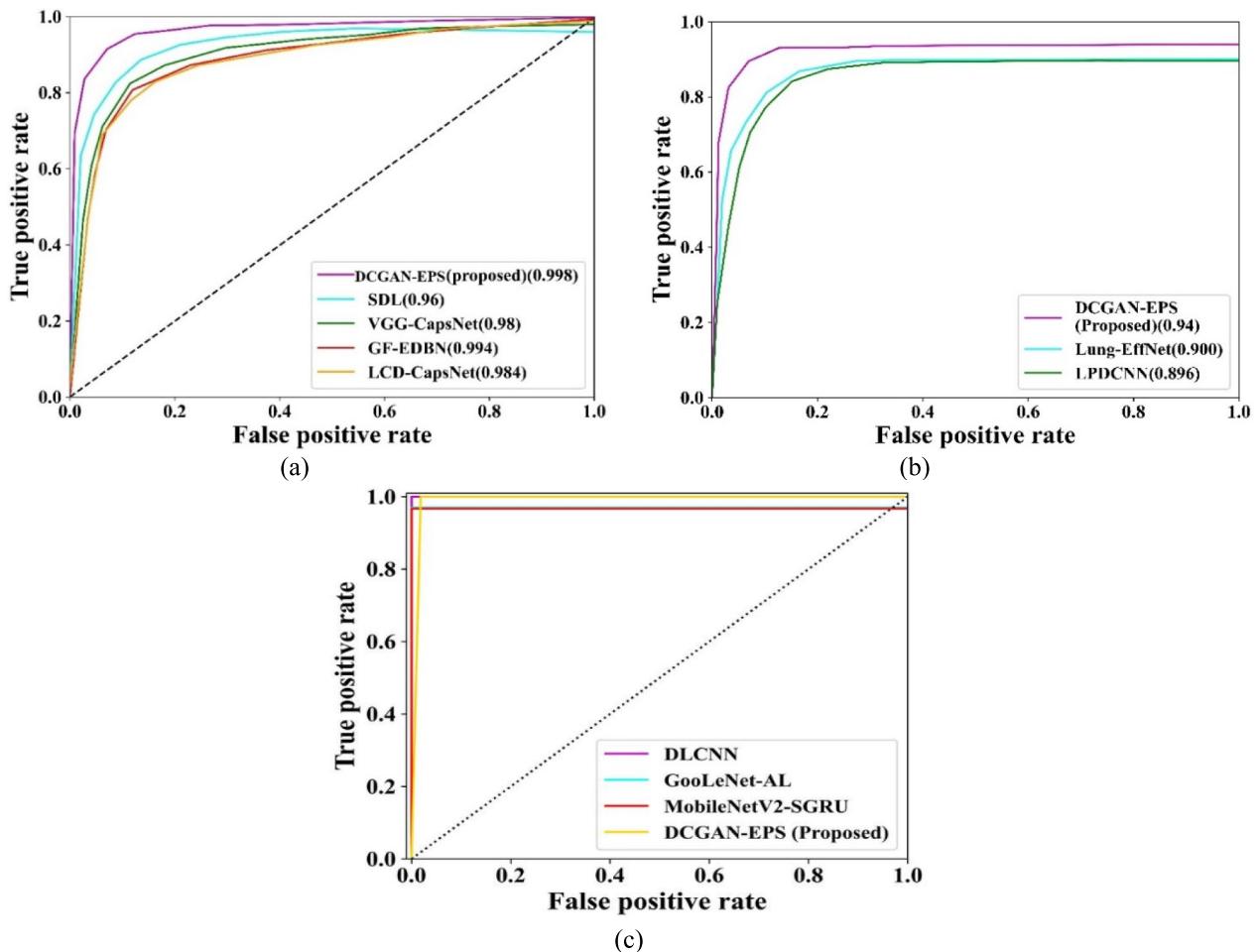


Fig. 8 AUC-ROC performance of proposed DCGAN-EPS on **a** Dataset A **b** Dataset B and **c** Dataset C

Table 4 Comparison of k-fold validation performance by using DCGAN-EPS

Datasets	K-fold	Fold-5	Fold-10	Fold-15	Fold-20
Dataset A	Accuracy (%)	99.56	99.92	99.94	99.92
	Precision (%)	99.56	99.95	99.71	99.94
	Recall (%)	99.72	99.94	99.97	99.94
	F1-Score (%)	99.65	99.89	99.71	99.92
	Specificity (%)	99.34	99.87	99.92	99.90
Dataset B	Accuracy (%)	99.0	99.82	99.85	99.42
	Precision (%)	99.88	99.33	99.06	99.25
	Recall (%)	99.42	99.56	99.05	99.53
	F1-Score (%)	99.90	99.89	99.87	99.47
	Specificity (%)	99.87	99.88	99.88	99.22

98.24% accuracy and 98.15% recall for Dataset B. Finally, at 100 epochs, DCGAN-EPS reaches 99.92% accuracy and 99.94% recall for Dataset A, and 99.4% accuracy and 99.5%

recall for Dataset B, consistently outperforming all other models in both datasets and across epochs.

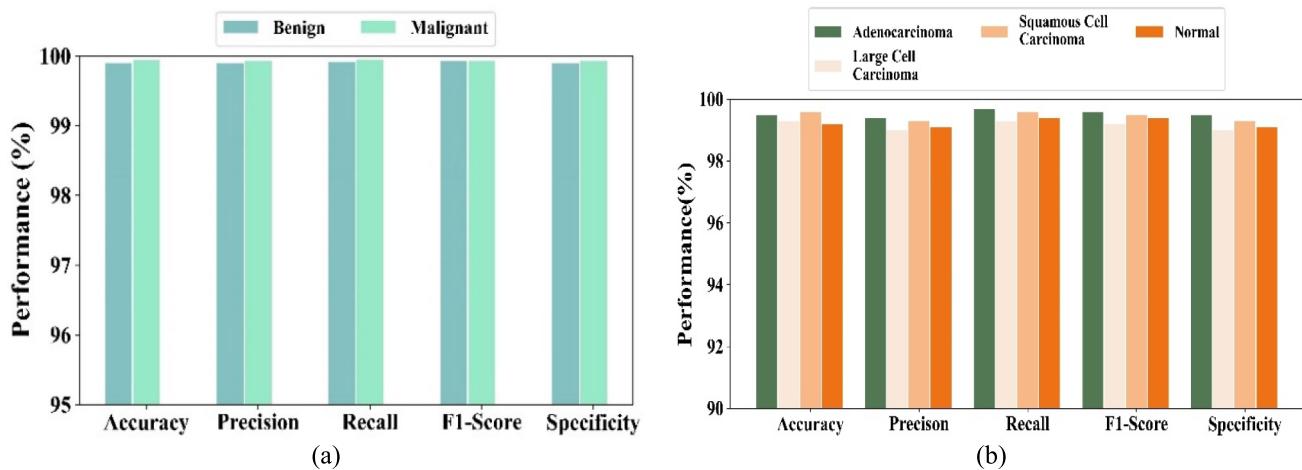
Figure 10 shows that DCGAN-EPS achieves superior balance performance compared to GAN [27] and DS-GAN [15], as indicated by a higher BR metric.

The y-axis represents the loss, and the x-axis shows the number of iterations. DCGAN-EPS demonstrates a rapid decline in loss, converging around the 20th epoch, which indicates its efficient optimization of the binary cross-entropy and adversarial losses. The faster convergence of DCGAN-EPS is attributed to its unique integration of a deep convolutional architecture with cross-max-out kernels and the EPS, which effectively explores the solution space and refines the learning process.

Despite this accelerated convergence, DCGAN-EPS also maintains a higher average accuracy across the tested models, highlighting its effectiveness and reliability in achieving an optimal solution in fewer epochs. While standard GANs struggle with balance due to mode collapse and DCGAN-EPS improves upon this with better

Table 5 Comparison of DCGAN-EPS with Magnitude-based weight pruning

Model Configuration	Dataset	Accuracy (%)	Inference Speed (ms)	Inference Memory Usage (MB)	Model Size (Parameters)
DCGAN-EPS without Pruning	Dataset A	99.96	150.4	250	2,350,000
	Dataset B	99.8	140.2	240	2,520,000
	Dataset C	99.96	145.2	245	2,260,000
DCGAN-EPS with 10% Pruning	Dataset A	99.95	120.3	230	1,984,000
	Dataset B	99.7	115.6	220	1,986,000
	Dataset C	99.95	118.3	225	1,962,000
DCGAN-EPS with 20% Pruning	Dataset A	99.94	100.5	210	1,868,000
	Dataset B	99.6	95.3	205	1,879,000
	Dataset C	99.94	98.6	208	1,846,000
DCGAN-EPS with 30% Pruning	Dataset A	99.93	85.7	190	1,784,000
	Dataset B	99.5	80.6	185	1,760,000
	Dataset C	99.93	82.4	187	1,748,000
DCGAN-EPS with 40% Pruning	Dataset A	99.92	70.1	170	1,643,000
	Dataset B	99.4	65.6	165	1,669,000
	Dataset C	99.92	68.15	168	1,689,000

**Fig. 9** Performance comparison on difference classes **a** Dataset A **b** Dataset B

training stability, DCGAN-EPS excels with its multiple adversarial components, offering the best balance. This makes DCGAN-EPS particularly suitable for analyzing LC detection, as its enhanced balance improves the quality of synthetic data and diagnostic accuracy.

4.5 Comparative analysis with state-of-art

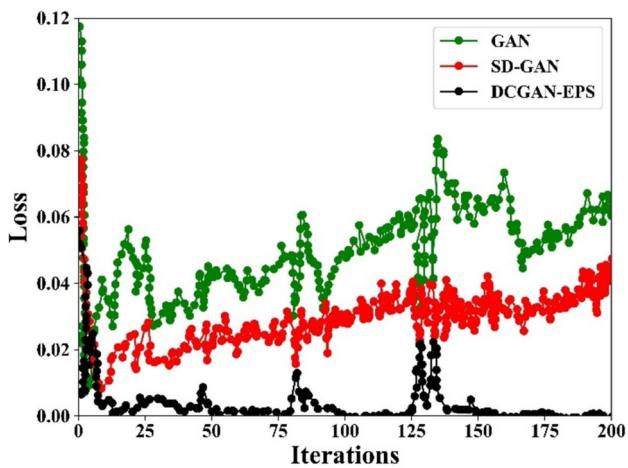
Comparative assessment with existing approaches involves evaluating the proposed method in comparison to established approaches using various performance metrics. These metrics include precision, accuracy, recall, F1-Score, specificity, and computation time. The comparative assessment is performed with previous techniques like ATCNN, DFF-CON,

SDL, EDL, WVDN, DL-LCD, LCD CapsNet, GF EDBN, VGG-CapsNet, CapsNet, LungEffNet, VGG19+SVM, VGG19+SVM, VGG16+SVM, VGG19+RF, VGG16+RF and LPDCNN.

Table 7 presents a comparative analysis of various LC methods on Dataset A, where the DCGAN-EPS approach significantly outperformed other methods, achieving 99.92% accuracy, 99.92% precision, 92.41% recall, and a 99.92% F1-score. The superior performance of DCGAN-EPS is attributed to its ability to effectively generate enhanced data representations through its GAN-based architecture, which helps improve the model's capability to generalize across complex data distributions. By employing an efficient generative and discriminative learning framework, DCGAN-EPS

Table 6 Comparison of GAN models under epochs

Epochs	Models	Dataset A		Dataset B	
		Accuracy (%)	Recall (%)	Accuracy (%)	Recall (%)
25	GAN-LSTM-3D	85.23	84.56	82.14	83.72
	GAN	80.89	82.45	81.54	80.67
	SD-GAN	84.15	83.21	83.45	82.29
	GAN-CNN	81.94	80.67	82.76	81.34
	DCGAN-EPS(Proposed)	95.51	95.11	96.16	96.75
50	GAN-LSTM-3D	88.12	87.65	86.48	85.78
	GAN	84.34	83.89	85.12	84.65
	SD-GAN	87.95	88.23	86.78	87.45
	GAN-CNN	85.67	84.90	84.56	85.23
	DCGAN-EPS(Proposed)	97.13	97.15	98.15	97.98
75	GAN-LSTM-3D	90.85	91.34	89.76	90.23
	GAN	87.98	86.45	88.56	87.34
	SD-GAN	91.23	90.89	90.12	89.45
	GAN-CNN	94.78	92.45	92.97	91.35
	DCGAN-EPS(Proposed)	98.43	98.65	98.24	98.15
100	GAN-LSTM-3D	93.54	94.12	92.78	93.23
	GAN	90.87	89.54	91.45	90.12
	SD-GAN	94.23	93.89	93.56	92.34
	GAN-CNN	92.45	91.67	91.78	92.12
	DCGAN-EPS(Proposed)	99.92	99.94	99.4	99.5

**Fig. 10** Balance performance of DCGAN-EPS using BR metric

reduces overfitting, leading to higher precision and recall, which contributes to its robust classification outcomes. Additionally, its advanced pixel-level synthesis and error propagation control further minimize false positives and false negatives, enhancing overall performance compared to traditional methods. Similarly, Table 8 shows the performance of various LC methods on Dataset B, where DCGAN-EPS achieved 99.4% accuracy, 99.2% precision, 99.5% recall and a 99.4% F1-score. Again, the model's GAN-based architecture plays a crucial role in providing a more accurate

representation of underlying data patterns. The effective use of deep convolutional layers allows DCGAN-EPS to extract intricate features, capturing both local and global structures in the images, which contributes to its higher recall and F1 score when compared to other techniques. Moreover, the model's ability to dynamically adapt to changing data patterns and its efficient training mechanism enables superior performance even in complex scenarios.

Table 9 outlines the performance metrics of various classification methods, demonstrating their effectiveness in terms of accuracy, precision, recall, F1-score, specificity, sensitivity, and Kappa statistics. The proposed method DCGAN-EPS, achieved the highest accuracy of 99.92% and impressive precision (99.89%), recall (99.91%), and F1-score (99.93%), showcasing its exceptional capability for accurate classification. The advantages of DCGAN-EPS include its ability to maintain high precision and recall simultaneously, which reduces the chances of false positives and negatives, thus ensuring reliable predictions. Additionally, its outstanding performance in Kappa statistics (99.925%) indicates strong agreement between predicted and observed classifications, enhancing its credibility. In comparison, the DLCNN [22] method also demonstrated strong performance with an accuracy of 99.64% and an F1-score of 99.64%, making it a robust alternative. Other methods, like (GoogLeNet-AL) [20] and UDCT [10], offered competitive precision values (99.45% and 98.7%, respectively), although they lacked complete recall

Table 7 Comparison of DCGAN-EPS method with existing methods on Dataset A

Methods	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Specificity (%)
ATCNN	95.97	—	95.84	—	96.89
DFF-CON	98.09	99.8	99.8	99.88	99.76
SDL	92.84	92.63	92.41	92.51	
EDL	99.34	99.12	99.18	99.15	99.24
WVDN	93.22	93.22	93.2	93.1	—
DL-LCD	97.88	94.94	97.63	96.42	—
LCD-CapsNet	94	95	94.5	94.5	99.07
GF-EDBN	99.42	—	98.49	99.37	98.31
VGG-CapsNet	98.61	99.07	98.16	98.61	99.07
DCGAN-EPS (proposed)	99.92	99.92	99.94	99.92	99.90

Table 8 Comparison of the DCGAN-EPS method with existing methods on Dataset B

Methods	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Specificity (%)
Lung-EffNet	94	94.98	92.25	93.61	—
VGG19+SVM	98.70	—	98.70	99.33	99
VGG16+SVM	97.40	—	97.73	98.52	99
VGG19+RF	91.55	—	91.55	95.58	96
VGG16+RF	89.94	—	89.93	94.65	98
LPDCNN	98.35	97.14	98.15	97.1	
DCGAN-EPS (Proposed)	99.4	99.2	99.5	99.4	99.2

Table 9 Comparison of the DCGAN-EPS method with existing methods on Dataset C

Methods	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Specificity (%)	Sensitivity (%)	Kappa (%)
DLCNN	99.64	96.77	99.04	99.64	—	—	99.38
GoogLeNet-AL	97.32	99.45	—	98.82	97.63	98.2	—
UDCT	96.82	98.7	97.5	98.24	98.4	—	—
MobileNetV2-SGRU	96.83	96.78	96.83	96.78	—	—	—
DCGAN-EPS (Proposed)	99.92	99.89	99.91	99.93	99.9	98.07	99.925

Table 10 Comparison of execution time and error rate of proposed DCGAN-EPS over three datasets

Methods	Execution Time (s)	Error rate
GF-EDBN	16	0.0057
EDL	2.27	—
DLCNN	—	0.0073
GoogLeNet-AL	18.23	—
UDCT	—	0.25
MobileNetV2-SGRU	475.1340	—
DCGAN-EPS (proposed)	2.18	0.0031

and sensitivity data. Overall, DCGAN-EPS stands out for its high performance across multiple metrics, indicating its potential for robust classification tasks, while methods like DLCNN also provide solid alternatives with significant advantages in specific scenarios.

Table 10 summarizes the execution time and error rates of various classification methods, illustrating their efficiency and accuracy. Among the methods, DCGAN-EPS stands out with the shortest execution time of 2.18 seconds and a low error rate of 0.0031, highlighting its efficiency and effectiveness. The EDL method follows closely, requiring only 2.27 seconds, but lacks an associated error rate. In contrast, MobileNetV2-SGRU (Tofighi et al. 2024a) has a significantly longer execution time of 475.1340 seconds, indicating a trade-off between processing speed and potentially other performance metrics not listed. Other methods like GF-EDBN and GoogLeNet-AL show error rates of 0.0057 and no reported error rates, respectively, while DLCNN and UDCT also provide error rates of 0.0073 and 0.25, respectively, without execution time data. Overall, the table emphasizes that the proposed DCGAN-EPS method not only excels in accuracy but also in execution speed, making it a compelling choice for efficient classification tasks.

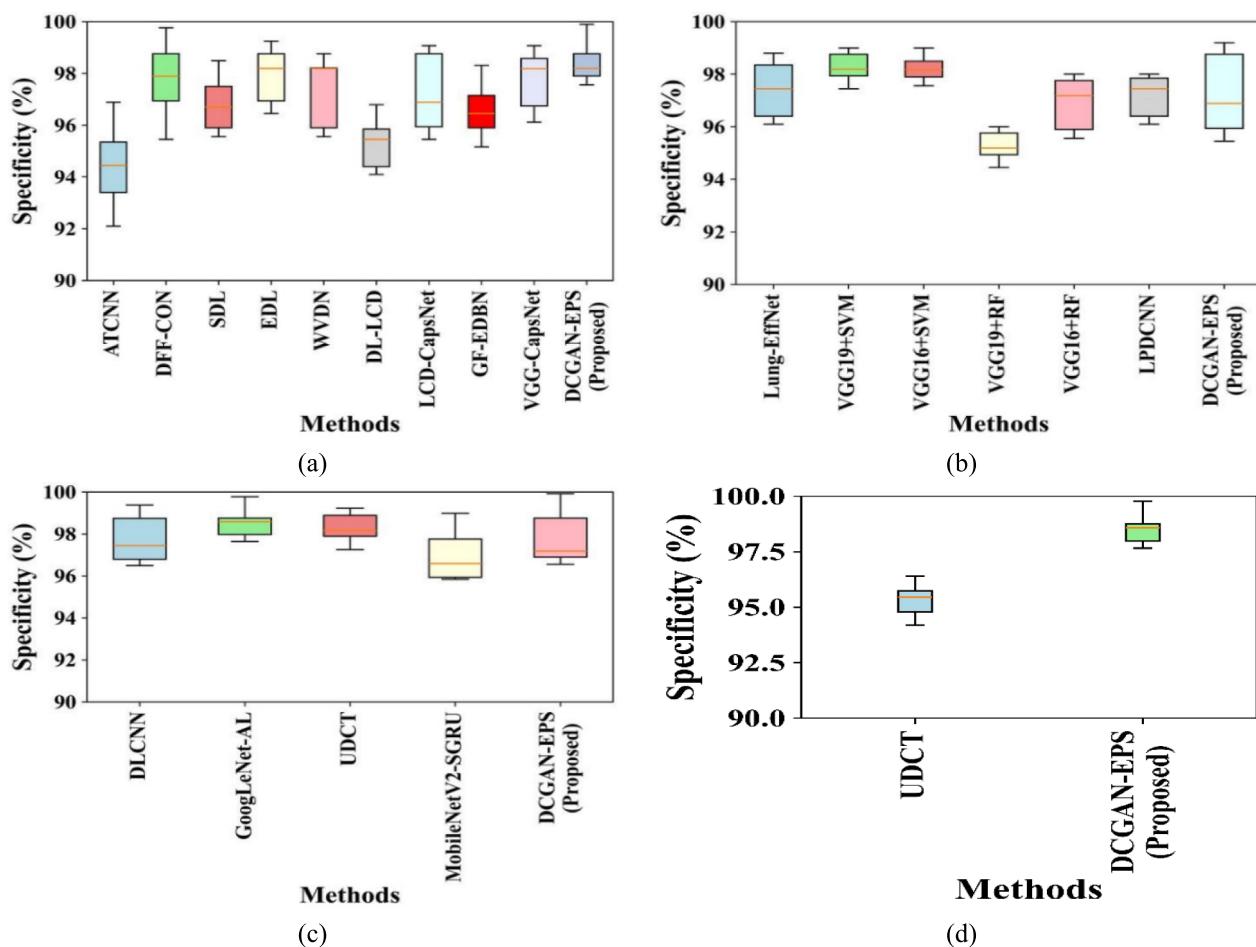


Fig. 11 Box-plot analysis for specificity comparison on **a** Dataset A, **b** Dataset B **c** Dataset C, and **d** overall dice-coefficient

In terms of specificity in Fig. 11a–c, the DCGAN-EPS method on Dataset A achieved an outstanding 99.90%, significantly higher than other approaches. The proposed model performs better because it minimizes false positives through the efficient learning of detailed data patterns, especially those associated with normal and abnormal cases, reducing the likelihood of incorrectly classifying normal cases as abnormal. The GAN-based architecture, coupled with precise feature extraction, ensures better discrimination between positive and negative samples.

On Datasets B and C, DCGAN-EPS again demonstrated superior performance with a specificity of 99.2% and 98.27%, outperforming other approaches. The higher specificity of DCGAN-EPS is due to its ability to learn and generate more representative features of normal cases, thus reducing false alarms and improving overall classification accuracy. By incorporating robust training techniques and improved generalization capabilities, the model maintains higher specificity compared to conventional LC methods.

Figure 11d presents a comparative box plot illustrating the specificity (%) of two methods: UDCT and the

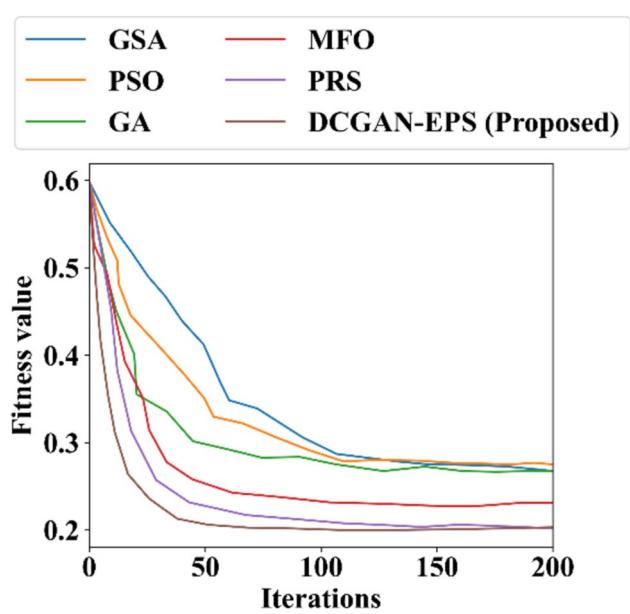


Fig. 12 Convergence comparison with various optimization approaches

proposed DCGAN-EPS. The DCGAN-EPS model achieves a higher specificity, with values ranging from approximately 98.5–99.8%, showcasing its ability to accurately identify true negative cases with minimal false positives. In contrast, the UDCT method exhibits a lower specificity, with a range between roughly 94.5% and 97.5%. The higher median and narrower range of specificity for DCGAN-EPS highlight its superior consistency and performance in distinguishing normal cases from abnormal ones, making it a more reliable model for classification tasks compared to UDCT.

Figure 12 compares the performance of several optimization algorithms like GSA (Gravitational Search Algorithm), PSO (Particle Swarm Optimization), GA (Genetic Algorithm), MFO (Moth Flame Optimization), PRS (Prism Refraction Search) [16].and the proposed DCGAN-EPS. In that, the DCGAN-EPS approach exhibits the most efficient performance, with its fitness value decreasing steadily and reaching approximately 0.18 by the 200th iteration. This improvement suggests that the DCGAN-EPS effectively optimizes the parameters crucial for LC recognition, thereby enhancing the accuracy of the recognition process. In contrast, the MFO algorithm, which achieves the second-lowest fitness value, stabilizes around 0.22, indicating its reasonable performance but still lagging behind the proposed method. The PSO and GA algorithms perform less effectively, with fitness values approximately at 0.25 and 0.3, respectively, demonstrating a slower convergence and less efficient parameter optimization. The EPS algorithm incorporated in the DCGAN-EPS is key to its superior performance. By leveraging advanced search strategies that simulate the behavior of light refraction, EPS allows for more effective exploration of the solution space, preventing premature convergence and enabling a more thorough search for optimal parameters. This capability results in enhanced recognition accuracy for LC when using the proposed model, illustrating the effectiveness of combining deep learning techniques with advanced optimization strategies in medical image analysis.

4.6 Statistical analysis

In this research, the ANOVA test was conducted to determine whether there are statistically significant differences in the performance of the LC classification models. The test evaluates whether the variance between the treatment groups (different models or techniques) is greater than the variance within the groups. According to the ANOVA results

in Table 11, the treatment effect (among columns) has a sum of squares (SS) of 0.01507, with 2 degrees of freedom (DF) and a mean square (MS) of 0.007617. The F-statistic is $F(3, 76) = 348.6$, and the p-value is less than 0.0001, which is significantly lower than the threshold of 0.05. This demonstrates that the differences in model performance are statistically significant, indicating that the proposed model outperforms others with a high degree of confidence. The residual variation across columns has an SS of 0.001733 with 75 degrees of freedom, reinforcing the reliability of the findings.

The ANOVA results show that the p-values are below 0.05, confirming the statistical significance of the DCGAN-EPS algorithm and demonstrating its superior performance compared to other methods.

4.7 Discussion

The two-stage DCGAN-EPS method integrates advanced preprocessing techniques with GAN-based architectures to enhance LC detection. Preprocessing involves AWDT and bilateral filtering to reduce noise while preserving key features like edges. The AU-Net, utilizing attention mechanisms and skip connections, effectively preserves essential information during segmentation. Feature vectors are extracted using WLBP, leading to clearer images for analysis. This comprehensive preprocessing combined with the DCGAN-EPS framework improves detection accuracy and ensures high-quality input data, while the LIME model provides interpretable, model-agnostic explanations, further enhancing the system's overall effectiveness.

The proposed DCGAN-EPS method has demonstrated superior performance compared to existing models. For instance, in comparison to baseline models, the DCGAN-EPS method achieved an accuracy of 99.92% and a recall of 99.94% on Dataset A at 100 epochs, and an accuracy of 99.40% with a recall of 99.50% on Dataset B. In contrast, other methods like GAN-LSTM-3D and SD-GAN achieved maximum accuracies of 93.54% and 94.23%, respectively, highlighting the DCGAN-EPS method's superior capability in both accuracy and recall. These results underscore the efficacy of combining GAN-based architectures with advanced preprocessing techniques. One notable limitation of the DCGAN-EPS method is its dependence on high-quality, well-annotated datasets, which restricts its applicability in situations where such data is scarce or challenging

Table 11 ANOVA test analysis of the proposed model

ANOVA table	SS	DF	MS	F(DFn, DFd)	P value
Treatment (among columns)	0.01507	2	0.007617	$F(3, 76) = 348.6$	$P < 0.0001$
Residual (in columns)	0.001733	75	2.32E-04	–	–
Overall	0.016803	77	–	–	–

to acquire. Additionally, the model experiences diminished performance in clinical environments with compromised imaging quality. Since the model is primarily trained on high-resolution images, which does not adequately account for the variability encountered in real-world clinical settings, potentially resulting in less accurate predictions when applied to lower-quality scans.

5 Conclusions and future works

LC diagnosis is often performed at advanced stages, complicating treatment and contributing to higher mortality rates. Early detection is recognized as critical in mitigating these risks. In this research, the DCGAN-EPS framework is introduced for effective LC detection and classification. The framework employs the AWDT for data preprocessing, followed by the IAU-Net for image segmentation. The WLBP is then utilized for feature extraction, leading to tumor classification through the EPS algorithm. The effectiveness of the proposed method is validated using two datasets. Remarkable performance is demonstrated, with an accuracy of 99.92%, precision of 99.94%, recall of 99.94%, F1-Score of 99.92%, and specificity of 99.90% achieved for Dataset A. For Dataset B, the method attains 99.4% accuracy, 99.2% recall, 99.5% precision, 99.4% F1-Score, and 99.2% specificity. These outcomes underscore the capability of the DCGAN-EPS framework in accurately detecting and classifying lung tumors from CT scan images. Future work will concentrate on enhancing the computational efficiency of the DCGAN-EPS framework and improving its performance under suboptimal data conditions for better clinical applicability. This will involve integrating generative models with unsupervised learning to effectively utilize unlabeled data and developing adaptive algorithms for low-quality imaging data. Advanced data augmentation, transfer learning, and hardware acceleration strategies will be employed to ensure robust performance across various clinical imaging scenarios. Additionally, the model will be optimized for real-time processing while maintaining high accuracy in lung cancer detection, enabling timely results without compromising diagnostic reliability.

Author contributions Both the authors have contributed equally to the work.

Funding This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability No datasets were generated or analysed during the current study.

Declarations

Conflict of interest The authors declare no competing interests.

Ethical approval All applicable institutional and/or national guidelines for the care and use of animals were followed.

Informed consent For this type of analysis, formal consent is not needed.

References

- Alshammari HH, Alkhiri H (2023) Optimized recurrent neural network mechanism for olive leaf disease diagnosis based on wavelet transform. *Alexandria Eng J* 78:149–161
- Angel Mary A, Thanammal KK (2024) Lung cancer detection via deep learning-based pyramid network with Honey Badger algorithm. *Meas Sens* 31:100993
- Bagheri Tofghi A, Ahmadi A, Mosadegh H (2024) Improving lung cancer detection via MobileNetV2 and stacked-GRU with explainable AI. *Int J Inf Technol*, pp 1–8
- Balachandran S, Ranganathan V (2022) Semantic context-aware attention UNET for lung cancer segmentation and classification. *Int J Imaging Syst Technol* 33:822–836
- Balci MA, Batrancea LM, Akgüller Ö, Nichita A (2023) A Series-based deep learning approach to lung nodule image classification. *Cancers* 15:843
- Bishnoi V, Goel N (2023) Tensor-RT-based transfer learning model for Lung Cancer Classification. *J Digital Imaging* 36:1364–1375
- Bushara A, Vinod Kumar R, Kumar S (2023) LCD-capsule network for the detection and classification of lung cancer on computed tomography images. *Multimedia Tools Appl* 82:37573–37592
- Gopinath A, Gowthaman P, Venkatachalam M, Saroja M (2023) Computer-aided model for lung cancer classification using cat optimized Convolutional Neural Networks. *Meas Sens* 30:100932
- Guo D, Liu Z, Li R (2023) Regraphgan: a graph generative adversarial network model for Dynamic Network Anomaly Detection. *Neural Netw* 166:273–285
- Gupta A, Kumar A, Rautela K (2024) UDCT: lung Cancer detection and classification using U-net and DARTS for medical CT images. *Multimed Tools Appl*, pp 1–21.
- Halder A, Dey D (2023a) Atrous convolution aided Integrated Framework for lung nodule segmentation and classification. *Biomed Signal Process Control* 82:104527
- Hari P, Singh MP (2023a) A lightweight convolutional neural network for disease detection of fruit leaves. *Neural Comput Appl* 35:14855–14866
- Hong L, Modirrousta MH, Hossein Nasirpour M, et al (2023) Gan-LSTM-3D: an efficient method for lung tumor 3D reconstruction enhanced by attention-based LSTM. *CAAI Trans Intell Technol*. 1–10
- Kalaiyarasi M, Janaki R, Sampath A, Ganage D, Chincholkar YD, Budaraju S (2023) Non-additive noise reduction in medical images using bilateral filtering and modular neural networks. *Soft Comput*, pp 1–10
- Kausar T, Lu Y, Kausar A, Ali M, Yousaf A (2023a) SD-gan: a style distribution transfer generative adversarial network for COVID-19 detection through X-ray images. *IEEE Access* 11:24545–24560

16. Kundu R, Chattopadhyay S, Nag S, Navarro MA, Oliva D (2024a) Prism refraction search: a novel physics-based metaheuristic algorithm. *J. Supercomput* 80:10746–10795
17. Lei H, Tian Z, Xie H et al (2023) Lac-Gan: Lesion attention conditional gan for ultra-widefield image synthesis. *Neural Netw* 158:89–98
18. Lu Z, Jia S, Li G, Jing S (2024) Neutron image denoising method based on adaptive new wavelet threshold function. *Nucl Instrum Methods Phys Res Sect A* 1059:169006
19. Lv Z, Zhang Z (2023a) Research on plant leaf recognition method based on multi-feature fusion in different partition blocks. *Digital Signal Process* 134:103907
20. Ma L, Wu H, Samundeeswari P (2024a) GoogLeNet-AL: a fully automated adaptive model for lung cancer detection. *Pattern Recognit*, Vol 155, p 110657
21. Movassagh AA, Alzubi JA, Gheisari M, Rahimi M, Mohan S, Abbas AA, Nabipour N (2023a) Artificial neural networks training algorithm integrating invasive weed optimization with differential evolutionary model. *J Amb Intell Human Comput*, Vol 14 pp 1–9
22. Musthafa MM, Manimozhi I, Mahesh TR, Guluwadi S (2024) Optimizing double-layered convolutional neural networks for efficient lung cancer classification through hyperparameter optimization and advanced image pre-processing techniques. *BMC Med Inform Decis Making* 24:142
23. Nahiduzzaman Md, Faisal Abdulrazak L, Arselene Ayari M, Khandakar A, Islam SMR (2024) A novel framework for lung cancer classification using lightweight convolutional neural networks and Ridge Extreme Learning Machine model with Shapley additive explanations (SHAP). *Expert Syst Appl* 248:123392
24. Pingale SV, Sutar SR (2023) Remora-based deep Maxout network model for network intrusion detection using convolutional neural network features. *Comput Electr Eng* 110:108831
25. Prasad U, Chakravarty S, Mahto G (2023) Lung cancer detection and classification using deep neural network based on hybrid metaheuristic algorithm. *Soft Comput*, Vol 28 pp1–24.
26. Raza R, Zulfiqar F, Khan MO, Arif M, Alvi A, Iftikhar MA, Alam T (2023) Lung-EffNet: lung cancer classification using EfficientNet from CT-scan images. *Eng Appl Artif Intell* 126:106902
27. Rezaei SR, Ahmadi A (2022b) A hierarchical GAN method with ensemble CNN for accurate nodule detection. *Int J Comput Assist Radiol Surg* 18:695–705
28. Saikia T, Kumar R, Kumar D, Singh KK (2022a) An automatic lung nodule classification system based on Hybrid Transfer Learning Approach. *SN Comput Sci* 3:272
29. Saleem MA, Le Thien N, Asdornwised W, Chaitusaney S, Javeed A, Benjapolakul W (2023) Sooty tern optimization algorithm-based deep learning model for diagnosing NSCLC TUMOURS. *Sensors* 23:2147
30. Sampangi Rama Reddy BR, Sen S, Bhatt R, Dhanetwal ML, Sharma M, Naaz R (2024) Stacked neural nets for increased accuracy on classification of Lung Cancer. *Meas Sens* 32:101052
31. Siddiqui EA, Chaurasia V, Shandilya M (2023) Detection and classification of lung cancer computed tomography images using a novel improved deep belief network with Gabor Filters. *Chemom Intell Lab Syst* 235:104763
32. Thangavel C, Palanichamy J (2024) Effective deep learning approach for segmentation of pulmonary cancer in the thoracic CT image. *Biomed Signal Process Control* 89:105804
33. Venkatesan N, Pasupathy S, Gobinathan B (2024) An efficient lung cancer detection using optimal SVM and improved weight-based beetle swarm optimization. *Biomed Signal Process Control* 88:105373
34. Wang Q, Liu Z, Zhang T, Alasmary H, Waqas M, Halim Z, Li Y (2023) Deep convolutional cross-connected kernel mapping support vector machine based on select dropout. *Info Sci* 626:694–709
35. Wang Y, Yang G, Li S, Li Y, He L, Liu D (2023) Arrhythmia classification algorithm based on multi-head self-attention mechanism. *Biomed Signal Process Control* 79:104206
36. Wankhade SSV (2023) A novel hybrid deep learning method for early detection of lung cancer using Neural Networks. *Healthc Anal* 3:100195
37. Yalvac B, Reubens N, Reniers B (2024) Early results of a remote dosimetry audit program for lung stereotactic body radiation therapy. *Phys Imaging Radiat Oncol* 29:100544
38. Zhao Y, Wang S, Zhang Y, Qiao S, Zhang M (2023) WRANet: Wavelet integrated residual attention U-Net Network for medical image segmentation. *Complex Intell Syst* 9:6971–6983

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.