IIETA International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Ingénierie des Systèmes d'Information
Vol. 31, No. 7, July, 2026, pp. 2449-2462
Journal homepage: http://iieta.org/journals/isi

An Anatomy-Aware Hierarchical Generative Adversarial Network Framework for Synthetic Lung Computed Tomography Image Generation and Cancer Classification

S. D. Barkath Unisha[1*], N. Anandakrishnan[2]

[1] PG & Research Department of Computer Science, Providence College for Women (Autonomous), Coonoor 643103, India
[2] PG & Research Department of Computer Science, Nilgiri College of Arts and Science (Autonomous), Nilgiris 643239, India

Corresponding Author Email: nishaphd2024@gmail.com

https://doi.org/10.18280/isi.310726

Received: 5 February 2026
Revised: 3 April 2026
Accepted: 16 April 2026
Available online: 31 July 2026

Keywords:
lung cancer diagnosis, lung computed tomography imaging, Generative Adversarial Network, synthetic image generation, data augmentation, medical image analysis, deep learning

ABSTRACT

Lung cancer diagnosis from computed tomography (CT) images remains challenging due to substantial variations in nodule morphology, limited annotated samples, and the difficulty of preserving anatomical consistency during synthetic image generation. Although Generative Adversarial Networks (GANs) have been widely explored for medical image augmentation, existing approaches often introduce structural artifacts and fail to maintain realistic lung anatomy. This study proposes an anatomy-aware Hierarchical Memory-efficient Lobe Generative Adversarial Network (HMLGAN) for generating high-quality synthetic lung CT images and improving downstream cancer classification. The proposed framework integrates lobe-specific segmentation, graph-based anatomical representation learning, and hierarchical adversarial refinement to preserve both local nodule characteristics and global pulmonary structures. A modified U-Net is first employed to obtain lobe-level Regions of Interest (ROI), followed by a Graph Encoder (GE) that captures anatomical relationships among segmented regions. The generated images are subsequently incorporated into a classification pipeline based on a modified AlexNet-Support Vector Machine (SVM) model. Experimental evaluations were conducted on the International Society for Optics and Photonics- American Association of Physicists in Medicine - National Cancer Institute (SPIE-AAPM-NCI) and Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI) lung CT datasets. The proposed framework achieved classification accuracies of 98.38% and 98.11%, respectively, demonstrating consistent improvements over representative deep learning (DL) baselines. These findings indicate that anatomy-guided synthetic image generation can enhance data diversity and support more reliable CT-based lung cancer analysis.

1. INTRODUCTION

Lung cancer remains one of the deadliest forms of cancer, characterized by the unconstrained proliferation of irregular cells within the lungs. Early detection is vital because it is the leading cause of cancer-related deaths globally, particularly when lung nodules that are often the earliest indicators of the disease are detected [1]. An earlier identification of lung nodules is essential for improving lung cancer outcomes. Imaging techniques like computed tomography (CT) and MRI are pivotal for early lung cancer detection (LCD) [2]. Among these, Low-Dose CT (LDCT) is favored due to its non-invasiveness, low radiation exposure, and high resolution, making it a standard tool for early screening. To automate the image analysis process and identify minute areas, machine learning (ML) techniques are used. In reference [3], an enhanced Fuzzy C-Means clustering technique was proposed to separate the extracted CT images. The pre-processing step utilized type-2 fuzzy and after segmentation, various ML classifiers were utilized for classification purposes. In another study [4], the authors introduced a Support Vector Machine (SVM) approach for the classification of nodules utilizing CT scans, as SVM excels in reducing the dimensions of the image and reconstructing the original images. In reference [5], ensemble ML models were used for the detection of early-stage lung cancer to improve their accuracy.

Despite their advantages, ML poses significant limitations. Feature extraction relies on pre-processing techniques like noise reduction and normalization to ensure accurate as well as meaningful features. In addition, ML methods are limited to supporting only small datasets. To handle large datasets, many deep learning (DL) based LCD (DL-LCD) methods were proposed. A DL-based computer-aided diagnosis (CAD) model that employed Moth Swarm Optimization (MSO) [6] was introduced for optimizing hyperparameters of LSTM, in turn improving the model's generalization ability for detecting lung cancer. transfer learning (TL)-based EfficientNet model [7] for lung cancer classification was introduced. The authors compared five variants of pre-trained EfficientNet and found the B1 variant to be a better model by fine-tuning the base model.

The above DL based methods directly extract features from images for LCD. However, segmentation before feature extraction allows the method to concentrate on extracting

features from precise Regions of Interest (ROI) such as lung nodules within an image. The combination of both segmentation and classification leverages the strengths of each other for more comprehensive insights and improved diagnostic abilities. In the study [8], a DL fusion model integrating ResNet and BiConvLSTM U-Net was introduced for segmenting lung cancer, with ResNet-34 functioning as the encoder within the U-Net framework. In the study [9], the authors developed a Convolutional Neural Network (CNN) based two-tiered attention U-Net model using CT images for lung cancer segmentation, where they combine multi-scale features from attention modules, the encoder, and decoder.

In the study [10], the authors proposed the UNETR network for segmentation and a self-supervised network for classification purposes. An advanced LCD model by integrating Modified U-Net [11] was introduced for both lobe segmentation as well as candidate nodule extraction, and Modified AlexNet-SVM for categorization. All the above models employed U-Net for lung cancer segmentation. But U-Net models often failed to capture nodule variability due to limited datasets, which leads to errors and missed detections in classification. To mitigate these issues, Generative Adversarial Network (GAN) based augmentation models are emerging as a solution in medical imaging, generating new or adapted samples to enhance DL model robustness. A GAN-based lung segmentation [12] using CT images was proposed to simplify the segmentation process and improve performance compared to existing methods. Similarly, in the study [13], an enhanced 3D pulmonary nodule detection method was proposed using Conditional GANs, leading to improved sensitivity and specificity in detecting nodules. In the study [14], a GAN-based technique for 3D lung cancer restoration using VGGNet for feature extraction and LSTM for segmentation leading to high-resolution tumor models.

However, the above GAN models face GPU memory limitations, particularly with high-resolution images, resulting in patchy artifacts and reduced quality. This highlights the need for more effective approaches to improve U-Net segmentation models' performance and reliability in detecting lung nodules. Despite the advancements in GANs, generating high-resolution images that maintain fine details of lung tissues and accurate textures remains a significant challenge.

1.1 Main contributions

The main contributions of this study are summarized as follows:

1. To solve the above issues, a novel Hierarchical Memory-efficient Lobe Generative Adversarial Network (HMLGAN) is proposed in this paper to synthesize high-resolution lung CT images with reduced artifacts.
2. Initially, the Modified U-Net [11] module is employed to segment the lung CT images into lobe-specific ROIs. The segmented images are used for ROI volume construction, which is the process of creating a volume to represent Structural Connectivity (SC) derived from the spatial and anatomical relationships among segmented lung regions. The same segmented images are used as input to discriminators in HMLGAN to check whether generated images are real or fake.
3. A Graph Encoder (GE) is introduced in HMLGAN to model structural relationships from the ROI volume, which supports the generator in generating segmented images.
4. A Lobe-Segment Generator (LSG) is proposed within HMLGAN to produce high-quality images. The generator utilizes the variance of the GE values obtained from the ROI volume.
5. An Anatomy-Aware Discriminator (A2D) based on Connectome-CNN (CCNN) is also proposed to find the closeness between the outputs of the LSG and the real segmented images produced by the Modified U-Net [11].
6. Finally, the high-quality synthetic images from HMLGAN are split into training and testing datasets. The training set is used to train the Modified AlexNet-SVM model [11], and the testing set is used to evaluate its classification performance.

The structure of the paper is organized as follows: Section 2 examines existing approaches for lung nodule detection and categorization. Section 3 describes the HMLGAN model. Section 4 details the experimental findings, and Section 5 presents the conclusion.

2. LITERATURE SURVEY

2.1 Image generation of lung cancer

A U-Net-based model incorporating Inception-Residual Blocks (IRB), Residual Mapping (RM), a Multi-Level Joint Discriminator (MLJD), and a Combination of Losses (CL) Unified Image Restoration and Multi-task Classification Network (UIRMC) [15] was introduced for denoising LDCT images using Least Square GAN (LSGAN). An Attention-Guided GAN (AGGAN) model [16] was suggested for generating high-quality synthetic CT images from low-dose cone-beam CT (CBCT) images using a built-in attention module.

A dual domain GAN (DU-GAN) model [17] was developed for denoising LDCT images; it makes use of GAN discriminators based on U-Nets to understand the global and local distinctions between normal-dose and denoised scans in the image and gradient domains. An improved version of the pix2pix model called Stylepix2pix [18] was developed for increasing the applicability of augmentation in lung disease detection.

Diffusion models were presented in the study [19] to address class imbalance in lung cancer CT classification using synthetic image augmentation. A diffusion-based DiffLung was presented in the study [20] to improve pathological lung tissue segmentation by generating synthetic CT images for underrepresented classes.

2.2 Classification of lung cancer

To identify and categorize lung cancer from CT scans, a weighted VGG deep network (WVDN) [21] was developed, which employed pre-processing and segmentation with K-Means Clustering as well as tuning and regularization of the hyperparameters.

An automatic CAD-based system (ACADS) [22] was suggested, which utilized U-Net for semantic segmentation and candidate nodule detection. A YOLOv5-CASP model [23] with improved Atrous Spatial Pyramid Pooling (ASPP) for enhancing the detection of small-lung nodules and Contextual Transformer (CoT) for obtaining more information from the lung CT images.

A DL-LCD model [24] was presented using CT and CXR images. It involves pre-processing using Contrast Limited Adaptive Histogram Equalization (CLAHE) and segmentation

using the honey badger algorithm. An improved YOLOv5 (iYOLOv5) model [25] was introduced for lung nodule detection to minimize loss and improve detection rate. A dual-attention model that integrated Squeeze and Excitation with Vision Transformer (SE-ViT) [26] was devised to focus on relevant features for lung nodule classification using Grad-CAM, which visualizes those features for effective classification. Table 1 shows the summary of the existing literature.

2.3 Research gap

Previous methods for lung tumor analysis using advanced models have faced several challenges, including difficulties in capturing complex lung features, insufficient generalization across varied conditions, and limitations in model accuracy and scalability. To overcome these challenges, this phase develops an HMLGAN technique designed to improve image generation quality.

Table 1. Comparative analysis of existing lung cancer classification methods

Ref No.	Method	Advantages	Limitations
		Image generation of lung cancer	
[15]	UIRMC	Artifact removal and feature preservation.	The noises visually similar to the structure of input images are not effectively removed.
[16]	AGGAN	Reduces severe artifacts present in synthetic CT images while preserving anatomical structure.	Processing of foreground and background attention maps requires significant memory and power.
[17]	DU-GAN	Remove streak artifacts while preserving edge information.	The model poses high computational cost during training.
[18]	Stylepix2pix	It produces tumours with intricate patterns, which results in far more realistic images.	Occasionally, Stylepix2pix generates spatially warped textures.
[19]	Diffusion Models	Poses the ability to generate rich synthetic data with lower variance.	The iterative noise reduction process in the diffusion approach causes computational overhead.
[20]	DiffLung	It generated additional synthetic CT images for less represented classes to mitigate class imbalance.	It relies on a slow and iterative denoising process, which consumes more memory and time.
		Classification of lung cancer	
[21]	WVDN	The model averted the issues of overfitting.	The parameter and memory usage are high.
[22]	ACADS	It reduced the occurrence of false positive rates.	It cannot detect less dense nodules located inside fats
[23]	YOLOv5-CASP	High accuracy due to effective contextual feature extraction.	This model may lead to higher false positives when a nodule is occluded by surrounding nodules.
[24]	DL-LCD	It resulted in lower misclassification rates with high reliability.	The standard honey badger algorithm may suffer from premature convergence.
[25]	iYOLOv5	Increases accuracy due to the ability to derive more intricate features.	It suffers from high computational demands.
[26]	SE-ViT	Focused on detecting every nodule even when the boundaries are unclear.	The model processes patch levels, so it cannot extract fine-grained features.

Note: Unified Image Restoration and Multi-task Classification Network (UIRMC), Low-Dose CT (LDCT), Generative Adversarial Network (GAN), Attention-Guided GAN (AGGAN), dual domain GAN (DU-GAN), weighted VGG deep network (WVDN), automatic CAD-based system (ACADS), DL based Lung Cancer Detection (DL-LCD), Squeeze and excitation with Vision Transformer (SE-ViT), computed tomography (CT).

3. PROPOSED METHODOLOGY

This part delivers a thorough overview of the HMLGAN with Modified U-Net and Modified AlexNet-SVM-based Deep Network (HMUMASDNet) model. Figure 1 shows the suggested HMUMASDNet architecture. The proposed HMUMASDNet model details the various stages involved in LCD, including segmentation, generation of high-quality images, and classification.

3.1 Modified U-Net for lobe segmentation

In this paper, Lobe segmentation utilizes Modified U-Net [11] in the initial phase of lung cancer categorization. In this phase, the model is designed to accurately segment lung lobes from CT scan images. During training, the Modified U-Net learns to predict masks for the lung lobes. Then, the model applies these predicted masks to extract the lobes from the test CT scan. The architecture is built on the U-Net model comprising an encoder, decoder, and skip connections. The encoder extracts essential features, while the decoder reconstructs the segmented image. The resultant outcome is a precise segmentation of lung lobes, enabling subsequent nodule detection and cancer classification.

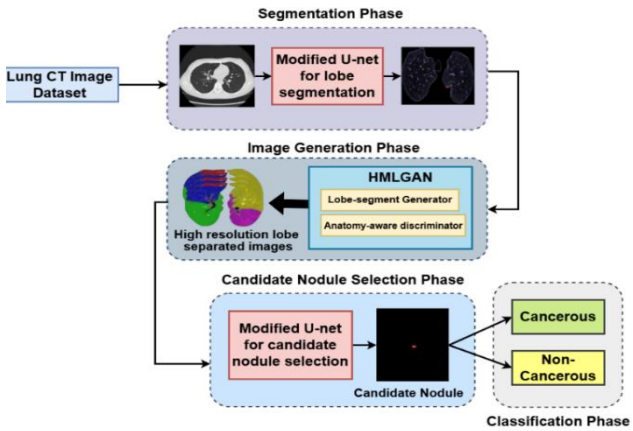

Figure 1. Structure of the proposed Modified U-Net and Modified AlexNet-SVM-based Deep Network (HMUMASDNet) technique

3.2 Hierarchical Memory-efficient Lobe Generative Adversarial Network

The HMLGAN model leverages a sophisticated architecture for generating a high-resolution lung CT image by capturing both local and global morphological features effectively.

It improves image generation quality and allows for the

creation of explainable anatomical variations by navigating specific paths in latent space.

3.2.1 Regions of Interest volume construction

The structural network of the lungs can be modelled as a graph in which the edges represent the SC derived from the spatial and anatomical relationships among segmented lung regions. The segmented lung CT images obtained from the Modified U-Net are divided into lobe-specific ROIs, where each ROI is treated as a graph node. Thus, the proposed GE approximates anatomical SC using ROI-based representations.

The process begins with estimating the initial SC of the lung, which captures the overall connectivity pattern among different ROIs. To incorporate regional anatomical variation, the total ROI volume was computed by the summation of the ROI volume of each image (i) with its predicted volume change, averaged across all ROIs as in Eq. (1):

$$V_{ROI_{total}} = \frac{1}{N}\sum_{i=1}^{N} \Delta V_{ROI}(i) + V_{ROI}(i) \tag{1}$$

where, $V_{ROI}(i)$ denotes the original volume of ROI of each image, N denotes the number of ROIs. Each node feature vector $F \in \mathbb{R}^{N\times(N+2)}$ is constructed from two components: the first N columns represent ROI-wise segmented feature representations, while the final two columns represent the disease label. In this modified approach, the segmented lung images serve as the input to the model. The matrix A represents the ROI-based SC derived from spatial relationships among segmented ROI regions and forms the edges of the graph.

3.2.2 Graph Encoder

The core of the model is a GCN, which updates each node's features based on a weighted combination of its neighboring nodes' features, allowing the network to capture anatomically meaningful relationships between distant lung regions. The graph data, consisting of F and A, then passes through two GCN layers with d hidden neurons that extract relevant structural information from the segmented images. Two additional one-layer GCNs compute the mean $\mu \in \mathbb{R}^{N\times d}$ and standard deviation $\sigma \in \mathbb{R}^{N\times d}$, forming the latent representation of each node. The input of GCN is the

$$H^{l}=G(F,A) \tag{2}$$

In Eq. (2), F is the feature of node $V_{ROI}(i)$, A is adjacency or edge matrix consists of relation between all V_{ROI} nodes. The GCN operation is defined as:

$$H^{l+1} = act(SH^{l}W^{l}) \tag{3}$$

In Eq. (3), the symmetrically normalized adjacency matrix $S = \widehat{D}^{-\frac{1}{2}}\hat{A}\widehat{D}^{-\frac{1}{2}}$, $\widehat{D}_{ii} = \sum_{j=1}^{N}\hat{A}_{ij}$, $\hat{A} = A + I$, act is the activation symbol (LeakyReLU). The activation function updates the node features, with W^{l} representing the layer weights and H^{l} the node features at layer l. H_0 is the initial node feature F. $\widehat{D}$ is a diagonal matrix with each diagonal element being the sum of the corresponding row in matrix $\hat{A}$ and I is the identity matrix. After two GCN layers, the network produces latent variables μ and σ through separate GCNs, as in Eq. (4):

$$\mu = act\left(SH^{2}W_{2\mu}\right); \sigma = act(SH^{2}W_{2\sigma}) \tag{4}$$

The latent representation is then re-parameterized to ensure it follows a Gaussian distribution, as in Eq. (5):

$$h' = \mu + \sigma \odot \epsilon \tag{5}$$

In Eq. (5), $\epsilon \in \mathcal{N}(0,1)$ is sampled from a Gaussian distribution and $h' \in \mathbb{R}^{N\times d}$. This latent representation h' generated from the ROI images encapsulates the SC and anatomical structure of the lungs. The latent space is used to generate high-resolution lung images in subsequent processes, effectively synthesizing detailed anatomical features based on the segmented input, bridging the gap between segmented image data and the SC.

3.2.3 Lobe-Segment Generator

Figure 2 depicts the architecture of the LSG, which utilizes the latent representation from the GE to generate high-resolution lung CT images. The generator architecture employs a CSF strategy to find local topological features of each lobe. As mentioned earlier, the latent representation h' from the GE serves as the input to the generator. The latent representation h' produced by the GE serves as the input to a Multi-Layer Perceptron (MLP). The MLP processes h' to generate a refined feature representation F' which is used to estimate the SC. This is vital for accepting the spatial and functional relationships within the lung structure. The initial SC is estimated using Eq. (6) and computed using Eq. (7).

$$F' = MLP(h' \parallel v) \tag{6}$$

$$L_0 = act\left(F' \otimes F'^{T}\right); act(x) = \frac{1}{1+e^{x}} \tag{7}$$

In Eq. (6), $F' \otimes F'^{T}$ captures pairwise relationships between ROI features, when the connectivity values are normalized between 0 and 1 by the sigmoid activation function act(x). This matrix L_0 reflects the foundational structural connections within the lung, which are further refined to enhance the accuracy of the synthetic images. The lobe segmentation is applied on L_0 and obtain L_1, an intermediate SC matrix. Various parts of L_1 include Right Upper Lobe Connectivities (L_{ru1}), Right Middle Lobe Connectivities (L_{rm1}), Right Lower Lobe Connectivities (L_{rl1}), Left Upper Lobe Connectivities (L_{lu1}) and Left Lower Lobe Connectivities (L_{ll1}) are represented in Eq. (8):

$$\begin{aligned} L_{ru1} &= right\ upper(L_0); \\ L_{rm1} &= right\ middle(L_0); \\ L_{rl1} &= right\ lower(L_0); \\ L_{lu1} &= left\ upper(L_0); \\ L_{ll1} &= left\ lower(L_0); \end{aligned} \tag{8}$$

The individual lobe connectivities are updated using the CSF strategy to refine local connectivity features. The updated lobe connectivities are then combined to adjust global topological features to ensure the generated SC uses the same CSF strategy. Each sub-matrix is processed by the CSF strategy and is regions of L'_1 which is represented in Eq. (9):

$$\begin{aligned} &L'_{ru1} = CSF(L_{ru1}); L'_{rm1} = CSF(L_{rm1}); L'_{rl1} = \\ &CSF(L_{rl1}); L'_{lu1} = CSF(L_{lu1}); L'_{ll1} = CSF(L_{ll1}) \end{aligned} \tag{9}$$

The processed sub-matrices are combined to form the rough matrix L_2 is defined in Eq. (10):

$$L_2 = unity\left(L'_{ru1}, L'_{rm1}, L'_{rl1}, L'_{lu1}, L'_{ll1}\right) \quad (10)$$

This matrix is then processed by the CSF Strategy to capture global topological features and is defined in Eq. (11):

$$L_3 = CSF(L_2) \quad (11)$$

The output $L_3 \in R^{N\times N}$ is the final generated SC and gives the generated images Z.

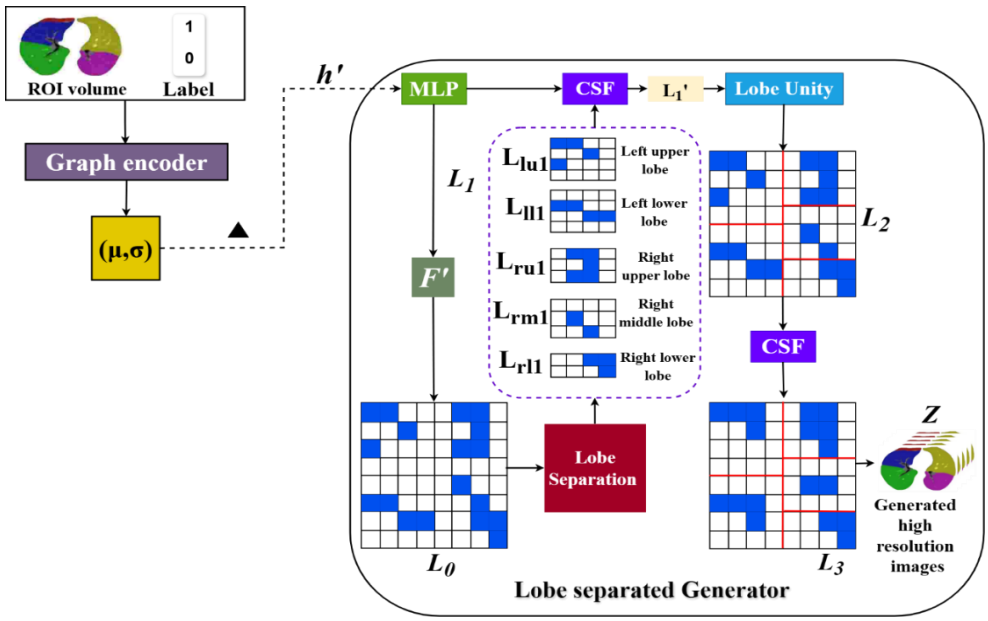

Figure 2. Lobe-Segment Generator (LSG)

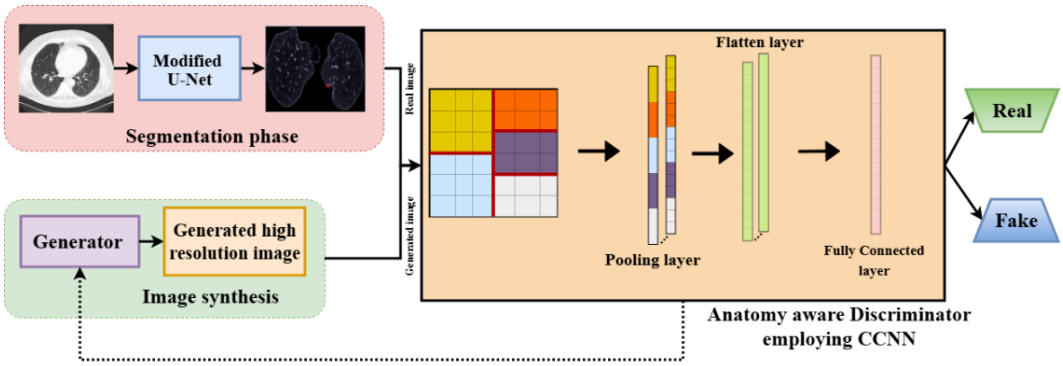

Figure 3. Outline of Generative Adversarial Network (GAN) employing an Anatomy-Aware Discriminator (A2D)

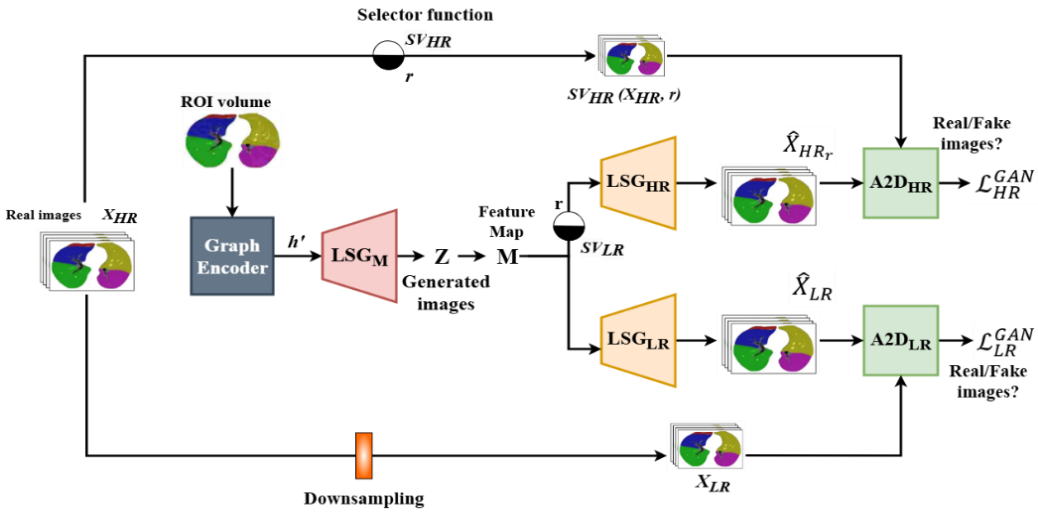

Figure 4. Architecture and workflow of Hierarchical Memory-efficient Lobe Generative Adversarial Network (HMLGAN) model

3.2.4 Anatomy-Aware Discriminator

Figure 3 demonstrates the process of A2D. Once the high-resolution images are generated, these are then fed into the A2D along with the original segmented images. The A2D then evaluates these segmented images against the generated high-resolution images and compares the anatomical and structural characteristics. A2D leverages a CCNN architecture to evaluate the anatomical plausibility of the generated high-resolution images by comparing them against the original segmented inputs. Unlike standard discriminators that rely on square convolutional filters, the CCNN adopts the specialized convolution strategy introduced by Meszlényi et al. [27].

Applying line-wise (1 × 499) and column-wise (499 × 1) filters to capture anatomy-driven structural dependencies. This architecture enables the discriminator to model region-wise connectivity patterns, extract high-dimensional structural fingerprints through successive convolutional layers (64 filters in the first layer, 128 in the second), and project them into a fully connected representation for robust discrimination.

3.2.5 Hierarchical structure

HMLGAN has two branches: one for generating low-quality images and the other for high-quality images. The low-resolution image is generated by branch LSG_{LR}, while the high-resolution image is produced by branch LSG_{HR}. In this process, LSG_M is the shared initial layers of the generator, which process the latent representation h' before it is passed to the respective branches for low-resolution and high-resolution image generation. These layers serve as a foundational block that refines the latent features before further processing by the LSG_{LR} block and the LSG_{HR} block. Figure 4 shows the hierarchical structure of the proposed HMLGAN. The low-resolution image $\hat{X}_{LR}$ is defined in Eq. (12):

$$\hat{X}_{LR} = LSG_{LR}\left(\underbrace{LSG_M(h')}_{\hat{M}}\right) \tag{12}$$

For the high-resolution image:

$$\hat{X}_{HR_r} = LSG_{HR}\left(\underbrace{SV_{LR}(LSG_M(h');r)}_{\hat{M}_r}\right) \tag{13}$$

In Eq. (13), SV_{LR} is a selector function that extracts a partial volume from the low-resolution image from slice r. Additionally, the model incorporates a High-Resolution A2D ($A2D_{HR}$) which evaluates the realism of high-resolution sub-volumes $\hat{X}_{HR}$, focusing on the local details. If any sub-volume fails the realism check, it is sent back to LSG_M for further refinement.

The loss function of $A2D_{HR}$ is derived as:

$$\begin{aligned}&\mathcal{L}_{HR}^{GAN}(LSG_M, LSG_{HR}, A2D_{HR}) = \\ &\min_{LSG_{HR},LSG_M}\max_{A2D_{HR}} \mathbb{E}_{r\sim U}[\mathbb{E}_{X\sim P_X}[\log A2D_{HR}(SV_{HR}(X_{HR};r),r)] \\ &+\mathbb{E}_{Z\sim P_Z}\left[\log(1 - A2D_{HR}(\hat{X}_{HR_r},r)\right]\end{aligned} \tag{14}$$

Similarly, the Low-Resolution A2D ($A2D_{LR}$) ensures that the global structure of the generated low-resolution image $\hat{X}_{LR}$ is maintained. If the global structure is incorrect, the image is sent back for refinement. The loss function of $A2D_{LR}$ is explained as:

$$\begin{aligned}&\mathcal{L}_{LR}^{GAN}(LSG_{LR}, LSG_M, A2D_{LR}) \\ &= \min_{LSG_{LR},LSG_M}\max_{A2D_{LR}} \mathbb{E}_{X\sim P_X}[log\, A2D_{LR}(X_{LR})] \\ &+ \mathbb{E}_{Z\sim P_Z}\left[log(1 - A2D_{LR}(\hat{X}_{LR})\right]\end{aligned} \tag{15}$$

In both Eqs. (14) and (15), the generator LSG_M iteratively refines the fake images until they successfully pass the discriminators' checks for realism and structural accuracy. The working principle of the HMLGAN model is detailed through a step-by-step procedure outlined in Algorithm 1.

Algorithm 1. HMLGAN model for high-resolution lung CT image synthesis

Input: Low-resolution lung CT image X_{LR}, Number of epochs

Output: High-resolution lung CT image X_{HR}

1. **Begin**
2. Initialize HMLGAN model parameters;
3. Initialize epoch = 0;
4. **while** (epoch < Maximum number of epochs) **do**

a) Generate the initial low-resolution image $\hat{X}_{LR}$ using the LSG as defined by Eq. (12)

b) **for** each sub-volume r in X_{LR} do:

i) Extract sub-volume $SV_{LR}(LSG_M(h');r)$;

ii) Generate high-resolution sub-volume $\hat{X}_{HR_r}$ from $SV_{LR}(LSG_M(h');r)$ using Eq. (13)

c) End for

d) Compute low-resolution adversarial loss $\mathcal{L}_{LR}^{GAN}$;

e) Update the weights of the low-resolution discriminator $A2D_{LR}$;

f) Update the weights of low-resolution generator LSG_{LR};

g) Compute high-resolution adversarial loss $\mathcal{L}_{HR}^{GAN}$;

h) Update the weights of the high-resolution discriminator $A2D_{HR}$;

i) Update the weights of high-resolution generator LSG_{HR};

j) Repeat the steps until the maximum number of epochs is reached;

5. **End while**
6. Return synthesized high-resolution lung CT image X_{HR};
7. **End**

3.3 Modified AlexNet-SVM for classification

Finally, the segmented images are input into the Modified AlexNet-SVM architecture for classification [11]. In this phase, the model classifies cancerous and non-cancerous patches extracted from CT scan slices. The model utilizes eight convolutional layers to abstract attributes, three max-pooling layers for dimensionality reduction, and the flattened feature maps are then processed by a sigmoid activation function. This integrated framework is termed the HMUMASDNet, highlighting its combination of HMLGAN for image synthesis, Modified U-Net for segmentation, and Modified AlexNet-SVM for classification.

4. EXPERIMENTAL RESULTS

This part evaluates the efficiency of both the HMLGAN and HMUMASDNet models on two different datasets described in Sections 4.3.1 and 4.3.2. The performance of the HMLGAN approach is evaluated by comparing it to the standard models like GAN, Deep Convolutional GAN (DCGAN), CycleGAN, and Conditional GAN (CGAN). On the other hand, the

performance of the HMUMASDNet technique is compared to the pre-trained models like ResNet50, InceptionV3, and InceptionResNetV2. In addition, a comparative evaluation is conducted to prove the efficacy of the proposed models against state-of-the-art models in terms of respective performance metrics across various benchmark datasets.

4.1 Dataset description

International Society for Optics and Photonics (SPIE)-American Association of Physicists in Medicine (AAPM)-National Cancer Institute (NCI) Lung Nodule Classification Challenge [28]: It was developed for a collaborative challenge organized by the SPIE, the AAPM, and the NCI. This dataset consists of 70 thoracic CT scans, with 22,489 images collected from patients screened for lung cancer. The dataset is segmented into two components: a calibration set and a test set. The calibration set includes 10 scans, with 5 scans containing malignant nodules (1721 images) and the other 5 containing benign nodules (1684 images). The test set comprises 60 scans. Within the 60 scans, 30 scans are benign (9398 images), and 29 are malignant (9281 images). One scan (with 405 images), which contains both benign and malignant nodules, is excluded from this study to avoid ambiguity. For experimental purposes, 2400 benign and 2400 malignant images are selected from the test set for testing. The remaining images, 6998 benign and 6881 malignant from the test set and calibration sets, are used for training (8682 benign and 8602 malignant). To enhance the diversity of the training data, image augmentation is performed using the HMLGAN. After augmentation, the number of training samples is increased to 9600 benign and 9600 malignant images. The split ratio of samples for the above training and testing is 80:20.

Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI): This dataset is publicly accessible [29]. It comprises 244,527 spiral CT images of the chest from 1,018 sets of 1,010 different patients (250-300 slices per patient). These images are annotated by four radiologists. Totally 13,806 CT slices (11,935 benign and 1,871 malignant) are selected for experimental purposes. The 393 benign and 374 malignant images are used for testing. The remaining images are augmented by HMLGAN to overcome challenges such as overfitting during model training. As a result, a balanced dataset of 22,792 CT images was obtained, with 11,935 benign and 10,857 malignant slices. An 80:20 split for training and validation. The 80:20 split at the slice level is processed with awareness of the patient level. This approach guarantees that there are no image slices from the same patient in the training, test set, or validation set to avoid potential leaks in the data and to obtain an unbiased assessment of the model performance.

4.2 Parameter settings

Table 2 outlines the parameters utilized for the simulation and evaluation of the performance of both the current and suggested approach. A grid search methodology was utilized to optimize the hyperparameters of the model. To ensure fair experimental comparison, all baseline models were trained using the same parameter settings under identical experimental conditions.

Table 2. Simulation parameters

Parameter	Range
GAN, DCGAN, CycleGAN, CGAN, HMLGAN	
Input Image size	512 × 512
Learning rate for generator	1×10^{-4}
Learning rate for discriminator	4×10^{-4}
Batch size	4
Epochs	200
Loss function	Cross-entropy
Adam Optimizer	$\beta_1 = 0, \beta_2 = 0.99$
Proposed HMLGAN	
Sub-volume	32 × 256
Size of the X_{LR}	64 × 256
ResNet50, InceptionV3, InceptionResNetV2 and proposed HMUMASDNet	
Input Image size	512 × 512
Batch size	32
Adam Optimizer	$\beta_1 = 0, \beta_2 = 0.99$
Learning rate	0.0001
Epochs	200
Activation function	ReLU
Loss function	Cross-entropy

Note: Generative Adversarial Network (GAN), Deep Convolutional GAN (DCGAN), Attention-Guided GAN (AGGAN), Hierarchical Memory-efficient Lobe Generative Adversarial Network (HMLGAN), Modified U-Net and Modified AlexNet-SVM based Deep Network (HMUMASDNet).

4.3 Performance metrics

The efficiency of both the developed and existing models is evaluated using the following metrics.

4.3.1 For image generation

Peak Signal-to-Noise Ratio (PSNR): PSNR evaluates the ratio of a signal's peak power (original image) to the noise that distorts its representation (synthesized or compressed image). The formula for PSNR is defined in Eq. (16):

$$PSNR = 10.\log_{10}\left(\frac{MAX^2}{MSE}\right) \tag{16}$$

Structural Similarity Index Measure (SSIM): SSIM evaluates image similarity through the analysis of structural detail, luminance, as well as contrast variations. The formula for SSIM is defined in Eq. (15):

$$SSIM = \frac{(2\mu_r\mu_g + C_1)(2\sigma_{rg} + C_2)}{(\mu_r^2 + \mu_g^2 + C_1)(\sigma_r^2 + \sigma_g^2 + C_2)} \tag{17}$$

where, μ_r and μ_g are the mean feature vectors, σ_r^2 and σ_g^2 are the variances, σ_{rg} is the covariance between real and generated image. C_1 and C_2 are constants that stabilize the division with a weak denominator.

Fréchet Inception Distance (FID): It computes the difference between the real image dataset and the produced image dataset's multivariate Gaussian distributions.

$$FID = \|\mu_r - \mu_g\|^2 + Tr(\sum_r + \sum_g - 2(\sum_r \sum_g)^{1/2}) \tag{18}$$

where, $\sum_r$ and $\sum_g$ are the covariance matrices for real and generated images, respectively. Tr refers to the trace of the matrix (sum of the diagonal elements).

Learned Perceptual Image Patch Similarity (LPIPS): It evaluates the perceptual similarity between generated and real

CT images. Lower LPIPS values indicate that the evaluated and reference images are highly similar in human perception.

$$d(x_r, x_g) = \sum_l \frac{1}{ht_l \times wt_l} \sum_{ht,wt} \left\| w_l \odot \left(\Phi_l(x_r) - \Phi_l(x_g) \right) \right\|_2^2 \quad (19)$$

where, $\Phi_l(x_r)$ and $\Phi_l(x_g)$ are the feature activations of the network for real and generated image. ht_l and wt_l are the spatial dimensions of the feature maps at layer l. w_l is a learned channel-wise weight vector. $\odot$ represents the Hadamard product.

4.3.2 For classification

Accuracy: The ratio of accurately classified malignant and benign cases to the total lung cancer images or patient cases. It is measured by

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (20)$$

In Eq. (20), TP represents True Positive, TN represents True Negative, FP represents False Positive, and FN represents False Negative.

Precision: The proportion of correctly identified malignant cases among all cases classified as malignant by the model. It is measured by Eq. (21):

$$Precision = \frac{TP}{TP + FP} \quad (21)$$

Recall: The percentage of true malignant cases accurately detected by the model. It is measured by Eq. (22):

$$Recall = \frac{TP}{TP + FN} \quad (22)$$

F1-score: The F1-score is defined as the harmonic mean of Precision and Recall, effectively balancing both measures to provide a comprehensive measure of performance. The determination is made according to Eq. (23).

$$F1 - score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (23)$$

Training time: The total duration required for training the model on the specified dataset is indicated.

Inference time: This time denotes the duration required to process a single CT image during inference.

Memory usage: The amount of RAM used during the training of the model.

4.4 Performance evaluation

4.4.1 Comparison of lung image synthesis models

Figures 5(a) and 5(b) illustrate the comparison of the suggested HMLGAN method against GAN, DCGAN, CycleGAN, and CGAN, focusing on PSNR, FID, SSIM, and LPIPS values. The proposed model exhibits a PSNR improvement of 195.60%, 150.68%, 106%, and 26.74% over GAN, DCGAN, CycleGAN, and CGAN models, respectively, in the SPIE-AAPM-NCI dataset. In terms of SSIM, the proposed HMLGAN achieves an increase of 67.80%, 59.68%, 26.92%, and 16.47% compared with the existing methods. Furthermore, the proposed HMLGAN model achieves a reduction in LPIPS of 61.68%, 55.91%, 42.66%, and 32.23% over GAN, DCGAN, CycleGAN, and CGAN models, respectively. Similarly, the FID values are reduced by 73.98%, 67.97%, 57.80%, and 44.21% compared with the baseline methods.

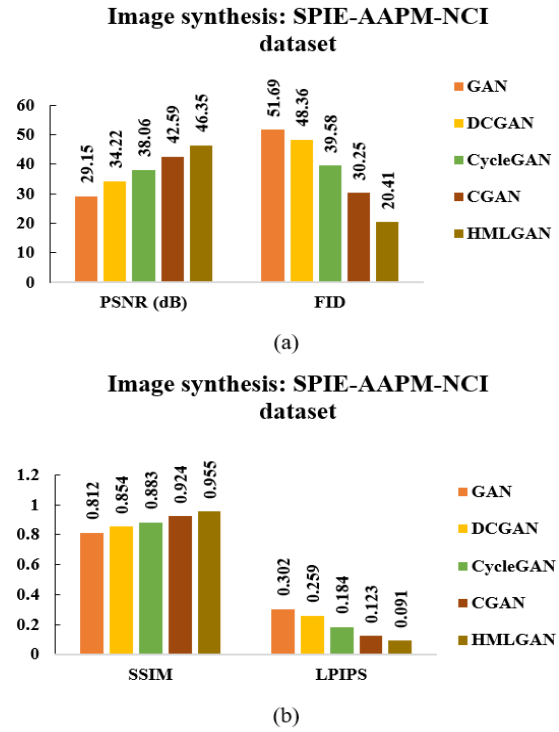

Figure 5. Comparison of the proposed model with existing image synthesis models on the SPIE-AAPM-NCI dataset
Note: International Society for Optics and Photonics (SPIE)- American Association of Physicists in Medicine (AAPM)- National Cancer Institute (NCI).

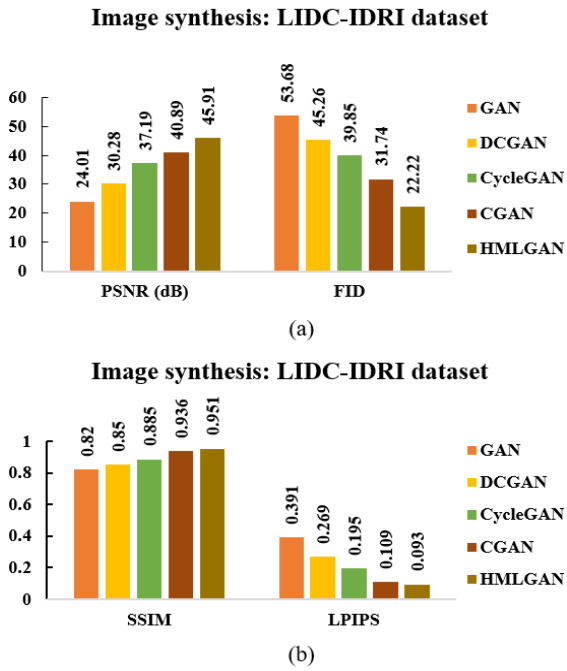

Figure 6. Comparison of proposed with existing image synthesis models on the LIDC-IDRI dataset
Note: Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI).

Figures 6(a) and 6(b) present the comparative analysis of the proposed HMLGAN framework using the LIDC-IDRI dataset. It achieves a PSNR improvement of 183.45%, 122.85%, 64.56%, and 21.86% over GAN, DCGAN, CycleGAN, and CGAN models. In terms of SSIM, the proposed HMLGAN method achieves an increase of 54.69%, 41.43%, 25.32%, and 11.86% compared with the existing methods. The proposed HMLGAN achieves a reduction in LPIPS of 63.18%, 57.47%, 45.59%, and 32.11% over the baseline models, respectively. The FID values are reduced by 73.78%, 67.61%, 56.85%, and 40.39% compared with the baseline approaches.

4.4.2 Comparison of lung disease classification

The experiments were conducted across five independent folds for both SPIE-AAPM-NCI and LIDC-IDRI datasets, and the corresponding fold-wise classification accuracies are presented in Table 3. The obtained results demonstrate consistent performance of the proposed HMUMASDNet model across all folds with minimal performance variation.

The confusion matrices for the testing set of SPIE-AAPM-NCI and LIDC-IDRI datasets using the proposed model are illustrated in Figures 7 and 8. Figure 9 demonstrates the performance evaluation of the proposed HMUMASDNet method against pre-trained models like ResNet50, InceptionV3, and InceptionResNetV2, focusing on classification metrics. In the SPIE-AAPM-NCI dataset, the proposed model significantly outperforms the pre-trained models. Specifically, it achieves an accurate improvement of 5.06%, 3.58%, and 3.21% over other models. For precision, it surpasses 7.38%, 5.27%, and 3.60% over others. Regarding recall, it is demonstrated by 5.85%, 5.18%, and 3.19% over other models. Finally, the F1-score improves by 9.06%, 7.61%, and 5.08% over other models. These results highlight that the proposed model consistently outperforms pre-trained models, demonstrating superior classification performance for LCD.

Figure 10 illustrates the evaluation of the suggested HMUMASDNet technique against pre-trained models like ResNet50, InceptionV3, and InceptionResNetV2, focusing on classification metrics in the LIDC-IDRI dataset. In the LIDC-IDRI dataset, the proposed model significantly outperforms the pre-trained models. Specifically, it achieves an accuracy improvement of 3.18%, 2.94%, and 1.42% over other models. In terms of precision, it outperforms by 8.72%, 6.85%, and 4.60% over others. Regarding recall, it shows an increase of 7.53%, 6.24%, and 5.21% over other models. Finally, for the F1-score, it improves by 7.38%, 5.12%, and 4.64% over others. These results highlight HMUMASDNet's effectiveness in achieving high classification performance for detecting lung cancer in the LIDC-IDRI dataset.

Table 3. Repeated experimental accuracy results across five folds for classification performance (%)

Dataset	Model	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Mean Accuracy
SPIE-AAPM-NCI	ResNet50	92.88	92.97	93.85	93.91	94.19	93.56
	InceptionV3	94.48	94.59	95.26	95.69	96.03	95.21
	InceptionResNetV2	94.58	94.76	95.39	95.84	96.48	95.41
	HMUMASDNet	98.57	98.69	98.92	98.52	98.45	98.63
LIDC-IDRI	ResNet50	93.97	94.58	94.85	95.44	95.81	94.93
	InceptionV3	94.35	94.69	94.98	95.35	96.23	95.12
	InceptionResNetV2	95.96	96.27	96.55	96.77	97.15	96.54
	HMUMASDNet	97.98	98.06	98.12	98.19	98.20	98.11

Note: International Society for Optics and Photonics (SPIE)- American Association of Physicists in Medicine (AAPM)- National Cancer Institute (NCI), Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI).

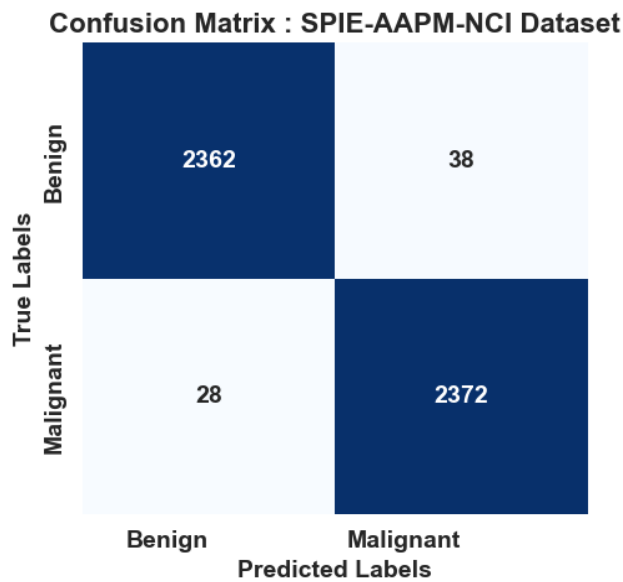

Figure 7. Confusion matrix of SPIE-AAPM-NCI dataset
Note: International Society for Optics and Photonics (SPIE)- American Association of Physicists in Medicine (AAPM)- National Cancer Institute (NCI).

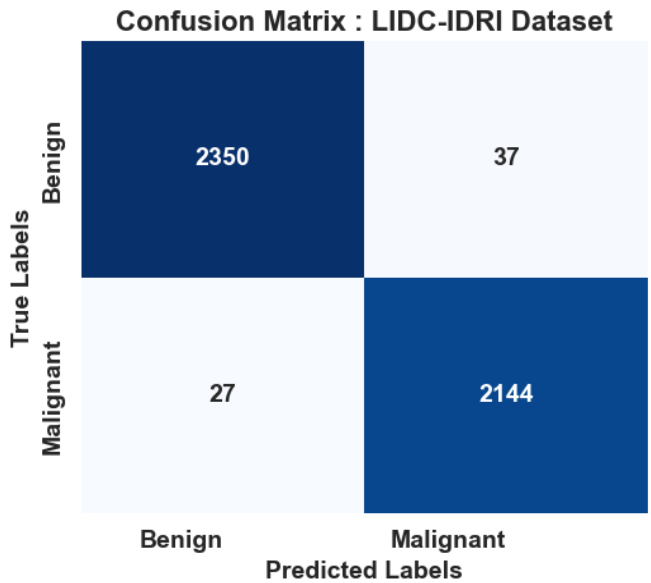

Figure 8. Confusion Matrix of LIDC-IDRI dataset
Note: Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI).

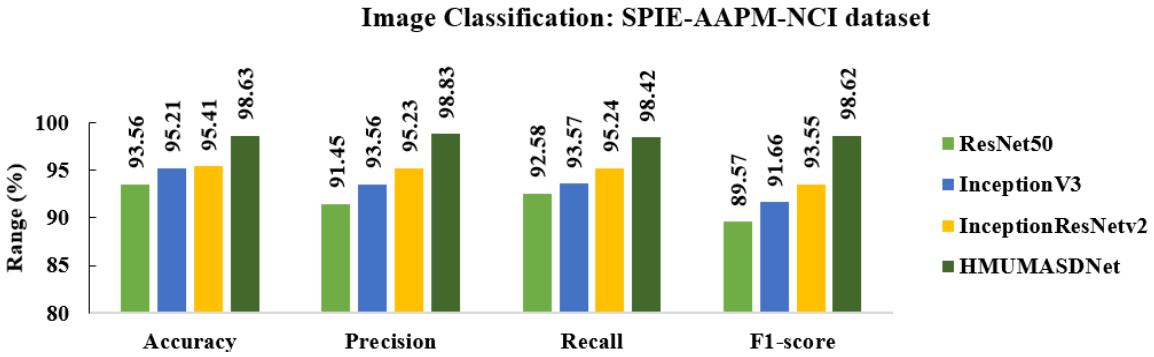

Figure 9. Comparison of proposed model with pre-trained classification models on SPIE-AAPM-NCI dataset
Note: International Society for Optics and Photonics (SPIE)- American Association of Physicists in Medicine (AAPM)- National Cancer Institute (NCI).

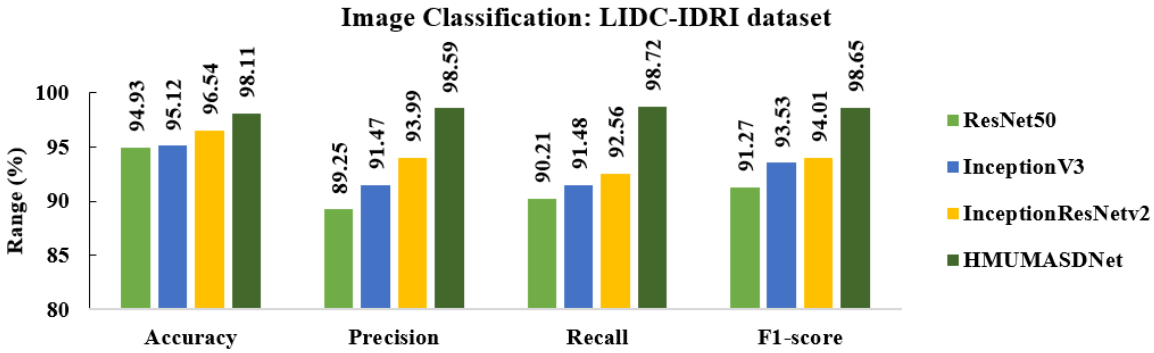

Figure 10. Comparison of proposed model with existing classification models on LIDC-IDRI dataset
Note: Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI).

Table 4. Computational efficiency of the proposed HMUMASDNet using both datasets

Datasets	Models	Training Time (sec)		Inference Time (sec) (for one instance)		Memory Usage (MB)		Training / Inference		Inference
		CPU	GPU	CPU	GPU	CPU	GPU	Parameters (M)	Model Size (MB)	GFLOPs
SPIE-AAPM-NCI dataset	ResNet50	7800	2536	4.8	2.6	5400	2126	25.6	98	4.1
	Inceptionv3	7015	2059	4.9	2.1	6238	2455	23.9	91	5.7
	InceptionResNetv2	6589	1789	3.2	1.3	7359	5445	55.9	215	13.2
	HMUMASDNet	5025	1584	3.9	1.4	8400	5630	19.4	78	4.5
LIDC-IDRI dataset	ResNet50	6214	590	6.9	3.6	5212	3231	25.6	98	4.1
	Inceptionv3	5412	454	5.5	2.8	5226	3154	23.9	91	5.7
	InceptionResNetv2	4513	657	4.2	1.5	5451	2362	55.9	215	13.2
	HMUMASDNet	3684	526	3.8	0.8	4840	1950	18.4	71	3.5

Note: International Society for Optics and Photonics (SPIE)- American Association of Physicists in Medicine (AAPM)- National Cancer Institute (NCI), Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI), Modified U-Net and Modified AlexNet-SVM based Deep Network (HMUMASDNet).

Table 5. Comparison of HMUMASDNet with and without hyperparameter optimization

Test Condition	Metrics	SPIE-AAPM-NCI	LIDC-IDRI
Without hyperparameter optimization	Accuracy (%)	97.25	97.37
	Precision (%)	96.25	97.19
	Recall (%)	97.59	97.44
	F1-score (%)	96.10	96.15
With hyperparameter optimization (random search)	Accuracy (%)	98.63	98.11
	Precision (%)	98.83	98.59
	Recall (%)	98.43	98.72
	F1-score (%)	98.62	98.65

Note: International Society for Optics and Photonics (SPIE)- American Association of Physicists in Medicine (AAPM)- National Cancer Institute (NCI), Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI), Modified U-Net and Modified AlexNet-SVM based Deep Network (HMUMASDNet).

4.4.3 Performance evaluation of classification using various configurations

Computational metrics are essential for assessing the efficacy of DL models; these include performance measures like recall, accuracy, precision, and F1-score. These include training time, inference time, and memory usage, which

require insights into the model's scalability and feasibility for real-world applications. This helps to determine the trade-offs among performance and resource consumption and enables a comparison of computational efficiency among different hardware environments and datasets. Table 4 presents the computational efficiency of the proposed HMUMASDNet approach across two datasets. The training time, inference time, and memory usage were assessed independently for both CPU and GPU environments. Also, it compares parameter count, FLOPs, and model size for the same models. The findings indicate that the speed of training and inference is significantly enhanced on the GPU in comparison to the CPU. Memory utilization improves in correlation with the size of the dataset, where the SPIE-AAPM-NCI dataset necessitates a greater amount of memory compared to the LIDC-IDRI dataset, which requires less memory. In addition, the computational complexity metrics demonstrate that the proposed HMUMASDNet model achieves reduced parameter complexity, lower FLOPs, and smaller model size compared with deeper pre-trained architectures. Table 5 provides the efficiency of HMUMASDNet, comparing results with and without the tuning of model hyperparameters. Prior to the hyperparameter tuning process, the model attained accuracy of 97.25% on the SPIE-AAPM-NCI dataset and 97.37% on the LIDC-IDRI dataset. Following the hyperparameter optimization conducted via grid search, the resulting accuracies were enhanced to 98.62% and 98.11%, respectively. Significant enhancements in precision, recall, and F1-score were observed across all datasets, indicating that optimal hyperparameters improved the model's ability to accurately classify lung cancer.

Table 6 evaluates the contribution of different augmentation strategies for lung cancer classification performance on the SPIE-AAPM-NCI and LIDC-IDRI datasets, respectively. The experiments compare classification performance using only real images, traditional augmentation techniques, GAN-based augmentation methods, and the proposed HMLGAN-based augmentation framework.

For traditional augmentation, several geometric transformations were applied, including random rotation within ±15° to ±30°, scaling from 0.8× to 1.2×, horizontal and vertical flipping with a probability of 0.5, and random cropping ranging from 70% to 100% of the original image size. The results indicate that the proposed HMLGAN-based augmentation consistently achieves the highest accuracy, precision, recall, and F1-score across both datasets.

4.4.4 Statistical validation

To validate the robustness and reliability of the proposed HMUMASDNet framework, statistical significance testing is computed in this section. A paired t-test analysis was performed between the proposed HMUMASDNet model and the baseline models, including ResNet50, InceptionV3, and InceptionResNetV2. The statistical results shown in Table 7 include mean difference, standard deviation, *t*-value, confidence intervals, and Sig. 2-tailed values. The obtained *p*-values (Sig. 2-tailed) are all less than 0.05, confirming that the performance improvements achieved by HMUMASDNet are statistically significant and not due to random variation. These additional analyses strengthen the reliability and robustness of the proposed framework.

Table 6. Comparison of HMUMASDNet on different augmentation strategies for test sets of datasets

Augmentation Type	Techniques	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
SPIE-AAPM-NCI dataset					
Only real images		94.14	93.52	93.31	93.41
Traditional augmentation	Rotation, Scaling, Flipping, Cropping	96.12	95.73	95.49	95.61
GAN-based augmentation	DCGAN	97.28	97.05	96.88	96.96
	CycleGAN	97.66	97.38	97.14	97.26
	CGAN	98.02	97.79	97.58	97.68
Proposed augmentation	HMLGAN	98.63	98.83	98.42	98.62
LIDC-IDRI dataset					
Only real images		88.82	88.15	88.94	88.04
Traditional augmentation	Rotation, Scaling, Flipping, Cropping	95.76	95.42	95.08	95.25
GAN-based augmentation	DCGAN	97.02	96.84	96.57	96.70
	CycleGAN	97.38	97.16	96.95	97.05
	CGAN	97.74	97.48	97.31	97.39
Proposed augmentation	HMLGAN	98.11	98.59	98.72	98.65

Note: International Society for Optics and Photonics (SPIE)- American Association of Physicists in Medicine (AAPM)- National Cancer Institute (NCI), Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI), Hierarchical Memory-efficient Lobe Generative Adversarial Network (HMLGAN), Generative Adversarial Network (GAN), Deep Convolutional GAN (DCGAN), Modified U-Net and Modified AlexNet-SVM based Deep Network (HMUMASDNet).

Table 7. Statistical significance analysis of accuracy for HMUMASDNet using paired t-test

Dataset	Compared Model	Mean Difference	Std. Dev	T-Value	Lower CI	Upper CI	Sig. (2-Tailed)
SPIE-AAPM-NCI	ResNet50 vs HMUMASDNet	5.07	0.647	17.52	4.27	5.87	0.00006
	InceptionV3 vs HMUMASDNet	3.42	0.761	10.04	2.47	4.36	0.00055
	InceptionResNetV2 vs HMUMASDNet	3.22	0.872	8.24	2.14	4.30	0.0011
LIDC-IDRI	ResNet50 vs HMUMASDNet	3.18	0.631	11.26	2.39	3.96	0.0003
	InceptionV3 vs HMUMASDNet	2.99	0.64	10.44	2.20	3.78	0.0004
	InceptionResNetV2 vs HMUMASDNet	1.57	0.369	9.52	1.11	2.03	0.0006

Note: International Society for Optics and Photonics (SPIE)- American Association of Physicists in Medicine (AAPM)- National Cancer Institute (NCI), Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI), Modified U-Net and Modified AlexNet-SVM based Deep Network (HMUMASDNet).

Table 8. Comparative evaluation of the developed HMLGAN approach with state-of-the-art models

Ref No.	Dataset	Techniques	Performance			
			PSNR (dB)	SSIM	FID	LPIPS
[15]	SPIE-AAPM-NCI	LSGAN	36.95	0.89	52.29	0.378
	LIDC-IDRI		44.40	0.98	54.12	0.456
[16]	SPIE-AAPM-NCI	AG-GAN	27.50	0.93	50.23	0.312
	LIDC-IDRI		31.84	0.95	52.13	0.367
[17]	SPIE-AAPM-NCI	DUGAN	22.30	0.74	45.41	0.298
	LIDC-IDRI		34.61	0.91	47.76	0.314
[18]	SPIE-AAPM-NCI	Stylepix2pix	14.97	0.49	36.41	0.234
	LIDC-IDRI		19.86	0.67	38.45	0.258
Our work	SPIE-AAPM-NCI	HMLGAN	46.35	0.99	20.41	0.091
	LIDC-IDRI		49.83	0.99	22.22	0.093

All models were retrained and tested under identical settings and dataset (refer to section 4.1, 4.2 & Table 2).

Note: International Society for Optics and Photonics (SPIE)- American Association of Physicists in Medicine (AAPM)- National Cancer Institute (NCI), Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI), Least Square GAN (LSGAN), Hierarchical Memory-efficient Lobe Generative Adversarial Network (HMLGAN).

Table 9. Comparative analysis of the proposed HMUMASDNet model with conventional models

1.1.1. Ref No.	1.1.2. Dataset	1.1.3. Techniques	1.1.4. Performance			
			1.1.5. Accuracy (%)	1.1.6. Precision (%)	1.1.7. Recall (%)	1.1.8. F1-Score (%)
1.1.9. [19]	1.1.10. SPIE-AAPM-NCI	1.1.11. WVDN	1.1.12. 92.64	1.1.13. 92.18	1.1.14. 91.94	1.1.15. 92.06
1.1.16.	1.1.17. LIDC-IDRI	1.1.18.	1.1.19. 93.22	1.1.20. 93.00	1.1.21. 93.00	1.1.22. 93.00
1.1.23. [20]	1.1.24. SPIE-AAPM-NCI	1.1.25. UIRMC	1.1.26. 82.46	1.1.27. 85.38	1.1.28. 84.72	1.1.29. 85.04
1.1.30.	1.1.31. LIDC-IDRI	1.1.32.	1.1.33. 83.80	1.1.34. 89.87	1.1.35. 88.75	1.1.36. 89.31
1.1.37. [21]	1.1.38. SPIE-AAPM-NCI	1.1.39. YOLOv5-CASP	1.1.40. 74.52	1.1.41. 75.84	1.1.42. 73.16	1.1.43. 74.47
	1.1.44. LIDC-IDRI		1.1.45. 76.60	1.1.46. 77.48	1.1.47. 75.92	1.1.48. 76.69
1.1.49. [22]	1.1.50. SPIE-AAPM-NCI	1.1.51. DL-LCD	1.1.52. 96.94	1.1.53. 94.16	1.1.54. 96.42	1.1.55. 95.28
1.1.56.	1.1.57. LIDC-IDRI	1.1.58.	1.1.59. 97.88	1.1.60. 94.94	1.1.61. 97.63	1.1.62. 96.42
1.1.63. [23]	1.1.64. SPIE-AAPM-NCI	1.1.65. iYOLOv5	1.1.66. 94.82	1.1.67. 95.12	1.1.68. 93.46	1.1.69. 94.28
1.1.70.	1.1.71. LIDC-IDRI	1.1.72.	1.1.73. 95.49	1.1.74. 95.49	1.1.75. 94.00	1.1.76. 94.74
1.1.77. [24]	1.1.78. SPIE-AAPM-NCI	1.1.79. SE-ViT	1.1.80. 84.68	1.1.81. 85.72	1.1.82. 86.11	1.1.83. 85.91
1.1.84.	1.1.85. LIDC-IDRI	1.1.86.	1.1.87. 86.30	1.1.88. 87.20	1.1.89. 87.60	1.1.90. 87.20
1.1.91. Our work	1.1.92. SPIE-AAPM-NCI	1.1.93. HMUMASDNet	1.1.94. 98.62	1.1.95. 98.83	1.1.96. 98.43	1.1.97. 98.63
	1.1.98. LIDC-IDRI		1.1.99. 98.11	1.1.100. 98.59	1.1.101. 98.72	1.1.102. 98.65

All models were retrained and tested under identical settings and dataset (refer to section 4.1, 4.2 & Table 2).

Note: International Society for Optics and Photonics (SPIE)- American Association of Physicists in Medicine (AAPM)- National Cancer Institute (NCI), Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI), weighted VGG deep network (WVDN), Unified Image Restoration and Multi-task Classification Network (UIRMC), DL based Lung Cancer Detection (DL-LCD), Squeeze and excitation with Vision Transformer (SE-ViT), computed tomography (CT), Hierarchical Memory-efficient Lobe Generative Adversarial Network (HMLGAN), Modified U-Net and Modified AlexNet-SVM based Deep Network (HMUMASDNet).

4.4.5 Comparative analysis of proposed and conventional models

The comparative analysis of existing image synthesis and lung cancer classification models demonstrates the effectiveness of the proposed frameworks. Tables 8 and 9 proposed approaches with existing state-of-the-art methods for image synthesis and classification.

As shown in Table 8, LSGAN achieved PSNR values of 36.95 dB and 44.40 dB with SSIM values of 0.89 and 0.98 on the SPIE-AAPM-NCI and LIDC-IDRI datasets, respectively. AG-GAN demonstrated moderate image synthesis capability with PSNR values of 27.50 dB and 31.84 dB, while DUGAN obtained comparatively improved structural similarity on the LIDC-IDRI dataset. Stylepix2pix produced comparatively lower PSNR and SSIM values due to reduced perceptual consistency in synthesized CT images. In contrast, the proposed HMLGAN framework achieved the highest PSNR values of 46.35 dB and 49.83 dB along with SSIM values of 0.99 on both datasets, indicating improved reconstruction quality and anatomical consistency of the generated lung CT images. Similarly, Table 9 summarizes the comparative classification performance of conventional models and the

proposed HMUMASDNet framework. The WVDN model demonstrated stable classification capability with accuracy values above 92%, while UIRMC and SE-ViT showed comparatively lower performance due to increased architectural complexity and reduced robustness across datasets.

YOLOv5-CASP achieved moderate detection performance, whereas DL-LCD and iYOLOv5 demonstrated improved classification capability on the LIDC-IDRI dataset. However, the proposed HMUMASDNet framework achieved the highest overall classification performance across both datasets.
Specifically, the proposed model achieved 98.62% accuracy, 98.83% precision, 98.43% recall, and 98.63% F1-score on the SPIE-AAPM-NCI dataset. Furthermore, on the LIDC-IDRI dataset, the proposed HMUMASDNet achieved 98.11% accuracy, 98.59% precision, 98.72% recall, and 98.65% F1-score, demonstrating improved robustness and classification effectiveness for LCD.

4.5 Discussion

Computationally, the proposed HMUMASDNet has excellent features, and practically it has good deployment parameters (as per Table 4). It achieves, in particular:

•Faster training and inference times, which facilitate the use of the HMUMASDNet in a real-time or near-real-time clinical context.

•Reducing memory requirements and model sizes to allow deployment in standard hospital hardware or even on edge devices.

•Reduced parameters and FLOPs, indicating a more efficient architecture that reduces computational expenses without compromising accuracy.

5. CONCLUSION

This study introduced the HMUMASDNet model, developed to enhance LCD through improved CT image synthesis and classification performance. Building on the advancements of the HMLGAN model, which effectively addressed limitations in traditional DL methods by enhancing high-resolution CT image synthesis and maintaining anatomical consistency, HMUMASDNet further refines these improvements. The experimental results demonstrate that HMLGAN achieves improved PSNR and SSIM values compared with existing GAN-based methods. Furthermore, the HMUMASDNet framework achieved classification accuracies of 98.62% on the SPIE-AAPM-NCI dataset and 98.11% on the LIDC-IDRI dataset. Statistical and ablation studies further indicate the effectiveness of the proposed framework in improving classification robustness. The results suggest that HMUMASDNet can serve as a promising framework for lung cancer classification.

REFERENCES

[1] Siegel, R.L., Miller, K.D., Wagle, N.S., Jemal, A. (2023). Cancer statistics, 2023. CA: A Cancer Journal for Clinicians, 73(1): 17-48. https://doi.org/10.3322/caac.21763

[2] Sieren, J.C., Ohno, Y., Koyama, H., Sugimura, K., McLennan, G. (2010). Recent technological and application developments in computed tomography and magnetic resonance imaging for improved pulmonary nodule detection and lung cancer staging. Journal of Magnetic Resonance Imaging, 32(6): 1353-1369. https://doi.org/10.1002/jmri.22383

[3] Than, J.C.M., Saba, L., Noor, N.M., et al. (2017). Lung disease stratification using amalgamation of Riesz and Gabor transforms in machine learning framework. Computers in Biology and Medicine, 89: 197-211. https://doi.org/10.1016/j.compbiomed.2017.08.014

[4] Baskar, S., Shakeel, P.M., Sridhar, K.P., Kanimozhi, R. (2019). Classification system for lung cancer nodule using machine learning technique and CT images. In Proceedings of the International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, pp. 1957-1962. https://doi.org/10.1109/ICCES45898.2019.9002529

[5] Thallam, C., Peruboyina, A., Raju, S.S.T., Sampath, N. (2020). Early stage lung cancer prediction using various machine learning techniques. In Proceedings of the International Conference on Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India, pp. 1285-1292. https://doi.org/10.1109/ICECA49313.2020.9297576

[6] Alamgeer, M., Mengash, H.A., Marzouk, R., et al. (2022). Deep learning-enabled computer-aided diagnosis model for lung cancer using biomedical CT images. Computers, Materials & Continua, 73(1): 1437-1448. https://doi.org/10.32604/cmc.2022.027896

[7] Raza, R., Zulfiqar, F., Khan, M.O., et al. (2023). Lung-EffNet: Lung cancer classification using EfficientNet from CT-scan images. Engineering Applications of Artificial Intelligence, 126: 106902. https://doi.org/10.1016/j.engappai.2023.106902

[8] Jalali, Y., Fateh, M., Rezvani, M., Abolghasemi, V., Anisi, M.H. (2021). ResBCDU-Net: A deep learning framework for lung CT image segmentation. Sensors, 21(1): 268. https://doi.org/10.3390/s21010268

[9] Chen, W., Yang, F.C., Zhang, X.R., Xu, X., Qiao, X. (2021). MAU-Net: Multiple attention 3D U-Net for lung cancer segmentation on CT images. Procedia Computer Science, 192: 543-552. https://doi.org/10.1016/j.procs.2021.08.056

[10] Said, Y., Alsheikhy, A.A., Shawly, T., Lahza, H. (2023). Medical images segmentation for lung cancer diagnosis based on deep learning architectures. Diagnostics, 13: 546. https://doi.org/10.3390/diagnostics13030546

[11] Naseer, I., Akram, S., Masood, T., Rashid, M., Jaffar, A. (2023). Lung cancer classification using Modified U-Net-based lobe segmentation and nodule detection. IEEE Access, 11: 60279-60291. https://doi.org/10.1109/ACCESS.2023.3285821

[12] Tan, J.X., Jing, L.L., Huo, Y.M., Li, L.H., Akin, O., Tian, Y.L. (2021). LGAN: Lung segmentation in CT scans using generative adversarial network. Computerized Medical Imaging and Graphics, 87: 101817. https://doi.org/10.1016/j.compmedimag.2020.101817

[13] Kryuchkov, M., Khanzhina, N., Osmakov, I., Ulyanov, P. (2021). CT images GAN-based augmentation with AdaIN for lung nodules detection. In International Conference on Machine Vision, SPIE, pp. 628-635. https://doi.org/10.1117/12.2587940

[14] Rezaei, S.R., Ahmadi, A. (2023). A GAN-based method for 3D lung tumor reconstruction boosted by a

knowledge transfer approach. Multimedia Tools and Applications, 82(28): 44359-44385. https://doi.org/10.1007/s11042-023-15232-0

[15] Chi, J.N., Wu, C.D., Yu, X.S., Ji, P., Chu, H. (2020). Single low-dose CT image denoising using a generative adversarial network with Modified U-Net generator and multi-level discriminator. IEEE Access, 8: 133470-133487. https://doi.org/10.1109/ACCESS.2020.3006512

[16] Gao, L.G., Xie, K., Wu, X.J., et al. (2021). Generating synthetic CT from low-dose cone-beam CT by using generative adversarial networks for adaptive radiotherapy. Radiation Oncology, 16: 1-16. https://doi.org/10.1186/s13014-021-01928-w

[17] Huang, Z.Z., Zhang, J.P., Zhang, Y., Shan, H.M. (2021). DU-GAN: Generative adversarial networks with dual-domain U-Net-based discriminators for low-dose CT denoising. IEEE Transactions on Instrumentation and Measurement, 71: 1-12. https://doi.org/10.1109/TIM.2021.3128703

[18] Toda, R., Teramoto, A., Kondo, M., Imaizumi, K., Saito, K., Fujita, H. (2022). Lung cancer CT image generation from a free-form sketch using style-based pix2pix for data augmentation. Scientific Reports, 12: 12867. https://doi.org/10.1038/s41598-022-16861-5

[19] Tabibian, M., Razmpour, T., Saha, R. (2025). Diffusion models vs. DCGANs for class-imbalanced lung cancer CT classification: A comparative study. Intelligence-Based Medicine, 13: 100336. https://doi.org/10.1016/j.ibmed.2025.100336

[20] Khiati, R.N., Brillet, P.Y., Ispas, R., Fetita, C. (2025). Diff-Lung: Diffusion-based texture synthesis for enhanced pathological tissue segmentation in lung CT scans. In 2025 IEEE 22nd International Symposium on Biomedical Imaging (ISBI), Houston, TX, USA, pp. 1-5. https://doi.org/10.1109/ISBI60581.2025.10980827

[21] Bishnoi, V., Goel, N. (2023). Tensor-RT-based transfer learning model for lung cancer classification. Journal of Digital Imaging, 36(4): 1364-1375. https://doi.org/10.1007/s10278-023-00822-z

[22] Harale, A.M., Bairagi, V.K. (2022). Detection of lung cancer pulmonary nodules using computer aided detection system. In 2022 6th International Conference On Computing, Communication, Control And Automation (ICCUBEA), Pune, India, pp. 1-7. https://doi.org/10.1109/ICCUBEA54992.2022.10011097

[23] Ji, Z.L., Wu, Y., Zeng, X.Y., An, Y.L., Zhao, L., Wang, Z.W. (2023). Lung nodule detection in medical images based on improved YOLOv5s. IEEE Access, 11: 76371-76387. https://doi.org/10.1109/ACCESS.2023.3296530

[24] Angel Mary, A., Thanammal, K.K. (2024). Lung cancer detection via deep learning-based pyramid network with honey badger algorithm. Measurement: Sensors, 31: 100993. https://doi.org/10.1016/j.measen.2023.100993

[25] Harale, A.M., Bairagi, V.K., Boonchieng, E., Bachute, M.R. (2024). Nodules detection in lungs CT images using improved YOLOV5 and classification of types of nodules by CNN-SVM. IEEE Access, 12: 140456-140471. https://doi.org/10.1109/ACCESS.2024.3466292

[26] Xue, X.Z., Ma, Y.H., Du, W.W., Peng, Y.H. (2025). Squeeze-and-excitation vision transformer for lung nodule classification. IEEE Access, 13: 24852-24866. https://doi.org/10.1109/ACCESS.2025.3529127

[27] Meszlényi, R.J., Buza, K., Vidnyánszky, Z. (2017). Resting state fMRI functional connectivity-based classification using a convolutional neural network architecture. Frontiers in Neuroinformatics, 11: 61. https://doi.org/10.3389/fninf.2017.00061

[28] The Cancer Imaging Archive. (2026). SPIE-AAPM Lung CT Challenge Dataset. https://www.cancerimagingarchive.net/collection/spie-aapm-lung-ct-challenge/.

[29] The Cancer Imaging Archive. (2026). LIDC-IDRI Dataset. https://www.cancerimagingarchive.net/collection/lidc-idri/.

NOMENCLATURE

h'	Reconstructed latent variable
$\mathcal{L}_{HR}^{GAN}$, $\mathcal{L}_{LR}^{GAN}$	Loss function for $A2D_H$ and $A2D_L$
M_r	Feature map associated with slice r
D_{ij}	Diagonal matrix
F'	Refined feature representation
H^l	Node features at layer l
H^{l+1}	Updated node features at layer $l+1$
L_0, L_1, L_2, L_3	SC matrices
$L_{ru1}, L_{rm1}, L_{rl1}, L_{lu1}, L_{ll1}$	Lobe-specific connectivity matrices
$L'_{ru1}, L'_{rm1}, L'_{rl1}, L'_{lu1}, L'_{ll1}$	Updated lobe-specific connectivity matrices
SV	Selector function
W^l	Layer weights in GCN
$\odot$	Element-wise multiplication operator
$\mathcal{L}$	Represent Loss of LSG and $A2D$
$A2D_{HR}$, $A2D_{LR}$	High-resolution and low-resolution Anatomy-Aware Discriminator
$CSF(.)$	Cross-Sectional Fusion operation
F	Node feature vector
I	Identity matrix
LSG_{HR}, LSG_{LR}	High-resolution and low-resolution Image generator
N	Number of ROIs
A	Adjacency matrix
S	Symmetrically Normalized Adjacency Matrix
$act(x)$	Activation Function
d	Number of hidden neurons in the GCN layer
μ, σ	Mean and standard deviation
ϵ	Gaussian noise
$\hat{X}_{LR}$, $\hat{X}_{HR}$	Low- and high-resolution image