Multimedia Tools and Applications (2025) 84:19065–19085
https://doi.org/10.1007/s11042-024-19801-9

UDCT: lung Cancer detection and classification using U-net and DARTS for medical CT images

Aakanksha Gupta[1] · **Ashwni Kumar**[1] · **Kamakshi Rautela**[2]

Received: 27 July 2023 / Revised: 25 June 2024 / Accepted: 27 June 2024 /
Published online: 13 July 2024

Abstract
Lung cancer is the most fatal disease in recent times. Early detection of the same is very crucial and challenging task. Therefore, proper diagnostic and treatment strategies should be employed for the automatic detection of nodules from Computed Tomography (CT) images. In this direction this paper introduces a novel UDCT, a new and effective framework for detecting and classifying lung cancer based on CT images. The proposed framework exploits Modified U-Net architecture that utilizes the traditional U-Net architecture for multi-scale feature extraction with the DARTS (Differentiable Architecture Search) which utilizes the concept of continuous relaxation of the architecture representation with efficient gradient descent search for efficient classification of lung cancer. In addition to this, the proposed framework, incorporates Multilevel Otsu thresholding for image preprocessing and to further segment the lung nodule from the CT scan images. Furthermore, the modified U-Net is capable to detect and classify lung nodule effectively and efficiently. To evaluate the performance of the proposed UDCT framework, extensive experiments are conducted on LIDC-IDRI and IQ-OTH/NCCD CT image datasets that provided state-of-the-art results. The proposed UDCT framework provided an accuracy of 95.01% on LIDC-IDRI dataset and 96.82% on IQ-OTH/NCCD dataset. An ablation study is also conducted that validated the efficiency and efficacy of the proposed UDCT framework.

Keywords U-net · DARTS · Lung cancer · CT images · Biomedical engineering · Segmentation · Thresholding

✉ Aakanksha Gupta
akanksha001phd20@igdtuw.ac.in

Ashwni Kumar
ashwanikumar@igdtuw.ac.in

Kamakshi Rautela
kamakshi_2k19phdec06@dtu.ac.in

[1] Department of ECE, Indira Gandhi Delhi Technical University for Women, Delhi, India

[2] Department of ECE, Delhi Technological University, Delhi, India

Abbreviations

CT	Computed Tomography
MRI	Magnetic resonance imaging
PET	Positron emission tomography
CXR	Chest radiography
DARTS	Differentiable Architecture Search
GPU	Graphics Processing Unit
LIDC-IDRI	Lung Image Database Consortium and Image Database Resource Initiative
AUC & ROC	Area Under Curve & Receiver Operating Characteristics

Variables

N & A	Normal & Abnormal classes
X & L	Threshold value & Grayscale level
p	One pixel value
Pq_a^o	Distributional probability, where 'o' is component value and $'q_a^o\ '$ is histogram.
$H_0^o(X)$ and $H_1^o(X)$	Distributional Probabilities of classes 'N' & 'A' respectively.
μ & σ	Sum and Variance of the classes.
$Z^{(u,v)}$	Output of concatenation approach, where $v-th$ is the convolution in $u-th$ layer module.
M(•), D(•), S(•) and [•]	Cascaded operation, Down-sampling function, Up-sampling function, and Channel concatenation operation respectively.
$b^{(g,h)}$	Collection of operations including convolution and max pooling
$\gamma^{(g,h)}$	Vector Quantity

1 Introduction

Lung cancer, a highly lethal disease, is responsible for approximately one out of every five cancer-related deaths. It is characterized by the presence of solid masses of tissue known as pulmonary tumor that develop in and around the lungs [1–3]. Lung cancer is a major cause of mortality for males, females, and transgender individuals worldwide, claiming around five million lives annually. The WHO (World Health Organization) estimates indicate that lung cancer causes millions of deaths each year. Unfortunately, more than 80% of lung cancer patients face a poor prognosis due to the advanced stage of the disease, which often renders surgery ineffective at the time of diagnosis [1, 4].

Detecting lung cancer early is crucial for effective treatment and saving lives since symptoms typically appear only when the disease has progressed significantly [5, 6]. Manual interpretation of numerous medical images used for lung cancer diagnosis can be time-consuming, labour-intensive, and prone to inaccuracies. Different methods such as X-rays, CT scans, MRI, PET, and breath analysis can be employed for lung cancer treatment [7, 8]. MRI and PET have limitations in identifying and staging lung cancer, while CXR (Chest radiography) and CT imaging techniques are prone to radiation exposure. However, CT imaging has become the preferred method for diagnosing lung cancer due to its ability to provide detailed information about the location and size of nodules [9]. Low-dose CT screening has been effective in detecting early-stage cancer tumours. In the past, image-processing

techniques have been used for diagnosing lung cancer, and machine-learning algorithms have played a role in analyzing CT scans and classifying lung nodules [9]. Feature extraction is a critical step in implementing machine learning, but existing methods often rely on manually created features, leading to lower accuracy. This limitation hinders computer-aided detection (CAD) systems from achieving optimal performance [10]. To address this issue, researchers have proposed the use of Artificial Intelligence, particularly deep learning. Deep learning algorithms can automatically learn important features from observed data during training, eliminating the need for manual feature creation and enabling end-to-end detection in CAD systems [11]. Convolutional Neural Networks (CNN) have shown superior performance compared to other deep learning networks in this context [12, 13].

Deep learning methods have shown promising results in classification and segmentation. Deep learning approaches, however, still face numerous challenges in the identification of lung cancer. First, the robustness of the prediction model is a challenging task as these models rely on large sets of annotated medical pictures that are contributed by seasoned radiologists, to meet their training requirements. Furthermore, great interpretability is necessary for the clinical application of deep learning; however, the learnt features cannot be satisfactorily explained by existing deep learning algorithms. Additionally, the existing approaches are not very applicable in clinical practice, mainly due to the inexperience of non-medical investigators in choosing more pertinent clinical outcomes. Therefore, the primary objectives of this study are to improve performance, reduce diagnostic time, and identify the location and characteristics of lung nodules. The goal is to develop a deep learning architecture specifically for early diagnosis and classification of lung cancer using deep learning techniques. In this direction, this paper presents a novel UDCT structure that exploits the concept of DARTS and U-Net with multi-level Otsu thresholding for easy detection of lung nodules. A basic block diagram representation is depicted in Fig. 1. Further, the proposed UDCT can effectively classify lung cancer into benign or malignant nodules.

The main Contributions of this work are:

i. This paper presents a lung cancer detection and classification framework namely, UDCT that efficiently detects and classifies lung cancer from CT images.
ii. The proposed framework exploits Multi-level Otsu thresholding on pre-processed images to efficiently extract features from CT images. Further, the extracted features are utilized by the U-Net architecture for the segmentation of CT images.
iii. Further, the proposed U-Net architecture is fused with DARTS (Differentiable Architecture Search) that uses a simplified gradient descent search in conjunction with the concept of continuous relaxation to describe architecture for efficient classification of lung cancer.
iv. To confirm the efficacy of the suggested UDCT framework, extensive experiments are conducted on LIDC-IDRI and IQ-OTH/NCCD CT image datasets that provided state-of-the-art results.
v. Also, an ablation study is carried out to prove the efficiency and efficacy of the suggested framework.

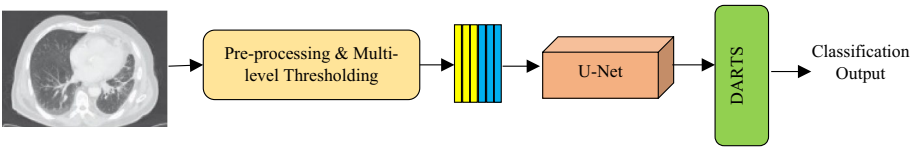

Fig. 1 Block Diagram for the proposed UDCT Framework

The organization of the rest of the paper is as follows: Section 2 presents the different state-of-the-art techniques for lung cancer detection, Section 3 presents the proposed methodology, Section 4 discusses the results obtained on the proposed UDCT framework. Finally, Section 5 concludes the work.

2 Related work

In [1], a thorough analysis was conducted on various CNN-based techniques and algorithms to analyse pulmonary nodules in medical images. The focus was on examining the methodologies employed, the datasets used for validation, and the detection outcomes of recently developed CNN-based systems for nodule classification. Three different architectures, each consisting of varying numbers of layers, were studied, leading to a fully connected layer and a SoftMax classifier. The LIDC-IDRI dataset was utilized for training and testing, as it yielded the best results in terms of accuracy and precision when combined with other architectures [8]. A state-of-the-art performance was achieved for both lung nodule identification and malignancy classification tasks using a cascade of 3D CNNs in a CADe & CADx system. However, if the CADe model fails to detect a malignant lesion, the subsequent CADx model is unable to classify that scan as malignant, resulting in a significant source of false negatives in the system [10].

To find pulmonary nodules in the lung's current CT pictures, a convolution neural network that has been genetically optimised is used [9] and also the detection rate and classification accuracy of pulmonary nodules can be effectively increased with the use of wavelet denoising on lung CT images. [13] You Only Look Once, Version 3 (YOLOv3), Single-Shot Detector (SSD), and Faster Region-Convolutional Neural Network (Faster-RCNN) are components of an automated deep learning system that is utilized for detection and Visual Geometry Group (VGG-16), Resnet 101, and Mobilenet for classifying lung nodules in CT scans. In this study, False Positive Rate (FPR) is reduced, and the accuracy is enhanced for classification and detection architectures. A hybrid approach that further aids in reducing the computational complexity of the classification is produced by combining Support Vector Machine with Feed-Forward Back Propagation Neural Network (FFBPNN) [14]. By using a hybrid classification algorithm, the classification accuracy is enhanced to 98.08%. To perhaps increase classification accuracy in the future, Deep Neural Network are used in contrast to FFBPNN. To identifying and categorizing lung cancer, a comprehensive framework is presented that includes a generative model designed to transform an unbalanced dataset which is smaller in size into a balanced one which is larger in size and ResNet50 which classifying CXR pictures into benign and malignant during the online phase by using a large balanced dataset [11]. In this, authors want to modify their system in the future to incorporate generative adversarial networks and other generative models. [30] Presents an optimization-driven technique for lung cancer classification. Lung cancer is classified using a Hierarchical Attention Network (HAN), Firefly Competitive Swarm Optimisation (FCSO), which combines the Firefly algorithm (FA) and Competitive Swarm Optimisation (CSO). The suggested FCSO-based HAN performs well, achieving a high accuracy rate of 91.3%, sensitivity of 88%, and specificity of 89.1%. The author addresses the convergence problems and computational overhead related to the multilevel thresholding process by utilizing the Levy flight and Gravitational Search Algorithm (LCGSA) based on Chaos theory. This technique is used to improve COVID-19 chest CT scan images

segmentation [31]. The author may also look into the future use of LCGSA in deep learning applications, such as CNNs, U-Nets, and LSTMs.

The authors of this paper addressed the scalability issue in architecture search by utilizing DARTS, a differential approach. This strategy depends on continuous or smooth relaxation, in contrast to traditional approaches that use evolutionary algorithms or reinforcement learning inside a discrete and non-smooth search space. The search was performed on multiple datasets including WikiText-2, CIFAR-10, Penn Treebank, and ImageNet. However, a limitation of this work is the potential discrepancies between the continuous architecture encoding and the resulting discrete architecture, which can impact the effectiveness of the current approach [15].

In another study, blood samples were taken from healthy individuals and those who had been diagnosed with breast cancer were subjected to metabolomics analysis using NMR spectroscopy and direct analysis in real time (DART)-MS. The new method introduced in this research improves the representation of disease classifications by allowing the expression of disease on a continuous scale rather than a binary scale, resulting in more accurate disease detection. However, DART-MS did not capture as much breast cancer-related variation [16]. In the same study, U-DARTS was employed to identify and categorize breast lesions in two datasets, achieving testing and validation accuracies of 91% and 98%, respectively. However, due to incomplete data, the study encountered low accuracy in one of the datasets [17].

Methods discussed above are mainly utilized for detecting and classifying lung cancer that may have certain limitations in a variety of areas, including accuracy, sensitivity, specificity, dataset availability, preprocessing techniques, and sometimes the classification models used. We present a novel UDCT framework with Multi-Level Otsu thresholding to address these issues and to improve image quality during preprocessing. Furthermore, to obtain higher accuracy, we combine U-NET and DARTS to improve lung cancer classification. Furthermore, different methods for the lung cancer detection are presented in Table 1.

3 Proposed methodology

This section describes the proposed UDCT framework (as depicted in Fig. 2) for the detection and classification of lung diseases. The proposed framework utilized a modified U-NET model which exploits DARTS architecture for efficient classification of lung cancer diseases. Initially, to improve the quality of the input images, an image enhancement technique is applied through pre-processing. The pre-processing step involves multilevel Otsu for better results. After pre-processing, the image is segmented using U-NET, and quantitative features are extracted. To enhance the classification accuracy, we fuse DARTS (Differentiable Architecture Search) with the U-NET model. The overall goal of this proposed work is to achieve improved results in the classification of lung diseases, particularly lung cancer.

3.1 Problem statement

In order to differentiate between benign and malignant nodules, computer-aided detection techniques for lung cancer require an examination of the visual properties of lung nodules on CT scans. Developing an automated detection method is becoming more and more crucial for early detection of lung cancer. It is incredibly challenging to build such a system using traditional approaches.

Table 1 Review based on some existing method for lung cancer detection

References	Approach	Evaluated Parameters
[1]	CNN-based techniques	–
[8]	Deep Convolutional Neural Networks (DCNN)	Accuracy = 80.02%, Precision = 84.51%, Recall = 83.45% and Dice Score Coefficient = 81.65%
[10]	Computer-Aided Detection (CADe) and Computer-Aided Diagnosis (CADx)	–
[9]	Wavelet analysis algorithm and Genetic Algorithm	Error = 0.01% which increased accuracy
[13]	Faster-RCNN, YOLOv3, SSD, VGG-16, Resnet 101, and Mobilenet	–
[14]	Feed-Forward Back Propagation Neural Network (FFBPNN)	Accuracy = 98.08%
[11]	Convolutional Variational Generative Auto-encoder and ResNet50	Accuracy = 98.91%, Area Under Curve (AUC) = 98.85%, Sensitivity = 98.46%, Precision = 97.72% and F1 score = 97.89%.
[15]	Differentiable Architecture Search (DARTS)	Search Cost (GPU Days) =4
[16]	NMR spectroscopy and Direct Analysis in real time Mass spectrometry (DART)-MS	AUC = 0.99
[17]	Modified Differentiable Architecture Search(U-DARTS)	Validation Accuracy = 98% and Testing Accuracy = 91%
[31]	Hierarchical Attention Network (HAN), Firefly Competitive Swarm Optimisation (FCSO) which combines the Firefly algorithm (FA) and Competitive Swarm Optimisation (CSO)	Accuracy = 91.3%, Sensitivity = 88%, and Specificity = 89.1%.
[32]	Levy flight and Chaos theory-based Gravitational Search Algorithm (LCGSA)	Accuracy = 92%, Precision = 1, Recall = 83% and Dice Coefficient = 90%

Deep-learning-based architectures have been extensively researched and developed in the last few years with the goal of being used to the healthcare industry. Thus, the goal of this work was to develop a deep learning-based, reliable, and accurate framework for lung cancer detection and classification. Furthermore, the pseudocode for the proposed methodology is presented in Algorithm 1.

Algorithm 1 Pseudocode for the proposed UDCT Framework for Lung Cancer detection and classification for CT Images.

Input: $I_{CT}(xxy)$; Thermal image with x=no. of rows and y= no. of column

Output: I_{UDART}; Output Image

Procedure UDCT

Initialize the image $I_{CT}(xxy)$

Apply Otsu multilevel thresholding (OMLT) on I_{CT} to separate object pixels from background pixels.

$I_{CT-OMLT} \xleftarrow[OMLT]{} I_{CT}$

For Normal (N) & Abnormal (A) class calculate thresholding value X

$N \leftarrow p$ if $0 \leq p < X$,

$A \leftarrow p$ if $X \leq p < L\text{-}1$

Where p is one of pixels of the grayscale image that can be represented in L grayscale levels

Apply UNet segmentation on I_{CT} image.

while $(x < UNet_{\text{iterations}})$ do

for each patch do

Update the image patches obtained through segmentation using eq (14)

end for

end while

Update I_{CT}

for i = 1 to n (n= no. of column in image)

$U(I_{CT}) = U(I_{CT-OMLT})$

end for

Use DARTS for classification of abnormal and normal images.

return archive

end procedure

3.2 Multilevel Thresholding using Otsu method

Image thresholding is a method used to convert an image, whether it is in full-color or grayscale, into a binary image where pixels are either classified as foreground or background. Its main purpose is to separate object pixels from background pixels. The determination of an appropriate threshold value plays a critical role in the segmentation process. Here, threshold value X is calculated as:

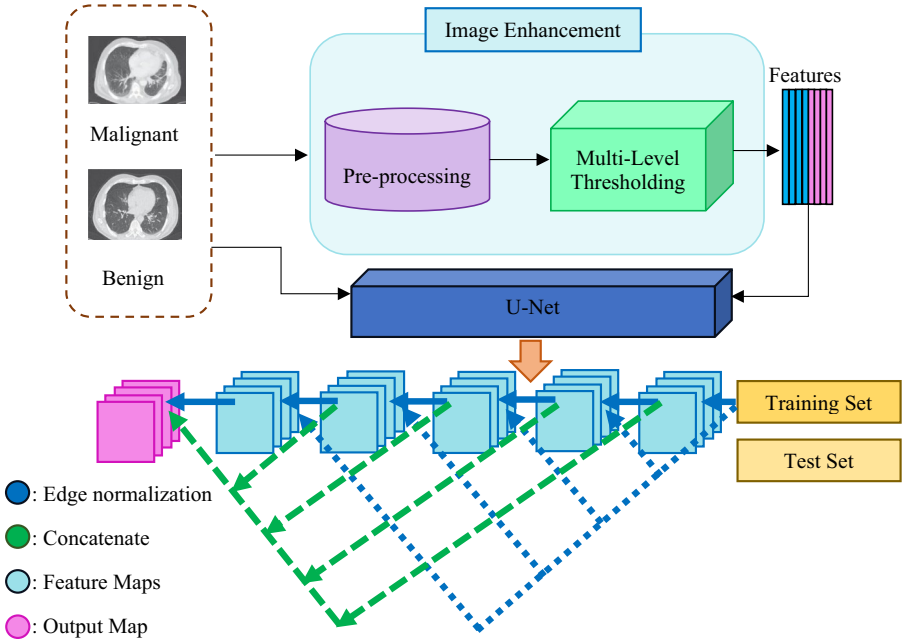

Fig. 2 Proposed UDCT Framework Methodology

$$\mathrm{N} \leftarrow \mathrm{p} \text{ if } 0 \leq \mathrm{p} < \mathrm{X}, \tag{1}$$

$$\mathrm{A} \leftarrow \mathrm{p} \text{ if } \mathrm{X} \leq \mathrm{p} < \mathrm{L} - 1$$

Where, N and A are the Normal & Abnormal classes when X is the threshold and p is one of pixels of the grayscale image that can be represented in L grayscale levels $L = \{0, 1, 2, \ldots, L-1\}$.

Equation (1) relate to the bi-level threshold and expanded to many sets like:

$$\mathrm{N} \leftarrow \mathrm{p} \text{ if } 0 \leq \mathrm{p} < X_1, \tag{2}$$

$$\mathrm{A} \leftarrow \mathrm{p} \text{ if } X_1 \leq \mathrm{p} < X_2,$$

$$N_n \leftarrow \mathrm{p} \text{ if } X_n \leq \mathrm{p} < X_{n+1},$$

$$A_m \leftarrow \mathrm{p} \text{ if } X_m \leq \mathrm{p} < \mathrm{L} - 1$$

In eq. 2, $\{X_1, X_2, \ldots.., X_n, X_{n+1}\}$ are different threshold levels. To determine the ideal threshold value, Otsu proposes various objective functions that must be maximised. For figuring out threshold values, Otsu is a well-known method. The Otsu method, which uses the greatest variance values of several classes as the only segmentation criterion in images, is preferred because of its ease of use and great efficacy. Applying the L designated intensity level for the grayscale image, the intensity value of image computation can be written as [18].

$$\mathrm{P}q_a^o = \frac{q_a^o}{P_{tot}}, \tag{3}$$

$$q_a^o \geq 0, \tag{4}$$

$$\sum_{a=0}^{P_{tot}} Pq_a^o = 1,$$

$$\mathrm{o} = \{1, \text{if Grayscale Image},$$

o = component, the value specified from I.

where,

$\mathrm{P}q_a^o$ = Distributional probability.

$\mathrm{P}_{\mathrm{tot}}$ = the total number of pixels in the image.

q_a^o =histogram (the number of pixels correlates to the intensity in o).

Now, Probability distributions for two classes (N and A) segmentation are given as [18]:

$$H_0^o(\mathrm{X}) = \sum_{a=0}^{X} Pq_a^o, H_1^o(\mathrm{X}) = \sum_{a=X+1}^{L} Pq_a^o \tag{5}$$

The mean and the variance is being calculated as representation mathematically in equation '6' and '9':

$$\mu_0^o = \sum_{a=0}^{X} \frac{aPq_a^o}{H_0^o(X)} \text{ and } \mu_1^o = \sum_{a=X+1}^{L} \frac{aPq_a^o}{H_1^o(X)} \tag{6}$$

$$\text{and } \mu_{tot}^o = H_0^o\mu_0^o + H_1^o\mu_1^o, \tag{7}$$

$$H_0^o + H_1^o = 1 \tag{8}$$

The total variation between the classes is computed after the calculation of the above values as,

$$\sigma^{2^o} = \sigma_1^o + \sigma_2^o \tag{9}$$

In the mathematical sense, number two does not represent exponent; it is a component of the Otsu variant operator in eq. (8). In the above, σ_1^o and σ_2^o are the variants of N and A which are obtained as:

$$\sigma_1^o = H_0^o\left(\mu_0^o + \mu_{tot}^o\right)^2 \& \sigma_2^o = H_1^o\left(\mu_1^o + \mu_{tot}^o\right)^2 \tag{10}$$

In (10), the value of $'\mu_{tot}^o{}'$ is find with the help of eq. (8).

For multilevel thresholding, the objective function can be defined by (9) because by the increase of number of thresholds there will be an exponential growth in processing time.

$$\theta(X) = \max\left(\left(\sigma^{2^o}(X)\right), 0 \leq \mathrm{X} \leq \mathrm{L} - 1\right. \tag{11}$$

Where, $\mathrm{X} = \{X_1, X_2, X_3 \ldots\ldots X_{k-1}\}$ there are multiple thresholds in the vector X. So, for various number of threshold values Variance, Probability Distribution and the Average values can be calculated by eqs. (11), (12) and (13) respectively.

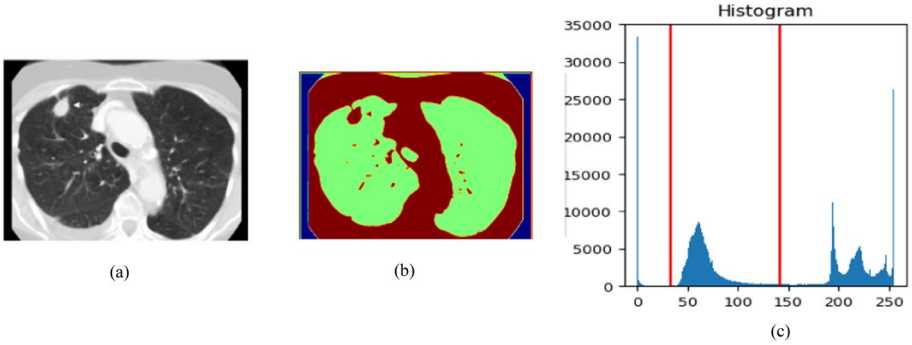

Fig. 3 Multilevel thresholding by using Otsu method, (**a**) Original lung image, (**b**) Result after Otsu thresholding, and (**c**) Equalized Histogram

$$\sigma^{2^o} = \sum_{c=0}^{k} H_c^o \left(\mu_c^o - \mu_{tot}^o\right) \tag{12}$$

$$H_0^o(X) = \sum_{c=1}^{X_1} Pq_c^o, H_{k-1}^o(X) = \sum_{c=X_{K+1}}^{L} Pq_c^o \tag{13}$$

$$\mu_0^o = \sum_{c=1}^{X_1} \frac{cPq_c^o}{H_0^o(X_1)}, H_{k-1}^o = \sum_{c=X_{k+1}}^{L} \frac{cPq_c^o}{H_1^o(X_k)} \tag{14}$$

Where, 'c' represents a class in the above equations.

Furthermore, the output obtained after multilevel thresholding on LIDC dataset is presented in Fig. 3. Multilevel thresholding is a technique that partitions an image containing various shades of gray into multiple distinct regions. This approach involves choosing multiple thresholds to separate the image into different brightness zones, representing various backgrounds and objects.

3.3 Image segmentation using U-net

The objective of segmentation is to separate individual elements from an entire image. Segmentation is a procedure that transforms original images into distinct regions or components. The accuracy of the resulting segmentation is used as a measure of how well the process was performed. The precision of the segmentation increases when the objects are recognized more accurately. In the field of biomedical imaging, the U-Net architecture is commonly utilized for image segmentation tasks. This architecture consists of two parts: a contracting side and an expanding side, which are symmetrical. The contracting side includes components such as dropout, pooling, convolutional, and Rectified Linear Unit (ReLU) layer. Each down sampling step in the process doubles the number of feature channels [19].

i. Down-sampling is performed in contracting path with the shallower, lower, and coarser grained characteristics are converted into the deeper, semantic, and finer grained features by this connection [20].
ii. The connection between contracting and expanding sides called skip connection enable the network to take in the required features from the relevant layer. Skip con-

nection bridges the shallower, lower, and coarse-grained features in the contracting path and the deeper, semantic, and fine-grained features in the expanding path.

iii. Up-sampling is performed in expanding path which transforms back Deeper, semantic, and fine-grained characteristics into shallower, lower, and coarse-grained features. At every section on the expanding side, the number of feature channels is cut in half [20].

Using the channel concatenation fusion approach, let $Z^{u,v}$ represent the $v-th$ convolution in the $u-th$ layer module. The output $z^{u,v}$ of $Z^{u,v}$can be calculated according to the following eq. [20]:

$$z^{u,v} = \begin{cases} \mathrm{M}\left(\mathrm{D}\left(z^{u-1,v}\right)\right) & \text{for u} > 0, \text{v} = 0 \\ \mathrm{M}\left(\left[\left[z^{u,l}\right]_{l=0}^{v-1}, \mathrm{S}\left(z^{u+1,v-1}\right)\right]\right) & \text{for u} = 0, \text{v} > 0 \\ \mathrm{M}\left(\left[\left[z^{u,l}\right]_{l=0}^{v-1}, \mathrm{S}\left(z^{u+1,v-1}\right), \mathrm{D}\left(z^{u-1,v}\right)\right]\right) & \text{for u} > 0, \text{v} > 0 \end{cases} \tag{15}$$

In the above equation, M(•) is a cascaded operation including convolution, ReLU activation function and batch normalization (BN), D(•) is a down- sampling function, S(•) is an up-sampling function, and [•] is the channel concatenation operation [20].

After multi-level thresholding, we have extracted following qualitative features with the following value: *Mean* = 1.925, *Standard Deviation* = 59.283, *Dissimilarity* = 0.6221, *Energy* = 24.171 *and Entropy* = 8.177 for LIDC-IDRI dataset. Furthermore, for IQ-OTH/NCCD dataset, *Mean* = 0.916, *Standard Deviation* = 50.302, *Dissimilarity* = 01.462, *Energy* = 23.059 *and Entropy* = 8.257 The different features shown with the help of Fig. 4. Along with the above network we explained DARTS (Differentiable Architecture Search) for searching the efficient architecture. Darts has had a significant impact on neural architecture search. The representation of DARTS is presented in Fig. 5. Earlier techniques relied on reinforcement learning and needed plenty of processing power. It required 3150 GPU days for evolution or 2000 Graphics Processing Unit (GPU) days for reinforcement learning [15].

We relax the search space to be continuous so that the architecture can be optimised via gradient descent with regard to the performance of the validation set. This allows us to search over a continuous universe of candidate architectures rather than a discrete set. This search is based on bilevel optimization and also relevant to recurrent structures as well as convolutional architectures. In the search space, a directed acyclic graph depicts the computation process for the proposed architecture as depicted in Fig. 5. A Directed acyclic graph with F nodes ($f_0, f_1, f_2 \ldots\ldots f_F$) arranged in an orderly manner.

In this work, we take B be a collection of operations which includes convolution and max pooling, where each operation denotes a function to be applied on $r(t)$ as $b(.)$. Here all possible operations are given by $b^{(g,h)}$ [15]:

$$b^{(g,h)}(r) = \sum\nolimits_{b \epsilon B} \frac{\exp\left(y_b^{(g,h)}\right)}{\sum_{b' \epsilon B} \exp\left(\gamma_{b'}^{(g,h)}\right)} r(t) \tag{16}$$

For a pair of nodes (g, h) a vector $\gamma^{(g,h)}$ of dimension $|B|$ is used as a parameter for the operation of mixing weights .After that, learning a set of continuous variables is all that

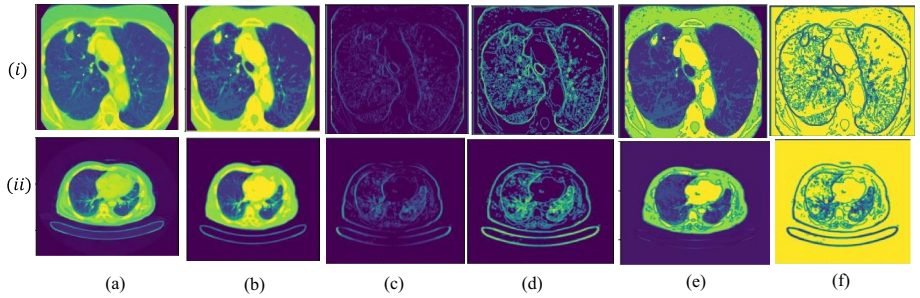

Fig. 4 Qualitative features extracted from the proposed U-NET approach for (*i*) LIDC-IDRI dataset, (*ii*) IQ-OTH/NCCD dataset. (**a**) Original source image from dataset, (**b**) mean value, (**c**) Dissimilarity, (**d**) Entropy, (**e**) Standard Deviation, and (**f**) Energy

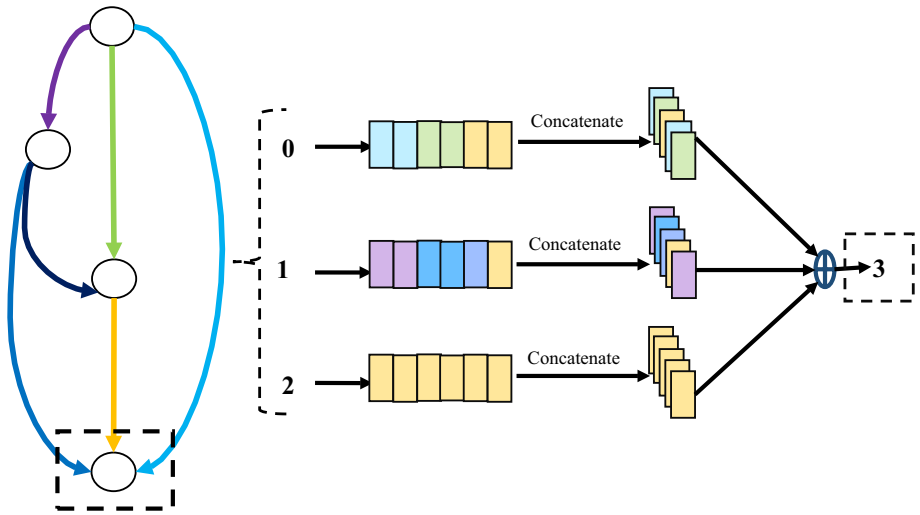

Fig. 5 Representation of DARTS

remains of the architectural search task $\gamma = (\gamma^{(g,h)})$ [17]. Based on each predecessor, each intermediate node is calculated by 'r(t)' mentioned in eq. 17 as follows:

$$\mathrm{r(t)} = \sum_{g<h} b^{(g,h)}(r(t)) \tag{17}$$

In this work, we have fused DARTS with UNET for getting better lung cancer classification results.

3.4 Dataset used

To assess the effectiveness of the proposed UDCT framework for the categorization of benign and malignant pulmonary nodules, the LIDC-IDRI public dataset of thoracic CT scans and IQ-OTH/NCCD dataset are used.

LIDC-IDRI stands for Lung Image Database Consortium and Image Database Resource Initiative. On a CT scan, lung nodules can be seen in several slices. Each case includes a clinical thoracic CT scan as well as the reports of four expert thoracic radiologists in an XML file [13]. Additionally, the XML files include semantic diagnostic elements that were noted by four qualified thoracic radiologists. The Lung Imaging Database Consortium (LIDC) initiative's goal is to assist a group of institutions in creating a database of spiral CT lung pictures and developing recommendations for a resource for spiral CT lung images. There are 1568 thoracic Computed Tomography (CT) scans with nodules in the dataset out of which 650 were malignant cases and 918 were benign cases [19]. There are three different types of nodules: ("nodule > or =3 mm," "nodule <3 mm," and "non-nodule > or =3 mm") [21].

The IQ-OTH/NCCD dataset, which stands for Iraq-Oncology Teaching Hospital/ National Center for Cancer Diseases, includes CT scans of both healthy persons and those with various stages of lung cancer diagnosis. It has 1190 images total, each of which shows slices from 110 CT scans. Three classes—normal, benign, and malignant—are used to group these situations. 40 cases are categorized as malignant, 15 as benign, and 55 as normal. DICOM is the original format used to store CT scan data.

4 Computational complexity

The computational complexity of the proposed UDCT framework algorithm is described in this subsection. The following describes the space and time complexity:

- **Time Complexity**

1. The implementation of UDCT algorithm desires $O(I_{CT}(x \times y))$ time where x and y represent the number of rows columns respectively in each image.
2. The OMLT calculation of each pixel requires

$$O\left(I_{CT}(OMLT \times x \times y)\right)$$

time.

3. The UNet calculations require

$$O\left(I_{CT}\left(U_{\text{iter}} \times OMLT \times x \times y\right)\right)$$

time where U_{iter} is the maximum number of iterations to simulate segmentation.

4. The algorithm requires $O(I_{UDART}) = O(I_{CT-DART}(O(I_{CT}(U_{\text{iter}} \times OMLT \times x \times y)))$ $I_{CT-DART}(O(I_{CT}(U_{\text{iter}} \times OMLT \times x \times y)))$ time to classify abnormal and abnormal type of cancer.

- **Space Complexity**

The UDCT algorithm's space complexity is determined at the initial stage, which at any given time requires space. Due to this, the UDCT algorithm's overall space complexity is $O(I_{CT}(x \times y))$.

5 Results and discussions

This section discusses the implementation details, and evaluation parameters with the experimental results obtained on LIDC-IDRI and IQ-OTH/NCCD CT image datasets.

5.1 Implementation details

We train our model from beginning to end in order to identify and extract image features from the lung images. The patch size is set to 9 and the batch size is set as 32. The input images are downsized to 124×124 resolution. A learning rate 0.0001 is used to train the entire model on the dataset. For the setup, there will be 100 epochs. Accuracy and loss curves are plotted to evaluate the effectiveness of the proposed model. Further, ThinkPad laptop that satisfies certain requirements is used for the experiment. It has an Intel Core i7 (11th generation) processor with integrated AMD Radeon Graphics, clock rates ranging from 1.90 GHz to 4.40 GHz, and 16 GB of RAM. The study's experiments were conducted on Windows 11 with Jupyter Notebook, Python 3.12.3, and the backend libraries for TensorFlow and Keras.

5.2 Evaluation parameters

Accuracy, Loss, Precision, Recall, Dice Coefficient, Specificity and F1-Score are the several parameters that are used to evaluate the proposed model's performance.

i) **Accuracy**: It is defined by the proportion of accurate predictions to all predictions [22]. Represented by

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{18}$$

ii) **Recall**: Ratio of the number of Positive samples that were correctly identified as Positive to all Positive samples. More positive samples are detected when the value of recall increases [22].

$$Recall = \frac{TP}{TP + FN} \tag{19}$$

iii) **Precision**: It shows the percentage of samples that were actually positive and those that were projected to be positive out of all the samples [22].

$$Precision = \frac{TP}{TP + FP} \tag{20}$$

iv) **Dice Coefficient**: It is calculated by dividing the combined pixels in the two images by the area of overlap, multiplied by 2.

$$DSC = \frac{2 * |D \cap E|}{|D| + |E|} \tag{21}$$

Where D is the expected set of pixels and E is the actual set of pixels.

xxii) Specificity: The ratio of real negative samples—both actual and predicted—to the total number of truly negative samples is used to compute specificity [22].

$$\text{Specificity} = \frac{TN}{TN + FP} \tag{22}$$

vi) F1-Score: The "F1-score" accounts for both false positives (FPs) and false negatives (FNs), striking a compromise between "precision" and "recall". Since each variable undermines the other, it penalizes extreme levels of "precision" and "recall" [22].

$$\text{F1} = \frac{2 * P * R}{P + R} \tag{23}$$

Where P and R represents Precision and Recall Respectively.

In the eqs. (18) to (22), *TP*, *TN*, *FN*, *and FP* represents true positive, true negative, false negative, and false positive, respectively.

5.3 Performance analysis on LIDC-IDRI and IQ-OTH/NCCD CT image datasets

The experimental findings for the proposed U-DARTS model are detailed in this section. The LIDC-IDRI dataset is used for the experiment. The sample contains a wide variety of lung nodules, which are separated into two groups: Benign (Unharmful) and Malignant (harmful). The experiment is done by using 80% training set and 20% testing set. Training and validation accuracy, as well as training and validation loss, are used to assess the proposed framework using the LIDC-IDRI dataset. Figures 6 and 7 shows training-validation loss and training-validation accuracy respectively for LIDC-IDRI and IQ-OTH/NCCD datasets. From these accuracy and loss curves it is evident that the as the number of epochs increases the accuracy increases whereas the loss decreases. Furthermore, the training and validation accuracy obtained for the proposed UDCT framework on LIDC-IDRI and IQ-OTH/NCCD is 93.5% and 95.01% and 94.8% 96.82% respectively. Furthermore, the training and validation losses attained are 0.32 and 0.28 for LIDC dataset and 0.38 and 0.25 respectively on IQ-OTH/NCCD.

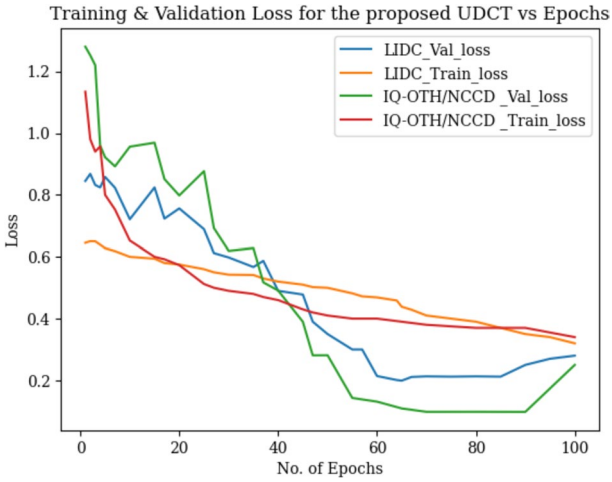

Fig. 6 Validation and Training Loss for the proposed UDCT Framework for LIDC and IQ-OTH Dataset

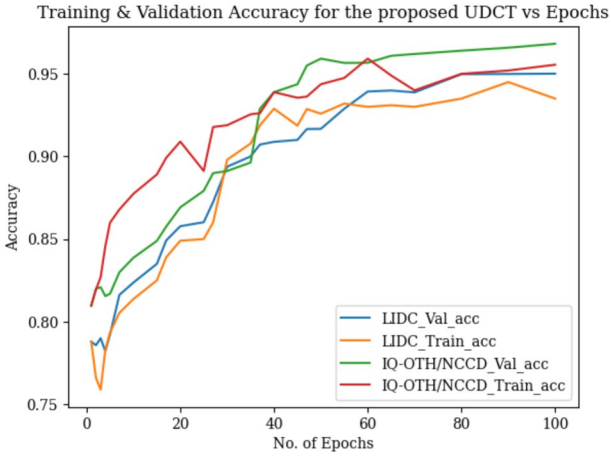

Fig. 7 Validation and Training Accuracy for the proposed UDCT Framework for LIDC and IQ-OTH Dataset

Tables 2 and 3 reports the comparison results obtained for the proposed UDCT framework with other state-of-the-art on LIDC-IDRI and IQ-OTH/NCCD CT image datasets for lung cancer detection and classification. The proposed work achieved an accuracy of 95.01%, Validation Loss of 0.28, Precision of 97.00%, Recall of 97.52%, Dice coefficient of 96.10%, Specificity of 99.09% and F1-Score of 96.01% on LIDC-IDRI Dataset. This proves that U-DARTS produces greater accuracy than the most effective approaches [24–29] presented in the Table 2. The proposed model offers relatively high accuracy when compared to the previous work because of these model adjustments, the usage of segmentation techniques, and the larger datasets as a result. Furthermore, a comparison for the proposed framework is made for the IQ-OTH/NCCD Dataset as reported in Table 3 which provided state-of-the-art results when compared with [33–38]. Also, Fig. 8(a) and (b) presents the confusion matrix for the proposed UDCT framework on LIDC and IQ-OOTH/NCCD datasets.

5.4 Ablation study

(i) An Ablation study is conducted to validate the performance of the proposed UDCT framework. In this section, we have performed three shuffles of the data, as depicted in Table 4, using fold 1, fold 2, and fold 3. This table illustrates how different evalu-

Table 2 Performance comparison with previous work done on LIDC-IDRI dataset

Model	Accuracy	Loss	Precision	Recall	Dice coefficient	Specificity	F1-Score
Wu et al. [23]	–	–	–	97.3%	80.2%	–	–
Yu et al. [24]	87.3%	–	77.1%	–	69.8%	–	–
Banu et al. [25]	94.66%	–	–	90.84%	90.35%	–	–
Bruntha et al. [26]	–	–	89.09%	81.08%	87.23%	–	–
Bruntha et al. [27]	–	–	96.75%	92.57%	95.7%	–	–
Bushara et al. [28]	94%	–	95%	94.5%	–	99.07%	94.5%
Shafi et al. [29]	94%	–	95%	94.5%	–	–	94.5%
UDCT	**95.01%**	**0.28**	**97.00%**	**97.52%**	**96.10%**	**99.09%**	**96.01%**

Table 3 Performance comparison with previous work done on *IQ-OTH/NCCD* dataset

Model	Accuracy	Loss	Precision	Recall	Dice coefficient	Specificity	F1-Score
Huseiny et al. [32]	94.38%	–	–	95.08%	–	93.70%	–
Mathews et al. [33]	95.00%	–	–	97.00%	–	98.00%	–
Suresha et al. [34]	–	–	98.40%	97.00%	–	–	97.70%
Al-Yasriy et al. [35]	93.54%	–	97.10%	95.71%	–	95.00%	96.40%
Gowda et al. [36]	96.50%	–	–	89.00%	–	94.00%	–
Santos et al. [37]	95.80%	–	–	99.20%	–	98.30%	–
Kareem et al. [38]	89.88%	–	98.55%	97.14%	–	97.50%	97.84%
UDCT	**96.82%**	**0.25**	**98.70%**	**97.50%**	**96.4%**	**98.40%**	**98.24%**

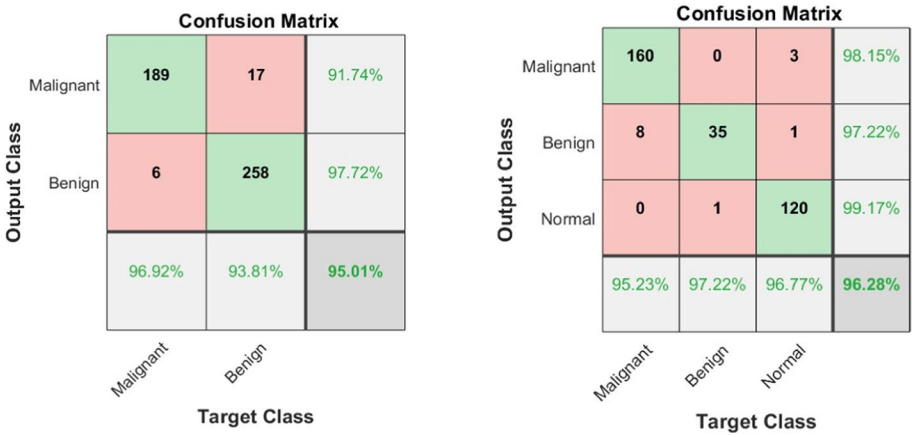

Fig. 8 Confusion Matrix for the proposed UDCT framework on (**a**) LIDC Dataset and (**b**) IQ-OTH Dataset

ation parameters vary across the different folds. The effectiveness of the method is demonstrated by employing a three-fold cross-validation approach in this study. Each fold of the cross-validation process allows us to assess the impact of the training samples on the accuracy, precision, recall, and Dice coefficient of the classification results by considering a specific fraction of the training data. Also, a comparison graph is depicted in Fig. 9 for three different folds with two classes each for one-fold.

(ii) An ablation is also conducted to study the influence batch size with optimizer utilized and learning rate. The results for the same are reported in Table 5. From these results we can infer that the best accuracy for 95.01% is obtained on Adam optimizer with a learning rate of 0.0001 for batch size of 32. Based on our results, we can say that the batch size and learning rate significantly affect the performance of the framework. Also, this experimental study aims at providing a better understanding of the batch size value which is 32 with the best suited optimizer and learning rate for the classification problems

Table 4 Ablation Study Results Obtained for the proposed UDCT Framework on LIDC-IDRI dataset

Fold	Classes	#Samples	Performance Measure			
			Accuracy	Precision	Recall	Dice Coefficient
Fold 1	Benign	918	0.909	0.912	0.873	0.901
	Malignant	650	**0.950**	0.961	0.966	0.948
Fold 2	Benign	918	0.910	0.904	0.908	0.890
	Malignant	650	0.928	0.950	0.951	0.922
Fold 3	Benign	918	0.904	0.935	0.870	0.907
	Malignant	650	0.948	**0.970**	**0.975**	**0.961**

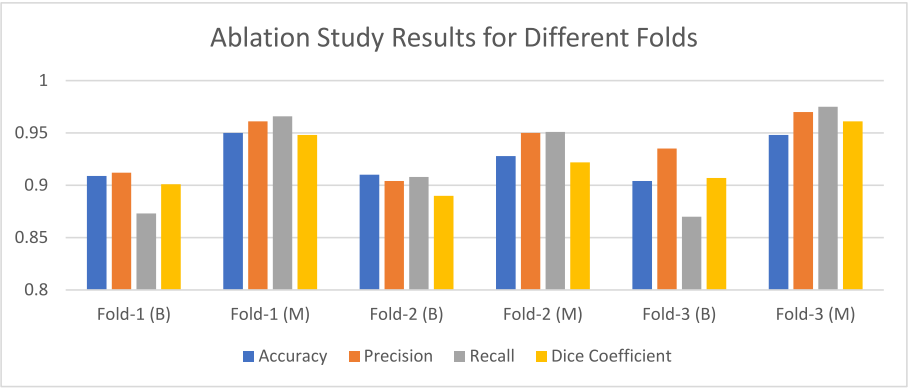

Fig. 9 Ablation Study Comparison Results for LIDC Dataset

Table 5 Ablation study to study the influence of different hyperparameters on the proposed UDCT Framework on LIDC Dataset

Batch Size	Optimizer			
	Adam		SGD	
	LR = 0.001	LR = 0.0001	LR = 0.001	LR = 0.0001
16	93.71%	92.2%	87.55%	89.64
32	94.21%	95.01%	89.90%	92.32
64	93.35%	94.00%	89.92%	92.35
128	93.87%	93.20%	91.23%	92.80

6 Conclusion

This paper presents a novel framework, UDCT for automatic detection of nodules on CT images. The proposed framework utilizes Multilevel Otsu thresholding for accurate segmentation of lung tumours from CT scans. Multilevel Otsu thresholding helps to separate lung tissue from the CT scans. Further, the proposed model incorporates U-Net network with robust multi-scale feature extraction capabilities. Also, DARTS-based classifier is incorporated which utilizes the concept of continuous relaxation of the architecture representation with efficient gradient descent search for efficient classification of lung cancer.

The proposed framework combines the knowledge of DARTS and U-Net with aims to improve lung nodule classification into benign and malignant classes. The proposed technique was evaluated and validated on the publicly available LIDC-IDRI dataset and IQ-OOTH/NCCD dataset and provided state-of-the-art results. The results reported provided an accuracy of 95.01%, a dice coefficient of 96.10%, and a recall rate of 97.52% on LIDC-IDRI dataset. Also, accuracy of 96.82%, a dice coefficient of 96.40%, and a recall rate of 97.50% on IQ-OOTH/NCCD dataset. Further, an ablation study is also conducted to validate the efficiency and efficacy of the proposed model. In future, we can extend our experiment on large scale dataset. Also, weak supervised learning algorithm can be incorporated that utilizes large number of incomplete, inaccurate, and ambiguous annotation data in the existing medical records to achieve efficient model training. Furthermore, bringing prior clinical knowledge into model training should be encouraged so as the model provides the best possible classification output.

Funding No funding was received to assist with the preparation of this manuscript.

Data availability The datasets generated during and/or analyzed during the current study are available through an online web repository via https://wiki.cancerimagingarchive.net/display/Public/LIDC-IDRI and https://www.kaggle.com/datasets/adityamahimkar/iqothnccd-lung-cancer-dataset.

Declarations

Conflict of interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Monkam P, Qi S, Ma H, Gao W, Yao Y, Qian W (2019) Detection and classification of pulmonary nodules using convolutional neural networks: a survey. IEEE Access 7:78075–78091. https://doi.org/10.1109/ACCESS.2019.2920980
2. Kumar V, Bakariya B (2021) Classification of malignant lung cancer using deep learning. J Med Eng Technol 45(2):85–93. https://doi.org/10.1080/03091902.2020.1853837
3. Fontana RS et al (1984) Early lung cancer detection: results of the initial (prevalence) radiologic and cytologic screening in the Mayo Clinic study. Am Rev Respir Dis 130(4):561–565. https://doi.org/10.1164/arrd.1984.130.4.549
4. Blandin Knight S, Crosbie PA, Balata H, Chudziak J, Hussell T, Dive C (2017) Progress and prospects of early detection in lung cancer. Open Biol 7(9):170070. https://doi.org/10.1098/rsob.170070
5. Inage T, Nakajima T, Yoshino I, Yasufuku K (2018) Early lung Cancer detection. Clin Chest Med 39(1):45–55. https://doi.org/10.1016/j.ccm.2017.10.003
6. D. S, B. R, M. T, P. P, D. T, and V. S (2022) Detection and Stage Classification of UNet Segmented Lung Nodules Using CNN," in 2022 5th International Conference on Multimedia, Signal Processing and Communication Technologies (IMPACT). 1–5. https://doi.org/10.1109/IMPACT55510.2022.10029182
7. Manser R et al (2013) Screening for lung cancer (review) screening for lung cancer. Cochrane Database Syst Rev (6). http://www.thecochranelibrary.com
8. Mohanapriya N, Kalaavathi B, senthil Kuamr T (2019) Lung Tumor Classification and Detection from CT Scan Images using Deep Convolutional Neural Networks (DCNN), in 2019 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE) pp. 800–805. https://doi.org/10.1109/ICCIKE47802.2019.9004247
9. Li G et al (2020) Study on the detection of pulmonary nodules in CT images based on deep learning. IEEE Access 8:67300–67309. https://doi.org/10.1109/ACCESS.2020.2984381
10. Ozdemir O, Russell RL, Berlin AA (2020) A 3D probabilistic deep learning system for detection and diagnosis of lung Cancer using low-dose CT scans. IEEE Trans Med Imaging 39(5):1419–1429. https://doi.org/10.1109/TMI.2019.2947595

11. Salama WM, Shokry A, Aly MH (2022) A generalized framework for lung Cancer classification based on deep generative models. Multimed Tools Appl 81(23):32705–32722. https://doi.org/10.1007/s11042-022-13005-9
12. Schneider J (2006) Tumor Markers in Detection of Lung Cancer pp. 1–41
13. Ahmed I, Chehri A, Jeon G, Piccialli F (2022) Automated pulmonary nodule classification and detection using deep learning architectures. IEEE/ACM Trans Comput Biol Bioinform:1–12. https://doi.org/10.1109/TCBB.2022.3192139
14. Nanglia P, Kumar S, Mahajan AN, Singh P, Rathee D (2021) A hybrid algorithm for lung cancer classification using SVM and neural networks. ICT Express 7(3):335–341. https://doi.org/10.1016/j.icte.2020.06.007
15. Liu H, Simonyan K, Yang Y (2018) DARTS: Differentiable Architecture Search. http://arxiv.org/abs/1806.09055
16. Gu H, Pan Z, Xi B, Asiago V, Musselman B, Raftery D (2011) Principal component directed partial least squares analysis for combining nuclear magnetic resonance and mass spectrometry data in metabolomics: application to the detection of breast cancer. Anal Chim Acta 686(1–2):57–63. https://doi.org/10.1016/j.aca.2010.11.040
17. Rautela K, Kumar D, Kumar V (2022) Dual-modality synthetic mammogram construction for breast lesion detection using U-DARTS. Biocybern Biomed Eng 42(3):1041–1050. https://doi.org/10.1016/j.bbe.2022.08.002
18. Nasir M, Farid MS, Suhail Z, Khan MH (2023) Optimal Thresholding for multi-window computed tomography (CT) to predict lung Cancer. Appl Sci 13(12):7256. https://doi.org/10.3390/app13127256
19. Tang S, Ma R, Li Q, Bai Y, Chen S (2021) Classification of benign and malignant pulmonary nodules based on the multiresolution 3D DPSECN model and Semisupervised clustering. IEEE Access 9:43397–43410. https://doi.org/10.1109/ACCESS.2021.3060178
20. Cao Y et al (2023) Segmentation of lung cancer-caused metastatic lesions in bone scan images using self-defined model with deep supervision. Biomed Signal Process Control 79:104068. https://doi.org/10.1016/j.bspc.2022.104068
21. Niranjan Kumar S et al. (2021) Lung Nodule Segmentation Using UNet, in 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS) 420–424. https://doi.org/10.1109/ICACCS51430.2021.9441977
22. Ibrahim DM, Elshennawy NM, Sarhan AM (2021) Deep-chest: multi-classification deep learning model for diagnosing COVID-19, pneumonia, and lung cancer chest diseases. Comput Biol Med 132:104348. https://doi.org/10.1016/j.compbiomed.2021.104348
23. Wu W, Gao L, Duan H, Huang G, Ye X, Nie S (2020) Segmentation of pulmonary nodules in CT images based on 3D-UNET combined with three-dimensional conditional random field optimization. Med Phys 47(9):4054–4063. https://doi.org/10.1002/mp.14248
24. Yu H, Li J, Zhang L, Cao Y, Yu X, Sun J (2021) Design of lung nodules segmentation and recognition algorithm based on deep learning. BMC Bioinform 22(S5):314. https://doi.org/10.1186/s12859-021-04234-0
25. Banu SF, Sarker MMK, Abdel-Nasser M, Puig D, Raswan HA (2021) AWEU-net: an attention-aware weight excitation U-net for lung nodule segmentation. Appl Sci 11(21):10132. https://doi.org/10.3390/app112110132
26. Bruntha PM, Dhanasekar S, Ahmed LJ, Govindaraj V, Pandian SIA, Abraham SS (2023) "Lung_RUNET: a Segmentation Framework for Lung Nodules," in 2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS) pp. 658–661. https://doi.org/10.1109/ICACCS57279.2023.10113093
27. Bruntha PM, Pandian SIA, Sagayam KM, Bandopadhyay S, Pomplun M, Dang H (2022) Lung_PAYNet: a pyramidal attention based deep learning network for lung nodule segmentation. Sci Rep 12(1):20330. https://doi.org/10.1038/s41598-022-24900-4
28. Bushara AR, Vinod Kumar RS, Kumar SS (2023) LCD-capsule network for the detection and classification of lung cancer on computed tomography images. Multimed Tools Appl 82(24):37573–37592
29. Shafi Imran et al. (2022) An effective method for lung cancer diagnosis from ct scan using deep learning-based support vector network. Cancers 14(21):5457
30. Spoorthi B, Mahesh S (2023) Firefly competitive swarm optimization based hierarchical attention network for lung cancer detection. Int J Image Graph 23(02):2350017
31. Rather SA, Das S (2023) Levy flight and chaos theory-based gravitational search algorithm for image segmentation. Mathematics 11(18):3913
32. Al-Huseiny MS, Sajit AS (2021) Transfer learning with GoogLeNet for detection of lung cancer. Indones J Electr Eng Comput Sci 22(2):1078–1086
33. Mathews AB (2022) Lung Cancer Segmentation and Classification Using Integration of Convolutional Neural Network & Unet Network Over CT Images: A Deep Learning Approach. Int J Manag Technol Soc Sci 7(1):520–534

34. Prabhu, Akshatha. "Lung Cancer Recognition: A Fusion Approach"
35. Al-Yasriy Hamdalla F et al. (2020) Diagnosis of lung cancer based on CT scans using CNN. IOP Conf Ser Mater Sci Eng. Vol. 928. No. 2. IOP Publishing
36. Gowda S, Jayachandran A (2022) Triple SVM integrated with enhanced random region segmentation for classification of lung tumors. Int J Adv Comput Sci Appl 13(10)
37. Santos DF (2023) Tackling Lung Cancer: Advanced Image Analysis and Deep Learning for Early Detection. Authorea Preprints
38. Kareem HF et al (2021) Evaluation of SVM performance in the detection of lung cancer in marked CT scan dataset. Indones J Electr Eng Comput Sci 21(3):1731