Received: March 14, 2024. Revised: May 29, 2024.

INASS

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

Detection of Lung Cancer Using Multi-Stage Image Processing and Advanced Deep Learning InceptiMultiLayer-Net Model

Syed Zaheer Ahammed[1,2] Radhika Baskar[3]* G Nalinipriya[4]

[1]*Department of Computational Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, India*
[2]*JNTUA College of Engineering Kalikiri, Andhra Pradesh, India*
[3]*Department of Electrical and Computer Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, India*
[4]*Department of IT, Saveetha Engineering College, Chennai, Tamilnadu, India*
* Corresponding author's Email: radhikabaskar@saveetha.com

Abstract: This study aims to improve early lung cancer detection by creating a sophisticated Computer-Aided Diagnosis (CAD) system. This system employs advanced image processing techniques such as adaptive dynamic histogram equalization (ADHE), Local Binary Pattern (LBP), and Tsallis thresholding to effectively reduce noise, analyze textures, and segment regions. It also includes the InceptiMultiLayer-Net (IML-Net), an advanced version of the Inception V3 architecture designed to capture complex features in medical images. The IML-Net includes a multiclass Error-Correcting Output Codes (ECOC) Support Vector Machine (SVM) classifier, which improves the system's ability to handle complex classification tasks. The system also employs statistical features such as mean, variance, energy, entropy, and correlation to fully describe the characteristics of segmented regions. With an impressive 99.573% accuracy in identifying lung cancer-affected regions, as well as a sensitivity of 99.46% and a specificity of 99.24%, this CAD system has significant potential as an early lung cancer detection tool. These findings highlight the system's ability to assist clinicians in making accurate diagnoses, ultimately improving patient outcomes significantly.

Keywords: Lung cancer, ADHE, Computer-aided diagnosis, SVM, Feature extraction, Image pre-processing, Segmentation, Texture analysis, CAD system.

1. Introduction

Lung cancer presents significant challenges to early detection, frequently resulting in delayed diagnoses and less effective treatment options. Late-stage lung cancer diagnosis is associated with severe health complications and a high mortality rate, so developing early detection techniques is critical to improving patient prognosis and increasing survival rates [1, 2].

Computed Tomography (CT) scans have become indispensable in the lung cancer diagnostic process, providing detailed information about the tumor's structure, size, and location. These scans are especially effective at detecting early-stage lung cancer, which is critical for initiating treatment and improving patient outcomes [3-5]. A newly proposed method aims to improve this diagnostic process by reducing the reliance on manual interpretation, allowing radiologists to make more precise and informed decisions, potentially saving lives [6].

The study describes a novel multistage methodology that combines advanced image processing techniques with cutting-edge deep learning (DL) models. This approach is based on the InceptiMultiLayer-Net (IML-Net), an advanced version of the Inception V3 architecture designed specifically to capture complex features in medical images. The IML-Net includes a multiclass Error-Correcting Output Codes (ECOC) Support Vector Machine (SVM) classifier, which improves the

model's ability to handle complex classification tasks. This integration significantly improves the accuracy and reliability of lung cancer detection, distinguishing it from other techniques.

The proposed method's technical superiority stems from its advanced integration of adaptive image processing and deep learning techniques to improve early lung cancer detection. The InceptiMultiLayer-Net (IML-Net) uses adaptive dynamic histogram equalization (ADHE) to improve contrast and visibility in CT scans, while Local Binary Pattern (LBP) and Tsallis thresholding refine texture analysis and region segmentation, respectively. This increases the granularity with which lung tissue can be examined, allowing for the detection of subtler abnormalities that would otherwise be missed by traditional methods. Furthermore, the analysis includes statistical features such as mean, variance, energy, entropy, and correlation, which provides a comprehensive quantitative assessment of the segmented regions and significantly enriches the diagnostic data available to clinicians. This multifaceted approach not only improves the system's accuracy and sensitivity, but also gives it the ability to handle a wide range of imaging scenarios, establishing a new standard for the precision and reliability of lung cancer diagnosis.

Section 1 emphasizes the importance of early detection and research objectives. Section 2 reviews recent studies, which serve as the basis for the proposed methodology. Section 3 describes the approach, which includes a block diagram, mathematical equations, and a new Deep Neural Network (DNN) model. Section 4 summarizes the research findings by calculating accuracy and sensitivity values and comparing existing methods. Section 5 concludes by emphasizing key findings, discussing implications, acknowledging limitations, and outlining potential next steps. This structured organization ensures coherence throughout the problem introduction, solution proposal, result analysis, and conclusion.

2. Literature survey

Nazir et al. [6] proposed a multiresolution rigid registration approach for the LIDC-IDRI dataset, which resulted in commendable accuracy and sensitivity rates of 98.2% and 96.4% in detection tasks, respectively. This method uses a hierarchical approach to align medical images at multiple scales, which improves the accuracy of detecting abnormalities in lung images. However, one significant limitation of this method is its reliance on the quality and consistency of the input data, which can vary between datasets, potentially affecting its performance in real-world applications.

Bushara et al. [7] proposed the Visual Geometry Group-Capsule Network (VGG-CapsNet), a novel fusion of the VGG and Capsule Network architectures. When tested on the LIDC-IDRI datasets, this innovative approach achieved an impressive 98.61% accuracy. VGG-CapsNet, which combines the strengths of the VGG and Capsule Network models, provides a promising solution for image classification tasks, particularly in the field of medical image analysis. However, one major limitation of this method is its computational complexity, which may impede its scalability and practical application.

Vijay et al. [8] proposed a hybrid classifier called the hybrid differential evolution (HDE) -based neural network, which combines differential evolution principles and neural network techniques. This novel approach achieved impressive performance metrics, including an accuracy of 96.39% and a sensitivity of 95.25 percent. By incorporating differential evolution into the neural network framework, the model can effectively optimize its parameters and improve its classification performance, especially on complex and high-dimensional datasets. A significant limitation of this method is its sensitivity to parameter selection and population size, which may necessitate fine-tuning for optimal performance across diverse datasets.

Balachandran et al. [9] proposed a context-aware attention UNET architecture that uses attention mechanisms to improve image contextualization. This novel approach achieved an impressive sensitivity of 99.15%, demonstrating its ability to accurately identify relevant features within medical images. By incorporating attention mechanisms into the UNET framework, the model can dynamically focus on informative regions, resulting in better segmentation performance, especially in areas with subtle distinctions or complex structures. One significant limitation of this method is its computational demand, as the addition of attention mechanisms may increase the model's complexity and resource requirements.

Wang et al. [10] proposed a multiple-scale residual network (MResNet), which uses residual units to improve the model's depth and accuracy. This method achieved impressive accuracy and sensitivity rates of 85.23% and 92.79%, respectively, demonstrating its ability to capture intricate patterns across different scales within the data. MResNet uses residual units to efficiently learn complex representations, resulting in improved classification and detection capabilities, especially in tasks

involving large and diverse datasets. A major limitation of this method is its susceptibility to vanishing gradients and optimization challenges, particularly when dealing with extremely deep architectures, which may impede its convergence and generalization performance on more complex datasets or tasks with limited training data.

Ramana et al. [11] proposed using Saliency capsule networks to detect lung cancer, with remarkable accuracy and sensitivity rates of 98.95% and 98.85%. This method uses capsule networks to capture salient features in medical images, allowing for precise identification of abnormalities that may indicate lung cancer. By incorporating Saliency capsule networks, the model can more effectively analyze the hierarchical relationships between image components, improving its ability to detect subtle patterns associated with malignancy. However, one significant limitation of this approach is its reliance on annotated training data, as obtaining accurately labeled medical images for model training can be time-consuming and potentially biased.

Mundher et al. [12] presented a novel Progressive Growing Channel Attentive Non-Local (ProCAN) network designed for lung cancer detection and classification. This novel technique achieved a remarkable accuracy rate of 95.28%. By incorporating non-local operations and channel attention mechanisms, ProCAN allows for the effective modeling of long-range dependencies as well as selective feature enhancement, increasing the model's discriminative power in identifying lung cancer-related abnormalities. This approach's computational complexity, especially with the inclusion of non-local operations, may increase training and inference time, making it unsuitable for real-time applications or deployment in resource-constrained settings. Furthermore, the performance of ProCAN may be sensitive to hyperparameters and dataset characteristics, necessitating careful optimization and validation to ensure consistent performance across different datasets and clinical scenarios.

Saihood et al. [13] proposed a hybrid approach to lung cancer detection that combines the water strider optimization algorithm (WSA) with OTSU thresholding and pre-processing with the Gray-Level Co-occurrence Matrix (GLCM). This fusion mode strategy aims to improve the accuracy of cancer detection in two-dimensional images. The method yielded promising results, with 94.4% accuracy and 91.6% sensitivity. By combining WSA for optimization, OTSU thresholding for segmentation, and GLCM for feature extraction, the model can effectively capture relevant image characteristics associated with lung cancer. A significant limitation of this method is its reliance on manual parameter tuning, especially for the WSA algorithm, which may necessitate extensive experimentation and domain knowledge. Furthermore, the approach's performance may vary depending on image quality and acquisition protocols. The proposed hybrid approach's limitation, particularly regarding the OTSU thresholding method, is its sensitivity to image noise and variability in image characteristics. OTSU thresholding assumes that the image has bimodal intensity distributions, which may not be true in all cases.

Halder et al. [14] introduced a cutting-edge framework called the 2-Pathway Morphology-based Convolutional Neural Network (2PMorphCNN), which is specifically designed for lung cancer detection. This novel approach demonstrated exceptional performance, with an accuracy of 96.10% and a sensitivity of 96.85%. By incorporating adaptive morphology techniques into the convolutional neural network architecture, the 2PMorphCNN effectively exploits the structural characteristics of medical images, allowing for more accurate feature extraction and classification. This method's primary limitation is its computational demand, especially when performing morphological operations. This computational overhead may result in increased processing time and resource requirements, potentially limiting scalability and practical application. Furthermore, the performance of the adaptive morphology approach may be influenced by the choice of appropriate structuring elements and parameter configurations, necessitating careful calibration and validation to ensure consistent performance across diverse datasets and clinical settings.

Aside from the works mentioned above, several notable methods have been developed to improve the prediction and classification of lung cancer outcomes. Aonpong et al. [15] proposed a novel methodology for predicting recurrence in non-small cell lung cancer patients by combining genotype information with radiomic signatures. While their approach shows promise for improving predictive accuracy, the developed signatures' generalizability across diverse patient populations is a potential limitation. Shen et al. [16] developed an interpretable deep hierarchical semantic convolutional neural network to enhance the interpretability of deep learning systems for lung nodule malignancy classification. However, the effectiveness of such models may be determined by the availability of large annotated training datasets. Choi et al. [17] also developed a deep learning model for predicting visceral pleural

invasion in lung cancer using CT scans, which performed similarly to experienced radiologists. However, rigorous validation across multiple patient cohorts is required to ensure the reliability and generalizability of this approach in clinical practice.

Whereas existing methods for lung cancer detection and classification have significant limitations such as computational complexity, sensitivity to parameter tuning, and reliance on annotated training data, the proposed method provides a compelling solution to these problems. The proposed framework addresses the shortcomings of traditional approaches by seamlessly integrating image processing techniques with advanced deep learning models, specifically the InceptiMultiLayer-Net (IML-Net) architecture and a multiclass Error-Correcting Output Codes (ECOC) SVM classifier. This comprehensive approach not only improves computational efficiency, but it also increases the model's ability to detect and classify lung cancer-related abnormalities. Furthermore, the use of the ECOC SVM classifier ensures that the model is versatile and adaptable across a wide range of datasets and clinical scenarios. Thus, the proposed method emerges as a significant advancement in the field, providing a strong and dependable tool for accurate lung cancer diagnosis and prognosis.

3. Methodology

Figure 1 depicts the proposed Computer-Aided Diagnosis (CAD) system. Initially, the system acquires medical images from the LIDC-IDRI dataset [18]. The Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI) [18] have been valuable resources for lung cancer diagnosis and research. LIDC-IDRI is a freely available database of thoracic CT scans with annotations from multiple radiologists. This dataset has helped to develop and validate various algorithms and methodologies for lung nodule detection, classification, and characterization. LIDC-IDRI has been used by researchers to train and evaluate machine learning and deep learning models for tasks like nodule segmentation, malignancy prediction, and treatment response assessment. LIDC-IDRI has made significant contributions to lung cancer imaging research by providing a standardized dataset with expert annotations, allowing for the development of more accurate and reliable computer-aided diagnostic systems.

Later the step converts the images to grayscale using a formula that mimics human colour perception. It improves image quality even further by using a Wiener filter to reduce noise and adaptive dynamic

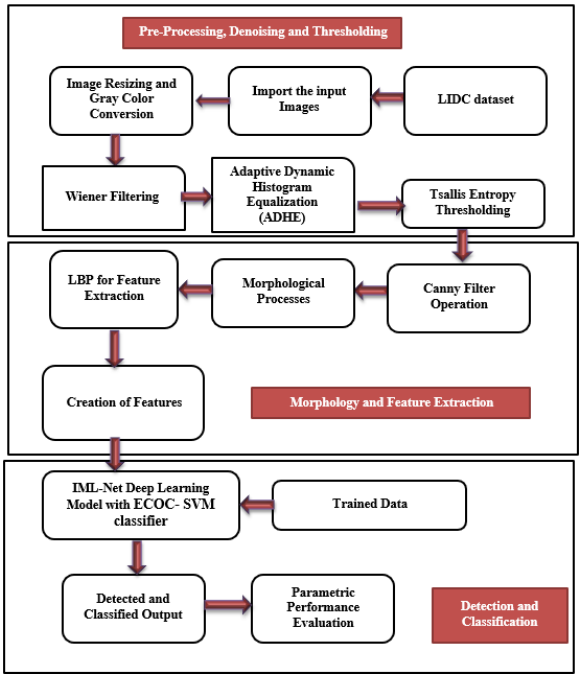

Figure. 1 Proposed system block diagram

histogram equalisation (ADHE) to improve contrast. Tsallis thresholding is then used to isolate regions of interest in images for effective image segmentation.

Moving on to the second stage, the CAD system refines the edge image obtained after Canny edge detection using morphological operations such as dilation and filling. LBPs are used to extract texture features, allowing for the characterization of local texture patterns within images. The CAD system then employs InceptiMultiLayer-Net (IML-Net), a DL model derived from the Inception V3 architecture, in the final stage. IML-Net incorporates a multiclass Error-Correcting Output Codes (ECOC) SVM classifier for handling complex classification tasks and multiple layers to capture intricate image features.

The entire process with necessary equations and elaboration is given below.

Obtain the LIDC-IDRI dataset, which is a collection of lung cancer-related medical images. This database contained approximately 250 lung CT scans, 80 % for training and 20 % for testing. To improve the visual representation of the images, convert the acquired RGB images to the appropriate color space, i.e. Grayscale. This pixel's grayscale value (L) can be calculated using the formula below.

$$L = 0.2989 \times R + 0.5870 \times G + 0.1140 \times B \quad (1)$$

The coefficients (0.2989, 0.5870, and 0.1140) represent the human perception of color intensity,

with green receiving the most weight, followed by red, and blue receiving the least.

We introduce the use of a Wiener filter to address the problem of noise reduction and smoothing irregularities or artefacts in colour images. Let I(x, y) represent the intensity value of a pixel at coordinates (x, y) within the grayscale image, and M(x, y) represent the final image after median filtering. The median filter works on the following principles:

$$M(x,y) = Wiener\big(I(x+i, y+j)\big) for\ i = -k\ to\ k, j = -k\ to\ k \quad (2)$$

Let's apply the outcome of equation 2 to Adaptive Dynamic Histogram Equalization (ADHE). Let H(L) represent the

local histogram for intensity level L in a neighborhood of size N. The local histogram can be calculated as follows:

$$H(L) = \Sigma\Sigma\, w(i,j) \times \delta\big(L - I(x+i, y+j)\big), for\ i = -\frac{N}{2} to \frac{N}{2}, j = -\frac{N}{2} to \frac{N}{2} \quad (3)$$

In equation 3, w(i, j) represents the weighting function for pixel (i, j), and δ denotes the Kronecker delta function, The CDF represents the cumulative sum of histogram values up to the given intensity level.

$$CDF(L) = \Sigma\,\Sigma\, w(i,j) \times \sum H(K), for\ K = 0\ to\ L, i = -\frac{N}{2} to \frac{N}{2}, j = -\frac{N}{2} to \frac{N}{2} \quad (4)$$

In equation 4, ∑ H(K) represents the cumulative sum of local histogram values up to intensity level K.Compute the new pixel value based on the local CDF and the original pixel value

$$M(x,y) = CDF\big(I(x,y)\big) \times (L_{\max} - L_{min}) + L_{min} \quad (5)$$

Here, $L_{\max}$ and L_{min} represent the maximum and minimum intensity levels

Apply the above steps for each pixel in the grayscale image to obtain the final ADHE-enhanced image. Use the Tsallis thresholding technique to segment regions of interest based on gray-level intensities in the pre-processed images. The objects of interest are separated from the background in this step. Tsallis thresholding is a statistical method for dividing an image into different regions based on gray-level distribution.

$$Hq = \frac{(1 - \Sigma[p(i)^q])}{(q-1)} \quad (6)$$

Here p(i) denotes the Probability based function for intensities of local Neighborhood.

Tsallis entropy values for different threshold values (T) within a specified range can be used to calculate the optimal threshold. The threshold T is chosen to maximise the Tsallis entropy Hq:

$$T_{optimal} = argmax(T) Hq(T) \quad (7)$$

For each pixel (x, y) in the pre-processed grayscale image

$$BW(x,y) = 1, if\ I(x,y) >= T_{optimal} \quad (8)$$

$$BW(x,y) = 0, if\ I(x,y) < T_{optimal} \quad (9)$$

The Sobel operator calculates the gradients Gx and Gy using convolution kernels for horizontal and vertical gradients, respectively.

$$Gx(x,y) = \sum_{i=-1}^{1}\sum_{j=-1}^{1} I(x+i, y+j) \cdot Sx(i,j) \quad (10)$$

$$Gy(x,y) = \sum_{i=-1}^{1}\sum_{j=-1}^{1} I(x+i, y+j) \cdot Sy(i,j) \quad (11)$$

Where Sx and Sy are the Sobel kernels for horizontal and vertical gradients.

The gradient magnitude (G) is computed as the Euclidean norm of the horizontal and vertical gradients.

$$G(x,y) = \sqrt{Gx^2(x,y) + Gy^2\ (x,y)} \quad (12)$$

For each pixel (x, y) in the gradient magnitude image:

Calculate the gradient direction (θ):

$$\boldsymbol{\theta(x,y)} = arctan(\tfrac{Gx(x,y)}{Gy(x,y)}) \quad (13)$$

Compare the gradient magnitude of the current pixel with its neighbors in the direction perpendicular to the gradient (θ). If it's a local maximum, retain it; otherwise, suppress it. The goal of edge tracking is to connect edge pixels to form continuous edge contours. We can use a connectivity-based approach to link strong edge pixels to potential edge pixels if they

share a common boundary to perform edge tracking by hysteresis. After edge tracking, we obtain a binary edge image where edge pixels are connected to form continuous edge contours. The binary image represents the detected edges in the original image.

Apply morphological image processing operations like dilation and filling on the final edge image obtained after Canny edge detection.

The dilation operation can be defined as:

$$DilatedImage(x,y) = \max\big(EdgeImage(x+i, y+j)\big) \quad (14)$$

for all (i,j) in the structuring element

Filling, also known as hole filling or morphological closing, is used to fill small gaps or holes within the connected edge regions.

$$FilledImage(x,y) = \min\big(DilatedImage(x+i, y+j)\big) \quad (15)$$

for all (i,j) in the structuring element

FilledImage will contain the filled and connected edge contours, allowing for further analysis or visualisation.

The LBP value for a pixel (x, y) can be calculated as follows.

Create a circular neighborhood of P evenly spaced points on a circle of radius R. The neighborhood points are denoted as (x_p, y_p), where p = 0, 1, ..., P-1. For each neighbor (x_p, y_p), calculate the difference between the intensity value of the central pixel (x, y) and the neighbor's intensity value:

$$D_p = ProcessedImage(x_p, y_p) - ProcessedImage(x,y) \quad (16)$$

Encode the differences as binary values by thresholding them:

If D_p >= 0, make the corresponding bit in the binary pattern to 1; otherwise, make it to value of 0.

for all neighbors to form an LBP pattern

$$LBP(x,y) = (B(P-1)B(P-2).B(1)B(0)) \quad (17)$$

Optionally, convert the binary LBP pattern to its decimal equivalent for compact representation if needed. Repeat the process for each pixel in the ProcessedImage to obtain the LBPImage. Once we have computed the LBP values for each pixel in the image, now calculate texture features like contrast, energy, and entropy. Contrast measures the local variations in pixel intensities and is calculated as the

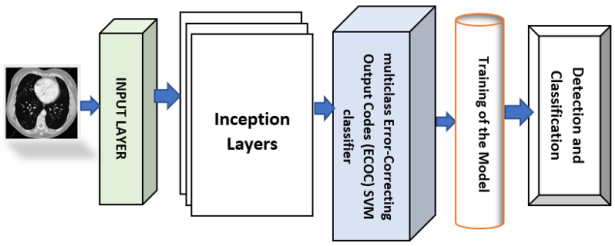

Figure. 2 InceptiMultiLayer Model

variance of the LBP values within a local neighborhood. For each pixel (x, y):

Calculate the local mean (μ) of LBP values within a neighborhood of size N

$$\mu(x,y) = \frac{\big(\Sigma\Sigma\ LBP(x+i, y+j)\big)}{(N \times N)}, for\ i = -\frac{N}{2} to \frac{N}{2}, j = -\frac{N}{2} to \frac{N}{2} \quad (18)$$

Calculate the contrast (C) as the variance of LBP values within the neighborhood:

$$C(x,y) = \frac{\Big(\Sigma\Sigma\ \big(LBP(x+i, y+j) - \mu(x,y)\big)^2\Big)}{(N \times N)}, for\ i = -\frac{N}{2} to \frac{N}{2}, j = -\frac{N}{2} to \frac{N}{2} \quad (19)$$

Calculate the energy (E) as the sum of squared LBP values within a neighborhood:

$$E(x,y) = \Sigma\Sigma\ \big(LBP(x+i, y+j)\big)^2, for\ i = -\frac{N}{2} to \frac{N}{2}, j = -\frac{N}{2} to \frac{N}{2} \quad (20)$$

Entropy quantifies the randomness or uncertainty of LBP values within a local neighborhood and is calculated as the Shannon entropy. For each pixel (x, y). Calculate the entropy (H) as

$$H(x,y) = -\Sigma\Sigma \begin{pmatrix} p\big(LBP(x+i, y+j)\big) \times \\ log2\Big(p\big(LBP(x+i, y+j)\big)\Big) \end{pmatrix} for\ i = -\frac{N}{2} to \frac{N}{2}, j = -\frac{N}{2} to \frac{N}{2} \quad (21)$$

where p(LBP) is the probability mass function of LBP values within the neighborhood

InceptiMultiLayer: The term "Incepti" is derived from Inception V3, the base architecture used in this model. "MultiLayer" reflects the modification involving multiple layers in the network as shown in Fig. 2.

Adding more convolutional layers, fully connected layers, or custom layers to the base architecture can be represented mathematically as:

$$Output = Activation(W \times Input + b) \quad (22)$$

Where Output: The output of the layer, Activation: The activation function (e.g., ReLU, sigmoid) applied element-wise, W: weight matrix, Input: input to the layer, b: bias vector

$$Output(i,j,k) = Average(Input(i \times PoolingSize: (i+1) \times PoolingSize, j \times PoolingSize: (j+1) \times PoolingSize, k)) \quad (23)$$

Where Output: Pooled output, Input: Input feature map, Pooling Size: Size of the pooling region.

IML-Net is a novel advancement in deep learning architecture, building on the foundation of Inception V3 with an increased number of layers to capture intricate features in complex datasets. This extended architecture not only improves feature representation, but it also includes a cutting-edge multiclass Error-Correcting Output Codes (ECOC) SVM classifier, which increases its ability to handle complex classification tasks. Furthermore, IML-Net incorporates the strategic use of average pooling as the output layer, a novel approach known for its effectiveness in image classification. This combination of architectural enhancements establishes IML-Net as a versatile and robust model, representing a significant step forward in addressing the challenges of sophisticated feature extraction and classification across domains.

The algorithm below describes the creation of a novel deep learning architecture called InceptiMultiLayer (IML-Net) for multi-class classification tasks, with a emphasis on improving feature extraction capabilities. In Step 1, the InceptiMultiLayer class is defined, which consists of a base model initialized with an InceptionV3 network and additional layers designed to improve feature representation. Step 2 introduces the ECOC_SVM class, which implements a multiclass Error-Correcting Output Codes (ECOC) Support Vector Machine (SVM) classifier, allowing for robust classification across multiple classes. Next, in Step 3, average pooling is used as the output layer to facilitate feature aggregation. In the main program, Steps 4 and 5 involve instantiating and training the IML-Net and ECOC_SVM models for classification tasks, respectively. Finally, in Step 6, average pooling is used to compute the model's final output. This algorithmic framework combines deep learning and traditional machine learning techniques to produce a comprehensive model capable of efficient and accurate multi-class classification, with a focus on improving feature representation using the InceptiMultiLayer architecture.

Algorithm: IML-Net

Step 1: Define InceptiMultiLayer with enhanced features

Step 1.1: Initialize InceptiMultiLayer: Let obj represent an instance of InceptiMultiLayer.

$$obj = InceptiMultiLayer(x, y) \quad (24)$$

Step 1.2: Create Base Model: Let $base\_model$ denote the base model created using InceptionV3 with input shape x.

$$base_{model} = InceptionV3(x) \quad (25)$$

Set $obj.base\_model$ to $base\_model$

Step 1.3: Add Additional Layers: Let A be the additional convolutional layer with a 3×3 kernel and 64 filters.

$$A = ConvolutionLayer(3,64) \quad (26)$$

Update $obj.base\_model$ by adding A to $base\_model$.

Step 1.4: Set Number of Classes: Let C represent the number of classes.

Step 1.5: Return InceptiMultiLayer Object: Return obj.

Step 2: Implement multiclass ECOC SVM classifier

Step 2.1: Initialize ECOC_SVM:
Let obj denote an instance of ECOC_SVM.

$$obj = ECOC\_SVM \quad (27)$$

Step 2.2: Train SVM Classifiers:
For each class i from 1 to C:
Train an SVM classifier svm_model_i using features and labels for class i.

$$svm_{modeli} = fitcecoc(features, labels == i) \quad (28)$$

Store svm_model_i in $obj.svm\_models[i]$

Step 2.3: Make Predictions: For each class i from 1 to C:

Predict class probabilities predictions$i$ using the trained SVM classifier svm_model$i$ and input data.

$$predictions i = predict(svm\_model i, input\_data) \quad (29)$$

Store predictions$i$ in the i th column of the predictions matrix.

Step 2.4: Return Predictions:

Return the predictions matrix.

Step 3: Implement average pooling as the output layer

Step 3.1: Perform Average Pooling: Let input_data represent the input feature matrix.

Calculate the mean of input_data along the spatial dimensions to obtain the output.

$$output = mean(input\_data) \quad (30)$$

Step 4: Instantiate InceptiMultiLayer (IML-Net): Initialize IML_Net with input shape x and C.

Step 5: Instantiate ECOC_SVM:

Initialize ECOC_SVMECOC_SVM with C.

Step 6: Train the Model: Train the SVM classifiers.

Step 7: Make Predictions:

Use the trained SVM classifiers to make predictions.

Step 8: Implement Average Pooling: Apply average pooling to the predictions to obtain the final output

The accuracy , sensitivity and Specificity are provided by following equations

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (24)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (25)$$

$$Specificity = \frac{TN}{TN + FP} \quad (26)$$

4. Results and discussion

The original CT scan image serves as the foundation for the entire process (see Fig. 3). This diagram depicts the initial input obtained from the LIDC-IDRI database [5], which served as the foundation for subsequent analysis.

Figure 4 discusses A fundamental preprocessing step is resizing the original image.

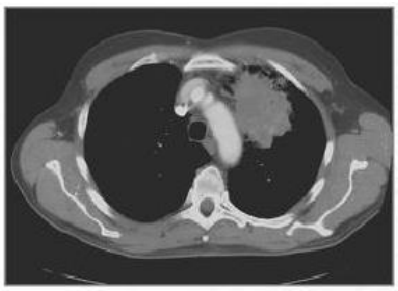

Figure. 3 Lung Cancer CT Image

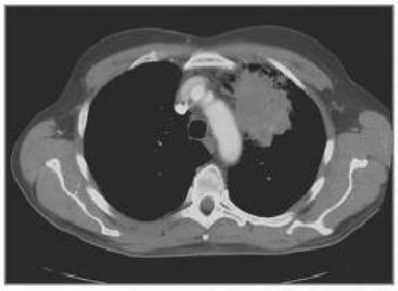

Figure. 4 Resized Image

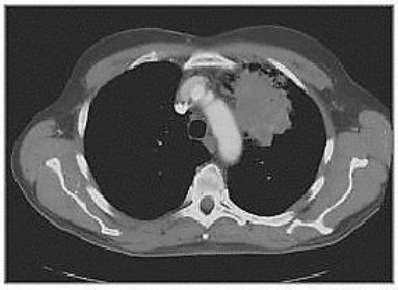

Figure. 5 Gray Color Conversion

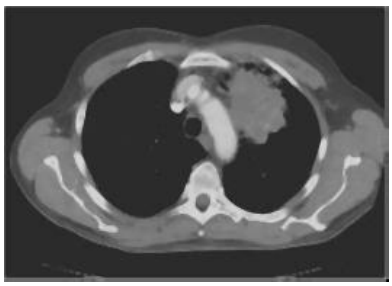

Figure. 6 Filtered Image using Wiener Filter

Figure. 7 Output of ADHE enhancement technique

Figure. 8 Thresholded image Using Tsallis Method

Grayscale conversion simplifies the image into a single-channel representation, as shown in Fig. 5. Figure 6 is the result of Weiner Filter. The use of a Wiener filter is required for noise reduction and image enhancement.

The image in Fig. 7 is the result of Adaptive Dynamic Histogram Equalisation (ADHE), which significantly improves image contrast and reveals finer details. Figure 8 shows how this Tsallis statistical method effectively separates objects of interest from the background, isolating potentially cancerous regions for further investigation. The Canny edge detection filter, depicted in Fig. 9, plays an important role in highlighting edges in the image.

Figure 10 is the end result of Dilation, as shown in Fig. 10, is an important morphological operation. This step is critical for obtaining accurate feature extraction. Filling, or morphological closing, as demonstrated in Fig. 11, is used to close gaps within connected edge regions.

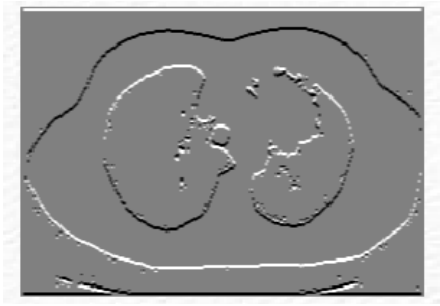

Figure. 9 Filtered image via Canny Filter

Figure. 10 Dilation Processed Image

Figure. 11 Outcome of Filled Image Process

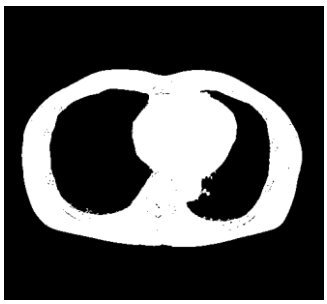

Figure. 12 Resultant Image of LBP

In a nutshell Figure 12 illustrates the final image that was obtained through the extraction of the Local Binary Patterns (LBP) feature. As a result of the fact that they describe local texture patterns, LBP values are an essential component of texture analysis.

Table 1 provides an in-depth look at the feature extraction process used on lung cancer image samples using the Local Binary Patterns (LBP) method. The table specifically provides results for essential features extracted from five sample images, such as entropy, contrast, and energy. These characteristics aid in characterising and quantifying textural patterns within images, which is useful for subsequent stages of analysis. The image samples that were utilized in this process of feature extraction were sourced from the Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI). This large database contains approximately 300 lung computed tomography (CT) scans, each of which represents a distinct clinical case.

Table 1. Extracted Features using LBP

	Sample Lung Cancer Images				
Features	**Sample Image 1**	**Sample Image 2**	**Sample Image 3**	**Sample Image 4**	**Sample Image 5**
Entropy	0.7532	0.7763	0.7821	0.7684	0.7342
Contrast	2.929	2.965	2.934	2.972	2.965
Energy	0.5923	0.5912	0.5917	0.5921	0.5917

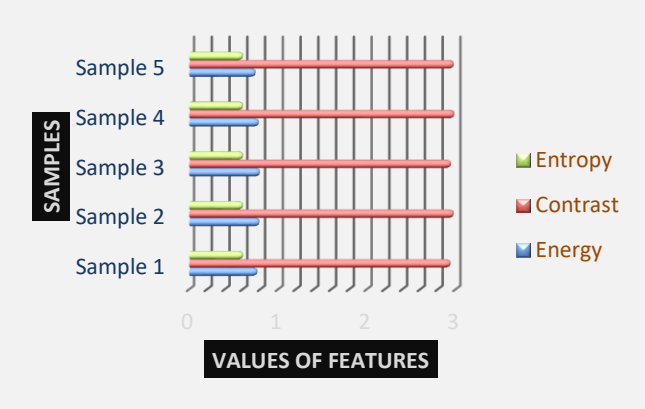

Figure. 13 Feature Plot

Table 2. Assessment of Accuracy Values

Techniques used	**Accuracy (%)**
multiresolution rigid [6]	98.2
VGG-CapsNet [7]	98.61
HDE [8]	96.39
MResNet [10]	85.23
Saliency capsule networks [11]	98.95
ProCAN [12]	95.28
WSA [13]	94.4
2PMorphCNN [14]	96.10
Proposed Method (IML-Net)	**99.573**

For model development and evaluation, an 80/20 data split was used in accordance with established best practises. Specifically, 80% of the dataset, or 240 lung CT scans, was used to train the model. This training phase served two purposes: fine-tuning the model's parameters to improve performance and uncovering latent patterns and relationships in the data to build a lung cancer predictive model. The remaining 20% of the dataset, corresponding to 60 lung CT scans, was set aside solely for assessing the trained model's performance.

Figure 13, titled "Feature Plot," depicts the extracted energy, contrast, and entropy values from lung cancer image samples.

Table 2 compares the accuracy values achieved by various techniques for lung cancer detection on the LIDC dataset. Notably, methods such as VGG-CapsNet and Saliency Capsule Networks achieve 98.61% and 98.95% accuracies, respectively, by combining deep learning architectures with novel

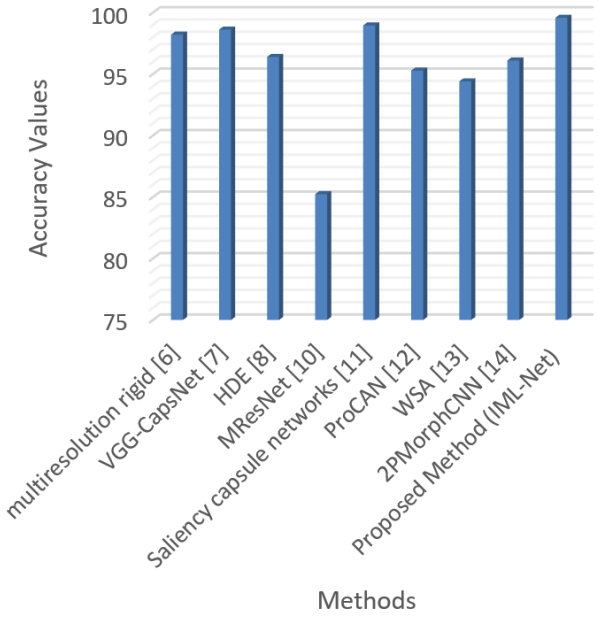

Figure. 14 Comparison plot of Accuracy

features such as capsule networks and saliency mechanisms. Meanwhile, the proposed method, IML-Net, outperforms all others with an accuracy of 99.573%, indicating the potential for interpretability-driven machine learning approaches. While techniques such as Multiresolution Rigid and HDE perform well with accuracies above 96%, MResNet falls behind with 85.23%, indicating varying degrees of effectiveness in combining different methodologies for lung cancer detection. These findings highlight advances in medical image analysis, demonstrating the potential of deep learning and novel techniques to improve detection accuracy for critical conditions such as lung cancer.

Figure 14 shows a comparison plot of accuracy values for various methods used in lung cancer detection, with methods represented on the x-axis and corresponding accuracy values in percentages plotted on the y-axis. The plot depicts the performance disparities between the techniques evaluated.

Notably, methods such as VGG-CapsNet, Saliency Capsule Networks, and the proposed IML-Net achieve remarkably high accuracies of over 98%. In contrast, MResNet stands out with a significantly lower accuracy of 85.23%, indicating potential limitations in its application for lung cancer detection.The plot effectively depicts the variability in performance across different methodologies, emphasizing the importance of using the right techniques to achieve maximum accuracy in medical image analysis tasks. Furthermore, the plot emphasizes the importance of novel approaches,

Table 3. Assessment for Sensitivity values

Techniques used	Sensitivity (%)
multiresolution rigid [6]	96.4
HDE [8]	95.25
context-aware attention UNET [9]	99.15
MResNet [10]	92.79
Saliency capsule networks [11]	98.85
WSA [13]	91.6
2PMorphCNN [14]	96.85
Proposed Method	**99.46**

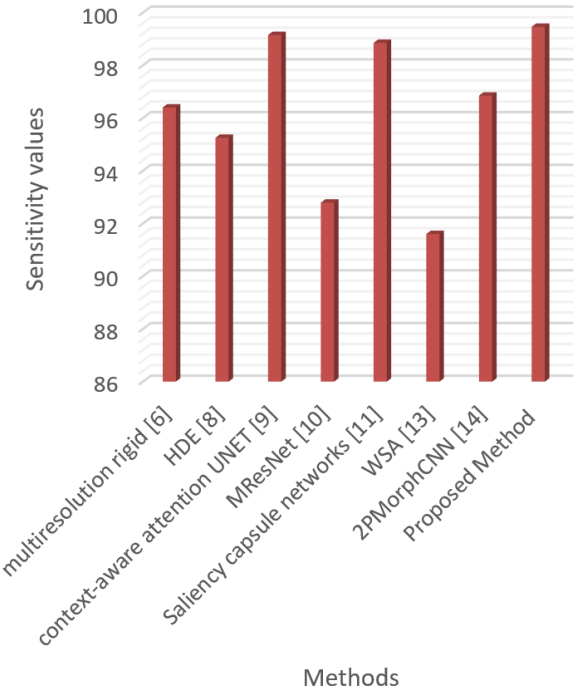

Figure. 15 Comparison Plot for Sensitivity

as evidenced by the superior performance of the proposed IML-Net, indicating its potential for practical implementation in clinical settings.

Table 3 shows a comparative analysis of sensitivity values for various techniques used in lung cancer detection using the LIDC dataset. Among these methods, the proposed approach stands out with a sensitivity of 99.46%, indicating that it can correctly identify cancerous regions in lung scans. This outperforms several other techniques, including context-aware attention UNET, demonstrating the efficacy of the proposed method in accurately detecting lung cancer. Its exceptionally high sensitivity highlights its potential as an advanced tool for precise cancer detection, emphasizing the importance of incorporating novel methodologies to improve diagnostic capabilities in clinical practice.

Figure 15 depicts a comparison plot of sensitivity values for various methods used in lung cancer detection, with methods represented on the x-axis and corresponding sensitivity values plotted on the y-axis. The plot depicts the efficacy of different techniques in correctly identifying cancerous regions in lung scans. Notably, the proposed method outperforms others, with sensitivity of 99.46%, as shown in Table 3. This high sensitivity indicates that the proposed method excels at accurately detecting cancerous regions, making it a promising candidate for clinical use. Furthermore, while other methods, such as context-aware attention UNET and Saliency Capsule Networks, exhibit high sensitivities, they fall short of the proposed method's capabilities. In contrast, techniques such as MResNet and WSA have lower sensitivity, indicating potential limitations in accurately identifying cancerous regions. As a result, the plot reinforces the proposed method's superiority in terms of sensitivity, highlighting its potential to improve diagnostic accuracy and patient outcomes in lung cancer detection.

The table 4 provides a comprehensive assessment of the specificity values for various techniques used in a given domain. Specificity, an important metric in fields such as medical imaging and pattern recognition, indicates a model's ability to correctly identify true negatives from all actual negatives. Across the techniques listed, we see a variety of specificity values, with notable performance differences. Contextualized attention UNET and WSA stand out with high specificity values of 98.81% and 99%, demonstrating their ability to accurately classify negative instances. However, MResNet has a significantly lower specificity of 72.89%, indicating potential limitations in its ability to detect true negatives effectively.

Table 4. Assessment for Specificity values

Techniques used	Specificity (%)
multiresolution rigid [6]	97.2
HDE [8]	96.12
context-aware attention UNET [9]	98.81
MResNet [10]	72.89
WSA [13]	99
2PMorphCNN [14]	95.17
Proposed Method	**99.24**

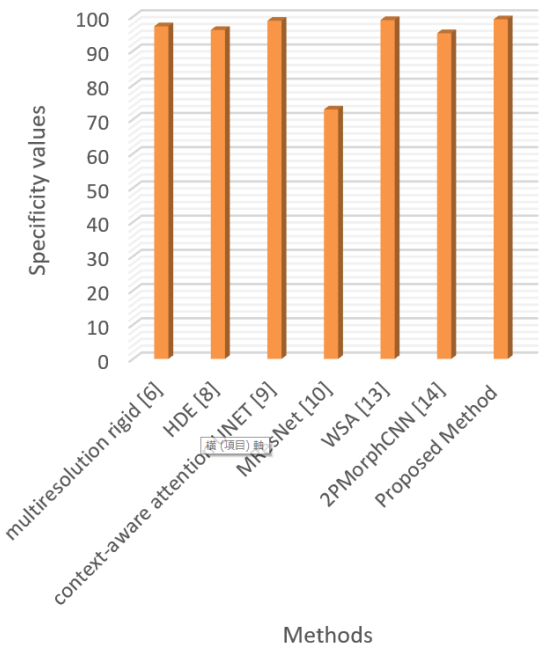

Figure. 16 Comparison Plot for Specificity

Table 5. Assessment for AUC

Techniques used	AUC (%)
HDE [8]	96.05
MResNet [10]	92.75
WSA [13]	93.1
2PMorphCNN [14]	99.36
Proposed Method	**99.76**

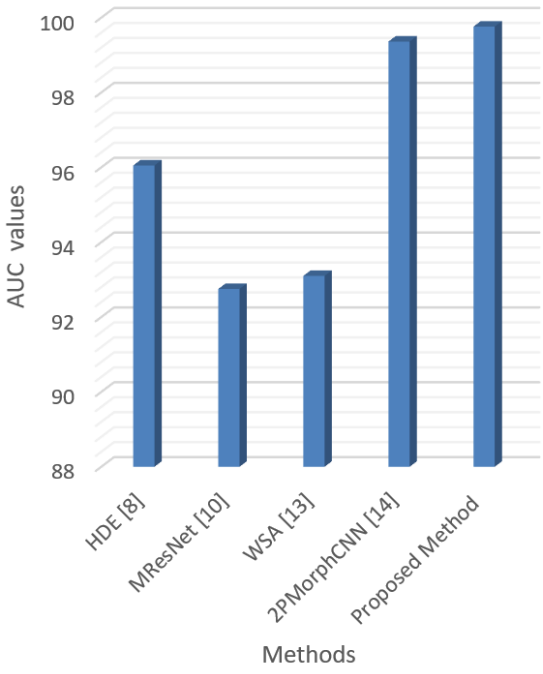

Figure. 17 Comparison Plot for AUC

The proposed method emerges as the table's frontrunner, with an impressive specificity of 99.24%, indicating superiority in accurately identifying true negatives when compared to the other techniques tested.

Moving on to the comparison plot in Fig. 16, we get a visual representation of how each technique performs in terms of specificity. The plot provides a clear overview of how each method performs in comparison to one another, making it easy to compare and identify trends. Techniques such as WSA and the proposed method emerge as top performers, displaying notably high specificity values and demonstrating their robustness in accurately classifying negative instances. In contrast, MResNet stands out as an outlier, with a lower specificity value than the rest, indicating areas for improvement. The plot is a useful tool for researchers and practitioners because it helps them identify effective methodologies for tasks that require precise negative instance classification. Overall, the table and plot provide detailed insights into the specificity performance of various techniques, with the proposed method demonstrating superior accuracy and reliability in detecting true negatives.

The evaluation for the Area Under the Curve (AUC) attained by different methods is presented succinctly in table 5. The name of each technique is given in the column "Techniques used," and the matching AUC value in percentage is shown in the column "AUC (%)".

Performance in the task assessed is shown by listing HDE [8], MResNet [10], WSA [13], and 2PMorphCNN [14] together with their corresponding AUC percentages. The proposed method is shown in the last row; it outperformed the other methods listed with an AUC of 99.76%.

The AUC values attained by each method are graphically shown in Fig. 17. Usually, each technique is shown as a line or a data point, which facilitates quick comparison of their effectiveness. Most likely, the various techniques are indicated by the x-axis, and the matching AUC values are shown by the y-axis.

One can see the relative performance of several approaches visually thanks to this plot. As seen by its higher position on the y-axis than the other data points or lines representing alternative techniques, the proposed method clearly outperforms the other techniques.

The table and figure taken together give a thorough summary of the AUC evaluation and provide both numerical and visual information about how various approaches performed on the assessed task.

5. Conclusion

Finally, this study presents a novel and comprehensive CAD system designed to detect lung cancer in medical images early. To effectively preprocess and segment medical images, the proposed methodology combines advanced image processing techniques such as adaptive dynamic histogram equalization (ADHE), Local Binary Pattern (LBP) feature extraction, and Tsallis thresholding. InceptiMultiLayer-Net (IML-Net), a deep learning model, provides a powerful framework for precise classification. It expands on the Inception V3 architecture by including a multiclass Error-Correcting Output Codes (ECOC) SVM classifier, which improves its ability to handle complex classification tasks. Our findings show that the CAD system performs admirably, with an accuracy of 99.573% in identifying lung cancer-affected regions with specificity of 99.24 %. It also has a high sensitivity of 99.46%.

This advancements in deep learning in future entail the development of even more sophisticated CAD systems capable of not only detecting lung cancer but also providing information about its stage and progression. Furthermore, the integration of real-time data and telemedicine applications may allow for remote monitoring and consultation, increasing the accessibility and effectiveness of lung cancer diagnosis and treatment. Technology advancement and its application in healthcare have the potential to significantly improve our ability to combat lung cancer and improve patient outcomes.

Notation List
I(x, y) represent the intensity value of a pixel at coordinates (x, y) within the grayscale image
M(x, y) represent the final image after median filtering.
H(L) represent the local histogram for intensity level L in a neighborhood of size N
w(i, j) represents the weighting function for pixel (i, j)
δ denotes the Kronecker delta function
CDF represents the cumulative sum of histogram values up to the given intensity level.
∑ H(K) represents the cumulative sum of local histogram values up to intensity level K.
$L_{\max}$ and L_{min} represent the maximum and minimum intensity levels
p(i) denotes the Probability based function for intensities of local Neighborhood
Gx and Gy : horizontal and vertical gradients
Sx and Sy are the Sobel kernels for horizontal and vertical gradients.
R: circle of radius
μ: local mean
C: Contrast
E: Energy
H: Entropy
p(LBP) is the probability mass function of LBP values within the neighborhood

Conflicts of Interest

The authors affirm that they are aware of no personal or financial conflicts of interest that might have affected the research described in this paper.

Author Contributions

Syed Zaheer Ahammed-Conceptualization, Data curation Formal analysis,

Radhika Baskar-Methodology, Software Writing, original draft, investigation, resources.

G.Nalinipriya-Writing—review and editing, visualization, supervision, project administration.

References

[1] M.G. Rojo, "State of the art and trends for digital pathology", *Studies in Health Technology and Informatics*, Vol. 179, pp. 15-28, 2012.

[2] A. Nibali, Z. He, D. Wollersheim, "Pulmonary nodule classification with deep residual networks", *International Journal of Computer Assisted Radiology and Surgery*, Vol. 12, No. 10, pp. 1799–1808, 2017.

[3] G. Tong, Y. Li, H. Chen, Q. Zhang, H. Jiang, "Improved U-NET network for pulmonary nodules segmentation", *Optik*, Vol. 174, pp. 460–469, 2018.

[4] D.J. Heineman, M.G. Ten Berge, J.M. Daniels, M.I. Versteegh, P.J. Marang-van de Mheen, M.W. Wouters, W.H. Schreurs, "The quality of staging non-small cell lung cancer in the Netherlands: data from the Dutch lung surgery audit", *The Annals of Thoracic Surgery*, Vol. 102, No. 5, pp. 1622–1629, 2016.

[5] S.G. Armato, G. McLennan, L. Bidaut, M.F. McNitt-Gray, C.R. Meyer, A.P. Reeves, C.I. Henschke, "Data from LIDC-IDRI", 2015, Available online: https://wiki.cancerimagingarchive.net/display/Public/LIDC-IDRI (accessed on 5 October 2022).

[6] I. Nazir, I. Haq, S. Alqahtani, M. Jadoon, M. Dahshan, "Machine Learning-Based Lung Cancer Detection Using Multiview Image

Registration and Fusion", *Journal of Sensors*, pp. 1-19., 2023.

[7] A. Bushara, R. Kumar, S. Kumar, "An ensemble method for the detection and classification of lung cancer using Computed Tomography images utilizing a capsule network with Visual Geometry Group", *Biomedical Signal Processing and Control*, Vol. 85, p. 104930, 2023.

[8] V.K. Gugulothu, S. Balaji, "An early prediction and classification of lung nodule diagnosis on CT images based on hybrid deep learning techniques", *Multimedia Tools and Applications*, Vol. 83, pp. 1041–1061, 2024.

[9] . Balachandran, V. Ranganathan, "Semantic context-aware attention UNET for lung cancer segmentation and classification", *International Journal of Imaging Systems and Technology*, Vol. 33, 2022.

[10] H. Wang, H. Zhu, L. Ding, et al., "A diagnostic classification of lung nodules using multiple-scale residual network", *Scientific Reports*, Vol. 13, p. 11322, 2023.

[11] K. Ramana, M.R. Kumar, K. Sreenivasulu, T.R. Gadekallu, S. Bhatia, P. Agarwal, S.M. Idrees, "Early Prediction of Lung Cancers Using Deep Saliency Capsule and Pre-Trained Deep Learning Frameworks", *Frontiers in Oncology*, Vol. 12, p. 886739, 2022.

[12] M. Al-Shabi, K. Shak, M. Tan, "ProCAN: Progressive growing channel attentive non-local network for lung nodule classification", *Pattern Recognition*, Vol. 122, C, Feb 2022.

[13] A. Saihood, H. Karshenas, A.R.N. Nilchi, "Deep fusion of gray level co-occurrence matrices for lung nodule classification", *PLoS ONE*, Vol. 17, No. 9, e0274516, 2022.

[14] A. Halder, S. Chatterjee, D. Dey, "Adaptive morphology aided 2-pathway convolutional neural network for lung nodule classification", *Biomedical Signal Processing and Control*, Vol. 72, 103347, 2022.

[15] P. Aonpong, Y. Iwamoto, X.-H. Han, L. Lin, Y.-W. Chen, "Genotype-Guided Radiomics Signatures for Recurrence Prediction of Non-Small Cell Lung Cancer", *IEEE Access*, Vol. 9, pp. 90244-90254, 2021.

[16] S. Shen, S.X. Han, D.R. Aberle, A.A. Bui, W. Hsu, "An interpretable deep hierarchical semantic convolutional neural network for lung nodule malignancy classification", *Expert Systems with Applications*, Vol. 128, pp. 84–95, 2019.

[17] H. Choi, H. Kim, W. Hong, J. Park, E.J. Hwang, C.M. Park, Y.T. Kim, J.M. Goo, "Prediction of visceral pleural invasion in lung cancer on CT: deep learning model achieves a radiologist-level performance with adaptive sensitivity and specificity to clinical needs", *European Radiology*, Vol. 31, No. 5, pp. 2866-2876, May 2021.

[18] S.G. Armato III, G. McLennan, L. Bidaut, M.F. McNitt-Gray, C.R. Meyer, A.P. Reeves, B. Zhao, D.R. Aberle, C.I. Henschke, E.A. Hoffman, E.A. Kazerooni, H. MacMahon, E.J. Van Beeke, D. Yankelevitz, A.M. Biancardi, P.H. Bland, M.S. Brown, R.M. Engelmann, G.E. Laderach, D. Max, R.C. Pais, D.P. Qing, R.Y. Roberts, A.R. Smith, A. Starkey, P. Batrah, P. Caligiuri, A. Farooqi, G.W. Gladish, C.M. Jude, R.F. Munden, I. Petkovska, L.E. Quint, L.H. Schwartz, B. Sundaram, L.E. Dodd, C. Fenimore, D. Gur, N. Petrick, J. Freymann, J. Kirby, B. Hughes, A.V. Casteele, S. Gupte, M. Sallamm, M.D. Heath, M.H. Kuhn, E. Dharaiya, R. Burns, D.S. Fryd, M. Salganicoff, V. Anand, U. Shreter, S. Vastagh, B.Y. Croft, "The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A completed reference database of lung nodules on CT scans", *Medical Physics*, Vol. 38, pp. 915–931, 2011.