Engineering Journal issn: 0125-8281

Predicting Lung Cancer Using ResNet-50 Deep Residual Learning Model using Relu-Memristor Activation Function

Sneha Khoria[a] **and Raghuraj Singh**[b,*]

Department of Computer Science & Engineering, Harcourt Butler Technical University, Kanpur, India
E-mail: [a]snehakhoria1@gmail.com, [b,*]raghurajsingh@rediffmail.com (Corresponding author)

Abstract. Lung cancer continues to stand as a prominent contributor to deaths caused due to cancer throughout the world, underscoring the paramount importance of precise and timely detection methodologies. In this paper, we introduce a pioneering strategy aimed at forecasting lung cancer by utilizing the power of deep residual learning, coupled with a meticulous examination of medical images at the pixel level. The given methodology uses convolutional neural networks (CNNs), specifically the Residual Network (ResNet) architecture, to extract intricate features from lung images, while also emphasizing the importance of raw pixel values as input features. Our experimental results showcase the promise of this novel approach, demonstrating remarkable predictive performance in lung cancer detection. By incorporating pixel values and deep residual learning, our model achieves a high degree of accuracy, sensitivity, and specificity, surpassing existing methodologies. This research makes important contribution in the advancement of early detection of lung cancer, ultimately enhancing the recovery chances of patient and potentially minimizing the burden of this devastating disease.

Keywords: Lung cancer, ResNet-50, CNN, accuracy.

ENGINEERING JOURNAL Volume 29 Issue 3
Received 20 August 2024
Accepted 5 March 2025
Published 31 March 2025
Online at https://engj.org/
DOI:10.4186/ej.2025.29.3.45

DOI:10.4186/ej.2025.29.3.45

1. Introduction

Lung cancer stands as a major challenge worldwide for public health due to its significant contribution to cancer-related fatalities year after year. The sheer magnitude of its impact cannot be overstated, as it leads to the loss of countless lives and imparts a heavy burden on healthcare systems worldwide [1]. In this context, the urgency to enhance our strategies for early detection and prediction of lung cancer becomes palpable, given that these twin pillars hold the potential to be transformative in improving patient outcomes and curtailing the mortality rates associated with this devastating disease [2].

Traditional methods for predicting lung cancer have primarily relied on statistical models and machine learning (ML) techniques [3]. While these models have provided useful information, they can struggle when dealing with the complexity of medical data [3]. Their ability to make accurate predictions can be affected by the intricate patterns and subtle relationships present in the data. Moreover, although statistical and machine learning models are valuable, they sometimes lack the structured approach needed for situations where accuracy and reliability are critical.

Recently, a promising breakthrough has emerged with the integration of medical imaging and machine learning (ML), creating a powerful synergy with vast potential [4]. This combination has produced a set of tools that greatly assist clinicians in making precise diagnoses and accurately categorizing lung tumours [5]. Central to this development is the innovative fusion of medical imaging with advanced ML algorithms, where each image becomes a detailed source of valuable insights into the underlying pathology [6]. One of the most crucial technologies driving this progress is computed tomography (CT) scans, which have become a cornerstone of modern medical imaging.

While human observation can detect some nuances in medical images, it is the power of machine learning algorithms that truly unlocks the potential of these extensive datasets, uncovering valuable insights. These algorithms are capable of thoroughly examining the images, identifying hidden patterns and features that might escape the attention of human experts. However, a crucial step in this process is the precise categorization of individual pixels within the image. By classifying the pixels as either "black" or "white," the algorithm effectively separates the image into regions of interest and background. This pixel-level classification forms the foundation for the subsequent stages of lung tumour diagnosis and classification [7].

Deep learning (DL) methods, particularly convolutional neural networks (CNNs), are very helpful in lung cancer detection [8]. These methods utilize large datasets of medical images, typically CT scans or X-rays, and employ deep neural networks to automatically identify lung nodules or tumors. By learning intricate patterns and features within these images, CNNs can provide highly accurate and efficient diagnoses. This technology holds great promise in revolutionizing lung cancer detection, potentially leading to earlier diagnoses, more effective interventions, and improved patient outcomes [9].

1.1. Motivation

To detect lung cancer accurately, and pattern or subtle anomaly shall be identified within these medical images. Traditional image processing methods often struggle to capture complex relationships, while manual segmentation is labour-intensive and subject to variability. ResNet-50, have shown remarkable capabilities in recognizing intricate patterns and relationships, making them well-suited for the task of pixel classification. By training ResNet-50 on large datasets of labelled images, we aim to leverage their potential to distinguish between cancerous and healthy lung tissue at a pixel level.

1.2. Objectives

This research seeks to achieve the following objectives:

1. Develop a robust classification framework using ResNet-50 for lung cancer detection in CT scans.
2. Design a ResNet-50architecture optimized for pixel-based classification, accounting for factors such as image resolution, complexity, and available computational resources.
3. Evaluation of the effectiveness of the ResNet-50model through quantitative metrics such as accuracy, sensitivity, specificity.
4. Compare the results of the ResNet-50-based approach with conventional image processing methods and manual segmentation techniques.

1.3. Organization

The rest of the research article is organized as follows: Section 2 given an idea of the state-of-the-art method for lung cancer classification. In the Section 3, the proposed methodology for pixel-based classification based on black and white calculations using ResNet-50. Section 4 discusses the results obtained from our approach and compares them to conventional methods. Section 5 summarizes the contributions and future research possibilities.

2. Literature Survey

In recent years, lung nodule classification has developed significantly. This is basically due to various DL methodologies [10-12]. These efforts have yielded highly promising results, marking substantial advancements in accuracy.

One noteworthy approach involves the utilization of Scale-Invariant Feature Transform (SIFT)-based techniques and Support Vector Machines (SVM), which have demonstrated impressive sensitivities and specificities. For instance, some classifiers achieved sensitivities as high as 86% and specificities as impressive

as 97% [13], while others reached values of 91.38% and 89.56% [10], respectively. Multi-scale CNN and multi-crop CNN methods have also contributed significantly, achieving accuracies in the range of 87-90.63%, sensitivities of 77-92.30%, and specificities of 89-93% [14], respectively. Moreover, deep-level semantic networks and multi-scale CNN models have showcased their capabilities with accuracy rates reaching 84.2% [15] and 86.84% [16], respectively.

In the year 2019, Zhang and colleagues (Zhang et al. [17]) introduced a pioneering ensemble learning approach, which harnessed the strength of multiple deep CNNs for the comprehensive analysis of CT images. Their groundbreaking effort yielded a noteworthy classification accuracy of 84%. Importantly, their proposed classifier outperformed other algorithms in the field, including SVM, multi-layer perceptron's, and random forests. This marked a significant leap forward in the quest for accurate lung nodule detection and classification. Another research was of Sahu et al. [18], who devised a lightweight multi-section CNN architecture. This innovation not only achieved a commendable classification accuracy of 93.2% but also streamlined the process of selecting representative cross-sections, a pivotal aspect in malignancy assessment. This development thus not only enhanced accuracy but also facilitated more effective clinical decision-making. Likewise, Ali et al. [19] made strides in the field by utilizing a transferable texture CNN, characterized by nine layers, designed for feature extraction and lung nodule classification achieving accuracy rates of 96.69% ± 0.72%, reflecting minimal errors and high recall rates. This impressive performance underscores the potential of deep learning techniques for highly precise nodule identification. Furthermore, Marques et al. [20] introduced a multi-task CNN tailored for the classification of malignancy in nodules, achieving an impressive Area Under the Curve (AUC) of 0.783. The automatic lung nodule classifier introduced by Thamilarasi et al. [21] was based on CNN and it secured an accuracy of 86.67%. Kawathekar et al. [22] took a ML approach to develop a lung nodule classifier, resulting in an accuracy of 94%. These advancements demonstrated the diversification of techniques employed for lung nodule classification, each contributing to the evolving landscape of accurate diagnosis and prognosis. Moreover, the continuous evolution of deep learning techniques has pushed the boundaries of lung nodule detection. Radford et al. [23] introduced the deep convolution Generative Adversarial Network (DCGAN), while Chuquicusma et al. [24] applied DCGAN for generating realistic lung nodules. These innovations opened new possibilities for data augmentation and enhanced training. Finally, Zhao et al. [25] made notable strides by utilizing Forward and Backward Generative Adversarial Network (F&BGAN) for lung nodule classification, achieving an accuracy of 95.24%. These remarkable results emphasize the remarkable benefits of using advanced DL techniques in achieving highly accurate and reliable lung nodule classifications.

Table 1. Summary of the Notable Methods.

Reference	Method	Accuracy
[10]	SIFT + SVM	Sensitivity: 91.38% Specificity: 89.56%
[13]	SIFT + SVM	Sensitivity: 86% Specificity: 97%
[14]	CNN	Accuracy: 87-90.63%, Sensitivity: 77-92.30%, Specificity: 89-93%
[15]	DSN	Accuracy: 84.2%
[16]	CNN	Accuracy: 86.84%
[17]	CNNs	Accuracy: 84%
[18]	CNN	Accuracy: 93.18%
[19]	CNN	Accuracy: 96.69%
[20]	CNN	AUC: 0.783
[21]	CNN	Accuracy: 86.67%
[22]	ML	Accuracy: 94%
[25]	F&BGAN	Accuracy: 95.24%

Table 1 provides a clear summary of different methods employed in lung nodule classification and their corresponding performance metrics.

In summary, recent advancements in both traditional and DL-based methods have proved to be quite effective in the classification of lung nodules. In the similar direction, this method is used for lung cancer prediction and detection method based on pixel intensity and ResNet-50 model.

3. Proposed Method

CT is very useful in diagnosing and monitoring lung diseases, including lung cancer. The recorded CT images are saved in a database.

As the first step, (Fig. 1) pre-processing is performed on the dataset to convert images are in a suitable format. This step includes normalization, noise removal, resizing, and other necessary operations to upgrade the dataset quality. After pre-processing, lung segmentation is carried out to isolate the lung region from the background. This segmentation step is crucial as it focuses the analysis on the relevant area of the image, i.e., the lungs.

Next, the segmented pixels of the lung image are classified into two categories: black and white pixels, based on intensity thresholds. White pixels generally represent areas of interest, such as lung nodules, while black pixels represent the background or non-relevant areas. If the number of white pixels exceeds a certain threshold (Th), it indicates the presence of potential lung nodules. Once the white pixels are identified, further evaluation is performed to classify them as benign or malignant.

DOI:10.4186/ej.2025.29.3.45

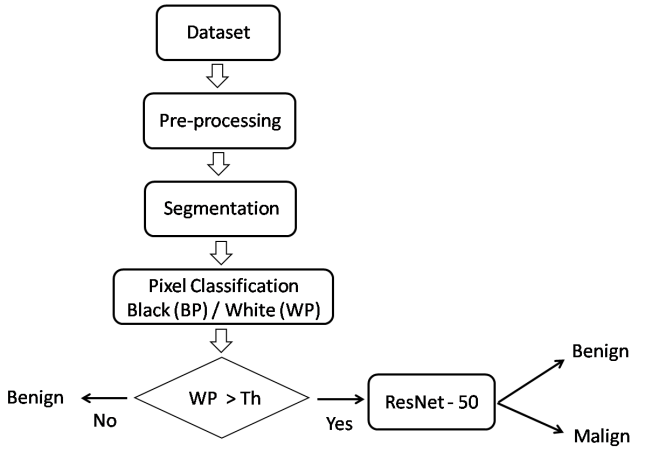


Fig. 1. Block diagram of the proposed lung cancer prediction method.

For the classification step, the system utilizes ResNet-50, an advanced version of CNNs. ResNet-50, with its 50 layers, captures complex features from the images, making it ideal for medical image classification tasks. It classifies the pixels based on learned patterns into two categories: benign (non-cancerous) and malignant (cancerous) nodules. ResNet-50 is an improvement over traditional CNNs because it incorporates residual connections that allow the network to effectively train even with a deep architecture, overcoming issues like vanishing gradients.

Figure 2 illustrates a more detailed diagram of the process, highlighting the crucial stages of the CNN.

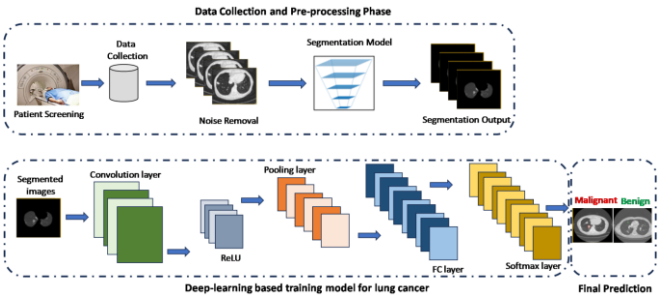


Fig. 2: Detailed diagram of the proposed lung cancer prediction method

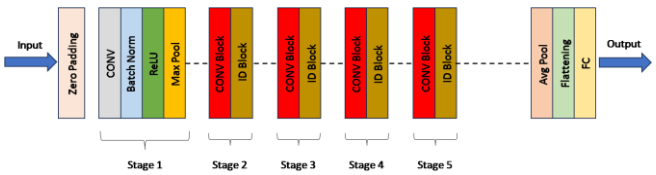


Fig. 3. Detailed diagram of ResNet 50 architecture.

Finally, Fig. 3 provides an in-depth view of the ResNet-50 architecture, including all its layers and residual connections. This detailed diagram demonstrates how the network processes input images through its various layers using the residual connections in order to minimize the gradient descent problem and enable effective training of deep networks. ResNet-50 allows it to learn hierarchical features at multiple scales which makes it quite fruitful in classification tasks such as lung nodule detection. The detail description of each process is detailed as follows:

3.1. Pre-Processing

Pre-processing lung images is a critical initial step in lung cancer detection. The primary aim of this step is to enhance image quality, reduce noise, and make the data suitable for subsequent deep learning analysis. To achieve this, a series of essential preprocessing steps are employed. These steps are resizing the images to a uniform resolution, normalizing pixel values for consistency, and applying contrast enhancement techniques to improve the visibility of lung structures. Noise reduction filters are used to eliminate the impact of noise, ensuring that the images are of the highest quality. Additionally, thresholding methods are applied to lung regions from the background and isolate specific structures of interest. Moreover, the extraction of the region of interest (ROI) containing the lungs and cropping to remove unnecessary borders or artifacts help streamline the analysis process. Histogram analysis and artifact removal are further employed to refine the data, while image registration techniques are used to align multiple images or modalities accurately. After preprocessing the lung image segmentation is performed as detailed in the nest sub-section.

3.2. Lung Image Segmentation

In CT scans, the lungs typically appear as darker regions due to the lower density of air within them, while the lighter areas inside the lungs usually represent blood vessels, airways, or other structures. The goal of lung segmentation is to extract the precise boundaries of the lung regions from each slice of the CT scan. This is crucial for various diagnostic purposes, such as detecting abnormalities or assessing lung diseases. The process of segmenting the lung regions demands great precision and attention to detail to ensure that no important structures, particularly those closely adjacent to the lung wall, are missed or incorrectly segmented. Even minor errors in segmentation can lead to misinterpretations of the scan and potentially affect clinical outcomes.

To achieve a high level of accuracy in lung region extraction, a multi-step approach is utilized. As outlined in [26], a sequence of six meticulously designed steps is followed to obtain the masked lung regions. These steps involve sophisticated techniques to identify, isolate, and refine the boundaries of the lungs while maintaining the integrity of surrounding structures. The careful execution of each of these steps ensure that the segmented lung regions are both precise and reliable. The steps are as follows:

Lung Image Segmentation Steps

Step 1: The initial stage involves transforming the image (I) into binary format, accomplished through the application of a thresholding technique defined by:

$$Binary\ I(i,j) = \begin{cases} 1 & \text{if } I(i,j) > T \\ 0 & \text{otherwise} \end{cases}$$

The threshold (T) is evaluated using Otsu method [27]. This threshold is applied for the binarization of the image.

Step 2: Eliminating the connected blobs at the border of the CT image is essential for accurate image classification, ensuring that these border-related regions do not interfere with the process.

Step 3: Labelling the image involves identifying connected regions where pixel neighbourhoods share the same intensity level. When this procedure is extended to the entire image, it results in the formation of various connected regions.

Step 4: The labels associated with the two most extensive regions, signifying both lungs, are retained, while tissues smaller than the expected lung size are eliminated.

Step 5: Executing the erosion operation, utilizing a disk with a radius of 2, is carried out in this step to distinguish pulmonary nodules connected to the lung wall from the blood vessels.

Step 6: Utilizing the closure operation, employing a disk with a radius of 10 as described in reference [28], serves the purpose of preserving the connection of nodules to the lung wall. This operation effectively eliminates small dark spots from the image while connecting small bright gaps.

In the subsequent phase, we undertake a pixel count of black and white pixels within the lung image. If the count of white pixels (WP) surpasses a predefined threshold (Th), the lung image is flagged as suspicious, warranting further in-depth investigation. This approach brings about dual advantages:

1. This approach enhances efficiency by ensuring that only relevant images, those with a higher concentration of white pixels beyond a certain threshold, are processed further in the classification task. This selective approach minimizes unnecessary computational overhead associated with processing all images, making the workflow more efficient.
2. Secondly, it takes into consideration that many lung images may contain smaller benign dots that do not require intricate classification. By excluding these benign dot-related images from the classification task, we can conserve computational resources and reduce false positives.

However, it is important to exercise caution when setting the threshold. It should be kept relatively low to avoid the possibility of missing any malignant nodules. Striking the right balance is crucial to ensure that the classification process remains sensitive enough to detect potentially harmful anomalies while still benefiting from the efficiency gains of this selective approach.

3.3. Feature Selection and Classification using ResNet-50

Using a ResNet-50 model for lung cancer detection involves employing CNN architecture to analyzes medical images, typically chest X-rays or CT scans, to identify potential signs of lung cancer. Here are the typical steps and details involved in using ResNet-50 for lung cancer detection:

Convolution Operation

In the realm of CNNs, the concept of convolution stands as a fundamental and pivotal mathematical operation. Its significance lies in its ability to transform raw input data, typically images, into a format that is more amenable to feature extraction and subsequent analysis. At its core, convolution involves the application of a set of learnable filters, often referred to as kernels, to the input image. These kernels serve as templates that slide across the input data, pixel by pixel, and perform a series of multiplicative operations. The essence of this process is to systematically scan the input image, emphasizing local patterns and structures. As each kernel traverses the image, it computes weighted sums of pixel values within its receptive field. This operation has a dual purpose: firstly, it facilitates the detection of relevant features, patterns, or edges within the image, and secondly, it helps in creating feature maps that represent these detected features. By learning the optimal kernel weights during the training process, CNNs can automatically identify and extract meaningful features, regardless of their position in the input data. Mathematically, the convolution operation can be represented as follows: Output feature map (O) can be written as

$$O = I * K + b \tag{1}$$

where, I is input map, the * represents the convolution operation, and b is an bias term. K is the kernel (also referred to as a filter), which is applied to the input I to extract features.

The Activation Function

An activation function in the context of ANN is a mathematical function applied to the output of a neuron or a node in a neural network. The primary purpose of an

activation function is to introduce non-linearity into the neural network.

Rectified Linear Unit (ReLU)

The ReLU, activation function is a critical component in artificial neural networks, including deep learning models like CNNs. It serves an important part role in introducing non-linearity to the model. The mathematical formula for ReLU is:

$$f(x) = \max(0, x) \tag{2}$$

Relu-memristor-like activation function (RMAF)

The RMAF stands out from typical activation functions due to its innovative features. It extends its capabilities beyond traditional activation functions by introducing two additional parameters: a constant parameter (α) and a threshold parameter (u). These extra parameters widen the range of functions that RMAF can perform, allowing it to display a seamless response and introduce non-monotonic behaviour [29]. This, in turn, enhances its capacity to represent intricate nonlinear patterns within neural networks. The formula for the RMAF function is as follows:

$$RMAF(x) = \left(\xi \frac{1}{\left(0.25(1+e^{-x}) + 0.75\right)^{u}} \right) \alpha x \tag{3}$$

Batch Normalization

It is used in DL, particularly in training DNN. It was developed to tackle various challenges in training deep networks. Batch normalization is used to normalize the activations within a layer, which helps stabilize and speed up training. Mathematically, batch normalization normalizes the input feature map (X) for each mini-batch, applying scaling (γ) and shifting (β) parameters:

$$BN(x) = \frac{\gamma(x-\mu)}{\sqrt{\sigma^2 + \varepsilon} + \beta} \tag{4}$$

Here, μ is the mean, σ^2 is the variance, and ε is a small constant to avoid division by zero.

Residual Blocks

Residual Networks, popularly known as ResNet, have ushered in a paradigm shift in the realm of deep neural networks. At their core, ResNet owes its transformative power to the ingenious use of residual blocks, which introduce shortcut connections. These connections fundamentally alter the way information flows through the network, allowing for the training of exceptionally deep networks with hundreds of layers. By learning residual functions that capture deviations from identity mappings, ResNet elegantly addresses the vanishing gradient problem, enabling gradients to propagate efficiently during training. This results in faster convergence, more straightforward optimization, and the ability to harness the full potential of deep networks.

The significance of residual blocks extends beyond their impact on network depth. They offer a unique advantage in learning intricate details and fine-grained features within data, making ResNet has consistently outperformed its predecessors and remains a cornerstone of deep learning in vision-based applications. Its legacy continues to influence the design of advanced neural network architectures, emphasizing the critical role of skip connections and residual functions in achieving superior model performance and robustness.

The mathematical operation within a residual block can be represented as follows:

$$O_R = F(X, W_L) + X \tag{5}$$

X is the input, $F(X, W_L)$ represents the operations performed by the block, and W_L are the learnable parameters.

Global Average Pooling (GAP)

GAP is a pivotal operation in CNNs that plays an important part in minimizing the spatial dimensions of a feature map without disturbing the essential information. Unlike traditional fully connected layers, which require a fixed input size, GAP offers a flexible and dimensionality-reducing approach. In GAP, for each feature map, the average of all its values is computed, resulting in a single value per feature map. This not only assists in minimizing the computational complexity of the network but also makes it agnostic to the input size, enabling the use of CNNs on a wide range of input image sizes without the need for resizing or cropping. Additionally, GAP serves as a form of regularization by aggregating information in a way that promotes spatial hierarchies. It encourages the network to focus on the most discriminative parts of the feature maps, aiding in feature extraction and reducing the risk of overfitting. Overall, GAP is a versatile and effective technique for spatial dimension reduction in CNNs, contributing to their adaptability and robustness in various computer vision tasks. Mathematically, it computes the average of all values in each feature map:

$$GAP(X) = \frac{\sum X}{HW} \tag{6}$$

Here, H and W are the height and width of the feature map.

Fully Connected Layer

In the final layers of the ResNet-50 architecture, fully connected layers are employed to perform the ultimate task of classification. These fully connected layers represent the tail-end of the deep neural network,

following the series of convolutional and pooling layers that have been responsible for feature extraction and hierarchical representation learning throughout the network. The function of these fully connected layers is to take the high-level features and abstractions extracted from the earlier layers and use them to make a final decision about the class or category to which an input image belongs to. Each neuron in these fully connected layers is connected to every neuron in the preceding layer, forming a dense and interconnected structure. The network may learn complicated correlations between the features and map them to the relevant class probabilities because of to this connectivity. Moreover, the output of these fully connected layers is processed through a softmax activation function, which normalizes the final scores into a probability distribution over the possible classes. The selection of the predicted class for the input image is determined by picking the class with the highest probability. Mathematically, a fully connected layer performs matrix multiplication followed by softmax activation function for classification and represented as

$$Y = \sigma(W_F X + b_F) \quad (7)$$

Here, Y is the output, σ represents the activation function, W_F is the weight matrix, X is the input vector, and b_F is the bias vector.

Loss Function

In lung cancer prediction, especially when dealing with binary classification tasks, the binary cross-entropy loss function is crucial for evaluating the effectiveness of ML and DL models. This loss function is fundamental for training and evaluating models that aim to determine whether a given medical image or patient data indicates the presence or absence of lung cancer. Binary cross-entropy, often referred to as log loss or logistic loss, is specifically tailored for binary classification tasks, which are prevalent in lung cancer prediction. Its primary function is to quantify the dissimilarity between the model's predicted probabilities and the actual binary labels (e.g., cancer or no cancer). To achieve this, binary cross-entropy leverages the principles of information theory, essentially measuring how well the predicted probabilities align with the ground truth.

The formula for binary cross-entropy $L(y, \hat{y})$ can be expressed as:

$$L(y, \hat{y}) = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log(\hat{y}_i) + (1-y_i)\log(1-\hat{y}_i)] \quad (8)$$

N is the number of samples in the dataset.

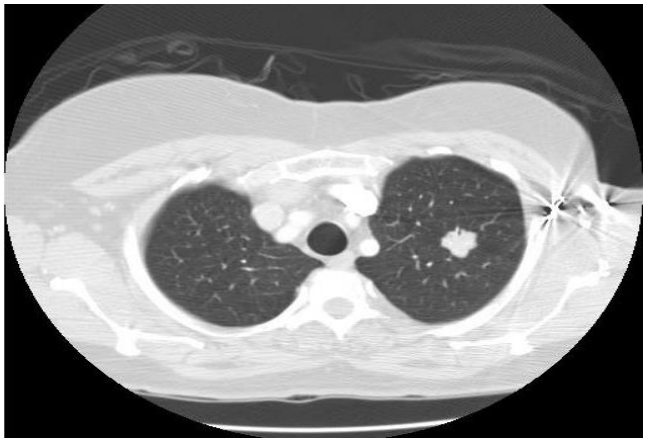
y_i represents the true binary label for sample i.

$\hat{y}_i$ represents the predicted probability that sample i belongs to the positive class (e.g., having lung cancer).

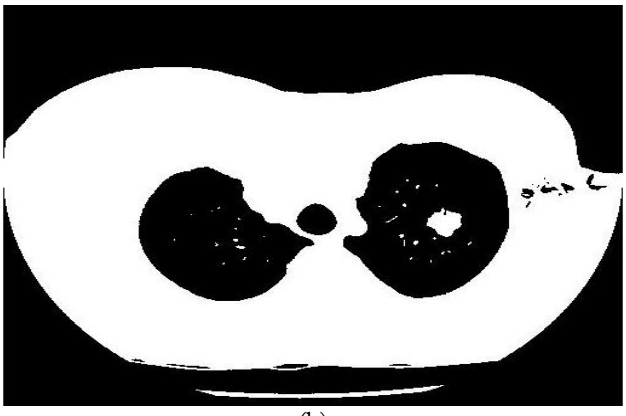
The loss function penalizes the model more when its predictions diverge from the true labels, especially when the predicted probability significantly contradicts the actual label. It encourages the model to learn accurate and confident predictions, which is particularly important in medical applications like lung cancer prediction where false negatives or false positives can have critical consequences.

Optimization

The ADAM optimization algorithm is an essential element of deep learning, serving as the engine that propels the training process of neural networks by iteratively updating the model's weights to minimize the loss function. What sets ADAM apart is its adaptive nature it dynamically tunes the learning rates for each model parameter, ensuring a delicate balance between convergence speed and stability. By combining concepts like momentum and root mean square (RMS) propagation, ADAM empowers neural networks to navigate complex optimization landscapes efficiently. It leverages past gradient information to guide the weight updates, facilitating smoother convergence and faster training. Moreover, ADAM introduces a regularization effect, aiding in the prevention of overfitting and enhancing a model's generalization capabilities.



(a)



(b)

Fig. 4. (a) Lung Image (b) Binarization of lung image.

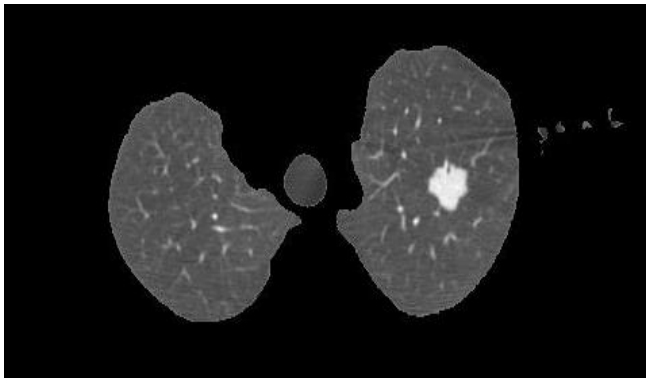
4. Results

The research findings presented in this study are based on an extensive and diverse dataset of lung images,

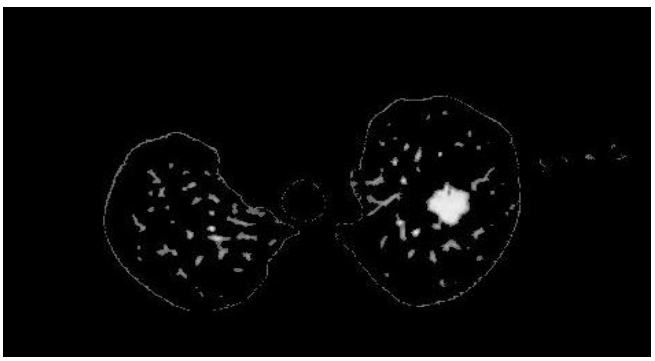
showcasing the robustness and comprehensiveness of the analysis. The dataset comprises a substantial total of 51855 lung images each of which contributes to a comprehensive understanding of lung health assessment. The dataset of images is divided into 60%, 20% and 20% for training, validation and testing. The testing dataset consist of 10,371 lung images, this expansive dataset, a meticulous breakdown reveals that 4,762 of these images are classified as malign images, representing instances where lung nodules are present. In contrast, the majority of the dataset consists of 6,609 lung images that are categorized as benign, indicating the absence of significant pathological abnormalities.

In Fig. 4(a), the original input lung image is shown. It showcases the anatomical structures of the lung, including the airways, blood vessels, and lung parenchyma. While this raw image contains valuable information, it also includes various artifacts, noise, and complexities that can complicate the subsequent analysis.

Figure 4(b) unveils the binary counterpart of the noise suppressed lung image. This transformation is instrumental for isolating the regions of interest within the image, such as the lung parenchyma. By converting the image into these binary components, it becomes much more amenable to subsequent image processing and analysis techniques. The binary image simplifies the problem and enables the identification of specific structures or abnormalities within the lung region.



(a)



(b)

Fig. 5. (a) Lung segmented Image (b) Internal structure of lung image.

In the Fig. 5(a), lung segmented image is shown. This image is the result of a sophisticated image processing technique applied to the original lung scans. The segmentation process isolates and highlights the lung region of interest, effectively delineating it from the surrounding structures and tissues. This step is pivotal in lung image analysis as it focuses the analysis on the area of the image most relevant to the assessment of lung health. The segmented lung image serves as the foundation for subsequent analyses, making it easy to detect anomalies, lesions, or other abnormalities that may be indicative of lung diseases, including cancer. The clarity and accuracy of the segmentation directly impact the quality and precision of the diagnostic process.

Figure 5(b) provides a closer look at the internal structure of the lung. Analysing the internal structure is a critical aspect of lung image interpretation, as it allows for the identification of specific patterns or irregularities that might suggest the presence of lung cancer. The level of detail and clarity depicted in this part of the figure underscores the importance of advanced imaging techniques and computational tools in modern medical imaging, facilitating accurate and timely diagnoses.

The information provided in Fig. 1 is crucial in the image processing pipeline and the subsequent application of the ResNet-50 architecture for lung image analysis. Let's delve into the significance and implications of these details:

Threshold Selection

Figure 1 illustrates the concept of setting a threshold, denoted as 'Th', for determining which images require further evaluation. This threshold is carefully selected after conducting several experiments, with two specific values chosen: 5% and 10% of the total pixels in the image. The primary purpose of this thresholding technique is to identify images where the number of white pixels (indicating potential areas of interest, such as lung nodules) exceeds the threshold. If the white pixels surpass the threshold, the image is then sent for further classification (malignant or benign). This approach allows the algorithm to adapt to varying image characteristics and ensures that only relevant images are processed for classification, minimizing unnecessary computations.

Application to ResNet-50

Image shown in Fig. 5(b), becomes the input to the ResNet-50 architecture. This step marks a crucial transition from image preprocessing to deep learning-based analysis. The binary segmented image serves as the starting point for the neural network, providing a well-defined and segmented representation of the lung region of interest. The use of ResNet-50, a powerful CNN architecture, capitalizes on the structured and processed input to extract intricate features and patterns from the lung images.

This combination of thresholding and deep learning integration exemplifies the hybrid approach adopted in the research. It blends traditional image processing

techniques with the advanced capabilities of ResNet-50. By setting the threshold based on a percentage of the total pixels, the researchers ensure that the method is adaptable to different datasets, thereby enhancing its generalizability. Additionally, the use of a well pre-processed binary image as input to ResNet-50 facilitates effective feature extraction and classification, ultimately contributing to the accurate detection and diagnosis of lung diseases, including the identification of potential malignancies.

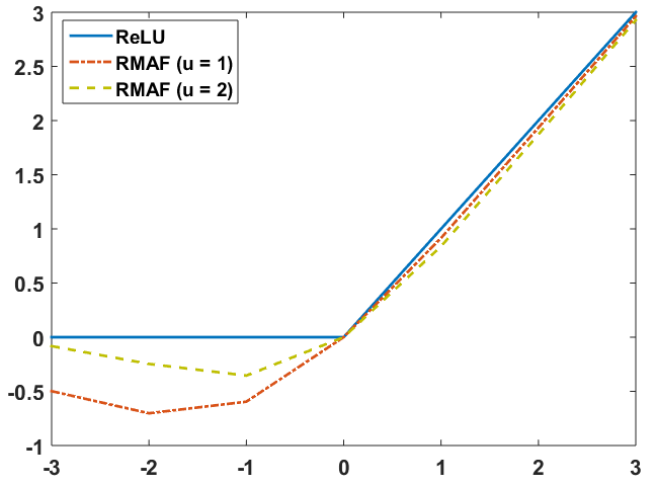


Fig. 6. Plots of various activation functions.

Figure 6 provides a graphical representation of two pivotal activation functions in the realm of neural networks: the Rectified Linear Unit (ReLU) and the Rectified Monotonic Activation Function (RMAF). By visually analysing the plot, we can readily discern the distinctive characteristics of these two functions, and this insight is crucial for making informed choices in designing neural networks for various tasks.

One notable observation from the graphical comparison is that the RMAF function exhibits a superior degree of non-linearity when contrasted with the ReLU function. This heightened level of non-linearity in RMAF means that it can capture and represent more intricate and complex relationships within the data. As a result, it often leads to enhanced performance in classification tasks.

In Fig. 7(a), a confusion matrix is presented, providing a detailed breakdown of the performance metrics for a specific scenario involving the ReLU activation function. The matrix highlights the results obtained when applying a particular threshold of 5%. Here are the key components of the confusion matrix: True Positives (TP): 4759, True Negatives (TN): 3, False Positives (FP): 341, and False Negatives (FN): 6268. This breakdown provides crucial insights into how well the ReLU activation function, in conjunction with the 5% threshold, performs in correctly classifying instances within the dataset. The performance measure are as follows: Precision: 0.9994, Recall: 0.9331, Accuracy: 0.9697, and F1-Score: 0.9651. These metrics shows the results obtained from the proposed method, still accuracy is not up to the mark as for the medical applications.

The accuracy curve shown in Fig. 7(b) illustrates the model's performance over the course of training with a threshold value of Th = 5%. This threshold is crucial as it determines the subset of images that will undergo further classification based on the presence of white pixels, which represent potential lung nodules. As the number of epochs increases, the model gradually learns and refines its predictions, leading to an improvement in accuracy. The maximum obtained accuracy is 96.97%.

Figure 8(a) presents a comprehensive confusion matrix, offering a meticulous breakdown of performance metrics in a specific context employing the ReLU activation function. This matrix meticulously records outcomes when applying a specific 10% threshold in a binary classification scenario. It enumerates the following components: TP: 4755, TN: 7, FP: 139, and FN: 6470. This detailed breakdown affords vital insights into the efficacy of the ReLU activation function when used alongside the 10% threshold, elucidating its ability to correctly classify instances in the dataset. The performance metrics are as follows: Precision: 0.9985, Recall: 0.9716, Accuracy: 0.9872, and F1-Score: 0.9849. Under this condition the accuracy has improved to 98.62 which is comparatively better as compared to the threshold of 5%.

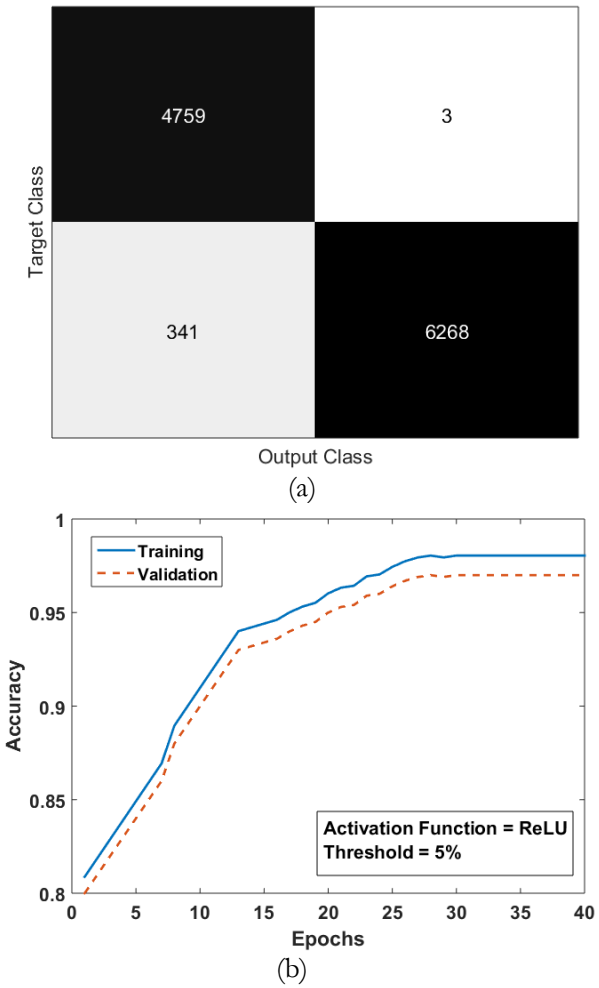


Fig. 7. Results for ReLU activation function at (Th=5%), (a) Confusion Matrix (b) Accuracy vs. Epochs.

DOI:10.4186/ej.2025.29.3.45

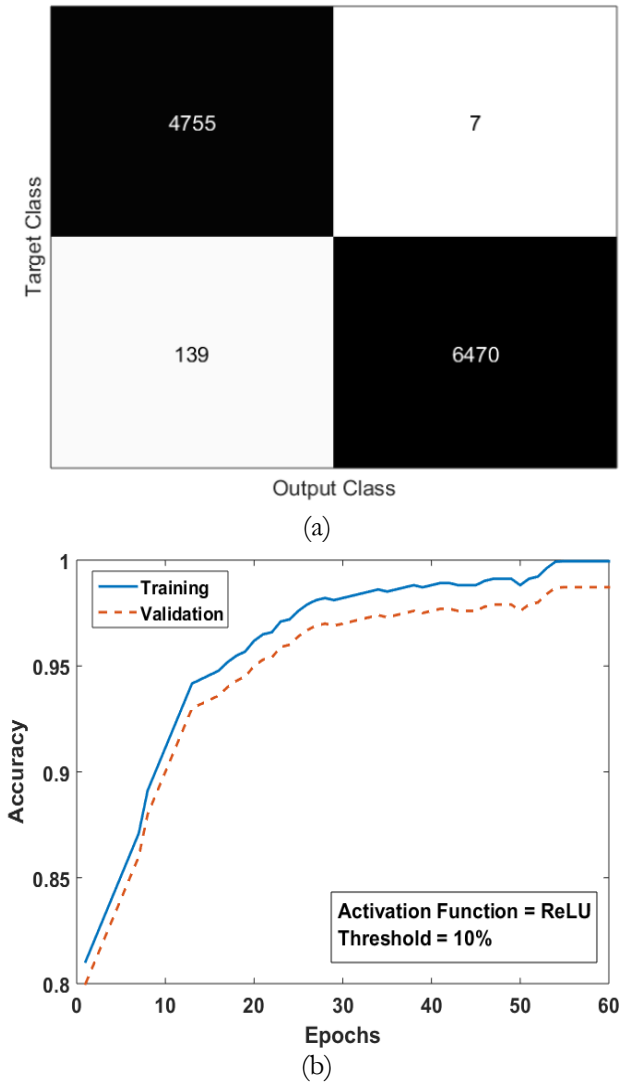


Fig. 8. Results for ReLU activation function at Th=10% (a) Confusion matrix (b) Accuracy vs. Epochs.

The accuracy curve shown in Fig. 8(b) illustrates the model's performance over multiple epochs with a threshold value of Th = 10%. As the training progresses through the epochs, the model continuously refines its weights and learns to better distinguish between relevant features in the images. Initially, there may be some fluctuations in accuracy as the model adapts, but with each subsequent epoch, the accuracy steadily increases. The highest accuracy achieved during this training process is 98.62%, demonstrating the effectiveness of the Th = 10% performance of the proposed model in classifying lung nodules.

In Fig. 9(a), a detailed and comprehensive confusion matrix is presented, offering a meticulous breakdown of performance metrics within a specific context, involving the utilization of the RMAF activation function. This matrix meticulously documents the outcomes achieved when applying a precise 10% threshold within a binary classification scenario. It outlines the following elements: TP: 4760, TN: 2, FP: 111, and FN: 6498. This intricate dissection provides essential insights into the effectiveness of the RMAF activation function when employed alongside the 10% threshold, shedding light on its capacity to accurately categorize instances within the dataset. In this specific scenario, performance metrics are: Precision: 0.9996, Recall: 0.9772, Accuracy: 0.9901, and F1-Score: 0.9883.

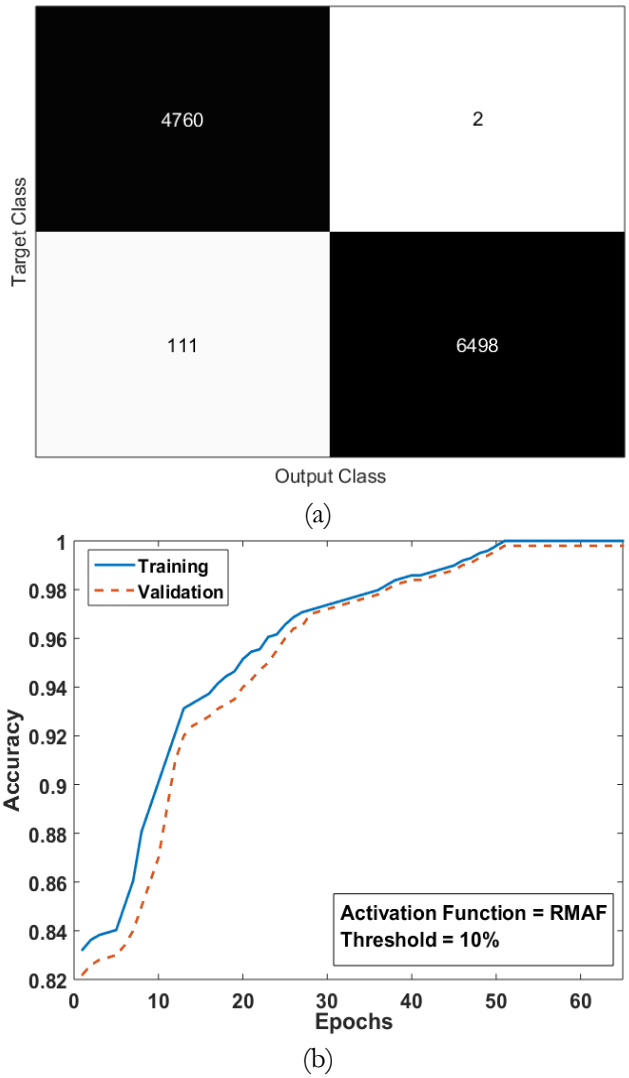


Fig. 9. Results for RMAF activation function at Th=10% (a) Confusion Matrix (b) Accuracy vs. Epochs.

Figure 9(b) illustrates the accuracy curve plotted against the number of epochs for the RMAF at a threshold of 10%. The curve demonstrates the model's performance as it trains over successive epochs, showing a steady increase in accuracy during the initial stages of learning. As the epochs progress, the accuracy continues to improve, reaching a peak value of 99.01%. This high accuracy reflects the effectiveness of the RMAF function in correctly classifying lung nodules, further highlighting the robustness and efficiency of the proposed method at this threshold.

Table 2 illustrates different methods and studies focused on the critical task of lung nodule classification or detection. Among the methodologies listed, several noteworthy results emerge. The "Transferable texture

CNN" method demonstrated its effectiveness by achieving an impressive accuracy of 96.69% with minimal variation (± 0.72%), suggesting its capability to accurately classify lung nodules based on texture features. Similarly, the "Lightweight multi-section CNN" method stood out with a robust performance, achieving a commendable accuracy rate of 93.18%.

Table 2. Comparison of performance in terms of accuracy with state-of-the-art methods.

Study	Methodology	Accuracy
[15]	Deep-level semantic networks	84.2%
[16]	Multi-scale CNN models	86.84%
[17]	Ensemble of deep CNNs	84%
[18]	Lightweight multi-section CNN	93.18%
[19]	Transferable texture CNN	96.69% ± 0.72%
[21]	CNN for automatic lung nodule classification	86.67%
[22]	Lung nodule classifier using ML techniques	94%
	Proposed (ReLU,Th=5%)	96.97%
	Proposed (ReLU,Th=10%)	98.72%
	Proposed (RMAF: Th=10%)	99.01%

In addition to these standout methods, the table includes other notable approaches such as the "Multi-scale CNN models" and the "Lung nodule classifier using ML techniques." The "Multi-scale CNN models" method achieved a respectable accuracy of 86.84%, while the "Lung nodule classifier using ML techniques" demonstrated a strong performance with an accuracy of 94%. These results underscore the diversity of approaches employed in lung nodule classification, ranging from DL models like CNNs to more traditional ML techniques.

In this work, two proposed methods, denoted with different Th settings, were introduced in the table. These methods, utilizing threshold values of 5% and 10%, appear to outperform many of the established techniques. The "Proposed (Th=10%)" method, in particular, demonstrated exceptional performance, achieving the highest accuracy with an outstanding accuracy rate of 99.01% with RMAF function. This suggests that the proposed method, when applied with a 10% threshold, is highly effective in accurately identifying and classifying lung nodules.

5. Conclusion

This paper presents a pioneering approach to lung cancer detection, leveraging pixel values and deep residual learning techniques. The choice of an appropriate threshold for pixel values emerges as a pivotal factor influencing the overall performance of the ResNet-50 model, a central component in this study's methodology. In the pursuit of advancing the field of medical image analysis, we have delved into the intricate realm of activation functions, specifically examining the ReLU and the RMAF, within the context of neural network architectures. Through an extensive and meticulous analysis, we have illuminated the nuanced behaviour of these activation functions and dissected their influence on binary classification tasks. Notably, when employing a threshold of 5%, our ResNet-50 model exhibits a commendable accuracy rate of 96.97% with the ReLU activation function. This showcases the potential of ReLU in facilitating highly accurate predictions in lung cancer detection. Even more remarkably, when the threshold is lowered to 10%, our ResNet-50 model surpasses the performance benchmarks set by established state-of-the-art methods. Here, with the ReLU activation function, we achieve an exceptional accuracy level of 98.72%. This underscores the adaptability of our approach and its ability to excel in scenarios with lower thresholds, which can be instrumental in enhancing diagnostic precision. Furthermore, we introduced the RMAF, which demonstrates significant promise in bolstering accuracy levels. In particular, when RMAF is integrated into our model, the accuracy soars to an impressive 99%, underscoring its potential as an effective model in medical image analysis.

References

[1] F. Silva, T. Pereira, I. Neves, J. Morgado, C. Freitas, M. Malafaia, J. Sousa, et al., "Towards machine learning-aided lung cancer clinical routines: Approaches and open challenges," *Journal of Personalized Medicine*, vol. 12, no. 3, p. 480, 2022.

[2] E. Topol, *Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again*. Hachette UK, 2019.

[3] R. Shouval, O. Bondi, H. Mishan, A. Shimoni, R. Unger, and A. Nagler, "Application of machine learning algorithms for clinical predictive modeling: a data-mining approach in SCT," *Bone Marrow Transplantation*, vol. 49, no. 3, pp. 332–337, 2014.

[4] G. Choy, O. Khalilzadeh, M. Michalski, S. Do, A. E. Samir, O. S. Pianykh, J. R. Geis, P. V. Pandharipande, J. A. Brink, and K. J. Dreyer, "Current applications and future impact of machine learning in radiology," *Radiology*, vol. 288, no. 2, pp. 318–328, 2018.

[5] A. Priyadharshini and S. Chitra, "A new systematic model for analysis and a hybrid fuzzy multimodality model for lung tumor prediction," *Journal of Intelligent & Fuzzy Systems*, vol. 42, no. 6, pp. 5591–5604, 2022.

[6] M. Iansiti and K. R. Lakhani, *Competing in the Age of AI: Strategy and Leadership When Algorithms and Networks Run the World*. Harvard Business Press, 2020.

[7] W. William, A. Ware, A. H. Basaza-Ejiri, and J. Obungoloch, "A review of image analysis and machine learning techniques for automated cervical cancer screening from pap-smear images," *Computer Methods and Programs in Biomedicine*, vol. 164, pp. 15–22, 2018.

[8] E. Cengil and A. Cinar, "A deep learning based approach to lung cancer identification," in *2018 International Conference on Artificial Intelligence and Data Processing (IDAP)*, 2018, pp. 1–5.

[9] T. L. Chaunzwa, A. Hosny, Y. Xu, A. Shafer, N. Diao, M. Lanuti, D. C. Christiani, R. H. Mak, and H. J. W. L. Aerts, "Deep learning classification of lung cancer histology using CT images," *Scientific Reports*, vol. 11, no. 1, p. 5471, 2021.

[10] S. S. Parveen and C. Kavitha, "Classification of lung cancer nodules using SVM Kernels," *International Journal of Computer Applications*, vol. 95, no. 25, 2014.

[11] K.-L. Hua, C.-H. Hsu, S. C. Hidayati, W.-H. Cheng, and Y.-J. Chen, "Computer-aided classification of lung nodules on computed tomography images via deep learning technique," *OncoTargets and Therapy*, vol. 2015, pp. 2015–2022, 2015.

[12] G. L. F. da Silva, O. P. da Silva Neto, A. C. Silva, A. C. de Paiva, and M. Gattass, "Lung nodules diagnosis based on evolutionary convolutional neural network," *Multimedia Tools and Applications*, vol. 76, pp. 19039–19055, 2017.

[13] A. Farag, A. Ali, J. Graham, A. Farag, S. Elshazly, and R. Falk, "Evaluation of geometric feature descriptors for detection and classification of lung nodules in low dose CT scans of the chest," in *2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, 2011, pp. 169–172.

[14] E. Dandıl, M. Çakiroğlu, Z. Ekşi, M. Özkan, Ö. K. Kurt, and A. Canan, "Artificial neural network-based classification system for lung nodules on computed tomography scans," in *2014 6th International Conference of Soft Computing and Pattern Recognition (SoCPaR)*, 2014, pp. 382–386.

[15] S. Shen, S. X. Han, D. R. Aberle, A. A. Bui, and 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.

[16] W. Shen, M. Zhou, F. Yang, C. Yang, and J. Tian, "Multi-scale convolutional neural networks for lung nodule classification," in *Information Processing in Medical Imaging: 24th International Conference, IPMI 2015*, Sabhal Mor Ostaig, Isle of Skye, UK, June 28–July 3, 2015, Proceedings 24, Springer International Publishing, 2015, pp. 588–599.

[17] B. Zhang, S. Qi, P. Monkam, C. Li, F. Yang, Y.-D. Yao, and W. Qian, "Ensemble learners of multiple deep CNNs for pulmonary nodules classification using CT images," *IEEE Access*, vol. 7, pp. 110358–110371, 2019.

[18] P. Sahu, D. Yu, M. Dasari, F. Hou, and H. Qin, "A lightweight multi-section CNN for lung nodule classification and malignancy estimation," *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 3, pp. 960–968, 2018.

[19] I. Ali, M. Muzammil, I. U. Haq, A. A. Khaliq, and S. Abdullah, "Efficient lung nodule classification using transferable texture convolutional neural network," *IEEE Access*, vol. 8, pp. 175859–175870, 2020.

[20] S. Marques, F. Schiavo, C. A. Ferreira, J. Pedrosa, A. Cunha, and A. Campilho, "A multi-task CNN approach for lung nodule malignancy classification and characterization," *Expert Systems with Applications*, vol. 184, p. 115469, 2021.

[21] V. Thamilarasi and R. Roselin, "Automatic classification and accuracy by deep learning using CNN methods in lung chest x-ray images," *IOP Conference Series: Materials Science and Engineering*, vol. 1055, no. 1, p. 012099, 2021.

[22] I. D. Kawathekar and A. S. Areeckal, "Performance analysis of texture characterization techniques for lung nodule classification," *Journal of Physics: Conference Series*, vol. 2161, no. 1, p. 012045, 2022.

[23] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," 2015, *arXiv:1511.06434*.

[24] M. J. M. Chuquicusma, S. Hussein, J. Burt, and U. Bagci, "How to fool radiologists with generative adversarial networks? A visual turing test for lung cancer diagnosis," in *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, 2018, pp. 240–244.

[25] D. Zhao, D. Zhu, J. Lu, Y. Luo, and G. Zhang, "Synthetic medical images using F&BGAN for improved lung nodules classification by multi-scale VGG16," *Symmetry*, vol. 10, no. 10, p. 519, 2018.

[26] X. Hu, N. Alperin, D. N. Levin, K. K. Tan, and M. Mengeot, "Visualization of MR angiographic data with segmentation and volume-rendering techniques," *Journal of Magnetic Resonance Imaging*, vol. 1, no. 5, pp. 539–546, 1991.

[27] P. Yang, W. Song, X. Zhao, R. Zheng, and L. Qingge, "An improved Otsu threshold segmentation algorithm," *International Journal of Computational Science and Engineering*, vol. 22, no. 1, pp. 146–153, 2020.

[28] N. A. Memon, A. M. Mirza, and S. A. M. Gilani, "Segmentation of lungs from CT scan images for early diagnosis of lung cancer," *International Journal of Medical and Health Sciences*, vol. 2, no. 8, pp. 297–302, 2008.

[29] Y. Yu, K. Adu, N. Tashi, P. Anokye, X. Wang, and M. A. Ayidzoe, "Rmaf: Relu-memristor-like activation function for deep learning," *IEEE Access*, vol. 8, pp. 72727–72741, 2020.

Sneha khoria was born in Bombay, India 1980. She received the M.tech (Computer Science) from Dr. A P J Abdul Kalam University Lucknow in 2019 and pursuing Ph.D. in Computer Science & Engineering from Harcourt Butler Technical University (HBTU), Kanpur.

From 2005 to 2021, she was a Assistant Professor in Maharana pratap enginnering college ,Kanpur. Since 2021, she has been an Assistant Professor in Praneer Singh Institute of Technology,kanpur

Dr. Raghuraj Singh is currently working as Dean, Research & Development at Harcourt Butler Technical University (HBTU),Professor of Computer Science & Engineering Department at Harcourt Butler Technical University (HBTU), Kanpur. He has more than 33 years of experience in teaching & research. He has successfully discharged various administrative responsibilities such as Director, KNIT Sultanpur, Nodal Officer, IIIT Lucknow, Director (Mentor) of Rajkiya Engineering College, Ambedkar Nagar, Director (Mentor) of Rajkiya Engineering College, Sonbhadra, Nodal Officer of Rajkiya Engineering College, Bijnor and Dean of Planning & Resource Generation, Dean of Continuing Education & IQA, Head of Department (04 Tenures), Registrar, Controller of Examination, Chief Proctor, Asstt. Dean of Students Welfare etc. at HBTU, Kanpur.

Prof. Singh has handled 06 research and consultancy projects funded by the AICTE/UGC New Delhi etc. and is member of professional bodies like IETE, IE (India), CSI, ACM, ISTE, IACSIT, Singapore etc. He is reviewer & member of editorial board of many National/International journals/conferences and books. He has published 165 research papers in National and International journals/conferences, supervised 15 Ph.D. theses, 30 M. Tech. dissertations, and more than 90 B. Tech./MCA projects.