Improved SegNet with Hybrid Classifier for Lung Cancer Segmentation and Classification

Rathod Dharmesh Ishwerlal, Reshu Agarwal, K.S. Sujatha
Research Scholar, Amity Institute of Information Technology (AIIT), Amity University, Noida, Uttar Pradesh, India[1]
Assistant Professor, Amity Institute of Information Technology(AIIT), Amity University, Noida, Uttar Pradesh, India[2]
Associate Professor, Department of Electrical and Electronics Engineering, JSS Academy of Technical Education, Noida, Uttar Pradesh, India[3]

***Abstract*—Prompt diagnosis is crucial globally to save lives, underscoring the urgent need in light of lung cancer's status as a leading cause of death. While CT scans serve as a primary imaging tool for LC detection, manual analysis is laborious and prone to inaccuracies. Recognizing these challenges, computational techniques, particularly ML and DL algorithms, are being increasingly explored as efficient alternatives to enhance the precise identification of cancerous and non-cancerous regions within CT scans, aiming to expedite diagnosis and mitigate errors. The proposed model employs Preprocessing to standardize image features, followed by segmentation using an Improved SegNet framework to delineate cancerous regions. Features like LGXP and MBP are then extracted, facilitating classification with a hybrid classifier which combines LSTM and LinkNet models. Implemented in Python, the model's performance is evaluated against conventional methods, showcasing superior accuracy, sensitivity, and precision. This framework promises to revolutionize LC diagnosis, enabling early intervention and improved patient outcomes.**

Keywords*—Improved SegNet; LGXP; MBN; LSTM; LinkNet; lung cancer

I. INTRODUCTION

Lung cancer remains one of the leading causes of cancer-related mortality worldwide, necessitating the development of advanced diagnostic tools for early detection and treatment planning [6] [7]. CT imaging is essential, particularly when it comes to lung cancer. Nonetheless, the manual assessment of CT scans by radiologists is laborious and susceptible to differences in interpretation between observers. In this context, automated methods for lung cancer segmentation and classification [9] have emerged as the promising avenues to enhance the diagnostic accuracy and streamline clinical workflows.

The primary objective of lung cancer segmentation [8] [13] is to delineate the boundaries of cancerous lesions within the lung parenchyma. This task is inherently challenging due to the variability in lesion size, shape, and intensity characteristics across different patients and disease stages. Various segmentation algorithms have been proposed, ranging from traditional techniques such as thresholding and region growing to more sophisticated deep learning-based approaches like U-Net [11] and Mask R-CNN. These methods leverage the spatial and contextual information present in CT images [10] to accurately segment lung tumors while minimizing false positives [14].

Once the lung tumors are segmented, the subsequent step involves the classification of these regions into malignant or benign categories [12]. Deep features that are directly learned from the raw image data and manually created features that are taken from segmented lesions are used to train classification algorithms. SVM, RF, and CNNs are among the commonly employed classifiers for this task. By using ML techniques, these classifiers can effectively discriminate between cancerous and non-cancerous regions, facilitating accurate diagnosis and treatment planning.

The development of robust segmentation and classification algorithms for lung cancer holds immense clinical significance [15]. Beyond aiding in early detection and accurate diagnosis, these automated methods have the potential to assist radiologists in monitoring disease progression, evaluating treatment response, and predicting patient outcomes. Furthermore, by reducing the reliance on manual interpretation and improving diagnostic efficiency, automated lung cancer detection systems can alleviate the burden on healthcare resources and improve patient care outcomes. However, these methods are arbitrary and may result in differences in observational quality. To solve this challenge, researchers have created computational algorithms that automate the interpretation of medical pictures for the goal of identifying lung cancer. The following are the proposed paper's primary contributions:

In this work, introduces an innovative approach for lung cancer segmentation by employing an Improved SegNet architecture. This segmentation method enhances the accuracy of identifying cancerous regions within CT images, thus facilitating precise diagnosis and treatment planning.

This remaining section of this document is organized as follows: Section II examines pertinent research on the categorization, segmentation, and associated techniques of lung cancer. Section III presented the proposed methodology. Results and discussion are explained in Section IV and summary of the paper’s contribution is given in below Section V.

II. Related Works

Before In 2024, Sangeetha S.K.B et al., [1] presented a MFDNN architecture designed to integrate various modalities in lung cancer diagnosis, including medical imaging, genomics, and clinical data. By addressing particular issues related to each area, this fusion strategy improves diagnostic precision. Reliability was further enhanced by integrating electronic health records and clinical data. The ethical implications of implementing AI in therapeutic settings were underlined, emphasizing the necessity of strict regulations and thorough validation. MFDNN achieves exceptional accuracy (92.5%) and excels in precision (87.4%) and recall (86.4%), with an F1-score of 86.2, surpassing established methods like CNN, ResNet and DNN. These findings demonstrate the vital role MFDNN plays in enabling the more rapid and precise identification of this serious illness, hence revolutionizing the diagnosis of lung cancer.

In 2024, Sampangi Rama Reddy B R et al., [2] introduced a novel architecture, the SNN, for detecting and classifying lung cancer utilizing CT scan data. Examining which NN models were most effective at lung cancer identification in its early stages was the main goal. Initially, lung regions were extracted employing image processing methods, followed by segmentation using the SNN. Various NN techniques were then employed for classification based on the features extracted from segmented images. Performance evaluation was conducted using F1-Measure, precision, accuracy, and recall metrics, showcasing a remarkable 96% classification accuracy in testing, surpassing existing methods. The proposed algorithm demonstrated clear potential for real-world clinical application, promising significant advancements in lung cancer diagnosis and treatment.

In 2024, Liangyu Li et al., [3] proposed a novel approach to enhancing lung cancer detection by integrating hybrid feature extraction methods and optimizing ML hyperparameters. Remarkable accuracy rates were obtained by integrating autoencoder-generated features and Haralick features with GLCM, and then using supervised machine learning approaches (specifically, SVM with varied kernels). In particular, SVM with Gaussian and RBF kernels work almost flawlessly, whereas SVM with polynomial kernel, GLCM with autoencoder, Haralick, and autoencoder features, achieves a remarkable 99.89% accuracy. These outcomes highlight how this strategy may greatly enhance prognostic and diagnostic tools for lung cancer treatment planning and decision-making.

In 2024, Fuli Zhang et al., [4] presented a novel two-step DL approach for enhancing NSCLC tumor segmentation from CT images. The method involves training separate segmentation networks for large and small tumor images, following an initial coarse segmentation step aimed at detecting lesion regions. Compared to other approaches, the suggested method obtains a much higher DSC of 0.80 ± 0.13 and has the lowest FPR, highest TPR, and HD95. Notably, the approach exhibits substantial improvement in segmenting tumors of both large and small sizes, particularly enhancing performance for small GTVs, which previously exhibited poor results with other methods. Overall, this two-step DL method demonstrates precise NSCLC tumor segmentation and holds promise for enhancing radiotherapy efficiency.

In 2024, Lavina Jean Crasta et al., [5] introduced a novel DL framework tailored for detecting and classifying lung cancer from input CT images. For segmentation and classification process, the proposed architecture consists of a 3D-VNet model and a 3D-ResNet model. The segmentation model attains an outstanding 99.34% DSC and significantly lowers false positives to 0.4% on the LUNA16 dataset. Furthermore, with 99.2% accuracy, 98.8% sensitivity, and 99.6% specificity, the classification model exhibits remarkable performance characteristics. The 3D-VNet network robustly and precisely defines lung nodules of various sizes and shapes, outperforming prior segmentation techniques. Metrics of the classification model show enhanced F1-Score, sensitivity, specificity, accuracy, and improved performance compared to existing techniques.

III. Proposed a Lung Cancer Segmentation and Classification Model

Lung cancer has a high death rate due to late-stage detection and few available treatment options. Efficient identification, planning, and tracking of lung cancer nodules using medical imaging, especially CT scans, depend on accurate segmentation and classification. The paper presents an approach for lung cancer segmentation and classification, which combines innovative techniques to improve accuracy and efficacy in medical imaging analysis.

1) Beginning with preprocessing stage, which involves applying Gaussian filtering to the input CT images. To enhance the image quality and lower noise, Gaussian filtering is applied, which increases the efficiency of the segmentation process that follows:

2) The segmentation stage employs an Improved SegNet architecture specifically tailored for lung cancer segmentation. The Improved SegNet architecture is chosen for its ability to accurately delineate cancerous regions within the CT images.

3) Discriminative features like LGXP and MBP are extracted from the segmented regions after segmentation. The ability to differentiate between cancerous and non-cancerous tissues depends on these features.

4) Finally, the classification process integrates a hybrid classifier, leveraging models such as LSTM and LinkNet, to enhance the classification accuracy, leading to more accurate identification and classification of lung cancer nodules. Fig. 1 depicts the overall process of the proposed model.

A. Preprocessing Phase

In the preprocessing stage of the suggested lung cancer segmentation and classification model described, Gaussian filtering is applied as the first method of processing for the input CT images.

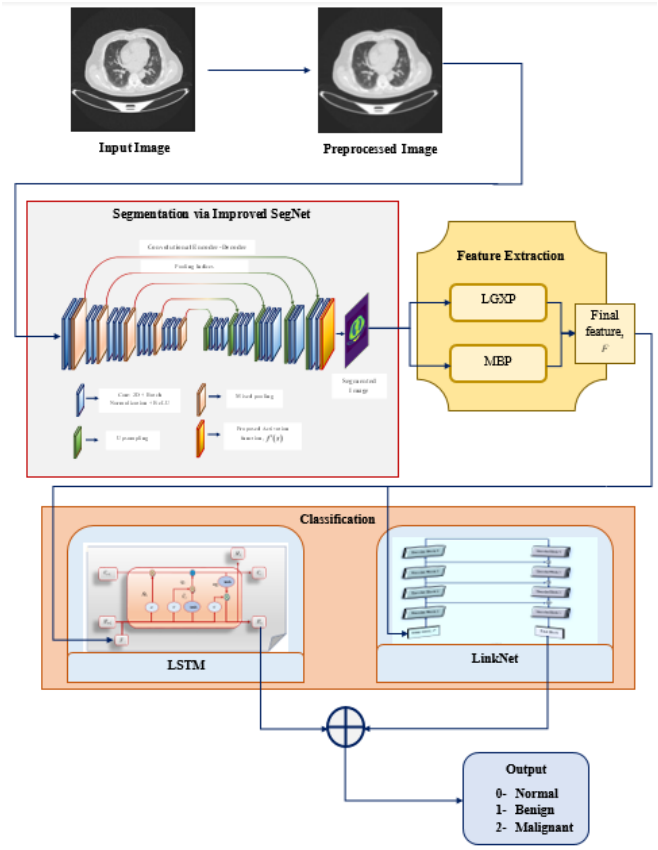

Fig. 1. Overall process of proposed model.

Gaussian filtering [16] is a widely used method in image processing for reducing noise and smoothing the image while preserving important features. By convolving the input image with a Gaussian kernel, high-frequency noise components are attenuated, resulting in a smoother and more visually coherent image. In order to improve the image quality of the CT images prior segmentation as well as feature extraction, this preprocessing procedure is essential. It helps to mitigate the impact of noise and artifacts present in the original images, thereby improving the performance of subsequent processing stages. Additionally, Gaussian filtering aids in standardizing the intensity distribution across different CT scans, which is beneficial for achieving consistent segmentation and classification results.

Eq. (1) is likely employed to calculate the values of individual components within the Gaussian smoothing filter. In this context, signifies the Gaussian function's value at coordinates, while signifies the standard deviation of the Gaussian kernel, influencing the extent or width of the Gaussian function's distribution. Normalization constant is denoted as these values are then used to convolve with the input CT images to perform Gaussian smoothing as a preprocessing step.

As a result, the preprocessed picture is represented as and is then put through another segmentation step.

$$F(x,y)=\frac{1}{P}e^{\frac{x^2+y^2}{2\sigma^2}} \quad (1)$$

B. Segmentation via Improved SegNet

One of the most important steps in medical image analysis is segmentation, which is dividing an image into sections or significant areas in order to retrieve pertinent data for further examination. The segmentation process in this work involves utilizing an Improved SegNet architecture tailored specifically for accurately delineating lung cancer nodules from preprocessed images.

The conventional SegNet model [17] follows an encoder-decoder structure with corresponding layers, which contains 13 layers. The encoder network gradually downsamples the input image, to capture hierarchical features, while the decoder network upsamples the feature maps to produce the segmentation mask. Max pooling operations are typically used in the encoder to perform downsampling, reducing computational burden and extracting dominant features. Despite its effectiveness for semantic segmentation tasks, the conventional SegNet model has limitations. One disadvantage is the loss of spatial information brought about by the encoder's max pooling operations, which might deteriorate segmentation mask quality, especially in applications that call upon accurate object localization. Additionally, the segmentation performance in areas with complex structures or textures may be affected by standard SegNet's inability to capture long-range relationships and contextual information across the image.

To address the limitations of the conventional SegNet model, the proposed Improved SegNet model incorporates several enhancements especially intended for lung cancer segmentation as well as classification. Replacing max pooling with a novel mixed pooling strategy [18], which combines max pooling and average pooling, which is expressed in Eq. (2).

$$\text{Mixed pooling} = \text{max pooling} + \text{Average pooling} \quad (2)$$

In a down sampling process known as "max pooling," the highest value falling inside a predetermined window is kept while the remainder is discarded. This method successfully reduces the spatial dimensions of the feature maps, collecting the most notable features and highlighting locations with significant activation. The calculation of max pooling value is derived in Eq. (3). Here, represents the input feature map, represents the result of max pooling, each channel, and the highest number during a pooling window of dimensions is selected by the maximum pooling function.

$$Max\ \ pooling, Y_{\max}\left(k_{ij}\right)=\max_{(p,q)\in R_{ij}} X\left(k_{ij}\right) \quad (3)$$

Instead, average pooling computes the mean value within the specified window, contrasting with the maximum value calculation described in Eq. (4). It represents the result of max pooling, each channel. Like max pooling, it helps in downsampling the feature maps, but instead of preserving only the most prominent features, it takes into account the overall intensity across the region. By lessening the effect of noise and outliers, this process can help to depict the features more smoothly.

$$Average\ \ pooling, Y_{avg}\left(k_{ij}\right)=\frac{1}{\left|R_{ij}\right|}\sum_{(p,q)\in R_{ij}} X\left(k_{ij}\right) \quad (4)$$

In traditional SegNet models for segmentation tasks, the softmax activation function [19] is commonly used at the output layer to produce probability distributions over different classes or regions within the segmented output. Eq. (5) shows the expression of softmax activation function. However, softmax tends to amplify the differences between input values, which can lead to gradients becoming extremely large during backpropagation, causing instability and slow convergence, especially in deep networks like SegNet. Additionally, softmax outputs tend to be overly confident, resulting in less nuanced representations of uncertainty in segmentation predictions.

$$f(x_i) = \frac{\exp(x_i)}{\sum_j \exp(x_j)} \tag{5}$$

To overcome these limitations, the proposed activation function combines softmax with the hard tanh activation function. The hard tanh function [19] introduces a saturation threshold, preventing gradients from exploding or vanishing and promoting more stable training, which is expressed in Eq. (6). By combining softmax with hard tanh, the proposed activation function aims to mitigate the amplification of differences in input values while preserving the probabilistic interpretation provided by softmax. The proposed improved activation function is expressed in Eq. (7). This combination enables the model to produce more calibrated probability estimates for each class or region in the segmentation output, leading to improved segmentation accuracy and better representation of uncertainty. The suggested activation function typically mitigates the adverse impacts of softmax in SegNet models. It enhances stability during training and yields more precise and nuanced segmentation results for lung cancer diagnosis as well as categorization using CT images. Fig. 2 shows the proposed SegNet model for segmentation process.

$$f(x) = \begin{pmatrix} -1 & if & x < -1 \\ x & if & -1 = x \le 1 \\ 1 & if & x > 1 \end{pmatrix} \tag{6}$$

$$f'(x) = \begin{pmatrix} -1 & if & x < -1 \\ \frac{\exp(x_i)}{\sum_j \exp(x_j)} & if & -1 = x \le 1 \\ 1 & if & x > 1 \end{pmatrix} \tag{7}$$

Finally, the output of the improved SegNet model is a pixel-wise segmentation mask highlighting lung cancer regions within CT images, which is represented in Fig. 2.

C. Feature Extraction Phase

Various strategies are utilized in the segmentation and classification process of the proposed lung cancer model to extract discriminative features from the segmented regions that are produced given SegNet-based segmentation.

The LGXP and MBP significantly enhance the classification process by providing robust and detailed texture representations crucial for distinguishing cancerous from non-cancerous regions. LGXP excels at detecting edges and capturing fine details by encoding the direction and magnitude of local gradients, making it particularly effective for identifying the boundaries of structures within the images. This method's emphasis on gradient information ensures resilience to noise, which is a common challenge. On the other hand, MBP captures intricate texture details through binary pattern encoding based on intensity differences; improve its robustness against noise and illumination variations.

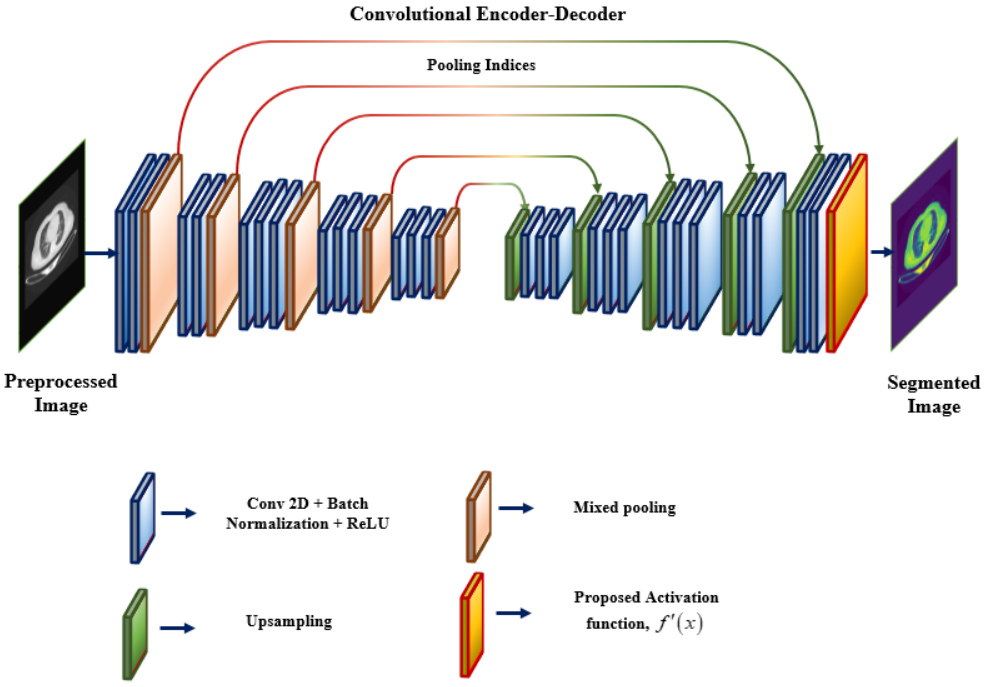

Fig. 2. Proposed SegNet model for segmentation process.

These methods were chosen for their complementary strengths in texture analysis. LGXP's focus on gradient information and MBP's detailed texture encoding together provide a comprehensive representation of lung tissue characteristics. Both methods have demonstrated effectiveness in texture analysis and image segmentation, making them suitable for the complex task of lung cancer detection. Their robustness to noise, illumination changes, and rotational variations ensures that the extracted features are reliable and consistent, leading to improved classification accuracy.

The features include LGXP and MBP features, which are described as follows:

1) LGXP feature: In the proposed model, the feature extraction method utilizes the LGXP descriptor [20], specifically designed to capture texture patterns within segmented lung cancer regions. Initially, the phases within the LGXP descriptor are quantized into distinct ranges, ensuring robustness to variations in Gabor phase. After quantizing each phase value, the LGXP operator interacts with the quantized phases of the central pixel and its neighboring pixels. Using XOR operations, the LGXP operator computes patterns between these phases to produce binary labels that are concatenated together to form a local pattern of the core pixel. Every neighborhood receives a thorough representation of local texture patterns when this process is performed for every neighborhood. Finally, the obtained binary labels are concatenated to generate a local pattern of the core pixel.

The equation for calculating the LGXP, as shown in Eq. (8), includes the quantization of phases and the application of XOR operations between the quantized phases of the central pixel and its neighboring pixels, where, $LGXP_{w,v}$ represents the LGXP descriptor for a specific scale (ω) and orientation (v). g_c and g_i denotes the phase at the central and neighboring pixel position in the Gabor phase map, respectively. $\oplus$ denotes the XOR operation.

$$LGXP_{\omega,v} = q\left(\varphi_{\omega,v}(g_c)\right) \oplus q\left(\varphi_{\omega,v}(g_i)\right) \tag{8}$$

2) MBP feature: In the proposed model, the MBP [21] operator is used for extracting texture features from segmented lung cancer regions, I_S . These features are computed by analyzing local neighborhoods within the segmented image, where the median intensity value serves as a reference point. By encoding each pixel's relationship with the local median intensity as a binary pattern, MBP features succinctly represent texture variations within segmented regions. With the help of these features, lung cancer areas in CT scans may be accurately segmented and classified, providing a thorough assessment of textural aspects.

By thresholding pixels against their median value within a neighborhood, usually 3x3, this operator maps intensities to localized binary patterns that improve resistance to noise and sensitivity to microstructure. The MBP at each pixel (i, j) is computed using a weighted sum over binary values determined by comparing pixel intensities to the median within the local patch. The formula to calculate the MBP at pixel (i, j) in a local patch of size Z (usually 3 × 3), according to Eq. (9). τ is the median that was calculated for the entire local patch.

$$MBP(i, j) = \sum_{k=0}^{Z-1} 2^k H\left(b_k - \tau\right) \tag{9}$$

When the threshold is set to the median, the number of possible binary patterns that result is limited to 256, guaranteeing that the binary pattern subset contains at least 5 one bits. The histogram of these patterns forms the texture descriptor, providing a compact representation of texture variations across the segmented image. As a result, each image is converted into a 256x1 vector that uses local patch median values to represent the MBP histogram. This approach enables comprehensive characterization of textural properties within lung cancer regions, facilitating accurate segmentation and classification in CT images.

Therefore, the feature extraction process aims to capture relevant texture information from the segmented lung cancer regions, facilitating the classification task by providing discriminative features for distinguishing between cancerous and non-cancerous tissues. These extracted features, $F = [LGXP, MBP]$ serve as input to the hybrid classifier in the final stage of the proposed model.

D. Classification Process

Features extracted using LGXP and MBP are fed into a hybrid classifier combining Long Short-Term Memory (LSTM) and LinkNet models. This integration of LGXP and MBP features provides the hybrid classifier with a rich and diverse set of features, enhancing its ability to accurately classify lung tissues. The hybrid classifier harnesses LSTM's ability to capture temporal dependencies and LinkNet's capacity to encode spatial context. Utilizing LSTM networks and LinkNet in the proposed model for lung cancer classification model leverages the complementary strengths of both architectures, addressing specific challenges in medical image analysis. LSTM networks are particularly adept at handling sequential data, which is crucial in the context of CT scans often taken in series over time. This capability allows LSTMs to capture the temporal progression of lung cancer, providing a dynamic perspective of tumor development. Additionally, LSTMs excel at retaining long-term dependencies, maintaining contextual information from previous slices, which enhances the accuracy of classification by ensuring continuity and reducing isolated misclassifications. On the other hand, LinkNet is specifically designed for fast and efficient classification tasks, incorporating residual connections that facilitate better gradient flow and deep feature learning without degradation. This makes LinkNet particularly well-suited for high-resolution medical images like CT scans. Moreover, LinkNet's architecture ensures computational efficiency, which is essential for handling large datasets, making the classification process scalable and practical. The outputs of both models are then combined to make a final

classification decision of segmented lung regions into three distinct classes based on their pathological characteristics. 'Normal' tissue is labeled as '0', indicating the absence of any abnormalities. 'Benign' conditions, which may include non-cancerous growths or abnormalities, are labeled as '1'. 'Malignant' tissue, indicative of lung cancer, is labeled as '2', highlighting the presence of cancerous growths within the lung regions. The hybrid approach combining LSTM and LinkNet results in a more comprehensive analysis of CT scans, improving the model's robustness and reliability. LSTM's ability to handle sequential and contextual information, coupled with LinkNet's precision in image segmentation, allows for a richer set of features for final classification. This dual approach enhances the model's capability to accurately distinguish between cancerous and non-cancerous regions. The synergy between LSTM and LinkNet leads to improved performance metrics, with the hybrid model expected to surpass conventional methods in terms of accuracy, sensitivity, and precision.

1) LSTM classifier: The LSTM [22] classifier forms a critical component of the lung cancer segmentation and classification model, facilitating the categorization of segmented lung regions into distinct classes based on their pathological characteristics. When operating within this model, the LSTM classifier receives input features F extracted from segmented lung regions. These features are structured into sequential data, with each sequence representing the feature vectors corresponding to a segmented lung region.

To effectively represent data sequences with long-range dependencies and overcome the vanishing gradient problem, an RNN type known as the LSTM architecture was developed. The model comprises memory cells responsible for retaining internal states and three types of gates namely, input, forget, and output. These gates govern the flow of information into, out of, and within the cells. Specifically, the forget gate selects data to be discarded from the cell state, the output gate manages the impact of the internal state on the cell's output, and the input gate regulates the quantity of new data stored in the cell state. Because of this architecture's capacity to recognize temporal patterns and learn from sequential input, long-term dependencies and the temporal patterns inherent in sequential data make LSTMs ideal for processing it.

The three gates of the LSTM design are responsible for managing the information flow via the memory cells. The forget gate, as specified in Eq. (10), regulates the extent to which past information is retained in memory. By applying the sigmoid function (σ) to a weighted sum of the prior short-term memory state (H_{t-1}), the present input state (F), and bias terms (S_{fg}), it calculates a forgetting factor (fg_t). The input gates, detailed in Eq. (11) and Eq. (12), produce an input factor (ig_t) and an alternate vector $(\tilde{C}_t)$ by employing the sigmoid function for gating and the hyperbolic tangent function $(\tanh)$ for the alternate vector. These gates regulate the amount of fresh information incorporated into the long-term memory state, as described in Eq. (13). Finally, the output gate, defined by Eq. (14), utilizes the sigmoid function applied to a weighted sum of the current input state F, the previous short-term memory state (H_{t-1}), and bias terms (S_{og}) to calculate an output factor (og_t). The final short-term memory output (H_t) is determined by the output factor (og_t) multiplied by the hyperbolic tangent of the long-term memory state (C_t) as per Eq. (15). Fig. 3 shows the architecture of LSTM classifier.

$$fg_t = \sigma\left(W_{fg}F + W_{fg}H_{t-1} + S_{fg}\right) \tag{10}$$

$$ig_t = \sigma\left(W_{ig}F + W_{ig}H_{t-1} + S_{ig}\right) \tag{11}$$

$$\tilde{C}_t = \tanh\left(W_C F + W_C H_{t-1} + S_C\right) \tag{12}$$

$$C_t = fg_t \times C_{t-1} + ig_t \times \tilde{C}_t \tag{13}$$

$$og_t = \sigma\left(W_{og}F + W_{og}H_{t-1} + S_{og}\right) \tag{14}$$

$$H_t = og_t \times \tanh\left(C_t\right) \tag{15}$$

Through its internal memory units, LSTM is capable of retaining information over extended time periods, thereby allowing the classifier to learn from past observations. Based on the learned temporal patterns and the information extracted from the input sequences, the LSTM classifier makes classification decisions for each segmented lung region.

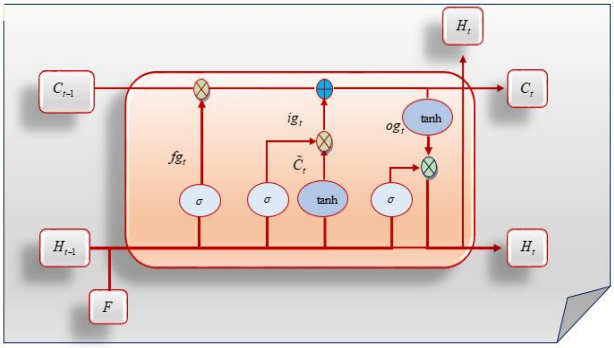

Fig. 3. Architecture of LSTM classifier.

2) LinkNet classifier: LinkNet [23] is a CNN architecture specifically designed for semantic segmentation tasks, where the goal is to classify each pixel in an image into predefined categories. Its architecture is characterized by its efficiency in capturing spatial context while maintaining computational efficiency. The key to its effectiveness lies in its encoder-decoder structure with skip connections. The structure of the LinkNet model is displayed in Fig. 4. In the LinkNet design, the encoder module is responsible for receiving the input image and extracting its hierarchical features, denoted as F. The encoder module primarily comprises multiple convolutional layers accompanied by pooling operations.

These techniques improve the feature maps' depth while progressively reducing their spatial dimensions. This systematic approach enables the network to extract features across various scales, spanning from intricate low-level details to overarching high-level semantics.

In contrast, the decoder module utilizes the features extracted by the encoder to generate the ultimate segmentation map. It usually consists of convolutional layers after upsampling layers, which gradually increase the feature maps' spatial dimensions while decreasing their depth. This procedure improves the segmentation map by reconstructing the spatial information that was lost during the encoding step. One way that LinkNet sets itself apart is by using skip connections to create interconnections between matching layers in the encoder as well as decoder modules. These skip connections enable the transmission of fine-grained details from the encoder to the decoder, allowing the network to efficiently mix low-level and high-level data. Consequently, LinkNet is capable of gathering both local information and global context, which makes it suitable for tasks such as lung cancer classification.

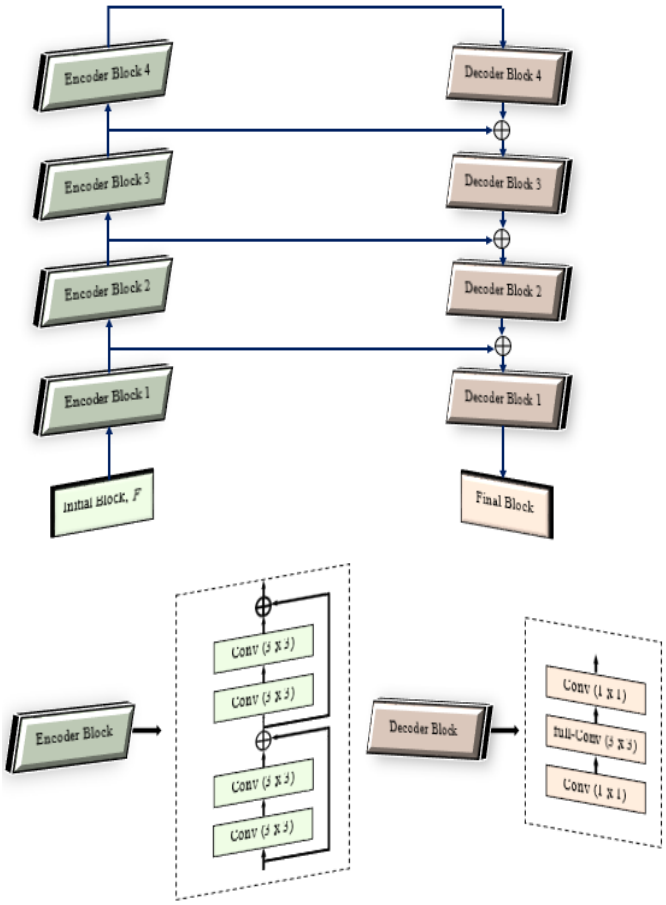

Fig. 4. Architecture of the LinkNet model.

IV. Results and Discussion

A. Simulation Procedure

The proposed lung cancer segmentation and classification were simulated using Python, specifically version "3.7." The simulation was conducted on a system equipped with an "Intel(R) Core(TM) i5-4210U CPU running at 1.70 GHz" and "8.00 GB" of RAM. Furthermore, the analysis for lung cancer segmentation and classification was conducted utilizing the IQ-OTHNCCD lung cancer dataset [24].

B. Dataset Description

The lung cancer dataset, collected from the IQ-OTH/NCCD, was compiled over three months in autumn 2019 at specialized medical centers. This extensive collection comprises CT scans acquired from individuals diagnosed with lung cancer at different stages, alongside scans from individuals without any such diagnosis. Notably, the slides from IQ-OTH/NCCD underwent thorough annotation by oncologists and radiologists at the respective centers. Containing 1190 images, which represent CT scan slices from 110 distinct cases, this dataset presents a varied collection of cases classified into normal, benign, and malignant categories. More precisely, the dataset comprises 40 cases diagnosed as malignant, 15 as benign, and 55 as normal. The original CT scans, acquired using Siemens' SOMATOM scanner, were stored in DICOM format.

The CT protocol employed specific parameters including a 120 kV voltage, a 1 mm slice thickness, and precise window settings spanning from 350 to 1200 HU for window width and from 50 to 600 HU for window center, along with a breath hold at full inspiration. To uphold privacy, all images underwent anonymization before analysis, with the oversight review board waiving the necessity for written consent. Additionally, the study received approval from the institutional review board at the collaborating medical centers. The dataset comprises scans, each containing anywhere from 80 to 200 slices, each offering a distinct view of the human chest. It is crucial to acknowledge the diversity within the 110 cases, encompassing variations in gender, age, education level, area of residence, and occupation. For instance, the subjects range from employees of Iraqi ministries to individuals working in agriculture or other occupations. Geographically, the majority of cases originate from regions in central Iraq, specifically the provinces of Baghdad, Wasit, Diyala, Salahuddin, and Babylon.

C. Performance Analysis

The study conducts a comparative assessment between LSTM+LinkNet and conventional models for lung cancer classification, focusing on various performance metrics including "Accuracy, False Negative Rate (FNR), Specificity, False Positive Rate (FPR), Precision, F-measure, Sensitivity, Matthews Correlation Coefficient (MCC), and Negative Predictive Value (NPV)." Furthermore, the LSTM+LinkNet model is evaluated against several conventional classifiers such as LSTM, LinkNet, SqueezeNet, PyramidNet, and MobileNet. Fig. 5 illustrates both the original images and their processed counterparts after applying Gaussian blur. Notably, the Gaussian blur technique effectively removes noise from the original images. This noise reduction significantly enhances the performance of subsequent segmentation and classification processes.

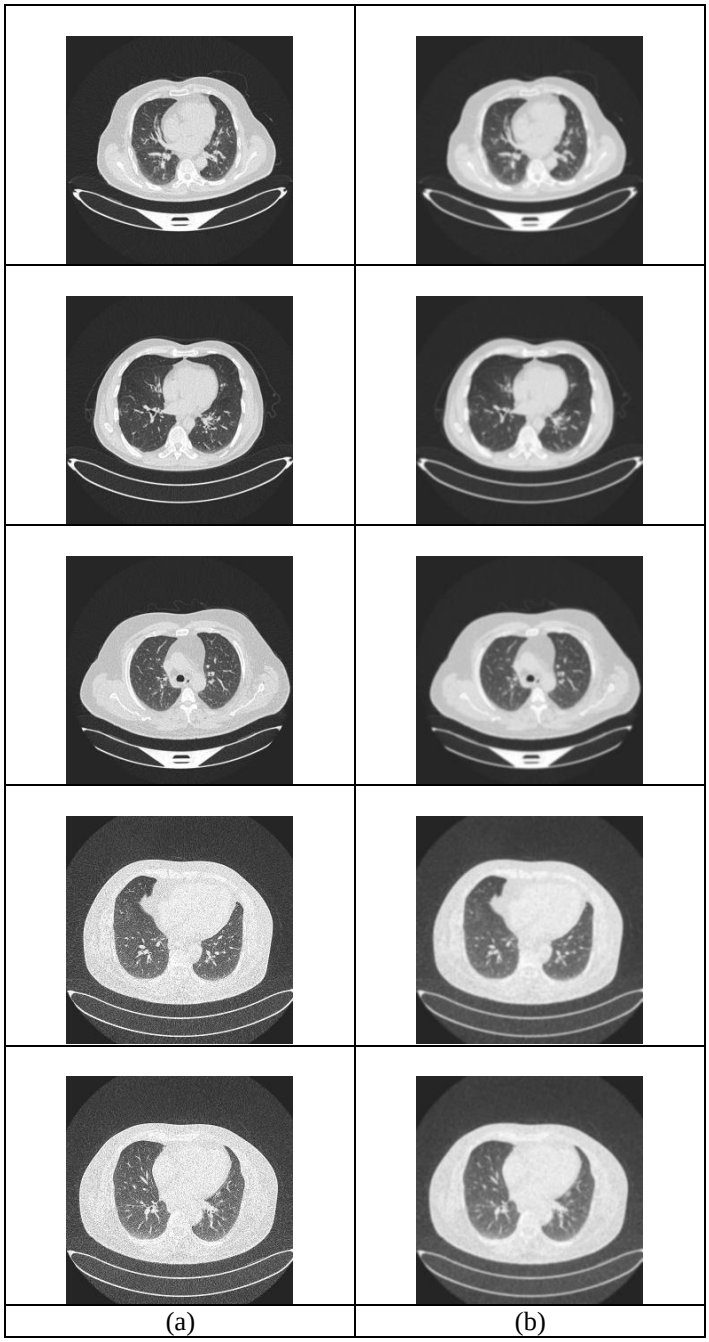

Fig. 5. Images for Lung Cancer Classification a) Original Images and b) Gaussian Blur using Pre-processed images.

D. Segmentation Analysis

In Fig. 6, original images and segmented images are presented, depicting the outcomes obtained using FCM, Conventional SegNet, K-Means, and Improved SegNet. Notably, Improved SegNet showcases superior segmentation results compared to the other methods. Table I presents a segmentation analysis comparing FCM, K-means, Conventional SegNet, and Improved SegNet based on various metrics. These metrics include Dice coefficient, Jaccard coefficient, and Segmentation Accuracy. Improved SegNet stands out as the top performer across all metrics, achieving the highest Dice coefficient of 0.884, Jaccard coefficient of 0.921, and Segmentation Accuracy of 0.933. In contrast, Conventional SegNet also demonstrates respectable performance, albeit lower than Improved SegNet, with scores of 0.800, 0.817, and 0.838 for Dice coefficient, Jaccard coefficient, and Segmentation Accuracy, respectively. FCM and K-means exhibit comparatively lower values across all metrics, indicating less accurate segmentation results. These results emphasize the efficacy of Enhanced SegNet in precisely segmenting regions, showcasing its potential for outperforming other methods assessed in the study in terms of image segmentation tasks.

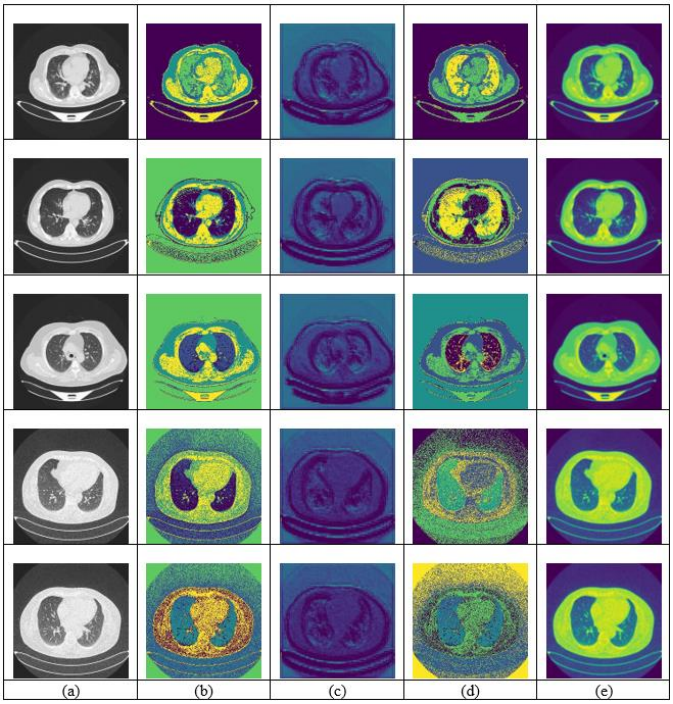

Fig. 6. Images for Lung Cancer Segmentation a) Original Images b) FCM c) Conventional SegNet d) K-means and e) Improved SegNet.

TABLE I. SEGMENTATION ANALYSIS ON IMPROVED SEGNET

Metrics	FCM	Conventional SegNet	K-means	Improved SegNet
Dice	0.635	0.800	0.614	0.884
Segmentation Accuracy	0.685	0.838	0.666	0.933
Jaccard	0.674	0.817	0.660	0.921

E. Comparative Study on various Metrics based on Classification

The study conducts a comparative assessment between LSTM+LinkNet and conventional models for lung cancer classification, focusing on various performance metrics including “Accuracy, False Negative Rate (FNR), Specificity, False Positive Rate (FPR), Precision, F-measure, Sensitivity, Matthews Correlation Coefficient (MCC), and Negative Predictive Value (NPV).” Furthermore, the LSTM+LinkNet model is evaluated against several conventional classifiers such as LSTM, LinkNet, SqueezeNet, PyramidNet, and MobileNet.

1) Analysis on positive measures: In the comparative analysis illustrated in Fig. 7, the LSTM+LinkNet approach is evaluated against LSTM, LinkNet, SqueezeNet, PyramidNet, and MobileNet for lung cancer classification in terms of positive measures. Achieving high positive metric values while minimizing negative ones is crucial for effective classification. The LSTM+LinkNet method consistently outperforms conventional approaches across these metrics. At 60% training data utilization, the LSTM+LinkNet method achieves an accuracy of 0.868, slightly surpassing LSTM, LinkNet, and SqueezeNet. However, as the training data increases, the superiority of LSTM+LinkNet becomes more pronounced. With 90% training data, it reaches an accuracy of 0.935, the highest among all methods, with LinkNet following at 0.911. In contrast, methods like LSTM, SqueezeNet, and PyramidNet, although competitive at lower training data levels, show less consistent accuracy improvements as the data increases.

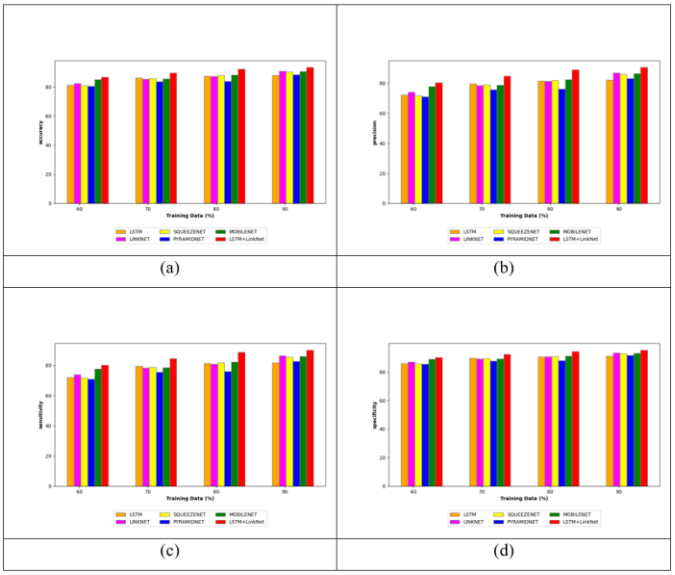

Fig. 7. Positive Metric assessment on LSTM+LinkNet and conventional methods.

Examining specificity metrics at 60% training data reveals several insights. LSTM has a specificity of 0.860, indicating its ability to correctly identify true negatives. LinkNet performs slightly better with a specificity of 0.870. SqueezeNet's specificity is 0.858, while PyramidNet's is 0.855, both closely trailing LinkNet. MobileNet outperforms these with a specificity of 0.889, demonstrating its proficiency in identifying true negatives. Notably, the LSTM+LinkNet model exceeds all conventional methods with a specificity of 0.901, highlighting its superior performance in correctly identifying true negatives in lung cancer classification tasks even with 60% of the training data.

2) Analysis on negative measure: In Fig. 8, a comparative analysis evaluates the LSTM+LinkNet approach against several other methods—LSTM, LinkNet, SqueezeNet, PyramidNet, and MobileNet—for lung cancer classification based on negative measures. When utilizing 80% of the training data, the LSTM+LinkNet method excels in minimizing the False Positive Rate (FPR). It achieves an impressively low FPR of 0.056, outperforming all other methods and effectively reducing false positives even with a significant portion of training data. In contrast, traditional methodologies like LSTM, LinkNet, SqueezeNet, and PyramidNet show slightly higher FPR values, ranging from 0.090 to 0.119.

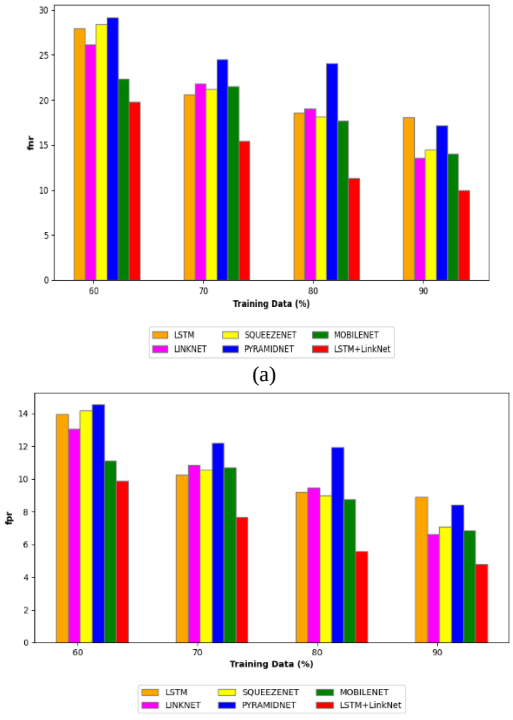

Fig. 8. Negative Metric assessment on LSTM+LinkNet and conventional methods.

While MobileNet demonstrates competitive performance with an FPR of 0.088, it still does not match the performance of the LSTM+LinkNet method. At 90% training data utilization, the models' FNR metrics reveal significant differences in lung cancer classification performance. LSTM has a relatively high FNR of 0.181, indicating a considerable proportion of missed positive cases. Conversely, LinkNet shows a lower FNR of 0.136, effectively reducing false negatives compared to LSTM. SqueezeNet and PyramidNet also perform well, with FNRs of 0.145 and 0.172, respectively, demonstrating their abilities to minimize missed positive cases. MobileNet stands out with an impressive FNR of 0.140, indicating its proficiency in reducing false negatives relative to other conventional methods. However, the LSTM+LinkNet model surpasses all these methods with a notably lower FNR of 0.100, underscoring its superior capability in reducing missed positive cases in lung cancer classification tasks.

3) Analysis on other measures: In Fig. 9, a comparative analysis evaluates the LSTM+LinkNet approach against several other methods—LSTM, LinkNet, SqueezeNet, PyramidNet, and MobileNet—for lung cancer classification based on different performance measures. With 60% of the

training data, the LSTM+LinkNet achieves the highest Negative Predictive Value (NPV) of 0.901, indicating its exceptional accuracy in predicting negative cases compared to the other methods. MobileNet follows closely with an NPV of 0.888, while LSTM, LinkNet, SqueezeNet, and PyramidNet show slightly lower NPV values, ranging from 0.854 to 0.869. As the percentage of training data increases, the LSTM+LinkNet consistently maintains its lead in NPV. At 90% training data, it reaches the highest NPV of 0.950, demonstrating its effectiveness in accurately identifying negative cases even with more extensive training data.

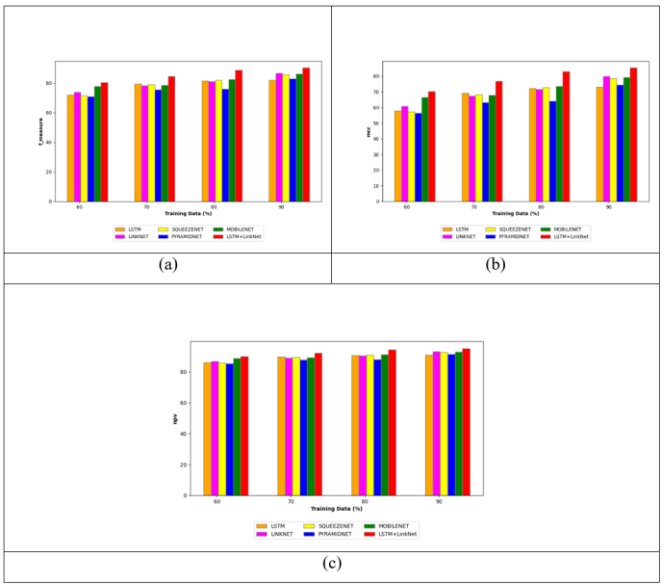

Fig. 9. Other Metric assessment on LSTM+LinkNet and conventional methods.

At 90% training data utilization, each lung cancer classification method exhibits distinct performance based on their F-measure values. The LSTM method achieves an F-measure of 0.821, while LinkNet performs competitively with an F-measure of 0.866, reflecting its effectiveness in correctly classifying lung cancer cases. SqueezeNet closely follows with an F-measure of 0.857, showcasing its balanced performance with significant training data. PyramidNet achieves an F-measure of 0.830, indicating a respectable balance between precision and recall. MobileNet shows improved performance compared to earlier stages, reaching an F-measure of 0.862. However, the LSTM+LinkNet method outperforms all others with the highest F-measure of 0.902, highlighting its superior ability to effectively classify lung cancer cases.

F. Statistical Analysis on Accuracy

The statistical evaluation comparing the LSTM+LinkNet model with LSTM, LinkNet, SqueezeNet, PyramidNet, and MobileNet for lung cancer classification is illustrated in Table II. Among the statistical metrics provided, the maximum statistical metric serves as a crucial indicator of the highest achievable performance level for each classification method in lung cancer detection. In this context, it represents the peak accuracy attained by each method across all trials. For instance, the LSTM+LinkNet attains the maximum accuracy of 0.935, suggesting that it can successfully classify lung cancer cases. When examining the median statistical metric for LSTM+LinkNet and conventional methods in lung cancer classification, distinct performance profiles emerge. With a accuracy of 91.1%, the LSTM+LinkNet model demonstrates its consistency in achieving high levels of accuracy in categorizing lung cancer cases. The LSTM model, with an accuracy of 87.0%, showcases a consistent capability in classifying lung cancer. Similarly, LinkNet and SqueezeNet exhibit competitive performance, with accuracies of 86.4% and 86.9% respectively, suggesting their efficacy in this domain. However, PyramidNet's slightly lower accuracy of 83.9% indicates a comparatively moderate performance level, albeit still within an acceptable range. MobileNet mirrors LSTM's performance with a median accuracy of 87.0%, further highlighting its stability in lung cancer classification tasks.

TABLE II. STATISTICAL ASSESSMENT ON ACCURACY FOR LUNG CANCER CLASSIFICATION

Methods	Minimum	Maximum	Mean	Median	Standard Deviation
LSTM	0.814	0.880	0.858	0.870	0.027
SqueezeNet	0.811	0.905	0.863	0.869	0.035
PyramidNet	0.806	0.886	0.842	0.839	0.029
LinkNet	0.826	0.911	0.866	0.864	0.031
MobileNet	0.852	0.908	0.875	0.870	0.022
LSTM+LinkNet	0.868	0.935	0.906	0.911	0.026

G. Computation Time Analysis

Table III describes the computation time analysis on LSTM+LinkNet and conventional methodologies for lung cancer classification. The computation time examination offers valuable insights into the efficiency of various approaches for lung cancer classification. Among the evaluated methods, the LSTM+LinkNet stands out with the shortest computation time of 62.978s, demonstrating its ability to swiftly process and classify lung cancer cases. In comparison, the LSTM method requires 71.849s, LinkNet takes 78.374s, SqueezeNet demands 85.243s, PyramidNet consumes 92.117s, and MobileNet necessitates the longest computation time at 97.842s. This data suggests that the LSTM+LinkNet method offers notable efficiency gains in processing time compared to several other methods for expedited lung cancer classification tasks.

TABLE III. ANALYSIS ON COMPUTATION TIME

Methods	Computation Time(s)
LSTM	71.849
LinkNet	78.374
SqueezeNet	85.243
PyramidNet	92.117
MobileNet	97.842
LSTM+LinkNet	62.978

V. CONCLUSION

In this work, the proposed method presents a novel and effective approach for lung cancer segmentation and classification, addressing critical challenges in medical imaging analysis. The suggested framework combines a hybrid classifier that integrates LSTM and LinkNet for classification, an Improved SegNet architecture for segmentation, and Gaussian filtering for preprocessing to achieve high accuracy and robustness in identifying and classifying lung cancer nodules from CT images. The experimental results illustrate how well the suggested strategy performs in comparison to current methods, highlighting its potential to help clinicians with lung cancer patients' early diagnosis, treatment planning, and follow-up. With 90% training data, it reaches an accuracy of 0.935, the highest among all methods. LSTM has a specificity of 0.860, indicating its ability to correctly identify true negatives. LinkNet performs slightly better with a specificity of 0.870. SqueezeNet's specificity is 0.858, while PyramidNet's is 0.855, both closely trailing LinkNet. MobileNet outperforms these with a specificity of 0.889, demonstrating its proficiency in identifying true negatives. Notably, the LSTM+LinkNet model exceeds all conventional methods with a specificity of 0.901, highlighting its superior performance in correctly identifying true negatives in lung cancer classification tasks even with 60% of the training data. As the training data increases, the superiority of LSTM+LinkNet becomes more pronounced. With its promising results and significant contributions to the field of computer-aided diagnosis, the proposed method represents a valuable tool for improving patient outcomes and advancing research in lung cancer detection and management.

REFERENCES

[1] Sangeetha S.K.B, Sandeep Kumar Mathivanan, P Karthikeyan, Hariharan Rajadurai, Basu Dev Shivahare, Saurav Mallik and Hong Qin, "An enhanced multimodal fusion deep learning neural network for lung cancer classification", Systems and Soft Computing, Volume 6, December 2024, 200068, doi : 10.1016/j.sasc.2023.200068.

[2] Sampangi Rama Reddy B R, Sumanta Sen, Rahul Bhatt, Murari Lal Dhanetwal, Meenakshi Sharma and Rohaila Naaz, "Stacked neural nets for increased accuracy on classification on lung cancer", Measurement: Sensors, Volume 32, April 2024, 101052, doi : 10.1016/j.measen.2024.101052.

[3] Liangyu Li, Jing Yang, Lip Yee Por, Mohammad Shahbaz Khan, Rim Hamdaoui, Lal Hussain, Zahoor Iqbal, Ionela Magdalena Rotaru, Dan Dobrotă, Moutaz Aldrdery and Abdulfattah Omar, "Enhancing lung cancer detection through hybrid features and machine learning hyperparameters optimization techniques", Heliyon, Volume 10, Issue 4, 29 February 2024, e26192, doi : 10.1016/j.heliyon.2024.e26192.

[4] Fuli Zhang, Qiusheng Wang, Enyu Fan, Na Lu, Diandian Chen, Huayong Jiang and Yanjun Yu, "Enhancing non-small cell lung cancer tumor segmentation with a novel two-step deep learning approach", Journal of Radiation Research and Applied Sciences, Volume 17, Issue 1, March 2024, 100775, doi : 10.1016/j.jrras.2023.100775.

[5] Lavina Jean Crasta, Rupal Neema and Alwyn Roshan Pais, "A novel Dezp Learning architecture for lung cancer detection and diagnosis from Computed Tomography image analysis", Healthcare Analytics, Volume 5, June 2024, 100316, doi : 10.1016/j.health.2024.100316.

[6] A. Angel mary and K.K. Thanammal, "Lung cancer detection via deep learning-based pyramid network with honey badger algorithm", Measurement: Sensors, Volume 31, February 2024, 100993, doi : 10.1016/j.measen.2023.100993.

[7] A. Gopinath, P. Gowthaman, M. Venkatachalam and M. Saroja, "Computer aided model for lung cancer classification using cat optimized convolutional neural networks", Measurement: Sensors, Volume 30, December 2023, 100932, doi : 10.1016/j.measen.2023.100932.

[8] Yongchun Cao, Liangxia Liu, Xiaoyan Chen, Zhengxing Man, Qiang Lin, Xianwu Zeng and Xiaodi Huang, "Segmentation of lung cancer-caused metastatic lesions in bone scan images using self-defined model with deep supervision", Biomedical Signal Processing and Control, Volume 79, Part 1, January 2023, 104068, doi : 10.1016/j.bspc.2022.104068.

[9] Rehan Raza, Fatima Zulfiqar, Muhammad Owais Khan, Muhammad Arif, Atif Alvi, Muhammad Aksam Iftikhar and Tanvir Alam, "Lung-EffNet: Lung cancer classification using EfficientNet from CT-scan images", Engineering Applications of Artificial Intelligence, Volume 126, Part B, November 2023, 106902, doi : 10.1016/j.engappai.2023.106902.

[10] Negar Maleki and Seyed Taghi Akhavan Niaki, "An intelligent algorithm for lung cancer diagnosis using extracted features from Computerized Tomography images", Healthcare Analytics, Volume 3, November 2023, 100150, doi : 10.1016/j.health.2023.100150.

[11] I. Naseer, S. Akram, T. Masood, M. Rashid and A. Jaffar, "Lung Cancer Classification Using Modified U-Net Based Lobe Segmentation and Nodule Detection", IEEE Access, vol. 11, pp. 60279-60291, 2023, doi: 10.1109/ACCESS.2023.3285821.

[12] R. Mahum and A. S. Al-Salman, "Lung-RetinaNet: Lung Cancer Detection Using a RetinaNet With Multi-Scale Feature Fusion and Context Module", IEEE Access, vol. 11, pp. 53850-53861, 2023, doi: 10.1109/ACCESS.2023.3281259.

[13] Hyung Min Kim, Taehoon Ko, In Young Choi and Jun-Pyo Myong, "Asbestosis diagnosis algorithm combining the lung segmentation method and deep learning model in computed tomography image", International Journal of Medical Informatics, Volume 158, February 2022, 104667, doi : 10.1016/j.ijmedinf.2021.104667.

[14] Gregory Z. Ferl, Kai H. Barck, Jasmine Patil, Skander Jemaa, Evelyn J. Malamut, Anthony Lima, Jason E. Long, Jason H. Cheng, Melissa R. Junttila and Richard A.D. Carano, "Automated segmentation of lungs and lung tumors in mouse micro-CT scans", iScience, Volume 25, Issue 12, 22 December 2022, 105712, doi : 10.1016/j.isci.2022.105712.

[15] Stine Hansen, Samuel Kuttner, Michael Kampffmeyer, Tom-Vegard Markussen, Rune Sundset, Silje Kjærnes Øen, Live Eikenes and Robert Jenssen, "Unsupervised supervoxel-based lung tumor segmentation across patient scans in hybrid PET/MRI", Expert Systems with Applications, Volume 167, 1 April 2021, 114244, doi: 10.1016/j.eswa.2020.114244.

[16] Kusrini, Muhammad Resa Arif Yudianto and Hanif Al Fatta, "The effect of Gaussian filter and data preprocessing on the classification of Punakawan puppet images with the convolutional neural network algorithm", International Journal of Electrical and Computer Engineering (IJECE), Vol. 12, No. 4, August 2022, pp. 3752-3761, ISSN: 2088-8708, doi: 10.11591/ijece.v12i4.pp3752-3761.

[17] Vijay Badrinarayanan, Alex Kendall and Roberto Cipolla, "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation", IEEE Transactions on Pattern Analysis and Machine, Intelligence (Volume: 39, Issue: 12, 01 December 2017), doi : 10.1109/TPAMI.2016.2644615.

[18] Brahim Ait Skourt, Abdelhamid El Hassani and Aicha Majda, "Mixed-pooling-dropout for convolutional neural network regularization", Journal of King Saud University – Computer and Information Sciences 34 (2022) 4756–4762, doi : 10.1016/j.jksuci.2021.05.001.

[19] Chigozie Enyinna Nwankpa, Winifred Ijomah, Anthony Gachagan, and Stephen Marshall, "Activation Functions: Comparison of Trends in Practice and Research for Deep Learning", arXiv:1811.03378v1 [cs.LG] 8 Nov 2018.

[20] Ma Xin and Jing Xiaojun, "Palm vein recognition method based on fusion of local Gabor histograms", The Journal of China Universities of Posts and Telecommunications, December 2017, 24(6), pp: 55–66, doi : 10.1016/S1005-8885(17)60242-5.

[21] Adel Hafiane, Kannappan Palaniappan and Guna Seetharaman, "Adaptive Median Binary Patterns for Texture Classification", 2014 22nd International Conference on Pattern Recognition, 2014.

[22] Haipeng Xiao, Miguel Angel Sotelo, Yulin Ma, Bo Cao, Yuncheng Zhou, Youchun Xu, Rendong Wang And Zhixiong Li, "An Improved LSTM Model for Behavior Recognition of Intelligent Vehicles", Special Section On Big Data Technology And Applications In Intelligent Transportation, IEEE Access, Volume 8, PP: 101514-101527, 2020.

[23] Abhishek Chaurasia; Eugenio Culurciello, "LinkNet: Exploiting encoder representations for efficient semantic segmentation," 2017 IEEE Visual Communications and Image Processing (VCIP), St. Petersburg, FL, USA, 2017, pp. 1-4, doi: 10.1109/VCIP.2017.8305148.

[24] https://data.mendeley.com/datasets/bhmdr45bh2/1.