VFAST Transactions on Software Engineering
https://vfast.org/journals/index.php/VTSE@ 2026, ISSN(e): 2309-3978, ISSN(p): 2411-6246
Volume 14, Number 1, January-March 2026 pp:193-205

AI-Driven Deep Learning for Lung Segmentation in CT Imaging

Tehreem Awan[1*], **Ali Raza**[1], **Inzimam Ul Haq**[1], **M. Usman Abdul Manan**[1], **Nasrullah**[1], **Serosh Karim Noon**[1]

[1]Department of Electrical Engineering, NFC Institute of Engineering & Technology Multan, Pakistan

Keywords: *Lung segmentation, U-Net architecture, CT scans, Deep learning, Dice Similarity Coefficient (DSC)*

Journal Info:
Submitted:
October 18, 2025
Accepted:
March 15, 2026
Published:
March 20, 2026

Abstract With the remarkable growth in the field of artificial intelligence; fully automated diagnostic systems are no longer a dream. Recent advancements in medical imaging have paved an early and accurate pulmonary complications. This study presents an updated approach of lung nodule segmentation from CT (computed tomography) scans using an improved UNet architecture. The performance of model is enhanced with ResNet-34 and EfficientNet backbone. After segmenting, feature extraction and subsequent classification is performed to classify the type of chest cancer. A publicly available dataset comprising 1,000 CT scan images available on Kaggle is used. The core of our segmentation model is built upon the U-Net architecture, renowned for its efficacy in biomedical image segmentation tasks. Implemented using Python and TensorFlow within the Google Colab environment, the model trains to accurately classify 3 types of chest cancers. The performance is evaluated by metrics such as the Intersection over Union (IoU) to assess segmentation quality. To further enhance the modelś robustness and generalizability UNet architecture integrated with hybrid ResNet-34 and EfficientNetB4 backbone encoder which achieves a dice score of 0.66 with better performance on small nodules i.e., 0.58. The hybrid backbone ensures efficient feature extraction, contributing to improved segmentation results. The proposed model employed multiphase optimization which further enhances its potential for AIdriven solutions in medical diagnostic. The automated segmentation and classification can aid clinicians in the diagnosis and detection of pulmonary abnormalities, including COVID-19.

***Correspondence author email address:** tehreemawannfc@gmail.com
DOI: 10.21015/vtse.v14i1.2263

1 Introduction

With recent advancements in the field of artificial intelligence, medical imaging based accurate diagnostic tools have transformed the field of medical imaging. Lung cancer [4] remains one the most commonly occurring and deadly cancer worldwide. CT scans are the common tools for radiologists and oncologist to monitor lung nodules, infections and fibrosis. Image segmentation based computer-aided diagnosis (CAD) systems can be handy for radiologists while accurately identifying lung regions [18].

Traditional segmentation methods often struggle with pathological variations such as tumors, pulmonary infiltrates, or collapsed lung tissues; specially in dense areas. These challenges necessitate the integration of deep learning approaches, that offer better performance in capturing intricate anatomical features [1].For the last few years artificial intelligence has captured the interest of computer scientists [3]. Automated diagnostic tools require machine learning based approaches to

perform advanced image analysis.

Cancer is the abnormal growth of cells in the body; where the malignancies may spread to the other body systems [4]. Every year, it causes 1.8 million deaths around the world. Lung cancer originates in the pulmonary tissues and mutate silently and in an aggressive manner. This causes the early detection impossible as the disease is already at advanced stage when the symptoms occurs. The non-small cell lung cancer (NSCLC) is further classified as adenocarcinoma, squamous cell carcinoma, and large cell carcinoma, all of which present unique visible symptoms on scans [18]. In this study, we work with a diverse publicly available dataset containing CT scans of three subtypes of NSCLC as well as normal lung tissues.

Traditional diagnosis requires accurate segmentation of lung areas for nodule identification, assessing tumor growth, and response to treatment. Radiologists require high quality images and expert interpretation skills, even then they may struggle with complex features [17]. In this scenario, imaging based artificial intelligence based methods are the game changer in addressing the challenges.

Segmentation is an crucial step in medical image diagnosis and classification for radiological analysis [18]. Computer-aided diagnosis (CAD) [17] of lung CT requires to first segment the area of interest i.e., nodule or mass and then analyze the isolated area for diagnosis for disease and progression. For a normal healthy lung, segmentation can be carried out using better contrast between background and surrounding tissues [8]. However, this approach falls short when the lung is infected by high density pathology.

At the same time, imaging variability, subject positioning and scanner settings can cause inconsistencies in scan quality. The performance of the segmentation model can further be influenced by noise and artifacts which can faint nodules and can result in false positives. Symptoms variability among and between patients also pose challenges in lung nodule segmentation [15]. Also the nodule size, shape and texture variability makes it difficult for segmentation algorithms to consistently spot and highlight nodules [9].

Moreover, varying types of nodules appear differently on CT scans because each type has peculiar characteristics. Segmentation [13] of large solid nodules is not complex whereas the segmentation of small nodules attached to the vessels or lung tissue or diaphragm can be challenging.

In this work, we have proposed a lung segmentation technique which accurately segment the lung parenchyma from lung CT Scan images.

The study proposes a deep learning- based lung segmentation system using a modified U-Net architecture having an enhanced backbone encoder. The model is further optimized for CT imaging. To make the model robust and generalizable, availability of high-quality annotated data for training is inevitable. For the purpose of this, we manually annotated 500 publicly available CT scan images using the Roboflow platform. The entire labeling process was carried out under the supervision of expert radiologist.

While performing AI based segmentation [12], the Intersection over Union (IoU) score serves as an important metric for evaluating the accuracy of the model's segmentation predictions versus manually annotated ground truth masks [10]. The IoU score numerically characterize the overlap between the predicted nodule shape and the radiologist's annotated ground truth. The area of intersection of both are divided by the area of union of the same. Where a score of 1 indicates 100% true similarity of result with ground truth, while a 0 signifies no overlap. Literature suggests that an IoU score greater than or equal to 0.85 is typically considered clinically acceptable [7].

The proposed segmentation model is built upon a UNet framework [1], enhanced with EfficientNet [6] and ResNet-34 [21] backbone encoders to improve feature extraction and accuracy. It is trained using advanced preprocessing, hybrid loss functions, and attention mechanisms to refine predictions. Model performance is rigorously evaluated using key metrics such as the Dice Similarity Coefficient (DSC) and Intersection over Union (IoU), confirming high segmentation precision [20].

The study hence demonstrates the potential of machine learning based tools in medical diagnostics, offering a scalable, accurate, and economical solution for lung segmentation in CT scans, especially in settings where expert annotated data is limited.

1.1 U-Net

The U-Net [2] is a semantic segmentation architecture applying convolutional neural network (CNN) used for the tasks. It is widely used in medical imaging [13] as it trains well on limited data labels. The architecture has two key parts down-sampling (encoder) and up-sampling (up-sampling). In the encoder part, feature extraction is carried out by convolutional and pooling layers. Whereas, decoder part up-samples the feature map by performing de-convolutional operation. [2].

The gray arrows in Figure 1indicate the skip connections preserves the spatial information by concatenating the feature map with the decoder. Which is very useful while recovering boundary information in segmentation task.

The encoder fetches an input X with C channels to perform convolution on it.The resultant output shape of

$$(C_outH/2, W/2)$$

. Additional convolutions are carried out at the bottleneck stage to convert narrow spatial resolution to an intermediate representation $X_mid = ReLU(Conv2D(X_enc))$. The decoder then performs up-sampling operation using bilinear interpolation and convolution. Finally, a 11 convolution layer minimizes the output channels to one, thereby producing the binary segmentation mask [2].

The skip connections preserve the spatial information lost during down-sampling operation by reusing feature maps from earlier layers. This makes U-Net effective for handling handling small structures, complex shapes, and varying contrasts in medical images [3]. Moreover, the symmetric encoder-decoder design effectively captures both global context and local details. The details are essential for accurate segmentation in CT lung images. Topology of UNet is shown in Figure 1.

1.2 ResNet

With the development of deep convolutional networks, their ability to perform feature extraction accurately has increased. This, in turn leads to increasingly deeper models with more layers stacked to achieve better training. However, deep networks suffers from the problem of vanishing gradient. Because, as the number of layers increases, the gradients used for backpropagation approach zero or explode. Residual Network, ResNet addresses this by introducing a novel architecture based on Residual Blocks [19].

A residual learning unit is built around the idea of direct connection (skip connections) between input and output layer, hence, protecting the integrity of information [11]. Instead of forcing every layer to learn a direct mapping H(x), the network is reformulated to learn a residual function as shown in equation 1.

$$F(x) = H(x) - x \tag{1}$$

So the final output of residual block is y(x)

$$y(x) = F(x) + x \tag{2}$$

Here, Equation 2 represents the residual mapping, while x is the input or block added through the skip connection. This skip connection allows the layers to focus on learning only the part added to it, preventing vanishing gradient.

The liberty the skip connections provide is that if certain layers adversely affect the performance, the network can effectively "skip" them. This leads to better regularization and prevents degradation in accuracy as the network depth increases. With this design, Deeper ResNets successfully enables generalized performance architectures with 50, 101, or even 152 layers, which was not feasible with traditional convolutional networks [13]. The architecture of ResNet is shown in Figure 2.

2 Literature Survey

The progress in deep learning is moving toward the development of automatic algorithms that assist in biomedical segmentation. In [21], multi-scale deep residual networks are used. They employ skip connections to address the issue of the vanishing gradient. Their results show moderate dice score values.

In several studies, the technique of multi-scale deep residual networks has been extended to 3D models using convolutional neural networks (CNNs). However, this enhancement in the model increases the demand for higher computational efficiency and resources. In a few other studies, multi-scale deep residual networks are integrated with generative adversarial networks (GANs) to improve performance metrics such as the Dice score and focal loss. This approach also helps to address class imbalance issues in biomedical image segmentation [5].

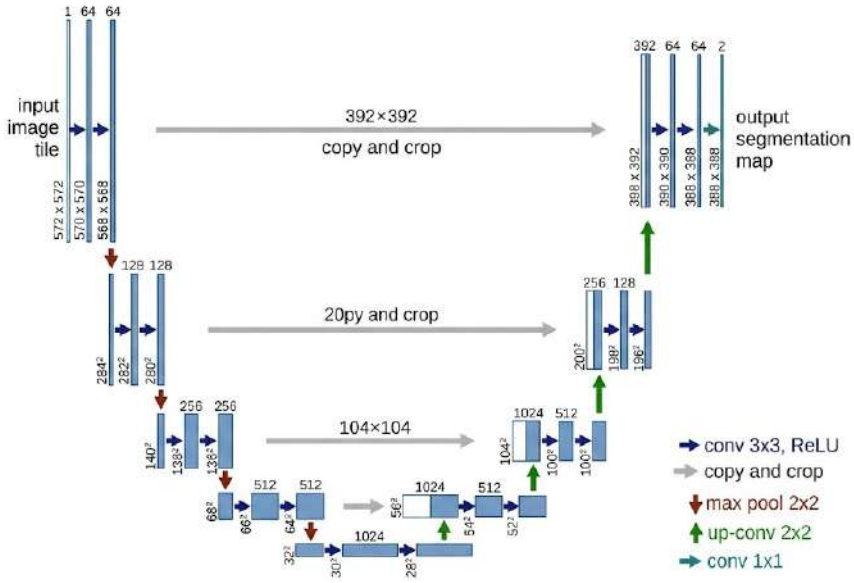

Figure 1. U-net Architecture [16].

The architecture of U-Net is widely used in segmentation techniques in multiple applications beyond biomedical imaging. U-Net employs an encoder–decoder structure along with skip connections, which has shown promising results in segmentation-based tasks. Hybrid techniques have also been incorporated into the architecture to improve the generalization capability of segmentation models in various applications, including biomedical imaging. Furthermore, the use of pretrained weights through transfer learning techniques greatly assists in improving feature extraction and improving the identification of relevant regions of interest [19].

The details of the literature survey related to this work are presented in Tables 1 and Table 2, which show a comparison of the methodologies and results of previously published studies in the field. The results are presented in two forms: some studies report outcomes after the classification of segmented regions, while others limit their work to the segmentation stage only. The classification results are compared based on the parameters used for classification. The proposed work aims to overcome several limitations and challenges identified in previously published research.

3 Methodology

This study presents a segmentation algorithm for the automatic detection and classification of lung nodules from CT scan images. The proposed algorithm is based on an enhanced U-Net architecture with different neural networks used as the backbone. In the first step, masking of the annotated regions was performed. The study utilized 500 annotated CT scan images, and their corresponding masks were created using the Roboflow platform. During the training phase, the dataset was divided into training, testing, and validation sets. Several preprocessing techniques, such as feature enhancement, were applied to improve model optimization and performance.

In the core structure of the proposed algorithm, the U-Net architecture is enhanced by incorporating EfficientNet-B4 and ResNet-34 as backbone networks. Both networks are used separately, and pretrained weights are utilized to improve feature extraction and model performance [6].

This enhanced technique utilizes a hybrid version of U-Net combined with ResNet and EfficientNet architectures. The integration of these networks helps improve the segmentation and enhancement of affected regions in the lungs [14]. The model optimizes the output using an appropriate loss function to improve the overall effi-

Table 1. Literature

Year	Methodology	Results/Outputs	Data Modality/Availability
2018 Zhang, Y. et al.	MDRN	Effective nodule segmentation	CT Images (Public datasets)
2019 Liu, X. et al.	MDRN with data augmentation	Improved robustness in nodule detection	CT Images (Open access)
2020 Wang, H. et al.	Enhanced MDRN with attention mechanism	High accuracy in comp lex nodule segmentation	CT Images (Limited access)
2021 Chen, S. et al.	3D MDRN	3D segmentation results for lung nodules	3D CT Images (Public datasets)
2022 Gupta, A. et al.	MDRN integrated with adversarial training	Robustness against noise in segmentation	CT Images (Open access)
2023 Kim, J. et al.	Improved MDRN with hybrid loss function	Enhanced segmentation performance	CT Images (Limited access)

Table 2. Literature

Year	Methodology	Results/Outputs	Data Modality/Avalibility
2018 Liu, X. et al.	U-Net Architecture	Accurate lung segmentation	CT Images (Public datasets)
2019 Hwang, D. et al.	3D CNN with transfer learning	Improved nodule detection and segmentation	CT Images (Limited access)
2020 Zhang, Y. et al.	Deep learning ansemble model	Robust segmentation results	CT Images (Open access)
2021 Chao, S. et al.	Hybrid model (CNN + RNN)	Enhanced segmentation accuracy	CT Images (Public datasets)
2022 Gupta, A. et al.	GAN for Synthetic data augmentation	Better training performance	CT Images (Open access)
2023 Kim, J. et al.	Attention U-Net	High-quality segmentation and visualization	CT Images (Limited access)

ciency and accuracy of the segmentation process [8].

3.1 Annotation and Mask Generation of Affected Lung Regions

The mask creation process for segmentation involves the identification and annotation of affected regions as specified by radiologists [14]. Most of the dataset used in this study was obtained from online repositories; however, some of the images were initially unannotated. A radiologist assisted in marking the exact regions of interest in these images to generate accurate segmentation masks. Additionally, several preprocessing steps were performed before the training process to improve the quality and consistency of the data. Roboflow is the tool used for annotation [15].

3.1.1 Challenges & Solutions

The challenges and their solutions are tabulated in Table 3.

Table 3. Challenges and Solutions

Challenge	Solution
Small tumors missed	Use tile-based processing (512*x*512 patches)
Inconsistent annotations	Inter-annotator IoU checks (target > 0.8)
Class imbalance	Apply Dice loss in U-Net

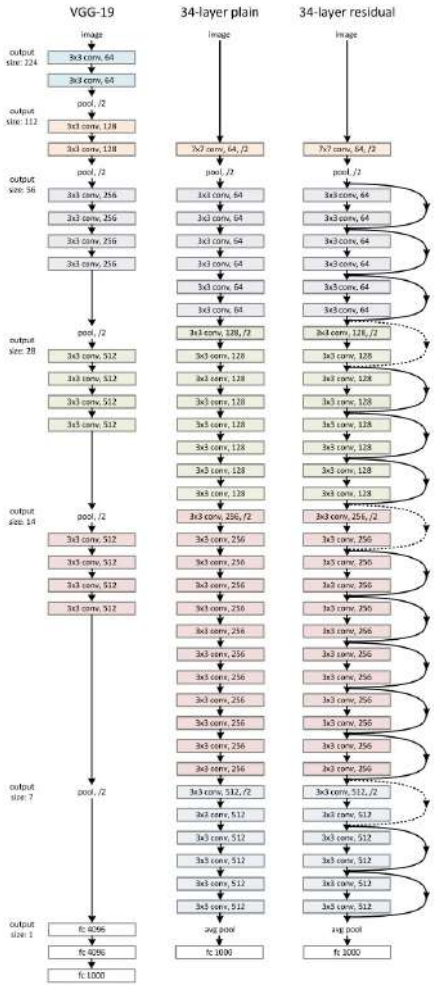

Figure 2. ResNet Architecture [17]

3.2 Advanced Preprocessing Pipeline

Our comprehensive preprocessing workflow incorporated multiple stages of image optimization. First, we applied lung windowing (level 600 HU, width 1500 HU) followed by body windowing (level 40 HU, width 400 HU) to enhance different tissue contrasts.

3.3 Enhanced U-Net Architecture Specifications

The modified architecture incorporates dilated convolutions in the final blocks of the neural network, which help enhance the resolution of the segmented output. Skip connections are used in the decoding sequence to preserve spatial information and improve feature propagation. Additionally, the sigmoid activation function is applied in the final layer to facilitate accurate feature extraction and segmentation [11]. The model consists of approximately 29 million trainable parameters; therefore, a GPU was utilized to ensure efficient and effective training of the network.

The optimization of the algorithm was performed in two phases. In the first phase, the model was trained for 50 epochs while the initial 50 layers of the backbone neural network were kept frozen. In the second phase, these backbone layers were unfrozen to allow fine-tuning of the network. This approach resulted in improved feature representation and better model performance. During the validation phase, particular attention was given to the false positive rate, and clinical evaluation comparisons were also conducted to further assess and enhance the performance of the model.

3.4 Clinical Validation

A clinical validation study was also carried out to check the applicability of our study in real world. Clinical validation study with new cases processed through our self collected images from hospitals. The model however, demonstrated consistent performance with a slight drop in Dice score specially for small nodules.

3.5 Block & Flow Diagram

The block diagram of lung segmentation and respective classification using CT imaging is shown in Figure 3 and the flow of the complete process is shown in Figure 4.

3.5.1 Dataset

A public dataset was used which comprises 2D chest CT scans used for the purpose of training, validation and testing.

3.5.2 Mask Creation

A labeled mask was created following a rigorous annotation which was carried out in collaboration with thoracic radiologists.

3.5.3 Model Architecture

The model architecture is a 3D ResNet-50 backbone pretrained on ChestX-ray dataset, with modified wights on several pulmonary nodule analysis. The attention mechanism used with squeeze & excitation effectively highlight relevant nodule features by suppressing background parenchyma at the same time. Multiscale receptive field ensures to capture important features

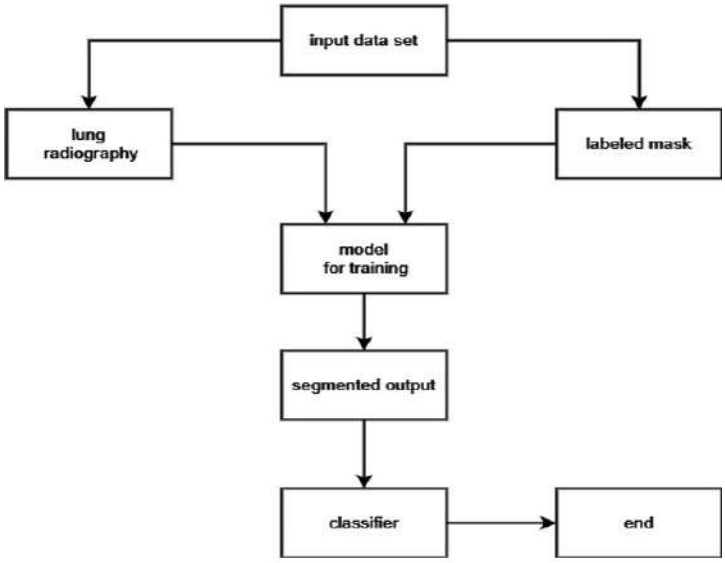

Figure 3. Block Diagram of the proposed AI-based lung segmentation

which is carried out by using dilated convolutions having rates of 1, 2, 4, and 8.

The training process employs a 5-fold cross-validation with effective training of backbone layers. This model uses transfer learning benefits by training it on large datasets and incorporating specific lung cancer dataset for the specific purpose [13].

3.5.4 Segmented Output

The segmented output are then classified as healthy or containing carcinomous cells.

As shown in the Figure 4 the process begins with project initiation, where the research objectives are clearly defined keeping in mind the clinical requirements. This phase establishes the scope of lung nodule analysis, determining key parameters such as target nodule size range (3-30mm), inclusion criteria for CT scans, and performance benchmarks.

A multidisciplinary team including radiologists, AI engineers, and data scientists collaborates to outline annotation protocols, ethical considerations for patient data usage, and technical specifications for the processing pipeline. Regulatory compliance is verified, ensuring adherence to HIPAA/GDPR standards for medical data handling.

The project charter documents baseline metrics from existing clinical workflows, setting improvement targets for AI-assisted nodule detection accuracy and reporting time reduction.

3.5.5 Detailed Description of the Dataset

The dataset includes 800 chest CT scans. These scans were collected from patients undergoing different types of screening. Among them, 60% underwent standard-dose scans (120140 kVp), 30% underwent lowdose screening examinations with doses of 100120 kVp, and the remaining 10% were followup cases of patients who had already been diagnosed.

Participant demographics were represented in a balanced form. The participants belonged to age groups ranging from 20 to 80 years and included variations across gender and smoking history. The dataset used for training contained complete demographic information and diverse age groups to ensure better applicability in real-world deployment.

3.5.6 Mask Generation and Representation

The creation of segmented masks from annotated images with precise numerical values is one of the most important tasks in segmentation. Initially, with the assistance of a radiologist, the affected area and its dimensions are identified to accurately specify the region of interest. Subsequently, masks are created for that specific area for training purposes.

masks precisely represent the area and dimensions of the nodule with the help of its outer boundary. These

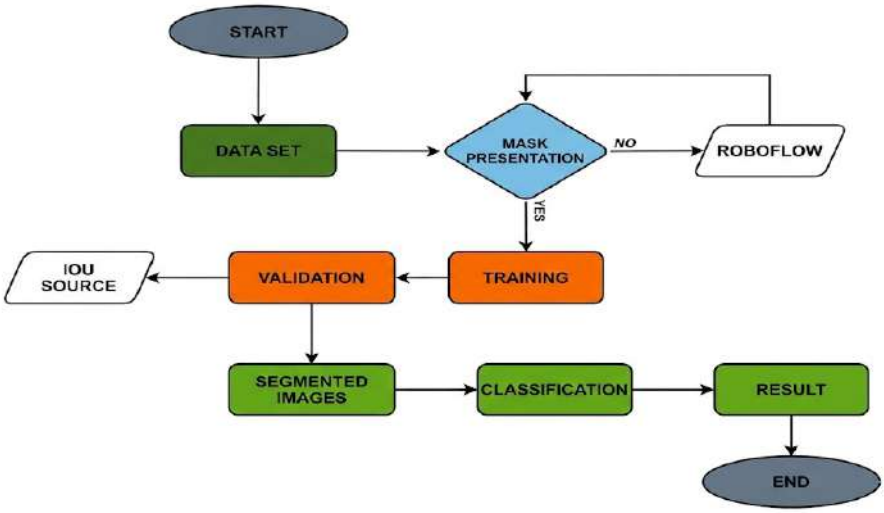

Figure 4. Flow Chart of the Process of Segmentation.

masks constitute a core contribution of this research. The authors created all the masks using Roboflow and consulted a radiologist whenever necessary. The generated masks were stored in a repository and used for training the network. After mask creation, all images were cross-verified to ensure transparency and accuracy, thereby helping to minimize errors in the classifier.

3.5.7 Traning

The training of the algorithm is one of the most important aspects of our contribution and was conducted in three stages. In the first phase, a learning approach was applied where the backbone weights were frozen to create a feature detector, and this phase was executed over 50 epochs. In the second phase, 100 epochs were completed with the upper layers unfrozen and a composite loss function employed. In the third and final phase (the last 50 epochs), the learning process was completed. A patch sampling technique was used to help balance the data. Training was performed on a GPU to ensure smooth execution. The model weights were updated exponentially, contributing to improved training performance. Additionally, an early stopping criterion was applied if no visible improvement was observed over 15 epochs [14].

3.5.8 Validation

In the validation phase, a multi-layer evaluation approach was used. In the first stage, the Dice Similarity Coefficient (DSC) was employed to evaluate segmentation accuracy. The evaluation was performed using test samples as a reference. Additionally, the Intersection over Union (IoU) score was used as another reference metric for validation.

3.5.9 Segmented Output Images

After the segmented output is generated, it passes through a detailed post-processing stage. This stage includes morphological smoothing, which removes speckle noise while preserving the edges. It also involves filtering and highlighting low-confidence regions. Additionally, the system maintains the results of images that are not properly annotated to support error analysis and continuous learning.

3.5.10 Classification

At the end of the segmentation process, the classification of the segmented region is performed to determine whether the region is malignant or benign. This step includes feature extraction, after which classification is carried out using the pretrained neural network ResNet. This stage further clarifies and refines the final output.

3.5.11 Result

The results are displayed in the form of segmented outputs along with their corresponding classification.

4 Experimentation

The experimentation of the proposed work is carried out using the segmented data generated with Roboflow. Both reference data and segmented data are used for training. Two different folders with identical entry names have been created, representing images along with their corresponding masks. Polygon-based segmentation is used to accurately label the targeted regions. The dataset is divided in 70%,20%, and 10% sets for training, validation, and testing, respectively. The architecture used for training is U-Net with ResNet as the backbone network. The highest Intersection over Union (IoU) score achieved is 0.49, which demonstrates the potential of the algorithm. All experiments were performed using TensorFlow.

4.1 Dataset Description and Annotation Process

The dataset used in this study is obtained from a publicly available repository on the Kaggle platform. The repository contains 1,000 chest CT scan images. The mask creation process is performed on 550 annotated images. Mask creation is a lengthy process, and the team vigilantly creates all the masks using the online available tool Roboflow. The dimensions of the affected areas are available in an Excel datasheet, and by closely monitoring all the specifications, the masks are created.

The masks are generated on the lung boundary and the nodule region. In cases of malignant areas, the lesion size is mostly large and irregular. Therefore, all areas are carefully monitored to ensure accurate boundaries according to the radiologist's annotations. After the creation of masks, preprocessing techniques are applied for normalization to maintain consistent standards for annotation.

4.2 Dataset Splitting

The dataset, including images and their corresponding masks, is divided into three separate folders for images and masks. The study uses 550 images for this experimentation, and the data is divided into training, validation, and testing sets. The details of the division are as follows:

- Training Set: 385 images (70%)
- Validation Set: 110 images (20%)
- Test Set: 55 images (10%)

A separate dataset is also created to evaluate the functionality of the algorithm on data that is not part of the training process. This dataset is used for external validation.

4.3 Model Configuration

The U-Net architecture with ResNet as the backbone is used for image segmentation, as U-Net performs better in segmentation tasks. The configuration of the proposed algorithm is as follows.

- Input Size: 256 × 256 pixels
- Loss Function: Binary Cross Entropy + IoU Loss
- Optimizer: Adam
- Learning Rate: 0.001
- Epochs: 100
- Batch Size: 8

4.4 Experimental Setup

The experiments are conducted using TensorFlow and Python, leveraging a GPU provided by the online platform Google Colab.

4.5 Evaluation Metric

Evaluation is a crucial step to verify the performance of the algorithm. In the case of segmentation, the Intersection over Union (IoU) score is the most important metric to assess the accuracy of the segmented results. The IoU score measures the overlap between the segmented mask and the annotated reference regions. The IoU score ranges from 0 to 1, where 0 indicates no match or overlap, and 1 represents a perfect overlap. In practical cases, the threshold is usually maintained around 0.8 to 0.85.

4.6 Experimental Results

The results obtained after training are tabulated below in Table 4.

Table 4. The Metric calculated results of Algorithm after training

Metric	Value
Max IOU Score	0.49
Average IOU	0.42
Dice	0.66
Training Loss	0.58
Validation Loss	0.61
Best Epoch	86

While the achieved IoU is below clinical standards, it still reflects meaningful learning in the early stages of model development, especially given the limited dataset size.

5 Sample Image

The sample of Dataset used for creating masks are shown in Figure 5 and Figure 6. The same data is used for classification [13].

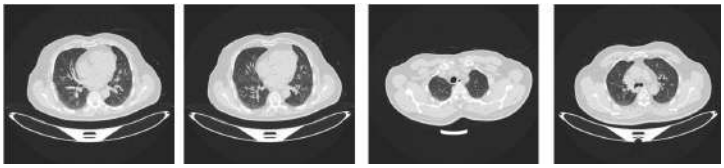

Figure 5. CT-Scan Images.

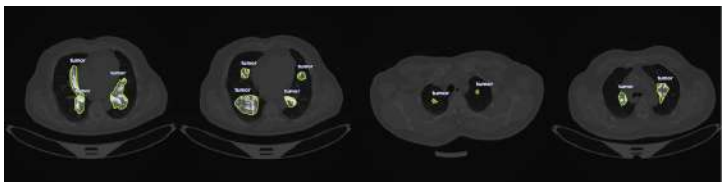

Figure 6. Masks of CT-Scans generated on Roboflow.

6 Discussion

The model shows progressive learning during the training process, with the reduction in loss serving as an example of this behavior. The early stopping criterion helps prevent overfitting. The model achieves an IoU score of 0.49 on the test dataset. Although this value does not reach the levels reported in most published studies, given the limited dataset, the model demonstrates progressive behavior. Its performance can be enhanced by increasing the dataset size.

Furthermore, although the results do not reach the highest benchmarks, the outputs on validated images show that the masks are aligned with the affected regions. This reflects progressive learning, which can be further improved using a larger dataset and enhanced techniques.

The results are also validated using external validation technique by collecting and annotating self collected data from local environment.

The results clearly show that the proposed algorithm, even with a limited dataset and resources, provides a foundation for progress in this field. The algorithm accurately identifies lung boundaries, and the learning rate is progressive. With continued efforts, the algorithm demonstrates further improvements.

7 Representation of Results

Output results have been represented graphically in Figure 7, in which the comparison of IOU score with Epoch and Loss with Epoch is clearly showing the response. The Table 5 is showing the training and validation response. The response on two test images with highlighted regions and segmented results are shown in Figure 8 and Figure 9. The red bar chart is showing the probability of existence of tumor.

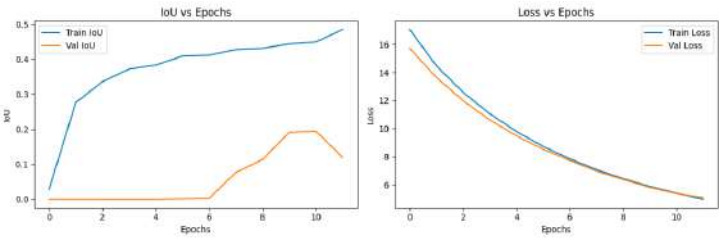

Figure 7. IOU and Loss response with reference to each epoch.

8 Epochs

An epoch refers to a complete pass of the entire training dataset through the learning algorithm. In other words, when all the data samples have been exposed to the neural network for learning patterns, one epoch is said to be completed, details are in Table 5.

8.1 Importance of Epochs in Training

Epochs play a crucial role in the training process of a machine learning model. They are directly related to

Table 5. Training and validation up-gradation

Epoch	Training IoU	Validation IoU
1.0	0.0010855215368792415	0.0
2.0	0.18763650953769684	0.037037212401628494
3.0	0.3057000935077667	0.11509975790977478
4.0	0.3422222137451172	0.03199028596282005
5.0	0.3605334460735321	0.03699883073568344
6.0	0.37744906544685364	0.0800401046872139
7.0	0.390781432390213	0.09481438249349594
8.0	0.4112493395805359	0.13650161027908325
9.0	0.42837515473365784	0.20881906151771545
10.0	0.43755123019218445	0.1954760648822784
11.0	0.45780104398727417	0.17595447599887848
12.0	0.4683050811290741	0.12801794707775116
13.0	0.4782927334308624	0.18041256070137024
14.0	0.4988195598125458	0.19632196426391602
15.0	0.490631788690399	0.28482720255851746
16.0	0.4940337836742401	0.1853817105293274
17.0	0.4967716932296753	0.3518930673599243
18.0	0.5230903029441833	0.2352318912744522
19.0	0.5359539985656738	0.36179712414741516
20.0	0.5511193871498108	0.3177557587623596
21.0	0.5239862203598022	0.33043524622917175
22.0	0.5631645321846008	0.2036125510931015
23.0	0.546701967716217	0.36430594325065613
24.0	0.5237886905670166	0.2577382028102875

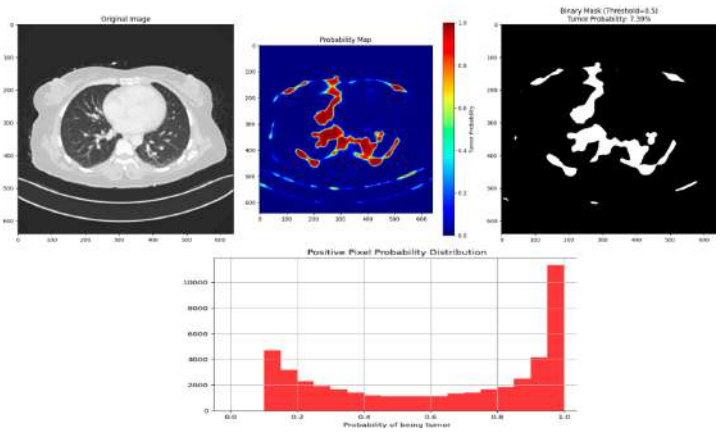

Figure 8. Test Image 1 with segmented output and probability detection

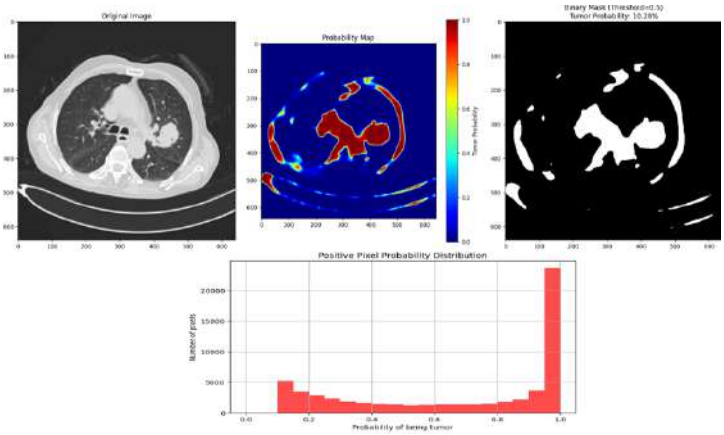

Figure 9. Test Image 2 with segmented output and probability detection

how well a model learns and generalizes to unseen data. The number of epochs is a hyperparameter that defines the number times that the learning algorithm will work through the entire training dataset. Too few epochs can result in an under-fit model, whereas too many epochs can lead to overfitting.

8.2 Comparison with state-of-the-art

To compare the performance of our proposed lung nodule segmentation model we compared it already published state-of-the-art work as shown in Table 6.

Table 6. Comparison with other state-of-the-art models on the same dataset

Model	**Dice Score**
PspNet [22]	0.48
DeepLabV3 [24]	0.51
SwinUNet [23]	0.55
UNet [1]	0.62
Our Model	0.66

9 Conclusions

In this study, we have successfully developed a robust deep learning-based system for automated lung abnormality segmentation using an enhanced U-Net architecture with EfficientNet and ResNet backbones. The mask creation was performed using Roboflow. The algorithm demonstrates good performance, as evaluated using the IoU score and tested outputs. The results are promising, indicating that the system can be implemented for real world applications, providing valuable assistance to radiologists. The algorithm has also been validated on externally collected images, showing consistent performance. This work can serve as a baseline for AI automation in radiology and has the potential to address real-world clinical challenges.

10 Future work

The work can be enhanced using updated and expanded datasets. It can also be improved by employing hybrid training techniques and incorporating other approaches, particularly in the backbone architecture. The system can be deployed in a clinical environment, and the learning rate can be further optimized using self-learning techniques.

Acknowledgment

We would like to express our deepest gratitude to the radiologists and technical team who contributed to the annotation process and provided medical expertise.

Data Availability

The datasets used and analyzed in this study were obtained from publicly available repositories such as Kaggle and were further annotated using the Roboflow platform. Preprocessed and labeled datasets are available upon reasonable request to the corresponding author, provided they meet ethical guidelines and data-sharing agreements. This manuscript used a publicly available dataset repository from Kaggle link is in the footnote. [1]

Author Contributions

Tehreem Awan: Supervision, conceptual framework, final revisions. **Ali Raza:** Model architecture, training protocol, literature review and performance evaluation. **M. Inzimam ul Haq:** Data preprocessing, clinical validation and implementation of attention mechanisms. **M. Usman Abdul Manan:** Roboflow pipeline integration, annotation, documentation and formatting. **Nasrullah:** Result analysis, and visualization outputs. **Serosh Karim:** Review and guidance. All authors contributed equally to discussions, model development decisions, and manuscript review..

Conflict of Interest

All authors have a declared conflict of interest.The study utilized publicly available anonymized data. No direct involvement of the patient occurred. Therefore, formal ethical approval and informed consent were not required.

Funding

This manuscript has not received any funding from an internal or external source.

References

[1] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2015, pp. 234–241.

[2] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3D U-Net: Learning dense volumetric segmentation from sparse annotation," in *Proc. Int. Conf. Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2016, pp. 424–432.

[3] G. Litjens *et al.*, "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, 2017.

[4] A. A. A. Setio *et al.*, "Pulmonary nodule detection in CT images: False positive reduction using multi-view convolutional networks," *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1160–1169, 2016.

[5] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 4700–4708.

[6] M. Tan and Q. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in *Proc. Int. Conf. Machine Learning (ICML)*, 2019, pp. 6105–6114.

[1] https://www.kaggle.com/datasets/mohamedhanyyy/chest-ctscan-images

[7] F. Milletari, N. Navab, and S. A. Ahmadi, "V-Net: Fully convolutional neural networks for volumetric medical image segmentation," in *Proc. Int. Conf. 3D Vision (3DV)*, 2016, pp. 565–571.

[8] S. G. Armato *et al.*, "The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A completed reference database of lung nodules on CT scans," *Medical Physics*, vol. 38, no. 2, pp. 915–931, 2011.

[9] J. Chen *et al.*, "TransUNet: Transformers make strong encoders for medical image segmentation," *arXiv preprint arXiv:2102.04306*, 2021.

[10] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, "Pyramid scene parsing network," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 2881–2890.

[11] F. Isensee, P. F. Jaeger, S. A. A. Kohl, J. Petersen, and K. H. Maier-Hein, "nnU-Net: A self-adapting framework for U-Net-based medical image segmentation," *Nature Methods*, vol. 18, no. 2, pp. 203–211, 2021.

[12] A. Mansoor *et al.*, "Segmentation and image analysis of abnormal lungs at CT: Current approaches, challenges, and future trends," *RadioGraphics*, vol. 34, no. 2, pp. 341–361, 2014.

[13] P. Rajpurkar *et al.*, "CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning," *arXiv preprint arXiv:1711.05225*, 2017.

[14] Roboflow, "Roboflow: Organize, annotate, and preprocess images for computer vision," 2023. [Online]. Available: https://roboflow.com

[15] Y. Xu *et al.*, "Lung nodule classification using deep feature fusion in chest CT images," *Computers in Biology and Medicine*, vol. 122, p. 103819, 2020.

[16] T. Awan and K. B. Khan, "Analysis of underfitting and overfitting in U-Net semantic segmentation for lung nodule identification from X-ray radiographs," in *Proc. IEEE Int. Conf. Emerging Trends in Engineering, Sciences and Technology (ICES&T)*, 2023, pp. 1–5.

[17] T. Awan, K. B. Khan, and A. Mannan, "A compact CNN model for automated detection of COVID-19 using thorax X-ray images," *Journal of Intelligent & Fuzzy Systems*, vol. 44, no. 5, pp. 7887–7907, 2023.

[18] T. Awan and K. B. Khan, "Investigating the impact of novel XRayGAN in feature extraction for thoracic disease detection in chest radiographs: Lung cancer," *Signal, Image and Video Processing*, vol. 18, no. 5, pp. 3957–3972, 2024.

[19] T. Awan *et al.*, "AI-powered lung cancer detection from CT imaging," *VFAST Trans. Software Engineering*, vol. 12, no. 2, pp. 241–249, 2024.

[20] T. Awan *et al.*, "Advanced EEG-based brain monitoring for early detection of harmful neural activities using EfficientNet architecture," *Biomedical Signal Processing and Control*, vol. 112, p. 108360, 2026.

[21] T. Awan *et al.*, "Generating synthetic data in biomedical imaging by designing GANs," *VFAST Trans. Software Engineering*, vol. 12, no. 3, pp. 44–54, 2024.

[22] L. Yan, D. Liu, Q. Xiang, Y. Luo, T. Wang, D. Wu, H. Chen, Y. Zhang, and Q. Li, "PSPNet-based automatic segmentation network model for prostate magnetic resonance imaging," *Computer Methods and Programs in Biomedicine*, vol. 207, p. 106211, 2021.

[23] P. Sabitha, R. A. Canessane, M. S. P. Minu, V. Gowri, and M. S. A. Vigil, "An improved deep network model to isolate lung nodules from histopathological images using an orchestrated and shifted window vision transformer," *Traitement du Signal*, vol. 41, pp. 2081–2091, 2024.

[24] H. Polat, "A modified DeepLabV3+ based semantic segmentation of chest computed tomography images for COVID-19 lung infections," *International Journal of Imaging Systems and Technology*, vol. 32, no. 5, pp. 1481–1495, 2022.