

## Research Article

# SegChaNet: A Novel Model for Lung Cancer Segmentation in CT Scans

Mehmet Akif Cifci 

*Dept. of Computer Engineering, Bandirma Onyedi Eylul University, Balikesir, Turkey*

Correspondence should be addressed to Mehmet Akif Cifci; [mcifci@bandirma.edu.tr](mailto:mcifci@bandirma.edu.tr)

Received 7 April 2022; Accepted 2 May 2022; Published 14 May 2022

Academic Editor: Yaodong Gu

Copyright © 2022 Mehmet Akif Cifci. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Accurate lung tumor identification is crucial for radiation treatment planning. Due to the low contrast of the lung tumor in computed tomography (CT) images, segmentation of the tumor in CT images is challenging. This paper effectively integrates the U-Net with the channel attention module (CAM) to segment the malignant lung area from the surrounding chest region. The SegChaNet method encodes CT slices of the input lung into feature maps utilizing the trail of encoders. Finally, we explicitly developed a multiscale, dense-feature extraction module to extract multiscale features from the collection of encoded feature maps. We have identified the segmentation map of the lungs by employing the decoders and compared SegChaNet with the state-of-the-art. The model has learned the dense-feature extraction in lung abnormalities, while iterative downsampling followed by iterative upsampling causes the network to remain invariant to the size of the dense abnormality. Experimental results show that the proposed method is accurate and efficient and directly provides explicit lung regions in complex circumstances without postprocessing.

## 1. Introduction

Lung cancer is the leading cause of cancer-related death globally. According to the World Health Organization (WHO) forecasts for 2019, cancer is the major or second leading cause of death in 112 of 183 nations, and it ranks third or fourth in an additional 23 countries before the age of 70 [1]. The 5-year survival rate is approximately 34% for patients with early-stage, resectable cancer, while the 5-year survival rate is less than 10% for unresectable cancer. Therefore, early identification and diagnosis of lung cancer are essential in enhancing patient treatment results. When images reveal the existence of a tumor, it is critical to histopathologically evaluate diagnostic samples collected by fiberoptic bronchoscopy following National Comprehensive Cancer Network guidelines [2].

The criterion for diagnosing lung cancer is a pathologist's diagnosis of biopsy tissue. On the other hand, diagnostic precision is less than 80% [3]. The four most prevalent malignant lung tumor subtypes are squamous carcinoma, adenocarcinoma, small-cell carcinoma, and undifferentiated carcinoma [4]. To make the optimal treatment decisions,

scientists must employ the proper techniques to diagnose distinct cancer subtypes during the biopsy. After the National Lung Screening Trial (NLST), the leading cancer research and training organization, developed a method to use a low-dose helical computed tomography (CT) to detect lung cancer nationwide in 2015, scientists were finally able to use low-dose CT to diagnose lung cancer nationwide. Furthermore, the Dutch-Belgian lung cancer screening trial, the world's second-largest randomized control trial, confirms the advantages of lung cancer screening [5]. However, implementing the US and prospective lung cancer screening in Europe would likely lead to many whole-slide histopathology images, biopsies, and excised tumors. Researchers have proposed many medical image analysis methods in CT scans to segregate the lung parenchyma region automatically [6]. For example, authors in [7] describe signal thresholding strategies based on contrast information for most methods. Because of their lower densities than the rest of the body, the lung region looks darker when framed by a denser region. The underlying framework of these approaches is straightforward and effective for normal lung segmentation. Still, they fall short when expanding the word "lung" to

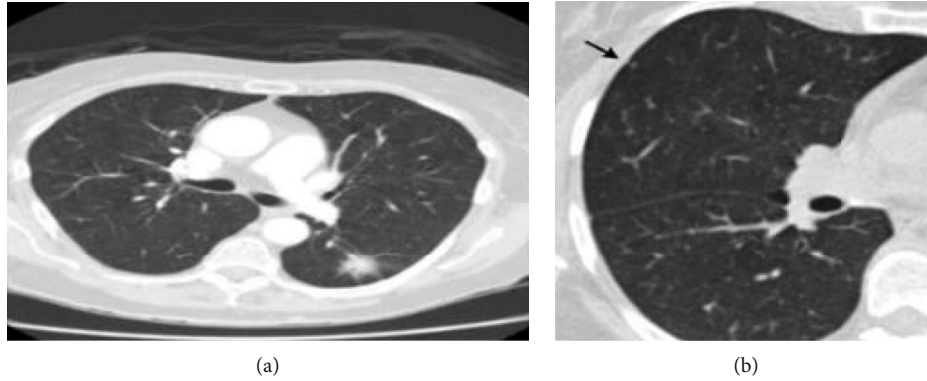


FIGURE 1: (a) Lung cancer nodule view. (b) An arrow indicates this 1.25 mm thick CT slice to contain approximately 2 mm long lung nodules.

include aberrant tissues, blood arteries, and normal lung tissues [8]. First, it is crucial to obtain a segmented region using iterative thresholding and then refine the obtained region using an opening-closing morphological operator. Significant success has been found in noninvasive therapy and clinical examination in the health industry, especially when utilizing medical image analysis [9]. Researchers use CT imaging for specific diagnostics, and hence, images such as X-rays, MRIs, and CTs are effective therapeutic methods. Because cancer is the deadliest disease, the value of CT images grows even further, given that 1.61 million individuals die from lung cancer each year, according to the World Health Organization [10]. Despite the progress made in medical imaging technology, including CT scans, lung cancer accounted for 76 million deaths in 2018, and about 13% of those newly diagnosed with lung cancer die from cancer [11]. Most cancer deaths occur in low- and middle-income nations, as over 70% of cancers are found in these regions [12]. The automated identification of lesions and automatic categorization of lung diseases have seen significant advancements in CT scans during the last two decades [13]. Machine learning (ML) systems have played a key role in processing images.

Recently, there has been a significant increase in the attention paid to deep learning (DL) in several domains, including image recognition and biomedical image analysis [14]. After retrieving CT scan features, image processing techniques are applied to the image data to assess whether the patient's cancer is benign or malignant. Figure 1(a) displays a lung cancer nodule view, and Figure 1(b) shows an arrow that indicates a 1.25 mm thick CT slice.

As seen in Figure 1, smaller nodules are in almost any CT scan indicative of problems. For example, the arrow on Figure 1 (b) indicates a tiny nodule to be malignant. Table 1 displays the causes and detection phases of lung cancer.

The objectives and justification of the work are as the following:

This study was aimed at developing a DL-based automated lung cancer tumor segmentation network utilizing CT scans.

TABLE 1: Lung cancer: causes and prevention.

(i) Detection of lung cancer is generally difficult because specialists cannot find the infected area until it reaches the next stage. As a result, the chance of survival for lung cancer, in 54% of detected cancers, not in the advanced stages, with early intervention, is only 4% [15].
(ii) The probability of increasing lung cancer diagnoses due to the number of cigarettes consumed and sometimes after drinking is proportional. As a result of harmful habits, a minor case of lung cancer may occur even in individuals without disease risk.
(iii) X-ray, CT, or MRI scans are performed to examine lung cancer and differentiate abnormal lung development. The best technique is CT, which experts can overlook when not in ML.

- (i) In current procedures, the lung images were taken and subjected to segmentation, to benefit from the support vector machine classifier [16]
- (ii) The present framework has a restriction as it could not predict the type, form, or size of the tumor, and it dealt with several pixels, which is not beneficial for the early detection of cancer. Now, when artificial neural network (ANN) develops a testing solution, it does not reveal some information as to why and how. This diminishes trust in the network
- (iii) Neural networks (NN) are a black box and have been confined in their capacity to perceive the potential causal links expressively. On the other hand, NN features deep networks with numerous hidden layers and is helpful for modeling complex systems. However, the training process is again more complex and constantly sensitive, creating a few complications
- (iv) It is expected that by applying this model, many current data mining and image processing procedures operate jointly in numerous ways. However, the fundamental drawback of the linear discriminant analysis approach is that it only differentiates the images having abnormalities

- (v) Because of the issues with classifiers, segmentation performs better in accuracy

The contribution of this work is as below:

- (i) This research presents a lung segmentation utilizing 3 traditional and a novel model. This study was aimed at developing a DL-based automated lung cancer tumor segmentation network utilizing CT scans
- (ii) Developed a methodology, SegChaNet, that achieves the state-of-the-art performance on numerous segmentation tasks dealing with lung cancer segmentation; of foremost importance is the increase acquired in the precision of tumor segmentation
- (iii) Using the channel attention module (CAM) to achieve the goal of managing the feature selection of the bottom-level feature map
- (iv) Consequently, this study reveals the capability of early detection of lung cancer, which is known to improve treatment outcomes

The rest of the paper is structured in the following way. The next section contains several examples from the literature and will contrast them with the present study. Next, we explain the methods in Section 3, including methods for the segmentation of lung cancer. Then, we have presented a detailed analysis of the proposed model in Section 4. In addition, the results and a discussion are in this section. Finally, we summarize the full text in Section 5 and present the future work.

## 2. Literature Scan

Numerous studies on lung segmentation have been conducted using conventional image processing techniques such as thresholding, edge detection, and clustering [17, 18]. When faced with chaotic input images, many image processing algorithms resort to primitive techniques and perform poorly. This project was aimed at developing an automated lung cancer tumor segmentation network based on deep learning using CT images. Researchers are examining lung segmentation using convolutional neural networks (CNNs). The researchers now want to enhance lung segmentation performance through complex image segmentation network topologies and other strategies such as attention modules. It has been demonstrated that the attention mechanism significantly improves performance on many DL-based activities.

As a result, our research focuses on segmentation techniques based on attention. For example, attention U-Net [19] outperforms U-Net by sandwiching an attention module with a fundamental structure between the current U-Net structure's contracting and growing channels. Similarly, the authors of [20–22] enhanced the structure of the corrective adversarial network, which was the first attempt to use adversarial learning for lung segmentation on chest X-rays by incorporating attention U-Net and focal Tversky loss for the generating network and binary cross-entropy, respectively.

Additionally, the researchers in [23, 24] used CC-Net as the backbone network, an image segmentation network based on the criss-cross attention module. Additional learning data were generated using image-to-image translation. XLSor is state-of-the-art software. This project was aimed at developing an automated lung cancer tumor segmentation network based on deep learning using CT images. Numerous lung image segmentation networks use a reference lung segmentation model based on chest X-ray images [25].

In their study, the authors in [26] proposed an algorithm for classifying lung nodules by considering the data learned from tissue, shape, and the D technique. This algorithm utilized a “gray level coformation matrix”-based surface identifier, a Fourier shape identifier, and CNNs to train the properties of nodes to depict the heterogeneity of nodules. However, the authors in [27] focused on computer-aided diagnosis (CAD), which they designed manually, so it is not ideal or sufficient for the solution.

U-Net employs the skip connection, which is symmetrical, to directly supervise and lose back propagate on semantic features instead of using these other procedures. The aspect of scale integration permits the implementation of various features from multiple sizes to support multiscale prediction. Because of this, U-Net is primarily used in the medical imaging industry [28]. The researchers in [29] suggested using a recurrent; the residual convolutional neural network (RRCNN) was introduced in conjunction with a U-Net to train more complex networks and extract deep semantic information. The results obtained using R2U-Net in the vascular, lung, and skin datasets are above expectations. As a result, the depth of the U-Net network varies according to the applications it serves. As a result, Liang et al. designed U-Net, which included smaller-scale U-Net levels in larger-scale U-Net levels, with deep supervision, which allows the network to self-select the appropriate amount of depth via training [30].

On the other hand, the researchers in [31] developed the cellular neural network algorithm to detect lung cancer symptoms. In this study, the authors employed X-ray images and thus the CNN algorithm to diagnose lung cancer. The authors in [32] focused on developing CNNs for lung cancer screening in CT scans. The researchers in [33] used multi-layered neurons with independent principal components to diagnose lung cancer.

In their study, the researchers in [34] performed a diagnostic classification of lung nodules using 3D-CNNs. They did not foresee overfitting in their model, so they had to include a retraining phase method to address issues linked to image label imbalance. Chon et al. [35] proposed a method that uses an ANN community chart to differentiate benign from cancerous lung nodules. The experimental results show that the scheme has 78.7% classification accuracy [36]. On the other hand, the authors in [37] utilized an ANN model with a classification accuracy of 92 percent in their study. Zhang et al. [38] presented a methodology with an accuracy rate of 75.01% using the Automatic coder (AC), a DL technique. Table 2 includes the comparison of the literature review in the related field.

TABLE 2: Comparison to the literature works.

References	Datasets	Method	Result (%)
Chaturvedi et al. 2019 [39]	LUNA 16	3D DL DÖ, V-Net architecture	Sensitivity: 96.5 FP:19.7
Chapaliuk et al., 2019 [40]	ACDC LUNGH	VGG16, ResNet50, and CNN	Sensitivity: 97.9 Accuracy: 93
Petrellis et al., 2018 [41]	UCI	Gaussian blur, Otsu thresholding	Sensitivity: 87 Accuracy: 97
Yuan et al., 2019 [42]	134 BT Shandong hospital	Watershed transform	Sensitivity: 88.8 Accuracy: 90
Cao et al., 2016 [43]	LUNA16	3D and 2D CNN	Precision: 87 Sensitivity: 99.1
Xie et al., 2019 [44]	LUNA16	2D CNN and RCNN	AUC: 95.4
Sun et al., 2017 [45]	LIDC-IDRI	CNN, deep belief network, and Boltzmann machine	Sensitivity: 82.2 AUC: 81.8
Huang et al., 2018 [46]	LIDC-IDRI	CNN, extreme learning machine, and deep transfer	Sensitivity: 91.6 Accuracy: 86.5
Pehrson et al. (2021) [47]	LIDC-IDRI	This study was aimed at developing a DL-based automated lung cancer tumor segmentation network utilizing CT scans	Sensitivity: 91.7 Accuracy: 93.8
Sharma et al., 2011 [48]	LIDC-IDRI	Diagnostic indicators	Accuracy: 80.1
Akram et al., 2012 [49]	LIDC-IDRI	Neurofuzzy	Accuracy: 95.5
Paulin et al. 2011 [50]	LIDC-IDRI	For MLP SVM training, the back propagation technique is employed	Accuracy: 83.6
JIA et al. 2007 [51]	NCA	For MLP SVM training, the back propagation technique is employed	Accuracy: 92.4

### 3. Methodology

**3.1. Data Acquisition.** Images of 46,698 CT scans, both with and without tumors, from the Cancer Imaging Archive (<https://wiki.cancerimagingarchive.net/display/Public/RIDER+Lung+CT>) were made available. 80% of the data were used for training and 10% for validation to obtain a reliable and accurate network, while the remaining 10% of the data were for testing. We have used different parameters for training segmentation models. We have applied the data augmentation steps to the data. On each patient’s scan, we employed intensity normalization based on the mean and standard deviation of the intensities. We implemented our models on a machine with an NVIDIA GeForce RTX 2080 8 GB and Intel Core i9-9980hk processor configuration. We utilized Python (3.8.12), the major programming language, for constructing the framework’s subsequent phases. The proposed architecture is implemented using the libraries Anaconda3 (64-bit), Jupyter Notebook, TensorFlow, and Keras.

As seen in Table 3, the dataset is split into train, test, and validation. The CT scan slices are as with tumors and without tumors.

**3.2. Preprocessing CT Scans.** Preprocessing CT scans dealing with high variation in data is a typical job performed during preprocessing. The first step was to truncate all the Hounsfield unit (HU) values between  $-1000$  and  $400$  [52]. The HU is a metric that measures the density of materials where the air has a value of  $-1000$  HU and bone has a value of  $1000$  HU; values outside the range of  $-1000-400$  is not considered

TABLE 3: Some distinct sorts of CT scan slices in the dataset.

Data	Number of patients	Tumor	Without tumor	Notes
Train	370	14,848	20,840	35,688
Test	90	4520	4740	9260
Validation	50	850	900	1750
Total	510	20,218	26,480	46,698

necessary for lobar segmentation [18]. We then normalized the CT scans to a zero mean with unit variance. To fit the images into a 3D model, we resized all the images and made them the same size, satisfying the memory constraints of the GPU. Two main approaches to this task achieve similar performance: downsampling CT images or applying a patch-based method by partitioning the images into overlapping patches. Since we needed to incorporate tracheal and bronchial segmentation as an auxiliary task into the segmentation, we opted for the downsampling technique, ensuring that the trachea and bronchi were present in each image. The size should be the same on all three axes to downsample the image [53]. If one axis is of better resolution than another, the performance along lower resolution axes is worse, as they contain less information than the other axes. All the CT scans, with their masks, were resized to  $128 \times 128 \times 128$  using linear and neighbor interpolation, respectively, as it was the most significant cube size shape we could fit into memory. What makes SegChaNet different from U-Net is that U-Net was initially designed for medical image



interpretation and segmentation, while we developed SegChaNet for lung image segmentation.

U-Net has a wide variety of industrial applications and has played an important part in the development of the image automation society. The architecture of SegChaNet is defined by its appealing and expansive components. The contracting path comprises several convolutional patches with  $3 \times 3$  filters and unity strides in both directions, which are preceded by ReLU layers. This path takes the input and extracts the key features, yielding a feature vector of a certain length. The second path uses information from the contractive path to creating an output segmentation map by copying and cropping the feature vector using upconvolutions. The exploit that connects the first and second channels is vital for this system. This connection enables the network to get very precise information from the contractive route, developing a segmentation mask close to the desired output. SegChaNet benefits from extremely asymmetric datasets because of its composition of big chunks of lung images. Due to the lungs' different geometries and image density distributions, it is difficult to distinguish them in clinical ultrasound images automatically.

SegChaNet resembles U-Net but as explained above, combining two different paths, and utilizing Binary cross Entropy, integrated with the Channel Attention Module, differs it from the rest. The encoder layers are similar to the convolutional layers of the U-Net.

For SegChaNet, nine distinct hyperparameters are used in training to determine the optimal configuration. These training hyperparameters are listed in Table 4. For training purposes, all trials were done in 500 epochs for training purposes. Training time variation was small across networks, with an average of 18 minutes.

**3.3. U-Net Model.** U-Net is a technique for automatically segmenting images. U-Net is a pixel-based image segmentation algorithm for various architectural and convolutional neural network layers. It outperforms traditional models in terms of success. It also performed well on datasets with a minimal number of photos. The term "U-Net" comes from its U-shaped design. U-Net is made up of encoding and decoding components. The content of the images is recorded in the code section. It is made up of top pooling layers. On the other hand, the decoding portion is the symmetric expansion route utilized to obtain exact localization via the use of delegated convolutions. The U-Net model is shown in Figure 2. Each of the  $32 \times 128 \times 128$  voxels included inside the input layer corresponds to a single channel. We did 1D convolution on the input by applying a  $1 \times 3 \times 3$ ,  $2 \times 3 \times 3$ , and  $1 \times 1 \times 1$  filter. Using max pooling, we perform one-dimensional max pooling. By using 1D upsampling, the voxel size was raised from 2 to 4. To model U-Nets, we use the rectified linear unit (ReLU) as the input and the sigmoid function as the output layer's activation function. Data augmentation is a strategy that may assist avoid overfitting and increase performance by increasing the quantity of data points. The left-to-right inversion and white-to-black inversion methods were used in this work to provide data augmentation. We used Adam as an optimizer because of its

TABLE 4: Hyperparameters used for training SegChaNet.

Exp.	SegChaNet	
	ILR	Minibatch
1	1e-4	2
2	1e-4	8
3	1e-4	12
4	1e-3	2
5	1e-3	4
6	1e-3	8
7	3e-3	4
8	3e-3	12
9	3e-3	8

adaptive moment estimation. To arrive at the ideal learning rate, various values were tried, including those between 0.3 and 0.1, with every successive change made along with the range. To optimize a solution, we must minimize the loss function, the difference between the projected result, and the correct answer. We calculated these equations to get the following:

$$\text{Loss function} = 2 - \{\text{IOU}(A, B) + \text{DSC}(A, B)\},$$

$$\text{IOU}(A, B) = \frac{|A \cap B|}{|A \cup B|}, \quad (1)$$

$$\text{DSC}(A, B) = \frac{2|A \cap B|}{|A| + |B|}.$$

In this case, the estimated value,  $B$ , is an estimation and the actual value,  $A$ , is accurate. The intersection over union (IOU) and dice similarity coefficient (DSC) are represented as the Jaccard Index (JI). The JI and the DSC are both methods of measuring class similarity. To determine the accuracy of contour demarcation, we use DSC. The ensemble learning for five inference results comprises 5-fold cross-validation; afterward, we obtained DSC for U-Net. The following example illustrates this principle in ensemble learning, for which we have found a pixel value of 1 for three inferences but a value of zero for two inferences. In this case, we selected a majority value of 1.

In Figure 2, the path has two distinctive sections: a contracting section and an expansive one. A typical convolutional architecture often features a contracting path. It is a  $2 \times 2$  max pooling with stride 2 and a ReLU followed by two unpadding convolutions such as  $3 \times 3$  convolutions. We have increased the amount of feature channels by a factor of two each time while performing downsampling. The enlarging path has upsampling, followed by  $2 \times 2$  convolutions.

**3.4. V-Net Model.** We applied the V-Net architecture using the TensorFlow and Keras frameworks in Anaconda, as stated in Data Acquisition. We have employed the V-Net, an autoencoder, for 3D image segmentation. Its residual connections aid networks to converge and perform well on small datasets. Figure 3 shows the model's architecture,

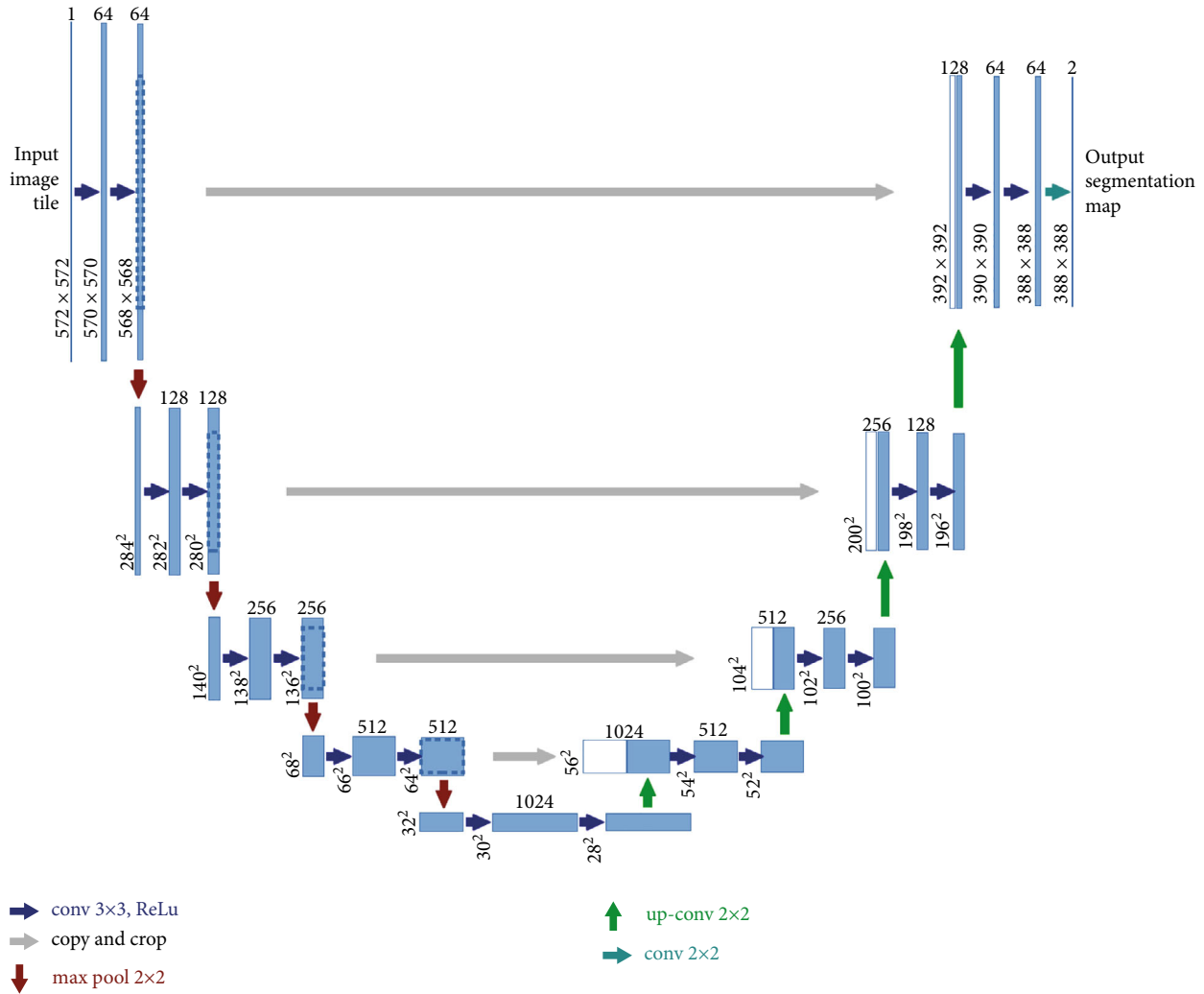


FIGURE 2: U-Net architecture.

where the input image is  $128 \times 128 \times 128 \times 1$  (height, width, depth, and channels) and passes through six encoder layers. As seen from Figure 3, V-Net, we used convolutional layers with a stride of  $2 \times 2 \times 2$  instead of the max pooling operation. The last layer before each output was a  $1 \times 1 \times 1$  convolution with a softmax activation function that gives us the probability that each voxel belongs to one of the classes. Various regularization techniques enhanced the model's generalizability, including dropout and batch normalization. Training the model on distinct but relevant tasks has been shown to increase a model's performance. The model's performance is measured using the DSC, a function of sensitivity and specificity typically used in segmentation studies, permitting a comparison of our model's performance to others. We compared our model to other models by researchers who released the original dataset. A  $p$  value obtained using a  $t$ -test, more diminutive than 0.05 between the models, was considered statistically significant for independent comparisons. Figure 3 shows V-Net architecture.

We illustrate our CNN using the schematic in Figure 3. To both find features in the data, we use convolutions, as

well as to reduce the resolution after each stage by employing stride suitable to the level. In the left portion of the network, we used a compression path to reduce the file size, and in the right portion, we decompressed the file until we reached its original size. We employed all of the convolutions with the proper amount of padding.

**3.5. SegChaNet Model.** Segmentation tasks generally concentrate on learning multiscale features and the integration of local and global contexts. However, computationally demanding training segmentation networks including these qualities and implementing these directly on volumetric data is difficult. To help promote multiscale learning, we integrate CAM convolutions into the network and place these convolutions at the very end of the feature extraction process. They work together with the output of the decoder layers to promote the learning of multiscale features. We use this connection to assist the training of network gradients as well. We maintained residual links between the decoding component and the alteration when we adjusted. The proposed network consists of four blocks of encoders and

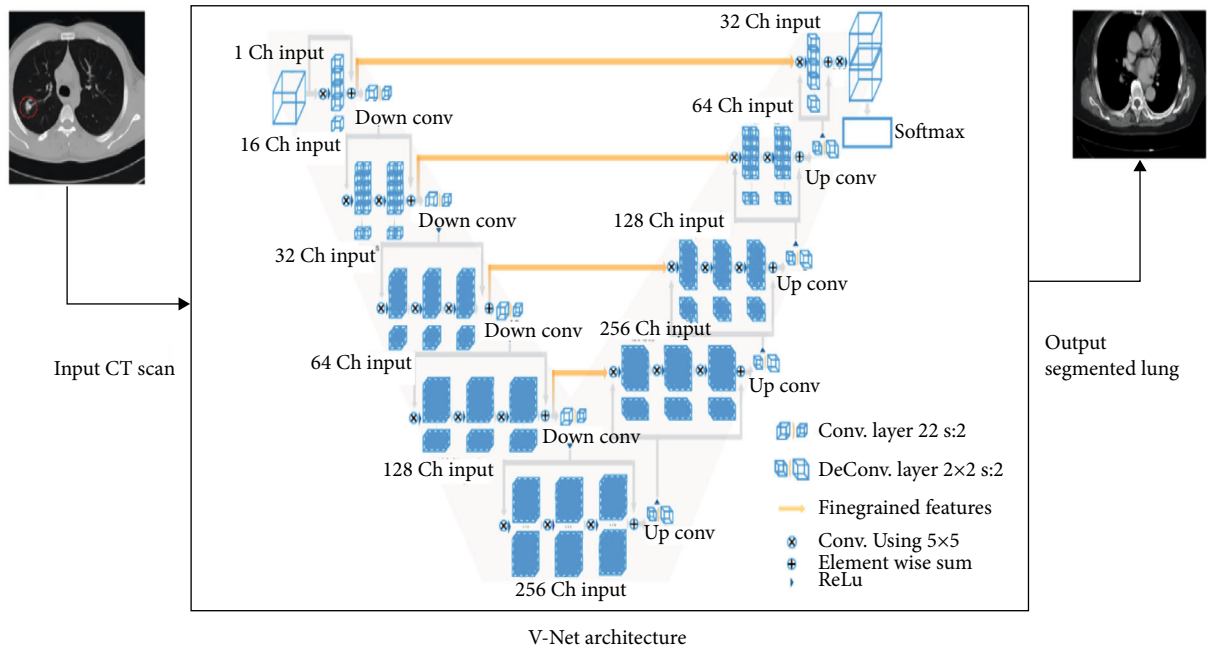


FIGURE 3: V-Net architecture.

decoders each using two 3D convolution layers with a  $3 \times 3 \times 3$  kernel size and a batch normalization and leaky ReLU activation function.

However similar it may seem, SegChaNet varies from U-Net. For example, we use the maximum linked component detection to extract the whole region, and SegChaNet additionally takes use of CAM. Besides, preprocessing steps have been applied before the SegChaNet model is applied. So, it necessarily comprises a preprocessing phase. Figure 4 displays the main novel architecture.

As seen in Figure 4, this decoder uses a  $2 \times$  upsampling method. We proposed the convolution of the augmentation module's CAM convolution layers such that the first layer utilizes a dilation rate of one, and each subsequent layer doubles the dilation rate. It is upsampled to scale each decoder's output, followed by all concatenated outputs.

**3.5.1. Channel Attention Module.** This section delves further into SegChaNet's model architecture (illustrated in Figure 4). Because of its higher performance, the pixel-level classification strategy is employed in medical image segmentation tasks based on the U-Net model. We included the cross-level feature fusion module (CFFM) in our end-to-end network architecture, SegChaNet, to create a merged network with two primary components and CAM. The CFFM enables the fusion of features from the top-level feature map into the feature map of the bottom layer, allowing the CAM to regulate feature selection in the bottom-level feature map. The second component is the U-Net, a two-part encoder-decoder architecture in which the first half is used to extract features and the second half is upsampled. Special caution is required when using a skip connection topology between the encoder and the decoder. Additionally, a skip link was used to connect the downsampled and upsampled feature maps. The CAM and

architecture used are shown in Figure 4. Because the high level includes much semantic information, it may help guide the low-level choices. Several studies have shown that attentiveness is vital in human perception. Recently, attention processes have been included in complicated sequences and transformation models for various activities. Channel attention modules, in particular, have been utilized to boost deep learning performance in a variety of computer vision applications. DL methods for image classification and semantic segmentation rely on continuous fundamental convolutional processes. Consequently, deep learning networks can only manage the local components of an image, resulting in the loss of global information.

A sophisticated method must be used to estimate the parameters. The parameters of the investigation were determined in three steps. Attenuation coefficients can be derived from the attenuation profile using a single-path model and a least squares estimator. The number, position, and amplitude of key pathways may be calculated from the impulse response using a basic peak detection approach. In simpler cases, this phase usually gives enough precision. To expand the number of pathways, parameters must be developed using either impulse response or amplitude and phase responses. To estimate parameters, complex procedures must be applied. The parameters of the investigation were determined in three steps. An initial estimate of the attenuation coefficients can be generated from the attenuation profile using least squares estimators. The number, position, and amplitude of key pathways may be calculated from the impulse response using a basic peak detection approach. In most cases, this is all that is required. To increase the number of pathways, the parameters must be further changed through an evolutionary process employing quality criteria such as impulse response or amplitude and phase response.

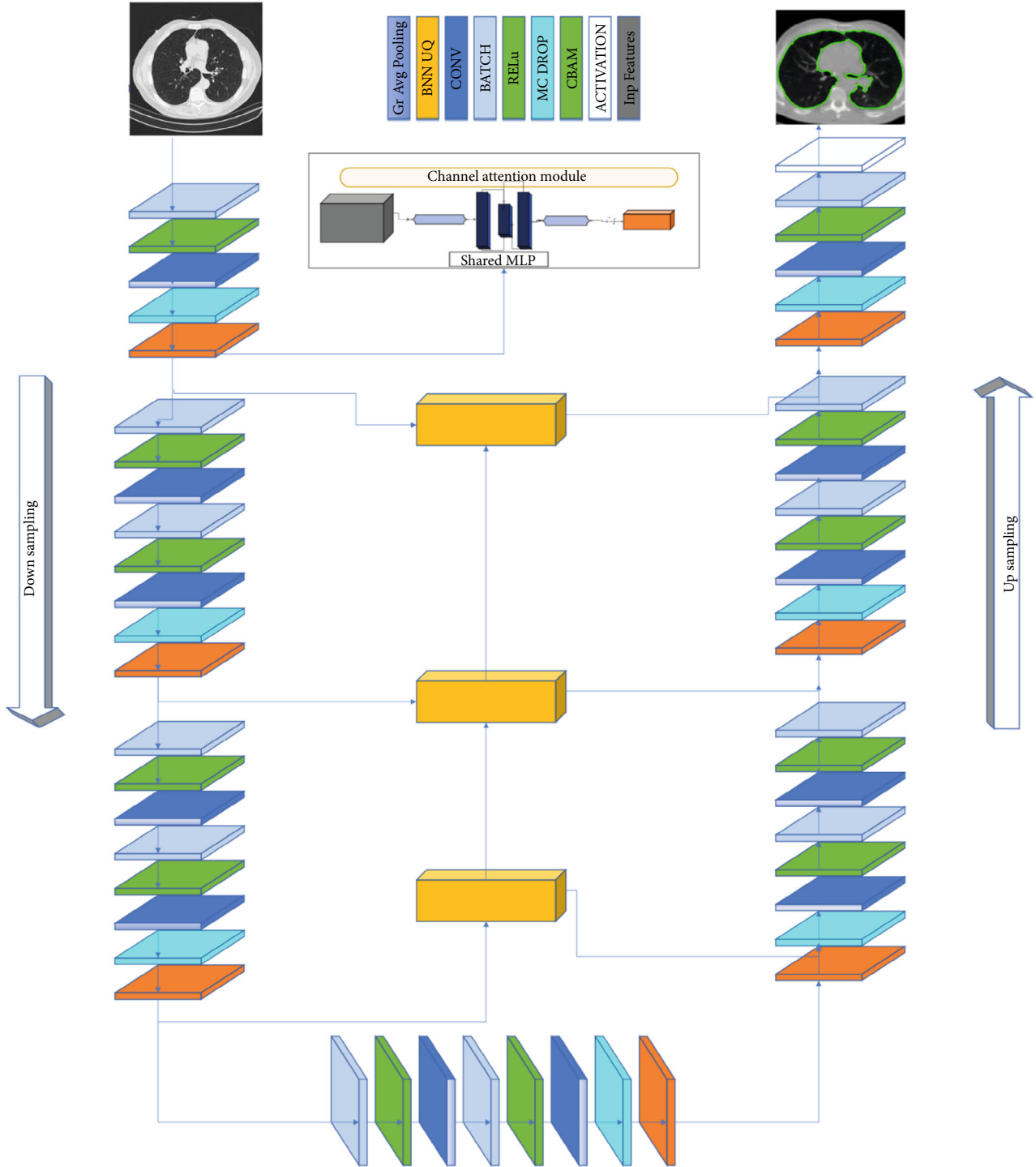


FIGURE 4: An encoder-stage-to-decoder stage residual connection network architecture. The encoder uses residual connections and 3D max pooling operations.

We have provided a more specific resolution information option. Furthermore, the SegChaNet model may learn the weight of each channel via CAM, resulting in attention in the channel domain. The CAM procedure is as follows:

$$X_{\text{cam}} = \text{CAM}(X_L, X_H), \quad (2)$$

where  $X_{\text{cam}}$  represents the output of the CAM module and  $X_L$  and  $X_H$  represent the low-level feature map and high-level feature map, respectively.

$$x_{\text{cam}}^{i,j,k} = x_{g_L}^{i,j,k} + \text{up}\left(x_H^{i,j,k}\right), \quad (3)$$



TABLE 5: U-Net method with/out CAM performance.

U-Net method	Dice (%)	Sensitivity	Specificity	Precision	F-measure	Accuracy (%)
Without CAM	88.61	97.45	93.12	93.01	93.21	93.21
With CAM	95.94	97.62	90.43	89.87	95.56	95.14

TABLE 6: The V-Net method with/out CAM performance.

V-Net method	Dice (%)	Sensitivity	Specificity	Precision	F-measure	Accuracy (%)
Without CAM	87.35	88.29	84.74	86.51	91.14	91.63
With CAM	95.75	96.96	89.77	89.21	94.91	94.48

TABLE 7: SegChaNet method with/out CAM performance.

SegChaNet	Dice (%)	Sensitivity	Specificity	Precision	F-measure	Accuracy (%)
Without CAM	96.81	93.79	90.19	92.15	96.89	96.47
With CAM	98.48	92.82	94.08	96.66	98.49	98.90

where  $x_{cam}^{i,j,k}$  represents the pixel value of the  $i$ th row,  $j$ th column, and  $k$ th channel in the output feature map of the CAM.

In addition,  $x_H^{i,j,k}$  is the pixel value of the  $i$ th row,  $j$ th column, and  $k$ th channel of the high-level feature map.  $\text{up}(\cdot)$  means the bilinear interpolation.

$$X_{gL} = \text{Conv1} \left( \left[ X_L' \times g \left( X_H^k \right) \right]_{k=0}^{k=c-1} \right). \quad (4)$$

Global pooling transforms the feature map  $X_{gL}$  and low-level feature map  $X_L$ . The  $3 \times 3$  convolution operation adjusts the feature map size and merges the channel features. Still, because the low-level and high-level features' feature map size and channel number are different, the low-level feature needs a  $3 \times 3$  convolution operation.

$$X_L' = \text{Conv3}(X_L). \quad (5)$$

Global pooling, convolution with a  $3 \times 3$  convolution kernel, and convolution with a  $1 \times 1$  convolution kernel which are known as  $g(\cdot)$ ,  $\text{conv3}(\cdot)$ , and  $\text{conv1}(\cdot)$ , respectively. The letter "c" denotes the total number of channels [54].

When global pooling is used, CAM gives an initial global context that serves as a guide for recognizing low-level features. To decrease the number of feature mappings in the encoder, we first apply three convolution filters to the low-level features. A global context is created by multiplying the low-level features by the high-level features using a  $1 \times 1$  convolution process. Finally, we have combined low-level and high-level features. Thus, the CAM module uses scale feature maps more effectively while delivering information to lower-level feature maps.

When it comes to the model's performance, as seen, Table 5 shows the performance of the U-Net method with and without CAM.

As seen in Table 5, accuracy of the U-Net method without CAM is 88.61.21 while with CAM is 95.94. Table 6 displays the methods and performance of the V-Net method with and without CAM.

From Table 6, the accuracy of the V-Net method without CAM is 87.35 while with CAM is 95.75. Table 7 shows the methods and performance of the novel method SegChaNet with and without CAM.

From Table 7, it is evident that the accuracy of the SegChaNet method without CAM is 96.81% while with CAM is 98.48. Figure 5 shows the performance of SegChaNet, illustrating the comparison of the novel model with the state-of-art ROC performance. Table 8 compares models' accuracy evaluated in DSC, JI.

Table 8 displays the results with CAM. As seen, SegChaNet is far better when compared with the other methods.

From Figure 5, it is evident that the AUC of V-Net, U-Net, and SegChaNet without CAM is lower when compared with used models plus CAM.

**3.6. Training Procedures.** We are creating a new model from scratch and will use the DSC, JI, and normalized surface distance to test it. We computed normalized surface distance as a combination of the model-predicted segmentations that overlap with the ground truth. In the segmentation process, we utilized these variables to determine segmentation accuracy. Segmentation networks' segmentation binary cross-entropy is commonly employed. As proposed, we employed binary cross-entropy and DSC loss functions to train the networks in this investigation. The loss value applied on SegChaNet is in Equation (2), proving that the overall loss is less sensitive to class imbalance. Our tests found greater segmentation accuracy when utilizing the binary cross-entropy over the individual loss. The equation in [55] is as follows:

$$\zeta(y, \hat{y}) = \zeta_{dc}(y, \hat{y}) + \zeta_{bcc}(y, \hat{y}). \quad (6)$$

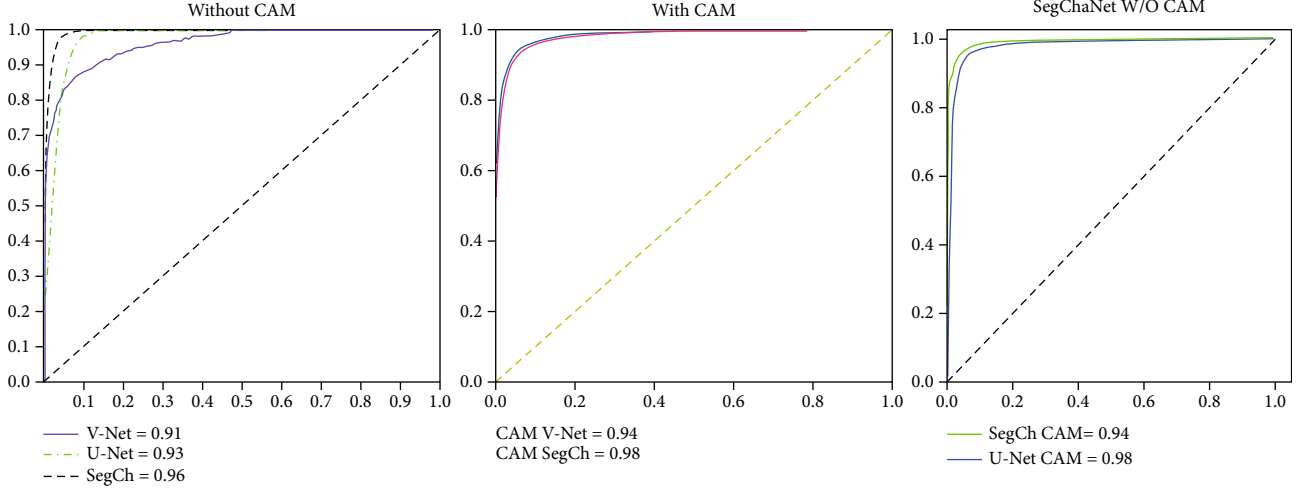


FIGURE 5: SegChaNet, U-Net, and V-Net model with and without CAM.

TABLE 8: Evaluation of segmentation accuracy in DSC, JI for 3 models.

Models	Training	Val	Testing	
	DSC	DSC	DSC	JI
V-Net	0.953	0.907	0.893	0.949
U-Net	0.955	0.893	0.911	0.956
SegChaNet	0.989	0.947	0.937	0.957

In Equation (6),  $\hat{y}$  denotes the model's output, and the ground truth labels are denoted by  $y$ . We use the two-class version of the DSC loss  $\zeta(y, \hat{y})$  proposed in [56], a fully connected DL model trained on 500 epochs with an Adam optimization learning rate of 0.0005. If the validation binary cross-entropy did not improve after 15 epochs, we would reduce the learning rate by 0.1 per ten epochs. We updated the dataset with additional random rotations and flipping to help minimize overfitting and improve the resilience of our technique to varied hippocampus shapes. Our research noticed that the segmentation masks became poorer during additional large rotation angles. So, to remain consistent, we reduced the rotation angles to fall within roughly 10 degrees.

**3.7. Evaluation Metrics.** The precision of segmentation directly affects the success or failure of the segmentation process. Therefore, three measurement variables, DSC, sensitivity, and specificity, are utilized to assess the accuracy of the suggested techniques. In addition, true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN) are also crucial in the assessment [57].

Dice: designed to evaluate the overlap rate of prediction results and ground truth. DSC ranges from 0 to 1, and the better-predicted result will have a more considerable DSC value.

$$DSC = \frac{2TP}{2TP + FP + FN}. \quad (7)$$

Sensitivity (also known as the true positive rate or the recall) indicates the percentage of genuine positives that are accurately detected.

$$Sensitivity = \frac{TP}{TP + FN}. \quad (8)$$

Specificity (also called the true negative rate) measures the proportion of correctly identified actual negatives.

$$Specificity = \frac{TN}{TN + FP}, \quad (9)$$

where TP denotes the true positive voxels, TN denotes the true negative voxels, FP denotes the false positive voxels, and FN denotes the false negative voxels. We use the 4-union (IOU), DSC value, and Hausdorff distance [58]. Here, we defined IOU as

$$IOU = \frac{TP}{FN + FP + TP} \quad (10)$$

To understand what this output means, think of a medical test proposed target capture, in which a group of clinicians manually label the target region, and the network prediction shows the areas around it.

The intersection and union of two sets are equivalent when defined as the IOU ratio [59]. Researchers use the IOU score widely to quantify pixel-level image segmentation performance in image recognition algorithms. IOU can have a value from 0 to 1. As the IOU increases, the overlap between the two zones will decrease and vice versa.

## 4. Results and Discussion

As demonstrated in Figures 2 and 3, we enhanced lung cancer segmentation in the proposed model SegChaNet by comparing it to two state-of-the-art models. Despite having almost identical encoder and decoder designs, U-Net

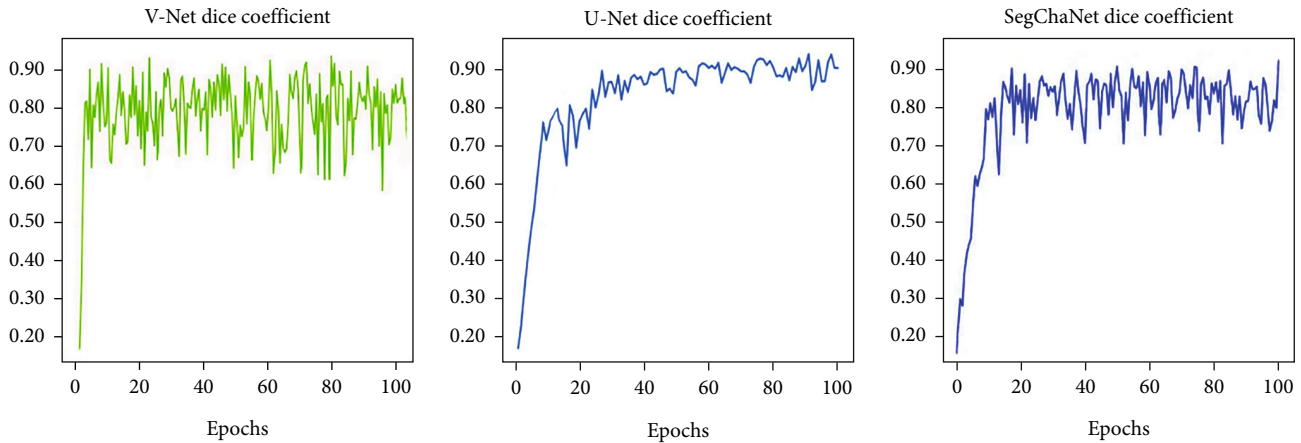


FIGURE 6: Applied model dice coefficient.

structures and their variants outperform non-U-Net structures. While decoders on other networks employ the addition operation, the skip connector and upsampling set of inputs are concatenated in the decoder. We lowered the data dimension to compensate for the value lost during the addition procedure. In addition, to impact the behavior of other network features, we paired SegChaNet's encoder and decoder with a large number of recurrent, residual convolutions. As a result, when we utilized SegChaNet to evaluate the CT scan dataset, we found that it was prone to overfitting. That is why we used CAM, which performs a function by correctly utilizing all of the feature information on each level. Unlike attention U-Net, which segments the esophagus with an air hole, the SegChaNet considers the air hole a boundary when segmenting. Furthermore, U-Net cannot identify small airways [60], required for esophageal analysis. This is due to the underutilization of complex features. However, neither can they use the features of several layers to guide feature selection, nor can they combine low-level and high-level features with driving feature selection at a lower level. On the other hand, CAM uses the channel to combine the features of many stages.

In contrast, CFFM selects low-level, fine-grained features based on high-level semantic features, resulting in the network producing features on the shape and border of the esophagus. As a result, we achieved our primary goal of enhancing segmentation effects and generalization capacities. Figure 6 depicts a comparison of the DSC coefficient performance of the employed models.

When compared to other works in the literature, such as segmentation techniques in [61], we can observe that, while they have profited from alpha attribute maximization, the result is not as expected, with an IOU of 74.18 percent and 84.80 percent. Furthermore, compared to 3D-U-Net neural in [62], they used 3D U-Net, and their DSC performance reached 95.30 percent. Speaking of which, the SegChaNet approach performs significantly better.

As seen from Figure 6, V-Net shows a good performance, but it fails in the final part, while the best performance is from the SegChaNet model. For the first 20

epochs, U-Net shows the worst performance. Figure 7 displays marked images using grad cam.

As seen in Figure 7, we have marked images using grad cam. Also, we utilized the DSC, JI, and normalized surface distance (NSD) with 4 mm performance to test various networks. As illustrated in Figure 8, we examined the segmentation performance of V-Net, U-Net, and U-Net and V-Net with CAM convolutions. The V-Net attained mean DSC scores of 95.75% without CAM, 87.35% for the training, validation, and test sets. At the same time, the U-Net results match the V-Net scores relatively well. CAM convolutions allowed U-Net to improve the scores to 88.61% and 95.94%. The proposed technique demonstrated 96.81% without CAM and 98.48% with CAM for the training, validation, and test sets.

Furthermore, our method outperformed other cutting-edge networks for both the JI and NSD metrics. This is highlighted clearly. Figure 7 shows the segmentation quality of the proposed method next to other methods so that it can be compared visually. When used to color coordinate ground truth with predictions, red represents the ground truth, while yellow, green, and cyan represent the V-Net, U-Net, and method provided in this article predictions. The top-right area of the mask in Figure 7 comprises small discontinuous portions; we demonstrated the advantage of CAM convolutions by effectively segmenting such small sections. However, when the learned qualities become more prevalent worldwide, the U-Net and the suggested method might effectively gather the regions. Figure 8 depicts the three-dimensional surface meshes of two different patients. We can observe how the SegChaNet model influences our conclusions in the method provided in columns 3 and 4. The bottom boundary of the red component of the hippocampus is about one hundred millimeters below the ground's surface. Figure 8 depicts SegChaNet's top and worst performance ratings.

Figure 8 displays the SegChaNet model's best and worst performance. As seen, the best score is about 95 while the worst one is 51.47.

For the segmentation of lung nodules using low-dose CT imaging, we investigated the imaging features of lung

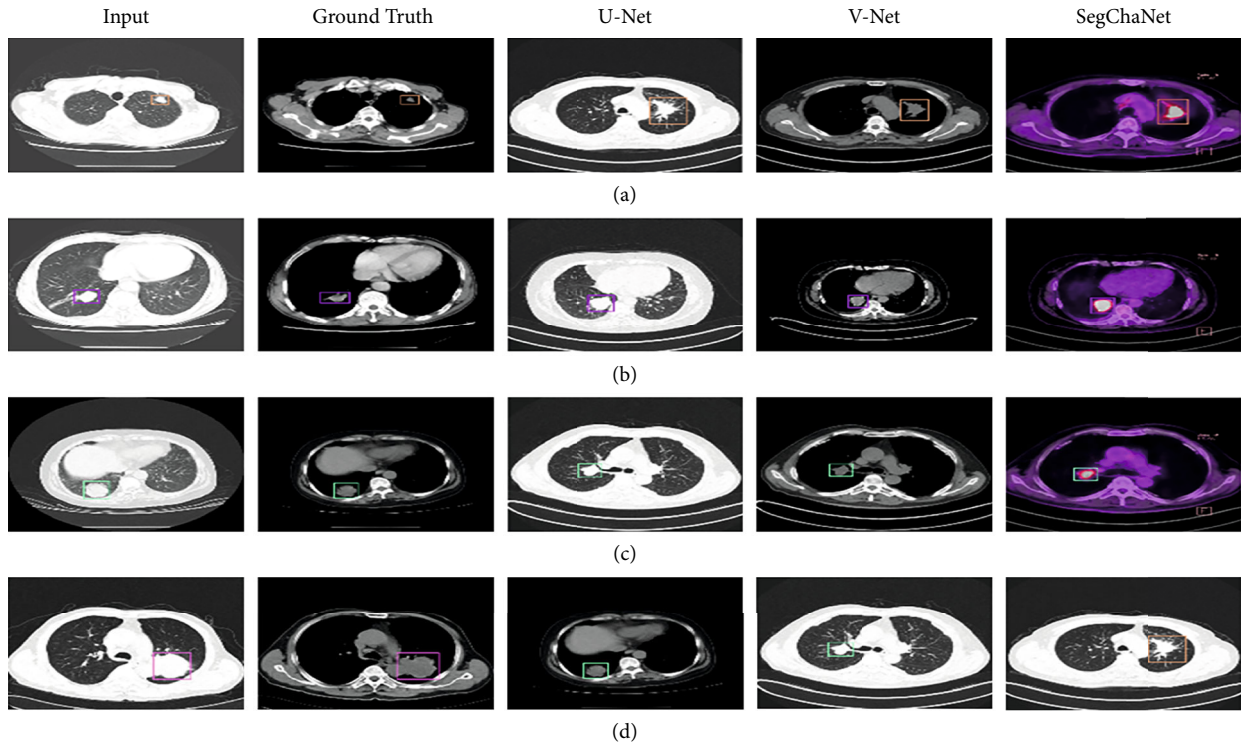


FIGURE 7: Marked images using grad cam.

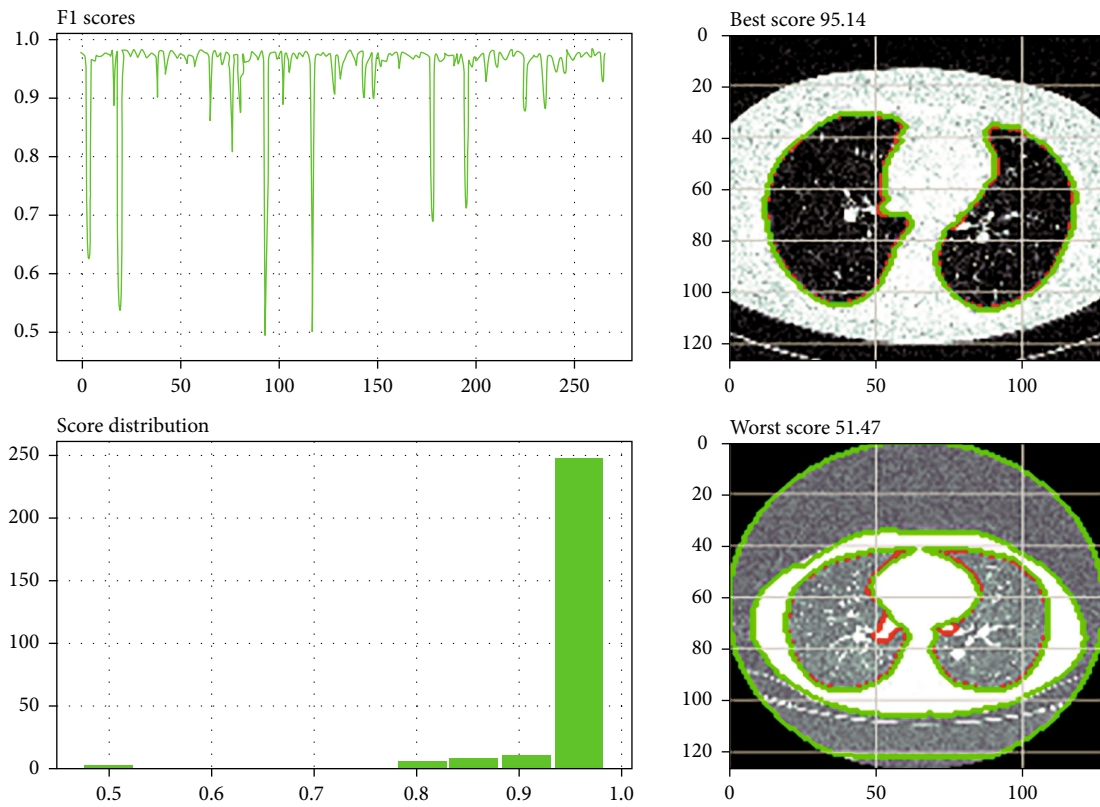


FIGURE 8: SegChaNet's best and worst performance scores.

TABLE 9: Comparison of the performance of the SegChaNet segmentation model with other state-of-the-art works.

Author (year) [reference]	Dataset	Accuracy	Qualitative analysis	Conclusion
Skourt, BA, et al. (2018)	MV, MI, and ME (union) auto: MV phase	0.95	The number of patients with nDSC 1 (within 1-millimeter uncertainty): 7. There are no discernible differences between b-spline and demons	With autocontouring, one may get sharp edges and corners.
Wouter R. P. H., et al. (2021)	All processes are done manually. Auto: dependent on the amount of artifacts, either ME or MI	0.76	NA	In the main breathing phase, there is good agreement between ITV and GTV with manual contours.
Mingjie X., et al. (2019)	Manuel: every stage auto: inherited from the MI phase	0.95	The number of patients that need manual adjustment	Good agreement between auto and manual contours.
Xu, M, et al. (2019)	Manual: all phases Auto: propagate from MI phase	0.92	Minimal. NA	Although autocontouring is precise, it produces bigger shapes.
Qinhua, H., et al. (2020)	Manual: every stage auto: propagated from ME phase	0.94	NA	Deformed contours agree well with physician-drawn contours.
Chiu, T. W., et al. (2021)	Manual: every stage auto: propagated from ME phase	0.63	The propagated IGTVs were mostly within the mIGTVs	The rigid body propagation method generates ITV within a 1 mm margin of error.
Jiang, J. et al. (2018)	Manual: ME (expert) and MI phase auto: ME phase	0.74	NA	The algorithm utilized generated more precise results. Results of segmentation differ from those in previously published papers.
SegChaNet (the proposed network)	Images of 46,698 CT scans both with and without tumors from the cancer imaging archive auto: ME phase	98.90	The applied preprocessing steps	A combined network with two primary components with many CAM has been utilized. We do not employ any manual contours.



nodules ranging in size from 5 to 32 mm in diameter in cancer patients undergoing low and normal voltage CT scans. This study was aimed at developing a DL-based automated lung cancer tumor segmentation network utilizing CT scan segmentation approaches combined with the assessment of segmentation uncertainty. Additionally, we investigated the influence of two widely used cost functions, dice and JI, on the model output's uncertainty measures. All three segmentation models perform well on the segmentation challenge, with dice-trained models marginally outperforming JI-trained models. The most precise methods for predicting nodule volumes were not the most repeatable, emphasizing the need to evaluate their precision and accuracy. Significant variations in algorithm performance were seen, especially in a sample of heterogeneous nodules, highlighting the need to use the same software throughout a longitudinal study.

In contrast to some recently proposed methods, we aimed to increase segmentation prediction variability by producing segmentation hypotheses with our novel method. Importantly, SegChaNet encompasses the entire space of possible segmentations rather than providing information about the models' uncertainty [63]. As is well known, manual, automated, and semiautomatic segmentation pose difficulties in terms of repeatability. The ability to accurately and consistently target lung cancer regions from surrounding tissues using CT images is critical for the clinical assessment of disease progression. The obtained features contribute to SegChaNet throughput and are subsequently decoded to provide an automatic network segmentation. In addition, these features are expected to encapsulate some of the intratumorally and peritumoral area's complex structural and functional geometry, some of which may be related to cancer survival. Overall, we show that SegChaNet, when combined with UQ, shows the best accuracy in lung cancer segmentation. When the existing V-Net models' results are studied, a novel SegChaNet model is proposed that capitalizes on the SegChaNet's superior qualities over the V-Net and U-Net models. On this dataset, three models were run, including the SegChaNet model, and the organ and tumor segmentation procedures were carried out independently. It has been shown that V-Net models can efficiently segregate organs and tumors using computerized images and that by handling the coding and decoding phases separately, more successful models may be developed than present V-Net models. Using medical imaging may be possible to develop more accurate models for multiorgan segmentation. The study may potentially be utilized as a reference for future SegChaNet models due to the success of the original SegChaNet model implementation and the beneficial contribution of the CAM architecture. Since the CAM design was applied solely to the output layer, tiny segmentation characteristics may be recorded. This is crucial for model design since each parameter is only effective when introduced to the appropriate model blocks. Table 9 compares the SegChaNet segmentation model's performance to that of other leading-edge studies.

Table 9 shows the performance of the SegChaNet segmentation model compared with other state-of-the-art works. Because of utilizing preprocessing stages and UQ

methods, SegChaNet demonstrates the highest accuracy and dice outcomes.

## 5. Conclusions

The primary goal of this work is to offer a practical approach to lung segmentation. This paper describes a novel method, SegChaNet, that compared two traditional models (V-Net and U-Net) for lung cancer segmentation in CT scan images. SegChaNet achieves its top performance by utilizing a U-Net framework with CAM convolutions in the network's deepest layer and substantial supervision in its decoder layer. The proposed architecture surpasses state-of-the-art methods in segmentation accuracy and demonstrates initial feasibility in autosegmentation of the lung. Furthermore, it is likely to perform with various datasets and apply to other segmentation challenges in medical imaging. Overall, the SegChaNet method without CAM is 96.81% while with CAM is 98.48%.

Future studies will focus on whether the segmentation results help clinicians in the clinical setting while treating cancer. Additionally, we believe SegChaNet still has potential for improvement in terms of computing costs.

## Data Availability

Data sharing is not applicable to this article as no datasets were generated during the current study.

## Conflicts of Interest

The author declares that there is no conflict of interest.

## References

- [1] H. Sung, J. Ferlay, R. L. Siegel et al., "Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: a Cancer Journal for Clinicians*, vol. 71, no. 3, pp. 209–249, 2021.
- [2] R. Adam, "Chemotherapy and surgery: new perspectives on the treatment of unresectable liver metastases," *Annals of Oncology*, vol. 14, p. ii13-ii16, 2003.
- [3] M. Ilie, C. Butori, S. Lassalle et al., "Optimization of EGFR mutation detection by the fully-automated qPCR-based Idylla system on tumor tissue from patients with non-small cell lung cancer," *Oncotarget*, vol. 8, no. 61, pp. 103055–103062, 2017.
- [4] X. Fan, X. Zhang, Z. Zhang, and Y. Jiang, "Deep learning-based identification of spinal metastasis in lung cancer using spectral CT images," *Scientific Programming*, vol. 2021, 7 pages, 2021.
- [5] V. Donati, G. Fontanini, M. Dell'Omodarme et al., "WVX expression in different histologic types and subtypes of non-small cell lung cancer," *Clinical Cancer Research*, vol. 13, no. 3, pp. 884–891, 2007.
- [6] F. Shariaty, M. Orooji, E. N. Velichko, and S. V. Zavjalov, "Texture appearance model, a new model-based segmentation paradigm, application on the segmentation of lung nodule in the CT scan of the chest," *Computers in Biology and Medicine*, vol. 140, article 105086, 2022.
- [7] N. S. Punn and S. Agarwal, "Modality specific U-Net variants for biomedical image segmentation: a survey," *Artificial Intelligence Review*, pp. 1–45, 2022.

- [8] Y. Yang and C. Yu, "Advances in silica based nanoparticles for targeted cancer therapy," *Nanomedicine: Nanotechnology, Biology and Medicine*, vol. 12, no. 2, pp. 317–332, 2016.
- [9] P. Savadjiev, J. Chong, A. Dohan et al., "Image-based biomarkers for solid tumor quantification," *European Radiology*, vol. 29, no. 10, pp. 5431–5440, 2019.
- [10] V. Bhavsar, S. Zhang, and D. Bhugra, "Globalization, migration, and mental health," *Oxford Textbook of Public Mental Health*, vol. 281, 2018.
- [11] S. Kumar, A. Srinivasan, and F. Nikolajeff, "Role of infrared spectroscopy and imaging in cancer diagnosis," *Current Medicinal Chemistry*, vol. 25, no. 9, pp. 1055–1072, 2018.
- [12] R. Osuala, K. Kushibar, L. Garrucho et al., "A review of generative adversarial networks in cancer imaging: new applications, new solutions," 2021, arXiv preprint arXiv:2107.09543.
- [13] A. González Rodríguez, *Early Detection of Lung Cancer through Nodule Characterization by Deep Learning (Bachelor's Thesis)*, Universidad Autónoma de Madrid, Spain, 2019.
- [14] S. Wang, J. Cao, and P. Yu, "Deep learning for spatio-temporal data mining: a survey," *IEEE transactions on knowledge and data engineering*, p. 1, 2020.
- [15] W. Fullick, *Human Infectious Disease and Public Health*, Oxford University Press, USA, 2019.
- [16] S. M. B. Netto, A. C. Silva, R. A. Nunes, and M. Gattass, "Automatic segmentation of lung nodules with growing neural gas and support vector machine," *Computers in Biology and Medicine*, vol. 42, no. 11, pp. 1110–1121, 2012.
- [17] A. Zotin, Y. Hamad, K. Simonov, and M. Kurako, "Lung boundary detection for chest X-ray images classification based on GLCM and probabilistic neural networks," *Procedia Computer Science*, vol. 159, pp. 1439–1448, 2019.
- [18] M. N. Saad, Z. Muda, N. S. Ashaari, and H. A. Hamid, "Image segmentation for lung region in chest X-ray images using edge detection and morphology," in *In 2014 IEEE International Conference on Control System, Computing and Engineering (ICCSCE 2014)*, pp. 46–51, IEEE, 2014, November.
- [19] X. Li, Y. Yang, Q. Zhao, T. Shen, Z. Lin, and H. Liu, "Spatial pyramid-based graph reasoning for semantic segmentation," in *In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 8950–8959, Seattle, WA, USA, June 2020.
- [20] M. Kim and B. D. Lee, "Automatic lung segmentation on chest X-rays using self-attention deep neural network," *Sensors*, vol. 21, no. 2, p. 369, 2021.
- [21] B. Maga, "Chest X-ray lung and heart segmentation based on minimal training sets. arXiv preprint arXiv:2101.08309," 2021.
- [22] S. Jadon, "A Survey of Loss Functions for Semantic Segmentation," in *In 2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, pp. 1–7, IEEE, 2020, October.
- [23] S. Minaee, Y. Y. Boykov, F. Porikli, A. J. Plaza, N. Kehtarnavaz, and D. Terzopoulos, "Image segmentation using deep learning: a survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, p. 1, 2021.
- [24] G. Zhang, G. Kang, Y. Wei, and Y. Yang, "Few-shot segmentation via cycle-consistent transformer," 2021, arXiv preprint arXiv:2106.02320.
- [25] H. Tang, X. Qi, D. Xu, P. H. Torr, and N. Sebe, "Edge guided GANs with semantic preserving for semantic image synthesis," 2020, arXiv preprint arXiv:2003.13898.
- [26] S. M. Naqi, M. Sharif, and A. Jaffar, "Lung nodule detection and classification based on geometric fit in parametric form and deep learning," *Neural Computing and Applications*, vol. 32, no. 9, pp. 4629–4647, 2020.
- [27] G. Savitha and P. Jidesh, "A holistic deep learning approach for identification and classification of sub-solid lung nodules in computed tomographic scans," in *Computers & Electrical Engineering*, vol. 84, article 106626, Elsevier, New York, United States, 2020.
- [28] Y. Han and J. C. Ye, "Framing U-net via deep convolutional framelets: application to sparse-view CT," *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, pp. 1418–1429, 2018.
- [29] M. Z. Alom, M. Hasan, C. Yakopcic, T. M. Taha, and V. K. Asari, *Recurrent Residual Convolutional Neural Network Based on U-Net (R2U-Net) for Medical Image Segmentation*, 2018.
- [30] L. Liang, I. Zharkov, F. Amjadi, H. R. V. Joze, and V. Pradeep, "Guidance network with staged learning for image enhancement," in *In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 836–845, Nashville, TN, USA, June 2021.
- [31] I. M. Nasser and S. S. Abu-Naser, "Lung cancer detection using artificial neural network," *International Journal of Engineering and Information Systems (IJEAIS)*, vol. 3, no. 3, pp. 17–23, 2019.
- [32] P. Kale and D. S. Yadav, "Lung cancer prognosis by implementing various evolutionary Image processing steps," *By the International journal of analytical and experimental modal analysis*, vol. 12, no. 9, pp. 1103–1117, 2020.
- [33] A. M. Mesleh, "Lung cancer detection using multi-layer neural networks with independent component analysis: a comparative study of training algorithms," *Jordan Journal of Biological Sciences*, vol. 10, no. 4, pp. 239–249, 2017.
- [34] Y. Gu, X. Lu, L. Yang et al., "Automatic lung nodule detection using a 3D deep convolutional neural network combined with a multi-scale prediction strategy in chest CTs," *Computers in Biology and Medicine*, vol. 103, pp. 220–231, 2018.
- [35] A. Chon, N. Balachandar, and P. Lu, *Deep Convolutional Neural Networks for Lung Cancer Detection*, Stanford University, 2017.
- [36] M. A. Khan, S. Rubab, A. Kashif et al., "Lungs cancer classification from CT images: an integrated design of contrast-based classical features fusion and selection," *Pattern Recognition Letters*, vol. 129, pp. 77–85, 2020.
- [37] V. Inturi, N. Shreyas, K. Chetti, and G. R. Sabareesh, "Comprehensive fault diagnostics of wind turbine gearbox through adaptive condition monitoring scheme," *Applied Acoustics*, vol. 174, p. 107738, 2021.
- [38] X. Zhang, L. Ning, Y. Hu et al., "Prognostic factors for primary localized gastrointestinal stromal tumors after radical resection: Shandong Gastrointestinal Surgery Study Group, Study 1201," *Annals of Surgical Oncology*, vol. 27, no. 8, pp. 2812–2821, 2020.
- [39] P. Chaturvedi, A. Jhamb, M. Vanani, and V. Nemade, "Prediction and classification of lung cancer using machine learning techniques," in *In IOP Conference Series: Materials Science and Engineering*, vol. 1099no. 1, p. 012059, IOP Publishing, 2021, March.
- [40] B. Chapaliuk and Y. Zaychenko, "Deep Learning Approach in Computer-Aided Detection System for Lung Cancer," in *In 2018 IEEE First International Conference on System*

- Analysis & Intelligent Computing (SAIC)*, pp. 1–4, IEEE, 2018, October.
- [41] N. Petrellis, “A review of image processing techniques common in human and plant disease diagnosis,” *Symmetry*, vol. 10, no. 7, p. 270, 2018.
- [42] Q. Yuan, M. Sui, C. Qin et al., “Migration, Transformation and Removal of Macrolide Antibiotics in the Environment: A Review,” *Environmental Science and Pollution Research*, vol. 29, no. 18, pp. 26045–26062, 2022.
- [43] H. Cao, H. Liu, E. Song et al., “Multi-branch ensemble learning architecture based on 3D CNN for false positive reduction in lung nodule detection,” *IEEE access*, vol. 7, pp. 67380–67391, 2019.
- [44] H. Xie, D. Yang, N. Sun, Z. Chen, and Y. Zhang, “Automated pulmonary nodule detection in CT images using deep convolutional neural networks,” *Pattern Recognition*, vol. 85, pp. 109–119, 2019.
- [45] W. Sun, B. Zheng, and W. Qian, “Automatic feature learning using multichannel ROI based on deep structured algorithms for computerized lung cancer diagnosis,” *Computers in Biology and Medicine*, vol. 89, pp. 530–539, 2017.
- [46] X. Huang, Q. Lei, T. Xie, Y. Zhang, Z. Hu, and Q. Zhou, “Deep transfer convolutional neural network and extreme learning machine for lung nodule diagnosis on CT images,” *Knowledge-Based Systems*, vol. 204, article 106230, 2020.
- [47] L. M. Pehrson, M. B. Nielsen, and C. Ammitzbøl Lauridsen, “Automatic pulmonary nodule detection applying deep learning or machine learning algorithms to the LIDC-IDRI database: a systematic review,” *Diagnostics*, vol. 9, no. 1, p. 29, 2019.
- [48] D. Sharma and G. Jindal, “Identifying lung cancer using image processing techniques,” *In International Conference on Computational Techniques and Artificial Intelligence (ICCTAI'2011)*, vol. 17, pp. , 2011872–880, 2011.
- [49] M. U. Akram, S. Khalid, A. Tariq, and M. Y. Javed, “Detection of neovascularization in retinal images using multivariate m-Mediods based classifier,” *Computerized Medical Imaging and Graphics*, vol. 37, no. 5–6, pp. 346–357, 2013.
- [50] F. Paulin and A. Santhakumaran, “Classification of breast cancer by comparing back propagation training algorithms,” *International Journal on Computer Science and Engineering*, vol. 3, no. 1, pp. 327–332, 2011.
- [51] T. Jia, D. Z. Zhao, J. Z. Yang, and X. Wang, “Automated detection of pulmonary nodules in HRCT images,” in *In 2007 1st International Conference on Bioinformatics and Biomedical Engineering*, pp. 833–836, IEEE, 2007, July.
- [52] T. Javaheri, M. Homayounfar, Z. Amoozgar et al., “CovidCT-Net: an open-source deep learning approach to diagnose covid-19 using small cohort of CT images,” *NPJ digital medicine*, vol. 4, no. 1, pp. 1–10, 2021.
- [53] M. B. Martell, M. Chen, K. Linton-Reid, J. M. Posma, S. J. Copley, and E. O. Aboagye, “Development of a multi-task learning V-Net for pulmonary lobar segmentation on computed tomography and application to diseased lungs,” 2021, arXiv preprint arXiv:2105.05204.
- [54] Q. Zheng, M. Yang, X. Tian, N. Jiang, and D. Wang, “A full stage data augmentation method in deep convolutional neural network for natural image classification,” *Discrete Dynamics in Nature and Society*, vol. 2020, 11 pages, 2020.
- [55] L. Qian, *Augmented Reality Assistance for Surgical Interventions Using Optical See-Through Head-Mounted Displays(Doctoral dissertation)*, The Johns Hopkins University, United States, 2020.
- [56] J. Ma, J. Chen, M. Ng et al., “Loss odyssey in medical image segmentation,” *Medical Image Analysis*, vol. 71, p. 102035, 2021.
- [57] H. G. Lee and S. C. Lee, “Morphological multi-cell discrimination for robust cell segmentation,” *IEEE Access*, vol. 8, pp. 49837–49847, 2020.
- [58] M. Hashemi, M. Akhbari, and C. Jutten, “Delve into multiple sclerosis (MS) lesion exploration: a modified attention U-net for MS lesion segmentation in brain MRI,” *Computers in Biology and Medicine*, vol. 145, p. 105402, 2022.
- [59] M. A. Rahman and Y. Wang, “Optimizing Intersection-over-Union in Deep Neural Networks for Image Segmentation,” in *In International symposium on visual computing*, pp. 234–244, Springer, Cham, 2016, December.
- [60] T. Pang, S. Guo, X. Zhang, and L. Zhao, “Automatic lung segmentation based on texture and deep features of HRCT images with interstitial lung disease,” *BioMed Research International*, vol. 2019, 8 pages, 2019.
- [61] J. Cai, L. Lu, Y. Xie, F. King, and L. Yang, “Improving deep pancreas segmentation in CT and MRI images via recurrent neural contextual learning and direct loss function,” 2017, arXiv preprint arXiv: 1707.04912.
- [62] A. C. F. D. Oliveira, *Segmentation of Lungs on CT: Tools to Aid Radiotherapy Planning(Doctoral dissertation)*, Universidade de Coimbra, Portugal, 2019.
- [63] Z. Xiao, B. Liu, L. Geng, F. Zhang, and Y. Liu, “Segmentation of lung nodules using improved 3D-UNet neural network,” *Symmetry*, vol. 12, no. 11, p. 1787, 2020.