ISSN: 1672 - 6553

JOURNAL OF DYNAMICS AND CONTROL

VOLUME 9 ISSUE 3: 93 - 108

ENHANCED ATTENTION RESIDUAL PARALLEL U-NETS (EARPU) FOR LUNG TUMOR SEGMENTATION

Dr. Farjana Farvin Sahapudeen[1], Dr. S. Krishna Mohan[2]

[1]Department of Computer Science and Engineering, Sastra Deemed University, Srinivasa Ramanujan Centre, Kumbakonam, Tamil Nadu, India

[2]Department of Mechanical Engineering, E.G.S. Pillay Engineering College, Nagapattinam, Tamil Nadu, India

DOI: https://doi.org/10.71058/jodac.v9i3007

ENHANCED ATTENTION RESIDUAL PARALLEL U-NETS (EARPU) FOR LUNG TUMOR SEGMENTATION

Dr. Farjana Farvin Sahapudeen[1*], Dr. S. Krishna Mohan[2]

[1]*Department of Computer Science and Engineering, Sastra Deemed University,*

Srinivasa Ramanujan Centre, Kumbakonam, Tamil Nadu, India

[2]*Department of Mechanical Engineering, E.G.S. Pillay Engineering College, Nagapattinam, Tamil Nadu, India*

Corresponding Author:farjanafarvin@src.sastra.ac.in

Abstract: Precise identification and classification of diseased tissue and its adjacent healthy structures are vital in the diagnosis of conditions like lung cancer. Achieving a more accurate diagnosis necessitates a substantial amount of data. Yet, physicians often encounter challenges in manually analyzing extensive and intricate CT scan images to extract essential information. While UNet-based architectures have demonstrated superior performance in image segmentation compared to other deep learning architectures, challenges arise in segmentation accuracy due to the low resolution of medical images and insufficient data. In this research, we propose a novel architectural design that addresses these issues by integrating four parallel UNETs through an attention residual network. To enhance performance, this architecture focuses on slicing the single image as four quadrant images and processing them individually rather than the entire image as a whole. This approach allows our model to capture intricate features of the images, as each image slice undergoes independent convolution and deconvolution through four parallel UNets. Ultimately, adhering to the attention residual network architecture, the UNet outputs are merged in a manner that amplifies the features of the image associated with the output through a skip connection. The suggested architecture demonstrated superior performance in terms of Dice score, achieving 91% on LIDI-IRDC, 89% on LUNA16, and 89% on Kaggle, compared to using a conventional U-Net or other U-Net variants.

Keywords: Deep learning, Parallel U-Nets, Residual blocks, Attention Units

1. Introduction:

Lungs are an important organ present in the human chest that transports oxygen throughout the body and eliminates carbon dioxide. In the real –world, detecting tumors in lungs from a huge number of CT images is a completely manual process that relies on the time and knowledge of medical professionals. The treatment of the disease also demands a considerable amount of time and relies heavily on the expertise of specialists

Lung cancer stands as the most perilous form of cancer, claiming lives globally. Countries categorized as middle and lower-income contribute to 50% of annual fatalities. In the United States, lung cancer ranks as the primary cause of cancer-related deaths for both men and women [1].The mortality rate reported by the National Center for Health Statistics reveals that in 2022, the United States documented 609,360 cancer cases. Among these, approximately 350 deaths per day were attributed to lung cancer, emerging as the foremost cause of cancer-related fatalities [2].

DOI: https://doi.org/10.71058/jodac.v9i3007

Various methods are employed to identify lung cancer, such as X-rays, blood tests, biopsies, and CT scans. Due to the diverse characteristics of pulmonary nodules, including variations in size, shape, location, and density, their identification can be challenging. The implementation of a computer-aided diagnostic (CAD) system proves beneficial in patient care, delivering rapid, accurate, and effective diagnoses. Diverse segmentation techniques are utilized by researchers to detect lung cancer nodules. CAD systems incorporating deep learning technologies can reduce the dependency on medical specialists for the identification and classification of lung cancer nodules, especially in tasks like segmentation, detection, and classification.

The U-Net architecture [3] consists of the encoder path and decoder path. The input of the encoder is down-sampled by a series of max pooling and convolution procedures. Final feature maps that have been down-sampled are sent to the decoder path, where they are up-sampled using a similar kind of convolution and max pooling procedure. Thus Noisy U-Net architecture has demonstrated superior performance in lung cancer segmentation [4], proving to be significantly faster than conventional methods.

Furthermore, the model is prone to losing information in specific areas of the image when the architecture uses the entire image as input. However, when dealing with a relatively large input image size, training the model demands a higher GPU memory capacity. The spatial information lost during downsampling is recovered by using lengthy skip connections to skip features from the contracting path to the expanding path [6].

To overcome the aforementioned problems, instead of training the U-Net with an single image, the U-Net model receives the slices of a medical image that have been divided into non-overlapping patches and their corresponding patches of ground truth[5].Any particular detail in the microscope images can be automatically segmented and measured using parallel U-Net[13]. In our proposed Enhanced Attention Residual Parallel U-Nets (EARPU-Net) residual networks are added to restore features that were lost during downsampling. These types of patch-wise image training can concentrate more on local information inside a patch and yield better localization results. Figure 1 shows the 4 quadrant image slices taken from the original lung image.

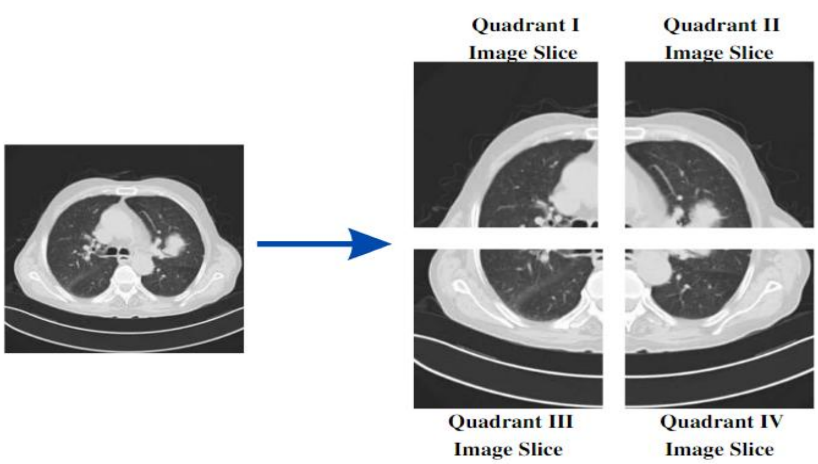


Fig.1. Lung cancer CT image and its slices

DOI: https://doi.org/10.71058/jodac.v9i3007

To summarize, the main contributions in this paper are as follows:

● We present an innovative architecture named Attention Residual Parallel U-Nets, which incorporates novel connections among parallel U-Nets through convolution networks. This design effectively addresses the semantic gap issue between encoder and decoder features by leveraging Attention and residual blocks.

● Our findings demonstrate that scaling the Enhanced Attention Residual Parallel U-Nets (EARPU) architecture enhances its performance, surpassing baseline U-Nets with constrained learnable parameters.

● Through comprehensive experiments, we showcase the proposed model's capability to adeptly learn intricate details of malignant lung features across a range of medical images sourced from diverse modalities.

The sessions are structured as follows: Section 2 provides a review of pertinent areas for additional investigation. In Section 3, we delineate the proposed work and introduce a depthwise encoder-decoder Attention Residual parallel U-Net architecture for binary biomedical image segmentation. Section 4 presents experimental results showcasing the competitive performance of our approach across five standard metrics. Concluding our findings, Section 5 highlights future directions for further research.

2. Related Work

In the examination of medical images U-Net is a crucial method that shows promising results for segmenting images. With limited training data, it can accurately segment images. U-Net variants [7] are widely employed in all main imaging modalities, including microscopes, X-rays, and CT and MRI scans.

The U-Net acts as the foundational model for segmentation, while the bi-directional convolutional long short-term memory (Bi-ConvLSTM) is selected for the extraction and fusion of inter-slice features. Two sequence segmentation strategies [8] can employ both inter-slice and intra-slice features concurrently, enhancing the overall segmentation outcome. Within the UNet++[9] architecture, the encoder and decoder sub-networks are linked by a series of nested, dense skip pathways. These re-imagined skip pathways are designed to minimize the semantic gap between the feature maps of the encoder and decoder pathways.

An attention-UNet [10] in the U-Net variations removes aspects that are unnecessary for the current task. The segmentation performance is greatly enhanced by repeatedly applying the attention gate after every layer, all without adding additional complexity to the model. Res U-Net [10] also comes in another version that tackles the issue of feature identity loss in deeper neural networks due to decreasing gradients in the weight vector. By using skip connections, which take the feature map from one layer and apply it to another layer deeper in the network, ResNet mitigates this issue. This tendency enables the network to perform better for deeper neural networks and better retain feature mappings in deeper neural networks.

DOI: https://doi.org/10.71058/jodac.v9i3007

Prior to combining the encoder and decoder features in Sharp U-Net[11], a depth wise convolution of the encoder feature map with a sharpening kernel filter is used, creating a sharpened intermediate feature map, instead of applying a simple skip connection. Dogan et.al [12] study delves into qualitative and quantitative assessments by employing the Mask R-CNN model to localize affected rough regions. Subsequently, segmentation is performed by refining the identified regions using the 3D U-Net.

Compared to single U-Net, parallel U-Net [13] combined with residual blocks produces promising results in locating the objects from images. Enhanced transformer-based structures integrated into CNN-based blocks to enhance feature extraction capabilities. Specifically designed for medical image segmentation [14], this network architecture incorporates the efficient P-Transformer and a fusion module. The P-Transformer extracts long-range dependencies related to distance, while the fusion module captures local information. The resulting fused features contribute to the U-shaped structure known as P-TransUNet. This model performs parallel weight extraction using convolution and transformer techniques, thereby enhancing features. Attention Residual U-Nets use attention blocks to separate the parts that need immediate treatment and also efficiently identify minimally impacted areas [18].Our proposed model focuses on the parallel connection of Attention Residual U-Nets for image slices to achieve effective segmentation.

3. Methodology

The Enhanced Attention Residual Parallel U-Nets (EARPU) architecture used for lung cancer segmentation is based on the standard U-Net architecture. The segmentation algorithm proposed consists of three processing stages: image preprocessing, image segmentation using the EARPU-Net model, and image post processing. A crucial step in the preprocessing stage involves extracting the region of interest (ROI), which helps eliminate the influence of surrounding organs on the lungs. The second stage of preprocessing is normalizing the input images to achieve a zero mean and a variance of 1. To address the shortage of training images, data augmentation is employed to augment the original dataset. The neighboring four slices of the original image, simultaneously inputted into the parallel U-Net, undergo artificial enhancement for increased training data variety through techniques such as random cropping, elastic deformations, and rotations.

3.1 Preprocessing

Due to memory constraints, computed tomographic (CT) images are scaled to 512 × 512 in order to decrease the size of the CT slices. In addition, before the images were fed into the model for training, they were normalized to reduce low contrast problems. Affected areas have to get more attention when training the neural network with sparse training data. Since augmentations alter CT scans, they may be utilized to reduce overfitting across different images from the existing image collection. When the dataset available is limited, data augmentation is also used to generate more data patches. Increased diversity in training data guarantees generalization of the model.

DOI: https://doi.org/10.71058/jodac.v9i3007

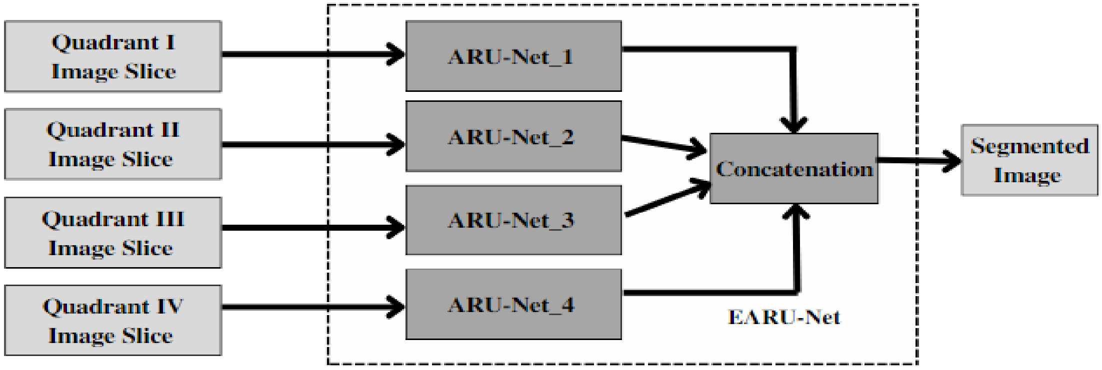


Fig. 2: Illustration of proposed Enhanced Attention Residual Parallel U-Nets(EARPU)

3.2 Lung image Segmentation

We evaluated our proposed model using the open dataset such as LIDC-IRDC dataset [15], Kaggle [17] dataset and LUNA16[16] for training and testing. The dimensions of each CT scan and its segmented 2D image is 572 x 572. We only chose slices ranging from 0 to 256 for slice 1, 256 to 572 for slice 2 in a horizontal orientation, and slices ranging from 0 to 256 for slice 3, 256 to 572 for slice 4 in a vertical orientation. We downsized every image to 256 × 256 because the sizes of these 2D slices varied across different dimensions and planes. As shown in Fig.2 The first slice represented the ‘Quadrant Slice 1’ slice, the second slice represented the ‘Quadrant Slice 2’ slice, , the third slice represented the ‘Quadrant Slice 3' slice, and finally the fourth slice represented the 'Quadrant Slice 4' slice.

The proposed EARP- UNet consists of four typical Attention residual U-Net that was similar to each other. All four input image slices are forwarded to Attention Residual U-Net 1, 2, 3, and 4 in a parallel manner. The resulting image slices are combined using a convolutional layer, which with residual blocks produces the segmented result. Image slice segmentation processes the image slice content in a detailed manner and also produces a faster response compared to a single U-Net. Residual blocks added to each U-Net resolve the vanishing gradient problem. Training in parallel U-Net uses F-fold cross-validation for improving the performance. Attention blocks of each U-Net promise segmenting the relevant and promising portions of the lungs using attention gates. Figure 3 depicts the single-attention residual U-Net of the parallel framework.

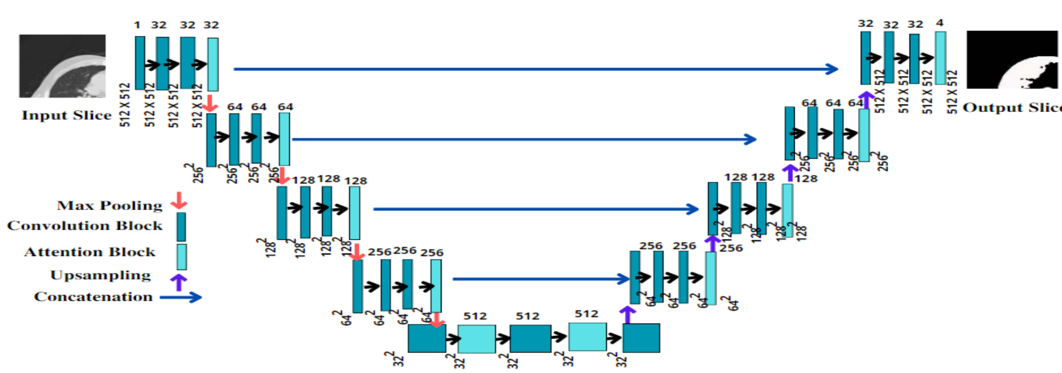


Fig. 3: Illustration of Single Attention Residual U-Net

DOI: https://doi.org/10.71058/jodac.v9i3007

The individual Attention Residual U-Net architectural design comprised a downsampling path and an upsampling path, with both the encoder and decoder adopting fully convolutional network architecture. During downsampling, a 3 × 3 convolution was applied twice, followed by a 2 × 2 max pooling operation that doubled the number of feature channels. On the other hand, the upsampling path involved a 2 × 2 up-convolution, leading to a halving of the feature channel count. This was followed by concatenation with the corresponding downsampling path feature map, and subsequently, two 3 × 3 convolutions were performed. Subsequently, the feature maps from all the distinct U-Net paths were integrated through a residual network, resulting in a unified and single segmented output.

The integration of EARPU-Net features through a residual network involves concatenating the feature maps from each U-Net, as illustrated in Fig.4. The resulting output from the stacked layers is then added to the feature maps of the central slices X2 and X3. Notably, the skip connection is exclusively employed for the feature maps of X2 and X3. The rationale behind this choice is based on the hypothesis that utilizing the skip connection solely for the central slices feature maps will effectively preserve and enhance relevant information, preventing the model from learning unnecessary features from the other slices.

$$H(x') = F(x') + (X2 + X3) \quad \textbf{(1)}$$

Equation (1) summarizes the summation of center slices weighted output as residual input to F(x’)that produces the output. Figure 4 shows the residual learning in EARPU-Nets.

DOI: https://doi.org/10.71058/jodac.v9i3007

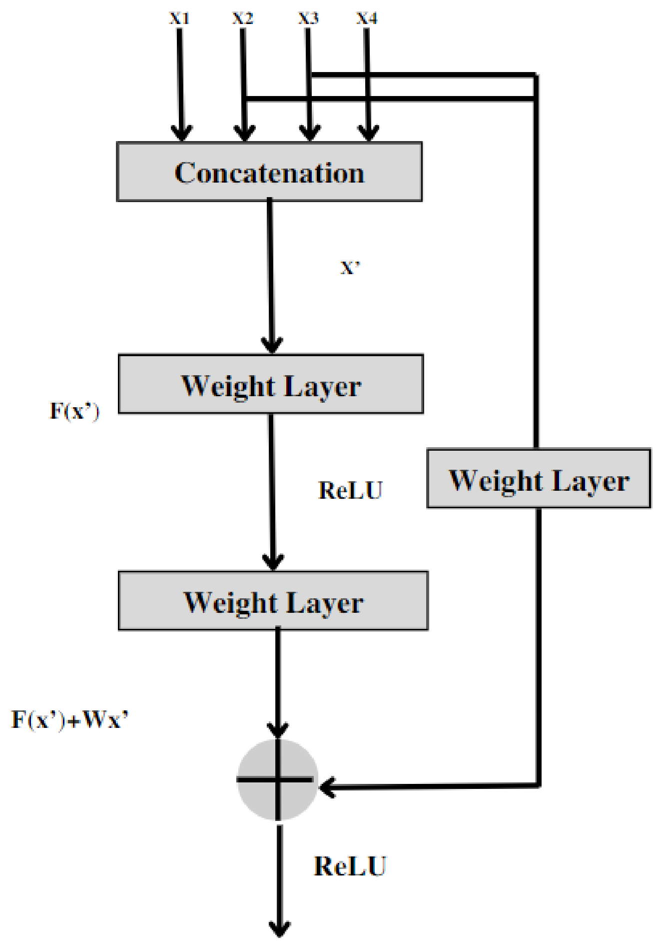


Fig. 4: The proposed building block of residual learning for four attention residual parallel U_Nets. X1, X2, X3, and X4 denote output features from the ARUNet_1, ARUNet_2, ARUNet_3, and ARUNet_4, respectively.

The resultant images features are added with central slice X2, X3 values to generate the output image. This type of fusion enables the EARPU-Nets to avoid the significant data losses occurring in training the models. Residual blocks helps in regenerating the important feature details existing in central slices X2,X3.Instead of depending on single central slice, two slices X2,X3 can elaborate the details of significant portions. The rectified linear unit (ReLU) was employed as the activation function for the input and hidden layers, while the sigmoid function was utilized for the output layer.

3.3 Post Processing

Convoluted image patches are post-processed for detailed review of clinical validation and integration. In order to recognize the infected regions after segmentation, more visual-contrasting techniques are applied. These kinds of segmentations improve volumetric analysis and quantify the affected regions. At different levels of training in terms of epochs, the resultant images are registered for further comparative study. Integrating segmented images helps medical experts detect abnormalities or initiate the diagnosis process.

4. Experiments and Results

DOI: https://doi.org/10.71058/jodac.v9i3007

Experiments run under Windows 7 Operating Systems with Google Colab. The LIDC-IDRI[15],LUNA16[16], Kaggle[17] dataset is utilized to train the suggested EARP model. The EARP U-Net models were created utilizing Python, an open-source programming language, Tensorflow 2.1 and GPU-supporting Keras API.

4.1 Data sets

We demonstrate and analyze the performance of the proposed image segmentation model on three datasets such as LIDC_IDRI [15], LUNA16 [16] and Kaggle [17]. Images collected from listed data sources used for training and testing the EARPU-Net model. The images feeded to individual U-Net model is 572X572.CT images collected from the LUNA16 [16] and Kaggle[17] are resized as 572X572 for training and testing purpose. Table 1 illustrates the medical images and their respective data sources utilized in these experiments.

Table 1: Medical Images used in the experiments

Input Dimension	Data Source
256 x 256	Kaggle Dataset[17]
512 x 512	LIDC-IDRI dataset[15]
256 x 256	Luna16 dataset[16]

4.2 Loss function

In training neural networks we want to reduce the loss function. The cross-entropy loss function is applied to both the encoder and decoder networks. Optimization is carried out using the adaptive moment estimation (Adam) optimizer throughout 120 epochs, with an initial learning rate set to 0.005 and a batch size of 8. To mitigate the risk of network overfitting, dropout is incorporated into the network layer. The cross-entropy function calculates how our proposed EARPU-Net model is defining the difference between the predicted probability and the ground truth value. In minimizing loss values, the model learns to assign higher probabilities for correctly classified samples and lower probabilities for incorrect samples. This type of minimization of cross-entropy loss values drastically improves the accuracy of the model. The following Equation (2) depicts the cross entropy loss used for our proposed EARPU-Net where P_i depicts true probability distribution and log P_i depicts the predicted probability distribution respectfully.

$$L = -1/(|P|)\sum€yilog(\sigma) \quad (2)$$

4.3 Performance metrics

We applied the Jaccard Index (JI) and the Dice Similarity Coefficient (DSC), Accuracy and Sensitivity for the performance evaluation. The Jaccard Index (JI) is calculated by dividing the union area of the anticipated and ground-truth images by the overlapping area between them.

The mathematical formulation of these scores is described by the equations (3) to (6).

$$\text{Dice Score} = \frac{2.TP}{2.TP+FP+FN} \quad (3)$$

DOI: https://doi.org/10.71058/jodac.v9i3007

$$\text{Jaccard Score} = \frac{TP}{TP+FP+FN} \tag{4}$$

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+FN+TN} \tag{5}$$

$$\text{Sensitivity} = \frac{TP}{TP+FN} \tag{6}$$

Another metric is the DSC, which is the ratio of the total number of pixels to two times the overlapping area of the anticipated and ground truth images. Accuracy is the ratio between the numbers of correct predictions with total number of predictions. Sensitivity refers to the accurate identification of positive samples among all the actual positive samples. This performance metric is calculated by taking the ratio of true positives to the sum of true positives and false negatives.

4.4 Results and Discussion

Comparison of various U-Net models for lung tumor segmentation using Dice Score, Accuracy, Sensitivity and Jaccard Score is displayed in Table 2.From Table 2, it is seen that the accuracy for the LIDC-IDRI dataset is higher than other models. The proposed Enhanced Attention Residual Parallel U-Net (EARPU-Net) improves Dice Score, Jaccard Score and Sensitivity when compared to simple U-Net, Attention U-Net, Attention Res_U-NetGD architectures.

It is observed from the results that the accuracy dramatically improves as the images are sliced as image patches. Enhanced Attention Residual Parallel U-Net (EARPU-Net) model requires fewer parameters and produces better results for Kaggle and LUNA16 dataset, comparing to non parallel models. Specifically, the developed EARPU-Net model achieves Dice Similarity Coefficient (DSC), Accuracy (ACC), Jaccard Similarity Coefficient (Jaccard SC), and Sensitivity (SEN) scores of 90%, 92%, 91.5%, and 92.8%, respectively. Notably, it demonstrates improvements of 1%, 0.5%, 0.5%, and 0.8% in DSC, ACC, SEN, and Jaccard SC scores, respectively, compared to the MARU-Net for the LIDC-IDRI dataset.

Table 2: Comparison of lung cancer segmentation performance of various U-Net models in CT images of LIDC-IDRI, Kaggle and LUNA 16 dataset

JOURNAL OF DYNAMICS AND CONTROL
DOI: https://doi.org/10.71058/jodac.v9i3007

Dataset	Models	Dice Score (%)	Jaccard Score (%)	Accuracy (%)	Sensitivity (%)
LIDC-IDRI[15]	U-Net[3]	81.5	89	88	87
	AU-Net[10]	86.8	91	89	88
	ARes-UNet[18]	87.5	91.5	89.6	90
	Proposed EARPU-Net	91	92	92.5	91.5
LUNA 16[16]	UNet[3]	82	85	87	85.7
	AUNet[10]	86	88	88.5	88
	ARes-UNet[18]	87.2	90.5	89	90
	Proposed EARPU-Net	89	91.5	91	90.7
Kaggle[17]	UNet[3]	81	85	86	86.5
	AUNet[10]	86	89	88.8	88
	ARes-UNet[18]	87.5	90	90.5	91
	Proposed EARPU-Net	89	91	92	92

The proposed Enhanced Attention Residual Parallel U-Net (EARPU-Net) model demonstrates substantial performance on the LUNA16 dataset, achieving Dice Similarity Coefficient (DSC), Accuracy (ACC), Jaccard Similarity Coefficient (Jaccard SC), and Sensitivity (SEN) scores of 89%, 91.5%, 91%, and 92.4%, respectively. This represents an improvement of 1.3%, 0.5%, 0.3%, and 0.9% in DSC, ACC, SEN, and Jaccard SC scores compared to the MARU-Net. Additionally, for the Kaggle dataset, the proposed model exhibits improvements of 0.8%, 0.57%, 0.7%, and 0.57% in DSC, ACC, SEN, and Jaccard SC scores relative to the MARU-Net. The details presented in Table 2 and above underscore the robustness of our suggested method in comparison to non-parallel U-Net models.

The Enhanced Attention Residual Parallel U-Net (EARPU-Net) proves effective in achieving accurate lung nodule segmentation. The results of the segmentation are presented in Fig.5, with column (A) displaying the original representation of the lungs. Column (B) illustrates the segmentation using U-Net, column (C) showcases lung nodule segmentation by Attention U-Net, and column (D) presents the predicted outcomes by the Enhanced Attention Residual Parallel U-Net (EARPU-Net).

Original Image / U-Net [3] / AU-Net [10] /ARes-UNet [18] /ARPU-Net

DOI: https://doi.org/10.71058/jodac.v9i3007

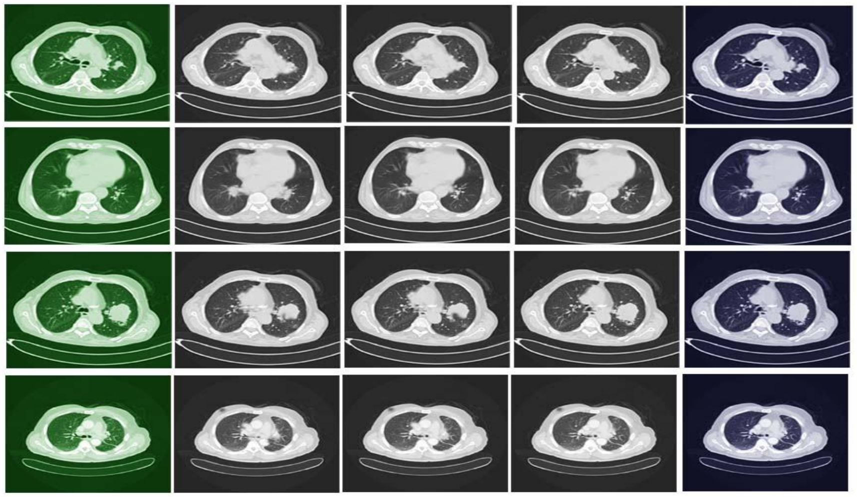


Fig. 5. Predictions generated by our proposed EARP model and state-of-the-art models from test data; (column -wise) 1. Original Image 2.UNet 3.AUnet 4.ARU-Net 5.EAPU-Net

Following the completion of the preprocessing steps for lung images, the original images are partitioned into four equal slices, and these slices are utilized in the training and validation processes. Images of diverse sizes, including small and irregularly shaped ones, are utilized for both training and testing. As indicated in Table 2, the recommended EARPU-Net model demonstrates outstanding overall performance. This proposed model adeptly segments image slices with meticulous detail and integrates the segmented outcomes through convolutional processes. The results, as depicted in the showcased images, are highly promising. In the context of computer-aided detection (CAD) systems, the imperative task involves detecting nodules and subsequently segmenting all identified nodules.

Using the LIDC-IRDI dataset, the proposed Enhanced Attention residual Parallel U-Net (EARPU-Net) had the highest training and validation accuracy of 91.5%, and 91%, respectively among the five segmentation models tested. The details of comparison of various model accuracy is depicted in Figure 6.Furthermore, to predict the accuracy of the binary segmented masks, Figure 7 displays the training accuracy and validation accuracy using the LUNA16 dataset.

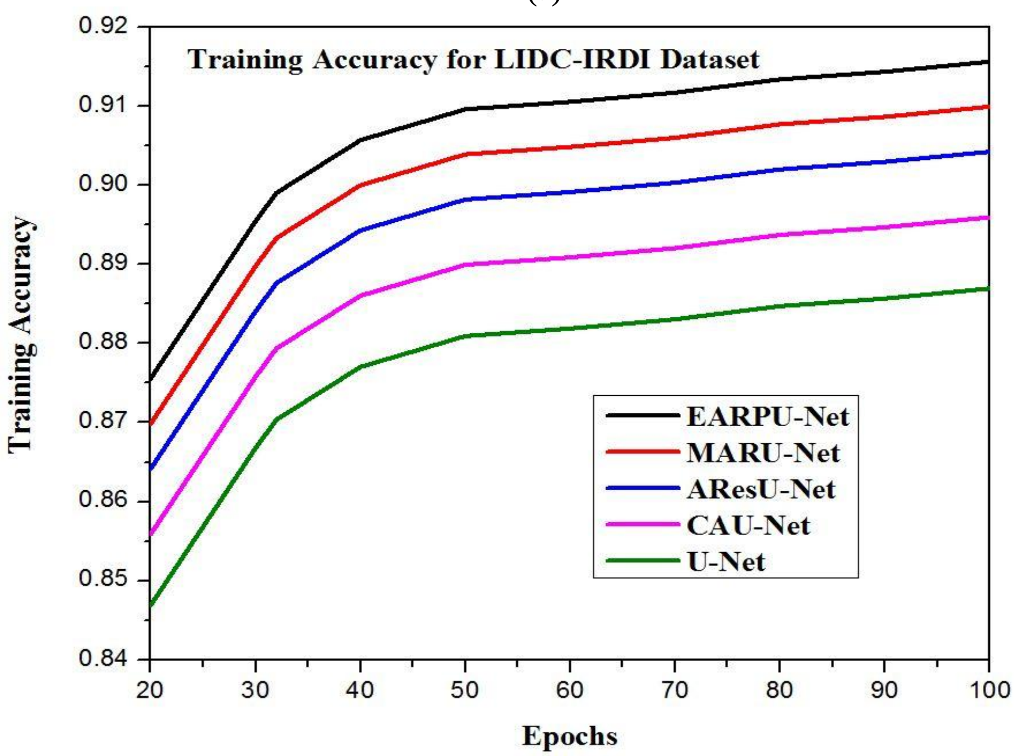
The proposed Enhanced Attention Residual Parallel U-Net (EARPU-Net) achieves the highest training and validation accuracy of 90% and 91%, respectively, when evaluated on the LUNA16 dataset test set, surpassing the performance of the other five tested segmentation models. Moreover, the proposed Enhanced Attention Residual Parallel U-Net (EARPU-Net) model achieved

DOI: https://doi.org/10.71058/jodac.v9i3007

the training accuracy and validation accuracy on the Kaggle derived dataset are 91 %, and 90.8%, respectively shown in Figure 8.

The results of our experiment, presented in Table 2, affirm that the EARP-U-Net model outperforms all other referenced models in performance metrics, including Dice Similarity Coefficient, Jaccard Index Score, Sensitivity, and Specificity. By considering both the tabulated results and the visual representation in Figure 5, it is evident that our proposed EARPU-Net model demonstrates superior performance.

(a)



DOI: https://doi.org/10.71058/jodac.v9i3007

(b)

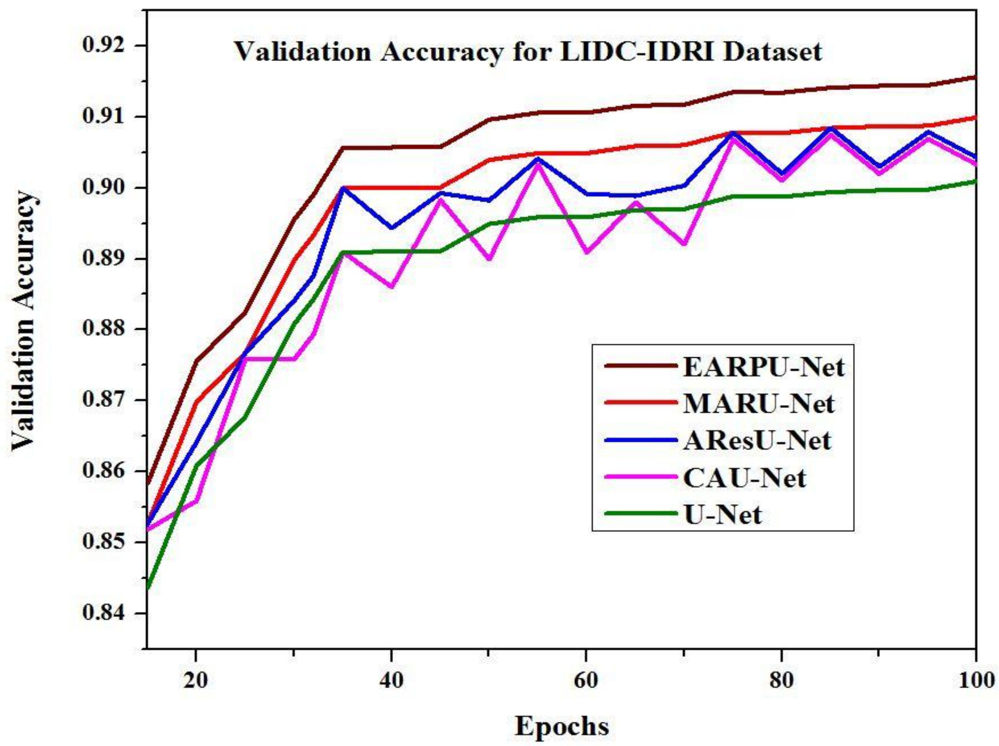
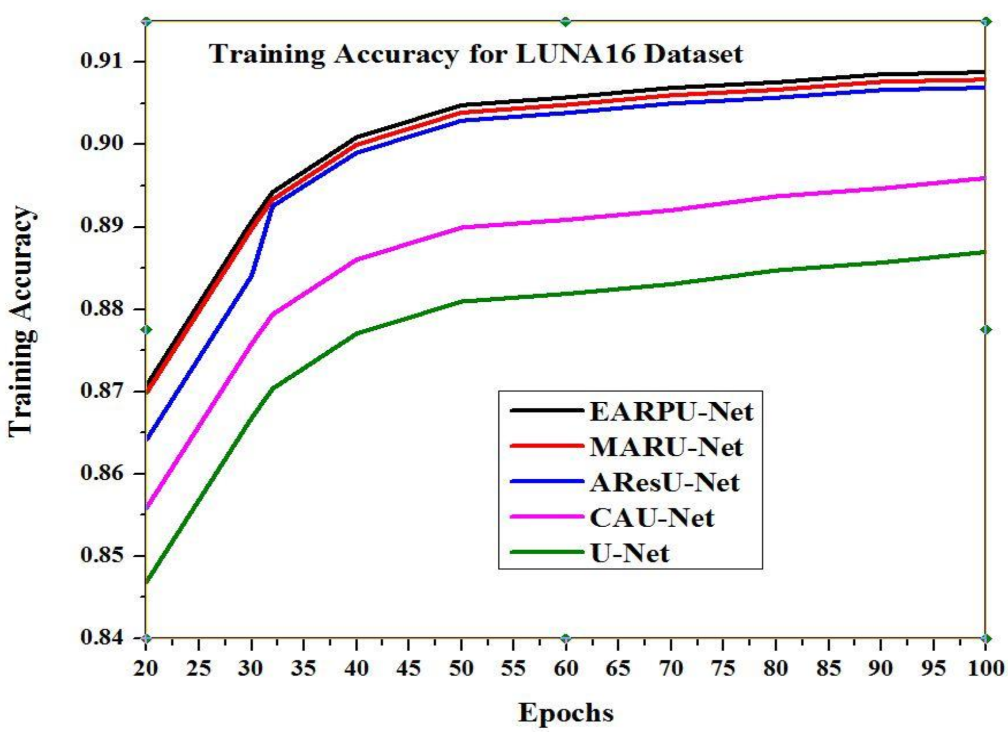


Fig.6. (a) Training Accuracy and (b) Validation Accuracy for Samples of the LIDC-IDRI[15] Dataset

(a)



DOI: https://doi.org/10.71058/jodac.v9i3007

(b)

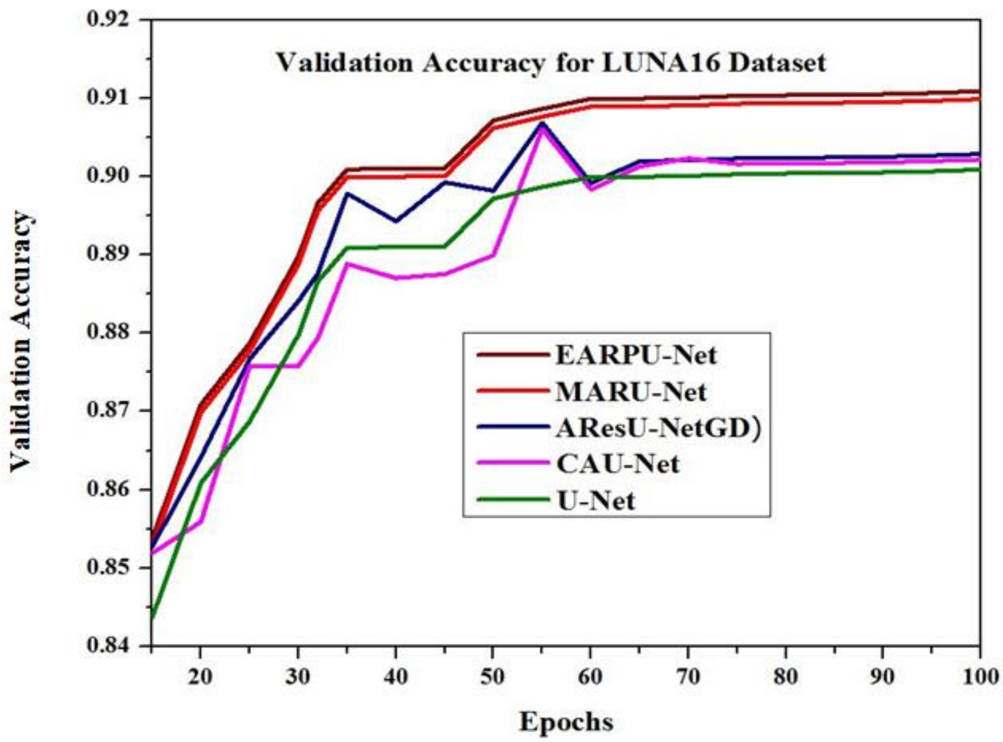
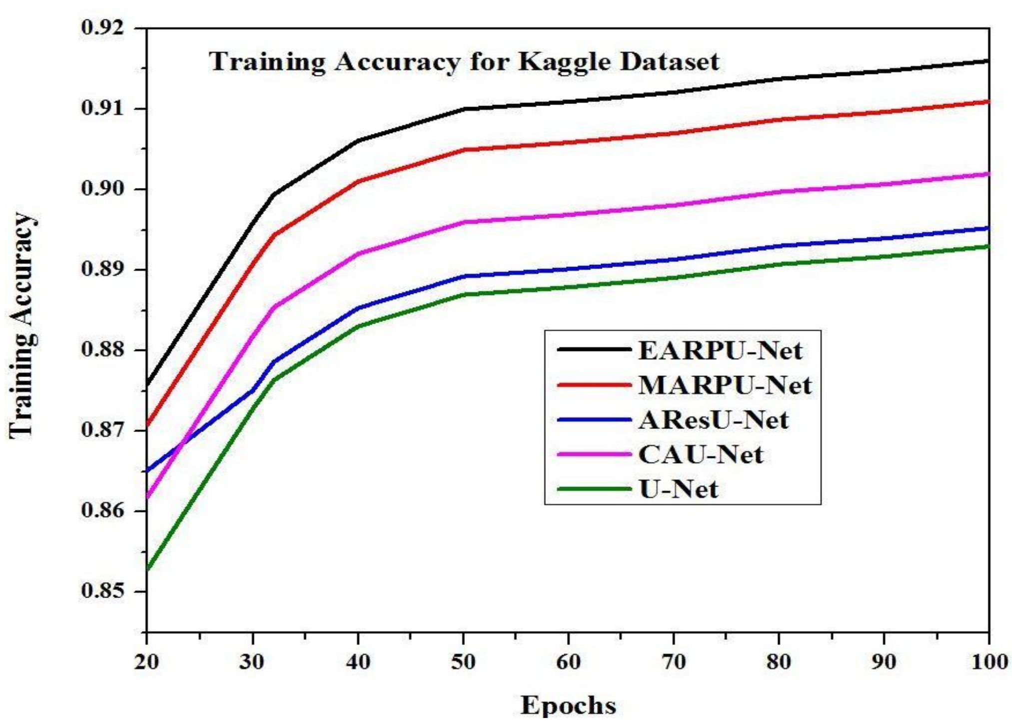


Fig. 6. (a) Training Accuracy and (b) Validation Accuracy for samples of LUNA16[16] Dataset.

(a)



(b)

DOI: https://doi.org/10.71058/jodac.v9i3007

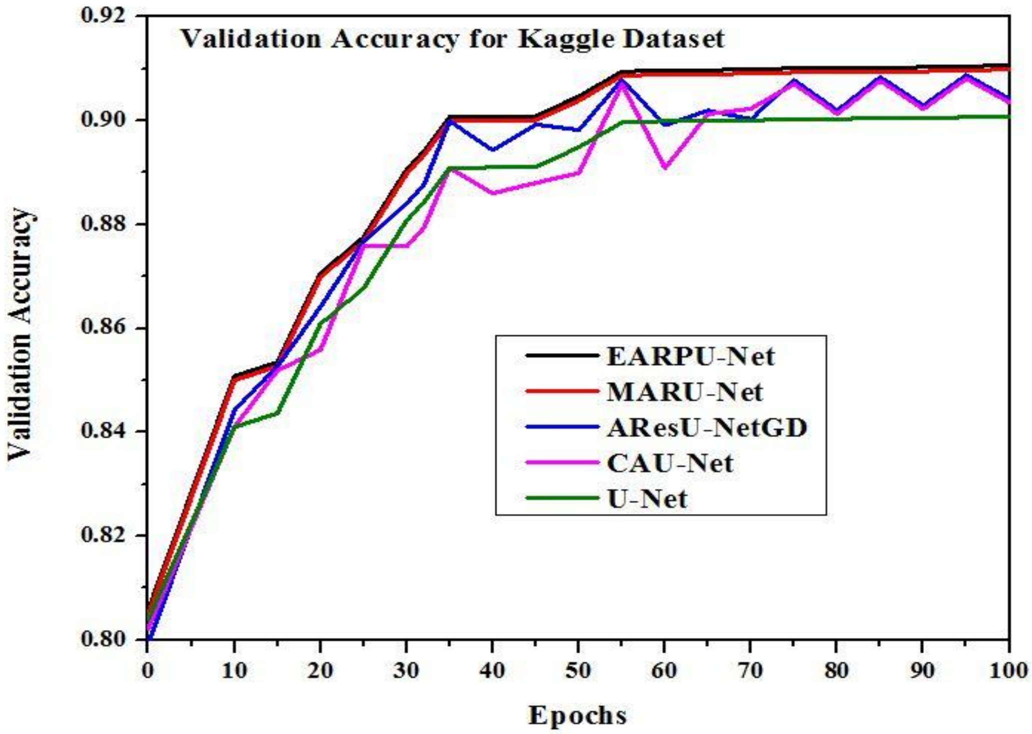


Fig. 7. (a) Training Accuracy and (b) Validation Accuracy for samples of Kaggle[17] Dataset.

5. Conclusion

In this study, we proposed the Enhanced Attention Residual Parallel U-Net (EARPU-Net) that is based on the U-shaped encoder-decoder structure for lung nodule segmentation. Our EARPU-Net generates global and local features of images in parallel mode on the edge of the lesion area to improve the segmentation performance. In our study we used cross entropy loss function in the training to guide the model to focus on the area of the interest so that the EARPU-Net model can discriminate details between affected region and non affected regions. Experiments on three datasets for multiple lung image segmentation tasks demonstrated that our EARPU-Net outperformed conventional U-Nets in terms of Dice score, achieving 91% on LIDI-IRDC[15], 89% on LUNA16[16], and 89% on Kaggle[17]. Additionally, advanced experiments further validated the effectiveness of the EARPU-Net model. In the future, we will focus on designing a more lightweight structure based on parallel computing in the medical field for further research.

Funding Statement: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References:

DOI: https://doi.org/10.71058/jodac.v9i3007

[1] Torre, Lindsey A., Rebecca L. Siegel, and Ahmedin Jemal. "Lung cancer statistics." *Lung cancer and personalized medicine: current knowledge and therapies* (2016): 1-19.

[2] Sung, Hyuna, 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* 71.3 (2021): 209-249.

[3] Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*. Springer International Publishing, 2015.

[4] Huang, Wenkai, and Lingkai Hu. "Using a noisy U-net for detecting lung nodule candidates." *IEEE Access* 7 (2019): 67905-67915.

[5] Lee, Bumshik, Nagaraj Yamanakkanavar, and Jae Young Choi. "Automatic segmentation of brain MRI using a novel patch-wise U-net deep architecture." *Plos one* 15.8 (2020): e0236493.

[6] Nodirov, Jakhongir, Akmalbek Bobomirzaevich Abdusalomov, and Taeg Keun Whangbo. "Attention 3D U-Net with Multiple Skip Connections for Segmentation of Brain Tumor Images." *Sensors* 22.17 (2022): 6501.

[7] Siddique, Nahian, et al. "U-Net and its variants for medical image segmentation: theory and applications." *arXiv preprint arXiv:2011.01118* (2020).

[8] Li, Jing, et al. "Study on strategy of CT image sequence segmentation for liver and tumor based on U-Net and Bi-ConvLSTM." *Expert Systems with Applications* 180 (2021): 115008.

[9] Zhou, Zongwei, et al. "Unet++: A nested u-net architecture for medical image segmentation." *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings 4*. Springer International Publishing, 2018.

[10] Oktay, Ozan, et al. "Attention u-net: Learning where to look for the pancreas." *arXiv preprint arXiv:1804.03999* (2018).

[11] Zunair, Hasib, and A. Ben Hamza. "Sharp U-Net: Depthwise convolutional network for biomedical image segmentation." *Computers in Biology and Medicine* 136 (2021): 104699.

[12] Dogan, Ramazan Ozgur, et al. "A two-phase approach using mask R-CNN and 3D U-Net for high-accuracy automatic segmentation of pancreas in CT imaging." *Computer Methods and Programs in Biomedicine* 207 (2021): 106141.

[13] Kong, Yan, et al. "Automated yeast cells segmentation and counting using a parallel U-Net based two-stage framework." *OSA Continuum* 3.4 (2020): 982-992.

[14] Chang, Keng-Chih, et al. "Locating waterfowl farms from satellite images with parallel residual u-net architecture." *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 2020.

[15] Armato III, Samuel G., 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* 38.2 (2011): 915-931.

[16] Setio, A. A. A., Traverso, A., de Bel, T., Berens, M. S., Bogaard, C. van den, Cerello, P., ... & van Ginneken, B. (2016). LUNA16: A challenge for computerized lung nodule detection. Retrieved from https://luna16.grand-challenge.org/

[17] Saeid, R. (n.d.). Lung Tumor Segmentation Dataset [Data set]. Kaggle. Retrieved from https://www.kaggle.com/datasets/rasoulisaeid/lung-cancer-segment.

[18] Mu, N., Lyu, Z., Rezaeitaleshmahalleh, M., Tang, J., & Jiang, J. (2023). An attention residual u-net with differential preprocessing and geometric postprocessing: Learning how to segment vasculature including intracranial aneurysms. *Medical image analysis*, *84*, 102697.