IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 8 Issue 4, April 2021
ISSN (Online) 2348 – 7968 | Impact Factor (2020) – 6.72
www.ijiset.com

Comparative Performance Analysis of Deep Convolutional Neural Network for Gastrointestinal Polyp Image Segmentation

Syed Qamrun Nisa and Amelia Ritahani Ismail

Department of Computer Science, Kulliyyah of Information and Communication Technology,
International Islamic University Malaysia, P.O. Box 10, 50728 Kuala Lumpur, Malaysia
syed.qamrunnisa@gmail.com

Abstract

Image segmentation is the most challenging and emerging field nowadays for medical image analysis. Polyp image segmentation is a difficult task due to the variations in the appearance and color intensity of the polyps in colonoscopy images. In this paper, we use a dataset of gastrointestinal polyp images for segmentation. The segmentation methods for gastrointestinal polyp images in this paper are based on three deep convolutional neural network models that are FCN, U-NET, and, hybrid Unet_Resnet. Data augmentation is applied to the dataset to increase the accuracy rate. The performance of the three models is evaluated by metrics that are Intersection of Union (IOU) and Dice Similarity Coefficient (DSC). The hybrid model, Unet_Resnet achieves higher IOU, and DSC of 0.75 and 0.86 respectively, which outperforms the other two models FCN and U-Net in gastrointestinal polyp image segmentation.

Keywords: *Images Segmentation, Gastrointestinal Polyp Images, Deep Convolutional Neural Network, IOU, DSC*

1. Introduction

Gastrointestinal polyps are the irregular development of cells in gastric and colonic mucosa. This irregular development is a gradual process and, in most cases, it does not cause symptoms until it reaches a large scale. Nevertheless, if polyps could be identified early, cancer is preventable and treatable [1]. Polyp segmentation is a difficult task due to differences in the form and color intensity of polyps in colonoscopy images[2]

Segmentation is one of the efficient and popular technique for image analysis, which is the procedure of partitioning of digital images into various segments to increase the quality of the images. Segmentation in medical images is used to extract the region of interest for medical image analysis [3]. Medical images have to go through several steps before the diagnosis of the disease. Initially, the images have to be acquired, and preprocessing has to be done and the data has to be stored in memory. It requires a huge amount of memory and processing time. In medical applications, in order to find the detail of the images, it is necessary to process the images [4].

Image analysis is a technique of obtaining information by measuring objects within an image. For image analysis, there are various image processing techniques such as preprocessing of images, edge detection, and segmentation of images[5].

2. Literature Review

Medical images play an essential role in diagnosis and monitoring the condition of the patient's health [6]. Medical images such as Magnetic Resonance Imaging (MRI), Computed Tomography (CT) scan, X-Rays, Ultrasound, and Positron-emission Tomography (PET) are used for a visual representation of the internal body for clinical studies, diagnosis, and treatment planning. Medical images contain rich features, these are the images with high resolution, massive amounts, and complex features [7]. The medical images have been used and stored continuously for diagnosis as well as research purposes [8].

Image analysis is the process to extract the information from images. Image analysis has become very useful for industrial purposes and in research because of its ability to process digital images and objectively, without disturbing the sample, analyze parameters such as size, color, distance, and a number of particles [9]. The various image processing techniques for image analysis are image pre-processing, image compression, edge detection, and image segmentation [5]. Image pre-processing is used to remove background noise to enhance data images prior to computational processing. Image compression is an application of data compression that reduces the redundancy of the image and to store and transmit data in an efficient form [5]. Edge detection is a technique that uses border recognition of linked objects or regions, and this is used to identify the discontinuity of the object [10]. Image segmentation is a technique in which an image is subdivided into multiple segments, each segment represents the information [11].

Image segmentation plays an important role to analyze the image. It is a fundamental stage to analyze the image and extract data from them [12]. Image segmentation is a process that divides an image into regions or segments which is used to locate boundaries and objects such as curves, and lines in images [13]. For image segmentation, commonly used segmentation techniques are Threshold-based image segmentation, Clustering-based image segmentation, Edge-based image segmentation, Region-based image segmentation, and Neural Network-based image segmentation [14]. Segmentation is an important step for the analysis and interpretation of medical images. The purpose of doing image segmentation is to increase the quality of medical images [3]. Segmentation is used to detect and extract the feature areas in the medical images [15]. Image segmentation extracts features from medical images for better and accurate medical diagnostics [14].

Wu, Jie, et al, 2008, proposed a segmentation algorithm based on a seeded region growing (SRG) algorithm for the automatic organ segmentation of MR images of abdominal. They have tested this algorithm on 12 sets of 3D MR images of abdominal. Initially, they extract texture features in the Region of Interest (ROI) for each pixel based upon the homogeneity criteria. Co-occurrence and semivariogram are the two texture features that have been analyzed. Co-occurrence is effective for biomedical image processing, whereas a semi-variogram is a commonly used scale for geostatistics dissimilarity. Then the automatic seeded region growing algorithm is applied to the feature spaces. The cost function is decreased with three factors to calculate the seed. The lower value is taken to acquire the threshold prior to the one producing 'explosion'. Few modifications are implemented to prevent over-segmentation, under-segmentation, and accelerate the evaluation. Seeded region growing algorithm is then implemented and the extraction of the right kidney is in the experiment [16]

Cao, et al, 2016, proposed advanced segmentation algorithms for whole heart CT series images. They improve the two traditional segmentation algorithms, one algorithm is region-based and another algorithm is an edge-based segmentation and then they apply these improved segmentation algorithms to the CT images of the whole heart. This study has evaluated the algorithms for segmentation on three features: performance of an algorithm, evaluation of segmentation results of medical images by subjective evaluation, and Jaccard coefficient and Dice coefficient deviation method are used to calculate the performance of segmentation. When these two advanced algorithms for segmentation are compared to the results of manual segmentation, the advanced segmentation algorithms demonstrate excellent performance in the accurate segmentation of the CT sequence images of the whole heart [17].

Recently, great progress has been achieved in the development of efficient and accurate algorithms of machine learning for biomedical image segmentation [16]. A rapid increase in deep learning-based methods has overwhelmed the machine learning community [18]. Segmentation of images is effectively addressed by different types of deep learning methods, including Artificial Neural Networks (ANNs), Recurrent Neural Networks (RNNs), and Convolutional Neural Networks (CNNs) [18]. CNN performs well on issues pertaining to biomedical images, convolutions are a significant aspect of CNN's performance in image analysis processes such as classification and segmentation [3].Convolutional Neural Network has different types of architectures in deep neural networks. Some of the Convolutional Neural Network architectures are FCN, VGG, AlexNet, ResNet, U-NET, and SegNet [19].

VGG, AlexNet, and ResNet are considered to be the most common architecture that performs successfully for image recognition and image classification tasks [20].

Fully Convolutional Network (FCN) is a segmentation approach that allows pixel-wise prediction based on the ground truth images. FCN extracts crucial feature maps and restores those feature maps to the image labels. The ability of FCN to represent feature is the main reason to achieve success in classification, object detection, and image segmentation. FCN will perform well only when a large number of the training dataset is available. In the field of medical imaging, data collection is expensive and is less available because of privacy and regulation concerns[21]. FCN provides less accurate results without enough training data [22] and results in inaccurate segmentation for small organs [23]. So FCN is extended to U-Net to do segmentation on medical images that perform well with comparatively less number of the training dataset [21].

SegNet is the CNN architecture for the image segmentation task, memory and efficient performance are the significant features of this architecture [24]. But U-Net shows a high speed of process during the training phase [25]. The U-Net performance is higher than the SegNet [26].

U-Net is one of the most popular Convolutional Neural Network architectures for medical image segmentation tasks [27]. U-Net network mainly consists of a convolutional layer, downsampling (contracting path), upsampling (expensive path), and activation function[28]. The architecture of U-Net is shown in Fig. 1.

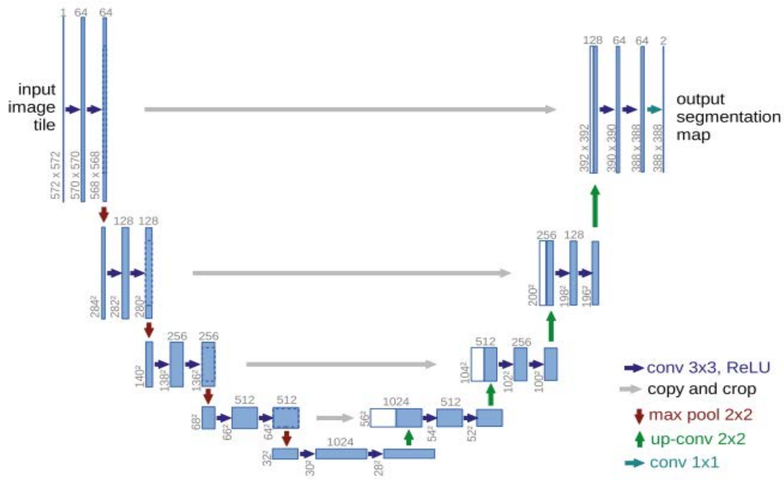

Fig. 1. Architecture of U-Net

In Fig.1. The U-Net architecture comprises of a downsampling/ contracting path (left side) and upsampling/ expansive path (right side). The downsampling consists of two 3x3 convolutional layers by two rectified linear unit (ReLU) activation function and 2x2 max pooling. At pooling, the number of features doubles. The bottleneck of the network is between downsampling and upsampling. It is made up of two convolutional layers. The upsampling consists of 2x2 convolution (up-convolution) that reduces the number of features, a concatenation with the correspondingly cropped feature map from the downsampling path, and it contains two 3x3 convolution layer followed by two ReLU activation function. A 1x1 convolution layer is used in the final layer to map each feature to the required number of classes[29]. This architecture has gained a huge interest in the segmentation of medical images, on the basis of which several variants have been created [23].
Habijan, Marija, et al, 2019, proposed a 3D U-Net architecture of the Convolutional Neural Network (CNN) method for the segmentation of the whole heart. They have used two convolutional neural networks based on 3D U-Net architecture, one architecture is used to locate the bounding box all over the heart and other architecture is used for segmentation. The analysis of their proposed method is carried out on 20 2D CT images. The dataset of 20 2D CT is divided into two sets which consist of 15 training images and 5 validation images. The 3D U-Net architecture comprises two parts: one is a contracting encoder part and the other is the consecutive expanding decoder part. The task of the contracting encoder part is to examine the entire image and the task of the decoder part is to generate segmentation in high resolution. Due to the less number of data set, the data augmentation process is used to maximize the number of data for the training which improves the performance. Dice coefficient is used during training to calculate the performance of the model. They have obtained a mean dice score of 89% for the segmentation of the whole heart [30].
Skourt, Abdelhamid El Hassani, et al. 2018, proposed a deep learning algorithm using U-NET architecture for the segmentation of lung CT images. The dataset they used for the experiment contains thoracic computed tomography (CT) scans of diagnostic and lung cancer with highlighted-up illustrated lesions. Before the experiment was performed, for the ground truth, manual segmentation of lung parenchyma was provided. After manual segmentation, cropping of images is done to eliminate the details which are not for their field of analysis. The performance of a network was evaluated by using the dice coefficient index used for the calculation of similarity. They obtain a dice-coefficient measure of 0.9502 for effective segmentation [31].

3. Methodology

The flowchart for the methodology of this study is shown in Fig. 2. The design of this methodology is categorized into the following phases; research design, data collection, testing, and performance analysis.

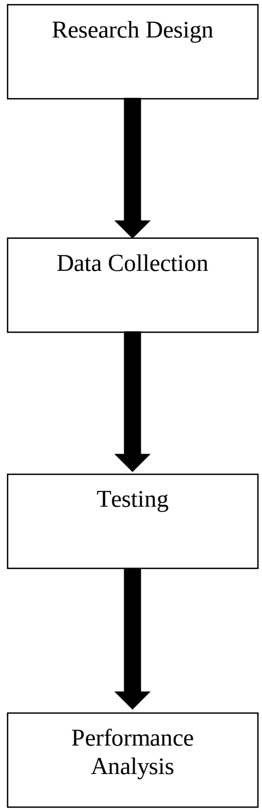

Fig. 2. Methodology Flow chart

3.1 Research Design

The first phase is to review and analyze the earlier studies that are examined by using queries 'image segmentation', 'medical images', 'polyp segmentation', and 'deep learning'. A literature review is essential to recognize the problem statement and identify the techniques that have been applied. Once the problem is identified, then the literature review is studied for the research objectives and questions. This phase covers the basic knowledge of image segmentation and enables the identification of the numerous techniques and algorithms of image segmentation and their advantages and disadvantages that are proposed for the segmentation of medical images using a deep convolutional neural network.

3.2 Data Collection

This phase describes the data gathering process. A Gastrointestinal Polyp image dataset is collected from 'Kvasir-SEG dataset https://datasets.simula.no/kvasir-seg/ '. After the collection of the dataset, ground truth is checked for each and every gastrointestinal polyp images. The dataset consists of 1000 images of gastrointestinal polyps and 1000 corresponding mask which is then used to illustrate the performance of the U-Net, FCN, and Unet-Resnet.

3.3 Testing

This phase includes designing, testing, and analysis of the proposed algorithm. The gastrointestinal polyp dataset is divided into three sets; 80% training set (800 images), 10% validation set (100 images), and 10% testing set (100 images). After dividing the dataset, data augmentation is applied to the training set (800 images) that increases the training dataset up to 13,600 images. We use three architectures of a deep convolutional neural network for the dataset to demonstrate image segmentation. This study will be able to examine the performance of the three algorithms after the testing phase is conducted. An analysis should be implemented after the testing to examine the accuracy and performance of the proposed algorithm. After analysis, a comparison will be conducted with other algorithms to analyze which algorithm performs best for image segmentation.

3.4 Performance Analysis

A performance comparison will be conducted between image segmentation techniques such as U-Net, FCN, Unet_Resnet. The performance of the proposed algorithm will be evaluated in terms of Dice Similarity Coefficient (DSC), and Intersection over Union (IOU) of image segmentation.

4. Results and discussion

.A Gastrointestinal Polyp image dataset is collected from 'Kvasir-SEG dataset'. The dataset contains 1000 images of gastrointestinal polyps and 1000 corresponding mask. The data set is divided into 80% training set (800 images), 10% validation set (100 images), and 10% testing set (100 images). After the division of dataset, data augmentation is applied on training set (800 images) that increases the training dataset up to 13,600 images which will increase the accuracy rate. Then the segmentation algorithm is applied on the dataset. In the paper we have applied U-Net, FCN and Unet-Resnet to perform segmentation on gastrointestinal polyp images.

This paper presents the performances of segmentation algorithm on the collected dataset. The performance is measured by Dice-coefficient, Intersection over Union (IOU), Precision and Recall. Dice-coefficient measures the degree of similarity between the mask and the predicted image. Dice coefficient > 0.7 means that the effect of segmentation is good. The formula for the dice coefficient is as follows

$$\text{Dice-coefficient} = 2 \times \frac{|X \cap Y|}{|X| + |Y|} \qquad (1)$$

Intersection over Union (IOU), also known as Jaccard index, is used to calculate the overlap between the mask and the predicted image, IOU > 0.5 is considered good. The formula for IOU is described as

$$\text{IOU} = \frac{|X \cap Y|}{|X \cup Y|}$$

$$= \frac{|X \cap Y|}{|X| + |Y| - |X \cap Y|} \qquad (2)$$

Precision is used to calculate the ratio between the correct prediction and the total prediction. The formula for precision is below

$$\text{Precision} = \frac{TRUE\ POSITIVE}{TRUE\ POSITIVE + FALSE\ POSITIVE} \qquad (3)$$

True positive is the number of predictions that the model predicted correctly. The denominator in the formula is the total number of predictions.

Recall is the ratio of the correct prediction to the total number of observations by the model. The formula for the recall is

$$\text{Recall} = \frac{TRUE\ POSITIVE}{TRUE\ POSITIVE + FALSE\ NEGATIVE} \qquad (4)$$

True positive in the numerator is the predictions correctly identified by the model and the denominator indicates the actual positive and the number of positives incorrectly predicted as negative by the model.

4.1 Results

The U-Net, FCN, Unet-Resnet models are used for the segmentation of the gastrointestinal polyp image dataset. The performance is measured by dice-coefficient and Intersection over union (IOU) as explained in equation (1) and equation

(2). Both the metrics range from 0-1. All three models are trained for 50 epochs. The experiment is carried out on GTX 1060, 6 GB RAM, and software framework python 3 with Tensor Flow 2.0.

Some examples of our results using U-Net, FCN, Unet-Resnet are shown below in Fig. 3.

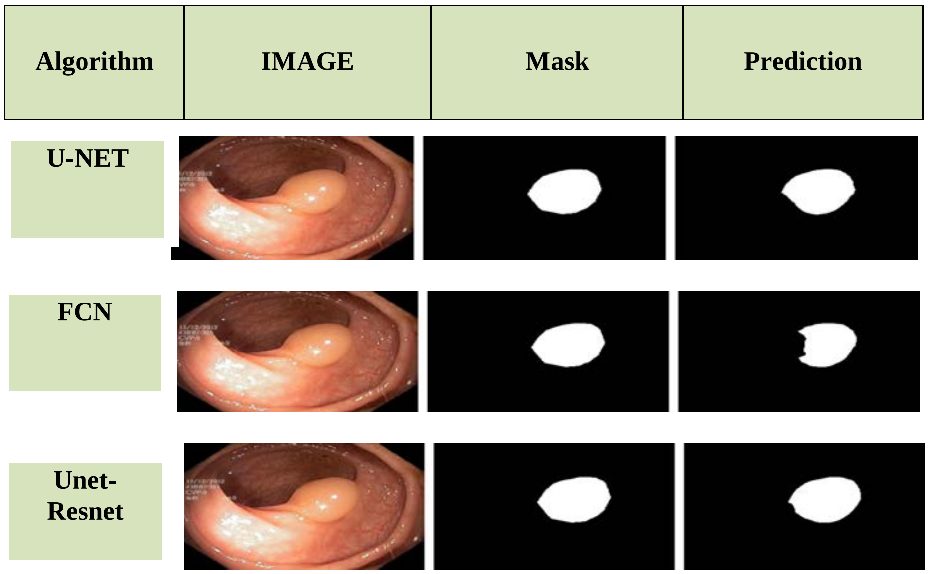

Fig. 3. Results of U-NET, FCN, Unet-Resnet

Fig.3. illustrates the result of three algorithms that shows that the prediction of algorithm Unet_Resnet is more accurate than the U-NET and FCN algorithms. Comparison of the performance for all three algorithms is shown below in Table 1.

Table 1. Comparison Table

Algorithm	Intersection of Union (IOU)	Dice similarity coefficient (DSC)	Training Duration (hours)
FCN	0.71	0.83	1.3
UNET	0.73	0.84	4
Unet_Resnet	0.75	0.86	2.5

Table 1 shows the performance comparison between three algorithms using IOU, DSC, and training duration time. FCN achieves less IOU and DSC than U-NET and Unet_Resnet, but takes less training hours than the other two algorithms. UNET achieves 0.73 IOU and 0.83 DSC which is less than Unet_Resnet, and also UNET takes around 4 hours to train the dataset. While Unet_Resnet takes less training duration that is, 2.5 hours, and Unet_Resnet achieves more IOU and DSC than the other two algorithms.

5. Conclusion

In this paper, we apply segmentation methods based on FCN, UNET, and Unet-Resnet for gastrointestinal polyp images of the Kvasir-SEG dataset. We calculate performance using Intersection of Unit (IOU) and Dice similarity Coefficient (DSC). IOU and DSC for U-NET are 0.73 and 0.84 respectively. Then we compared U-NET results with other algorithms, FCN, and UNet_Resnet. FCN achieved 0.71 IOU and 0.83 DSC and Unet_Resnet achieved 0.75 IOU AND 0.86 DSC. Unet_Resnet achieves higher performance, which outperforms the other two methods in gastrointestinal polyps segmentation

References

[1] M. Billah, S. Waheed, and M. M. Rahman, "An Automatic Gastrointestinal Polyp Detection System in Video Endoscopy Using Fusion of Color Wavelet and Convolutional Neural Network Features," vol. 2017, pp. 1–10, 2017.

[2] M. Akbari *et al.*, "Polyp Segmentation in Colonoscopy Images Using Fully Convolutional Network," 2018.

[3] R. Merjulah and J. Chandra, "Segmentation Technique for Medical Image Processing : A Survey combines the concept of computational and mathematical," no. Icici, pp. 1055–1061, 2017.

[4] M. Vardhana, N. Arunkumar, S. Lasrado, and E. Abdulhay, "ScienceDirect Convolutional neural network for bio-medical image segmentation with hardware acceleration," *Cogn. Syst. Res.*, vol. 50, pp. 10–14, 2018.

[5] P. Muthuselvi, S., & Prabhu, "DIGITAL IMAGE PROCESSING TECHNIQUES – A SURVEY," no. October, 2017.

[6] A. Mehmood, "Enhancement of CT images for visualization," pp. 3–4, 2019.

[7] W. Bo, W. Ying, and C. Lijie, "Fuzzy clustering recognition algorithm of medical image with multi-resolution feature," no. June, pp. 1–12, 2018.

[8] S. Sharma and A. Aggarwal, "Computer Methods in Biomechanics and Biomedical Engineering : Imaging & Visualization Content-based retrieval of biomedical images using orthogonal Fourier-Mellin moments moments," *Comput. Methods Biomech. Biomed. Eng. Imaging Vis.*, vol. 1163, pp. 1–11, 2018.

[9] G. Petterson, "3 . IMAGE ANALYSIS TECHNIQUES Department of Ecology," vol. 2, pp. 23–39, 2001.

[10] Y. M. Y. Abdallah and T. Alqahtani, "Research in Medical Imaging Using Image Processing Techniques," 2019.

[11] W. Khan, "Image Segmentation Techniques : A Survey," vol. 1, no. 4, 2013.

[12] V. Gurusamy and S. Kannan, "Review on image segmentation techniques," no. October 2014, 2015.

[13] M. S. Abirami and T. Sheela, "Analysis of Image Segmentation Techniques for Medical Images," pp. 1–6, 2014.

[14] S. Shirly and K. Ramesh, "Review on 2D and 3D image segmentation," pp. 150–160, 2019.

[15] and U. K. Gupta, Kapil Kumar, Namrata Dhanda, "A Comparative Study of Medical Image Segmentation Techniques for Brain Tumor Detection," *2018 4th Int. Conf. Comput. Commun. Autom.*, pp. 1–4, 2018.

[16] M. V Wu, J., Poehlman, S., Noseworthy, M. D., & Kamath, "Texture Feature based Automated Seeded Region Growing in Abdominal MRI Segmentation," pp. 1187–1191, 2008.

[17] W. Cao, J. Li, J. Liu, and P. Zhang, "Two Improved Segmentation Algorithms for Whole Cardiac CT Sequence Images," *2016 9th Int. Congr. Image Signal Process. Biomed. Eng. Informatics*, no. 32, pp. 346–351, 2016.

[18] S. Ghosh, N. Das, I. Das, and U. Maulik, "Understanding Deep Learning Techniques for Image Segmentation," vol. 52, no. 4, 2019.

[19] M. Mittal, M. Arora, T. Pandey, and L. M. Goyal, *Image Segmentation Using Deep Learning Techniques in Medical Images*. 2020.

[20] A. Khan, A. Sohail, U. Zahoora, and A. S. Qureshi, "A Survey of the Recent Architectures of Deep Convolutional Neural Networks," pp. 1–70, 2020.

[21] J. Lai, H. Zhu, and X. Ling, "Segmentation of Brain MR Images by Using Fully Convolutional Network and Gaussian Mixture Model with Spatial Constraints," vol. 2019, 2019.

[22] L. Bi, D. Feng, and J. Kim, "Dual-Path Adversarial Learning for Fully Convolutional Network (FCN) -Based Medical Image Segmentation," *Vis. Comput.*, vol. 34, no. 6, pp. 1043–1052, 2018.

[23] M. H. Hesamian, W. Jia, X. He, and P. Kennedy, "Deep Learning Techniques for Medical Image Segmentation : Achievements and Challenges," vol. 3, pp. 582–596, 2019.

[24] Z. Alom *et al.*, "The History Began from AlexNet : A Comprehensive Survey on Deep Learning Approaches," 2018.

[25] M. Pashaei and H. Kamangir, "Review and Evaluation of Deep Learning Architectures for Efficient Land Cover Mapping with UAS Hyper-Spatial Imagery : A Case Study Over a Wetland," 2020.

[26] T. Takumi and A. Sumika, *Automatic Multi-class Classification of Tiny and Faint Printing Defects Based on Semantic Segmentation*. 2020.

[27] Z. Zhang, C. Wu, S. Coleman, and D. Kerr, "Highlights based on U-Net architecture is proposed for medical image ☐ The modified Inception-Res module vanishing or redundant computation during network training . tasks and make evaluation with other segmentation methods .," *Comput. Methods Programs Biomed.*, p. 105395, 2020.

[28] Y. Yang, C. Feng, and R. Wang, "AUTOMATED SEGMENTATION OF PULMONARY LOBES USING COORDINATION-GUIDED DEEP NEURAL NETWORKS," *Expert Syst. Appl.*, p. 113419, 2020.

[29] O. Ronneberger, P. Fischer, and T. Brox, "U-Net : Convolutional Networks for Biomedical," pp. 234–241, 2015.

[30] M. Habijan, H. Leventic, I. Galic, and D. Babin, "W hole Heart Segmentation from CT images Using 3D U-Net architecture," pp. 121–126, 2019.

[31] A. El, E. L. Hassani, A. Majda, and B. Ait, "ScienceDirect ScienceDirect ScienceDirect Lung CT CT Image Image Segmentation Segmentation Using Using Deep Deep Neural Neural Networks Networks," *Procedia Comput. Sci.*, vol. 127, pp. 109–113, 2018.