International Journal on Perceptive and Cognitive Computing (IJPCC) Vol 9, Issue 2 (2023)
https://doi.org/10.31436/ijpcc.v9i2.408

Comparison of U-Net's Variants for Segmentation of Polyp Images

Amelia Ritahani Ismail, Syed Qamrun Nisa
Department of Computer Science, Kulliyyah of ICT,
International Islamic University Malaysia

*Corresponding author: amelia@iium.edu.my
(Received: 17rd May 2023; Accepted: 15th June 2023; Published on-line: 28th July 2023

Abstract— Medical image analysis involves examining pictures acquired by medical imaging technologies in order to address clinical issues. The aim is to increase the quality of clinical diagnosis and extract useful information. Automatic segmentation based on deep learning (DL) techniques has gained popularity recently. In contrast to the conventional manual learning method, a neural network can now automatically learn image features. One of the most crucial convolutional neural network (CNN) semantic segmentation frameworks is U-net. It is frequently used for classification, anatomical segmentation, and lesion segmentation in the field of medical image analysis. This network framework's benefit is that it not only effectively processes and objectively evaluates medical images, properly segments the desired feature target, and helps to increase the accuracy of medical image-based diagnosis.

Keywords— segmentation, medical images, deep learning, Convolutional Neural Network.

I. Introduction

With the rising use of computed tomography (CT) and magnetic resonance imaging (MR) for diagnosis, treatment planning, and clinical investigations, using computers to assist radiological specialists in clinical diagnosis and treatment planning has essentially become important. With the creation of CNNs, significant advances in the field of medical image analysis have been made in recent years. One of the most difficult problems in medical imaging is dealing with image segmentation. Many different clinical applications require segmentation of medical images, such as brain segmentation, cardiac ventricle segmentation, abdominal organ segmentation, and cell segmentation in medical imaging. Because of their flexible architectures, the U-Net architecture and its version are one of the methods for segmenting medical pictures using CNNs. Features are learned in the first layer of the U-Net architecture (the analysis path), and segmentation is performed in the second half of the layer of U-Net architecture (the synthesis path). End-to-end training of the layers of the network yields good segmentation results. Skip connections are utilized to allow dense feature mappings from the analysis path to propagate to the corresponding layers in the synthesis section of the network, which improves performance dramatically because the quantity of features in the first part of the network is reduced due to convolutions and pooling.

II. Related Work

In medical imaging, medical image segmentation is one of the crucial tasks as it facilitates the precise delineation as well as the extraction of regions of interest from complex medical images, which need medical experts' validation. The field of medical image segmentation has undergone a gain importance with the emergence of deep learning techniques, particularly Convolutional Neural Networks (CNNs).

CNNs are a class of deep learning architectures that are frequently employed for computer vision tasks including image recognition and segmentation. CNNs are composed of several layers, including convolutional layers, pooling layers, and fully connected layers, which are inspired by the visual processing systems of the human brain. Convolutional layers scan the input image with filters to pick up regional patterns and features. The feature maps are down sampled by the pooling layers that will reduce spatial dimensions while preserving crucial data. The fully linked layers will finally aggregate the retrieved features to produce the predictions. Due to their capacity to automatically identify important features from the data, CNNs have demonstrated exceptional success in a variety of fields, including medical image segmentation [1]. The architecture of CNN is demonstrated in Figure 1.

https://doi.org/10.31436/ijpcc.v9i2.408

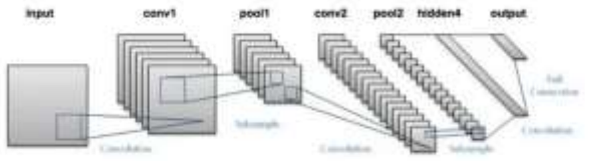

Figure 1. Architecture of CNN

The convolutional neural network architecture known as U-Net was created especially for the segmentation of biological images. One of the most known models for medical image segmentation tasks was put forth by Ronneberger [2]. To gather contextual information and condense spatial dimensions, the U-Net design uses a contracting path that employs a series of convolutional and pooling layers. An expansive path that combines up-convolutions (transpose convolutions) and skip connections to create a fine-grained segmentation map comes after the contracting path. The model may incorporate low-level and high-level data resulting in precise localization of objects and structures in medical pictures. As an example, segmenting organs and tumours in MRI and CT images has shown the U-Net architecture to be very effective in a variety of medical segmentation tasks [2]. Figure 2. Illustrates the architecture of U-Net.

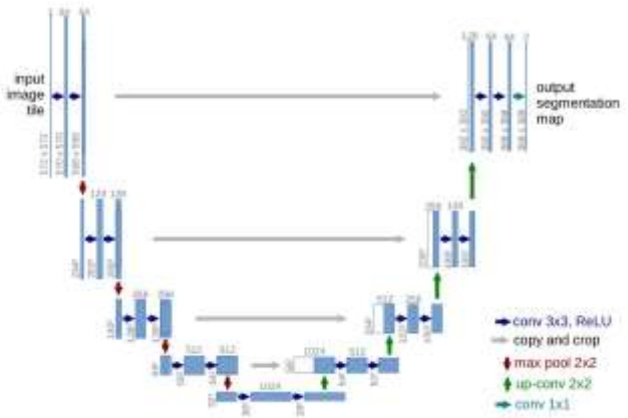

Figure 2. Architecture of U-Net

The U-Net's distinctive architecture, which features an expanding and contracting path, makes it easier to localise details while still effectively gathering contextual data. To get cutting-edge findings, researchers have successfully deployed U-Net-based CNNs to a variety of medical imaging modalities, including MRI, CT, and ultrasound [3] [4]. Numerous experiments have shown that its distinctive design, which incorporates a contracting path to gather contextual information and an expansive path for precise localization has performs better than typical CNNs [3],[4].

To improve the results of medical picture segmentation, researchers have combined CNNs with the U-Net architecture. Combining CNN's feature extraction capabilities with U-Net's accurate localization results in increased segmentation accuracy by combining the best aspects of both methods [5],[6]. Deep learning techniques' use in multi-modal medical image segmentation has also attracted a lot of attention. Models can better capture complimentary features by combining data from many imaging modalities, producing segmentation results that are more reliable and precise [7],[8].

Despite the effectiveness of deep learning techniques, medical image segmentation frequently faces issues if there is an insufficient number of labelled data. To efficiently use the data that is currently accessible and to draw knowledge from different fields, researchers have proposed transfer learning approaches and data augmentation methodologies [5],[9]. Another issue includes the necessity for interpretable models and the generalisation of models across various datasets. To improve medical image segmentation for diverse clinical applications, additional research is required to solve these problems and investigate novel architectures [9], [10] for better segmentation results.

The Multiresolution U-Net (Multires U-Net), among other architecture has become a well-known deep learning model, revolutionising the field of medical image segmentation. Incorporating multiple resolution paths to efficiently collect both local and global features, the Multires U-Net expands upon the original U-Net architecture. The model effectively learns hierarchical representations by fusing contracting and expanding routes, making it easier to precisely segment complex structures. The Multires U-Net is very useful in medical imaging applications, such as tumour segmentation, organ delineation, and lesion detection, because to its robustness to complex anatomical changes and diseases [11]. The Multires U-Net is still being investigated and improved by scientists and medical professionals, leading to significant developments in the field of medical picture segmentation [12]. The Multires U-Net can give cutting-edge results in a variety of applications, including cell detection, cardiac analysis, and brain segmentation, thanks to its innovative design, which enables it to adapt successfully to a wide range of medical imaging data. It is unquestionably advancing the field of medical picture segmentation as the community of medical imaging continues to tap into its potential [11],[12].

Another deep learning architecture called the Double U-Net has demonstrated incredible promise in the segmentation of medical images. Double U-Net [11], takes

https://doi.org/10.31436/ijpcc.v9i2.408

the problems presented by the scarcity of annotated medical data and makes use of dual paths to increase segmentation accuracy. It is made up of two U-Net paths that cooperate to improve segmentation. The input image is processed by the first pathway, which creates preliminary segmentation predictions. The second pathway then refines these predictions using the data from the first pathway to produce more exact segmentation boundaries. The Double U-Net overcomes data scarcity and emerges as a potent tool for many medical imaging tasks, showing substantial potential for tumour diagnosis, organ segmentation, and anomaly recognition in medical pictures. This is accomplished by effectively utilising the knowledge from both paths [13].

Due to its capacity to manage complicated anatomical features and solve data constraints, the Double U-Net has become a well-known deep learning approach in the field of medical image segmentation. The Double U-Net architecture was used by [11] to segment the pancreas in abdominal CT scans, demonstrating the architecture's potential. Dual paths are used to allow the model to learn both local and global information, producing segmentation masks that are more precise and refined. Additionally, the architecture can utilise the limited annotated data, which is a common difficulty in medical imaging, to its most potential thanks to the collaborative nature of the two paths. Thus so far it has demonstrated tremendous promise in a variety of medical picture segmentation tasks, demonstrating its versatility and potential for further applications in the medical industry [13].

Another segmentation approach for medical image segmentation, uses a dual U-Net with Resnet encoder [14]. Two U-Net architectures are combined, in which one U-Net uses pre-trained resnet as an encoder. And another U-Net in the architecture is built from the scratch. The segmentation results shows that dual U-Net performs better than other existing segmentation architectures.

III. Methodology

The segmentation dataset is derived from the Kvasir-SEG dataset, which may be found at https://datasets.simula.no/kvasir-seg/. Images of gastrointestinal polyps are included in the Kvasir-SEG collection. Following the acquisition of the dataset, gastrointestinal polyp images are examined for ground truth. The gastrointestinal polyp collection includes 1000 pictures and masks. The dataset is then divided into three sets: training, validation, and testing. Data augmentation is done to the training set to increase the data, which helps to improve segmentation performance. MultiresUnet and DoubleUnet are the methods used in this paper to segment gastrointestinal polyp pictures. The architectures of MultiresUnet and DoubleUnet are depicted in Figure 3 and Figure 4, respectively.

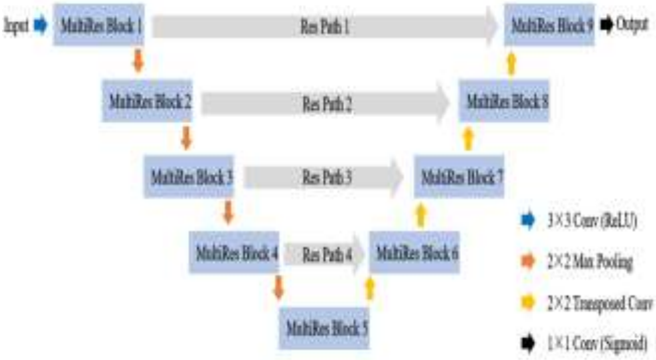

Figure 3. MultiresUnet [15]

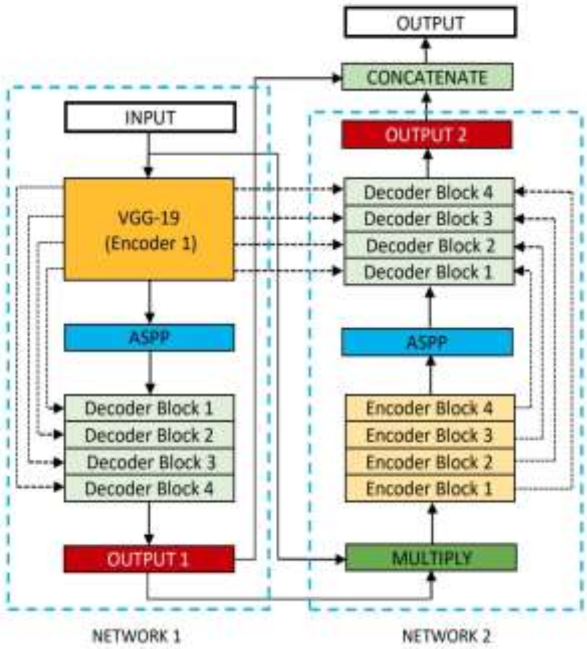

Figure 4. DoubleUnet[16]

IV. Results and Discussion

From the "Kvasir-Seg dataset," a dataset consisting of polyps found in the gastrointestinal tract is compiled. It is made up of a thousand photos of polyps found in the digestive tract and the masks that correlate to them. The dataset is divided into three categories: training data (800 images; 80 percent), validation data (100 images; 10 percent), and testing data (100 images; 10 percent). After

https://doi.org/10.31436/ijpcc.v9i2.408

the dataset has been split, data augmentation is applied to the remaining 800 photos of the training set. This brings the total number of images in the training dataset up to 13,600, which in turn increases the accuracy rate. Following that, the dataset containing gastrointestinal polyps is subjected to the segmentation process. In the research, we performed segmentation on the Gastrointestinal polyp dataset using MultiresUnet and DoubleUnet. These two algorithms were used. Intersection over Union, also known as IOU, and Dice Similarity Coefficient, or DSC, are two metrics that are utilized to assess the effectiveness of a segmentation method.

Intersection over Union (IOU) is utilized to measure the overlap between the predicted image and mask. Formula for IOU is shown below:

$$IOU = A \cap B / A \cup B$$

Dice Similarity Coefficient (DSC) determines the similarity between the predicted image and the mask. Formula for DSC shown below:

DSC = 2 * (number of common elements) / (number of elements in set A + number of elements in set B)

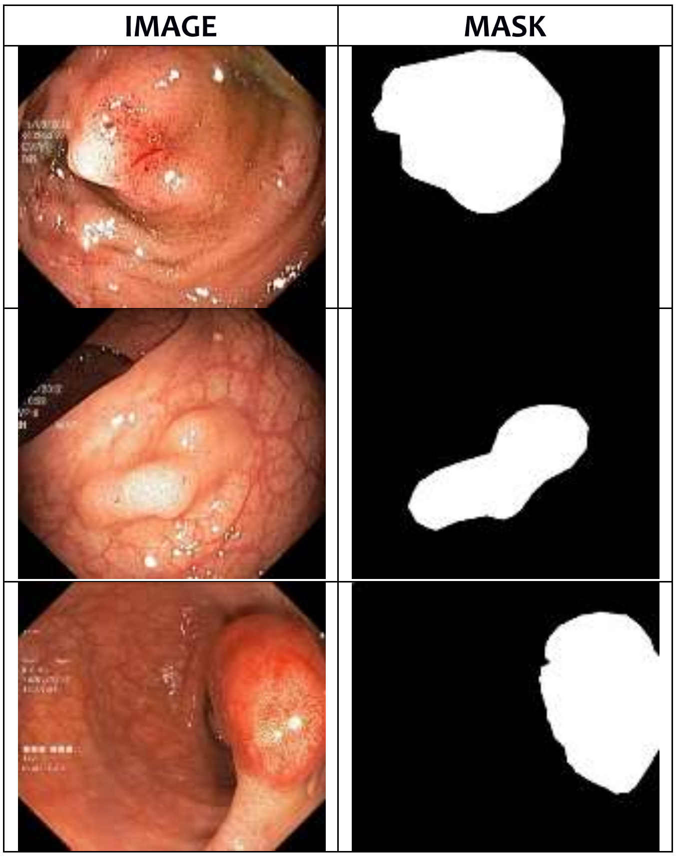

Figure 5. Sample Dataset

Figure 5 shows the sample of gastrointestinal polyp dataset. The left column shows the original image of gastrointestinal polyps and right column shows the mask of the original image.

A. Results

MultiresUnet and DoubleUnet algorithms are implemented to perform segmentation on the dataset of gastrointestinal polyps. The evaluation metrics for the calculation of performance ranges from 0 to 1. The algorithms are trained for 50 epochs. Figure 6 demonstrates the segmentation results using MultiresUnet and DoubleUnet algorithms.

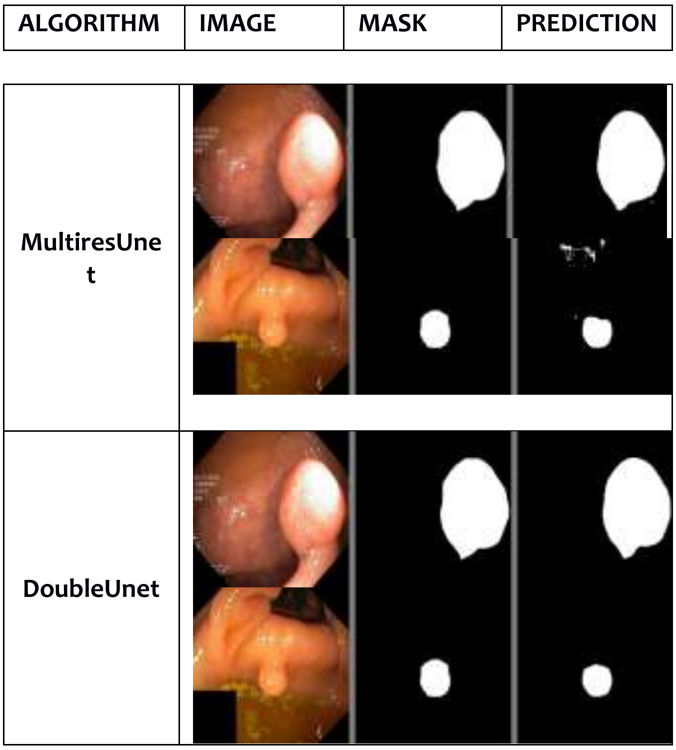

Figure 6. Results of segmentation using MultiresUnet and DoubleUnet

Figure 6 demonstrates the segmentation results on gastrointestinal polyp images by employing MultiresUnet and DoubleUnet algorithms. It indicates that the DoubleUnet algorithm predicts better than the MultiresUnet algorithm. The performance of both the algorithms is compared in Table I shown below.

TABLE I. PERFORMANCE COMPARISON

Algorithm	Dice Similarity Coefficient (DSC)	Intersection over Union (IOU)
MultiresUnet	0.83540	0.76553
DoubleUnet	0.86137	0.79241

Table I illustrates the comparison of the performance between the MultiresUnet and DoubleUnet algorithms using Dice Similarity Coefficient and Intersection over Union metrics. MultiresUnet achieves DSC of 0.83540 and IOU of 0.76553. DoubleUnet achieves DSC of 0.86137 and IOU of 0.79251. It indicates that the DoubleUnet performs segmentation much better than the MultiresUnet.

V. CONCLUSIONS

In this study, the segmentation of the gastrointestinal polyp dataset obtained from the Kvasir-Seg dataset is carried out with the assistance of two different algorithms: MultiresUnet and DoubleUnet. The Dice Similarity Coefficient (DSC) and the Intersection over Union (IOU) are two metrics that are used to evaluate the effectiveness of various segmentation techniques. According to the performance result, the DoubleUnet is capable of achieving higher levels of DSC and IOU than the MultiresUnet. In the future, one of our primary focuses will be to work toward optimizing the characteristics of design in order to make it more efficient.

ACKNOWLEDGMENT

The authors would like to thank the Kulliyyah of ICT IIUM

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

REFERENCES

[1] Y. LeCun, Y. Bengio, & G. Hinton, "Deep learning". Nature, 521(7553), 436-444, 1998.

[2] O. Ronneberger, P. Fischer, & T. Brox, "U-Net: Convolutional networks for biomedical image segmentation". In International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI) (pp. 234-241). Springer, 2015.

[3] H. Jiang, C. Ma, Y. Zhang, & H. Xie, "Automatic liver tumor segmentation using a U-Net based deep learning framework". Journal of Computer Assisted Tomography, 42(5), 841-848, 2018.

[4] B. Zhou, C. Zhao, Y. Huang, & Y. Wang, "U-Net++: A nested U-Net architecture for medical image segmentation". IEEE Transactions on Medical Imaging, 39(5), 1856-1867, 2020.

[5] J. Nalepa, M. Marcinkiewicz, & M. Kawulok, "Data augmentation for brain-tumor segmentation: a review," Frontiers in computational neuroscience, 13, 83, 2019.

[6] J. Walsh, A. Othmani, M. Jain, & S. Dev, "Using U-Net network for efficient brain tumor segmentation in MRI images". Healthcare Analytics, 2, 100098, 2022.

[7] E. Smistad, T. L. Falch, M. Bozorgi, A. C. Elster, & F. Lindseth, "Medical image segmentation on GPUs–A comprehensive review". Medical image analysis, 20(1), 1-18, 2015.

[8] X. Jia, Y. Liu, Z. Yang, & D. Yang, "Multi-modality self-attention aware deep network for 3D biomedical segmentation". BMC Medical Informatics and Decision Making, 20(3), 1-7, 2020.

[9] S. Lazendić, J. Janssens, S. Huang, & A. Pižurica, "On Interpretability of CNNs for Multimodal Medical Image Segmentation". In 2022 30th European Signal Processing Conference (EUSIPCO) (pp. 1417-1421). IEEE, 2022.

[10] K. D. Apostolidis, & G. A. Papakostas, "A survey on adversarial deep learning robustness in medical image analysis". Electronics, 10(17), 2132, 2021.

[11] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, & J. Liang, "UNet++: A Nested U-Net Architecture for Medical Image Segmentation," In Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support (pp. 3-11). Springer, Cham, 2018.

[12] F. Isensee, J. Petersen, A. Klein, D. Zimmerer, P. F. Jaeger, S. A. Kohl, & K. H. Maier-Hein, "nnU-Net: A Self-Configuring Method for Deep Learning-Based Biomedical Image Segmentation," Nature Methods, 18(2), 203-211, 2021.

[13] E. K. Fishman, A. L. Yuille, L. Xie, W. Shen, Y. Wang, "An abdominal CT scan pancreas segmentation fixed-point model." (Pages 8561–8571) in Advances in Neural Information Processing Systems, , 2018.

[14] S. Q. Nisa, & A. R. Ismail, "Dual U-Net with Resnet Encoder for Segmentation of Medical Images". International Journal of Advanced Computer Science and Applications, 13(12), 2022.

[15] N. Ibtehaz, & M. S. Rahman, "MultiResUNet: Rethinking the U-Net architecture for multimodal biomedical image segmentation". Neural networks, 121, 74-87, 2022.

[16] D. Jha, M. A. Riegler, D. Johansen, P. Halvorsen, & H. D. Johansen, "Doubleu-net: A deep convolutional neural network for medical image segmentation". In 2020 IEEE 33rd International symposium on computer-based medical systems (CBMS) (pp. 558-564). IEEE, 2020.