Assessing the Performance of U-Net in Three Dimensional Medical Image Segmentation

Fatma Khenaifer, Ilyes Benaissa, Athmane Zitouni and Zine-Eddine Baarir

December 8, 2023

Assessing the Performance ofU-Net in 3D Medical Image Segmentation

1st Khenaifer Fatma
VSC Lab. Dept. of Electrical Engineering
*University of Mohamed Khider*Biskra, Algeria
fatma.khenaifer@univ-biskra.dz

2nd Ilyes Benaissa
VSC Lab. Dept. of Electrical Engineering
*University of Mohamed Khider*Biskra, Algeria
ilyes.benaissa@univ-biskra.dz

3rd Athmane Zitouni
VSC Lab. Dept. of Electrical Engineering
University of Mohamed Khider Biskra, Algeria
athmane.zitouni@univ-biskra.dz

4th Baarir Zinne-eddine
VSC Lab. Dept. of Electrical *Engineering*
University of Mohamed Khider Biskra, Algeria
zineeddine.baarir@univ-biskra.dz

.

Abstract—**3D medical data plays a vital role in the field of healthcare, particularly in disease diagnosis and surgical planning, as it allows for precise identification, segmentation, and visualization of lesions and organs. However, organ segmentation in medical images is a challenging task due to the diverse features of organs, such as variations in size, shape, and location. To address this challenge, deep learning techniques have been widely utilized for medical image segmentation. One of the most prominent and successful algorithms used for this purpose is U-Net, which was proposed in 2015 demonstrated significant achievements in various medical segmentation tasks and has gained popularity among researchers and scientists. Its effectiveness and versatility have led to its continuous development and improvement in recent years. With over 70,000 citations, U-Net has become a widely adopted algorithm for medical image segmentation. To evaluate the performance of U-Net and ensure more accurate and detailed segmentation results; it is commonly trained on 3D datasets containing deformations of organs, such as the spleen which is obtained from the Medical Segmentation Decathlon. The evaluation of the algorithm's capabilities often relies on metrics like the Accuracy and f1-score, that is a widely accepted measure for assessing the quality of medical image segmentation.**

Keywords—Deep Learning, Three-Dimensional Medical Image, Semantic Segmentation, U-Net .

I. Introduction

Image segmentation in medical imaging refers to the process of dividing images into regions or parts corresponding to specific organs or tissues; it is a crucial task in various medical applications. Medical images, particularly three-dimensional ones, are complex and information-rich, making their segmentation challenging. Thanks to recent advancements in artificial intelligence, precisely deep learning, which has shown effective results compared to traditional learning methods. There has been a need for robust techniques to reduce time and guide professionals in various healthcare applications.[1]

Image segmentation techniques are rapidly evolving and becoming more accurate. These algorithms can be applied to any type of image. Widely used medical imaging modalities include computed tomography (CT), magnetic resonance imaging (MRI), and positron emission tomography (PET)[2]. These techniques provide visualization of tissues and internal organs. Semantic segmentation, as a preliminary step in image processing, enhances the quality of visualization through modeling, providing physicians and surgeons with a general understanding of specific regions or different body parts.[3]

Despite the effectiveness of deep learning-based medical image segmentation, the method has several limitations, including limited data and low resolution, leading to inaccurate segmentation that does not meet clinical requirements [4]. Consequently, several approaches have been developed to improve the performance of medical image segmentation. The U-Net algorithm is the most widely used one [2]; its popularity has extended to the segmentation of three-dimensional medical data. This model is evaluated based on the segmentation of three-dimensional spleen images using the Accuracy and F1-score, which is one of the most valuable metrics in medical image analysis.

II. Overview of U-Net Architecture

The U-Net algorithm is one of the common techniques in medical image segmentation that relies on convolutional layers. The current U-Net architecture goes through several changes over the past years. The narrative begins with the EM segmentation challenge in 2012. Cireseam et al, were from the first researchers who focused on medical segmentation methods using convolutional networks[1]. This eventually led to the emergence of the U-Net algorithm in 2015[5], introduced by Rennebeger et al. This method is proved to be a superior solution compared to traditional CNN-based approaches, and achieved promising results in the medical field. U-Net segmentation is useful in medical image analysis where it is used to identify or evaluate structures within medical images. It can be applied to all types of biomedical images. The model's architecture can be divided into two parts: the encoding path and the decoding path.

The encoding path aims to capture the contextual information of the input data and consists of convolutional neural network (CNN) blocks. Each block consists of two consecutive 3x3 convolutions followed by a ReLU activation function. Max pooling operations are applied to reduce the spatial dimensions of the image. This process is repeated multiple times in each block, gradually leading to a bottleneck layer. This layer also includes 3x3 convolutions followed by activation functions[1]. The consecutive convolution operations help capture fine-grained features. The bottleneck layer contains all the important information in the image.[1][6]

In the decoding path, the model performs upsampling to restore the feature maps to their original resolution. It consists of a 2x2 convolution layer (up-sampling the feature maps) followed by two consecutive 3x3 convolutions with activation functions. To mitigate the loss of important information, the algorithm uses skip connections between the encoding and decoding units. After the upsampling process, an additional 3x3 convolution is applied to refine the results[6][1][5]. The skip connections help to bypass reconstruction errors.

Overall, the illustrated U-net architecture in figure 1 enables the accurate segmentation of structures within medical images by leveraging both local and global contextual information.

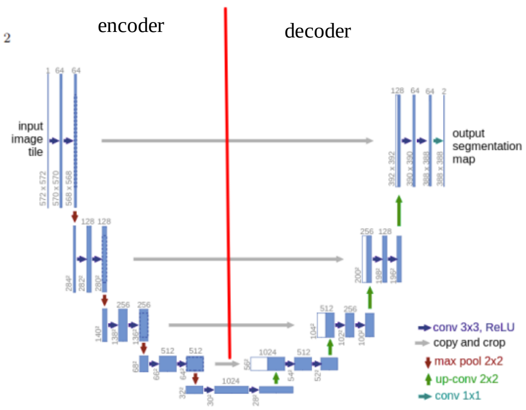

Fig. 1. U-Net Architecture [1]

III. Dataset for Segmenting Medical Images

The Medical Segmentation Decathlon (MSD) is a compilation of datasets for medical image segmentation. It includes 2,633 three-dimensional images in total, gathered from various sources, modalities, and anatomies of interest [11]. The available dataset is large in size, so we focus on partitioning the spleen among 10 different body organs.

The spleen is an important organ in the immune system and is responsible for several functions, such as blood filtration, fighting against microbes, and producing lymphocytes. Therefore, the process of delineating the boundaries of the spleen in medical imaging is crucial for medical diagnosis. It allows doctors to detect abnormal conditions such as cysts, and can help in protecting the spleen from damage during radiation therapy for cancer patients. A dataset of 61 three-dimensional volumes obtained using computed tomography (CT) imaging is used. The spleen data set comprised of patients undergoing chemotherapy treatment, for liver metastases at Memorial Sloan Kettering Cancer Center [7][8].The dataset aims to enhance research in the medical field and improve diagnosis.

IV. Evaluation Techniques

The performance evaluation of the segmentation model on medical images is necessary to ensure accurate assessment or compare different techniques for segmentation processes. Several measures have been derived from the confusion matrix, where true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) represent the values. The most important and commonly used measures for evaluating model performance are [1][6]:

Accuracy: It is a simple measure that represents the proportion of cases that have been correctly classified. It can be calculated using the following formula:

$$Accuracy = \frac{(\mathrm{TP} + TN)}{(TP + TN + FP + FN)} \qquad (1)$$

F1 Score: It is a more complex measure that combines precision and recall, which represent the algorithm's ability to accurately segment structures or shapes in an image. F1 Score is calculated using the following formula:

$$\text{F1 Score} = \frac{2 * (Precision * Recall)}{(Precision + Recall)} \qquad (2)$$

Precision and recall ,this is calculated using the formula:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \qquad (3)$$

$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \qquad (4)$$

In summary, F1 Score provides a balanced assessment of both precision and recall, capturing the algorithm's ability to accurately classify positive cases while minimizing false positives and false negatives.

V. Implementation Specifics

Due to the settings of medical imaging and the heterogeneous nature of medical image components, the spatial resolution of dimensions is not always the same. Additionally, the volumetric image segmentation process requires high computational resources compared to 2D images [9]. Therefore, the approach of separating the volumetric image with dimensions (x, y, z) into a series of 2D slices is adopted, with division along the x-axis. The resulting 28162 images that have (y, z) dimensions taking x-axis as volumetric axis, and the figure 2 shows an illustration of the process followed.

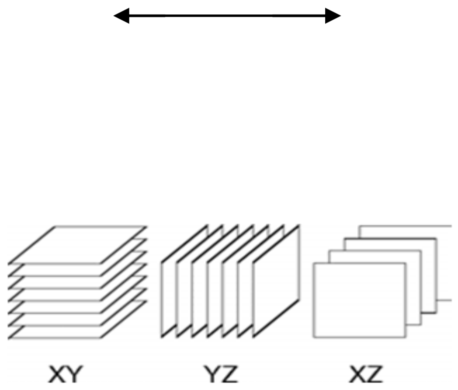

Fig. 2. Dividing three-dimensional data into 2D slices

After the separation and obtaining the 2D slice sequences, the data is trained using the U-Net algorithm. Upon completion of the training process, the volumetric image is reconstructed using a variety of techniques. Data segmentation provides an organized representation of distinct regions and identifies targeted areas, enabling visualization techniques to transform segmented data into three dimensional visual representations.This helps medical experts in addressing specific medical problems and understand medical data [12]. Both CT and MRI scans produce a three-dimensional volume with density data in it. Slices are a common way todisplay the output.

VI. THE RESULTS

3D spleen images are used to evaluate the performance of the traditional 2D U-Net model. We were able to obtain a Dice Score of 0.840 with 89% accuracy through 80 epochs; we used stop-loss function for the binary cross entropy loss to prevent model's overfit. To have a better analysis of the results, the values for each of F1-Score (Dice Score), Accuracy, and Cross Entropy Loss are given in Fig.1, Fig. 2, and Fig. 3 respectively.

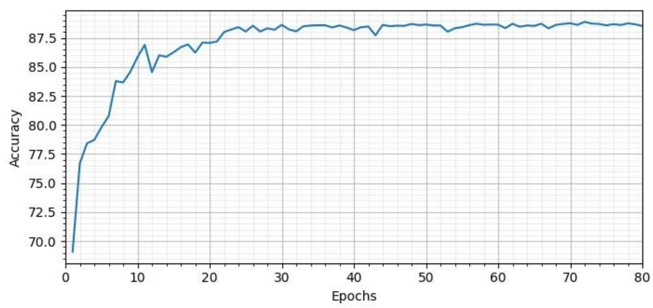

Fig. 3. Accuracy values of training 3D spleen images with 2D based U-Net

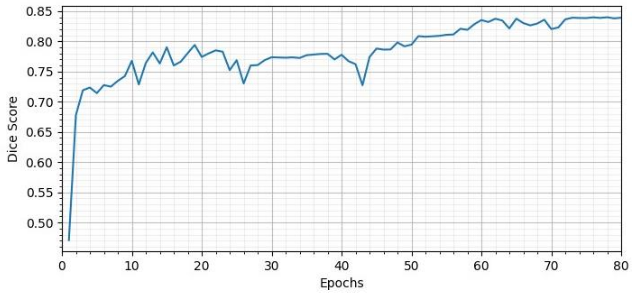

Fig. 4. F1-Score values of training 3D spleen images with 2D based U-Net

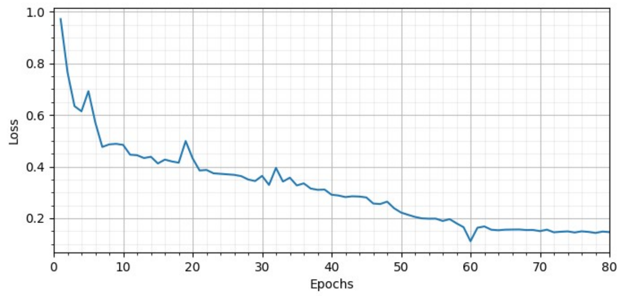

Fig. 5. Cross Entropy Loss values of training 3D spleen images with 2D based U-Net

Some random slices of the 3D spleen images were extracted from the original dataset for the test. After the training, the best saved model is used on the extracted original 3D images; however, it is important to mention that each volume image used for the validation has to be divided into slices for full prediction, in order to pass it through the model because it is used only for 2D images. To make sure that the slices are successfully loaded to the RAM we plotted the Coronal, Sagittal, and Axial planes of one of the slices of the extracted image alongside their mask as shown in Fig. 6 and Fig. 7. Respectively.

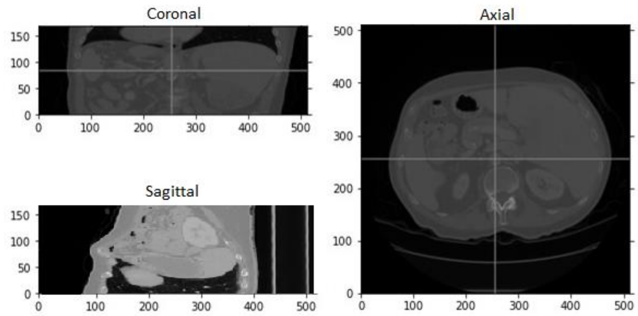

Fig. 6. Different angles for the MSD spleen image used for prediction

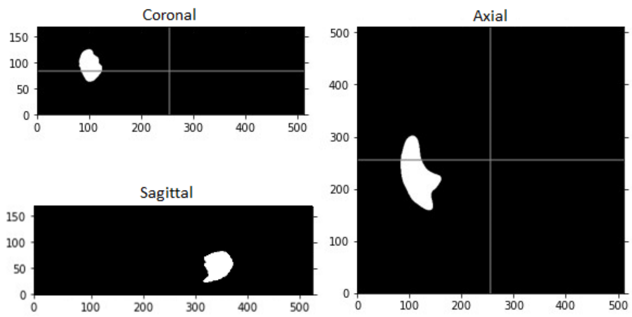

Fig. 7. Different angles for the MSD spleen mask used to test model's prediction

We can notice the overall performance of U-Net in the different planes of the given slice in Fig. 8. Especially when it comes to edge features detection.

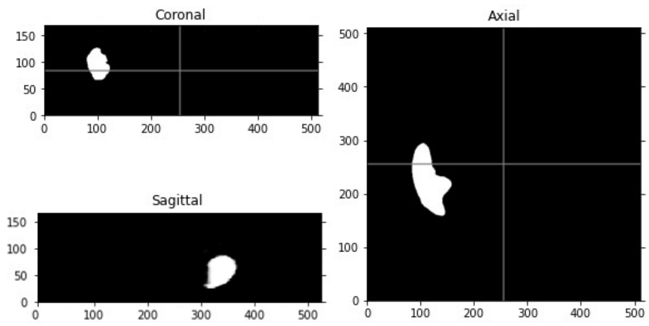

Fig. 8. Predicted spleen mask on different angles for the MSD spleen

After predicting all planes of the volumetric image with a 2D model, we rebuilt those masks of all different planes into one 3D predicted segmentation label. Then we visualized this volume using 3D Slicer software. The final volumetric image is presented in Fig. 9.

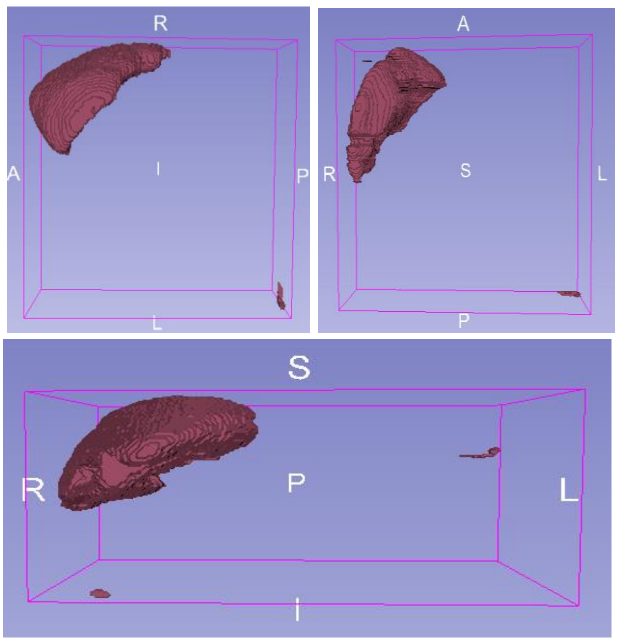

Fig. 8. Different angles of the reconstructed volume based on predicted slices of U-Net model

VII. CONCLUSION

Data segmentation is a technique that allows for the organized representation and identification of specific regions of interest. In the context of medical data, such as spleen images obtained from computed tomography (CT), segmentation plays a crucial role in analyzing and evaluating the organ's characteristics. While the U-Net model has proven its effectiveness by evaluating its performance using 2D slice-based methods, it may not meet the desired level of performance when dealing with 3D medical data volumes. segmented data is converted into 3D visual representations using advanced visualization techniques, by incorporating three orthogonal planes—axial, sagittal, and coronal—can gain a more comprehensive image of the spleen's size, shape, and position. This enables them to better assess and identify abnormalities or deformities present in the organ. By harnessing the power of 3D visualization, medical professionals can enhance their ability to interpret complex structures within the spleen. This in turn, assists in accurate diagnoses, treatment planning, and monitoring of patients' conditions. The visual representation of 3D segmented data facilitates, and improves communication and collaboration among medical teams. Ultimately, leading to better patient care and outcomes.

REFERENCES

[1] Azad, R., Aghdam, E. K., Rauland, A., Jia, Y., Avval, A. H., Bozorgpour, A., ... & Merhof, D. (2022). Medical image segmentation review: The success of u-net. arXiv preprint arXiv:2211.14830.

[2] Engineering, J. O. H. (2023). Retracted: U-Net-Based Medical Image Segmentation. Journal of healthcare engineering, 2023, 9890389.

[3] Liu, X., Song, L., Liu, S., & Zhang, Y. (2021). A review of deep-learning-based medical image segmentation methods. Sustainability, 13(3), 1224.

[4] Esteva, A., Chou, K., Yeung, S., Naik, N., Madani, A., Mottaghi, A., ... & Socher, R. (2021). Deep learning-enabled medical computer vision. NPJ digital medicine, 4(1), 5.

[5] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18 (pp. 234-241). Springer International Publishing.

[6] Siddique, N., Paheding, S., Elkin, C. P., & Devabhaktuni, V. (2021). U-net and its variants for medical image segmentation: A review of theory and applications. Ieee Access, 9, 82031-82057.

[7] Antonelli, M., Reinke, A., Bakas, S., Farahani, K., Kopp-Schneider, A., Landman, B. A., ... & Cardoso, M. J. (2022). The medical segmentation decathlon. Nature communications, 13(1), 4128.

[8] http://medicaldecathlon.com/

[9] Yu, Q., Xia, Y., Xie, L., Fishman, E. K., & Yuille, A. L. (2019). Thickened 2D networks for efficient 3D medical image segmentation. arXiv preprint arXiv:1904.01150.

[10] A. Schuchter, G. Petrarolo and D. Grießner, "Real-Time Voxel-based

[11] A. Schuchter, G. Petrarolo and D. Grießner, "Real-Time Voxel-based

[12] 2D to 3D CT Visualisation Framework for Volumetric CapacityEstimation," 2022 E-Health and Bioengineering Conference (EHB),Iasi, Romania, 2022, pp. 01-04, doi:10.1109/EHB55594.2022.9991516.