

# BRAIN TUMOR SEGMENTATION BASED ON U-NET WITH IMAGE DRIVEN LEVEL SET LOSS

Pham Van Truong<sup>\*</sup>, Tran Thi Thao

Hanoi University of Science and Technology, No.1 Dai Co Viet, Hai Ba Trung, Ha Noi

<sup>\*</sup>Email: [truong.phamvan@hust.edu.vn](mailto:truong.phamvan@hust.edu.vn)

Received: 20 December 2020; Accepted for publication: 9 September 2021

**Abstract.** Brain tumor segmentation plays a vital role in treatment planning and clinical assessment. However, manual segmentation of the magnetic resonance imaging (MRI) is time consuming and labor intensive. So, an automatic approach for brain tumor segmentation is of high demand and attracts more attention from research community. This paper presents a deep learning-based approach for fully automatic segmentation and detection of brain tumors from MRI images. The approach aims at utilizing the U-Net as an architecture of the approach to capture fine and soars information from input images. Especially for training the neural network, we propose to employ a new loss function that combines Level set loss with Dice loss functions instead of using commonly used cross-entropy loss or dice loss functions. The level set loss is inspired from the Mumford-Shah functional for unsupervised task. Meanwhile, the Dice loss function measures the similarity between the predicted mask and desired mask. The proposed approach is applied to segment brain tumors from MRI images and compared with other works when applying to a database of approximately 4,000 image slices acquired from brain MRI scans. Experimental results show that the proposed approach achieves high performance in terms of Dice coefficient and Intersection over Union scores.

**Keywords:** Brain Tumor Segmentation, U-net, Mumford-Shah loss, Level set method, Activate Contour Model.

**Classification numbers:** 4.2.3, 4.7.4.

## 1. INTRODUCTION

It has been reported in many studies that brain tumor remains one of the leading causes of death for cancer patients [1]. According to a report from the American Cancer Society, in 2020, 23,890 new brain cancer cases were identified [2]. However, the treatment of a brain tumor depends on several factors such as the patient's age, the tumor type, and its location. Brain magnetic resonance imaging (MRI), which clearly capture the soft tissues of the brain, is a commonly used to take brain tumor images. Since tumors can grow and spread to nearby healthy tissue, it leads to challenges for diagnosis as well as for treatment. Therefore, accurate detection and segmentation of brain tumor regions from MRI images play an important role in the treatment of brain tumors, especially in the early stages.

Manual segmentation of brain tumors is a crucial procedure and can be considered the best method. However, it is tedious, time consuming and depends on the experience and skills of the physicians. In addition, manual segmentation results also have inter-rater variance [3]. Therefore, automatic brain tumor segmentation from MRI is in high demand to help doctors better diagnose and treat. A vast number of automatic methods for brain tumor segmentation have been proposed in the literature, such as active contour [4, 5], machine learning-based methods [6-9], etc. Among them, machine learning-based methods are considered as a promising approach since they use features extracted from MRI images, which can learn from images during the training stage. In recent years, thanks to the great successes of deep learning, convolutional neural network (CNN)-based approaches have shown outstanding performance in image segmentation [10]. In the field of medical image segmentation, the work proposed by Long *et al.* namely the Fully Convolutional Network (FCN) [11] - a CNN based architecture, has attracted a lot of research [12]. In particular, the U-Net proposed by Ronneberger *et al.* [13] has become the most well-known structure for medical image segmentation [7, 14]. To improve the performance of segmentation using the U-Net, in addition to improving the architecture, performing different loss functions for U-Net training can be considered as a potential approach. In the case of the U-Net, some loss functions have been used, such as Dice loss (Dice), cross-entropy (CE) [13], Tversky loss [15]. However, such loss functions lack of boundary constraints, leading to unexpected results near object boundaries [16].

In the present study, inspired by the performance of the U-Net [17] architecture for medical images, we propose an approach to segment brain tumors from MRI scans. In more detail, the proposed approach is an end-to-end approach, in which we explore the architecture of the U-Net in order to inherit its advantages for segmenting brain tumors. To train the network, we propose a new loss function combining Level set loss [16] with Dice loss. The proposed loss function offers some advantages over other commonly used dice loss and cross-entropy loss [16] when handling the pixel-wise fittings between the regions inside and outside the prediction maps with those in the ground truths. In addition, the proposed loss function included boundary constraints, which can help the proposed approach to obtain satisfactory results near the desired boundary, especially in the case of tumors with weak boundaries and inhomogeneity [13].

The segmentation process can be summarized as follows: First, for the training stage, the set of image pairs including the images and their corresponding labels (ground truths) - is fed into the U-Net neural network. The parameters of the network are updated so that the proposed loss is minimized. The loss represents the difference between the ground truth and the corresponding network output at each epoch. The parameter updating is iterated until the optimization process converged. Then, the model weight where the optimization converged is saved for evaluation and inference. For the testing phase, the test images are fed into the network in a feedforward manner, and the model weight saved from the training phase is used to predict the output map. The network output map is called the prediction or segmentation map.

The rest of the paper is structured as follows: in Section 2, the architecture of the U-Net is presented. In Section 3, the proposed approach is described in detail. In Section 4, some experimental results are presented, including a comparison with other approaches. Conclusions along with discussion of future work are given in Section 5.

## **2. BACKGROUND**

### **2.1. U-Net Architecture**

The U-Net neural network architecture was proposed by Ronneberger *et al.* [13] for image segmentation task. Although originally proposed for segmenting neuronal structures in electron microscopic stacks, the U-Net architecture becomes the typical model and the baseline for various FCN-based networks for image segmentation, especially for medical images. The general structure of the U-Net is shown in Fig. 1, where it can be seen that the U-Net consists of two main parts, encoder and decoder. The encoder used to extract image features and capture image contexts consists of convolutional layers, activations, and max-pooling layers. The decoder, responsible for enabling precise localization, includes up-sampling layers and transposed convolutional layers. To avoid losing spatial information due to pooling operations, Ronneberger *et al.* introduced a skip connections mechanism that allows the relevant features to be recovered from the encoder.

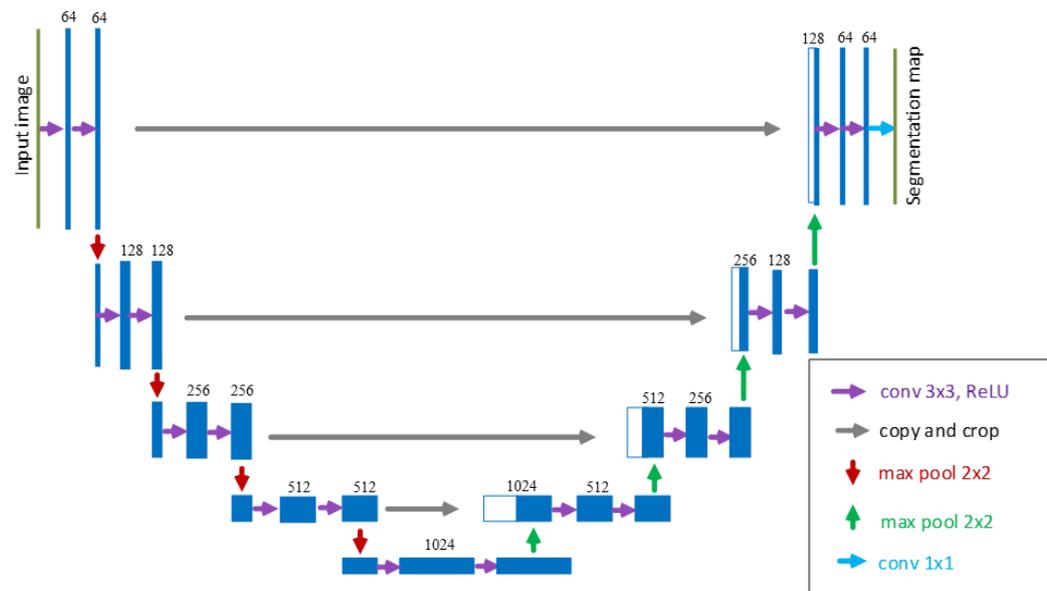


Figure 1. General structure of the U-Net, adapted from [13].

### 3. METHODOLOGY

#### 3.1. The proposed method

In this paper, we propose an approach to segment brain tumors from MRI images using the U-Net baseline and the proposed loss function. Figure 2 shows the general pipeline of the proposed method. First, the training images and their ground truth masks are fed into the U-Net neural network. The network is trained using the proposed loss function that presents the difference between the ground truth and corresponding network output. During minimizing the loss function, parameters of the network are updated at each epoch. The parameter updating is iterated until the optimization process is converged. Then, the model weight at which the optimization is converged is used to predict the output maps of the testing images. For the evaluation of segmentation results, two metrics namely Dice similarity coefficient (DSC) and the Intersection over Union (IoU) are computed.

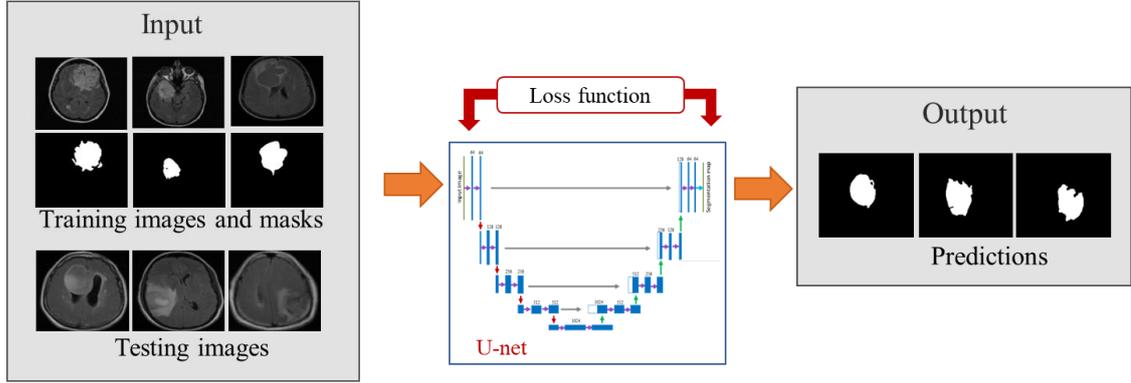


Figure 2. Main steps of the proposed method for segmentation of brain tumor.

### 3.2. The proposed image-driven Level Set based Loss function

In common neural networks for image segmentation, the binary cross entropy (BCE) and Dice losses are usually used to train the network. However, those losses often lack pixel-wise fitting and boundary constraints, leading to undesired results near the boundary [16]. To solve this problem, we propose to introduce an image-driven level set term into the Dice loss. In particular, we combine the Dice loss with the Level set loss inspired by the work of Kim and Yeh [18]. The proposed loss function for brain tumor segmentation is expressed as:

$$L_{DLS} = L_{Dice} + \lambda L_{LS} \quad (1)$$

where  $L_{Dice}$  is the Dice loss,  $L_{LS}$  is the level set loss, and  $\lambda$  is a hyper-parameter that controls the importance of the level set term. The two loss terms are defined as follows:

Let  $S_i$  be the predicted probability of pixel  $i$  belonging to the desired object (brain tumor region in this study) in the binary segmentation map, and  $G_i$  denotes the corresponding semantic label, the Dice loss can be expressed as:

$$L_{Dice} = - \frac{2 \sum_{i=1}^P G_i S_i}{\sum_{i=1}^P G_i + \sum_{i=1}^P S_i + \varepsilon} \quad (2)$$

with  $\varepsilon$  is a smooth factor used for numerical stability, and  $P$  is the number of pixels of input image  $I$ .

The level set loss is inspired by the Mumford-Shah functional [18] for the unsupervised task under the level set framework described in [19]. Denote  $\theta$  be the trainable network parameters, the level set loss is defined as:

$$L_{LS}(\mathbf{x}; \theta) = \frac{1}{P} \left[ \int_{\Omega} |I(\mathbf{x}) - c_1|^2 S(\mathbf{x}; \theta) d\mathbf{x} + \int_{\Omega} |I(\mathbf{x}) - c_2|^2 (1 - S(\mathbf{x}; \theta)) d\mathbf{x} + \beta \int_{\Omega} |\nabla S(\mathbf{x}; \theta)| d\mathbf{x} \right] \quad (3)$$

where  $I(\mathbf{x})$  is the given image;  $S(\mathbf{x}; \theta)$  refers to the prediction;  $\mathbf{x}$  is a pixel of the given image;  $P$  is the number of pixel,

$$c_1(\theta) = \frac{\int_{\Omega} I(\mathbf{x})S(\mathbf{x};\theta)d\mathbf{x}}{\int_{\Omega} S(\mathbf{x};\theta)d\mathbf{x}}; \quad c_2(\theta) = \frac{\int_{\Omega} I(\mathbf{x})(1-S(\mathbf{x};\theta))d\mathbf{x}}{\int_{\Omega} (1-S(\mathbf{x};\theta))d\mathbf{x}} \quad (4)$$

are the mean intensity values of the regions inside and outside the boundary of predicted map; and  $\beta$  is a hyper-parameter of the boundary term  $\int_{\Omega} |\nabla S(\mathbf{x};\theta)|d\mathbf{x}$ . The boundary term exploits the boundary information of the prediction and can be recognized as a regularization term for the desired object boundary so that the boundary can be smooth, and undesired small objects can be omitted.

In the discrete form, the above equation can be expressed as:

$$L_{LS}(\mathbf{x};\theta) = \frac{1}{W \times H} \left[ \sum_{i=1}^W \sum_{j=1}^H (I_{i,j} - c_1)^2 S_{i,j}(\theta) + \sum_{i=1}^W \sum_{j=1}^H (I_{i,j} - c_2)^2 (1 - S_{i,j}(\theta)) + \beta \sum_{i=1}^W \sum_{j=1}^H \sqrt{(\nabla S_{x_{i,j}}(\theta))^2 + (\nabla S_{y_{i,j}}(\theta))^2} + \varepsilon \right] \quad (5)$$

where  $W$  is image width,  $H$  is image height, and  $\varepsilon(\varepsilon > 0)$  is used to avoid zeros.

$$c_1 = \frac{\sum_{i=1}^W \sum_{j=1}^H I_{i,j} S_{i,j}(\theta)}{\sum_{i=1}^W \sum_{j=1}^H S_{i,j}(\theta)}; \quad c_2 = \frac{\sum_{i=1}^W \sum_{j=1}^H I_{i,j} (1 - S_{i,j}(\theta))}{\sum_{i=1}^W \sum_{j=1}^H (1 - S_{i,j}(\theta))} \quad (6)$$

### 3.3. Neural Network Training

The Adam algorithm is used to optimize the trainable parameters of the model with an initial learning rate of 0.001. The network is trained for 150 epochs with a batch size of 32. The hyper-parameters  $\lambda_{LS}$  and  $\beta$  are adjusted as  $10^{-3}$  and 1 for all experiments, respectively. During training, data augmentation is performed with some basic operations like rotation up to 20 degrees, vertical and horizontal flipping.

## 4. EVALUATIONS AND RESULTS

### 4.1. Dataset

To evaluate the performance of the proposed approach, we used a database including manual segmentation and tumor masks provided by Buda *et al.* [7] and available for download at [20]. The dataset consists of images with and without tumors acquired from 110 LGG patients. The dataset is obtained from the Cancer Genome Atlas and the Cancer Imaging Archive [21] that includes the data acquired from various groups of Grade II (50 patients), Grade III (58 patients), and with unknown's tumor grades (2 patients). In implementation, the data is split into training and testing at a ratio of 80 and 20, respectively. While training the neural network, 10 % of the training data portion is used for validation to assess the performance of the method.

### 4.2. Evaluation metrics

For quantitative evaluation of image segmentation, we use the Dice similarity index (DSC), and Intersection over Union (IoU). The DSC measures the overlap between automatic and manual segmentation [22], and is expressed as:

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

where  $TP$ ,  $FP$ ,  $FN$  denote the number of true positives, false positives and false negatives, respectively.

The Intersection over Union, also called Jaccard index, is used to measure the similarity between the manual and automatic segmentation maps, and is defined as:

$$\text{IoU} = \frac{TP}{TP + FN + FP} \quad (8)$$

### 4.3 Results and performance evaluation

In this section, we show some representative segmentations of brain tumors and provide quantitative scores by the proposed approach. In addition, we also compare the results when training the network with other loss functions. The comparative losses include the binary cross entropy (BCE) loss, Dice loss, and their combination (BCE-Dice) loss.

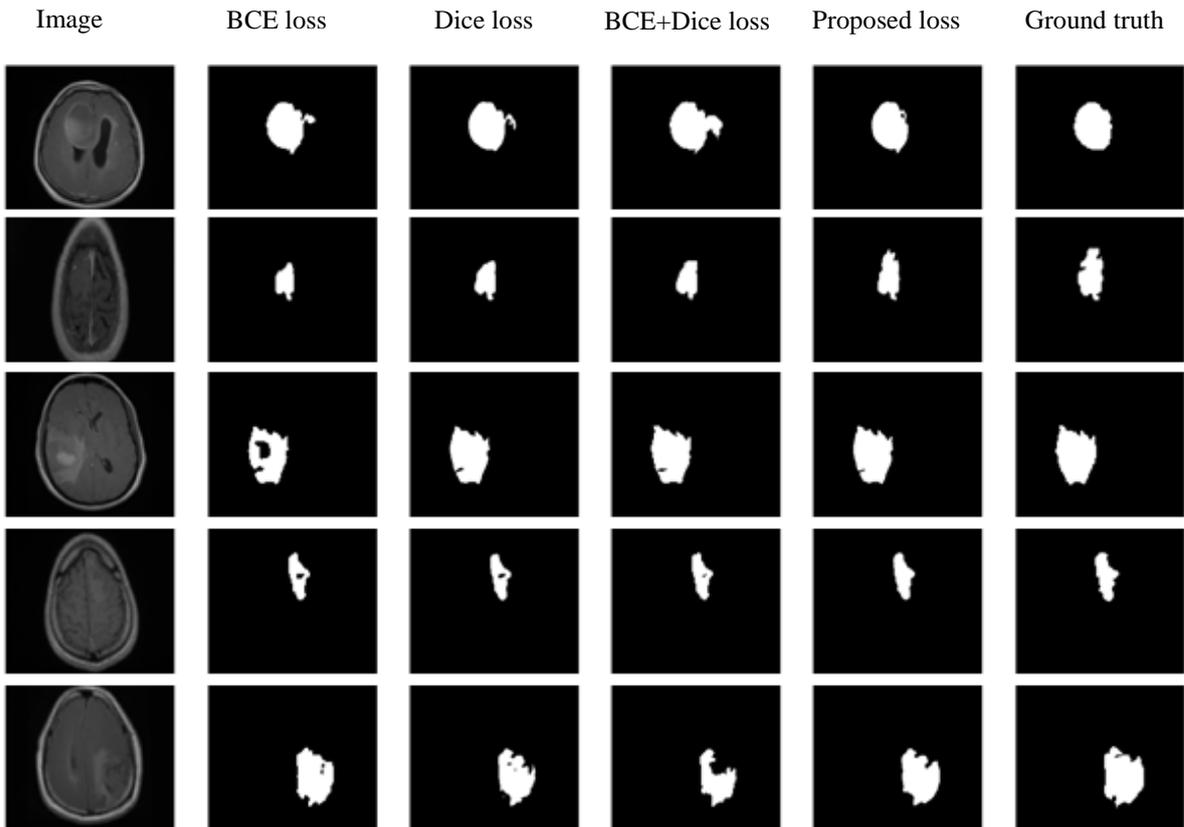


Figure 3. Representative segmentation results on the brain tumor dataset by the proposed loss function in comparison with different loss functions including Binary cross entropy (BCE), Dice, and combination of BCE and Dice losses.

To demonstrate advantages of the proposed method when using our hybrid loss in comparison with other loss functions, we show some representative segmentations and comparison in Fig. 3. Segmentation results when training by BCE loss, Dice loss, and the combination BCE-Dice loss are respectively presented in the second, third, and fourth columns of Fig. 3. The corresponding ground truths are given in the last column of this figure. It can be seen from the last two columns of this figure, especially the fifth and sixth columns, that the segmentations using the proposed loss are best agreed with the ground truths.

For quantitative evaluation, we provide the evaluation scores including both DSC and IoU metrics in Table 1. As we can see from this table, the scores obtained by the proposed approach are with the highest DSC and IoU values, demonstrating the performance of the proposed method.

We conduct another experiment to compare the performance of the proposed approach and other state of the arts. In more detail, we reimplement the Attention UNet [17], Nested UNet [23], and Multires UNet [24] and apply the models with the same training and test images. The comparison results are reported in Table 2. It can be observed from this Table, the proposed approach archives higher DSC and IoU values, which shows the advantages of the proposed approach.

*Table 1.* The average Dice similarity coefficient (DSC) and Intersection over Union (IoU) by U-Net architecture when trained by different losses

| Loss function | DSC          | IoU          |
|---------------|--------------|--------------|
| BCE           | 0.877        | 0.783        |
| Dice          | 0.880        | 0.788        |
| BCE-Dice      | 0.880        | 0.789        |
| Proposed      | <b>0.916</b> | <b>0.848</b> |

*Table 2.* Comparison of average Dice similarity coefficient (DSC) and Intersection over Union (IoU) between prediction results and ground truths on test data by the proposed approach with other approaches

| Method              | DSC   | IoU   |
|---------------------|-------|-------|
| Attention UNet [17] | 0.875 | 0.785 |
| Nested UNet [23]    | 0.904 | 0.826 |
| Multires UNet [24]  | 0.897 | 0.817 |
| Proposed            | 0.916 | 0.848 |

#### 4. CONCLUSION

In this paper, a method for automatic brain tumor segmentation from MRI images is presented, in which a loss function created by integrating a Level set loss function into Dice loss function is proposed. We have applied the proposed approach to segment brain tumors from MRI images. We also compare the performance of the proposed loss function with other loss functions when applying them to train the U-Net network using the benchmark of MRI brain tumor datasets. The experimental results show good performance of the network when trained

by the proposed loss function. In future works, we plan to extend the proposed loss function to multiphase segmentation tasks such as left ventricle, skin lesions, and brain segmentation. In addition, some other information related to image intensity and histogram can be incorporated in the loss for better segmentation performance of natural images such as CamVid and BSDS500 databases.

**Acknowledgments:** This research is funded by the Hanoi University of Science and Technology (HUST) under project number T2020-PC-017.

**CRedit authorship contribution statement.** Van-Truong Pham: Introduction, Background, Methodology, Funding acquisition. Thi-Thao Tran: Evaluations and Results, Conclusions.

**Declaration of competing interest.** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## REFERENCES

1. De Angelis L. M. - Brain Tumors, *New England Journal of Medicine* **344** (2) (2001) 114-123.
2. Siegel R., Miller K., and Jemal A. - Cancer statistics, 2020 (US statistics), *CA A Cancer J. Clin.* **70** (1) (2020) 7-30.
3. Menze B. H., et al. - The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS) *IEEE Trans Med Imaging* **34** (10) (2015) 1993-2024.
4. Sachdeva J., Kumar V., Gupta I., Khandelwal N., and Ahuja C. K. - A novel content-based active contour model for brain tumor segmentation, *Magnetic Resonance Imaging* **30** (5) (2012) 694-715.
5. Shyu K. K., Pham V. T., Tran T. T., and Lee P. L. - Unsupervised active contours driven by density distance and local fitting energy with applications to medical image segmentation, *Mach. Vis. Appl.* **23** (6) (2012) 1159-1175.
6. Havaei M., Guizard N., Larochelle H., and Jodoin P. - Deep Learning Trends for Focal Brain Pathology Segmentation in MRI, *Machine Learning for Health Informatics*, 2016, pp. 125-148.
7. Buda M., Saha A., and A. Mazurowski M. - Association of genomic subtypes of lower-grade gliomas with shape features automatically extracted by a deep learning algorithm, *Comp. in Bio. and Med.* **109** (2019) 218-225.
8. Zhang J., Zeng J., Qin P., and Zhao L. - Brain tumor segmentation of multi-modality MR images via triple intersecting U-Nets, *Neurocomputing* **421** (2021) 195-209.
9. Liu Z., Chen L., Tong L., Zhou F., Jiang Z., and Zhang Q., et al. - Deep learning based brain tumor segmentation: A survey, arXiv:2007.09479, 2020, [online] (2020) Available: <http://arxiv.org/abs/2007.09479>.
10. Pham V. T., Tran T. T., Wang P. C., Chen P. Y., and Lo M. T. - EAR-UNet: A deep learning-based approach for segmentation of tympanic membranes from otoscopic images, *Artificial Intelligence In Medicine* **115** (2021) 1-12.
11. J. Long, E. Shelhamer, and T. Darrell - Fully convolutional networks for semantic segmentation, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 3431-3440.

12. Ninh Q. C., Tran T. T., Tran T. T., Tran T. A. X., and Pham V. T. - Skin Lesion Segmentation Based on Modification of SegNet Neural Networks, Proceedings of the 6th NAFOSTED Conference on Information and Computer Science (NICS), 2020 Ha Noi, pp. 575-578.
13. Ronneberger O., Fischer P., and Brox T. - U-net: Convolutional networks for biomedical image segmentation, Proceedings of the Int. Conf. Med. Image Comput. Comput.-Assist. Intervent., 2015, pp. 234-241.
14. Chen W., Liu B., Peng. S., Sun J., and Qiao X. - S3D-UNet: Separable 3D U-Net for Brain Tumor Segmentation, Proceedings of the International MICCAI Brainlesion Workshop, 2018, pp. 358-368.
15. Salehi S. S. M., Erdogmus D., and Gholipour A. - Tversky loss function for image segmentation using 3D fully convolutional deep networks, Proceedings of the International Workshop on Machine Learning in Medical Imaging, Springer, 2017, pp. 379-387.
16. Chen X., M. Williams B., R. Vallabhaneni S., Czanner G., Williams R., and Zheng Y. - Learning Active Contour Models for Medical Image Segmentation, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2019. 11623-11640.
17. Oktay O., Schlemper J., Folgoc L. L., Lee M., Heinrich M., Misawa K., *et al.* - Attention U-Net: Learning Where to Look for the Pancreas, Proceedings of the 1st Conf. Med. Imaging with Deep Learn., 2018. Available: <https://arxiv.org/abs/1804.03999>.
18. Kim B. and Ye J. C. - Mumford–Shah loss functional for image segmentation with deep learning, IEEE Transactions on Image Processing **29** (2019) 1856-1866.
19. Chan T. F. and Vese L. A. - Image segmentation using level sets and the piecewise-constant Mumford-Shah model, Tech. Rep. 0014, Computational Applied Math Group, 2000. Citeseer.
20. "Brain-segmentation-pytorch Kaggle," [Online], Available: [www.kaggle.com/mateuszbuda/brain-segmentation-pytorch/data](http://www.kaggle.com/mateuszbuda/brain-segmentation-pytorch/data) (access date: August, 29, 2020),
21. TCGA-LGG - the cancer imaging archive (TCIA) public access -cancer imaging, [Online]. Available: <https://wiki.cancerimagingarchive.net/display/Public/TCGA-LGG> (access date: August, 29, 2020),
22. Pham V. T., Tran T. T., Wang P. C., and Lo M. T., Tympanic Membrane Segmentation in Otosopic Images based on Fully Convolutional Network with Active Contour Loss, Signal, Image and Video Processing, **15** (3) (2021) 519-527.
23. Zhou Z., Siddiquee M. M. R., Tajbakhsh N., and Liang - J. UNet++: A nested U-net architecture for medical image segmentation, Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support, 2018, pp. 3-11.
24. Ibtihaz N. and Rahman M. S. - Multiresunet: Rethinking the u-net architecture for multimodal biomedical image segmentation, Neural Networks **121** (2020) 74-87.