



# Deep Learning Systems for Automated Segmentation of Brain Tissues and Tumors in MRIs

Dante Mújica-Vargas<sup>1</sup>(✉), Manuel Matuz-Cruz<sup>2</sup>, Eduardo Ramos-Díaz<sup>3</sup>,  
and Jean Marie Vianney Kinani<sup>3</sup>

<sup>1</sup> Departamento de Ciencias Computacionales, TecNM/CENIDET,  
Cuernavaca-Morelos, México  
[dante.mv@cenidet.tecnm.mx](mailto:dante.mv@cenidet.tecnm.mx)

<sup>2</sup> Departamento de Sistemas Computacionales, TecNM/ITTapachula,  
Tapachula, Chiapas, México

<sup>3</sup> Instituto Politécnico Nacional-UPIIH,  
San Agustín Tlaxiaca-Hidalgo, México

**Abstract.** A pair of fully automatic brain tissue and tumor segmentation frameworks are introduced in current paper, these consist of a parallel and cascade architectures of a specialized convolutional deep neural network designed to develop binary segmentation. The main contributions of this proposal imply their ability to segment Magnetic Resonance Imaging (MRI) of the brain, of different acquisition modes without any parameter, they do not require any preprocessing stage to improve the quality of each slice. Experimental tests were developed considering BraTS 2017 database. The robustness and effectiveness of this proposal is verified by quantitative and qualitative results.

**Keywords:** Magnetic resonance imaging of the brain · Tissues and tumor segmentation · Convolutional deep neural network

## 1 Introduction

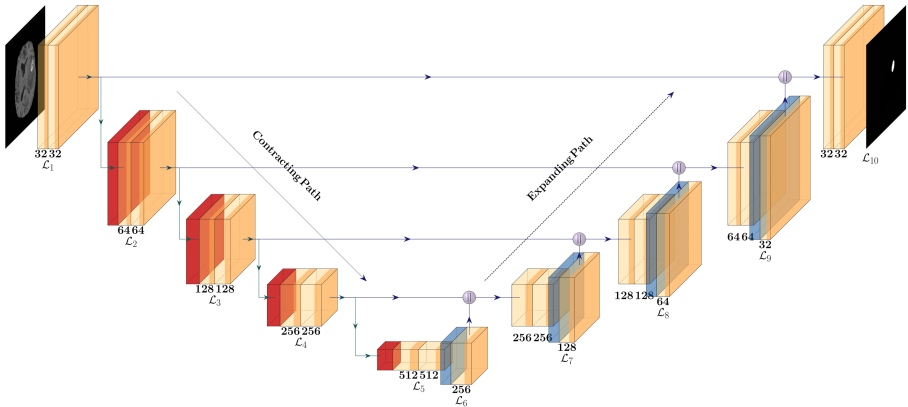
Magnetic resonance imaging is a medical modality used to guide the diagnosis process and the treatment planning. To do so, it needs to develop the images or slices segmentation, in order to detect and characterize the lesions, as well as to visualize and quantify the pathology severity. Based on their experiences and knowledge, medical specialists make a subjective interpretation of this type of images; in other words, a manual segmentation is performed. This task is long, painstaking and subject to human variability. Brain tumor segmentation is one of the crucial steps for surgery planning and treatment evaluation. Despite the great amount of effort being put to address this challenging problem in the past two decades, segmentation of brain tumor remains to be one of the most challenging tasks in medical image analysis. This is due to both the intrinsic nature of the tumor tissue being heterogeneous and the extrinsic problems with

unsatisfactory image quality of clinical MRI scans. For example, the tumor mass of a glioma patient, the most common brain tumor, often consists of peritumoral edema, necrotic core, enhancing and non-enhancing tumor core. In addition to the complicated tumor tissue pattern, the MRI images can be further corrupted with a slowly varying bias field and/or motion artifacts, etc. Brain MRIs in most cases do not have well-defined limits between the elements that compose them; in addition, they include non-soft tissues, as well as artifacts that can hinder segmentation. Despite all these inherent conditions, numerous automatic algorithms or techniques have been developed and introduced in state-of-the-art. Among approaches exclusively designed to segment the brain tissues stand out those that are based on the paradigm of fuzzy clustering as well as all its variants [1–4, 28]. With the same purpose, hybrid methods based on combinations of different paradigms of machine learning and optimization algorithms have also been presented, e.g. [5–7]. On the other hand, methods designed to segment brain tumors or other abnormalities have also been introduced, among which one can refer to [8–14, 29]. For this task it is possible to affirm that in the state-of-the-art the proposals based on Deep Learning are the most novel and have the best results. The majority of these proposals yielded a high performance in the image processing task, specifically when these were brain magnetic resonance images. Nevertheless, after the pertinent analysis it was noted that most methods of them suffer from one or more challenges such as: training need, special handcrafted features (local or global), sensitive to initializations, many parameters that require a tuning, various processing stages, designed to segment just  $T1$ -weighted brain MRI images, among others. In this research paper, we concentrate on brain tissues and tumors segmentation in magnetic resonance images. For that purpose we introduce a system consisting in a cascade of U-Net models enhanced with a fusion information approach, this proposal can be considered an extension to the one presented in [29], as well as a formalization of theoretical concepts. The introduced proposal has the following special features in contrast with those above-mentioned: 1) it is able to segment RMIs with different relaxation times such as  $T1$ ,  $T2$ ,  $T1ce$  and  $Flair$ , 2) it is easily adaptable to segment brain tumors, 3) it does not require the initialization of any parameter, such as the number of regions in which the slice will be segmented, 4) it does not require any preprocessing stage to improve the segmentation quality of each slice and 5) it does not need various processing stages to increase its performance. The rest of this paper is organized as follow. A mathematical formulation of U-Net is introduced in detail in Sect. 2. The parallel and cascade architectures of Convolutional Neural Networks are introduced in Sect. 3. Experimental results and a comparative analysis with other current methods in the literature are presented in Sect. 4. In the final section the Conclusions are drawn and future work is outlined.

## 2 U-Net: Convolutional Networks for Biomedical Image Segmentation

Deep neural networks have shown remarkable success in domains such as image classification [6,15], object detection and localization [17–19], language processing applications [20,21], speech recognition [22], medical image processing [16,23,24], among many others. In the deep learning paradigm, Convolutional Neural Networks (CNNs) outstands as the major architecture, they deploy convolution operations on hidden layers for weight sharing and parameter reduction, in order to extract local information from grid-like input data. In simple words, they process de information by means of hierarchical layers in order to understand representations and features from data in increasing levels of complexity.

For biomedical image segmentation, U-Net is one of the most important architectures that have been introduced in the state of the art [25]. It is a fully convolutional neural network model originally designed to develop a binary segmentation; that is, the main object and the background of the image. This network is divided into two parts, in the first part, the images are subjected to a downward sampling (Contracting Path or Encoder), by means of convolution operations with a kernel of  $3 \times 3$  each followed by a rectified linear unit (ReLU) and a maximum grouping layer of  $2 \times 2$ . The next part of the model (Expanding Path or Decoder) consists of layers of deconvolution and convolution with  $2 \times 2$  kernel, finally the output will correspond to a specific class of objects to be segmented. The processing implies nineteen convolutions ( $\mathcal{C}$ ), four subsamplings ( $\mathcal{S}$ ), four upsamplings ( $\mathcal{U}$ ) and four mergings ( $\mathcal{M}$ ). To understand the model variant that we propose to segment brain tumors from BraTS 2017, it can be see Fig. 1.



**Fig. 1.** U-Net model.

In the **Contracting Path**: *Layer 1* is a convolution layer which takes input image  $X$  of size  $240 \times 240$  and convolves it with 32 filters of size  $3 \times 3$ ; producing 32 feature maps of  $240 \times 240$ . In order to make the output of linear operation

nonlinear, it is advisable to output of convolution operation is passed through a Rectified Linear Units (ReLU) activation function  $\sigma(x) = \max(0, x)$ , in this way:

$$\mathcal{C}1_{i,j}^k = \sigma \left( \sum_{m=0}^2 \sum_{n=0}^2 w_{m,n}^k * X_{i+m,j+n} + b^k \right) \quad (1)$$

where  $\mathcal{C}1^k$  stands for  $k$ -th output feature map in  $\mathcal{C}1$  layer,  $(m, n)$  are the indices of the  $k$ -th kernel (filter), while that  $(i, j)$  are the indices of output.  $\mathcal{C}1_{i,j}^k$  is convolved with 32 kernels of size  $3 \times 3$ ; in the same way, the output is rectified:

$$\mathcal{C}2_{i,j}^k = \sigma \left( \sum_{d=0}^{31} \sum_{m=0}^2 \sum_{n=0}^2 w_{m,n}^{k,d} * \mathcal{C}1_{i+m,j+n}^d + b^k \right) \quad (2)$$

In *Layer 2*, the output of the convolution layer  $\mathcal{C}2^k$  is fed to max-pooling layer  $\mathcal{S}1^k = \max \text{Pool}(\mathcal{C}2^k)$ . For each feature map in  $\mathcal{C}2^k$ , max-pooling performs the following operation:

$$\mathcal{S}1_{i,j}^k = \max \left( \begin{array}{cc} \mathcal{C}2_{2i,2j}^k & \mathcal{C}2_{2i+1,2j}^k \\ \mathcal{C}2_{2i,2j+1}^k & \mathcal{C}2_{2i+1,2j+1}^k \end{array} \right) \quad (3)$$

where  $(i, j)$  are the indices of  $k$ -th feature map of output, and  $k$  is the feature map index. The output consists in 32 feature maps with size of  $120 \times 120$ , which implies one-half the size of the input image ( $X/2$ ).  $\mathcal{S}1^k$  is convolved with 64 kernels of size  $3 \times 3$ , and the result is rectified:

$$\mathcal{C}3_{i,j}^k = \sigma \left( \sum_{d=0}^{63} \sum_{m=0}^2 \sum_{n=0}^2 w_{m,n}^{k,d} * \mathcal{S}1_{i+m,j+n}^d + b^k \right) \quad (4)$$

By means of 64  $3 \times 3$  filters  $\mathcal{C}3^k$  is convolved, and then the results is rectified:

$$\mathcal{C}4_{i,j}^k = \sigma \left( \sum_{d=0}^{63} \sum_{m=0}^2 \sum_{n=0}^2 w_{m,n}^{k,d} * \mathcal{C}3_{i+m,j+n}^d + b^k \right) \quad (5)$$

*Layers 3, 4 and 5* a similar process than *Layer 2*. A max-pooling filtering followed that two convolutional stages along with their respective linear rectification. In brief, for *Layer 3*:

$$\mathcal{S}2_{i,j}^k = \max \left( \begin{array}{cc} \mathcal{C}4_{2i,2j}^k & \mathcal{C}4_{2i+1,2j}^k \\ \mathcal{C}4_{2i,2j+1}^k & \mathcal{C}4_{2i+1,2j+1}^k \end{array} \right) \quad (6)$$

with size ( $X/4$ ) is convolved with  $128 \times 128$  kernels with size  $3 \times 3$  kernel in order to obtain  $128 \times 128$  feature maps:

$$\mathcal{C}5_{i,j}^k = \sigma \left( \sum_{d=0}^{127} \sum_{m=0}^2 \sum_{n=0}^2 w_{m,n}^{k,d} * \mathcal{S}2_{i+m,j+n}^d + b^k \right) \quad (7)$$

$$\mathcal{C}6_{i,j}^k = \sigma \left( \sum_{d=0}^{127} \sum_{m=0}^2 \sum_{n=0}^2 w_{m,n}^{k,d} * \mathcal{C}5_{i+m,j+n}^d + b^k \right) \quad (8)$$

For *Layer 4*, the max-pooling filtering on  $\mathcal{C}6^k$  is calculated as:

$$\mathcal{S}3_{i,j}^k = \max \left( \begin{array}{cc} \mathcal{C}6_{2i,2j}^k & \mathcal{C}6_{2i+1,2j}^k \\ \mathcal{C}6_{2i,2j+1}^k & \mathcal{C}6_{2i+1,2j+1}^k \end{array} \right) \quad (9)$$

The result with size  $(X/8)$  is convolved with  $3 \times 3$  size  $256 \times 256$  kernels to obtain same number of rectified feature maps:

$$\mathcal{C}7_{i,j}^k = \sigma \left( \sum_{d=0}^{255} \sum_{m=0}^2 \sum_{n=0}^2 w_{m,n}^{k,d} * \mathcal{S}3_{i+m,j+n}^d + b^k \right) \quad (10)$$

$$\mathcal{C}8_{i,j}^k = \sigma \left( \sum_{d=0}^{255} \sum_{m=0}^2 \sum_{n=0}^2 w_{m,n}^{k,d} * \mathcal{C}7_{i+m,j+n}^d + b^k \right) \quad (11)$$

Last layer in contracting path takes as input  $\mathcal{C}8^k$ , and performances a down-sampling to reduce the feature maps into  $15 \times 15$  size (i.e.  $X/16$ ), the filter is given as:

$$\mathcal{S}4_{i,j}^k = \max \left( \begin{array}{cc} \mathcal{C}8_{2i,2j}^k & \mathcal{C}8_{2i+1,2j}^k \\ \mathcal{C}8_{2i,2j+1}^k & \mathcal{C}8_{2i+1,2j+1}^k \end{array} \right) \quad (12)$$

A  $3 \times 3$  pair of kernels are considered to extract the deepest features, for that purpose  $\mathcal{S}4^k$  is convolved with 512 kernels, and the result is once again convolved with other 512 kernels, in both cases it is necessary to consider a linear rectification to avoid negative numbers, these operations are given as:

$$\mathcal{C}9_{i,j}^k = \sigma \left( \sum_{d=0}^{511} \sum_{m=0}^2 \sum_{n=0}^2 w_{m,n}^{k,d} * \mathcal{S}4_{i+m,j+n}^d + b^k \right) \quad (13)$$

$$\mathcal{C}10_{i,j}^k = \sigma \left( \sum_{d=0}^{511} \sum_{m=0}^2 \sum_{n=0}^2 w_{m,n}^{k,d} * \mathcal{C}9_{i+m,j+n}^d + b^k \right) \quad (14)$$

In the **Expanding Path**: *Layer 6* develops an un-pooling process by means of nearest interpolation with an up-sampling factor of  $\uparrow_2$  for rows and columns. In short, it repeats twice each row and column of the  $k$ -th feature map:

$$\mathcal{U}1_{i,j}^k = \begin{bmatrix} \mathcal{C}10_{2i,2j}^k & \mathcal{C}10_{2i,2j}^k \\ \mathcal{C}10_{2i,2j}^k & \mathcal{C}10_{2i,2j}^k \end{bmatrix} \quad (15)$$

$\mathcal{U}1^k$  makes that the  $k$ -th feature map increase its size to  $X/8$ . *Layers 6* and *4* are merged such as a concatenation process, i.e.:

$$\mathcal{M}1_{i,j}^k = [\mathcal{U}1_{i,j}^k; \mathcal{C}8_{i,j}^k] \quad (16)$$

$\mathcal{M}1^k$  consists of  $768 \times 768$  feature maps. In *Layer 7*,  $\mathcal{M}1^k$  is in the first instance convolved with  $256 \times 256$  kernels with a size of  $3 \times 3$ , each of the results is rectified in order to avoid negative numbers:

$$\mathcal{C}11_{i,j}^k = \sigma \left( \sum_{d=0}^{255} \sum_{m=0}^2 \sum_{n=0}^2 w_{m,n}^{k,d} * \mathcal{M}1_{i+m,j+n}^d + b^k \right) \quad (17)$$

$$\mathcal{C}12_{i,j}^k = \sigma \left( \sum_{d=0}^{255} \sum_{m=0}^2 \sum_{n=0}^2 w_{m,n}^{k,d} * \mathcal{C}11_{i+m,j+n}^d + b^k \right) \quad (18)$$

Afterwards, (18) is up-sampling with a  $\uparrow_2$  factor in order to increase the feature maps to an  $X/4$  size:

$$\mathcal{U}2_{i,j}^k = \begin{bmatrix} \mathcal{C}12_{2i,2j}^k & \mathcal{C}12_{2i,2j}^k \\ \mathcal{C}12_{2i,2j}^k & \mathcal{C}12_{2i,2j}^k \end{bmatrix} \quad (19)$$

Outflows of *Layers 7* and *3* are concatenated as:

$$\mathcal{M}2_{i,j}^k = [\mathcal{U}2_{i,j}^k; \mathcal{C}6_{i,j}^k] \quad (20)$$

$\mathcal{M}2^k$  consists of  $384 \times 384$  feature maps. *Layers 8* and *9* follow a similar processing than *Layer 7*; but with a decreasing of feature maps number. In this regard, for *Layer 8*:

$$\mathcal{C}13_{i,j}^k = \sigma \left( \sum_{d=0}^{127} \sum_{m=0}^2 \sum_{n=0}^2 w_{m,n}^{k,d} * \mathcal{M}1_{i+m,j+n}^d + b^k \right) \quad (21)$$

$$\mathcal{C}14_{i,j}^k = \sigma \left( \sum_{d=0}^{127} \sum_{m=0}^2 \sum_{n=0}^2 w_{m,n}^{k,d} * \mathcal{C}13_{i+m,j+n}^d + b^k \right) \quad (22)$$

$$\mathcal{U}3_{i,j}^k = \begin{bmatrix} \mathcal{C}14_{2i,2j}^k & \mathcal{C}14_{2i,2j}^k \\ \mathcal{C}14_{2i,2j}^k & \mathcal{C}14_{2i,2j}^k \end{bmatrix} \quad (23)$$

$\mathcal{U}3^k$  makes that all feature maps increase their size to  $X/2$ .  $\mathcal{U}3^k$  is concatenated with  $\mathcal{C}4^k$ , the output implies  $192 \times 192$  features maps:

$$\mathcal{M}3_{i,j}^k = [\mathcal{U}3_{i,j}^k; \mathcal{C}4_{i,j}^k] \quad (24)$$

For *Layer 9*:

$$\mathcal{C}15_{i,j}^k = \sigma \left( \sum_{d=0}^{63} \sum_{m=0}^2 \sum_{n=0}^2 w_{m,n}^{k,d} * \mathcal{M}3_{i+m,j+n}^d + b^k \right) \quad (25)$$

$$\mathcal{C}16_{i,j}^k = \sigma \left( \sum_{d=0}^{63} \sum_{m=0}^2 \sum_{n=0}^2 w_{m,n}^{k,d} * \mathcal{C}15_{i+m,j+n}^d + b^k \right) \quad (26)$$

$$\mathcal{U}4_{i,j}^k = \begin{bmatrix} \mathcal{C}16_{2i,2j}^k & \mathcal{C}16_{2i,2j}^k \\ \mathcal{C}16_{2i,2j}^k & \mathcal{C}16_{2i,2j}^k \end{bmatrix} \quad (27)$$

Last up-sampling process  $\mathcal{U}4^k$  makes that all feature maps increase their size to  $X$ .  $\mathcal{U}4^k$  is concatenated with  $\mathcal{C}2^k$ :

$$\mathcal{M}4_{i,j}^k = [\mathcal{U}4_{i,j}^k; \mathcal{C}2_{i,j}^k] \quad (28)$$

$\mathcal{M}4^k$  implies  $96 \times 96$  features maps, these are convolved with  $32 \times 32$  kernels with size  $3 \times 3$ :

$$\mathcal{C}17_{i,j}^k = \sigma \left( \sum_{d=0}^{31} \sum_{m=0}^2 \sum_{n=0}^2 w_{m,n}^{k,d} * \mathcal{M}4_{i+m,j+n}^d + b^k \right) \quad (29)$$

$$\mathcal{C}18_{i,j}^k = \sigma \left( \sum_{d=0}^{31} \sum_{m=0}^2 \sum_{n=0}^2 w_{m,n}^{k,d} * \mathcal{C}17_{i+m,j+n}^d + b^k \right) \quad (30)$$

Last convolution layer convolves  $\mathcal{C}18^k$  with 32 kernels of size  $1 \times 1$ , the linear activation function is replaced by a Sigmoid one,  $\sigma(x) = \frac{1}{1 + \exp^{-x}}$ ; this function lets to highlight the region of interest (region to be segmented) in the segmented image  $\check{X}$ , since it squashes to  $\mathcal{C}18^k$  into the range  $[0, 1]$ . In this way:

$$\check{X}_{i,j} = \sigma \left( \sum_{d=0}^{31} \sum_{m=0}^1 \sum_{n=0}^1 w_{m,n}^{k,d} * \mathcal{C}18_{i+m,j+n}^d + b^k \right) \quad (31)$$

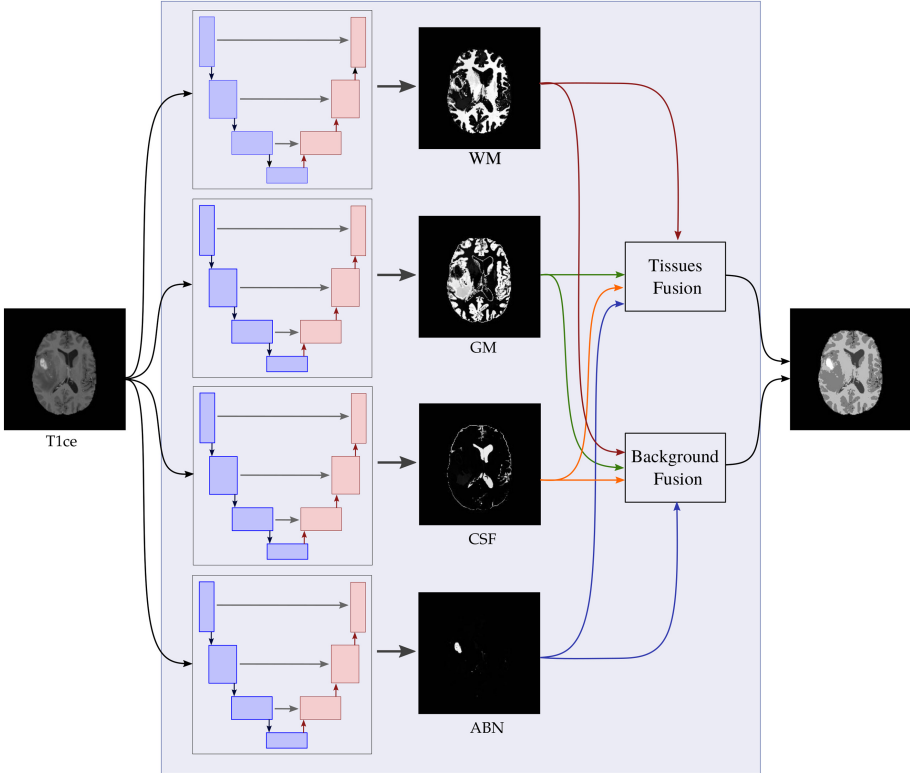
The segmented image  $\check{X}$  has a size  $240 \times 240$  such as the input image  $X$ . Expressions (1) to (31) let us to depict how CNNs follow a hierarchical processing to extract specific features in order to detect and segment specific regions; in this case, brain tumors.

### 3 Deep Learning Systems for Brain Image Segmentation

#### 3.1 Proposed Parallel System for Brain Tissues Segmentation

Conventionally, it may be assumed that different tissues can be found in a MRI slice: (1) White Matter (WM), (2) Gray Matter (GM), (3) Cerebral Spinal Fluid (CSF) and (4) Abnormalities (ABN). Nevertheless, it should be clarified that depending on the slice, not all regions may be present or the magnitude of their presence will be variant. In order to develop an automatic soft tissues recognition and their segmentation we suggest the system depicted in Fig. 2; it is basically comprised by four U-Nets models trained to work on a specific soft. After each tissue segmentation, it is necessary to perform the joint area representation by means of a method that determines appropriately the target and background region. From following rules [26], it is possible to perform the fusion of the segmented tissues, the background fusion, as well as the fusion of the detected and regions:

1. If  $R^1$  and  $R^2$  do not intersect, two regions are formed in the representation of joint area, that is  $R_1^3 = R^1$  and  $R_2^3 = R^2$ .
2. If  $R_1$  and  $R_2$  are partly intersected, three regions are formed in the representation of joint area, that is  $R_1^3 = R^1 \cap R^2$ ,  $R_2^3 = R^1 - R_1^3$ ,  $R_3^3 = R^2 - R_1^3$ .
3. If there is an area that is completely include, such as  $R^1 \subset R^2$ , then two regions are formed in the representation of joint area, that is  $R_1^3 = R^2$  and  $R_2^3 = R^2 - R_1^3$ .



**Fig. 2.** Parallel deep learning system for tissues segmentation.

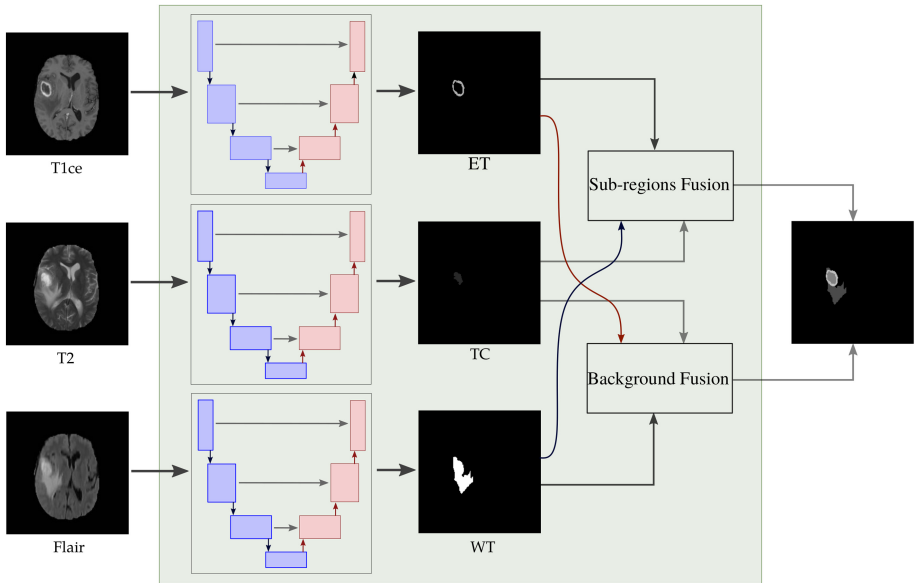
The operation of proposed scheme is quite intuitive, in the first instance any slice of a study must be entered into the system, then a binary segmentation is developed by each U-Net model. That is, all of them have to identify the pixels that correspond to the tissue for which it was trained, and therefore must be able to segment it. After that, the binary segmented images are merged in order to obtain the final segmentation. Two remarks must be stated: (1) Depending on the slice number, the different tissues should appear; in this situation, if the input image does not contain certain specific tissue, the U-Net in charge



of segmenting it will return the corresponding label to the background of the image as a result. (2) If the study corresponds to a healthy patient, then there will be no abnormality or tumor, in the same way as in the previous remark, the result should be the label of the image background. This adaptive capacity of the proposed scheme allows it to be able to segment all slices of a complete medical study, automatically and without human assistance.

### 3.2 Proposed Cascade System for Brain Tumors Segmentation

To address the automatic brain tumor segmentation in multimodal MRI of high-grade (HG) glioma patients, a variation of the proposed previous system should be considered, since the information provided by the tissues is not sufficient to detect and segment correctly a brain tumor. In this regard, other modalities T1, T1 contrast-enhanced, T2 and Flair must be taking into account. How it was stated in [27], Enhancing Tumor (ET) structure is visible in T1ce, whereas the Tumor Core (TC) is in T2, and Whole Tumor (WT) by means of FLAIR. In view of this, the proposal to detect an segment these sub-regions is illustrated in Fig. 3.



**Fig. 3.** Cascade deep learning system for tumor segmentation.

As can be seen, these is an U-Net trained for segmenting a specific glioma sub-region; then, the binary outputs are fused to reshape whole brain tumor, based on three rules stated in Subsect. (3.1). It is necessary to point out that although both methods are based on the same deep neural network, they cannot

be used alternately, since the first model must be trained to identify and segment the soft tissues of the brain, while the second is trained to detect and segment the sub-regions of the gliomas and therefore reconstruct a brain tumor. As is well known, CNNs from the input information but with a specific objective.

## 4 Experimental Setup

### 4.1 Data

In this research paper, it was considered BraTS2017 database. In specific, 210 pre-operative MRI scans of subjects with glioblastoma (GBM/HGG) were used for training and 75 scans of subjects with lower grade glioma (LGG) were used for testing the proposed system. Each study has RMIs in modalities T1, T1ce, T2 and Flair, as well as their respective ground truth images, for each modality there are 155 images of 8-bits with a size of  $240 \times 240$  pixels.

### 4.2 Training

In order to obtain the best results in the test phase it is suggested for BraTS2017 database: a) gray-scale of 8– bits, b) TIFF image format, c) Adaptive Moment Estimation (ADAM) optimization method, d) 1000 epochs and e) learning rate of 0.001.

### 4.3 Evaluation

In order to evaluate quantitative and objectively the image segmentation performance as well as the robustness three metrics were considered in this study. To measure the segmentation accuracy, we used the Misclassification Ratio (MCR), which is given by:

$$MCR = \frac{\text{misclassified pixels}}{\text{overall number of pixels}} \times 100 \quad (32)$$

where, the values can ranges from 0 to 100, a minimum value means better segmentation. Dice Similarity Coefficient is used to quantify the overlap between segmented results with ground-truth; it is expressed in terms of true positives ( $TP$ ), false positives ( $FP$ ), and false negatives ( $FN$ ) as:

$$Dice = \frac{2 \cdot TP}{2 \cdot TP + FP + FN} \quad (33)$$

where  $TP + FP + TN + FN =$  number of brain tissue pixels in a brain MR image. In this metric a higher value means better agreement with respect to ground-truth. In addition to stated metrics, the Intersection-Over-Union (IOU) metric was also considered. This is defined by:

$$IOU = \frac{TP}{TP + FP + FN} \quad (34)$$

The IOU metric takes values in  $[0, 1]$  with a value of 1 indicating a perfect segmentation.

## 5 Results and Discussion

### 5.1 Tissues Segmentation

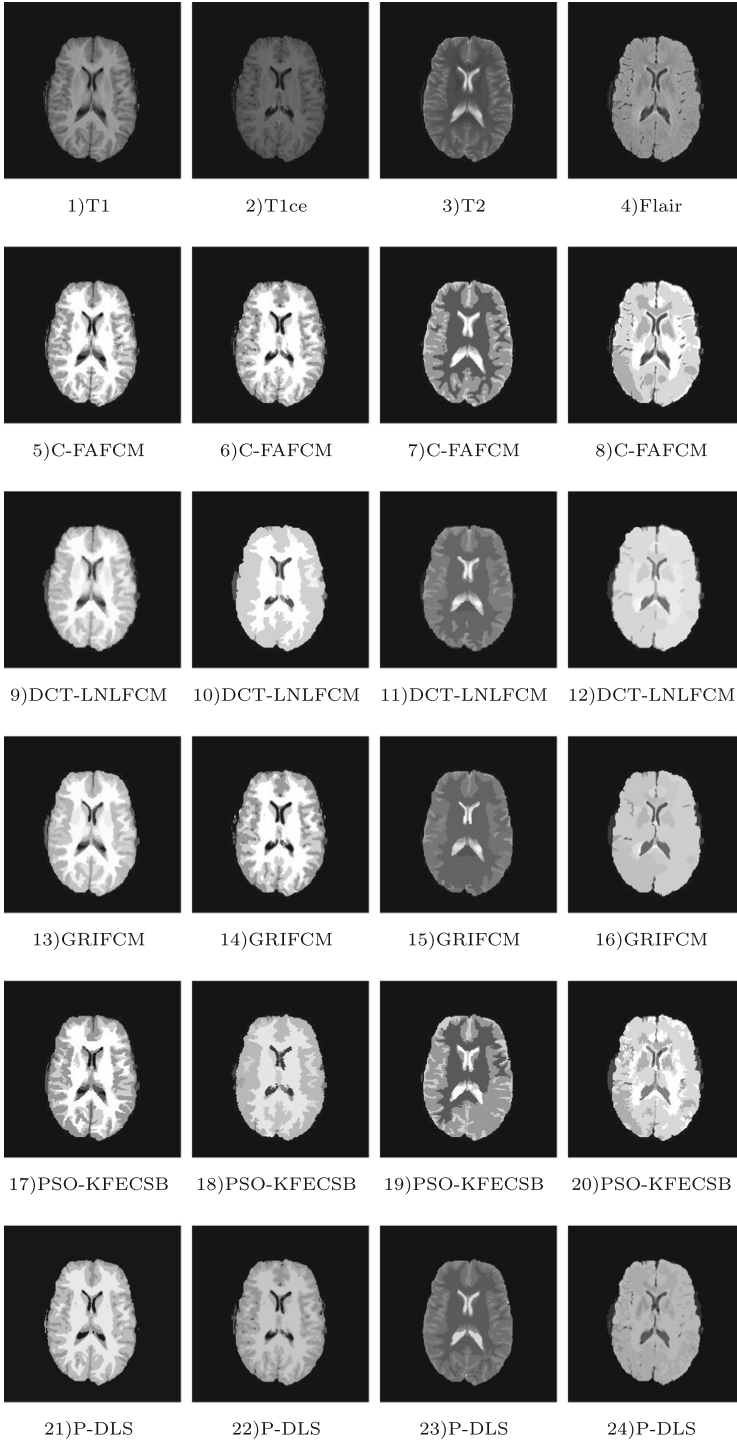
A convincing way to know the true performance of the proposed method is to subject it to the task of tissues segmentation of real brain magnetic resonance images. In this regard, the first experiment is related with the segmentation of images with modalities T1, T1ce, T2 and Flair taken from the BraTS-2017 database; specifically, the Low-grade Gliomas Brats17\_TCIA\_420\_1 study.

The performance of the proposed Parallel and Cascade Deep Learning Systems (for convenience they will be identified as P-DLS and C-DLS, respectively) is compared with other methods designed to segment brain tissues, which were mentioned previously in the introductory section, such as the Chaotic Firefly Integrated Fuzzy C-Means (C-FAFCM) [2], Discrete Cosine Transform Based Local and Nonlocal FCM (DCT-LNLFCM) [4], Generalized Rough Intuitionistic Fuzzy C-Means (GRIFCM) [3], Particle Swarm Optimization - Kernelized Fuzzy Entropy Clustering with Spatial Information and Bias Correction (PSO-KFECSB) [7]. All of them were implemented in the MATLAB R2018a environment, while for ours we used CUDA+CuDNN+TensorFlow+ Keras, that is, conventional frameworks and libraries for Deep Learning, as well as a GPU Nvidia Titan X.

The quantitative evaluation was done considering the MCR, Dice and IOU metrics. A summary of these is presented in Table 1. The numerical results reveal a superior performance of the segmentation method proposed in all the metrics considered, as well as all exposition modalities.

**Table 1.** Average performance on Brats17\_TCIA\_420\_1 study.

Modality	Metric	C-FAFCM	DCT-LNLFCM	GRIFCM	PSO-KFECSB	P-DLS
T1	MCR	8.042	8.991	9.274	10.181	<b>7.132</b>
	Dice	0.899	0.892	0.879	0.827	<b>0.915</b>
	IOU	0.911	0.905	0.880	0.840	<b>0.924</b>
T1ce	MCR	8.937	10.412	10.402	11.698	<b>7.388</b>
	Dice	0.891	0.864	0.853	0.814	<b>0.908</b>
	IOU	0.895	0.868	0.853	0.818	<b>0.911</b>
T2	MCR	9.294	9.500	9.638	10.487	<b>7.491</b>
	Dice	0.886	0.846	0.876	0.805	<b>0.907</b>
	IOU	0.894	0.855	0.877	0.814	<b>0.914</b>
Flair	MCR	9.743	10.797	9.015	11.321	<b>7.621</b>
	Dice	0.872	0.800	0.862	0.783	<b>0.899</b>
	IOU	0.880	0.809	0.863	0.792	<b>0.905</b>



**Fig. 4.** Tissues segmentation sample results of Brats17\_TCIA\_420\_1 study.

A sample image and the segmentation provided by all algorithms evaluated in this experiment are depicted in Fig. 4, it is possible to see that just the proposed algorithm was able to segment images with different modalities. On the other hand, all the other methods presented problems of loss of information in the segmented regions, and in some cases they were not even able to segment the images in the 4 established regions. In brief, a good segmentation of the images in these modalities can guarantee the identification and segmentation of brain tumors.

## 5.2 Tumor Segmentation

Second experiment is related with Multimodal Brain Tumor Segmentation, with the aim to detect and segment the sub-regions: ET, WT and TC. In this respect, the proposed cascade system is used (see Fig. 3). In order to make a comparison of system performance, some methods based on deep learning are considered; particularly, Multi-dimensional Gated Recurrent Units for Brain Tumor Segmentation (MD-GRU-BTS) [10], Masked V-Net (MV-N) [11], 3D Deep Detection-Classification Model (3DDD-CM) [12], Ensembles of Multiple Models and Architectures (EMMA) [13] and Deep U-Net (DU-N) [14]. The quantitative evaluation is made from the studies CBICA\_ATX\_1 and TCIA.639\_1 considering the metrics established previously, a summary of the obtained results is given in Table 2 and Table 3, respectively.

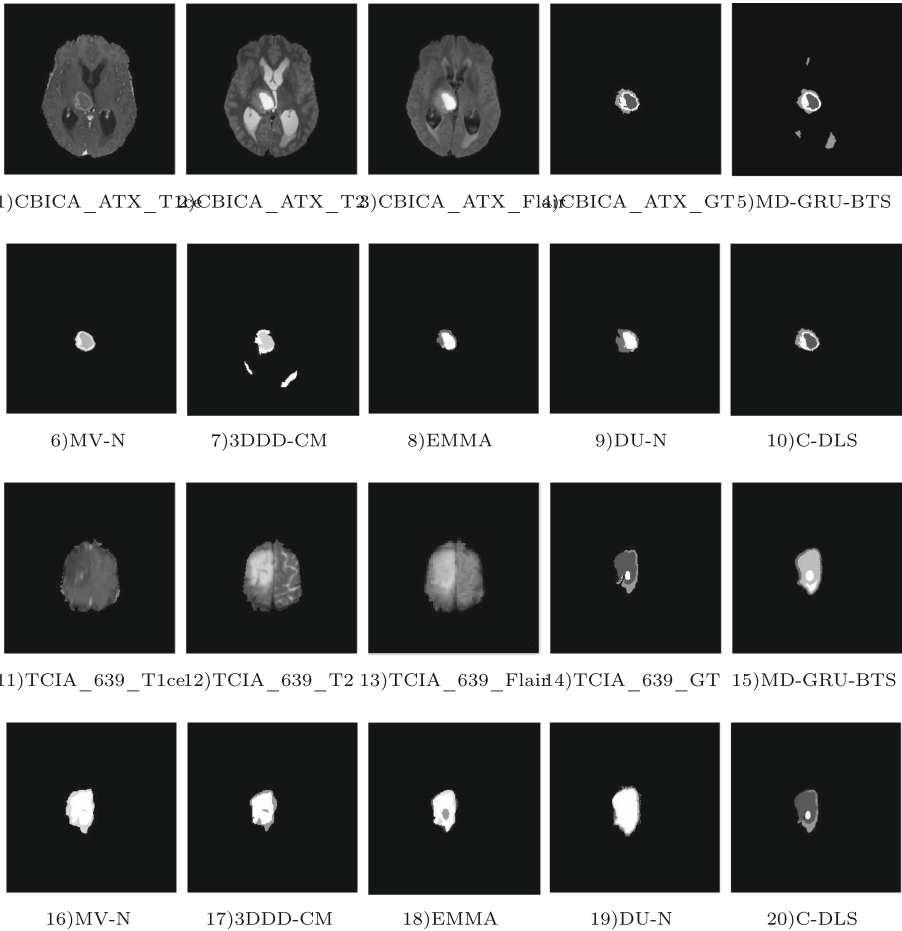
**Table 2.** Average performance on the CBICA\_ATX\_1 study.

Approach	ET			WT			TC		
	MCR	Dice	IOU	MCR	Dice	IOU	MCR	Dice	IOU
MD-GRU-BTS	10.762	0.833	0.872	12.213	0.797	0.822	10.409	0.841	0.856
MV-N	11.531	0.793	0.811	12.639	0.757	0.774	10.756	0.831	0.834
3DDD-CM	14.028	0.768	0.860	14.657	0.734	0.784	12.687	0.783	0.792
EMMA	15.451	0.709	0.745	15.633	0.675	0.702	13.748	0.732	0.740
DU-N	12.553	0.755	0.816	13.855	0.755	0.767	12.457	0.778	0.790
C-DLS	<b>10.182</b>	<b>0.863</b>	<b>0.882</b>	<b>11.511</b>	<b>0.835</b>	<b>0.846</b>	<b>9.585</b>	<b>0.876</b>	<b>0.889</b>

**Table 3.** Average performance on the TCIA.639\_1 study.

Approach	ET			WT			TC		
	MCR	Dice	IOU	MCR	Dice	IOU	MCR	Dice	IOU
MD-GRU-BTS	11.809	0.810	0.842	12.440	0.768	0.776	12.648	0.826	0.843
MV-N	11.934	0.789	0.830	13.279	0.753	0.765	14.010	0.820	0.846
3DDD-CM	13.183	0.747	0.787	15.404	0.704	0.704	13.141	0.771	0.784
EMMA	12.981	0.689	0.704	14.851	0.654	0.681	11.153	0.772	0.819
DU-N	14.238	0.753	0.784	15.901	0.752	0.756	18.270	0.763	0.774
C-DLS	<b>10.227</b>	<b>0.841</b>	<b>0.850</b>	<b>11.877</b>	<b>0.826</b>	<b>0.829</b>	<b>10.119</b>	<b>0.875</b>	<b>0.869</b>

The numerical results reveal a superior performance of the segmentation method proposed in all the metrics considered for three tumor sub-regions. In brief, for CBICA\_ATX.1 study the proposed system obtained  $9.585 \leq MCR \leq 11.511$ ,  $0.835 \leq Dice \leq 0.876$  and  $0.846 \leq IOU \leq 0.889$ ; whilst for other study results showed a slight change, i.e.  $10.119 \leq MCR \leq 11.877$ ,  $0.875 \leq Dice \leq 0.826$  and  $0.829 \leq IOU \leq 0.869$ . Although all the methods evaluated are based on the use of deep neural networks, the fusion of information that we consider in our proposal helped to increase its performance. To ratify and illustrate the quantitative results obtained, it can be seen in Fig. 5, samples 71 of the first study and 127 of the second study, where a segmentation closer to ground-truth can be seen. It can also be seen that some of the comparative methods could not correctly identify and segment the three sub-regions considered.



**Fig. 5.** Tumor segmentation sample results of CBICA\_ATX.1 and TCIA.639.1 studies.

## 6 Conclusions and Future Improvements

Based on the U-Net CNN, it was introduced a parallel system to detect and segment tissues on RMIs with different modalities; considering the same idea, a cascade version was proposed to detect and segment brain tumors. Both systems were enhanced by means three fusion rules in order to do a better job. The first proposal was able to segment different modalities of real RMI images degraded inherent noise densities. Instead, the second proposal was able to detect and segment the ET, WT and TC sub-regions of a brain tumor, with a superior performance than all comparative methods. As future work, we will consider proposing a deep neural network that can perform both tasks simultaneously.

**Acknowledgments.** The authors thank to CONACYT, as well as Tecnológico Nacional de México/Centro Nacional de Investigación y Desarrollo Tecnológico for their financial support trough the project so-called “Controlador Difuso para ajuste de coeficientes de rigidez de un modelo deformable para simulación en tiempo real de los tejidos del hígado humano”.

## References

1. Ganesh, M., Naresh, M., Arvind, C.: MRI brain image segmentation using enhanced adaptive fuzzy K-means algorithm. *Intell. Autom. Soft Comput.* **23**(2), 325–330 (2017)
2. Ghosh, P., Mali, K., Das, S.K.: Chaotic firefly algorithm-based fuzzy C-means algorithm for segmentation of brain tissues in magnetic resonance images. *J. Vis. Commun. Image Represent.* **54**, 63–79 (2018)
3. Namburu, A., Samayamantula, S.K., Edara, S.R.: Generalised rough intuitionistic fuzzy c-means for magnetic resonance brain image segmentation. *IET Image Proc.* **11**(9), 777–785 (2017)
4. Singh, C., Bala, A.: A DCT-based local and non-local fuzzy C-means algorithm for segmentation of brain magnetic resonance images. *Appl. Soft Comput.* **68**, 447–457 (2018)
5. Narayanan, A., Rajasekaran, M.P., Zhang, Y., Govindaraj, V., Thiagarajan, A.: Multi-channeled MR brain image segmentation: a novel double optimization approach combined with clustering technique for tumor identification and tissue segmentation. *Biocybern. Biomed. Eng.* **39**(2), 350–381 (2019)
6. Pham, T.X., Siarry, P., Oulhadj, H.: Integrating fuzzy entropy clustering with an improved PSO for MRI brain image segmentation. *Appl. Soft Comput.* **65**, 230–242 (2018)
7. Senthilkumar, C., Gnanamurthy, R.K.: A Fuzzy clustering based MRI brain image segmentation using back propagation neural networks. *Cluster Comput.* **22**(5), 12305–12312 (2018). <https://doi.org/10.1007/s10586-017-1613-x>
8. Angulakshmi, M., Lakshmi Priya, G.G.: Brain tumour segmentation from MRI using superpixels based spectral clustering. *J. King Saud Univ. Comput. Inf. Sci.* 1–12 (2018). <https://www.sciencedirect.com/science/article/pii/S1319157817303476>

9. Charron, O., Lallement, A., Jarnet, D., Noblet, V., Clavier, J.-B., Meyer, P.: Automatic detection and segmentation of brain metastases on multimodal MR images with a deep convolutional neural network. *Comput. Biol. Med.* **95**, 43–54 (2018)
10. Andermatt, S., Pezold, S., Cattin, P.: Multi-dimensional gated recurrent units for brain tumor segmentation. In: International MICCAI BraTS Challenge. Pre-Conference Proceedings, pp. 15–19 (2017)
11. Marcel, C., et al.: Masked V-Net: an approach to brain tumor segmentation. In: International MICCAI BraTS Challenge. Pre-Conference Proceedings, pp. 42–49 (2017)
12. Hu, Y., Xia, Y.: Automated brain tumor segmentation using a 3D deep detection-classification model, pp. 94–99 (2017)
13. Kamnitsas, K., et al.: Ensembles of multiple models and architectures for robust brain tumour segmentation. In: Crimi, A., Bakas, S., Kuijff, H., Menze, B., Reyes, M. (eds.) *BrainLes 2017*. LNCS, vol. 10670, pp. 450–462. Springer, Cham (2018). [https://doi.org/10.1007/978-3-319-75238-9\\_38](https://doi.org/10.1007/978-3-319-75238-9_38)
14. Kim, G.: Brain tumor segmentation using deep u-net. In: International MICCAI BraTS Challenge. Pre-Conference Proceedings, pp. 154–160 (2017)
15. Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet classification with deep convolutional neural networks. *Commun. ACM* **60**(6), 84–90 (2017)
16. Qi, R., et al.: Pointnet: deep learning on point sets for 3D classification and segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 652–660 (2017)
17. Girshick, R., Donahue, J., Darrell, T., Malik, J.: Region-based convolutional networks for accurate object detection and segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **38**(1), 142–158 (2016)
18. Uijlings, J.R.R., van de Sande, K.E.A., Gevers, T., Smeulders, A.W.M.: Selective search for object recognition. *Int. J. Comput. Vis.* **104**(2), 154–171 (2013). <https://doi.org/10.1007/s11263-013-0620-5>
19. Ren, S., He, K., Girshick, R., Sun, J.: Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **39**(6), 1137–1149 (2017)
20. Sutskever, I., Vinyals, O., Le, V.: Sequence to sequence learning with neural networks. In: Advances in Neural Information Processing Systems, pp. 3104–3112 (2014)
21. Bahdanau, D., Cho, K., Bengio, Y.: Neural machine translation by jointly learning to align and translate. arXiv preprint [arXiv:1409.0473](https://arxiv.org/abs/1409.0473) (2014)
22. Saon G., et al.: English conversational telephone speech recognition by humans and machines. arXiv preprint [arXiv:1703.02136](https://arxiv.org/abs/1703.02136) (2014)
23. Tran, P.V.: A fully convolutional neural network for cardiac segmentation in short-axis MRI. arXiv preprint [arXiv:1604.00494](https://arxiv.org/abs/1604.00494) (2016)
24. Andermatt, S., Pezold, S., Amann, M., Cattin, P. C.: Multi-dimensional gated recurrent units for automated anatomical landmark localization. arXiv preprint [arXiv:1708.02766](https://arxiv.org/abs/1708.02766) (2017)
25. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) *MICCAI 2015*. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). [https://doi.org/10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28)
26. Gong, S., Liu, C., Ji, Y., Zhong, B., Li, Y., Dong, H.: Advanced Image and Video Processing Using MATLAB. MOST, vol. 12. Springer, Cham (2019). <https://doi.org/10.1007/978-3-319-77223-3>



27. Menze, B.H., et al.: The multimodal brain tumor image segmentation benchmark (BRATS). *IEEE Trans. Med. Imaging* **34**(10), 1993–2024 (2015)
28. Mújica-Vargas, D.: Redescending intuitionistic fuzzy clustering to brain magnetic resonance image segmentation. *J. Intell. Fuzzy Syst.* **39**(1), 1097–1108 (2020)
29. Mújica-Vargas, D., Martínez, A., Matuz-Cruz, M., Luna-Alvarez, A., Morales-Xicohtencatl, M.: Non-parametric brain tissues segmentation via a parallel architecture of CNNs. In: Carrasco-Ochoa, J.A., Martínez-Trinidad, J.F., Olvera-López, J.A., Salas, J. (eds.) *MCPR 2019. LNCS*, vol. 11524, pp. 216–226. Springer, Cham (2019). [https://doi.org/10.1007/978-3-030-21077-9\\_20](https://doi.org/10.1007/978-3-030-21077-9_20)