Automatic identification of brain tumors in medical images

George Laurențiu APETRIA
Artificial Intelligence and Optimization
Faculty of Computer Science - UAIC
Iași, 2024

***Abstract*—This paper presents an innovative approach in the context of image classification problems in the medical field. Moreover, the study exposes an additional component to the classification one, namely a segmentation module of the classified images.**

In order to successfully solve the stated problem, I proposed a solution based on the architecture of a convolutional neural network aimed at detection or classification and furthermore, segmentation of brain tumors.

The method involves a robust architecture and consists of two main phases which, in turn, respect a modularized structure, constituting the processing pipelines. In the first phase I used a simple binary annotation of the images from the data set. In this way, the presence or absence of the tumor can be reflected to avoid the use of segmentation-tagged images that require a subjective intervention of a medical specialist. Later, in the second phase of the whole mechanism, the input data consisting of pre-processed images were fed into the deep learning model through which was obtained the final classification. If, following classification, the presence of a tumor is indicated, then it is segmented based on the representation of the features generated by the convolutional neural network. The BraTS 2020 data set containing different types of gliomas is the starting point for training the proposed method. In the following, the obtained results present the performance of the proposed approach in terms of accuracy, precision, recall and Dice similarity coefficient. The developed model achieved a remarkable accuracy value of 96.65% for the classification of images from the test data set, and in addition the Dice similarity coefficient recorded values of up to 89.428% in tumor segmentation.

I. Introduction

A series of approaches based on different methods and techniques aimed at predicting and segmenting tumors were proposed. However, these involve certain disadvantages, among which I mention: the need for the intervention of a specialist in the medical field whose subjectivity can decisively influence the results obtained, the processing time much too long for considerable volumes of data, but also the choice of an appropriate feature extractor in different situations.

In order to overcome the problems related to the subjectivity of human operators, the high processing time and the difficulty of extracting some representative features through classical technical means, I proposed a tumor classification and segmentation approach based on a convolutional neural network architecture without using image masks. In this way, no human influence is exerted on the computing system. The innovative aspect of the approached method consists in the method of training the convolutional neural network. This is done without using labels in the form of segmented images, but in the form of 2 binary annotations (values 0 or 1). It should be noted that the images that were selected to form the training data set were first subjected to a pre-processing step in order to match the input format of the neural network, but also to be able to obtain notable results. In this step, cutting operations (without removing useful information), normalization, and rescaling were performed. The classification of the images is the first task of the work, and its importance is revealed because, following this process, the images that are placed in the ”no tumor” instance class are removed from the processing pipeline.

After predicting the existence of the tumor based on the architecture of the convolutional neural network, the second task followed - tumor segmentation. In this sense, I built for each instance of the class ”with tumor” an image by combining the gradients of the last layer of features. First, the 32 feature maps resulting from the last convolution layer of the network were retrieved, and then the gradients corresponding to the 32 feature maps were computed. Then the calculation of the average, respectively the global maximum value of each gradient was performed and for simplicity only these values were stored in 2 distinct vectors at the expense of the feature maps. I multiplied these vectors with all the components obtained from the last convolution layer (the 32 feature maps) to obtain the masks of the test images, and finally I applied a color map to the images to highlight the region affected by the tumor from the entire cerebral surface. The last stage of the proposed method was post-processing and consisted of two operations: removing small regions around the corners of the MRI scan and correcting the values of some pixels. The last operation was performed by applying a threshold function and the opening and erosion morphological operations.

As mentioned in the abstract, the values of the metrics of interest - accuracy, precision, recall and Dice similarity coefficient - reached remarkable values, comparable to those obtained in other studies of the same type.

The paper is organized in the following sections: in section II are described the related works; in section III the proposed approach to solve the problem is described step by step. In section IV is presented the ablation study. Section V exposes the discussion based on the project. Last but not least, within section VI is stated the conclusion.

II. Related Works

A. Comparative presentation of current achievements on the same topic

The problem of classifying and segmenting tumors located in the brain has been addressed by many researchers who have proposed various ways to solve them based on algorithms from the spectrum of automatic learning and which have proven their effectiveness. The works presented briefly in the following represent one of the main milestones that form the basis of the current research.

Initially, the most known and used classifiers were supervised ones. This could also be influenced by the number of instances used in the training process, a defining criterion for differentiating between the decision tree method - Random Forest and support vector machines, by comparison with convolutional neural networks. Lefkovits et al. [1] developed a model using a Random Forest classifier. After the extraction of the features and the optimal selection of the most representative ones, the extractions of the first-order operators of the characteristics (mean, standard deviation, maximum, minimum, median, Sobel, gradient), higher-order operators (Laplacian, difference of Gaussians, entropy, curvatures, kurtosis – sharpness coefficient, asymmetry), texture features (Gabor filter – linear filter used for texture analysis) and spatial context features. This set of characteristics was finally analyzed in the sense of a decision on the parameters of the algorithm, taking into account the variability of the accuracy of the results obtained.

Szabo et al. [2] designed a new method to be able to segment low-grade gliomas from images obtained on the basis of MRI scans; they extracted 104 morphological features and Gabor wavelet features, and the classifier used was a Random Forest type. Moreover, in order to perform output regularization, the authors performed a neighborhood-based post-processing.

Zhang et al. [3] presented a method that was divided into a sequence of three major steps: pre-processing and feature generation (minimum, maximum, mean, median, gradient, Gabor wavelet features). Afterwards the Random Forest method was trained to classify the normal pixels from the detected instances as positive. Finally, in the post-processing stage, morphological methods were used to regularize the shapes of the detected lesions.

Behadure et al. [4] proposed an innovative method that incorporates the Barkeley waveform transform to convert the spatial shape into the time domain frequency and an SVM classifier.

Ayachi et al. [5] were able to surprisingly implement the segmentation problem in the form of a classification problem. They classified pixels into two categories: normal and abnormal based on several features that take into account intensity and texture. The final step of the classification was done by means of an SVM classifier.

Kwon et al. [6] proposed a segmentation modality whereby all the different tissues in a patient's brain are distinctly highlighted. Thus, they were able to compose a spatial probability map for each tissue type.

Also, Menze et al. [7] used spatial regularization within the probabilistic model. Within the generative probabilistic model were used: an atlas with images containing healthy brain structures and an atlas of latent brain tumors. These were combined, resulting in a series of image sequences, and segmentation of brain tumors was thus possible.

Jayachandran et al. [8] performed a binary classification - the substituted classes being normal and abnormal - for MRI images. Their approach was based on fuzzy logic and involved the use of a hybrid kernel SVM classifier.

Gabor waveforms were the basis of a tumor classification study by Liu et al. [9]. Feature extraction was performed using the previously mentioned concept, and the tumor classification step involved the use of Gabor filters and an SVM classifier.

Long at the forefront of the scientific world, segmentation research is still active, and interest in this technical branch shows a growing trend. The remarkable performances in the analysis and restoration of medical images have been confirmed again in the last period thanks to the studies carried out [10], [11]. The new trend in research centers is pixel-based segmentation, once again proving the importance of deep learning methods [12].

B. Analysis of the types of existing applications in the category of the topic addressed

Convolutional neural networks are currently widely used due to the great advantages brought both from a structural point of view and from the point of view of the results provided and the time required to obtain them. The level of diversity of the implemented models is vast, and in the current section only a part of them will be presented. Considering the fact that the personal choice on how to solve the problem of classification and segmentation falls into the same class of models as those below, the level of performance achieved will be taken into account in the following exposition, the metrics of interest being the accuracy obtained within the data classification process and Dice similarity coefficient for segmentation task.

Lyksborg et al. [13] designed a binary convolutional neural network that enables whole tumor identification. Following its detection, the segmentation result is smoothed, so that later, a multi-class convolutional neural network highlights the subregions of the tumor.

Dataset pre-processing as well as augmentation can be crucial steps when it comes to automatic segmentation. Although this is not common in CNN-based segmentation methods, [14] analyzed the possibility of using intensity normalization and augmentation of the input data in the designed model. Ultimately this proved to be very effective for brain tumor segmentation in MRI images.

The simultaneous use of local features as well as more general contextual features were considered in the architecture of a convolutional neural network by Havaei et al. [15]. Structurally cascading the output of a classical convolutional neural

network and another convolutional neural network allowed the additional information needed for the second component to be obtained.

Also, a cascade architecture was designed by Madhupriya et al. [16]. They used a convolutional neural network and a probabilistic neural network based on a comparison sketch of different models. An important detail from the implementation of their model is the overlapping use of cores of different sizes: 3x3, respectively 7x7.

Zhao et al. [17] proposed a method based on the complete integration of a fully connected convolutional neural network and conditional random fields. This is done at the expense of adopting conditional random fields as a post-processing step of fully connected convolutional neural networks.

Segmentation of multimodal MRI images was the topic addressed by Wang et al [18]. They implemented a cascade architecture of fully connected convolutional neural networks, thus enabling the hierarchical highlighting of segmented regions: whole tumor, tumor core and contrast-enhanced tumor. The decomposition of the original problem is attributed to the cascade structure and in this way the multi-class segmentation is solved by three binary segmentations. Each of the three networks has well-defined characteristics, but the common goal is to reduce the number of false-positive cases. This aspect is achievable thanks to the increased number of layers containing anisotropic and dilated convolution filters, but also thanks to multi-view fusion.

Section segmentation using deep learning models with fully connected neural networks and conditional random fields integrated as recurrent neural networks was the topic addressed by Zhao et al. [19]. This was implemented using axial, coronal and sagittal views on the data and moreover the fusion between the segmentation results obtained in the three different views was achieved.

Dong et al. [20] developed a new fully connected convolutional network for the purpose of performing 2D image segmentation, based on a U-Net architecture [21]. The importance of obtaining notable results led to the use of a comprehensive data augmentation technique, with increased segmentation accuracy being obtained. In addition, in this paper the authors applied a "soft Dice" loss function (directly using the probabilities indicated by the network instead of applying a threshold function on them to convert them into a binary mask).

Hybrid convolutional neural networks have proven in recent years to be able to speed up the training process without significantly affecting the subsequent performance obtained. Therefore, Sajid et al [22] proposed a hybrid convolutional neural network architecture that uses a novel approach. This is patch-based and takes into account both local and general or contextual information. In order to eliminate the overfitting phenomenon, the proposed network deals with the problem by using a dropout regularizer, but also by normalizing batches. Thus, when predicting the output class, the correlation between the accuracy obtained for the training and testing datasets should indicate a valid training of the network.

Thaha et al. [23] developed a hybridized convolutional neural network aimed at tumor segmentation. Optimizing the "loss" was achieved through the Bat algorithm. This automatic segmentation method involves skull stripping and image enhancement techniques used beforehand in the pre-processing phase.

Convolutional neural networks can also be used to process 3D images. Such a model was proposed by Kamnistsas et al. [24]. They used a multi-scale 3D convolutional neural network with a depth of 11 layers. The solution method assumed a more complex architecture composed of two parallel convolutional pathways capable of processing input data at multiple scales. The goal was to achieve a large receptive field for the final classification. The calculation cost was not neglected either, being kept at low values.

A network with an increased number of layers was designed by Mengqiao et al. [25]. They proposed an approach based on a deeper 3D convolutional neural network composed of twenty-two levels. The central idea of the authors was to use an increased number of cascaded convolution layers with small kernels, making it possible to build a deeper convolutional neural network architecture.

Architectures of convolutional neural networks based on the encoder-decoder model represent directions of interest in the field of machine learning. Myronenko et al. [26] designed a model based on an automatic coding architecture. The researchers' decision to add an additional branch to the encoder endpoint was consistent with the need to reconstruct the original image. Thus, the architecture of their model became closer to that of the autoencoder. The additional guidance and the regularization of the coding component due to the limited number of instances in the training data set were the reasons why the self-coding branch was resorted to. It should also be mentioned that in order to better group the characteristics of the encoding component the authors have used the variational autoencoder approach.

Also in machine learning, a new architecture called Residual Cyclic Unpaired Encoder Decoder Network (RescueNet) was released. Such a model was also proposed by Nema et al. [27]. Their goal was brain tumor segmentation, and a specific type of training based on unpaired generative adversarial networks was used to train the RescueNet network. In the post-processing step, a scale-invariant algorithm was chosen to improve the accuracy of the results obtained.

Method	Data set	Dice similarity coefficient
Lyksborg et al. [13]	BraTS 2014	79.9%
Pereira et al. [14]	BraTS 2013	84%
Havaei et al. [15]	BraTS 2012	82%
Wang et al. [18]	BraTS 2017	87%
Zhao et al. [19]	BraTS 2012	80%
Dong et al. [20]	BraTS 2015	86%
Kamnistsas et al. [24]	BraTS 2016	85%
Myronenko et al. [26]	BraTS 2018	81%
Nema et al. [27]	BraTS 2018	94%

TABLE I
SUMMARY OF PERFORMANCE RESULTS OBTAINED IN RELATED WORKS

III. Proposed approach

The proposed method contains two main phases. In the first phase, the preparation (pre-processing) and augmentation of the data set is executed, and then follows the training of the convolutional neural network. The second phase is the testing phase in which the pre-processing of the test image is initially performed, and then the classification is carried out to determine whether the tumor exists and finally the tumor is segmented according to the features extracted from the model. The end of the implementation brings into the spotlight the ways of post-processing the previously obtained segmented images in order to provide results that indicate high values of the metrics of interest (accuracy and Dice similarity coefficient); they should be comparable to those obtained by other methods, both common and uncommon from a structural and methodological point of view, by using data sets similar in structure and type of component data. Each step will be detailed below.

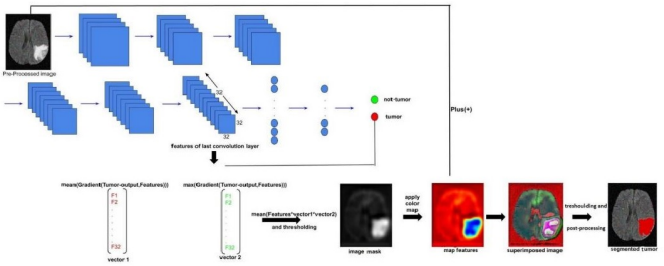

Fig. 1. Illustration of the sequence of processing steps of an image from an MRI scan to obtain tumor segmentation

A. The training data set

The initial data set or rather its import is the starting point in terms of project development. The proposed application was tested and evaluated, using the BraTS 2020 data set provided within BraTS (Brain Tumor Segmentation) competition in conjunction with MICCAI ("Medical Image Computing and Computer Assisted Intervention"). On the occasion of the previously mentioned events, the latest methods and models for solving contemporary are usually launched, including the segmentation of brain tumors.

The data set brings together a number of 369 NifTI (Neuroimaging Informatics Technology Initiative, .nii extension) elements, components that represent nothing more than multimodal MRI scans containing glioms (the most common type of brain tumor) of high level (High Grade Glioma or HGG in number of 293 items) or reduced (Low Grade Glioma or LGG in number of 76 items).

By multimodal images it is understood that there are several channels of information, respectively 4, as can be seen in figure 2. The 4 volumes presented are T1, T1ce (T1 contrast enhanced), T2, respectively FLAIR (Fluid-Attenuated Inversion Recovery), the last image, the one presenting the expected result (ground truth) being made by a specialist in the field of

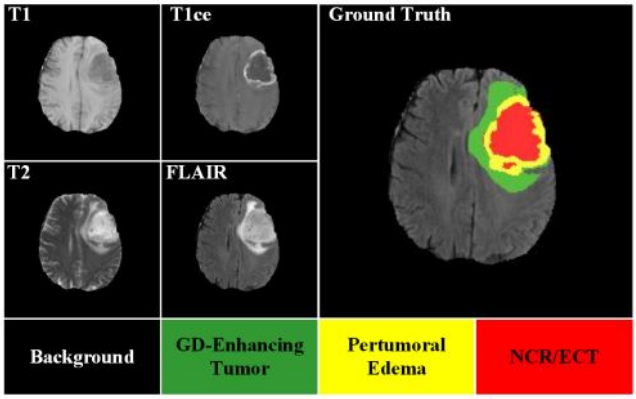

Fig. 2. Highlighting the multimodal structure of data set elements

radiology and showing the manual segmentation of the tumor. Each type of image provides different information about the tumor. The T1ce channel highlights the necrotic area (accumulation of dead cells), but also the breakdown of the blood-brain barrier. Considering the fact that for the T1 channel there is an improved variant, namely T1ce, no more details will be provided related to it. The T2 channel highlights the area of peritumoral edema (accumulation of water or other fluid). The FLAIR channel highlights the area of peritumoral edema, but also areas that do not belong to any category (edema, necrosis or blood-brain barrier). Thus, for training the network, the FLAIR visualization channel was used due to the possibilities of segmenting the tumor in its entirety. Moreover, each FLAIR instance complies with a size standard of 240x240x155, which means a 3D conformation, with multiple viewing possibilities: axial, coronal, sagittal. Considering that the neural network will be designed to process 2D images, only one section will be selected from the initial image, namely the one with the number 90 from axial view as it highlights the most details due to the maximum surface area. Hence, the data set to be augmented and pre-processed gathers a number of 369 elements, each consisting of a 2D image of size 240x240. Before cropping of the images below 192x152 dimensions is performed and thus an enormous number of 0 intensity pixels (black pixels) are eliminated, the useful information not being damaged, a fact guaranteed by the structure images from the original data set.

B. Augmentation phase

Data augmentation is a method of reducing overfitting and generating more training data based on the initial data set. The improvement methods applied to the training data set are: horizontal or vertical flipping, rotation ($\pm 90°$), noise addition, shifting by 20px in vertical/horizontal direction, Gaussian blurring, intensity level changes in the range (0.2, 0.8), image shear with a range limit of 0.2 and zooming in/out in the interval [0.7, 1.0].

Example of data set augmentation methods:

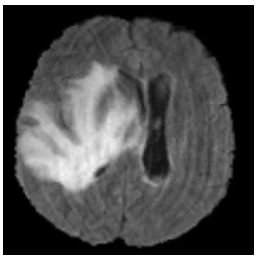

Fig. 3. Original FLAIR image

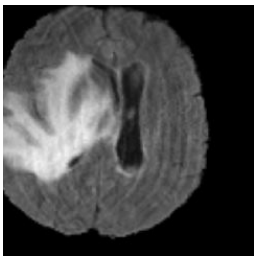

Fig. 4. Shifting the image

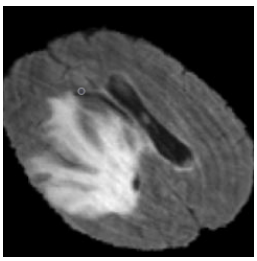

Fig. 5. Rotation in the range +90°

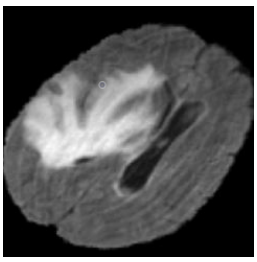

Fig. 6. Rotation in the range -90°

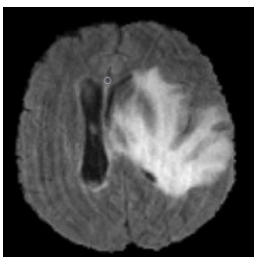

Fig. 7. Horizontally Flip

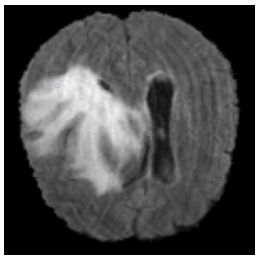

Fig. 8. Vertically Flip

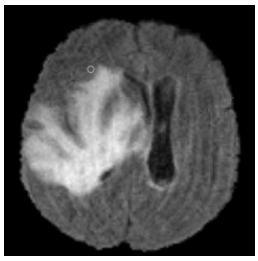

Fig. 9. Adding random noise

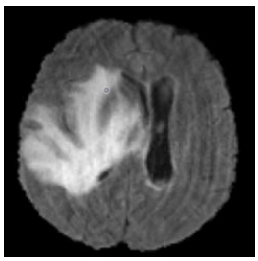

Fig. 10. Gaussian blur

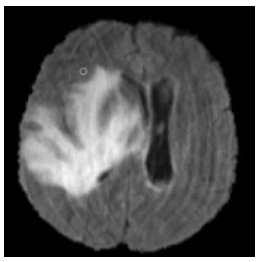

Fig. 11. Inensity change

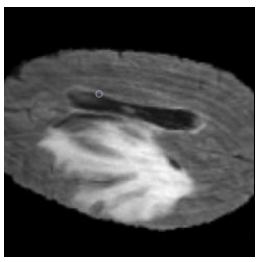

Fig. 12. Image shearing

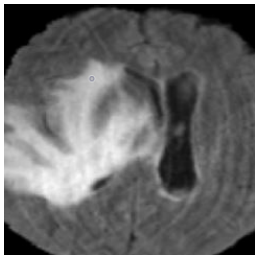

Fig. 13. Zooming in

C. Pre-processing phase

One of the most important steps to be performed in the pre-processing steps is image normalization. This refers to resizing the pixel values so that they are within a narrow range. One of the main reasons why this is done is the propagation of aspect gradients of utmost importance within project. The normalization step involves applying a function to all images in the training data set and consists of subtracting the mean μ from each input image i, the result being divided by the standard deviation σ. After the normalization stage, from a visual point of view, no major changes can be distinguished in the level of the images. Thus, histograms are used to highlight the changes made to the images.

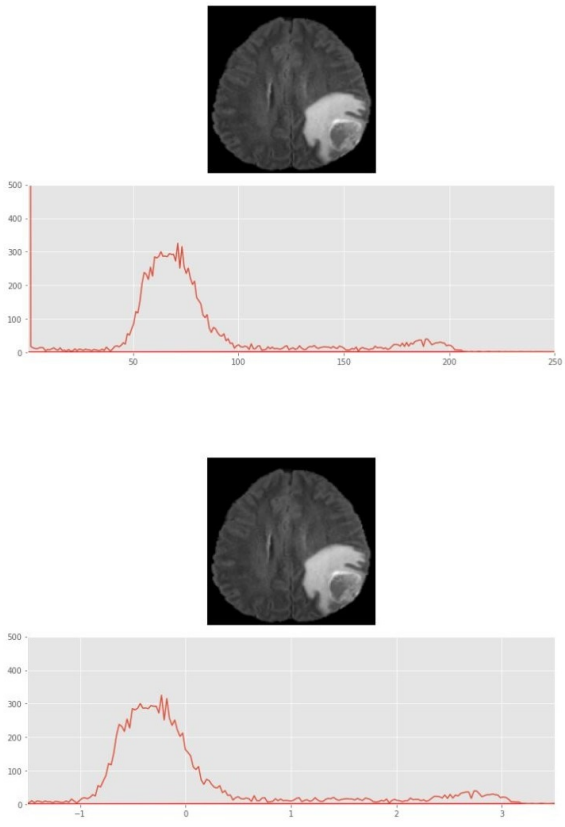

Fig. 14. The results obtained after the normalization of images

D. Convolutional Neural Network

Due to their efficiency, in recent years, convolutional neural networks have become increasingly used in the field of high-precision technical research. The operations performed on the input data are assigned to the neurons that compose the layers of the network arranged in a logical sequence.

The convolutional neural network is trained to make the connection between each image and the corresponding category by detecting a number of feature maps based on results (parameters) that are propagated in order from a layer convolutional to the next.

In this way, the instances obtained in the data augmentation phase are needed. Based on the 2D structure of the image, the local features in different regions of the input maps are determined. Non-linear activation is carried out by means of ReLU, respectively sigmoid functions. The pooling operation is used to reduce the spatial dimension of the representation map by substitution. The purpose of the final fully connected layers is to provide classification given previously obtained features. Due to outstanding processing capabilities, we used a graphics processing unit through which the training process was accelerated, and the required time was 9 time less.

E. Classification component

Obviously, before performing tumor segmentation, its existence should be verified to prevent unncessary consumption of resources. In this regard, I created a model of a convolutional neural network to classify the instances with help of 2 neurons in the output layer. So a binary classification will be made.

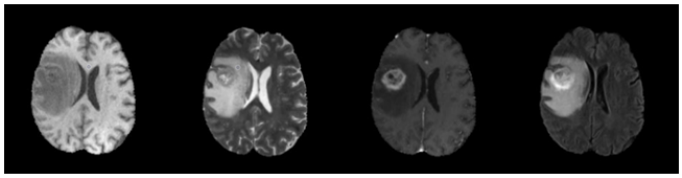

Fig. 15. Different ways of viewing MRI images (from left to right the modalities are displayed: T1, T2, T1ce and FLAIR)

Each type of MRI image is provided through one of the 4 channels indicated. In figure 15 are highlighted varied tissue contrasts in different sequences. Therefore, medical imaging involves a high degree of adaptability being possible to visualize regions of interest in the human brain. Each visualization contains essential information that can lead to improved model performance based on the region of interest. However, considering that we want to segment the tumor in its entirety, we chose the FLAIR visualization modality.

The neural network architecture brings together a number of 18 layers, being designed to primarily achieve 2D image classification. Table 1 shows the main architecture of the convolutional neural network, as well as the values of characteristic dimensions of each integral component. The inputs of the network for the first convolution layer are constituted by the FLAIR modalities, and each subsequent layer takes the feature maps produced by the previous layer as input.

	I	II	III	IV	V	VI
L1	Conv	3x3	1x1	512	-	512x128x128
L2	Activation	-	-	-	-	512x128x128
L3	Max pooling	2x2	2x2	-	-	512x64x64
L4	Conv	3x3	1x1	256	-	256x64x64
L5	Activation	-	-	-	-	256x64x64
L6	Conv	3x3	1x1	128	-	128x64x64
L7	Activation	-	-	-	-	128x64x64
L8	Max pooling	2x2	2x2	-	-	128x32x32
L9	Conv	3x3	1x1	64	-	64x32x32
L10	Activation	-	-	-	-	64x32x32
L11	Conv	3x3	1x1	32	-	32x32x32
L12	Activation	-	-	-	-	32x16x16
L13	Max pooling	2x2	2x2	-	-	-
L14	Flatten	-	-	-	8192	-
L15	FC	-	-	-	32	-
L16	Activation	-	-	-	32	-
L17	FC	-	-	-	2	-
L18	Activation	-	-	-	2	-

TABLE II
TABULAR ILLUSTRATION OF THE MAIN ARCHITECTURE COMPONENT OF THE CONVOLUTIONAL NEURAL NETWORK, WHERE "#" MEANS THE "NUMBER OF" AND "-" MEANS THE ABSENCE OF THE COMPONENT

The meaning of the table column starting from the 2nd column to the right is as follows:

- I - Operation type
- II - Filter dimension
- III - Stride
- IV - Number of filters
- V - Fully Connected units
- VI - Outputs

The network consists of 5 convolutional layers, containing 512, 256, 128, 64, and 32 filters of the same size 3x3, respectively. Within each convolution step, the movement of the window over the input map is carried out with a stride of 2 units and a padding with additional elements of type "SAME" for so that the resulting image has the same dimensions as the input. At the end of each convolution step, I applied the ReLU nonlinear activation function.

Pooling layers are arranged after each convolution and their purpose is to reduce the image sizes. By means of the maximum convolution (max pooling), the maximum value is selected from a window specified using pl x pl dimensions. The pl parameter was configured to the value 2, a step size equal to 2, and the padding use was also of the "SAME" type.

In the flattening layer, the output feature maps of the previous layer are reshaped, and the output neurons are modified to obtain a one-dimensional vector. Considering the configuration of the implemented network, namely the fact that the last aggregation layer provides a matrix of size 32x16x16, i.e. 32 feature maps of size 16x16, a feature vector of size 8192 resulted.

Finally, the last fully connected layer contains 32 features that are used to predict the target class.

In other words, for the classification task we assigned each image a "tumor" or "non-tumor" label based on the expected result (ground truth). This was done to simplify the following segmentation task as it is desired to avoid segmenting images that do not actually contain a tumor. Therefore, it was tried as much as possible to eliminate the detected false positive instances in order to record a good result in the segmentation stage.

F. The segmentation component

Following the classification of the image and the detection of the existence of the tumor, the second phase of the processing method followed, namely tumor segmentation. The vast majority of approaches from the same methodological class regarding tumor segmentation involve the need for the expected result (ground truth), a fact that inherently leads to a subjective intervention by a specialist in the field of radiology. Moreover, all the operations performed by it bring with it a high latency in the detection process and do not guarantee a 100% accurate result.

Unlike the methods of previous generations, the approach proposed in the personal application solves the previously stated problems (the problem of labeled images by the subjective intervention of the medical specialist). This is possible thanks to the capabilities of convolutional neural networks to extract the optimal features from the input data.

The features extracted from the last convolution layer, figure 16, are multiplied by a gradient according to the particularities of these identified features to be able to convey additional information. By means of the obtained information, the pixels that are assumed to be exactly the tumor requiring segmentation are detected, the selection criterion being the light intensity level.

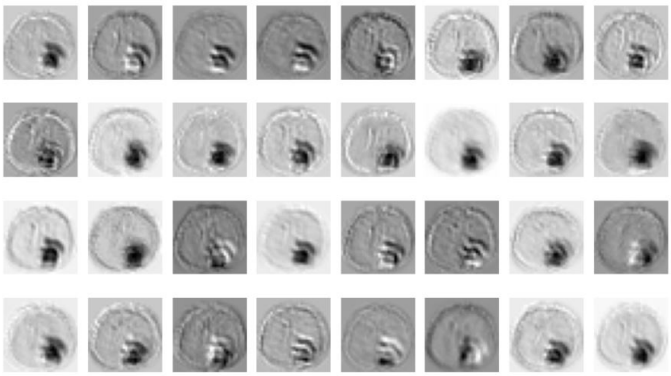

Fig. 16. Examples of tumor features extracted from the last convolutional layer of the CNN

After generating the 32 gradients corresponding to the 32 feature maps extracted from the last convolution layer, we determined the global mean and maximum for each of them. The determined values were stored in 2 vectors of size 32, at the expense of a matrix of size (32, 32, 32) to obtain the importance of the neuron's weights, but also to minimize the feature maps, which will imply a decrease in the time of processing.

We then multiplied each feature map with the previously stored gradients in the specified components-by-component form (component1 x filter1, component2 x filter2, ..., component32 x filter32). This was done to extract more relevant and

detailed spatial information, which would specify the pixels characterized by high intensity values and which will represent the segmented tumor. I thus obtained 32 meaningful images based on the 32 feature maps, but only one final image is required for each initial instance. In this sense, I computed the average of the 32 images on which I applied a treshold function to obtain the mask containing the tumor pixels. Considering the impossibility of determining the intensity changes that appear within an RGB image, I used a grayscale representation of the obtained mask. I than applied a color mask over the original image to highligt the affected area on the brain surface.

Segmentation improvement was finally achieved by applying a treshold function to the image and post-processing operations.

G. Post-processing phase

The post-processing of the images consisted of removing small regions around the corners of the MRI scans and correcting the values of some pixels by applying a treshold function, but also by using a morphological technique: opening, erosion and smoothing.

H. User interface

The interface of application presents a modularized layout structure, independent at each processing stage, but unitary in appearance. The interactivity level is supported by minimal screens in the form of interactive fields which allow the visualization of data after performing the processing. This offers the possibility to receive instant feedback in relation to the processing carried out. Therefore, the application interface encompasses all the elements required for a complete processing flow, as illustrated in Figure 17.

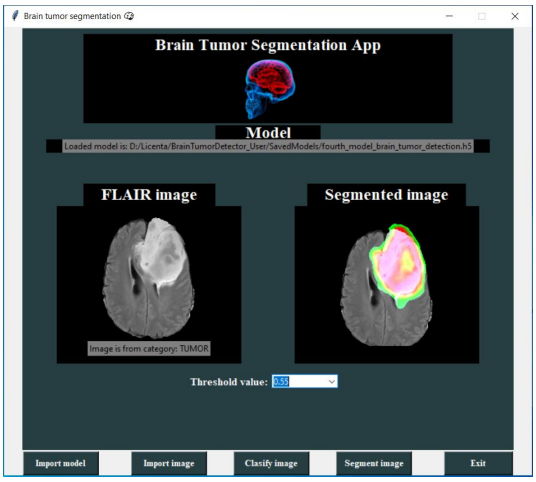

Fig. 17. Application interface

IV. Ablation Study

A. Comparisons in different scenarios, types of data, data regimes

Considering the parameters of the platform on which the application was developed, it must be stated that the resource requirement on which it will be executed in user mode can be much lower. Obviously, the positive influence on the amount of resources is exerted by the lack of the need to train a convolutional neural network – a costly operation in terms of time, energy and processing capacity, being only a trained, already existing model is used. It should also be noted that a computer system with reduced capabilities will provide the desired results in a longer time.

Hardware specifications of the computer used in the development process:

- System manufacturer: LENOVO
- Model: 82B5
- CPU: AMD Ryzen 7 4800H with Radeon Graphics – 16 CPU $\sim$ 2.9 GHz
- RAM: 16.384 MB
- Disk capacity: 476.94 GB
- Graphic card: NVIDIA GeForce GTX 1650 Ti
- Totoal memory approx.: 12.088 MB
- Operating system: Windows 10 Pro 64-bit

There are two possibilities for running the developed application depending on the purpose: for the purpose of further development, the application can be run locally or in a virtual environment that allows the execution of python scripts as well as some extensions necessary for an efficient and pleasant interaction between the system and user. From this category can be specified Project Jupyter and Google Colaboratory, platforms intended for students and researchers, offering computing resources, including GPUs and possibilities of content distribution, all without the need for prior configurations. In the test version, the application can be run locally.

The case of local execution requires extra attention. Version differences between drivers can cause major problems that are hard to detect. Regarding the hardware resources of the station used for the execution of the application, it is desirable that it has parameters as close as possible to those previously indicated. In the impossibility of reaching the mentioned values, it is recommended to comply with the minimum specifications that correspond to the standard indicated by the ”University of Massachusetts Amherst Information Technology”: CPU Intel Core i5 6th generation or newer, memory 8 GB RAM [28].

If the execution of the development application cannot be carried out locally and there is a connection to the Internet network, a virtual environment will be used, the necessary resources being provided to obtain results in a short time. The differences between the online platforms and the station used for development are not noticeable in user mode. The important aspect that must be specified in the given situation is how to configure the data import step, respectively the packages and libraries required at run-time. The running of the application will no longer be possible if the specifications regarding the versions of the libraries that expose functions required in data flow processing are not met.

B. Evaluation of the robustness of the proposed approach in comparison with other implementations

The biggest advantage, obviously, of convolutional neural networks is represented by how to extract features from images. Unlike previous methods, training neural networks can be done using instances that did not require prior processing by a subjective human operator. In fact, the method of training the network implies a double advantage, namely:

- the processing is carried out 100% by computer, and the results cannot be distorted by the actions of a human, who, regardless of the level of training and experience in the medical field, will not be able to perform a perfect segmentation of a tumor;
- the complete elimination of the radiologist involved in tumor segmentation from the entire scientific process, which will greatly shorten the time needed to obtain results.

Moreover, after carrying out a comparative study between convolutional neural networks and those with fully connected layers, there were values indicating a trend of decreasing accuracy for the validation data in the last training epochs of the network with fully connected layers [29]. The dataset used for training the networks was provided by MINST (Modified National Institute of Standards and Technology database) and contained 60,000 28x28 grayscale instances of the 10 digits, and the validation dataset gathered a number of 10,000 images.

For a training duration of 5 epochs, a batch size of 128 elements, and a 70:30 training/test distribution, the following accuracy values were obtained for the convolutional neural network: 99.19% in case of training data set and 99.63% for validation data set. In addition, by comparison with the results indicated after the training of the network with fully connected layers, in figure 18 the stable trend can be observed including during the last training epochs [29].

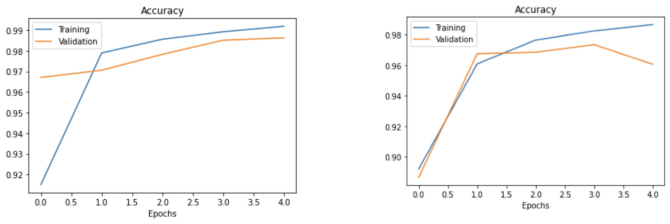

Fig. 18. Comparative exposition of the results obtained following the training process of two different models: convolutional neural network (left), respectively neural network with fully connected layers (right)

V. Discussion based on project

A. Original ideas, new solutions

First of all, the methodology for treating the problem of brain tumor segmentation represents the innovative core of the entire work. Convolutional neural networks stand out in the context of machine learning due to their remarkable data processing capabilities completely independent of any subjective factor.

The original character of the implementation is supported by the complementarity between the methodology approached and the management mode at the hardware level of the processing performed. Given the increased need for convolutional neural networks with regard to the number and diversity of training data sets, the use of a stable and affordable system in terms of hardware resources is required. This is a key criterion to be able to obtain qualitatively consistent results. However, processing using a CPU can be successfully substituted by using a graphics processing component. Thus, the final results are generated in an incomparably shorter time.

B. Presentation of test data/metrics

Figure 21 which illustrates the ROC curve represents one of the most important performance evaluation indicators. The binary classification evaluation indicators. The binary classification evaluation results support the high level of performance of the developed model. The ROC curve underlines the fact that the classification process performed as expected for both the ”with tumor” and the ”without tumor” class instances. The performance achieved in the data classification stage facilitates the transition to the next stage, namely the segmentation stage, taking into account the fact that instances from the ”no tumor” class have been eliminated.

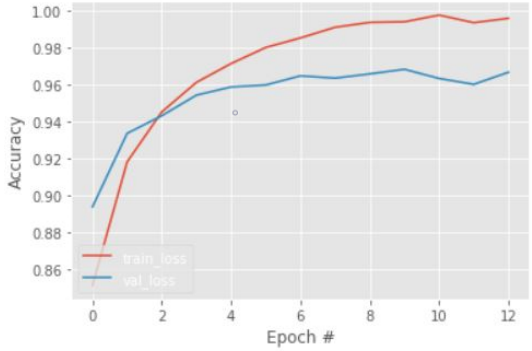

Fig. 19. Accuracy values recorded over several training epochs using the training and validation sets respectively)

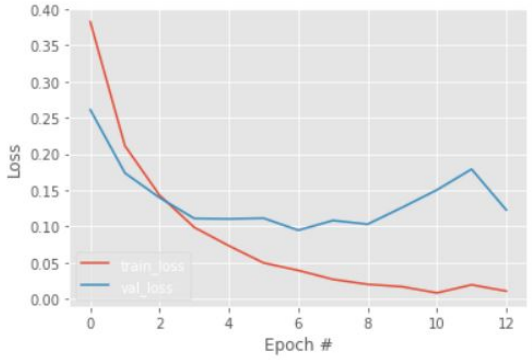

Fig. 20. Loss function values recorded over several training epochs using the training and validation sets respectively)

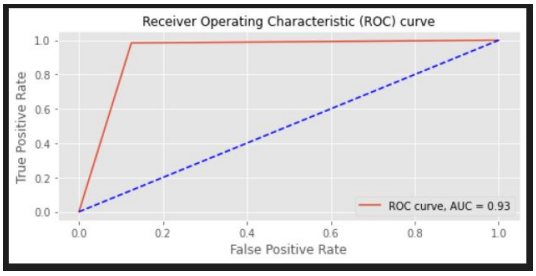

Fig. 21. ROC curve corresponding to the binary model implemented for the purpose of brain tumor detection)

The experimental results were evaluated using different types of performance indicators: precision, recall and accuracy for the classification task, respectively the Dice similarity coefficient for the segmentation task.

- Precision: represents the percentage of relevant results and is defined as follows:

$$Precision = \frac{TP}{TP + FP} \tag{1}$$

- Recall: represents the total percentage of relevant results correctly classified by means of the proposed algorithm and is defined as follows:

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

- Accuracy: formally, accuracy can be defined as:

$$Accuracy = \frac{TP + TN}{Total} \tag{3}$$

TP = True Positive; FP = False Positive; TN = True Negative

- The Dice similarity coefficient: represents the overlap between the segmentation performed using deep learning model and the one performed using the deep learning model and the one performed manually and is computed as follows:

$$Dicesimilaritycoefficient = 2 \cdot \frac{|G \cap S|}{|G| + |S|} \tag{4}$$

where G and S represents the expected output and the output provided by the convolutional neural network, respectively.

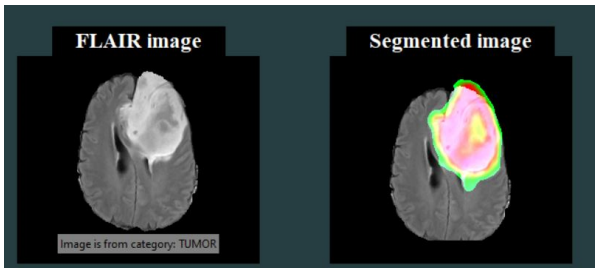

Fig. 22. Final result obtained from segmentation in the test application - instance 1)

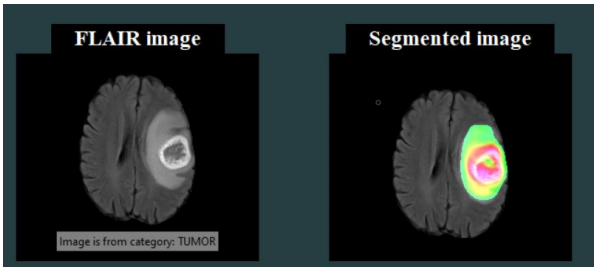

Fig. 23. Final result obtained from segmentation in the test application - instance 2)

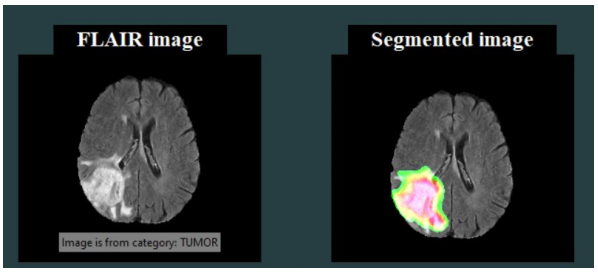

Fig. 24. Final result obtained from segmentation in the test application - instance 3)

VI. Conclusion

A. Highlighting personal contributions/solutions

Personal contributions are revealed in the optimization of the values of the neural network parameters so that it is possible to obtain results comparable to those provided by other specialized studies within the same issue. Moreover, I used GPU processing in a complementary manner, an aspect specified only in certain studies. Thus, the time required to train the neural network was much less, compared to the situation where the CPU would have been used for the same process.

Achieving compatibility between packages required in deployment and testing required additional time and detailed analysis. Considering the diversity of libraries, but also of computing systems, a series of prototype tests were carried out to reach the optimal conformation.

Last but not least, at the level of the user interface, the implementation was carried out entirely by own contribution, using specific tools.

B. Comparison with other similar projects

Notations with Roman numerals denote other studies used for comparison with the personal model:

- I - Single-Path MLDeepMedic [30]
- II - U-NET
- III - RescueNet [27]
- IV - Cascaded Anisotropic CNNs [18]
- V - K-means and FCM [31]
- VI - B.T.S. Based on Deep Learning's Feature Representation [32]
- VII - Proposed personal approach

Method	Data set	Segmentation performance
I	BraTS 2017	DC 79.73%
II	BraTS 2017	DC 80%
III	BraTS 2017	DC 95%
IV	BraTS 2017	DC 87%
V	https://radiopedia.org/	DC 56.4%
VI	BraTS 2017	DC 82.35%
VII	BraTS 2020	DC 89.428%

TABLE III
COMPARATIVE PRESENTATION OF THE VALUES OF THE DICE SIMILARITY COEFFICIENT OBTAINED BY DIFFERENT METHODS BASED ON THE ARCHITECTURE OF A CONVOLUTIONAL NEURAL NETWORK ON DATA SETS PROVIDED WITHIN THE BRATS COMPETITIONS

REFERENCES

[1] L. Lefkovits; S. Lefkovits, M. F. Vaida, "Brain tumor segmentation based on random forest", Mem. Sci. Sect. Rom. Acad. 2016, 1, 83–93.

[2] Z. Szabó, Z. Kapa, L. Lefkovits, A. Gyor, "Automatic segmentation of low-grade brain tumor using a random forest classifier and Gabor features", In Proceedings of the 14th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discove, Huangshan, China, 28–30 July 2018; pp. 1106–1113.

[3] J. Zhang, K. K. Ma, M. H. Er, V. Chong, "Tumor segmentation from magnetic resonance imaging by learning via one-class support vector machine", In Proceedings of the 7th International Workshop on Advanced Image Technology (IWAIT'04), Singapore, 12–13 January 2004; pp. 207–211.

[4] N. B. Bahadure, A. K. Ray, H. P Thethi, "Image analysis for MRI based brain tumor detection and feature extraction using biologically inspired BWT and SVM", Int. J. Biomed. Imaging 2017, 2017, 1–12. [CrossRef] [PubMed]

[5] R. Ayachi, N. Ben Amor, "Brain tumor segmentation using support vector machines", In Symbolic and Quantitative Approaches to Reasoning with Uncertainty, C. Sossai, G. Chemello, Eds. ECSQARU 2009, Lecture Notes in Computer Science, Springer: Berlin/Heidelberg, Germany, 2009, Volume 5590.

[6] D. Kwon, R. T. Shinohara, H. Akbari, C. Davatzikos, "Combining generative models for multifocal glioma segmentation and registration", In Proceedings of the International Conference on Medical, Boston, MA, USA, 14–18 September 2014.

[7] B. H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, "The multimodal brain tumor image segmentation benchmark (BRATS)", IEEE Trans. Med. Imaging 2014, 34, 1993–2024. [CrossRef] [PubMed]

[8] A. Jayachandran, G. K. Sundararaj, "Abnormality segmentation and classification of multi-class brain tumor in MR images using fuzzy logic-based hybrid kernel SVM", Int. J. Fuzzy Syst. 2015, 17, 434–443. [CrossRef]

[9] Y. H. Liu, M. Muftah, T. Das, L. Bai, "Classification of MR tumor images based on Gabor wavelet analysis", J. Med. Biol. Eng. 2012, 32, 22–28. [CrossRef]

[10] D. Shen, G. Wu, H.-I. Suk, "Deep learning in medical image analysis", Annu. Rev. Biomed. Eng. 2017, 19, 221–248. [CrossRef] [PubMed]

[11] A. Qayyum, S. M. Anwar, M. Awais, M. Majid, "Medical image retrieval using deep convolutional neural network", arXiv 2017, arXiv:1703.08472. [CrossRef]

[12] L. Jonathan, E. Shelhamer, T. Darrell, "Fully convolutional networks for semantic segmentation", In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015.

[13] M. Lyksborg, O. Puonti, M. Agn, R. Larsen, "An ensemble of 2d convolutional neural networks for tumor segmentation", In Image Analysis, Springer: New York, NY, USA, 2015, pp. 201–211.

[14] S. Pereira, A. Pinto, V. Alves, C. A. Silva, "Brain tumor segmentation using convolutional neural networks in MRI Images", IEEE Trans. Med. Imaging 2016, 35, 1240–1251. [CrossRef] [PubMed]

[15] M. Havaei, A. Davy, D. Warde-Farley, A. Biard, "Brain tumor segmentation with deep neural networks", Med. Image Anal. 2017, 35, 18–31. [CrossRef] [PubMed]

[16] G. Madhupriya, N. M. Guru, S. Praveen, B. Nivetha, "Brain tumor segmentation with deep learning technique", In Proceedings of the 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, India, 3–5 April 2019, pp. 758–763.

[17] X. Zhao, Y. Wu, G. Song, Z. Li, "Brain tumor segmentation using a fully convolutional neural network with conditional random fields", In International Workshop on Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries, Springer: Cham, Switzerland, 2016, pp. 75–87.

[18] G. Wang, W. Li, S. Ourselin, T. Vercauteren, "Automatic brain tumor segmentation using cascaded anisotropic convolutional neural networks", In International MICCAI Brainlesion Workshop, Springer: Cham, Switzerland, 2017; pp. 178–190.

[19] X. Zhao, Y. Wu, G. Song, Z. Li, "A deep learning model integrating FCNNs and CRFs for brain tumor segmentation", Med. Image Anal. 2018, 43, 98–111. [CrossRef] [PubMed]

[20] H. Dong, G. Yang, F. Liu, Y. Mo, "Automatic brain tumor detection and segmentation using u-net based fully convolutional networks", In Annual Conference on Medical Image Understanding and Analysis, Springer: Cham, Switzerland, 2017, pp. 506–517.

[21] O. Ronneberger, P. Fischer, T Brox, "U-net: Convolutional networks for biomedical image segmentation", In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5–9 October 2015, pp. 234–241.

[22] S. Sajid, S. Hussain, A. Sarwar, "Brain tumor detection and segmentation in MR images using deep learning", Arab. J. Sci. Eng. 2019, 44, 9249–9261. [CrossRef]

[23] M. M. Thaha, K. P. M. Kumar, B. S. Murugan, S. Dhanasekeran, "Brain tumor segmentation using convolutional neural networks in MRI images", J. Med. Syst. 2019, 43, 294. [CrossRef] [PubMed]

[24] K. Kamnitsas, E. Ferrante, S. Parisot, C. Ledig, "DeepMedic for brain tumor segmentation", In Proceedings of the International Workshop on Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries, Athens, Greece, 17 October 2016, pp. 138–149.

[25] M. Wang, J. Yang, Y. Chen, H. Wang, "The multimodal brain tumor image segmentation based on convolutional neural networks", In Proceedings of the 2017 2nd IEEE International Conference on Computational Intelligence and Applications (ICCIA), Beijing, China, 8–11 September 2017.

[26] A. Myronenko, "3D MRI brain tumor segmentation using autoencoder regularization", In International MICCAI Brainlesion Workshop, Springer: Cham, Switzerland, 2018, pp. 311–320.

[27] S. Nema, A. Dudhane, S. Murala, S. Naidu, "RescueNet: An unpaired GAN for brain tumor segmentation", Biomed. Signal Process. Control 2020, 55, 101641. [CrossRef]

[28] C. Misra, J. Mileski, "Recommended & minimum computer configurations for students (Windows)", University of Massachusetts Amherst.

[29] P. Mahajan, "Fully connected vs convolutional neural networks", Medium.

[30] S. Chen, C. Ding, M. Liu, "Dual-force convolutional neural networks for accurate brain tumor segmentation", Pattern Recognit, 2019, 88, 90–100. [CrossRef]

[31] M. A. Almahfud, R. Setyawan, C. A. Sari, D. R. I. M. Setiadi, "An effective MRI brain image segmentation using joint clustering (K-Means and Fuzzy C-Means)", In Proceedings of the 2018 International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), Yogyakarta, Indonesia, 21–22 November 2018, pp. 11–16.

[32] I. Aboussaleh, J. Riffi, A. M. Mahraz, H. Tairi, "Brain tumor segmentation based on deep learning's feature representation", C. Petitjean, MDPI, Journal of Imaging, vol. 7, no. 12, Basel, Switzerland, 8 Dec 2021.