

Article

# A Double-Stage 3D U-Net for On-Cloud Brain Extraction and Multi-Structure Segmentation from 7T MR Volumes

Selene Tomassini , Haidar Anbar, Agnese Sbrollini , MHD Jafar Mortada, Laura Burattini \*   
and Micaela Morettini \* 

Department of Information Engineering, Università Politecnica delle Marche, Via Brecce Bianche 12, 60131 Ancona, Italy; s.tomassini@pm.univpm.it (S.T.); s1101956@studenti.univpm.it (H.A.); a.sbrollini@staff.univpm.it (A.S.); s1101958@studenti.univpm.it (M.J.M.)

\* Correspondence: l.burattini@univpm.it (L.B.); m.morettini@univpm.it (M.M.)

**Abstract:** The brain is the organ most studied using Magnetic Resonance (MR). The emergence of 7T scanners has increased MR imaging resolution to a sub-millimeter level. However, there is a lack of automatic segmentation techniques for 7T MR volumes. This research aims to develop a novel deep learning-based algorithm for on-cloud brain extraction and multi-structure segmentation from unenhanced 7T MR volumes. To this aim, a double-stage 3D U-Net was implemented in a cloud service, directing its first stage to the automatic extraction of the brain and its second stage to the automatic segmentation of the grey matter, basal ganglia, white matter, ventricles, cerebellum, and brain stem. The training was performed on the 90% (the 10% of which served for validation) and the test on the 10% of the Glasgow database. A mean test Dice Similarity Coefficient (DSC) of 96.33% was achieved for the brain class. Mean test DSCs of 90.24%, 87.55%, 93.82%, 85.77%, 91.53%, and 89.95% were achieved for the brain structure classes, respectively. Therefore, the proposed double-stage 3D U-Net is effective in brain extraction and multi-structure segmentation from 7T MR volumes without any preprocessing and training data augmentation strategy while ensuring its machine-independent reproducibility.



**Citation:** Tomassini, S.; Anbar, H.; Sbrollini, A.; Mortada, M.J.; Burattini, L.; Morettini, M. A Double-Stage 3D U-Net for On-Cloud Brain Extraction and Multi-Structure Segmentation from 7T MR Volumes. *Information* **2023**, *14*, 282. <https://doi.org/10.3390/info14050282>

Academic Editor: Amar Ramdane-Cherif

Received: 31 March 2023

Revised: 5 May 2023

Accepted: 8 May 2023

Published: 10 May 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** brain extraction; brain multi-structure segmentation; cloud computing; deep learning; double-stage 3D U-Net; neuroradiology; 7T magnetic resonance; volume measure analysis

## 1. Introduction

Magnetic Resonance (MR) is a radiological imaging modality of pivotal importance in diagnostics. The brain is the organ most frequently studied by MR, as it allows to obtain the greatest sensitivity for the characterization of brain structures by combining radio waves and strong magnetic fields [1,2]. MR imaging is an effective way to diagnose neurological diseases and conditions by detecting structural or connectivity alterations, such as the Grey Matter (GM) atrophy in Alzheimer's disease, the shrinkage of the Brain Stem (BS) structures (e.g., substantia nigra) in Parkinson's disease, the presence of lesions in the White Matter (WM) in multiple sclerosis and the abnormal connectivity of the cortico-cerebellar-striatal-thalamic loop in schizophrenia [2,3]. Nowadays, the advancements in instrumentation technology join improved acquisition methodologies. The emergence of 7T scanners, in particular, has increased the MR imaging resolution to a sub-millimeter level, making the visualization of such brain structures more evident [4,5]. Despite the potentialities, these innovative technologies come with new technical challenges, such as more pronounced radiofrequency field non-uniformities, larger spatial distortion near air-tissue interfaces, and more susceptibility artifacts.

Brain structure segmentation is an important step in MR imaging diagnostics for monitoring the presence of anatomical alterations by isolating specific brain areas and, thus, allowing a region-by-region quantitative analysis [2]. Manual segmentation is the

gold standard for brain structure segmentation in MR [6]. It is necessary for providing the Ground Truth (GT), requiring experienced operators to first define the region of interest and then draw boundaries surrounding it [7]. Although being the most accurate, manual segmentation is only feasible for small collections of data, as it is time-consuming, because performed slice-by-slice, and labor-intensive, due to the noisy yet complex tissue edges [8]. Moreover, its results are difficult to reproduce, as even experienced operators exhibit significant variability with respect to their previous delineations [6,9]. It may also happen that high-resolution MR images, such as 7T ones, no longer have a crisp boundary of the region of interest. As a consequence, slight variations in the selection of pixels may lead to errors [7]. Automatic segmentation techniques have recently aroused great interest for their use in both research and clinical applications. However, most such techniques require labeled MR images obtained through manual segmentation and, thus, experience similar constraints as mentioned earlier. Additional challenges with automatic segmentation techniques include the poor contrast between brain areas, the complex anatomical environment of the brain, and the wide variations in size, shape, and texture found in the brain tissue of subjects. Lack of consistency in source data acquisition may also result in such variations. Consequently, most existing approaches based on clustering, watershed, and machine learning share the problem of a lack of global applicability, which limits their usage to a limited number of applications. Deep Learning (DL)-based algorithms are capable of processing unenhanced data by extracting the salient features automatically, thus eliminating the need for manually-extracted features [10]. DL-based brain structure segmentation seems to be currently the most promising, thanks to the rapid increase of hardware capabilities together with computational and memory resources that have largely reduced the execution time [7,9].

Over the past few years, researchers have reported in the literature on various brain structure automatic segmentation techniques of different accuracy and degrees of complexity [6,8,11–13]. Some researchers, in particular, developed DL-based algorithms for brain structure segmentation from 1.5T and 3T MR volumes, and 3D Convolutional Neural Network (CNN) was the neural architecture used most predominantly. In 2019, Sun et al. [14] proposed a spatially-weighted 3D U-Net for the automatic segmentation of the brain structures into WM, GM, and cerebrospinal fluid from T1-weighted MR volumes of the MRBrainS13 and MALC12 databases, then extended to multi-modal MR volumes. In the same year, Wang et al. [15] proposed a 3D CNN including recursive residual blocks and a pyramid pooling module for the automatic segmentation of the brain structures into WM, GM, and cerebrospinal fluid from T1-weighted MR volumes of the CANDI and IBSR databases. One year later, Bontempi et al. [16] proposed a 3D CNN trained in a weakly-supervised fashion by exploiting a large database collected from the Centre for Cognitive Neuroimaging of the University of Glasgow and composed of T1-weighted MR volumes. Again in 2020, Ramzan et al. [17] proposed a 3D CNN with residual learning and dilated convolution operations for the automatic segmentation of the brain structures into nine different classes, including WM, GM, and cerebrospinal fluid from T1-weighted MR volumes of the ADNI, MRBrainS18, and MICCAI 2012 databases. In 2022, Laiton-Bonadiez et al. [18] injected T1-weighted MR sub-volumes of the Mindboggle-101 database into a set of successive 3D CNN layers free of pooling operations for extracting the local information. Later, they sent the resulting feature maps to successive self-attention layers for obtaining the global context, whose output was later dispatched to the decoder composed mostly of up-sampling layers. However, there is still a severe lack of brain structure automatic segmentation techniques for 7T MR volumes compared to lower field MR volumes. To the authors' best knowledge, the only DL-based algorithm for brain structure segmentation from unenhanced 7T MR volumes is the one proposed by Svanera et al. [19]. They pretrained a 3D CNN on the Glasgow database in a weakly-supervised fashion by taking advantage of training data augmentation strategies. Additionally, they took into account two different collections of data for exploring the condition of limited data availability. However, they directed their research to focus more on the demonstration of the practical

portability of a pretrained neural architecture with a fine-tuning procedure involving very few MR volumes rather than on effective performance evaluation and analysis.

Thus, the focus of this research is on developing a novel DL-based algorithm for on-cloud brain extraction and multi-structure segmentation from 7T MR volumes without any preprocessing and training data augmentation strategy. To this aim, a double-stage 3D U-Net was designed and implemented in a scalable GPU cloud service, directing its first stage to the automatic extraction of the brain by removing the background and stripping the skull, and its second stage to the automatic segmentation of the GM, Basal Ganglia (BG), WM, VENtricles (VEN), CereBellum (CB), and BS.

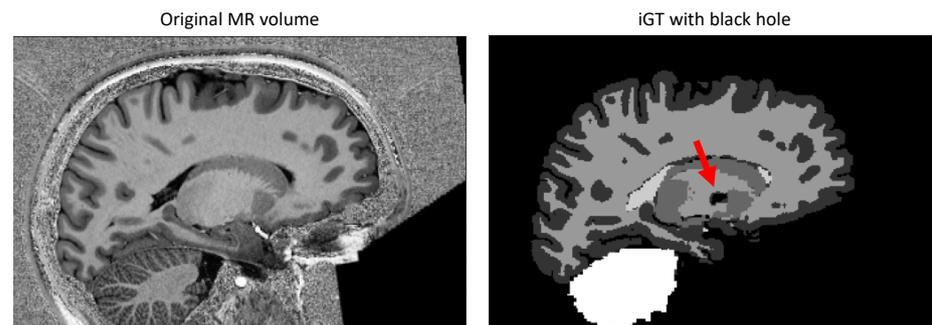
## 2. Data and Methodology

### 2.1. Data Labeling and Division

Data used in this research come from the Glasgow database (<https://search.kg.ebrains.eu/instances/Dataset/2b24466d-f1cd-4b66-afa8-d70a6755ebea>, accessed on 2 January 2023), which was collected at the Imaging Centre of Excellence of the Queen Elizabeth University Hospital in Glasgow and publicly released by Svanera et al. [19]. The database includes 142 out-of-the-scanner T1-weighted MR volumes of  $256 \times 352 \times 224 \text{ mm}^3$ , obtained with an MP2RAGE sequence at  $0.63 \text{ mm}^3$  isotropic resolution, acquired by a Siemens 7T Terra Magnetom scanner with a 32-channel head coil, and belonging to 76 healthy subjects. Neck cropping was the only preprocessing performed by the data providers using the INV2 volume obtained during the acquisition. Together with MR volumes, a multi-class segmentation mask is also included in the database. The segmented classes are (0, 1, 2, 3, 4, 5, 6) for, respectively, the background, GM, BG, WM, VEN, CB, and BS [19]. Once selected, all MR volumes were stored as compressed NIFTI files without applying any further preprocessing.

Due to the considerable time cost and expertise required to produce manual annotations on such a database, the inaccurate GTs (iGTs), made available together with MR volumes by Svanera et al. [19], were exploited, and corrections were then applied. The automatic data labeling procedure accounts for an upper branch dealing with GM and WM automatic segmentation, and a lower branch dealing with BG, VEN, CB, and BS automatic segmentation. In the upper branch, AFNI-3dSeg proposed by Cox et al. [20] was used, followed by geometric and clustering techniques as seen in Fracasso et al.'s [21] research. In the lower branch, FreeSurfer v6 proposed by Fischl et al. [22] was used, with preliminary denoising of MR volumes as in O'Brien et al. [23]. Then, the two branches were combined, and a manual correction was carried out to reduce the major errors (e.g., CB wrongly labeled as GM) using ITK-SNAP as in Yushkevich et al. [24]. In those cases in which the iGTs came with black holes (Figure 1) due to inaccuracies in the automatic data labeling procedure, an additional correction was performed by applying a morphological operation, called dilation, to both increase the object area and fill the black holes. The appropriateness of such a correction was confirmed by an expert neurosurgeon.

Data division was performed not from a data-level perspective but from a subject-level one, being careful not to include data belonging to the same subject in training, validation, and test sets, in order to avoid a biased prediction. Thus, MR volumes were partitioned into 90% (128 MR volumes, 62 subjects) for training, 10% of which served for validation, and the remaining 10% (14 MR volumes, 14 subjects) for testing.



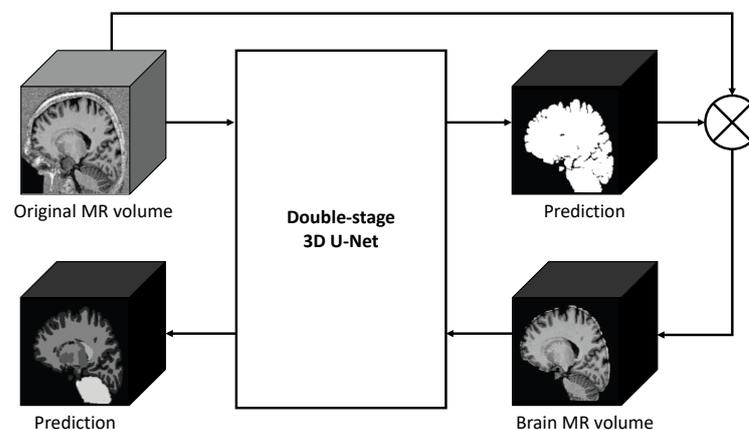
**Figure 1.** Example of one inaccurate Ground Truth (iGT) with a black hole (pointed out by the red arrow) to be corrected.

## 2.2. Double-Stage 3D U-Net

Entire unenhanced 7T MR volumes were processed to take advantage of both the global and local spatial information of MR, conducting the analysis in two learning stages, both accomplished by the same neural architecture, as displayed in Figure 2.

The first learning stage is directed to the automatic extraction of the brain by removing the background and stripping the skull. To fulfill this learning stage, the multi-class segmentation mask was adjusted by giving the background the value of 0 and giving all six brain structures the value of 1. Then, the original MR volumes were injected into the double-stage 3D U-Net.

The second learning stage is directed to the automatic segmentation of the brain structures into GM, BG, WM, VEN, CB, and BS at once. To accomplish this learning stage, the multi-class segmentation mask was kept unaltered. The original MR volumes were multiplied with the predicted masks obtained from the first learning stage, and the resulting brain MR volumes were later injected into the double-stage 3D U-Net.

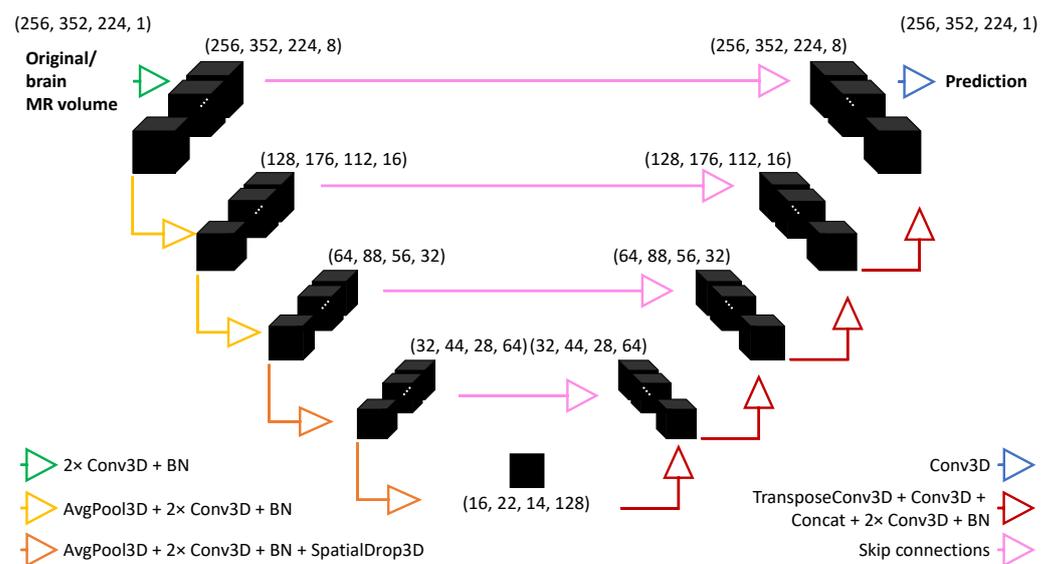


**Figure 2.** Workflow of the double-stage 3D U-Net.

### 2.2.1. Neural Architecture

A double-stage 3D U-Net based on the standard U-Net neural architecture proposed by Ronneberger et al. [25] was designed. Architecturally, it consists of a down-sampling path and an up-sampling path, as depicted in Figure 3. The down-sampling path is made up of five stadiums. The first stadium consists of two  $3 \times 3 \times 3$  Convolution (Conv3D) layers with Rectified Linear Unit (ReLU) activation function followed by a Batch Normalization (BN) layer, used to accelerate the training by reducing the internal covariate shift [26]. The second stadium consists of a 3D Average Pooling (AvgPool3D) layer with a stride of 2, used to look at the complete extent of the input by smoothing it out, thus smoothly extracting the features. The third stadium consists of two Conv3D layers with a ReLU activation function followed by a BN layer. The fourth and fifth stadiums are analogous to the second and third

ones with the only addition of a 3D Spatial Dropout (SpatialDrop3D) layer with a dropout rate of 0.5, in order to reduce the overfitting effect. A SpatialDrop3D layer was introduced in the neural architecture because of the reduced training size, in order to improve the generalization performance by preventing activations from becoming strongly correlated and, thus, avoiding overtraining. The SpatialDrop3D layer, indeed, drops entire 3D feature maps in place of individual elements. If adjacent voxels within 3D feature maps are strongly correlated, a 3D regular dropout will not regularize the activations. The SpatialDrop3D, instead, will help promote the independence between 3D feature maps. The filter sizes of the Conv3D layers in each stadium of the down-sampling path are 8, 16, 32, 64, and 128, respectively. The up-sampling path is made up of five stadiums as well. Differently from the standard U-Net neural architecture of [25], in the first four stadiums, a 3D Transposed Convolution (TransposeConv3D) layer was used in place of the 3D up-sampling layer, followed by one Conv3D layer with ReLU activation function. The TransposeConv3D layer served to up-sample the volumes by increasing the size, height, and width of their inputs. Then, a Concatenation (Concat) layer was added for skip connections, followed by two Conv3D layers with ReLU activation function and a BN layer. The last stadium consists of a Conv3D layer with Softmax activation function, as it assigns probabilities to each class by squashing the outputs to real values between 0 and 1, with a sum of 1 [27]. It has 2 (i.e., [0, 1]) and 7 (i.e., [0, 1, 2, 3, 4, 5, 6]) output neurons for, respectively, the first and second learning stage of the double-stage 3D U-Net. Details of the double-stage 3D U-Net neural architecture are also summarized in Table 1.



**Figure 3.** Neural architecture of the double-stage 3D U-Net. Conv3D refers to 3D Convolution layer, BN refers to Batch Normalization layer, AvgPool3D refers to 3D Average Pooling layer, SpatialDrop3D refers to 3D Spatial Dropout layer, TransposeConv3D refers to 3D Transposed Convolution layer, and Concat refers to Concatenation layer.

**Table 1.** Details of the double-stage 3D U-Net neural architecture. Conv3D refers to the 3D Convolution layer, BN refers to the Batch Normalization layer, AvgPool3D refers to the 3D Average Pooling layer, SpatialDrop3D refers to the 3D Spatial Dropout layer, TransposeConv3D refers to the 3D Transposed Convolution layer, and Concat refers to the Concatenation layer.

Layer	Output Shape	Number of Parameters
Input	(None, 256, 352, 2, 24, 1)	0
Conv3D	(None, 256, 352, 22, 4, 8)	224
Conv3D	(None, 256, 352, 22, 4, 8)	1736
BN	(None, 256, 352, 22, 4, 8)	32
AvgPool3D	(None, 128, 176, 11, 2, 8)	0
Conv3D	(None, 128, 176, 11, 2, 16)	3472
Conv3D	(None, 128, 176, 11, 2, 16)	6928
BN	(None, 128, 176, 11, 2, 16)	64
AvgPool3D	(None, 64, 88, 56, 16)	0
Conv3D	(None, 64, 88, 56, 32)	13,856
Conv3D	(None, 64, 88, 56, 32)	27,680
BN	(None, 64, 88, 56, 32)	128
AvgPool3D	(None, 32, 44, 28, 32)	0
Conv3D	(None, 32, 44, 28, 64)	55,360
Conv3D	(None, 32, 44, 28, 64)	110,656
BN	(None, 32, 44, 28, 64)	256
SpatialDrop3D	(None, 32, 44, 28, 64)	0
AvgPool3D	(None, 16, 22, 14, 64)	0
Conv3D	(None, 16, 22, 14, 128)	221,312
Conv3D	(None, 16, 22, 14, 128)	442,496
BN	(None, 16, 22, 14, 128)	512
SpatialDrop3D	(None, 16, 22, 14, 128)	0
TransposeConv3D	(None, 32, 44, 28, 64)	65,600
Conv3D	(None, 32, 44, 28, 64)	32,832
Concat	(None, 32, 44, 28, 128)	0
Conv3D	(None, 32, 44, 28, 64)	221,248
Conv3D	(None, 32, 44, 28, 64)	110,656
BN	(None, 32, 44, 28, 64)	256
TransposeConv3D	(None, 64, 88, 56, 32)	16,416
Conv3D	(None, 64, 88, 56, 32)	8224
Concat	(None, 64, 88, 56, 64)	0
Conv3D	(None, 64, 88, 56, 32)	55,328
Conv3D	(None, 64, 88, 56, 32)	27,680
BN	(None, 64, 88, 56, 32)	128
TransposeConv3D	(None, 128, 176, 11, 2, 16)	4112
Conv3D	(None, 128, 176, 11, 2, 16)	2064
Concat	(None, 128, 176, 11, 2, 32)	0
Conv3D	(None, 128, 176, 11, 2, 16)	13,840
Conv3D	(None, 128, 176, 11, 2, 16)	6928
BN	(None, 128, 176, 11, 2, 16)	64
TransposeConv3D	(None, 256, 352, 22, 4, 8)	1032
Conv3D	(None, 256, 352, 22, 4, 8)	520
Concat	(None, 256, 352, 22, 4, 16)	0
Conv3D	(None, 256, 352, 22, 4, 8)	3464
Conv3D	(None, 256, 352, 22, 4, 8)	1736
BN	(None, 256, 352, 22, 4, 8)	32
Conv3D	(None, 256, 352, 22, 4, 2/7)	18/63
		Total: 1,456,890/1,456,935
		Trainable: 1,456,154/1,456,199
		Non-trainable: 736

## 2.2.2. Experimental Setup and Learning Process

The double-stage 3D U-Net was developed in Python, exploiting the Pro version of Google Colab to take advantage of the cloud storage and computing power of the

Google servers. The GPU hardware acceleration (NVIDIA Tesla P100 with 16 GB of video RAM) and high system RAM (34 GB) settings were chosen. The Keras library built on a TensorFlow backend was also used.

The combination of both neural network and training parameters that led to the best performance on validation data was considered. Thus, the double-stage 3D U-Net was trained by scratch for 50 epochs, fixing the batch size to 1 and the learning rate to 0.001. The RMSprop was utilized as an optimizer because, during training, it uses an adaptive mini-batch learning rate that changes over time. A combination of Weighted Dice Loss (WDL) and Categorical Cross Entropy (CCE) was used as loss function. The combination of WDL, which is a region-based loss, and CCE, which is a distribution-based loss, allows to simultaneously minimize dissimilarities between two distributions while minimizing the mismatch or maximizing the overlapping regions between desired and predicted outputs [12,28]. Multiple loss functions and a weighting strategy were used here to minimize the problems coming from the highly imbalanced sizes of brain structures. The early stopping callback with a patience of 5 was also used to further minimize the overfitting effect. The weights that led to the lowest validation loss were saved and then used to evaluate the performance of the double-stage 3D U-Net on test data.

### 2.3. Performance Evaluation and Volume Measure Analysis

A comparison between iGTs and predictions was performed to evaluate the performance of the double-stage 3D U-Net on test data. In this guise, the three metrics adopted by the MICCAI MRBrainS18 Challenge were taken into account, being the most commonly used in the context of semantic segmentation, namely Dice Similarity Coefficient (DSC), Volumetric Similarity (VS), and Hausdorff Distance 95% percentile ( $HD_{95}$ ). In addition to the DSC, the Accuracy (ACC), loss, weighted DSC, and mean DSC were monitored in both training and validation phases to provide a better perspective on the behavior of the double-stage 3D U-Net throughout the learning process. Focusing on DSC, VS, and  $HD_{95}$ , DSC is an overlap-based metric useful to predict the similarity index between iGTs and predictions by comparing the pixel-wise agreement between the couple. It is also used as an index of spatial overlap, where a value of 1 indicates the perfect overlap [13]. It is computed as in Equation (1), where L refers to the iGT pixels, and S refers to the prediction pixels:

$$DSC(S, L) = \frac{|S \cap L|}{|S| + |L|}. \quad (1)$$

VS is not an overlap-based metric but rather a measure that considers the volumes of the segments to indicate the similarity [29]. Despite there are several definitions for this metric, VS is typically defined as  $1 - VD$ , where VD is the Volumetric Distance. Mathematically, VS is defined as the absolute volume difference divided by the sum of the compared volumes, as reported in Equation (2), where FN stands for False Negative, FP stands for False Positive, and TP stands for True Positive:

$$VS = 1 - \frac{|FN - FP|}{2TP + FP + FN}. \quad (2)$$

$HD_{95}$  is one of the most commonly used boundary-base metrics, essential for calculating the distance between iGTs and predictions. Basically, it calculates the maximum of all shortest distances for all points from one object boundary to the other. Small values represent a high segmentation accuracy. Specifically, 0 refers to a perfect segmentation (distance of 0 to the reference boundary), and no fixed upper bound exists. It is computed as in Equation (3), where L refers to the iGT and S refers to the prediction [29]:

$$HD_{95} = \max\{K_{s \in S}^{th} \min_{l \in L} \|S - L\|, K_{s \in S}^{th} \min_{l \in L} \|L - S\|\}. \quad (3)$$

In addition, the volume measures of each prediction and corresponding iGT were analyzed for evaluating the goodness of the predictions of the double-stage 3D U-Net

on test data. To do so, the volumes ( $volume_{es}$ ) of each prediction and corresponding iGT were calculated by automatically counting the number of voxels ( $number_{vox}$ ) inside, respectively, the brain, GM, BG, WM, VEN, CB, and BS, and multiplying it by the voxel volume ( $volume_{vox}$ ), expressed in  $cm^3$ , according to Equation (4):

$$volume_{es} = number_{vox} \times volume_{vox}. \quad (4)$$

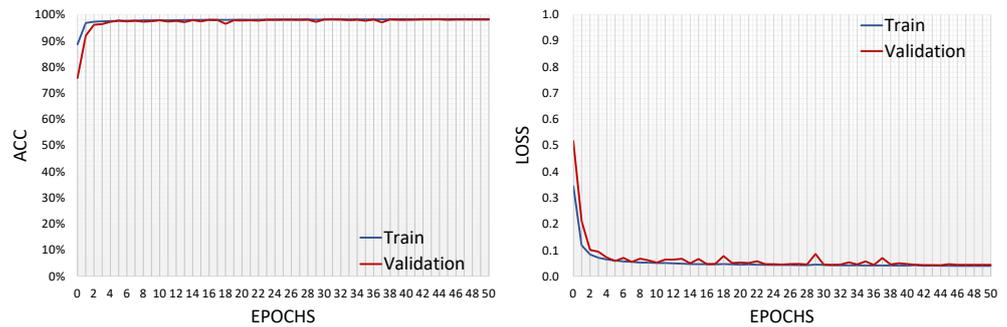
A comparison between the volume distributions computed by the iGT and the prediction related to the same subject was also performed. Specifically, the volume distributions computed by iGTs and predictions were calculated and reported in terms of 50th (median) [25th; 75th] percentiles. Then, non-normal volume distributions computed by iGTs and predictions were statistically compared by means of Mean Absolute Error (MAE, %) and paired Wilcoxon rank-sum test, setting 0.05 as the statistical level of significance ( $P$ ).

### 3. Results

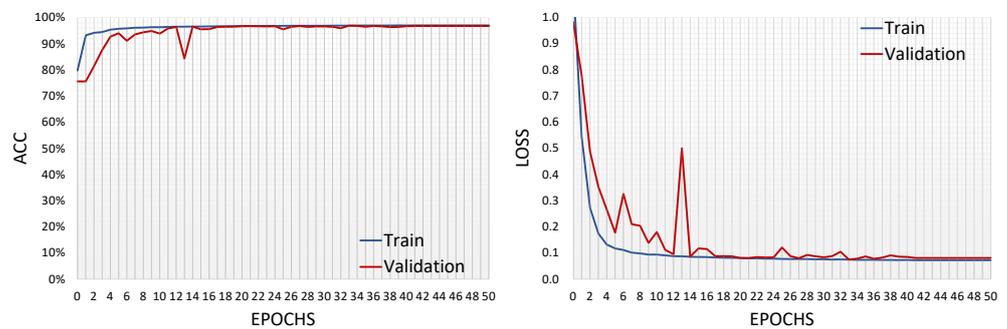
The behavior of the double-stage 3D U-Net throughout the learning process is reported in Table 2, in terms of ACC, loss, weighted DSC, mean DSC, and DSC (computed for each brain structure class). The trend of ACC and loss across the epochs in both training and validation phases is also depicted in Figure 4 for the first learning stage and in Figure 5 for the second learning stage. For the first learning stage, the lowest validation loss value (0.043) was reached in the 44th epoch. For the second learning stage, the lowest validation loss value (0.075) was reached in the 33rd epoch. Due to the highly imbalanced sizes of brain structures in the second learning stage, the trend of weighted DSC and mean DSC across the epochs in both training and validation phases was also monitored, and it is displayed in Figure 6.

**Table 2.** Behavior of the double-stage 3D U-Net in both training and validation phases, in terms of ACCuracy (ACC), loss, weighted Dice Score Coefficient (DSC), mean DSC, and DSC (computed for each brain structure class). GM refers to Grey Matter, BG refers to Basal Ganglia, WM refers to White Matter, VEN refers to VENTricles, CB refers to CereBellum, and BS refers to Brain Stem.

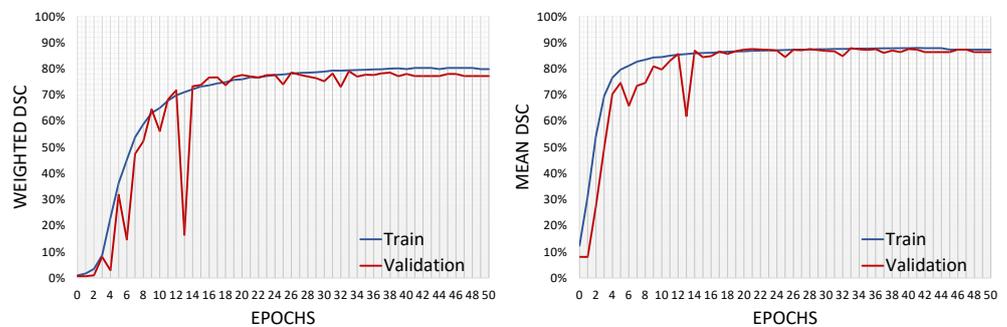
Metrics	Class	First Learning Stage		Second Learning Stage	
		Training	Validation	Training	Validation
ACC (%)	All	98.31	98.24	96.95	96.93
Loss (–)	All	0.04	0.04	0.08	0.08
Weighted DSC (%)	All	-	-	79.41	79.09
Mean DSC (%)	All	-	-	87.63	87.91
	GM	-	-	86.62	87.46
	BG	-	-	80.42	80.54
DSC (%)	WM	-	-	91.46	92.53
	VEN	-	-	82.22	82.05
	CB	-	-	88.81	88.48
	BS	-	-	87.09	86.55



**Figure 4.** Trend of ACCuracy (ACC) and loss across the epochs in both training and validation phases of the first learning stage.



**Figure 5.** Trend of ACCuracy (ACC) and loss across the epochs in both training and validation phases of the second learning stage.

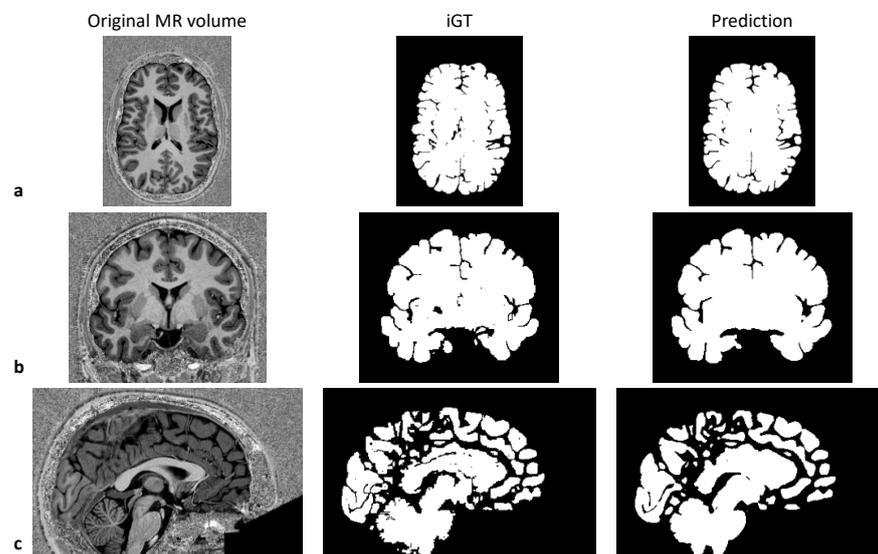


**Figure 6.** Trend of weighted Dice Score Coefficient (DSC) and mean DSC across the epochs in both training and validation phases of the second learning stage.

The test performance of the double-stage 3D U-Net in the automatic extraction of the brain (i.e., first learning stage) and segmentation of the brain structures into six different classes at once (i.e., second learning stage) is reported in Table 3. Since unenhanced 7T MR volumes reserved for testing are the same 14 for both learning stages, DSC, VS, and HD<sub>95</sub> values of the eight total classes (background, brain, GM, BG, WM, VEN, CB, and BS) were computed and expressed as mean ± standard deviation. The qualitative outcome of the double-stage 3D U-Net in the automatic extraction of the brain and segmentation of the brain structures into GM, BG, WM, VEN, CB, and BS at once is provided in Figure 7 and Figure 8, respectively. The qualitative outcome of the double-stage 3D U-Net in the automatic segmentation of the above-mentioned brain structure classes is also displayed in Figure 9, with a different color for each predicted class. Eventually, the volume measures of each prediction and corresponding iGT are reported in Table 4, together with the volume distributions computed by the iGT and the prediction related to the same subject.

**Table 3.** Test performance of the double-stage 3D U-Net in automatically extracting the brain and segmenting its structures into Grey Matter (GM), Basal Ganglia (BG), White Matter (WM), VENTricles (VEN), CereBellum (CB), and Brain Stem (BS) at once, in terms of Dice Score Coefficient (DSC), Volumetric Similarity (VS), and Hausdorff Distance 95% percentile (HD<sub>95</sub>). Values are reported as mean  $\pm$  standard deviation.

Class	Learning Stage	DSC (%)	VS (%)	HD <sub>95</sub> (mm)
Background	First	98.78 $\pm$ 0.22	99.75 $\pm$ 0.25	2.74 $\pm$ 0.68
Brain	First	96.33 $\pm$ 0.51	99.27 $\pm$ 0.67	3.36 $\pm$ 0.54
GM	Second	90.24 $\pm$ 1.04	98.61 $\pm$ 1.33	1.15 $\pm$ 0.21
BG	Second	87.55 $\pm$ 0.83	94.88 $\pm$ 1.82	2.94 $\pm$ 0.31
WM	Second	93.82 $\pm$ 0.87	98.38 $\pm$ 1.51	1.03 $\pm$ 0.11
VEN	Second	85.77 $\pm$ 4.16	96.91 $\pm$ 2.11	2.15 $\pm$ 0.94
CB	Second	91.53 $\pm$ 1.96	96.87 $\pm$ 2.05	5.93 $\pm$ 1.73
BS	Second	89.95 $\pm$ 2.63	97.46 $\pm$ 1.36	2.92 $\pm$ 0.91

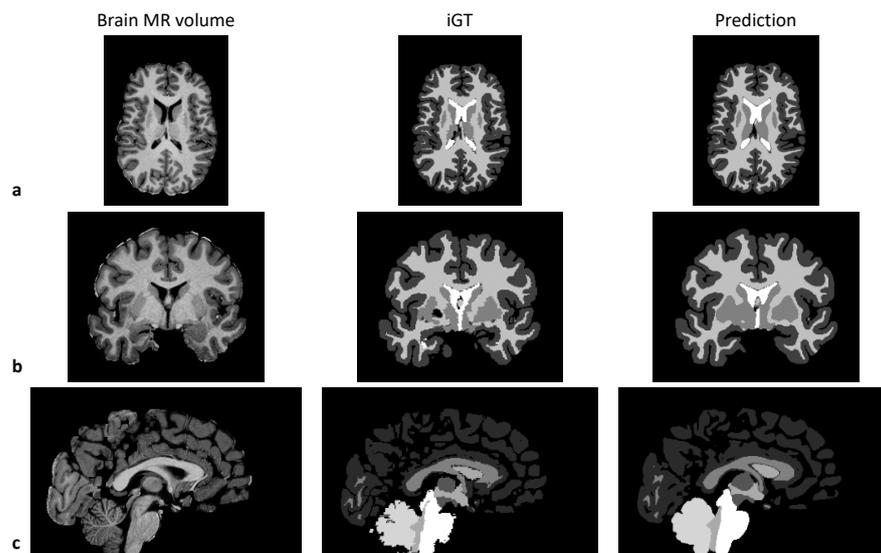


**Figure 7.** Qualitative outcome of the double-stage 3D U-Net in the automatic extraction of the brain. Mid-axial (a), mid-coronal (b), and mid-sagittal (c) slices of original MR volume together with corresponding inaccurate Ground Truth (iGT) and prediction are reported.

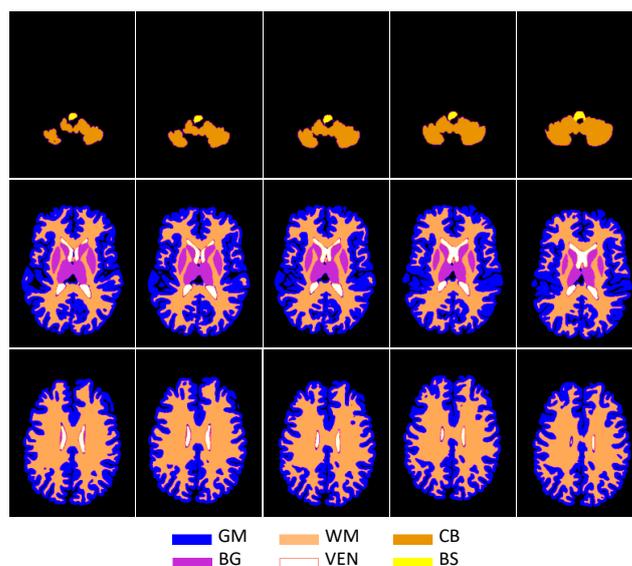
**Table 4.** Measures of the automatically extracted volume of the brain and segmented volumes of the Grey Matter (GM), Basal Ganglia (BG), White Matter (WM), VENTricles (VEN), CereBellum (CB), and Brain Stem (BS) in terms of 50th (median) [25th; 75th] percentiles. The Mean Absolute Error (MAE, %) between distributions of volumes computed by prediction and corresponding iGT is also reported.

Class	Learning Stage	iGT (cm <sup>3</sup> )	Prediction (cm <sup>3</sup> )	MAE (%)
Brain	First	1269 [1152; 1312]	1253 [1162; 1313]	1.02 [0.83; 1.73]
GM	Second	624 [581; 663]	630 [589; 672]	2.11 [0.55; 3.66]
BG	Second	46 [42; 47]	50 [48; 54] *	11.72 [8.69; 14.29]
WM	Second	444 [385; 461]	421 [386; 446]	2.45 [1.28; 4.39]
VEN	Second	16 [15; 21]	16 [15; 21]	6.56 [0; 8]
CB	Second	109 [104; 115]	110 [106; 113]	5.83 [4.27; 10]
BS	Second	17 [16; 18]	17 [16; 18]	5.72 [5.26; 7.14]

\*:  $p < 0.05$  (paired Wilcoxon rank-sum test).



**Figure 8.** Qualitative outcome of the double-stage 3D U-Net in the automatic segmentation of the brain structures into six different classes at once. Mid-axial (a), mid-coronal (b), and mid-sagittal (c) slices of original MR volume together with corresponding inaccurate Ground Truth (iGT) and prediction are reported.



**Figure 9.** Qualitative outcome of the double-stage 3D U-Net in the automatic segmentation of the Grey Matter (GM), Basal Ganglia (BG), White Matter (WM), VENtricles (VEN), CereBellum (CB), and Brain Stem (BS), with a different color for each predicted class, on a bunch of axial MR slices (15 out of total 224) of the same MR volume.

#### 4. Discussion

In this research, a novel DL-based algorithm for on-cloud brain extraction and multi-structure segmentation from unenhanced 7T MR volumes was developed by taking advantage of a double-stage 3D U-Net, the first stage of which was directed to automatically extract the brain by removing the background and stripping the skull, and the second served for the automatic segmentation of the GM, BG, WM, VEN, CB, and BS.

During the learning process, the ACC increased smoothly till it reached a value above 98% (first learning stage) and 97% (second learning stage) while the loss decreased steadily and under a value of 0.04 and 0.08 in both training and validation phases of the double-stage 3D U-Net without going to either overfitting or underfitting (Table 2, Figures 4 and 5).

The reasons for the small decrease in both training and validation ACC (Table 2) from, respectively, 98.31% and 98.24% (first learning stage) to, respectively, 96.95% and 96.93% (second learning stage) rely on the different segmentation tasks. In the first learning stage of the double-stage 3D U-Net, the task was to automatically extract the brain, thus automatically segmenting the entire MR volume into two major classes (background and brain). In the second learning stage of the double-stage 3D U-Net, instead, the task was to automatically segment the entire brain volume (as extracted in the first learning stage) into six further brain structures (GM, BG, WM, VEN, CB, and BS). Here, the complexity of the six brain structures, due to both overlapping and interference, made their segmentation much harder, especially at the boundaries and edges, which led to an expected yet quite low decrease in both training and validation ACC (−1.38% and −1.33%, respectively). Because of imbalanced data, the trend in weighted DSC and mean DSC was also monitored for the second learning stage (Table 2 and Figure 6). When monitoring weighted DSC and mean DSC in both training and validation phases of the double-stage 3D U-Net, the DSC of the GM, BG, WM, VEN, CB, and BS were also taken into account to make sure that all six brain structures were getting automatically segmented properly. All six brain structures were automatically segmented with a DSC higher than 80%, and the brain structures with the highest DSC are the WM, CB, and GM in both training and validation phases of the double-stage 3D U-Net (Table 2). These three brain structures are the ones that were segmented with the highest DSC also in the test phase of the double-stage 3D U-Net (Table 3). All eight total classes achieved DSCs higher than 86%, VS values higher than 95%, and HD<sub>95</sub> values lower than 6 mm, and the brain structures with the highest DSC, highest VS, and lowest HD<sub>95</sub> in the test phase of the double-stage 3D U-Net are the WM and GM (Table 3). The reason may rely on the fact that the strong T1 contrast that is present between fluid and more solid anatomical structures is likely to make the delineation of brain structures such as WM and GM easier [27,30]. Eventually, the analysis of volumes revealed the goodness of the predictions of the double-stage 3D U-Net, considering that all classes (except for the BG class) present volume distributions computed by predictions that are not statistically different from volume distributions computed by iGTs, with median MAE lower than 10%. As for the BG class, the reason for the statistical difference between volume distributions computed by predictions and corresponding iGTs may rely on its complex structure and specific location in the brain, which makes its delineation challenging. Moreover, in high-resolution MR images like 7T ones, BG no longer has a crisp boundary and, thus, slight voxel variations may lead to errors [7]. However, the associated median MAE is lower than 12% and, thus, still very close to 10%.

The proposed double-stage 3D U-Net for on-cloud brain extraction and multi-structure segmentation digests entire 7T MR volumes at once, avoiding the drawbacks of the tiling process. Moreover, it preserves the two-scale (i.e., global and local) analysis (that is a peculiar characteristic of manual segmentation [19]). Instead, the publicly-available non-DL tools for brain structure automatic segmentation, such as Statistical Parametric Mapping 12 (SPM12, [www.fil.ion.ucl.ac.uk/spm](http://www.fil.ion.ucl.ac.uk/spm) accessed on 31 March 2023) [31] and FMRIB Software Library (FSL, [www.fmrib.ox.ac.uk/fsl](http://www.fmrib.ox.ac.uk/fsl) accessed on 31 March 2023) [32], emulate these two steps using atlases to gain global clues and, for most of them, gradient methods for the local processing. In SPM12, the brightness information and position of voxels, along with tissue probability maps, are considered, and the construction of appropriate priors is recommended. In FSL, the GM, WM, and cerebrospinal fluid segmentation is performed using an FMRIB Automated Segmentation Tool, which works on the extracted brain and uses the Markov random field model along with the expectation-maximization algorithm. Instead, the subcortical segmentation is performed using an FMRIB Integrated Registration and Segmentation Tool, which provides a deformation model-based segmentation. SPM12 and FSL are sufficiently resilient with respect to noise and artifacts introduced at the acquisition stage, and have performed consistently across different collections of data [6,8]. However, they are still far from being accepted at par with manual segmentation. Firstly, the quality of these approaches is limited by the accuracy of the pairwise registration

method [8]. Secondly, image contrast, gross morphological deviation, high noise levels, and high spatial signal bias may lead to erroneous segmentation of brain structures [6]. Relying on priors, SPM12 and FSL are prone to erroneous results or simply fail in the presence of abnormal contrast or gross morphological alterations [8]. Thirdly, image artifacts due to poor subject compliance may systematically skew the results [8]. Eventually, despite the fact that such automated tools return a higher number of brain structures, it is often the case that, depending on the final application, having too many labels is not always useful, thus re-clustering is needed. Moreover, the higher the number of labels, the less accurate the available segmentation. Therefore, DL-based brain structure segmentation turns out to be more useful for MR protocols that are missing proper anatomical data or in those cases in which it is difficult to achieve high-quality anatomical structure registration, thanks to its applicability directly to data together with its property to automatically and adaptively learn spatial hierarchies of image features from low- to high-level patterns [33].

In the literature, Svanera et al. [19] were the only researchers to develop a DL-based algorithm for brain structure segmentation from unenhanced 7T MR volumes, as addressed in Section 1. Since they mentioned the results only graphically while reporting the overall test performance across the classes and used a different data division protocol, a fair quantitative comparison between their findings and the ones achieved in this research could not be provided, but a methodological comparison could be performed. Like Svanera et al. [19], a neural architecture able to deal with the full spatial information contained in the analyzed data was designed in this research, as 3D neural architectures can find voxel relationships in the three anatomical planes, thus maximizing the use of the intrinsic spatial nature of MR imaging. The Glasgow database was considered, and entire unenhanced 7T MR volumes, instead of sub-volumes, were processed. Additionally, the iGT automatic procedure was exploited for data labeling, although an additional neurosurgeon-approved correction was performed in this research to eliminate the inaccuracies (i.e., black holes). Differently from Svanera et al. [19], who pretrained a 3D CNN on the Glasgow database in a weakly-supervised fashion to demonstrate its practical portability with a fine-tuning procedure involving very few MR volumes of two different collections of data, this research was directed toward the effective performance evaluation and analysis of the proposed double-stage 3D U-Net by monitoring its behavior throughout the learning process (Table 2, Figures 4–6), assessing its test performance on eight total classes (Table 3, Figures 7–9), and evaluating the goodness of the predictions by means of a volume measure analysis (Table 4). The 3D U-Net was called ‘double-stage’ because the analysis was conducted in two learning stages for the automatic extraction of the brain and segmentation of the brain structures into six different classes at once, respectively. The 3D U-Net was also customized with respect to the standard 3D U-Net neural architecture of Ronneberger et al. [25] by adding two SpatialDrop3D layers in the fourth and fifth stadiums of the down-sampling path, in order to lift the generalization performance. It has been found that the SpatialDrop3D layer contributed to improving the performance without the need for any training data augmentation strategy by extending the dropout value across the entire feature map. Moreover, a TransposeConv3D layer was used in the first four stadiums of the up-sampling path because it is a convolution layer and, thus, it has trainable kernels. It has also been found that the combination of WDL and CCE as loss functions helped in overcoming the problem of data imbalance. In addition, the proposed double-stage 3D U-Net was developed in a scalable GPU service running entirely in the cloud. To allow so, a publicly-available database, already compliant with ethical and regulatory issues, was used. The uploading of entire 7T MR volumes in the compressed NIFTI format took just a few seconds. Although the training of the proposed double-stage 3D U-Net took approximately 6 hours to be completed, only 10 to 20 s served to generate the predictions on test data, which is a perfectly acceptable time in terms of execution efficiency. Moreover, no special network bandwidth requirements were necessary for on-cloud training, validating, testing, and visualizing. Cloud computing was chosen in this research because one of the main challenges facing radiological imaging analysis is the development of benchmarks that allow

methodologies to be compared under common standards and measures. Thus, the cloud can contribute to creating such benchmarks [34]. Additionally, the scalable and distributed computational resources of cloud computing have the potential to increase the execution speed while keeping the costs low [35]. The pivotal component of the cloud, in fact, is the analysis platform, which supports a wide spectrum of data queries and cost-effectiveness solutions without the surcharge of purchasing and maintaining additional setups. In case the proposed double-stage 3D U-Net would need to be tested on a private collection of data, the neural network weights could be exploited to produce predictions in a local environment, with no need for defacing data/anonymizing facial features. Furthermore, if more brain structures other than the ones segmented in this research would need to be analyzed, the training of only the second learning stage of the proposed double-stage 3D U-Net could be re-performed after simply modifying the number of output classes (and, eventually, the loss function weights).

In this research, unenhanced 7T MR volumes were processed by performing no preliminary quality check, and the proposed double-stage 3D U-Net was demonstrated to be effective in brain extraction and multi-structure segmentation from raw, noisy input data. Accordingly, it can be assumed that it is highly generalizable to the quality of data, even though this cannot yet be demonstrated due to the lack of availability of suitable MR volumes.

In fact, MR volumes with annotations of seven classes are really difficult to achieve, and the results found can state only to the treated database. However, being a DL-based brain extraction and multi-structure segmentation algorithm, it is likely to generalize well to heterogeneous data coming from scanners of different manufacturers and/or acquired with 3T and lower field strengths, as DL is able to adaptively learn directly from data (i.e., it is fully data-driven). Moreover, in case of such heterogeneity, preliminary scan registration and intensity normalization may be enclosed in the pipeline, in order to align multiple brain structures for verifying their spatial correlation in anatomical terms and reduce the intensity variation caused by the use of different scanners [27].

One limitation is that, due to the unavailability of other openly-accessible collections of 7T MR volumes, performance was evaluated using test data selected from the same database used for training and validation. However, as mentioned before, we were careful not to mix data belonging to the same subject to avoid a biased prediction, which frequently happens when data belonging to the same subject end up in training, validation, and test sets.

Another limitation is that the choice to analyze entire unenhanced 7T MR volumes made it impossible to increase the batch size to a value greater than 1 because of the technical constraints linked to the hardware capabilities. This resulted in undercutting the advantages that larger batch sizes (without exceeding, otherwise the generalization may decrease) could carry, such as faster convergence. However, hardware capabilities are now experiencing rapid empowerment, so it will soon be possible to manage this limitation. A further limitation relies on the choice to investigate only one sequence. This choice, however, not only limited the scanning time but also avoided the need for sequence alignment while reducing distortion. Moreover, T1-weighted MR imaging is extremely useful in analyzing brain structures from an anatomical point of view [30]. For instance, the presence of brain shrinkage and anomalies on subcortical structures caused by neurodegeneration can be appreciated easily from a T1-weighted MR volume [27].

In the future, therefore, further analysis on the effective reproducibility of the proposed double-stage 3D U-Net for multi-site data will be conducted. In addition, future investigations would also demonstrate an optimization in the delineation of BG boundaries and edges. Eventually, it might be appropriate to extend the analysis to pathological MR volumes to monitor the anatomical abnormalities of the brain structures involved the most in neurological pathologies, especially neurodegenerative ones.

## 5. Conclusions

The double-stage optimized 3D U-Net proposed in this research is powerful for use in brain extraction and multi-structure segmentation from 7T MR volumes without any preprocessing and training data augmentation strategy. Furthermore, it ensures a machine-independent reproducibility of its implementation and has the potential to be integrated into any decision-support system, thanks to the cloud nature and, thus, machine-independent reproducibility.

**Author Contributions:** Conceptualization, S.T., H.A. and A.S.; methodology, S.T., H.A. and A.S.; software, S.T., H.A. and M.J.M.; validation, S.T., H.A., A.S., M.J.M., L.B. and M.M.; formal analysis, S.T.; investigation, S.T. and H.A.; resources, L.B.; data curation, S.T. and H.A.; writing—original draft preparation, S.T.; writing—review and editing, H.A., A.S., M.J.M., L.B. and M.M.; visualization, S.T.; supervision, L.B. and M.M.; project administration, L.B. and M.M. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Informed Consent Statement:** Informed consent was obtained from all subjects employed in collecting data for the Glasgow database [19].

**Data Availability Statement:** Data analyzed in this research are openly available in <https://search.kg.ebrains.eu/instances/Dataset/2b24466d-f1cd-4b66-afa8-d70a6755ebea> (accessed on 2 January 2023) under the need for account generation. For research purposes, the developed algorithm will be released free of charge to the scientific community by contacting the corresponding authors (L.B. and M.M.).

**Acknowledgments:** The authors thank Svanera et al. [19] for sharing data. The authors also thank Consortium GARR, the Italian National Research & Education Network, for promoting this research.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Haq, E.U.; Huang, J.; Kang, L.; Haq, H.U.; Zhan, T. Image-based state-of-the-art techniques for the identification and classification of brain diseases: A review. *Med Biol. Eng. Comput.* **2020**, *58*, 2603–2620. [CrossRef] [PubMed]
2. Zhao, X.; Zhao, X.M. Deep learning of brain magnetic resonance images: A brief review. *Methods* **2021**, *192*, 131–140. [CrossRef] [PubMed]
3. Tomassini, S.; Sernani, P.; Falcionelli, N.; Dragoni, A.F. CASPAR: Cloud-based Alzheimer’s, schizophrenia and Parkinson’s automatic recognizer. In Proceedings of the IEEE International Conference on Metrology for Extended Reality, Artificial Intelligence and Neural Engineering, Rome, Italy, 26–28 October 2022; pp. 6–10.
4. Keuken, M.C.; Isaacs, B.R.; Trampel, R.; Van Der Zwaag, W.; Forstmann, B. Visualizing the human subcortex using ultra-high field magnetic resonance imaging. *Brain Topogr.* **2018**, *31*, 513–545. [CrossRef] [PubMed]
5. Helms, G. Segmentation of human brain using structural MRI. *Magn. Reson. Mater. Phys. Biol. Med.* **2016**, *29*, 111–124. [CrossRef]
6. González-Villà, S.; Oliver, A.; Valverde, S.; Wang, L.; Zwiggelaar, R.; Lladó, X. A review on brain structures segmentation in magnetic resonance imaging. *Artif. Intell. Med.* **2016**, *73*, 45–69. [CrossRef]
7. Haque, I.R.I.; Neubert, J. Deep learning approaches to biomedical image segmentation. *Inform. Med. Unlocked* **2020**, *18*, 100297. [CrossRef]
8. Singh, M.K.; Singh, K.K. A review of publicly available automatic brain segmentation methodologies, machine learning models, recent advancements, and their comparison. *Ann. Neurosci.* **2021**, *28*, 82–93. [CrossRef]
9. Despotović, I.; Goossens, B.; Philips, W. MRI segmentation of the human brain: Challenges, methods, and applications. *Comput. Math. Methods Med.* **2015**, *2015*, 450341. [CrossRef]
10. Tomassini, S.; Falcionelli, N.; Sernani, P.; Burattini, L.; Dragoni, A.F. Lung nodule diagnosis and cancer histology classification from computed tomography data by convolutional neural networks: A survey. *Comput. Biol. Med.* **2022**, *146*, 105691. [CrossRef]
11. Fawzi, A.; Achuthan, A.; Belaton, B. Brain image segmentation in recent years: A narrative review. *Brain Sci.* **2021**, *11*, 1055. [CrossRef]
12. Minaee, S.; Boykov, Y.; Porikli, F.; Plaza, A.; Kehtarnavaz, N.; Terzopoulos, D. Image segmentation using deep learning: A survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *44*, 3523–3542. [CrossRef] [PubMed]
13. Krithika alias AnbuDevi, M.; Suganthi, K. Review of semantic segmentation of medical images using modified architectures of U-Net. *Diagnostics* **2022**, *12*, 3064. [CrossRef]
14. Sun, L.; Ma, W.; Ding, X.; Huang, Y.; Liang, D.; Paisley, J. A 3D spatially weighted network for segmentation of brain tissue from MRI. *IEEE Trans. Med Imaging* **2019**, *39*, 898–909. [CrossRef] [PubMed]

15. Wang, L.; Xie, C.; Zeng, N. RP-Net: A 3D convolutional neural network for brain segmentation from magnetic resonance imaging. *IEEE Access* **2019**, *7*, 39670–39679. [[CrossRef](#)]
16. Bontempi, D.; Benini, S.; Signoroni, A.; Svanera, M.; Muckli, L. CEREBRUM: A fast and fully-volumetric Convolutional Encoder-decodeR for weakly-supervised sEgmentation of BBrain strUctures from out-of-the-scanner MRI. *Med. Image Anal.* **2020**, *62*, 101688. [[CrossRef](#)]
17. Ramzan, F.; Khan, M.U.G.; Iqbal, S.; Saba, T.; Rehman, A. Volumetric segmentation of brain regions from MRI scans using 3D convolutional neural networks. *IEEE Access* **2020**, *8*, 103697–103709. [[CrossRef](#)]
18. Laiton-Bonadiez, C.; Sanchez-Torres, G.; Branch-Bedoya, J. Deep 3D neural network for brain structures segmentation using self-attention modules in MRI images. *Sensors* **2022**, *22*, 2559. [[CrossRef](#)]
19. Svanera, M.; Benini, S.; Bontempi, D.; Muckli, L. CEREBRUM-7T: Fast and fully volumetric brain segmentation of 7 Tesla MR volumes. *Hum. Brain Mapp.* **2021**, *42*, 5563–5580. [[CrossRef](#)]
20. Cox, R.W. AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Comput. Biomed. Res.* **1996**, *29*, 162–173. [[CrossRef](#)]
21. Fracasso, A.; van Veluw, S.J.; Visser, F.; Luijten, P.R.; Spliet, W.; Zwanenburg, F.; Dumoulin, S.O.; Petridou, N. Lines of Baillarger in vivo and ex vivo: Myelin contrast across lamina at 7 T MRI and histology. *NeuroImage* **2016**, *133*, 163–175. [[CrossRef](#)]
22. Fischl, B. FreeSurfer. *NeuroImage* **2012**, *62*, 774–781. [[CrossRef](#)] [[PubMed](#)]
23. O'Brien, K.R.; Kober, T.; Hagmann, P.; Maeder, P.; Marques, J.; Lazeyras, F.; Krueger, G.; Roche, A. Robust T1-weighted structural brain imaging and morphometry at 7T using MP2RAGE. *PLoS ONE* **2014**, *9*, e99676. [[CrossRef](#)] [[PubMed](#)]
24. Yushkevich, P.A.; Piven, J.; Hazlett, H.C.; Smith, R.G.; Ho, S.; Gee, J.C.; Gerig, G. User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *NeuroImage* **2006**, *31*, 1116–1128. [[CrossRef](#)]
25. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional networks for biomedical image segmentation. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Proceedings, Part III 18, Munich, Germany, 5–9 October 2015; Springer: Berlin/Heidelberg, Germany, 2015; pp. 234–241.
26. Ioffe, S.; Szegedy, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In Proceedings of the International Conference on Machine Learning, Lille, France, 7–9 July 2015; pp. 448–456.
27. Tomassini, S.; Sbröllini, A.; Covella, G.; Sernani, P.; Falcionelli, N.; Müller, H.; Morettini, M.; Burattini, L.; Dragoni, A.F. Brain-on-Cloud for automatic diagnosis of Alzheimer's disease from 3D structural magnetic resonance whole-brain scans. *Comput. Methods Programs Biomed.* **2022**, *227*, 107191. [[CrossRef](#)] [[PubMed](#)]
28. Sugino, T.; Kawase, T.; Onogi, S.; Kin, T.; Saito, N.; Nakajima, Y. Loss weightings for improving imbalanced brain structure segmentation using fully convolutional networks. *Healthcare* **2021**, *9*, 938. [[CrossRef](#)] [[PubMed](#)]
29. Taha, A.A.; Hanbury, A. Metrics for evaluating 3D medical image segmentation: Analysis, selection, and tool. *BMC Med. Imaging* **2015**, *15*, 1–28. [[CrossRef](#)]
30. Le Bihan, D. How MRI makes the brain visible. In *Make Life Visible*; Springer: Singapore, 2020; pp. 201–212.
31. Ashburner, J.; Barnes, G.; Chen, C.C.; Daunizeau, J.; Flandin, G.; Friston, K.; Kiebel, S.; Kilner, J.; Litvak, V.; Moran, R.; et al. *SPM12 Manual*; Wellcome Trust Cent. Neuroimaging: London, UK, 2014; Volume 2464.
32. Jenkinson, M.; Beckmann, C.F.; Behrens, T.E.; Woolrich, M.W.; Smith, S.M. FSL. *NeuroImage* **2012**, *62*, 782–790. [[CrossRef](#)]
33. Zhang, F.; Breger, A.; Cho, K.I.K.; Ning, L.; Westin, C.F.; O'Donnell, L.J.; Pasternak, O. Deep learning based segmentation of brain tissue from diffusion MRI. *NeuroImage* **2021**, *233*, 117934. [[CrossRef](#)]
34. Kagadis, G.C.; Kloukinas, C.; Moore, K.; Philbin, J.; Papadimitroulas, P.; Alexakos, C.; Nagy, P.G.; Visvikis, D.; Hendee, W.R. Cloud computing in medical imaging. *Med. Phys.* **2013**, *40*, 070901. [[CrossRef](#)]
35. Erfannia, L.; Alipour, J. How does cloud computing improve cancer information management? A systematic review. *Inform. Med. Unlocked* **2022**, *33*, 101095. [[CrossRef](#)]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.