

Article

# Investigating the Behaviour of Machine Learning Techniques to Segment Brain Metastases in Radiation Therapy Planning

Gloria Gonella <sup>1,\*</sup> , Elisabetta Binaghi <sup>2</sup>, Paola Nocera <sup>2,3</sup>  and Cinzia Mordacchini <sup>3</sup>

<sup>1</sup> Department of Theoretical and Applied Science, University of Insubria, 21100 Varese, Italy

<sup>2</sup> Department of Physics, University of Milan, 20122 Milan, Italy

<sup>3</sup> C.S. Health Physics, ASST dei Sette Laghi, 21100 Varese, Italy

\* Correspondence: ggonella@uninsubria.it

Received: 15 July 2019; Accepted: 9 August 2019; Published: 14 August 2019



**Featured Application:** Software tools for brain metastases segmentation and active support in radiotherapy workflow.

**Abstract:** This work aimed to investigate whether automated classifiers belonging to feature-based and deep learning may approach brain metastases segmentation successfully. Support Vector Machine and V-Net Convolutional Neural Network are selected as representatives of the two approaches. In the experiments, we consider several configurations of the two methods to segment brain metastases on contrast-enhanced T1-weighted magnetic resonance images. Performances were evaluated and compared under critical conditions imposed by the clinical radiotherapy domain, using in-house dataset and public dataset created for the Multimodal Brain Tumour Image Segmentation (BraTS) challenge. Our results showed that the feature-based and the deep network approaches are promising for the segmentation of Magnetic Resonance Imaging (MRI) brain metastases achieving both an acceptable level of performance. Experimental results also highlight different behaviour between the two methods. Support vector machine (SVM) improves performance with a smaller training set, but it is unable to manage a high level of heterogeneity in the data and requires post-processing refinement stages. The V-Net model shows good performances when trained on multiple heterogeneous cases but requires data augmentations and transfer learning procedures to optimise its behaviour. The paper illustrates a software package implementing an integrated set of procedures for active support in segmenting brain metastases within the radiotherapy workflow.

**Keywords:** MRI brain segmentation; brain metastases; machine learning; features extraction; convolutional neural network; medical software

## 1. Introduction

Brain metastases (BMs) are one of the most common neurological neoplasms, and their incidence is increasing with the availability of advanced imaging techniques such as Magnetic Resonance Imaging (MRI) [1,2]. By visual inspection of MR scans, physicians can accurately examine and identify pathological tissues thanks to the high spatial resolution and contrast and the enhanced signal differentiation. In the clinical practice, MRI has been confirmed as a significant approach supporting diagnosis, surgical planning, follow up and therapy. In order to exploit this potential, intelligent techniques should complement image acquisition and visualisation tools, addressing relevant issues such as cancer detection. Extensive research has already been developed to introduce computer-aided detection and segmentation methods in neuro-oncology clinical studies. Available methods include image-based methods and machine learning-based methods [3,4]. The image-based methods use

image data and image processing techniques to detect and delineate lesions. Exemplary techniques include tissue classification based on Raman spectroscopy [5], a colour-coded map from quantitative optical coherence tomography (OCT) for differentiating cancer from noncancer in human brain tissues [6], watershed segmentation algorithms [7], active contour algorithms [8], and region growing segmentation algorithms [9]. Supervised machine learning (ML) approaches have been successfully applied to circumvent the problem of explicitly and analytically describing the specific segmentation procedure and related parameters, lying to a learning stage the charge of inducing the classifier from supervised data available. The proposed techniques make use of a single image or multispectral pattern and are interactive or fully automatic [3,10–16]. Among the most promising methods, we found the support vector machine (SVM) [17,18], discriminative models based on Random Forest and logistic regression [19,20]. Cai et al. [21] and Verma et al. [22] proposed classification methods based on SVM to classify brain neoplasms and their sub-components using multidimensional patterns obtained by a high number of MRI modalities. Ruan et al. [23] proposed SVM to segment lesions using a lower number of modalities. Bauer et al. [24] adopt a hybrid method, based on SVM and hierarchical regularisation, to segment tumour and healthy tissues, including sub-compartments. Random Forest-based methods are proposed by Zikic et al. [25] to identify brain tumour sub-compartments from multi-modal images and by Geremia et al. [26] that generate synthetic tumour images to train a discriminative regression forest algorithm using different groups of features.

Although the problem of segmenting all or part of the brain in MRI imagery does continue to be investigated attempting to satisfy the high accuracy demand in diversified clinical and neuroimaging application, only a few studies have applied such ML approaches on BM detection and segmentation [27]. BM requires specific approaches given their small size and multiplicity and the stringent requirements of RT clinical practice in which usually segmentation procedures are inserted. In this application context, MRI T1c images are usually the only modality considered and fast segmentation is required for a rapid clinical workflow.

In their review, Perez et al. [28] describe methods that use hand-crafted templates or blob based complex procedures. These works report performances strongly affected by feature extraction methods, which becomes more and more complex to improve results. Machine learning algorithms such as SVM and Random Forest are proposed in combination with spatial regularisation procedures and Gaussian processes to refine patch-based segmentation [29]. Despite the achievements obtained, automated segmentation of brain metastases and lesions, in general, remains an unsolved problem due to normal anatomical variations in brain morphology, variations in acquisition parameters of MRI scanners, the heterogeneous appearance of pathology.

Recent BM segmentation studies propose the use of deep learning methods using different convolution neural network (CNN) architectures and different MRI sequences as input data. Losch et al. have been among the first that investigated the use of deep networks to detect and segment BM [30]. Their work compares several types of spatial inputs and network topology to find performances comparable to conventional state-of-the-art models. The importance of database quality has also experimented. Growick et al. propose the use of a CNN based on the GoogLeNet architecture for automatic BM detection and segmentation. The retrospective study focusses on 156 patients with brain metastases from several primary cancers and makes use of multi-sequence MR images, including pre- and post-gadolinium T1-weighted and FLAIR scans. Results obtained are good, but several limitations because of limited sample size and false-positive results near vascular structures, are highlighted.

The use of multi-sequence MRI limits the applicability in the clinical domain. Charron et al. [27] studied the influence of MRI modalities together with the impact of the number of segmented classes, both on the detection and segmentation of brain metastases in MRI imagery. To do this, they used a modified version of DeepMedic neural network proposed by Kamnitsas et al. [31] and data augmentation strategies. The combined use of different MRI modalities outperformed the performances of the network when using single modalities. In single modality, best performances are obtained using

volumetric T1c scans. Liu et al. investigate the use of deep learning convolutional neural network (CNN) algorithm trained on both gliomas and BM [32] on T1c.

In radiation therapy (RT), the workflow for BM treatment requires fast and accurate detection and delineation of tumour volume of small size on MRI scans. Contrast-enhanced T1 (T1c) magnetic resonance imaging (MRI) is generally the only imaging modality adopted. Segmentation accomplished through a complete manual tracing is still the standard routine although it has shown to be a time-consuming and labour-intensive process affected by high intra- and inter-observer variability. In this context, ML techniques would be very helpful, actively supporting human experts in tracing the boundaries of the pathological tissues with varying degrees of automation and improving the efficiency of the overall radiotherapy clinical workflow.

Promising results obtained by recent studies in automated BM segmentation [4,27,28,33] lead us to further investigate ML techniques derived from both conventional feature-based learning and deep learning approaches and measure and compare their capability to compute spatially accurate and stable results. Indeed, both approaches have their advantages and disadvantages.

In the design of an automated segmentation procedure, the feature extraction phase plays a major role. Deep learning methods offer the advantage of automatically learn hierarchical image representations which often outperform the most effective features extracted by well-assessed procedures. However, good performances are obtained provided that a large annotated training sets are available. This requirement is especially critical in medical image analysis, in particular in BM segmentation where annotated consistent training data are difficult to collect and at best of our knowledge, no public datasets of diagnosed patients with metastatic brain tumours are available. Conventional ML models have shown excellent performances in MRI brain tumours segmentation studies even when using a much smaller training set. However, rarely these methods alone offer the opportunity to cope with the complete segmentation task and are usually complemented with pre- and post-processing procedure to overcome limitations of hand-craft features and improve spatial consistency in classification results.

In line with this general perspective, the first objective of this study was to investigate whether BM segmentation may be approached successfully by two supervised ML classifiers belonging to feature-based and deep learning approaches, respectively. SVM and V-Net Convolutional Neural Network model [34] are selected from the literature as representative of the two approaches. Their performances, already assessed in several biomedical image segmentation studies, were evaluated and compared under critical conditions imposed by specific clinical radiotherapy domain which requires fast segmentation of multiple lesions of small size on single T1c magnetic resonance image.

A second objective of the present work was the design of a software package implementing an integrated set of procedures for active support in BM segmentation within the radiotherapy workflow. It was designed to be user-friendly, with a right trade-off between automation and interactivity.

The solutions investigated in this paper are an extension of those presented in a previous work [35] from which we inherit the idea of using segmentation procedures based on SVM. In the present work, experiments have been extended by using an enlarged dataset and by developing and evaluating a second segmentation procedure based on deep learning approach.

The remainder of the paper is organised as follows. The next section illustrates previous works. Section 2 introduces the conceptual segmentation frameworks based on SVM and V-Net model, respectively. Section 3 describes the experimental evaluation and comparison of their performances using in-house collected datasets of different size. Section 4 illustrates the main features of the software package and the use of the system in clinical studies. Section 5 reports the discussion and conclusions.

## 2. Interactive BM Segmentation Based on Supervised Learning Models

In line with the above considerations, the present work is aimed to design an interactive system for active support in the delineation of BM in clinical RT studies. It is designed to operate in a clinical setting to reduce the workload of health-care professionals but leaving them full control of the process.

It is then conceived semi-automatic but requiring limited user interaction in an attempt to facilitate the insertion in current clinical practice.

The system implements a segmentation procedure hierarchically structured in three phases:

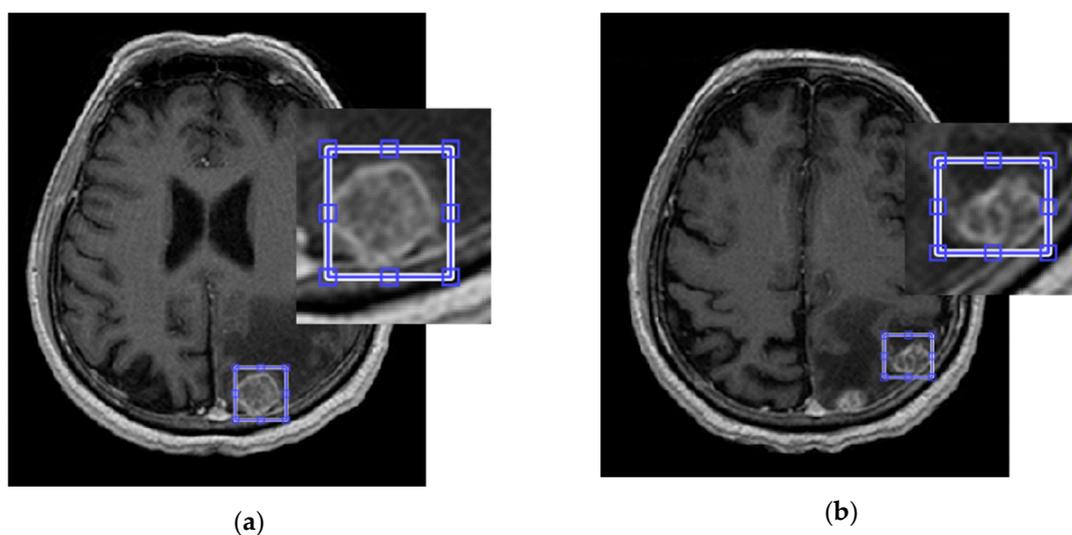
- Volume-of-interest (VoI) specification
- Automated segmentation by supervised learning models
- Segmentation refinement

In regard to the second phase, we investigated whether automated segmentation may be approached successfully by conventional, feature-based machine learning and deep learning models exploiting mutual advantages, limitations and synergies among them. We mainly studied the potential of the two methods in optimising the balance between accuracy and demand for training data. The following subsections describe the three phases of the proposed segmentation procedure.

### 2.1. VoI Specification

Conceptually, BM segmentation task includes both detection and delineation. Automated procedures cope with one or both of the two sub-tasks usually. In our context, the proposed interactive segmentation procedure limits the role of the automated solutions to the delineation task, lying to users the manual detection of BM in the T1c imagery. The advantage of this strategy is a simplification of the automated segmentation task for the learning model selected.

In this preliminary step, a user specifies a volume of interest (VOI) by drawing a rectangular region on one slice of the input volume and selecting first and last slices in such a way that the entire pathological area is bounded within the specified parallelepiped (see Figure 1).



**Figure 1.** (a,b) show two different source slices of T1c MR scans with the corresponding Volume slice perimeter bounding Brain metastases (BM) areas.

### 2.2. Supervised Classification of Pathological and Healthy Tissues

In the second stage of the segmentation procedure, a supervised machine learning model is applied to the selected sub-image. The supervised learning task is aimed to perform a hard-binary categorisation, labelling voxels within the selected volume as Brain Metastasis (BM) and Healthy tissue (H). During the training phase, the classifier learns an approximation for the true input-output relationship based on a given training set of examples constituted by  $N$  input-output pairs  $\{(x_i, y_i)\}$ ,  $i = 1, \dots, N$  where  $x_i$  is the input vector and  $y_i \in \{BM, H\}$  is a supervised label denoting the membership in the metastasis or healthy class.

We conceived and conducted two sets of experiments. The first used a conventional supervised classifier built on the top of a hand-crafted feature extraction procedure. In our previous works, we deal with MRI brain tumour segmentation using several methods selected from state-of-the-art classifiers in the field of MRI segmentation. In particular, we investigated the use of Fuzzy connectedness and Graph Cut for glial tumour segmentation [12] and SVM for meningioma and edema segmentation [13]. Fuzzy Connectedness and Graph Cut methods are interactive, asking experts to provide accurate initialisation information for each image processed. Results obtained by these methods were accurate but strongly influenced by the prior knowledge provided by the users or by ancillary methods. In the RT domain, where a large number of images are needed to be handled, they can be laborious and time-consuming. We have shown that trained SVM allows complete delineation of meningioma and edema tissues and accurate volume estimation by processing both volumetric and non-volumetric imagery in a few minutes, without requiring manual selection of example voxels or seeds. Performances obtained were good, confirming the results obtained in other studies [10]. Approaches based on Random Forest show similar characteristic and can also be suitable in the RT domain. Bauer et al. in their review [10] investigate the behaviour of SVM and Random Forest segmentation. The two methods showed comparable performances. It is worth to note, however, that evaluations were performed on different datasets and different evaluation metrics, making difficult the comparison. Statnikov et al. [36] found that, both on average and in the majority of data used in the study, Random Forests exhibit larger classification error than SVM when processing microarray datasets. In the present work, the choice fell on SVM, not because the best that can be made but as it has been extensively used in MRI image segmentation and can be considered representative of the conventional learning approaches.

The second experiment was conducted using the V-Net convolutional neural network model [34] selected among the deep learning models oriented to MRI volumetric images [37] and properly adapted for our application context. To make the work self-contained, we briefly outline the basic concepts of both learning models investigated.

### 2.2.1. BM Delineation Using SVM

SVM is a classification algorithm based on kernel methods [17,38] mapping the input patterns into a high dimensional feature space. Classes which are non-linearly separable in the original space can be linearly separated in the higher dimensional feature space.

Let  $\{(x_i, y_i)\}$  be a supervised training set of elements for a two-class classification problem, with  $x_i \in X \subseteq R^n$  and  $y_i \in \{-1, 1\}$ . Considering the case of linearly separable data, the solution to the classification problem consists of the construction of the decision function:

$$f(x) = \text{sgn}(g(x)) \text{ with} \tag{1}$$

$$g(x) = w^t x + b \tag{2}$$

that can correctly classify an input pattern  $x$  not necessarily belonging to the training set.

SVM classifier defines the hyperplane that causes the largest separation between the decision function values for the “borderline” examples from the two classes. This hyperplane can be found by minimising the cost function, as follow:

$$J(W) = \frac{1}{2} \|W\|^2 \quad \text{subject to} \tag{3}$$

$$W^T X_i + b \geq +1 \quad \text{for } y_i = +1 \tag{4}$$

$$\text{or } W^T X_i + b \leq -1 \quad \text{for } y_i = -1 \tag{5}$$

The extension to the nonlinear classification is based on the function  $g' = W^T \varphi(X) + b$ , in which the non-linear operator  $\varphi(\cdot)$  hyperparameters.

In this case, the SVM cost function to be minimized is

$$J(W, \xi) = \frac{1}{2} \|W\|^2 + C \sum_{i=1}^l \xi_i \tag{6}$$

$$y_i(w^t \varphi(X_i) + b) \geq +1 - \xi_i \text{ with } \xi_i \geq 0, i = 1, 2, \dots, l \tag{7}$$

Suykens [18] proposed a new formulation of SVM by adding an LS term in the original formulation of the cost function. This modification significantly reduces computational complexity.

The trained SVM classifier receives in input multidimensional patterns, in the form of vectors of measured features and assigns labels to corresponding T1c MR elements. Multidimensional input patterns are composed of T1c voxel intensities and corresponding textural and contextual features extracted from the MR scan. The literature proposes different sets of features related to the supervised classification of MRI data. Features are selected in function of the MRI channels used, and the classifiers adopted [3,10]. Based on our previous works, in addition to image intensities from the T1c MR scan, we consider features describing neighbour relationships and texture [39]. The feature set adopted is described in the next section. As detailed in Section 3, several strategies were conceived to conduct the learning stage. Results obtained were analysed, and the resulting optimal configuration has been used in the comparison analysis.

### 2.2.2. BM Delineation Using V-Net

The V-Net model is a Fully Convolutional Neural Network proposed by Milletari et al. [34] for volumetric medical image segmentation. The name V-Net comes from the fact that the network can be drawn with the symmetric shape like the letter V. Salient aspects of the model proposed are the use of volumetric convolution to overcome the slice-by-slice processing of input image volumes and the use of a novel objective function based on Dice coefficient maximisation, optimised during training. Both these aspects are significant for our segmentation task, which is based on volumetric data characterised by a strong imbalance between the number of voxels belonging to the pathological area and background.

In our experiment, we adopt the V-Net configuration proposed for Lung Tumour Segmentation in [40] (see Figure 2). The inputs of the network are T1c MRI sub-volumes by dimension  $64 \times 64 \times 64$ . The V-Net has an increasing number of convolutional filters (16, 32, 64, 128, 256) with a total of 116 layers and 128 connections. The output layer is a custom Dice loss layer [41].

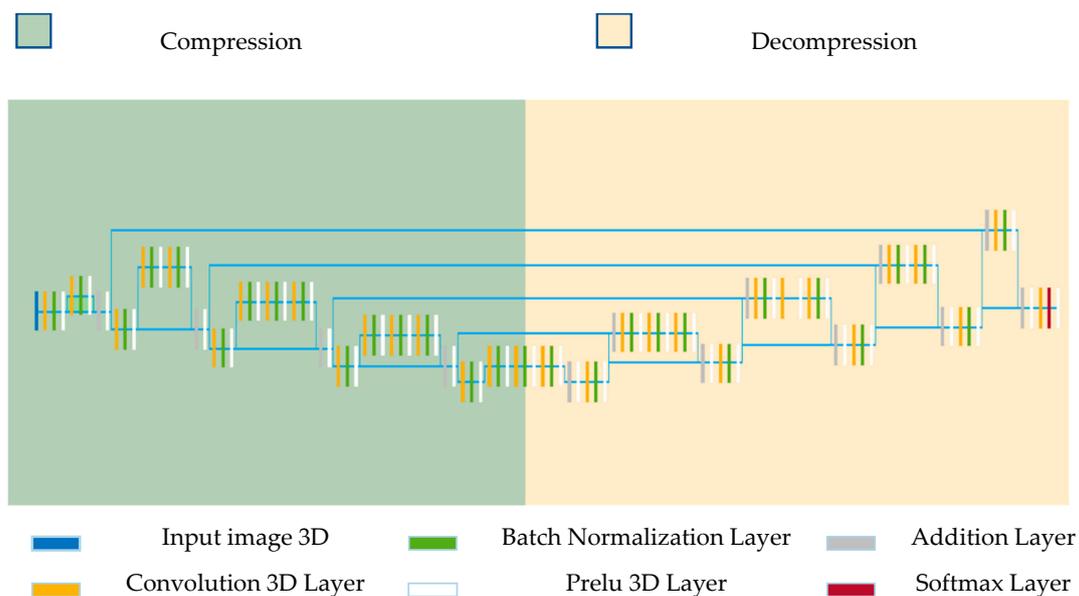


Figure 2. The architecture of 3D convolutional V-Net.

Figure 2 shows a schematic representation of V-Net architecture adopted, which is composed of two parts implementing compression and decompression paths. In the left part, we found a deep residual learning strategy: all the outputs of the convolutional layers, after non-linearities processing, are added to the output of the last convolutional layer [42].

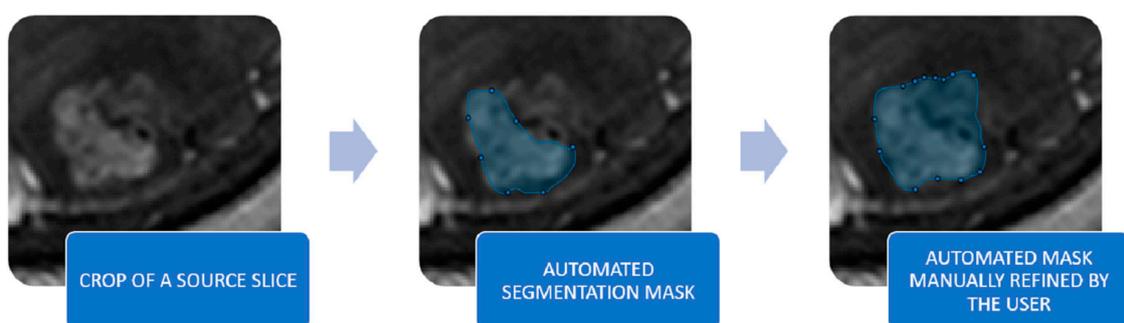
In each stage, the compression part of the network computes features two times higher in number than the one of the previous layers. In the decompression part, the network extracts the features and expands the spatial support of the lower resolution feature maps. The last convolutional layer implements two outputs of the same size of the input image volume. A voxel-wise soft-max operation is applied, converting the two outputs as probabilistic segmentation of background and BM regions.

Results obtained in several segmentation studies show how necessary it is to use large-scale datasets for effective application of deep learning methods [11,43]. Whereas in our context, as in many other biomedical domains, manually segmented reference volumes are not easy to obtain, our strategy includes data augmentation and transfer learning tasks to compensate for the limited dataset available and reaching the double goal of increasing robustness and generalisation capability of the network respectively, as described in Section 3.

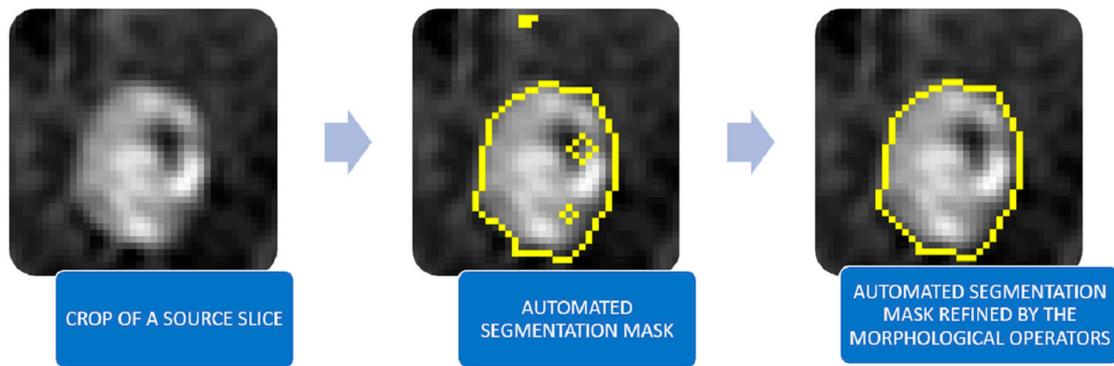
### 2.3. Segmentation Refinement

In our study, we conceived automated segmentation as an intermediate task within a typical workflow in radiotherapy planning. The active decision support of the machine learning procedure reduces user-interaction that is limited to two phases: a preliminary “detection” phase with which to specify the VoI, as illustrated in Section 2.1, and a post-processing phase aimed to refine the identified segments. Refinement is made necessary by commission and omission errors that are most likely to occur when lesions are of small size, and the tumour area presents inhomogeneity due to necrosis and active parts. Our strategy provides both for computer-aided manual editing as illustrated in Section 4 and an automated procedure based on the use of Morphological Operators to refine the segmented masks in an attempt to reduce omission and commission errors and making the segmented tumour area more compact. For each selected slice, Opening and Closing Operators with spherical shape apply consequently. The Opening Operator removes from the binary input image all the connected components that have a lower number of pixels than a set value and outputs a new binary image. The Closing Operator closes holes present in the image and returns the closed binary image. We performed three different tests, using for all a disk-shaped structuring element, aimed to tune parameters values of the Morphological Operators and decide the order of application. In the first test, only the opening morphological operator (Open) was applied by varying the radius; in the second test, only the morphological closing operator (Close) was applied, varying the radius; in the third test, we applied both operators in a different sequence. The best result was obtained by applying first the opening morphological operator with a radius of 5 and then the closing operator with a radius equal to 10.

Figures 3 and 4 illustrate the refinement of automated segmentation results performed manually by the user and by the Morphological Operators, respectively.



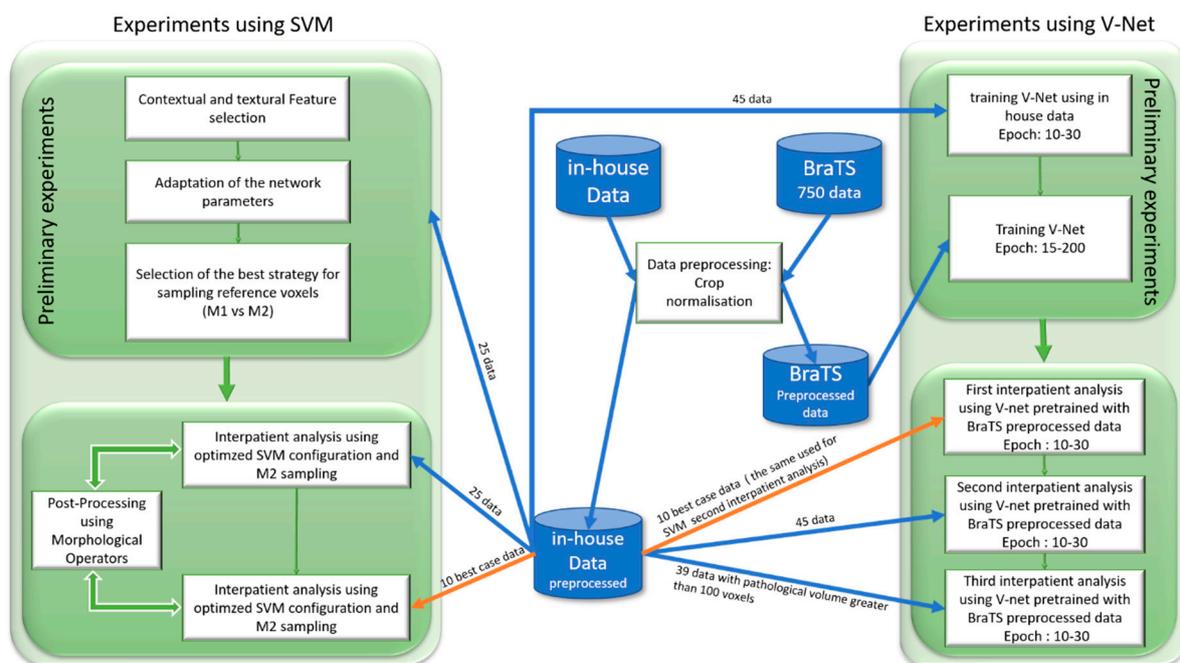
**Figure 3.** From left to right: crop of a source T1c MRI slice, superimposition of the automated segmentation mask, superimposition of the mask manually refined by the user.



**Figure 4.** From left to right: crop of a source T1c MRI slice, superimposition of the automated segmentation mask, superimposition of the mask automatically refined by Morphological Operators.

### 3. Experiments

We conceived and conducted several experiments to systematically investigate the behaviour and quantitatively assess performances of the selected machine learning models for automated support in the BM segmentation within the RT workflow (see Figure 5).

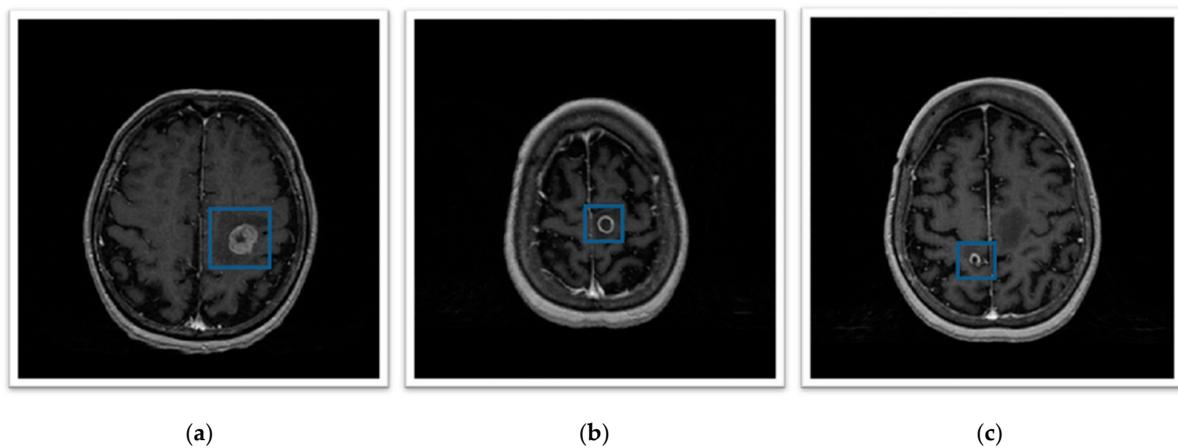


**Figure 5.** Flowchart illustrating the experimental design.

The experiments used two data sets: one in-house collected and one public, created for the Multimodal Brain Tumour Image Segmentation (BraTS) challenge [44]. In-house data are used to train and test SVM classifier and to fine-tune and test V-Net network pre-trained using data from BraTS.

#### 3.1. In-House Data Acquisition

The acquired dataset is composed of 45 T1c volumetric MR scans. Volumes are acquired using a 3D sequence characterised by 0.9 mm isotropic voxels, the pixel spacing of 0.47 mm and the slice thickness of 2.67 mm. The tumour cases considered are heterogeneous in terms of shape, position and intensity level (see Figure 6).



**Figure 6.** (a–c) Show three source slices of T1c MR Volumes with the delimitation of the lesion area.

### 3.2. Evaluation Metrics

We assessed the accuracy of segmentation results by comparing the spatial distribution of the masks obtained by the automated segmentation with that of the masks obtained through a manual segmentation of the T1c images performed by radiologists. Metrics adopted for the accuracy is described below, according to Bouix et al. [45]. The minimal problem of assessing the agreement between two binary maps B1 and B2, representing reference and segmented data respectively, is obtained in terms of number of voxels at which both B1 and B2 score “1” (True Positive  $Tp$ ) or “0” (True Negative  $Tn$ ), the number of voxels at which B1 scores “0” and B2 scores “1” (False Positive  $Fp$ ) and vice-versa (False Negative  $Fn$ ). Several similarity indexes could be defined. We use Dice ( $DSC$ ) [46], Precision and ( $P$ ) and Recall ( $R$ ) indexes (Olson and David, 2008) defined as follows:

$$DSC = \frac{2Tp}{(2Tp + Fn + Fp)} \quad (8)$$

$$P = \frac{Tp}{(Tp + Fp)} \quad (9)$$

$$R = \frac{Tp}{(Tp + Fn)} \quad (10)$$

The  $DSC$  index has been used broadly in the field of segmentation as a measure of spatial overlap,  $P$  and  $R$  indexes allow to measure under- and over-estimations [45].

### 3.3. Experiments with SVM

Performances of the SVM-based segmentation were evaluated by generating several segmentation procedures obtained by different SVM configurations and using different learning strategies. We used 25 reference data corresponding to those acquired and available in our clinical domain at the start of the study.

We developed two preliminary tasks. First, we defined the optimal feature set that composes patterns in input to the SVM. For this task, the SVM has been configured as a soft-margin least square (LS) model with a linear kernel. Contextual and textural features have been analysed systematically in order to determine the combination that is most appropriate for the current classification task. In particular, several configurations of the segmentation procedure have been experimented initially providing in input only intensity values of central voxel and neighbour voxels. Different neighbourhoods have been considered including incrementally neighbours along with voxel faces, corners and edges up to a maximum of 26 voxels. In a second step, an enlarged feature set has been considered adding

textural features to the best neighbourhood configuration. The following set of features has been finally selected (see Figure 7):

- intensities from T1c scan
- first-order texture features: mean ( $\mu$ ), variance ( $\sigma$ ), skewness ( $\gamma$ ), kurtosis ( $\beta$ ) and entropy ( $H$ )
- intensities in 26 neighbourhood voxels

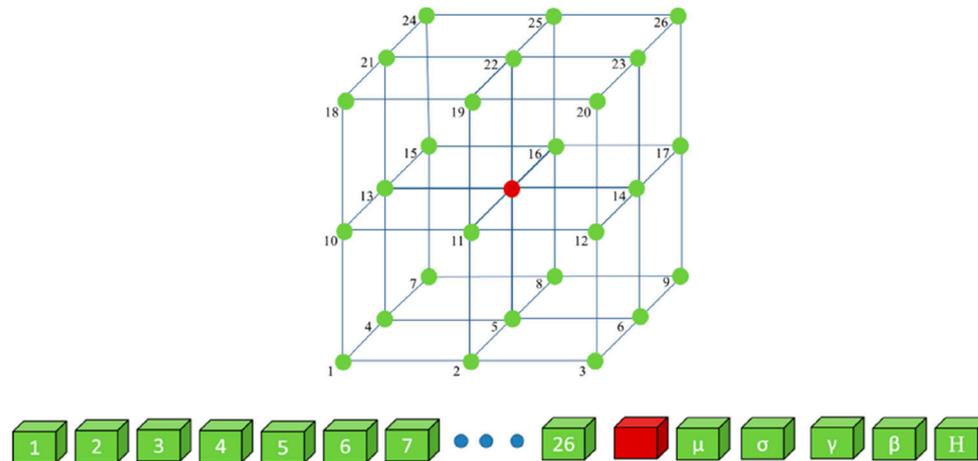


Figure 7. Input pattern vector composed of contextual and textural features.

Feature values have been normalised to have zero mean and unit variance.

The second step was to determine the value of the internal parameters of the model. Different types of kernels were tested, such as linear, quadratic, cubic, fine-medium-coarse gaussian. Given the results obtained, we confirmed the configuration of SVM as a soft-margin LS model with a linear kernel.

With this optimised configuration, we conducted the first set of experiments by varying the modality for selecting training samples from the VoIs extracted from the available in-house dataset. In the first modality (M1) data are extracted from the reference masks selecting elements within the overall VoI under study. In the second modality (M2), the random selection was limited within a region built around the contour of the tumour reference masks. After trial and error procedures, the edge width was tuned to 8 pixels. These experiments have been conducted in intra-patient modality selecting training and test sets from the reference masks of the same VoI and built by randomly selecting elements in the proportion of 70% and 30% respectively. An equal number of elements labelled BM and H was extracted. In the M1 modality, before the random extraction of BM elements, contour elements of BM regions were reinserted.

Table 1 shows the numerical results obtained in terms of mean values of  $DSC$ ,  $P$  and  $R$  indexes over 25 cases under study. SVM, trained according to M2 strategy, slightly prevails with a mean  $DSC$  value equal to 0.878.  $P$  and  $R$  values highlight a significant reduction of omission and commission errors. These results can be explained by the fact that metastases have little extensions and a high level of heterogeneity occurs in the internal part of the pathology due to the presence of necrosis and/or active parts (see Figure 6). This high level of heterogeneity is difficult to learn during training and, as seen in our experimental context, SVM shows better behaviour when trained on relatively homogeneous regions.

**Table 1.** Dice (*DSC*), Precision (*P*), Recall (*R*) values obtained using M1 and M2 strategy and tested on the overall cases under study.

		M1	M2
<i>DSC</i>	Mean	0.808	0.878
	Var	0.008	0.003
	Min	0.549	0.757
	Max	0.908	0.963
<i>P</i>	Mean	0.824	0.884
	Var	0.006	0.003
	Min	0.648	0.749
	Max	0.927	0.963
<i>R</i>	Mean	0.796	0.873
	Var	0.012	0.003
	Min	0.476	0.764
	Max	0.923	0.963

Using the configuration described above and with the M2 voxel extraction strategy, we quantified segmentation performances by conducting a second set of experiments aimed to investigate the ability of the SVM to generalise under increasing levels of heterogeneity seen in training.

First, 25 SVM models are trained from one case and tested on all the cases under study. Table 2 shows the results obtained by the best configuration. We compute accuracy values obtained with and without the use of Morphological Operators to isolate their contribution within the overall segmentation procedure.

**Table 2.** Dice (*DSC*), Precision (*P*), Recall (*R*) values obtained by the segmentation procedure trained on one case according to M2 strategy, with and without the use of Morphological Operators (MO) and tested on the overall 25 cases under study.

		SVM	SVM + MO
<i>DSC</i>	Mean	0.701	0.693
	Var	0.011	0.035
	Min	0.462	0
	Max	0.844	0.897
<i>P</i>	Mean	0.747	0.696
	Var	0.026	0.047
	Min	0.437	0
	Max	0.997	0.997
<i>R</i>	Mean	0.737	0.769
	Var	0.035	0.057
	Min	0.41	0
	Max	0.983	0.990

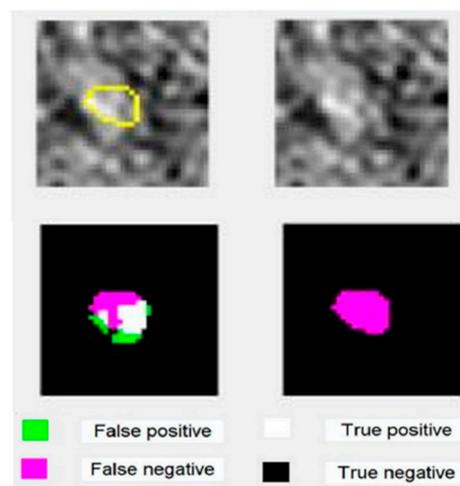
Secondly, in order to measure the ability of the SVM to segment under an increased level of heterogeneity, we select the 10 cases where the SVM gave the best results in intra-patient analysis. The Series of 6 cases extracted from the 10 cases, previously selected are considered for training, for a total of 120 learning stages. Table 3 shows accuracy values obtained by the best SVM, with and without the use of Morphological Operators, tested on the overall 25 cases.

**Table 3.** Dice (*DSC*), Precision (*P*), Recall (*R*) values obtained by the segmentation procedure with data from a set of VoIs belonging to different scans with and without the use of Morphological Operators (MO) and tested on the overall cases under study.

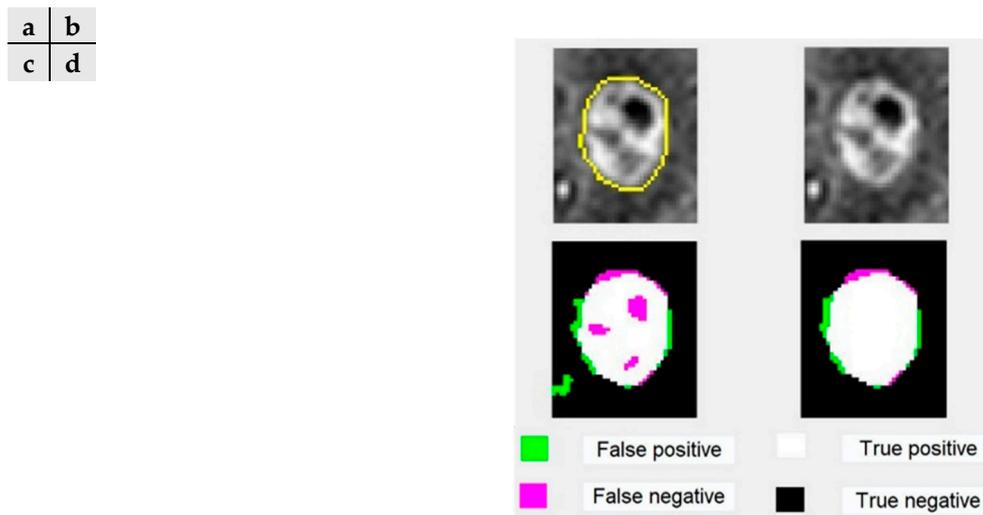
		SVM	SVM + MO
<i>DSC</i>	Mean	0.653	0.66
	Var	0.008	0.028
	Min	0.39	0
	Max	0.77	0.82
<i>P</i>	Mean	0.681	0.641
	Var	0.017	0.025
	Min	0.278	0
	Max	0.968	0.881
<i>R</i>	Mean	0.71	0.762
	Var	0.026	0.035
	Min	0.482	0
	Max	0.955	0.976

Looking at values in Tables 2 and 3 in more detail, we notice that the application of Morphological Operators is not always advantageous. However, when studying individual cases, we have noticed that under-estimation and over-estimation errors occur systematically when the pathology occupies a very small volume (under the 100 elements), and it is inserted in a highly heterogeneous context. Figure 8 illustrates an example with a slice (Slice 1) including with a metastasis remarkable small. The refinement accomplished by the Morphological Operators deletes all the true positive elements identified by the SVM classifier. On the contrary, the segmentation masks of the larger pathological area in the slice (Slice 2) shown in Figure 9, indicate that the segmentation strategy benefits from the allied use of SVM and Morphological Operators. Table 4 lists the numerical results of the cases illustrated in Figures 8 and 9.

a	b
c	d



**Figure 8.** First row, from left to right: (a) crop of a source slice (Slice 1) of T1c MR Volume with superimposed the contour of metastasis reference mask (dimension: 83 elements), (b) Slice of the corresponding VoI. Second row from left to right: (c) Segmentation mask produced by SVM, (d) refinement by the Morphological Operators.



**Figure 9.** First row, from left to right: (a) crop of a source slice (Slice 2) of T1c MR Volume with superimposed the contour of metastasis reference mask (dimension: 644 elements), (b) Slice of the corresponding VoI. Second row from left to right: (c) Segmentation mask produced by SVM, (d) refinement by the Morphological Operators.

**Table 4.** Dice (*DSC*), Precision (*P*), Recall (*R*) values obtained by the segmentation procedure when processing slices in Figures 8 and 9.

		DSC	P	R
Slice 1	SVM	0.556	0.656	0.482
	SVM + MO	0	0	0
Slice 2	SVM	0.885	0.898	0.873
	SVM + MO	0.940	0.926	0.953

Automatic segmentations were evaluated qualitatively through visual inspection. The complete strategy, including the combined use of SVM and Morphological Operators, have been judged satisfactory. The limitations of the segmentation procedure, inherent to specific cases, as illustrated above, are considered acceptable and manageable with interactive phases devoted to manual refinements of the automated results.

### 3.4. Experiments with V-Net

In these experiments, we consider an enlarged in-house data set composed of 45 collected MRI scans made available for this second phase of the study. These data are used to fine-tune the V-Net model after a pre-training procedure based on the use of BraTS data, including 750 4D annotated volumes, each representing a stack of 3D images. Each 4D volume has size  $240 \times 240 \times 155 \times 4$ , where the first three dimensions correspond to height, width, and depth of a 3D volumetric image. The fourth dimension corresponds to different scan modalities. For all the experiments developed, we pre-processed both in-house and public BraTS data before their use in training, according to the following strategy. To reduce the amount of data to be processed, we cropped original volumes to sub-regions containing a significant amount of brain and tumour tissues. Then, we normalise each data independently by subtracting the mean and dividing by the standard deviation of the cropped brain region; then, we removed the outliers and, finally, we applied a data resizing in the interval  $[0,1]$ . Data are augmented by randomly rotating and reflecting image patches to make training more robust. Finally, we randomly extract patches of  $64 \times 64 \times 64$  voxels sizes from the cropped data.

Initially, we conduct experiments aimed to assess the utility of transfer learning. V-Net was trained directly using the preprocessed in-house collected data. We considered several learning strategies for

training V-Net by varying the number of epochs with a range from 10 to 30 and splitting the data set into 82% for training, 6% for validation and 12% for test. We obtained poor, unacceptable results with accuracy close to zero.

We proceeded then, by pre-training V-Net using pre-processed T1c scans of BraTS data. During the experiment, the number of epochs varied from 15 to 200 and the preprocessed dataset was split again into 82% for training, 6% for validation and 12% for test. The DSC value obtained is equal to 0.53. Performances obtained would not be acceptable for the final segmentation task, but should be considered acceptable for pre-training.

The V-Net architecture, pre-trained on glioma segmentation task, was refined using in-house collected data with different learning strategies. Several configurations of V-Net-based procedure were considered by varying epoch.

We conceived three experiments addressing the following main questions:

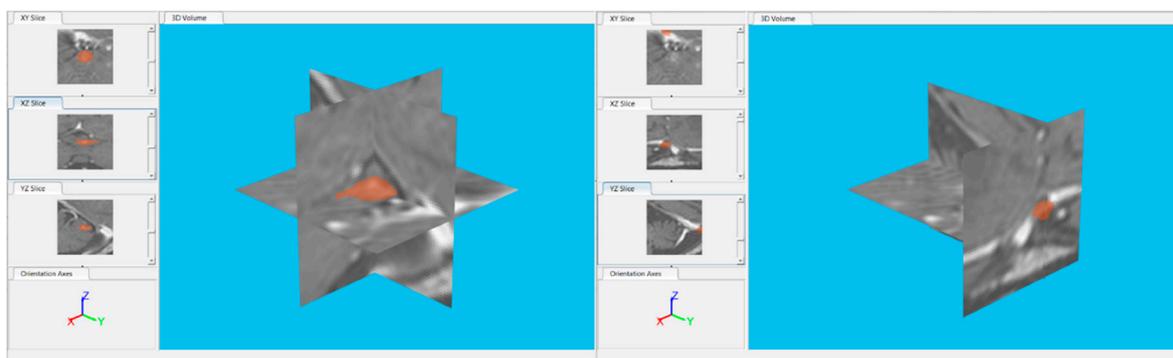
- How did the V-Net compare with SVM when trained on the same data set?
- How did performances could be optimised increasing the number of training data?
- How did dimension of lesions influence accuracy of segmentation results?

To address the first question, we train and test V-Net with the same strategy used in the second interpatient analysis conducted with SVM (see Section 3.3). Results obtained are shown in Table 5. Comparing these results with those obtained by the SVM, shown in Table 3, we found that SVM segmentation prevails.

**Table 5.** Dice (*DSC*), Precision (*P*), Recall (*R*) values obtained by V-Net segmentation trained according to the second inter-patient analysis conducted for SVM (see Section 3.3).

	<i>P</i>	<i>R</i>	<i>D</i>
<b>Mean</b>	0.63	0.46	0.51
<b>Var</b>	0.17	0.11	0.13
<b>Min</b>	0.56	0.39	0.46
<b>Max</b>	0.68	0.55	0.59

To address the second question, we used an enlarged dataset composed of 45 T1c volumes, of which 17 for test and the remaining 28 used in 4-fold cross-validation. Results obtained were still poor with accuracy values under the value of 0.50 for all the indexes used. Moving deeper into the experiment conducted and analysing case-by-case performances obtained, we found that lower accuracy values occur when the pathology occupies a very small volume (under the 100 elements), and it has a highly heterogeneous context. V-Net has difficulties in identifying a very small lesion, as illustrated in Figure 10, where the segmented mask has a spatial distribution quite different from the reference mask.



**Figure 10.** 3D visualisation of reference mask (left) and V-Net mask superimposed on a very small lesion in MRI scan under study.

We conducted a third experiment by using a reduced set of data obtained by the initial in-house data set, by eliminating cases in which pathology occupies volumes under 100 voxels.

The final dataset is composed of 39 T1c volumes of which 12 for test and the remaining 27 used in 3-fold cross-validation. Results obtained are shown in Table 6.

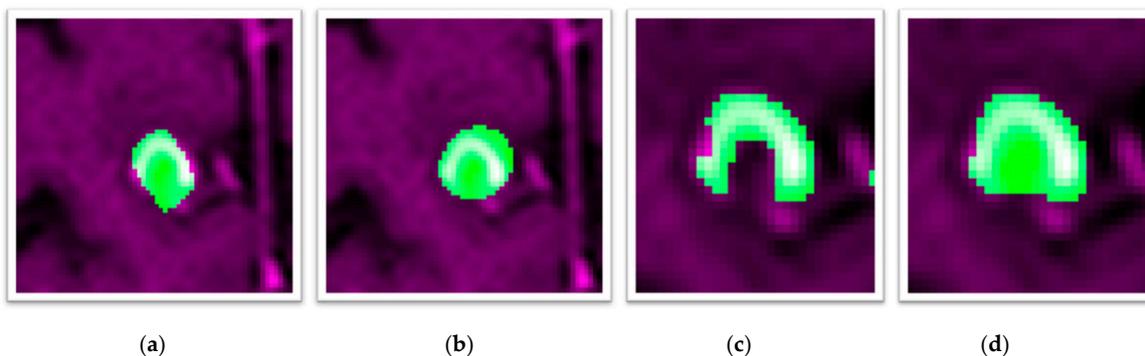
**Table 6.** Dice (*DSC*), Precision (*P*), Recall (*R*) values obtained by V-Net segmentation trained on the refinement dataset.

	<i>P</i>	<i>R</i>	<i>D</i>
<b>Mean</b>	0.857	0.537	0.641
<b>Var</b>	0.021	0.017	0.010
<b>Min</b>	0.523	0.298	0.455
<b>Max</b>	1.000	0.706	0.802

Comparing performances obtained by the SVM configuration optimised for inter-patient analysis, shown in Table 2, we noticed that performances of the two classifiers are in this case comparable.

Analysing in detail, the segmentations performed we notice that the masks produced by V-Net have a higher level of spatial consistency than those produced by SVM, making the use of MO ineffective. As an example, Figure 11 shows results produced by SVM and V-Net highlighting differences in the computed segmentation masks.

As regard hyperparameters and execution time, the pre-training task takes about 150 h on an NVIDIA™ 1080Ti with 11 GB of RAM, and the training with the in-house data takes about 7 h. In training we used the stochastic Adam optimizer, the shuffle in every epoch, InitialLearnRate =  $1 \times 10^{-3}$ , LearnRateSchedule = piecewise, LearnRateDropPeriod = 5, RateDropFactor = 0.97.



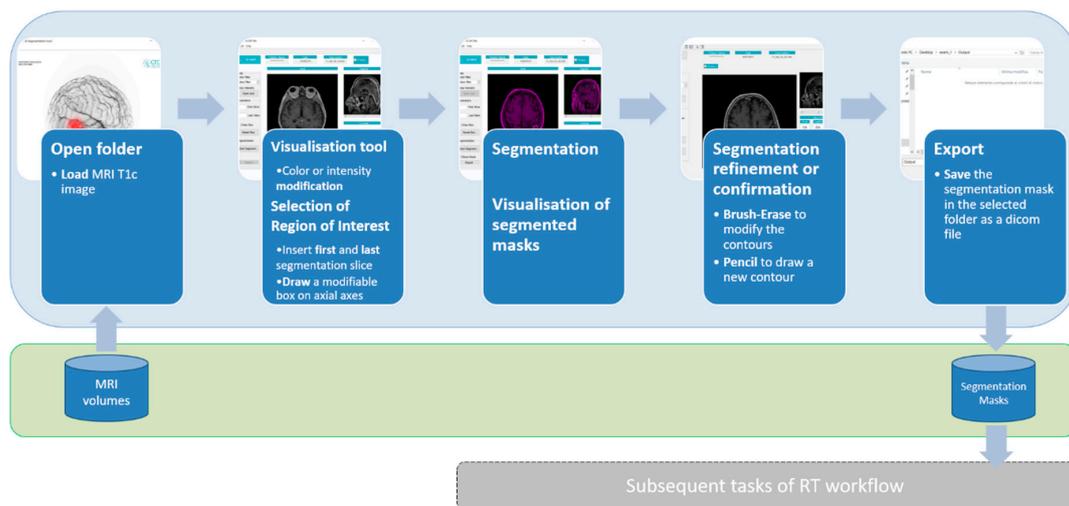
**Figure 11.** Crop of the source slice of T1c MR Volume with superimposed reference mask (a), segmentation mask produced by V-Net (b), segmentation mask produced by Support vector machine (SVM) without (c) and with MO (d). The source slice is the last image on the right shown in Figure 6c.

#### 4. The Software System: Graphical User Interface (GUI) and Sample Program Runs

The above-illustrated computational procedures are integrated into a unified framework to support the segmentation of BM within the RT workflow (see Figure 12).

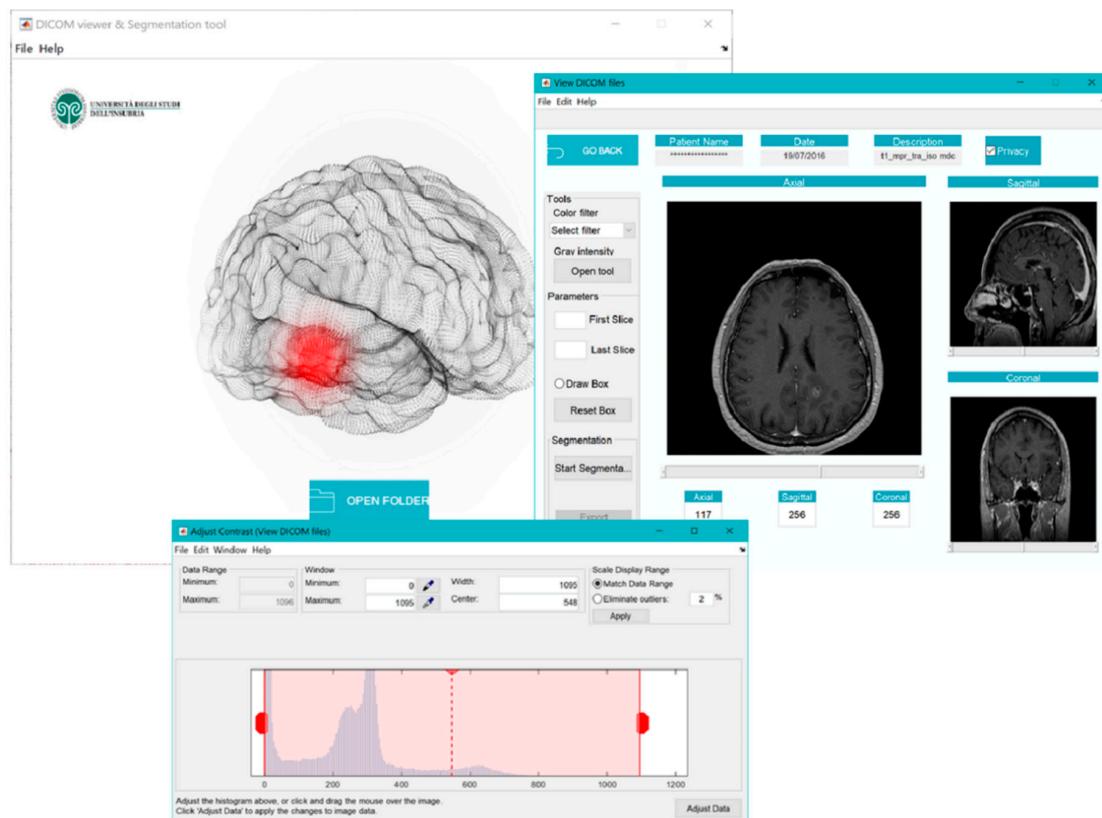
Both the SVM-based and deep learning-based procedures have been implemented in different system configurations to extend the evaluation of the automated results obtained through an accurate field test with novel clinical subjects.

The software design started with the collection and analysis of requirements to outline the user model and operational conditions. We assessed cognitive and perceptual processes, attitudes and limitations involved in visual inspection and analysis tasks. Operation requirements concerning hardware facilities and input and output devices are inherited from protocols in use in neuroradiology domain. The multi-window system was implemented in MATLAB (MathWorks®).



**Figure 12.** The workflow of the functional structure of the software system implementing the segmentation procedure.

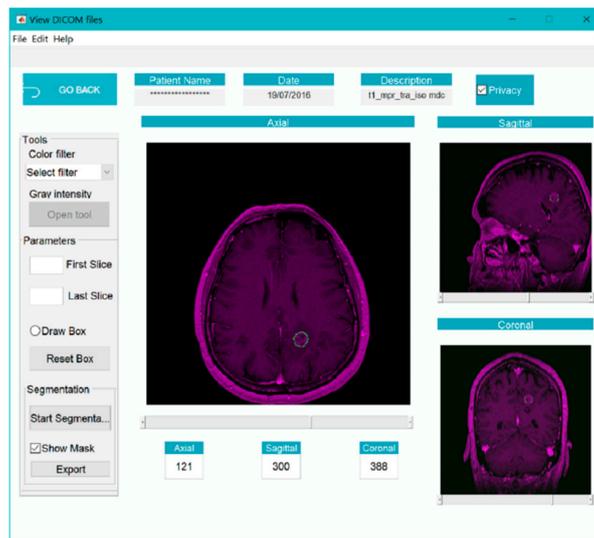
Figure 13 shows the system interfaces documenting the initial phases of the session and input/output facilities. Visualisation tools are available for a reliable inspection of the selected MRI scan. By default, the three orthogonal planes of the corresponding central slice are simultaneously visualised. Several colour maps can be easily selected to modify the appearance of the axial slice.



**Figure 13.** Screenshots of the system interfaces documenting the initial input/output phases.

On the left part of the GUI we found options made available by the system to develop segmentation task. The user selects the range of slices that must be processed by inserting the numbers of first and last slices and proceed in the VoI specification (check on “Drow Box”) by using zooming facilities if the case.

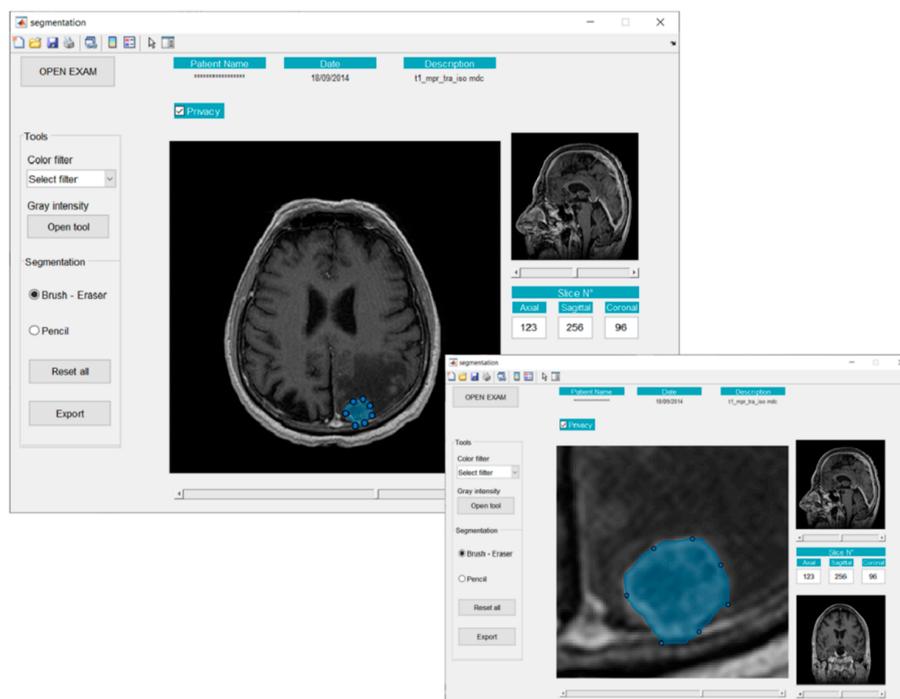
Segmentation procedure is run, and the contour of mask obtained is superimposed on the selected slices in all the projections as illustrated in Figure 14.



**Figure 14.** Screenshot of the system interfaces documenting the visualisation of the segmentation masks computed by SVM.

The system allows the user to export the segmentation mask, to refine the segmentation results, to proceed in a further segmentation in the same volume or to return to the initial phase by unchecking “show Mask” command.

Figure 15 shows options made available by the system to validate the automated segmentation. The segmented mask is shown superimposed on corresponding T1c source slice. The user is allowed to modify the contour by checking on “Brush-Eraser” command or delineating the entire segment in case the automated result occupies a wrong position (check Pencil).



**Figure 15.** Screenshots of the System interfaces documenting the validation of the computed segmentation masks.

The software is preliminary installed at the radiotherapy unit of the Hospital ASSTsettelaghi, Varese (Italy) for a field test and then made available online with links in our institutional website.

## 5. Discussion and Conclusions

As seen in our experimental context, brain metastases can be automatically segmented using automated processing of T1c scans of MR images. The main goal of the proposed procedure was to facilitate the segmentation task within the radiation therapy workflow, actively supporting radiologist in the delineation of lesions. Our results showed that both a feature-based and a deep network approach are promising for the segmentation of MRI brain metastases achieving both an acceptable, although a not very good level of performance. Nevertheless, it is worth to note that SVM and V-Net achieved best results under quite different conditions. In the inter-patient analysis, SVM appears to be unable to manage a high level of heterogeneity in training, producing best results when trained only on individual cases, plus omitting heterogeneous voxels belonging to necrosis within the pathological area. On the contrary, V-Net model achieved an acceptable level of accuracy when trained on multiple cases and selecting voxels from the entire pathological area. Moreover, the two optimised models show comparable performances, but segmentation results are qualitatively different in many cases, as illustrated in Figure 11. Differences are emphasized when segmenting highly heterogeneous lesions. In line with results obtained in previous works focused on MRI brain tumour segmentation, we found that SVM can capture complex multivariate relationships in the data, but classification results may suffer from spatial inconsistencies. The use of post-processing refinement stages is usually suggested, as in our strategy that contemplates the use of morphological operators. V-Net segmentations are more homogeneous on average and results obtained confirm that no additional processing is required.

We identified some limitations of the study. A very small number of in-house data were available, and the interpretation of the results would inevitably have to allow for that. However, as a general consideration, if the use of the deep network, pre-trained on neighbouring but not equivalent data and then fine-tuned with a very small dataset, allowed us to obtain acceptable performances, it is reasonable to think that it is possible to improve these results significantly when a broader dataset becomes available. For SVM, instead, it is difficult to foresee that there will be a benefit from the use of a wider training set having shown limitations in learning from varied heterogeneous data.

Moreover, both the classifiers made under-estimation and over-estimation errors when the pathology occupies a small volume (under the 100 elements) and has a heterogeneous context. These errors are because of misclassifications between the vessel and lesion voxels.

Probably, the simultaneous use of different MRI modalities could improve performances than when using single MRI modalities as experimented in [27], but the use of multimodal procedure would make it difficult the implementation of the automated segmentation in the radiotherapy clinical practice. Both the learning models showed advantages and drawbacks, but in the light of a great number of data available, a deep learning algorithm appears preferable.

The main conclusions of our experimental work can be summarised as follows:

- SVM shows good performances with a smaller training set, but it is unable to manage a high level of heterogeneity in the data and requires post-processing refinement stages
- The V-Net model shows good performances when trained on multiple heterogeneous cases but requires data augmentations and transfer learning procedures to optimise performances
- Users have found the segmentation strategy implemented either with SVM + Morphological Operator or V-Net, satisfactory and have considered misclassification errors acceptable and manageable with interactive phases devoted to manual refinements of the automated results.

The automated procedure is installed in the RT workflow to measure benefits in supporting brain metastasis delineation. Taking data is going and retraining of the segmentation procedure will be possible.

**Author Contributions:** These authors contributed equally to this work.

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

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

## References

1. Cagney, D.N.; Martin, A.M.; Catalano, P.J.; Redig, A.J.; Lin, N.U.; Lee, E.Q.; Wen, P.Y.; Dunn, I.F.; Bi, W.L.; Weiss, S.E.; et al. Incidence and prognosis of patients with brain metastases at diagnosis of systemic malignancy: A population-based study. *Neuro Oncol.* **2017**, *19*, 1511–1521. [[CrossRef](#)] [[PubMed](#)]
2. Greenberg, H.S.; Chandler, W.F.; Sandler, H.M. *Brain Tumors*; Contemporary Neurology Series; Oxford University Press: New York, NY, USA, 1999; ISBN 978-0-19-512958-8.
3. Gordillo, N.; Montseny, E.; Sobrevilla, P. State of the art survey on MRI brain tumor segmentation. *Magn. Reson. Imaging* **2013**, *31*, 1426–1438. [[CrossRef](#)] [[PubMed](#)]
4. Liu, Y.; Stojadinovic, S.; Hrycushko, B.; Wardak, Z.; Lu, W.; Yan, Y.; Jiang, S.B.; Timmerman, R.; Abdulrahman, R.; Nedzi, L.; et al. Automatic metastatic brain tumor segmentation for stereotactic radiosurgery applications. *Phys. Med. Biol.* **2016**, *61*, 8440–8461. [[CrossRef](#)] [[PubMed](#)]
5. Jermyn, M.; Mok, K.; Mercier, J.; Desroches, J.; Pichette, J.; Saint-Arnaud, K.; Bernstein, L.; Guiot, M.C.; Petrecca, K.; Leblond, F. Intraoperative brain cancer detection with Raman spectroscopy in humans. *Sci. Transl. Med.* **2015**, *7*, 274ra19. [[CrossRef](#)] [[PubMed](#)]
6. Kut, C.; Chaichana, K.L.; Xi, J.; Raza, S.M.; Ye, X.; McVeigh, E.R.; Rodriguez, F.J.; Quinones-Hinojosa, A.; Li, X. Detection of Human Brain Cancer Infiltration ex vivo and in vivo Using Quantitative Optical Coherence Tomography. *Sci. Transl. Med.* **2015**, *7*, 292ra100. [[CrossRef](#)]
7. Bleau, A.; Leon, L. Watershed-Based Segmentation and Region Merging. *Comput. Vis. Image Underst.* **2000**, *77*, 317–370. [[CrossRef](#)]
8. Ilunga Mbuyamba, E.; Cruz-Duarte, J.; Avina-Cervantes, J.; Rodrigo Correa-Cely, C.; Lindner, D.; Chalopin, C. Active Contours Driven by Cuckoo Search Strategy for Brain Tumour Images Segmentation. *Expert Syst. Appl.* **2016**, *56*. [[CrossRef](#)]
9. Subudhi, B.N.; Thangaraj, V.; Sankaralingam, E.; Ghosh, A. Tumor or abnormality identification from magnetic resonance images using statistical region fusion based segmentation. *Magn. Reson. Imaging* **2016**, *34*, 1292–1304. [[CrossRef](#)]
10. Bauer, S.; Wiest, R.; Nolte, L.P.; Reyes, M. A survey of MRI-based medical image analysis for brain tumor studies. *Phys. Med. Biol.* **2013**, *58*, R97–R129. [[CrossRef](#)]
11. Akkus, Z.; Galimzianova, A.; Hoogi, A.; Rubin, D.L.; Erickson, B.J. Deep Learning for Brain MRI Segmentation: State of the Art and Future Directions. *J. Digit. Imaging* **2017**, *30*, 449–459. [[CrossRef](#)]
12. Pedoia, V.; Balbi, S.; Binaghi, E. Fully Automatic Brain Tumor Segmentation by Using Competitive EM and Graph Cut. In Proceedings of the 18th International Conference on Image Analysis and Processing, Genoa, Italy, 7–11 September 2015; Volume 9279, pp. 568–578.
13. Binaghi, E.; Pedoia, V.; Balbi, S. Meningioma and peritumoral edema segmentation of preoperative MRI brain scans. *Comput. Methods Biomech. Biomed. Eng. Imaging Vis.* **2018**, *6*, 362–370. [[CrossRef](#)]
14. Sachdeva, J.; Kumar, V.; Gupta, I.; Khandelwal, N.; Ahuja, C.K. Segmentation, feature extraction, and multiclass brain tumor classification. *J. Digit. Imaging* **2013**, *26*, 1141–1150. [[CrossRef](#)]
15. Bergner, N.; Romeike, B.F.M.; Reichart, R.; Kalff, R.; Krafft, C.; Popp, J. Tumor margin identification and prediction of the primary tumor from brain metastases using FTIR imaging and support vector machines. *Analyst* **2013**, *138*, 3983–3990. [[CrossRef](#)]
16. Glotsos, D.; Tohka, J.; Ravazoula, P.; Cavouras, D.; Nikiforidis, G. Automated diagnosis of brain tumours astrocytomas using probabilistic neural network clustering and support vector machines. *Int. J. Neural Syst.* **2005**, *15*, 1–11. [[CrossRef](#)]
17. Vapnik, V. *The Nature of Statistical Learning Theory*; Springer: New York, NY, USA, 2000; Available online: <https://www.springer.com/gp/book/9780387987804> (accessed on 14 July 2019).
18. Suykens, J. *Least Squares Support Vector Machines*; Suykens, J.A.K., Ed.; World Scientific: River Edge, NJ, USA, 2002; ISBN 978-981-238-151-4.
19. Ho, T.K. Random decision forests. In Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, QC, Canada, 14–16 August 1995; Volume 1, pp. 278–282.

20. Wong, K.C.; Chen, J.; Zhang, J.; Lin, J.; Yan, S.; Zhang, S.; Li, X.; Liang, C.; Peng, C.; Lin, Q.; et al. Early Cancer Detection from Multianalyte Blood Test Results. *iScience* **2019**, *15*, 332–341. [[CrossRef](#)]
21. Cai, H.; Verma, R.; Ou, Y.; Lee, S.; Melhem, E.R.; Davatzikos, C. Probabilistic Segmentation of Brain Tumors Based on Multi-Modality Magnetic Resonance Images. In Proceedings of the 2007 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, Arlington, VA, USA, 12–15 April 2007; pp. 600–603.
22. Verma, R.; Zacharaki, E.I.; Ou, Y.; Cai, H.; Chawla, S.; Lee, S.K.; Melhem, E.R.; Wolf, R.; Davatzikos, C. Multiparametric tissue characterization of brain neoplasms and their recurrence using pattern classification of MR images. *Acad. Radiol.* **2008**, *15*, 966–977. [[CrossRef](#)]
23. Ruan, S.; Lebonvallet, S.; Merabet, A.; Constans, J. Tumor Segmentation from a Multispectral Mri Images by Using Support Vector Machine Classification. In Proceedings of the 2007 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, Arlington, VA, USA, 12–15 April 2007; pp. 1236–1239.
24. Bauer, S.; Nolte, L.P.; Reyes, M. Fully automatic segmentation of brain tumor images using support vector machine classification in combination with hierarchical conditional random field regularization. In Proceedings of the 14th International Conference on Medical Image Computing and Computer-Assisted Intervention, Toronto, ON, Canada, 18–22 September 2011; Volume 14, pp. 354–361.
25. Zikic, D.; Glocker, B.; Konukoglu, E.; Criminisi, A.; Demiralp, C.; Shotton, J.; Thomas, O.M.; Das, T.; Jena, R.; Price, S.J. Decision forests for tissue-specific segmentation of high-grade gliomas in multi-channel MR. In Proceedings of the 15th International Conference on Medical Image Computing and Computer-Assisted Intervention, Nice, France, 1–5 October 2012; Volume 15, pp. 369–376.
26. Geremia, E.; Menze, B.; Prastawa, M.; Weber, M.-A.; Criminisi, A.; Ayache, N. Brain Tumor Cell Density Estimation from Multi-modal MR Images Based on a Synthetic Tumor Growth Model. In Proceedings of the Second International Conference on Medical Computer Vision: Recognition Techniques and Applications in Medical Imaging, Nice, France, 5 October 2012.
27. Charron, O.; Lallement, A.; Jarnet, D.; Noblet, V.; Clavier, J.B.; Meyer, P. Automatic detection and segmentation of brain metastases on multimodal MR images with a deep convolutional neural network. *Comput. Biol. Med.* **2018**, *95*, 43–54. [[CrossRef](#)]
28. Perez, U.; Arana, E.; Moratal, D. Brain Metastases Detection Algorithms in Magnetic Resonance Imaging. *IEEE Lat. Am. Trans.* **2016**, *14*, 1109–1114. [[CrossRef](#)]
29. Wachinger, C.; Golland, P. Atlas-Based Under-Segmentation. In Proceedings of the 17th International Conference on Medical Image Computing and Computer-Assisted Intervention, Boston, MA, USA, 14–18 September 2014; Volume 17, pp. 315–322.
30. Losch, M. Detection and Segmentation of Brain Metastases with Deep Convolutional Networks. Master's Thesis, KTH Royal Institute of Technology in Stockholm, Stockholm, Sweden, 2015.
31. Kamnitsas, K.; Ledig, C.; Newcombe, V.F.J.; Simpson, J.P.; Kane, A.D.; Menon, D.K.; Rueckert, D.; Glocker, B. Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Med. Image Anal.* **2017**, *36*, 61–78. [[CrossRef](#)]
32. Liu, Y.; Stojadinovic, S.; Hrycushko, B.; Wardak, Z.; Lau, S.; Lu, W.; Yan, Y.; Jiang, S.B.; Zhen, X.; Timmerman, R.; et al. A deep convolutional neural network-based automatic delineation strategy for multiple brain metastases stereotactic radiosurgery. *PLoS ONE* **2017**, *12*, e0185844. [[CrossRef](#)]
33. Grøvik, E.; Yi, D.; Iv, M.; Tong, E.; Rubin, D.; Zaharchuk, G. Deep learning enables automatic detection and segmentation of brain metastases on multisequence MRI. *J. Magn. Reson. Imaging* **2019**. [[CrossRef](#)]
34. Milletari, F.; Navab, N.; Ahmadi, S.A. V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. In Proceedings of the 2016 Fourth International Conference on 3D Vision (3DV 2016), Stanford, CA, USA, 25–28 October 2016; pp. 565–571.
35. Gonella, G.; Binaghi, E.; Nocera, P.; Mordacchini, C. Semi-Automatic Segmentation of MRI Brain Metastases combining Support Vector Machine and Morphological Operators. In Proceedings of the 11th International Conference on Neural Computation Theory and Applications, Vienna, Austria, 17–19 September 2019.
36. Statnikov, A.; Wang, L.; Aliferis, C.F. A comprehensive comparison of random forests and support vector machines for microarray-based cancer classification. *BMC Bioinform.* **2008**, *9*, 319. [[CrossRef](#)]
37. Litjens, G.; Kooi, T.; Bejnordi, B.E.; Setio, A.A.A.; Ciompi, F.; Ghafoorian, M.; van der Laak, J.A.W.M.; van Ginneken, B.; Sánchez, C.I. A Survey on Deep Learning in Medical Image Analysis. *Med. Image Anal.* **2017**, *42*, 60–88. [[CrossRef](#)]

38. Schölkopf, B.; Smola, A.J. *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*; MIT Press: Cambridge, MA, USA, 2002; ISBN 978-0-262-19475-4.
39. Tuceryan, M.; Jain, A.K. Texture Analysis. In *Handbook of Pattern Recognition and Image Processing (Vol. 2)*; Academic Press, Inc.: Orlando, FL, USA, 1994; pp. 207–248, ISBN 978-0-12-7745461-9.
40. 3D Deep Learning: Lung Tumor Segmentation-File Exchange-MATLAB Central. Available online: <https://www.mathworks.com/matlabcentral/fileexchange/71521-3-d-deep-learning-lung-tumor-segmentation> (accessed on 14 July 2019).
41. Sudre, C.H.; Li, W.; Vercauteren, T.; Ourselin, S.; Cardoso, M.J. Generalised Dice overlap as a deep learning loss function for highly unbalanced segmentations. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*; Springer: Cham, Switzerland, 2017; Volume 10553, pp. 240–248.
42. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. *arXiv* **2015**, arXiv:1512.03385.
43. ImageNet Classification with Deep Convolutional Neural Networks BibSonomy. Available online: <https://www.bibsonomy.org/bibtex/2886c491fe45049fee3c9660df30bb5c4/albinzehe> (accessed on 14 July 2019).
44. The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)-IEEE Journals & Magazine. Available online: <https://ieeexplore.ieee.org/document/6975210> (accessed on 14 July 2019).
45. Bouix, S.; Martin-Fernandez, M.; Ungar, L.; Nakamura, M.; Koo, M.-S.; McCarley, R.W.; Shenton, M.E. On Evaluating Brain Tissue Classifiers without a Ground Truth. *NeuroImage* **2007**, *36*, 1207–1224. [[CrossRef](#)]
46. Dice, L.R. Measures of the Amount of Ecologic Association Between Species. *Ecology* **1945**, *26*, 297–302. [[CrossRef](#)]



© 2019 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 (<http://creativecommons.org/licenses/by/4.0/>).