Segmentation-Enhanced Registration Methods to Support Image-Guided Treatments of Brain Tumors

von

Luca Canalini

Dissertation
zur Erlangung des akademischen Grades eines
Doktors der Ingenieurwissenschaften
- Dr.-Ing. -
vorgelegt im Fachbereich 3
Mathematik und Informatik
der Universität Bremen
November 2023

Gutachter:
Prof. Dr. Horst Hahn, Universität Bremen
Prof. Dr. Mattias Heinrich, Universität zu Lübeck

Datum des Promotionskolloquiums: 12. April 2024

I confirm that this doctoral thesis is my own work and I have documented all sources and material used.

Augsburg, 03.11.2023

This work was conducted at Fraunhofer Institute for Digital Medicine MEVIS

Acknowledgments

This thesis is based on my work, but it would not be possible without the support of several people. First of all, I would like to thank Horst Hahn, my PhD supervisor, for his guidance and recommendations to improve my thesis. I would also like to thank Jan Klein, my scientific mentor at Fraunhofer MEVIS, for his continuous support and interesting discussions about my work. I am very grateful to both of them for accepting me to work at Fraunhofer MEVIS, where I had the opportunity to interact with experts from different fields of medical imaging and to meet amazing and supportive colleagues. In this context, I would like to thank Ron Kikinis for his advice during the early stages of my PhD. I would also like to thank all my colleagues, especially Constantin, Annika, Stefan, Alessa, and Christoph, for their support and for sharing good times in the office and on business trips. Besides my colleagues at Fraunhofer MEVIS, I would also like to thank Dorothea Miller for supporting my research and showing me a clinical application of 3D ultrasound in neurosurgery. A special thanks goes to Prof. Mattias Heinrich for reviewing my thesis.

I would also like to thank all my colleagues from the TRABIT project, funded by the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreement No 765148. I am grateful to have been part of this European project and to have exchanged ideas and experiences with people from all over Europe. Thank you also for the good time we spent together in the different retreats. A special mention goes to Ezequiel, with whom I shared my research path since our Master's degree, and to Amnah, my TRABIT colleague at Fraunhofer MEVIS, for her support during the new chapter of my life in Bremen. Furthermore, I would like to thank all academic, clinical, and industrial partners who supported my research in the TRABIT project, especially Jan Kirschke and Benedikt Wiestler from Klinikum Rechts der Isar and Nuno Pedrosa de Barros and Diana Sima from icometrix, for their guidance and advice during my secondments.

Finally, I would like to thank my friends in Bremen and Augsburg: Without you, my journey in Germany would not be the same. And thanks to my family and friends in Italy, who always believed in me.

Abstract

This thesis presents *segmentation*-enhanced *registration* methods to support *image-guided treatments* of brain tumors. The solutions are proposed for magnetic resonance imaging (MRI) and ultrasound acquisitions.

Multiparametric MRI data are acquired before and after neurosurgery. To accurately detect any pathological remnants or regrowth, a comparison of successive acquisitions is often performed, which can be further improved by the application of automatic registration methods. Such solutions are required to address the lack of one-to-one correspondence between pre- and post-operative acquisitions, which is often caused by the presence of pathology, and the need to identify the optimal MRI sequence to guide the registration process. This thesis proposes two automatic solutions for registering pre- and post-operative volumes, an iterative method and a deep learning-based approach. In both algorithms, segmentation masks can be used to select the voxels of non-corresponding pathological tissues, whose contribution is discarded from the registration process. A quantitative analysis of the impact of pathology exclusion on the registration methods is conducted. Furthermore, an evaluation of the influence of multiple MRI sequences on the registration result is presented. In addition, a deep learning method for the segmentation of resection cavities in multi-parametric MRI volumes is proposed. The influence of different MRI sequences on the segmentation of these structures is evaluated. The masks generated by the segmentation method are used in the aforementioned registration solutions.

During tumor resection procedures, multiple intraoperative ultrasound volumes are acquired at different stages to provide continuous imaging of the ongoing surgery. However, due to the *brain shift*, which is a displacement of brain tissue from its pre-operative position, there is a need to compare successive acquisitions to track the changes during the surgery. Automatic registration methods can be employed to facilitate this task, but they have to account for the lack of correspondence between successive acquisitions. Therefore, this thesis proposes two iterative segmentation-enhanced registration methods to compensate for the brain shift in intra-patient ultrasound volumes. The first method involves automatically segmenting masks of corresponding healthy structures, which are then used to guide the registration of

successive acquisitions. In the second method, masks of non-corresponding resection cavities are segmented, and subsequent ultrasound volumes are registered by excluding the contribution of the segmented structures from the search for correspondences.

Kurzfassung

Diese Arbeit präsentiert *Registrierungsmethoden* im Kontext der bildgeführten Behandlung von Hirntumoren, die durch die zusätzliche Nutzung von *Segmentierungen* hinsichtlich ihrer Genauigkeit verbessert werden konnten. Die Lösungen werden für Ultraschall- und Magnetresonanztomographie-Volumina vorgeschlagen.

Multiparametrische Magnetresonanztomographie-Daten werden vor und nach der neurochirurgischen Behandlung erfasst. Um Tumorreste und erneutes Tumorwachstum besser identifizieren zu können, ist es vorteilhaft, longitudinale Aufnahmen visuell zu vergleichen. Automatische Registrierungsmethoden können diese Aufgabe erleichtern. In diesem Kontext schlägt diese Arbeit zwei Arten von automatischen Methoden für die Registrierung von prä- und postoperativen Volumina vor: eine iterative Lösung und ein *Deep Learning* Ansatz. Die Registrierung von multiparametrischen, longitudinalen Daten ist aufgrund i) des Fehlens einer Entsprechung zwischen prä- und postoperativen Aufnahmen und ii) der Notwendigkeit, die beste Sequenz zur Führung des Registrierungsprozesses auszuwählen, eine Herausforderung. Aus diesem Grund führt diese Arbeit eine quantitative Analyse durch, die untersucht, wie der Ausschluss von Pathologien Registrierungsmethoden beeinflusst. Darüber hinaus wird eine Bewertung des Einflusses mehrerer MRT-Sequenzen auf das Registrierungsergebnis präsentiert. Ferner wird eine Deep-Learning-Methode zur Segmentierung von Resektionshöhlen in multiparametrischen MRT-Volumina beschrieben. Der Einfluss verschiedener MRT-Sequenzen auf die Segmentierung dieser Strukturen wird bewertet. Die Ergebnisse der vorgeschlagenen Segmentierungsmethode werden in den Registrierungslösungen angewandt.

Mehrere intraoperative Ultraschallvolumina werden zu verschiedenen Zeitpunkten der Tumorresektionsverfahren zur Verfügung gestellt, um eine aktualisierte Bildgebung der laufenden Operation zu liefern. Aufgrund von *Brain Shift* besteht die Notwendigkeit, aufeinanderfolgende Aufnahmen visuell zu vergleichen. Diese Aufgabe kann durch automatische Registrierungsmethoden verbessert werden, die das Fehlen einer Entsprechung zwischen aufeinanderfolgenden Aufnahmen berücksichtigen müssen. Daher werden in dieser Arbeit zwei iterative, durch Segmentierung geführte Registrierungsmethoden vorgeschlagen, um den Brain Shift von Ultraschallvolumina

des gleichen Patienten zu kompensieren. In der ersten Lösung werden Masken von entsprechenden gesunden Strukturen automatisch segmentiert und dann verwendet, um die Registrierung aufeinanderfolgender Aufnahmen zu anzuleiten. Die zweite Methode besteht ebenfalls aus zwei Schritten: Masken von nicht entsprechenden Resektionshöhlen werden segmentiert und aufeinanderfolgende Ultraschallvolumina werden registriert, indem der Beitrag der segmentierten Strukturen von der Suche nach Entsprechungen ausgeschlossen wird.

Contents

1. Introduction

Medical images are obtained at multiple stages of brain tumor treatments. The visual inspection of successive acquisitions helps differentiate pathology from healthy tissues [1] and, thus, contributes to providing more accurate care [2]. In neurosurgery, intraoperative US (iUS) volumes obtained towards the end of the resection can be visually compared with iUS data acquired at earlier stages of the procedure, to precisely identify the remaining pathological tissues [3]. Moreover, the visual examination of pre- and postoperative magnetic resonance imaging (MRI) data can help distinguish between pathology and imaging artifacts induced by post-surgical treatments, such as radiotherapy [4]. Yet, finding correspondences in the visual inspection of subsequent acquisitions is challenging: The conformation and position of brain structures in successive images get modified by brain shift and post-surgical treatments [5; 6]. As a solution, image-based registration algorithms can ease the visual examination by automatically establishing one-to-one correspondences in subsequent acquisitions [3; 7; 8; 9]. However, image registration of neurosurgical data is hampered by the fact that a one-to-one correspondence cannot be guaranteed for all the voxels. Tumors have no matching tissues in successive acquisitions [10], since they are only visible in preoperative stages, not in the surgery's late stages or after the removal. A strategy to solve this challenge is to limit the search for correspondences, conducted by registration algorithms, to brain tissues in common between successive images. In this context, this thesis proposes *segmentation-enhanced registration* algorithms, that utilize masks of pathological or healthy tissues to specify the voxels in which correspondences between two successive scans are guaranteed. The proposed methods are applied to register pre- and post-operative MRI acquisitions and intraoperative ultrasound volumes. Moreover, since multiple sequences are usually obtained in MRI acquisitions, this thesis also quantitatively investigates the influence of multiple MRI sequences on the registration of pre- and post-operative acquisitions, and on an automatic method for segmentation resection cavities.

The following introduction provides the clinical and technical information for better comprehending the solutions proposed in the thesis. First, section 1.1 gives an overview of image-guided neurosurgery for brain tumors, and of the multiple imaging acquisitions obtained in brain tumor treatments. Second, sections 1.2 and 1.3 introduce provide the main concepts of image registration and image segmentation, to better comprehend the segmentation-enhanced registration methods described later in this thesis. Moreover, since several methods make use of convolutional neural networks, an introduction to deep learning is given in section 1.4. Lastly, an overview of the image modalities used in the experiments described in this thesis is provided in section 1.5.

1.1. Image-Guided Neurosurgery

Neurosurgery is the main treatment for brain tumors (1.1.1), and comprehend all those procedures aimed at the removal of pathological tissues [11; 12]. Image-guided neurosurgery integrates the use of imaging data at the different surgical steps [13], including preoperative planning, resection, and postoperative steps (1.1.2).

1.1.1. Brain Tumors

Brain tumors are abnormal masses of cells growing in the human cranium [14]. They can be distinguished between primary tumors, that originate from the brain, or secondary (metastatic) ones, generated in other parts of the body and then spreading to the human brain. Intracranial tumors can be extremely debilitating for diagnosed patients [15], who are exposed to a wide variety of symptoms, including headache, nausea, vision problems, inability to accomplish simple tasks, hearing problems, etc. The most common treatments to mitigate or eliminate the negative effects of tumors include surgical resection, radiotherapy, and chemotherapy [16; 17]. A combination of them can also be used. In all the stages, medical images are utilized to support the treatment's decisions.

1.1.2. Multiple Imaging Acquisitions in Neurosurgery

The first step in neurosurgery is surgical planning. It consists of delineating the tissues that have to be removed and is conducted on brain images acquired before surgery [18]. MRI acquisitions are commonly acquired since they provide a good

image contrast between tumors and healthy tissues. Several MRI sequences are usually obtained for a better analysis of the pathology [4]. Because accurate and complete removal of the pathological tissues correlates with high survival rates of the operated patients [19; 17], preoperative data can also be made available during surgery for making tumor resection as precise as possible [18]. This is achieved by utilizing neuronavigation systems, consisting of computer-assisted techniques to guide (or navigate) surgeons during neurosurgical procedures [20]. Commercial systems utilize electromagnetic or optical technology to track the positions of specific landmarks placed on the head of the operated patients. Then, the corresponding locations are identified in the image data, and an image-to-patient registration is computed [3]. Thanks to this transformation, neuronavigation systems register preoperative image data to the patient's head observable in the surgical scene [21]. The positions of surgical tools can also be tracked and mapped to the preoperative data, and each pin-pointed location within the skull of the patient can be related to its corresponding area in the imaging data acquired before resection [22]. However, the use of preoperative data during resection is degraded due to a phenomenon known as *brain shift* [23]. It consists of anatomical changes caused by the surgical procedure, which modify the conformation and the positions of brain tissues, especially those proximal to the resection area [24]. As a consequence, the pre-surgical images do not anymore show what is actually observed in the surgical scene. The longer the surgery is performed, the more evident brain shift effects become, and, thus, the less reliable preoperative imaging is to support the removal of tumors. As a solution, intraoperative images can be acquired [25; 26].

Intraoperative images provide an updated view of the surgical scenario and, thus, a better identification of the remaining tumor can be achieved. Multimodal data, including magnetic resonance imaging data and ultrasound images, can be acquired to identify the pathology [27]. In particular, ultrasound (US) imaging demonstrated to be a valid option for supporting neurosurgical procedures, because it provides a good image contrast between pathological and healthy tissues and is relatively fast to be acquired [28]. Thanks to these advantages, US data can be obtained at multiple intraoperative stages [29], to repeatedly have an up-to-date imaging visualization [25]. However, images obtained towards the end of the resection are difficult to interpret [3]. In fact, in the late stages of neurosurgical procedures, several artifacts affect the image quality of the acquired data [3]. They are due, for example, to possible inflammations around the resection area, or saline solutions

used to fill the cavity left by the removal [30]. To improve the interpretation of intraoperative images obtained at later stages of the resection, they can be visually compared with earlier acquisitions [29]. Nevertheless, the visual comparison of subsequent intraoperative data is challenging: The resection procedure removes and deforms the brain tissues, which have different positions and conformations in subsequent acquisitions. Only experienced neurosurgeons can mentally compensate for the anatomical changes caused by the brain shift.

After the surgical removal is concluded, postoperative MRI acquisitions are usually obtained to locate any pathological remnants [31], and plan postsurgical adjuvant treatments, such as radiotherapy. To better identify tumor regrowth and possible residuals in postoperative data, a visual inspection of preoperative and postoperative MRI data can be carried out [4]. However, this task is challenging. In preoperative data, the pathology is present, whereas in postoperative images the resection cavity is visible. Furthermore, the brain tissues surrounding the removal point are deformed due to the changes in the resection cavity, which tends to modify its shape and conformation in postoperative stages, also due to radiotherapy treatments.

1.1.3. Image-Based Registration in Brain Tumor Treatments

In brain tumor treatments, there is often the need of comparing intra- and postoperative data with acquisitions obtained before surgery. However, the visual inspection of subsequent acquisitions is time-consuming and mentally exhausting, especially for non-experienced medical personnel. As a solution, the search for visual correspondence between subsequent acquisitions can be supported by neuronavigation systems. They rigidly register subsequent acquisitions by using the corresponding locations in the surgical scene and in the imaging data. However, these systems often compute only a rigid transformation, which is not accurate enough to model the local deformations induced by the surgical resection. Deformable image-based registration algorithms compute voxel-wise transformations and can generate more realistic transformations to make the visual inspection of subsequent acquisitions more precise. They aim to reduce the dissimilarity between the two images, by establishing one-to-one correspondences between two subsequent acquisitions. The template (moving) image is registered to be as close as possible to the reference (fixed) image (more details in Section 1.2).

The design of automatic methods registering neurosurgical subsequent data is challenging, due to the fact there is no guarantee of one-to-one correspondence for all the voxels in subsequent acquisitions. Segmentation masks can help to guarantee that the search for correspondences is conducted on brain tissues that remain visible in the images to register.

1.2. Segmentation-Based Image Registration

This thesis proposes segmentation-enhanced registration methods to improve the visual comparison of intraoperative ultrasound and pre- and post-operative MRI volumes. The following chapter provides an overview of the main concepts of image registration and is based on the theory presented in [32; 33]. It will be useful to better comprehend the solutions proposed in the thesis.

1.2.1. Introduction to Image Registration

Pair-wise image registration algorithms aim to transform two sets of data to the same reference system by establishing correspondences between the images (see Fig. 1.1). In case of volumetric data, let $\mathcal{R}$, $\mathcal{T} : \mathbb{R}^3 \to \mathbb{R}$ denote a reference and a template volume. The goal of image registration is to generate a transformation $y : \Omega \to \mathbb{R}^3$ that aligns the volumes $\mathcal{R}$ and $\mathcal{T}$ on the field of view $\Omega \subset \mathbb{R}^3$ such that $\mathcal{R}(x)$ and $\mathcal{T}(\mathrm{y}(x))$ are similar for $\mathrm{x} \in \Omega$. The distance measure is used to compute the dissimilarity (or similarity, according to the formulation of the loss function) between the warped template ($\mathcal{T}(\mathrm{y}(x))$) and reference ($\mathcal{R}(x)$) image. It has typically the form:

$$D(\mathcal{R}, \mathcal{T}(y)) = \int_{\Omega} d(\mathcal{R}(x), \mathcal{T}(y(x)) \, \mathrm{d}x \tag{1.1}$$

Image registration can be modeled as the minimization of a loss function $\mathcal{J}$, composed of the distance measure. The deformation computed by an image registration algorithm represents one of its minimizers:

$$\mathcal{J}(\mathcal{R}, \mathcal{T}, y) = \mathcal{D}(\mathcal{R}, \mathcal{T}(y)) \tag{1.2}$$

In the registration process, the transformation parameters composing y are updated to reduce the value of the loss function. An optimization algorithm is used

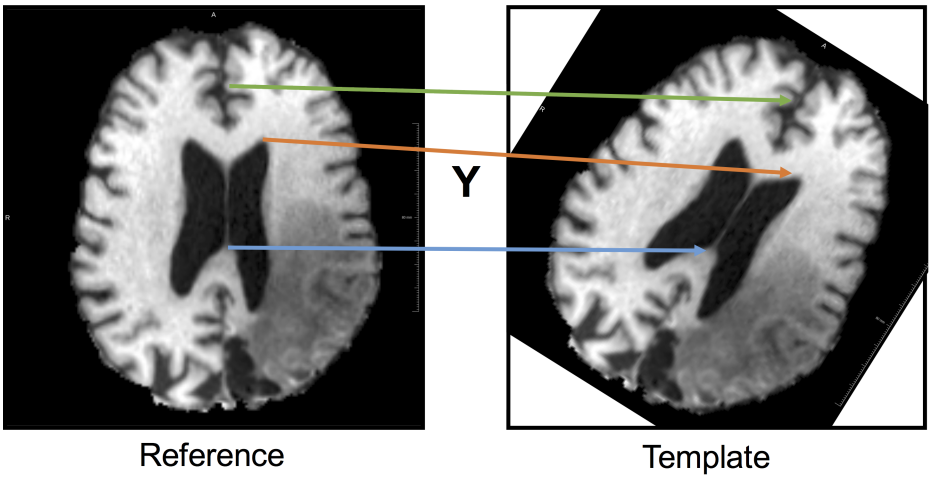


Figure 1.1.: Image registration. Image registration aims to map each point of the reference image to the corresponding one in the template data by computing a transformation y.

to guide the minimization process: The steps of updating the transformation parameters and computing the distance measure are repeated until some convergence criteria are met.

Transformation Model

Image registration algorithms can be classified into two groups, rigid and non-rigid (deformable), according to the degrees of freedom of the transformation y computed to register the template volume to the reference one. Rigid registration includes those transformations that can be expressed with a limited number of parameters, such as translation, rotation, scaling, and shear. It is also defined as parametric registration and provides a global deformation because the same transformation is applied to each voxel of the template image. Rigid registration algorithms cannot solve complex scenarios, in which image voxels have to be locally transformed. For this purpose, non-rigid solutions are more suited, because they can locally deform different areas of the template volume. These transformations have a large number of parameters and, thus, this type of registration is also called non-parametric. Their output is composed of a dense deformation field, which contains unique deformation

parameters for each voxel position.

Regularization Term Deformable image registration is an ill-posed problem: When searching for a deformation field, multiple solutions exist, and some of them might not be anatomically unrealistic [34]. The search for deformation fields can be limited to plausible transformations by introducing in the loss function an extra term, the so-called regularizer. The regularization term usually favors the smoothness of the deformation field and discourages unrealistic transformations [35]. The loss function for deformable image registration in which a regularizer term $\mathcal{R}$ is utilized is defined as follows:

$$\mathcal{J}(\mathcal{R}, \mathcal{T}, \mathrm{y}) = \mathcal{D}(\mathcal{R}, \mathcal{T}(\mathrm{y})) + \alpha R(\mathrm{y}) \tag{1.3}$$

The hyperparameter α controls the strength of the regularization term.

Feature-Based Registration

Image registration algorithms can also be divided into intensity- and feature-based methods, depending on how the distance measure is computed [36]. In intensity-based registration, the distance measure is directly computed on the intensity values of the two images. For example, the sum of squared differences computes a pixel-wise intensity difference between the fixed (reference) and moving (template) image [37]. Besides, other distance measures compute correspondences on features extracted from one or both images. In features-based registration, features taken into account for computing the distance measure can be of different types, such as landmarks, segmentation, or results of some image filtering applied to the images.

1.2.2. Segmentation-Enhanced Registration in Brain Tumor Treatments

The registration of subsequent brain volumes acquired at different stages of tumor treatments is challenging because corresponding structures are often not available in both acquisitions. In fact, the pathology visible in preoperative data gets removed, whereas the resection cavity is observable in postoperative data. Segmentation-enhanced algorithms are particularly suited to deal with the lack of correspondence between subsequent neurosurgical acquisitions. Anatomical or pathological

structures can be segmented in one or both images, and the generated masks are then used in the registration process. The use of the segmentation can be dual, as Fig. 1.2 clarifies. First, the segmentation of the pathology can help to determine the voxels on which the similarity measure shouldn't be computed. Masks of the non-corresponding structures are used to exclude from the computation of the distance measure the contribution of non-matching tissues (see Fig. 1.2b). Therefore, the algorithm can focus on the registration of the corresponding healthy structures, for which a one-to-one correspondence is guaranteed. By indicating the non-correspondences as Σ, the distance measure 1.1 can be expressed as follows:

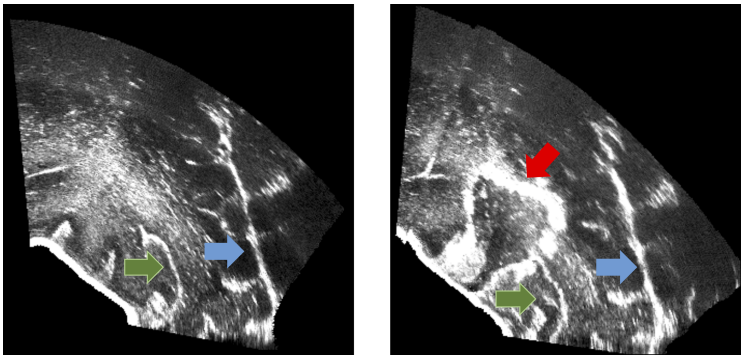
$$D(\mathcal{R}, \mathcal{T}(\mathrm{y})) = \int_{\Omega \setminus \Sigma} d(\mathcal{R}(x), \mathcal{T}(\mathrm{y}(x)) \, \mathrm{d}x \tag{1.4}$$

Second, the registration of subsequent images can also be performed by solely using features extracted from corresponding structures. Anatomical matching elements can be segmented, and the registration only uses the generated masks to compute the distance measure (see Fig. 1.2c).

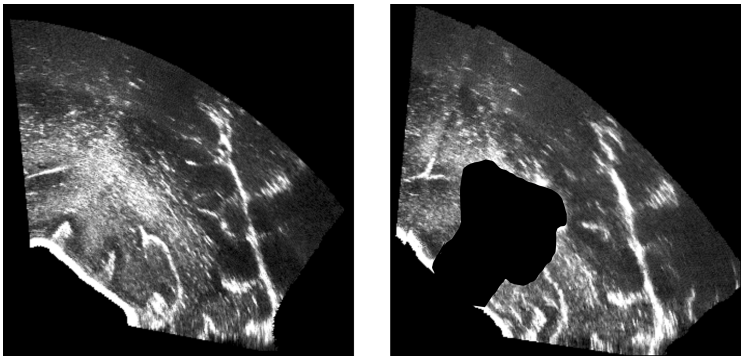
This thesis proposes and evaluates different segmentation-driven deformable registration methods. It demonstrates how automatic image segmentation algorithms can support the registration of subsequent neurosurgical images.

1.3. Image Segmentation

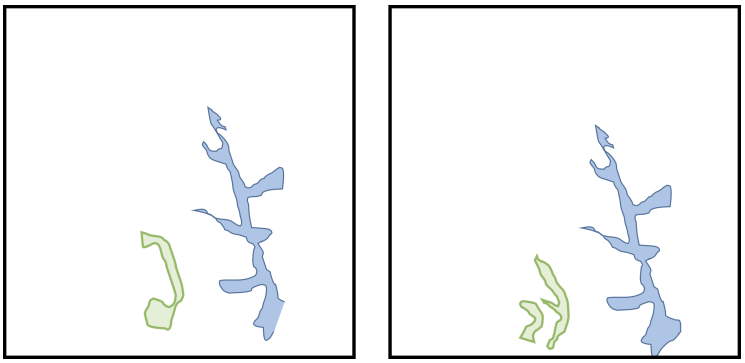
As the previous section explains, image registration algorithms for neurosurgical data can benefit from using segmentation masks to compute the distance measure only on specific structures. Image segmentation is the process of delineating specific structures in imaging data. It can be seen as a voxel-wise classification, in which a class is assigned to each voxel of an image. In particular, the image gets divided into different regions: The pixels identified by the segmentation process are labeled as foreground, whereas the rest is the background. Image segmentation is used to make the images more understandable, and the identification of specific objects clearer (see Fig. 1.3).



(a) Corresponding (blue and gree arrows) and non-corresponding structures (red arrow).



(b) Exclusion of the non-corresponding elements.



(c) Masks of corresponding elements.

Figure 1.2.: Complementary uses of segmentation to support image registration of neurosurgical ultrasound data. Registration algorithms can use the original intensity volumes and exclude the contribution of the non-corresponding elements (Fig. 1.2b) from the computation of the distance measure. Besides, in other scenarios, the distance measure can focus only on the masks of corresponding structures (Fig. 1.2c).

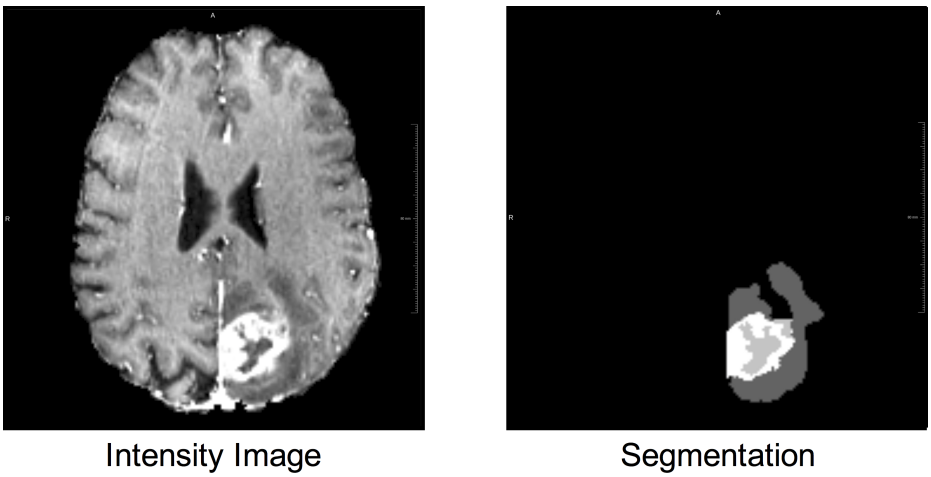


Figure 1.3.: Image segmentation. Image segmentation aims to classify each voxel of the input image as background or foreground. In this example, the pathological tissues of an MRI image are segmented in different labels, whereas the rest of the volume is classified as background.

1.4. Convolutional Neural Networks

The last decade has been witnessing a tremendous increase in the use of convolutional neural networks (CNNs) to solve numerous tasks in medical imaging. The solutions presented in this thesis to tackle segmentation and registration tasks in neurosurgical procedures also utilize CNNs. Therefore, an introduction to CNNs, based on [38; 39], is here provided.

1.4.1. Artifical Neural Networks

CNNs are a particular type of artificial neural network (ANN) used in image processing and recognition tasks. ANNs are computational systems inspired by the way the biological nervous system works. Traditional machine learning techniques make use of hand-crafted features to learn how to solve a specific task. The main characteristic of ANNs is the ability to automatically define and extract such features from the input data. The feature extraction is possible thanks to the so-called *artificial neurons*, computational nodes having incoming and outgoing connections with other neurons. In the basic structure of an ANN (visible in Fig. 1.4), an input vector is passed to the input layer, which distributes it to hidden layers. In ANNs, each neuron in a hidden layer is connected to the neurons of the adjacent layers. A weight

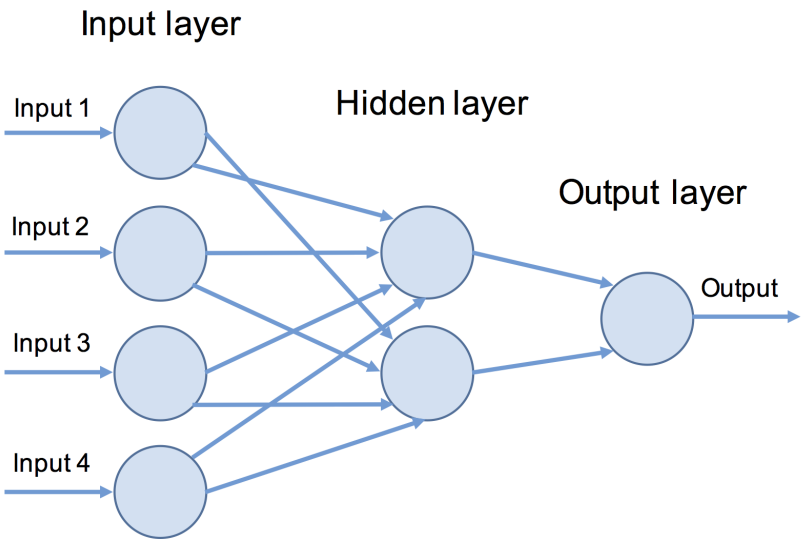


Figure 1.4.: General scheme for a three layered ANN.

is associated with every connection to a node so that the neuron can perform a weighted sum of the incoming connections. The sum is then passed to an activation function. An additional input, called bias, has a value of 1 and associated weight and is added to the sum. The final weighted sum is then passed to the next neuron. Thus, ANNs are composed of a large number of interconnected neurons, whose parameters (or weights) are learned from input data and continuously adapted in a learning phase [39].

The values of weights of the different neurons are repetitively adjusted thanks to a learning (or training) process, carried out to minimize a loss function, which quantitatively quantifies how close the generated outcome is to the desired output. The training process can be divided into two main phases. First, in the forward phase, images are passed through the network, in which input layers initially compute weighted sums of the pixel values. Their output is then passed to hidden layers, which compute weighted sums of the incoming values and, thus, extract features at different levels. Eventually, a certain output (or prediction) is delivered as a result of the output layers. The generated outcome represents a non-linear combination of the image values and the weights of the neurons of the model. Second, in the backward phase, the parameters of the models are automatically updated to improve the generated output and minimize the loss function, using back-propagation algorithms and modified versions of stochastic gradient descent. This technique is repeated for several batches of images until the convergence of the loss function.

When the network is composed of multiple hidden layers, whose weights are learned in the training phase, the training process is called deep learning.

Limitations of Artificial Neural Networks

Two major limitations of ANN methods applied for image processing tasks rise when the input image size increases [39]. First, the processing of large images leads to high computational costs, due to the fact that large input images require more hidden layers and more weights to be trained. The second problem is related to a phenomenon known as overfitting, which happens when the network ineffectively learns, leading to poor performance. It is due to several reasons, including the high number of parameters. The fewer parameters used for training, the fewer chances of overfitting are encountered. Thus, large ANNs are also more likely to overfit. The difficulties of the use of the ANNs in image processing have been solved by the introduction of convolutional neural networks.

1.4.2. Convolutional Neural Networks

In CNNs, neurons have connections only to some of the neurons of the adjacent layers [39]. Consequently, the number of training weights is smaller and the computational issues of ANN methods are reduced. This allows the creation of larger models, able to solve more complex medical imaging tasks, without the exponential increase of model parameters.

CNNs extract spatially independent features [38]. In solving a specific task involving an image, it is not taken into consideration where the various objects are located in the image. The region of an image influencing the output of a specific neuron is limited to a neighborhood and takes the name of the receptive field [40].

Convolutional Neural Networks Layers

Convolutional neural networks are composed of three main types of layers: convolutional layers, pooling layers, and fully-connected layers [39]. In image processing, convolution is a mathematical operation that calculates for each pixel a weighted sum of its neighborhood. The so-called kernel is the matrix that is convolved with an image. According to the values (weights) of this matrix, different features are

extracted from the image. A convolutional layer is composed of many kernels extracting specific features from the input image. The learning parameters of a convolutional layer are the weights that compose the matrices of the kernels. According to the values composing the kernels, the convolutional layers can identify specific patterns in the input image. The size of a convolutional layer is influenced by the convolution matrix, with a size of k x k, but also by the input channels n and the desired output m: The final dimension of a convolutional layer can be expressed as k x k x m x n [40]. In CNNs, stacked convolutional layers are utilized, so that a feature hierarchy is created. Shallow layers are usually used to extract low-level features such as edges, whereas deeper layers are responsible for more complex combinations of the previous features.

Pooling layers are applied after convolutional ones for shrinking the dimensions of input activation maps. They are responsible for reducing the number of parameters and their complexity. They also allow features to be extracted at different scales of the input image.

Fully connected layers contain neurons that are connected to all the neurons of the previous layer.

Supervised and Unsupervised Learning

The goal of the training phase is to minimize a differentiable loss function, that evaluates the generated output w.r.t. the desired outcome. There are two main learning paradigms for CNNs, supervised and unsupervised learning. In the first paradigm, the learning process is conducted to minimize the differences between the automatically generated outputs and the desired outcome (the so-called ground truth). A measure of the differences between the desired and generated outputs is computed by a metric included in the loss function. Thus, in a supervised paradigm, ground truth annotations are available for each input sample of the training dataset, to evaluate the outcome of the algorithm. A drawback of this paradigm is that manual annotation of medical data requires time and effort, and is not always possible. This is especially true for image registration ground truth, where the identification of ground truth deformation fields is not always feasible [34]. On the contrary, unsupervised learning differs from the first paradigm because the training set has no need for annotated labels. In this paradigm, the learning process is successful whether the network can reduce or increase a loss function based on a metric (not

computed on ground truth), as, for example, the distance measure of an image registration framework. In this context, the loss function is based on the computation of the similarity between two images to register, for which no ground truth is needed.

When training convolutional neural networks, there are two different approaches to providing input images. The data can be supplied using its original dimensions. However, the larger the images are, the more computationally expensive the training process becomes. The use of large images also means an increasing number of parameters to process them. Thus, as a solution, the size of the images can be reduced. A rescaling of the original size can be performed. Another possibility, commonly used in CNN dedicated to image segmentation, is to perform a patch-based learning process. Instead of providing each image as a whole, it is divided into smaller parts (patches). The corresponding annotated data are also accordingly divided, and a batch of patches is fed to the CNN in every training iteration.

1.4.3. U-net

The organization of the layers of a CNN is defined as architecture. Different CNN architectures have been proposed according to the diverse image processing tasks. One of the most successful CNN architectures for image processing tasks (especially for image segmentation) is the U-Net [41]. This subsection describes the original U-net architecture and introduces the applications for which modified versions of this method have been used in this thesis.

The U-Net has been proposed in 2015 to tackle 2D image segmentation tasks and achieved state of art in many challenges [41]. It consists of an encoder-decoder architecture and, to better localize and segment objects, features from the contracting path are merged with the upsampled output. A convolutional layer can learn to combine the features coming from different depths. The segmentation of the images is performed by using an overlapping strategy of multiple patches. This patch strategy allows the application of the U-Net also to large images, which couldn't be fit to a GPU.

U-Net Architecture

The architecture of the U-Net consists of a contracting path (encoder) and an expansive path (decoder) (see Fig. 1.5). The first one comprehends repetitions of the

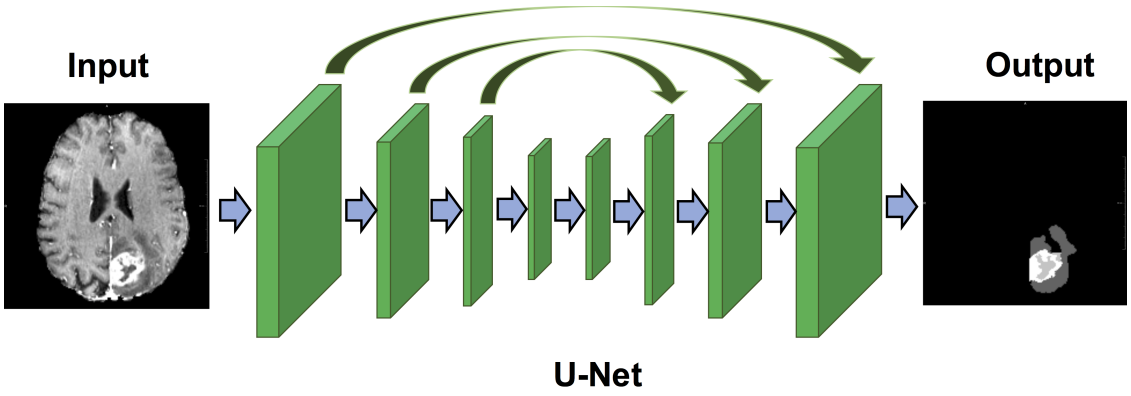


Figure 1.5.: Visualization of the general architecture of the U-net method used for segmentation of pathological tissues in brain images. The U-Net has an encoder-decoder architecture.

combination of two 3x3 convolutions, each followed by a rectified linear unit and a 2x2 max pooling operation with stride 2 for reducing the dimension. At each level, the channels of the features are doubled. The expansive path is composed of an upsampling of the feature map followed by a 2x2 convolution halving the feature channels, a concatenation with features maps obtained from the contracting path, and two 3x3 convolutions followed by a ReLu. Each component of the final feature vector of size 64 is mapped to the desired number of classes. The total number of convolutional layers is equal to 23.

U-Net: Beyond Segmentation

Following the success of the U-Net in 2D medical segmentation, a 3D version of this architecture has been proposed [42]. Moreover, in recent years the applications of modified versions of this architecture have been expanding to other medical imaging fields, including image registration [43]. This thesis uses modified implementations of the 3D architecture of the U-Net to tackle segmentation and registration tasks. More details about the specific modifications of the original architecture are provided in each chapter.

1.5. Medical Imaging Modalities

Images of the human brain are essential in many steps of intracranial tumor treatments [44; 45]. In particular, MRI and US imaging are useful to support neuro-

surgical procedures [29]. This thesis proposes image registration and segmentation solutions for MRI and iUS. An introduction to these image modalities is here provided.

1.5.1. Magnetic Resonance Imaging

MRI is a non-invasive imaging technique providing detailed information about the human anatomy, function, and metabolism. MRI technique generates cross-sectional images of the body without using ionizing (and, thus, dangerous) radiations. The human body is largely composed of water molecules, each with two atoms of hydrogen containing a proton and an electron [46]. Since protons are particularly sensitive to magnetic excitation, MRI scanners use a combination of static magnetic fields and radio waves to influence their orientation in the human body. After excitation, the protons tend to return to their status of equilibrium through different *relaxation* processes and produce radio waves. By measuring the relaxation signals emitted in a body area after a magnetic excitement, MRI images can be created.

In MRI scanners, a magnet generates a strong and homogeneous magnetic field and produces a strong magnetization in the tissues of a scanned patient. Then, a gradient system delivers linear and orthogonal magnetic fields, which are needed for spatial localization. In particular, magnetic excitements in X and Y dimensions select a specific slice of the volume of interest. Third, a radio frequency (RF) system selectively excites the protons within the selected slice. The coil used for the RF signal is also used to capture the relaxation of the nuclei when the RF magnetic field is turned off. The image contrasts observed in MRI images for various structures are influenced by the fact that human tissues have different hydrogen compositions, which influence their relaxation properties.
The RF coils are continuously turned on and off. Repetition Time (TR) is the amount of time between successive pulse sequences applied to the same slice. Time to Echo (TE) is the time between the delivery of the RF pulse and the receipt of the echo signal. The variation of TE and TR leads to different MRI protocols. In brain tumor treatments, the most common acquired MRI protocols are the T1-weighted (T1), contrast-enhanced T1 (T1-CE), T2-weighted (T2), and T2 Fluid Attenuated Inversion Recovery (FLAIR) [47]. Each protocol is useful to highlight specific characteristics of the human brain [48]. For instance, T1-CE is obtained by the injection of contrast-enhancing liquid (usually Gadolinium-based contrast),

which accumulates in the tumor tissues and, thus, enhances their image contrast in MRI images. Besides, the FLAIR protocol is useful to better visualize tumor-induced edema, which appears hyper-intense in the MRI images.

1.5.2. Ultrasound Imaging

Medical ultrasonography is an imaging technique based on the emission and reflection of ultrasound waves, that is, sound waves with a frequency between 20 kHz up to several gigahertz [49]. To comprehend how ultrasound images are formed, an ultrasound emitter positioned in front of the object to visualize can be taken as an example. Both items are contained in a medium, with a different acoustic impedance than the object. The emitter usually is a piezoelectric transducer that converts electric signals into ultrasound waves. The waves travel in the medium until they reach the object of interest. If the encountered element has an acoustic impedance greater than the first substance, the incoming waves are reflected and the so-called echo is formed. Thus, an outgoing wave departs from the reflecting element, and a portion of the original ultrasound waves will continue traveling. The backward signal is then detected by a receiver, which often is the same device used to emit the signals. The receiving piezoelectric instrument converts the incoming sound waves into voltage signals. These are then converted into digital signals, which are used to generate images.

Different types of US images can be generated, and one of the most used for medical diagnosis is the B-mode, where B stands for brightness [50]. It is a cross-sectional imaging showing boundaries of organs and tissues within the body [51]. It is generated from the US waves reflected at anatomical boundaries and the scattering caused by small irregularities between the different tissues. Each echo is visible as a point in the image. The brightness of each point displays the strength and amplitude of the echo. The position of each point in the image corresponds to the distance of the related object's point from the ultrasound transducer. The information about the location is obtained by using the pulse-echo sequence. Each pulse-echo sequence is used to generate a line in the B-mode image, showing the contours of the anatomical structures positioned on the wave trajectory [51].

Bidimensional ultrasound images are formed by collecting many B-mode lines [51], acquired by multiple transducers on the same probe. A linear probe, composed of an array of transducers, can be taken as an example. When one of the transducers is active, the others are off. It emits the ultrasound wave, and then receives the

produced echoes. The first line of the image is formed based on the returning echoes. This procedure is then repeated for all the transducers in the array until the 2D image is created. The usual time for a complete sweep is around 1/30th of a second. This allows the formation of images almost in real time and, thus, the application of such imaging techniques is suitable also for those medical scenarios in which medical images should be fast obtained, such as neurosurgery.

3D Ultrasound

One of the main difficulties of using 2D data as visualization support in neurosurgical operations is the need to mentally integrate bidimensional images into a 3D surgical scene. This task requires mental effort and time. 3D imaging represents a better option for visualizing the anatomical and pathological structures in neurosurgical procedures [52].

The most utilized techniques to reconstruct three-dimensional US images are mechanical scanning, free-hand scanning with position tracking, and free-hand without sensing. In free-hand scanning with position tracking, a transducer acquiring 2D US scans is converted into a 3D scanning by tracking its position and orientation [53]. A sensor, whose position can be tracked, is attached to the extremity of the US probe and, thus, the localization of the moving transducer is possible. There are two main tracking technologies, respectively based on magnetic and optical systems. An optical system detects the position and orientation of an object by identifying and tracking visual markers on its surface. An infrared light emitter determines the position of an object via active (transmitting its infrared signal) or inactive (reflected infrared light) markers. Inactive markers are usually colored with a highly reflective material. Multiple cameras positioned at different locations detect the marker from several perspectives, and its positions and orientations can be then triangulated. By tracking multiple markers, the location and orientation of each 2D acquisition become available. Eventually, several bidimensional images are used to build the 3D volume. The operator handling the ultrasound probe has to be careful in moving the transducer at an appropriate speed, to ensure that the following 2D images have no too large gap.

1.6. About This Thesis

The thesis presents automatic segmentation-enhanced registration solutions to improve neurosurgical procedures for tumor resection. It is divided into two parts, each related to methods proposed for a different image modality. A short overview of the two parts is given in this chapter. At the end of the thesis, a conclusion section is provided to discuss the proposed solutions and future work.

Part I: Magnetic Resonance Imaging The first part of the thesis focuses on the registration of subsequent multisequence magnetic resonance imaging data acquired in intracranial tumor treatments. Two types of automatic methods for registering pre- and postoperative MRI data are proposed: an iterative method and a CNN-based solution. The impact of excluding the pathological tissues from the distance measure in the registration algorithms is investigated. Moreover, an analysis of the influence of different MRI sequences on the registration performances is provided.
Furthermore, a segmentation method based on the 3D U-Net is proposed to automatically delineate the resection cavities in post-operative MRI data. It is investigated how much influence the different modalities have on the training process.

Part II: Ultrasound The second part proposes two automatic segmentation-based registration methods to register intraoperative ultrasound volumetric data acquired at different stages of neurosurgical procedures. The first solution is based on two steps. First, an automatic method based on 3D U-Net is developed for segmenting brain healthy structures. Second, an iterative method solely using the generated masks registers subsequent US data.
The second approach registers US volumes by excluding the contribution of the resection cavities. An automatic segmentation method based on 3D U-Net to segment the resection cavities in during- and postoperative acquisitions is developed. Then, thanks to the generated masks, these structures are excluded from the computation of the distance measures of the iterative method registering the US acquisitions.

Part I.

Image Registration and Segmentation Methods for Magnetic Resonance Imaging Volumes Acquired in Brain Tumor Treatments

2. Magnetic Resonance Imaging in Intracranial Tumor Treatments

Magnetic resonance imaging provides essential support at multiple stages of brain tumor treatments. Before neurosurgery, surgical planning is conducted on MRI volumes. At the end of the procedure, additional data are acquired to spot any tumor remnants and evaluate the surgical resection. Subsequently, MRI acquisitions are obtained at multiple post-operative stages to identify any pathological regrowth and measure the target volume for post-surgical treatments, such as radiation therapy [6].

Subsequent MRI volumes can be visually compared to improve intracranial tumor treatments. For example, the visual inspection of pre- and post-operative MRI data helps differentiate tumor recurrence from lookalike tissues, which can be induced by postsurgical treatments [54]. However, the comparison of pre- and post-operative MRI data is difficult, due to anatomical shifts induced by the resection and post-surgical treatment. Automatic image-based registration algorithms can facilitate the visual comparison of subsequent MRI acquisitions.

2.1. Image-Based Registration and Segmentation of MRI Volumes

The automatic registration of MRI volumes acquired at different stages of brain tumor treatments is challenging since tumors locally and heavily deform the surrounding intracranial tissues. Deformable registration solutions are preferable to rigid approaches to accommodate these anatomical changes since they can generate local and more realistic deformations. However, the design of a non-rigid registration algorithm is hindered by non-corresponding pathological tissues. In pre-operative data, the tumor is visible, whereas in post-surgical acquisitions the resection cavity is

formed. Image-based algorithms registering MRI acquisitions must have a strategy to deal with the lack of one-to-one correspondence between pre- and post-operative images.
Another challenge in the design of image registration algorithms is related to the fact that multiple MRI sequences are acquired to differentiate pathological tissues from healthy structures precisely. The most common ones are T1, T1-CE, T2, and FLAIR [55; 31; 4]. Registration algorithms are highly influenced by the choice of the MRI sequences used to compute the correspondences between the reference and the moving image. An evaluation of the influence of the different sequences on the image registration outcomes is advisable. The choice of the MRI sequence is also important for other image-processing tasks, such as image segmentation.

Chapter 3 presents an iterative solution and a convolutional neural network-based method for the registration of pre- and post-operative MRI data to support neurosurgical procedures for tumor resection. An analysis of the effects of the pathological tissues on both algorithms is conducted. Each method is designed by excluding or including the contribution of the pathological tissues in the computation of the distance measure: The voxels belonging to the pathology are specified by segmentation masks. Besides, an evaluation of the influence of the different MRI sequences on the outcomes of both registration methods is provided.

Chapter 4 further investigates the exclusion of the pathology from the regularization term in the registration of subsequent MRI volumes. An iterative method, which discards the contribution of the pathological tissues from the regularizer, is proposed to register pre- and post-operative MRI acquisitions. Segmentation masks of pathological tissues are utilized to specify the voxels, whose contribution is discarded from the computation of the regularization term.

Chapter 5 presents a deep learning-based method to segment the resection cavity in post-operative data and evaluates the influence of different MRI sequences on the algorithm's outcome. Similarly to registration methods, image segmentation algorithms applied on post-operative MRI volumes are also influenced by the different MRI sequences. Thus, the chapter evaluates which sequence leads to the most accurate results. Besides, the segmentation masks obtained by the proposed method are also utilized to support the registration methods in 3.

3. Quantitative Evaluation of the Influence of Multiple MRI Sequences and of Pathological Tissues on the Registration of Longitudinal Data Acquired During Brain Tumor Treatment

The material of this chapter is available as an open-access journal paper published in Frontiers in Neuroimaging.

Canalini, L., Klein J., Waldmannstetter D., Kofler F., Cerri S., Hering A., Heldmann S., Schlaeger S., Menze B.H., Wiestler B., Kirschke J. and Hahn H.K. (2022) Quantitative evaluation of the influence of multiple MRI sequences and of pathological tissues on the registration of longitudinal data acquired during brain tumor treatment. Front. Neuroimaging 1:977491. doi: 10.3389/fnimg.2022.977491

Abstract

Purpose Registration methods facilitate the comparison of multiparametric magnetic resonance images acquired at different stages of brain tumor treatments. Image-based registration solutions are influenced by the sequences chosen to compute the distance measure and the lack of image correspondences due to the resection cavities and pathological tissues. Nonetheless, an evaluation of the impact of these input parameters on the registration of longitudinal data is still missing.

Methods and Experiments This work evaluates the influence of multiple sequences (T1, T2, T1-CE, and FLAIR) and the exclusion of the pathological tissues on the non-rigid registration of pre- and post-operative images. We here investigate two types of registration methods, an iterative approach and a convolutional neural network solution based on a 3D U-Net. We employ two test sets to compute the mean target registration error (mTRE) based on corresponding landmarks.

Results In the first set, markers are positioned exclusively in the surroundings of the pathology. The methods employing T1-CE achieve the lowest mTREs, with an improvement of up to 0.8 mm for the iterative solution. The results are higher than the baseline when using the FLAIR sequence. When excluding the pathology, lower mTREs are observable for most of the methods. In the second test set, corresponding landmarks are located in the entire brain volumes. Both solutions employing T1-CE obtain the lowest mTREs, with a decrease of up to 1.16 mm for the iterative method, whereas the results worsen using the FLAIR. When excluding the pathology, an improvement is observable for the CNN method using T1-CE.

Conclusion Both approaches utilizing the T1-CE sequence obtain the best mTREs, whereas the FLAIR is the least informative to guide the registration process. Besides, the exclusion of pathology from the distance measure computation improves the registration of the brain tissues surrounding the tumor. Thus, this work provides the first numerical evaluation of the influence of these parameters on the registration of longitudinal magnetic resonance images, and it can be helpful for developing future algorithms.

3.1. Introduction

In neurosurgery for tumor resection, a pre-operative MRI acquisition is obtained to plan the surgical removal. After neurosurgery, MRI images are also acquired at follow-up stages to identify any pathology recurrence [6]. The identification of pathological tissues in post-operative acquisitions can be improved by comparing MRI images obtained at subsequent stages of neurosurgical treatments [54]. Registration algorithms are used to establish correspondences for a precise visual inspection between the subsequent MRI scans [56]. Mass effects, pathology resection, and tumor regrowth produce large deformations in the close-to-tumor regions [57]. To accommodate these changes, rigid registration algorithms are not accurate enough [58]. Instead, nonrigid registration solutions are a better option, because they generate deformations fields that can locally register brain areas.
Several methods to register pre- and post-operative MRI images are already available. The authors in [59] propose a solution to register corresponding healthy tissues of longitudinal images. Furthermore, the same authors [60] develop a method to register pre-operative MRI data with any stage of images acquired after tumor resection. By estimating missing correspondences, their algorithm encourages the accommodation of the tissues surrounding the tumor. Another solution is proposed by [55], in which the authors register T1 MRI images by excluding pathological tissues from the computation of the distance measure (see Eq. 3.2 and Eq. 3.1 for a better explanation). Furthermore, the authors in [57] propose a semi-automatic method to register pre-operative, post-operative, and follow-up images of individual patients. Their approach first semi-automatically segments brain contours, ventricles, enhanced tissues, and resection cavities in the pre- and post-operative images. In the second step, T1-CE volumes and the masks are used as input to a registration method. Besides, in [31] the authors propose a method to register pre-operative and post-recurrence brain tumor images. The acquisitions are registered by excluding the pathological tissues from the image-correspondence term. T1-CE and T1 MRI sequences are used to guide the registration process. One of the few algorithms based on deep learning to register longitudinal MRI data is proposed by [61]. 3D T1 images are registered by excluding the segmentation of pathological tissues. A 5-level 3D U-Net model is trained on the registration of inter-patient data. In the test phase,

they use volumes coming from 18 longitudinal studies, each having 2 follow-up acquisitions. Moreover, the work proposed by [62] is based on a convolutional neural network (CNN). The authors propose a joint segmentation-registration solution to i) automatically segment pathological tissues in the moving and fixed images and ii) register the pair of multiparametric images by excluding the automatically segmented structures from the computation of the distance measure. T1, T2, T1-CE, and FLAIR sequences are all used as input.

Multiple MRI sequences are acquired at subsequent stages of the neurosurgical treatment to better identify pathological tissues [55; 31; 4]. The recommended minimum requirements in neurosurgery include T1-CE, T1, T2, and FLAIR [47]. The standard for T1-CE and T1 images is usually to acquire high-resolution 3D isotropic volumes, whereas for T2-weighted 2D acquisitions are obtained [63; 47]. Image-based registration algorithms using high-resolution images are likely to obtain better results than those utilizing lower resolution images, such as FLAIR and T2w acquisitions. Nevertheless, the already proposed solutions utilize different MRI protocols to guide the registration process. A numerical evaluation of the influence of multiple sequences on the registration of longitudinal MRI data is still missing.

Furthermore, image-based registration algorithms rely on the fact that corresponding structures can be found in the pairs of images to be registered. This assumption is not valid for the registration of longitudinal MRI acquisitions acquired during tumor resection. The pathological tissues visible in pre-operative acquisitions are removed and are not observable in post-removal images. Many of the proposed registration methods tackle this problem by excluding the contribution of pathological tissues from the correspondences computation [60; 55; 31]. In fact, it is commonly assumed that the outcome of the registration process improves if only corresponding (healthy) structures are taken into account. However, no exhaustive evaluation of the influence of the exclusion of the pathological tissues has been done yet.

First, this work aims to evaluate the influence of different MRI sequences on the registration of longitudinal MRI data. To the best of our knowledge, it is the first time that this analysis has been performed. More details are given in section 3.2.3. Second, this work quantitatively analyzes the effects of excluding (and including) the pathological tissues from the computation of distance measure used for registration of longitudinal MRI data (more details in section 3.2.3). Two registration approaches are proposed for performing the aforementioned experiments, an iterative method,

and a CNN-based solution. The exclusion of the pathological tissues in the CNN method is performed only during training, whereas the iterative method excludes them during the registration process.

3.2. Material and Methods

3.2.1. Datasets

For the volume of each dataset, four different MRI sequences are available: native T1, T1-CE, T2, and FLAIR. Each case is normalized using the same preprocessing [63]: Every volume is skull-stripped, noise corrected, rigidly registered to an atlas reference volume, and interpolated to $1mm^3$ voxel resolution. The images have a size of 240 × 240 × 155 and are downsampled to 160 × 160 × 160 to be input to the registration methods.

Munich Dataset

This set includes two or more consecutive post-operative acquisitions of 66 patients, acquired at the Klinikum Rechts der Isar during 2015 and 2020 [64]. From this dataset, we choose a subset of 57 patients to only include volumes characterized by four MRI sequences. The original acquisitions for each patient include an isotropic T1 (voxel size of 1 mm^3) before and after contrast, axial T2 (voxel size of 0.72 × 0.72 mm^2), an isotropic FLAIR (voxel size 1 mm^3). The volumes are available after normalization performed according to [63]. The pathological tissues are automatically segmented [64]. To generate a robust brain tumor segmentation, we use an iterative process. First, we generate binary segmentation masks using five segmentation algorithms [65; 66; 67; 68; 69] developed within the scope of the BraTS challenge [70; 71; 72; 63] using BraTS Toolkit *btk*) [73]. Second, we fuse the segmentation masks using equally weighted majority via *btk* [73]. Third, a visual inspection is conducted to correct the fused segmentation masks. This approach promises to achieve a higher segmentation quality than a pure manual delineation [74] while saving valuable expert radiologists' time.

BraTS 2015: Validation Set

BraTS 2015 dataset includes a mixture of pre-operative and follow-up MRI images. A subset of the BraTS 2015 training dataset is chosen [75], to include only longi-

tudinal studies. This subset includes 45 pairs of images, each composed of a pre- and post-operative acquisition. According to [63], the original acquisitions for the images sets include a T1 image (1–6 mm slice thickness), a T1-CE image (voxel size of 1 mm^3 for most patients), a T2 image (with 2–6 mm slice thickness), a FLAIR image (2–6 mm slice thickness). The ground truth masks of the pathological tissues are also available. Moreover, the resection cavity, not originally segmented in the original ground truth, has been manually segmented in a previous work [76].

BraTS 2015: Test Set

A subset of the BraTS 2015 test set is selected, to include only longitudinal studies. It has 59 pairs of images of different patients. The acquisition details of this dataset are the same as described in the previous subsection (*BraTS 2015: Validation Set*) [63]. The masks of the pathological tissues (edema, enhancing tumor, necrosis non-enhancing tumor) have been already segmented in the original dataset. Moreover, the resection cavities in the post-operative volumes are manually segmented by two raters. An example of the finally available structures is shown in the fifth column of Fig. 3.1. To compute the registration results, six landmarks have been manually acquired for each pair of longitudinal acquisitions. First, for each pre-operative scan, landmarks are acquired near the tumor (within 40mm). Second, corresponding markers are obtained in post-operative images. The landmarks are acquired on anatomical structures, such as brain sulci and gyri, and the midlines of the brain. An example of the annotated landmarks is available in Fig. 3.2. One or two raters annotated them and an experienced neuroradiologist evaluated them clinically. The baseline mean target registration error (mTRE) is 2.92 mm.

BraTS-Reg Challenge Dataset

The dataset includes 140 pairs of pre and post-operative MRI volumes [4]. The time window between all pairs of pre-operative and follow-up volumes is in the range of 27 days-37 months. This dataset comprises already pre-processed image sets collected in affiliated and public institutions. Although no information about the acquisition parameters is provided by the challenge organizers, it is likely to assume that the data have been acquired following the standard protocols mentioned in [63] and already described in the previous section (*BraTS 2015: Validation Set*). Several raters manually annotated 6 to 50 corresponding landmarks between the pre- and post-operative volumes. For each pre-operative scan, landmarks are acquired near

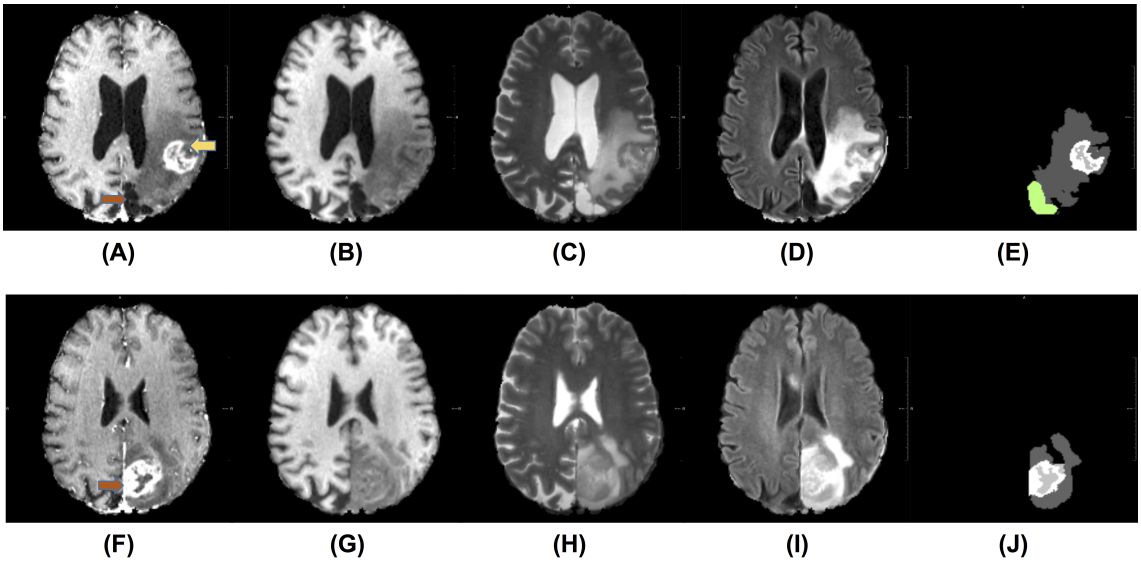


Figure 3.1.: **Available MRI sequences.** The figures in the first row show example slices of a post-operative acquisition, whereas in the second row images of the corresponding pre-operative volume are displayed. Each volume comprehends four sequences: T1 gadolinium contrast-enhanced T1-CE in (A) and (F), T1 in (B) and (G), T2 (C) and (H), and FLAIR in (D) and (I). Each sequence is useful to spot a particular component of the pathological tissues. For example, in T1-CE the enhanced tissue is observable, whereas, on the FLAIR sequence, the edema is well visible. The masks of the pathological tissues available for this work are visible in subfigures (E) and (J). In (E), the resection cavity is colored green. The pre-operative tumor and the corresponding resection cavity are indicated by the orange arrows. In after-surgery acquisitions, pathological tissues can also be present, as in this example (the post-operative tumor is pointed by the yellow arrow).

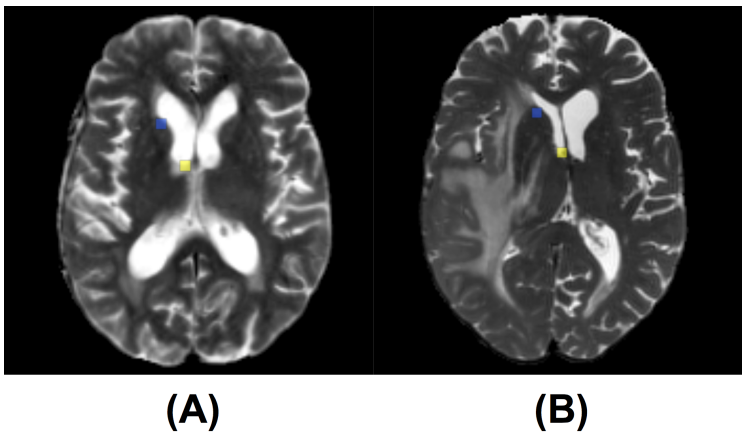


Figure 3.2.: **Example of annotated landmarks on Test Set.** Subfigures A and B correspond to post and pre-operative MRI acquisitions of the same patient. The corresponding landmarks are visualized with the same colors.

the tumor (within 30mm) and far from the tumor (beyond 30mm). Thus, matching points are obtained in post-operative images. The landmarks are anatomical structures, such as blood vessel bifurcations, the anatomical shape of the cortex, and anatomical landmarks of the midline of the brain [4]. An example of the annotated landmarks in BraTS-Reg is available in Fig. 3.3. After a rigid pre-registration between the pre- and post-operative volumes, the baseline mTRE is 3.62 mm. The automatic algorithm already applied for 3.2.1 is here utilized to segment the pathological structures in every volume of this dataset.

3.2.2. Methods

This work investigates two different types of registration methods, an iterative solution, and a CNN-based approach. The following design concepts are valid for both approaches.

The reference (post-operative) and template (pre-operative) images can be modeled as functions $\mathcal{R}$, $\mathcal{T} : \mathbb{R}^3 \to \mathbb{R}$. The goal of the proposed image registration approaches is to generate a deformation $\mathrm{y} : \Omega \to \mathbb{R}^3$ that aligns the two images $\mathcal{R}$ and $\mathcal{T}$ on the field of view $\Omega \subset \mathbb{R}^3$ such that $\mathcal{R}(x)$ and $\mathcal{T}(\mathrm{y}(x))$ are similar for $\mathrm{x} \in \Omega$. The deformation field represents a minimizer of the cost function:

$$\mathcal{J}(\mathcal{R}, \mathcal{T}, \mathrm{y}) = \mathcal{D}(\mathcal{R}, \mathcal{T}(\mathrm{y})) + \alpha R(\mathrm{y}) \tag{3.1}$$

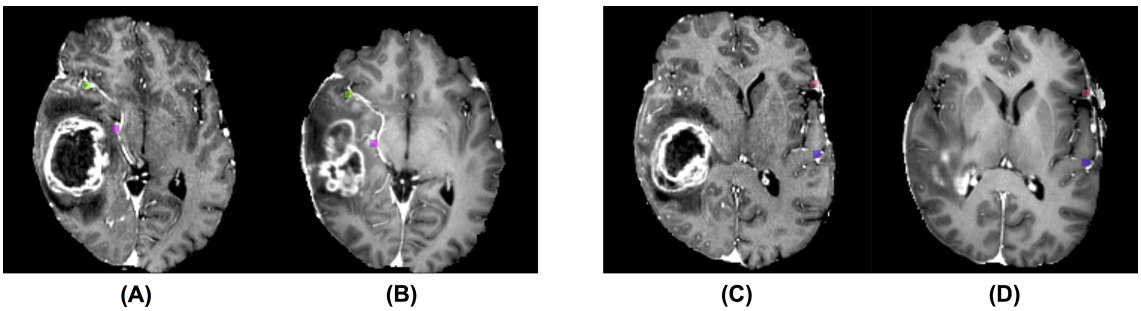


Figure 3.3.: **Example of annotated landmarks on BraTS-Reg dataset.** Subfigures A and C show different axial images of the post-operative acquisition for the same patient, whereas B and D correspond to the corresponding locations in the pre-operative image. The corresponding landmarks are visualized with the same colors.

The first term $\mathcal{D}$ is a distance measure computing the difference between the reference $\mathcal{R}$ and the warped template image $\mathcal{T}(\mathrm{y})$ summing pointwise distances. A challenge of registering longitudinal MRI data is that a one-to-one correspondence between the two images is not guaranteed due to the resection of pathological tissues. We tackle the problem by adding the possibility of excluding the contribution of the pathology of the fixed image from the computation of the distance measure. By indicating the pathological tissues as Σ, the distance measure is computed as follows:

$$D(\mathcal{R}, \mathcal{T}(\mathrm{y})) = \int_{\Omega \setminus \Sigma} d(\mathcal{R}(x), \mathcal{T}(\mathrm{y}(x)) \, \mathrm{d}x \tag{3.2}$$

In our settings, $\mathcal{R}$ and $\mathcal{T}$ respectively refer to the post-operative and pre-operative volumes. Thus, Σ corresponds to the pathological masks of the post-operative image. More details about the distance measure are available in Eq. 4.2.

The second term R in Eq. 3.1 is the regularizer, which limits the possible deformations that can be computed during the minimization process. The hyperparameter α controls the strength of the minimization term. More details about the regularization chosen in this work are available in Eq. 3.4 and Eq. 3.5.

Multi-level deep learning method

The deep learning registration method used in this work is based on the solution proposed by [43]. It is a multi-level variational image registration approach, that combines deformation fields computed at different image scales. Since the solution

already achieved good results in other medical imaging fields, we want to evaluate how this method performs on longitudinal MRI data.

The solution has been originally proposed to register lung data, where also corresponding masks and landmarks were available to compute additional similarity terms. In our case, only intensity brain volumes are available. Therefore, we use only the intensity images to compute the similarity. Due to the lack of consistent intensity profiles in the MRI acquisitions [4], and the presence of pathological tissues, the normalized gradient fields (NGF) measure is chosen as distance metric [77]. The use of the NGF is based on the observation that two images are considered similar if intensity changes occur at the same locations. Instead of computing the magnitude of the image gradient ($\nabla\mathcal{R}(x)$ and $\nabla\mathcal{T}(x)$ respectively for the reference and template image), the normalized gradient field is utilized [77]. The goal of image registration based on NGF is to align them by reducing the difference between the normalized gradient fields computed for the reference and the template image. It is defined as follows,

$$\mathrm{NGF}(\mathcal{R},\mathcal{T}) = \frac{1}{2}\int_{\Omega} 1 - \left(\frac{\langle\nabla\mathcal{R}(x),\nabla\mathcal{T}(x)\rangle_{\varepsilon_R\varepsilon_T}}{\|\nabla\mathcal{T}(x)\|_{\varepsilon_T}\,\|\nabla\mathcal{R}(x)\|_{\varepsilon_R}}\right)^2 dx \tag{3.3}$$

where $\langle x,\mathrm{y}\rangle_{\varepsilon} := x^{\top}\mathrm{y}+\varepsilon$, $\|x\|_{\varepsilon} := \sqrt{\langle x,x\rangle_{\varepsilon^2}}$ and $\varepsilon_R,\varepsilon_T > 0$ are the so-called edge-parameters controlling influence of noise in the images. Their value has been empirically chosen. Moreover, we modified the original CNN implementation by introducing the possibility of using masks of the pathological tissues as external input during the training and validation phases, to exclude their contribution from the correspondence computation (see Eq. 3.2). In the test phase, no mask of the pathological tissues is needed. Moreover, we also added the possibility of using input images characterized by two MRI sequences, whereas the original implementation accepted only one-channel images.

When searching for the best solution to the minimization term, multiple solutions may exist. However, not all the possible minima of the objective function represent good and realistic registration solutions. The R term in the objective function (see Eq. 3.1) favors a smoother deformation field y. As in the original architecture, we utilize the curvature regularizer,

$$R(\mathrm{y}) = \int_{\Omega}\sum_{k=1}^{3}\|\Delta\mathrm{y_k}(x)\|^2 dx \tag{3.4}$$

which penalizes deformation fields having too large second derivatives. To limit even further the viable solutions, another regularization term

$$\mathcal{V}(\mathrm{y}) = \int_{\Omega} \psi(\det \nabla \mathrm{y}(x))dx \tag{3.5}$$

is added to the objective function, where $\psi(t) = (t-1)^2/t$ for $t > 0$ and $\psi(t) := \infty$ for $t \leq 0$. The volume change control is used to discourage foldings in the deformation field y that may be generated during the minimization of the cost function. Folding in the deformation field represents an unrealistic transformation that the minimization process may lead to. The hyperparameter controlling the influence of this extra term on the loss function is γ, thus the final term to be added is $\gamma\mathcal{V}(\mathrm{y})$.

The proposed solution is based on a 3D U-Net architecture that takes as input the concatenation of the 3D fixed (follow-up) and the moving (pre-operative) image and provides as output a 3D dense deformation field with a resolution identical to the images [78]. The following description of the CNN architecture is based on what was reported in [78]. The network consists of three levels starting with 16 filters in the first layer, doubled after each downsampling step. 3D convolutions are used in both the encoder and decoder paths with a kernel size of 3 followed by an instance normalization and a ReLU layer. In the encoder path, the feature map downsampling steps use average pooling with a stride of 2. In the decoder path, the upsampling steps use transposed convolution with $2 \times 2 \times 2$ filters and half the number of filters than the previous step. The final layer uses a $1 \times 1 \times 1$ convolution filter to map each 16-component feature vector to a three-dimensional displacement.

The datasets *BraTS 2015: Validation Set* and *Munich Dataset* are used as the training sets. Each CNN model is trained for 40 epochs. The loss weighting parameters are set as follows: $\alpha = 0.1$, $\gamma = 0.01$. The learning rate is set to 0.001. The values of these parameters are different from the original implementation and were empirically modified. Besides using individual sequences, input characterized by two features (i.e., two MRI sequences) are also utilized to train the CNN method. The combinations of MRI sequences used to train and test the CNN solution are available in 3.2. The use of data input characterized by two features is a novelty with respect to the original architecture, where only one-channel images have been used [78].

The registration is performed on three levels ($L = 3$) by using images at different scales. The deformation field is initially computed on the coarsest level and the

images are downsampled by a factor equal to 2^{L-1}. On a finer level, the previously computed deformation fields are utilized as an initial guess by warping the moving image. At each level, the moving and fixed images are downsampled. Due to graphical memory issues, the finest resolution of the registered images is $160 \times 160 \times 160$, which is also the size of the generated deformation field [78]. The final deformation field is then upsampled to the original size of the input images.

Multi-level iterative method

The iterative solution utilized in this work is a variational image registration approach [32]. This method has been already used in the neurosurgical context [79] and, in this work, we evaluate it on longitudinal MRI data. The registration can be considered as an iterative optimization algorithm where the search for the correct registration between two images corresponds to an optimization process aimed at finding a global minimum of an objective function. The objective function has to be minimized for each image pair and the minimization process is performed according to a discretize-then-optimize paradigm. The objective function to be minimized includes a distance measure, quantifying the similarity between the warped template image and the reference one, and a regularizer, which favors the smoothness of the computed deformation fields. NGF is here used as a distance measure (see Equation 3.3), and a curvature regularizer is utilized (see Equation 3.4). The method also allows to mask the pathological tissues out from distance measure computation (see Equation 3.2). However, differently from the CNN method, these segmentations are needed in the test phase. Moreover, in the iterative method, the choice of the optimal transformation parameters is conducted by using the quasi-Newton l-BGFS [80], due to its speed and memory efficiency.

The iterative method performs a non-parametric registration that, as for the deep learning method, is performed on three levels ($L = 3$) by always using images at different scales. On the finest level, the volumes have a size of $160 \times 160 \times 160$. The deformation field obtained in output from the iterative method is then upsampled to the size of the original images. The stopping criteria for the optimization process are empirically defined: the minimal progress, the minimal gradient, and the relative one, the minimum step length are set to 0.001, and the maximum number of iterations is set to 100. The loss weighting parameter is empirically set to $\alpha = 0.1$. The registration algorithm is used to register the volumes of the test set.

Diffeomorphic registration ANTs method

The Symmetric Diffeomorphic registration method ANTs represents a standard registration algorithm for MRI data [81; 82]. Thus, we aim to evaluate how well the proposed methods perform in comparison to a standard approach. ANTs is applied by using the original size of the volumes ($240 \times 240 \times 155$). As suggested in [83], the Symmetric Normalization (SyN) transformation model of ANTs is utilized, and cross-correlation is used as the distance measure since NGF is not available. The T1-CE is used to guide the registration, since it is supposed to be the sequence with the higher original resolution and, thus, the one leading to better registration results. Moreover, the masking of the distance measure is performed, to reduce the negative effects of the non-corresponding tissues on the registration results.

3.2.3. Experiments

Influence of different MRI sequences

This work aims to numerically analyze the influence of multiple MRI sequences on the registration of longitudinal data by evaluating the performances of two types of methods, the iterative solution, and the deep learning-based approach. The CNN method is trained on four individual sequences (T1-CE, T1, T2, and FLAIR). In the inference process, the models are then used to register the corresponding sequences in the test sets. Besides, the iterative method is applied to different individual sequences. Moreover, we also train the CNN solution using input volumes characterized by two distinct sequences. This experiment aims to verify whether multiple MRI sequence input is better than only one to train the neural network models. The deformation field computed on an individual or multiple sequences can then be applied to the other MRI acquisitions of the same patient.

The nonparametric Wilcoxon signed-rank test is utilized to verify whether there is a statistically significant difference between the results of each registration solution (iterative method and trained CNN models) and the baselines of the two test sets. This analysis tests whether the median of the differences between the two paired results is zero. The data distribution of the baseline registration errors is not normal according to the One-sample Kolmogorov-Smirnov test [84].

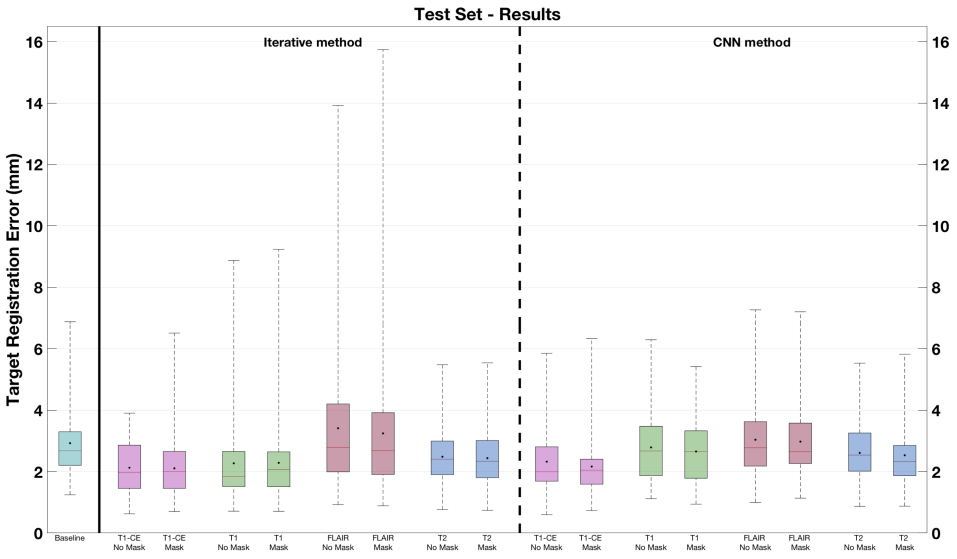


Figure 3.4.: **Target registration errors on the test set.** The box plots related to the solutions where no masking procedure is performed are indicated as *No Mask*. In each box plot, the red line and the point respectively indicate the median and mean values.

Effects of excluding the pathological tissues from the distance measure

This work also evaluates the influence of pathology on the registration process. Thus, masks are used to exclude the contribution of the pathological tissues from the distance measure computation. For the CNN method, the segmentation of the pathological tissues is used as extra input only during training, to compute the distance measure on the healthy tissues (see Eq. 3.2). In the inference process, no mask is required. Four additional models are trained, each for a different MRI sequence, without masking the distance measure. The iterative method also has the possibility of excluding the pathological tissues from the distance measure. However, it requires the segmentation of pathological tissues when applied to the test set. The iterative method excluding and including the pathological tissues is also utilized for each of the four MRI sequences.

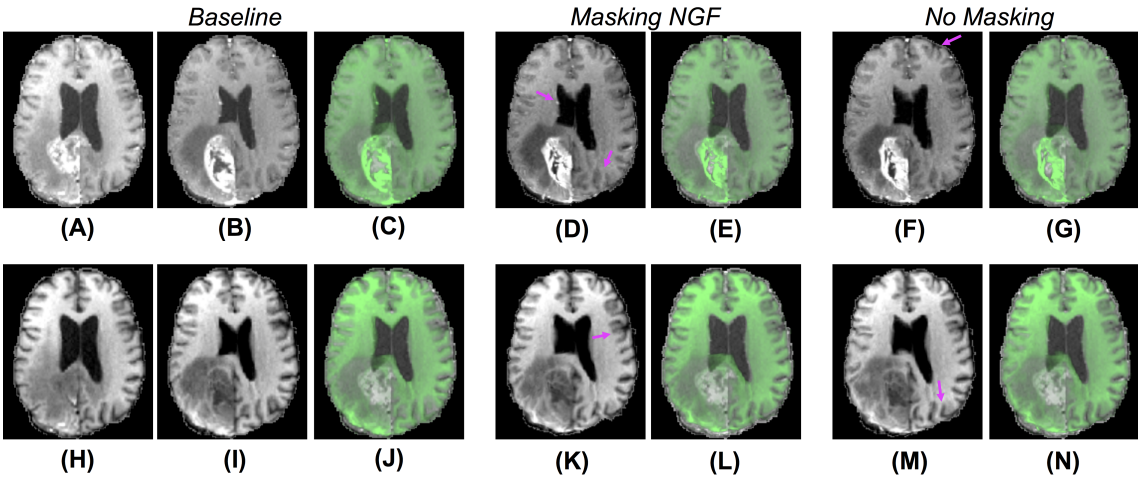


Figure 3.5.: **Comparison of qualitative results for CNN models trained on T1-CE and T1 MRI sequences.** Each row refers to the results for the CNN models trained on different sequences (from the top to the bottom, T1-CE, and T1). The post- and pre-operative images are visualized in the first and second columns, and the initial overlay between the two acquisitions is visible in the third column. The last four columns display the warped moving volumes and the overlays between the fixed image (post-operative) and the warped moving images, respectively for the models excluding and including the pathology in the distance measure. The purple arrows point to locations where improvements are observable. In Subfigs. D, and K a better overlap of the lateral ventricles is visible. Moreover, in Subfigs. F, K, and D the sulci are better registered.

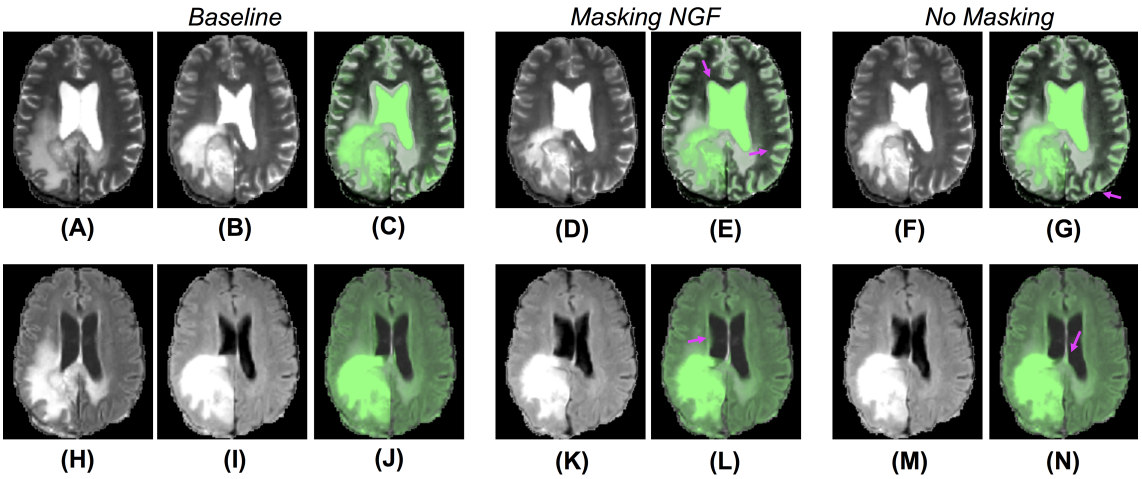


Figure 3.6.: **Comparison of qualitative results for CNN models trained on T2 and FLAIR MRI sequences.** Each row refers to the results for the CNN models trained on different sequences (from the top to the bottom, T2 and FLAIR). The post- and pre-operative images are visualized in the first and second columns and the initial overlay between the two cases is in the third column. The last four columns display the warped moving volumes and the overlays between the fixed image (post-operative) and the warped moving images, respectively for the models excluding and including the pathological tissues in the distance measure. The purple arrows point to locations where improvements are observable. In Subfigs. E, L, and N, a better registration of the lateral ventricles is visible. Besides, in Subfigs. G and E the sulci are more aligned.

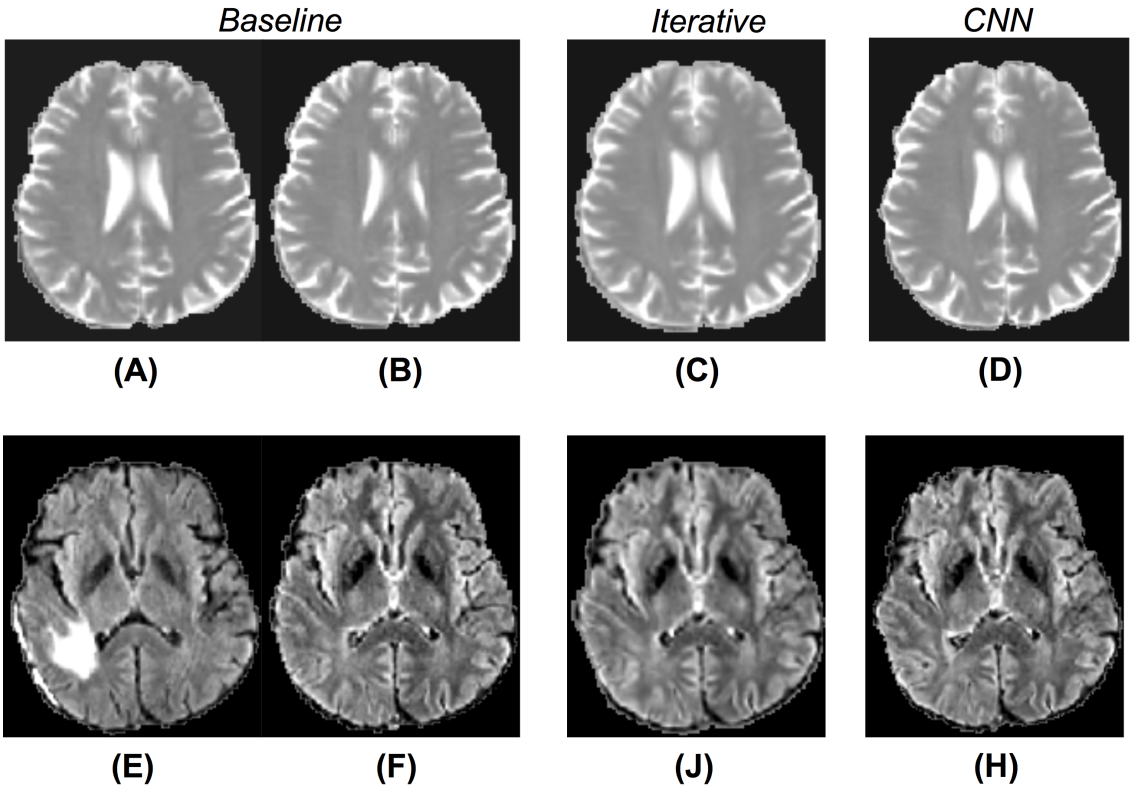


Figure 3.7.: **Qualitative results obtained by CNN and iterative solutions masking the distance measure.** The first row is related to methods using the T2 MRI sequence, whereas the last one shows example results for solutions using the FLAIR sequence. The first two columns show the corresponding slices of the pre- and post-operative volumes, rigidly registered in the pre-processing step. The third column presents warped moving images obtained by the iterative method, the fourth column shows the results for the CNN models.

3.3. Results

The mean target registration errors are computed to evaluate the outcome of the different registration algorithms. For each matching landmark (see Fig. 3.2 for an example), the Euclidean distance between its position in the reference image and its position in the moving image is computed. Then, for each patient set, the mean Euclidean distance among all the landmarks is calculated. Thus, the mean value of the distances of all the image sets of a test dataset (i.e., the mTRE) is estimated (please refer to Appendix III for more details). The proposed methods' output is the warped moving image and the deformation field. The latter is applied to the landmarks and, if the value obtained after registration is lower than the initial baseline, the warped moving images are supposed to be better registered to the corresponding reference images. The results shown in this paper are computed on two different test sets, *BraTS 2015: Test Set* and *BraTS-Reg Challenge Dataset*.

Table 3.1 shows the mTREs obtained by the proposed solutions. For the CNN models trained without masking, the lowest mTRE (2.32 mm) is obtained by the model using the T1-CE, whereas the highest value is achieved when the FLAIR sequence is used (3.04 mm). Besides, the iterative method using this sequence also achieves the highest mTRE (3.41 mm). The lowest mTRE is obtained with the T1-CE (2.13mm). A further comparison of the results obtained on different MRI sequences by the two methods is visible in Fig. 3.4. In both cases, the FLAIR sequence leads to higher median TREs, with a large range of results. On the contrary, the T1-CE and T2 sequences help to lower the median TREs and limit the ranges of values. Besides, the Wilcoxon test is utilized considering the target registration errors. In both methods, the null hypothesis cannot be accepted for the models trained on T1-CE ($p < 0.000001$), and on T2 ($p < 0.01$). Fig. 3.5 shows the qualitative results of CNN models trained on T1-CE and T1 MRI sequences, and Fig. 3.6 those of the methods using the T2 and FLAIR MRI sequences. The qualitative results for the iterative method using T2 and FLAIR MRI are visible in the third column of Fig. 3.7. In this figure, the visual results of the iterative method using T2 or FLAIR are provided.

Moreover, for what concerns the results obtained by discarding the contribution of the pathology, Table 3.1 shows that the lowest and highest mTREs obtained by the CNN without masking the pathological are reduced when these tissues are excluded from the distance measure computation (respectively, 2.16 mm and 2.98 mm). When

the tumor is excluded, an mTRE of 3.24 mm is achieved by the iterative method using FLAIR. The lowest mTRE in our experiments is obtained when registering the T1-CE MRI sequence (2.11 mm). A comparison between each method excluding or not the pathological tissues from the distance measure is also available in Fig. 3.4. Besides, the fourth column in Fig. 3.7 shows the registration results for two CNN models trained on longitudinal data by excluding the pathological tissues from the distance measure. Furthermore, in Table 3.2, the lowest mTRE (2.41 mm) is obtained by CNN trained with T1-CE and T2 sequences. On the contrary, the highest mTRE of 2.82 mm is achieved by the solution using FLAIR and T2 sequences.
The CNN models trained on individual sequences, as well as the iterative methods, are also applied to the BraTS-Reg dataset. The mTRE results are available in Table 3.3 and Fig. 3.8. The results of both methods are lower when using the T1-CE and the T2 sequences, whereas the highest mTREs are achieved on the FLAIR sequence. Fig. 3.8 shows that, when using the FLAIR sequence, both methods lead to a range of values even higher than the baseline. Moreover, when the iterative method uses the T1 sequence, some cases also have larger TRE than before registration. On the contrary, when using T2 and T1-CE sequences, smaller ranges of values are achieved. When comparing the CNN models masking the pathological tissues and those not excluding them in Table 3.3, we can observe a lower value only for the networks trained on T1-CE. However, higher mTREs are obtained by the CNN methods trained on FLAIR and T2. Besides, almost no difference can be seen between the sections related to the traditional method (Iterative vs Iterative masks). A more detailed overview is observable in Fig. 3.8, comparing for each sequence the box plots labeled as *No Mask* and *Mask*. Besides, the second section of Table 3.2 provides the results obtained on the BraTS-reg dataset by the CNN models trained on multiparametric input. The numerical result obtained by using T1-CE and T2 achieves the lowest mTRE obtained by the CNN method on the BraTS-Reg dataset.

According to our experiments, the T1-CE sequence and the masking of the pathological tissues from the distance measure lead to the lowest mTREs. The ANTs algorithm also uses these settings (see 3.2.2) and the mTRE obtained by this method on the BraTS-Reg dataset is 2.84 mm, whereas on the Test set the final value is equal to 2.37 mm.

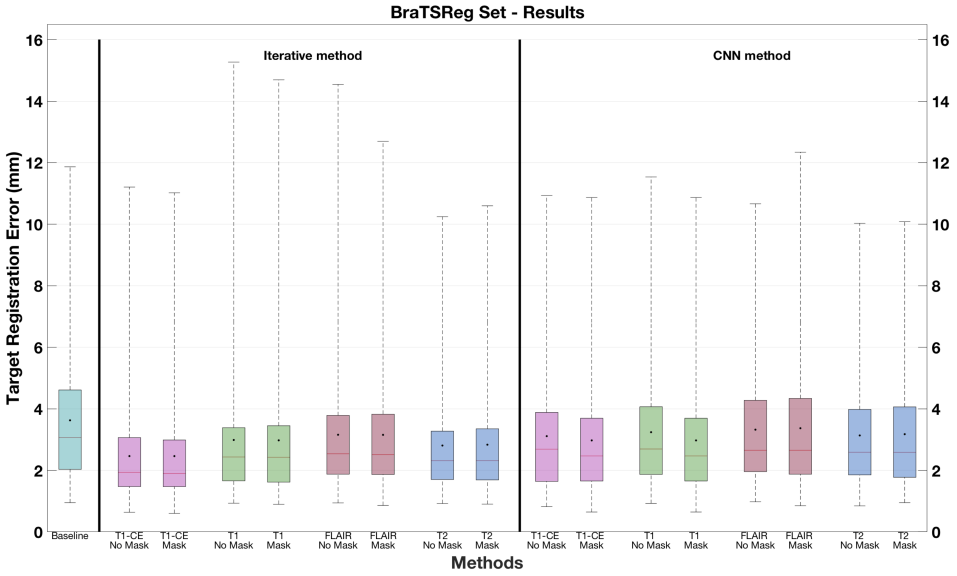


Figure 3.8.: **Target registration errors on the BraTS-Reg dataset.** The box plots related to the solutions where no masking procedure is performed are indicated as *No Mask*. In each box plot, the red line and the point respectively indicate the median and mean values.

3.4. Discussion

None of the already proposed methods analyzed the influence of different sequences on the registration of longitudinal MRI data. However, our experiments show that the choice of the MRI sequence has a strong impact on the outcome of the registration of longitudinal data. This is evident by analyzing the results obtained by two different types of registration methods, namely an iterative and CNN method, proposed for the task of non-rigidly registering longitudinal MRI data. Our experiments show that the T1-CE sequence is the best choice for designing the registration algorithms, leading to better mTREs. This outcome could be explained by the fact that T1-CE images used in neurosurgery usually have a higher resolution than other MRI protocols. Moreover, thanks to the contrast enhancement, the better image contrast of anatomical tissues in this protocol could also be responsible for better registration results. On the contrary, the FLAIR is the worst to guide the registration process: This couldn't be predicted from the original acquisition parameters of the test sets, since the T1-weighted, T2, and FLAIR images are acquired with comparable resolution. These findings are also visible in Figs. 3.4 and 3.8, where the boxplots related to the FLAIR sequence present higher median and mean TREs than the other sequences, and a range of values higher than the baseline. Besides, all the multi-sequences trained CNN models to improve the baseline mTRE of the test set, but none leads to an improvement in terms of registration accuracy. The deformation fields computed on T1-CE sequence can then be employed to warp the other acquisitions characterized by the other MRI protocols. Besides, the model trained on T1-CE and T2 sequences outperforms that method trained solely on T1-CE in the BraTS-Reg dataset (refer to Table 3.2 and Table 3.3).

Our experiments also show that computing the distance measure on non-corresponding elements negatively impacts the registration of the longitudinal MRI data. Yet, the influence of the masking procedure differently affects the brain tissues, depending on their positions relative to the pathological tissues. The exclusion of the pathology from the computation of the distance measure has a positive effect on the mTREs of Table 3.1. However, the exclusion of these tissues has almost no influence on the registration of the BraTS-Reg dataset, as shown in 3.3, except for the CNN trained on the T1-CE. where landmarks are also positioned far away from the pathology.

Furthermore, the iterative solution using downsampled volumes outperforms the CNN approach, and the standard method ANTs, which utilizes original-size images

in both test sets. In fact, the best improvement is obtained by the iterative method using T1-CE on the BraTS-Reg dataset, where the initial mTRE is reduced by 1.16 mm. In the test set, the initial value is reduced by 0.81 mm by this method. Instead, the CNN approach is outperformed by the standard method in the BraTS-Reg set. Besides, in the test set, the CNN method achieves better results than the standard solution. The CNN method trained on the T1-CE sequence obtains an improvement of 0.65 mm on the BraTS-Reg set and of 0.76 mm on the test set.

3.4.1. Limitations

The iterative method is not affected by memory issues as the CNN solution. Thus, original resolution images could be utilized to validate the iterative method. Nevertheless, this method has been evaluated by using the resampled volumes as input, which could be suboptimal for the accuracy outcome. Although it is not uncommon to use lower resolution images to speed up the registration process, an improvement in the registration results might be achieved by using original size data. Nevertheless, the iterative solution using downsampled volumes already achieves the best results in our experiments.

Due to memory issues, the input data to the CNN had to be downsampled. By reducing the input size, the information stored in the original images gets lost. Less information can also be responsible for the poorer performance of the CNN solution, which is based on a learning process. To overcome the memory issues related to the 3D CNN solution, a 2.5 dimension approach could be used [85]. It has already been demonstrated to provide good registration results for 3D data and could help to improve the registration results by using larger input images. In this work, up to two sequences could be used to train the CNN solution, due to memory limitations. By using more powerful hardware, larger combinations of MRI protocols could include up to four sequences. It would be interesting to investigate whether the performance of the CNN method could improve. Moreover, a larger and more heterogeneous dataset could help to improve the performance of the deep learning method.

Another limitation is related to the landmarks provided in the BraTS-Reg dataset. Although this set provides more image pairs and landmarks than the Test Set, no information about the distance of the landmarks from the tumors is shared with the public. If the landmarks would be divided into two groups according to their distance from the pathology, it would be interesting to validate how the proposed

Table 3.1.: **Mean Target Registration Errors for the test set.** The mean TREs obtained by CNN models trained with and without masking procedure are visible in the second and third columns. The results achieved by the iterative method are in the fourth and fifth columns.

	mTRE (mm)			
Baseline	2.92			
ANTs (T1-CE, mask)	2.37			
MRI sequence	CNN	CNN Mask	Iterative	Iterative Mask
T1	2.79	2.65	2.27	2.29
T1-CE	2.32	**2.16**	2.13	**2.11**
FLAIR	3.04	2.98	3.41	3.24
T2	2.61	2.53	2.48	2.44

methods perform in the different brain areas for this dataset.

3.4.2. Conclusions

To the best of our knowledge, our work provides the first quantitative analysis of the influence of different MRI sequences on the registration of longitudinal MRI data. We also evaluate how much impact the exclusion of the pathological tissues has on the registration of pre- and post-operative data. To conduct our experiments, a multi-level deep learning solution and an iterative method are proposed for the registration pre- and post-operative MRI data acquired in the neurosurgical context. A few changes have been made to the original CNN implementation i) to accept multiparametric images and ii) to mask specific tissues out of the distance measure. Our experiment showed that the best sequence to guide the registration process is the T1-CE. For the CNN solution, the combination of T1-CE and T2 sequences also leads to good results. The best-performing solution in our experiments is provided by the iterative method, using the T1-CE sequence.

Acknowledgements

This work was funded by the H2020 Marie-Curie ITN TRABIT (765148) project.

Table 3.2.: **Mean Target Registration Errors of the multisequence trained models on the test set (first table) and BraTS-Reg dataset (second table).**

mTRE (mm) - Test set	
Sequence	CNN Mask
T1-CE + T2	**2.41**
T1-CE + T1	2.45
T1 + T2	2.48
T1-CE + FLAIR	2.43
T1 + FLAIR	2.61
FLAIR + T2	2.82

mTRE (mm) - BraTS-Reg set	
Sequence	CNN Mask
T1-CE + T2	**2.94**
T1-CE + T1	3.17
T1 + T2	3.58
T1-CE + FLAIR	3.15
T1 + FLAIR	3.23
FLAIR + T2	3.03

Table 3.3.: **Mean Target Registration Errors for the BraTS-Reg Dataset.** The mean TREs obtained by CNN models trained with and without masking procedure are visible in the second and third columns. The results achieved by the iterative method are in the fourth and fifth columns.

	mTRE (mm)			
Baseline	3.62			
ANTs (T1-CE, mask)	2.84			
MRI sequence	CNN	CNN Mask	Iterative	Iterative Mask
T1	3.24	3.24	2.99	2.98
T1-CE	3.11	**2.97**	**2.46**	**2.46**
FLAIR	3.32	3.37	3.16	3.15
T2	3.13	3.17	2.81	2.83

4. Deformable Registration of Pre- and Post-operative MRI Volumes by Utilizing a Voxel-wise Weighting of the Regularizer

Abstract

Purpose Deformable registration of pre- and post-operative MRI brain acquisitions is hindered by non-corresponding pathological tissues. Besides excluding their contribution from the distance measure, the constraints imposed by the regularizer on the deformations of the pathological tissues could be eliminated to improve the registration of corresponding healthy structures. This work evaluates how the exclusion of pathological tissues from the computation of the regularizer affects the registration results.

Methods and Experiments An iterative deformable method is proposed to register pre- and post-operative contrast-enhanced T1-weighted (T1-CE) volumes. Masks of pathological tissues are used to identify the voxels of non-corresponding pathological structures, whose contribution is discarded from the calculation of the regularizer. This work compares this method with the same solution not excluding pathology from the registration process.

Results A public dataset of 140 pairs of MRI volumes is utilized. Corresponding landmarks acquired in pre-and post-operative acquisitions are available to evaluate registration algorithms. Masks of pathological tissues are automatically segmented by a freely available method. The deformable solution excluding the pathological tissues from the objective function reduces the initial mean target registration error (mTRE) from 4.25 mm to 2.38 mm. The mTRE obtained by the approach not excluding the pathology is 2.43 mm.

Conclusion The exclusion of the pathological tissues from the regularizer computation is beneficial to the registration results. Future algorithms could start from these results for investigating additional ways of weighting the regularizer.

4.1. Introduction

Non-corresponding pathological tissues in pre- and post-operative MRI volumes negatively affect the image-based registration of subsequent acquisitions, due to the lack of one-to-one correspondence between the reference and template images [4]. In subsequent MRI acquisitions, tumors visible in pre-operative scans are removed; in post-operative volumes, they are no longer observable and resection cavities get formed. Image-based algorithms must have a strategy to reduce the negative effects of pathological tissues on the registration of subsequent MRI images acquired in brain tumor treatments.

A solution for dealing with the lack of correspondence between pre- and post-operative MRI acquisitions is to exclude the contribution of the pathological tissues from registration algorithms [31]. A commonly utilized approach is to compute the distance measure by excluding the contribution of the non-corresponding tissues [86; 60; 55; 31]. For this purpose, pathological tissues can be manually or automatically segmented and the generated masks are used to identify the voxels belonging to the non-corresponding structures. The contribution of these voxels is excluded from the distance measure computation. An analysis of the benefits of masking the pathological tissues out of the distance measure is available in the previous chapter (please refer to 3).

In deformable registration algorithms, the objective function to be minimized is not only composed of the distance measure but of a regularization term too, which penalizes too large or unrealistic deformations. When discarding the contribution of the pathology from the distance measure, the goal of a registration algorithm becomes to correctly register the healthy tissues. To further improve the registration of these structures, the contribution of non-corresponding tissues to the regularizer could be adapted to the surrounding tissues, or discarded by using a weighting parameter for the regularizer. By doing so, a more precise and accurate registration of the healthy corresponding structures could be achieved. The idea of locally adapting the amount of deformation to different types of tissues is not new in literature [87; 88; 89; 90; 43; 91; 92]. The authors in [91] propose an inter-patient deformable registration method for T1 MRI volumes, which computes a dense deformation field to register brain tissues. They introduce a voxel-based weighting term to locally

control the values of the regularization term, allowing some tissues to contribute more to the regularizer computation. The motivation behind the weighting parameter is to generate more realistic transformations. In their approach, some tissues are considered more as a stiff material, and others can be more freely deformed. The weighting term has large values where little deformation is expected, and small values in areas that should be more deformed. The degree of smoothness of the deformation fields is achieved by using the segmentation of brain tissues. Three classes are defined: cerebrospinal fluid, grey matter, and white matter. White and gray matter are deformed as stiff material, therefore the value of the weighting parameters is set equal to one. For the CSF, the value is 0, thus it does not contribute to the computation of the regularization. The authors in [92] also propose a grid-powered solution for the registration of MRI volumes with pathology to a healthy atlas. Similarly to the previous approach, they propose a solution in which the computation of the regularization term can be adapted according to the brain tissues. After segmenting healthy tissues and brain tumors, they accordingly modulate the value of the weighting parameter. To deal with the non-corresponding tissues of the pathology, their contribution is excluded from the computation of the regularizer: The value of the weighting parameter in these areas assumes the same value as the surrounding tissues. They utilize a dataset of 22 T1-weighting MRI volumes of different patients and qualitatively demonstrate how their pathology-aware solution improves the registration outcome. The authors in [89] propose a variational image registration approach with spatially varying regularization. They introduce a spatially varying parameter applied to an elastic regularizer, which can assume different positive values. By using segmentation masks of the objects to register, the factor multiplying the regularizer assumes different values. The weighting parameter controls how much the regularization values computed in correspondence with the voxels of different areas contribute to the final calculation. They use this approach to produce more realistic deformations.

Whereas the previous chapter 3 investigates the exclusion of pathology solely from the computation of the distance measure, this chapter proposes a method discarding the pathological tissues also from the computation of the regularization term. The goal of this work is to analyze whether the registration results are positively affected by the exclusion of the pathological tissues from the regularizer.

4.2. Method

The fixed and moving images can be modeled as functions $\mathcal{F}$, $\mathcal{M} : \mathbb{R}^3 \to \mathbb{R}$. The goal of the proposed image registration approaches is to generate a deformation $y(x) : \Omega \to \mathbb{R}^3$ that aligns the two images $\mathcal{F}$ and $\mathcal{M}$ on the field of view $\Omega \subset \mathbb{R}^3$ such that $\mathcal{F}(x)$ and $\mathcal{M}(y(x))$ are similar for $x \in \Omega$. The proposed method is a variational image registration approach based on [32], in which the registration of two volumes corresponds to the minimization of a discretized objective function $\mathcal{J}(\mathcal{F}, \mathcal{M}, y(x))$. This work focuses on a non-rigid registration solution for intra-patient MRI volumes. Its goal is to generate a dense deformation field y that locally deforms each voxel of the template image. The search for the correct solution corresponds to the minimization of an objective function, which aims at finding a deformation field that can reduce the distance measure computed between the reference and the warped template image. Expressing as $\mathcal{D}$ the distance measure, the objective function is defined as

$$\mathcal{J}(\mathcal{F}, \mathcal{M}, y) = \mathcal{D}(\mathcal{F}, \mathcal{M}(y(x))) \tag{4.1}$$

The deformation field should reduce the dissimilarity between the reference and warped moving images. In the proposed solution, the normalized gradient field (NGF) is used as a distance measure:

$$\mathrm{NGF}(R, T) = \frac{1}{2} \int_\Omega 1 - \left(\frac{\langle \nabla R(x), \nabla T(x) \rangle_{\varepsilon_R \varepsilon_T}}{\|\nabla T(x)\|_{\varepsilon_T} \, \|\nabla R(x)\|_{\varepsilon_R}} \right)^2 dx \tag{4.2}$$

where $\langle x, y \rangle_\varepsilon := x^\top y + \varepsilon$, $\|x\|_\varepsilon := \sqrt{\langle x, x \rangle_{\varepsilon^2}}$ and $\varepsilon_R, \varepsilon_T > 0$ are the so-called edge-parameters controlling influence of noise in the images.

Deformable registration is an ill-posed problem, where multiple solutions to the minimization process exist. In this context, limiting the number of possible solutions and penalizing unrealistic deformations is necessary. This is achieved by introducing additional regularization terms. In this work, two additional regularization terms are utilized. Thus, the objective function can be written as

$$\mathcal{J}(\mathcal{F}, \mathcal{M}, y(x)) = \mathcal{D}(\mathcal{F}, \mathcal{M}(y)) + \alpha \mathcal{R}(y) + \gamma \mathcal{V}(y) \tag{4.3}$$

The first regularization term is the curvature regularizer

$$\mathcal{R}(y(x)) = \int_{\Omega} \sum_{k=1}^{3} \|\Delta y(x)\|^2 dx \tag{4.4}$$

which penalizes deformation fields having too large second derivatives. To limit even further the available solutions, another regularization term

$$\mathcal{V}(y(x)) = \int_{\Omega} \psi(\det \nabla y(x)) dx \tag{4.5}$$

is added to the objective function, where $\psi(t) = (t-1)^2/t$ for $t > 0$ and $\psi(t) := \infty$ for $t \leq 0$. The volume change control is used to reduce foldings in the deformation field y(x) that may be generated during the minimization of the cost function. Folding in the deformation field represents an unrealistic transformation that the minimization process may lead to. The two hyperparameters γ and α control the influence of the regularization terms on the loss function.

Registration Parameters

Multi-level approaches register images at different scales (L) and are demonstrated to be useful to avoid local minima in the optimization of the objective function, and to speed up computational runtimes [78; 43; 93; 94]. The deformation field is initially computed on the coarsest level and the images are downsampled by a factor equal to 2^{L-1}. On a finer level, the previously computed deformation fields are utilized as an initial guess by warping the moving image. At each level, the moving and fixed images are downsampled. Our solution proposes a multi-level approach in which the registration is performed on three levels ($L = 3$).

Besides, the values of the stopping criteria for the optimization process are empirically set: the minimal progress, the minimal gradient, the relative one, and the minimum step length are set to 0.001, and the maximum number of iterations is set to 100. The loss weighting parameters are empirically set to $\alpha = 1$ and $\gamma = 1$.

The choice of the optimal transformation is conducted by using the quasi-Newton l-BGFS [80], due to its speed and memory efficiency.

4.2.1. Voxel-Based Weighting of the Regularization Term

This work proposes a voxel-based weighting parameter for the regularizer. By expressing the sum of the regularization terms 4.4 and 4.5 as

$$\mathcal{S}(y) = \alpha\mathcal{R}(y) + \gamma\mathcal{V}(y) \tag{4.6}$$

a voxel-based parameter $\rho(x) : \Omega \to \mathbb{R}$ multiplying $\mathcal{S}(y)$ is introduced to provide a local weighting of the regularizer ($\rho(x)\mathcal{S}(y)$). In registration approaches, in which no weighting is applied, the parameter multiplying the regularization terms is fixed and constant ($\rho = 1$). In the proposed method, its value is spatially dependent on x and can be locally controlled according to the voxel positions.

This work aims at excluding the contribution of the pathological tissues from the regularizer computation. To achieve this, the value of the weighting parameter $\rho(x)$ is set to zero in correspondence with voxels belonging to the pathological tissues and one in the other brain areas. The two classes of voxels (pathology and healthy tissues) are determined using segmentation masks of the reference images, automatically obtained by a freely available method (see next section for more details).

The objective function optimized by the proposed method assumes the following expression:

$$\mathcal{J}(\mathcal{F}, \mathcal{M}, y) = \mathcal{D}(\mathcal{F}, \mathcal{M}(y)) + \rho_y\mathcal{S}(y) = \mathcal{D}(\mathcal{F}, \mathcal{M}(y)) + \rho(x)(\alpha\mathcal{R}(y) + \gamma\mathcal{V}(y)) \tag{4.7}$$

The solution is proposed in two variants. In the first approach, the value of the hyperparameter is set to zero for the voxels included in the segmentation masks of the tumor. In the second one, the value of the parameter is set to one for each voxel ($\rho = 1$).

4.3. Dataset

This work utilizes the dataset released for the BraTS-Reg challenge, organized at the ISBI 2022 and MICCAI 2022 conferences [4]. The dataset includes 140 pairs of MRI volumes obtained before and after neurosurgical resection. Each pair includes two volumes, one acquired before surgery and the second after resection. Each case is characterized by four MRI sequences: native T1 (T1), contrast-enhanced T1-weighted (T1-CE), T2-weighted (T2), and T2 Fluid Attenuated Inversion Recovery (FLAIR). The set also includes the corresponding landmarks in each pair of volumes. These landmarks are obtained on anatomical locations such as blood vessel bifurcations, the anatomical shape of the cortex, and anatomical landmarks of the midline

of the brain. The total number of landmarks varies from case to case and across all cases in the range of 6-50 per scan. The volumes of each pair have previously preregistered to a standard reference [4]. The experiments conducted in this chapter are related to deformable solutions: As a normal practice with deformable solutions, an intra-patient rigid transformation is utilized to pre-align each pair of volumes. The baseline mean target registration error (mTRE) after rigid registration is 3.62 mm.

Several sequences are available. The previous chapter 3 demonstrated that the T1-CE is the most informative to guide the registration process in an iterative method. Thus, this is the sequence utilized in the following experiments.

Automatic Segmentation of pathological tissues

This work investigates the exclusion of pathological tissues from the computation of the regularization and distance measure terms. These issues are not provided in the public dataset. To segment them, the automatic method (BraTS Toolkit *btk*) [73]) described in the previous chapter 3 is utilized (please refer to section 3.2.1 for more details).

4.4. Evaluation

Both methods are evaluated and compared by computing the registration results on corresponding landmarks acquired on the 140 pairs of the training set of the BraTS-Reg dataset [4]. The deformation fields output by the proposed method are used to register the landmarks of the post-operative acquisitions. The goal of the deformable registration algorithm is to reduce the baseline mTRE (please refer to Appendix III for more details). Besides, as already described in [86], the non-parametric Wilcoxon signed-rank test is used to check the statistically significant difference between the results of each registration solution (iterative method and trained CNN models) and the baseline. This analysis tests whether the median of the differences between the two paired results is zero. The data distribution of the baseline registration errors is not normal according to the One-sample Kolmogorov-Smirnov test [84].

Moreover, the plausibility of the deformation fields generated by the two different approaches is computed. For a deformation field to be plausible, no folding should be present. However, image registration is an ill-posed problem, in which unwanted and unrealistic solutions might also be generated [95]. To measure the presence of

foldings, the Jacobian determinant of the deformation field ($det(J(y))$, also indicated as $det(J)$) can be computed [96; 97]. It provides a local measure for volume change. A value of 1 indicates no change. If $det(J) > 1$, there is a volume expansion, while if its value is less than 1, there is a local contraction. On the other hand, if $det(J) < 0$, a folding has occurred. Thus, the percentage of voxels in which $det(J) < 0$ is computed as a measure of the plausibility of the deformation fields [4].

4.5. Results

The baseline mTRE is reduced to 2.43 mm by the method not excluding the pathological tissues, whereas the solution discarding the pathology from the registration process achieves an mTRE value of 2.38 mm. The boxplots in Fig. 4.1 provide a more detailed comparison between the baseline registration results and the outcome of the proposed methods. According to the Wilcoxon test, the null hypothesis cannot be accepted for both methods (p-value $< 10^{-21}$).

Qualitative results are provided in Figs 4.2, 4.3. In particular, Fig. 4.2 shows an example of the registration results for the two automatic methods. The first row shows two corresponding slices of pre- and postoperative images before deformable registration. The slices of the preoperative images after being warped by the deformation fields obtained by the two methods are available in the second row. Fig. 4.3 shows how the landmarks registration is improved before and after applying the method excluding the contribution of the pathological tissues. The landmarks of the template and reference images are colored respectively green and purple. The two subfigures in the first column display the baseline location of the corresponding landmarks from two different views. The second column shows the registration of the landmarks after the application of the computed deformation field: The new positions of the purple markers are obtained after applying the deformation field generated by the registration method excluding the pathological tissues.

The percentage of foldings for both methods is $< 0.1\%$.

4.6. Discussion and Conclusion

This work investigates how much influence the exclusion of the pathology from the computation of the regularizer has on an iterative method non-rigidly registering pre- and post-operative MRI volumes. Two solutions, based on the same method,

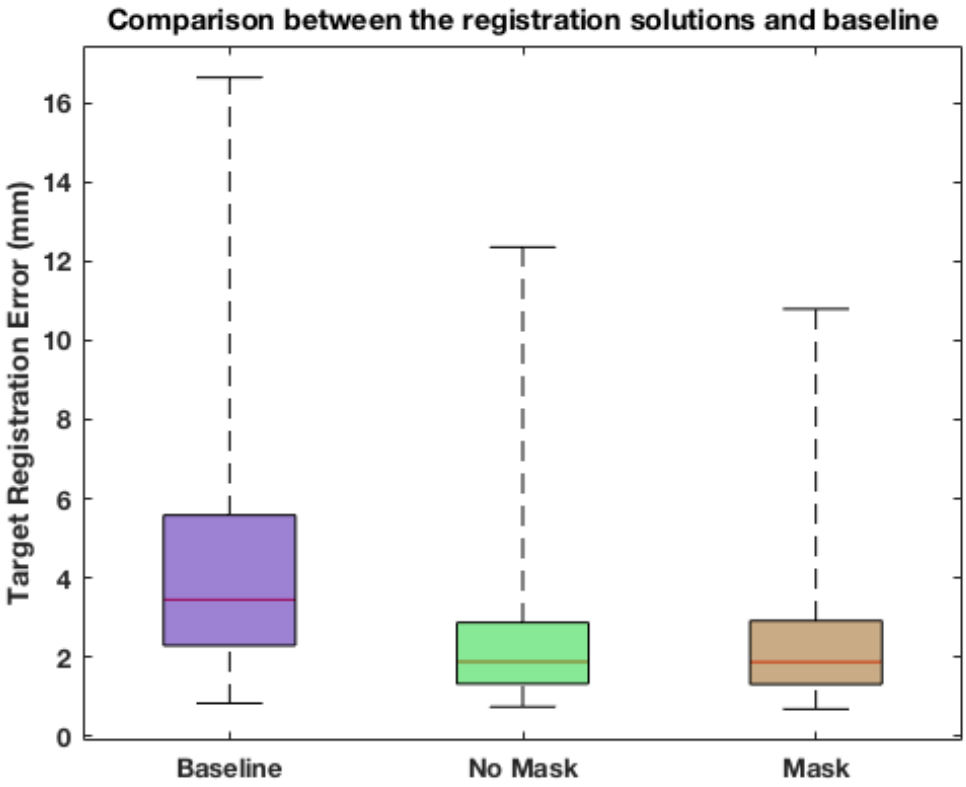
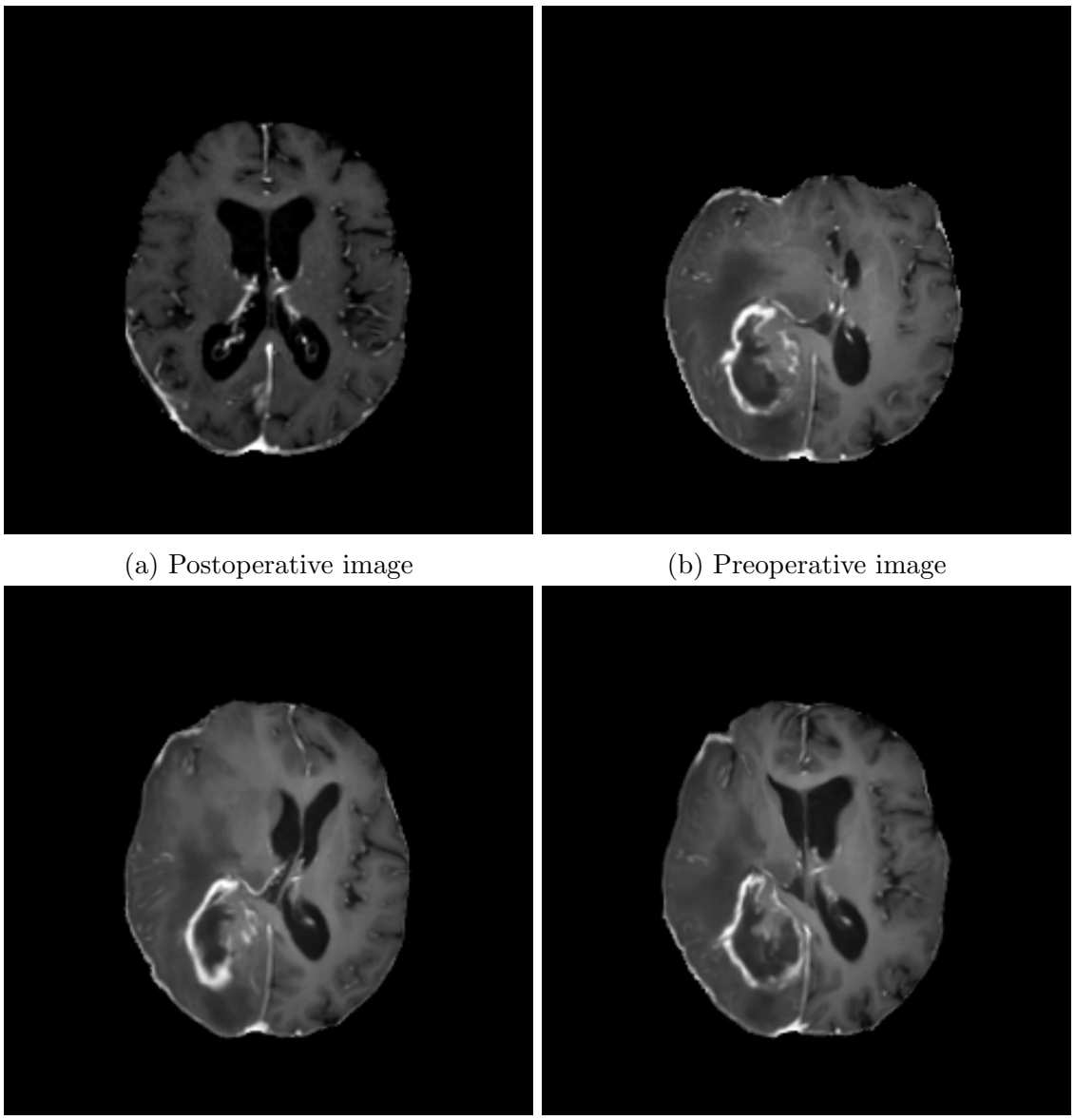


Figure 4.1.: Registration results visualized with boxplots. The baseline mTRE is compared to the method not excluding the pathological tissues and the one excluding them from the regularization terms.



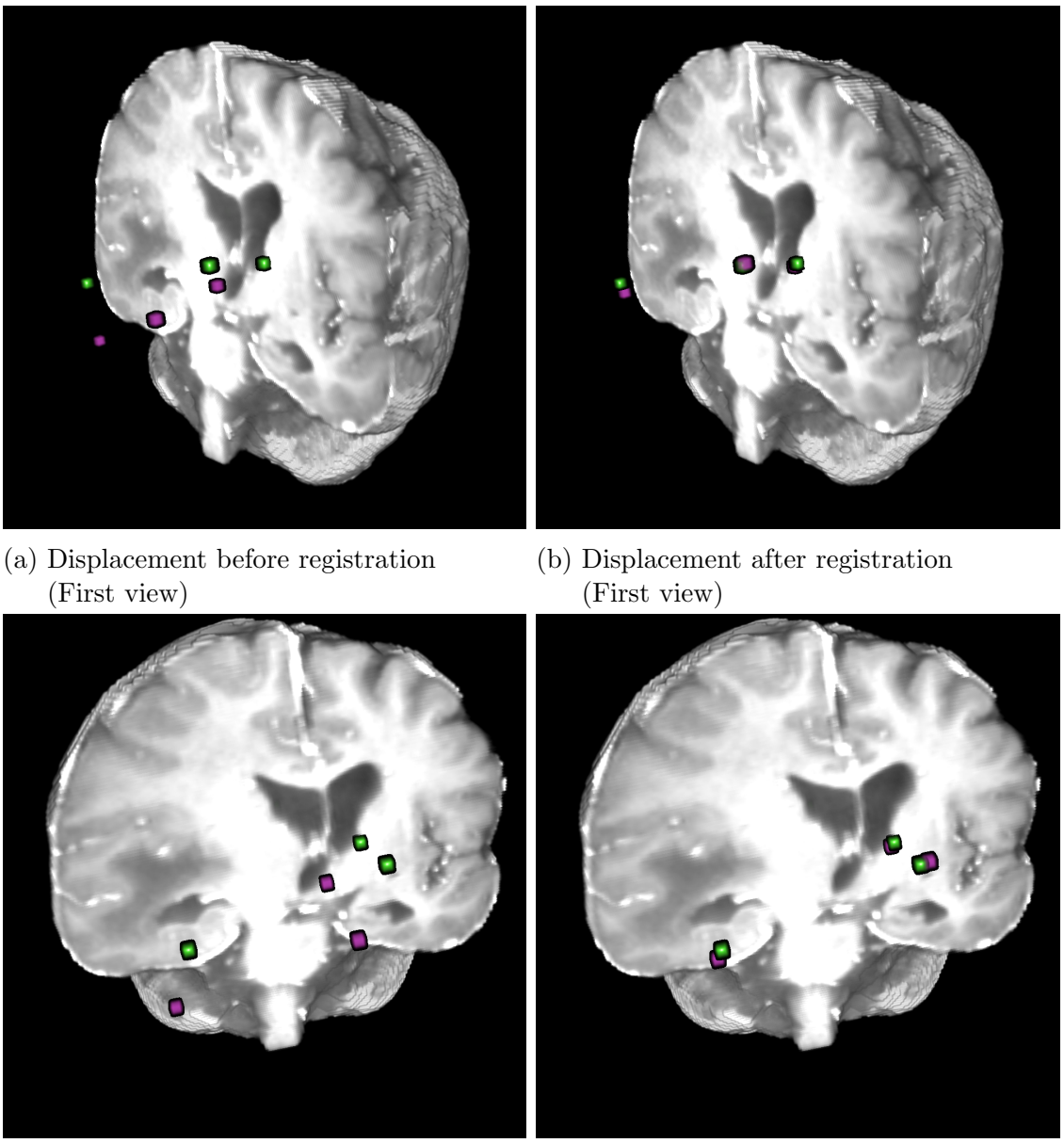
(a) Postoperative image

(b) Preoperative image

(c) Warped preoperative image (Method No Mask)

(d) Warped preoperative image (Method Mask)

Figure 4.2.: Registration results for the two registration methods. Subfigures a) and b) show the post- and pre-operative images (respectively, reference and template images). In subfigure b) it can be observed how much deformation the tumor produces in the surrounding tissues: For example, the lateral ventricles appear compressed by the pathological tissues. The second row shows the corresponding warped preoperative slices obtained after applying the proposed methods. In subfigures c) and d), the solutions lead to a better registration: The lateral ventricles, the brain parenchyma, and the sulci are more aligned to the post-operative image. Moreover, comparing d) versus c) (for example, the lateral ventricles), the improvements in the second subfigure appear to be more evident.



(a) Displacement before registration (First view)

(b) Displacement after registration (First view)

(c) Displacement before registration (Second view)

(d) Displacement after registration (Second view)

Figure 4.3.: 3D visualization of landmarks before and after registration. Corresponding landmarks of the reference and template images are colored green and purple; A partial volume of the template brain parenchyma is also shown to give spatial information about their location. Two different views of the same landmarks and brain are given for a better understanding of their displacement before and after registration. The first column shows the landmarks' location before applying any deformable registration. In the second column, reference landmarks are registered by applying the deformation field computed by the method excluding the contribution of the pathology. In subfigures b) and d), landmarks appear closer than their initial location in the first column. Thus, the proposed method leads to an improvement in the registration results.

are applied by including and excluding the contribution of the pathology from the registration process. According to the experiments presented in this work, the initial mTRE is improved if the pathology is completely discarded from the registration method (see Fig. 4.1). By discarding the pathology from the regularizer, the plausibility of the generated deformation field is not negatively influenced. From the visual results in Fig. 4.2, it can also be observed how the solution excluding the contribution of the pathological tissues from the registration process leads to a more accurate registration of healthy brain tissues.

In future work, additional classes of healthy brain tissues could be considered in the weighting of the regularization terms. In fact, the weighting parameter proposed in this work is not limited to a binary mask: Values different from 0 and 1 can also be utilized. It might be beneficial for some tissues to have more influence than others on the regularization term since they could get more realistically deformed [91; 92]. To accomplish this, masks of healthy tissues should be first obtained.

Other works already investigated the adaptation of the regularizer to different tissues in CNN-based solutions. In future experiments, the weighting of the regularization term could also be tested in deep learning-based approaches for the registration of subsequent MRI acquisitions.

5. Comparison of Different Automatic Solutions for Resection Cavity Segmentation in Post-operative MRI Volumes Including Longitudinal Acquisitions

The material of this chapter is available as a conference paper published in Proceedings Volume 11598, Medical Imaging 2021: Image-Guided Procedures, Robotic Interventions, and Modeling, and presented at the SPIE Medical Imaging 2021 conference.

Canalini, L., Klein, J., de Barros, N. P., Sima, D. M., Miller, D., Hahn, H. (2021). Comparison of different automatic solutions for resection cavity segmentation in post-operative MRI volumes including longitudinal acquisitions. Proceedings Volume 11598, Medical Imaging 2021: Image-Guided Procedures, Robotic Interventions, and Modeling. SPIE. doi: 10.1117/12.2580889

Abstract

Purpose Glioblastoma multiforme (GBM) represents the most common primary brain tumor. After its neurosurgical resection, radiation therapy (RT) is indicated as a post-operative adjuvant treatment to limit possible regrowth. In RT planning, the target volume is carefully decided including also the resection cavity originating from neurosurgery. The treatment is split into consecutive sessions to reduce side effects. Throughout RT the pathological tissues often change their size and shape, and the target volume defined in the initial planning has to be updated. Magnetic resonance imaging (MRI) data can be acquired before each session to manually redefine the resection cavity contours. However, the manual segmentation of this structure is a tedious and time-consuming task.

Method In this work, we compare five deep-learning solutions to automatically segment the resection cavity in postoperative MRI. The proposed methods are based on the same 3D U-Net, widely used to tackle segmentation tasks. We use a dataset of post-operative MRI volumes including also longitudinal acquisitions. Each case counts four MRI sequences and comprises the ground truth of the corresponding resection cavity. Four solutions are trained each with a different MRI sequence. Besides, a method designed with all the available sequences is also presented.

Results In this work, we compare all the solutions to find which one obtains the best DICE index computed between the automatically generated masks and ground truth. Our experiments show that the method trained only with the T1 weighted contrast-enhanced MRI sequence achieves the best results, with a median DICE index of 0.81. The 3D U-Net trained on the FLAIR sequence achieves the lowest value 0.44. The solutions trained on all the MRI protocols obtain a DICE of 0.77.

Purpose The results obtained in this work show that the choice of the MRI sequence has a strong effect on the training of deep learning solutions to segment resection cavities. Thus, future solutions aiming to segment this structure should take our results into account to obtain the best outcome.

5.1. Introduction

Glioblastoma multiforme (GBM) is the most common type of intracranial tumor and has a very poor prognosis [17]. Tumor resection is indicated as the first treatment [11]. However, given the infiltrating nature of GBM, a maximal excision is often unachievable. Therefore, radiotherapy (RT) is normally performed as a post-operative treatment [98]. It uses high-energy radiations on a target volume specified in RT planning, to destroy pathological cells and limit tumor recurrence. However, radiation exposure can also damage healthy tissues. To limit side effects, the treatment is usually split into several sessions in which a small dose of the radiation is delivered [17]. The healthy tissue can recover faster from a small fraction of the dose, reducing RT side effects. Besides, careful planning of the target volume is important to limit the radiation dose only to a precise structure. In intracranial RT planning, the target volume also includes the resection cavity's contours [98]. However, throughout the RT, the anatomical volumes often change. This is also true for the resection cavity, which can go under severe modifications regarding its shape, size, and intensity [99; 6]. An example of the possible alterations is visible in Fig. 5.1, showing two subsequent acquisitions for the same patient. The second and third columns are respectively related to the FLAIR and T1 MRI sequences and show how the intensities of a resection cavity can differ in two consecutive acquisitions. Since it is important to focus the radiation therapy only on the structures of interest, an update of the target volume becomes necessary. Besides computed tomography (CT) scans, post-operative MRI is a valid alternative to obtain updated images of the resection cavity. By observing this data, the contours of the cavity can be manually modified, and the target volume updated. However, manual segmentation is tedious work for physicians [100], who would benefit from having an automatic method to accomplish this task. Despite the importance of the resection cavity in the postsurgical phases, very few solutions have been proposed to automatically delineate its contours. The authors in [100] introduced an automatic method to automatically segment the resection cavity in post-operative MRI. They demonstrated that a convolutional neural network (CNN) can be a valid alternative to manual segmentation. They used multi-sequence MRI volumes to train their solution. However, they didn't use post-operative MRI volumes acquired at different

sessions of the same radiotherapy treatment. An automatic solution tackling the segmentation of the resection cavity also in longitudinal MRI acquisitions is still missing. In this work, we investigate five different automatic solutions to segment resection cavities in post-operative MRI volumes which also include longitudinal studies. Every method is based on the same CNN based on the 3D U-Net [42], which has become a standard approach in automatic segmentation tasks in medical imaging. Four of the proposed solutions are trained with a distinct MRI sequence, whereas only one approach uses all the available MRI sequences together. A comparison between all the trained models is performed to check which one obtains the best results. This could give a better understanding of which MRI sequences may be the most informative to segment the resection cavity in post-operative MRI.

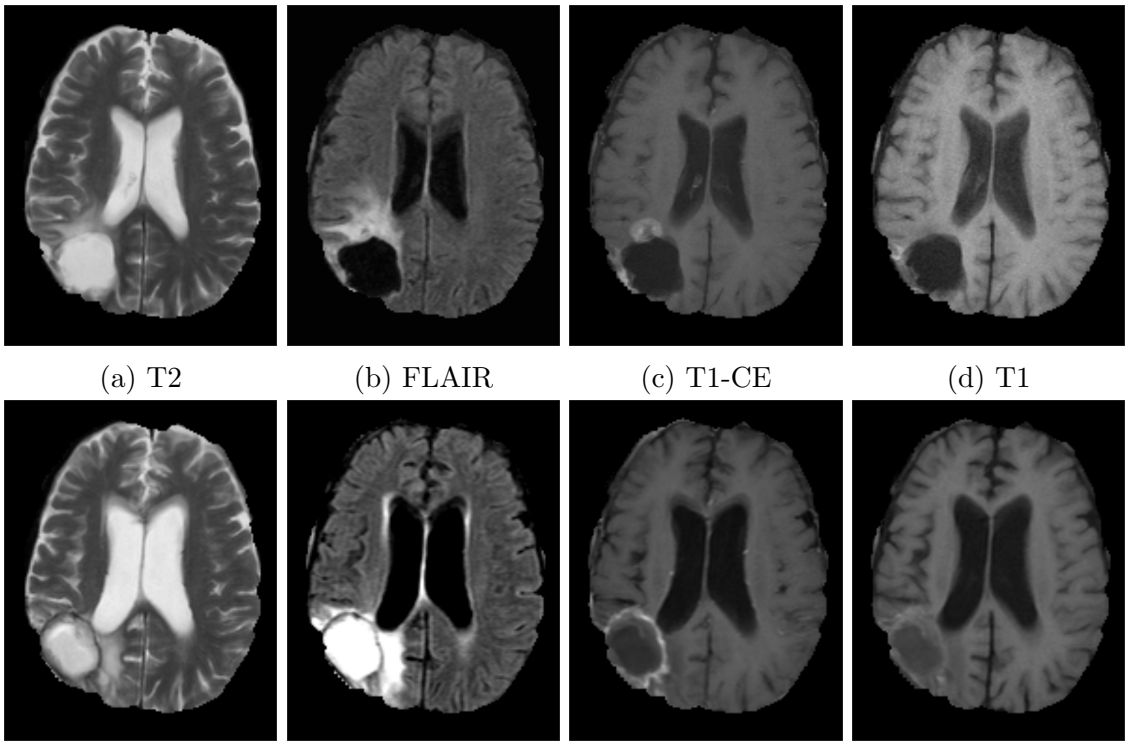
5.2. Material and Methods

5.2.1. Data

In this work, we utilized the data coming from BraTS 2015 [101], a public dataset including a mixture of pre and post-operative MRI images. The volumes come already skull-stripped, co-registered to the same anatomical template, and interpolated to 1 $mm\hat{3}$ voxel resolution. For our experiments, we selected 47 post-operative volumes, in which the resection cavity is clearly visible. This data is related to 14 different patients in which high-grade gliomas have been resected. The utilized dataset contains also longitudinal studies for eight patients who have been scanned subsequent times. An example of the volumes used for our work is visible in Fig. 5.1, in which we can observe the same patient scanned in two subsequent MRI acquisitions. The original challenge focused on the segmentation of other tumor tissues but not on the resection cavity, which therefore wasn't originally segmented. Thus, we manually annotated the structures of interest by looking at the resection cavity on the different MRI sequences. Figs. 5.3 and 5.4 show two examples of the ground truth annotated for this work (highlighted with green contours).

5.2.2. Methods

We here propose five different solutions based on the same CNN architecture. We utilized a 3-leveled 3D U-Net [42] with a receptive field size of $44 \times 44 \times 44$ voxels. The patches containing the background are a larger quantity than the ones including



(a) T2 (b) FLAIR (c) T1-CE (d) T1

(e) T2 (f) FLAIR (g) T1-CE (h) T1

Figure 5.1.: Resection cavity changes in two subsequent MRI acquisitions of the same patient. The first row contains four images, each showing different MRI sequences (T2 weighted, FLAIR, T1 weighted, and T1 contrast-enhanced) acquired during the same session. The second row shows the same sequences acquired in a subsequent session.

resection cavities labels. Thus, to speed up the training procedure, the patches' composition has been modified: 20 percent of them include only background voxels, whereas 80 percent contain at least a foreground voxel. Moreover, as a loss function, we used the Tversky Loss to weigh recall higher than precision (a = 0.2, b = 0.8) [102]. It is used to give more penalty on the false negative predictions so that a larger amount of voxels will be segmented as foreground. Even if the false positive predictions may increase, a lower number of foreground voxels will be missed. All the proposed solutions are trained with a batch of size 5 and the best model is saved every 100 iterations based on the Jaccard coefficient computed on the validation data. Four of the proposed solutions are trained each with a different MRI sequence, and only one has been trained with all 4 sequences together. Besides, we utilized a five cross-validation procedure to test our algorithms. The 47 cases are split into five disjoint groups, three of them composed of nine volumes each, and two with ten volumes. Each of these disjoint groups represents a test set and the remaining volumes are used to train and validate our solutions. Thus, each proposed method is trained five times and tested always on a disjoint test set. Furthermore, in the inference process, we apply connected component analysis to keep only the largest segmented mask. The Tversky loss function could lead to an over-segmentation of the structure of interest. However, only one resection cavity is present per volume, so we discard smaller structures that may be wrongly segmented.

5.3. Results

We compute the DICE indices between the ground truth and the automatically generated masks [103]. The results of the different approaches are visible in Fig. 5.2, where five box plots show the median DICE values. As we can observe, the model trained with only the T1-CE MRI sequence obtained a median DICE of 0.81 in our experiments. On the contrary, the 3D U-net trained only on the FLAIR sequence achieved the lowest score median DICE of 0.44. The method using all four modalities together reaches a median DICE index of 0.79, which is slightly lower than the approach trained with only T1-CE, but higher than the other solutions (0.73 and 0.77 are the median DICE values for the solutions trained with T1 and T2). Besides, the Wilcoxon sign-rank test is performed between the outperforming method and each of the other solutions to verify if the differences are statistically significant. The results of this test are available in Fig. 5.2. The test shows that

there are statistically significant differences with a p-value<0.01 between the best solution and methods trained only with T1, FLAIR, and T2. Besides, there is also a statistically significant difference with a p-value<0.05 between the outperforming solution and the model trained with all the sequences. The visual results for the solutions trained with a single MRI sequence are available in Fig. 5.3. Thus, T1-CE seems to be the most informative sequence to train the proposed method. This may be because the resection cavities observed on T1 contrast-enhanced images usually have uniform characteristics among different acquisitions. In fact, in T1-CE images, the cerebral spinal fluid (CSF) within the resection cavity often appears hypointense and delimited by a hyperintense border highlighted by the contrast enhancement. Thus, it may be easier for the solution trained with only T1-CE to identify unique features to segment the structures of interest. On the FLAIR sequence, the cavities can assume instead very different characteristics even in longitudinal sequences, mainly due to tumor recurrences or radiotherapy treatment [6]. As an example, the second column in Figure 1 shows the same resection cavity observed on the FLAIR sequence during two subsequent acquisitions: The CSF within the resection cavity has very different intensities between the two images. Thus, the method trained only with the FLAIR sequence may not be able to find a specific set of features to correctly characterize the resection cavity. The variability in the intensity could also explain the segmentation errors observed in the second column of Fig. 5.3. The segmented structure (highlighted in orange) includes both hypo and hyperintense parts, mainly because the resection cavity can be characterized by both of them.

Besides, the multisequence approach performs worse than the solution designed with only the T1-CE sequence. This is surprising because the method trained with multiple MRI sequences should learn more features, potentially leading to a better outcome. However, there may be some MRI sequences that negatively affect the task, especially considering what has already been discussed for the FLAIR MRI sequence. The visual results for this solution are available in Fig. 5.4.

5.4. Discussion

To the best of our knowledge, we proposed the first comparison between different solutions aimed at segmenting the resection cavity in post-operative MRI including also longitudinal studies. We compare five approaches based on the same 3D U-Net architecture, each trained on five different MRI sequence combinations. Four meth-

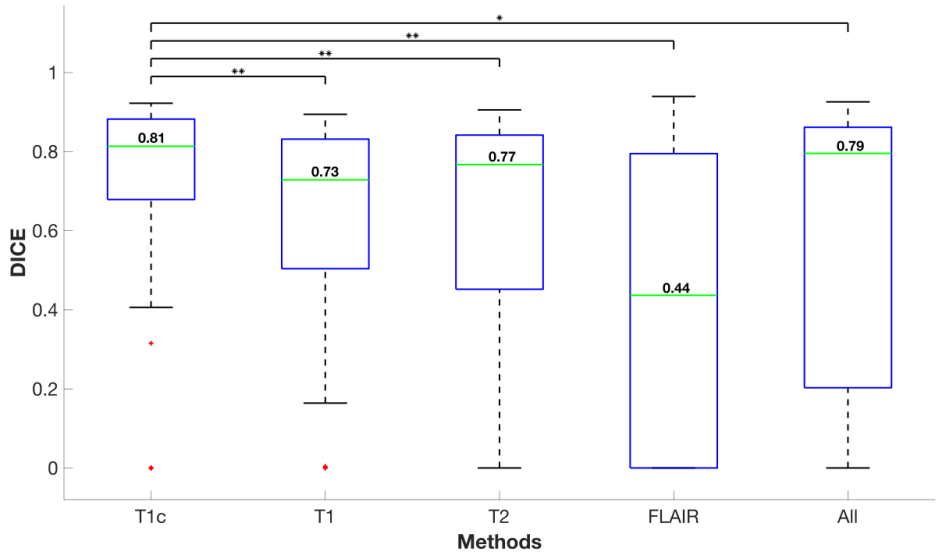
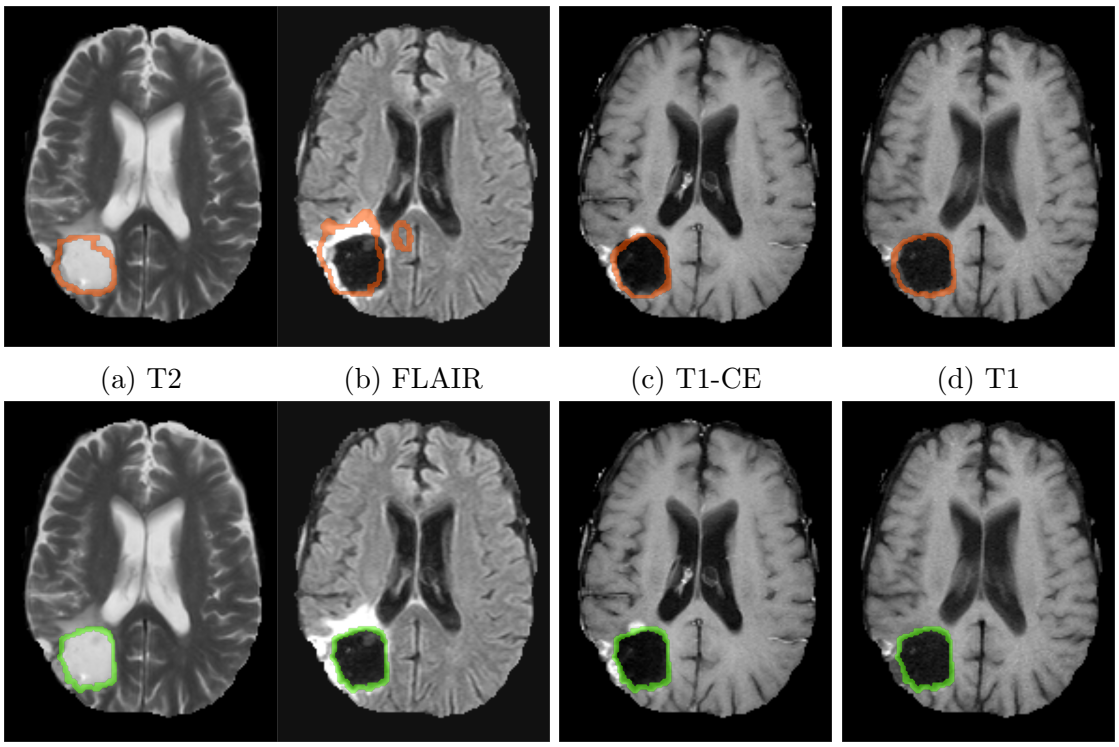


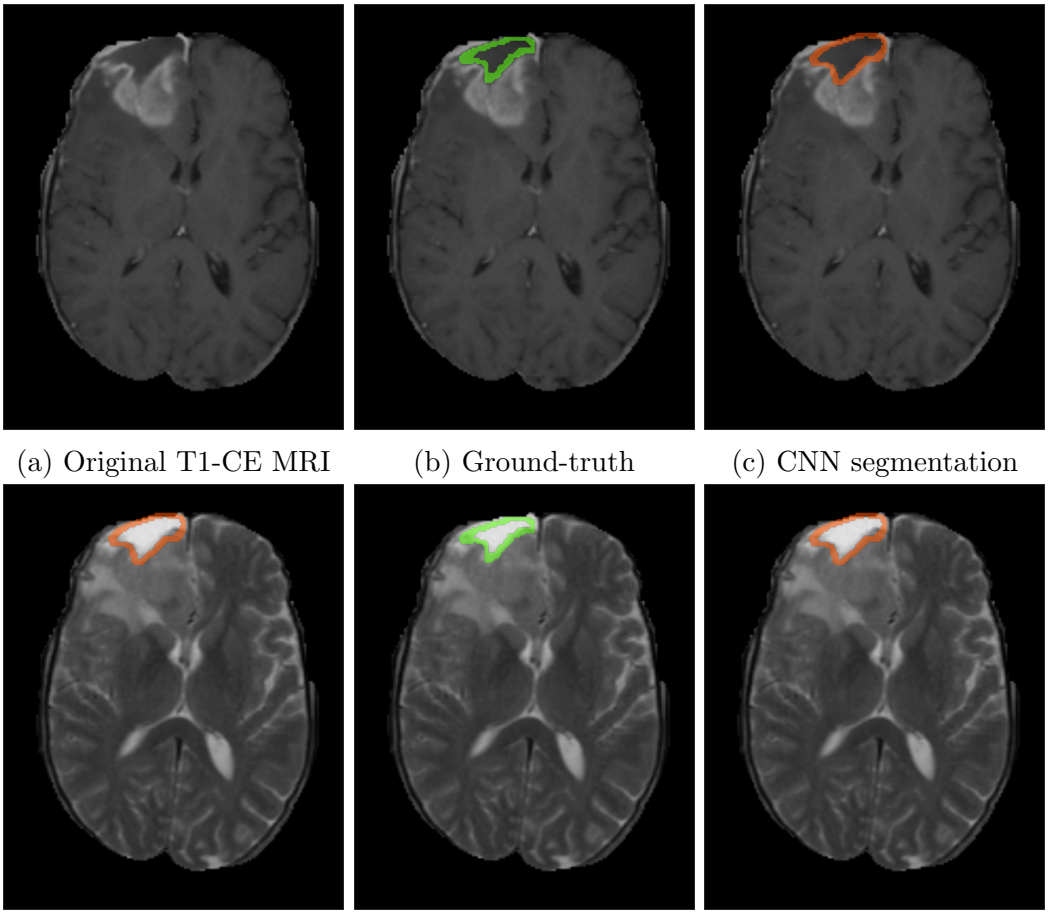
Figure 5.2.: Figure 2: DICE results for the five solutions. The methods are listed on the x-axis, whereas the y-axis reports the DICE values. The results of every solution are summarized in the corresponding box plot. The median DICE indices obtained on the whole dataset are indicated by a green line (with the corresponding values written on top of it), whereas the outliers are highlighted with plus-symbols in red color. The model trained only with the T1-CE sequence is the outperforming solution and shows a statistically significant difference in comparison to the other methods (according to the Wilcoxon sign rank test). The lines on top of the graph relate the T1-CE method with the other solutions. The asterisk on top of each line indicates if the statistically significant difference between the solution trained with T1-CE and another method is verified with a p-value < 0.05 (one asterisk) or < 0.01 (two asterisks).



(a) T2 (b) FLAIR (c) T1-CE (d) T1

(e) T2 (f) FLAIR (g) T1-CE (h) T1

Figure 5.3.: Figure 3: Resection cavity segmentation results for the four solutions trained on a different MRI sequence. From left to right, the first row shows the visual segmentation results obtained by the methods trained on T2, FLAIR, T1, and T1-CE sequences. The second row displays the corresponding ground truth observed in the four different MRI sequences. The automatically generated segmentation of the resection cavity is highlighted by an orange border (first row). The ground truth is highlighted in green in the second row.



(a) Original T1-CE MRI (b) Ground-truth (c) CNN segmentation

(d) Original T2 MRI (e) Ground-truth (f) CNN segmentation

Figure 5.4.: Figure 4: Resection cavity segmentation results for the four solutions trained on a different MRI sequence. From left to right, the first row shows the visual segmentation results obtained by the methods trained on T2, FLAIR, T1, and T1-CE sequences. The second row displays the corresponding ground truth observed in the four different MRI sequences. The automatically generated segmentation of the resection cavity is highlighted by an orange border (first row). The ground truth is highlighted in green in the second row.

ods are designed only with a single MRI sequence, whereas one approach is trained with all four different sequences. Our experiments show that training performed with only T1 post-contrast MRI sequence achieves the best results, also when compared with the multi-sequence approach. Besides, the outperforming method shows a statistically significant difference compared to the other methods. On the contrary, the solution trained only with the FLAIR sequence achieves the lowest DICE score. This work is not being or has not been submitted for publication or presentation elsewhere.

5.5. Conclusions

We have proposed a comparison between five different solutions based on the same 3D U-Net architecture to segment the resection cavity in post-operative MRI volumes including longitudinal studies. Each method is trained with a different combination of MRI sequences. Our experiments showed that the solution trained with only the T1-CE obtains the best results. This represents a good starting point for further investigations into resection cavity segmentation, where very few automatic solutions have been proposed so far. Instead of using all the available MRI sequences, the next solutions may utilize only the most informative ones. For future work, different deep learning architectures could be tested to verify which one would be the best to segment the desired output. Besides, multi-label approaches could be investigated, in which other pathological tissues (for instance edema and active tumor) are segmented together with the resection cavity.

Acknowledgements

This work was funded by the H2020 Marie-Curie ITN TRABIT (765148) project.

Part II.

Segmentation-Enhanced Registration of Intraoperative Ultrasound Volumes

6. Intraoperative Ultrasound Imaging in Neurosurgery

Intraoperative ultrasound (iUS) imaging can assist neurosurgeons in the resection of brain tumors [104; 105]. It represents a lower-cost solution compared to intraoperative MRI, and US volumes are relatively fast to be acquired. Thanks to these advantages, iUS data can be acquired at different stages of the neurosurgical procedures 6.1.

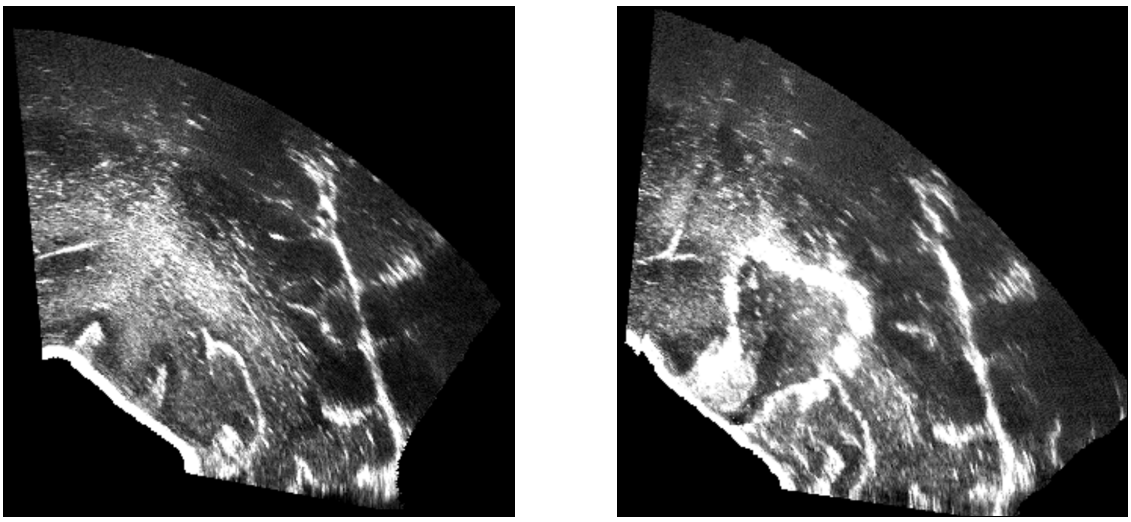
An initial iUS acquisition is usually obtained just after opening the skull, before the tumor resection starts, to better identify and localize the pathology [106]. Thanks to a neuronavigation system, these data can also be intraoperatively observed. However, the more the surgery advances, the more unreliable this initial acquisition becomes due to brain shift effects: The tissues observed in the US data have a different position and conformation from what is observed in the surgical scene. Therefore, additional intraoperative US images can be required to provide an updated visualization [106; 107; 108]. Images acquired during and at the end of the resection help to understand how much pathology has still to be removed. However, the image quality of US data acquired at the late stages of the resection degrades due to brain shift and the saline solution utilized to fill the cavity [52]. To ease the understanding of these images, they can be visually compared with the US data acquired just after the opening of the dura mater, in which brain shift effects are not present yet. Neuronavigation systems can register intraoperative images acquired at different surgical phases. However, these devices are prone to technical inaccuracies, which affect the registration procedure from the beginning of the resection [23]. Moreover, the commonly available neuronavigation systems only compute a rigid transformation, which is not sufficient to compensate for the anatomical changes caused by brain shifts.

6.1. Registration of iUS Volumes

Image registration algorithms make the visual inspection of subsequent iUS acquisitions more precise and, thus, improve neurosurgery outcomes for tumor resection. Image registration methods usually assume that a one-to-one correspondence is guaranteed for each voxel of the images to be registered. However, this assumption is often not verified when dealing with iUS. In fact, the registration of iUS is extremely challenging due to the lack of correspondence between successive acquisitions (see Fig. 6.1 as an example). In early acquisitions obtained after the opening of the dura mater, the tumor is still intact, whereas, at later stages of the procedure, pathology is removed and the resection cavity is formed. Therefore, a method aiming to register iUS should have a strategy for dealing with the lack of matching structures between subsequent iUS acquisitions. It also has to provide a deformable transformation to accommodate the brain shift effects. In the last years, the research about US registration has been pushed by the release of two public datasets including intraoperative subsequent US acquisitions with corresponding landmarks [109; 29; 110]. The authors in [111; 112] proposed a siamese network to detect corresponding landmarks between subsequent iUS. The detected landmarks were then used to estimate an affine transformation to register the US data. Besides, the authors in [113] present a GAN-based method, in which a generator calculates an affine transformation whereas the discriminator learns the similarity measure utilized to guide the generator during training.

Chapter 7 proposes a *segmentation-based solution* for registering intraoperative US volumes. It offers a combination of rigid and deformable registration steps. Besides, it tackles the lack of correspondence between subsequent acquisitions by computing the distance measure of the objective function solely on matching healthy structures. First, the solution automatically segments sulci and falx cerebri in subsequent acquisitions by using a 3D U-net-based method. Second, the segmented structures are used as input to the registration method. The deformation field obtained by registering the masks is then applied to the original US volumes.

Chapter 8 introduces a *segmentation-enhanced solution* to register iUS volumes. A combination of rigid and deformable steps is utilized. The method can be seen as complementary to the first solution 7 because it registers the original US acquisitions



(a) iUS at early neurosurgical stages

(b) iUS at late neurosurgical stages

Figure 6.1.: Example images of intrapatient intraoperative ultrasound volumes acquired before (a) and after (b) tumor resection. In neurosurgery, the pathology is removed and the resection cavity gets formed. The resection cavity has no specific counterpart in the initial acquisitions. In the process of registering intrapatient volumes, it would be beneficial to exclude the contribution of non-corresponding elements of the resection cavity and rely only on the structures which remain visible in subsequent acquisitions.

by excluding the contribution of non-corresponding structures from the distance measure computation. The resection cavities are not visible in US acquisitions obtained just after opening the dura mater, but only at the intermediate and late stages of the surgery. By excluding their contribution from the distance measure computation, the minimization of the objective function can focus on the healthy structures, for which a one-to-one correspondence is available.

7. Segmentation-based Registration of Ultrasound Volumes for Glioma Resection in Image-guided Neurosurgery

The material of this chapter is available as an open-access journal paper published in the International Journal of Computer Assisted Radiology and Surgery.

Canalini, L., Klein, J., Miller, D., Kikinis, R. (2019). Segmentation-based registration of ultrasound volumes for glioma resection in image-guided neurosurgery. Int. J. CARS, 14(10), 1697–1713. doi: 10.1007/s11548-019-02045-6

Abstract

Purpose In image-guided surgery for glioma removal, neurosurgeons usually plan the resection on images acquired before surgery and use them for guidance during the subsequent intervention. However, after the surgical procedure has begun, the preplanning images become unreliable due to the brain shift phenomenon, caused by modifications of anatomical structures and imprecisions in the neuronavigation system. To obtain an updated view of the resection cavity, a solution is to collect intraoperative data, which can be additionally acquired at different stages of the procedure to provide a better understanding of the resection. A spatial mapping between structures identified in subsequent acquisitions would be beneficial. We propose here a fully automated segmentation-based registration method to register ultrasound (US) volumes acquired at multiple stages of neurosurgery.

Methods We chose to segment sulci and falx cerebri in US volumes, which remain visible during resection. To automatically segment these elements, first we trained a convolutional neural network on manually annotated structures in volumes acquired before the opening of the dura mater, and then we applied it to segment corresponding structures in different surgical phases. Finally, the obtained masks are used to register US volumes acquired at multiple resection stages.

Results Our method reduces the mean target registration error (mTRE) between volumes acquired before the opening of the dura mater and during resection from 3.49 mm ($\pm$ 1.55 mm) to 1.36 mm ($\pm$ 0.61 mm). Moreover, the mTRE between volumes acquired before opening the dura mater and at the end of the resection is reduced from 3.54 mm ($\pm$ 1.75 mm) to 2.05 mm ($\pm$ 1.12 mm).

Conclusion The segmented structures demonstrated to be good candidates to register US volumes acquired at different neurosurgical phases. Therefore, our solution can compensate for brain shift in neurosurgical procedures involving intraoperative US data.

7.1. Introduction

In brain surgery for tumor removal, neurosurgeons usually plan the intervention on pre-surgical images. The most widely used modality for neurosurgery planning is magnetic resonance imaging [27; 114; 21]. To help physicians with the resection, neuronavigation systems can be used to link preplanning data positions to patients' head locations. By tracking fiducial markers placed on the patient's skull and surgical tools, an optical system computes an image-to-patient transformation. Consequently, by pin-pointing an intracranial location, neurosurgeons can obtain the same position in the preplanning images. However, initialization inaccuracies of the neuronavigation system may invalidate the image-to-patient transformation, affecting the quality of these images since the beginning of the resection [23]. Additionally, after resection starts, the preplanning data become even more unreliable due to the brain shift phenomenon: Structures observed in preplanning images don't remain in the same conformation and position during tumor removal [23]. As a consequence, the probability that pathological elements are missed increases, reducing the survival rates of the operated patients [115; 116]. To overcome this problem, intraoperative images can be acquired [117]: They provide an updated view of the ongoing procedure and hence compensate for the brain shift effects. A solution is represented by intraoperative magnetic resonance imaging (iMRI) [118]. It is demonstrated to be a good option [119] since its high image quality provides good contrast in anatomical tissue even during the resection [120]. However, the high costs of iMRI and the architectural adaptations required in the operating room seem to prevent this modality from being deployed more widely. A valid alternative is given by intraoperative ultrasound (iUS) [121; 122; 123; 124]. Some authors reported that for certain grades of glioma, iUS is equal or even superior to iMRI in providing good contrast between tumor and adjacent tissues [106; 125]. Moreover, US represents a lower-cost solution compared to MRI. In our work, we focus on intraoperative 3D ultrasound used in neurosurgical procedures.

The more the resection advances, the more the initial acquisition of iUS becomes unreliable due to increased brain shift effects. Therefore, an update of the intraoperative imaging may be required. In [126], the authors acquired US volumetric data in subsequent phases of glioblastoma resections in 19 patients and compared the ability

to distinguish tumors from adjacent tissues at three different steps of the procedure. According to their observations, the 3D images acquired after opening the dura, immediately before starting the resection (we indicate this phase as before resection), are highly accurate for delineating tumor tissue. This ability reduces during resection, i.e., after most of the resection has been performed but with residual tumor, and after resection, i.e., when all the detected residual tumor has been removed. In fact, the resection procedure itself is responsible for creating small air bubbles, debris, and blood. Besides this, a blood clotting-inducing material, commonly used during neurosurgical procedures, causes several image artifacts [106; 108]. Successive studies regarding other types of tumor resection confirmed the degradation of image quality in US during resection [52]. Therefore, it would be helpful to combine US images acquired during and after resection with higher-quality data obtained before resection. Such a solution may also be beneficial to improve the registration of intraoperative data with higher-quality preplanning MRI images. In fact, instead of combining directly degraded US data with preplanning imaging, it would be useful to register first the pre-surgical MRI data with US volumes acquired before resection, in which few anatomical modifications occurred. Afterward, intraoperative US data acquired at the first stage of the surgery (which therefore has a higher quality) may be registered to subsequent US acquisitions, and then the preplanning data could be registered to those by utilizing a two-step registration [127]. In this context, neuronavigation systems could be used to co-register intraoperative images acquired at different surgical phases. However, these devices are prone to technical inaccuracies, which affect the registration procedure from the beginning of the resection [23]. Moreover, the available neuronavigation systems usually offer only a rigid registration, which is not sufficient to address anatomical changes caused by brain shift. In our work, we propose a deformable method to improve the registration of US volumes acquired at different stages of brain surgery.

Few solutions have been proposed to improve the US–US registration during tumor resection in neurosurgery. In [128], the authors studied the performance of the entropy-based similarity measures joint entropy (JE), mutual information (MI), and normalized mutual information (NMI) to register ultrasound volumes. They conducted their experiments with two volumes of an US calibration phantom and two volumes of real patients, acquired before the opening of the dura mater. Different rigid transformations were applied on each volume, and the target registration error (TRE) was used as an evaluation metric. The accuracy of the registration

was examined by comparing the induced transformation to move the original images to the deformed ones, with the transformation defined by the entropy-based registration method. In both of the datasets, NMI and MI outperformed JE. In another work [129], the same authors developed a non-rigid registration based on free-form deformations using B-splines and using normalized mutual information as a similarity measure. Two datasets of patients were used, where for each case a US volume was acquired before the opening of the dura, and one after (but before the start of tumor resection). To assess the quality of the registration, the correlation coefficient was computed within the overlap of both volumes and before and after registration. Furthermore, these authors segmented the volumetric extension of the tumor with an interactive multi-scale watershed method and measured the overlap before and after the registration. One limitation of the aforementioned two studies is that no experiment is conducted on volumes acquired at different stages of the surgical procedure, but only before the resection actually starts. In a real scenario, neurosurgeons use intraoperative data to find residual tumors after the first resection, which is conducted after the opening of the dura mater.

One of the first solutions to register US data obtained at subsequent surgical phases utilized an intensity-based registration method to improve the visualization of volumetric US images acquired before and after resection [130]. The results are computed for 16 patients with different grades of brain supratentorial tumors located in various lobes. Half of the cases were first operations, and half were re-operations. Pre-resection volumes were acquired on the dura mater, or either directly on the cortex (or tumor) or on a dura repair patch. The post-resection ultrasound was used to find any residual tumor. The authors used mutual information as similarity measure for rigid registration. In the further non-rigid transformation, the correlation coefficient objective function was used. To correctly evaluate their findings, for each of the 16 cases, a neuroradiologist chose 10 corresponding anatomic features across US volumes. The initial mean Euclidean distance of 3.3 mm was reduced to 2.7 mm with a rigid registration, and to 1.7 mm with a non-rigid registration. The quality of the alignment of the pre- and post-resection ultrasound image data was also visually assessed by a neurosurgeon. Afterward, an important contribution to neurosurgical US–US registration came by the release of the BITE dataset [109], in which pre- and post-resection US data are publicly available with relative landmarks to test registration methods. One of the first studies involving the BITE dataset came from [108]. The authors proposed an algorithm for non-rigid REgistration

of ultraSOUND images (RESOUND) that models the deformation with free-form cubic B-splines. Normalized cross-correlation was chosen as similarity metric, and for optimization, a stochastic descendent method was applied to its derivative. Furthermore, they proposed a method to discard non-corresponding regions between the pre- and post-resection ultrasound volumes. They were able to reduce the initial mTRE from 3.7 to 1.5 mm with a registration average time of 5 s. The same method has been then used in [127]. In a compositional method to register preoperative MRI to post-resection US data, they applied the RESOUND method to first register pre- and post-resection US images. In another solution [10], the authors aimed to improve the RESOUND algorithm. They proposed a symmetric deformation field and an efficient second-order minimization for a better convergence of the method. Moreover, outlier detection to discard non-corresponding regions between volumes is proposed. The BITE mean distance is reduced to 1.5 mm by this method. Recently, another method to register pre- and post-resection US volumes was proposed by [3]. The authors presented a landmark-based registration method for US–US registration in neurosurgery. Based on the results of the 3D SIFT algorithm [131], image features were found in image pairs and then used to estimate dense mapping through the images. The authors utilized several datasets to test the validity of this method. A private dataset of nine patients with different types of tumors was acquired, in which 10 anatomical landmarks were selected per case, in both pre- and post-resection volumes: For this set, they were able to reduce the mTRE from 3.25 mm to 1.54 mm. Then, they applied the same method to the BITE dataset and reduced the initial mean error to 1.52 mm. Moreover, they tested their approach on the more recent RESECT dataset [106]. By using the same method on the pre- and post-resection volumes, the mTRE was reduced from 3.55 to 1.49 mm.

Our solution proposes a segmentation-based registration approach to register US volumes acquired at different stages of neurosurgical procedures and compensate brain shift. A few approaches already applied segmentation methods on US data to then register MRI and iUS [132; 133]. Our solution represents the first segmentation-based method aimed at US–US volumes registration. Our approach includes a deep-learning-based method, which automatically segments anatomical structures in subsequent US acquisitions. We chose to segment the hyperechogenic structures of the sulci and falx cerebri, which remain visible during the resection and thus represent good corresponding elements for further registration. In the following step, parametric and nonparametric methods use the generated masks to register US volumes

acquired at different surgical stages. Our solution reduces the initial mTRE for US volumes acquired at subsequent acquisitions in both RESECT and BITE datasets.

7.2. Material and Methods

7.2.1. Datasets

We used two different public datasets to validate our segmentation-based registration method. Most of our experiments are conducted on the RESECT dataset [29], including clinical cases of low-grade gliomas (Grade II) acquired on adult patients between 2011 and 2016 at St. Olavs University Hospital, Norway. There is no selection bias, and the dataset includes tumors at various locations within the brain. For 17 patients, B-mode US-reconstructed volumes with good coverage of the resection site have been acquired. No blood clotting agent, which causes well-known artifacts, is used. US acquisitions are performed at three different phases of the procedure (before resection, during, and after resection), and different US probes have been utilized. This dataset is designed to test intra-modality registration of US volumes and two sets of landmarks are provided: one to validate the registration of volumes acquired before, during, and after resection, and another set that increases the number of landmarks between volumes obtained before and after resection. Regarding both sets, the reference landmarks are taken in the volumes acquired before resection and then utilized as references to select the corresponding landmarks in US volumes acquired during and after tumor removal. In the RESECT dataset, landmarks have been taken in the proximity of deep grooves and corners of sulci, convex points of gyri, and vanishing points of sulci. The number of landmarks of the first and second sets can be, respectively, found in the second column of Tables 7.3 and 7.4.

In addition, the BITE dataset is also utilized to test our registration framework [109]. It contains 14 US-reconstructed volumes of 14 different patients with an average age of 52 years old. The study includes four low-grade and ten high-grade gliomas, all supratentorial, with the majority in the frontal lobe (9/14). For 13 cases, acquisitions are obtained before and after tumor resection. Ten homologous landmarks are obtained per volume, and initial mTRE is provided. The quality of BITE acquisitions is lower with respect to RESECT dataset, mainly because a blood clotting agent is used, creating large artifacts [29].

7.2.2. Methods

We used MeVisLab [1] for implementing (a) an annotation tool for medical images, (b) a 3D segmentation method based on a convolutional neural network (CNN), and (c) a registration framework for three-dimensional data.

Manual Segmentation of Anatomical Structures

The first step of our method consists of the 3D segmentation of anatomical structures in different stages of US acquisitions. Both RESECT and BITE datasets are used to test registration algorithms and no ground truth is provided for validating segmentation methods. Therefore, we decided to conduct a manual annotation of the structures of interest in the US volumes acquired before resectioning of RESECT dataset. Pathological tissue was excluded from the manual annotation since it is progressively removed during resection and correspondences could not be found in volumes acquired at subsequent stages. On the contrary, we focused on other hyperechogenic (with an increased response—echo—during ultrasound examination) elements such as the sulci and falx cerebri. We consider these elements valid correspondences because the majority of them have a high chance of remaining visible in different stages of the procedure.

The manual segmentations were performed on a web-based annotation tool. As shown in Fig. 7.1, each RESECT volume can be simultaneously visualized on three different projection planes (axial, sagittal, and coronal). The segmentation task is accomplished by contouring each structure (yellow contour in the first frame of Fig. 7.1) of interest on the axial view. The drawn contours are then projected onto the other two views (blue overlay in the second frames of Fig. 7.1) so that a better understanding of the segmentation process is possible by observing the structures in different projections. The annotation process can be accomplished very easily and smoothly, and 3D interpolated volumes can be then obtained by rasterizing the drawn contours. As shown in Fig. 7.1, the contours are well defined in the axial view but several elements are not correctly included if considering the other two views. This is a common issue that we found in our annotation, which would require much time and effort to be corrected. However, we decided to have a maximum annotation time of 2 h per volume. The obtained masks correctly include the major structures of interest, but some elements such as minor sulci are missing. Despite the sparseness

[1] https://www.mevislab.de/mevislab/

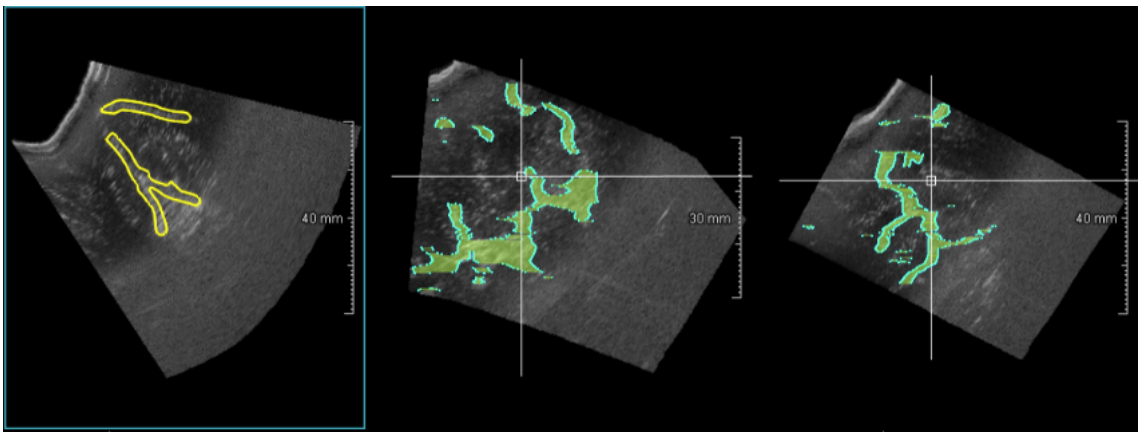


Figure 7.1.: Web-based annotation tool. While contouring the structures of interest on the axial view (yellow line in the left frame), the segmentation process can be followed in real-time on the other two views of US volumes. The annotation tool is accessible by common web browsers, and it has been used to obtain and then review the manual annotation

of our dataset, we expect our training set to be good enough to train our model to segment more refined structures of interest [42; 134].

The manual annotation was performed by the main author of this work (L.C.), who has two years of experience in medical imaging and almost one year in US imaging for neurosurgery. Then, a neurosurgeon with many years of experience in the use of US modality for tumor resection reviewed and rated the manual annotations, by taking into account the sparseness of the dataset. According to the defined criteria, each volume could be rated with a point between 4 and 1. More precisely, a point equal to 1 means that the main structures (falx cerebri and major sulci) are correctly segmented, and only minor changes should be made to exclude parts of no interest (i.e., slightly over-segmented elements). A point equal to 2 indicates that the main structures are correctly segmented, but major corrections should be made to exclude structures of no interest. A point equal to 3 indicates that main structures were missed in the manual annotations, which, however, are still acceptable. A score of 4 means that a lot of major structures are missing; therefore, annotation for the volume of interest cannot be accepted. The neurosurgeon evaluated the annotations by looking at the projected structures on the sagittal and coronal views of the drawn contours. Table 7.1 shows the results of the rating process for the volumes of interest.

Table 7.1.: Rating of the manual annotations

Volumes	**1**	**2**	**3**	**4**	**6**	**7**	**12**	**14**	**15**	**16**	**17**	**18**	**19**	**21**	**24**	**25**	**27**
Ranking	2	2	3	2	2	3	2	3	2	3	2	2	2	2	2	2	2

After the contours of the main structures of interest were manually drawn, the neurosurgeons rated them according to the criterion defined in the session "Manual segmentation of anatomical structures". The criterion is defined by taking into account the sparseness of the manual annotations. A point equal to 4 is given to the annotations where many of the main structures of interest are missing. On the contrary, if minor structures of interest (i.e., minor sulci) are missing but the major ones are correctly included, the best point of 1 is given

Segmentation

A convolutional neural network (CNN) aimed at volumetric segmentation is trained on manual annotations. More details about CNNs are provided in section 1.4.2. In particular, we utilized the original 3D U-net [42] architecture. A more detailed description of this original implementation is available 1.4.3. A few modifications were made to the original implementation: (a) The analysis and synthesis paths have two resolution steps and (b) before each convolution layer of the upscaling path a dropout with a value of 0.4 is used in order to prevent the network from overfitting. The training is conducted with a patch size of $30 \times 30 \times 30$, padding of $8 \times 8 \times 8$, and a batch size of 15 samples. The learning rate was set to 0.001, and the best model was saved according to the best Jaccard index computed on 75 samples every 100 iterations. The architecture modifications, as well as the training parameters, were chosen by conducting several experiments and selecting those providing the best results. As training, validation, and test sets, we split the seventeen volumes acquired before resection, which we annotated in the manual annotation. The split has been done as follows: The training set includes the volumes from 1 to 15, the validation one the volumes from 16 to 21, and the test one the volumes 24, 25, and 27.

After having found the best model to segment anatomical structures in pre-resection US volumes, we applied it to segment ultrasound volumes acquired at different surgical phases.

Registration

The masks automatically segmented by our trained model are used to register US volumes. The proposed method is a variational image registration approach based on [32]: The registration process can be seen as an iterative optimization algorithm where the search for the correct registration between two images corresponds to an optimization process aimed at finding a global minimum of an objective function. The minimization of the objective function is performed according to the discretize-then-optimize paradigm [32]: The discretization of the various parameters is followed by their optimization. The objective function to be minimized is composed of a distance measure, which quantifies the similarity between the deformed template image and the reference one, and a regularizer, which penalizes undesired transformations. In our approach, the binary 3D masks generated by the previous step are used as input for the registration task, which can be seen as a mono-modality intensity-based problem. Therefore, we chose the sum of squared differences (SSD) as a similarity measure, which is usually suggested to register images with similar intensity values. Moreover, to limit the possible transformations in the deformable step, we utilized the elastic regularizer, which is one of the most commonly used [32]. In our method, the choice of the optimal transformation parameters has been conducted by using the quasi-Newton l-BGFS [80], due to its speed and memory efficiency. The stopping criteria for the optimization process were empirically defined: the minimal progress, the minimal gradient, and the relative one, the minimum step length was set equal to 0.001, and the maximum number of iterations equal to 100.

Our registration method aims to provide a deformable solution to compensate for anatomical changes happening during tumor resection. As commonly suggested for methods involving non-rigid registration tasks [32], the proposed solution includes an initial parametric registration used then to initialize the nonparametric one. First of all, the parametric approach utilizes the information provided by the optical tracking systems as an initial guess. Based on this pre-registration, a two-step approach is conducted, including a translation followed then by a rigid transformation. In this stage, to speed up the optimization process, the images are registered at a resolution one-level coarser compared to the original one. Then, the information computed during the parametric registration is utilized as the initial condition for the nonparametric step. In this stage, to reduce the chance of reaching a local minimum, a multilevel technique is introduced: the images are registered at three different scales, from a third-level to one-level coarser. As the output of the registration step,

the deformed template image is provided.

7.3. Evaluation

7.3.1. Segmentation

In Table 7.1, we can see that no annotation received the best score of 1, but all of them have some imperfections. However, none of the manually annotated masks was scored with 4. Consequently, we can consider our annotations as a sparse ground truth in which only the main hyperechogenic structures of interest are included. Regarding this, CNNs trained on a sparse dataset already proved to be able to segment more refined and numerous structures with respect to the sparse training set [42; 134]. Therefore, we expect our annotations to be good enough to train the CNN model in order to generate meaningful structures for guiding the further registration step. In fact, the registration step will give important feedback about the quality of the generated masks: For our purposes, the segmented structures are meaningful if they correctly guide the registration method. In addition to this, an analysis of the segmentation results will be provided, as described in the following section.

Regarding US volumes acquired before resection, no ground truth is available for the structures not contained in the manual annotations. Consequently, the DICE coefficients are computed by including only the automatically segmented elements with correspondences to manual annotations and by discarding elements having no counterpart in manually annotated data. This measure is useful to verify whether the main structures of interest are correctly segmented by the trained model. As further information, we also provide the DICE coefficients computed without excluding any structure. These values would be useful for a deeper analysis of our algorithm but, as aforementioned, they may not be so indicative for our purposes due to the sparseness of our dataset. Furthermore, the automatically generated masks should also include more refined elements than the original ground truth. To verify this, a first visual assessment of the generated masks is performed. Moreover, the over-segmented elements are expected to have a mean intensity value as close as possible to one of the manually annotated structures. To verify this, we compared the mean intensity values of the manual annotations and the automatically generated masks.

Regarding US volumes acquired during and after resection, no manual annotation

was obtained, so no DICE index could be computed. Therefore, to be sure that structures of interest are correctly segmented, we show that the masks of the three stages of US data segmented by our trained model (a) are strongly correlated in terms of volume extension by computing the Pearson correlation coefficient and (b) include structures with a mean intensity value similar to the manual annotations. Secondly, we conduct a visual inspection of the results, which is helpful to verify whether or not corresponding anatomical structures are segmented in these stages.

Given the fact that our annotations are not publicly available, only a qualitative comparison is made with respect to other methods which also proposed a US segmentation solution in the context of neurosurgery [132; 133; 134; 135].

7.3.2. Registration

The transformations and deformation fields computed in the parametric and non-parametric steps are applied to the landmarks contained in the datasets. The TRE values before and after registration are provided per patient, with the measure of the closest and farthest couple of points, and mean TRE (mTRE) and standard deviation values are also given per each set of landmarks. For more details about the computation of the mTRE, please refer to Appendix III. A visual inspection of the registration results is also provided, in which the initial registration based on the information from the optical tracker can be compared with the results obtained by our method. Moreover, a comparison with previous solutions is provided. A few methods have been proposed to register BITE volumes [108; 127; 10; 3], but none of them except one [3] provided a generalized solution able to register volumes of both datasets (BITE and RESECT). On the contrary, our method provides an approach valid for both two datasets. For the RESECT dataset, the authors of [3] proposed a solution only for volumes acquired before and after resection. Our approach is the first one to be applied to the volumes acquired before and during the resection of RESECT dataset; therefore, no comparison is available for this specific set. The capture range of our method is also computed. We define the capture range as the largest initial misalignment within which our algorithm still converges to a solution for 80% of the cases. To evaluate it, we started the registration from multiple starting misalignments and we checked whether or not the method converged to a solution. Then, we computed the value of the capture range by using as distance measure the mTRE computed on the available landmarks.

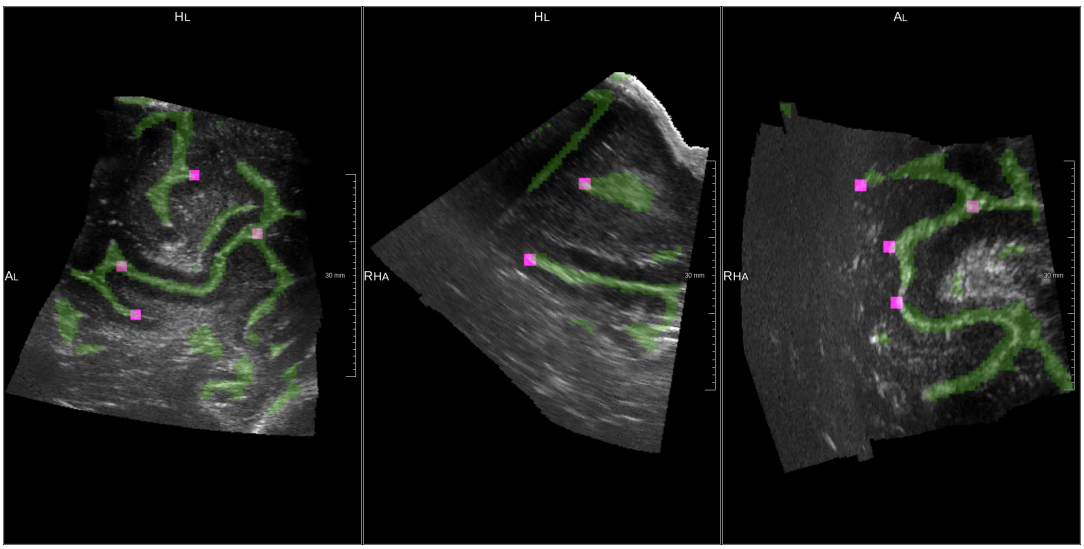


Figure 7.2.: Segmentation and landmarks. Original intensity volumes where the generated masks (in green) and RESECT landmarks (purple squares) are overlaid. In RESECT dataset, landmarks have been taken in the proximity to deep grooves and corners of sulci, convex points of gyri, and vanishing points of sulci. We chose to segment sulci and falx cerebri, and therefore, we can see how the landmarks are closely located to the segmented structures

7.4. Results

7.4.1. Segmentation

Figure 7.2 shows an example of a segmented structure in a volume acquired before resection. It can be seen that the generated masks cover the locations where landmarks were acquired. In fact, we decided to segment sulci and falx cerebri, which are the anatomical elements taken into account to acquire the majority of the landmarks in the RESECT dataset.

Regarding US data acquired before resection, Table 7.2 (a) provides the DICE coefficients computed between the manually segmented structures and the corresponding masks generated by our trained model. In Table 7.2 (b), the DICE coefficients for the whole set of generated masks (without excluding the elements not included in the manual annotation) are given. Furthermore, the first and third bars in Fig. 7.3 show that the structures automatically segmented in pre-resection volumes have a mean intensity value very similar to those chosen in the manual annotations. A similar consideration can be made for the elements considered as background (second

Table 7.2.: DICE coefficients for volumes acquired before surgery

Volumes	**1**	**2**	**3**	**4**	**6**	**7**	**12**	**14**	**15**	**16**	**17**	**18**	**19**	**21**	**24**	**25**	**27**
(a)																	
Dice %	68	62	57	76	71	56	78	76	78	61	62	70	70	63	74	68	69
(b)																	
Dice %	62	46	28	59	50	46	67	67	63	53	45	35	61	42	58	44	51

(a) Refers to the DICE coefficient computed by considering only the structures with a counterpart in the manual annotations. The method shows evidence of being able to properly segment the anatomical structures considered in the manual annotations. (b) Refers to the DICE values for the whole set of the automatically segmented structures

and fourth bars in Fig. 7.3). Qualitative results also confirm this evidence. Figure 7.4 shows four examples of automatically generated masks in comparison with the corresponding manual annotations. In most cases, our method correctly segments refined elements which were not included in the manual annotation due to timing restrictions (see "Manual segmentation of anatomical structures" in Section 7.2.2). Violet squares highlight some examples of these structures. Though, in several cases, the neural network wrongly segments pathological tissue which we excluded from the manual annotations (see blue squares in Fig. 7.4 (d)).

For the volumes acquired during and after resection, a strong correlation exists between the extension of their masks segmented by the neural network and of the volumes before resection. The Pearson coefficient between the masks of US data acquired before and during resection has a value of 0.90, and a value of 0.91 for those of pre- and post-removal. As for the US data acquired before resection, Fig. 7.3 shows that the anatomical structures segmented at the different stages have a mean intensity similar to the manual annotation (last four bars). Therefore, we can state that our segmentation method, applied to volumes acquired at different stages, segment structures related to each other in terms of volume extension and mean intensity values. Then, visual results in Figs. 7.5 and 7.6 confirm the evidence of the quantitative results, showing that our model trained on a stage of US correctly segments analogous elements in volumes acquired at different stages. However, qualitative results in Fig. 7.6 also show that our method often detects resection cavities, which have no corresponding structures in the pre-resection volumes.

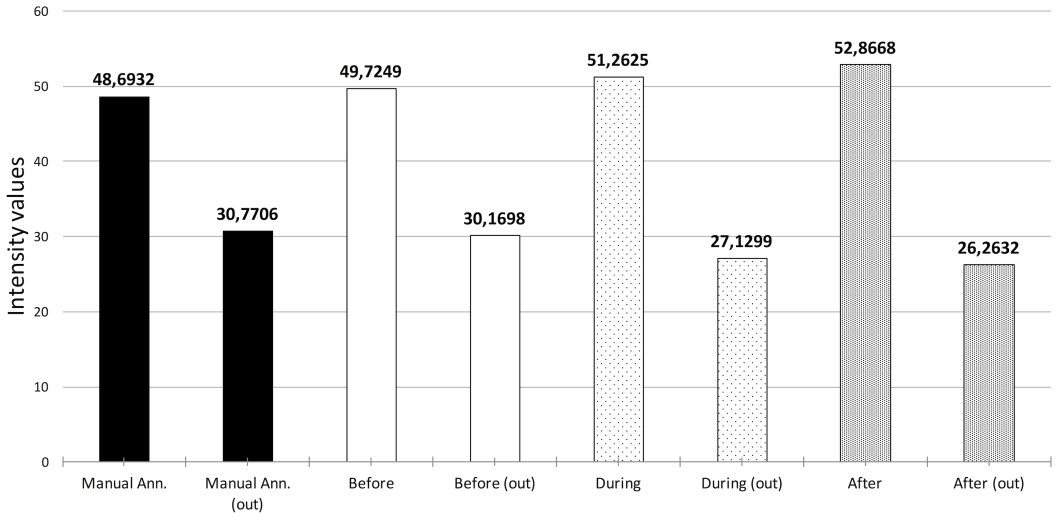


Figure 7.3.: Intensity values of the masked ultrasound volumes. This graph presents the mean intensity values of the masked ultrasound volumes (first, third, fifth, and seventh bars) acquired at the three stages, and the mean intensity values of the area excluded by the segmentation (second, fourth, sixth, eighth bars). For the volumes acquired before resection, volumes masked with manual annotation and elements segmented by the neural network are compared (first four bars). The masked volumes have in all the cases a similar mean value, higher than the excluded areas. This is meaningful since our elements of interest are the bright (hyperechogenic) structures in the US. On the contrary, the even-numbered bars have a similar mean intensity value, lower than the chosen structures. We are not interested in hypoechogenic structures, which look darker in the US acquisitions

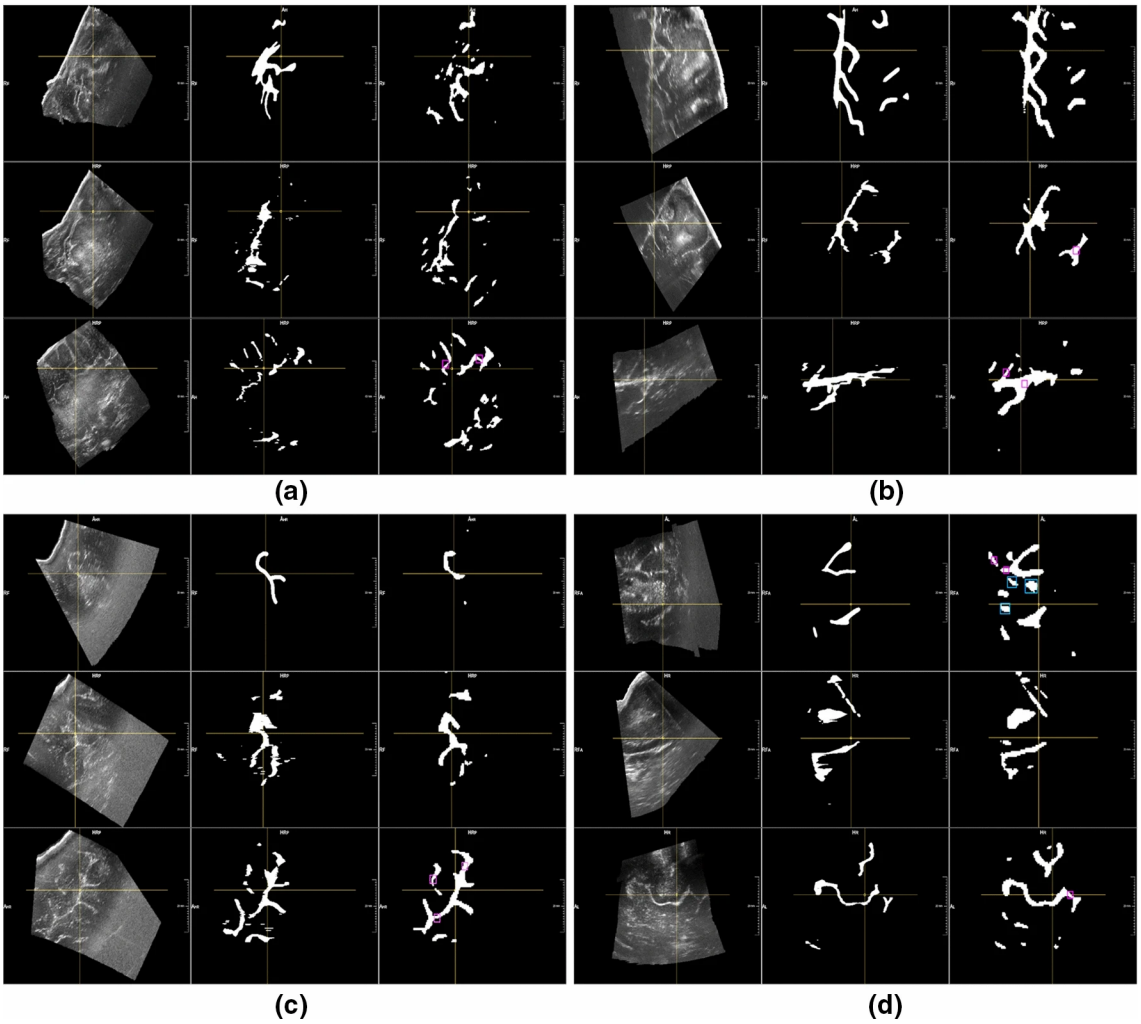


Figure 7.4.: Segmentation of ultrasound volumes acquired before resection. In each example, the axial, sagittal, and coronal views are shown in the first, second, and third rows, respectively. In the first column, the original ultrasound volume is exhibited, in the second column, the manual annotation performed on the axial view and projected in the other two views is shown, in the third column, the segmentation result obtained by the 3D U-net for the same volume of interest is displayed. In each example, a pointer (intersection of yellow crossing lines) highlights the same volume position in the three views. Our method correctly segments the main structures. Moreover, structures wrongly not included in the manual annotations are correctly detected by the trained neural network (purple squares). However, in image (d), pathological tissue correctly excluded in the original masks is wrongly segmented by our method (blue squares in axial view)

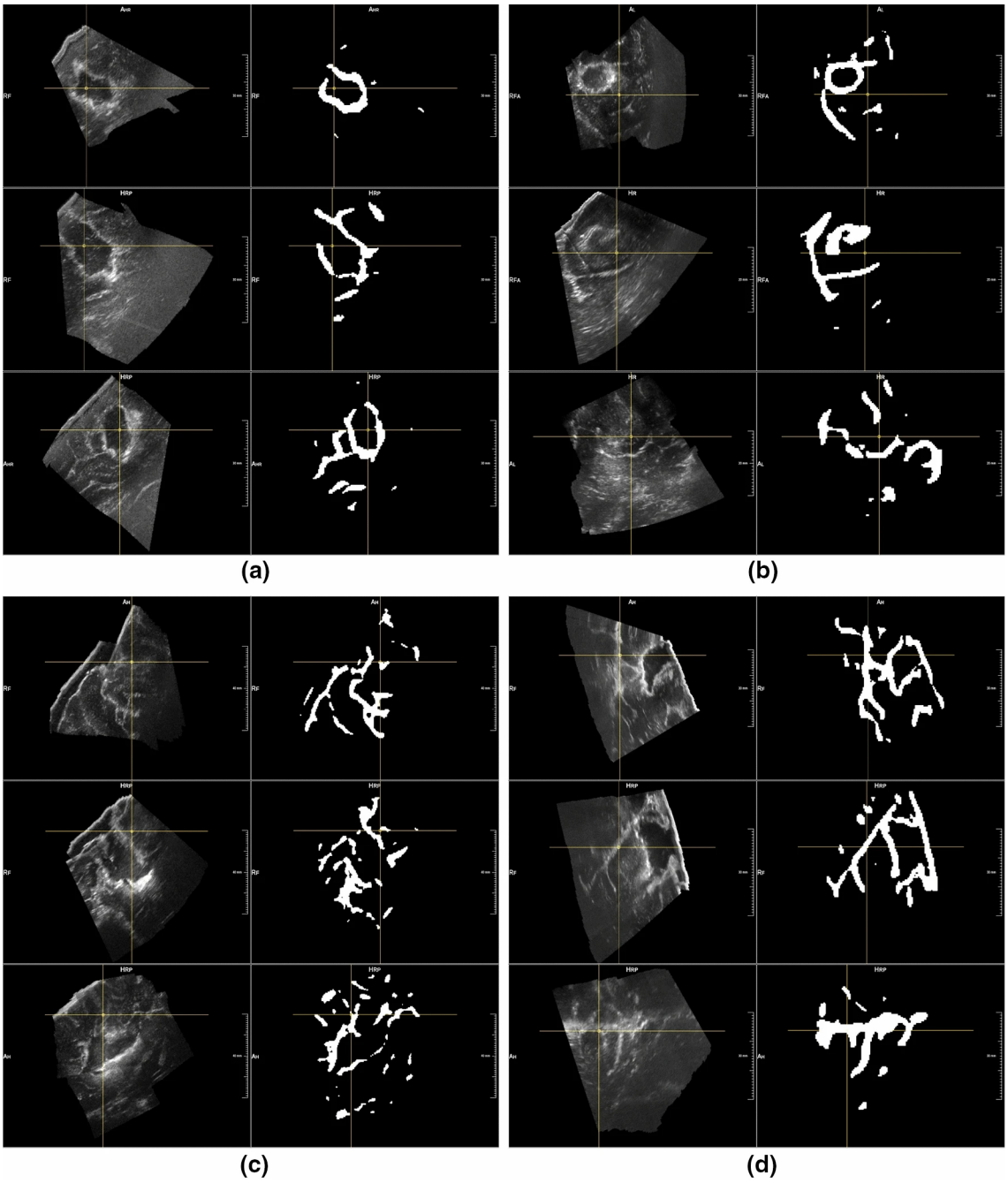


Figure 7.5.: Segmentation of ultrasound volumes acquired during resection. After having trained the neural network on the stage before resection, we applied it to ultrasound volumes acquired during resection. This figure shows four examples of segmentation results, each containing one intensity volume together with the generated mask. It appears clear how the main hyperechogenic structures are correctly included in the segmentation

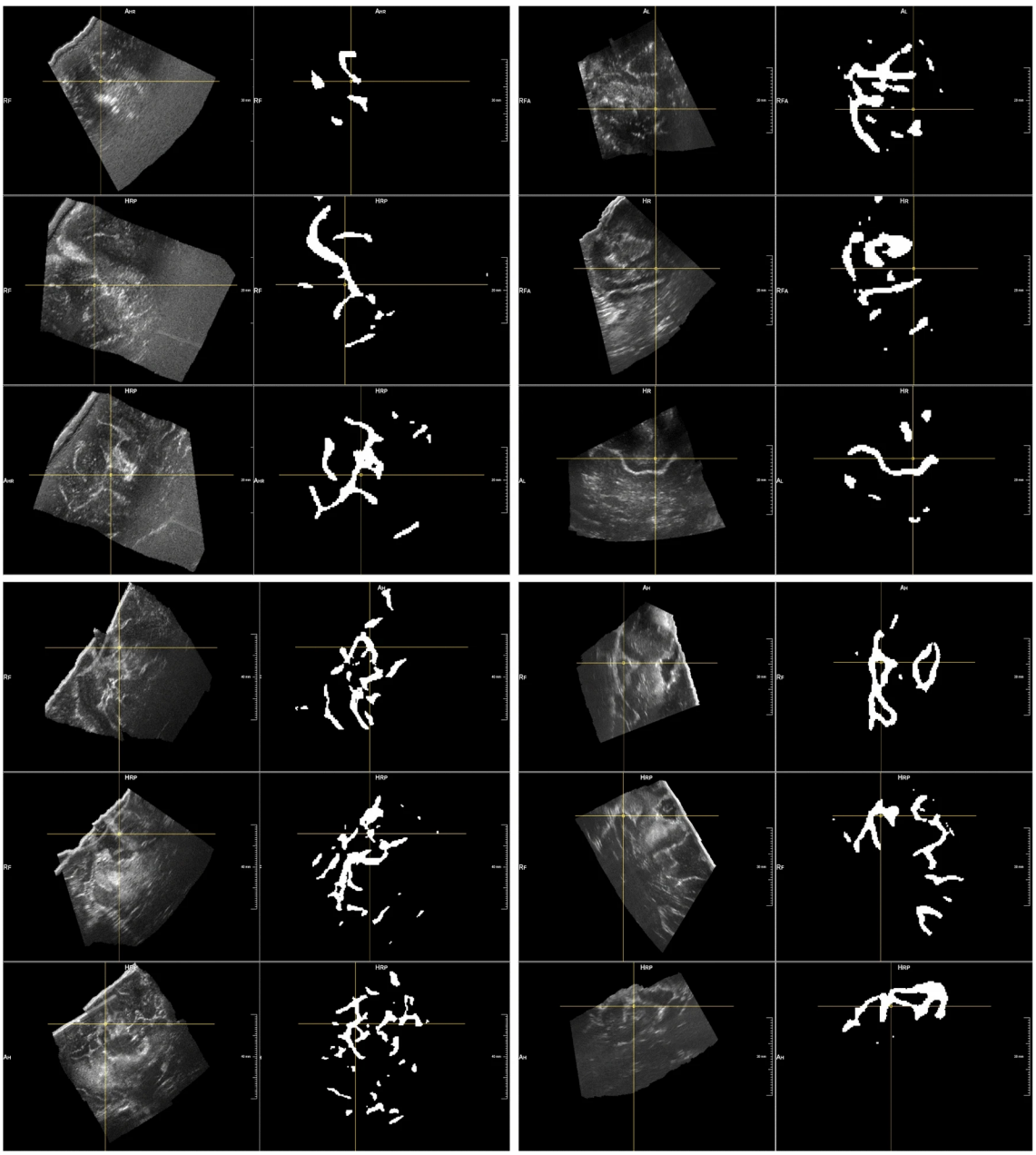


Figure 7.6.: Segmentation of ultrasound volumes acquired after resection. After having trained the neural network on the stage before resection, we applied it to ultrasound volumes acquired after resection. This figure shows four examples of segmentation results, each containing one intensity volume together with the generated mask. It appears clear how the main hyperechogenic structures are correctly included in the segmentation. In the last two examples (second row), we see how resection cavities (appearing hyperechogenic on US) are segmented by the 3D U-net, even if they have no counterparts in the pre-resection stage

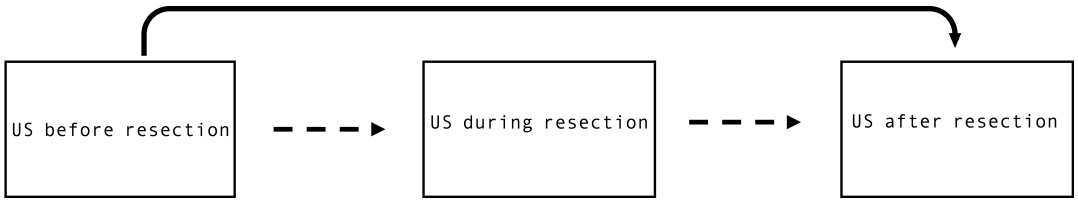


Figure 7.7.: Registration of different US volume pairs. Instead of registering directly pre-resection US data with those after resection (continuous line), a two-step method (dotted arrows) is proposed by including the US volumes acquired during resection. The final transformation from before to after resection volumes is obtained by concatenating two different registrations results (US before resectioning to US during resection + US during resection to US after resection)

7.4.2. Registration

The mean time required by the registration tasks is given in Table 7.6, together with the mean time required by each volume to be segmented by the trained model. All experiments are made on a computer equipped with an Intel Core i7 and a GeForce GTX 1080 (8 GB).

By relying on the automatically generated masks in the segmentation step, we registered the US volumes acquired at different surgical stages. First, the volumes acquired before and during resection are registered. Then, our algorithm is applied to volumes acquired during and after resection. The computed deformation fields are applied to the landmarks provided in the RESECT dataset, and the results after registration are shown in Table 7.3 (for volumes acquired before and during resection) and in Table 7.4 (volumes acquired before and after resection). Regarding the results in Table 7.4, the registration of the landmarks is performed by concatenating two different transformations: the one computed before–during US volumes together with the one for volumes acquired during and after resection (see Fig. 7.7 for a more detailed description).

As can be seen in both tables, both parametric and nonparametric methods reduce the initial mean registration errors provided in the RESECT dataset. In Table 7.3, it can be noticed that the proposed methodology improves the initial mTRE by more than 2 mm, by decreasing the mean errors for each patient. For the second registration task, our method reduces the mean registration error by nearly 1.5 mm (Table 7.4). Visual examples in Figs. 7.8 and 7.9 also confirm the numerical results.

Table 7.3.: Registration errors on RESECT dataset

		Mean distance (range) in mm before versus during (RESECT dataset)		
Patient	**Number of landmarks**	**Mean initial distance**	**Mean distance after parametric registration**	**Mean distance after nonparametric registration**
1	34	2.32 (1.49–3.29)	0.93 (0.31–1.73)	0.89 (0.22–1.57)
2	16	3.10 (1.79–5.19)	1.54 (0.37–3.58)	1.69 (0.71–4.19)
3	17	1.93 (0.67–3.02)	1.20 (0.36–2.47)	1.14 (0.24–2.45)
4	19	4.00 (3.03–5.22)	0.89 (0.31–1.86)	0.83 (0.24–1.65)
6	21	5.19 (2.60–7.18)	1.95 (0.63–3.71)	1.80 (0.58–3.61)
7	22	4.69 (0.94–8.16)	2.50 (1.24–5.78)	2.39 (1.15–5.86)
12	24	3.39 (1.74–4.81)	1.57 (0.43–3.20)	1.58 (0.44–3.36)
14	22	0.71 (0.42–1.59)	0.52 (0.09–1.15)	0.52 (0.12–0.93)
15	21	2.04 (0.85–2.84)	0.80 (0.28–1.44)	0.73 (0.18–1.31)
16	19	3.19 (1.22–4.53)	1.52 (0.95–2.21)	1.40 (0.75–2.43)
17	17	6.32 (4.65–8.07)	2.93 (1.67–4.46)	2.51 (1.14–4.03)
18	23	5.06 (1.55–7.44)	1.75 (0.70–3.04)	1.29 (0.46–2.81)
19	21	2.06 (0.42–3.40)	1.93 (0.20–3.19)	1.33 (0.48–2.67)
21	18	5.10 (3.37–5.94)	1.27 (0.19–3.53)	1.22 (0.19–3.46)
24	21	1.76 (1.16–2.65)	0.89 (0.18–2.17)	0.81 (0.08–2.07)
25	20	3.60 (2.19–5.02)	3.56 (2.09–5.14)	2.27 (1.04–3.92)
27	16	4.93 (3.61–7.01)	0.77 (0.24–1.35)	0.71 (0.19–1.26)
		Mean value ± SD	Mean value ± SD	Mean value ± SD
		3.49 ± 1.55	1.56 ± 0.82	1.36 ± 0.61

Mean registration errors between ultrasound volumes acquired before and during resection. Original distances are compared to the results obtained with our segmentation-based registration

Table 7.4.: Registration errors on RESECT dataset

		Mean distance (range) in mm before vs. after (RESECT Dataset)		
Patient	**Number of landmarks**	**Mean initial distance**	**Mean distance after parametric registration**	**Mean distance after nonparametric registration**
1	13	5.80 (3.62–7.22)	2.69 (0.93–4.08)	2.67 (0.75–4.18)
2	10	3.65 (1.71–6.72)	2.32 (0.90–4.25)	2.18 (0.55–3.93)
3	11	2.91 (1.53–4.30)	1.63 (0.82–2.48)	1.53 (0.82–2.25)
4	12	2.22 (1.25–2.94)	1.05 (0.46–1.95)	1.06 (0.30–2.05)
6	11	2.12 (0.75–3.82)	1.91 (0.47–3.06)	1.88 (0.24–2.93)
7	18	3.62 (1.19–5.93)	2.29 (0.92–4.13)	2.08 (0.70–3.93)
12	11	3.97 (2.58–6.35)	1.60 (0.54–4.73)	1.44 (0.61–4.51)
14	17	0.63 (0.17–1.76)	0.63 (0.11–1.84)	0.57 (0.09–1.52)
15	15	1.63 (0.62–2.69)	0.85 (0.12–2.13)	0.88 (0.23–2.38)
16	17	3.13 (0.82–5.41)	2.40 (0.61–4.70)	2.14 (0.79–4.35)
17	11	5.71 (4.25–8.03)	3.82 (2.36–6.68)	3.40 (1.91–6.28)
18	13	5.29 (2.94–9.26)	2.19 (1.14–4.32)	1.51 (0.65–2.92)
19	13	2.05 (0.43–3.24)	4.00 (1.42–14.27)	3.97 (0.91–15.29)
21	9	3.35 (2.34–5.64)	1.23 (0.29–3.20)	1.18 (0.28–3.16)
24	14	2.61 (1.96–3.41)	0.86 (0.18–2.26)	0.79 (0.13–2.02)
25	12	7.61 (6.40–10.25)	5.75 (4.39–8.34)	3.88 (2.74–6.07)
27	12	3.98 (3.09–4.82)	3.77 (2.22–5.10)	3.76 (2.24–5.30)
		Mean value ± SD	Mean value ± SD	Mean value ± SD
		3.54 ± 1.75	2.29 ± 1.37	2.05 ± 1.12
Other methods		**mTRE after registration**		
[3]		1.49 mm		

Mean registration errors between ultrasound volumes acquired before and after resection. Original distances are compared to the results obtained with our segmentation-based registration. Moreover, a comparison is made with a previous method proposed to solve this task

The images show the fixed volumes with the related segmentation (in red), together with the mask of the moving volumes (in green). By comparing the overlay before and after registration, we highlight the registration improvements by coloring in yellow the correct overlay of the two masks. Regarding the results on the RESECT dataset, only those obtained for volumes acquired before and after resection can be compared with another solution [3] (see Table 7.4).

Our segmentation-based registration method is then applied to the BITE dataset, directly registering volumes acquired before and after resection. The results are available in Table 7.5, with a comparison to previously proposed solutions (last section of Table 7.5). As it can be seen, also for this dataset the initial mTRE is reduced by both parametric and nonparametric registration approaches.

7.5. Discussion

The manual annotations, even if sparse, are good enough to train the CNN model to segment the anatomical structures of interest, as shown by the DICE coefficients in Table 7.2. Moreover, Fig. 7.4 shows that automatically generated segmentations are more precise than manual annotations, with better contour refinement and a larger number of identified structures. However, some pathological tissues are wrongly segmented by our method (see Fig. 7.4 (d)). This may be due to the fact that in US data the glioma of grade II appears as hyperechogenic structures, with an intensity similar to the elements of interest. In future work, we could consider separately segmenting pathological tissue and then exclude it during registration. A similar consideration can be made for the resection cavities in volumes acquired during and after resection, which appear as bright as sulci and are wrongly segmented by the proposed method (Fig. 7.6). Furthermore, from a qualitative comparison with other segmentation methods involving US data, we can highlight some advances in our approach. First of all, with respect to [132; 135], a higher number of anatomical structures are included in our manual annotations. Therefore, the potential range of clinical scenarios in which our method could be applied might be wider. Secondly, a trained neurosurgeon has clinically validated the manual annotations (Table 7.1). This is not the case for other segmentation-based methods [134; 133], in which no precise rating of the manual masks is provided.

The second important contribution of this work is the registration of US volumes acquired at different surgical stages. First of all, the segmentation method gives

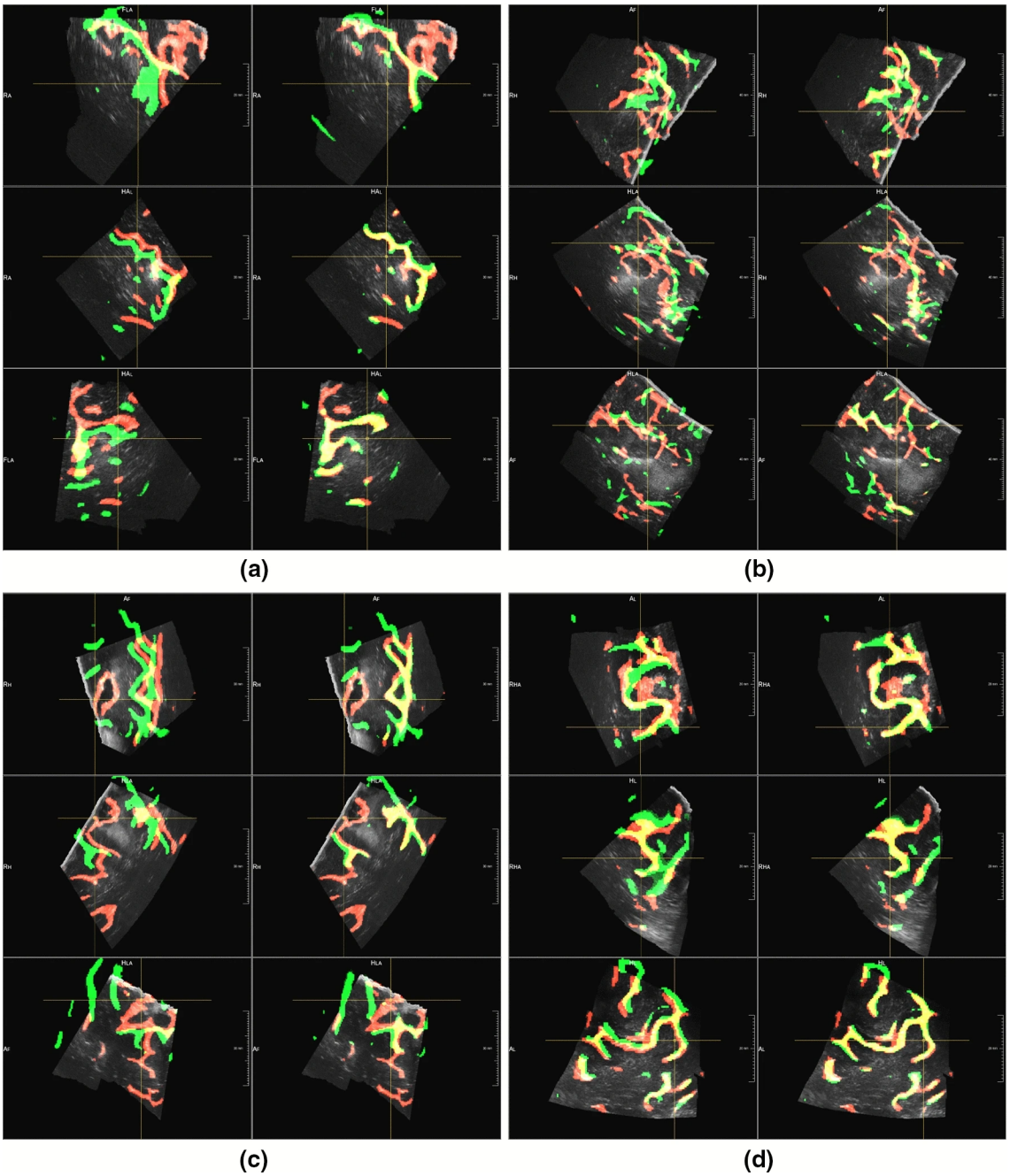


Figure 7.8.: Registration results for before and during resection volumes. The images show four examples of registration by combining fixed volumes (during resection) with its segmented structures (in red) and the segmented elements of moving volumes acquired before resection (in green). In the first column of each example, we show the segmentation overlay according to the original information. The second column displays the overlay of the segmented structures after registration. By highlighting in yellow the correct overlap of segmented structures, we can see how the structures are more aligned after the performed registration

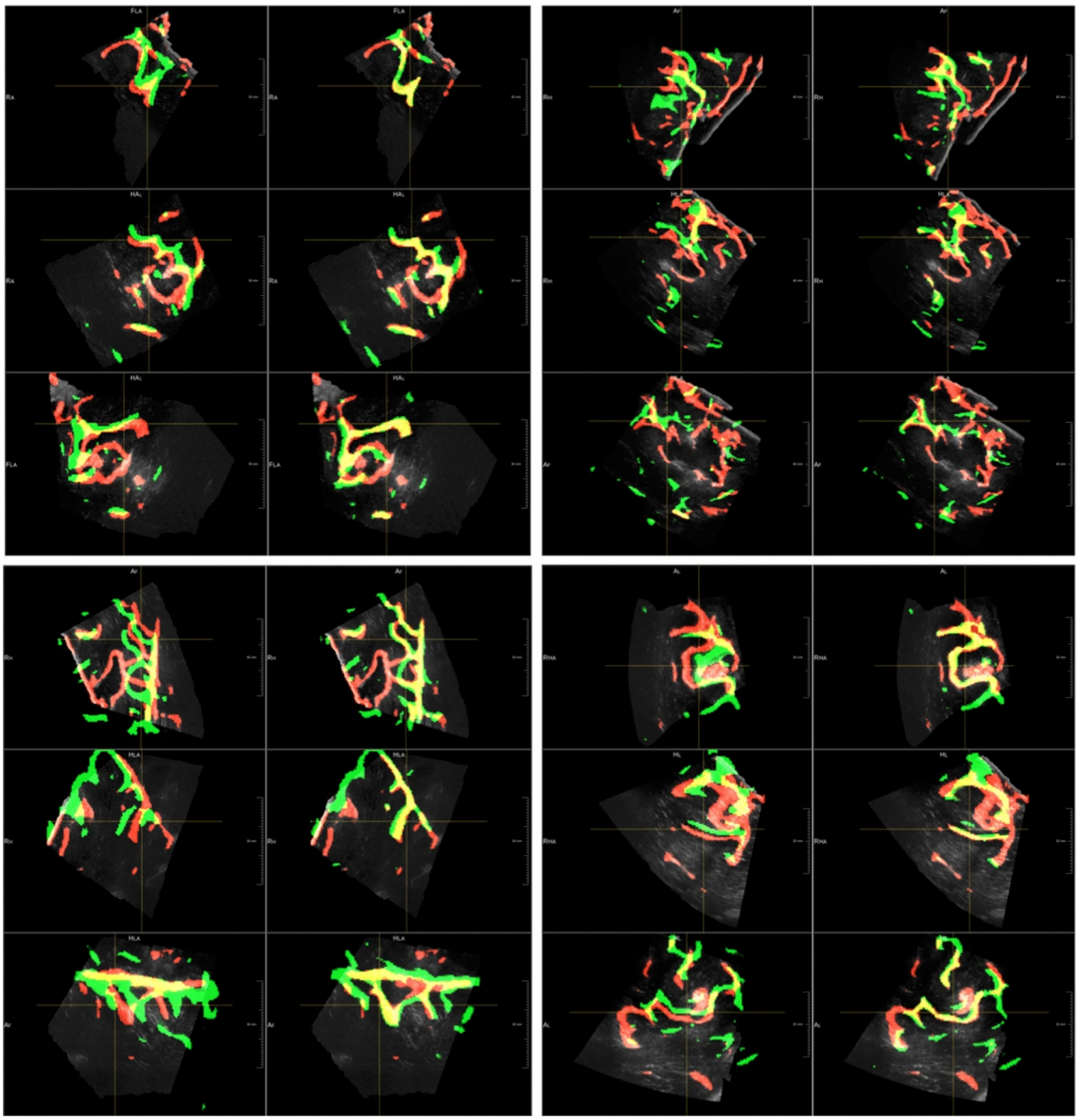


Figure 7.9.: Registration results for before and after resection volumes. These images show four examples of registration by combining fixed volumes (after resection) with its segmented structures (in red) and the segmented elements of moving volumes acquired before resection (in green). In the first columns of each example, we show the segmentation overlay according to the original information. The second column displays the overlay of the segmented structures after registration. By highlighting in yellow the correct overlap of segmented structures, we can see how the structures are more aligned after the performed registration

Table 7.5.: Registration errors on BITE dataset

Mean distance (range) in mm before vs. after (BITE dataset)			
Patient	**Mean initial distance**	**Mean distance after parametric registration**	**Mean distance after nonparametric registration**
2	2.30 (0.57–5.42)	1.97 (0.69–4.84)	1.70 (0.51–4.70)
3	3.40 (0.0–5.09)	3.12 (0.43–4.73)	1.45 (0.25–3.48)
4	4.60 (2.96–5.88)	3.62 (2.43–4.81)	2.20 (1.02–3.62)
5	4.11 (2.58–5.52)	3.68 (2.39–5.04)	1.32 (0.46–2.96)
6	2.26 (1.36–3.10)	1.33 (0.56–2.00)	1.17 (0.30–1.64)
7	3.87 (2.60–5.07)	1.27 (0.83–2.36)	1.17 (0.75–1.75)
8	2.51 (0.67–3.93)	2.11 (0.54–3.35)	2.18 (1.05–3.43)
9	2.21 (1.00–4.59)	1.99 (0.64–4.52)	1.95 (0.58–4.54)
10	3.86 (0.98–6.68)	3.81 (1.91–6.22)	3.43 (1.53–5.69)
11	2.74 (0.44–8.22)	2.74 (0.51–7.55)	2.39 (0.35–7.42)
12	10.54 (7.85–13.04)	10.88 (8.28–13.34)	11.08 (8.36–13.72)
13	1.62 (1.33–2.21)	0.73 (0.39–1.43)	0.75 (0.30–1.67)
14	2.19 (0.59–3.99)	1.60 (0.60–2.81)	1.43 (0.41–2.24)
	Mean value ± SD	Mean value ± SD	Mean value ± SD
	3.55 ± 2.28	2.98 ± 1.80	$2.48 \pm 2{,}67$
Other methods	**mTRE after registration**		
[108]	1.50 mm		
[127]	1.50 mm		
[10]	1.50 mm		
[3]	1.54 mm		

Mean registration errors between ultrasound volumes acquired before and after resection. In this dataset, a fixed number of markers (10 per case) is provided. Original distances are compared to the results obtained with our segmentation-based registration. Moreover, a comparison with previously proposed solutions (with corresponding mTREs) is provided

Table 7.6.: Mean time in seconds per task

Mean time (in s) per each task		
Segmentation (inference)	**Total registration time (US before–during)**	**Total registration time (US during–after)**
1.28	28.55	29.40

With segmentation, we indicate the inference process in which the trained model generates the mask of a volume given in input. The other two values are related to the registration tasks, including the time of both the parametric and nonparametric approaches

evidence of being able to generate meaningful masks to guide the registration task. In fact, the proposed registration method is able to reduce the mTREs of three sets of volumes from two different datasets (Table 7.3, 7.4, 7.5) by using the corresponding anatomical structures previously segmented. From numerical and visual results, we can notice that even if minor corresponding segmented elements are missing in volume pairs, our method is able to reduce the initial registration errors. However, in the case of volumes acquired after removal, resection cavities may be segmented by our method due to their intensity similar to the sulci. Consequently, the mTRE in Table 7.4 is reduced less with respect to Table 7.3, since these structures have no or few corresponding elements in volumes acquired in previous steps. This is a limiting factor of our registration method, which is completely based on the masks generated by our trained model. In future work, we could try to segment such structures and exclude them during the registration. Only another work [3] focused on the registration of US volumes acquired before and after the resection of RESECT dataset (Table 7.4). The mTRE obtained by the aforementioned approach is better than our method, which, however, is the first one to provide results for the volumes obtained before and during the resection of the RESECT dataset. In this set of volumes, our registration performs quite well, reducing the initial mTRE to 1.36 mm.

Regarding the BITE dataset, our algorithm improves the initial registration (see Table 7.5), proving not to be over-tuned on RESECT dataset. Note that in contrast to our approach, all other methods compared in Table 7.5 have only been tested on the BITE dataset. Thus, the results may be over-tuned on this limited set of volumes and the approaches could lack generalization. On the contrary, our solution is the second one after [3] to propose a more generalized method, which has been tested on registering the volumes of both RESECT and BITE datasets. Therefore, our method is validated on a larger number of US acquisitions, providing a more

generalized solution. Nevertheless, there might be some reasons why a few other approaches have smaller average mTREs for the BITE dataset (last section of Table 7.5). First of all, a numerical impacting factor for our results comes from case 12, where the TRE increases from 10.54 up to 11.08 mm, affecting the overall result. The capture range of our method is too low to register this volume pair, which has a very large initial misalignment. In future work, we could improve the results by performing an initial registration which could increase the capture range of our method. Moreover, the limited improvement obtained by our method might be due to the lower quality of the BITE dataset compared to the RESECT volumes, which are used for training the segmentation approach. Since our registration method is based on the generated masks, it is almost impossible for the registration method to converge to the right solution if the segmented masks are not accurate enough.

The total time required by each task of our method is visible in Table 7.6: The segmentation step requires 1.28 s and 28.55 s (before/during) and 29.40 s (during/after) that are needed to register the generated 3D masks. In addition to this, we should also take into account the time to reconstruct 3D US volumes from 2D images, which is a few seconds [106]. Considering the increase in the brain shift over time and the average duration of a neurosurgical procedure [136], our algorithm is fast enough to register US volumes and therefore provides a meaningful solution for brain shift. Nevertheless, in future work, we could optimize our algorithm to speed up the registration step.

7.6. Conclusion

To the best of our knowledge, our solution is the first one to propose a segmentation-based registration method that registers US volumes acquired at different surgical stages. Our approach provides some important contributions. Regarding the segmentation step, a model based on a 3D U-Net has been trained on a large number of anatomical structures, whose manual annotations have been validated by an experienced neurosurgeon. Even if the training is performed on a sparse set of annotations, the proposed solution can automatically segment hyperechogenic elements in US volumes. Moreover, the segmented anatomical structures prove to be meaningful elements that can guide the registration of US volumes acquired in the neurosurgical context. In fact, for two different datasets of US volumes acquired at different surgical stages, the initial mTREs are correctly reduced, demonstrating

that our solution is not over-tuned for a specific dataset. Moreover, our work is the first one to be applied also on the US volumes of RESECT dataset acquired during resection, for which no previous work has been published.

Acknowledgements

This work was funded by the H2020 Marie-Curie ITN TRABIT (765148) project.

8. Enhanced Registration of Ultrasound Volumes by Segmentation of Resection Cavity in Neurosurgical Procedures

The material of this chapter is available as an open-access journal paper published in the International Journal of Computer Assisted Radiology and Surgery.

Canalini, L., Klein, J., Miller, D., Kikinis, R. (2020). Enhanced registration of ultrasound volumes by segmentation of resection cavity in neurosurgical procedures. Int. J. CARS, 15(12), 1963–1974. doi: 10.1007/s11548-020-02273-1

Abstract

Purpose Neurosurgeons can have a better understanding of surgical procedures by comparing ultrasound images obtained at different phases of the tumor resection. However, establishing a direct mapping between subsequent acquisitions is challenging due to the anatomical changes happening during surgery. We propose here a method to improve the registration of ultrasound volumes, by excluding the resection cavity from the registration process.

Methods The first step of our approach includes the automatic segmentation of the resection cavities in ultrasound volumes, acquired during and after resection. We used a convolution neural network inspired by the 3D U-Net. Then, subsequent ultrasound volumes are registered by excluding the contribution of the resection cavity.

Results Regarding the segmentation of the resection cavity, the proposed method achieved a mean DICE index of 0.84 on 27 volumes. Concerning the registration of the subsequent ultrasound acquisitions, we reduced the mTRE of the volumes acquired before and during resection from 3.49 to 1.22 mm. For the set of volumes acquired before and after removal, the mTRE improved from 3.55 to 1.21 mm.

Conclusion We proposed an innovative registration algorithm to compensate the brain shift affecting ultrasound volumes obtained at subsequent phases of neurosurgical procedures. To the best of our knowledge, our method is the first to exclude automatically segmented resection cavities in the registration of ultrasound volumes in neurosurgery.

8.1. Introduction

In the neurosurgical planning for tumor resection, preoperative magnetic resonance imaging (MRI) data are usually acquired [29; 137]. Subsequently, neuronavigation systems are utilized to make these images available during the tumor resection. Through a rigid transformation computed between the surgical scene and the MRI data, neurosurgeons are able to map any intracranial position to the preoperative data. This is beneficial for the surgery outcome since it decreases the risk of tumor residuals and increases the survival rate of the operated patients. However, image-guided surgery based only on preoperative data has some limitations [23]. In the early stages of the procedure, inaccuracies in the neurosurgical setting can degrade the rigid registration computed by the neuronavigation systems. Moreover, during the ongoing procedure, several anatomical modifications take place and the observed surgical scene modifies compared to the preoperative data. In the early stages of the surgery, the opening of the dura mater is responsible for the leakage of cerebrospinal fluid that heavily modifies the brain structure. Additionally, the resection of the tumor leads to other anatomical modifications, with no counterpart in the preoperative data. All these effects combined together are denoted as brain shift [23]. This phenomenon impedes a correct mapping between preoperative data and surgical scene. Consequently, the probability of missing pathological tissue in the resection increases, reducing the survival rates of the operated patients [138; 139].

To compensate for the brain shift, intra-operative images can be acquired to provide an update of the resection scene [140]. The most common intraoperative solutions are given by MRI and ultrasound (US) modality. Intraoperative MRI (iMRI) data give a good image contrast between healthy and pathological tissues [141; 120], but it has the disadvantages of being expensive, requiring special adaptation in the operating room (OR), and being relatively long to be acquired. A valid alternative is given by intraoperative US (iUS), which is inexpensive, fast, and practical to obtain [123; 142]. However, the understanding of the US data can be challenging [30], in particular, if compared to the image quality obtained by MRI modality. To overcome this problem, neuronavigation systems can provide a direct mapping between preoperative MRI and iUS. By observing the same structures in two different imaging modalities, a better understanding of the tissues is also possible.

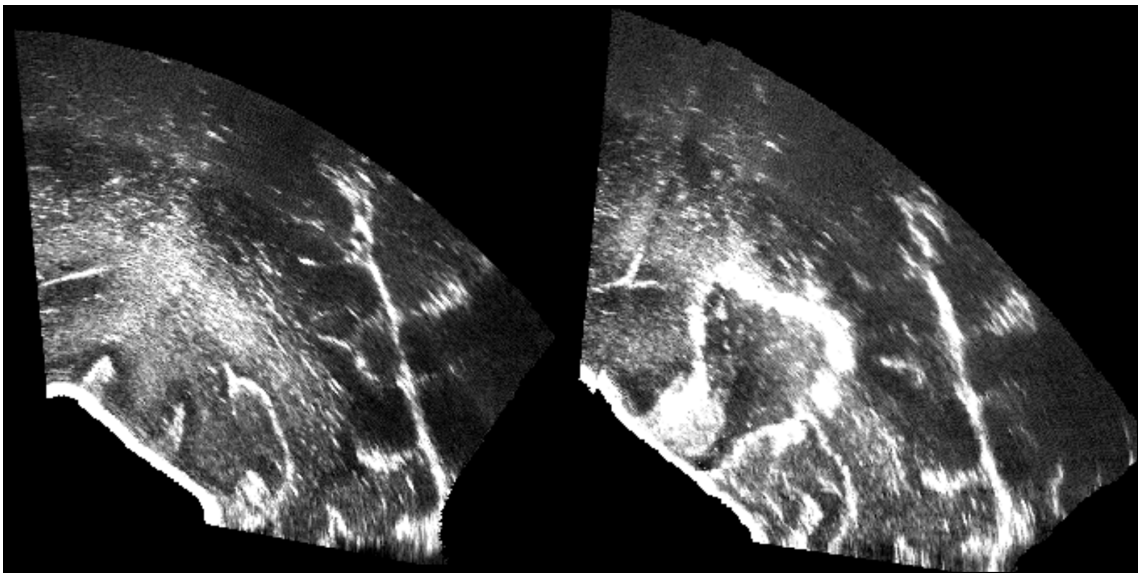
Additionally, a source of artifacts negatively affecting iUS is related to the resection cavity (RC), which appears in the ongoing procedure [30]. To perform US acquisitions after a first resection, a saline water solution is used to fill the operative cavity. When the US probe is used, the attenuation of sound waves in tissue is higher than in the saline water solution used for coupling. Consequently, hyperechoic artifacts appear, especially at the bottom of the resection cavity. They negatively affect the interpretation of the images, since they can be wrongly seen as pathological tissue. Therefore, toward the end of the resection, it becomes extremely important to identify these artifacts. As a solution, US volumes can be obtained at different time points of the resection, without a delay in the surgical procedure [137]. By tracking the US probe, neuronavigation systems compute a registration of the US volumes acquired at different phases of the resection. Then, the US data obtained at the end of the surgery can be compared with the early-stage acquisitions [30], in which the artifacts related to the resection did not appear yet. Thus, image interpretation becomes easier. Nonetheless, the direct comparison between subsequent phases is challenging due to the brain shift, which can only be compensated by a non-rigid solution [29; 143]. The registration provided by the neuronavigation systems is often not accurate, since it does not take into account anatomical modifications [29]. Therefore, this task is an open issue and a solution is still needed.

The registration of US volumes acquired at different resection stages is challenging since the brain undergoes anatomical modifications, such as the emerging resection cavity, which have no counterpart in the data acquired at the beginning of the surgery (Fig. 8.1). A few registration solutions, which take into account the missing correspondences between the different acquisitions, have been already proposed. In [108], the authors suggested a non-rigid registration algorithm that models the deformation field with free-form cubic B-splines. In the cost function, the similarity metric is based on the normalized cross-correlation (NCC). They also introduced a method to suppress the effect of non-corresponding regions between the pre- and post-resection ultrasound volumes. The outlier detection is based on the standard deviation of the NCC gradients. The same approach has been applied in [127]. The authors in [10] advised an improvement compared to [108]. They proposed a symmetric deformation field and an efficient second-order minimization for a better convergence of the method. Moreover, an outlier detection to discard non-corresponding regions between volumes is proposed. Their approach starts from the one applied in [108] and adds feature to improve the accuracy of outlier detection.

Another solution considering also the resection cavity has been proposed by [144] to tackle the registration of pre- and intra-operative MRI images. Their framework is based on the Demons algorithm using an anisotropic diffusion smoother. The resected tissues in intra-operative data are detected with a level set method and then integrated into the smoother as a diffusion sink.

Furthermore, the effect of the outliers can be reduced by feature-based methods, limiting the registration only to corresponding elements. In [3], the authors proposed a feature-based registration approach where corresponding image features are computed in US pairs. Afterward, they used them to estimate dense mapping through the images. The authors utilized several datasets to test the validity of their approach. Besides, the authors in [79] proposed a segmentation-based method. In the first step of their approach, they introduced an automatic segmentation of corresponding elements (sulci and falx cerebri) in subsequent US acquisitions. Then, their registration solution is only based on the automatically generated masks. In this way, they discarded the non-corresponding elements by focusing on structures available in the different acquisitions. A similar approach has been proposed in [145], where the authors applied a distance transform on the generated segmentation of the sulci and falx cerebri. Then, the transformed masks were registered.

We here propose a solution to register subsequent US volumes by discarding the non-corresponding elements from the registration process. In our approach, a 3D convolution neural network is utilized to segment the cavities generated with the tumor removal. Then, in the following step, the generated masks are used to discard the contribution of the resection cavity from the registration process. We expect the proposed registration approach to achieve better results than the same method not excluding the resection cavity. Regarding the segmentation step, to the best of our knowledge only the authors in [146; 147] proposed a solution for this task. In [147], they described a method based on a 2D U-Net to segment the resection cavity in US volumes. Besides, in [146] they also demonstrated that the 3D architecture achieves better results than a 2D approach. The mean time required from their 3D neural network to process a volume is around 5 min, making the application of their method not feasible in a clinical scenario. For both their methods, the authors used a private ground truth to train their neural network.



(a) (b)

Figure 8.1.: Ultrasound images acquired before (a) and after (b) resection. During the ongoing resection, the tumor is gradually removed and the cavity of the resection appears. The resection cavity is usually filled with saline water, which appears usually hypoechogenic in US acquisitions. The extension of the resection is usually recognizable by a hyperechogenic contour. By comparing images a, b, we can notice that other anatomical elements remain visible, even if deformed due to the brain shift effects. On the contrary, the resection cavity has no specific counterpart in the initial acquisitions. In the process of registering the two images, it would be beneficial to exclude the non-corresponding elements of the resection cavity and rely only on the structures that remain visible among subsequent acquisitions

8.2. Material and Methods

8.2.1. Datasets

We utilized two publicly available datasets containing US acquisitions obtained at different stages of tumor resection [29; 143]. In both datasets, manually annotated landmarks are given for testing registration algorithms. The RESECT dataset [29] includes US volumes acquired at three different stages of the resection. As mentioned in the original publication, we indicate with before resection the acquisition at the beginning of the neurosurgery, when the dura mater is typically intact. After an initial resection is performed, a further acquisition is performed, and we refer to it as during resection. At the end of the resection, neurosurgeons verify if any pathological tissue is left, and the performed acquisitions are referred to as after removal. To have more details about the initial mean target registration error (mTRE) and the number of landmarks per each pair of volumes, please refer to Tables 8.2 and 8.3. The BITE dataset [143] was released before the RESECT one and contains volumes acquired before and after resection. Ten landmarks are provided per each pair of volumes, and initial mTRE is provided in Table 8.4. The quality of the images of the BITE dataset is lower compared to the more recent RESECT dataset, as observed by [29]. Moreover, as observed by [146], the acquisition protocols of this dataset differ from the one in the RESECT data (more details in Fig. 8.7).

8.2.2. Manual Annotations

No ground truth for resection cavity segmentation is provided in the aforementioned datasets. Thus, we decided to manually annotate 27 volumes of the RESECT dataset, acquired during and after resection. To manually annotate the resection cavities in the US volumes, we utilized MEVIS Draw [1] (see Fig. 8.2). With this tool, the volume of interest can be visualized in three main projections and the user can choose the more appropriate one for performing the manual segmentation (Fig. 8.2a). If the annotation is executed on non-adjacent slices, an interpolation automatically fills in the contour on slices not yet annotated. This reduces the time needed for the annotation procedure, making 3D segmentation very efficient. Furthermore, the user has always the chance to observe the manual annotation in three

[1] https://www.mevis.fraunhofer.de/en/research-and-technologies/image-and-data-analysis/mevis-draw.html

Table 8.1.: DICE coefficient for the segmentation of the resection cavity

Training set																
Volume	**2 a**	**2 d**	**3 a**	**4 a**	**7 a**	**7 d**	**12 d**	**15 a**	**16 a**	**16 d**	**17 a**	**18 a**	**19 a**	**21 a**	**24 a**	**25 d**	**27 a**
DICE	0.91	0.91	0.85	0.88	0.89	0.7	0.88	0.82	0.88	0.82	0.91	0.90	0.95	0.95	0.86	0.87	0.92

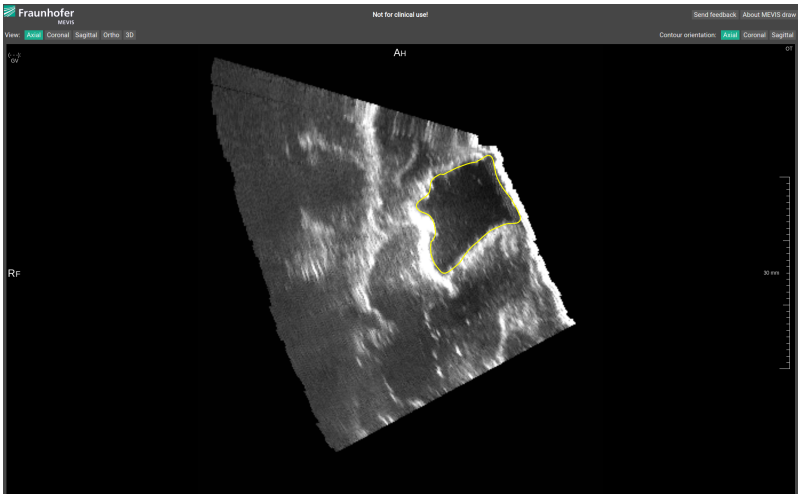
	Validation set					Test set				
Volume	**1 a**	**14 a**	**17 d**	**19 d**	**21 d**	**6 a**	**6 d**	**12 a**	**18 d**	**25 a**
DICE	0.76	0.88	0.75	0.75	0.71	0.88	0.88	0.88	0.26	0.86

The first part of the table is related to the training set, whereas the second one is for the validation and test sets. The second row of each table indicates the RESECT US volumes of interest: Each volume is indicated by a number, specifying its related case in the dataset, followed by a letter. The letter indicates if the volume of interest is related to the acquisition performed during (d) or after (a) resection. For example, 4 a is used for the volume belonging to Case 4 acquired after resection. The third row indicates the computed DICE indices

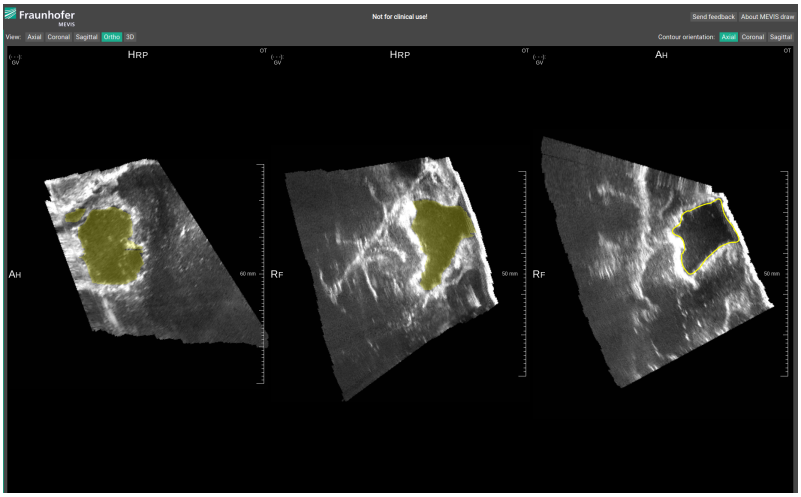
different views (b), in which the drawn contours are projected. In case a modification is needed, the user can modify the manual annotation in any of the three views. Then, the interpolation is processed again. Two raters (L.C. and D.M.) annotated the resection cavities in the volumes of RESECT specified in Table 8.1 (Fig. 8.3). The author L.C. has two years of experience with ultrasound data, and the co-author D.M. is a neurosurgeon with a long experience in the use of the US for tumor resection [79]. The masks generated by the intersection of the two manual segmentations are available at the following link.

8.2.3. Automatic Segmentation

We used the clinically revised masks as a ground truth to train a convolution neural network (CNN). CNNs are a class of artificial neural networks that automatically and adaptively learn spatial hierarchies of features from images in order to solve specific tasks (classification, segmentation, etc.) [148]. CNN is generally composed of three types of blocks: convolution, pooling, and fully connected layers. In a training phase, the generated outputs are compared with the ground truth, and a loss function is defined to measure the performance of the CNN. Thanks to a back-propagation algorithm, the trainable parameters are iteratively changed to minimize the loss function and reduce the difference between generated outputs and ground truth. A stochastic gradient descent optimization algorithm is usually used. More details about CNNs are provided in section 1.4.2. For the segmentation task, we utilized a neural network whose architecture is based on the 3D U-net [42]. It consists of an analysis path to capture anatomical context and a symmetric synthesis



(a)



(b)

Figure 8.2.: Annotation tool. The contour of the structure of interest is drawn on a specific view (a). After having drawn contours on a limited number of images, an interpolation automatically computes the contours on the remaining slices. In the process of segmenting the element of interest on a specific projection (a), the corresponding contour can be visualized in the other two views (b)

(a) (b) (c)

Figure 8.3.: Manual segmentation. Two raters annotated the masks of interest. As an example, (a) shows the image of an original intensity volume. In (b) (first rater) and (c) (second rater), the two different versions of the manual segmentation of the corresponding resection cavity are provided

path to precisely locate the structures of interest. An image is given in input to the contracting (analysis) path. Here, each layer has two convolutions followed by an activation function and a max pooling: Feature maps are extracted from the image passing through the first part of the network. Then, in the expanding (synthesis) path, the image is upsized to the original size. Each layer consists of an up convolution followed by two convolutions and an activation function. In the synthesis path, at each level, the upsized image is combined with the corresponding high-resolution features extracted in the contracting path. In the output of the neural network, a volumetric mask is generated and an error between the generated mask and ground truth is computed. More details are available in section 1.4.3.

With respect to the original implementation, some modifications were empirically made: The analysis and synthesis paths have three resolution steps. Before each convolution layer of the upscaling path, a dropout with a value of 0.3 is used to prevent the network from overfitting. Additionally, we utilized the Tversky loss function [149] which penalizes the false negatives and thus advantages the over-segmentation of the foreground structures. We decided to train with this loss function because it has been shown to achieve good results in the case of an unbalanced dataset. The training of the CNN is conducted with patches of size $48 \times 48 \times 48$, a padding of $20 \times 20 \times 20$, and a batch size of 5 samples. The learning rate was set to 0.0005. Additionally, we divided the annotated volumes into three groups (more details in Table 8.1). The first set (training set) includes 17 volumes (60 percent of the total

volumes) and is used to train the neural network. The second one is disjoint from the training set and includes 5 volumes (20 percent of the dataset): It is utilized to validate the trained algorithm on volumes different from the one used for training. Then, after the best model has been saved, the network is applied on a test set (5 volumes) not used during validation and training. To evaluate the effects of differences in the segmentation on the registration results, two 3D U-Net models are trained, each using the annotation manually segmented by a different rater.

As commonly happens in medical images, in our dataset the positive voxels of the foreground are in minor quantity than the background ones, and this may slow the learning process of the neural network. To speed up the training, we decided to alter the composition of the patches used to train the neural network. Thus, during the training and validation processes, we decided to feed the network with 20 percent of patches including only the background. The rest of the 80 percent includes patches containing at least a pixel with the foreground label. In our experiments, the Tversky loss had often the effect of producing an over-segmentation of structures of interest. The resection cavity is usually a closed isolated structure, with no dispersed elements. Therefore, we applied a connected component analysis on the automatically generated masks, to keep only the biggest segmented structure, corresponding to the resection cavity.

8.2.4. Registration

In the proposed solution, the volume acquired before resection represents the template (moving) image which is deformed to match the reference (fixed) one, respectively, the data acquired during or after resection. The proposed solution is based on [32], where the registration of two volumes is treated as an iterative optimization algorithm. In this scenario, the correct registration of the two images corresponds to the global minimum of a discretized objective function: This includes a distance measure, determining the similarity between the deformed template and the reference image, and a regularizer, which penalizes undesired transformations. In addition to the moving and fixed images, the proposed method uses an additional input (mask). The parts excluded from the masks are usually those not in common in both acquisitions. Thus, the contribution to the distance measure is limited to the areas for which the segmentation is available.

In our solution, we use the quasi-Newton l-BGFS [80] to guide the optimization process and the normalized gradient field (NGF) as the distance measure. The

stopping criteria for the optimization process are empirically decided: The minimal progress, the minimal gradient, and the relative one, the minimum step length was set equal to 0.001, and the maximum number of iterations is equal to 100. The algorithm takes in input two original intensity volumes. Besides, the automatically generated masks are provided as a third input. The area of the resection cavity is excluded from the computation of the distance measure: This represents the only element, not in common between the two input images, and its exclusion may improve the registration output. Moreover, as suggested by [128; 130], the background outside the US beam is also discarded. The proposed approach is initialized by a rigid registration, followed by a deformable approach. In the first step, the volumes are directly registered at one fixed level coarser than the original scale, in order to improve the speed of the algorithm. A transformation matrix is obtained and then utilized for initializing the deformable registration. In the second step, we utilized the curvature regularizer to limit the range of possible transformations [32]. To avoid local minima in the optimization and to speed the computation, the volumes are registered from a third coarse level, in which computations are cheap, until the fine level. Besides, to estimate the effects of the inter-variation in the segmentation on the registration results, we test the method with the two versions of the masks generated by the two 3D U-Net models. We would like to check how much the differences in the masks to be excluded affect the registration results.

In our experiments, we also tested a solution not excluding the resection cavity from the registration.

8.3. Results

8.3.1. Segmentation

Table 8.1 shows the DICE coefficients for the segmentation of the resection cavity for the 3D U-Net model obtaining the best mean value. Overall, the mean DICE coefficient is 0.84. Visual results are available in Figs. 8.4 and 8.5, respectively, for case 2 acquired during resection and for case 27 acquired after resection. Each figure shows the related volume in three projections. To maintain a good tradeoff between the visibility of the surrounding anatomical structures and visualization of the mask of the resection cavity, we decided to highlight the element of interest with a border in green (ground truth) and purple (automatic segmentation). Figure 8.6

shows the segmentation results for the volume 18 acquired during resection (18 d in Table 8.1), for which the worst DICE coefficient (0.32) has been obtained. In the figure, we show the overlay between ground truth and automatic segmentation.

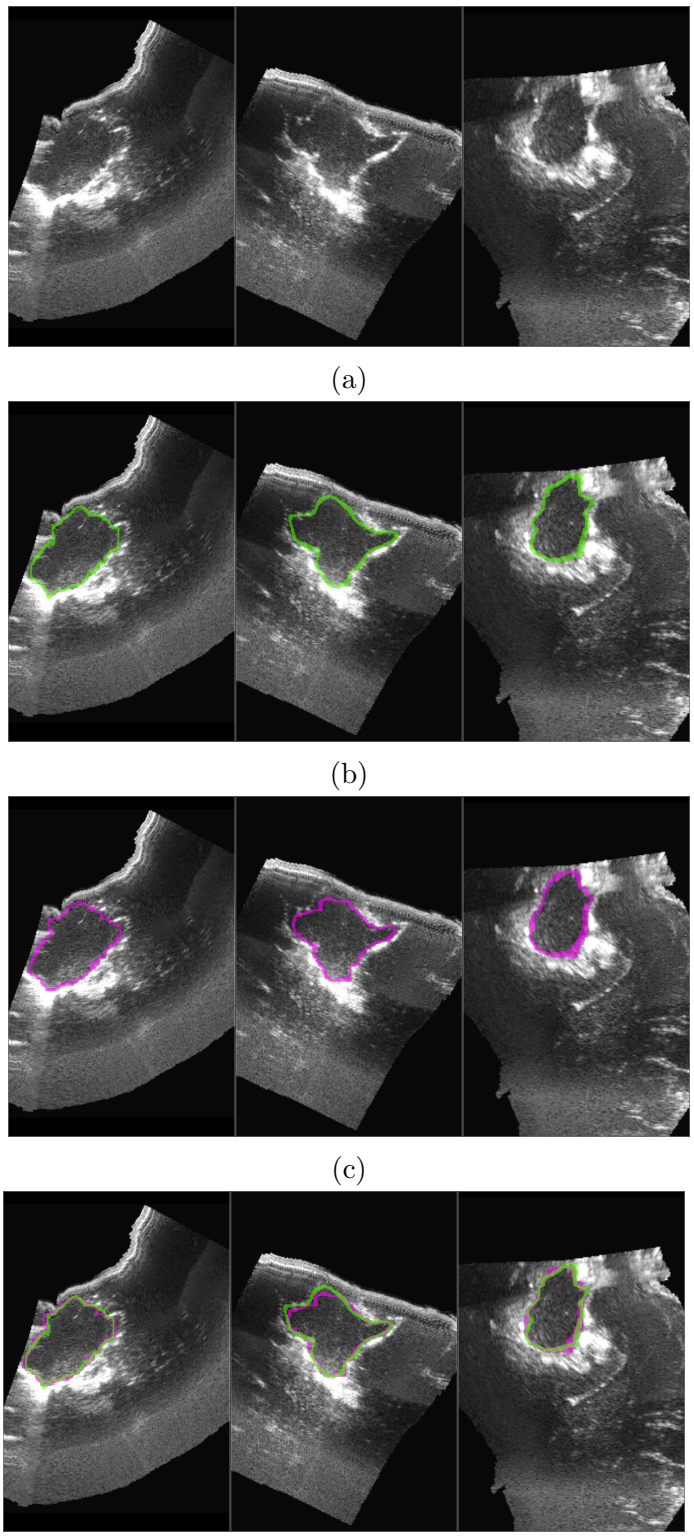
After having trained the neural network on the RESECT dataset volumes, we applied it to volumes after the resection of the BITE dataset. As observed by [146], the acquisition methodologies and the quality of the volumes are different between the two datasets. The cavity is sometimes partially visible in the BITE dataset, whereas on the contrary in the RESECT volumes, it is always completely observable and usually surrounded by a bright border. In Fig. 8.7, we show an example of a BITE case segmented by our methodology.

The proposed 3D U-Net requires a mean time of 4.86 s to generate the mask of interest.

8.3.2. Registration

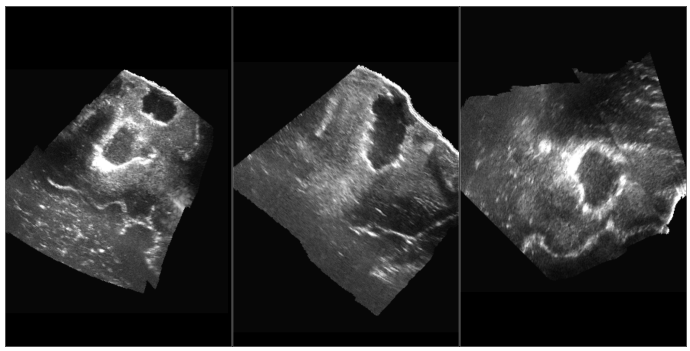
After having registered the volumes, the deformation field is applied to the related landmarks and the mTRE is computed (please refer to Appendix III for more details). The mTRE obtained excluding the masks generated by the algorithm trained with the segmentation of the first rater is 1.25 mm and 1.24 mm (before-during and before-after registration, respectively), whereas the results by the method discarding the resection cavity segmented with the 3D U-Net designed with the masks of the second rater are 1.22 mm and 1.21 mm. In Tables 8.2 and 8.3, we report the detailed results of the second method. Table 8.2 is related to the volumes before and during resection, whereas Table 8.3 refers to those before and after tumor removal. In both tables, the fourth column is related to the results obtained by our solution without the exclusion of the resection cavity, whereas in the last one, we report the registration results achieved by discarding it. For the volumes acquired during and before resection, we achieved an mTRE of 1.22 mm by excluding the resection cavity and 2.57 mm without excluding it. Instead, the corresponding results for the volumes acquired after and before removal are of 1.21 mm and 3.53 mm. For a better comparison with previously proposed methods, each table also contains a second section in which the mTRE achieved by other methods is listed.

We conducted a statistical test to determine whether the two proposed methods (with and without masking) show a statistically significant difference. The data are not normally distributed, and thus, we decided to use the nonparametric Wilcoxon signed rank test. It tests the null hypothesis that two related paired samples (the

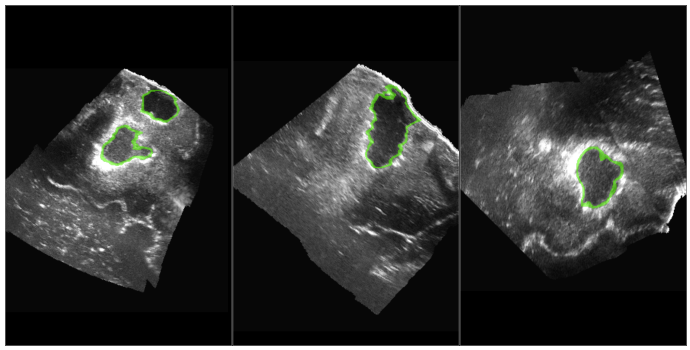


(a) (b) (c) (d)

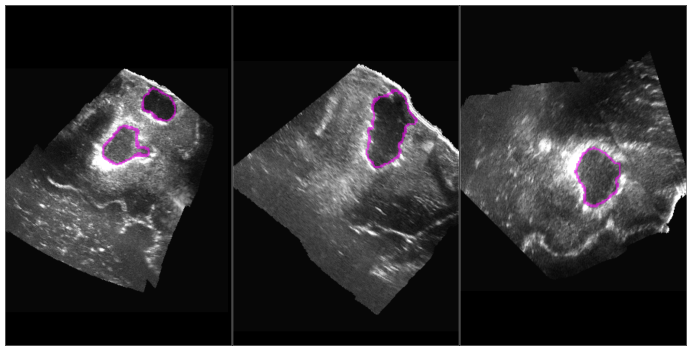
Figure 8.4.: Example of automatic segmentation in a volume acquired during resection. (a) The original intensity volume, on which a manual annotation has been drawn (in green) (b). The automatically computed mask is visible in (c) (in pink), whereas a direct comparison between the ground truth and the generated masks is given in (d)



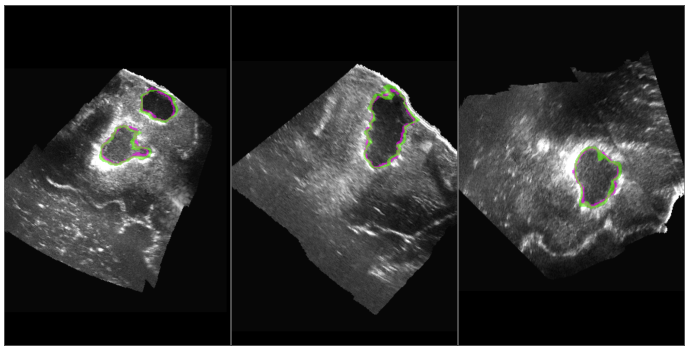
(a)



(b)



(c)



(d)

Figure 8.5.: Example of automatic segmentation in a volume acquired after resection. (a) The original intensity volumes, on which a manual annotation has been drawn (in green) (b). The automatically computed mask is visible in (c) (in pink), whereas a direct comparison between the ground truth and the generated masks is given (d)

Table 8.2.: Registration results for volumes acquired before and during resection (RESECT dataset)

Patient ID	**# Of landmarks**	**Mean initial distance (mm)**	**Without masking**	**With masking**
1	34	2.32 (1.49–3.29)	2.32 (1.49–3.29)	0.64 (0.20–1.72)
2	16	3.10 (1.79–5.19)	1.57 (0.56–4.07)	1.50 (0.32–4.20)
3	17	1.93 (0.67–3.02)	0.86 (0.18–2.14)	0.77 (0.29–1.39)
4	19	4.00 (3.03–5.22)	0.82 (0.22–2.38)	0.80 (0.20–2.40)
6	21	5.19 (2.60–7.18)	7.17 (0.54–12.58)	5.17(0.25–10.36)
7	22	4.69 (0.94–8.16)	1.95 (0.47–5.82)	1.98 (0.32–6.17)
12	24	3.39 (1.74–4.81)	2.01 (0.08–8.66)	0.84 (0.16–2.02)
14	22	0.71 (0.42–1.59)	0.49 (0.11–1.19)	0.41 (0.09–1.11)
15	21	2.04 (0.85–2.84)	6.37 (2.03–11.15)	0.60 (0.16–1.29)
16	19	3.19 (1.22–4.53)	11.23(9.20–13.26)	1.26 (0.06–3.40)
17	17	6.32 (4.65–8.07)	1.76 (0.33–4.70)	1.49 (0.25–3.69)
18	23	5.06 (1.55–7.44)	1.25 (0.26–3.98)	1.18 (0.34–3.76)
19	21	2.06 (0.42–3.40)	2.06 (0.20–6.98)	0.96 (0.12–2.76)
21	18	5.10 (3.37–5.94)	4.54 (0.51–9.63)	1.11 (0.18–3.91)
24	21	1.76 (1.16–2.65)	0.96 (0.22–2.72)	0.67 (0.17–1.44)
25	20	3.60 (2.19–5.02)	0.55 (0.15–1.61)	0.55 (0.18–1.61)
27	16	4.93 (3.61–7.01)	1.06 (0.05–2.92)	0.87 (0.15–2.19)
Mean value $\pm$ SD	20.0 ± 4.8	3.49 ± 1.55	2.57 ± 2.93	1.22 ± 1.20
Other methods				Mean value
[79]				1.36

The table shows the registration results (mTRE registration errors in mm, the range (min–max) of the distances in parenthesis, and the standard deviation) obtained by our solution. The different cases (first column) come with landmarks, which are specified in the second column. In the third column, we show the initial registration. We compare the results obtained without (fourth column) and by applying (fifth column) the exclusion of the resection cavity previously segmented. A second section (other methods) of the table shows the mean TRE obtained by previously proposed methods

Table 8.3.: Registration results for volumes acquired before and after resection (RESECT dataset)

Patient ID	**# Of landmarks**	**Mean initial distance (mm)**	**Without masking**	**With masking**
1	13	5.80 (3.62–7.22)	4.88 (1.02–7.44)	1.03 (0.18–3.23)
2	10	3.65 (1.71–6.72)	3.28 (1.46–6.16)	3.90 (2.36–5.88)
3	11	2.91 (1.53–4.30)	4.47 (1.68–6.42)	1.15 (0.34–1.93)
4	12	2.22 (1.25–2.94)	2.57 (0.38–4.84)	0.61 (0.13–1.17)
6	11	2.12 (0.75–3.82)	2.38 (0.19–6.01)	1.41 (0.26–4.70)
7	18	3.62 (1.19–5.93)	3.41 (0.52–5.62)	2.03 (0.13–4.88)
12	11	3.97 (2.58–6.35)	7.85 (5.77–11.03)	0.79 (0.32–1.92)
14	17	0.63 (0.17–1.76)	0.50 (0.19–1.04)	0.46 (0.15–0.98)
15	15	1.63 (0.62–2.69)	0.60 (0.21–1.48)	0.58 (0.19–1.25)
16	17	3.13 (0.82–5.41)	6.04 (3.51–8.84)	0.92 (0.28–2.27)
17	11	5.71 (4.25–8.03)	3.33 (0.90–6.17)	1.10 (0.23–4.89)
18	13	5.29 (2.94–9.26)	4.04 (0.69–6.63)	1.13 (0.37–3.13)
19	13	2.05 (0.43–3.24)	2.53 (0.50–8.60)	1.10 (0.41–2.40)
21	9	3.35 (2.34–5.64)	1.99 (0.52–3.47)	1.80 (0.84–3.39)
24	14	2.61 (1.96–3.41)	4.80 (3.17–7.12)	0.87 (0.35–2.42)
25	12	7.61 (6.40–10.25)	6.73 (3.29–9.20)	1.21 (0.15–5.65)
27	12	3.98 (3.09–4.82)	0.60 (0.17–1.79)	0.53 (0.14–0.90)
Mean value $\pm$ SD	12.9 ± 2.6	3.55 ± 1.76	3.53 ± 2.12	1.21 ± 0.66
Other methods				Mean value
[3]				1.49
[79]				2.05
[145]				1.27
[150]				1.92

The table shows the registration results (mTRE registration errors in mm, the range (min–max) of the distances in parenthesis, and the standard deviation) obtained by our solution. The different cases (first column) come with landmarks, which are specified in the second column. In the third column, we show the initial registration. We compare the results obtained without (fourth column) and by applying (fifth column) the exclusion of the resection cavity previously segmented. A second section (other methods) of the table shows the mean TRE obtained by previously proposed methods

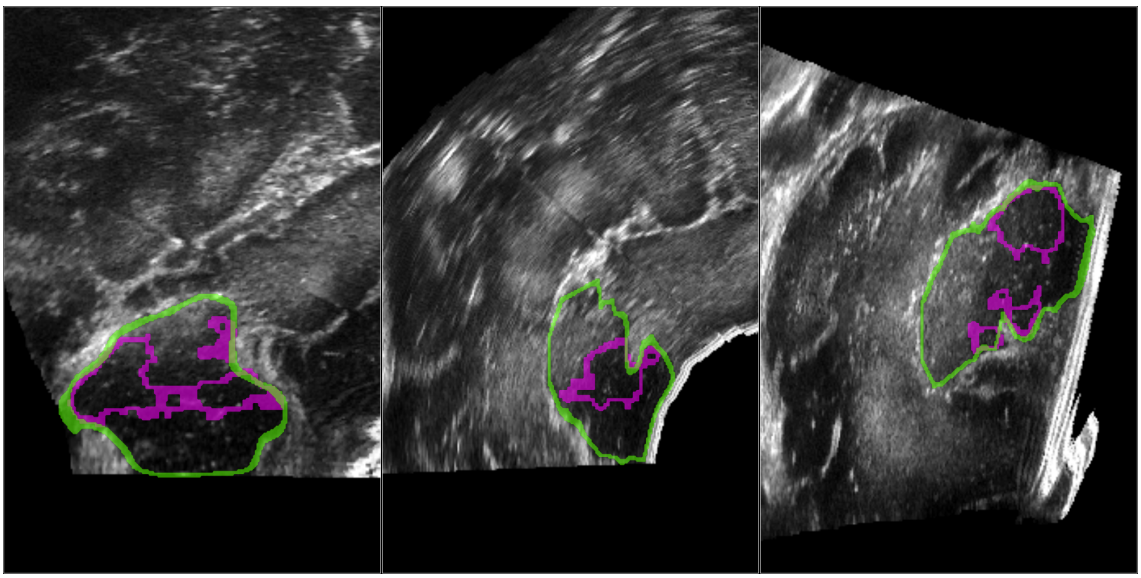


Figure 8.6.: Worst-case segmentation. This figure is a visual inspection of volume 18 during resection for which we obtained the worst DICE index. The green contour is related to the ground truth, whereas the purple one to the automatic segmentation. The generated mask is smaller compared to the manual annotation

results of the two algorithms) come from the same distribution. For both studies (before-during and before-after), we verified that the null hypothesis cannot be accepted (p-value < 0.001), meaning that there is a statistical difference between the two methods. Besides, we conducted the same statistical analysis for the two registration methods using the masks generated by the two different 3D U-Net models. The test fails to reject the null hypothesis (p-value > 0.6) for both registration tasks.

Moreover, we tested our final solution also on the BITE dataset. The related results are available in Table 8.4.

The visual results for the registration task are shown in Figs. 8.8 and 8.9, displaying the data in three projections. The volumes before resection are colored in purple and are overlaid on the volumes during (Fig. 8.8) and after (Fig. 8.9) resection, shown in gray intensity. Each figure contains two sub-images displaying the overlay of the two volumes of interest, before (a) and after (b) having applied our registration algorithm. The difference between a and b is related to the volume before resection (in purple), which is deformed according to the deformation field computed by the registration algorithm.

The registration step requires a mean of 49.67 s to register the volumes of interest.

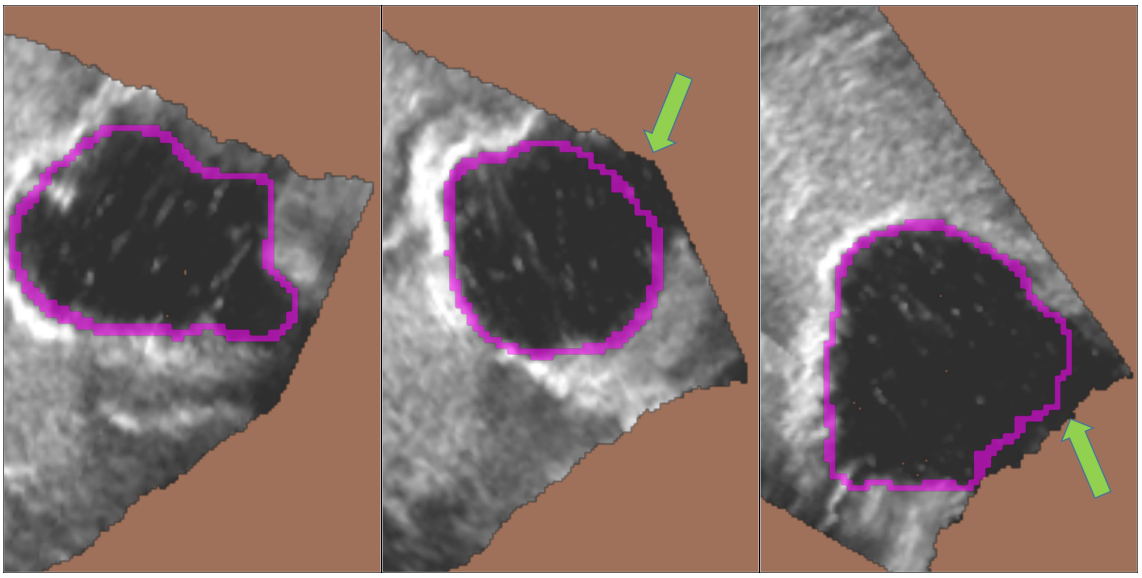
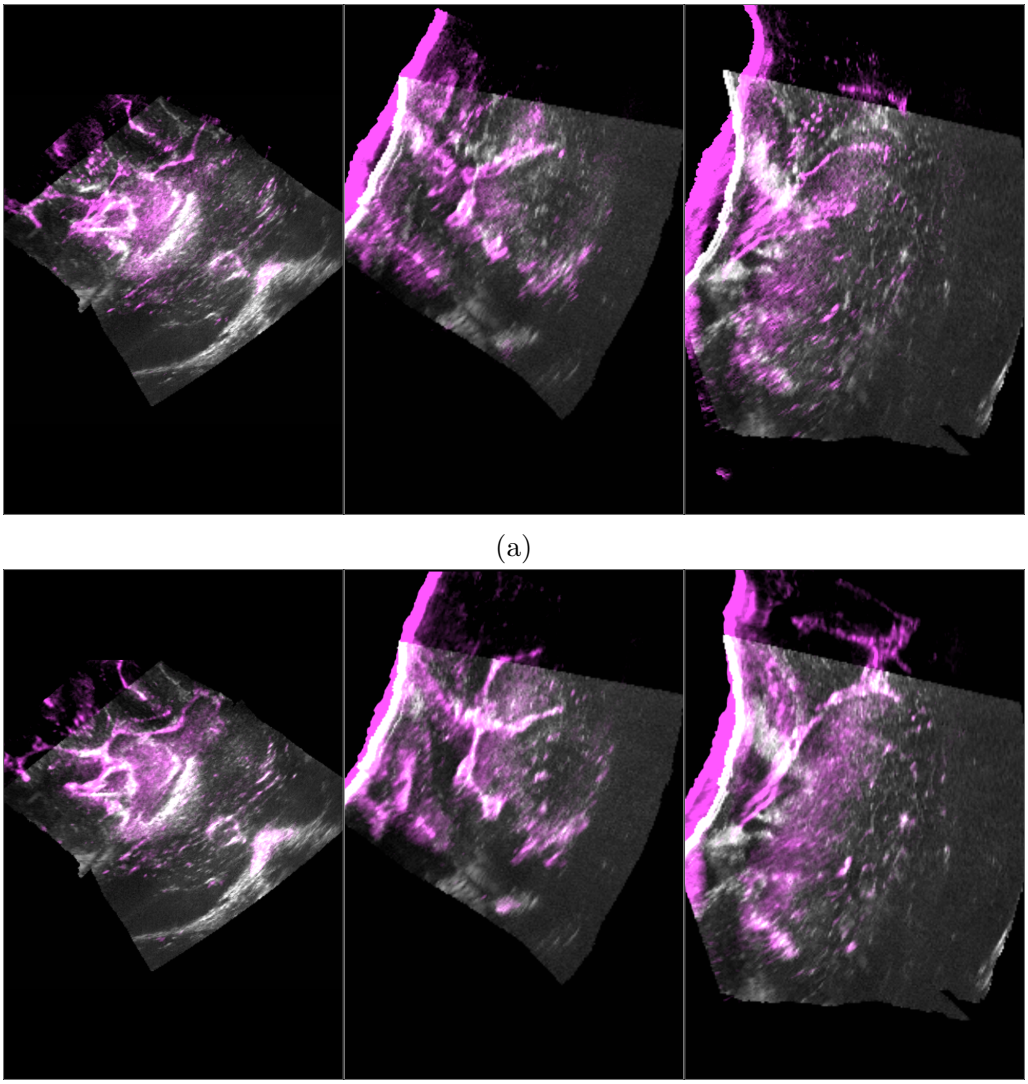


Figure 8.7.: Segmentation results on resection cavity with an open border. This figure shows the result of the automatic segmentation (purple border) on a BITE dataset volume. The background surrounding the US cone is colored orange. By looking at the positions pointed by the green arrows, we can see how a side of the resection cavity is not surrounded by a hyperintense contour: The US probe is directly inserted in the RC, and this side of the cavity has no enhanced border. The acquisition procedure of the volumes included in the BITE dataset differs from the RESECT one, in which the resection cavities are always surrounded by a hyperintense contour. Our segmentation algorithm correctly segments the structure of interest when it is surrounded by a hyperintense element. However, it under-segments the part of the cavity where no hyperintense border is available (the hypointense parts pointed by the arrows should be included within the purple border)

Table 8.4.: Registration results for volumes acquired before and after resection (the BITE dataset)

Patient ID	**Mean initial distance (mm)**	**With masking**
2	2.30 (0.57–5.42)	1.70 (0.65–4.11)
3	3.40 (0.0–5.09)	1.49 (0.25–3.91)
4	4.60 (2.96–5.88)	5.34 (3.09–8.89)
5	4.11 (2.58–5.52)	1.17 (0.41–2.32)
6	2.26 (1.36–3.10)	1.08 (0.38–2.38)
7	3.87 (2.60–5.07)	1.23 (0.54–2.09)
8	2.51 (0.67–3.93)	1.21 (0.45–2.32)
9	2.21 (1.00–4.59)	1.57 (0.26–4.22)
10	3.86 (0.98–6.68)	1.18 (0.44–2.26)
11	2.74 (0.44–8.22)	2.29 (0.20–7.49)
12	10.54 (7.85–13.04)	10.79 (7.68–13.34)
13	1.62 (1.33–2.21)	0.71 (0.25–1.76)
14	2.19 (0.59–3.99)	1.17 (0.34–3.10)
Mean value $\pm$ SD	3.55 ± 2.28	2.38 ± 2.78
Other methods		Mean value
[108]		1.50
[127]		1.50
[10]		1.50
[3]		1.54
[79]		2.48

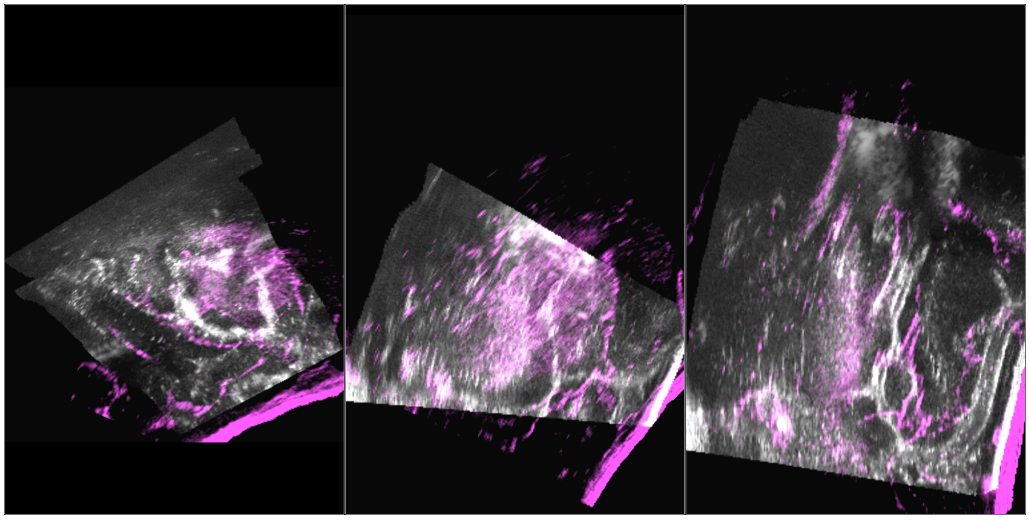
The table shows the registration results (mTRE registration errors in mm, the range (min–max) of the distances in parenthesis, and the standard deviation) obtained by our solution excluding the resection cavity previously segmented. A second section (other methods) of the table shows the mean mTRE obtained by previously proposed approaches. The methods [127; 108; 10] were tested only on the BITE dataset, whereas [3] also on the RESECT dataset, but only on the set of volumes acquired after resection



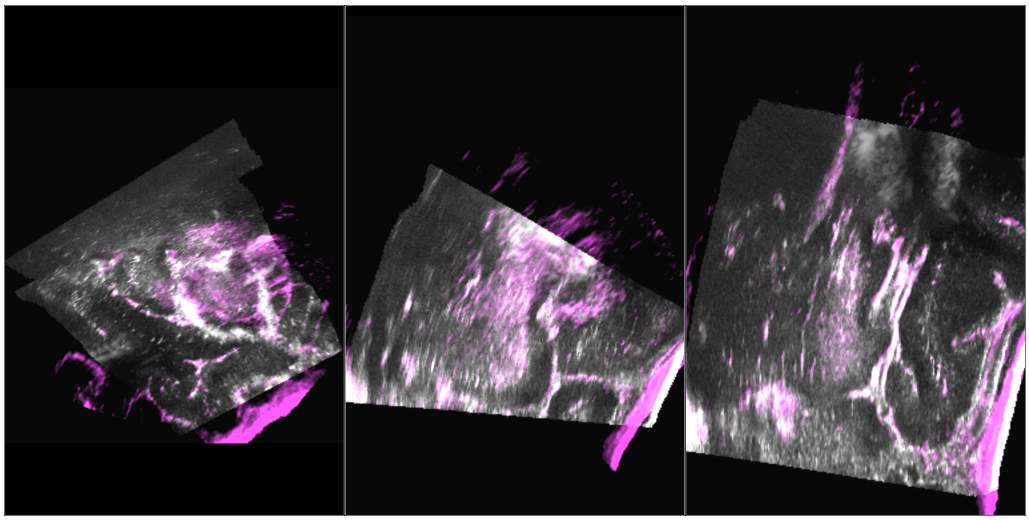
(a)

(b)

Figure 8.8.: Registration of volumes acquired before and during resection. The two figures show the volumes before resection (in purple) overlaid on the volumes during resection (in gray). (a) The situation before applying our registration algorithm; (b) the overlay of the volumes after having deformed the moving image



(a)



(b)

Figure 8.9.: Registration of volumes acquired before and after resection. The two figures show the volumes before resection (in purple) overlaid on the volumes after resection (in gray). (a) The situation before applying our registration algorithm; (b) the overlay of the volumes after having deformed the moving image

8.4. Discussion

Regarding the segmentation approach, our solution can achieve a mean DICE index of 0.84 by comparing the automatically generated masks with the clinically validated ground truth. The visual results in Figs. 8.4 and 8.5 give evidence that the 3D U-Net achieves good results in segmenting the resection cavities in US volumes. The only exception is given by volume 18 during resection, for which our solution obtained the worst DICE index. This is confirmed also by visual inspection in Fig. 8.6: The resection cavity is under-segmented in comparison to the manual annotation. The resection cavities are usually hypoechogenic structures contoured by a hyperechogenic border. However, for the case in Fig. 8.6, the resection cavity has intensity characteristics slightly different compared to the majority of the other volumes. Thus, a possible cause for this failure may be related to the intensity of the internal part of the resection cavity, which includes also a partial hyperechogenic area. Nevertheless, for the other volumes of the test set, we obtained DICE values in line with the rest of the dataset (Table 8.1). Our algorithm has been also applied to the BITE volumes. Our automatic method can segment the volumes of interest, but it usually underestimates the parts where a hyperintense border is missing. In this dataset, the US probe is usually inserted inside the resection cavity, whereas the volumes of the RESECT dataset are usually acquired from a position outside it. Consequentially, in the BITE dataset, the hyper-intense borders surrounding the cavity are not visible on the sides of the volume in which the acquisition probe has been interested (Fig. 8.7), whereas, in the RESECT dataset, the resection cavity is usually visible as a closed structure. The other algorithms proposed to segment the resection cavity used a privately defined ground truth to train their solutions. Therefore, a numerical comparison based on the DICE index is not possible. Regarding the time required to process a single volume, the solution proposed in [146] requires around 5 min. It may be due to (1) the deeper architecture of the neural network that they used, which requires more time to process a single volume, and (2) the sampling method they chose, a sliding window approach with large patches. On the contrary, our algorithm is faster, requiring less than 5 s in the inference process.

Regarding the registration results, the proposed solution registers the volumes of interest by reducing the overall mTRE for all the sets of volumes taken into account. For the RESECT dataset, we are able to reduce the initial mTRE for all

the volumes of both sets. Figures 8.8 and 8.9 also give impressions that the proposed registration algorithm produces a better alignment of the volumes, compared to the initial alignment. If we compare the results obtained by using the masks generated by the two 3D U-Net models trained on different ground truths, we can observe that the results differ from each other: Changes in the segmentation have an impact on the registration outcome. However, these results also show that the differences in the masks are negligible on the registration results (less than 0.1 mm in terms of mTRE), as long as the segmentation includes the resection cavity. Moreover, from the numerical results obtained in the RESECT dataset (Tables 8.2 and 8.3), the exclusion of the resection cavity from the registration process (fifth column) provides better results than the case in which it is not excluded (fourth column). As expected, by discarding the non-corresponding elements from the registration process, the algorithm can focus on the elements in common and therefore obtain better results.

Compared to previous approaches, for both the sets of the RESECT dataset, our algorithm achieves better results than the solutions compared in Tables 8.2 and 8.3. This is true for the algorithms proposed to register both sets [146; 79], but also for methods considering only the registration of volumes before-after resection [79; 150]. Additionally, our final method has been also tested on the BITE dataset (Table 8.4), in which it can reduce the mTRE of each pair of volumes. In the comparison of our approach with previous solutions, the algorithms [108; 127; 10] were tested only on the BITE dataset: Even if they obtained slightly better results than our solution, they lack generalization. On the contrary, our method has been tested on a larger set of volumes, providing a better generalization. Moreover, the overall mTRE is improved with respect to [79] (Table 8.4). However, the solution proposed by [3] achieves a better mTRE. Our results are mostly affected by the registration of volume 12, for which the initial error is only slightly reduced. The starting condition for the registration of this case is the worst of all the sets of volumes, and our method may be affected by the initial registration information provided by the optical tracking system. Combining the time required by the segmentation of the resection cavity and the registration step, the proposed algorithm can register two volumes in less than 55 seconds. This small delay to the neurosurgical procedure may be tolerable, especially if a better understanding of the surgical scene after the registration can be achieved.

8.5. Conclusions

We presented here an automatic algorithm for 3D segmentation of resection cavities in US volumes, acquired in the neurosurgical procedures for tumor removal. For this specific method, we manually annotated a ground truth that has been made publicly available. Besides, we proposed a novel method to register US volumes acquired in the neurosurgery context. In our solution, the resection cavities are excluded from the registration thanks to the automatic segmentation method, reducing the impact of non-corresponding elements in the computation of the distance measure. Our experiments show that the registration results are only slightly influenced by the differences in the masks, as long as they include the resection cavity to be excluded. On the contrary, we show that by omitting the exclusion of the resection cavities, a worsening of the results is obtained. To the best of our knowledge, it is the first time that the resection cavities are taken into account to improve the registration of US volumes in neurosurgical tumor resection. Moreover, the registration results obtained in the RESECT dataset are the lowest in comparison with the other methods in the literature (Tables 8.2 and 8.3).

In future work, we plan to manually annotate the resection cavity in the volumes of the BITE dataset, to propose a more generalized solution. Moreover, the registration method based on the exclusion of the resection cavity could be also utilized for the inter-modality registration of intraoperative US volumes and preoperative MRI.

Acknowledgements

This work was funded by the H2020 Marie-Curie ITN TRABIT (765148) project. Moreover, Prof. Dr. Kikinis is supported by NIH Grants NIH/NIBIB P41EB015902, NIH/NIBIB P41EB015898, NIH/NCI 1R01CA235589-01A1, Contract number 19X037Q from Leidos Biomedical Research under Task Order HHSN26100071 from NCI.

Part III.

Concluding Remarks

This thesis investigated the segmentation-enhanced registration of intrapatient imaging data acquired at multiple stages of brain tumor treatments. Chapters 3 and 4 proposed methods registering pre- and post-operative magnetic resonance imaging volumes, whereas Chapters 7 and 8 focused on the registration of intra-operative ultrasound acquisitions. In Chapters 3 and 5, I also conducted a quantitative analysis to find the most informative sequence for designing registration and segmentation solutions for multiparametric MRI data.

In this chapter, I will review the findings reported in the thesis. The text is divided into two parts, each respectively covering the results presented for MRI and US data. For each modality, a summary of the proposed methods is provided, and ideas for future research directions are suggested.

9.1. Pre- and Post-operative MRI Acquisitions

In the first part of the thesis, three automatic methods were proposed for registering pre- and post-operative MRI volumes acquired in brain tumor treatments (see Chapters 3 and 4). These solutions have in common the exclusion of the contribution of non-corresponding pathological tissues from the registration process. The results demonstrated that the discard of pathological tissues from the objective function is beneficial for the registration of subsequent acquisitions.

Additionally, the experiments presented in Chapter 3 quantitatively assessed the influence of four MRI sequences on the registration of MRI volumes. Furthermore, Chapter 5 proposed a convolution neural network-based solution to segment resection cavities in post-operative data and compared the results obtained by models trained on different sequences.

9.1.1. Image Registration Enhanced by Segmentation of Non-corresponding Structures

In Chapter 3, I presented a deep-learning solution and an iterative method for registering pre- and post-operative MRI volumes. I quantitatively evaluated the impact of excluding non-corresponding pathological tissues from the distance measure of the registration methods by using the segmentation of pathological tissues. To do

so, I designed two versions of each solution, one masking out and the other one including the pathology in the distance measure. In the iterative method, the masking procedure is applied on the test set, whereas, for the deep learning solution, the exclusion of the pathological tissues is performed only during the training and validation phase, and not in the test set. The results show that excluding pathological tissues is beneficial to the registration outcome.

In the proposed methods, the pathological tissues excluded from the distance measure belong to the reference, post-operative volumes. The decision to use masks defined on the reference images is based on the assumption that pathological tissues in post-operative data are located in the same intracranial areas occupied by the tumors in pre-operative acquisitions. However, the methods do not take into account that tumors and edema in pre-operative images could be characterized by a larger number of voxels than those included in post-operative pathology. In future work, the masking procedure could also include the pathological tissues defined in the warped moving images.

As additional future work, it would also be interesting to evaluate how the proposed solutions perform on longitudinal imaging data including only post-operative cases. For this purpose, an additional dataset would be necessary.

Exclusion of the Pathological Tissues from the Regularization Term

Although the pathology is excluded from the search of correspondences in Chapter 3, it is still taken into account in the computation of the regularizer, which discourages large transformations and improves the plausibility of the deformation fields. However, when the pathological tissues are discarded from the distance measure, the registration method solely focuses on correctly registering corresponding healthy tissues. Consequentially, the inclusion of voxels of the pathology in the regularizer can be considered redundant since the only interest should be in regularizing the deformations for voxels of healthy voxels. Besides, the registration of corresponding structures could benefit from excluding the contribution of voxels belonging to the pathology from the computation of the regularizer, since more plausible deformations could be achieved. Chapter 4 proposed a registration method in which the contribution of the voxels belonging to the pathological tissues is also discarded from the computation of the regularization term. The experiments showed the benefits of the proposed method w.r.t. the same solution not excluding the pathology.

The adjustment of the impact of the pathology on the regularizer has already been investigated in the literature for the registration of brain images, but it is the first time that its complete exclusion has been evaluated. As a possible future direction, a weighting parameter for the regularization could take into account healthy tissues too, and not only exclude the contribution of the pathology. As shown by other groups, more realistic transformations could be achieved for the brain tissues.

9.1.2. Influence of Different MRI Sequences

Chapter 3 investigated how four different MRI sequences impact the registration of pre- and post-operative MRI acquisitions. The proposed registration solutions are designed on T1-CE, T1, FLAIR, and T2 sequences. The results showed that the methods using the T1-CE achieved the best results. Although this outcome was somewhat expected due to the higher image resolution of this protocol, this finding has been quantitatively demonstrated for the first time in this thesis. The worst sequence for both registration approaches is the FLAIR.

The deep learning method is further trained and tested on the combination of two sequences. The experiments showed that the use of multi-features input images doesn't improve the results. This finding could be somehow surprising since it is expected that learning-based solutions, trained on more input features (in this case, MRI sequences), can more precisely adapt their model parameters to minimize the loss function and, thus, better solve a specific task. Differently from what one could anticipate, more features didn't lead to an improvement in the outcome. In future research, more combinations of MRI sequences could be tested for the deep learning method.

Influence of MRI Sequences on the Segmentation of the Resection Cavity

Chapter 5 analyzed the effects of four sequences on the design of a deep learning method, based on the 3D U-Net, for the segmentation of the resection cavity in post-operative MRI acquisitions. Other research groups proposed methods to segment this anatomical structure in brain images, and, in particular, a deep learning-based method was recently presented in [100]. In their work, all the available sequences (T1-CE, T1, FLAIR, T2) were used as input, and the authors didn't analyze which sequence would be the most informative to train a deep-learning method. My experiments concluded that the T1-CE sequence is preferable for training and testing

the 3D U-Net. My work also demonstrated that using multiparametric input data, composed of all the available sequences, leads to a worsening of the DICE values w.r.t. the same model trained on the T1-CE.

Based on my experiments, it appears that the T1-CE sequence is the preferred choice. This could be attributed to its higher resolution, as has been observed in previous registration tasks.

In future work, it would be valuable to assess the impact of different combinations of MRI sequences on the 3D U-net-based method's training and testing, hopefully on a larger dataset. Additionally, multiple deep-learning architectures and methods could be tested [151].

9.2. Intraoperative Ultrasound Acquisitions

The registration of subsequent intraoperative US acquisitions is tackled by two automatic segmentation-enhanced solutions in Chapters 7 and 8.

9.2.1. Image Registration Based on Segmentation Masks of Corresponding Anatomical Structures

Chapter 7 presented a segmentation-based registration method in which only corresponding tissues are considered for guiding the registration process. First, sulci and falx cerebri are automatically segmented in successive US intrapatient acquisitions with a solution based on a 3D U-net. The masks are then used to compute the distance measure in the iterative registration method. The deformation fields generated by registering the segmentation masks are applied to deform the corresponding US acquisitions. The idea of utilizing segmentation to guide the registration of US volumes is not a novelty in the field of US registration. However, my solution was the first to demonstrate that the use of automatically generated segmentation masks of falx cerebri and sulci represents a suitable strategy for the registration of subsequent US data acquired in neurosurgery.

A limitation of the proposed solution might be related to the metric employed for estimating the registration results in the utilized datasets. The mean target registration error is computed on corresponding landmarks annotated on brain structures, such as falx cerebri and sulci. These also coincide with the anatomical elements

that are automatically delineated by the segmentation method and whose masks are used to guide the registration process. The results computed with this metric might be influenced by the fact that the proposed solution might only focus on correctly registering those anatomical elements, on which landmarks are also annotated. To ensure that the US volumes are correctly registered, a visual inspection of the registration was conducted in this thesis. Yet, in the future, it would be interesting to also quantitatively evaluate the registration results in brain locations different from the automatically segmented anatomical structures. To this purpose, additional landmarks could be acquired in sites different from the automatically segmented structures, possibly also on new data cohorts. A challenge in this regard was, and remains, the limited availability of publicly available datasets, on which additional results could be computed.

A second observation about the proposed solution regards the automatically segmented structures to guide the registration process. The goal of my work was to demonstrate that the idea of only using segmentation masks is a valid option for registering subsequent US data. Since the masks are used to guide the registration method, the quality of the generated masks could have a strong influence on the final registration outcome. In future work, it would be interesting to evaluate whether masks obtained by multiple raters or other automatic methods would affect the registration outcome. It could be expected that better quality segmentation would lead to an improvement in the results of the segmentation-based registration of US data.

9.2.2. Image Registration Excluding the Contribution Non-Corresponding Anatomical Structures

The second solution is a segmentation-enhanced method, in which resection cavities are first automatically segmented by a solution based on a 3D U-Net in intra- and post-operative US acquisitions. Second, the generated masks are utilized in the registration process to exclude from the computation of the distance measure the contribution of the cavities, for which one-to-one correspondences are not guaranteed. The experiments of Chapter 8 demonstrated how the exclusion of automatically generated masks of non-corresponding structures is beneficial for registering subsequent US data: The method discarding the contribution of resection cavities achieves better results than the same solution including these structures in the distance measure computation. By comparing the approaches presented in Chap-

ters 8 and 7, one could observe how the use of the automatically generated masks is complementary: in the first solution, corresponding structures are used, whereas the method presented in Chapter 8 utilizes segmentation of non-corresponding elements.

This work aimed to validate the effects of excluding the contribution of the resection cavities from the registration process, automatically segmented by a deep learning method. As for the first solution proposed for US acquisitions, the quality of the segmentation masks used in the registration could be a very important parameter affecting the registration results. Even if it was out of the scope of the proposed work, I partly evaluated this topic in the conducted investigation by comparing the same segmentation 3D U-Net-based method trained on ground truth generated by two different raters. Nevertheless, it would be interesting to more extensively evaluate how masks generated from additional methods would affect the registration outcome. In this regard, additional works have been recently conducted [152; 153; 154; 155].

9.3. Future Directions

This thesis demonstrated that segmentation-enhanced solutions are suitable methods for the registration of neurosurgical subsequent acquisitions in which anatomical correspondences are not guaranteed. I confirmed this finding by focusing on the registration of intraoperative US acquisitions and pre- and post-operative MRI volumes obtained during brain tumor treatments. To confirm this result more comprehensively, future solutions dealing with the registration of intracranial acquisitions of other image modalities could also analyze the effects of discarding non-corresponding tissues through segmentation masks. I believe one could expect that the exclusion of the pathology from the registration process would also be beneficial for other image modalities.

My experiments quantitatively demonstrated that the choice of the MRI sequence strongly impacts the design of registration methods dealing with multiparametric MRI data. The T1-CE sequence should be preferred to guide the two registration methods, i.e., the iterative solution and the deep learning-based approach. A similar analysis has also been conducted for the segmentation method, based on the 3D U-Net, for automatically delineating the resection cavity: In this case, the model

using only the T1-CE sequence achieves a higher DICE index than employing all the available sequences. Therefore, from my point of view, image-based methods dealing with multiparametric MRI data should first concentrate on choosing the most appropriate MRI sequence. To extensively investigate this topic, future works could explore how different combinations of MRI sequences, in addition to those utilized in the methods presented in this thesis, affect the registration and segmentation of multiparametric data.

Bibliography

[1] D. Avola, L. Cinque, A. Fagioli, G. Foresti, and A. Mecca, "Ultrasound Medical Imaging Techniques: A Survey," *ACM Comput. Surv.*, vol. 54, pp. 1–38, Apr. 2021.

[2] P. Risholm, A. J. Golby, and I. William M. Wells, "Multi-Modal Image Registration for Pre-Operative planning and Image Guided Neurosurgical Procedures," *Neurosurg. Clin. N. Am.*, vol. 22, p. 197, Apr. 2011.

[3] I. Machado, M. Toews, J. Luo, P. Unadkat, W. Essayed, E. George, P. Teodoro, H. Carvalho, J. Martins, P. Golland, S. Pieper, S. Frisken, A. Golby, and W. Wells, "Non-rigid registration of 3D ultrasound for neurosurgery using automatic feature detection and matching," *Int. J. CARS*, vol. 13, pp. 1525–1538, Oct. 2018.

[4] B. Baheti, D. Waldmannstetter, S. Chakrabarty, H. Akbari, M. Bilello, B. Wiestler, J. Schwarting, E. Calabrese, J. Rudie, S. Abidi, M. Mousa, J. Villanueva-Meyer, D. S. Marcus, C. Davatzikos, A. Sotiras, B. Menze, and S. Bakas, "The Brain Tumor Sequence Registration Challenge: Establishing Correspondence between Pre-Operative and Follow-up MRI scans of diffuse glioma patients," *arXiv*, Dec. 2021.

[5] S. Bayer, A. Maier, M. Ostermeier, and R. Fahrig, "Intraoperative Imaging Modalities and Compensation for Brain Shift in Tumor Resection Surgery," *Int. J. Biomed. Imaging*, vol. 2017, June 2017.

[6] S. Bette, J. Gempt, T. Huber, C. Delbridge, B. Meyer, C. Zimmer, J. S. Kirschke, and T. Boeckh-Behrens, "FLAIR signal increase of the fluid within the resection cavity after glioma surgery: generally valid as early recurrence marker?," *J. Neurosurg.*, vol. 127, pp. 417–425, Aug 2017.

[7] M. Nawaz, T. Nazir, M. Masood, A. Mehmood, R. Mahum, M. A. Khan, S. Kadry, and O. Thinnukool, "Analysis of Brain MRI Images Using Improved CornerNet Approach," *Diagnostics (Basel).*, vol. 11, p. 1856., Oct. 2021.

[8] A. Fawzi, A. Achuthan, and B. Belaton, "Brain Image Segmentation in Recent Years: A Narrative Review," *Brain Sciences*, vol. 11, Aug. 2021.

[9] M. Cenek, M. Hu, G. York, and S. Dahl, "Survey of Image Processing Techniques for Brain Pathology Diagnosis: Challenges and Opportunities," *Front. Rob. AI*, vol. 0, 2018.

[10] H. Zhou and H. Rivaz, "Registration of Pre- and Postresection Ultrasound Volumes With Noncorresponding Regions in Neurosurgery," *IEEE J. Biomed. Health Inf.*, vol. 20, pp. 1240–1249, Apr. 2016.

[11] R. Stupp, M. Brada, M. J. van den Bent, J.-C. Tonn, G. Pentheroudakis, ESMO Guidelines Working Group, A. Cervantes, G. Pentheroudakis, E. Felip, N. Pavlidis, R. A. Stahel, F. Cardoso, E. Senkus-Konefka, D. Arnold, P. Rougier, A. Horwich, C. Sessa, N. Colombo, L. Licitra, C. Buske, M. Ladetto, S. Peters, P. Casali, U. Keilholz, N. Cherny, and J. Herrstedt, "High-grade glioma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up," *Ann. Oncol.*, vol. 25, p. 3, Sept. 2014.

[12] D. A. Hardesty and N. Sanai, "The Value of Glioma Extent of Resection in the Modern Neurosurgical Era," *Front. Neurol.*, vol. 0, 2012.

[13] R. C. Miner, "Image-Guided Neurosurgery," *Journal of Medical Imaging and Radiation Sciences*, vol. 48, pp. 328–335, Dec. 2017.

[14] "Adult Central Nervous System Tumors Treatment (PDQ®)–Patient Version," July 2022. [Online; accessed 4. Jul. 2022].

[15] L. M. DeAngelis, "Brain Tumors," *N. Engl. J. Med.*, vol. 344, pp. 114–123, Jan. 2001.

[16] S. Sathornsumetee and J. N. Rich, "New approaches to primary brain tumor treatment," *Anti-Cancer Drugs*, vol. 17, pp. 1003–1016, Oct. 2006.

[17] M. Lara-Velazquez, R. Al-Kharboosh, S. Jeanneret, C. Vazquez-Ramos, D. Mahato, D. Tavanaiepour, G. Rahmathulla, and A. Quinones-Hinojosa,

"Advances in Brain Tumor Surgery for Glioblastoma in Adults," *Brain Sci.*, vol. 7, p. 166, Dec. 2017.

[18] D. A. Orringer, A. Golby, and F. Jolesz, "Neuronavigation in the surgical management of brain tumors: current and future trends," *Expert Rev. Med. Devices*, vol. 9, pp. 491–500, Sept. 2012.

[19] N. Sanai, M.-Y. Polley, M. W. McDermott, A. T. Parsa, and M. S. Berger, "An extent of resection threshold for newly diagnosed glioblastomas," *J. Neurosurg.*, vol. 115, pp. 3–8, July 2011.

[20] O. Ganslandt, "Neuronavigation : concept, techniques and applications.," *Neurol. India*, vol. 50, p. 244, Jan. 2002.

[21] R. D. Bucholz, K. R. Smith, K. A. Laycock, and L. L. McDurmont, "Three-Dimensional Localization: From Image-Guided Surgery to Information-Guided Therapy," *Methods*, vol. 25, pp. 186–200, Oct. 2001.

[22] S. Başarslan and C. Göçmez, "NEURONAVIGATION: A REVOLUTIONARY STEP OF NEUROSURGERY AND ITS EDUCATION," *Medical Journal of Mustafa Kemal University*, vol. 5, pp. 24–31, Mar. 2015.

[23] I. J. Gerard, M. Kersten-Oertel, K. Petrecca, D. Sirhan, J. A. Hall, and D. L. Collins, "Brain shift in neuronavigation of brain tumors: A review," *Med. Image Anal.*, vol. 35, pp. 403–420, Jan. 2017.

[24] S. Narasimhan, J. A. Weis, M. Luo, A. L. Simpson, R. C. Thompson, and M. I. Miga, "Accounting for intraoperative brain shift ascribable to cavity collapse during intracranial tumor resection," *J. Med. Imaging (Bellingham).*, vol. 7, p. 031506., May 2020.

[25] O. Bozinov, J.-K. Burkhardt, C. M. Fischer, R. A. Kockro, R.-L. Bernays, and H. Bertalanffy, "Advantages and Limitations of Intraoperative 3D Ultrasound in Neurosurgery. Technical note," in *Intraoperative Imaging*, pp. 191–196, Wien, Austria: Springer, Vienna, Aug. 2010.

[26] D. Kuhnt, M. H. A. Bauer, and C. Nimsky, "Brain Shift Compensation and Neurosurgical Image Fusion Using Intraoperative MRI: Current Status and Future Challenges," *CRB*, vol. 40, no. 3, 2012.

[27] V. M. Tronnier, M. M. Bonsanto, A. Staubert, M. Knauth, S. Kunze, and C. R. Wirtz, "Comparison of intraoperative MR imaging and 3D-navigated ultrasonography in the detection and resection control of lesions," *Neurosurg. Focus*, vol. 10, pp. 1–5, Feb. 2001.

[28] L. Dixon, A. Lim, M. Grech-Sollars, D. Nandi, and S. Camp, "Intraoperative ultrasound in brain tumor surgery: A review and implementation guide," *Neurosurg. Rev.*, vol. 45, pp. 2503–2515, Aug. 2022.

[29] Y. Xiao, M. Fortin, G. Unsgård, H. Rivaz, and I. Reinertsen, "REtroSpective Evaluation of Cerebral Tumors (RESECT): A clinical database of pre-operative MRI and intra-operative ultrasound in low-grade glioma surgeries," *Med. Phys.*, vol. 44, pp. 3875–3882, July 2017.

[30] T. Selbekk, A. S. Jakola, O. Solheim, T. F. Johansen, F. Lindseth, I. Reinertsen, and G. Unsgård, "Ultrasound imaging in neurosurgery: approaches to minimize surgically induced image artefacts for improved resection control," *Acta Neurochir.*, vol. 155, pp. 973–980, June 2013.

[31] D. Kwon, M. Niethammer, H. Akbari, M. Bilello, C. Davatzikos, and K. M. Pohl, "PORTR: Pre-Operative and Post-Recurrence Brain Tumor Registration," *IEEE Trans. Med. Imaging*, vol. 33, pp. 651–667, Dec. 2013.

[32] J. Modersitzki, *FAIR: Flexible Algorithms for Image Registration (Fundamentals of Algorithms)*. Society for Industrial and Applied Mathematics, Nov. 2009.

[33] J. V. Hajnal and D. L. G. Hill, *Medical Image Registration (Biomedical Engineering) (English Edition)*. Boca Raton, FL, USA: CRC Press, June 2001.

[34] X. Chen, A. Diaz-Pinto, N. Ravikumar, and A. F. Frangi, "Deep learning in medical image registration," *Prog. Biomed. Eng.*, vol. 3, p. 012003, Feb. 2021.

[35] D. Ruan, J. A. Fessler, M. Roberson, J. Balter, and M. Kessler, "Nonrigid registration using regularization that accommodates local tissue rigidity," in *Proceedings Volume 6144, Medical Imaging 2006: Image Processing*, vol. 6144, pp. 346–354, SPIE, Mar. 2006.

[36] S. Boda, *Feature-Based Image Registration*. PhD thesis, May 2009.

[37] M. B. Hisham, S. N. Yaakob, R. A. A. Raof, A. B. A. Nazren, and N. M. Wafi, "Template Matching using Sum of Squared Difference and Normalized Cross Correlation," in *2015 IEEE Student Conference on Research and Development (SCOReD)*, pp. 100–104, IEEE, Dec. 2015.

[38] S. Albawi, T. A. Mohammed, and S. Al-Zawi, "Understanding of a convolutional neural network," in *2017 International Conference on Engineering and Technology (ICET)*, pp. 1–6, IEEE, Aug. 2017.

[39] K. O'Shea and R. Nash, "An Introduction to Convolutional Neural Networks," *ArXiv*, Nov. 2015.

[40] A. Hänsch, "Implications of Dataset Heterogeneity on Deep Learning Performance in Medical Image Segmentation," July 2021. [Online; accessed 5. Jul. 2022].

[41] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, pp. 234–241, Cham, Switzerland: Springer, Nov. 2015.

[42] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016*, pp. 424–432, Cham, Switzerland: Springer, Oct. 2016.

[43] A. Hering, S. Häger, J. Moltz, N. Lessmann, S. Heldmann, and B. van Ginneken, "CNN-based lung CT registration with multiple anatomical constraints," *Med. Image Anal.*, vol. 72, p. 102139, Aug. 2021.

[44] J. Zhang, K. Chen, D. Wang, F. Gao, Y. Zheng, and M. Yang, "Editorial: Advances of Neuroimaging and Data Analysis," *Front. Neurol.*, vol. 0, 2020.

[45] M. C. Mabray, R. F. Barajas, Jr., and S. Cha, "Modern Brain Tumor Imaging," *Brain Tumor Research and Treatment*, vol. 3, p. 8, Apr. 2015.

[46] D. B. Plewes and W. Kucharczyk, "Physics of MRI: A primer," *J. Magn. Reson. Imaging*, vol. 35, pp. 1038–1054, May 2012.

[47] B. M. Ellingson, M. Bendszus, J. Boxerman, D. Barboriak, B. J. Erickson, M. Smits, S. J. Nelson, E. Gerstner, B. Alexander, G. Goldmacher, W. Wick, M. Vogelbaum, M. Weller, E. Galanis, J. Kalpathy-Cramer, L. Shankar, P. Jacobs, W. B. Pope, D. Yang, C. Chung, M. V. Knopp, S. Cha, M. J. van den Bent, S. Chang, W. K. Al Yung, T. F. Cloughesy, P. Y. Wen, and M. R. Gilbert, "Editor's choice: Consensus recommendations for a standardized Brain Tumor Imaging Protocol in clinical trials," *Neuro-oncol.*, vol. 17, p. 1188, Sept. 2015.

[48] J. E. Villanueva-Meyer, M. C. Mabray, and S. Cha, "Current Clinical Brain Tumor Imaging," *Neurosurgery*, vol. 81, p. 397, Sept. 2017.

[49] J. E. Aldrich, "Basic physics of ultrasound imaging," *Crit. Care Med.*, vol. 35, p. Suppl, May 2007.

[50] J. A. Jensen, "Medical ultrasound imaging," *Prog. Biophys. Mol. Biol.*, vol. 93, pp. 153–165, Jan. 2007.

[51] K. Martin, "Introduction to B-mode imaging," in *Diagnostic Ultrasound: Physics and Equipment*, pp. 1–3, Cambridge, England, UK: Cambridge University Press, June 2010.

[52] T. Selbekk, A. S. Jakola, O. Solheim, T. F. Johansen, F. Lindseth, I. Reinertsen, and G. Unsgård, "Ultrasound imaging in neurosurgery: approaches to minimize surgically induced image artefacts for improved resection control," *Acta Neurochir.*, vol. 155, no. 6, p. 973, 2013.

[53] A. Fenster, G. Parraga, and J. Bax, "Three-dimensional ultrasound scanning," *Interface Focus*, vol. 1, p. 503, Aug. 2011.

[54] R. Verma, E. I. Zacharaki, Y. Ou, H. Cai, S. Chawla, S.-K. Lee, E. R. Melhem, R. Wolf, and C. Davatzikos, "Multi-parametric tissue characterization of brain neoplasms and their recurrence using pattern classification of MR images," *Acad. Radiol.*, vol. 15, p. 966, Aug 2008.

[55] X. Han, S. Bakas, R. Kwitt, S. Aylward, H. Akbari, M. Bilello, C. Davatzikos, and M. Niethammer, "Patient-Specific Registration of Pre-operative and Post-recurrence Brain Tumor MRI Scans," *Brainlesion : glioma, multiple sclerosis,*

stroke and traumatic brain injuries. BrainLes (Workshop), vol. 11383, p. 105, 2019.

[56] D. Waldmannstetter, F. Navarro, B. Wiestler, J. S. Kirschke, A. Sekuboyina, E. Molero, and B. H. Menze, "Reinforced redetection of landmark in pre- and post-operative brain scan using anatomical guidance for image alignment," *Biomedical Image Registration9th International Workshop, WBIR 2020, Portorož, Slovenia, December 1–2, 2020, Proceedings*, vol. 12120, p. 81—90, May 2020.

[57] A. van der Hoorn, J.-L. Yan, T. J. Larkin, N. R. Boonzaier, T. Matys, and S. J. Price, "Validation of a semi-automatic co-registration of MRI scans in patients with brain tumors during treatment follow-up," *NMR Biomed.*, vol. 29, pp. 882–889, Jul 2016.

[58] A. Klein, J. Andersson, B. A. Ardekani, J. Ashburner, B. Avants, M.-C. Chiang, G. E. Christensen, D. L. Collins, J. Gee, P. Hellier, J. H. Song, M. Jenkinson, C. Lepage, D. Rueckert, P. Thompson, T. Vercauteren, R. P. Woods, J. J. Mann, and R. V. Parsey, "Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration," *Neuroimage*, vol. 46, pp. 786–802, Jul 2009.

[59] N. Chitphakdithai and J. S. Duncan, "Pairwise registration of images with missing correspondences due to resection," in *2010 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pp. 1025–1028, IEEE, Apr. 2010.

[60] N. Chitphakdithai, V. L. Chiang, and J. S. Duncan, "Non-rigid Registration of Longitudinal Brain Tumor Treatment MRI," *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, vol. 2011, p. 4893, 2011.

[61] Y. Lao, V. Yu, E. Chang, W. Yang, and K. Sheng, "Deformable alignment of longitudinal postoperative brain GBM scans using deep learning," in *Proceedings Volume 11313, Medical Imaging 2020: Image Processing*, vol. 11313, pp. 133–138, SPIE, Mar. 2020.

[62] T. Estienne, M. Lerousseau, M. Vakalopoulou, E. Alvarez Andres, E. Battistella, A. Carré, S. Chandra, S. Christodoulidis, M. Sahasrabudhe, R. Sun, C. Robert, H. Talbot, N. Paragios, and E. Deutsch, "Deep Learning-Based Concurrent Brain Registration and Tumor Segmentation," *Front. Comput. Neurosci.*, vol. 0, 2020.

[63] B. H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby, Y. Burren, N. Porz, J. Slotboom, R. Wiest, L. Lanczi, E. Gerstner, M. Weber, T. Arbel, B. B. Avants, N. Ayache, P. Buendia, D. L. Collins, N. Cordier, J. J. Corso, A. Criminisi, T. Das, H. Delingette, C. Demiralp, C. R. Durst, M. Dojat, S. Doyle, J. Festa, F. Forbes, E. Geremia, B. Glocker, P. Golland, X. Guo, A. Hamamci, K. M. Iftekharuddin, R. Jena, N. M. John, E. Konukoglu, D. Lashkari, J. A. Mariz, R. Meier, S. Pereira, D. Precup, S. J. Price, T. R. Raviv, S. M. S. Reza, M. Ryan, D. Sarikaya, L. Schwartz, H. Shin, J. Shotton, C. A. Silva, N. Sousa, N. K. Subbanna, G. Szekely, T. J. Taylor, O. M. Thomas, N. J. Tustison, G. Unal, F. Vasseur, M. Wintermark, D. H. Ye, L. Zhao, B. Zhao, D. Zikic, M. Prastawa, M. Reyes, and K. Van Leemput, "The multimodal brain tumor image segmentation benchmark (brats)," *IEEE Transactions on Medical Imaging*, vol. 34, pp. 1993–2024, Oct 2015.

[64] K. J. Paprottka, S. Kleiner, C. Preibisch, F. Kofler, F. Schmidt-Graf, C. Delbridge, D. Bernhardt, S. E. Combs, J. Gempt, B. Meyer, C. Zimmer, B. H. Menze, I. Yakushev, J. S. Kirschke, and B. Wiestler, "Fully automated analysis combining [18F]-FET-PET and multiparametric MRI including DSC perfusion and APTw imaging: a promising tool for objective evaluation of glioma progression," *Eur. J. Nucl. Med. Mol. Imaging*, vol. 48, pp. 4445–4455, Dec 2021.

[65] Y.-X. Zhao and et al, "Multi-view semi-supervised 3d whole brain segmentation with a self-ensemble network," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 256–265, Springer, 2019.

[66] X. Feng, N. Tustison, and C. Meyer, "Brain tumor segmentation using an ensemble of 3d u-nets and overall survival prediction using radiomic features," in *International MICCAI Brainlesion Workshop* (A. Crimi, S. Bakas, H. Kuijf,

F. Keyvan, M. Reyes, and T. van Walsum, eds.), (Cham), pp. 279–288, Springer, 2019.

[67] F. Isensee, P. Kickingereder, W. Wick, M. Bendszus, and K. H. Maier-Hein, "No New-Net," in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, pp. 234–244, Cham, Switzerland: Springer, Jan. 2019.

[68] R. McKinley, R. Meier, and R. Wiest, "Ensembles of densely-connected cnns with label-uncertainty for brain tumor segmentation," in *International MICCAI Brainlesion Workshop* (A. Crimi, S. Bakas, H. Kuijf, F. Keyvan, M. Reyes, and T. van Walsum, eds.), (Cham), pp. 456–465, Springer, 2019.

[69] R. McKinley, M. Rebsamen, R. Meier, and R. Wiest, "Triplanar ensemble of 3d-to-2d cnns with label-uncertainty for brain tumor segmentation," in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries* (A. Crimi and S. Bakas, eds.), (Cham), pp. 379–387, Springer International Publishing, 2020.

[70] S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J. Kirby, J. Freymann, K. Farahani, and C. Davatzikos, "Segmentation labels and radiomic features for the pre-operative scans of the tcga-gbm collection. the cancer imaging archive (2017)," 2017.

[71] S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J. S. Kirby, J. B. Freymann, K. Farahani, and C. Davatzikos, "Advancing the cancer genome atlas glioma mri collections with expert segmentation labels and radiomic features," *Scientific data*, vol. 4, p. 170117, 2017.

[72] S. Bakas, M. Reyes, A. Jakab, S. Bauer, M. Rempfler, A. Crimi, R. T. Shinohara, C. Berger, S. M. Ha, M. Rozycki, *et al.*, "Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the brats challenge," *arXiv preprint arXiv:1811.02629*, 2018.

[73] F. Kofler, C. Berger, D. Waldmannstetter, J. Lipkova, I. Ezhov, G. Tetteh, J. Kirschke, C. Zimmer, B. Wiestler, and B. H. Menze, "Brats toolkit: Translating brats brain tumor segmentation algorithms into clinical and scientific practice," *Frontiers in Neuroscience*, vol. 14, 2020.

[74] F. Kofler, I. Ezhov, F. Isensee, F. Balsiger, C. Berger, M. Koerner, J. Paetzold, H. Li, S. Shit, R. McKinley, S. Bakas, C. Zimmer, D. Ankerst, J. Kirschke, B. Wiestler, and B. H. Menze, "Are we using appropriate segmentation metrics? identifying correlates of human expert perception for cnn training beyond rolling the dice coefficient," 2021.

[75] "BraTS 2015 - MICCAI BRATS 2017," Dec 2021. [Online; accessed 13. Dec. 2021 at `https://sites.google.com/site/braintumorsegmentation/home/brats2015`].

[76] L. Canalini, J. Klein, N. P. de Barros, D. M. Sima, D. Miller, and H. Hahn, "Comparison of different automatic solutions for resection cavity segmentation in postoperative MRI volumes including longitudinal acquisitions," in *Medical Imaging 2021: Image-Guided Procedures, Robotic Interventions, and Modeling*, vol. 11598, pp. 558–564, SPIE, Feb 2021.

[77] E. Haber and J. Modersitzki, "Intensity Gradient Based Registration and Fusion of Multi-modal Images," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2006*, pp. 726–733, Berlin, Germany: Springer, 2006.

[78] A. Hering, B. van Ginneken, and S. Heldmann, "mlVIRNET: Multilevel Variational Image Registration Network," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019*, pp. 257–265, Cham, Switzerland: Springer, Oct 2019.

[79] L. Canalini, J. Klein, D. Miller, and R. Kikinis, "Segmentation-based registration of ultrasound volumes for glioma resection in image-guided neurosurgery," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 14, pp. 1697–1713, Oct. 2019.

[80] D. C. Liu and J. Nocedal, "On the limited memory BFGS method for large scale optimization," *Math. Program.*, vol. 45, pp. 503–528, Aug. 1989.

[81] B. B. Avants, C. L. Epstein, M. Grossman, and J. C. Gee, "Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain," *Med. Image Anal.*, vol. 12, pp. 26–41, Feb. 2008.

[82] Antsx, "ANTsPy," July 2022. [Online; accessed 28. Jul. 2022].

[83] Y. Ou, H. Akbari, M. Bilello, X. Da, and C. Davatzikos, "Comparative Evaluation of Registration Algorithms in Different Brain Databases With Varying Difficulty: Results and Insights," *IEEE Trans. Med. Imaging*, vol. 33, pp. 2039–2065, June 2014.

[84] A. N. Pettitt and M. A. Stephens, "The Kolmogorov-Smirnov Goodness-of-Fit Statistic with Discrete and Grouped Data," *Technometrics*, pp. 205–210, Apr. 2012.

[85] A. Hering, S. Kuckertz, S. Heldmann, and M. P. Heinrich, "Memory-efficient 2.5D convolutional transformer networks for multi-modal deformable registration with weak label supervision applied to whole-heart CT and MRI scans," *Int. J. CARS*, vol. 14, pp. 1901–1912, Nov. 2019.

[86] L. Canalini, J. Klein, D. Waldmannstetter, F. Kofler, S. Cerri, A. Hering, S. Heldmann, S. Schlaeger, B. H. Menze, B. Wiestler, J. Kirschke, and H. K. Hahn, "Quantitative evaluation of the influence of multiple MRI sequences and of pathological tissues on the registration of longitudinal data acquired during brain tumor treatment," *Front. Neuroimaging*, vol. 1, Sept. 2022.

[87] S. Kabus, A. Franz, and B. Fischer, "On Elastic Image Registration with Varying Material Parameters," in *Bildverarbeitung für die Medizin 2005*, pp. 330–334, Berlin, Germany: Springer, 2005.

[88] S. Kabus, B. Fischer, and A. Franz, "Variational Image Registration Allowing for Discontinuities in the Displacement Field," in *Image Processing Based on Partial Differential Equations*, pp. 363–377, Berlin, Germany: Springer, 2007.

[89] S. Kabus, A. Franz, and B. Fischer, "Variational Image Registration with Local Properties," in *Biomedical Image Registration*, pp. 92–100, Berlin, Germany: Springer, 2006.

[90] A. Hering, S. Kuckertz, S. Heldmann, and M. P. Heinrich, "Enhancing Label-Driven Deep Deformable Image Registration with Local Distance Metrics for State-of-the-Art Cardiac Motion Tracking," in *Bildverarbeitung für die Medizin 2019*, pp. 309–314, Wiesbaden, Germany: Springer Vieweg, Feb. 2019.

[91] R. Stefanescu, X. Pennec, and N. Ayache, "Grid Enabled Non-Rigid Registration with a Dense Transformation and A Priori Information," *Springer*, vol. 2879, pp. 804–811, 2003.

[92] R. Stefanescu, O. Commowick, G. Malandain, P.-Y. Bondiau, N. Ayache, and X. Pennec, "Non-rigid Atlas to Subject Registration with Pathologies for Conformal Brain Radiotherapy," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2004*, pp. 704–711, Berlin, Germany: Springer, 2004.

[93] T. C. W. Mok and A. C. S. Chung, "Large Deformation Diffeomorphic Image Registration with Laplacian Pyramid Networks," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, pp. 211–221, Cham, Switzerland: Springer, Sept. 2020.

[94] M. Meng, L. Bi, D. Feng, and J. Kim, "Non-iterative Coarse-to-Fine Registration Based on Single-Pass Deep Cumulative Learning," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2022*, pp. 88–97, Cham, Switzerland: Springer, Sept. 2022.

[95] B. Fischer and J. Modersitzki, "Ill-posed medicine—an introduction to image registration," *Inverse Prob.*, vol. 24, p. 034008, May 2008.

[96] A. D. Leow, I. Yanovsky, M.-C. Chiang, A. D. Lee, A. D. Klunder, A. Lu, J. T. Becker, S. W. Davis, A. W. Toga, and P. M. Thompson, "Statistical properties of Jacobian maps and the realization of unbiased large-deformation nonlinear image registration," *IEEE Trans. Med. Imaging*, vol. 26, pp. 822–832, June 2007.

[97] A. Hering, L. Hansen, T. C. W. Mok, A. C. S. Chung, H. Siebert, S. Häger, A. Lange, S. Kuckertz, S. Heldmann, W. Shao, S. Vesal, M. Rusu, G. Sonn, T. Estienne, M. Vakalopoulou, L. Han, Y. Huang, P.-T. Yap, M. Brudfors, Y. Balbastre, S. Joutard, M. Modat, G. Lifshitz, D. Raviv, J. Lv, Q. Li, V. Jaouen, D. Visvikis, C. Fourcade, M. Rubeaux, W. Pan, Z. Xu, B. Jian, F. De Benetti, M. Wodzinski, N. Gunnarsson, J. Sjölund, D. Grzech, H. Qiu, Z. Li, A. Thorley, J. Duan, C. Großbröhmer, A. Hoopes, I. Reinertsen, Y. Xiao, B. Landman, Y. Huo, K. Murphy, N. Lessmann, B. Van Ginneken, A. V. Dalca, and M. P. Heinrich, "Learn2Reg: comprehensive multi-task medical image registration challenge, dataset and evaluation in the era of deep learning," *IEEE Trans. Med. Imaging*, p. 1, Oct. 2022.

[98] M. Niyazi, M. Brada, A. J. Chalmers, S. E. Combs, S. C. Erridge,

A. Fiorentino, A. L. Grosu, F. J. Lagerwaard, G. Minniti, R.-O. Mirimanoff, U. Ricardi, S. C. Short, D. C. Weber, and C. Belka, "ESTRO-ACROP guideline "target delineation of glioblastomas"," *Radiother. Oncol.*, vol. 118, pp. 35–42, Jan. 2016.

[99] S. Mehta, S. R. Gajjar, K. R. Padgett, D. Asher, R. Stoyanova, J. C. Ford, and E. A. Mellon, "Daily Tracking of Glioblastoma Resection Cavity, Cerebral Edema, and Tumor Volume with MRI-Guided Radiation Therapy," *Cureus*, vol. 10, p. e2346., Mar. 2018.

[100] E. Ermiş, A. Jungo, R. Poel, M. Blatti-Moreno, R. Meier, U. Knecht, D. M. Aebersold, M. K. Fix, P. Manser, M. Reyes, and E. Herrmann, "Fully automated brain resection cavity delineation for radiation target volume definition in glioblastoma patients using deep learning," *Radiat. Oncol.*, vol. 15, pp. 1–10, Dec. 2020.

[101] B. H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby, Y. Burren, N. Porz, J. Slotboom, R. Wiest, L. Lanczi, E. Gerstner, M.-A. Weber, T. Arbel, B. B. Avants, N. Ayache, P. Buendia, D. L. Collins, N. Cordier, J. J. Corso, A. Criminisi, T. Das, H. Delingette, Ç. Demiralp, C. R. Durst, M. Dojat, S. Doyle, J. Festa, F. Forbes, E. Geremia, B. Glocker, P. Golland, X. Guo, A. Hamamci, K. M. Iftekharuddin, R. Jena, N. M. John, E. Konukoglu, D. Lashkari, J. A. Mariz, R. Meier, S. Pereira, D. Precup, S. J. Price, T. R. Raviv, S. M. S. Reza, M. Ryan, D. Sarikaya, L. Schwartz, H.-C. Shin, J. Shotton, C. A. Silva, N. Sousa, N. K. Subbanna, G. Szekely, T. J. Taylor, O. M. Thomas, N. J. Tustison, G. Unal, F. Vasseur, M. Wintermark, D. H. Ye, L. Zhao, B. Zhao, D. Zikic, M. Prastawa, M. Reyes, and K. Van Leemput, "The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)," *IEEE Trans. Med. Imaging*, vol. 34, pp. 1993–2024, Dec. 2014.

[102] S. S. M. Salehi, D. Erdogmus, and A. Gholipour, "Tversky Loss Function for Image Segmentation Using 3D Fully Convolutional Deep Networks," in *Machine Learning in Medical Imaging*, pp. 379–387, Cham, Switzerland: Springer, Sept. 2017.

[103] L. R. Dice, "Measures of the Amount of Ecologic Association Between Species on JSTOR," *Ecology*, vol. 26, pp. 297–302, July 1945.

[104] I. Reinertsen, D. L. Collins, and S. Drouin, "The Essential Role of Open Data and Software for the Future of Ultrasound-Based Neuronavigation," *Front. Oncol.*, vol. 0, 2021.

[105] P. Shetty, U. Yeole, V. Singh, and A. Moiyadi, "Navigated ultrasound-based image guidance during resection of gliomas: practical utility in intraoperative decision-making and outcomes," *Neurosurg. Focus*, vol. 50, p. E14, Jan. 2021.

[106] Y. Xiao, M. Fortin, G. Unsgård, H. Rivaz, and I. Reinertsen, "REtroSpective Evaluation of Cerebral Tumors (RESECT): A clinical database of pre-operative MRI and intra-operative ultrasound in low-grade glioma surgeries.," *Med. Phys.*, vol. 44, pp. 3875–3882, May 2017.

[107] O. M. Rygh, T. Selbekk, S. H. Torp, S. Lydersen, T. A. N. Hernes, and G. Unsgaard, "Comparison of navigated 3D ultrasound findings with histopathology in subsequent phases of glioblastoma resection," *Acta Neurochir.*, vol. 150, pp. 1033–1042, Sept. 2008.

[108] H. Rivaz and D. L. Collins, "Near Real-Time Robust Non-rigid Registration of Volumetric Ultrasound Images for Neurosurgery," *Ultrasound Med. Biol.*, vol. 41, pp. 574–587, Feb. 2015.

[109] L. Mercier, R. Del Maestro, K. Petrecca, D. Araujo, C. Haegelen, and D. Collins, "Online database of clinical MR and ultrasound images of brain tumors.," *undefined*, 2012.

[110] I. Reinertsen, D. L. Collins, and S. Drouin, "The Essential Role of Open Data and Software for the Future of Ultrasound-Based Neuronavigation," *Front. Oncol.*, vol. 10, Feb. 2021.

[111] A. Pirhadi, H. Rivaz, M. O. Ahmad, and Y. Xiao, "Robust Ultrasound-to-Ultrasound Registration for Intra-operative Brain Shift Correction with a Siamese Neural Network," in *Simplifying Medical Ultrasound*, pp. 85–95, Cham, Switzerland: Springer, Sept. 2021.

[112] A. Pirhadi, S. Salari, M. O. Ahmad, H. Rivaz, and Y. Xiao, "Robust landmark-based brain shift correction with a Siamese neural network in ultrasound-guided brain tumor resection," *Int. J. CARS*, pp. 1–8, Oct. 2022.

[113] M. Wodzinski and A. Skalski, "Adversarial Affine Registration for Real-Time Intraoperative Registration of 3-D US-US for Brain Shift Correction," in *Simplifying Medical Ultrasound*, pp. 75–84, Cham, Switzerland: Springer, Sept. 2021.

[114] A. K. Petridis, M. Anokhin, J. Vavruska, M. Mahvash, and M. Scholz, "The value of intraoperative sonography in low grade glioma surgery," *Clin. Neurol. Neurosurg.*, vol. 131, pp. 64–8, Apr. 2015.

[115] N. F. Marko, R. J. Weil, J. L. Schroeder, F. F. Lang, D. Suki, and R. E. Sawaya, "Extent of resection of glioblastoma revisited: personalized survival modeling facilitates more accurate survival prediction and supports a maximum-safe-resection approach to surgery," *J. Clin. Oncol.*, vol. 32, pp. 774–782, Mar. 2014.

[116] T. J. Brown, M. C. Brennan, M. Li, E. W. Church, N. J. Brandmeir, K. L. Rakszawski, A. S. Patel, E. B. Rizk, D. Suki, R. Sawaya, and M. Glantz, "Association of the Extent of Resection With Survival in Glioblastoma: A Systematic Review and Meta-analysis," *JAMA Oncol.*, vol. 2, pp. 1460–1469, Nov. 2016.

[117] J. P. Almeida, K. L. Chaichana, J. Rincon-Torroella, and A. Quinones-Hinojosa, "The value of extent of resection of glioblastomas: clinical evidence and current approach," *Curr. Neurol. Neurosci. Rep.*, vol. 15, no. 2, p. 517., 2015.

[118] K. P. L., T. M. K. J., S. O. E., T. L.-P. M. P., V. O. J. J., and H. Van Santbrink, "Intraoperative MRI-guided resection of glioblastoma multiforme: a systematic review.," *Lancet Oncol.*, vol. 12, pp. 1062–1070, Aug. 2011.

[119] C. Nimsky, O. Ganslandt, B. von Keller, J. Romstöck, and R. Fahlbusch, "Intraoperative High-Field-Strength MR Imaging: Implementation and Experience in 200 Patients1," *Radiology*, Oct. 2004.

[120] S. Mittal and P. M. Black, "Intraoperative magnetic resonance imaging in neurosurgery: the Brigham concept," in *Medical Technologies in Neurosurgery*, pp. 77–86, Wien, Austria: Springer, Vienna, 2006.

[121] G. Unsgaard, O. M. Rygh, T. Selbekk, T. B. Müller, F. Kolstad, F. Lindseth, and T. A. N. Hernes, "Intra-operative 3D ultrasound in neurosurgery," *Acta Neurochir.*, vol. 148, pp. 235–253, Mar. 2006.

[122] G. Unsgaard, A. Gronningsaeter, S. Ommedal, and T. A. Nagelhus Hernes, "Brain Operations Guided by Real-time Two-dimensional Ultrasound: New Possibilities as a Result of Improved Image Quality," *Neurosurgery*, vol. 51, p. 402, Aug. 2002.

[123] G. Unsgaard, S. Ommedal, T. Muller, A. Gronningsaeter, and T. A. N. Hernes, "Neuronavigation by intraoperative three-dimensional ultrasound: initial experience during brain tumor resection," *Neurosurgery*, vol. 50, pp. 804–812, Apr. 2002.

[124] B. Saß, D. Zivkovic, M. Pojskic, C. Nimsky, and M. H. A. Bopp, "Navigated Intraoperative 3D Ultrasound in Glioblastoma Surgery: Analysis of Imaging Features and Impact on Extent of Resection," *Front. Neurosci.*, vol. 16, 2022.

[125] P. D. Leroux, T. C. Winter, M. S. Berger, L. A. Mack, K. Wang, and J. P. Elliott, "A comparison between preoperative magnetic resonance and intraoperative ultrasound tumor volumes and margins," *J. Clin. Ultrasound*, vol. 22, pp. 29–36, Jan. 1994.

[126] O. M. Rygh, T. Selbekk, S. H. Torp, S. Lydersen, T. A. N. Hernes, and G. Unsgaard, "Comparison of navigated 3D ultrasound findings with histopathology in subsequent phases of glioblastoma resection," *Acta Neurochir.*, vol. 150, no. 10, pp. 1033–1042, 2008.

[127] H. Rivaz and D. L. Collins, "Deformable registration of preoperative MR, pre-resection ultrasound, and post-resection ultrasound images of neurosurgery," *Int. J. CARS*, vol. 10, pp. 1017–1028, July 2015.

[128] M. M. J. Letteboer, M. A. Viergever, and W. J. Niessen, "Rigid registration of 3D ultrasound data of brain tumours," *International Congress Series*, vol. 1256, pp. 433–439, June 2003.

[129] M. M. J. Letteboer, P. W. A. Willems, M. A. Viergever, and W. J. Niessen, "Non-rigid Registration of 3D Ultrasound Images of Brain Tumours Acquired

during Neurosurgery," in *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2003*, pp. 408–415, Berlin, Germany: Springer, 2003.

[130] L. Mercier, D. Araujo, C. Haegelen, R. F. Del Maestro, K. Petrecca, and D. L. Collins, "Registering Pre- and Postresection 3-Dimensional Ultrasound for Improved Visualization of Residual Brain Tumor," *Ultrasound Med. Biol.*, vol. 39, pp. 16–29, Jan. 2013.

[131] M. Toews and I. William M. Wells, "Efficient and Robust Model-to-Image Alignment using 3D Scale-Invariant Features," *Med. Image Anal.*, vol. 17, p. 271, Apr. 2013.

[132] J. Nitsch, J. Klein, P. Dammann, K. Wrede, O. Gembruch, J. H. Moltz, H. Meine, U. Sure, R. Kikinis, and D. Miller, "Automatic and efficient MRI-US segmentations for improving intraoperative image fusion in image-guided neurosurgery," *NeuroImage: Clinical*, vol. 22, p. 101766, Jan. 2019.

[133] J. Rackerseder, M. Baust, R. Göbl, N. Navab, and C. Hennersperger, "Initialize Globally Before Acting Locally: Enabling Landmark-Free 3D US to MRI Registration," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018: 21st International Conference, Granada, Spain, September 16-20, 2018, Proceedings, Part I*, pp. 827–835, Berlin, Germany: Springer-Verlag, Sept. 2018.

[134] J. Rackerseder, R. Göbl, N. Navab, and C. Hennersperger, "Fully Automatic Segmentation of 3D Brain Ultrasound: Learning from Coarse Annotations," *arXiv*, Apr. 2019.

[135] J. Nitsch, J. Klein, J. H. Moltz, D. Miller, U. Sure, R. Kikinis, and H. Meine, "Neural-network-based automatic segmentation of cerebral ultrasound images for improving image-guided neurosurgery," in *Proceedings Volume 10951, Medical Imaging 2019: Image-Guided Procedures, Robotic Interventions, and Modeling*, vol. 10951, pp. 418–424, SPIE, Mar. 2019.

[136] C. Trantakis, M. Tittgemeyer, J.-P. Schneider, D. Lindner, D. Winkler, G. Strauß, and J. Meixensberger, "Investigation of time-dependency of intracranial brain shift and its relation to the extent of tumor removal using intra-operative MRI," *Neurol. Res.*, vol. 25, pp. 9–12, Jan. 2003.

[137] G. Unsgård and F. Lindseth, "3D ultrasound-guided resection of low-grade gliomas: principles and clinical examples," *Neurosurg. Focus*, vol. 47, p. E9., Dec. 2019.

[138] N. F. Marko, R. J. Weil, J. L. Schroeder, F. F. Lang, D. Suki, and R. E. Sawaya, "Extent of Resection of Glioblastoma Revisited: Personalized Survival Modeling Facilitates More Accurate Survival Prediction and Supports a Maximum-Safe-Resection Approach to Surgery," *J. Clin. Oncol.*, vol. 32, pp. 774–782, Feb. 2014.

[139] M. D. Timothy J. Brown, "Association of the Extent of Resection With Survival in Glioblastoma: A Systematic Review and Meta-analysis," *JAMA Oncol.*, vol. 2, pp. 1460–1469, Nov. 2016.

[140] J. P. Almeida, K. L. Chaichana, J. Rincon-Torroella, and A. Quinones-Hinojosa, "The Value of Extent of Resection of Glioblastomas: Clinical Evidence and Current Approach," *Curr. Neurol. Neurosci. Rep.*, vol. 15, pp. 1–13, Feb. 2015.

[141] P. L. Kubben, K. J. ter Meulen, O. E. Schijns, M. P. ter Laak-Poort, J. J. van Overbeeke, and H. v. Santbrink, "Intraoperative MRI-guided resection of glioblastoma multiforme: a systematic review," *Lancet Oncol.*, vol. 12, pp. 1062–1070, Oct. 2011.

[142] A. S. Jakola, G. Unsgård, K. S. Myrmel, R. Kloster, S. H. Torp, L. M. Sagberg, S. Lindal, and O. Solheim, "Surgical strategies in low-grade gliomas and implications for long-term quality of life," *Journal of Clinical Neuroscience*, vol. 21, pp. 1304–1309, Aug. 2014.

[143] L. Mercier, R. F. Del Maestro, K. Petrecca, D. Araujo, C. Haegelen, and D. L. Collins, "Online database of clinical MR and ultrasound images of brain tumors," *Med. Phys.*, vol. 39, pp. 3253–3261, June 2012.

[144] P. Risholm, E. Samset, I.-F. Talos, and W. Wells, "A Non-rigid Registration Framework That Accommodates Resection and Retraction," in *Information Processing in Medical Imaging*, pp. 447–458, Berlin, Germany: Springer, 2009.

[145] L. Canalini, J. Klein, D. Miller, and R. Kikinis, "Registration of Ultrasound Volumes Based on Euclidean Distance Transform," in *Large-Scale Annotation*

of Biomedical Data and Expert Label Synthesis and Hardware Aware Learning for Medical Imaging and Computer Assisted Intervention, pp. 127–135, Cham, Switzerland: Springer, Oct. 2019.

[146] F.-X. Carton, M. Chabanas, F. Le Lann, and J. H. Noble, "Automatic segmentation of brain tumor resections in intraoperative ultrasound images using U-Net," in *Journal of Medical Imaging, Vol. 7, Issue 3*, vol. 7, p. 031503, SPIE, Feb. 2020.

[147] F.-X. Carton, J. H. Noble, and M. Chabanas, "Automatic segmentation of brain tumor resections in intraoperative ultrasound images," in *Proceedings Volume 10951, Medical Imaging 2019: Image-Guided Procedures, Robotic Interventions, and Modeling*, vol. 10951, pp. 211–217, SPIE, Mar. 2019.

[148] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, "Convolutional neural networks: an overview and application in radiology," *Insights Imaging*, vol. 9, pp. 611–629, Aug. 2018.

[149] S. S. M. Salehi, D. Erdogmus, and A. Gholipour, "Tversky loss function for image segmentation using 3D fully convolutional deep networks," *arXiv*, June 2017.

[150] D. Drobny, M. Ranzini, S. Ourselin, T. Vercauteren, and M. Modat, "Landmark-Based Evaluation of a Block-Matching Registration Framework on the RESECT Pre- and Intra-operative Brain Image Data Set," in *Large-Scale Annotation of Biomedical Data and Expert Label Synthesis and Hardware Aware Learning for Medical Imaging and Computer Assisted Intervention: International Workshops, LABELS 2019, HAL-MICCAI 2019, and CuRIOUS 2019, Held in Conjunction with MICCAI 2019, Shenzhen, China, October 13 and 17, 2019, Proceedings*, pp. 136–144, Berlin, Germany: Springer-Verlag, Oct. 2019.

[151] T. C. Arnold, R. Muthukrishnan, A. R. Pattnaik, N. Sinha, A. Gibson, H. Gonzalez, S. R. Das, B. Litt, D. J. Englot, V. L. Morgan, K. A. Davis, and J. M. Stein, "Deep learning-based automated segmentation of resection cavities on postsurgical epilepsy MRI," *NeuroImage : Clinical*, vol. 36, 2022.

[152] B. Behboodi, F.-X. Carton, M. Chabanas, S. De Ribaupierre, O. Solheim, B. K. R. Munkvold, H. Rivaz, Y. Xiao, and I. Reinertsen, "RESECT-SEG:

Open access annotations of intra-operative brain tumor ultrasound images," *arXiv*, July 2022.

[153] M. Sharifzadeh, H. Benali, and H. Rivaz, "Segmentation of Intraoperative 3D Ultrasound Images Using a Pyramidal Blur-Pooled 2D U-Net," in *Lesion Segmentation in Surgical and Diagnostic Applications*, pp. 69–75, Cham, Switzerland: Springer, Mar. 2023.

[154] A. Weld, A. Agrawal, and S. Giannarou, "Ultrasound Segmentation Using a 2D UNet with Bayesian Volumetric Support," in *Lesion Segmentation in Surgical and Diagnostic Applications*, pp. 63–68, Cham, Switzerland: Springer, Mar. 2023.

[155] A. Qayyum, M. Mazher, S. Niederer, and I. Razzak, "Segmentation of Intra-operative Ultrasound Using Self-supervised Learning Based 3D-ResUnet Model with Deep Supervision," in *Lesion Segmentation in Surgical and Diagnostic Applications*, pp. 55–62, Cham, Switzerland: Springer, Mar. 2023.

Publications

The doctoral thesis is partially based on the following publications.

- **Canalini, L.**, Klein, J., Gerken, A., Heldmann, S., Hering, A., Hahn, H. K. (2023). Iterative Method to Register Longitudinal MRI Acquisitions in Neurosurgical Context. Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. Springer. doi: 10.1007/978-3-031-33842-7_23

  (3rd Place at the ISBI 2022 BraTS-Reg Challenge)

- **Canalini, L.**, Klein J., Waldmannstetter D., Kofler F., Cerri S., Hering A., Heldmann S., Schlaeger S., Menze B.H., Wiestler B., Kirschke J. and Hahn H.K. (2022) Quantitative evaluation of the influence of multiple MRI sequences and of pathological tissues on the registration of longitudinal data acquired during brain tumor treatment. Front. Neuroimaging 1:977491. doi: 10.3389/fnimg.2022.977491

- **Canalini, L.**, Klein, J., de Barros, N. P., Sima, D. M., Miller, D., Hahn, H. (2021). Comparison of different automatic solutions for resection cavity segmentation in post-operative MRI volumes including longitudinal acquisitions. Proceedings Volume 11598, Medical Imaging 2021: Image-Guided Procedures, Robotic Interventions, and Modeling. SPIE. doi: 10.1117/12.2580889

- **Canalini, L.**, Klein, J., Miller, D., Kikinis, R. (2020). Enhanced registration of ultrasound volumes by segmentation of resection cavity in neurosurgical procedures. Int. J. CARS, 15(12), 1963–1974. doi: 10.1007/s11548-020-02273-1

- **Canalini, L.**, Klein, J., Miller, D., Kikinis, R. (2019). Segmentation-based registration of ultrasound volumes for glioma resection in image-guided neurosurgery. Int. J. CARS, 14(10), 1697–1713. doi: 10.1007/s11548-019-02045-6

- **Canalini, L.**, Klein, J., Miller, D., Kikinis, R. (2019). Registration of Ultrasound Volumes Based on Euclidean Distance Transform. Large-Scale Annotation of Biomedical Data and Expert Label Synthesis and Hardware Aware Learning for Medical Imaging and Computer Assisted Intervention, (127-135), doi: 10.1007/978-3-030-33642-4_14

  (1^{st} Place at the MICCAI 2019 CuRIOUS Challenge: US-US Registration)

Author Contributions For each of the above publications, Luca Canalini was the main contributor to the development of the evaluated algorithms, the design and execution of the study, and the writing of the manuscript.

Appendix

Evaluation of the Accuracy of Image Registration Algorithms

An important aspect of image registration regards the way the results are analyzed and computed. In this context, the accuracy of the registration is a standard performance measure taken into consideration to evaluate registration algorithms in medical applications [97]. It can be measured using metrics that quantitatively express the improvement of the proposed solution w.r.t. a baseline value. Different metrics exist, but one of the most common includes anatomical landmarks. Pairs of landmarks are located in corresponding locations in the two images to register. Given that each point has its coordinates, the Euclidean distance between two points can be computed as:

$$D((x_1, y_1, z_1), (x_2, y_2, z_2)) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2} \tag{.1}$$

Given two volumetric images F and M, the minimum number of pairs of corresponding landmarks is three, but a larger quantity is preferred. For each pair p, the Euclidean distance can be computed (D_p). If the number of pairs of landmarks obtained in two images is N, the target registration error (TRE) is calculated as the mean of all the D_i computed for each pair acquired in two images F and M.

$$TRE = \frac{\sum_{i=1}^{N} D_i}{N} \tag{.2}$$

If landmarks are acquired in M pairs of images, a mean TRE (mTRE) can be computed:

$$mTRE = \frac{\sum_{i=1}^{M} TRE_i}{M} \tag{.3}$$