Technische Universität München
TUM School of Computation, Information and Technology

Medical Image Registration for Brain Tumor Surgery using Deep Learning

Diana Waldmannstetter

Vollständiger Abdruck der von der TUM School of Computation, Information and Technology der Technischen Universität München zur Erlangung einer

Doktorin der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitz:

Prof. Dr.-Ing. Werner Hemmert

Prüfende der Dissertation:

1. Prof. Dr. Bjoern H. Menze
2. Assistant Prof. Adrian V. Dalca, Ph.D.

Die Dissertation wurde am 19.12.2023 bei der Technischen Universität München eingereicht und durch die TUM School of Computation, Information and Technology am 26.09.2024 angenommen.

ABSTRACT

Image registration is vital in medical image analysis for integrating data from diverse imaging modalities, enabling precise temporal comparison, and facilitating accurate treatment planning and guidance. It addresses anatomical variability, aids in disease localization and quantification, and plays a crucial role in follow-up evaluation and disease monitoring.

This thesis addresses the challenge of medical image registration in general and especially in the context of brain tumor resection. Despite recent advances, aligning images before and after surgery remains complex due to tumor-induced deformations. The focus of this dissertation lies on the development of efficient image registration methods, particularly employing Deep Learning (DL) approaches.

Two core publications form the foundation of this work. The first approach enhances alignment by employing Deep Reinforcement Learning (DRL) for landmark redetection, incorporating patient-specific anatomical guidance. This methodology opens avenues for refining anatomical guidance and analyzing temporal changes before and after tumor resection, contributing to a more efficient model for understanding anatomical transformations. The second approach introduces the simultaneous optimization of multiple image similarity metrics in an unsupervised Convolutional Neural Network (CNN). This method demonstrates the potential benefits of combining different metrics during training, resulting in enhanced quantitative accuracy and visual registration quality.

Additionally, the thesis discusses challenges and recent advances in the registration of baseline and follow-up images of brain tumors and explores the value of suitable evaluation metrics, aiming to contribute towards a more clinically relevant assessment of registration errors.

ZUSAMMENFASSUNG

Die Bildregistrierung ist in der medizinischen Bildanalyse von entscheidender Bedeutung, um Daten aus verschiedenen bildgebenden Modalitäten zusammenzuführen, präzise Vergleiche zeitübergreifend zu ermöglichen und die Planung und Durchführung von Behandlungen zu unterstützen. Sie befasst sich mit anatomischer Variabilität, unterstützt die Lokalisierung und Quantifizierung von Krankheiten und spielt eine entscheidende Rolle bei der Nachsorge und Überwachung von Erkrankungen.

Diese Arbeit befasst sich mit der Herausforderung der medizinischen Bildregistrierung im Allgemeinen und insbesondere im Kontext der Resektion von Hirntumoren. Trotz jüngster Fortschritte zeigt sich die Registrierung von Bildern vor und nach der Operation aufgrund tumorbedingter Deformationen als komplexer Sachverhalt. Der Fokus dieser Dissertation liegt auf der Entwicklung effizienter Methoden der Bildregistrierung, insbesondere unter Verwendung von Deep Learning (DL) Ansätzen.

Zwei Kernveröffentlichungen bilden das Fundament dieser Arbeit. Der erste Beitrag verbessert die Bildregistrierung durch den Einsatz von Deep Reinforcement Learning (DRL) zur Redetektierung von Landmarken unter Einbeziehung patientenspezifischer Anatomie. Diese Methodik eröffnet Möglichkeiten zur Verfeinerung der anatomischen Anleitung und zur Analyse zeitlicher Veränderungen vor und nach der Tumorresektion, um ein effizienteres Modell zum Verständnis anatomischer Transformationen zu entwickeln. Im zweiten Beitrag wird die simultane Optimierung mehrerer Ähnlichkeitsmetriken in einem unüberwachten Convolutional Neural Network (CNN) vorgestellt. Diese Methode zeigt das Potenzial der kombinierten Verwendung verschiedener Metriken während des Trainings, was eine verbesserte quantitative Genauigkeit und visuelle Registrierungsqualität zur Folge hat.

Darüber hinaus diskutiert die Arbeit Herausforderungen und jüngste Fortschritte bei der Registrierung von Ausgangs- und Folgebildern von Hirntumoren und sondiert den Wert geeigneter Bewertungsmetriken mit dem Ziel, zu einer klinisch relevanten Bewertung von Registrierungsfehlern beizutragen.

PUBLICATIONS

This publication-based dissertation is based on the following two peer-reviewed first-author publications.

[1] Diana Waldmannstetter, Fernando Navarro, Benedikt Wiestler, Jan S Kirschke, Anjany Sekuboyina, Ester Molero, and Bjoern H Menze. "Reinforced redetection of landmark in pre-and post-operative brain scan using anatomical guidance for image alignment." In: *Biomedical Image Registration: 9th International Workshop, WBIR 2020, Portorož, Slovenia, December 1–2, 2020, Proceedings 9*. Springer. 2020, pp. 81–90.

[2] Diana Waldmannstetter, Benedikt Wiestler, Julian Schwarting, Ivan Ezhov, Marie Metz, Spyridon Bakas, Bhakti Baheti, Satrajit Chakrabarty, Daniel Rueckert, Jan S Kirschke, et al. "Primitive Simultaneous Optimization of Similarity Metrics for Image Registration." In: *International MICCAI Brainlesion Workshop*. Springer. 2023, pp. 57–68.

The following peer-reviewed papers have been published during the course of this dissertation, but are not included in this thesis.

[1] Luca Canalini, Jan Klein, Diana Waldmannstetter, Florian Kofler, Stefano Cerri, Alessa Hering, Stefan Heldmann, Sarah Schlaeger, Bjoern H Menze, Benedikt Wiestler, et al. "Quantitative evaluation of the influence of multiple MRI sequences and of pathological tissues on the registration of longitudinal data acquired during brain tumor treatment." In: *Frontiers in Neuroimaging* 1 (2022), p. 977491.

[2] Ivan Ezhov, Kevin Scibilia, Katharina Franitza, Felix Steinbauer, Suprosanna Shit, Lucas Zimmer, Jana Lipkova, Florian Kofler, Johannes C Paetzold, Luca Canalini, et al. "Learn-Morph-Infer: a new way of solving the inverse problem for brain tumor modeling." In: *Medical Image Analysis* 83 (2023), p. 102672.

[3] Xiaobin Hu, Rui Guo, Jieneng Chen, Hongwei Li, Diana Waldmannstetter, Yu Zhao, Biao Li, Kuangyu Shi, and Bjoern Menze. "Coarse-to-fine adversarial networks and zone-based uncertainty analysis for NK/T-cell lymphoma segmentation in CT/PET images." In: *IEEE journal of biomedical and health informatics* 24.9 (2020), pp. 2599–2608.

[4] Florian Kofler, Christoph Berger, Diana Waldmannstetter, Jana Lipkova, Ivan Ezhov, Giles Tetteh, Jan Kirschke, Claus Zimmer, Benedikt Wiestler, and Bjoern H Menze. "Brats toolkit: translating brats brain tumor segmentation algorithms into clinical and scientific practice." In: *Frontiers in neuroscience* (2020), p. 125.

[5] Fernando Navarro, Anjany Sekuboyina, Diana Waldmannstetter, Jan C Peeken, Stephanie E Combs, and Bjoern H Menze. "Deep reinforcement learning for organ localization in CT." In: *Medical Imaging with Deep Learning*. PMLR. 2020, pp. 544–554.

[6] Carolin M Pirk, Pedro A Gómez, Ilona Lipp, Guido Buonincontri, Miguel Molina-Romero, Anjany Sekuboyina, Diana Waldmannstetter, Jonathan Dannenberg, Sebastian Endt, Alberto Merola, et al. "Deep learning-based parameter mapping for joint relaxation and diffusion tensor MR Fingerprinting." In: *Medical Imaging with Deep Learning*. PMLR. 2020, pp. 638–654.

[7] Carolin M Pirkl, Matteo Cencini, Jan W Kurzawski, Diana Waldmannstetter, Hongwei Li, Anjany Sekuboyina, Sebastian Endt, Luca Peretti, Graziella Donatelli, Rosa Pasquariello, et al. "Learning residual motion correction for fast and robust 3D multiparametric MRI." In: *Medical Image Analysis* 77 (2022), p. 102387.

[8] Carolin M Pirkl, Ilona Lipp, Guido Buonincontri, Miguel Molina-Romero, Anjany Sekuboyina, Diana Waldmannstetter, Jonathan Dannenberg, Valentina Tomassini, Michela Tosetti, Derek K Jones, et al. "Deep learning-enabled diffusion tensor MR fingerprinting." In: *Proc. 27th Annu. Meet (ISMRM)*. 2019, p. 1102.

[9] Carolin Pirkl, Matteo Cencini, Jan W Kurzawski, Diana Waldmannstetter, Hongwei Li, Anjany Sekuboyina, Sebastian Endt, Luca Peretti, Graziella Donatelli, Rosa Pasquariello, et al. "Residual learning for 3D motion corrected quantitative MRI: Robust clinical T1, T2 and proton density mapping." In: *Medical Imaging with Deep Learning*. 2021.

[10] Yu Zhao, Yuan Liu, Yansheng Kan, Anjany Sekuboyina, Diana Waldmannstetter, Hongwei Li, Xiaobin Hu, Xiaozhi Zhao, Kuangyu Shi, and Bjoern Menze. "Spatial-frequency non-local convolutional lstm network for prcc classification." In: *Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part VI 22*. Springer. 2019, pp. 22–30.

MANUSCRIPTS

The following manuscripts have been written during the course of this doctoral thesis. Respective latest versions are provided in the appendix. They are related to the content of this dissertation, but are not part of the cumulative thesis.

[1] Bhakti Baheti, Diana Waldmannstetter, Satrajit Chakrabarty, Hamed Akbari, Michel Bilello, Benedikt Wiestler, Julian Schwarting, Evan Calabrese, Jeffrey Rudie, Syed Abidi, et al. "The brain tumor sequence registration challenge: establishing correspondence between pre-operative and follow-up MRI scans of diffuse glioma patients." In: *arXiv preprint arXiv:2112.06979* (2021).

[2] Diana Waldmannstetter, Benedikt Wiestler, Julian Schwarting, Ivan Ezhov, Marie Metz, Spyridon Bakas, Bhakti Baheti, Satrajit Chakrabarty, Jan S Kirschke, Rolf A Heckemann, et al. "Framing image registration as a landmark detection problem for better representation of clinical relevance." In: *arXiv preprint arXiv:2308.01318* (2023).

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to everyone who has contributed to the success of this thesis. I am more than grateful for all your support during the last few years.

First and foremost, I am deeply thankful to my supervisor, Prof. Dr. Bjoern Menze, who provided me with an outstanding research environment. Thank you for your guidance and expertise, all the inspiring and supporting discussions and the freedom to explore all the project ideas I was passionate about.

Special thanks go to PD Dr. Benedikt Wiestler, for all the insights into clinical routines, answering all my medical questions, at any time and always with a smile. Bene, I really appreciate your passion for this research and all your hard work. I also thank Prof. Dr. Jan Kirschke in his role as project team leader of the RadioTumor project, it was always a pleasure to work with you.

A huge thanks goes to all my colleagues and friends from the research groups in Munich and Zurich. From Amir, Anjany, Berkan, Bran, Cagdas, Carina, Carolin, Chinmay, Dhritiman, Fernando, Florian, Giles, Ibrahim, Ivan, Izabella, Jana, Johannes, John, Judith, Kaiyuan, Lucas, Michal, Oliver, Rami, Sandeep, Sebastian, Supro, Tamaz, Timo, Xiaobin, Yu to Yusuf. Thank you for the great atmosphere, your friendship and support. Also, I would like express my sincere acknowledgement to all my coauthors and collaborators all over the world for the great team work.

Further, I acknowledge my funding source, the Deutsche Forschungsgemeinschaft (DFG) through TUM International Graduate School of Science and Engineering (IGSSE), GSC 81.

Finally, and most importantly, to my dear family. For all your unconditional and unwavering love, support and trust. You have truly carried me through the last years. I cannot thank you enough.

CONTENTS

ACRONYMS

CNN	Convolutional Neural Network
CT	Computed Tomography
DQN	Deep Q-Network
DIR	Deep Image Registration
DL	Deep Learning
DRL	Deep Reinforcement Learning
DRN	Deformable Registration Network
DVF	Displacement Vector Field
MDP	Markov Decision Process
MRI	Magnetic Resonance Imaging
NN	Neural Network
RL	Reinforcement Learning
TRE	Target Registration Error

Part I

INTRODUCTION AND BACKGROUND

1 INTRODUCTION

Since the very first application of image registration in the medical field in 1895, when a needle was successfully removed from a hand by guiding the surgeon through aligning a radiograph's plate, some time has passed and significant development has been made in the field of medical image registration [49]. Starting with image-guided surgery using a stereotactic frame for aligning a patient's radiographic images to the actual anatomy [97, 114], image registration is nowadays an inherent part of the analysis and processing of images from medical imaging routines like Magnetic Resonance Imaging (MRI) or Computed Tomography (CT).

Such routines are a mandatory step for diverse purposes, including - but not limited to - the diagnoses and treatment of brain diseases and pathologies in neurosciences [122]. Due to the absence of ionizing radiation and the ability to picture the different soft tissue structures with high contrast, especially the MRI technique is widely used for the clarification of pathological processes like tumor growth, while recent algorithmic developments in terms of Deep Learning (DL) are getting more and more included into clinical practice [1]. Meanwhile, there has been developed a variety of different image registration methods and algorithms for various tasks in the medical field, including both optimization-based and DL-based approaches [14, 40, 92].

Particularly for the treatment of highly aggressive and invasive primary malignant brain tumors like gliomas, an early diagnosis is crucial [18]. The primary objective of brain tumor diagnosis is to acquire relevant clinical information about the location and type of a tumor. This information, derived from clinical imaging, guides and regulates subsequent interventions, ensuring accurate diagnosis and therapy. Typical treatment options for brain tumors include surgery, chemotherapy and radiotherapy, while tumor resection represents the "gold standard" treatment [19, 84].

For the planning of such brain tumor resection surgeries, neuroradiologists and neuro-surgeons are manually evaluating the produced MRI scans of the patients. Therefore, pre-operative scans are analyzed for surgery preparation and are then compared with intra-operative scans taken during the resection process for ensuring optimal resection results. A similar procedure is employed for assessing a surgery's outcome by evaluating post-operative against pre-operative images. Typically, attending clinicians perform this procedure manually, which is a difficult and time-consuming process. This form of

comparison involves the manual alignment of MR scans, which is then constituting a registration process.

CONTRIBUTIONS

The objective of this thesis is to assist clinicians in enhancing the registration process, specifically in the context of aligning pre- to post-operative and pre- to intra-operative images. This is achieved through the implementation of DL approaches and computer-aided algorithms for image analysis, aimed at optimizing and expediting the alignment procedures. Hence, the core of this thesis adds to the subsequent peer-reviewed publications:

1. **Reinforced Redetection of Landmark in pre- and post-operative Brain Scan using Anatomical Guidance for Image Alignment** [130]

   The proposed approach involves identifying landmarks (points of interest) in both pre- and post-operative images using Reinforcement Learning (RL). An artificial agent autonomously learns the optimal path to a target through feedback from the environment. Anatomical guidance is integrated, restricting the agent's search space to structures unaffected by surgery. By individually defining landmarks for each patient, patient-specific representations are ensured, enhancing image alignment.

2. **Primitive Simultaneous Optimization of Similarity Metrics for Image Registration** [132]

   An unsupervised Deformable Registration Network (DRN) is enhanced through the incorporation of a composite of image similarity metrics, concurrently optimized within the loss function. The approach's performance is assessed on two demanding datasets encompassing pre- to post-operative and pre- to intra-operative brain tumor images. The simultaneous optimization of similarity metrics is shown to improve registration accuracy and quality.

The task of re-identifying points of interest in pre- and post-operative images poses challenges attributed to anatomical changes resulting from tumor resection and tissue displacement. Conventional image registration techniques, particularly those relying on iterative optimization, frequently encounter limitations in proximity to the tumor. However, these specific locations are crucial for the evaluation of tumor progression. Given the already demonstrated adaptability of RL agents to specific environments, we leveraged this capability and expanded the approach by integrating anatomical guidance. This augmentation enables the model to navigate and account for the

pathologies and the altered tissue environment resulting from the presence of a tumor, see Chapter 7.

While the re-identification of landmarks can streamline the image registration process, it is particularly valuable for the initial alignment phase. This initial alignment is commonly succeeded by a deformable registration step to achieve optimal overlay of differing structures. Traditional deformable registration algorithms can take a lot of time to succeed. Integrating DL has proven to be highly beneficial in the development of Deep Image Registration (DIR) approaches, facilitating an accelerated registration process. Registration always involves measuring the similarity between the respective images in order to evaluate the degree of alignment that has already been achieved. Despite being widely used in the context of segmentation, the simultaneous optimization of similarity metrics is not as common in image registration. Hence, we investigate the concurrent optimization of registration metrics inside a DRN in order to improve the effectiveness of image registration. The potential of this method is demonstrated on pre- to post-operative and pre- to intra-operative brain tumor images, see Chapter 8.

ORGINIZATION

This thesis is organized as follows. In Chapter 1, a general introduction to the field of medical image registration is given, followed by the description of the main contributions of this dissertation. Chapter 2 gives an overview of the theoretical background of image registration, while the Chapters 3 and 4 explain the fundamental concepts of DL and DRL that are relevant to this thesis. In Chapter 5, some background on the handling of pathologies in medical image registration is provided. Chapter 6 concludes the first part with remarks on a new way of evaluating image registration. In the Chapters 7 and 8, two peer-reviewed publications are presented as the main contributions in this thesis. The dissertation is closed with a discussion of this work in Chapter 9, followed by concluding remarks and an outlook towards possible future research directions, in Chapter 10. Appendices A and B provide two manuscripts that are not subject to evaluation, but have been written within the scope of this dissertation and complement its content thematically.

2 IMAGE REGISTRATION

The application scope of image registration is extensive, not only in the medical field. This thesis focuses on the registration of MR images of patients afflicted with a brain tumor, placing particular emphasis on image scans affected by surgical interventions. Image registration is a crucial process in the field of image analysis and computer vision, involving the alignment and integration of multiple images to establish spatial correspondences between them. In the following section, the fundamental concepts of intensity-based image registration are described, encompassing several key principles:

- Spatial Transformation
- Similarity Metrics
- Optimization
- Diffeomorphism

Section 2.2 then gives an overview of the wide range of applications in the medical field.

2.1 FUNDAMENTAL CONCEPTS

The general idea behind registration is the transformation of different image data into a common space, specifically into the same coordinate system. One of the images is denoted as the moving or source image, while the other one is termed as the target or fixed image. Therefore, the source image is transformed to achieve alignment with the target image. We also denote this transformation as the *warping* of one image onto another. The reference frame within the target image is fixed, while transformation is applied to the source image for alignment with the target [48].

2.1.1 *Spatial Transformation*

Image registration involves the application of spatial transformations to align images. These transformations include translation, rotation, scaling, shearing and deformation, enabling the adjustment of images to a shared coordinate system. Transformations can be subdivided in two main groups: *linear* and *nonlinear (deformable)*. Linear transformations are further partitioned into *rigid* and *affine* alignment. A transformation is called *rigid* when only translations and rotations are used.

When the transformation maps parallel lines onto parallel lines, it is denoted as *affine* [80].

Linear Transformation

The individual linear transforms are specified as follows, expressed through matrix multiplication ($y = Mx$) with mapping m for the 3D case [6]:

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ 1 \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ 1 \end{bmatrix} \tag{1}$$

TRANSLATION Transforming an image point x using a translation by t quantities (see Fig. 1a):

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & t_1 \\ 0 & 1 & 0 & t_2 \\ 0 & 0 & 1 & t_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ 1 \end{bmatrix} \tag{2}$$

ROTATION The rotation transform is defined separately for each axis with respective radians θ (see Fig. 1b):

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta_x & \sin\theta_x & 0 \\ 0 & -\sin\theta_x & \cos\theta_x & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ 1 \end{bmatrix} \tag{3}$$

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\theta_y & 0 & \sin\theta_y & 0 \\ 0 & 1 & 0 & 0 \\ -\sin\theta_y & 0 & \cos\theta_y & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ 1 \end{bmatrix} \tag{4}$$

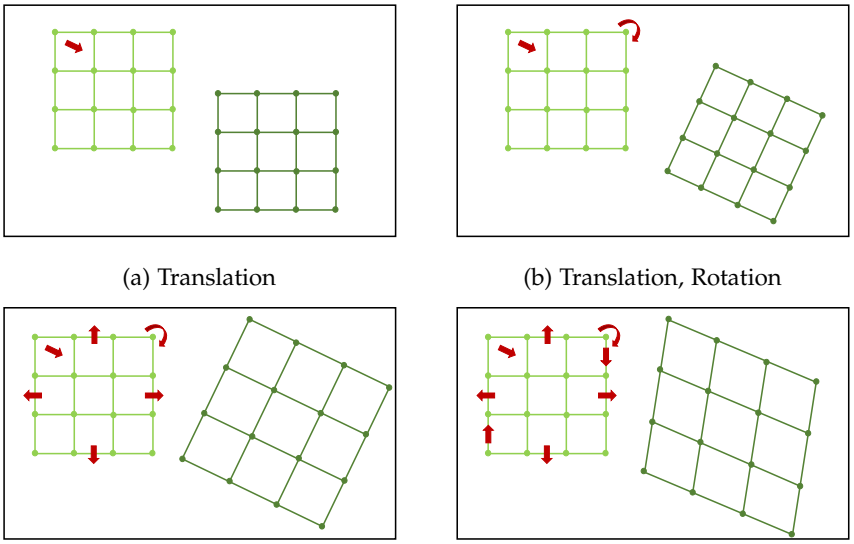
$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\theta_z & \sin\theta_z & 0 & 0 \\ -\sin\theta_z & \cos\theta_z & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ 1 \end{bmatrix} \tag{5}$$

SCALE Incorporating scaling into the transformation process extends the rigid mappings towards similarity transforms (see Fig. 1c):

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ 1 \end{bmatrix} = \begin{bmatrix} s_1 & 0 & 0 & 0 \\ 0 & s_2 & 0 & 0 \\ 0 & 0 & s_3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ 1 \end{bmatrix} \tag{6}$$

SHEAR The shearing transform allows for distorting the shape of an image by shifting its content along a particular direction, leading to parallel translation and a fully affine transform (see Fig. 1d):

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & s_1 & s_2 & 0 \\ 0 & 1 & s_3 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ 1 \end{bmatrix} \tag{7}$$



(a) Translation (b) Translation, Rotation

(c) Translation, Rotation, Scale (d) Translation, Rotation, Scale, Shear

Figure 1: Linear Transformations.

Nonlinear Transformation

Rigid and affine registration are commonly employed as a preliminary step for image pre-alignment before starting a nonlinear/deformable registration [92, 96]. This approach aims to reduce the search space in preparation for the subsequent application of deformable transformation, a process that may require a substantial number of parameters. Nonlinear or deformable image registration

involves the spatial alignment of images in a way that allows for local deformations, see Fig. 2. Unlike rigid or affine transformation, which involves only global transforms as described above, nonlinear transformation allows for more complex and localized adjustments. Within deformable registration, the transformation has the flexibility to vary throughout the image, enabling the representation of spatial variations or distortions [5]. This is especially beneficial in medical imaging, where tissues or organs may show non-uniform deformations due to factors like organ motion, alterations in shape or shifts in the patient's position [113].

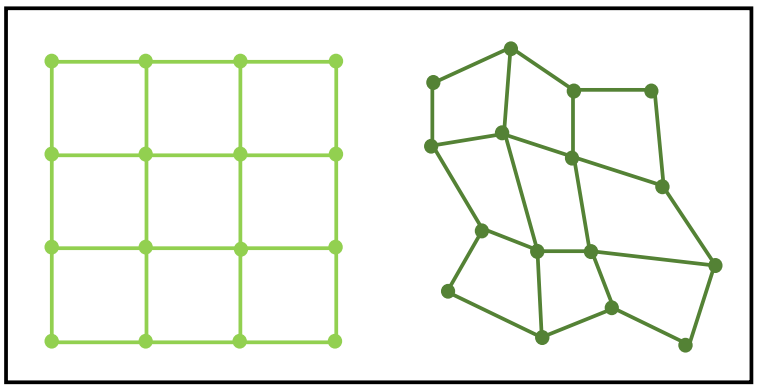


Figure 2: Nonlinear (deformable) transformation, allowing for highly flexible image adjustments.

Nonlinear deformations often involve a high degree of freedom, with a high amount of parameters enabling the transformation [39, 91]. Regularization techniques, such as adding a penalty term to the optimization objective, help in reducing the complexity of the model. Without regularization, deformable registration processes might result in discontinuities or abrupt changes in the deformation field. Such regularization techniques are specifically important in medical image registration settings, since otherwise the transformation might result in so-called *foldings* inside the deformed images, leading to an inconsistent topology of the transforming field. Regularization can therefore be helpful to enforce smoothness and to ensure that the deformation is changing gradually and consistently across the image, leading to physiologically reasonable image registration [24, 100].

2.1.2 *Similarity Metrics*

For evaluating the success of registration, the degree of similarity or dissimilarity between images needs to be measured. A variety of metrics is employed to assess the similarity between source and target image [98, 124]. Selecting an appropriate similarity metric depends on the characteristics of the images and the specific requirements of

the registration task. A crucial consideration in employing similarity measures are the modalities engaged in the registration process. When registering images within the same modality, *Normalized Cross Correlation (NCC)*, *Mean Squared Error (MSE)* and *Sum of Squared Differences (SSD)* are frequently utilized [124].

The *NCC* calculates the normalized cross-correlation between corresponding pixel/voxel intensities in the fixed and moving image. The result is a value between -1 and 1, with 1 indicating a perfect match:

$$NCC = \frac{\sum_{i,j}(I_{fixed}(i,j) \cdot I_{moving}(i,j))}{\sqrt{\sum_{i,j}(I_{fixed}(i,j))^2 \cdot \sum_{i,j}(I_{moving}(i,j))^2}}, \tag{8}$$

with $I_{fixed}(i,j)$ and $I_{moving}(i,j)$ representing the intensity value of the fixed and the moving image at position (i,j), respectively.

The *MSE* computes the average of the squared differences between corresponding pixel/voxel intensities in the fixed and moving images:

$$MSE = \frac{1}{N}\sum_{i,j}(I_{fixed}(i,j) - I_{moving}(i,j))^2, \tag{9}$$

with N denoting the total number of pixels in the images.

The *SSD* evaluates the squared differences between corresponding pixel/voxel intensities in the fixed and moving images and sums up these squared differences:

$$SSD = \sum_{i,j}(I_{fixed}(i,j) - I_{moving}(i,j))^2, \tag{10}$$

For the alignment of multi-modal images, *Mutual Information (MI)* [128] and *Normalized Mutual Information (NMI)* are widely adopted measures [124].

MI measures the amount of information shared between the intensities of the two images:

$$MI(X,Y) = \sum_{i=1}^{N}\sum_{j=1}^{M} p_{XY}(x_i,y_j) \cdot \log\left(\frac{p_{XY}(x_i,y_j)}{p_X(x_i)\cdot p_Y(y_j)}\right), \tag{11}$$

with the random variables X and Y, representing the intensities in the fixed and moving images, respectively. $p_{XY}(x_i,y_j)$ denotes the joint probability mass function of X and Y, while $p_X(x_i)$ and $p_Y(y_j)$ are the respective probability mass functions.

NMI normalizes *MI* by dividing it by the geometric mean of the entropies of X and Y:

$$NMI(X,Y) = \frac{MI(X,Y)}{\sqrt{H(X)\cdot H(Y)}}, \tag{12}$$

with $H(X)$ and $H(Y)$ denoting the entropy of X and Y, respectively.

The choice of a similarity metric is crucial in image registration, since it directly influences the accuracy and robustness of the registration process [25]. The similarity metric defines how well the registration algorithm measures the similarity or dissimilarity between corresponding points or regions in the images [98]. This involves to find a balance between sensitivity to relevant image features, robustness to noise as well as computational efficiency. Therefore, choosing an appropriate similarity metric requires careful consideration.

2.1.3 *Optimization*

In order to find the optimal parameters for the spatial transformation, an optimization process is involved. Specific algorithms are employed to minimize an objective function, i.e. the difference between the respective images, measured by the chosen similarity metric [112, 119]. The goal is to find the set of transformation parameters that either maximizes similarity or minimizes dissimilarity. A variety of different algorithms can be applied in order to efficiently search for the optimal set of parameters. Gradient-based methods, such as gradient descent, are commonly used [32, 138, 146]. These algorithms iteratively update the registration parameters in a way that reduces the objective function until a minimum is reached. Optimization can involve local or global search strategies. Local optimization methods focus on finding the best parameters close to the initialized parameters, while global optimization methods aim to explore a broader parameter space to find the global optimum [15, 74]. Balancing between local and global optimization, while circumventing potential entrapment in local minima or maxima, is of great importance and constitutes a major challenge.

2.1.4 *Diffeomorphism*

Image registration often involves the alignment of images that are obtained under varying conditions or at different time points. In order to guarantee a smooth deformation, *diffeomorphism* is introduced. Diffeomorphic transformations provide a bijective mapping between points in the source and the target image, where the mapping itself and its inverse should be differentiable. Therefore, diffeomorphic transforms have the ability to enforce a seamless and continuous deformation from one image to another. Mathematically, a diffeomorphic transformation $\phi_t(x, y)$ can be expressed through a vector field or flow field, describing how the image points are deformed over time. This time-varying velocity field $v_t(x, y)$ can then be defined as:

$$\frac{d\phi_t}{dt} = v_t(\phi_t(x, y)), \tag{13}$$

with time point t and image coordinates (x, y). Integrating this equation with respect to time, yields the diffeomorphic transformation:

$$\phi_t(x, y) = \int_0^t v_s(\phi_s(x, y))\, ds, \tag{14}$$

with the identity transform $\phi_0(y, x) = (x, y)$ at time $t = 0$.

Examples of such transformations encompass diffeomorphic B-spline transformations [27, 103] and diffeomorphic demons algorithms [126, 127].

2.2 MEDICAL IMAGE REGISTRATION

Image registration plays a crucial role in various medical applications, contributing to improved diagnosis, treatment planning and medical research. This includes, for example, applications related to multi-modal image fusion [77, 90, 95], image-guided interventions [82, 105, 143], brain mapping [122, 139] and neuro-surgery [35, 87], radiation therapy planning [65, 66, 102], monitoring disease progression [22, 47, 78], and cardiac image registration [45, 81, 109].

3

DEEP LEARNING FOR MEDICAL IMAGE REGISTRATION

Deep Learning (DL) has played a crucial role in advancing the field of medical image analysis, transforming the way healthcare practitioners interpret and leverage information from medical imaging data [76, 133]. DL approaches, especially in the form of Neural Networks (NNs) and CNNs offer advantages, such as end-to-end learning, the ability to handle complex deformations, and improved performance over traditional methods [52, 148]. The following sections briefly introduce the fundamental principles of DL and its applications to medical image registration.

3.1 NEURAL NETWORKS

(Artificial) NNs are inspired by the principles of neural networks in biology, especially by the interconnected neurons in brains. The key principles of NNs are described subsequently [33, 118].

NEURON *Neurons* are the basic units of a NN. Every neuron receives input, undergoes a computational process and generates an output. Thus, the output of an artificial neuron can formally be expressed as:

$$f(x) = \phi\left(\sum_n (w_n \cdot x_n) + b\right), \tag{15}$$

where ϕ denotes the activation function, w_n represents the weights associated with the inputs x_n and b is the bias term. A visualization is provided in Fig. 3.

WEIGHTS The *weights* of the network are parameters linked to the connections between neurons, determining the strengths of these connections.

BIAS *Biases* are parameters providing an additional degree of freedom, allowing the network to adjust the output even when the input is zero.

ACTIVATION FUNCTION The *activation function* transforms the linear combination of inputs into a non-linear output, determining whether a neuron is activated. Commonly employed activation functions are defined below [31, 108].

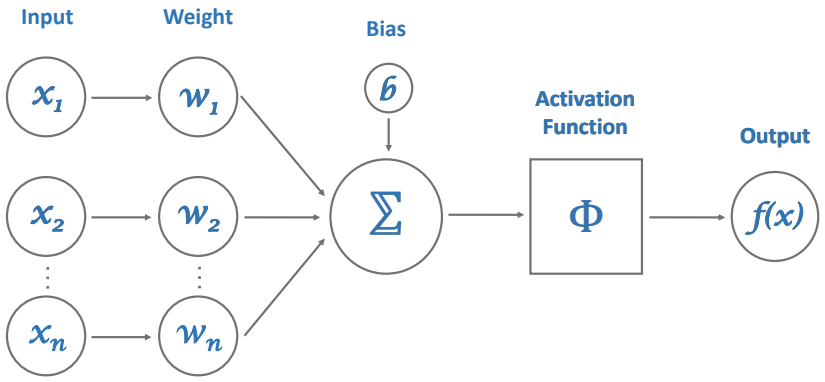


Figure 3: Schematic scheme of an artificial neuron, constituting the fundamental element of a Neural Network (NN). The computational process to generate an output $f(x)$ from n inputs involves the application of weights w_n, a bias b and an activation function ϕ.

The **Sigmoid** function is one of the most widely employed activation functions. It transforms values into the range $[0, 1]$ and is therefore often used for binary classification tasks.

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{16}$$

The **Hyperbolic Tangent** function is similar to the sigmoid function but with a wider output range that is symmetric to the origin $[-1, 1]$.

$$\tanh(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} \tag{17}$$

Compared to other activation functions, the **ReLU** [46] function – representing the *Rectified Linear Unit* function – performs more computationally efficient because not all neurons are activated simultaneously, but only a subset is activated at the same time.

$$\text{ReLU}(x) = \max(0, x) \tag{18}$$

The **Leaky ReLU** [141] function is similar to ReLU, but allows a small, non-zero gradient for negative inputs.

$$\text{Leaky ReLU}(x) = \max(\alpha x, x) \tag{19}$$

The **Softmax** function is combined of multiple sigmoid functions and converts raw output scores (logits) into probabilities. Hence, it is commonly used for multi-class classification problems.

$$\text{Softmax}(x_i) = \frac{e^{x_i}}{\sum_j e^{x_j}} \tag{20}$$

LAYERS An *input layer* receives the input data and forwards it to *hidden layers*, where the relationship between inputs and outputs is established and the computations take place. In the final *output layer*, the network's output is generated.

BACKPROPAGATION The *backpropagation* algorithm [104] is an optimization process that is iteratively adjusting a network's weights and biases in order to minimize the difference between the predicted output and the actual target values (ground truth) for a given set of training examples. By computing the gradient vector of the error of the cost function C, the error diminishes systematically until the anticipated outputs are achieved. The gradient descent algorithm is iteratively adjusting the model's parameters θ in the direction of the steepest decrease in the function [53]. This can formally be expressed as:

$$\theta^{i+1} = \theta^i - \eta \nabla C(\theta^i), \tag{21}$$

$$\theta_n^{i+1} = \theta_n^i - \eta \frac{\partial C}{\partial \theta_n^i}, \tag{22}$$

with η denoting the learning rate, determining the size of the steps taken during the current iteration i of the optimization process [104].

3.2 CONVOLUTIONAL NEURAL NETWORKS

Convolutional Neural Networks (CNNs) share a close resemblance to Neural Networks, as they consist of neurons with adaptable weights and biases. The fundamental distinction lies in the CNN architecture, which operates under the implicit assumption that the inputs resemble images. This assumption enables the encoding of specific properties within the architecture [120]. CNNs have achieved remarkable results in various fields associated with medical image analysis [4, 142, 144]. One of the most advantageous features of CNNs is their ability to significantly decrease the number of parameters in artificial NNs due to the inherent invariance of convolutions to translation. This allows for the development of larger models for addressing complex tasks [3, 120]. The main concepts of CNNs are briefly explained in this section subsequently [3, 110, 120]. A visualization of the basic structure is provided in Fig. 4.

CONVOLUTIONAL LAYER The fundamental component of CNNs is the *convolutional layer*, wherein convolution operations are applied to the input data. Convolutions entail the systematic movement of a small window, referred to as a kernel or filter, across the input data. This involves element-wise multiplication and summation, resulting in the production of the output.

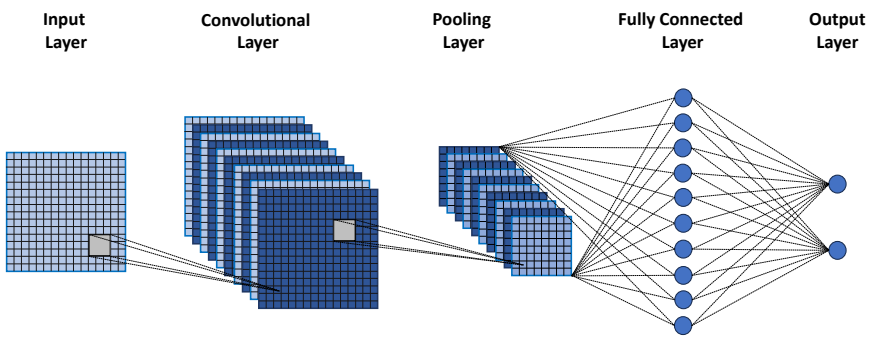


Figure 4: Basic structure of a Convolutional Neural Network (CNN). For generating an output from the received input image data, a CNN operates through the use of convolutional layers for hierarchical feature extraction, pooling layers for downsampling, and fully connected layers for learning the dependencies between preceding and succeeding layers.

KERNEL / FILTER A *kernel*, also referred to as *filter*, denotes a small matrix employed for the convolutional operation. It extracts particular features from the input data. Constituting a collection of learnable parameters, the kernel serves the purpose of facilitating the CNN in detecting features within the input. The convolution operation involves systematically traversing the kernel across the input data, performing element-wise multiplication with overlapping input values, and aggregating the results through summation to produce the output feature map.

STRIDE The parameter that specifies the step size by which the kernel moves between input iterations is known as the *stride* of the convolutional layer. This parameter dictates the extent of pixel displacement between consecutive applications of the kernel. A larger stride reduces the spatial dimensions of the output, while a smaller stride preserves more spatial information.

PADDING *Padding* encompasses the addition of supplementary pixels around the input data. This practice is employed to mitigate the rapid reduction of spatial dimensions during convolution. Involving padding techniques can help to avoid loosing information at the edges.

POOLING LAYER The main objective of a *pooling layer* is to decrease the spatial dimensions of the resulting feature map and involves down-sampling to decrease complexity for subsequent layers. Several pooling techniques are available, with max and average pooling being commonly used methods. Max pooling divides the image into rectangular sub-regions and outputs the maximum value within each

of those individual neighborhoods. In the case of average pooling, the layer outputs the average values.

FULLY CONNECTED LAYER The *fully connected layer* resembles the arrangement of neurons in a CNN. Consequently, each node in such a layer establishes direct connections with every node in both the preceding and succeeding layers.

OUTPUT LAYER / FEATURE MAP The final layer in a CNN generates the *output* of the network, also denoted as *feature map*, undergoing a non-linear activation function, as already described in Section 3.1.

3.3 APPLICATIONS TO MEDICAL IMAGE REGISTRATION

In the recent decades, multiple conventional iterative registration methods and toolboxes have been developed, such as *Optical Flow* [56], *Demons* [121], *Elastix* [68] and *Advanced Normalization Tools (ANTs)* [7]. Meanwhile, DL has gained significant attention in the field of medical image registration due to its ability to automatically detect complex image features and to solve intricate problems. DL models have therefore already been applied to various aspects of medical image registration [11, 14, 52, 123].

DL-based 3D medical image registration can be divided into different categories, depending on the type of method employed and the training regimen. Two main categories are *supervised registration* and *unsupervised registration* [140].

In *supervised registration*, a network is usually trained using some sort of *ground truth* for generating the transformation parameters [14]. Such ground truth can, for example, be provided in the form of a real or artificial reference Displacement Vector Field (DVF) for each training image sample [16, 41, 101, 111]. The error computed by the loss (cost) function is then the difference between the reference DVF and the predicted DVF. This kind of training scheme has several drawbacks. The ground truth DVFs have to be pre-produced by conventional registration algorithms, which can be very time consuming. Also, this might lead to a systematic bias introduced by the chosen algorithms. Another approach for supervised training is to incorporate ground truth labels for guiding the learning process, such as segmented organs [58, 59, 147] and corresponding landmark points [60]. The loss function then calculates the (distance) error between the reference and the predicted labels. Especially in the field of medical imaging, this can be challenging, since clinical experts need to be involved to generate those labels.

Hence, opting for *unsupervised registration* has strong benefits, as it involves inputting only pairs of source and target images and allows

the model to learn the deformation from them. Unsupervised models need to be trained using a loss function usually incorporating a similarity term for optimal image matching and a regularization term for ensuring smoothness in the deformation field. This kind of DL model has already been applied to a variety of tasks in medical imaging [9, 17, 26, 29, 36, 67, 71, 107].

4

DEEP REINFORCEMENT LEARNING FOR MEDICAL IMAGE REGISTRATION

Besides the already described methods for Deep Image Registration (DIR), another way to make the aligment of images learnable is Reinforcement Learning (RL) [116], especially Deep Reinforcement Learning (DRL) [88, 99], where RL is combined with DL techniques. Therefore, DRL extends RL by incorporating deep neural networks to handle complex and high-dimensional input spaces [51, 106]. In the following sections, a brief overview of the theoretical background behind RL and DRL is given and applications to the field of medical image registration are discussed.

4.1 REINFORCEMENT LEARNING

The fundamental components of RL can be described as follows [117]:

- **Agent:** The entity that is responsible for carrying out actions within a given environment.
- **Environment:** The external system or context with which the agent is interacting.
- **Actions:** The collection of all potential decisions or moves that the agent can execute inside the environment.
- **State:** The current condition or arrangement of the environment as observed by the agent.
- **Rewards:** Numeric value representing the immediate feedback towards the agent following a specific action in a particular state of the environment.
- **Policy:** The strategy for decision-making that the agent follows based on a collection of rules.

RL is a machine learning approach in which a so-called artificial agent interacts with its environment and learns to make decisions based on feedback to the interaction. Through rewards or penalties, the agent is then able to refine the policy progressively in order to optimize the accumulation of total rewards. This learning process involves exploration, decision-making, and continuous adaptation of the agent's policy based on the outcomes observed in the environment. A schematic visualization of this process is provided in Fig. 5. Basically, the agent needs to start with trial and error first, and

refines its strategy gradually following the observations when interacting with the environment. Therefore, RL is particularly useful in settings where only limited training data is available due to the principle of self-teaching.

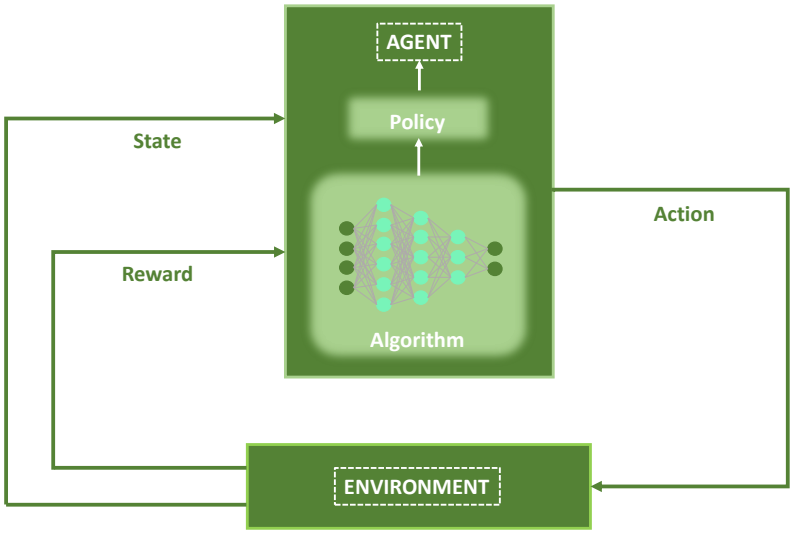


Figure 5: Schematic visualization of the basic structure of a Reinforcement Learning (RL) approach. An artificial agent learns an optimal policy by interacting with its environment and receiving rewards in the context of specific state-action situations.

4.2 MARKOV DECISION PROCESSES

RL is closely related to Markov Decision Processs (MDPs) [125]. MDPs are mathematical frameworks for representing decision-making problems in a sequential fashion [13]. Such a model refers to a stochastic process and is defined by a system of states, actions, transition probabilities and rewards. Essential for the model is the Markov property, which refers to the assumption that the future state depends solely on the current state and action and is independent of any preceding behaviour of the process.

In many RL tasks, where discrete actions are part of the scenario, MDPs are employed as models. In the beginning, the agent has no knowledge about the next state and the rules upon which the rewards are based. Therefore, the agent needs to explore various states in order to find the optimal strategy. The transition function T, representing the probability distribution of all possible future states, is then defined as

$$T : S \times A \times S \to [0, 1], \tag{23}$$

with action $a \in A$ and state $s \in S$, while the probability of transitioning to state s', subsequent to executing action a within state s, is

defined as $T(s, a, s')$ at a specific time step t and the reward function can then be denoted as $R(s, a, s')$ [125].

4.3 TYPES OF REINFORCEMENT LEARNING APPROACHES

There are two main approaches of RL: *model-based* and *model-free* [38]. Selecting an algorithm from one of these categories needs evaluation of the task's specific requirements, the inherent complexity of the environment, and the available computational resources [94].

In *model-based* RL, the agent needs to learn a model of its environment through interacting with it. Therefore, gaining knowledge about the surrounding system trough exploration and experience is essential for these kinds of algorithms. Based on the learned model, the agent can then develop the ultimate strategy for deciding on actions, allowing the agent to make information-based decisions without the need of actually performing an action.

In contrast, *model-free* RL algorithms do not need to build a predefined model of the environment in order to develop a policy. Here, the agent performs actions based on trial and error and updates the strategy accordingly by evaluating the received rewards.

4.4 Q-LEARNING

One of the most popular model-free algorithms is *Q-Learning* [135, 136]. It is value-based, meaning that during training, a value function is incorporated in order to acquire knowledge regarding the relative importance of different states and to subsequently make decisions based on this learned information. The so-called Q-values $Q(s, a)$ are then stored in a lookup table (Q-table), referring to the anticipated future reward associated with a specific state-action combination. While the agent is interacting with its environment, the Q-table is iteratively updated accordingly, gaining knowledge from this exploration through rewards. In order to learn the optimal solution $Q(s, a)$ of the Q-value function, Q-learning exploits the Bellman equation [10]:

$$Q(s, a) = R(s, a) + \gamma \max_{a'} Q(s', a'), \tag{24}$$

where s and a denote the current and s' and a' the subsequent state and action, respectively. $Q(s, a)$ refers to the respective Q-value, $R(s, a)$ represents the reward and γ is the discount factor, specifying the importance attributed to future rewards.

4.5 DEEP Q-LEARNING

In DRL, RL is coupled with DL methods like CNNs, enabling the training of neural networks to learn hierarchical representations and han-

dling large amounts of input data [38]. The basic Q-learning configuration, as explained above, might not be sufficient for solving complex tasks involving high-dimensional state-action spaces. Therefore, deep NNs are incorporated – referred to as neural fitted Q-learning – so that the Q-function can be approximated through parameterization, resulting in a network $Q(s, a, \theta)$ with parameters θ [99]. A target network is introduced for stabilizing the training convergence, basically copying the Q-network and periodically updating for matching the parameters of the Q-network [88]. Then, the loss function for a Deep Q-Network (DQN) can be defined as:

$$L_{DQN} = \left(r + \gamma \max_{a'} Q(s', a'; \theta_n^-) - Q(s, a; \theta_n) \right)^2 \tag{25}$$

with the target network $Q(s', a'; \theta_n^-)$ that is updated with the Q-network parameters every nth iteration.

The DQN has been further developed towards *double* and *dueling* DQNs. In *double* DQN, the overestimation bias of a regular DQN is reduced by employing the Q-network for selecting actions and the target network for estimating values independently [125]. In *dueling* DQN, a dueling network architecture is introduced in order to segregate the estimation of state values and action advantages, enhancing learning efficiency [134].

4.6 APPLICATIONS TO MEDICAL IMAGE REGISTRATION

In the context of RL-based image registration, the respective images are created within a specified environment. The agent is trained to iteratively produce the ultimate transformation. RL has already been used for various tasks in medical image registration. This includes multi-modal rigid [75, 79] and affine registration [57]. Also, strategies towards deformable registration [70, 79, 93, 115], as well as the use of multiple agents instead of a single one [85], have been explored. Furthermore, DRL has successfully been used for the localization of landmarks in medical images, facilitating image alignment [2, 42–44, 129].

5 DEALING WITH PATHOLOGIES

In medical image registration, a primary challenge arises from the inherent nature of having to contend with diverse pathologies such as lesions or tumors. Moreover, special complexity is introduced by tumor resection, when transformations need to be performed between baseline images before surgery and follow-up images after surgery [21, 34, 37, 50, 72, 73]. In the subsequent sections, specific factors causing these difficulties are described. Fig. 6 provides sample visualizations.

5.1 DIFFERENCES IN IMAGING

Various factors contribute to the heightened difficulty of image registration in the presence of tumors, all sharing the common consequence of introducing a lack of knowledge between images. This can be caused by two main factors due to differences in imaging.

DIFFERENCES IN MODALITIES Different imaging modalities can depict tumors in varying ways, and these modalities often differ in terms of resolution and contrast [28]. The integration of information from modalities with distinct characteristics demands advanced registration techniques to guarantee precise alignment.

IMAGE ARTIFACTS Changes associated with tumors can introduce artifacts into images, such as blurring or distortion [64, 86]. These artifacts may impede the identification of corresponding features and compromise the accuracy of the registration process.

5.2 MISSING CORRESPONDENCES

Besides the differences in imaging, another major problem is the lack of correspondences, significantly complicating the registration process. This is especially the case when aligning pre-, intra- and postoperative images [23, 73], caused by the following issues.

INTENSITY VARIATIONS Tumors frequently show intensity variations in contrast to healthy tissue. These variations introduce difficulty in discerning shared features for registration purposes. Therefore, the inconsistency in intensity changes add an additional layer of complexity to the registration process.

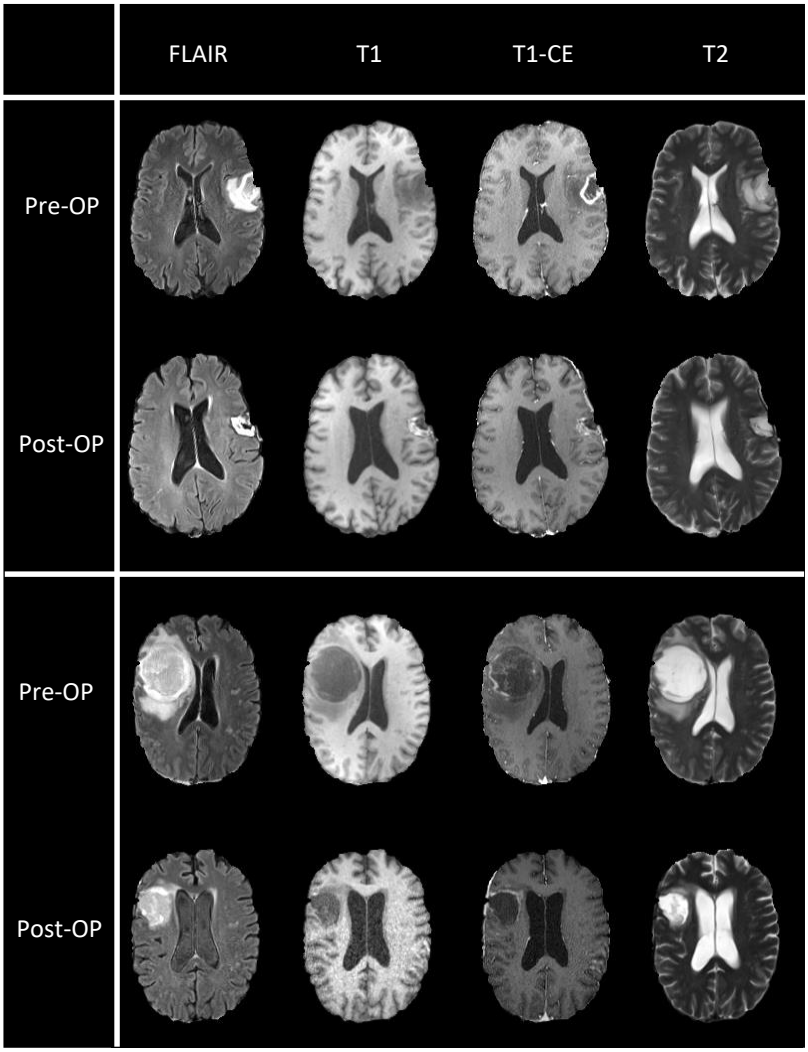


Figure 6: Pre-operative (Pre-OP) and post-operative (Post-OP) MR scans of two glioma patients, including four sequences each: native T1-weighted (T1), contrast-enhanced T1 (T1-CE), T2-weighted (T2), and T2 Fluid Attenuated Inversion Recovery (FLAIR). Missing correspondences due to tissue deformation and intensity shifts, induced by the tumor, complicate image registration.

TISSUE DEFORMATION Tumors have the potential to induce local or global deformations in adjacent tissues, and these deformations can exhibit a non-linear nature, posing challenges in establishing a direct mapping between images. Especially traditional linear registration methods are often insufficient to adequately capture these deformations.

CHANGES IN ANATOMY Furthermore, tumors can modify the anatomical structure of tissues, leading to compression or expansion of adjacent structures. Achieving accurate registration in the presence of these changes necessitates algorithms capable of managing non-rigid (deformable) transformations.

Overcoming these challenges frequently requires the advancement and implementation of sophisticated image registration techniques, including non-rigid registration methods adept at accommodating complex deformations. Furthermore, incorporating domain-specific knowledge, such as information about tumor characteristics and anatomy, can augment the precision of image registration in the context of pathologies.

An attempt to encourage the development, collection and evaluation of such specified algorithms has been made for the registration of brain MR images before and after tumor resection through the *Brain Tumor Sequence Registration (BraTS-Reg) Challenge* [8]. In this study, multiple algorithms have been collected and analyzed in the context of pre- and post-operative image registration, giving a broad overview of the current state of the art in this research field. The latest version of the manuscript (Dec. 2023) is provided in Appendix A.

6

ASSESSING REGISTRATION QUALITY

Evaluating the performance of image registration is challenging due to the lack of a *gold standard* or a definitive *ground truth* for comparison, for the majority of image registration applications. This is especially a problem in medical image registration, since accurate registration results are crucial for clinical decision-making processes such as diagnosis and treatment planning. In routine clinical workflows, the conventional approach for appraising the fidelity of image alignment involves clinicians visually inspecting the images and performing a qualitative comparative analysis. Given the impracticality of relying on visual inspection for a systematic evaluation of image registration, more quantitative methodologies are needed. This chapter provides an overview of frequently employed methods for quantifying image registration accuracy and discusses associated challenges and limitations.

6.1 QUANTITATIVE METRICS FOR ACCURACY EVALUATION

The most commonly used methodologies for quantifying the accuracy of image registration include overlap-based [9, 54, 55] and landmark-based metrics [12, 83, 89, 137].

OVERLAP-BASED METRICS Metrics based on overlap measure the precision of image registration by evaluating the alignment of corresponding segmentations between the transformed images. These metrics analyze the intersection or correspondence of regions within the registered images, usually represented through respective segmented areas. The following are some prevalent metrics that rely on overlap of aligned segmentations:

Dice Coefficient:

$$\mathrm{DSC}(X, Y) = \frac{2 \times |X \cap Y|}{|X| + |Y|}, \tag{26}$$

where $|X|$ denotes the size of set X, $|Y|$ is the size of set Y, and $|X \cap Y|$ represents the size (number of pixels or voxels) of the intersection between the sets X and Y [30].

Jaccard Index (Intersection over Union):

$$J(X, Y) = \frac{|X \cap Y|}{|X \cup Y|}, \tag{27}$$

where $|X \cap Y|$ represents the size of the intersection between the sets X and Y, and $|X \cup Y|$ denotes the size of the union of sets X and Y [63].

LANDMARK-BASED METRICS Landmark-based metrics evaluate registration accuracy by analyzing the spatial positions of predetermined landmarks or corresponding points in both the reference and the transformed images. These metrics furnish a quantitative assessment of the registration performance, particularly in scenarios where accurate alignment of specific anatomical or structural points is crucial. Commonly used evaluation metrics based on landmarks are the following:

Target Registration Error (TRE):

$$\mathrm{TRE}(T) = \mathrm{Distance}_{\mathrm{Euclidean}}(T_{\mathrm{ref}}, T_{\mathrm{trans}}), \tag{28}$$

where the *Euclidean distance* refers to the straight-line distance between corresponding landmarks in the reference and transformed images and is defined as:

$$D_E(P_i, Q_i) = \sqrt{(x_{P_i} - x_{Q_i})^2 + (y_{P_i} - y_{Q_i})^2 + (z_{P_i} - z_{Q_i})^2}, \tag{29}$$

for a pair of landmarks P_i and Q_i in 3D space.

Hausdorff Distance:

$$D_H(P, Q) = \max\left(\sup_{p \in P} \inf_{q \in Q} d(p, q), \sup_{q \in Q} \inf_{p \in P} d(q, p)\right), \tag{30}$$

for two sets of landmarks P and Q, measuring the maximum distance from each point in one set to the closest point in the other set [61].

6.2 CHALLENGES AND LIMITATIONS

Despite the increasing advancement of novel image registration methods, encompassing a diverse array of broadly applicable as well as highly specialized algorithms, progress in the development of measures to systematically assess the quality of registration has stagnated in recent years.

Highly effective algorithms are often characterized by minimal Target Registration Error (TRE) differences, frequently falling within a sub-resolution range. When interpreting these TRE values, it becomes paramount to consider both inter- and intra-rater reliability. This consideration is crucial for distinguishing genuine performance enhancements from random improvements in fitting the test set [69]. An approach towards a more clinically relevant representation of image registration accuracy has therefore been proposed in [131], enabling the incorporation of use-case-specific accuracy requirements. The latest version of the manuscript (Dec. 2023) is provided in Appendix B.

Part II

ACADEMIC CONTRIBUTIONS

7

REINFORCED REDETECTION OF LANDMARK IN PRE- AND POST-OPERATIVE BRAIN SCAN USING ANATOMICAL GUIDANCE FOR IMAGE ALIGNMENT

Authors: Diana Waldmannstetter, Fernando Navarro, Benedikt Wiestler, Jan S. Kirschke, Anjany Sekuboyina, Ester Molero, Bjoern H. Menze

In: Biomedical Image Registration: 9th International Workshop, WBIR 2020, Portorož, Slovenia, December 1–2, 2020, Proceedings 9. Springer. [130]

Abstract: Re-identifying locations of interest in pre- and post-operative images is a hard identification problem, as the anatomical landscape changes dramatically due to tumor resection and tissue displacement. Classical image registration techniques oftentimes fail in vicinity of the tumor, where the enclosing structures are massively altered from one scan to another. Still, locations nearby the tumor or the resection cavity are the most relevant for evaluating tumor progression patterns and for comparing pre- and post-operative radiomic signatures. We address this issue by exploring a Reinforcement Learning (RL) approach. An artificial agent is self-taught to find the optimal path towards a target driven by a feedback signal from the environment. Incorporating anatomical guidance, we restrict the agent's search space to surgery-unaffected structures only. By defining landmarks for each patient individually, we aim to obtain a patient-specific representation of its differential radiomic features across different time points for enhancing image alignment. Estimated landmarks reach a remarkable mean distance error around 3 mm. In addition, they show a high agreement with expert annotations on a challenging dataset of MR scans from the brain before and after tumor resection.

Contributions of thesis author: Project idea and concept, data collection and preparation, implementation and experimental design, data analysis and evaluation, manuscript preparation and writing.

Reinforced Redetection of Landmark in Pre- and Post-operative Brain Scan Using Anatomical Guidance for Image Alignment

Diana Waldmannstetter[1,2(✉)], Fernando Navarro[1,4], Benedikt Wiestler[3], Jan S. Kirschke[3], Anjany Sekuboyina[1,3,4], Ester Molero[1], and Bjoern H. Menze[1,2,4]

[1] Department of Informatics, Technical University of Munich, Munich, Germany
diana.waldmannstetter@tum.de
[2] Munich School of BioEngineering, Munich, Germany
[3] Department of Neuroradiology, Klinikum rechts der Isar, Munich, Germany
[4] TranslaTUM - Central Institute for Translational Cancer Research, Munich, Germany

Abstract. Re-identifying locations of interest in pre- and post-operative images is a hard identification problem, as the anatomical landscape changes dramatically due to tumor resection and tissue displacement. Classical image registration techniques oftentimes fail in vicinity of the tumor, where the enclosing structures are massively altered from one scan to another. Still, locations nearby the tumor or the resection cavity are the most relevant for evaluating tumor progression patterns and for comparing pre- and post-operative radiomic signatures. We address this issue by exploring a Reinforcement Learning (RL) approach. An artificial agent is self-taught to find the optimal path towards a target driven by a feedback signal from the environment. Incorporating anatomical guidance, we restrict the agent's search space to surgery-unaffected structures only. By defining landmarks for each patient individually, we aim to obtain a patient-specific representation of its differential radiomic features across different time points for enhancing image alignment. Estimated landmarks reach a remarkable mean distance error around 3 mm. In addition, they show a high agreement with expert annotations on a challenging dataset of MR scans from the brain before and after tumor resection.

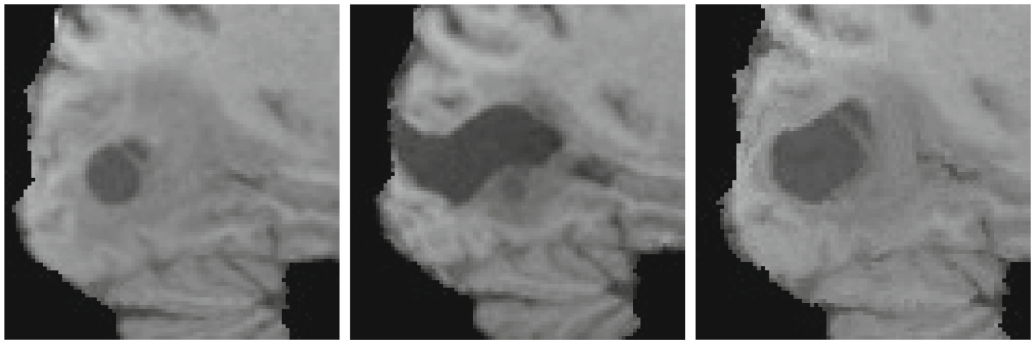
Keywords: Reinforcement Learning · Image registration · Image alignment · Differential radiomics · Brain tumor

1 Introduction

The most effective treatment for progression delay in aggressive primary brain tumors is tumor resection, usually followed by radiation therapy or chemotherapy [4]. When evaluating the post-operative scans, the areas that show signs of

Ž. Špiclin et al. (Eds.): WBIR 2020, LNCS 12120, pp. 81–90, 2020.
https://doi.org/10.1007/978-3-030-50120-4_8

tumor re-growth are compared to the same areas in the pre-operative scans. As there is almost always a shift in brain tissue as well as tumor- and resection-induced intensity changes, conventional image registration techniques oftentimes fail when it comes to map the differing structures, see Fig. 1. We aim to evaluate local patterns with anatomical guidance for a better adaption in this task. We perform re-identification of landmarks for making use of quantitative radiomic approaches, since radiomics intend to improve image analysis by extracting large amounts of quantitative features [8]. In order to detect reference points before and after tumor resection, we use individual landmarks for each patient, representing locations prone to progression. Redetecting these landmarks automatically in follow-up scans may simplify future image alignment for the same patient. Therefore, we define multiple patient-specific landmarks around the tumor and take a first step towards differential radiomic feature extraction and therefore, a more precise alignment of the resection-affected regions.



(a) pre-operative (b) post-operative (c) registered

Fig. 1. 2D zoom into tumor-/resection-affected regions in (a) a pre-operative scan, (b) the corresponding post-operative scan and (c) the result of a standard image registration

Recent literature shows a variety of approaches towards localization of anatomical landmarks in medical images. Li *et al.* [9] developed a patch-based CNN for landmark localization combining regression and classification for the detection of both single and multiple landmarks simultaneously by involving Principle Component Analysis (PCA). Zheng *et al.* [19] evolved a two-step approach combining a shallow network with a deep network for efficient landmark detection. Another two-stage approach was proposed by Zhang *et al.* [18] comprising a patch-based CNN regression model followed by another CNN for predicting landmarks in an end-to-end manner. For the first time, Ghesu *et al.* [6] introduced Deep RL for localizing anatomical landmarks using Q-learning. Their method is further developed by exploring multiple scales in [5,7]. In [11], Maicas *et al.* adopt this method and extend it to the more complex detection of breast lesions. There, adaptive bounding boxes are leveraged to train the agent. Alansary *et al.* [2] presented a multi-scale strategy by iteratively training their agent using action steps with different sizes on multiple scales. Additionally, they evaluate several Q-Learning approaches as there are Double, Dueling

and Dueling Double Q-Learning. A variant of their approach towards automatic view planning is shown in [1] and an extension towards the detection of multiple landmarks simultaneously in [15].

The recent success of RL in the field of landmark localization [2,5–7,15] in combination with the ability of RL agents to adapt to a specific environment, encouraged us to transfer this approach for landmark redetection to pre- and post-operative brain images. Furthermore, RL has the benefit of being able to perform on limited training data, which is crucial for our task. Based on the approach in [2], we further develop the method to consider anatomical guidance and propose the following contributions: First, we present two RL-based agents. A baseline agent and an extended version of it under anatomical guidance, improving the agent's ability to adapt to the issue of altering tissue structures by integrating patient-specific anatomy into our model. Second, we evaluate our approach on a challenging dataset of MR scans before and after tumor resection, provided by the BraTS challenge [12] and TCIA [3], achieving results comparable to an expert performance. Additionally, we provide annotations for this data.

In the following, we present how we utilize Q-Learning in RL (Sect. 2) and introduce our extension (Sect. 3), before demonstrating the performance of our approach on a complex data set.

2 Deep Reinforcement Learning Using Q-Learning

In RL, an artificial agent is self-taught by interacting with an environment. In every step, the agent retrieves a reward from its environment after executing an action. The final goal of the agent is to find an optimal policy, guiding the agent from any given state to the target by maximizing future rewards. This can be formulated as a sequential decision process. RL then is modeled as a *Markov Decision Process* (MDP), which defines the interaction between the agent and its environment. The agent executes an action $a \in A$ at state $s \in S$, returning a reward signal $r \in R$ at each time step t [14].

Finding the optimal policy is described by the action-value function $Q(s, a)$, which is optimized during training and gives the maximum expected discounted future reward, where the accumulated discounted reward after τ time steps is defined as

$$R_\tau = \sum_{\tau=0}^{\infty} \gamma^\tau r_{t+\tau+1}, \tag{1}$$

with the discount rate $\gamma \in [0, 1]$ for weighting immediate and future rewards [17]. Using the Bellman optimality equation, the action-value function can be solved recursively [14]:

$$Q(s, a) = \mathbb{E}\left[r + \gamma \max_{a'} Q(s', a')\right], \tag{2}$$

where s' and a' are the possible subsequent state and action.

Mnih *et al.* [13] developed the Deep Q-Network (DQN), which approximates

$$Q(s,a) \approx Q(s,a;\theta), \tag{3}$$

using a CNN with the network parameters θ. For stability reasons, a target network $Q(\theta^-)$ is introduced. It estimates the actual Q-network iteratively by updating the parameters of the target network only every nth iteration with the steadily updated Q-network parameters. The loss function reads:

$$L_n(\theta_n) = E_{s,a,r,s'}\left[\left(r + \gamma \max_{a'} Q(s',a';\theta_n^-) - Q(s,a;\theta_n)\right)^2\right] \tag{4}$$

Experience replay technique [10] is added, training the network using randomly sampled minibatches from experiences the agent has already gained. This is stored in an experience replay memory.

The DQN was further improved by Wang *et al.* [16], separating the network into two partitions. One handles the state-value function $V(s)$ and the other one deals with the advantage function $A(s,a)$, see Fig. 2. Both are then combined by an aggregation layer to provide a single Q-function

$$Q(s,a) = V(s) + A(s,a). \tag{5}$$

Here, estimating the state-value function is essential in every time step, while this is not necessary for the advantage function. Consequently, the dueling network learns the state-value function more accurately, thus improving the network performance with increasing number of actions.

3 Anatomically Guided RL Agent

Similar to [2], we make use of a Deep Q-Network with dueling architecture, shown in Fig. 2. Each state in our image environment is modeled as a 3D patch centered around the current location of the agent, see Fig. 3(a). Hence, the agent sees a different part of its environment in every time step. Due to the experience replay technique, we define an experience buffer storing the last four patches, which the network can see in one iteration, enhancing the agent's robustness.
We define the action space with the six actions *right, left, forward, backward, up, down*. This results in two actions along each axis in positive and negative direction, $a \in A = \{+x, -x, +y, -y, +z, -z\}$, thus moving the agent by one voxel. The reward is defined similar to [6], calculating the relative change in the distance to the position of the target landmark. Furthermore, we make use of a search strategy operating on multiple scales [2,5,7] for more robustness and efficiency.

A key feature of our approach is the anatomical guidance. We return a negative reward $r = -1$ when the agent steps inside the surgery-affected regions. Therefore, we provide the segmentation mask to the agent, so the agent learns to stay in unaffected structures only, see Fig. 3(b). Since this guides the agent to move towards the target without touching the immense tissue changes inside the most affected regions, this leads to a policy that is more generalizable to altering brain tissue.

4 Experimental Setup

We evaluate our approach on a challenging dataset provided by the BraTS challenge [12] and TCIA [3]. We use MR image data from 10 patients with brain tumors, with one scan before and one scan after tumor resection, comprising 20 image volumes in total. All images are skull-stripped, rigidly co-registered and interpolated to a common resolution of $1\,\mathrm{mm}^3$, while the initial resolution is in the range of 3–8 mm for most sequences. The dataset includes the image volumes and their corresponding segmentation masks of the tumor- and resection-affected regions in the pre- and post-operative scans, respectively. For each patient, 3 landmarks in the post-operative scan are annotated by a clinical expert, in varying distances up to 4 cm around the resection-affected region. The same expert redetected the landmarks in the corresponding pre-operative scan for generating ground truth annotations.

Training and Testing. Before training, we crop an initialization box of size $50 \times 50 \times 50$ voxels around the target in the training image, see Fig. 3(a). When training under anatomical guidance, we exclude the resection mask, see Fig. 3(b). Then, we randomly initialize the agent inside this region and sample a patch of size $15 \times 15 \times 15$ voxels, which follows the agent in every step. For every patient and landmark individually, we train on the respective post-operative scan and test on the corresponding pre-operative scan, generating patient-specific models. Similar to [2], we define the terminal state in training as the point, when the distance between the agent and the target landmark is less or equal to 1 mm. During testing, the agent is stopped, when it is oscillating around the same location.

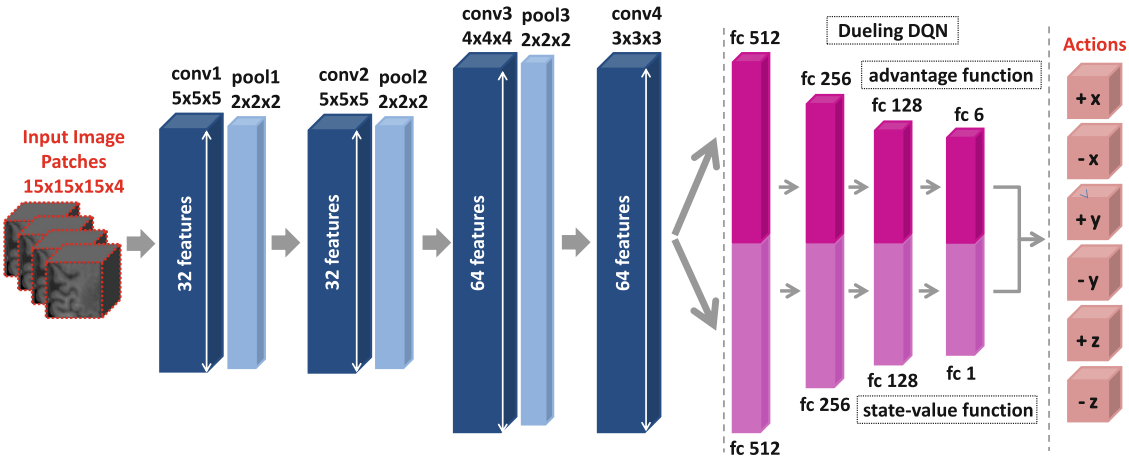


Fig. 2. Our network architecture with a dueling DQN. A 3D patch is sampled around the current position of the agent and fed to the network, consisting of convolutional (conv) layers alternating with pooling (pool) layers, followed by a dueling DQN with fully connected (fc) layers. The network outputs the Q-value for the six possible actions, whereof the agent selects the one with the highest value.

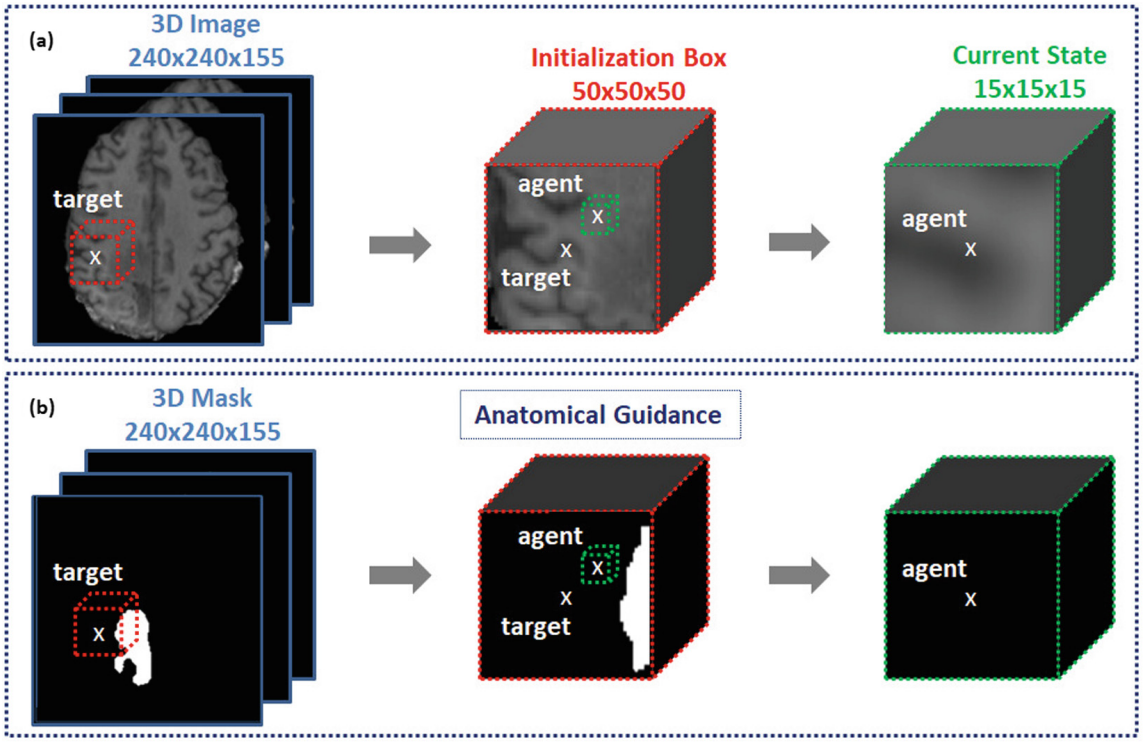
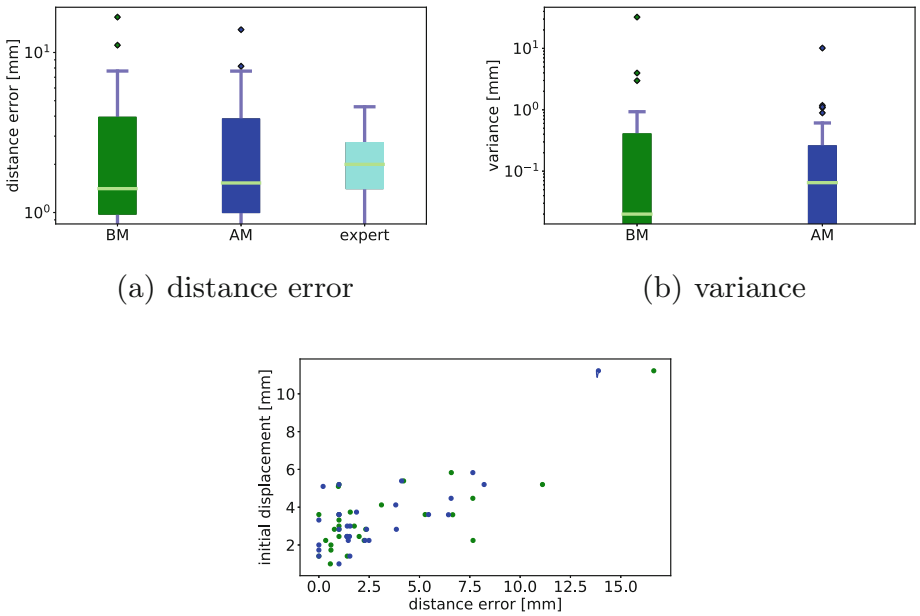


Fig. 3. (a) Patch generation. A box of size $50 \times 50 \times 50$ is sampled around the target for initialization reasons. Then, a patch of size $15 \times 15 \times 15$ is extracted around the current position of the agent, representing the current state. (b) Anatomical Guidance. We provide the segmentation mask to the agent for excluding this region during training. When stepping inside the masked region, a negative reward is returned, so that the agent is guided to avoiding affected structures.

Experiments. We use two different agents for each experiment. One is trained on the baseline method without anatomical guidance, whereas the other one is anatomically guided. Due to the lack of a validation set, we tune the model on the respective training image. For each experiment, we then choose the model that is performing best on the corresponding test image. Although this might lead to some underestimation regarding the distance error measurements, it makes sure that we provide the same conditions for the two agents in the different experiments, leading to comparable results. During evaluation, we define 20 fixed starting points, typically converging to slightly different final endpoints. Since we initialize the training agent inside the initialization box, we use the same box for testing and select the starting points from there. For each of the 20 evaluation runs per experiment, we calculate the Euclidean distance in mm between the final location of the agent and the true landmark, which gives us the distance error, and calculate the mean, for producing comparable results. Subsequently, for the sake of simplicity, we refer to this mean distance error simply as distance error.

5 Results

Quantitative results can be observed from Fig. 4(a) and (b), showing the distance errors in mm for the baseline method (BM), the extended method using anatomical guidance (AM), as well as another expert's annotation for comparison. Therefore, we take the landmark annotations of a second expert, when performing the redetection task manually, and calculate the Euclidean distance to the ground truth annotations, giving us the distance errors of an expert. For further comparisons, we calculate additional measurements on the distance errors, the mean and the median distance as well as the normalized mean and median, respectively, see Table 1. All measurements are calculated on landmark level. Qualitative results are presented in Fig. 5, showing a sample redetection for two different landmarks, where both methods achieve high precision with a distance error of 0 mm.



(a) distance error (b) variance

(c) initial displacement

Fig. 4. Results for the baseline method (BM, green), the extended method with anatomical guidance (AM, dark blue) and an expert annotation for comparison (light blue). (a) shows the distance mean errors in mm and (b) the variances in the distance errors due to the multiple starting points. (c) shows the relation between the initial expert annotation displacements and the distance errors. The dots represent the respective offsets of the initial expert annotation for the training and test images in relation to the corresponding distance errors of BM and AM. (Color figure online)

RL vs Expert. The lowest distance errors for both methods are close to 0 mm, representing a perfect redetection. The highest lie above 1 cm. Due to the 20 starting points, we achieve variances in the distance errors, tending to increase with growing errors, see Fig. 4(a) and (b). High variances are caused by some

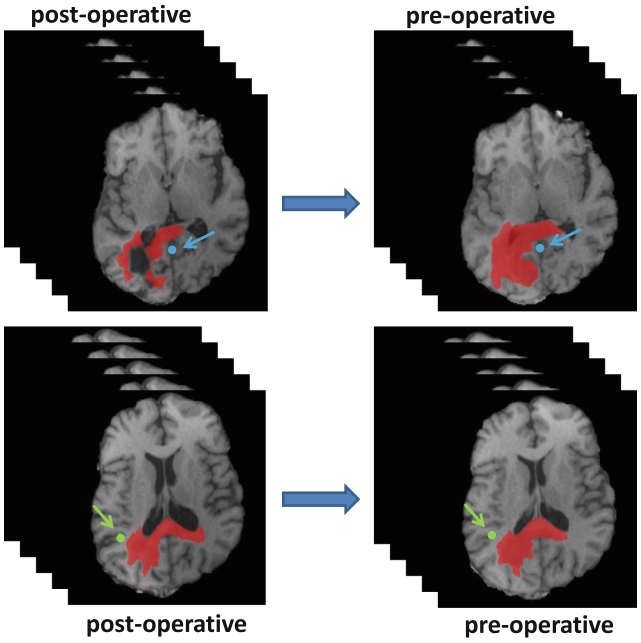


Fig. 5. Sample redetection for two different landmarks (top, bottom) in the same patient. The masked region is marked in red. The landmarks in blue and green, respectively. (Color figure online)

outliers, where the agent gets lost in the environment. However, it is remarkable that we achieve a variance of 0 in some experiments, which means that the agent navigates towards the same target from every starting point, demonstrating high robustness. As Fig. 4(c) shows, the distance errors from BM and AM both scale with the initial displacements between the ground truth annotations in the training and test images, when annotated by an expert. That means, a larger offset between the initial expert landmark annotation in the training and the test image results in larger distance errors. This makes sense, since larger initial annotation displacements are linked to larger tissue changes. Nevertheless, the majority ranges within smaller errors from 0–4 mm. From Table 1, we observe that the mean of all distance errors is lowest for the comparison expert, while both BM and AM show high agreement with it. Still, the median of all distance errors is smaller for BM and AM. A normalization with the initial displacements leads to similar mean errors of both RL methods and the comparison expert, see Table 1.

Benefits of Anatomical Guidance. Figure 4(a) and (b) as well as Table 1 show that AM performs more robust than BM, since the outliers have slightly smaller mean distance errors and variances. Hence, incorporating anatomical guidance outperforms the baseline agent in average, while showing high agreement with the comparison expert's annotations. Our approach achieves noticeable performance with an average distance error below 3 mm. Moreover, anatomical guidance provides potential to incorporate additional anatomical information.

Table 1. Calculations on the distance errors for BM, AM and an expert annotation.

	Mean distance [mm]	Median distance [mm]	Normalized mean	Normalized median
BM	3.05	1.41	0.79	0.57
AM	2.82	1.53	0.74	0.64
Expert	2.18	2.0	0.75	0.72

6 Conclusion

In this work, we presented a RL framework for landmark redetection in a challenging dataset of pre- and post-operative brain scans. We evaluated two RL agents: a basic one exploring the full environment and an extended one guided by the resection anatomy for finding the optimal path towards the target landmark. Overall, both approaches showed good results in terms of speed and accuracy, while the agent under anatomical guidance performs better in average. Therefore, this approach allows to further develop the guidance by anatomical structures, especially in analyzing the connection between different time points before and after tumor resection, for generating a more representative and efficient model of the anatomical changes. For further automatization, the segmentation masks can be produced using some segmentation framework, which would be of minimal additional effort here and would be needed to be done once for training only. Additionally, we will invest in finetuning our approach towards a more robust redetection for eliminating outliers. Moreover, we will further investigate in generating a dense representation of patient-specific differential radiomics by localizing multiple landmarks simultaneously, ideally incorporating the spatial relationships between tumor structures, resection region and landmarks.

Acknowledgements. Supported by Deutsche Forschungsgemeinschaft (DFG) through TUM International Graduate School of Science and Engineering (IGSSE), GSC 81.

References

1. Alansary, A., et al.: Automatic view planning with multi-scale deep reinforcement learning agents. In: Frangi, A.F., Schnabel, J.A., Davatzikos, C., Alberola-López, C., Fichtinger, G. (eds.) MICCAI 2018. LNCS, vol. 11070, pp. 277–285. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-00928-1_32
2. Alansary, A., Oktay, O., Li, Y., et al.: Evaluating reinforcement learning agents for anatomical landmark detection. Med. Image Anal. **53**, 156–164 (2019)
3. Clark, K., Vendt, B., Smith, K., et al.: The cancer imaging archive (TCIA): maintaining and operating a public information repository. J. Digit. Imag. **26**(6), 1045–1057 (2013)
4. DeAngelis, L.: Brain tumors. New Engl. J. Med. **344**(2), 114–123 (2001)

5. Ghesu, F.C., Georgescu, B., Grbic, S., Maier, A.K., Hornegger, J., Comaniciu, D.: Robust multi-scale anatomical landmark detection in incomplete 3D-CT data. In: Descoteaux, M., Maier-Hein, L., Franz, A., Jannin, P., Collins, D.L., Duchesne, S. (eds.) MICCAI 2017. LNCS, vol. 10433, pp. 194–202. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-66182-7_23
6. Ghesu, F.C., Georgescu, B., Mansi, T., Neumann, D., Hornegger, J., Comaniciu, D.: An artificial agent for anatomical landmark detection in medical images. In: Ourselin, S., Joskowicz, L., Sabuncu, M.R., Unal, G., Wells, W. (eds.) MICCAI 2016. LNCS, vol. 9902, pp. 229–237. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46726-9_27
7. Ghesu, F., et al.: Multi-scale deep reinforcement learning for real-time 3D-landmark detection in CT scans. IEEE TPAMI **41**(1), 176–189 (2017)
8. Lambin, P., Rios-Velazquez, E., Leijenaar, R., et al.: Radiomics: extracting more information from medical images using advanced feature analysis. Eur. J. Cancer **48**(4), 441–446 (2012)
9. Li, Y., et al.: Fast multiple landmark localisation using a patch-based iterative network. In: Frangi, A.F., Schnabel, J.A., Davatzikos, C., Alberola-López, C., Fichtinger, G. (eds.) MICCAI 2018. LNCS, vol. 11070, pp. 563–571. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-00928-1_64
10. Lin, L.J.: Self-improving reactive agents based on reinforcement learning, planning and teaching. Mach. Learn. **8**(3–4), 293–321 (1992)
11. Maicas, G., Carneiro, G., Bradley, A.P., Nascimento, J.C., Reid, I.: Deep reinforcement learning for active breast lesion detection from DCE-MRI. In: Descoteaux, M., Maier-Hein, L., Franz, A., Jannin, P., Collins, D.L., Duchesne, S. (eds.) MICCAI 2017. LNCS, vol. 10435, pp. 665–673. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-66179-7_76
12. Menze, B., Jakab, A., Bauer, S., et al.: The multimodal brain tumor image segmentation benchmark (BRATS). IEEE Trans. Med. Imag. **34**(10), 1993–2024 (2015)
13. Mnih, V., Kavukcuoglu, K., Silver, D., et al.: Human-level control through deep reinforcement learning. Nature **518**(7540), 529 (2015)
14. Sutton, R., Barto, A.: Introduction to Reinforcement Learning, 1st edn. MIT Press, Cambridge (1998)
15. Vlontzos, A., Alansary, A., Kamnitsas, K., Rueckert, D., Kainz, B.: Multiple landmark detection using multi-agent reinforcement learning. In: Shen, D., et al. (eds.) MICCAI 2019. LNCS, vol. 11767, pp. 262–270. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-32251-9_29
16. Wang, Z., Schaul, T., Hessel, M., Van Hasselt, H., Lanctot, M., De Freitas, N.: Dueling network architectures for deep reinforcement learning. In: International Conference on Machine Learning, pp. 1995–2003 (2016)
17. Watkins, C., Dayan, P.: Q-learning. Mach. Learn. **8**(3–4), 279–292 (1992)
18. Zhang, J., Liu, M., Shen, D.: Detecting anatomical landmarks from limited medical imaging data using two-stage task-oriented deep neural networks. IEEE Trans. Image Process. **26**(10), 4753–4764 (2017)
19. Zheng, Y., Liu, D., Georgescu, B., Nguyen, H., Comaniciu, D.: 3D deep learning for efficient and robust landmark detection in volumetric data. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9349, pp. 565–572. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24553-9_69

8

PRIMITIVE SIMULTANEOUS OPTIMIZATION OF SIMILARITY METRICS FOR IMAGE REGISTRATION

Authors: Diana Waldmannstetter, Benedikt Wiestler, Julian Schwarting, Ivan Ezhov, Marie Metz, Spyridon Bakas, Bhakti Baheti, Satrajit Chakrabarty, Daniel Rueckert, Jan S. Kirschke, Rolf A. Heckemann, Marie Piraud, Bjoern H. Menze†, Florian Kofler†

† Equal contribution

In: International MICCAI Brainlesion Workshop. Springer. 2023. [132]

Abstract: Even though simultaneous optimization of similarity metrics is a standard procedure in the field of semantic segmentation, surprisingly, this is much less established for image registration. To help closing this gap in the literature, we investigate in a complex multi-modal 3D setting whether simultaneous optimization of registration metrics, here implemented by means of primitive summation, can benefit image registration. We evaluate two challenging datasets containing collections of pre- to post-operative and pre- to intra-operative Magnetic Resonance (MR) images of glioma. Employing the proposed optimization, we demonstrate improved registration accuracy in terms of Target Registration Error (TRE) on expert neuroradiologists' landmark annotations.

Contributions of thesis author: Project conceptualization, data collection and preparation, Algorithmic design and implementation, experimental design and concept, data analysis and evaluation, manuscript preparation and writing.

Primitive Simultaneous Optimization of Similarity Metrics for Image Registration

Diana Waldmannstetter[1,2(✉)], Benedikt Wiestler[3], Julian Schwarting[3], Ivan Ezhov[1,4], Marie Metz[3], Spyridon Bakas[5,6,7], Bhakti Baheti[5,6,7], Satrajit Chakrabarty[8], Daniel Rueckert[9,10], Jan S. Kirschke[3], Rolf A. Heckemann[11], Marie Piraud[12], Bjoern H. Menze[2], and Florian Kofler[1,3,4,12]

[1] Department of Informatics, Technical University of Munich, Munich, Germany
diana.waldmannstetter@tum.de
[2] Department of Quantitative Biomedicine, University of Zurich, Zurich, Switzerland
[3] Department of Diagnostic and Interventional Neuroradiology, School of Medicine, Klinikum rechts der Isar, Technical University of Munich, Munich, Germany
[4] TranslaTUM - Central Institute for Translational Cancer Research, Technical University of Munich, Munich, Germany
[5] Center for Artificial Intelligence and Data Science for Integrated Diagnostics (AI2D) and Center for Biomedical Image Computing and Analytics (CBICA), University of Pennsylvania, Philadelphia, PA, USA
[6] Department of Pathology and Laboratory Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA
[7] Department of Radiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA
[8] Department of Electrical and Systems Engineering, Washington University in St. Louis, St. Louis, MO, USA
[9] Artificial Intelligence in Healthcare and Medicine, Technical University of Munich, Munich, Germany
[10] Department of Computing, Imperial College London, London, UK
[11] Department of Medical Radiation Sciences, University of Gothenburg, Gothenburg, Sweden
[12] Helmholtz AI, Helmholtz Zentrum München, Munich, Germany

Abstract. Even though simultaneous optimization of similarity metrics is a standard procedure in the field of semantic segmentation, surprisingly, this is much less established for image registration. To help closing this gap in the literature, we investigate in a complex multi-modal 3D setting whether simultaneous optimization of registration metrics, here implemented by means of primitive summation, can benefit image registration. We evaluate two challenging datasets containing collections

B. H. Menze and F. Kofler—Equal contribution.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/978-3-031-76160-7_6.

U. Baid et al. (Eds.): BrainLes 2023/SWITCH 2023, LNCS 14668, pp. 57–68, 2024.
https://doi.org/10.1007/978-3-031-76160-7_6

of pre- to post-operative and pre- to intra-operative Magnetic Resonance (MR) images of glioma. Employing the proposed optimization, we demonstrate improved registration accuracy in terms of *Target Registration Error (TRE) o* on expert neuroradiologists' landmark annotations.

Keywords: Registration · Brain Tumor · Similarity Metric · Loss Function · Glioma

1 Introduction

The standard treatment for glioma is an *operative procedure (OP)* aiming to remove the tumor in full. Clinicians measure the success of surgery by comparing pre- and post-operative scans. Furthermore, this comparison is required for subsequent treatment planning, such as radiation therapy. Image registration techniques can enhance this process by providing a direct overlay of the differing structures. Enormous tissue shift and consequential missing correspondences mark a common side effect of tumor resection and pose a major challenge in registering pre- to post-operative images. The same holds true for intra-operative imaging in order to keep track of the surgery progress. A fast and accurate image registration method is beneficial for the precise estimation of tumor resection. Additionally, image registration is an important part of the preprocessing for segmentation algorithms [15,16]. Moreover, registration can be part of the segmentation algorithm itself [11]. Recent advances in the field of *Deep Learning (DL)* also benefited medical image registration. Methods like *VoxelMorph (VM)* [6], *LapIRN* [19], and *TransMorph* [9] have demonstrated that unsupervised deformable image registration is a promising alternative to frameworks based on iterative optimization algorithms like *Symmetric Normalization (SyN)* from *Advanced Normalization Tools (ANTs)* [3], *Free-Form Deformation (FFD)* [21] or toolboxes like Elastix [13]. *DL* methods provide comparable performance while significantly improving processing time. Within the scope of the *Brain Tumor Sequence Registration Challenge (BraTS-Reg)* [4], there has also been considerable development of registration algorithms for MR brain images before and after tumor resection [8,12,17,18,25]. Unsupervised *Deformable Registration Network (DRN)* can nicely register healthy brain scans [6,9,19]; however, they frequently struggle with large pathologies. Therefore, *Instance-Specific Optimization (IO) p* proved to be advantageous to mitigate major deformations [18,25]. Existing deep learning registration algorithms usually make use of a single similarity metric, often coupled with a smoothing regularization based on the deformation field, to be optimized in the loss function [6,19]. However, there is only limited literature investigating the combination of multiple similarity metrics to improve registration results [1,2,7,10,22,23,26].

In this work, we extend an unsupervised *DRN* using a combination of image similarity metrics, which are optimized simultaneously in the loss function. We benchmark the performance of our approach against the baseline on two challenging datasets of pre- to post-operative as well as pre- to intra-operative brain

tumor images. Furthermore, we compare our approach to established reference methods, achieving competitive results. Following [6,18,25], we opt for instance-specific optimization in addition to the *DRN*. We demonstrate that the simultaneous optimization of two similarity metrics can improve registration accuracy in terms of *TRE* on expert landmark annotations.

2 Methods

Our workflow consists of three steps. First, we train a *Deformable Registration Network (DRN)*, followed by *Instance-Specific Optimization (IO)* at test time. Both *DRN* and *IO* optimize the same loss function. We then evaluate the registration performance in terms of *Target Registration Error (TRE)*.

2.1 Unsupervised Deformable Registration

We start our approach with a *DRN*, similar to [6]. Given two 3D images, a source image X and a target image Y, the network models the following function $f_\theta(X, Y) = u$, where the displacement field u, aligning X and Y, is defined bidirectionally, leading to $u_{x,y} = f_\theta(X, Y)$ and $u_{y,x} = f_\theta(Y, X)$ with θ being a set of learning parameters. X and Y are then warped with the respective displacement field using a spatial transform. While many *Convolutional Neural Network (CNN)*-based models are applicable here, like [6], we opt for a *U-Net*-like architecture with encoder, decoder and skip connections. The network architecture is based on an early implementation of *VM*, see GitHub implementation. The loss function comprises two main components, which are an image similarity metric and a smoothness regularizer for the displacement field. In addition to optimizing a single similarity metric as part of the loss function, we opt for optimizing multiple metrics simultaneously. An overview of the workflow is shown in Fig. 1.

2.2 Proposed Combined Loss Function

We define the objective function L for bidirectional training as follows:

$$L = \sum_{n=1}^{N} \left(L_{Sim_n}^{forward} \cdot \omega_n + L_{Sim_n}^{backward} \cdot \omega_n \right) + L_{Reg} \cdot \lambda \tag{1}$$

where N is the number of similarity metrics L_{Sim}, L_{Reg} is the regularization term, and the respective weights are denoted by ω_n and λ.

2.3 Instance-Specific Optimization

Since the registration results of a plain *DRN* might not always be sufficient, especially in the case of pathologies, *IO* is added in order to further refine the registration. Similar to [6], we take the output displacement field of *DRN* as initialization for a gradient-descent based iterative optimization on each test scan individually, following the initial training process. This technique optimizes the same loss function that is used for *DRN*. An overview of the complete method is given in Fig. 1.

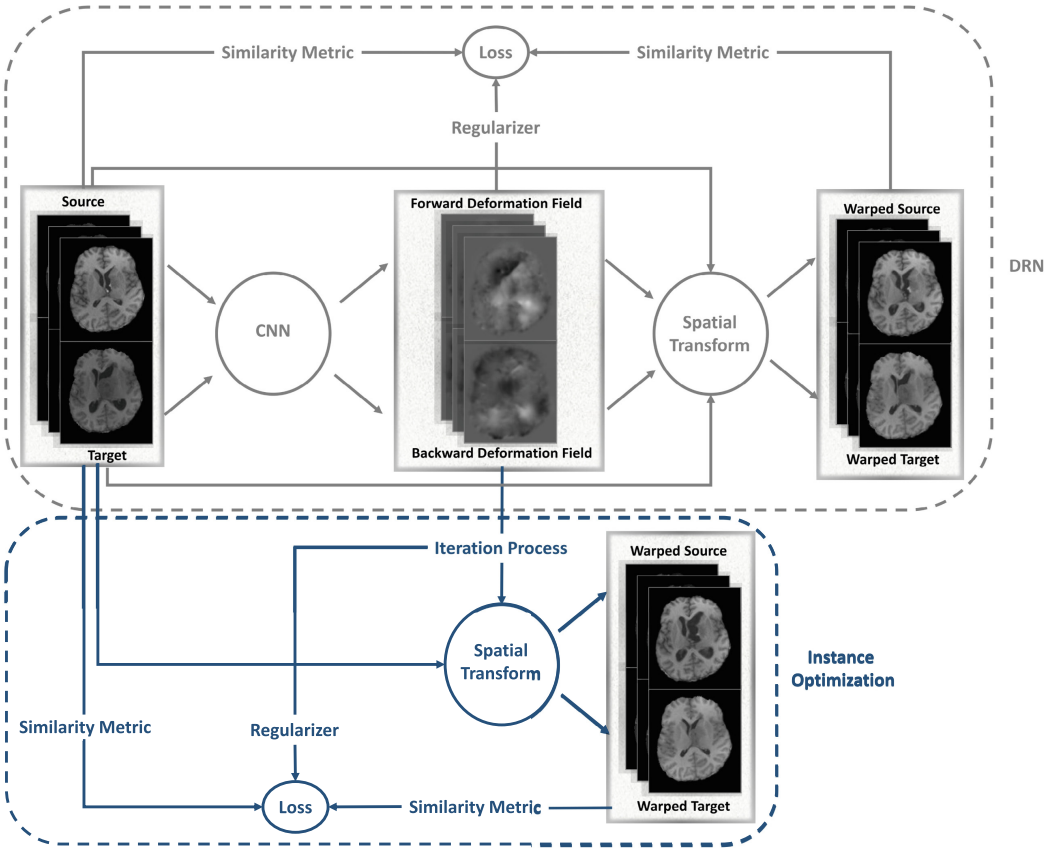


Fig. 1. Overview of the workflow. An unsupervised deformable registration network (*DRN*, top) is combined with *Instance-Specific Optimization (IO)* (bottom) for iterative refinement at test time using the output deformation field of the trained *DRN*. *DRN* as well as *IO* are trained bidirectionally, providing both forward and backward deformation fields. For both modules, the loss function combines either a single or multiple image similarity metrics with a smoothness regularization on the deformation field.

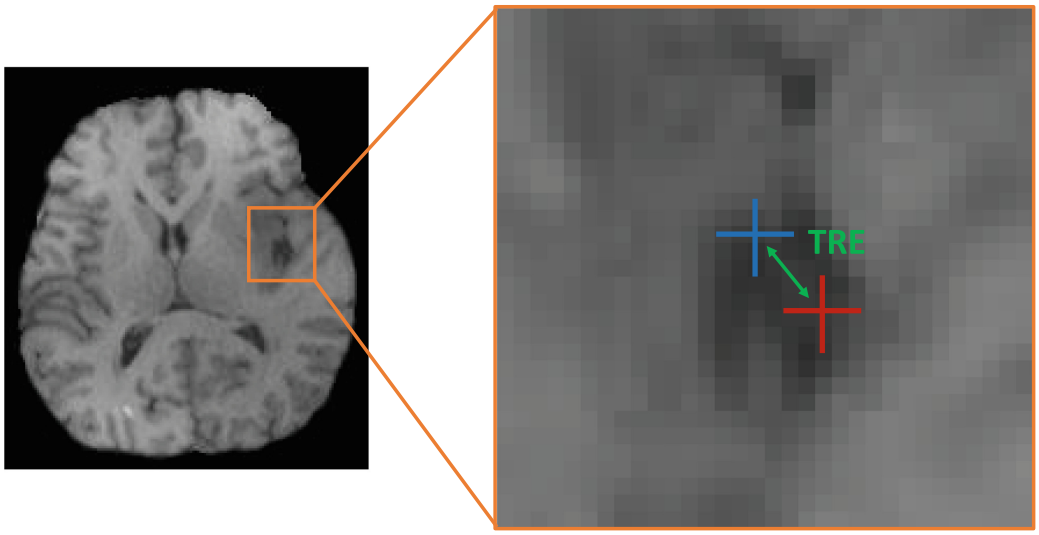


Fig. 2. Illustration of *Target Registration Error (TRE)* evaluation between expert landmark annotation (red) and warped landmark after image registration (blue). The green arrow indicates the Euclidean distance between the two points. For simplicity, the illustration is in 2D, while the actual evaluation is performed in 3D space. (Color figure online)

2.4 Evaluation

We evaluate all experiments by calculating the mean *TRE* between the expert landmark annotations and the warped landmarks on the respective test sets. We use the Euclidean distance to calculate the *TRE*, as illustrated in Fig. 2.

3 Experiments and Results

We perform experiments using different loss functions on two brain *Magnetic Resonance Imaging (MRI)* datasets. Additionally, we implement two reference methods for comparison.

3.1 Data and Pre-processing

Training and evaluation are performed on two brain tumor datasets comprising 419 paired exams in total:

Pre-Post-OP (PP): Contains 300 pairs of brain MR exams with one timepoint before and one timepoint after tumor resection. It is comprised of the official training+validation dataset of *BraTS-Reg* [4] with 160 cases and three datasets collected at *Klinikum rechts der Isar (TUM)* with 49, 30, and 61 cases, respectively.

Pre-Intra-OP (PI): Contains 119 pairs of brain MR exams with one timepoint before and one timepoint during tumor surgery and originates again from *Klinikum rechts der Isar (TUM)*.

Each exam comprises either three or four of the MR sequences *native T1-weighted (T1)*, *T2 Fluid Attenuated Inversion Recovery (FLAIR)*. We preprocess the data using the *BraTS toolkit* [14], which includes reorientation into the same coordinate system, rigid co-registration to a brain atlas as well as brain extraction. Additionally, intensity normalization is applied to all images. For testing, we use 30 cases for *PP* and 20 cases for *PI*. For each test set, six landmarks annotated by clinical experts are provided for each case, accumulating to 180 landmarks on 30 patients for *PP* and 120 landmarks on 20 patients for *PI*.

3.2 Similarity Metrics

We demonstrate the simultaneous optimization strategy on the similarity metrics *MSE* and *NCC*, which are widely used for medical image registration tasks [6].

3.3 Training and Testing

We perform multi-channel training and testing for *DRN* and *IO*. For *PP*, this includes all four sequences, while for *PI*, the sequences *T1*, *T1CE* and *FLAIR* are available. To overcome this imbalance, we train separate networks for *PP* and *PI*. We split the *PP* dataset into training (80%), validation (10%), and test (10%)

and select the model for testing that shows a minimal loss on the validation set. Since we have much fewer cases available for the experiments on the *PI* dataset, we waive the validation set here and decide on a fixed number of 600 training epochs in all experiments. At test time, the *DRN*'s output deformation field serves as input to the 30 iterations of *IO*. For each dataset, we compare the combined loss of *MSE+NCC* against the individual addends, both for the initial *DRN* training as well as the *IO*. Therefore, we exhaustively explore loss combinations for *DRN* and *IO*, as depicted in Table 1.

The implementation is inspired by [6] using *Pytorch* 1.8.1 [20]. We use learning rates 1*e*-4 and 1*e*-3 for *DRN* and *IO*, respectively. Image similarity metrics are weighted equally in the loss function, summing up to 1 for the similarity term. Smoothness regularization weight is set to 1.0 for all experiments for a fair comparison.

3.4 Reference Methods

We compare our approach with two established methods: a deformable *SyN* registration by *ANTs* [3], and *VM* [6], both using the respective T1-weighted sequences. We implement two variants of *VM* [6], available at [5], that employ *MSE* and *NCC* as respective image similarity metrics. Therefore, we use default learning rate 1*e*-4 and recommended smoothness regularization weights of 0.02 and 1.5 for *MSE* and *NCC*, respectively. Likewise, *ANTs SyN* [3] is implemented using default parameters. Since we opt for deformable registration without prior rigid/affine alignment, we apply the same for *VM* and *ANTs SyN*.

Table 1. Mean *TRE* in *mm* with standard deviation using different losses for *DRN* and *DRN+IO* on the Pre-Post-OP and Pre-Intra-OP test sets. Shown are the results on all possible combinations of loss functions. The best result in each category is highlighted. Statistical comparisons are provided in Table 2. Hit rate curves [24] are shown in Fig. 3.

Loss [DRN]	Loss [IO]	Pre-Post-OP	Pre-Intra-OP
MSE	(not applied)	2.40 ± 1.56	3.48 ± 3.45
MSE	MSE	2.21 ± 1.56	3.27 ± 3.46
MSE	NCC	1.81 ± 1.45	3.24 ± 3.39
MSE	MSE+NCC	$\mathbf{1.76 \pm 1.36}$	$\mathbf{3.17 \pm 3.39}$
NCC	(not applied)	2.18 ± 1.52	3.31 ± 2.68
NCC	MSE	2.06 ± 1.50	2.96 ± 2.78
NCC	NCC	1.80 ± 1.44	2.71 ± 2.79
NCC	MSE+NCC	$\mathbf{1.77 \pm 1.36}$	$\mathbf{2.68 \pm 2.72}$
MSE+NCC	(not applied)	2.19 ± 1.49	3.07 ± 2.96
MSE+NCC	MSE	2.07 ± 1.48	2.94 ± 3.05
MSE+NCC	NCC	1.80 ± 1.44	2.86 ± 2.95
MSE+NCC	MSE+NCC	$\mathbf{1.73 \pm 1.34}$	$\mathbf{2.76 \pm 2.89}$

3.5 Results

Table 1 shows quantitative results on both datasets. The *TREs* indicate that for *PP*, *DRN* only with *NCC* loss achieves the lowest mean error. When coupling *DRN* with *IO*, for all possible combinations, mean *TRE* is always lowest when using *MSE+NCC* loss during *IO*. On *PI*, lowest mean *TRE* for *DRN* only is achieved when combining *MSE* and *NCC* in the loss function. When adding *IO*, *DRN* with *NCC* loss coupled with *IO* using *MSE+NCC* loss shows lowest mean *TRE*. Statistical comparisons based on *Paired Samples T-Tests* are provided in Table 2.

Table 2. *P-values* and 95% *Confidence Intervals (CI)* of *Paired Samples T-Tests* on the *TRE* of competitive methods for the Pre-Post-OP and the Pre-Intra-OP datasets. Methods are given by declaring the used losses for *DRN* and *IO*, respectively.

Method 1	Method 2	p-value [CI] (Pre-Post-OP)	p-value [CI] (Pre-Intra-OP)
DRN (NCC)	DRN (MSE+NCC)	0.85 [−0.09; 0.07]	0.05 [0.00; 0.47]
DRN (MSE)+ IO (NCC)	DRN (MSE)+ IO (MSE+NCC)	0.13 [−0.01; 0.10]	0.23 [−0.04; 0.16]
DRN (NCC)+ IO (NCC)	DRN (NCC)+ IO (MSE+NCC)	0.22 [−0.02; 0.08]	0.50 [−0.06; 0.13]
DRN (MSE+NCC)+ IO (NCC)	DRN (MSE+NCC)+ IO (MSE+NCC)	**0.03 [0.01; 0.13]**	**0.03 [0.01; 0.19]**

Table 3. Comparison with reference methods. Mean *TRE* in *mm* with standard deviation on the Pre-Post-OP and Pre-Intra-OP datasets for the best *DRN+IO* compared to *VM* and ANTs *SyN*. The proposed method is on par with established reference registration methods.

Method	Pre-Post-OP	Pre-Intra-OP
DRN+IO	1.73 ± 1.34	2.68 ± 2.72
VM (MSE)	2.98 ± 2.11	3.66 ± 2.90
VM (NCC)	2.09 ± 1.54	2.37 ± 2.37
ANTs SyN	2.02 ± 1.49	1.81 ± 1.17

Moreover, we have evaluated the experiments with respect to their *hit rate* curves according to [24], see Fig. 3, showing the behavior of the different methods with respect to registration accuracy with increased tolerance.

Evaluating performance with respect to running time, we determine that the runtime mainly depends on the implementation of the respective similarity metric. In our case, *NCC* loss is significantly slower than *MSE*. When combining

the two metrics in the loss function, there is barely any difference in runtime compared to *NCC* alone, since the runtime of *MSE* is overall negligible. This finding will probably change when the respective implementation differs and/or other similarity metrics are used.

Table 3 shows a comparison of the best *DRN+IO* with results achieved by reference methods *VM* and *ANTs*. For *PP*, *DRN+IO* shows lowest mean *TREs*, while for *PI*, this is the case for *ANTs SyN*.

Our findings are in line with how experts visually perceive the registration quality: In a blinded evaluation of registration methods, three expert radiologists independently picked the *MSE+NCC* registration as their favorite. Asked to provide reasoning for their choice, experts cited better registration quality around ventricles and fewer artifacts, such as unrealistically deformed tissue. Figure 4 shows sample qualitative results on *PI*, illustrating these findings.

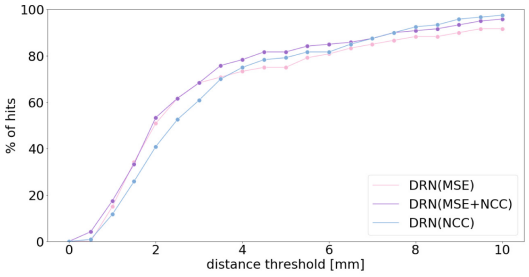
4 Discussion

The contribution of this work is to simultaneously optimize multiple image similarity metrics for image registration tasks in a *DL* setting. For demonstration purposes, we combine the well-established metrics *MSE* and *NCC*. We evaluate *DRNs* with and without *IO* in two challenging multi-modal 3D registration settings, namely pre- to post-operative and pre- to intra-operative glioma *MRI*.

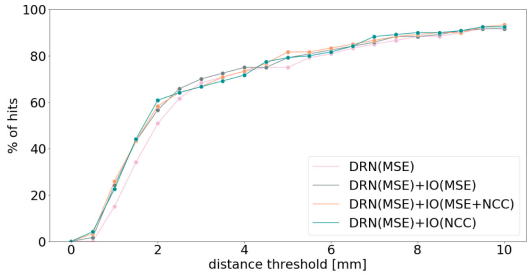
Table 1 shows that combining *MSE* and *NCC* losses consistently improves *IO* performance in both datasets. When applied to *DRN* only, the combination still performs best for *PI* while showing comparable results to *NCC* for *PP*. Furthermore, Table 3 illustrates that our method achieves competitive results compared to established reference methods. Moreover, for *PP*, the proposed approach outperforms all reference methods. For *PI*, *ANTs SyN* performs best. A potential explanation might be lower registration quality introduced by the *FLAIR* images. When using *T1* images only for training and testing, the results are getting more comparable, see Table S1 in the supplementary material. Also, *DRN* as well as *VM* would likely benefit from a bigger training dataset. When performing an additional experiment using the Elastix toolbox [13] for single- vs multi-metric registration, we can observe that combining two metrics improves performance for *PI*, see Table S2 in the supplementary material.

Combining *MSE* and *NCC* can help to improve the alignment of the source image to the target image. The qualitative assessment by clinical experts supports these findings.

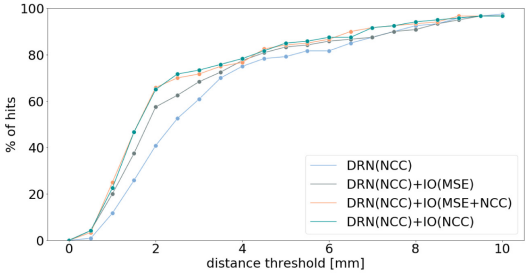
There are several limitations to this study. The evaluation is performed on rather small test sets of 20 for *PI* and 30 for *PP*. The used datasets are fairly specific, so for generalization purposes, an extension toward other registration tasks promises to yield additional insights. Here, we focus on multi-sequence training on all available MR sequences using a 4D implementation of the *NCC* loss. Training on different sequences separately will possibly indicate increased understanding on the respective influences. Performing cross validation for model



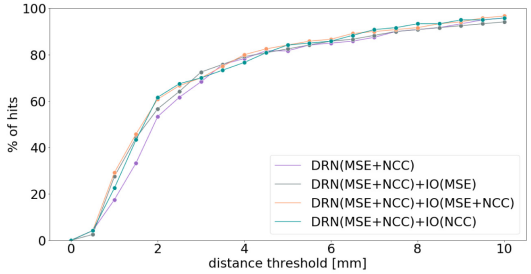
(a) PI: Percentage of *hits* for three *DRN* using different losses.



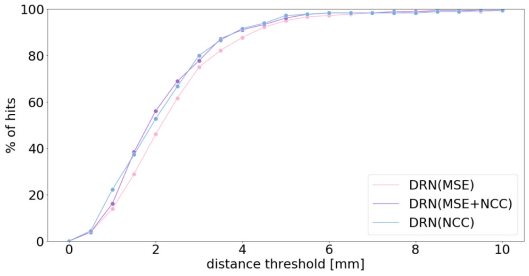
(b) PI: Percentage of *hits* for *DRN* using *MSE* and different *IO* losses.



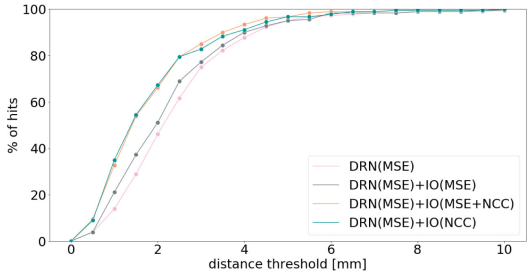
(c) PI: Percentage of *hits* for *DRN* using *NCC* and different *IO* losses.



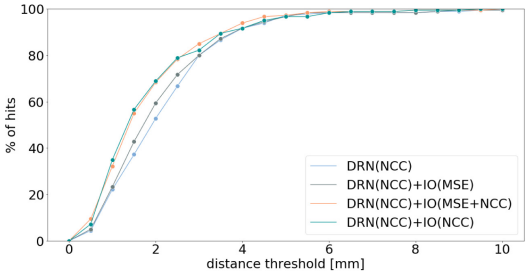
(d) PI: Percentage of *hits* for *DRN* using *MSE+NCC* and different *IO* losses.



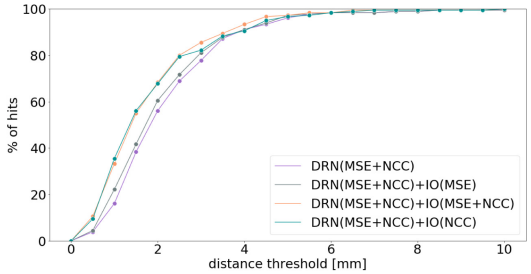
(e) Hits for three *DRN* using different losses.



(f) PP: Percentage of *hits* for *DRN* using *MSE* and different *IO* losses.



(g) PP: Percentage of *hits* for *DRN* using *NCC* and different *IO* losses.



(h) PP: Percentage of *hits* for *DRN* using *MSE+NCC* and different *IO* losses.

Fig. 3. Hit rate curves [24] on PI and PP, respectively. *Hit* denotes that a warped landmark lies within a certain distance threshold to the respective expert landmark annotation. Here, evaluation thresholds are set every $0.5mm$, *hit* percentages in between are interpolated.

selection might improve stability in the results. Even though the regularization could be fine-tuned, we use a fixed weight on the deformation field in our approach.

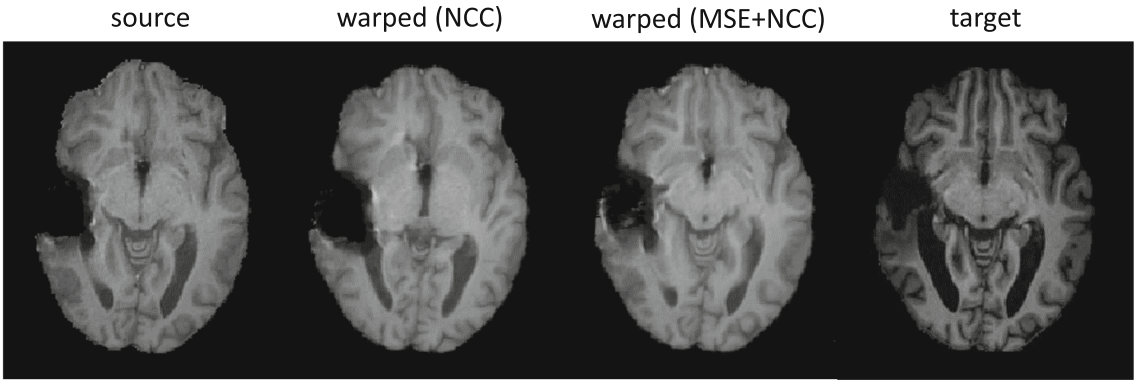


Fig. 4. Sample registration results on the Pre-Intra-OP dataset. Registration with combined losses *MSE+NCC* shows improved registration in comparison to registration with *NCC* only, especially nearby the tumor and the resection cavity, respectively.

5 Conclusion and Future Work

We propose an optimization strategy – here implemented by means of a simple summation – that benefits registration performance with regard to *TRE*. Moreover, we conduct extensive comparisons against established reference methods. Future work should investigate the addition of further similarity metrics such as *Mutual Information*. Besides, introducing different weights for the individual parts of the loss function (instead of equally weighting) might further enhance registration performance.

Acknowledgements. Supported by Deutsche Forschungsgemeinschaft (DFG) through TUM International Graduate School of Science and Engineering (IGSSE), GSC 81. BM, BW and FK are supported through the SFB 824, subproject B12. BM acknowledges support by the Helmut Horten Foundation. Finally, we acknowledge Andreas Poschenrieder and Anna Valentina Lioba Eleonora Claire Javid Mamasani for eye-opening insights.

References

1. Avants, B., et al.: Multivariate analysis of structural and diffusion imaging in traumatic brain injury. Acad. Radiol. **15**(11), 1360–1375 (2008)
2. Avants, B.B., Tustison, N.J., Song, G., Cook, P.A., Klein, A., Gee, J.C.: A reproducible evaluation of ants similarity metric performance in brain image registration. Neuroimage **54**(3), 2033–2044 (2011)
3. Avants, B.B., Tustison, N., Song, G., et al.: Advanced normalization tools (ANTS). Insight J. **2**(365), 1–35 (2009)
4. Baheti, B., et al.: The brain tumor sequence registration challenge: establishing correspondence between pre-operative and follow-up MRI scans of diffuse glioma patients. arXiv preprint arXiv:2112.06979 (2021)
5. Balakrishnan, G., Zhao, A., Sabuncu, M., Guttag, J., Dalca, A.V.: voxelmorph: learning-based image registration. https://github.com/voxelmorph/voxelmorph

6. Balakrishnan, G., Zhao, A., Sabuncu, M.R., Guttag, J., Dalca, A.V.: Voxelmorph: a learning framework for deformable medical image registration. IEEE Trans. Med. Imaging **38**(8), 1788–1800 (2019)
7. Boukellouz, W., Moussaoui, A.: Evaluation of several similarity measures for deformable image registration using T1-weighted MR images of the brain. In: 2017 5th International Conference on Electrical Engineering-Boumerdes (ICEE-B), pp. 1–5. IEEE (2017)
8. Canalini, L., Klein, J., Gerken, A., Heldmann, S., Hering, A., Hahn, H.K.: Iterative method to register longitudinal MRI acquisitions in neurosurgical context. In: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries, pp. 262–272. Springer, Cham (2023)
9. Chen, J., Frey, E.C., He, Y., Segars, W.P., Li, Y., Du, Y.: Transmorph: transformer for unsupervised medical image registration. Med. Image Anal. **82**, 102615 (2022)
10. Ferrante, E., Dokania, P.K., Marini, R., Paragios, N.: Deformable registration through learning of context-specific metric aggregation. In: Wang, Q., Shi, Y., Suk, H.-I., Suzuki, K. (eds.) MLMI 2017. LNCS, vol. 10541, pp. 256–265. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-67389-9_30
11. Fidon, L., et al.: A dempster-shafer approach to trustworthy AI with application to fetal brain MRI segmentation. arXiv preprint arXiv:2204.02779 (2022)
12. Großbröhmer, C., Siebert, H., Hansen, L., Heinrich, M.P.: Employing convexadam for brats-reg. In: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries, pp. 252–261. Springer, Cham (2023)
13. Klein, S., Staring, M., Murphy, K., Viergever, M.A., Pluim, J.P.: Elastix: a toolbox for intensity-based medical image registration. IEEE Trans. Med. Imaging **29**(1), 196–205 (2009)
14. Kofler, F., et al.: Brats toolkit: translating brats brain tumor segmentation algorithms into clinical and scientific practice. Front. Neurosci. 125 (2020)
15. Kofler, F., et al.: Are we using appropriate segmentation metrics? Identifying correlates of human expert perception for CNN training beyond rolling the dice coefficient. Mach. Learn. Biomed. Imaging **2**(May 2023 issue), 27–71 (2023)
16. Kofler, F., et al.: blob loss: instance imbalance aware loss functions for semantic segmentation. In: International Conference on Information Processing in Medical Imaging, pp. 755–767. Springer, Cham (2023)
17. Meng, M., Bi, L., Feng, D., Kim, J.: Brain tumor sequence registration with non-iterative coarse-to-fine networks and dual deep supervision. In: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries, pp. 273–282. Springer, Cham (2023)
18. Mok, T.C., Chung, A.: Robust image registration with absent correspondences in pre-operative and follow-up brain MRI scans of diffuse glioma patients. In: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries, pp. 231–240. Springer, Cham (2023)
19. Mok, T.C.W., Chung, A.C.S.: Large deformation diffeomorphic image registration with laplacian pyramid networks. In: Martel, A.L., et al. (eds.) MICCAI 2020. LNCS, vol. 12263, pp. 211–221. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-59716-0_21
20. Paszke, A., et al.: PyTorch: an imperative style, high-performance deep learning library. In: Wallach, H., Larochelle, H., Beygelzimer, A., d'Alché Buc, F., Fox, E., Garnett, R. (eds.) Advances in Neural Information Processing Systems 32, pp. 8024–8035. Curran Associates, Inc. (2019). http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf

21. Rueckert, D., Sonoda, L.I., Hayes, C., Hill, D.L., Leach, M.O., Hawkes, D.J.: Non-rigid registration using free-form deformations: application to breast MR images. IEEE Trans. Med. Imaging **18**(8), 712–721 (1999)
22. Uss, M.L., Vozel, B., Abramov, S.K., Chehdi, K.: Selection of a similarity measure combination for a wide range of multimodal image registration cases. IEEE Trans. Geosci. Remote Sens. **59**(1), 60–75 (2020)
23. Wachs, J., Stern, H., Burks, T., Alchanatis, V.: Multi-modal registration using a combined similarity measure. Appl. Soft Comput. **52**, 159–168 (2009)
24. Waldmannstetter, D., et al.: Framing image registration as a landmark detection problem for better representation of clinical relevance. arXiv preprint arXiv:2308.01318 (2023)
25. Wodzinski, M., Jurgas, A., Marini, N., Atzori, M., Müller, H.: Unsupervised method for intra-patient registration of brain magnetic resonance images based on objective function weighting by inverse consistency: Contribution to the brats-reg challenge. In: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries, pp. 241–251. Springer, Cham (2023)
26. Zhou, J., Liu, Q.: A combined similarity measure for multimodal image registration. In: 2015 IEEE International Conference on Imaging Systems and Techniques (IST), pp. 1–5. IEEE (2015)

Part III

CONCLUDING REMARKS

9 DISCUSSION

Despite many recent advances in the field of medical image registration, transforming images with pathologies is still a major challenge. This is particularly the case for the alignment of brain tumor images before and after surgery, since the tumor itself as well as the resection procedure can cause immense tissue shift and deformation. The goal of this thesis is to present efficient methods for performing image registration on images of the brain undergoing tumor resection. A special focus is put on the development of DL approaches, especially in the context of aligning image landmarks using DRL and the simultaneous optimization of similarity metrics in an unsupervised CNN. Two peer-reviewed first-author publications constitute the core of this publication-based thesis, presented in the Chapters 7 and 8. The proposed approaches address challenges related to the alignment of pre-, post- and intra-operative brain tumor scans from two different aspects, as described below.

REDETECTION OF LANDMARKS USING ANATOMICAL GUIDANCE

Focusing on the redetection of landmarks for enhancing the alignment of pre- and post-operative brain scans, a framework based on DRL has been designed. Therefore, the specific anatomical particularities related to tumor resection have been incorporated by extending an artificial RL agent towards patient-specific anatomical guidance. When being compared to the baseline approach without involving the anatomy, the proposed approach demonstrated superior performance w.r.t. the average accuracy and overall robustness in finding the respective target landmark. Furthermore, the presented methodology opens avenues for further refinement of anatomical guidance, particularly in the analysis of temporal changes before and after tumor resection. This, in turn, facilitates the development of a more representative and efficient model for understanding anatomical transformations.

SIMULTANEOUS OPTIMIZATION OF IMAGE SIMILARITY METRICS

Finding a suitable similarity metric when performing image registration is a crucial step in the development of any approach for image alignment and requires careful consideration. The proposed approach on the simultaneous optimization of multiple image similarity metrics illustrated the potential to benefit the registration process due to the combined strengths of the different metrics, when being applied during the training of an unsupervised CNN. The findings on quanti-

tative accuracy enhancement have been confirmed through the visual evaluation by clinical experts, demonstrating the effectiveness of the presented technique for improving image alignment in the context of tumor resection.

Complementing the contents that are described above, a study on state-of-the-art algorithms for the registration of pre- and post-operative images of brain tumors has been conducted in the *Brain Tumor Sequence Registration (BraTS-Reg) Challenge*, that is described in Appendix A. The scope of this challenge was to collect image registration algorithms that are specified on the alignment of brain image data, comprising baseline and follow-up scans before and after tumor resection. The collected algorithms have been evaluated using a combination of different evaluation metrics on handcrafted expert landmark annotations, giving a broad overview of the methods' capabilities and weaknesses. Therefore, it serves as a public benchmark for deformable image registration, dealing with highly demanding imaging data. The results of this study are prepared to be published within the scope of a journal paper in the future. An earlier version of the manuscript is available as an *arXiv preprint* [8].

Inspired by the results of the *BraTS-Reg Challenge*, the value of suitable evaluation metrics for measuring the accuracy of image registration has been explored in the study provided in Appendix B, proposing a novel way of registration evaluation towards a more clinically relevant assessment of target registration errors. By incorporating inter-rater variability, the introduced method aims to compute *hit rates* of evaluated algorithms, addressing the inherent label noise in image annotations and allowing for case-specific evaluation of a method's clinical relevance. The findings of this study are meant to be formally published in a journal paper at a later date. An earlier version of the manuscript is available as an *arXiv preprint* [131].

The focus of this thesis lies in the development of techniques aimed at improving image alignment in the context of tumor pathologies and resection procedures. The integration of DL into image registration has demonstrated notable effectiveness and efficiency, contributing significant advancements also to the general field of medical image registration. However, an inherent limitation of DL methods in this domain is their reliance on extensive training data, especially for specific cases involving images with substantial deformations caused by tumors. This limitation was particularly evident in the proposed simultaneous optimization approach [132], particularly concerning the dataset involving pre- and intra-operative brain scans. Here, a larger volume of training data is believed to be crucial for optimal performance.

Moreover, a meaningful evaluation of registration results remains an ongoing challenge, constituting a primary limitation in assessing

the value and significance of image alignment methods, as they are presented in this thesis. Achieving consistent and reproducible analysis of registration outputs is imperative for establishing actual clinical relevance when evaluating image registration algorithms. This necessitates the incorporation of diverse quantitative and qualitative validation metrics, aiming for a comprehensive assessment of significance in the context of medical image alignment.

10 CONCLUSION AND OUTLOOK

This dissertation is centered on advancing image registration, both in a general sense and specifically within the challenging context of anatomical variations resulting from brain tumor resection. The primary contributions in this thesis leverage Reinforcement Learning, Deep Learning methodologies, and a concurrent optimization technique of similarity metrics.

DL has already demonstrated its efficacy in enhancing deformable image registration, particularly in expediting this intricate and time-consuming process [9, 20]. Future work should therefore focus on further exploring DL techniques, particularly on the incorporation of dynamic changes in brain anatomy. This entails, for example, automating anatomical guidance through the application of segmentation frameworks [62, 145]. Additionally, a dense representation of landmark annotations would further propel the development of patient-specific models, while also ensuring a more reliable and robust evaluation of image registration.

A main contribution of this thesis lies in the simultaneous optimization of image similarity metrics to enhance image alignment in a DL framework. While the proposed method has demonstrated its potential by combining two metrics, the inclusion of additional metrics and an extension towards case-specific weight adaptation remain subjects for future investigation.

From a clinical perspective, a thorough evaluation and analysis of the proposed methods across larger patient cohorts with diverse brain pathologies are imperative. This is essential for establishing methodologies that are broadly applicable, robust to unforeseen changes in a patient's anatomy, and consistently reliable in their outcomes.

Part IV

APPENDIX

A

THE BRAIN TUMOR SEQUENCE REGISTRATION (BRATS-REG) CHALLENGE: ESTABLISHING CORRESPONDENCE BETWEEN PRE-OPERATIVE AND FOLLOW-UP MRI SCANS OF DIFFUSE GLIOMA PATIENTS

The manuscript has been written during the course of this thesis, but is not subject to evaluation. The latest version of the manuscript is available as an *arXiv preprint* [8].

B

FRAMING IMAGE REGISTRATION AS A LANDMARK DETECTION PROBLEM FOR LABEL-NOISE-AWARE TASK REPRESENTATION (HITR)

The manuscript has been written during the course of this thesis, but is not subject to evaluation. The latest version of the manuscript is avaliable as an *arXiv preprint* [131].

C

LICENSES

On the following pages, the licenses permitting the use of the publications in this thesis are provided.

SPRINGER NATURE LICENSE
TERMS AND CONDITIONS

Jan 02, 2025

This Agreement between Ms. Diana Waldmannstetter ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

License Number	5744071158077
License date	Mar 08, 2024
Licensed Content Publisher	Springer Nature
Licensed Content Publication	Springer eBook
Licensed Content Title	Reinforced Redetection of Landmark in Pre- and Post-operative Brain Scan Using Anatomical Guidance for Image Alignment
Licensed Content Author	Diana Waldmannstetter, Fernando Navarro, Benedikt Wiestler et al
Licensed Content Date	Jan 1, 2020
Type of Use	Thesis/Dissertation
Requestor type	academic/university or research institute
Format	print and electronic
Portion	full article/chapter
Will you be translating?	no

Circulation/distribution	1 - 29
Author of this Springer Nature content	yes
Title of new work	Medical Image Registration for Brain Tumor Surgery using Deep Learning
Institution name	Technical University of Munich
Expected presentation date	Jul 2024
Requestor Location	Ms. Diana Waldmannstetter [redacted] [redacted]
Billing Type	Invoice
Billing Address	Ms. Diana Waldmannstetter [redacted] [redacted]
Total	0.00 EUR

Terms and Conditions

Springer Nature Customer Service Centre GmbH Terms and Conditions

The following terms and conditions ("Terms and Conditions") together with the terms specified in your [RightsLink] constitute the License ("License") between you as Licensee and Springer Nature Customer Service Centre GmbH as Licensor. By clicking 'accept' and completing the transaction for your use of the material ("Licensed Material"), you confirm your acceptance of and obligation to be bound by these Terms and Conditions.

1. Grant and Scope of License

1. 1. The Licensor grants you a personal, non-exclusive, non-transferable, non-sublicensable, revocable, world-wide License to reproduce, distribute, communicate

to the public, make available, broadcast, electronically transmit or create derivative works using the Licensed Material for the purpose(s) specified in your RightsLink Licence Details only. Licenses are granted for the specific use requested in the order and for no other use, subject to these Terms and Conditions. You acknowledge and agree that the rights granted to you under this License do not include the right to modify, edit, translate, include in collective works, or create derivative works of the Licensed Material in whole or in part unless expressly stated in your RightsLink Licence Details. You may use the Licensed Material only as permitted under this Agreement and will not reproduce, distribute, display, perform, or otherwise use or exploit any Licensed Material in any way, in whole or in part, except as expressly permitted by this License.

1. 2. You may only use the Licensed Content in the manner and to the extent permitted by these Terms and Conditions, by your RightsLink Licence Details and by any applicable laws.

1. 3. A separate license may be required for any additional use of the Licensed Material, e.g. where a license has been purchased for print use only, separate permission must be obtained for electronic re-use. Similarly, a License is only valid in the language selected and does not apply for editions in other languages unless additional translation rights have been granted separately in the License.

1. 4. Any content within the Licensed Material that is owned by third parties is expressly excluded from the License.

1. 5. Rights for additional reuses such as custom editions, computer/mobile applications, film or TV reuses and/or any other derivative rights requests require additional permission and may be subject to an additional fee. Please apply to journalpermissions@springernature.com or bookpermissions@springernature.com for these rights.

2. Reservation of Rights

Licensor reserves all rights not expressly granted to you under this License. You acknowledge and agree that nothing in this License limits or restricts Licensor's rights in or use of the Licensed Material in any way. Neither this License, nor any act, omission, or statement by Licensor or you, conveys any ownership right to you in any Licensed Material, or to any element or portion thereof. As between Licensor and you, Licensor owns and retains all right, title, and interest in and to the Licensed Material subject to the license granted in Section 1.1. Your permission to use the Licensed Material is expressly conditioned on you not impairing Licensor's or the applicable copyright owner's rights in the Licensed Material in any way.

3. Restrictions on use

3. 1. Minor editing privileges are allowed for adaptations for stylistic purposes or formatting purposes provided such alterations do not alter the original meaning or intention of the Licensed Material and the new figure(s) are still accurate and representative of the Licensed Material. Any other changes including but not limited to, cropping, adapting, and/or omitting material that affect the meaning, intention or moral rights of the author(s) are strictly prohibited.

3. 2. You must not use any Licensed Material as part of any design or trademark.

3. 3. Licensed Material may be used in Open Access Publications (OAP), but any

such reuse must include a clear acknowledgment of this permission visible at the same time as the figures/tables/illustration or abstract and which must indicate that the Licensed Material is not part of the governing OA license but has been reproduced with permission. This may be indicated according to any standard referencing system but must include at a minimum 'Book/Journal title, Author, Journal Name (if applicable), Volume (if applicable), Publisher, Year, reproduced with permission from SNCSC'.

4. STM Permission Guidelines

4. 1. An alternative scope of license may apply to signatories of the STM Permissions Guidelines ("STM PG") as amended from time to time and made available at https://www.stm-assoc.org/intellectual-property/permissions/permissions-guidelines/.

4. 2. For content reuse requests that qualify for permission under the STM PG, and which may be updated from time to time, the STM PG supersede the terms and conditions contained in this License.

4. 3. If a License has been granted under the STM PG, but the STM PG no longer apply at the time of publication, further permission must be sought from the Rightsholder. Contact journalpermissions@springernature.com or bookpermissions@springernature.com for these rights.

5. Duration of License

5. 1. Unless otherwise indicated on your License, a License is valid from the date of purchase ("License Date") until the end of the relevant period in the below table:

Reuse in a medical communications project	Reuse up to distribution or time period indicated in License
Reuse in a dissertation/ thesis	Lifetime of thesis
Reuse in a journal/ magazine	Lifetime of journal/magazine
Reuse in a book/ textbook	Lifetime of edition
Reuse on a website	1 year unless otherwise specified in the License
Reuse in a presentation/ slide kit/poster	Lifetime of presentation/slide kit/poster. Note: publication whether electronic or in print of presentation/slide kit/poster may require further permission.
Reuse in conference proceedings	Lifetime of conference proceedings
Reuse in an annual report	Lifetime of annual report
Reuse in training/CME materials	Reuse up to distribution or time period indicated in License
Reuse in newsmedia	Lifetime of newsmedia
Reuse in coursepack/ classroom materials	Reuse up to distribution and/or time period indicated in license

6. Acknowledgement

6. 1. The Licensor's permission must be acknowledged next to the Licensed Material in print. In electronic form, this acknowledgement must be visible at the same time as the figures/tables/illustrations or abstract and must be hyperlinked to the journal/book's homepage.

6. 2. Acknowledgement may be provided according to any standard referencing system and at a minimum should include "Author, Article/Book Title, Journal name/Book imprint, volume, page number, year, Springer Nature".

7. Reuse in a dissertation or thesis

7. 1. Where 'reuse in a dissertation/thesis' has been selected, the following terms apply: Print rights of the Version of Record are provided for; electronic rights for use only on institutional repository as defined by the Sherpa guideline (www.sherpa.ac.uk/romeo/) and only up to what is required by the awarding institution.

7. 2. For theses published under an ISBN or ISSN, separate permission is required. Please contact journalpermissions@springernature.com or bookpermissions@springernature.com for these rights.

7. 3. Authors must properly cite the published manuscript in their thesis according to current citation standards and include the following acknowledgement: '*Reproduced with permission from Springer Nature*'.

8. License Fee

You must pay the fee set forth in the License Agreement (the "License Fees"). All amounts payable by you under this License are exclusive of any sales, use, withholding, value added or similar taxes, government fees or levies or other assessments. Collection and/or remittance of such taxes to the relevant tax authority shall be the responsibility of the party who has the legal obligation to do so.

9. Warranty

9. 1. The Licensor warrants that it has, to the best of its knowledge, the rights to license reuse of the Licensed Material. **You are solely responsible for ensuring that the material you wish to license is original to the Licensor and does not carry the copyright of another entity or third party (as credited in the published version).** If the credit line on any part of the Licensed Material indicates that it was reprinted or adapted with permission from another source, then you should seek additional permission from that source to reuse the material.

9. 2. EXCEPT FOR THE EXPRESS WARRANTY STATED HEREIN AND TO THE EXTENT PERMITTED BY APPLICABLE LAW, LICENSOR PROVIDES THE LICENSED MATERIAL "AS IS" AND MAKES NO OTHER REPRESENTATION OR WARRANTY. LICENSOR EXPRESSLY DISCLAIMS ANY LIABILITY FOR ANY CLAIM ARISING FROM OR OUT OF THE CONTENT, INCLUDING BUT NOT LIMITED TO ANY ERRORS, INACCURACIES, OMISSIONS, OR DEFECTS CONTAINED THEREIN, AND ANY IMPLIED OR EXPRESS WARRANTY AS TO MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. IN NO EVENT SHALL LICENSOR BE LIABLE TO YOU OR ANY OTHER PARTY OR ANY OTHER PERSON OR FOR ANY SPECIAL, CONSEQUENTIAL, INCIDENTAL, INDIRECT,

PUNITIVE, OR EXEMPLARY DAMAGES, HOWEVER CAUSED, ARISING OUT OF OR IN CONNECTION WITH THE DOWNLOADING, VIEWING OR USE OF THE LICENSED MATERIAL REGARDLESS OF THE FORM OF ACTION, WHETHER FOR BREACH OF CONTRACT, BREACH OF WARRANTY, TORT, NEGLIGENCE, INFRINGEMENT OR OTHERWISE (INCLUDING, WITHOUT LIMITATION, DAMAGES BASED ON LOSS OF PROFITS, DATA, FILES, USE, BUSINESS OPPORTUNITY OR CLAIMS OF THIRD PARTIES), AND WHETHER OR NOT THE PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. THIS LIMITATION APPLIES NOTWITHSTANDING ANY FAILURE OF ESSENTIAL PURPOSE OF ANY LIMITED REMEDY PROVIDED HEREIN.

10. Termination and Cancellation

10. 1. The License and all rights granted hereunder will continue until the end of the applicable period shown in Clause 5.1 above. Thereafter, this license will be terminated and all rights granted hereunder will cease.

10. 2. Licensor reserves the right to terminate the License in the event that payment is not received in full or if you breach the terms of this License.

11. General

11. 1. The License and the rights and obligations of the parties hereto shall be construed, interpreted and determined in accordance with the laws of the Federal Republic of Germany without reference to the stipulations of the CISG (United Nations Convention on Contracts for the International Sale of Goods) or to Germany´s choice-of-law principle.

11. 2. The parties acknowledge and agree that any controversies and disputes arising out of this License shall be decided exclusively by the courts of or having jurisdiction for Heidelberg, Germany, as far as legally permissible.

11. 3. This License is solely for Licensor's and Licensee's benefit. It is not for the benefit of any other person or entity.

Questions? For questions on Copyright Clearance Center accounts or website issues please contact springernaturesupport@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777. For questions on Springer Nature licensing please visit https://www.springernature.com/gp/partners/rights-permissions-third-party-distribution

Other Conditions:

Version 1.4 - Dec 2022

Questions? customercare@copyright.com.

SPRINGER NATURE LICENSE
TERMS AND CONDITIONS

Jan 02, 2025

This Agreement between Diana Waldmannstetter, Technical University of Munich ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

License Number	5940741225527
License date	Jan 02, 2025
Licensed Content Publisher	Springer Nature
Licensed Content Publication	Springer eBook
Licensed Content Title	Primitive Simultaneous Optimization of Similarity Metrics for Image Registration
Licensed Content Author	Diana Waldmannstetter, Benedikt Wiestler, Julian Schwarting et al
Licensed Content Date	Jan 1, 2024
Type of Use	Thesis/Dissertation
Requestor type	academic/university or research institute
Format	print and electronic
Portion	full article/chapter
Will you be translating?	no
Circulation/distribution	1 - 29

Author of this Springer Nature content	yes
Title of new work	Medical Image Registration for Brain Tumor Surgery using Deep Learning
Institution name	Technical University of Munich
Expected presentation date	Jan 2025
The Requesting Person / Organization to Appear on the License	Diana Waldmannstetter, Technical University of Munich
Requestor Location	Ms. Diana Waldmannstetter [redacted] [redacted]
Billing Type	Invoice
Billing Address	Ms. Diana Waldmannstetter [redacted] [redacted]
Total	0.00 EUR

Terms and Conditions

Springer Nature Customer Service Centre GmbH Terms and Conditions

The following terms and conditions ("Terms and Conditions") together with the terms specified in your [RightsLink] constitute the License ("License") between you as Licensee and Springer Nature Customer Service Centre GmbH as Licensor. By clicking 'accept' and completing the transaction for your use of the material ("Licensed Material"), you confirm your acceptance of and obligation to be bound by these Terms and Conditions.

1. Grant and Scope of License

1. 1. The Licensor grants you a personal, non-exclusive, non-transferable, non-sublicensable, revocable, world-wide License to reproduce, distribute, communicate to the public, make available, broadcast, electronically transmit or create derivative

works using the Licensed Material for the purpose(s) specified in your RightsLink Licence Details only. Licenses are granted for the specific use requested in the order and for no other use, subject to these Terms and Conditions. You acknowledge and agree that the rights granted to you under this License do not include the right to modify, edit, translate, include in collective works, or create derivative works of the Licensed Material in whole or in part unless expressly stated in your RightsLink Licence Details. You may use the Licensed Material only as permitted under this Agreement and will not reproduce, distribute, display, perform, or otherwise use or exploit any Licensed Material in any way, in whole or in part, except as expressly permitted by this License.

1. 2. You may only use the Licensed Content in the manner and to the extent permitted by these Terms and Conditions, by your RightsLink Licence Details and by any applicable laws.

1. 3. A separate license may be required for any additional use of the Licensed Material, e.g. where a license has been purchased for print use only, separate permission must be obtained for electronic re-use. Similarly, a License is only valid in the language selected and does not apply for editions in other languages unless additional translation rights have been granted separately in the License.

1. 4. Any content within the Licensed Material that is owned by third parties is expressly excluded from the License.

1. 5. Rights for additional reuses such as custom editions, computer/mobile applications, film or TV reuses and/or any other derivative rights requests require additional permission and may be subject to an additional fee. Please apply to journalpermissions@springernature.com or bookpermissions@springernature.com for these rights.

2. Reservation of Rights

Licensor reserves all rights not expressly granted to you under this License. You acknowledge and agree that nothing in this License limits or restricts Licensor's rights in or use of the Licensed Material in any way. Neither this License, nor any act, omission, or statement by Licensor or you, conveys any ownership right to you in any Licensed Material, or to any element or portion thereof. As between Licensor and you, Licensor owns and retains all right, title, and interest in and to the Licensed Material subject to the license granted in Section 1.1. Your permission to use the Licensed Material is expressly conditioned on you not impairing Licensor's or the applicable copyright owner's rights in the Licensed Material in any way.

3. Restrictions on use

3. 1. Minor editing privileges are allowed for adaptations for stylistic purposes or formatting purposes provided such alterations do not alter the original meaning or intention of the Licensed Material and the new figure(s) are still accurate and representative of the Licensed Material. Any other changes including but not limited to, cropping, adapting, and/or omitting material that affect the meaning, intention or moral rights of the author(s) are strictly prohibited.

3. 2. You must not use any Licensed Material as part of any design or trademark.

3. 3. Licensed Material may be used in Open Access Publications (OAP), but any such reuse must include a clear acknowledgment of this permission visible at the

same time as the figures/tables/illustration or abstract and which must indicate that the Licensed Material is not part of the governing OA license but has been reproduced with permission. This may be indicated according to any standard referencing system but must include at a minimum 'Book/Journal title, Author, Journal Name (if applicable), Volume (if applicable), Publisher, Year, reproduced with permission from SNCSC'.

4. STM Permission Guidelines

4. 1. An alternative scope of license may apply to signatories of the STM Permissions Guidelines ("STM PG") as amended from time to time and made available at https://www.stm-assoc.org/intellectual-property/permissions/permissions-guidelines/.

4. 2. For content reuse requests that qualify for permission under the STM PG, and which may be updated from time to time, the STM PG supersede the terms and conditions contained in this License.

4. 3. If a License has been granted under the STM PG, but the STM PG no longer apply at the time of publication, further permission must be sought from the Rightsholder. Contact journalpermissions@springernature.com or bookpermissions@springernature.com for these rights.

5. Duration of License

5. 1. Unless otherwise indicated on your License, a License is valid from the date of purchase ("License Date") until the end of the relevant period in the below table:

Reuse in a medical communications project	Reuse up to distribution or time period indicated in License
Reuse in a dissertation/ thesis	Lifetime of thesis
Reuse in a journal/ magazine	Lifetime of journal/magazine
Reuse in a book/ textbook	Lifetime of edition
Reuse on a website	1 year unless otherwise specified in the License
Reuse in a presentation/ slide kit/poster	Lifetime of presentation/slide kit/poster. Note: publication whether electronic or in print of presentation/slide kit/poster may require further permission.
Reuse in conference proceedings	Lifetime of conference proceedings
Reuse in an annual report	Lifetime of annual report
Reuse in training/CME materials	Reuse up to distribution or time period indicated in License
Reuse in newsmedia	Lifetime of newsmedia
Reuse in coursepack/ classroom materials	Reuse up to distribution and/or time period indicated in license

6. Acknowledgement

6. 1. The Licensor's permission must be acknowledged next to the Licensed Material in print. In electronic form, this acknowledgement must be visible at the same time as the figures/tables/illustrations or abstract and must be hyperlinked to the journal/book's homepage.

6. 2. Acknowledgement may be provided according to any standard referencing system and at a minimum should include "Author, Article/Book Title, Journal name/Book imprint, volume, page number, year, Springer Nature".

7. Reuse in a dissertation or thesis

7. 1. Where 'reuse in a dissertation/thesis' has been selected, the following terms apply: Print rights of the Version of Record are provided for; electronic rights for use only on institutional repository as defined by the Sherpa guideline (www.sherpa.ac.uk/romeo/) and only up to what is required by the awarding institution.

7. 2. For theses published under an ISBN or ISSN, separate permission is required. Please contact journalpermissions@springernature.com or bookpermissions@springernature.com for these rights.

7. 3. Authors must properly cite the published manuscript in their thesis according to current citation standards and include the following acknowledgement: '*Reproduced with permission from Springer Nature*'.

8. License Fee

You must pay the fee set forth in the License Agreement (the "License Fees"). All amounts payable by you under this License are exclusive of any sales, use, withholding, value added or similar taxes, government fees or levies or other assessments. Collection and/or remittance of such taxes to the relevant tax authority shall be the responsibility of the party who has the legal obligation to do so.

9. Warranty

9. 1. The Licensor warrants that it has, to the best of its knowledge, the rights to license reuse of the Licensed Material. **You are solely responsible for ensuring that the material you wish to license is original to the Licensor and does not carry the copyright of another entity or third party (as credited in the published version).** If the credit line on any part of the Licensed Material indicates that it was reprinted or adapted with permission from another source, then you should seek additional permission from that source to reuse the material.

9. 2. EXCEPT FOR THE EXPRESS WARRANTY STATED HEREIN AND TO THE EXTENT PERMITTED BY APPLICABLE LAW, LICENSOR PROVIDES THE LICENSED MATERIAL "AS IS" AND MAKES NO OTHER REPRESENTATION OR WARRANTY. LICENSOR EXPRESSLY DISCLAIMS ANY LIABILITY FOR ANY CLAIM ARISING FROM OR OUT OF THE CONTENT, INCLUDING BUT NOT LIMITED TO ANY ERRORS, INACCURACIES, OMISSIONS, OR DEFECTS CONTAINED THEREIN, AND ANY IMPLIED OR EXPRESS WARRANTY AS TO MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. IN NO EVENT SHALL LICENSOR BE LIABLE TO YOU OR ANY OTHER PARTY OR ANY OTHER PERSON OR FOR ANY SPECIAL, CONSEQUENTIAL, INCIDENTAL, INDIRECT, PUNITIVE, OR EXEMPLARY DAMAGES, HOWEVER CAUSED, ARISING

OUT OF OR IN CONNECTION WITH THE DOWNLOADING, VIEWING OR USE OF THE LICENSED MATERIAL REGARDLESS OF THE FORM OF ACTION, WHETHER FOR BREACH OF CONTRACT, BREACH OF WARRANTY, TORT, NEGLIGENCE, INFRINGEMENT OR OTHERWISE (INCLUDING, WITHOUT LIMITATION, DAMAGES BASED ON LOSS OF PROFITS, DATA, FILES, USE, BUSINESS OPPORTUNITY OR CLAIMS OF THIRD PARTIES), AND WHETHER OR NOT THE PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. THIS LIMITATION APPLIES NOTWITHSTANDING ANY FAILURE OF ESSENTIAL PURPOSE OF ANY LIMITED REMEDY PROVIDED HEREIN.

10. Termination and Cancellation

10. 1. The License and all rights granted hereunder will continue until the end of the applicable period shown in Clause 5.1 above. Thereafter, this license will be terminated and all rights granted hereunder will cease.

10. 2. Licensor reserves the right to terminate the License in the event that payment is not received in full or if you breach the terms of this License.

11. General

11. 1. The License and the rights and obligations of the parties hereto shall be construed, interpreted and determined in accordance with the laws of the Federal Republic of Germany without reference to the stipulations of the CISG (United Nations Convention on Contracts for the International Sale of Goods) or to Germany´s choice-of-law principle.

11. 2. The parties acknowledge and agree that any controversies and disputes arising out of this License shall be decided exclusively by the courts of or having jurisdiction for Heidelberg, Germany, as far as legally permissible.

11. 3. This License is solely for Licensor's and Licensee's benefit. It is not for the benefit of any other person or entity.

Questions? For questions on Copyright Clearance Center accounts or website issues please contact springernaturesupport@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777. For questions on Springer Nature licensing please visit https://www.springernature.com/gp/partners/rights-permissions-third-party-distribution

Other Conditions:

Version 1.4 - Dec 2022

Questions? customercare@copyright.com.

BIBLIOGRAPHY

[1] Mahmoud Khaled Abd-Ellah, Ali Ismail Awad, Ashraf AM Khalaf, and Hesham FA Hamed. "A review on brain tumor diagnosis from MRI images: Practical implications, key achievements, and lessons learned." In: *Magnetic resonance imaging* 61 (2019), pp. 300–318.

[2] A. Alansary, O. Oktay, Y. Li, et al. "Evaluating Reinforcement Learning Agents for Anatomical Landmark Detection." In: *Medical Image Analysis* (2019).

[3] Saad Albawi, Tareq Abed Mohammed, and Saad Al-Zawi. "Understanding of a convolutional neural network." In: *2017 international conference on engineering and technology (ICET)*. Ieee. 2017, pp. 1–6.

[4] Syed Muhammad Anwar, Muhammad Majid, Adnan Qayyum, Muhammad Awais, Majdi Alnowami, and Muhammad Khurram Khan. "Medical image analysis using convolutional neural networks: a review." In: *Journal of medical systems* 42 (2018), pp. 1–13.

[5] John Ashburner and Karl J Friston. "Spatial transformation of images." In: *Human brain function* (1997), pp. 43–58.

[6] John Ashburner and Karl J Friston. "Rigid body registration." In: *Statistical parametric mapping: The analysis of functional brain images* (2007), pp. 49–62.

[7] Brian B Avants, Nicholas J Tustison, Gang Song, Philip A Cook, Arno Klein, and James C Gee. "A reproducible evaluation of ANTs similarity metric performance in brain image registration." In: *Neuroimage* 54.3 (2011), pp. 2033–2044.

[8] Bhakti Baheti, Diana Waldmannstetter, Satrajit Chakrabarty, Hamed Akbari, Michel Bilello, Benedikt Wiestler, Julian Schwarting, Evan Calabrese, Jeffrey Rudie, Syed Abidi, et al. "The brain tumor sequence registration challenge: establishing correspondence between pre-operative and follow-up MRI scans of diffuse glioma patients." In: *arXiv preprint arXiv:2112.06979* (2021).

[9] Guha Balakrishnan, Amy Zhao, Mert R Sabuncu, John Guttag, and Adrian V Dalca. "VoxelMorph: a learning framework for deformable medical image registration." In: *IEEE transactions on medical imaging* 38.8 (2019), pp. 1788–1800.

[10] Richard E Bellman and Stuart E Dreyfus. *Applied Dynamic Programming*. 1962.

[11] Subrato Bharati, M Mondal, Prajoy Podder, and VB Prasath. "Deep Learning for Medical Image Registration: A Comprehensive Review." In: *International Journal of Computer Information Systems & Industrial Management Applications* 14 (2022).

[12] Jiří Borovec, Jan Kybic, Ignacio Arganda-Carreras, Dmitry V Sorokin, Gloria Bueno, Alexander V Khvostikov, Spyridon Bakas, I Eric, Chao Chang, Stefan Heldmann, et al. "ANHIR: automatic non-rigid histological image registration challenge." In: *IEEE transactions on medical imaging* 39.10 (2020), pp. 3042–3052.

[13] Richard J Boucherie and Nico M Van Dijk. *Markov decision processes in practice*. Vol. 248. Springer, 2017.

[14] Hamid Reza Boveiri, Raouf Khayami, Reza Javidan, and Alireza Mehdizadeh. "Medical image registration using deep neural networks: a comprehensive review." In: *Computers & Electrical Engineering* 87 (2020), p. 106767.

[15] James M Calvin, Craig J Gotsman, and Cuicui Zheng. "Global optimization for image registration." In: *AIP Conf Proc*. Vol. 2070. 2019, p. 20008.

[16] Xiaohuan Cao, Jianhua Yang, Jun Zhang, Dong Nie, Minjeong Kim, Qian Wang, and Dinggang Shen. "Deformable image registration based on similarity-steered CNN regression." In: *Medical Image Computing and Computer Assisted Intervention-MICCAI 2017: 20th International Conference, Quebec City, QC, Canada, September 11-13, 2017, Proceedings, Part I 20*. Springer. 2017, pp. 300–308.

[17] Xiaohuan Cao, Jianhuan Yang, Li Wang, Zhong Xue, Qian Wang, and Dinggang Shen. "Deep learning based intermodality image registration supervised by intra-modality similarity." In: *Machine Learning in Medical Imaging: 9th International Workshop, MLMI 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 16, 2018, Proceedings 9*. Springer. 2018, pp. 55–63.

[18] Maria Geraldina Castro, Rachel Cowen, Ian K Williamson, Alexandria David, Maximiliano J Jimenez-Dalmaroni, Xiuqin Yuan, Alireza Bigliari, Julie C Williams, Jianying Hu, and Pedro R Lowenstein. "Current and future strategies for the treatment of malignant brain tumors." In: *Pharmacology & therapeutics* 98.1 (2003), pp. 71–108.

[19] Sreenivasa R Chandana, Sujana Movva, Madan Arora, and Trevor Singh. "Primary brain tumors in adults." In: *American family physician* 77.10 (2008), pp. 1423–1430.

[20] Junyu Chen, Eric C Frey, Yufan He, William P Segars, Ye Li, and Yong Du. "Transmorph: Transformer for unsupervised medical image registration." In: *Medical image analysis* 82 (2022), p. 102615.

[21] Kanglin Chen, Alexander Derksen, Stefan Heldmann, Marc Hallmann, and Benjamin Berkels. "Deformable image registration with automatic non-correspondence detection." In: *International Conference on Scale Space and Variational Methods in Computer Vision*. Springer. 2015, pp. 360–371.

[22] Nicha Chitphakdithai, Veronica L Chiang, and James S Duncan. "Tracking metastatic brain tumors in longitudinal scans via joint image registration and labeling." In: *Spatio-temporal Image Analysis for Longitudinal and Time-Series Image Data: Second International Workshop, STIA 2012, Held in Conjunction with MICCAI 2012, Nice, France, October 1, 2012. Proceedings 2*. Springer. 2012, pp. 124–136.

[23] Nicha Chitphakdithai and James S Duncan. "Non-rigid registration with missing correspondences in preoperative and postresection brain images." In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 2010, pp. 367–374.

[24] Delia Ciardo, Marta Peroni, Marco Riboldi, Daniela Alterio, Guido Baroni, and Roberto Orecchia. "The role of regularization in deformable image registration for head and neck adaptive radiotherapy." In: *Technology in cancer research & treatment* 12.4 (2013), pp. 323–331.

[25] Steffen Czolbe, Paraskevas Pegios, Oswin Krause, and Aasa Feragen. "Semantic similarity metrics for image registration." In: *Medical Image Analysis* 87 (2023), p. 102830.

[26] Adrian V Dalca, Guha Balakrishnan, John Guttag, and Mert R Sabuncu. "Unsupervised learning of probabilistic diffeomorphic registration for images and surfaces." In: *Medical image analysis* 57 (2019), pp. 226–236.

[27] Mathieu De Craene, Gemma Piella, Oscar Camara, Nicolas Duchateau, Etelvino Silva, Adelina Doltra, Jan D'hooge, Josep Brugada, Marta Sitges, and Alejandro F Frangi. "Temporal diffeomorphic free-form deformation: Application to motion and strain estimation from 3D echocardiography." In: *Medical image analysis* 16.2 (2012), pp. 427–450.

[28] Dante De Nigris, D Louis Collins, and Tal Arbel. "Multi-modal image registration based on gradient orientations of minimal uncertainty." In: *IEEE transactions on medical imaging* 31.12 (2012), pp. 2343–2354.

[29] Bob D De Vos, Floris F Berendsen, Max A Viergever, Hessam Sokooti, Marius Staring, and Ivana Išgum. "A deep learning framework for unsupervised affine and deformable image registration." In: *Medical image analysis* 52 (2019), pp. 128–143.

[30] Lee R Dice. "Measures of the amount of ecologic association between species." In: *Ecology* 26.3 (1945), pp. 297–302.

[31] Bin Ding, Huimin Qian, and Jun Zhou. "Activation functions and their characteristics in deep neural networks." In: *2018 Chinese control and decision conference (CCDC)*. IEEE. 2018, pp. 1836–1841.

[32] Eustace M Dogo, OJ Afolabi, NI Nwulu, Bhekisipho Twala, and CO Aigbavboa. "A comparative analysis of gradient descent-based optimization algorithms on convolutional neural networks." In: *2018 international conference on computational techniques, electronics and mechanical systems (CTEMS)*. IEEE. 2018, pp. 92–99.

[33] AD Dongare, RR Kharde, Amit D Kachare, et al. "Introduction to artificial neural network." In: *International Journal of Engineering and Innovative Technology (IJEIT)* 2.1 (2012), pp. 189–194.

[34] Fotis Drakopoulos and Nikos P Chrisochoides. "Accurate and fast deformable medical image registration for brain tumor resection using image-guided neurosurgery." In: *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization* 4.2 (2016), pp. 112–126.

[35] G Eggers, J Mühling, and R Marmulla. "Image-to-patient registration techniques in head surgery." In: *International journal of oral and maxillofacial surgery* 35.12 (2006), pp. 1081–1095.

[36] Théo Estienne, Maria Vakalopoulou, Stergios Christodoulidis, Enzo Battistela, Marvin Lerousseau, Alexandre Carre, Guillaume Klausner, Roger Sun, Charlotte Robert, Stavroula Mougiakakou, et al. "U-ReSNet: Ultimate coupling of registration and segmentation with deep nets." In: *Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part III 22*. Springer. 2019, pp. 310–319.

[37] Tingting Feng, Yisong Lv, and Binjie Qin. "Region matching for pre-operative and post-operative brain images." In: *2015 IEEE 10th Conference on Industrial Electronics and Applications (ICIEA)*. IEEE. 2015, pp. 951–954.

[38] Vincent François-Lavet, Peter Henderson, Riashat Islam, Marc G Bellemare, Joelle Pineau, et al. "An introduction to deep reinforcement learning." In: *Foundations and Trends® in Machine Learning* 11.3-4 (2018), pp. 219–354.

[39] Qi Fu, Xin Xie, Yingjie Xu, Jing Zuo, Xi Yang, Wenlong Xia, Jusheng An, Manni Huang, Hui Yan, and Jianrong Dai. "Applying Multi-Metric Deformable Image Registration for Dose Accumulation in Combined Cervical Cancer Radiotherapy." In: *Journal of Personalized Medicine* 13.2 (2023), p. 323.

[40] Yabo Fu, Yang Lei, Tonghe Wang, Walter J Curran, Tian Liu, and Xiaofeng Yang. "Deep learning in medical image registration: a review." In: *Physics in Medicine & Biology* 65.20 (2020), 20TR01.

[41] Yabo Fu, Yang Lei, Jun Zhou, Tonghe Wang, S Yu David, Jonathan J Beitler, Walter J Curran, Tian Liu, and Xiaofeng Yang. "Synthetic CT-aided MRI-CT image registration for head and neck radiotherapy." In: *Medical Imaging 2020: Biomedical Applications in Molecular, Structural, and Functional Imaging*. Vol. 11317. SPIE. 2020, pp. 572–578.

[42] F.C. Ghesu, B. Georgescu, S. Grbic, A.K. Maier, J. Hornegger, and D. Comaniciu. "Robust multi-scale anatomical landmark detection in incomplete 3D-CT data." In: *MICCAI*. Springer. 2017, pp. 194–202.

[43] F.C. Ghesu, B. Georgescu, T. Mansi, D. Neumann, J. Hornegger, and D. Comaniciu. "An artificial agent for anatomical landmark detection in medical images." In: *MICCAI*. Springer. 2016, pp. 229–237.

[44] F.C. Ghesu, B. Georgescu, Y. Zheng, S. Grbic, A. Maier, J. Hornegger, and D. Comaniciu. "Multi-Scale Deep Reinforcement Learning for Real-Time 3D-Landmark Detection in CT Scans." In: *IEEE TPAMI* (2017).

[45] Fabian Gigengack, Lars Ruthotto, Martin Burger, Carsten H Wolters, Xiaoyi Jiang, and Klaus P Schafers. "Motion correction in dual gated cardiac PET using mass-preserving image registration." In: *IEEE transactions on medical imaging* 31.3 (2011), pp. 698–712.

[46] Xavier Glorot, Antoine Bordes, and Yoshua Bengio. "Deep sparse rectifier neural networks." In: *Proceedings of the fourteenth international conference on artificial intelligence and statistics*. JMLR Workshop and Conference Proceedings. 2011, pp. 315–323.

[47] Vladlena Gorbunova, Pechin Lo, Haseem Ashraf, Asger Dirksen, Mads Nielsen, and Marleen de Bruijne. "Weight preserving image registration for monitoring disease progression in lung CT." In: *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2008: 11th International Conference, New York, NY, USA, September 6-10, 2008, Proceedings, Part II 11*. Springer. 2008, pp. 863–870.

[48] Arthur Ardeshir Goshtasby. *2-D and 3-D image registration: for medical, remote sensing, and industrial applications*. John Wiley & Sons, 2005.

[49] Joseph V Hajnal and Derek LG Hill. *Medical image registration*. CRC press, 2001.

[50] Xu Han, Spyridon Bakas, Roland Kwitt, Stephen Aylward, Hamed Akbari, Michel Bilello, Christos Davatzikos, and Marc Niethammer. "Patient-specific registration of pre-operative and post-recurrence brain tumor MRI scans." In: *International MICCAI Brainlesion Workshop*. Springer. 2018, pp. 105–114.

[51] Cosmo Harrigan. "Deep reinforcement learning with regularized convolutional neural fitted Q iteration." In: *differences* 14 (2016), p. 1.

[52] Grant Haskins, Uwe Kruger, and Pingkun Yan. "Deep learning in medical image registration: a survey." In: *Machine Vision and Applications* 31 (2020), pp. 1–18.

[53] Robert Hecht-Nielsen. "Theory of the backpropagation neural network." In: *Neural networks for perception*. Elsevier, 1992, pp. 65–93.

[54] Pierre Hellier, Christian Barillot, Isabelle Corouge, Bernard Gibaud, Georges Le Goualher, D Louis Collins, Allan Evans, Grégoire Malandain, Nicholas Ayache, Gary E Christensen, et al. "Retrospective evaluation of intersubject brain registration." In: *IEEE transactions on medical imaging* 22.9 (2003), pp. 1120–1130.

[55] Alessa Hering, Lasse Hansen, Tony CW Mok, Albert CS Chung, Hanna Siebert, Stephanie Häger, Annkristin Lange, Sven Kuckertz, Stefan Heldmann, Wei Shao, et al. "Learn2Reg: comprehensive multi-task medical image registration challenge, dataset and evaluation in the era of deep learning." In: *IEEE Transactions on Medical Imaging* (2022).

[56] Berthold KP Horn and Brian G Schunck. "Determining optical flow." In: *Artificial intelligence* 17.1-3 (1981), pp. 185–203.

[57] Jing Hu, Ziwei Luo, Xin Wang, Shanhui Sun, Youbing Yin, Kunlin Cao, Qi Song, Siwei Lyu, and Xi Wu. "End-to-end multimodal image registration via reinforcement learning." In: *Medical Image Analysis* 68 (2021), p. 101878.

[58] Shunbo Hu, Lintao Zhang, Guoqiang Li, Mingtao Liu, Deqian Fu, and Wenyin Zhang. "Infant brain deformable registration using global and local label-driven deep regression learning." In: *Machine Learning in Medical Imaging: 10th International Workshop, MLMI 2019, Held in Conjunction with MICCAI 2019, Shenzhen, China, October 13, 2019, Proceedings 10*. Springer. 2019, pp. 106–114.

[59] Yipeng Hu, Eli Gibson, Nooshin Ghavami, Ester Bonmati, Caroline M Moore, Mark Emberton, Tom Vercauteren, J Alison Noble, and Dean C Barratt. "Adversarial deformation regularization for training image registration neural networks." In: *Medical Image Computing and Computer Assisted Intervention–MICCAI 2018: 21st International Conference, Granada, Spain, September 16-20, 2018, Proceedings, Part I*. Springer. 2018, pp. 774–782.

[60] Yipeng Hu, Marc Modat, Eli Gibson, Wenqi Li, Nooshin Ghavami, Ester Bonmati, Guotai Wang, Steven Bandula, Caroline M Moore, Mark Emberton, et al. "Weakly-supervised convolutional neural networks for multimodal image registration." In: *Medical image analysis* 49 (2018), pp. 1–13.

[61] Daniel P Huttenlocher, Gregory A. Klanderman, and William J Rucklidge. "Comparing images using the Hausdorff distance." In: *IEEE Transactions on pattern analysis and machine intelligence* 15.9 (1993), pp. 850–863.

[62] Fabian Isensee, Philipp Kickingereder, Wolfgang Wick, Martin Bendszus, and Klaus H Maier-Hein. "Brain tumor segmentation and radiomics survival prediction: Contribution to the brats 2017 challenge." In: *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: Third International Workshop, BrainLes 2017, Held in Conjunction with MICCAI 2017, Quebec City, QC, Canada, September 14, 2017, Revised Selected Papers 3*. Springer. 2018, pp. 287–297.

[63] Paul Jaccard. "The distribution of the flora in the alpine zone. 1." In: *New phytologist* 11.2 (1912), pp. 37–50.

[64] Pranita Balaji Kanade and PP Gumaste. "Brain tumor detection using MRI images." In: *Brain* 3.2 (2015), pp. 146–150.

[65] Michael R Kaus, Kristy K Brock, Vladimir Pekar, Laura A Dawson, Alan M Nichol, and David A Jaffray. "Assessment of a model-based deformable image registration approach for radiation therapy planning." In: *International Journal of Radiation Oncology* Biology* Physics* 68.2 (2007), pp. 572–580.

[66] Marc L Kessler. "Image registration and data fusion in radiation therapy." In: *The British journal of radiology* 79.special_issue_1 (2006), S99–S108.

[67] Boah Kim, Jieun Kim, June-Goo Lee, Dong Hwan Kim, Seong Ho Park, and Jong Chul Ye. "Unsupervised deformable image registration using cycle-consistent cnn." In: *Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part VI 22*. Springer. 2019, pp. 166–174.

[68] Stefan Klein, Marius Staring, Keelin Murphy, Max A Viergever, and Josien PW Pluim. "Elastix: a toolbox for intensity-based medical image registration." In: *IEEE transactions on medical imaging* 29.1 (2009), pp. 196–205.

[69] Florian Kofler, Johannes Wahle, Ivan Ezhov, Sophia J Wagner, Rami Al-Maskari, Emilia Gryska, Mihail Todorov, Christina Bukas, Felix Meissen, Tingying Peng, et al. "Approaching peak ground truth." In: *2023 IEEE 20th International Symposium on Biomedical Imaging (ISBI)*. IEEE. 2023, pp. 1–6.

[70] Julian Krebs, Tommaso Mansi, Hervé Delingette, Li Zhang, Florin C Ghesu, Shun Miao, Andreas K Maier, Nicholas Ayache, Rui Liao, and Ali Kamen. "Robust non-rigid registration through agent-based action learning." In: *Medical Image Computing and Computer Assisted Intervention- MICCAI 2017: 20th International Conference, Quebec City, QC, Canada, September 11-13, 2017, Proceedings, Part I 20*. Springer. 2017, pp. 344–352.

[71] Dongyang Kuang. "Cycle-consistent training for reducing negative jacobian determinant in deep registration networks." In: *Simulation and Synthesis in Medical Imaging: 4th International Workshop, SASHIMI 2019, Held in Conjunction with MICCAI 2019, Shenzhen, China, October 13, 2019, Proceedings 4*. Springer. 2019, pp. 120–129.

[72] Dongjin Kwon, Marc Niethammer, Hamed Akbari, Michel Bilello, Christos Davatzikos, and Kilian M Pohl. "PORTR: Pre-operative and post-recurrence brain tumor registration." In: *IEEE transactions on medical imaging* 33.3 (2013), pp. 651–667.

[73] Dongjin Kwon, Ke Zeng, Michel Bilello, and Christos Davatzikos. "Estimating patient specific templates for pre-operative and follow-up brain tumor registration." In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 2015, pp. 222–229.

[74] Zhang Li, Lucas J van Vliet, Jaap Stoker, and Frans M Vos. "A hybrid optimization strategy for registering images with large local deformations and intensity variations." In: *International journal of computer assisted radiology and surgery* 13 (2018), pp. 343–351.

[75] Rui Liao, Shun Miao, Pierre de Tournemire, Sasa Grbic, Ali Kamen, Tommaso Mansi, and Dorin Comaniciu. "An artificial agent for robust image registration." In: *Proceedings of the AAAI conference on artificial intelligence*. Vol. 31. 1. 2017.

[76] Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen Awm Van Der Laak, Bram Van Ginneken, and

Clara I Sánchez. "A survey on deep learning in medical image analysis." In: *Medical image analysis* 42 (2017), pp. 60–88.

[77] Manhua Liu, Fan Li, Hao Yan, Kundong Wang, Yixin Ma, Li Shen, Mingqing Xu, Alzheimer's Disease Neuroimaging Initiative, et al. "A multi-model deep convolutional neural network for automatic hippocampus segmentation and classification in Alzheimer's disease." In: *Neuroimage* 208 (2020), p. 116459.

[78] Marco Lorenzi, Nicholas Ayache, Giovanni Frisoni, Xavier Pennec, et al. "4D registration of serial brain's MR images: a robust measure of changes applied to Alzheimer's disease." In: *Spatio Temporal Image Analysis Workshop (STIA), MICCAI*. Vol. 1. Citeseer. 2010.

[79] Kai Ma, Jiangping Wang, Vivek Singh, Birgi Tamersoy, Yao-Jen Chang, Andreas Wimmer, and Terrence Chen. "Multimodal image registration with deep context reinforcement learning." In: *Medical Image Computing and Computer Assisted Intervention-MICCAI 2017: 20th International Conference, Quebec City, QC, Canada, September 11-13, 2017, Proceedings, Part I 20*. Springer. 2017, pp. 240–248.

[80] JB Antoine Maintz and Max A Viergever. "A survey of medical image registration." In: *Medical image analysis* 2.1 (1998), pp. 1–36.

[81] Timo Makela, Patrick Clarysse, Outi Sipila, Nicoleta Pauna, Quoc Cuong Pham, Toivo Katila, and Isabelle E Magnin. "A review of cardiac image registration methods." In: *IEEE Transactions on medical imaging* 21.9 (2002), pp. 1011–1021.

[82] Primoz Markelj, Dejan Tomaževič, Bostjan Likar, and Franjo Pernuš. "A review of 3D/2D registration methods for image-guided interventions." In: *Medical image analysis* 16.3 (2012), pp. 642–661.

[83] Kasper Marstal, F Berendsen, Niels Dekker, Marius Staring, and Stefan Klein. "The continuous registration challenge: evaluation-as-a-service for medical image registration algorithms." In: *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*. IEEE. 2019, pp. 1399–1402.

[84] Swapnil Mehta, Aatman Shah, and Henry Jung. "Diagnosis and treatment options for sequelae following radiation treatment of brain tumors." In: *Clinical neurology and neurosurgery* 163 (2017), pp. 1–8.

[85] Shun Miao, Sebastien Piat, Peter Fischer, Ahmet Tuysuzoglu, Philip Mewes, Tommaso Mansi, and Rui Liao. "Dilated FCN for multi-agent 2D/3D medical image registration." In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 32. 1. 2018.

[86] EH Middlebrooks, CJ Frost, IS Tuna, IM Schmalfuss, M Rahman, and A Old Crow. "Reduction of motion artifacts and noise using independent component analysis in task-based functional MRI for preoperative planning in patients with brain tumor." In: *American Journal of Neuroradiology* 38.2 (2017), pp. 336–342.

[87] Michael I Miga, Tuhin K Sinha, David M Cash, Robert L Galloway, and Robert J Weil. "Cortical surface registration for image-guided neurosurgery using laser-range scanning." In: *IEEE Transactions on medical Imaging* 22.8 (2003), pp. 973–985.

[88] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fidjeland, Georg Ostrovski, et al. "Human-level control through deep reinforcement learning." In: *nature* 518.7540 (2015), pp. 529–533.

[89] Keelin Murphy, Bram van Ginneken, J Reinhardt, Sven Kabus, Kai Ding, Xiang Deng, and J Pluim. "Evaluation of methods for pulmonary image registration: The EMPIRE10 study." In: *Grand Challenges in Medical Image Analysis* 2010 (2010), pp. 11–22.

[90] Thomas O'Donnell, Shmuel Aharon, Sandra Simon Halliburton, Alok Gupta, Gareth Funka-Lea, and Richard D White. "Multi-modality model-based registration in the cardiac domain." In: *Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No. PR00662)*. Vol. 2. IEEE. 2000, pp. 790–791.

[91] Seungjong Oh and Siyong Kim. "Deformable image registration in radiation therapy." In: *Radiation oncology journal* 35.2 (2017), p. 101.

[92] Francisco PM Oliveira and Joao Manuel RS Tavares. "Medical image registration: a review." In: *Computer methods in biomechanics and biomedical engineering* 17.2 (2014), pp. 73–93.

[93] Jorge Onieva Onieva, Berta Marti-Fuster, María Pedrero de la Puente, and Raúl San José Estépar. "Diffeomorphic lung registration using deep cnns and reinforced learning." In: *International Workshop on Reconstruction and Analysis of Moving Body Organs*. Springer. 2018, pp. 284–294.

[94] Aske Plaat, Walter Kosters, and Mike Preuss. "High-accuracy model-based reinforcement learning, a survey." In: *Artificial Intelligence Review* (2023), pp. 1–33.

[95] Srinivasu Polinati and Ravindra Dhuli. "A review on multimodel medical image fusion." In: *2019 International Conference on Communication and Signal Processing (ICCSP)*. IEEE. 2019, pp. 0554–0558.

[96] Michał Posiewnik and Tomasz Piotrowski. "Utility of deformable image registration for adaptive prostate cancer treatment. Analysis and comparison of two commercially available algorithms." In: *Zeitschrift für Medizinische Physik* 32.3 (2022), pp. 369–377.

[97] CLARKE RH. "On a method of investigating the deep ganglia and tracts of the central nervous system (cerebellum)." In: *Bmj* 2 (1906), pp. 1799–1800.

[98] Qolamreza R Razlighi, Nasser Kehtarnavaz, and Siamak Yousefi. "Evaluating similarity measures for brain image registration." In: *Journal of visual communication and image representation* 24.7 (2013), pp. 977–987.

[99] Martin Riedmiller. "Neural fitted Q iteration–first experiences with a data efficient neural reinforcement learning method." In: *Machine Learning: ECML 2005: 16th European Conference on Machine Learning, Porto, Portugal, October 3-7, 2005. Proceedings 16*. Springer. 2005, pp. 317–328.

[100] S Riyahi-Alam, Marta Peroni, Guido Baroni, and Marco Riboldi. "Regularization in deformable registration of biomedical images based on divergence and curl operators." In: *Methods of information in medicine* 53.01 (2014), pp. 21–28.

[101] Marc-Michel Rohé, Manasi Datar, Tobias Heimann, Maxime Sermesant, and Xavier Pennec. "SVF-Net: learning deformable image registration using shape matching." In: *Medical Image Computing and Computer Assisted Intervention- MICCAI 2017: 20th International Conference, Quebec City, QC, Canada, September 11-13, 2017, Proceedings, Part I 20*. Springer. 2017, pp. 266–274.

[102] Julian G Rosenman, Elizabeth P Miller, and Tim J Cullip. "Image registration: an essential part of radiation therapy treatment planning." In: *International Journal of Radiation Oncology* Biology* Physics* 40.1 (1998), pp. 197–205.

[103] Daniel Rueckert, Paul Aljabar, Rolf A Heckemann, Joseph V Hajnal, and Alexander Hammers. "Diffeomorphic registration using B-splines." In: *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2006: 9th International Conference, Copenhagen, Denmark, October 1-6, 2006. Proceedings, Part II 9*. Springer. 2006, pp. 702–709.

[104] David E Rumelhart, Paul Smolensky, James L McClelland, and G Hinton. "Sequential thought processes in PDP models." In: *Parallel distributed processing: explorations in the microstructures of cognition* 2 (1986), pp. 3–57.

[105] Frank Sauer. "Image registration: enabling technology for image guided surgery and therapy." In: *2005 IEEE engineering in medicine and biology 27th annual conference*. IEEE. 2006, pp. 7242–7245.

[106] Hugo Scurto, Bavo Van Kerrebroeck, Baptiste Caramiaux, and Frédéric Bevilacqua. "Designing deep reinforcement learning for human parameter exploration." In: *ACM Transactions on Computer-Human Interaction (TOCHI)* 28.1 (2021), pp. 1–35.

[107] Wei Shao, Linda Banh, Christian A Kunder, Richard E Fan, Simon JC Soerensen, Jeffrey B Wang, Nikola C Teslovich, Nikhil Madhuripan, Anugayathri Jawahar, Pejman Ghanouni, et al. "ProsRegNet: A deep learning framework for registration of MRI and histopathology images of the prostate." In: *Medical image analysis* 68 (2021), p. 101919.

[108] Sagar Sharma, Simone Sharma, and Anidhya Athaiya. "Activation functions in neural networks." In: *Towards Data Sci* 6.12 (2017), pp. 310–316.

[109] Dinggang Shen, Hari Sundar, Zhong Xue, Yong Fan, and Harold Litt. "Consistent estimation of cardiac motions by 4D image registration." In: *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2005: 8th International Conference, Palm Springs, CA, USA, October 26-29, 2005, Proceedings, Part II 8*. Springer. 2005, pp. 902–910.

[110] Sandro Skansi and Sandro Skansi. "Convolutional neural networks." In: *Introduction to Deep Learning: from logical calculus to artificial intelligence* (2018), pp. 121–133.

[111] Hessam Sokooti, Bob De Vos, Floris Berendsen, Boudewijn PF Lelieveldt, Ivana Išgum, and Marius Staring. "Nonrigid image registration using multi-scale 3D convolutional neural networks." In: *Medical Image Computing and Computer Assisted Intervention- MICCAI 2017: 20th International Conference, Quebec City, QC, Canada, September 11-13, 2017, Proceedings, Part I 20*. Springer. 2017, pp. 232–239.

[112] Guoli Song, Jianda Han, Yiwen Zhao, Zheng Wang, and Huibin Du. "A review on medical image registration as an optimization problem." In: *Current Medical Imaging* 13.3 (2017), pp. 274–283.

[113] Aristeidis Sotiras, Christos Davatzikos, and Nikos Paragios. "Deformable medical image registration: A survey." In: *IEEE transactions on medical imaging* 32.7 (2013), pp. 1153–1190.

[114] Edward A Spiegel, Henry T Wycis, M Marks, and Arnold J Lee. "Stereotaxic apparatus for operations on the human brain." In: *Science* 106.2754 (1947), pp. 349–350.

[115] Shanhui Sun, Jing Hu, Mingqing Yao, Jinrong Hu, Xiaodong Yang, Qi Song, and Xi Wu. "Robust multimodal image registration using deep recurrent reinforcement learning." In: *Asian conference on computer vision*. Springer. 2018, pp. 511–526.

[116] R.S. Sutton and A.G. Barto. *Reinforcement Learning: An Introduction*. MIT Press, 1998.

[117] Richard S Sutton and Andrew G Barto. *Reinforcement Learning: An Introduction*. MIT Press, 2018.

[118] Kenji Suzuki. *Artificial neural networks: methodological advances and biomedical applications*. BoD–Books on Demand, 2011.

[119] Camillo Jose Taylor and Arvind Bhusnurmath. "Solving image registration problems using interior point methods." In: *Computer Vision–ECCV 2008: 10th European Conference on Computer Vision, Marseille, France, October 12-18, 2008, Proceedings, Part IV 10*. Springer. 2008, pp. 638–651.

[120] Jonas Teuwen and Nikita Moriakov. "Convolutional neural networks." In: *Handbook of medical image computing and computer assisted intervention*. Elsevier, 2020, pp. 481–501.

[121] J-P Thirion. "Image matching as a diffusion process: an analogy with Maxwell's demons." In: *Medical image analysis* 2.3 (1998), pp. 243–260.

[122] Arthur W Toga and Paul M Thompson. "The role of image registration in brain mapping." In: *Image and vision computing* 19.1-2 (2001), pp. 3–24.

[123] Nicholas J Tustison, Brian B Avants, and James C Gee. "Learning image-based spatial transformations via convolutional neural networks: A review." In: *Magnetic resonance imaging* 64 (2019), pp. 142–153.

[124] JN Ulysses and A Conci. "Measuring similarity in medical registration." In: *IWSSIP 17th Int. Conf. on Systems, Signals and Image Processing*. 2010.

[125] Martijn Van Otterlo and Marco Wiering. "Reinforcement learning and markov decision processes." In: *Reinforcement learning: State-of-the-art*. Springer, 2012, pp. 3–42.

[126] Tom Vercauteren, Xavier Pennec, Aymeric Perchant, and Nicholas Ayache. "Non-parametric diffeomorphic image registration with the demons algorithm." In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 2007, pp. 319–326.

[127] Tom Vercauteren, Xavier Pennec, Aymeric Perchant, and Nicholas Ayache. "Diffeomorphic demons: Efficient non-parametric image registration." In: *NeuroImage* 45.1 (2009), S61–S72.

[128] Paul Viola and William M Wells III. "Alignment by maximization of mutual information." In: *International journal of computer vision* 24.2 (1997), pp. 137–154.

[129] A. Vlontzos, A. Alansary, K. Kamnitsas, D. Rueckert, and B. Kainz. "Multiple Landmark Detection using Multi-Agent Reinforcement Learning." In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 2019, pp. 262–270.

[130] Diana Waldmannstetter, Fernando Navarro, Benedikt Wiestler, Jan S Kirschke, Anjany Sekuboyina, Ester Molero, and Bjoern H Menze. "Reinforced redetection of landmark in pre-and post-operative brain scan using anatomical guidance for image alignment." In: *International Workshop on Biomedical Image Registration*. Springer. 2020, pp. 81–90.

[131] Diana Waldmannstetter, Benedikt Wiestler, Julian Schwarting, Ivan Ezhov, Marie Metz, Spyridon Bakas, Bhakti Baheti, Satrajit Chakrabarty, Jan S Kirschke, Rolf A Heckemann, et al. "Framing image registration as a landmark detection problem for better representation of clinical relevance." In: *arXiv preprint arXiv:2308.01318* (2023).

[132] Diana Waldmannstetter, Benedikt Wiestler, Julian Schwarting, Ivan Ezhov, Marie Metz, Spyridon Bakas, Bhakti Baheti, Satrajit Chakrabarty, Daniel Rueckert, Jan S Kirschke, et al. "Primitive Simultaneous Optimization of Similarity Metrics for Image Registration." In: *International MICCAI Brainlesion Workshop*. Springer. 2023, pp. 57–68.

[133] Jian Wang, Hengde Zhu, Shui-Hua Wang, and Yu-Dong Zhang. "A review of deep learning on medical image analysis." In: *Mobile Networks and Applications* 26 (2021), pp. 351–380.

[134] Ziyu Wang, Tom Schaul, Matteo Hessel, Hado Hasselt, Marc Lanctot, and Nando Freitas. "Dueling network architectures for deep reinforcement learning." In: *International conference on machine learning*. PMLR. 2016, pp. 1995–2003.

[135] Christopher JCH Watkins and Peter Dayan. "Q-learning." In: *Machine learning* 8 (1992), pp. 279–292.

[136] Christopher John Cornish Hellaby Watkins. "Learning from delayed rewards." In: (1989).

[137] Philippe Weitz, Masi Valkonen, Leslie Solorzano, Johan Hartman, Pekka Ruusuvuori, and Mattias Rantalainen. "ACROBAT-Automatic Registration of Breast Cancer Tissue." In: *10th Internatioal Workshop on Biomedical Image Registration*. 2022.

[138] Austin Williams, Noah Walton, Austin Maryanski, Sandra Bogetic, Wes Hines, and Vladimir Sobes. "Stochastic gradient descent for optimization for nuclear systems." In: *Scientific Reports* 13.1 (2023), p. 8474.

[139] Adam Wittek, Karol Miller, Ron Kikinis, and Simon K Warfield. "Patient-specific model of brain deformation: Application to medical image registration." In: *Journal of biomechanics* 40.4 (2007), pp. 919–929.

[140] Haonan Xiao, Xinzhi Teng, Chenyang Liu, Tian Li, Ge Ren, Ruijie Yang, Dinggang Shen, and Jing Cai. "A review of deep learning-based three-dimensional medical image registration methods." In: *Quantitative Imaging in Medicine and Surgery* 11.12 (2021), p. 4895.

[141] Bing Xu, Naiyan Wang, Tianqi Chen, and Mu Li. "Empirical evaluation of rectified activations in convolutional network." In: *arXiv preprint arXiv:1505.00853* (2015).

[142] Xujing Yao, Xinyue Wang, Shui-Hua Wang, and Yu-Dong Zhang. "A comprehensive survey on convolutional neural network in medical image analysis." In: *Multimedia Tools and Applications* 81.29 (2022), pp. 41361–41405.

[143] Michael C Yip, David G Lowe, Septimiu E Salcudean, Robert N Rohling, and Christopher Y Nguan. "Tissue tracking and registration for image-guided surgery." In: *IEEE transactions on medical imaging* 31.11 (2012), pp. 2169–2182.

[144] Hang Yu, Laurence T Yang, Qingchen Zhang, David Armstrong, and M Jamal Deen. "Convolutional neural networks for medical image analysis: state-of-the-art, comparisons, improvement and perspectives." In: *Neurocomputing* 444 (2021), pp. 92–110.

[145] Ramy A Zeineldin, Mohamed E Karar, Oliver Burgert, and Franziska Mathis-Ullrich. "Multimodal CNN networks for brain tumor segmentation in MRI: a BraTS 2022 challenge solution." In: *arXiv preprint arXiv:2212.09310* (2022).

[146] Chuanlei Zhang, Minda Yao, Wei Chen, Shanwen Zhang, Dufeng Chen, and Yuliang Wu. "Gradient descent optimization in deep learning model training based on multistage and method combination strategy." In: *Security and Communication Networks* 2021 (2021), pp. 1–15.

[147] Zhenyu Zhu, Yiqin Cao, Chenchen Qin, Yi Rao, Di Lin, Qi Dou, Dong Ni, and Yi Wang. "Joint affine and deformable three-dimensional networks for brain MRI registration." In: *Medical Physics* 48.3 (2021), pp. 1182–1196.

[148] Jing Zou, Bingchen Gao, Youyi Song, and Jing Qin. "A review of deep learning-based deformable medical image registration." In: *Frontiers in Oncology* 12 (2022), p. 1047215.