



Data Article

GBM-Reservoir: Brain tumor (Glioblastoma Multiforme) MRI dataset collection with ground truth segmentation masks

Naida Solak^{a,b,c}, André Ferreira^{a,e,f,g}, Gijs Luijten^{a,b,c,d},
Behrus Puladi^{f,g}, Victor Alves^e, Jan Egger^{a,b,c,d,*}

^a Institute for AI in Medicine (IKIM), University Hospital Essen (UKE), Ruhrgebiet, Essen, Germany

^b Graz University of Technology (TU Graz), Graz, Styria, Austria

^c Computer Algorithms for Medicine Laboratory (Café Lab), Graz, Styria, Austria

^d Center for Virtual and Extended Reality in Medicine, University Medicine Essen, Essen, Germany

^e Center Algoritmi / LASI, University of Minho, Braga, Portugal

^f Institute of Medical Informatics, University Hospital RWTH Aachen, Aachen, Germany

^g Department of Oral and Maxillofacial Surgery, University Hospital RWTH Aachen, Aachen, Germany



ARTICLE INFO

Article history:

Received 18 December 2024

Revised 2 January 2025

Accepted 6 January 2025

Available online 8 January 2025

Dataset link: [GBM-Reservoir: Dataset and Segmentations \(Original data\)](#)

Keywords:

Brain tumor segmentation

Data augmentation

Registration

BraTS

Deep learning

ABSTRACT

In this article, we present a brain tumor database collection comprising 23,049 samples, with each sample including four different types of MRI brain scans: FLAIR, T1, T1ce, and T2. Additionally, one or two segmentation masks (ground truth) are provided for each sample. The first mask is the raw output from the registration process and is provided for all samples, while the second mask, provided particularly for synthetic samples, is a post-processed version of the first, designed to simplify interpretation and optimize it for network training. These samples have been acquired via registration process of 438 samples available at the moment of registration from the original dataset provided by the BraTS 2022 Challenge. Registering each pair of existing brain scans results in two additional scans that retain a similar brain shape while featuring varying tumor locations. Consequently, by registering all possible pairs, a dataset originally consisting of n samples can be expanded to n^2 samples. The original dataset was collected from different institutions under standard clinical conditions, but with different equipment and

* Corresponding author.

E-mail address: jan.egger@uk-essen.de (J. Egger).

imaging protocols. As a result, the image quality is heterogeneous, reflecting the diversity of clinical practices across institutions. This dataset can be utilized for various tasks, such as developing fully automated segmentation algorithms for new, unseen brain tumor cases, particularly through deep learning-based approaches, since ground truth is provided for each sample.

© 2025 The Authors. Published by Elsevier Inc.

This is an open access article under the CC BY-NC license (<http://creativecommons.org/licenses/by-nc/4.0/>)

Specification Table

Subject	Computer Vision and Pattern Recognition
Specific subject area	Dataset Augmentation via 3D Image Registration
Type of data	3D Image (.nii.gz format), Pre-processed (Co-registration to a template, resampling, skull-stripping, registration).
Data collection	The real-life brain scans of Glioblastoma Multiforme patients are augmented via a process of registration where new samples are created from the original BraTS 2022 dataset.
Data source location	Figshare.
Data accessibility	Repository name: Figshare [1] Data Identification Number: 10.6084/m9.figshare.28001450 Original BraTS Dataset: Kaggle [2], Cancer Imaging Archive [3].
Related research articles	André Ferreira, Naida Solak, Jianning Li, Philipp Dammann, Jens Kleesiek, Victor Alves, Jan Egger. Title: Enhanced Data Augmentation Using Synthetic Data for Brain Tumour Segmentation [4]. DOI: 10.1007/978-3-031-76163-8_8 .

1. Value of the Data

- The brain tumor scans together with the corresponding segmentations can serve as an evaluation set for automatic brain tumor segmentation algorithms.
- Researchers can use the scans and corresponding segmentations in order to train deep learning algorithms.
- Researchers can use the collection as basis for data augmentation to further increase the collection without having to do the registration themselves, saving computer resources and time.
- The generated brain tumor scans and their corresponding segmentation masks serve as valuable resources for educational purposes, addressing the common challenges associated with acquiring medical data.

2. Background

The brain tumor dataset was created using image registration to create a more extensive and diverse training set for developing neural network models, addressing the scarcity of annotated medical data due to privacy constraints and time-intensive labeling [5,6]. This approach ensures that the dataset contains a broader range of imaging variations, improving the model's ability to generalize and learn from diverse tumor presentations. The theoretical background for this dataset is grounded in medical image processing and deep learning, specifically within the context of improving the performance of state-of-the-art neural networks in automated tumor detection and segmentation [7–10]. The dataset was used in the related research article [4] to enhance the performance of a neural network model in classifying and segmenting brain tumors from medical imaging due to increased size and variation of data, which ultimately contributed

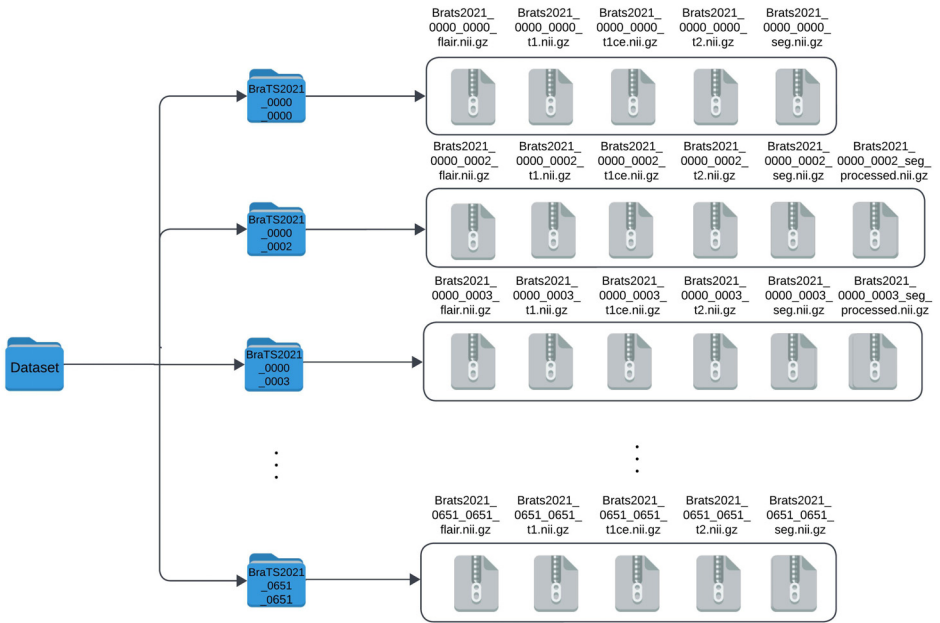


Fig. 1. Scheme representing a folder structure of the dataset collection, comprising 23049 folders with 5 or 6 files each (in NIfTI format *.nii.gz*) - native (T1), post-contrast T1-weighted (T1ce), T2-weighted (T2), and T2 Fluid Attenuated Inversion Recovery (FLAIR) volumes, along with one raw mask and optionally, one post-processed mask.

to improving the accuracy and robustness of the neural network. This data article complements the referred paper by providing the extended dataset, which is crucial for reproducing the results and further refining the model in future research on brain tumor detection.

3. Data Description

In Fig. 1, the structure of the dataset is illustrated. The dataset includes 23049 folders with 5 or 6 files per folder. These files represent different types of MRI scans: FLAIR, T1, T2, and T1ce, and 1 or 2 ground truth masks in addition. All files are stored in the compressed NIfTI format (*.nii.gz*).

An example of the file naming convention used in the dataset is "Brats2021_0000_0002_flair.nii.gz". This filename represents a combination of two MRI scans: 0000 and 0002, which were obtained through the registration process. The prefix "Brats2021" indicates that the data is part of the BraTS 2021 dataset (which has been used as a dataset for BraTS 2022 Challenge), and the order of numbers indicates that the first scan (0000) has been warped into the spatial domain of the second scan (0002). The sample numbers are non-sequential (e.g., 0000, 0002, etc.) because only training samples from the original BraTS 2022 dataset were used to generate this dataset. If a file name contains two identical numbers (e.g. "Brats2021_0000_0000_flair.nii.gz"), it corresponds to the original "BraTS2021_0000_flair.nii.gz" sample. For such samples, only a raw segmentation mask is provided since the labels already align with the original BraTS dataset and a mask is immediately suitable for training. The suffix "_flair" denotes that the MRI scan corresponds to the FLAIR modality. In addition to the FLAIR scan, there are corresponding files for other MRI modalities: T1, T1ce, T2 and in addition "seg" (raw output from a registration process) and "seg_processed" (post-processed mask) for segmentation masks. This systematic naming convention facilitates the identification

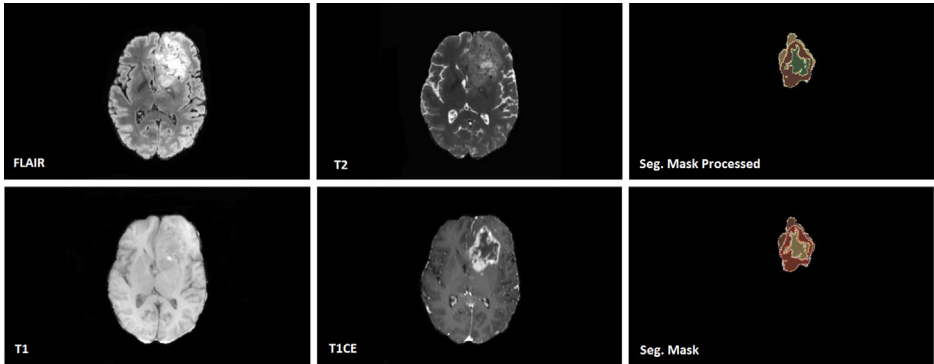


Fig. 2. An example of a dataset sample includes the following components: Fluid Attenuated Inversion Recovery - FLAIR, native - T1-weighted, contrast-enhanced - T1CE, and T2-weighted MRI scans, along with two segmentation masks (raw and post-processed). This specific sample corresponds to "BraTS2021_0000_0002" case.

and organization of MRI scans and their corresponding segmentation masks generated from the registration process.

Fig. 2 provides an example of a data sample that includes four MRI scans of different types. The dataset was generated using a data augmentation method known as registration, applied to the BraTS 2022¹ dataset.

3.1. The BraTS 2022 Dataset

The (training) dataset of the BraTS challenge [11–16] represents a collection of the mpMRI scans that have been acquired from different institutions under standard clinical conditions, but with different equipment and imaging protocols. As a result, the image quality is heterogeneous and it reflects diverse clinical practice across different institutions. The BraTS 2022 dataset is divided in training, validation and testing subsets, where the ground truth (approved by expert neuroradiologists [17]) is provided only for training set, and a testing dataset is kept hidden at all times. To generate the dataset presented in this paper, only the training BraTS 2022 dataset has been used because it includes ground truth masks (which are needed to train machine learning models) [18].

The mpMRI scans, acquired with different protocols and various scanners from multiple institutions, include:

1. Native T1 volumes;
2. Post-contrast (Gadolinium) T1 volumes;
3. T2-weighted (T2) volumes;
4. Fluid Attenuated Inversion Recovery (T2-FLAIR) volumes.

All the scans are pre-processed and the pre-processing pipeline is publicly available through the Cancer Imaging Phenomics Toolkit (CaPTk) and Federated Tumor Segmentation Tool (FeTS). This routine includes:

1. Conversion of the DICOM files to the NiftI file format [19];

¹ The Brain Tumor Segmentation Challenge (BraTS) has been jointly organized by the Radiological Society of North America (RSNA), the American Society of Neuroradiology (ASNR), and the Medical Image Computing and Computer Assisted Interventions (MICCAI) society. This challenge focuses on evaluating computational algorithms for segmenting histologically distinct brain tumor subregions. The size of the datasets provided by the organizers has increased each year, from just 50 samples in the inaugural challenge in 2012 to 438 samples at the time the dataset presented in this paper was created.

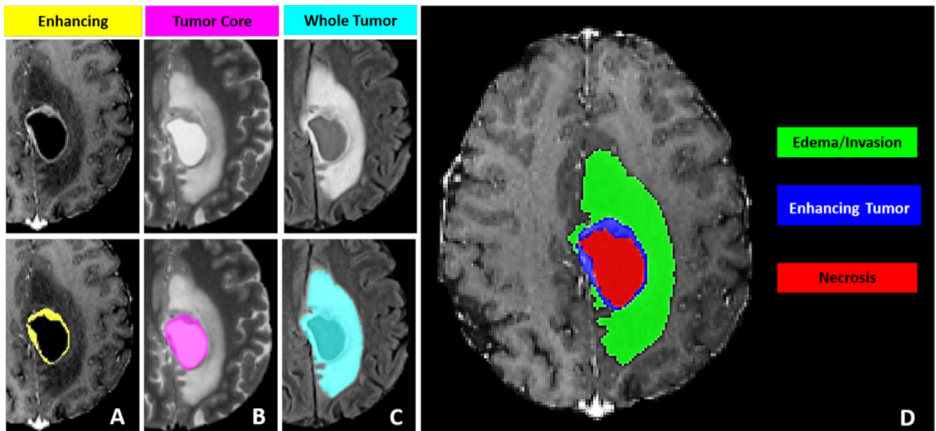


Fig. 3. Glioma sub-regions considered in the BraTS 2022 Challenge. The image panels A-C denote the regions considered for the performance evaluation: (A) Enhancing tumor (yellow) surrounding the cystic/necrotic core; (B) Tumor Core (magenta) - a union of necrosis and the enhancing tumor; (C) Whole Tumor (blue) encompasses necrosis/cystic core (red), enhancing tumor (blue), and edema (green). (D) The individual segmentations are combined into final tumor sub-region labels: the enhancing tumor (blue), the necrotic/cystic core (red), and edema/invasion (green).

2. Co-registration to the same anatomical template (SRI24) [20];
3. Resampling to a uniform isotropic resolution (1mm^3);
4. Skull-stripping.

All volumes are segmented using the STAPLE fusion of previous top-ranked algorithms [11] (DeepScan [21], DeepMedic [10], and nnUNet [22] all trained on the BraTS 2020 dataset) and later manually refined by volunteer neuroradiologists of varying experience. All the segmentation refinements were approved by experienced board-certified attending neuro-radiologists, with more than 15 years of experience working with gliomas [18]. The tumor sub-regions (as shown in Fig. 3 [18]) are the following:

1. **Enhancing tumor - EH:** areas with higher intensity values in T1Gd scans in comparison to T1 and healthy white matter in T1Gd;
2. **Tumor core - TC:** the bulk of the tumor that is supposed to be surgically removed. The TC entails the ET and the necrotic parts (NCR) of the tumor. The intensity is usually lower in T1Gd in that parts than in T1 scans;
3. **Complete tumor extent/whole tumor - WT:** the whole malignancy that entails the TC and the peritumoral edematous (ED) tissue. The intensity of the area is usually higher in T2-FLAIR volume.

However, other delineation criteria could be set as well, which would result in different tumor sub-regions [18].

4. Experimental Design, Materials and Methods

Inspired by the ongoing Brain Tumor Sequence Registration (BraTS-Reg) Challenge [23], also organized by MICCAI² and also the winning solution for the AutoImplant Challenge 2020 implemented by Ellis and Aizenberg [24], the training dataset was enhanced by generating new samples through registration. This approach aimed to introduce variations in tumor locations while

² For more information on BraTS-Reg Challenge, please visit: <https://www.med.upenn.edu/cbica/brats-reg-challenge/>

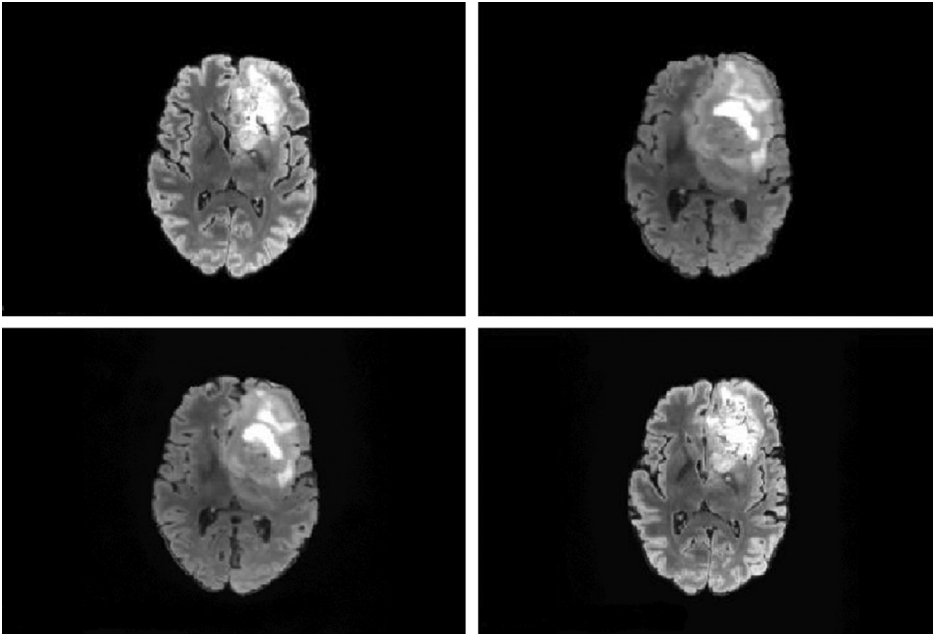


Fig. 4. Example of augmentations produced via registration. Images in the first row show the original samples from the BraTS21 dataset and images in the second row show the synthetically created samples with similar brain shape and structure to the original samples.

preserving consistent brain structures, as previously demonstrated in medical imaging studies for generating robust datasets [25,26]. Each image in the dataset can be registered with every other image, resulting in a warped version of the target image. The Advanced Normalization Tools (ANTs) package³, a widely used software for image normalization and registration in neuroimaging research [27,28], was utilized for the image registration process [29]. It is a software package used for normalizing data to a template and it provides different scripts (such as `antsRegistrationSyNQuick.shg`) which allow for the application of different transformations, including rigid, affine, non-linear, and their combinations, as outlined in previous studies [30,31]. Typically, one image is designated as the moving image, while the other serves as the fixed image. The moving image is warped into the fixed image's space by applying the computed transformation, and the inverse transformation is then used to warp the fixed image into the moving image's space. This methodology aligns with approaches used in other biomedical image registration studies, where the goal is to enhance dataset variability and improve model generalization [28,32,33].

4.1. Data augmentation via registration

In Fig. 4, we illustrate an example of data augmentation through registration on the BraTS 2022 dataset. The non-linear registration process is performed between the brain images of Subject 0000 (top left) and Subject 0002 (top right). This results in the generation of two additional scans: the bottom left image displays an artificially created brain MRI scan that maintains a similar shape to Subject 0000 (top left) while exhibiting content akin to that of Subject 0002

³ For more information on ANTs tool, please visit a GitHub repository: <https://github.com/ANTsX/ANTs>

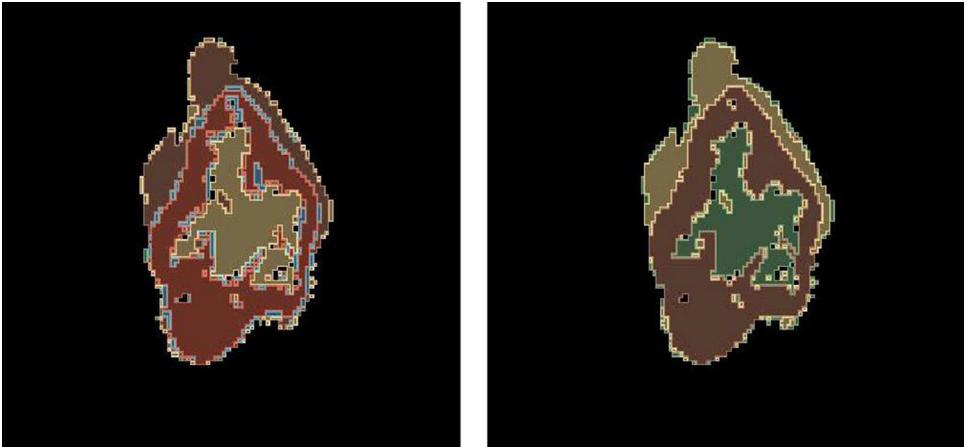


Fig. 5. Example of a segmentation mask with 6 labels produced via registration (left) and a post-processed mask with 3 labels only, for data consistency and easier understanding⁴.

(top right). Conversely, the bottom right image features a newly generated brain MRI scan with a shape resembling that of Subject 0002 and content similar to Subject 0000. This registration process can be systematically applied to each image pair within the BraTS 2022 dataset [34].

Additionally, as shown in Fig. 5, the same transformations - and their inverses - must be applied to the corresponding segmentation masks. As a result, we obtain a mask with six regions, as some pixels do not correspond to any of the three original labels (e.g., background pixels). Exceptions are samples corresponding to the original BraTS samples (e.g., "BraTS2021_0000_0000" corresponding to the original "BraTS2021_0000" sample), where the registration masks retain their original three labels and are immediately suitable for network training. However, for newly generated samples (e.g., "BraTS2021_0000_0002"), raw masks require post-processing before they can be used for training (as presented in [4]). To address this, we provide a second mask for each synthetic sample, where the raw mask has been refined by merging and discarding certain labels from the six regions. This approach aligns with the logic of the original BraTS dataset, which contains only three labels. However, for those interested in working directly with the raw data, the original registration masks are also available.

The process of registration was applied to a training data subset consisting of 438 samples, each comprising a segmentation mask and four imaging modalities: FLAIR, T1, T1CE, and T2. Due to the extensive time required for registration and the significant storage resources necessary for data management and model training, the final number of registered samples totaled 23,049. Completing the registration of these cases took over two weeks.

Limitations

While the GBM-Reservoir dataset provides a valuable resource for developing and testing deep learning models for brain tumor segmentation, it is important to acknowledge its limitations:

- **Synthetic Nature of Augmented Data:** The dataset relies on non-linear registration to generate synthetic samples. Although this process increases data diversity, it does not perfectly

⁴ Label colors are randomly assigned by the 3D Slicer visualization tool and do not indicate corresponding labels in these two masks.

replicate the variations observed in real-world patient data. The synthetic samples may introduce artifacts or biases.

- **Computational Overhead:** The registration process used to create this dataset is computationally intensive, requiring significant processing power and time. This limitation may affect reproducibility for users with limited access to high-performance computing resources.
- **Heterogeneity of Source Data:** While the BraTS dataset includes multi-institutional scans, the quality and variability of the original data depend on the acquisition equipment and protocols. These variations may influence the quality of the generated synthetic data.
- **Post-Processing Challenges:** The raw segmentation masks generated during the registration process contain six labels, which usually require post-processing to simplify to the three-label format (consistent with the original BraTS dataset). While we provide both raw and post-processed masks, users must be cautious when working with the raw data to avoid inconsistencies.
- **Exclusion of External Modalities:** The dataset is limited to the imaging modalities (FLAIR, T1, T1CE, T2) provided in the BraTS challenge. It does not include other imaging types or modalities that may enhance tumor characterization, such as diffusion-weighted imaging or perfusion MRI.

By acknowledging these limitations, we hope to provide researchers with a clearer understanding of the dataset's scope and potential constraints, encouraging further validation and improvements in related research.

Ethics Statement

The datasets, described in the article are adapted from public collections, which are licensed under **CC BY-NC-SA 4.0** and End User License Agreement (EULA). According to the usage notes provided in their website, we make the adapted datasets public under the same licenses as the original collections. The authors have read and follow the ethical requirements for publication in Data in Brief and confirm that the current work does not involve human subjects, animal experiments, or any data collected from social media platforms.

Data Availability

[GBM-Reservoir: Dataset and Segmentations \(Original data\)](#) (figshare)

CRedit Author Statement

Naida Solak: Data curation, Writing – original draft; **André Ferreira:** Data curation, Writing – original draft; **Gijs Luijten:** Writing – original draft; **Behrus Puladi:** Supervision; **Victor Alves:** Supervision; **Jan Egger:** Writing – original draft, Supervision.

Acknowledgement

We acknowledge support by the Open Access Publication Fund of the University of Duisburg-Essen. André Ferreira thanks the Fundação para a Ciência e Tecnologia (FCT) Portugal for the grant 2022.11928.BD. We thank all the organizers of the BraTS challenge for allowing the publication of this dataset.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Supplementary Material

Supplementary material associated with this article can be found, in the online version, at [10.1016/j.dib.2025.111287](https://doi.org/10.1016/j.dib.2025.111287)

References

- [1] N. Solak, A. Ferreira, G. Luijten, B. Puladi, V. Alves, J. Egger, GBM-Reservoir: dataset and segmentations, Figshare (2024), doi:[10.6084/m9.figshare.28001450](https://doi.org/10.6084/m9.figshare.28001450).
- [2] D. Schettler, BraTS 2021 Task 1 Dataset, 2021, Kaggle. <https://www.kaggle.com/datasets/dschettler8845/brats-2021-task1/data>.
- [3] U. Baid, S. Ghodasara, S. Mohan, M. Bilello, E. Calabrese, E. Colak, et al., RSNA-ASNR-MICCAI-BraTS-2021 Dataset, Cancer Imaging Arch. (2023), doi:[10.7937/jc8x-9874](https://doi.org/10.7937/jc8x-9874).
- [4] A. Ferreira, N. Solak, J. Li, P. Dammann, J. Kleesiek, V. Alves, J. Egger, Enhanced data augmentation using synthetic data for brain tumour segmentation, Springer Nat. Switzerland (2024) 79–93, doi:[10.1007/978-3-031-76163-8_8](https://doi.org/10.1007/978-3-031-76163-8_8).
- [5] J. Egger, T. Kapur, A. Fedorov, S. Pieper, J.V. Miller, H. Veeraraghavan, B. Freisleben, A.J. Golby, C. Nimsky, R. Kikinis, GBM Volumetry using the 3D slicer medical image computing platform, Sci. Rep. 3 (1) (2013) 1364, doi:[10.1038/srep01364](https://doi.org/10.1038/srep01364).
- [6] M. Visser, D.M.J. Müller, R.J. van Duijn, M. Smits, N. Verburg, E.J. Hendriks, R.J.A. Nabuurs, J.C.J. Bot, R.S. Eijgelaar, M.G. Witte, M.B. van Herk, F. Barkhof, F. Barkhof, P.C. de Witt Hamer, J.C. de Munck, Inter-rater agreement in glioma segmentations on longitudinal MRI, NeuroImage Clin. 22 (2019), doi:[10.1016/j.nicl.2019.101727](https://doi.org/10.1016/j.nicl.2019.101727).
- [7] O. Ronneberger, P. Fischer, T. Brox, U-Net: convolutional networks for biomedical image segmentation, in: MICCAI 2015, 9351, Springer, Cham, 2015, pp. 234–241, doi:[10.1007/978-3-319-24574-4_28](https://doi.org/10.1007/978-3-319-24574-4_28).
- [8] J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 3431–3440, doi:[10.1109/CVPR.2015.7298965](https://doi.org/10.1109/CVPR.2015.7298965).
- [9] K. Kamnitsas, L. Chen, C. Ledig, D. Rueckert, B. Glocker, Multiscale 3D convolutional neural networks for lesion segmentation in brain MRI, Proc. MICCAI Ischemic Stroke Lesion Segmentation Challenge (2015).
- [10] K. Kamnitsas, C. Ledig, V. Newcombe, J. Simpson, A. Kane, D. Menon, D. Rueckert, B. Glocker, Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation, Med. Image Anal. 36 (2016), doi:[10.1016/j.media.2016.10.004](https://doi.org/10.1016/j.media.2016.10.004).
- [11] U. Baid, S. Ghodasara, M. Bilello, S. Mohan, E. Calabrese, E. Colak, K. Farahani, J. Kalpathy-Cramer, F. Kitamura, S. Pati, L. Prevedello, J. Rudie, C. Sako, R. Shinohara, T. Bergquist, R. Chai, J. Eddy, J. Elliott, W. Reade, S. Bakas, The RSNA-ASNR-MICCAI BraTS 2021 Benchmark on Brain Tumor Segmentation and Radiogenomic Classification (2021), [10.48550/arXiv.2107.02314](https://arxiv.org/abs/10.48550/arXiv.2107.02314)
- [12] B.H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby, Y. Burren, N. Porz, J. Slotboom, R. Wiest, et al., The multimodal brain tumor image segmentation benchmark (BRATS), IEEE Trans. Med. Imaging 34 (10) (2014) 1993–2024, doi:[10.1109/TMI.2014.2377694](https://doi.org/10.1109/TMI.2014.2377694).
- [13] S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J.S. Kirby, J.B. Freymann, K. Farahani, C. Davatzikos, Advancing the Cancer Genome Atlas Glioma MRI collections with expert segmentation labels and radiomic features, Sci. Data 4 (1) (2017) 1–13, doi:[10.1038/sdata.2017.117](https://doi.org/10.1038/sdata.2017.117).
- [14] S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J. Kirby, J. Freymann, K. Farahani, C. Davatzikos, Segmentation labels for the pre-operative scans of the TCGA-GBM collection, Cancer Imaging Arch. (2017), doi:[10.7937/K9/TCIA.2017.KLXWJJ1Q](https://doi.org/10.7937/K9/TCIA.2017.KLXWJJ1Q).
- [15] S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J. Kirby, J. Freymann, K. Farahani, C. Davatzikos, Segmentation labels and radiomic features for the pre-operative scans of the TCGA-LGG collection, Cancer Imaging Arch. (2017), doi:[10.7937/K9/TCIA.2017.GJQ7R0EF](https://doi.org/10.7937/K9/TCIA.2017.GJQ7R0EF).
- [16] A. Karargyris, et al., Federated benchmarking of medical artificial intelligence with MedPerf, Nat. Mach. Intell. 5 (7) (2023) 799–810, doi:[10.1038/s42256-023-00652-2](https://doi.org/10.1038/s42256-023-00652-2).
- [17] E. Calabrese, J.E. Villanueva-Meyer, J.D. Rudie, A.M. Rauschecker, U. Baid, S. Bakas, S. Cha, J.T. Mongan, C.P. Hess, The University of California San Francisco preoperative diffuse glioma MRI dataset, Radiol. Artif. Intell. 4 (6) (2022), doi:[10.1148/ryai.220058](https://doi.org/10.1148/ryai.220058).
- [18] S. Bakas, K. Farahani, M.G. Linguraru, U. Anazodo, et al., The Brain Tumor Segmentation Challenge (2022 Continuous Updates & Generalizability Assessment), Zenodo, 2022, doi:[10.5281/zenodo.6362180](https://doi.org/10.5281/zenodo.6362180).
- [19] R.W. Cox, J. Ashburner, H. Brenman, K. Fissell, C. Haselgrove, C.J. Holmes, J.L. Lancaster, D.E. Rex, S.M. Smith, J.B. Woodward, S. Strother, A (sort of) new image data format standard: NiFTI-1, 10th Annual Meeting of the Organization for Human Brain Mapping, 2004.
- [20] T. Rohlfing, N. Zahr, E. Sullivan, A. Pfefferbaum, The SRI24 multichannel atlas of normal adult human brain structure, Hum. Brain Mapp. 31 (2009) 798–819, doi:[10.1002/hbm.20906](https://doi.org/10.1002/hbm.20906).
- [21] R. McKinley, R. Meier, R. Wiest, Ensembles of densely-connected CNNs with label-uncertainty for brain tumor segmentation, BrainLes@MICCAI, 2018, doi:[10.1007/978-3-030-11726-9_40](https://doi.org/10.1007/978-3-030-11726-9_40).

- [22] F. Isensee, P. Jaeger, S. Kohl, J. Petersen, K. Maier-Hein, nnU-Net: a Self-configuring method for deep learning-based biomedical image segmentation, *Nat. Methods* 18 (2021) 1–9, doi:[10.1038/s41592-020-01008-z](https://doi.org/10.1038/s41592-020-01008-z).
- [23] B. Baheti, D. Waldmannstetter, H. Akbari, M. Bilello, B. Wiestler, J. Schwarting, S. Abidi, M. Mousa, E. Calabrese, J. Rudie, J. Villanueva-Meyer, D.S. Marcus, C. Davatzikos, A. Sotiras, B. Menze, S. Bakas, The Brain Tumor Sequence Registration (BraTS-Reg) Challenge, Zenodo, 2022, doi:[10.5281/zenodo.6362420](https://doi.org/10.5281/zenodo.6362420).
- [24] D. Ellis, M. Aizenberg, Deep Learning Using Augmentation via Registration: 1st Place Solution to the AutoImplant 2020 Challenge(2020). 10.1007/978-3-030-64327-0_6
- [25] B.B. Avants, N.J. Tustison, G. Song, P.A. Cook, A. Klein, J.C. Gee, A reproducible evaluation of ANTs similarity metric performance in brain image registration, *NeuroImage* 54 (3) (2011) 2033–2044, doi:[10.1016/j.neuroimage.2010.09.025](https://doi.org/10.1016/j.neuroimage.2010.09.025).
- [26] A. Klein, J. Andersson, B.A. Ardekani, J. Ashburner, B. Avants, M.-C. Chiang, G.E. Christensen, D.L. Collins, J. Gee, P. Hellier, J.H. Song, M. Jenkinson, C. Lepage, D. Rueckert, P. Thompson, T. Vercauteren, R.P. Woods, J.J. Mann, R.V. Parsey, Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration, *NeuroImage* 46 (3) (2009) 786–802, doi:[10.1016/j.neuroimage.2008.12.037](https://doi.org/10.1016/j.neuroimage.2008.12.037).
- [27] N.J. Tustison, B.B. Avants, Explicit B-spline regularization in diffeomorphic image registration, *Front. Neuroinf.* 7 (2013) 39, doi:[10.3389/fninf.2013.00039](https://doi.org/10.3389/fninf.2013.00039).
- [28] Y. Ou, H. Akbari, M. Bilello, X. Da, C. Davatzikos, Comparative evaluation of registration algorithms in different brain databases with varying difficulty: results and insights, *IEEE Trans. Med. Imaging* 33 (2014), doi:[10.1109/TMI.2014.2330355](https://doi.org/10.1109/TMI.2014.2330355).
- [29] B. Avants, N. Tustison, G. Song, Advanced normalization tools (ANTs), *Insight J* 1–35 (2008), doi:[10.54294/uvnhin](https://doi.org/10.54294/uvnhin).
- [30] J. Ashburner, K.J. Friston, Unified segmentation, *NeuroImage* 26 (3) (2005) 839–851, doi:[10.1016/j.neuroimage.2005.02.018](https://doi.org/10.1016/j.neuroimage.2005.02.018).
- [31] M.A. Yassa, C.E.L. Stark, A quantitative evaluation of cross-participant image registration techniques for MRI studies of aging, *Front. Aging Neurosci.* 1 (2009) 11, doi:[10.1016/j.neuroimage.2008.09.016](https://doi.org/10.1016/j.neuroimage.2008.09.016).
- [32] Y. Ou, A. Sotiras, N. Paragios, C. Davatzikos, DRAMMS: Deformable registration via attribute matching and mutual-saliency weighting, *Med. Image Anal.* 18 (2) (2014) 235–247, doi:[10.1016/j.media.2010.07.002](https://doi.org/10.1016/j.media.2010.07.002).
- [33] J.E. Iglesias, M.R. Sabuncu, Multi-atlas segmentation of biomedical images: a survey, *Med. Image Anal.* 24 (1) (2015) 205–219, doi:[10.1016/j.media.2015.06.012](https://doi.org/10.1016/j.media.2015.06.012).
- [34] N. Solak, Improved Automatic Brain Tumor Segmentation Facilitated by Intensive Data Augmentation, Master ThesisTU Graz Repository, Library and Archives(2023) 1–120. 10.3217/5972s-bev43

DuEPublico

Duisburg-Essen Publications online

UNIVERSITÄT
DUISBURG
ESSEN

Offen im Denken

ub

universitäts
bibliothek

This text is made available via DuEPublico, the institutional repository of the University of Duisburg-Essen. This version may eventually differ from another version distributed by a commercial publisher.

DOI: 10.1016/j.dib.2025.111287

URN: urn:nbn:de:hbz:465-20260108-122959-2



This work may be used under a Creative Commons Attribution - NonCommercial 4.0 License (CC BY-NC 4.0).