CSIT Conference 2025, Yerevan, Armenia, September 22 - 26

LandMatch: A collaborative annotation Tool for Landmark Matching and Benchmarking 3D volumes Registration in Longitudinal Studies

Gayane Kharatyan*†, Amel Aissaoui*, Wafae Labriji* and Lotfi Chaari*
*IRIT, Université de Toulouse, CNRS, Toulouse INP, UT3, Toulouse, France.
e-mail: firstname.lastname@irit.fr
†French University in Armenia (UFAR), Yerevan, Armenia
e-mail: I19.Kharatyan.Gayane@etud.ufar.am

***Abstract*—Accurate registration is an important step in longitudinal medical image studies. However, the scarcity of ground-truth annotations limits the reproducibility and benchmarking of the state-of-the-art methods. We introduce a lightweight, collaborative annotation tool for identifying the corresponding points in 3D longitudinal scans. Designed for radiologists to easily use, the tool enables a precise and scalable evaluation of registration algorithms through ground-truth landmarks. We demonstrate its application on the brain MRI longitudinal dataset BraTS-Reg and report benchmark metrics across standard registration methods.**

***Keywords*—Medical Image Analysis, Longitudinal Imaging, Image Registration, Landmark Annotation, landmarks matching, Benchmarking.**

I. INTRODUCTION

Longitudinal studies play a vital role in tracking disease progression, assessing treatment efficacy, and guiding clinical decisions [1]–[3]. A core challenge in these analyses is image registration across time points. This step estimates the spatial correspondence of each point from baseline to follow-up scan (see Figure 1), thus aligning the anatomical structures despite local deformations.

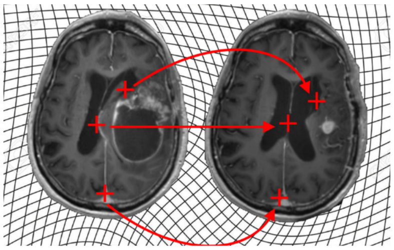


Fig. 1. Points displacement between two time-points [4].

Accurate registration ensures that the observed changes reflect actual biological progression rather than misalignment artifacts, which is critical for quantitative longitudinal analysis. Traditional registration methods [5] are typically optimized for healthy organ anatomy and often fail to accommodate complex deformations due to pathological changes, especially in oncology. However, structural changes over time—such as tumor growth, cavity resorption, and tissue deformation—pose significant challenges (see Figure 2).

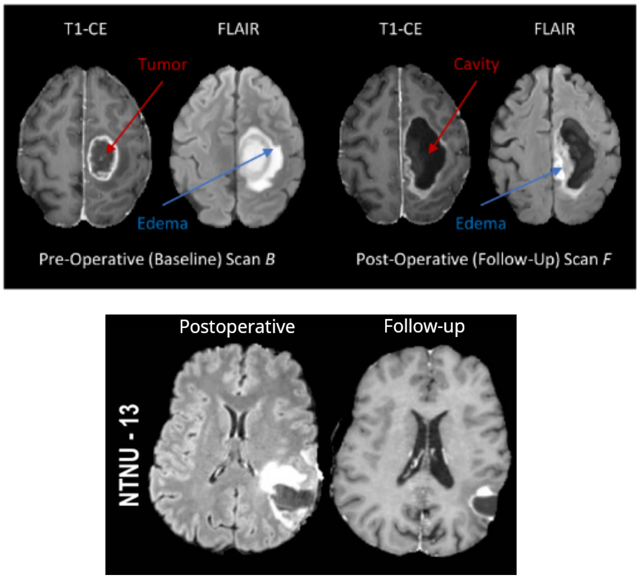


Fig. 2. Examples of two longitudinal MRI scan pairs. The first row shows a preoperative baseline scan from the BraTS-Reg Challenge dataset [4] with a visible tumor and surrounding edema, alongside its corresponding follow-up scan where a resection cavity has appeared in place of the tumor, with noticeable changes in the edema pattern. The second row shows a post-operative baseline scan from the St. Olavs University Hospital dataset [2], acquired after tumor resection, followed by a later scan showing partial cavity resorption and post-surgical evolution of surrounding tissue.

Although recent advances have introduced more sophisticated image registration algorithms [6], their evaluation in clinical settings remains limited. Most validation efforts rely on visual inspection or the use of synthetic datasets, where a known deformation is applied to enable quantitative comparison with ground truth. However, synthetic data often lack

https://doi.org/10.51408/csit2025_54

realism and scalability, while visual inspection of real data is inherently subjective. Moreover, the scarcity of publicly available annotated datasets [4] significantly hinders objective benchmarking, reproducibility, and supervised training of task-specific registration models, given the high cost and effort required for expert annotation.

To address this gap, we propose LandMatch, a lightweight and collaborative annotation tool specifically designed to generate ground-truth point correspondences across 3D volumes acquired at different time-points. This tool allows radiologists to manually identify anatomically consistent landmarks and provides an integrated framework to evaluate state-of-the-art registration algorithms using standardized metrics.

This paper is organized as follows. Section II gives an overview of the state-of-the-art annotation tools. Section III explains the motivation and system design. Section IV introduces the proposed annotation tool and its capabilities. Section V demonstrates its use on the well-known BraTS-Reg dataset, and finally, section VI provides a discussion and concludes the paper.

II. RELATED WORK

Several annotation platforms have emerged to facilitate medical image labeling, focusing primarily on segmentation tasks. For example, MONAI Label [7] supports semi-automatic segmentation with model-in-the-loop feedback in environments like 3D Slicer. Similarly, QuickDraw [8] combines segmentation and visualization tools to accelerate annotation in multimodal MRI datasets. While effective, these tools are not tailored for longitudinal correspondence annotation.

Classical tools such as ITK-SNAP [9] and 3D Slicer [10] allow for basic annotation and visualization but lack built-in support for collaborative workflows or direct support for temporal correspondence annotation. Recent efforts, such as SVRDA [11], target slice-to-volume registration and allow landmark annotation in web environments. The Uncertainty-Aware Annotation Protocol [12] and BrainMorph [13] demonstrate that corresponding points can effectively serve as a foundation for evaluating deformable registration, though these contributions are algorithmic rather than tool-centric.

To the best of our knowledge, no existing open-source system is explicitly designed to enable radiologists to annotate corresponding points in longitudinal 3D scans and evaluate registration algorithms quantitatively and collaboratively.

III. MOTIVATION AND DESIGN OBJECTIVES

High-quality ground truth annotations for image registration are vital yet rarely available, especially for longitudinal scans containing pathologies causing dramatic changes in tissue appearance (as illustrated earlier in Figure 2). To support reproducible research and scalable evaluation, we designed our tool with the following key objectives:

- **Ease of Annotation:** Intuitive interface for rapid annotation of corresponding points in 3D volumes.
- **Collaborative Workflow:** Supports loading, merging, and refining annotations from multiple annotators.
- **Low-latency Visualization:** Smooth rendering with support for different modalities (such as T1, T2, T1c, and FLAIR in case of MRI scans) and all anatomical planes.
- **Customizability:** Easily adaptable to a wide range of medical imaging datasets, including MRI, CT, and PET, for any anatomical region, with support for up to four imaging modalities simultaneously.

IV. TOOL OVERVIEW

The tool is implemented in Python and can be deployed on local machines without specialized hardware. It comprises two main components: the annotation interface and the evaluation engine.

A. Annotation Interface

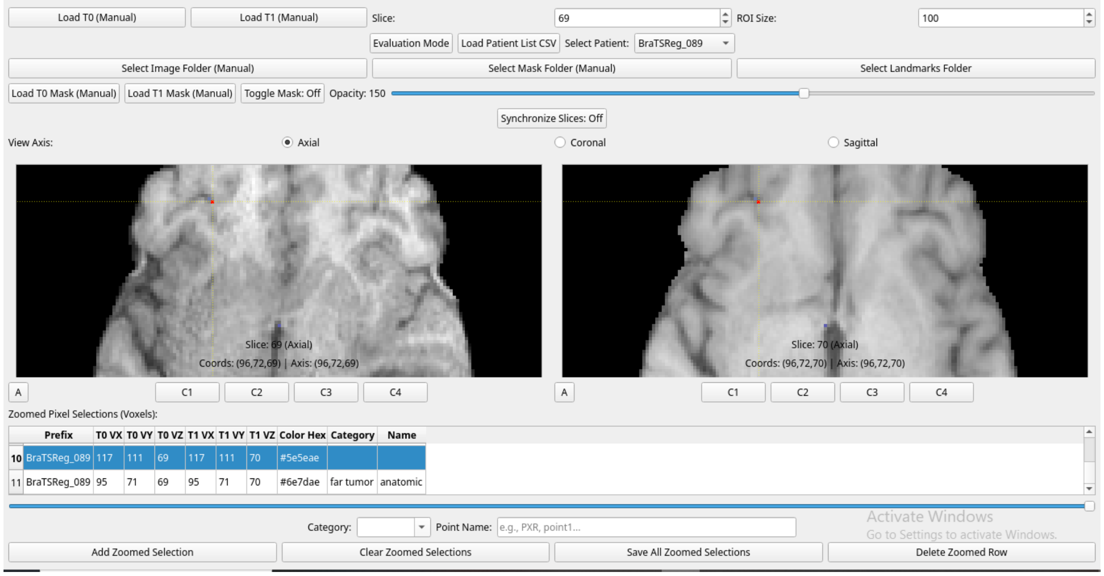
The first component of the tool is the annotation interface (Figure 3), where users provide a CSV file listing patient identifiers along with the corresponding file paths to the 3D image volumes acquired at two different timepoints. The tool displays the paired volumes side-by-side, allowing annotators to view and mark corresponding landmarks. All annotated pairs are listed for review and can be exported for further use.

1) Scan Viewing and Navigation: The interface offers intuitive and flexible navigation through 3D medical images. Users can load volumetric scans (with the well-known NIFTI Format) of two time-points simultaneously and explore them in a synchronized and non-synchronized way. It also offers synchronized orthogonal views (axial, sagittal, and coronal), each dynamically linked via a central crosshair. Interacting with the crosshair in one plane automatically updates the corresponding slices in the other two, ensuring that all views remain centered on the same spatial location. This coordinated navigation enables efficient exploration of complex anatomical structures.

The tool provides flexible visualization capabilities tailored to multi-modal imaging. Users can explore anatomical correspondence across modalities using a synchronized layout, where identical slice positions are displayed for each modality. To enhance interpretability, pre-existing segmentation masks can be overlaid on the scans with customizable transparency.

2) Manual landmarks annotation: Landmarks annotation for registration is represented as matching point pairs defined within two 3D image volumes acquired at two different time-points. The user begins by selecting a point of anatomical interest in the first volume (timepoint 1), which is recorded as the source coordinate. The corresponding location is then selected in the second volume (timepoint 2), completing the pair. Landmarks can be adjusted by dragging the control point in any of the orthogonal slice views. Once confirmed, the landmark pair is stored and visualized in both the list and the image viewers.

The tool supports real-time cross-linked navigation to facilitate accurate matching, with the same anatomical region visible across all views and time-points. Users may assign



Slice: 69

ROI Size: 100

Select Patient: BraTSReg_089

Toggle Mask: Off

Opacity: 150

Synchronize Slices: Off

View Axis: ● Axial ○ Coronal ○ Sagittal

Slice: 69 (Axial)
Coords: (96,72,69) | Axis: (96,72,69)

Slice: 70 (Axial)
Coords: (96,72,70) | Axis: (96,72,70)

Zoomed Pixel Selections (Voxels):

	Prefix	T0 VX	T0 VY	T0 VZ	T1 VX	T1 VY	T1 VZ	Color Hex	Category	Name
10	BraTSReg_089	117	111	69	117	111	70	#5e5eae		
11	BraTSReg_089	95	71	69	95	71	70	#6e7dae	far tumor	anatomic

Category:

Point Name: e.g., PXR, point1...

Fig. 3. Annotation interface displaying paired MRI volumes and landmark list.

labels to landmark pairs and can load, edit, or remove previously defined annotations. Annotations from multiple experts can be merged to support collaborative workflows and inter-observer analysis. Additionally, landmark sets are stored in a CSV format to be used in the quantitative evaluation of image registration methods.

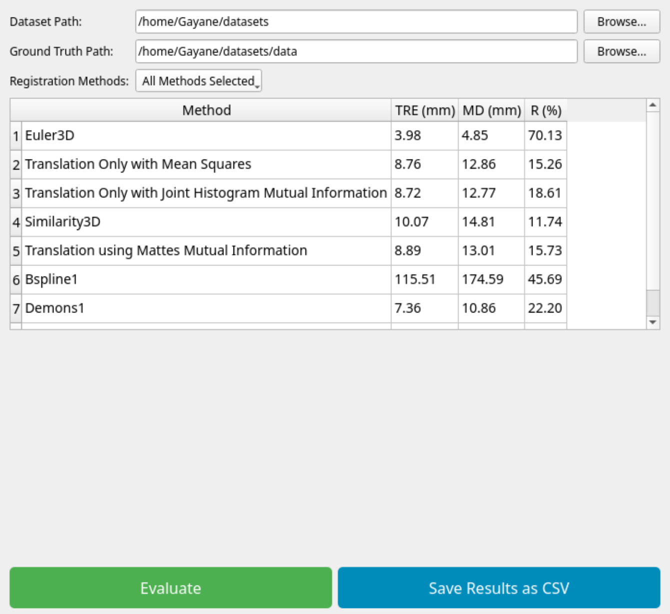
B. Evaluation Engine

The evaluation engine applies a list of off-the-shelf, most used registration algorithms [14] to the given dataset and computes quantitative metrics such as Target Registration Error (TRE), Manhattan Distance (MD), and Robustness (R). The engine supports any dataset with a compatible structure and can integrate additional algorithms or metrics.

V. TOOL DEMONSTRATION

We demonstrated the tool using sample cases from the 2022 Brain Tumor Sequence Registration (BraTS-Reg) Challenge [4]. Data consists of multi-modality brain MRI scans of patients with glioblastomas for two time-points: preoperative and post-operative. For each patient, four co-registered scans are available: T1-weighted, T1-weighted contrast-enhanced (T1CE) with gadolinium contrast agent; T2-weighted; and FLAIR. In Figure 3, we show a sample case from BraTS-Reg with the annotated landmarks.

When used to benchmark standard registration methods, the tool produced consistent and informative metrics. Table I shows the comparative performance of different algorithms [14] across several longitudinal cases, illustrating variations in registration accuracy under pathological deformations. These metrics help identify failure cases, assess method robustness, and support algorithm selection tailored to specific clinical scenarios. The built-in evaluation engine thus not only facilitates reproducible benchmarking but also accelerates method development by providing direct and interpretable feedback on alignment quality.



Dataset Path: /home/Gayane/datasets

Ground Truth Path: /home/Gayane/datasets/data

Registration Methods: All Methods Selected

	Method	TRE (mm)	MD (mm)	R (%)
1	Euler3D	3.98	4.85	70.13
2	Translation Only with Mean Squares	8.76	12.86	15.26
3	Translation Only with Joint Histogram Mutual Information	8.72	12.77	18.61
4	Similarity3D	10.07	14.81	11.74
5	Translation using Mattes Mutual Information	8.89	13.01	15.73
6	Bspline1	115.51	174.59	45.69
7	Demons1	7.36	10.86	22.20

Fig. 4. Evaluation engine interface showing registration methods and evaluation metrics

VI. CODE AND DATA AVAILABILITY STATEMENTS

LandMatch is an open-source project, developed and tested on both Linux and Windows operating systems. Installation instructions are provided in the project's GitHub repository. The source code is currently available for review upon request and will be released publicly at https://github.com/GayaneKharatyan/LandMatch under the Apache-2.0 license.

TABLE I

REGISTRATION PERFORMANCE ON BRATS-REG SAMPLES USING THE ANNOTATED LANDMARKS WITH STANDARD REGISTRATION METHODS.

Method	**TRE (mm)**	**MD (mm)**	**R (%)**
Euler3D	3.98	4.85	70.13
Translation Only with Mean Squares	8.76	12.86	15.26
Translation Only with Joint Histogram Mutual Information	8.72	12.77	18.61
Similarity3D	10.07	14.81	11.74
Translation using Mattes Mutual Information	8.89	13.01	15.73
Bspline1	115.51	174.59	45.69
Demons1	7.36	10.86	22.20
Demons2	7.44	10.96	19.57

VII. CONCLUSION

We introduced a lightweight and extensible annotation tool for generating ground-truth landmarks in longitudinal medical imaging. While demonstrated here on brain tumor MRI, the tool is modality-agnostic and supports CT, PET, and other imaging formats, making it adaptable to a wide range of diseases. It provides a reproducible and scalable solution for benchmarking registration algorithms. The pipeline is designed to handle multimodal imaging inputs, jointly leveraging information from multiple channels to enhance landmark annotation and evaluation.

In future work, we plan to integrate machine learning based keypoint suggestions to accelerate the annotation process, as well as segmentation models to automatically generate masks when these are not available in the dataset. Additional developments include extending support to more formats (such as DICOM) and integrating more sophisticated registration algorithms such as [15] in the evaluation engine.

We also aim to formalize annotation guidance through a standardized protocol defining clearly identifiable anatomical landmarks with salient features. Such a protocol will promote accurate localization, facilitate quality control (e.g., detecting artifacts, verifying acquisition consistency), and ensure reproducibility, thereby reducing inter-annotator variability. While this protocol will be is designed for brain imaging, its structure can be adapted to develop organ-specific guidelines for other anatomical regions and imaging contexts.

REFERENCES

[1] J. E. van Timmeren, J. Bussink, P. Koopmans, R. J. Smeenk, and R. Monshouwer, “Longitudinal image data for outcome modeling”, *Clinical Oncology*, vol. 38, p. 103610, 2025.

[2] S. Cepeda, L. T. Luppino, A. Pérez-Núñez, O. Solheim, S. García-García, M. Velasco-Casares, A. Karlberg, L. Eikenes, R. Sarabia, I. Arrese, et al., “Predicting regions of local recurrence in glioblastomas using voxel-based radiomic features of multiparametric postoperative MRI”, *Cancers*, 2023.

[3] W. Labriji, A. Aissaoui, J.-Y. Tourneret, L. Chaari, E. M. Cohen-Jonathan and S. Ken, “Longitudinal MRI radiomics and MRSI metabolic signatures for progression-free survival prediction in glioblastoma”, in *Proc. ESMRMB, 41st Annual Meeting*, Marseille, France, 2025, in press.

[4] B. Baheti et al., “The Brain Tumor Sequence Registration (BraTS-Reg) Challenge: Establishing correspondence between pre-operative and follow-up MRI scans of diffuse glioma patients”, 2024, [Online]. Available: https://arxiv.org/abs/2112.06979.

[5] M. A. Viergever, J. B. A. Maintz, S. Klein, K. Murphy, M. Staring, and J. P. W. Pluim, “A survey of medical image registration”, *Medical Image Analysis*, vol. 33, pp. 140–144, 2016.

[6] J. Chen, Y. Liu, S. Wei, Z. Bian, S. Subramanian, A. Carass, J. L. Prince, and Y. Du, “A survey on deep learning in medical image registration: New technologies, uncertainty, evaluation metrics, and beyond”, *Medical Image Analysis*, vol. 100, p. 103385, 2025.

[7] A. Diaz-Pinto, S. Alle, V. Nath, Y. Tang, A. Ihsani, M. Asad, F. Pérez-García, P. Mehta, W. Li, M. Flores, H. R. Roth, T. Vercauteren, D. Xu, P. Dogra, S. Ourselin, A. Feng, and M. J. Cardoso, “MONAI Label: A framework for AI-assisted interactive labeling of 3D medical images”, *Medical Image Analysis*, vol. 95, p. 103207, 2024.

[8] D. Syomichev, P. Gopinath, G.-L. Wei, E. Chang, I. Gordon, A. Seifu, R. Pemmaraju, N. Peri, and J. Purtilo, “QuickDraw: Fast visualization, analysis and active learning for medical image segmentation”, *HCI International 2025 Posters, Communications in Computer and Information Science*, vol. 2522, pp. 357–364, 2025.

[9] P. A. Yushkevich, J. Piven, H. C. Hazlett, R. G. Smith, S. Ho, J. C. Gee, and G. Gerig, “User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability”, *NeuroImage*, vol. 31, no. 3, pp. 1116–1128, 2006. [Online]. Available: http://www.itksnap.org/

[10] A. Fedorov, R. Beichel, J. Kalpathy-Cramer, J. Finet, J.-C. Fillion-Robin, S. Pujol, C. Bauer, D. Jennings, F. Fennessy, M. Sonka, J. Buatti, S. Aylward, J. V. Miller, S. Pieper, and R. Kikinis, “3D Slicer as an image computing platform for the quantitative imaging network”, *Magnetic Resonance Imaging*, vol. 30, no. 9, pp. 1323–1341, 2012.

[11] W. Luo, A. T. Bagur, P. Aljabar, G. Ralli, and S. M. Brady, “SVRDA: A web-based dataset annotation tool for slice-to-volume registration”, 2023, [Online]. Available: https://arxiv.org/abs/2311.15536.

[12] L. Peter, D. C. Alexander, C. Magnain, and J. E. Iglesias, “Uncertainty-aware annotation protocol to evaluate deformable registration algorithms”, *IEEE Transactions on Medical Imaging*, vol. 40, no. 8, pp. 2053–2065, 2021.

[13] A. Q. Wang, R. Saluja, H. Kim, X. He, A. Dalca, and M. R. Sabuncu, “BrainMorph: A foundational keypoint model for robust and flexible brain MRI registration”, *Machine Learning for Biomedical Imaging*, vol. 3, pp. 181–215, May 2025.

[14] H. J. Johnson, R. Beare, Z. Yaniv, and B. C. Lowekamp, “SimpleITK image-analysis notebooks: A collaborative environment for education and reproducible research”, *Digit. Imaging*, vol. 38, no. 8, pp. 1788–1800, 2018.

[15] G. Balakrishnan, A. V. Dalca, A. Zhao, M. R. Sabuncu, and J. Guttag, “Voxelmorph: A learning framework for deformable medical image registration”, *IEEE Transactions on Medical Imaging*, vol. 38, no. 8, pp. 1788–1800, 2019.