Published in final edited form as:

Nat Methods. 2021 August ; 18(8): 845–846. doi:10.1038/s41592-021-01218-z.

CloudReg: automatic terabyte-scale cross-modal brain volume registration

Vikram Chandrashekhar¹, Daniel J. Tward², Devin Crowley¹, Ailey K. Crow³, Matthew A. Wright⁴, Brian Y. Hsueh^{3,4}, Felicity Gore^{3,4}, Timothy A. Machado^{3,4}, Audrey Branch⁵, Jared S. Rosenblum⁶, Karl Deisseroth^{3,4}, Joshua T. Vogelstein¹

¹Center for Imaging Science, Johns Hopkins University, Baltimore, MD, USA.

²Department of Computational Medicine and Neurology, University of California Los Angeles, Los Angeles, CA, USA.

³CNC Program, Stanford University, Stanford, CA, USA.

⁴Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford, CA, USA.

⁵Department of Psychological and Brain Sciences, Johns Hopkins University, Baltimore, MD, USA.

⁶Neuro-Oncology Branch, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA.

To the Editor

Imaging methods such as magnetic resonance imaging (MRI), micro-computed tomography (microCT) and light-sheet microscopy (LSM) of cleared tissue samples can generate intact anatomic and molecular whole-brain data. However, each modality produces unique artifacts based on the physical principles of the technique, including intensity inhomogeneity due to magnetic field bias in MRI or microscope optics in LSM and beam hardening in microCT^{1, 2, 3}. These artifacts and the size of the datasets generated pose a substantial challenge in data handling, cross-modal image registration, and analysis. Visualization and anatomically relevant analysis of high-resolution, multi-field-of-view (mFOV) datasets require preprocessing to remove artifacts, stitching into a complete volume, and registration to a reference atlas^{3,4}. Each step presents specific challenges. First, stitching acquired fields of view (FOVs) into a complete volume is computation and time intensive. Second,

jovo@jhu.edu .
Author contributions

V.C. conceived the study, drafted the original manuscript, performed data analysis, built a reproducible pipeline and revised the manuscript. D.J.T. conceived the study, revised the manuscript, performed data analysis and supervised data analysis. D.C. ported registration algorithm to Python, performed data analysis and revised the manuscript. A.K.C., M.A.W., B.Y.H., F.G., T.A.M. and A.B. provided CLARITY/iDISCO/SHIELD LSM data and revised the manuscript. J.S.R. helped draft the original manuscript, revised the manuscript, provided data and placed landmarks for determining registration accuracy. K.D. and J.T.V. conceived the study, supervised data analysis and revised the manuscript.

Competing interests

The authors declare no competing interests.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41592-021-01218-z>.

preprocessing requires correcting artifacts unique to each modality and sample. Third, registration methods are often intramodal, have manual components, or are limited by artifacts introduced by specimen preparation and imaging^{5,6}. Finally, visualization of these terabyte-scale datasets on a local machine can be computationally intensive⁷.

To address these challenges, we present CloudReg, an automatic, cross-modal, cloud-based pipeline that performs local and global intensity correction, stitching, nonlinear image registration and interactive online visualization through Neuroglancer (<https://github.com/google/neuroglancer>)^{4,8,9}. We combine state-of-the-art, open-source tools with algorithms for distributed local intensity correction and cross-modal registration developed in this work. We applied CloudReg to various datasets acquired with *in vivo* human brain MRI, *ex vivo* macaque brain MRI, *ex vivo in situ* mouse brain microCT, and cleared mouse and rat brains imaged with LSM.

The CloudReg pipeline is launched from a local machine and runs automatically in the cloud to perform intensity correction, stitching, registration and upload, allowing visualization with Neuroglancer (Fig. 1 and Supplementary Note). All the data are stored in the cloud and routed through a content delivery network and firewall to facilitate efficient and secure visualization. With our deployment of Neuroglancer, the resulting visualization of terabyte-scale datasets can be shared via a URL and accessed from anywhere with a web browser and internet connection (Supplementary Fig. 1)⁹.

We initially developed CloudReg using high-resolution, LSM-imaged CLARITY mouse brain data and used the Allen Reference Atlas (ARA) Common Coordinate Frame Version 3 (CCFv3) as the reference atlas (Fig. 1)¹⁰. Tissue clearing procedures and optics of LSM introduce sample-specific artifacts and intensity inhomogeneity in the imaged samples, which we corrected with our mFOV-based preprocessing algorithm. Intensity inhomogeneity, in particular, makes automatic, intensity-based registration a challenge. To minimize this per-FOV artifact, we developed a parallelized intensity correction algorithm. To compute this correction, we uniformly subsample the mFOV data in voxel space in all three dimensions, compute the mean across subsampled FOVs in parallel, and apply the N4 bias correction algorithm to the resulting mean FOV (Supplementary Fig. 2). Our intensity correction algorithm accounts for differences in tissue scattering from different clearing methods by estimating the intensity correction directly from the data. These preprocessed data are then stitched using Terasticher⁴. To minimize intensity inhomogeneity at the whole-brain scale, we apply the N4 bias correction algorithm to the whole stitched volume directly (Supplementary Fig. 2).

The fully preprocessed sample is then registered to a corresponding reference atlas. To enable registration of tissue samples that contain intensity inhomogeneity, we developed a spatially varying polynomial intensity transform estimation procedure, building on the expectation-maximization large deformation diffeomorphic metric mapping (EM-LDDMM) algorithm that we previously developed (Fig. 1, Supplementary Fig. 3 and Supplementary Video 1)⁸. Our extension of EM-LDDMM built into CloudReg enables cross-modal registration of a diversity of brain volume samples with artifacts, tears and deformations (Supplementary Fig. 4). Our spatially varying intensity transform is a polynomial function of

image intensity and can therefore estimate local non-monotonic mappings of intensity from one sample to another at every voxel in the image, whereas mutual information operates on histograms of the global image intensity.

Cloud-based services allow scalability, lower up-front costs, and ease of infrastructure setup and maintenance, but may be limited by long-term cost. With our pipeline running on Amazon Web Services, uploading to and downloading from cloud storage does not have an associated cost since both occur on a computing instance in the same region as our storage. Use of file formats that minimize the number of separate objects stored can further mitigate costs.

CloudReg can accurately correct intensity, stitch, register, and visualize terabyte-scale brain volumes with artifacts and tears (Supplementary Fig. 5). CloudReg is immediately applicable to brain volumes spanning a variety of species—including mouse, rat, monkey and human—and imaging modalities. Extensive documentation is available for deployment of CloudReg (<https://cloudreg.neurodata.io>)¹¹.

Data availability

The datasets in this study are available from the corresponding author on reasonable request.

Code availability

The CloudReg pipeline is open-source and available under an Apache 2.0 license at <https://github.com/neurodata/CloudReg> and at <https://doi.org/10.5281/zenodo.4949737> (ref. ¹¹).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This work was supported by R01 AG066184/AG/NIA NIH HHS/United States and by the National Science Foundation (NSF) under NSF Award Number EEC-1707298. The authors would also like to thank Microsoft Research for supporting this work. V.C. was supported by UPENN/NIH grant 133284. D.J.T. was supported by the Kavli Neuroscience Discovery Institute, the Karen Toffler Charitable Trust through the Toffler Scholar Program, and the NIH (U19MH114821). A.K.C. was supported by DARPA grant W911NF-14-2-0013 and NIMH TR01 R01 MH099647. M.A.W. was supported by NIDDK grant K08MH113039, DARPA grant W911NF-14-2-0013 and NIMH TR01 R01 MH099647. B.Y.H. was supported by DARPA grant W911NF-14-2-0013 and NIMH TR01 R01 MH099647. F.G. was supported by NARSAD Young Investigator Award from BBRF, K99 from NIDA (1K99DA050662-01), DARPA grant W911NF-14-2-0013 and NIMH TR01 R01 MH099647. T.A.M. was supported by the AP Giannini Foundation, a Stanford Dean's Fellowship, NIH/NINDS (K99-NS116122), DARPA grant W911NF-14-2-0013 and NIMH TR01 R01 MH099647. A.B. was supported by National Institute of Health (NIH) (grant P01AG009973 and the Johns Hopkins University Kavli Neuroscience Discovery Institute Postdoctoral Fellowship. J.S.R. was supported in part by the intramural program of the NCI at the NIH. K.D. was supported by DARPA grant W911NF-14-2-0013 and NIMH TR01 R01 MH099647. J.T.V. was supported by R01 AG066184/AG/NIA NIH HHS/United States, EEC-1707298, UPENN/NIH grant 133284 and Microsoft Research.

References

1. Zhuo J & Gullapalli R *PMR Radiographics* 26, 275–297 (2006). [PubMed: 16418258]
2. Barrett JF & Keat N *Radiographics* 24, 1679–1691 (2004). [PubMed: 15537976]
3. Tomer R, Ye L, Hsueh B & Deisseroth K *Nat. Protoc* 9, 1682–1697 (2014). [PubMed: 24945384]

4. Bria A & Iannello G BMC Bioinformatics 13, 316 (2012). [PubMed: 23181553]
5. Goubran M et al. Nat. Commun 10, 5504 (2019). [PubMed: 31796741]
6. Renier N et al. Cell 165, 896–910 (2016). [PubMed: 27153497]
7. Husz ZL, Burton N, Hill B, Milayev N & Baldock RA BMC Bioinformatics 13, 122 (2012). [PubMed: 22676296]
8. Tward D et al. Front. Neurosci 14, 52 (2020). [PubMed: 32116503]
9. Vogelstein JT et al. Nat. Methods 15, 846–847 (2018). [PubMed: 30377345]
10. Wang Q et al. Cell 181,936–953.e20 (2020). [PubMed: 32386544]
11. Chandrashekhar V & Tward DJ (2021). *CloudReg* (version 1.0). Zenodo 10.5281/zenodo.4949737 (2021).

Fig. 1 |. CloudReg.
a, Overview of the pipeline. **b**, Example outputs. Each row demonstrates registration of brain imaging data of various species to the corresponding atlas using CloudReg. The data from the autofluorescence channel is used for samples imaged with a light-sheet microscope (LSM). ARA CCFv3, Allen Reference Atlas Common Coordinate Frame version 3; CLARITY, clear lipid-exchanged anatomically rigid imaging/immunostaining-compatible tissue hydrogel; COLM, CLARITY-optimized light-sheet microscopy; GB, gigabyte; iDISCO, immunolabeling-enabled three-dimensional imaging of solvent-cleared organs; MB, megabyte; microCT, micro-computed tomography; TB, terabyte.