Machine Learning Enabled Brain Segmentation for Small Animal Neuroimaging Registration

Hendrik J. Klug[1,2], Berkan Lafci[2], Zhiva Skachokova[2], Markus Rudin[2], Daniel Razansky[2], Horea-Ioan Ioanas[2,3]

[1]Department of Information Technology and Electrical Engineering, ETH Zurich, Switzerland
[2]Institute for Biomedical Engineering, ETH and University of Zurich, Switzerland
[3]Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, United States of America

ETH zürich
Universität Zürich UZH

Abstract

In biomedical imaging, between-subject comparability is attained at the voxel level via image registration. This process requires several data preparation steps, of which tissue segmentation remains challenging in preclinical research. In preclinical neuroimaging, the process of brain extraction is the simplest tissue segmentation process, yet, remains insufficiently accurate in extant registration workflows. On account of this, current solutions rely on human brain extraction library adaptations, which introduce artefacts via rostrocaudal cropping. Further, advanced workflows aware of brain extraction shortcomings may forego this stem entirely, and insted introduce registration optimizations aiming to function without prior brain extraction — which may nonetheless remain susceptible to non-brain-tissue hyperintensities.

Here, we present a deep learning framework for multi-contrast MRI brain tissue classification, and benchmark its performance with respect to improving state-of-the-art registration workflows.

Objectives

- Develop a brain extraction framework based on "imperfect prior" data from a registration workflow which provides medium-accuracy registration without employing brain extraction.
- Analyze classification performance.
- Integrate this framework into a state-of-the-art small animal registration workflow [1].
- Verify whether pre-registration brain extraction improves the overall registration quality of the workflow.

Workflow Integration

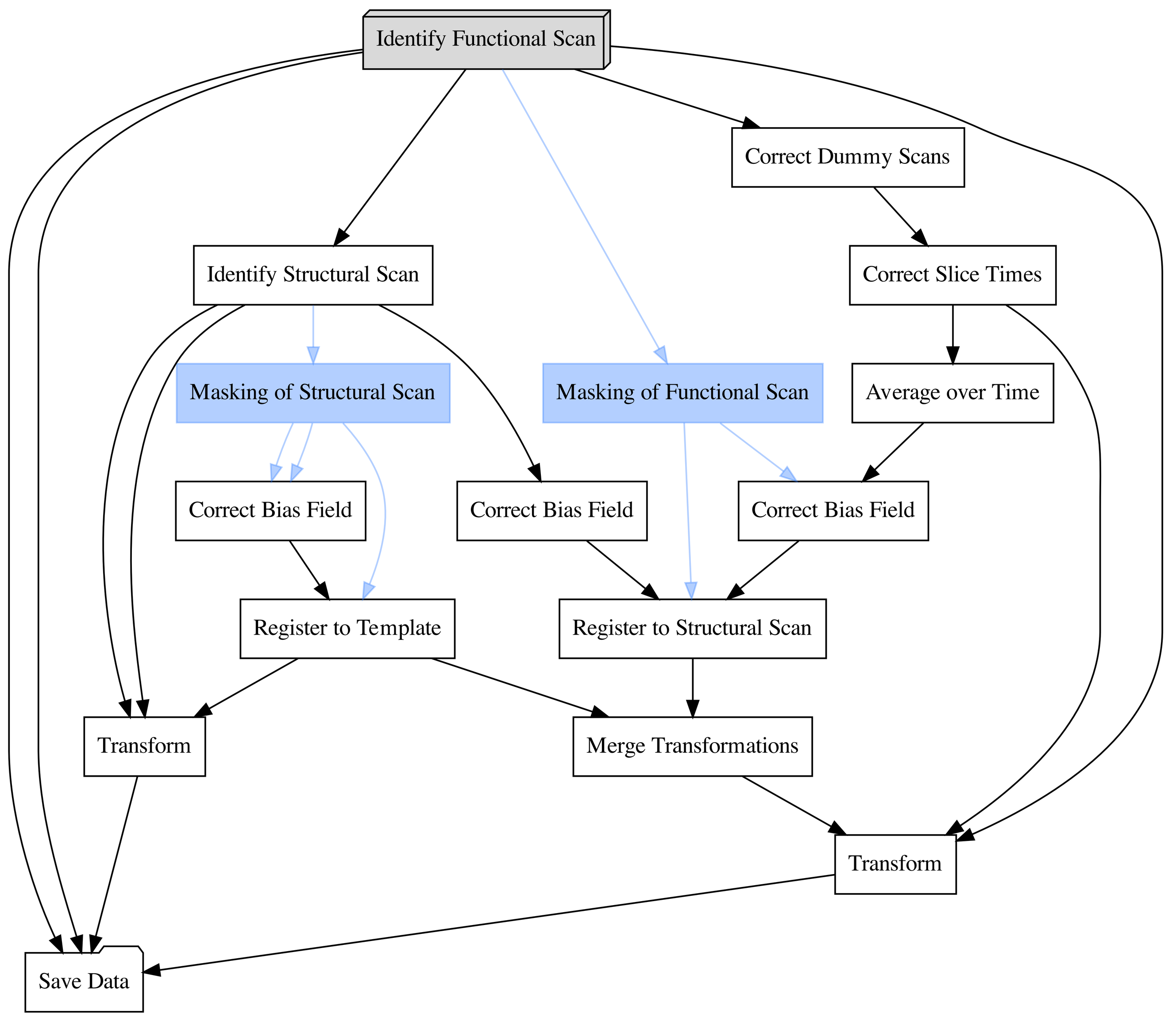


Figure 1: "SAMRI Generic Masked" workflow, a variant of the "SAMRI Generic" workflow [1], which includes two additional nodes (shown in blue) providing the workflow with both the masked data and the binary mask as produced by the MLEBE [2] brain extraction framework developed for this study.

Figure 2: The MLEBE package takes as input the unprocessed data and produces as outputs the masked data and with the binary mask. Both are used to restrict the similarity metric computation area in the "SAMRI Generic Masked" workflow, derived from the SAMRI [1] package.

Brain extraction framework

- The brain tissue classification is performed via a trained 3D U-Net [3].
- The training dataset consists of priors obtained from the "SAMRI Generic" workflow SAMRI [1], which registers unmasked brain scans to a template.
- Some data optimization steps were added to create the training dataset:
  - Expert operators blacklisted inaccurate registration results.
  - The transformed data was normalized (substracting the whole image mean and dividing by the standard deviation, fig. 4).
  - Annotations extending beyond the experimental data range were set to non-brain i.e. 0 (fig. 4).
  - Some random transformations were performed to augment the data — e.g. rotations of up to 20°, random bias field addition, and horizontal as well as vertical flips.

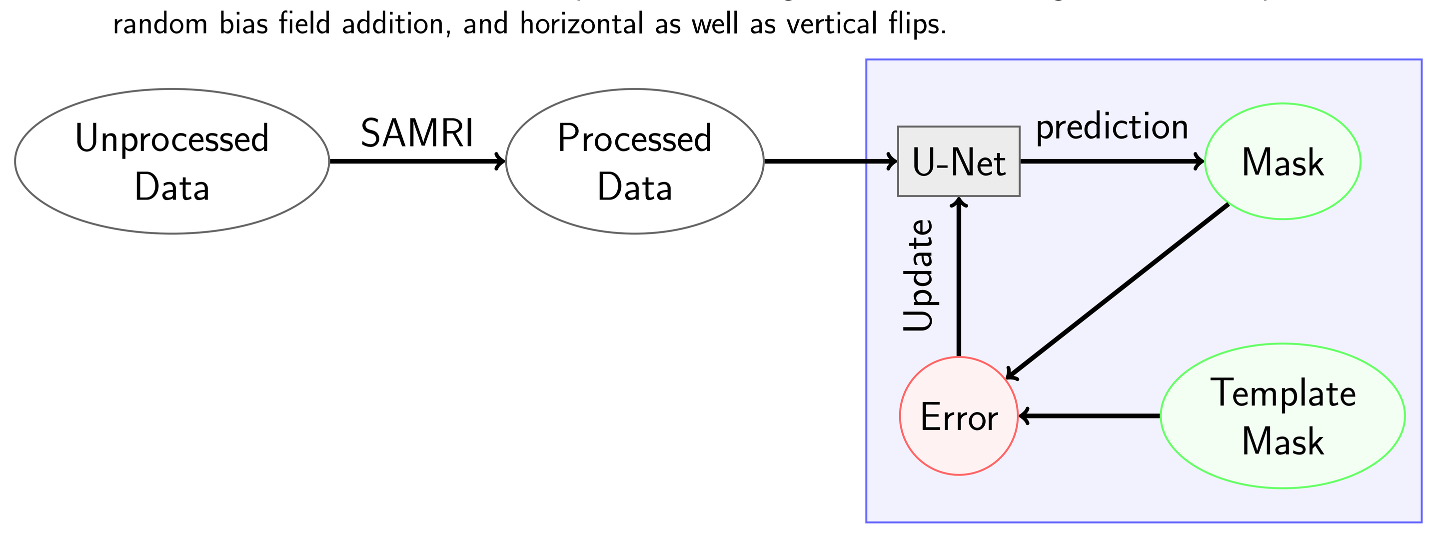


Figure 3: Flowchart depicting the training process of the U-Net model. Outside the blue box are the processing steps that map the data to the template reference space (i.e. create the training dataset) Inside the blue box are the training steps of the U-Net model. The model predicts a mask for each preprocessed input 3D image, which is then compared to the template mask using the Dice score. The parameters of the model are then updated such that the Dice score is maximized.

Figure 4: Comparison of an unprocessed slice (left) and a preprocessed slice (right) where the volume was normalized (see intensities) and the mask adapted to the NIfTI data (see blue overlay).

Results

We evaluate the effects of our classifier on a full-fledged registration workflow via the benchmarking algorithms from [1]. Additionally, we show a qualitative comparison between data processed with our Masked workflow and data processed with the Generic workflow (fig. 5).

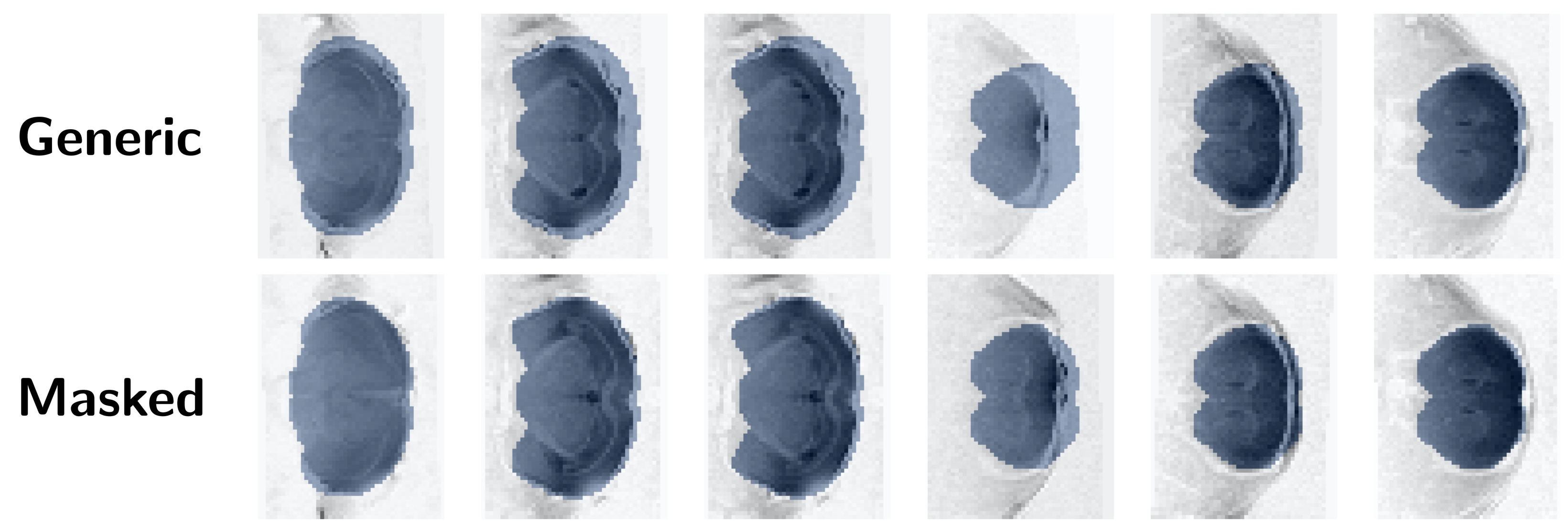


Figure 5: **The masked workflow accurately transforms the image to the template space.** The comparison of the Generic (first row) and the Masked (second row) workflow shows that the Masked workflow accurately maps the brain region to the template and prevents the inclusion of non-brain voxels.

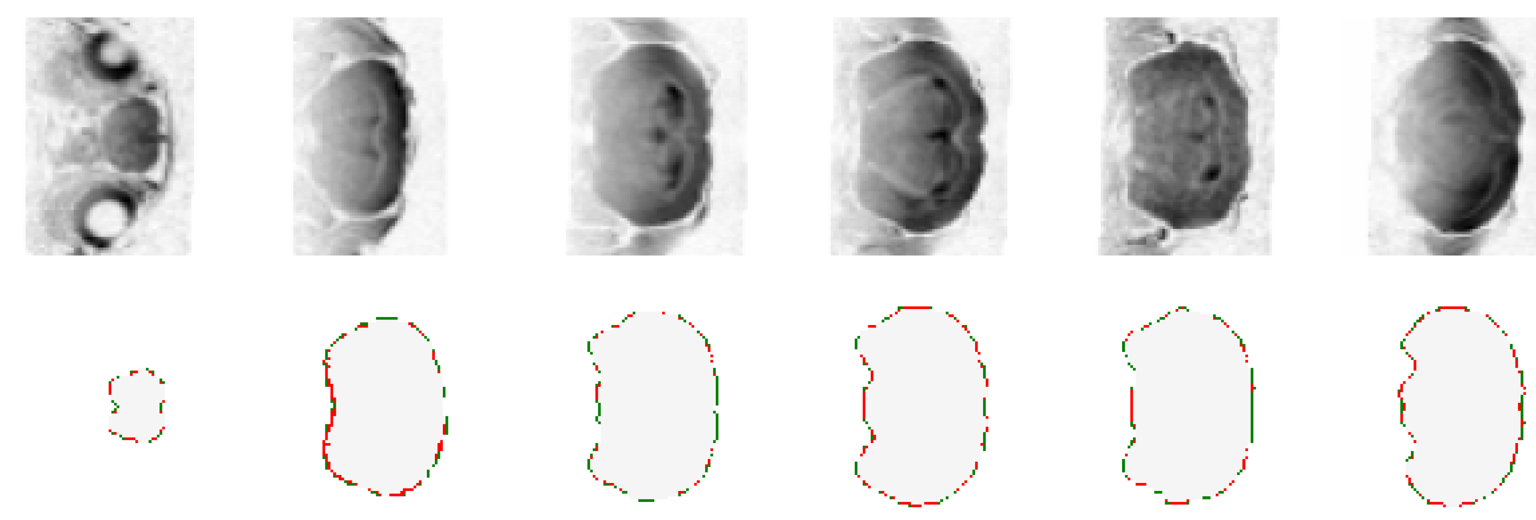
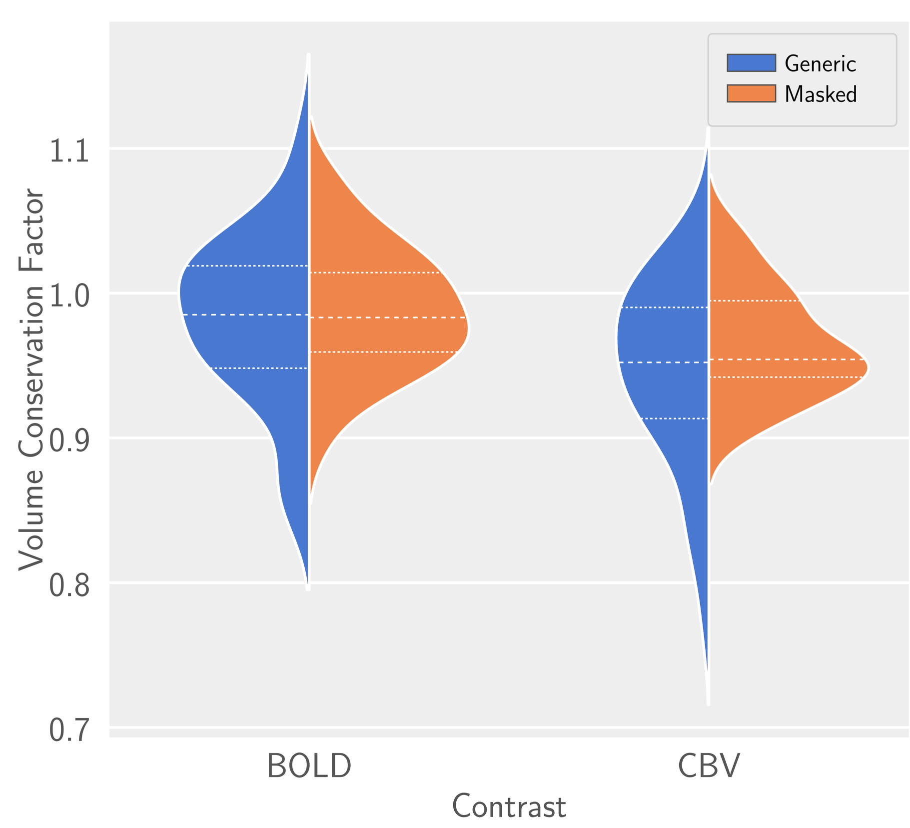
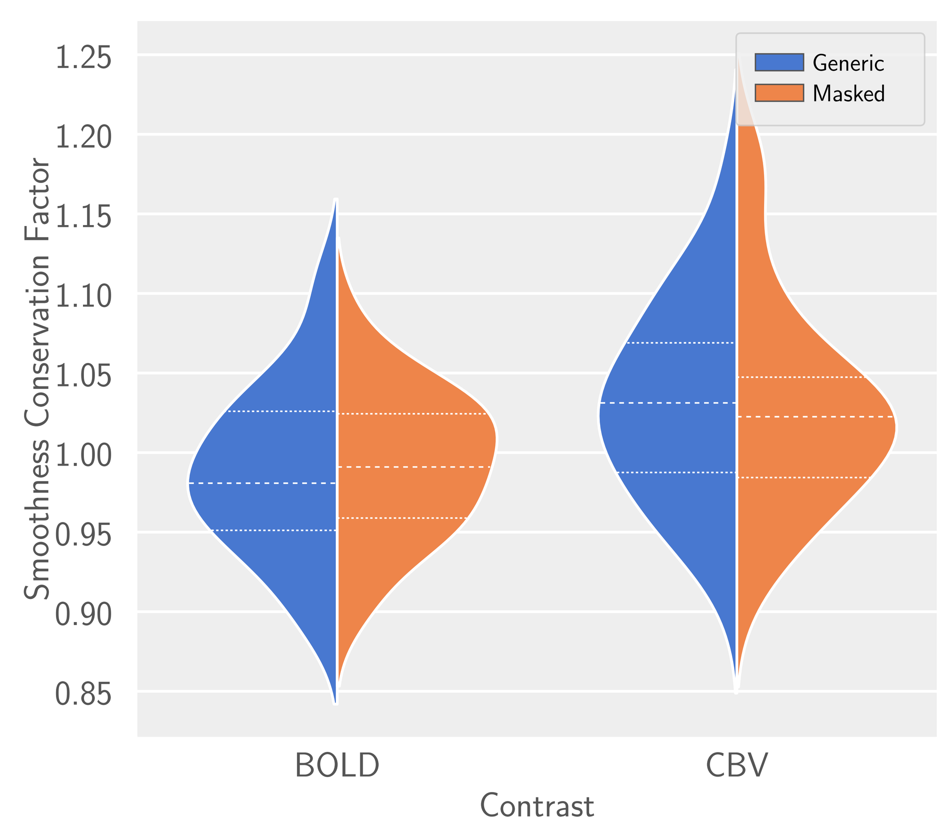


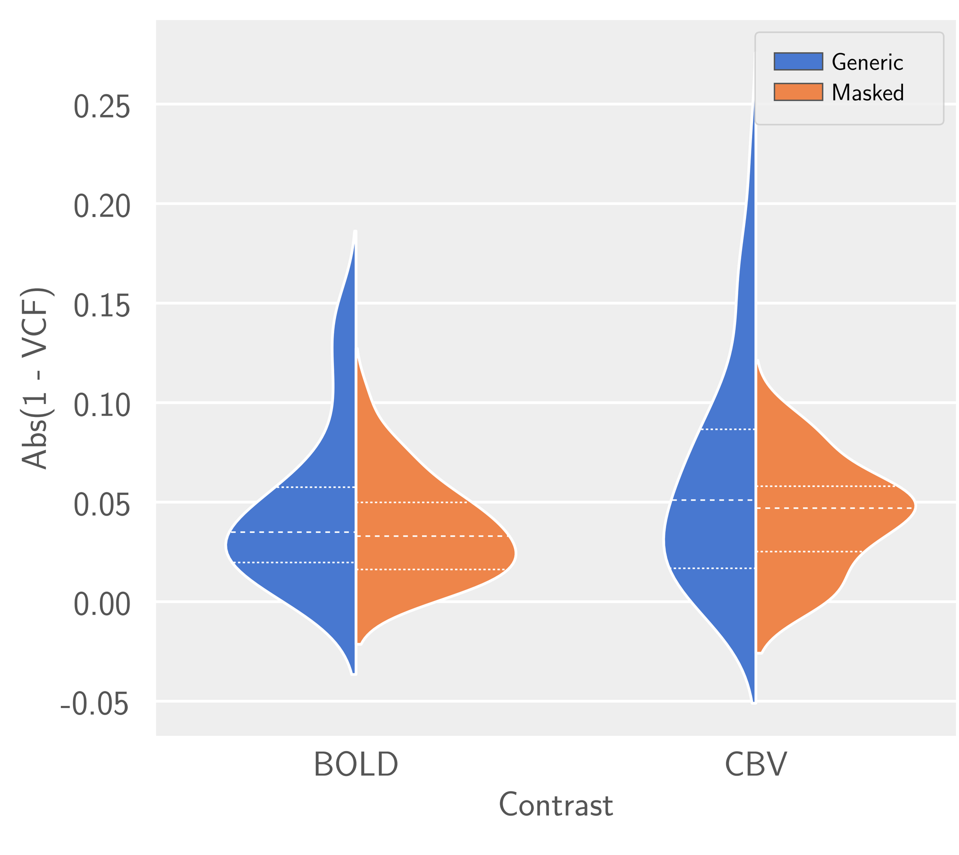
Figure 6: **The Classifier predicts a similar mask to the ground truth.** Randomly picked plots from the test set illustrate the predictions of the classifier (shown in red) in contrast to the template mask (shown in green). The overlap of the predicted mask and the template mask is shown in grey.



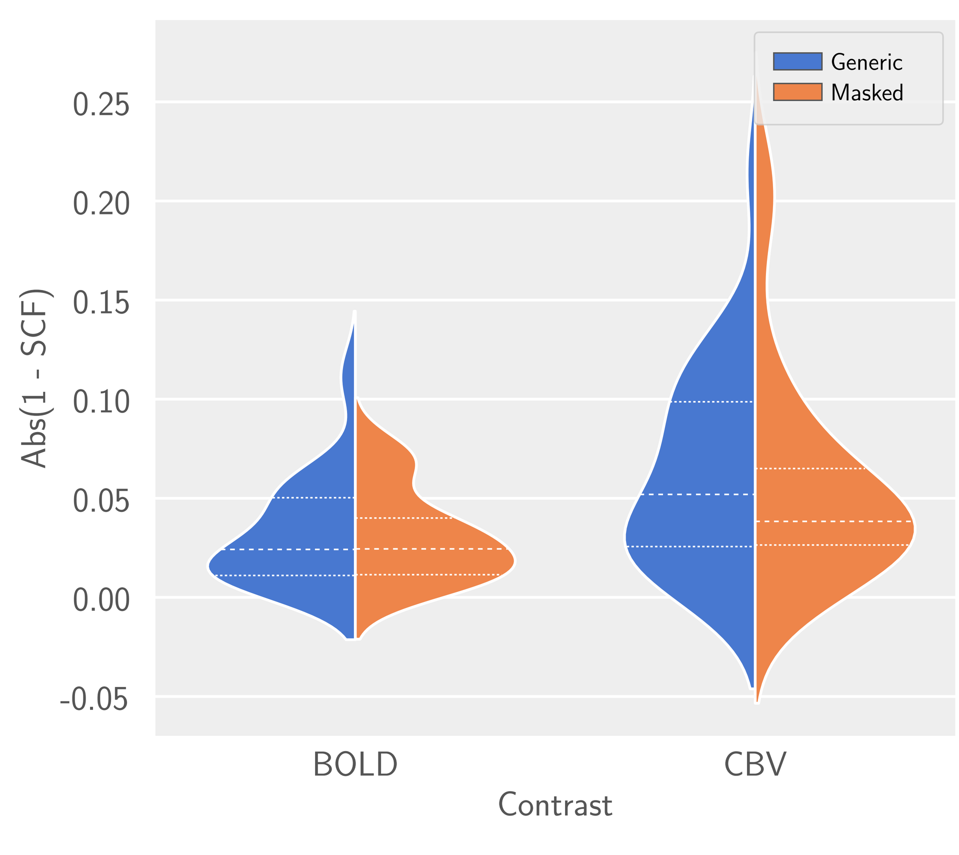
(a) Comparison of the VCF across workflows and functional contrasts.



(b) Comparison of the SCF across workflows and functional contrasts.



(c) Comparison of the distributions of the absolute VCF errors, across workflows and functional contrasts.



(d) Comparison of the distributions of the absolute SCF errors, across workflows and functional contrasts.

Figure 7: **Both the SAMRI Generic and the Masked workflow optimally and reliably conserve volume and smoothness, the latter showing values that are closely distributed to 1.** Plots showing the distribution of two target metrics in the first row, together with the respective distributions of the absolute distances to 1 in the second row. Dashed lines in the colored distribution densities indicate the sample mean and dashed lines the inner quartiles.

Results and Discussion

- Introducing MLEBE-based brain extraction to the SAMRI workflow considerably improves registration accuracy — with respect to volume conservation, smoothness conservation, and qualitative evaluation.
- The brain extraction classifier is trained with data that is preprocessed by the SAMRI Generic registration workflow such that large data collections can be leveraged with no manual labeling required. The classifier can thus be incrementally perfected with additional data that can be obtained via the novel SAMRI Masked workflow.
- The FOSS distribution model [4] for the classifier, workflow, and the article, allows users to easily take advantage of the library and reexecure the steps described herein.
- We make all functions publicly available through the mlebe Python package [2], including those used for masking in the workflow and those used to train the classifier [5].

References

[1] H.-I. Ioanas, M. Marks, V. Zerbi, M. F. Yanik, and M. Rudin, "An optimized registration workflow and standard geometric space for small animal brain imaging," *NeuroImage*, vol. 241, p. 118386, Nov. 2021.

[2] H. Klug, "github, mlebe." url: https://github.com/Jimmy2027/MLEBE.

[3] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," *arXiv:1505.04597 [cs]*, May 2015.

[4] H.-I. Ioanas and M. Rudin, "Reproducible self-publishing for Python-based research," EuroSciPy, Aug. 2018.

[5] Horea-Ioan Ioanas, Zhiva Kalinova Skachokova, B. J. Saab, and M. Rudin, "Preprocessed data used for the mlebe package.," 2020.

DOI: 10.5281/zenodo.5226215

EMIM 2021

hendrik.klug@gmail.com