

INTRODUCTION

Background

- Sharing of structural MRI important for open science.
- Facial anatomy may reveal participant identity.
- De-identification improves privacy but may alter anatomy.
- Atlas-based parcellation accuracy affected.

Motivation

- Reliability of open-source atlas-based software after anonymisation.
- Quantifying de-identification effects on GIF parcellation.

Aim

- Evaluation of the effect of de-identification on GIF parcellation fidelity.
- Identification of best de-identification method for open-source atlas-based software release.

METHODS

Data

- Atlas priors: SABRE 101 T1-weighted scans.
- Evaluation: ADNI-3, 150 participants (50 healthy; 50 mild cognitive impairment; 50 dementia).

Experimental design

- De-identification methods: **PyDeface**, **MiDeFace**, **mri_reface**, **SynthStrip**.
- Defaced atlas priors generated from SABRE scans.
- **GIF parcellation** tested with **raw ADNI scans** and **defaced ADNI scans**, across **156 regions**.
- Tested on **raw SABRE priors + raw ADNI scans**.

Evaluation

- **Dice**, **volume similarity**, and **95th percentile Hausdorff distance**.
- **Volume-weighted** metrics over **5 aggregated regions + overall**.
- **Wilcoxon signed-rank tests** with **FDR** correction.

RESULTS

Key Findings

- GIF parcellation remained consistent across methods.
- **PyDeface performed best**, with the highest Dice and lowest HD95 across all aggregated regions.
- **Raw input outperformed defaced input**.
- mri_reface and SynthStrip showed larger drops with defaced input.

CONCLUSIONS

- **GIF robust to de-identification.**
- Best de-identification method **depends on atlas-based software** and **experiment aim**.
- **Method-specific de-identification recommended** for open-source release of atlas-based tools.

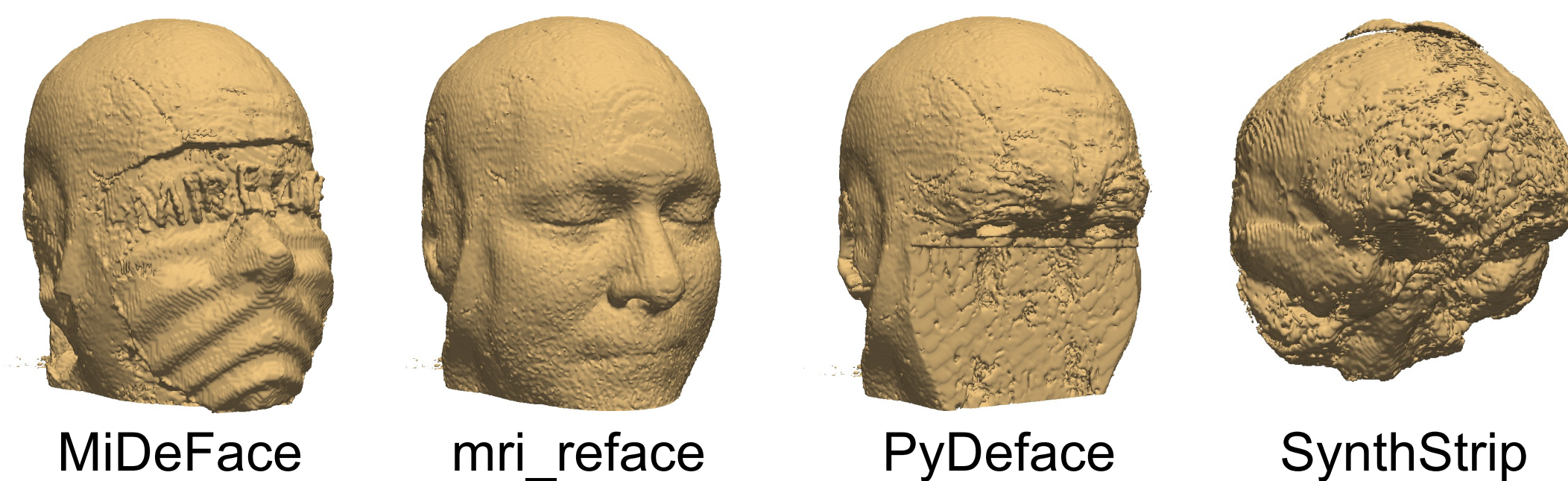


Figure 1: Example 3D volume outputs produced by each de-identification method.

Table 1: Parcellation performance across de-identification methods and preprocessing configurations. Bold values indicate the best performance for each metric.

Pairwise median difference in Dice Coefficient PyDeface vs Others (Defaced Atlas + Raw Input)						
Method	Full Brain	White Matter	Grey Matter	CSF	Basal Ganglia	Hippocampus
MiDeFace	0.023 (0.976 vs 0.953) ***	0.014 (0.987 vs 0.973) ***	0.031 (0.964 vs 0.933) ***	0.021 (0.979 vs 0.956) ***	0.014 (0.990 vs 0.975) ***	0.012 (0.989 vs 0.976) ***
mri_reface	0.028 (0.976 vs 0.947) ***	0.017 (0.987 vs 0.970) ***	0.039 (0.964 vs 0.924) ***	0.029 (0.979 vs 0.948) ***	0.018 (0.990 vs 0.971) ***	0.012 (0.989 vs 0.976) ***
synthstrip	0.060 (0.976 vs 0.916) ***	0.027 (0.987 vs 0.961) ***	0.067 (0.964 vs 0.898) ***	0.086 (0.979 vs 0.895) ***	0.036 (0.990 vs 0.953) ***	0.018 (0.989 vs 0.971) ***

Pairwise median difference in Hausdorff Distance 95th Percentile PyDeface vs Others (Defaced Atlas + Raw Input)						
Method	Full Brain	White Matter	Grey Matter	CSF	Basal Ganglia	Hippocampus
MiDeFace	-0.434 (0.029 vs 0.475) ***	-0.288 (0.000 vs 0.288) ***	-0.564 (0.056 vs 0.663) ***	-0.007 (0.001 vs 0.008) ***	-0.000 (0.000 vs 0.000) ***	-0.000 (0.000 vs 0.000) **
mri_reface	-0.620 (0.029 vs 0.655) ***	-0.333 (0.000 vs 0.333) ***	-0.768 (0.056 vs 0.837) ***	-0.837 (0.001 vs 0.837) ***	-0.000 (0.000 vs 0.000) ***	-0.000 (0.000 vs 0.000) *
synthstrip	-1.639 (0.029 vs 1.696) ***	-0.703 (0.000 vs 0.703) ***	-1.205 (0.056 vs 1.290) ***	-3.458 (0.001 vs 3.458) ***	-0.592 (0.000 vs 0.592) ***	0.000 (0.000 vs 0.000) **

Figure 3: Pairwise median differences in Dice (top) and Hausdorff Distance 95th Percentile (bottom) between the best performing PyDeface and other methods on raw subject inputs across the full brain and distinct anatomical regions.

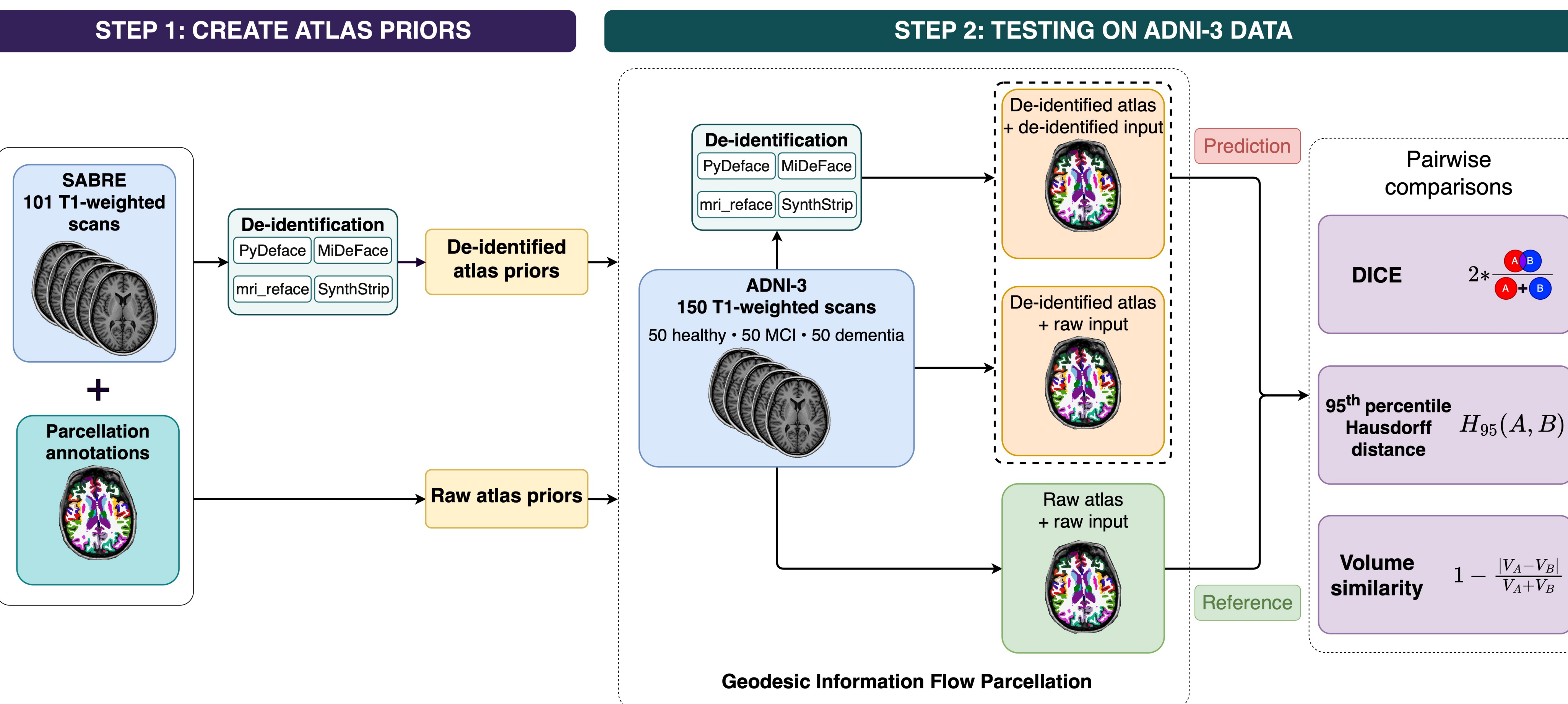


Figure 2: Experiment Pipeline



Contact Details:
Stylianos Charalampous
rmapss1@ucl.ac.uk