Supplemental Information

Table I: **Method Comparison:** Mean (standard deviation) of the Dice Similarity Coefficient (DSC), Hausdorff Distance (HD), and mean surface distance (MSD) for the left and right thalamus validated against Rater #1 and Rater #2. Compared are FreeSurfer (Fischl et al., 2002), FSL FIRST (Patenaude et al., 2011), Non-Local Spatial STAPLE (Asman and Landman, 2013) using either the Neuromorphometrics atlas (NMM-NLSS) or our own thalamus atlas set (MA-NLSS). The $*$ and $\dagger$ indicate a statistically significant difference between the algorithm and RAFTS with $p < 0.05$ and $p < 0.01$, respectively.

| | | Ground Truth Rater #1 | | Ground Truth Rater #2 | |
		Left	Right	Left	Right
DSC	FreeSurfer	$0.799\,(\pm0.057)^{\dagger}$	$0.823\,(\pm0.024)^{\dagger}$	$0.747\,(\pm0.063)^{\dagger}$	$0.755\,(\pm0.047)^{\dagger}$
	FSL FIRST	$0.821\,(\pm0.031)^{*}$	$0.843\,(\pm0.027)^{\dagger}$	$0.776\,(\pm0.038)^{*}$	$0.785\,(\pm0.034)^{\dagger}$
	NMM-NLSS	$0.795\,(\pm0.028)^{\dagger}$	$0.797\,(\pm0.035)^{\dagger}$	$0.736\,(\pm0.049)^{\dagger}$	$0.739\,(\pm0.062)^{\dagger}$
	MA-NLSS	$0.859\,(\pm0.057)^{*}$	$0.874\,(\pm0.016)^{\dagger}$	$0.787\,(\pm0.058)^{\dagger}$	$0.811\,(\pm0.036)^{*}$
	RAFTS w/o Prior	$0.842\,(\pm0.028)^{\dagger}$	$0.816\,(\pm0.047)^{\dagger}$	$0.794\,(\pm0.031)^{*}$	$0.730\,(\pm0.069)^{\dagger}$
	RAFTS w/o DTI	$0.841\,(\pm0.041)^{*}$	$0.825\,(\pm0.046)^{\dagger}$	$0.775\,(\pm0.026)^{\dagger}$	$0.740\,(\pm0.064)^{\dagger}$
	RAFTS	**0.882** (±0.033)	**0.890** (±0.013)	**0.818** (±0.052)	**0.827** (±0.041)
HD (voxels)	FreeSurfer	$10.79\,(\pm2.63)^{\dagger}$	$10.50\,(\pm1.75)^{\dagger}$	$10.63\,(\pm2.14)^{\dagger}$	$10.85\,(\pm1.80)^{\dagger}$
	FSL FIRST	$10.94\,(\pm1.22)^{\dagger}$	$11.62\,(\pm1.75)^{\dagger}$	$11.01\,(\pm0.85)^{\dagger}$	$11.39\,(\pm0.84)^{\dagger}$
	NMM-NLSS	$10.75\,(\pm3.44)^{\dagger}$	$10.58\,(\pm1.79)^{\dagger}$	$10.45\,(\pm3.07)^{\dagger}$	$10.28\,(\pm1.41)^{\dagger}$
	MA-NLSS	$5.61\,(\pm1.85)^{\dagger}$	$6.43\,(\pm2.58)^{*}$	$10.79\,(\pm2.63)^{\dagger}$	$10.50\,(\pm1.75)^{\dagger}$
	RAFTS w/o Prior	$5.26\,(\pm1.46)^{\dagger}$	$5.94\,(\pm1.55)^{\dagger}$	$7.73\,(\pm0.83)$	$8.48\,(\pm2.11)^{\dagger}$
	RAFTS w/o DTI	$5.05\,(\pm1.20)^{*}$	$4.93\,(\pm0.78)^{*}$	$7.83\,(\pm1.27)$	$7.30\,(\pm2.19)$
	RAFTS	**4.05** (±1.07)	**4.23** (±0.79)	**7.07** (±1.24)	**7.04** (±1.75)
MSD (voxels)	FreeSurfer	$1.410\,(\pm0.300)^{\dagger}$	$1.296\,(\pm0.111)^{\dagger}$	$1.729\,(\pm0.333)^{\dagger}$	$1.722\,(\pm0.272)^{\dagger}$
	FSL FIRST	$1.456\,(\pm0.201)^{\dagger}$	$1.333\,(\pm0.158)^{\dagger}$	$1.769\,(\pm0.193)^{\dagger}$	$1.751\,(\pm0.215)^{\dagger}$
	NMM-NLSS	$1.578\,(\pm0.166)^{\dagger}$	$1.575\,(\pm0.169)^{\dagger}$	$1.986\,(\pm0.221)^{\dagger}$	$1.996\,(\pm0.355)^{\dagger}$
	MA-NLSS	$0.895\,(\pm0.289)$	$0.856\,(\pm0.061)^{\dagger}$	$1.322\,(\pm0.267)$	$1.248\,(\pm0.142)$
	RAFTS w/o Prior	$1.079\,(\pm0.183)^{\dagger}$	$1.294\,(\pm0.328)^{\dagger}$	$1.378\,(\pm0.156)$	$1.840\,(\pm0.382)^{\dagger}$
	RAFTS w/o DTI	$1.101\,(\pm0.294)^{*}$	$1.236\,(\pm0.336)^{\dagger}$	$1.526\,(\pm0.162)^{*}$	$1.800\,(\pm0.376)^{\dagger}$
	RAFTS	**0.829** (±0.174)	**0.802** (±0.054)	**1.236** (±0.259)	**1.219** (±0.209)