Automated Hippocampal Segmentation using new standardized manual segmentations from the Harmonized Hippocampal Protocol

Anker, Cecilie; Pai, Akshay ; Sørensen, Lauge; Lyksborg, Mark; Larsen, Rasmus; Conradsen, Knut; Nielsen, Mads

Publication date:
2014

Automated Hippocampal Segmentation using new standardized manual segmentations from the Harmonized Hippocampal Protocol

Authors: Cecilie Anker, Akshay Pai, Lauge Sørensen, Mark Lyksborg, Knut Conradsen, Rasmus Larsen, Mads Nielsen.

Background: Hippocampal volumetry derived from structural magnetic resonance imaging (MRI) has been endorsed by the Alzheimer's disease (AD) diagnostic guidelines as a radiological marker of disease progression. Among the top performing automated hippocampal segmentation methods are multi-atlas segmentation methods, which rely on manual annotations. In this study, we investigate a combination of such method with annotations from a new Harmonized Hippocampal Protocol (HHP). We compare its capabilities to a FreeSurfer method and verify its impact on segmentation and diagnostic group separation capabilities.

Methods: 40 manual HHP hippocampal annotations (12 normal control (NC), 11 mild cognitive impairment (MCI), 17 AD) were transformed to a common segmentation space. The corresponding 1.5T MRIs were preprocessed using FreeSurfer. An automated Non-Local Patch-based segmentation technique (N-L Patch) was used to segment the left- and right hippocampus, separately. All 40 HHP annotations were used as atlases during pre-selection, but only the 9 most similar contributed to the final segmentation. Leave-one-out cross-validation was performed on the 40 atlases, and the corresponding DICE-scores with the manual annotations were calculated.

A standardized ADNI dataset containing 1.5T MRIs from 504 subjects (169 NC, 234 MCI, 101 AD) at baseline and month 12 was segmented using the method described above and atrophy rate calculated as percentage volume change was estimated.

Results: Mean ($\pm$sd) cross-validation DICE-scores of the 40 atlases segmented using N-L Patch and cross-sectional FreeSurfer were 0.868 ($\pm$0.019) and 0.781 ($\pm$0.031), respectively. A paired t-test between N-L Patch and FreeSurfer DICE-scores showed significance ($p<0.001$).

Statistics in terms of AUC and Cohens' D were used to evaluate differences in atrophy rates between diagnostic groups for N-L Patch and FreeSurfer segmentations. N-L Patch performed significantly better in separating AD from NC and AD from MCI, Table 1.

Conclusions: Including the HHP labels in a multi-atlas segmentation method resulted in better segmentation consensuses with the new hippocampal label standard than a state-of-the-art method, FreeSurfer. Furthermore, N-L Patch yielded significantly better group separation than FreeSurfer in separating AD from NC and AD from MCI. This illustrates the longitudinal robustness of segmentations when annotations from the new hippocampal label standard are included in automated segmentation methods.

	AD	**MCI**	**NC**
	mean (std)	mean (std)	mean (std)
N-L Patch	-4.23 (3.07)	-2.38 (3.28)	-0.86 (2.46)
FreeSurfer	-4.29 (5.32)	-3.69 (5.48)	-1.39 (5.41)
	AD vs. NC	**AD vs. MCI**	**MCI vs. CN**
	AUC	AUC	AUC
N-L Patch	0.80*	0.66**	0.65
FreeSurfer	0.69	0.53	0.67
	AD vs. NC	**AD vs. MCI**	**MCI vs. CN**
	Cohens'D	Cohens'D	Cohens'D
N-L Patch	1.21**	0.58*	0.53
FreeSurfer	0.54	0.11	0.42

Table 1: Statistics based on atrophy (%) of 504 subjects between baseline and month 12. N-L Patch: Non-local Patch-based segmentation with HHP atlases, FreeSurfer: cross-sectional FreeSurfer. DeLong test is done to compare AUC's between methods and bootstrap test is done to compare Cohens' D's between methods. * indicates significance of N-L Patch with p-value <0.01 and $>=0.001$; and ** indicates significance of N-L Patch with p-value <0.001.