Fully automatic deep-learning based segmentation of Hippocampal subfields from standard resolution T1W MRI in Alzheimer's Disease

Joules R[1], Wolz R[12]

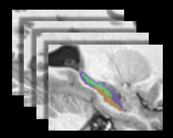
[1]IXICO plc, London UK, [2]Department of Computing, Imperial College, London, UK

This work demonstrates hippocampal subfield segmentations from standard resolution T1 MRI can provide clinically informative measures of volume beyond those available with whole hippocampal segmentations.

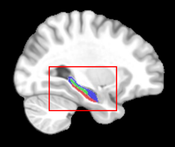
Hippocampal subfield volumes are affected at different points in the Alzheimer's disease Trajectory [1] and can provide improved sensitivity to disease progression and intervention efficacy compared to whole hippocampal volume alone. Typically, subfield volume estimation requires a high resolution T2 MRI in addition to standard resolution T1 to ensure accurate delineation.

Here we train an AI method to segment the hippocampus into larger subfields (CA1-3, CA4+DG, and Subiculum) from standard resolution T1W MRI alone to assess utility as biomarkers compared to whole hippocampal volume, in the absence of a high resolution T2 MRI.

Model Training



We employed a training set of 25 T1W MRI at 1mm^3 voxel size [2] with manual segmentations of Hippocampus delineated into 3 Subregions:
- Cornu Ammonis 1-3 (CA),
- Cornu Ammonis 4 +Dentate gyrus (DG)
- Subiculum (SUB)

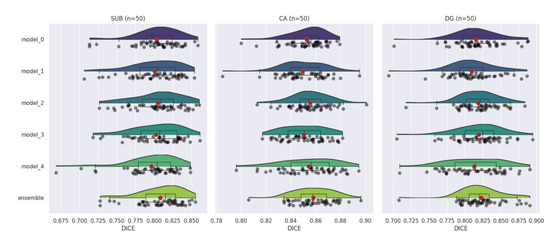


All training images were registered to an MNI template, and a mean atlas generated for each ROI, permitting estimation of bounding box dimensions and placement during inference.

To mitigate learnt laterality biases, and to extend the limited training set, models were trained on a single laterality per ROI, with right laterality segmentations and anatomical images reflected to appear on the left, resulting in two training samples per ROI subject.

The dataset was employed in training a nn-Unet [3] model with a generalised Dice focal loss. Within a 5-fold cross validation loop, data was split into a training + validation (0.75 ratio) and test sets (20%). For each cross-validation fold, 5 models were trained with different validation / train data splits and augmenter initialisation. Output segmentations were generated from the median softmax of the five trained models, mitigating the impact of poor performance from a single model.

For each cross-validation fold, an ROI segmentation was obtained from individual trained models and the median softmax output. Employing the median provided a minor improvement in DICE scores, and notably mitigated the poor performance of any single model.



Segmentation	DICE (n=50)	SUB	CA	DG
Individual models (n=5)	mean	0.802	0.853	0.815
	Std	0.033	0.018	0.33
Median softmax	mean	**0.808**	**0.858**	**0.822**
	std	0.031	0.016	0.030

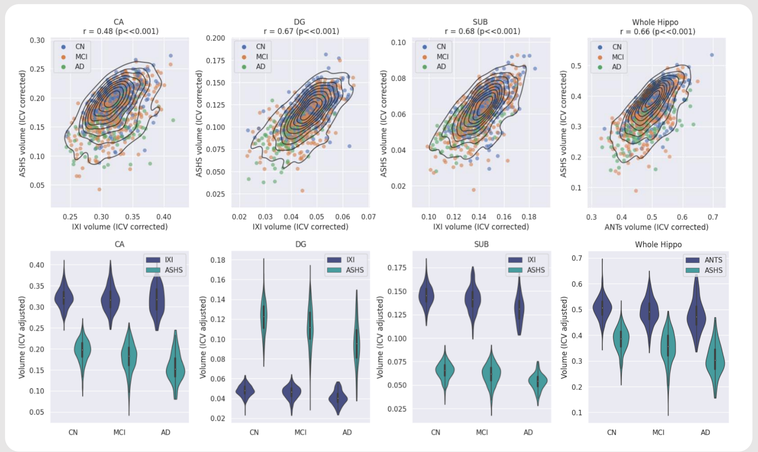
Evaluation was performed on **1055 visits identified from the ADNI database** with QC passing hippocampal subfield volumes, as computed by the University of Pennsylvania using the ASHS method [4] applied to high-resolution T2W MRI and standard resolution T1 MRI. Small region volumes, as provided generated by ASHS, were grouped into 3 larger hippocampal subfields labels: the **Cornu Ammonis (CA), Subiculum (SUB) and Dentate Gyrus (DG).**

The standard resolution (1mm iso) ADNI T1W MR images were processed with the proposed hippocampal subfield segmentation models (IXI) and compared to the available ASHS volumes. T1 images were also processed with ANTs [5] to obtain independent estimates of whole hippocampal volumes.

Evaluation Dataset			
Subjects		632	
Visits		1055	
Diagnosis at baseline*	**CN**	**MCI**	**AD**
# Subjects	298	263	71
# Male	123 (41%)	152 (58%)	38 (54%)
Mean Age at baseline	74.7	74.8	76.2
# visits	477	469	125

*Baseline defined as the earliest of the available scans with ASHS analysis

Cross-sectional volumes, normalised by intracranial volume, were significantly correlated between the proposed method (IXI) and ASHS and show expected trends between arms.



Diagnostic groups were compared using normalised subfield volume with logistic regressions, correcting for age and sex. The SUB and DG volumes reported greater group separation than the whole Hippocampus when using only T1W MRI.

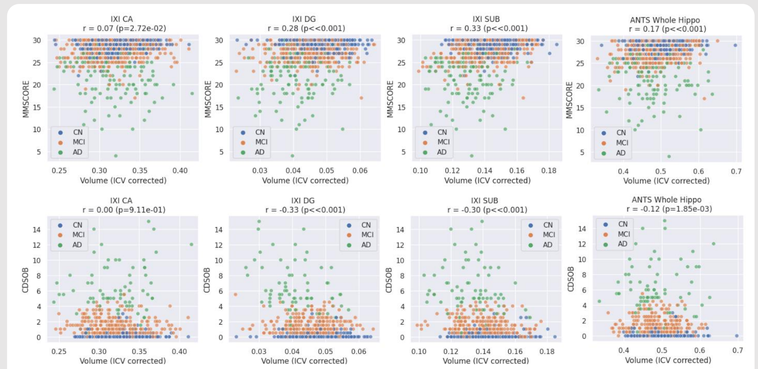
Comparison (n)	Method	z- score CA	SUB	DG	HIP
CN – MCI (477 – 469)	ASHS	**5.49***	**4.14***	**5.60***	**5.56***
	Proposed	1.13	**3.08***	**3.45***	-
	ANTs	-	-	-	2.73
CN – AD (477 – 125)	ASHS	**9.53***	**9.20***	**9.80***	**9.92***
	Proposed	**2.82***	**8.40***	**9.13***	-
	ANTs	-	-	-	**5.33***
MCI – AD (469 – 125)	ASHS	**6.26***	**6.63***	**7.01***	**6.81***
	Proposed	**1.88***	**6.24***	**6.96***	-
	ANTs	-	-	-	**3.32***

* Significant after Bonferroni correction (α 0.05)

The method presented here enables the extraction of hippocampal subfield volumes from routinely acquired data, providing additional biomarkers that may not be possible for on-going collection where patient burden prohibits collection of an additional high-resolution T2W MRI or for retrospective data without a high-resolution T2.

In the absence of a high-resolution T2W MRI, hippocampal subfields volumes, as estimated from T1W MRI, can provide an informative measure of disease progression with greater than the whole hippocampus alone. The subiculum and dentate gyrus reported the greatest sensitivity to group separation and correlation to cognitive scores, with the subiculum notably reporting sensitivity comparable to that ASHS. Further development is required with validation on additional datasets.

Significant correlations, correcting for age and sex, were observed between **cognitive scores (MMSE and Sum of Boxes) and subfield volumes** for both methods. For the proposed method (IXI), both the DG and SUB region reported a significant relationship to cognitive scores, greater than observed in the whole hippocampus alone; with the SUB correlation comparable to ASHS.

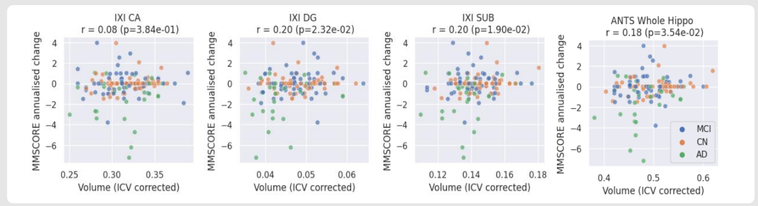


(Pearson r values in graph computed without partialing out age and sex)

Region	Method	MMSE (n=1065) r	MMSE (n=1065) CI 95%	CD – SOB (n=706) r	CD – SOB (n=706) CI 95%
CA	ASHS	**0.370***	[0.32, 0.42]	**-0.368***	[-0.43, -0.3]
	Proposed	0.073	[0.01, 0.13]	0.003	[-0.08, 0.07]
DG	ASHS	**0.393***	[0.34, 0.44]	**-0.411***	[-0.47, -0.35]
	Proposed	**0.280***	[0.22, 0.34]	**-0.325***	[-0.39, -0.26]
SUB	ASHS	**0.341***	[0.29, 0.39]	**-0.315***	[-0.38, -0.24]
	Proposed	**0.330***	[0.27, 0.38]	**-0.296***	[-0.36, -0.22]
Hipp	ASHS	**0.400***	[0.35, 0.45]	**-0.410***	[-0.47, -0.35]
	ANTS	**0.172***	[0.11, 0.23]	-0.111	[-0.18, -0.03]

* Significant after Bonferroni correction (α 0.05)

Trend correlations were noted between T1 derived baseline hippocampal volumes and the annualized MMSE score change within a 2-year period for amyloid positive, and MMSE > 21 subjects (n=137). This suggests that hippocampal subfield volumes do not offer an advantage over whole hippocampal volumes as an enrichment / stratification tool to identify stable from declining subjects .



[1] Hett et al; *Multimodal Hippocampal Subfield Grading For Alzheimer's Disease Classification*; Scientific Reports; 2019; *9*(1), p.13845.
[2] www.nitrc.org/projects/mni-hisub25
[3] Isensee et al; *nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation*; Nature Methods; 2021; *18*(2), pp.203-211
[4] Yushkevich et al; *Automated Volumetry and Regional Thickness Analysis of Hippocampal Subfields and Medial Temporal Cortical Structures in Mild Cognitive Impairment*; Human Brain Mapping; 2014; *36*(1), pp.258-287.
[5] Tustison et al; *ANTsX neuroimaging-derived structural phenotypes of UK Biobank*; medRxiv; 2023; pp.2023-01