



Can SPHARM-Based Features from Automated or Manually Segmented Hippocampi Distinguish Between MCI and TLE?

Michael Liedlgruber¹, Kevin Butz^{2,3}, Yvonne Höller², Georgi Kuchukhidze², Alexandra Taylor², Aljoscha Thomschevski^{2,3}, Ottavio Tomasi⁴, Eugen Trinka^{2,3}, and Andreas Uhl^{1(✉)}

¹ Department of Computer Sciences, University of Salzburg, Salzburg, Austria
uhl@cosy.sbg.ac.at

² Department of Neurology,
Christian Doppler Medical Centre and Centre for Cognitive Neuroscience,
Paracelsus Medical University, Salzburg, Austria

³ Spinal Cord Injury and Tissue Regeneration Center Salzburg, Salzburg, Austria

⁴ Department of Neurosurgery, Paracelsus Medical University, Salzburg, Austria

Abstract. Spherical Harmonics (SPHARM), when computed from hippocampus segmentations, have been shown to be useful features for discriminating patients with mild cognitive impairment (MCI) from healthy controls. In this paper we use this approach to discriminate patients with temporal lobe epilepsy (TLE) from healthy controls and for the first time, we aim to discriminate TLE patients from MCI ones. When doing so, we assess the impact of (i) using three different automated hippocampus segmentation techniques and (ii) three human raters with different qualification providing manual segmentation labels. We find that (only a fusion of) the considered automated segmentation tools deliver segmentation data which can finally be used to discriminate TLE from MCI, but for discriminating TLE from healthy controls automated techniques do not help. Further, the qualification of human segmenters has a decisive impact on the outcome of subsequent SPHARM-based classification, especially for distinguishing TLE from healthy controls (which is obviously the more difficult task).

Keywords: TLE · MCI · Spherical Harmonics · Hippocampus segmentation

1 Introduction

Mild cognitive impairment (MCI) is a condition of cognitive deterioration that is difficult to classify as normal aging or as a prodromal stage to dementia. Neuropsychological tests alone are highly valuable but not sufficient to determine MCI or early stages thereof, since they are not sensitive enough for patients with subjective complaints and no significant and clinically detectable deficits [33].

Further, the diagnosis of temporal lobe epilepsy (TLE) was, and still currently is, based on clinical assessment and electro-encephalographic (EEG) examination, which is sometimes inconclusive [31].

However, both diseases, i.e., MCI as well as TLE, need to be treated and handled adequately, in order to prevent massive memory decline or risks due to seizures [7]. MCI and TLE occur also as a co-morbid disease, when patients with MCI encounter seizures; moreover, the presentation of cognitive disorders in the elderly warrants a differential diagnosis of MCI or TLE, since seizures can mimic confusional episodes of MCI or Alzheimer's Disease without visible seizure activity with scalp electroencephalography [23]. Treating these seizures can restore cognitive functioning of these patients. However, diagnostics based on invasive electroencephalography bear significant risk and costs.

From a structural point of view, the hippocampus is an area of the brain that links MCI and TLE as it has been found that both diseases affect the hippocampal structure in some way [14]. It is therefore worth evaluating techniques for the diagnosis of these conditions that are based on distinctive features of this brain structure. Segmentation of the hippocampi is, of course, a prerequisite for such approaches. The hippocampus is atrophic in mild cognitive impairment and dementia [11], and it is sclerotic in specific subtypes of epilepsy [28], specifically in TLE. Thus volumetry alone is not sufficient for a discrimination which calls for shape-based features to be investigated.

For structural characterisation, the amount of time it takes an expert to segment the hippocampus is a significant obstacle. In a high-resolution magnetic resonance image, a specialist has to trace the contour of the formation in each slice and review the result in several dimensions. Even after a certain training, this might still take about up to one hour per hippocampus - i.e. two hours per patient. This motivates the use of automated segmentation techniques. A large variety of techniques and algorithms for automated hippocampus segmentation have been published over the last years, some of them targeted to specific disease or deformation classes (see e.g. [2, 10, 17, 20, 29, 30, 34, 35, 39, 40]). The classical state-of-the-art algorithms for automated hippocampus segmentation [5, 24] are based on multi-atlas segmentation (MAS [15]), with recent upcoming deep learning-based techniques (e.g. [36] – see also [1] for a review on deep learning-based brain MRI segmentation techniques).

Interestingly, in a recent large-scale study on algorithms for computer-aided diagnosis of dementia based on structural MRI [4], 6 out of 15 considered techniques (including the best performing algorithms) still relied on FreeSurfer segmentations (see Sect. 3) which is the majority of algorithms based on segmentation in this comparison. This massive employment of FreeSurfer in large-scale studies, although not any more being considered as state-of-the-art, underpins the need for publicly available and easy-to-use segmentation tools. This also holds true for the upcoming deep learning-based tools. Recent work [26] also demonstrates that employing publicly available and cost free segmentation tools were not able to reproduce MCI vs. control group classification results relying on a custom (private) hippocampus segmentation tool [13].

In this paper, we closely follow a shape-based approach originally used to distinguish hippocampi affected by MCI from those of a healthy control group by employing spherical harmonics coefficients (SPHARM) as potentially discriminating features [13, 26, 32]. However, as major original contribution, we investigate if this approach is suited to differentiate hippocampi affected by TLE from (i) those affected by MCI (which has never been investigated at all with any technique) and from (ii) those of a healthy control group, respectively (which has been done relying on manual segmentations only [9, 18, 19, 21]). In this context, we investigate the impact of using different hippocampus segmentation approaches: Three cost-free and pre-compiled out-of-the-box hippocampus segmentation software packages as well as three segmentations independently conducted by human raters with different qualifications (the availability of which can be considered an extremely rare asset). This work discriminates itself from our earlier work [27] by not relying on the two hippocampi separately, thus not looking into lateralisation effects, leading to larger datasets and thus increased statistical result significance.

Section 2 briefly explains how we obtain SPHARM coefficients used to compose feature vectors subject to subsequent classification. In Sect. 3, we first explain the experimental setup in detail, specifically including the hippocampus segmentation variants and SPHARM coefficient selection strategy employed. Subsequently, classification results are shown and described. Section 4 concludes the paper by discussing the observed results.

2 Spherical Harmonics Descriptors in Structural MCI Characterisation

The features used for classification are based on Spherical Harmonics (SPHARM). These are a series of functions which are used to represent functions defined on the surface of a sphere. Once a 3D object has been mapped onto a unit sphere, it is also possible to describe that object in terms of coefficients for the basis function of SPHARM. In other words, the SPHARM coefficients can be used as shape descriptors. In this work we follow the approach described in [3] in order to obtain coefficients for the hippocampi voxel volumes.

Once a voxel volume for a hippocampus has been obtained, either by automatic or manual segmentation, we first fix the topology of the voxel objects. This is necessary since, in order to map a 3D object to a sphere, the respective voxel object must exhibit a spherical topology. Moreover, the voxel dimensions are adjusted to obtain isotropic voxel. Since the voxel volumes resulting from MRI scans yield voxels with a size depending on the scan parameters (i.e. often non-isotropic), we resample the data such that we end up with voxel cubes each having a side length of 1 mm. Sometimes the segmented data consists of one large, and two or more disconnected smaller voxel compounds, respectively. In such a case we determine all voxel compounds and remove all but the largest one. This step removes small spurious voxel masses which occur quite frequently

especially in manual segmentation and hinder a mapping of the voxel object to a sphere.

Based on the resampled and fixed voxel volumes, we generate 3D objects. While other implementations create objects based on triangular faces, we decided to use quadrilaterals since these more naturally correspond to voxels (see Fig. 1(a)). The 3D objects are then mapped onto a unit sphere during the initial parameterisation, which is followed by a constrained optimisation (described in more detail in [3]). The optimised parameterisation is then used to compute the SPHARM coefficients (see Fig. 1(b) and (c)).

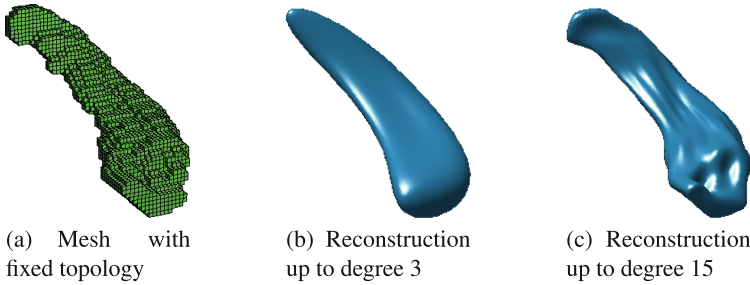


Fig. 1. Basic principle visualised.

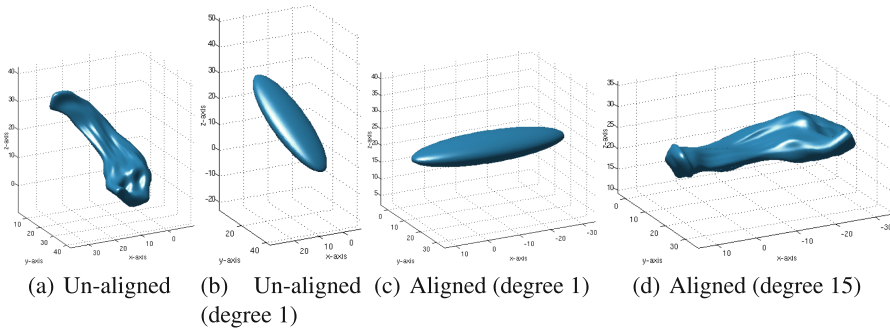


Fig. 2. The process of re-aligning the hippocampus. (a) original orientation (SPHARM reconstruction up to degree 15), (b) the same reconstruction but only to degree 1, (c) same as (b) but re-aligned to axes, and (d) the final object, same as (a) but re-aligned to axes.

We are mainly interested in shape differences and thus we want to ignore orientation differences which are found e.g. in malrotated hippocampi. Instead of using e.g. rotation invariant SPHARM representations [16] we resort to a classical alignment procedure, as invariant representation sometimes suffer from a lower extent of discriminativeness. For alignment, we compute a reconstruction of the hippocampus object up to SPHARM degree 1 (based on a triangulated sphere).

This results in an ellipsoid which is aligned with the main orientation of the 3D object. Using PCA we determine the principal axes of the ellipsoid and rotate the object such that all hippocampal volumes are always in the coordinate system of the principal axes (i.e. co-aligned). After the re-alignment we recompute the SPHARM coefficients up to degree 15 for each re-aligned object and use them for the subsequent classification process. Figure 2 shows the process of re-aligning a hippocampal 3D object.

Computing SPHARM coefficients up to degree J , we obtain a total of N complex coefficients per hippocampus, where N is computed as

$$N = 3 * (J + 1)^2. \quad (1)$$

Extracting SPHARM coefficients up to degree 15 we end up with 768 coefficients per hippocampus. We have used a custom MATLAB implementation following the *SPHARM-PDM* code (<https://www.nitrc.org/projects/spharm-pdm>).

The final feature vectors F_i available for the feature selection process are composed of the absolute coefficient values for the left and right hippocampi:

$$F_i = (|C_{i,1}^l|, \dots, |C_{i,N}^l|, |C_{i,1}^r|, \dots, |C_{i,N}^r|), \quad (2)$$

where $C_{i,n}^l$ and $C_{i,n}^r$ denote the n -th coefficient for the left and right hippocampus of subject i , respectively. These feature vectors contain 1536 coefficients in total out of which subsets can be selected for actual classification.

3 Experiments

3.1 Experimental Settings

Data. In this work we use 58 T1-weighted MRI volumes, a data set that has been acquired at the Department of Neurology, Paracelsus Medical University Salzburg, including patients with mild cognitive impairment (MCI, 20 subjects), temporal lobe epilepsy (TLE, 17 subjects), and a healthy control group (CG, 21 subjects). These data are a subset of a larger study [25]. We defined patients with amnesic MCI according to level three of the global deterioration scale for aging and dementia described by [12]. Diagnosis/ground truth w.r.t. MCI and TLE was based on multimodal neurological assessment, including imaging (high resolution 3 T magnetic resonance tomography, and single photon emission computed tomography with Hexamethylpropylenaminooxim), electroencephalography, and neuropsychological testing.

Hippocampus Segmentations. Manual segmentations have been performed by 3 experienced raters (one senior neurosurgeon – Rater1 – and two junior neuroscientists supervised by a senior neuroradiologist – Rater2 & Rater3) on a Wacom Cintiq 22HD graphic tablet device (resolution 1920×1200) using a DTK-2200 pen and employing the 32-bit 3DSlicer software for Windows (v. 4.2.2-1 r21513) to delineate hippocampus voxels for each slice separately. The

raters independently used consensus on anatomical landmarks/boarders of the hippocampus based on Henry Duvernoy’s hippocampal anatomy [22]. The procedure used was to depict the hippocampal outline in the view of all planes in the following order: sagittal – coronal – axial with subsequent cross line control through all planes.

For automated hippocampus segmentation, in contrast to most of the algorithms presented in literature, e.g. [40], all three employed hippocampus segmentation software packages are already pre-compiled and available for free [25]:

*FreeSurfer(FS)*¹ is a popular set of tools which allow an automated labelling of subcortical structures in the brain [10]. Such a subcortical labelling is obtained by using the volume-based stream which consists of five stages [10]. The result is a label volume, containing labels for various different subcortical structures (e.g. hippocampus, amygdala, and cerebellum). FreeSurfer is a highly popular tool in hippocampal analysis to assess clinical hypotheses [6, 17, 20, 35, 38] or to compare to newly proposed segmentation techniques (e.g. [29, 30, 40]). The winning algorithm in a recent large-scale study on algorithms for computer-aided diagnosis of dementia based on structural MRI [4] was based on FS segmentation, as well as 5 other out of 15 considered techniques in this work.

AHEAD (Automatic Hippocampal Estimator using Atlas-based Delineation²) is specifically targeted at an automated segmentation of hippocampi [34] and employs multiple atlases and statistical learning method. After an initial rigid registration step, a deformable registration is carried out using the Symmetric Normalisation algorithm. From the result of these steps, the volume is normalised to the atlas. The hippocampus segmentation from the atlas is then warped back to the input volume. Based on multiple atlases and a statistical learning method, the final segmentation is obtained.

Although *BrainParser(BP)*³ is usually able to label various different subcortical structures, we use a version of BrainParser which is specifically tailored to hippocampus segmentation. After re-orienting the input volume to the coordinate system of the included, pre-trained atlas, skull stripping is performed. This is followed by computing an affine transform between the input volume and the reference brain volume. Then a deformable registration between the input and the reference volume is carried out. Then, according to the trained atlas, the input volume is labelled.

We also fuse the segmentation results using voxel-based majority voting (abbreviated M.V. in the results – a voxel is active in the fused volume if at least two raters or segmentation tools marked that voxel as belonging to a hippocampus) and STAPLE [37] (abbreviated as STA). Since for human raters there is hardly a difference between majority voting and STAPLE, the latter results are not shown.

¹ v. 51.0, available at <http://surfer.nmr.mgh.harvard.edu>.

² v. 1.0, available at <http://www.nitrc.org/projects/>.

³ available at <http://www.nitrc.org/projects/>.

Feature Selection, Classification, and Evaluation Protocol. Features used for actual classification are selected from the feature vectors F_i according to the degree J of their coefficients. The strategy “Cumula J ” selects all coefficients with degree $\leq J$. Thus, for example, according to Eq. 1 for $J = 5$, Cumula5 employs 108 coefficients of degrees $J = 1 \dots 5$.

For the classification of the features we use the Support Vector Machines (SVM) classifier [8] with a linear kernel. The choice for this classifier has been made since the classifier is known to be able to cope very well with high-dimensional features. However, this classifier does not guarantee that a large feature set containing a smaller one exhibits better classification when trained with the larger feature set compared to the smaller one.

To come up with classification accuracy estimation, we apply leave-one-out-cross-validation (LOOVC) for the feature vectors.

3.2 Experimental Results

Tables 1 and 2 display the overall classification accuracy in percent for our test data set. Results are shown up to $J = 7$. Depending on the underlying segmentations, for higher values of J , classification accuracy either decreases (for the automated techniques and low quality human segmentations) or does not increase any more (for high quality human segmentations).

The first impression of the results to discriminate TLE from CG (Table 1) is that there are no bold numbers in the left half of the table, i.e. the employment of automated segmentation tools does not lead to any decent classification results using the considered SPHARM approach and most results are hardly superior to random guessing.

Table 1. Classification results: temporal lobe epilepsy (TLE) vs. healthy control group (CG) (overall accuracy $\geq 75\%$ in bold).

	Automated segmentation					Human segmentation			
	STA	M.V	AHEAD	BP	FS	M.V	Rater1	Rater2	Rater3
Cumula1	55.6	55.3	67.5	56.4	56.4	53.9	66.7	43.6	51.3
Cumula2	50.0	31.6	47.5	56.4	43.6	59.0	79.5	66.7	64.1
Cumula3	50.0	52.6	52.5	46.2	38.5	64.1	76.9	61.5	66.7
Cumula4	47.2	36.8	62.5	66.7	41.0	66.7	79.5	46.2	51.3
Cumula5	50.0	42.1	42.5	64.1	48.7	64.1	79.5	64.1	53.9
Cumula6	61.1	60.5	42.5	69.2	43.6	74.4	79.5	43.6	76.9
Cumula7	47.2	57.9	47.5	71.8	43.6	69.2	82.1	53.9	74.4

Looking at the results of applying individual automated segmentation tools, BP is clearly the best technique with increasing classification rates for increasing J (which indicates that also coefficients representing finer detail are still useful),

while FS and AHEAD based results are on a comparable level and worse than BP. Applying Majority Voting (M.V.) or STAPLE (STA) to the individual segmentations does not really help, indicating that these are too different to benefit from fusion strategies.

The situation is different for the results when basing SPHARM classification on human rater segmentations. There is a clear trend that Rater1, being the most qualified human rater, provides segmentations that lead to sensible classification results (increasing up to 82% for $J = 7$). Rater3 achieves some useful results as well, but lower ($\leq 77\%$) and for $J = 6, 7$ only, while Rater2 stays below 67%. Overall, the differences among the human raters in terms of achieved classification accuracy are considerable and applying fusion techniques does not lead to useful results (i.e. better than Rater2 and Rater3 in most settings, but clearly inferior to the Rater1 results).

Table 2. Classification results: temporal lobe epilepsy (TLE) vs. mild cognitive impairment (MCI) (overall accuracy $\geq 75\%$ in bold).

	Automated segmentation					Human segmentation			
	STA	M.V	AHEAD	BP	FS	M.V	Rater1	Rater2	Rater3
Cumula1	80.6	84.2	73.7	57.9	52.6	75.7	56.8	70.3	62.2
Cumula2	75.0	71.1	68.4	63.2	65.8	67.6	81.1	70.3	81.1
Cumula3	75.0	60.5	71.1	42.1	60.5	56.8	81.1	81.1	73.0
Cumula4	72.2	79.0	65.8	50.0	60.5	59.5	83.8	70.3	67.6
Cumula5	63.9	57.9	73.7	47.4	65.8	62.2	81.1	75.7	62.2
Cumula6	55.6	68.4	68.4	42.1	65.8	62.2	86.5	81.1	62.2
Cumula7	52.8	65.8	71.1	44.7	63.2	56.8	83.8	78.4	70.3

Considering the results to discriminate TLE affected hippocampi from those affected by MCI (see Table 2), we see a different picture. Automated segmentation tools at least deliver useful segmentations for classification when fusion techniques are applied (for small values of J only, again indicating that only coarser details contribute useful information for classification). Also the relation among the individual tools is different, with the clear ranking as AHEAD best, followed by FS, and BP worst (note that this is almost inverse to the classification task before).

Concerning manual segmentations, we notice a somewhat similar behaviour as before. Again, Rater1 (with the highest qualification) is the best, while for this classification task, segmentations of Rater2 also lead to quite decent classification results for higher values of J . Except for $J = 2$, Rater3 segmentations are the worst. Again, fusion of human segmentation results does not reach the classification results of the individual segmentations, except for $J = 1$ (but on a moderate level).

4 Discussion and Conclusion

The obtained results indicate that discriminating TLE patients from healthy controls is far more difficult than discriminating TLE patients from MCI patients. In the former case, only human segmentation (and only that provided by the most qualified human rater) leads to SPHARM coefficients that can be used to discriminate the two classes with reasonable accuracy. In this case, contrasting to the case of discriminating MCI affected patients from healthy controls [26], also applying fusion among the human segmentations does not work properly (obviously the segmentations are too different to benefit from fusion [25]). For discriminating TLE patients from healthy controls, the considered automated segmentation tools do not lead to sensible classification results under the employed SPHARM-methodology, neither applied individually, nor under fusion techniques (which might also explain the non-existence of corresponding publications).

The situation is different for the second case (i.e. discriminating TLE patients from MCI patients). While employing individual automated tools does not work, fusing the corresponding segmentations and using the lower order SPHARM coefficient for classification is successful. Obviously, the shape difference is so fundamental, that it is reflected in the coarse grain SPHARM details of the automated tools segmentations' in some complementary manner. The best qualified human rater's segmentations again lead to best classification results, but also the two other human raters achieve useful accuracy for some settings. Thus, the fundamental differences are also grasped by the less qualified human raters segmentations. As in the first case, human segmentations seem to be too different to provide useful results under fusion (see also [25] for corresponding segmentation result analysis confirming this assumption).

Acknowledgments. This work has been partially funded by the Austrian Science Fund (FWF) under Project No. KLI12-B00.

References

1. Akkus, Z., Galimzianova, A., Hoogi, A., Rubin, D., Erickson, B.: Deep learning for brain MRI segmentation: state of the art and future directions. *J. Digit. Imaging* **30**(4), 449–459 (2017)
2. Babakchian, S., Chew, N., Green, A., et al.: Automated and manual hippocampal segmentation techniques: a comparison of results and reproducibility. *Neurology* **80**, P06.053 (2013)
3. Brechühler, C., Gerig, G., Kübler, O.: Parametrization of closed surfaces for 3-D shape description. *Comput. Vis. Image Underst.* **61**(2), 154–170 (1995)
4. Bron, E., et al.: Standardized evaluation of algorithms for computer-aided diagnosis of dementia based on structural MRI: the CADDementia challenge. *NeuroImage* **111**, 562–579 (2015)

5. Cardoso, J., Leung, K., et al.: STEPS: similarity and truth estimation for propagated segmentations and its application to hippocampal segmentation and brain parcellation. *Med. Image Anal.* **17**(6), 671–684 (2013)
6. Cherbuin, N., Anstey, K.J., Réglade-Meslin, C., Sachdev, P.S.: In vivo hippocampal measurement and memory: a comparison of manual tracing and automated segmentation in a large community-based sample. *PLoS ONE* **4**, 1–10 (2009)
7. Dodrill, C.: Neuropsychological effects of seizures. *Epilepsy Behav.* **5**(Suppl. 1), 22–24 (2004)
8. Duda, R.O., Hart, P.E., Stork, D.G.: *Pattern Classification*, 2nd edn. Wiley, Hoboken (2000)
9. Esmaeilzadeh, M., Soltanian-Zadeh, H., Jafari-Khouzani, K.: Mesial temporal lobe epilepsy lateralization using SPHARM-based features of hippocampus and SVM. In: *Proceedings of Medical Image 2012, SPIE Proceedings*, vol. 8314 (2012)
10. Fischl, B., et al.: Automatically parcellating the human cerebral cortex. *Cereb. Cortex* **14**(1), 11–22 (2004)
11. Fotuhi, M., Do, D., Jack, C.: Modifiable factors that alter the size of the hippocampus with ageing. *Nat. Rev. Neurol.* **8**, 189–202 (2012)
12. Gauthier, S., et al.: Mild cognitive impairment. *Lancet* **367**(9518), 1262–1270 (2006)
13. Geradin, E., Chetelat, G., et al.: Multidimensional classification of hippocampal shape features discriminates Alzheimer’s disease and mild cognitive impairment from normal aging. *Neuroimage* **47**(4), 1476–1486 (2009)
14. Höller, Y., Trinka, E.: What do temporal lobe epilepsy and progressive mild cognitive impairment have in common? *Front. Syst. Neurosci.* **8**, 58 (2014)
15. Iglesias, J., Sabancu, M.: Multi-atlas segmentation of biomedical images: a survey. *Med. Image Anal.* **24**, 205–219 (2015)
16. Kazhdan, M., Funkhouser, T., Rusinkiewicz, S.: Rotation invariant spherical harmonic representation of 3D shape descriptors. In: *Eurographics/ACM SIGGRAPH Symposium on Geometry Processing*, pp. 156–164, June 2003
17. Kim, H., Chupin, M., Colliot, O., et al.: Automatic hippocampal segmentation in temporal lobe epilepsy: impact of developmental abnormalities. *Neuroimage* **59**, 3178–3186 (2012)
18. Kim, H., Bernhardt, B.C., Kulaga-Yoskovitz, J., Caldaïrou, B., Bernasconi, A., Bernasconi, N.: Multivariate hippocampal subfield analysis of local MRI intensity and volume: application to temporal lobe epilepsy. In: Golland, P., Hata, N., Barillot, C., Hornegger, J., Howe, R. (eds.) *MICCAI 2014. LNCS*, vol. 8674, pp. 170–178. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-10470-6_22
19. Kim, H., Mansi, T., Bernasconi, N.: Disentangling hippocampal shape anomalies in epilepsy. *Front. Neurol.* **4**, 131 (2013)
20. Kim, J., Choi, D., et al.: Evaluation of hippocampal volume based on various inversion time in normal adults by manual tracing and automated segmentation methods. *Investig. Magn. Reson. Imaging* **19**(2), 67–75 (2015)
21. Kohan, Z., Azami, R.: Hippocampus shape analysis for temporal lobe epilepsy detection in magnetic resonance imaging. In: *Proceedings of Medical Image 2016, SPIE Proceedings*, vol. 9788, p. 97882T (2016)
22. Kuzniecky, R., Jackson, G.: *Magnetic Resonance in Epilepsy*. Raven Press, New York (1995)
23. Lam, A.D., Deck, G., et al.: Silent hippocampal seizures and spikes identified by foramen ovale electrodes in Alzheimer’s disease. *Nat. Med.* **23**(6), 678–680 (2017)

24. Leung, K., Barnes, J., et al.: Automated cross-sectional and longitudinal hippocampal volume measurement in mild cognitive impairment and Alzheimer's disease. *NeuroImage* **51**(4), 1345–1359 (2013)
25. Liedlgruber, M., et al.: Pathology-related automated hippocampus segmentation accuracy. In: Maier-Hein, F.K., Deserno, L.T., Handels, H., Tolxdorff, T. (eds.) *Bildverarbeitung für die Medizin 2017. I*, pp. 128–133. Springer, Heidelberg (2017). https://doi.org/10.1007/978-3-662-54345-0_31
26. Uhl, A., et al.: Hippocampus segmentation and SPHARM coefficient selection are decisive for MCI detection. In: Maier, A., Deserno, L.T., Handels, H., Maier-Hein, F.K., Palm, C., Tolxdorff, T. (eds.) *Bildverarbeitung für die Medizin 2018. I*, pp. 239–244. Springer, Heidelberg (2018). https://doi.org/10.1007/978-3-662-56537-7_65
27. Liedlgruber, M., et al.: Lateralisation matters: discrimination of TLE and MCI based on SPHARM description of hippocampal shape. In: *Proceedings of the 31st IEEE International Symposium on Computer-Based Medical Systems (CBMS 2018)*, pp. 129–134, June 2018
28. Malmgren, K., Thom, M.: Hippocampal sclerosis - origins and imaging. *Epilepsia* **53**, 19–33 (2012)
29. Morey, R., Petty, C., et al.: A comparison of automated segmentation and manual tracing for quantifying hippocampal and amygdala volumes. *Neuroimage* **45**(3), 855–866 (2009)
30. Pardoe, H., Pell, G., Abbott, D., Jackson, G.: Hippocampal volume assessment in temporal lobe epilepsy: how good is automated segmentation? *Epilepsia* **50**(12), 2586–2592 (2009)
31. Renzel, R., Baumann, C., Poryazova, R.: EEG after sleep deprivation is a sensitive tool in the first diagnosis of idiopathic generalized but not focal epilepsy. *Clin. Neurophysiol.* **127**(1), 38 (2016)
32. Shen, L., Saykin, A.J., Chung, M.K., Huang, H.: Morphometric analysis of hippocampal shape in mild cognitive impairment: An imaging genetics study. In: *Proceedings of the 7th IEEE International Conference on Bioinformatics and Bioengineering* (2007)
33. Stewart, R.: Mild cognitive impairment—the continuing challenge of its “real-world” detection and diagnosis. *Arch. Med. Res.* **43**, 609–14 (2012)
34. Suh, J.W., Wang, H., Das, S., Avants, B., Yushkevich, P.A.: Automatic segmentation of the hippocampus in T1-weighted MRI with multi-atlas label fusion using open source software: evaluation in 1.5 and 3.0T ADNI MRI. In: *Proceedings of the International Society for Magnetic Resonance in Medicine Conference (ISMRM 2011)* (2011)
35. Tae, W.S., Kim, S.S., Lee, K.U., Nam, E.C., Kim, K.W.: Validation of hippocampal volumes measured using a manual method and two automated methods (FreeSurfer and IBASPM) in chronic major depressive disorder. *Neuroradiology* **50**, 569–581 (2009)
36. Thyreau, B., Sato, K., Fukuda, H., Taki, Y.: Segmentation of the hippocampus by transferring algorithmic knowledge for large cohort processing. *Med. Image Anal.* **43**, 214–228 (2018)
37. Warfield, S., Zou, K., Wells, W.: Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation. *IEEE Trans. Med. Imaging* **23**(7), 903–921 (2004)

38. Wenger, E., Martensson, J., Hoack, H., et al.: Comparing manual and automatic segmentation of hippocampal volumes: reliability and validity issues in younger and older brains. *Epilepsia* **35**(8), 4236–4248 (2014)
39. Winston, G., Cardoso, M., et al.: Automated hippocampal segmentation in patients with epilepsy: available free online. *Epilepsia* **54**(12), 2166–2173 (2013)
40. Zarpalas, D., Gkontra, P., Daras, P., Maglaveras, N.: Accurate and fully automatic hippocampus segmentation using subject-specific 3D optimal local maps into a hybrid active contour model. *IEEE J. Transl. Eng. Health Med.* **2**, 1–16 (2014)