NORHA: A NORmal Hippocampal Asymmetry deviation index based on one-class novelty detection and 3D shape features

Autor/es ditelliano/s: Iarussi, Emmanuel *(Universidad Torcuato Di Tella. Escuela de Negocios. Laboratorio de Inteligencia Artificial*)

Fecha de publicación: 2023

Es una versión del artículo publicado en Brain Topography (ISSN 1573-6792)

¿Cómo citar la versión final?
Deangeli, D., Iarussi, F., Külsgaard, H. et al. NORHA: A NORmal Hippocampal Asymmetry Deviation Index Based on One-Class Novelty Detection and 3D Shape Features. Brain Topogr (2023). https://doi.org/10.1007/s10548-023-00985-6

NORHA: A NORmal Hippocampal Asymmetry deviation index based on one-class novelty detection and 3D shape features

Duilio Deangeli[1,2*], Francisco Iarussi[1], Hernán Külsgaard[1,2], Delfina Braggio[1,2], Juan Pablo Princich[3], Mariana Bendersky[3,4], Emmanuel Iarussi[2,5], Ignacio Larrabide[1,2] and José Ignacio Orlando[1,2]

[1]Yatiris, PLADEMA, UNICEN, Tandil, Buenos Aires, Argentina.
[2]CONICET, Buenos Aires, Argentina.
[3]ENyS, CONICET-HEC-UNAJ, Florencio Varela, Buenos Aires, Argentina.
[4]Normal Anatomy Department, UBA, CABA, Argentina.
[5]Laboratorio de Inteligencia Artificial, Universidad Torcuato Di Tella, CABA, Argentina.

*Corresponding author(s). E-mail(s): ddeangeli@plademe.exa.unicen.edu.ar;

Abstract

Radiologists routinely analyze hippocampal asymmetries in magnetic resonance (MR) images as a biomarker for neurodegenerative conditions like epilepsy and Alzheimer's Disease. However, current clinical tools rely on either subjective evaluations, basic volume measurements, or disease-specific models that fail to capture more complex differences in normal shape. In this paper, we overcome these limitations by introducing NORHA, a novel NORmal Hippocampal Asymmetry deviation index that uses machine learning novelty detection to objectively quantify it from MR scans. NORHA is based on a One-Class Support Vector Machine model learned from a set of morphological features extracted from automatically segmented hippocampi of healthy subjects. Hence, in test time, the model

automatically measures how far a new unseen sample falls with respect to the feature space of normal individuals. This avoids biases produced by standard classification models, which require being trained using diseased cases and therefore learning to characterize changes produced only by the ones. We evaluated our new index in multiple clinical use cases using public and private MRI datasets comprising control individuals and subjects with different levels of dementia or epilepsy. The index reported high values for subjects with unilateral atrophies and remained low for controls or individuals with mild or severe symmetric bilateral changes. It also showed high AUC values for discriminating individuals with hippocampal sclerosis, further emphasizing its ability to characterize unilateral abnormalities. Finally, a positive correlation between NORHA and the functional cognitive test CDR-SB was observed, highlighting its promising application as a biomarker for dementia.

Keywords: Hippocampus, Normal asymmetries, Machine learning, Novelty Detection

1 Introduction

The Hippocampus is a brain structure located in the medial temporal lobe that plays roles in the process of cognition, navigation and memory consolidation (Marshall and Born, 2007). Changes in its morphology are expected to occur naturally due to ageing and mostly involve alterations in volume and bilateral displacements either inwards or outwards, among others (Lucarelli et al, 2013). Neurodegenerative conditions such as Alzheimer's Disease (AD), focal cortical dysplasia (FCD) and epilepsy also tend to perturb the morphology and volume of the entire hippocampus (Moodley and Chan, 2014). AD, for instance, has been observed to generate progressive hippocampal loss and shape changes (Shi et al, 2009). On the other hand, epilepsy is known to produce unilateral or bilateral hippocampal volume loss (Bernasconi et al, 2003).

Hippocampal asymmetries have been reported, and traditionally the right hippocampus is considered to be larger than the left one (Shi et al, 2009; Pedraza et al, 2004). Differences in thickness and volume of the layers of the different subfields were also observed between the right and left hippocampus in normal subjects (Lister et al, 2006). On the other hand, AD is known to produce striking changes in the hippocampus, which explains the memory deficits that are the hallmark of this disease. These alterations are typically in terms of volume, shape and also symmetry. Some authors report that the hippocampal R>L asymmetry may be increased in AD (Shi et al, 2009), while others postulate that such asymmetry disappears or may even be reversed in some patients (Geroldi et al, 2000). In the same way that there is a clinical continuum between Mild Cognitive Impairment (MCI) and AD, it is postulated that there are imaging biomarkers, particularly related to hippocampal anatomy, that

could predict the conversion of the former into the latter. However, visual evaluations of hippocampal asymmetries are not exempt from human error, and it is arduous to define the exact boundaries between normal and pathological asymmetry. Moreover, current clinical tools for objectively quantifying normal asymmetry patterns do not go beyond volumetric measurements (Manjón and Coupé, 2016), or are limited to disease-specific models (Fu et al, 2021). Using only volumetric features fails to capture morphological asymmetries beyond mere differences in size (Richards et al, 2020). On the other hand, disease-specific models cannot be extrapolated to the discrimination of other conditions (Fu et al, 2021).

To overcome these limitations, in this study, we introduce NORHA, a NORmal Hippocampal Asymmetry deviation index that can characterize healthy asymmetries in the hippocampus using a novelty detection machine learning model. Given a set of MR scans from a normal cohort, 3D meshes of both hippocampi are obtained using a deep learning-based segmentation algorithm. Features describing their shape are then automatically extracted and used to characterize the differences in morphology between the hippocampi. These vectors are subsequently used for training a One-class Support Vector Machine (OC-SVM) model, a novelty detection algorithm that learns a geometrical boundary that describes the training samples. During test time, the resulting model is used to quantify how much a new, unknown input sample deviates from normal patterns of hippocampal asymmetry by quantifying the signed distance to this boundary. This value then acts as an index that can be used to measure deviations from normal anatomical differences and applied in clinical studies to monitor this property in multiple cohorts.

To the best of our knowledge, NORHA is the first learnable index to objectively quantify normal morphological shape asymmetries between a single subject's own hippocampi.

While automatically characterizing hippocampal asymmetries with machine learning has been explored for automated classification of diseases such as AD and MCI in the past (Fu et al, 2021), differences were only studied between a healthy and a diseased population and not among healthy subjects. This renders models that can automatically identify a disease by analyzing hippocampal asymmetry but cannot be used to quantify asymmetries in healthy subjects or to determine when the changes are abnormal, e.g. when applied to a disease unseen during training. On the contrary, our model is, by definition, tailored to quantify asymmetries in normal individuals, and it is expected to recognize deviations produced by any disease.

We empirically evaluated this model using multiple publicly available and in-house databases. We observed that our NORHA deviation index increases in subjects with temporal lobe epilepsy due to left/right hippocampal sclerosis, which is characterized by notorious unilateral differences between both hippocampi. Furthermore, we observed a gradual increment of the index with the MCI-to-AD progression, indicating that the model is sensitive to progressive unilateral shape variations. Finally, to study potential clinical applications, we

analyzed the relationship between NORHA and a cognitive score measured on control subjects, MCI and AD patients, observing a positive correlation.

In summary, our method contributes to state of the art from multiple perspectives:

- It is the first approach for automatically characterizing deviations in normal anatomical asymmetries of the hippocampi beyond using only volumetric comparison;
- Compared to standard classification models, ours requires only healthy subjects for training, making it able to quantify alterations produced by any condition;
- We provide a comprehensive assessment of potential applications for our index, using both public and private MRI datasets;
- Finally, we show that NORHA can effectively quantify unilateral asymmetries in healthy subjects and individuals with dementia or hippocampal sclerosis.

The remainder of this paper is organized as follows. Section 2 describes the datasets used in our experiments and the proposed method. Section 3 presents the results of our experiments, while Section 4 discusses potential clinical implications. Finally, Section 5 presents the summary and conclusions of this work.

2 Materials and Methods

2.1 Subject data

We conducted our experiments using volumetric T1 structural MRIs of both healthy and diseased patients from multiple datasets, both public and private. Public sets included OASIS [1] (LaMontagne et al, 2019), IXI [2] and ADNI [3] (Mueller et al, 2005), while two private datasets were used. The first one was acquired at the Hospital de Alta Complejidad En Red "El Cruce" Dr. Néstor Carlos Kirchner (Florencio Varela, Argentina) and the second one at the Instituto de Oncología Ángel H. Roffo (Buenos Aires, Argentina). A summary of their acquisition parameters and image characteristics is included in Table 1. The acquired resolution was never modified in this study. None of these datasets included information about the handedness of the individuals.

2.1.1 OASIS

OASIS is a widely used public set of MR data created by the Alzheimer's Disease Research Centre from the Washington University. We made use of the T1w images from OASIS-3, which is the latest release of the database. It contains a total number of 2363 T1w scans from 1098 participants with ages ranging

[1] https://www.oasis-brains.org/
[2] http://brain-development.org/ixidataset/
[3] http://adni.loni.usc.edu/

from 42 to 95 years, all of them with a voxel size of $1 \times 1 \times 1$ mm^3, including cognitively normal individuals (Normal Controls, NC), and patients with early-stage AD dementia. The images were obtained using different scanners, namely Vision 1.5T, TIM Trio 3T and BioGraph mMR PET-MR 3T from Siemens (Erlangen, Germany). Further technical details about the datasets are available in (LaMontagne et al, 2019).

2.1.2 IXI

IXI is a publicly accessible dataset of MR scans acquired at three sites, namely the Hammersmith and the Guy's Hospitals and the Institute of Psychiatry. We used images from these three institutions in our experiments. We only used T1w volumes, all of them with a dimension of $256 \times 256 \times 128$ voxels, obtained from subjects without known pathologies with ages ranging from 20 to 86 years. The Hammersmith Hospital subset contains 171 T1w scans, acquired using a Philips Intera 3T scanner at a voxel resolution of $0.9375 \times 0.9375 \times 1.2$ mm^3, with a Field of View (FoV) $= 256 \times 256$ mm^2, with a TR $= 9.6$ ms, TE $= 4.6ms.$, 208 phase encoding steps, echo train length of 208, reconstruction diameter of 240.0 mm, and a flip angle of 8°. The Guy's Hospital set includes 297 T1w images obtained using a Philips Intera 1.5T scanner device at a voxel resolution of $0.93 \times 0.93 \times 1.2$ mm^3, FoV $= 256 \times 256$ mm^2, with a TR $= 9.8$ ms, TE $= 4.6$ ms, 192 phase encoding steps, echo train length of 0, reconstruction diameter of 240.0 mm, and a flip angle of 8°. Finally, the Institute of Psychiatry subset includes 71 images retrieved with a GE 1.5T system. Details about the scan parameters were not publicly available at the time of preparing this manuscript.

2.1.3 ADNI

We used the public available version 1 of ADNI, namely ADNI1 [4], which was created between years 2004 and 2009. This set includes MR scans from 199 subjects, 53 NC, 71 with mild cognitive impairment (MCI) and 33 with AD, with their associated age and their Clinical Dementia Rating Scale–Sum of Boxes (Lynch et al, 2006) (CDR–SB) scores, Mini-mental state examination (Folstein et al, 1975) (MMSE) and Alzheimer's Disease Assessment Scale (Rosen et al, 1984) - Cognitive Subscale 11 (ADAS11). As previously observed by Nettiksimmons et al (2014) and (Ezzati et al, 2020), ADNI's MCI group is highly heterogeneous, with cognitive scores suggesting a common path to AD but biological biomarkers such as hippocampal changes contradicting this evidence in certain subgroups (Nettiksimmons et al, 2014). Further details about this variability can be found in the corresponding studies (Nettiksimmons et al, 2014; Ezzati et al, 2020).

T1w images covering the entire brain were obtained with scanners from three different manufacturers (GE, Philips and Siemens), acquired with 1.2 mm slice thickness, 160 sagittal slices, FoV $= 192 \times 192$ mm^2, and a 192×192

[4] www.adni-info.org

scan matrix, voxel size of $1.2 \times 0.9375 \times 0.9375$ mm^3, TR = 3000 ms, and TE = 3.55 ms.

2.1.4 Hospital El Cruce dataset (HEC)

This dataset corresponds to 101 T1w MR scans acquired at Hospital de Alta Complejidad En Red "El Cruce" Dr. Néstor Carlos Kirchner (Florencio Varela, Argentina) in the context of two major clinical research projects. Images correspond to 53 acquisitions from control subjects and 48 scans from patients with temporal lobe epilepsy associated with unilateral right and left hippocampal sclerosis (HSR and HSL, respectively), who were enrolled to be part of a study about phenotyping brain structures. Patients, on the other hand, were retrospectively enrolled between 2014 and 2019 based on clinical records from the Epilepsy Unit at the hospital. All patients gave informed consent prior to inclusion in the respective clinical trials. All prospective studies were approved by the Ethics Committee of the Hospital in accordance with the tenets of the Declaration of Helsinki.

Structural images consist of a 3D T1w fast field echo (FFE) sequence, with 180 slices of 1-mm isotropic resolution, TE = 3.3 ms, TR = 2300 ms, TI = 900 ms, flip angle = 9°, FoV = $240 \times 240 \times 180$ mm^3, voxel size of $0.75 \times 0.75 \times 1$ mm^3, acquired using a Philips Achieva (Best, The Netherlands) 3T scanners with 8-channel head coil.

2.1.5 Instituto de Oncología Ángel H. Roffo dataset (ROFFO)

This dataset contains scans for 83 subjects without known pathology, all of them recruited at the Instituto de Oncología Ángel H. Roffo. Participants gave informed consent to take part in a study focused on analyzing the brain in epilepsy, which was approved by the institutional ethics review board of the Institute in agreement with the tenets of the Declaration of Helsinki.

The T1w images were acquired with a 3.0T Siemens Trio MRI scanner, with a TR = 2 ms, TE= 3.7 ms, inverted angle = 80, FoV in plane = 214×214 mm^2 and matrix of size 240×240, coding phase in antero-posterior direction and from left to right, block thickness = 128 mm, Nav = 1 (average number of signals), voxel size = $0.89 \times 0.89 \times 1.0$ mm^3, acquisition of bandwidth = 191.5 Hz/pixel, and parallel image (SENSE factor = 8). The images were reconstructed with an intra-plane interpolation of factor = 2 in each dimension.

2.2 Hippocampus segmentation

Left and right hippocampi were automatically segmented on all T1w images for subsequent analysis. Each scan was manually aligned and normalized to standard space, using the Montreal Neurological Institute (MNI) standard T1 template as a reference. This model has a dimension of $181 \times 217 \times 181$ voxels with a voxel size of $1 \times 1 \times 1$ mm^3. Preprocessing was performed using the Statistical Parametric Mapping tool (SPM12) [5]. After

[5] https://www.fil.ion.ucl.ac.uk/spm/

Table 1 Acquisition parameters and image characteristics on each dataset.

Dataset		Resolution (voxel size, mm)	FoV (mm)	TR (ms)	TE (ms)	Image dimension (voxels)
OASIS		1 × 1 × 1	256 × 256	9.7	4.0	256 × 256 × 176
ADNI		0.9375 × 0.9375 × 1.2	192 × 192	3000	3.55	256 × 256 × 166
IXI	IOT	0.9375 × 0.9375 × 1.2	-	-	-	256 × 256 × 128
	GUY	0.93 × 0.93 × 1.2	256 × 256	9.8	4.6	256 × 256 × 128
	HH	0.9375 × 0.9375 × 1.2	256 × 256	9.6	4.6	256 × 256 × 128
ROFFO		0.89 × 0.89 × 1	214 × 214	2	3.7	256 × 256 × 192
HEC		0.75 × 0.75 × 1	240 × 240	2500	3.3	320 × 320 × 180

alignment, voxelwise segmentation masks of the hippocampi were obtained using hippmapp3r [6] (Goubran et al, 2020), a method that uses a 3D convolutional neural network with a U-shaped architecture for segmenting these structures in the patients' images from all datasets.

Triangular meshes were automatically retrieved from each hippocampus using the marching cubes algorithm implemented in the Isosurface function of AFNI/SUMA (Saad et al, 2004). As the overall shape and topology of the structures can be degraded due to spatial interpolation and voxelization artifacts, meshes were smoothed using Laplacian smoothing from the Trimesh library [7] (Dawson-Haggerty et al, 2019). A smoothing factor of 0.2 was empirically chosen after observing that the resulting meshes did not show artifacts and were faithful representations of the hippocampus' anatomy.

2.3 Dataset integration

Images from all the datasets described in Section 2.1 were combined into a single initial set of 3619 potentially eligible scans. All scans were automatically segmented and analyzed. Hippocampus meshes with a volume smaller than 1500 mm^3 (which is anatomically highly unlikely) were assumed as failed segmentations and, therefore, not included in the final dataset (376 in total, approximately 11% of the original set). Figure 1 depicts flow diagrams illustrating the data preparation process. Our final dataset included 3619 scans, whose demographic characteristics are summarized in Table 2.

In this study, we aimed to address the potentially confounding variables of sex and age by creating a training dataset that covers a wide range of ages. The distribution of gender was randomized and analyzed separately.

The final dataset was separated into a training set and a series of test sets, as indicated in Table 3. The training set was used for learning our deviation from the normal hippocampal asymmetry index from a healthy population. Hence, it included all the healthy subjects from ROFFO (83) and IXI (539)

[6] https://hippmapp3r.readthedocs.io
[7] https://trimsh.org/

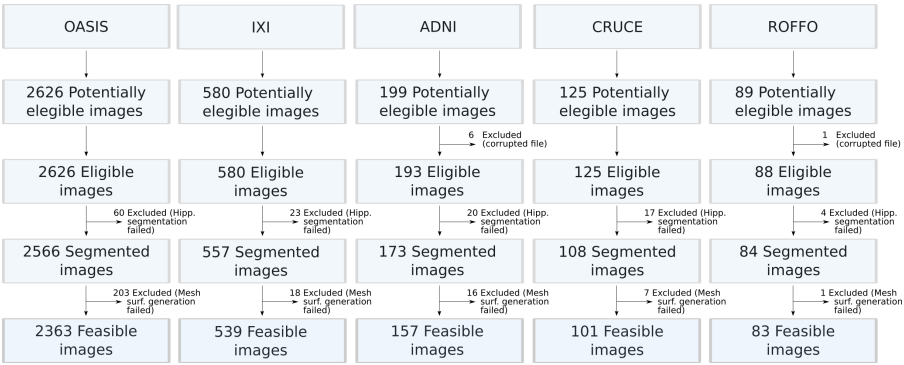

Fig. 1 Flow diagrams illustrating our data preparation protocol. Side arrows indicate exclusion of images due to selection criterion failure.

Table 2 Summary of demographic characteristics of the final set used in our experiments. AD: Alzheimer's Disease, MCI: Mild Cognitive Impairment, HSR: Right-sided Hippocampal Sclerosis, HSL: Left-sided Hippocampal Sclerosis, NC: Normal Controls. M: male. F: female.

Parameter	AD (N = 179)	MCI (N = 71)	HSR (N = 16)	HSL (N = 32)	NC (N = 2945)
Age [years]	74 ± 9	75 ± 8	35 ± 9	34 ± 10	60 ± 15
Gender [M/F]	86/93	42/29	10/6	15/17	1684/1261

and 70% of the subjects from the healthy subset of OASIS. The test sets were designed to be used in the evaluation of the proposed index under different settings. TEST-OASIS comprises all subjects with AD and the remaining 30% of control individuals from OASIS. TEST-ADNI and TEST-HEC, on the other hand, include all subjects from ADNI and HEC, respectively. Notice that the training set includes ages spanning a wider range than the individual test sets. This was done on purpose to ensure that our normal asymmetry index captures the anatomical variations of healthy individuals of multiple ages.

Table 3 Training and test sets characteristics. AD: Alzheimer's Disease, MCI: Mild Cognitive Impairment, HSR: Right-sided Hippocampal Sclerosis, HSL: Left-sided Hippocampal Sclerosis, NC: Normal Controls.

Partitions		Gender (M/F)	Mean age [min-max]	NC	AD	MCI	HSL	HSR	Total
Train		1284/838	60 [19-95]	1500 (OASIS) + 539 (IXI) + 83 (ROFFO)	-	-	-	-	2122
Test	OASIS	434/429	65 [45-88]	717	146	-	-	-	863
	ADNI	74/83	75 [56-89]	53	33	71	-	-	157
	HEC	43/58	32 [18-63]	53	-	-	32	16	101

2.4 Characterization of hippocampal asymmetries using shape features

Figure 2 shows the feature extraction process applied to characterize asymmetries between hippocampi. After image segmentation (Section 2.2), we automatically calculate the shape and morphological features of both hippocampi from their corresponding meshes, being $v_R^{(i)}$ and $v_L^{(i)}$ the i-th feature from the right and left hippocampus feature vector, respectively. Differences in features between both structures are computed using distance metrics d such as the Euclidean and the Mahalonobis ones. An asymmetry vector with 14 elements is constructed by concatenating the resulting distances altogether. These arrays are then used in training time to learn our model and subsequently applied in test time to predict the distance with respect to the normal cohort. Features rationale and interpretation are explained in the following section.

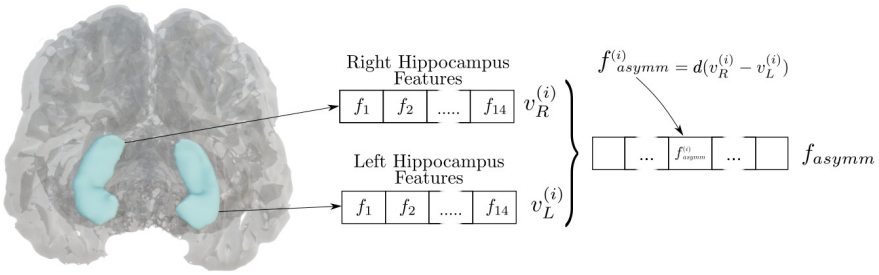

Fig. 2 Representation of feature extraction process. For the i-th hippocampi, morphological features are calculated for the right and left hippocampus (v_R and v_L). Differences between them are accounted using distance metrics d, i.e. Mahalanobis and Euclidean distances. Finally, a vector of 14 features f_{asymm}, contains the values, which describe the asymmetry in this subject.

2.4.1 Morphological features

Morphological features represent the overall 3D shape of the hippocampus based on its triangulated surface mesh representation, are then used to describe it, and are used to quantify similarities and differences between neighbors.

The following 10 morphological features were collected from each hippocampus:

- Sphericity: is a measure of how closely the shape of the hippocampus resembles that of a perfect sphere. The value ranges from 0 to 1, with 1 indicating a perfect sphere. Hence, less elongated and more round hippocampi will result in high sphericities, while the elongated ones will be characterized by a lower value.

- Compactness: represent how compact the shape of the hippocampus is relative to a sphere. The value ranges from 0 to $\frac{16}{\pi}$, where $\frac{16}{\pi}$ indicates a perfect sphere.
- Quadratic Compactness: represents how compact the shape of the hippocampus is relative to a sphere, but this time computed using the quadratic volume. The value ranges from 0 to 1, where a value of 1 indicates a perfect sphere.
- Elongation: characterizes the degree of lengthening of the hippocampus. It is computed as the square root of the ratio between the minor and major principal components of the segmented structure in the voxel space. The feature ranges from 0 (maximal elongation, e.g. a 1D line) to 1 (circular, non-elongated structure).
- Flatness: ratio between the largest and smallest principal components of the rectangle enclosing the hippocampus. Thus, it indicates how much the hippocampus matches a perfectly plane surface. The values range between 1 (spherical object) and 0 (a flat object).
- Spherical Disproportion: measures how much the hippocampus differs from a sphere with its same volume. It is obtained by taking the quotient of the surface area of the hippocampus and the surface area of this sphere. Thus, the value is ≥ 1, with 1 being a perfect sphere.
- Surface Volume Ratio: area of the hippocampus divided by the hippocampal volume. A smaller value indicates a more compact (sphere-like) shape.
- Maximum 2D Diameter: the distance between the first and the last axial slices of the segmented volume. The bigger the value is, the larger the hippocampus is in the axial plane.
- Maximum 3D Diameter: the largest pairwise Euclidean distance between mesh vertices. It can be used as a proxy for the overall size of the hippocampus, as large values of this feature correspond to structures in which the vertices are distant from one another.
- Major Axis: corresponds to the largest axis length of the boundaries of the hippocampus, as obtained from an enclosing ellipsoid. Elongated hippocampi will have larger major axes than non-elongated ones.

We used Pyradiomics [8] to automatically compute these values for each hippocampus. Features representing the asymmetry between the left and right counterparts were obtained by computing the Euclidean distance between each other's vectors.

2.4.2 Volumetric features

The volume of each hippocampus was extracted directly from the 3D segmentation obtained from the MRI scans. We estimated two volumetric asymmetry measurements, namely raw volume differences (obtained as the volume of the left hippocampus minus the right one) and normalized volume differences (raw value difference divided by the maximum volume).

[8] http:/PyRadiomics.readthedocs.io/en/latest/

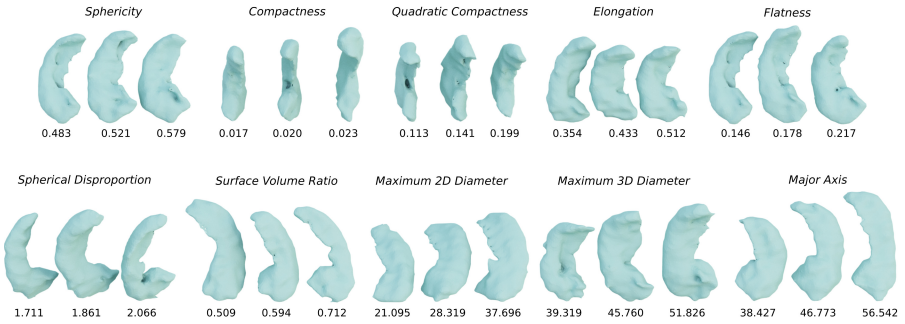

Fig. 3 Qualitative examples of different hippocampi and their corresponding morphological features. For each characteristic, we include a triplet of hippocampi with the minimum, median and maximum values.

2.4.3 ShapeDNA (Reuter et al, 2006)

ShapeDNA is a technique widely applied for characterizing non-rigid 3D meshes (Lian et al, 2013). It is obtained by computing the normalized eigenvalues of the Laplace-Beltrami operator (Reuter et al, 2006), and results in a vector of $\sim$ 50 eigenvalues for each analyzed structure. We followed the same approach as in (Wachinger et al, 2016), consisting in first extracting a ShapeDNA vector with 50 features from each hippocampus and then taking the Euclidean and Mahalanobis distances between them, which is a way of quantifying their differences. The covariance matrix for the Mahalanobis distance was obtained from the training set of healthy subjects and reused in test time for extracting the features. This process results in a length 2 vector. These 2 values were placed in the 13th and 14th positions of the final asymmetry vector.

2.5 Learning a normal hippocampal asymmetry deviation index with one-class support vector machines

Novelty detection algorithms aim at learning the common properties of a set of training samples to discriminate if any new given input belongs or not to the previously observed population. These models are usually applied for anomaly detection in medical imaging by training them using normal subjects. These models are then applied to samples with unknown diagnostic, for instance, to determine if they fall within or close to the healthy population or not. The key advantage of this family of methods over traditional supervised learning models is that they do not require diseased samples for training but only control individuals.

The One-Class Support Vector Machine (OC-SVM) is a novelty detection model (Schölkopf et al, 1999) that is widely applied for this purpose. This algorithm learns a hyperdimensional compact boundary that encloses the set of normal training samples and automatically determines if a new unseen sample is normal or not, depending on its location with respect to this boundary. Each

training sample corresponds, in our case, to a feature vector representing the asymmetry between the hippocampi of a normal subject. These observations are then mapped into a higher dimensional feature space by projecting them using a γ parametrized Radial Basis Function (RBF), a translation-invariant kernel that makes features grouped within a circular area. This ensures the parameters of the learned boundary match those from a hypersphere enclosing the majority of the training samples. The number of observations kept outside this boundary is controlled by the hyperparameter ν, which determines the fraction of the normal data that will be classified as outliers. After training, new observations can be classified either as typical (inliers) or atypical (outliers) by measuring their signed distance with respect to the boundary. In this study, we propose to use this distance as a deviation from the normal hippocampal asymmetry index. In this setting, small or negative values of the index correspond to more typical anatomical differences, while large positive values indicate that the hippocampal asymmetries are likely abnormal.

Data standardization was applied before training the model to avoid features with large ranges to absorb the contributions of the ones with smaller ranges. We applied a robust scaler alternative, consisting of subtracting the median of the feature instead of the mean, thus reducing the effect of outliers. Then, the resulting value was divided by the range between the third (75th) and first (25th) quartiles.

We used OC-SVM classifier with an RBF kernel. Hyperparameters γ and ν were experimentally calibrated using 5-fold cross-validation on the training set. Since our training data only contains healthy samples, hyperparameters were adjusted by optimizing the F1-score for detecting pseudo anomalies artificially generated by manipulating the normal data on each of the validation folds. Assuming that each i-th feature has a normal distribution with mean μ_i and standard deviation σ_i, we created fake anomalous samples by replacing each i-th feature with a random value in the range $[2\sigma_i, 3\sigma_i]$. This allows us to create degenerated samples that differ from the original data and, therefore, should lie outside the limits of our model. We used this approach to augment our validation folds with 50 new artificial samples, resulting in a total number of 475 samples per fold. As these points have a different distribution with respect to the majority of the training data, we expect our algorithm to classify them as outliers. The OC-SVM model that performed the best during this stage ($\nu = 0.2$ and $\gamma = 0.001$) was considered optimal and used for evaluation on new unseen subjects in the test phase.

2.6 Statistical analysis and evaluation metrics

Differences in the predicted asymmetry indices of diseased and healthy subjects were statistically assessed using Wilcoxon signed-rank tests, with $\alpha = 0.05$ (R statistical software package (Team et al, 2013)). Correlation analysis between the index and the CDR-SB values was performed using the Pearson correlation coefficient (R). As one of the representative scores used to assess recognition

and social function (O'Bryant et al, 2008), CDR-SB rates the degree of deterioration in the subject's cognitive status in 6 different domains of functioning, including memory, orientation, judgment and problem-solving, community affairs, home and hobbies, and personal care, using five values per domain (0: None, 0.5: Questionable, 1: Mild, 2: Moderate, 3: Severe). The CDR-SB score values used in this study were obtained by summing each of the domain box scores, therefore ranging from 0 to 18. Receiver Operating Characteristic (ROC) curves and their areas under the curve (AUC), as implemented in Scikit-Learn, were applied to study the classification performance of the predicted NORHA index when trained using different sets of features and applied to differentiate between healthy and diseased subjects. The 95% Confidence interval (95%CI) was calculated using bootstrapping with $N = 1000$ random samples. The Star tool (Vergara et al, 2008) was used to assess the statistical significance of the differences between curves. Finally, t-distributed stochastic neighbor embeddings (t-SNE) (Van der Maaten and Hinton, 2008) were used to visually assess the distribution of samples in the feature space, using the Scikit-Learn implementation (Pedregosa et al, 2011). This representation preserves distances between pairs of samples and reveals proximity relationships and feature clusters, providing an intuition on how these are arranged in a higher dimensional space.

3 Results

We evaluate the proposed NORHA deviation index in multiple scenarios based on different alternative tests.

First, we analyze the statistical variations in the distribution of NORHA values when computed from healthy subjects and individuals suffering from different neurodegenerative conditions, discriminating by their corresponding source dataset too. The obtained results are depicted in Figure 4. Patients with HSL and HSR report significantly higher NORHA indices compared to NC, MCI and AD individuals ($p \approx 0.0$). No statistically significant differences are observed between the sets of control subjects from OASIS, ADNI and HEC ($p > 0.360$) and between patients with MCI and healthy individuals ($p > 0.766$). AD subjects from OASIS show significant differences with respect to controls from the same original database ($p = 0.035$). However, they do not show significant differences with respect to controls from ADNI and HEC or to MCI individuals ($p > 0.768$). AD patients from ADNI, on the other hand, exhibit statistically significant deviations from NC ($p > 0.003$) and MCI ones ($p = 0.035$). Finally, the differences in NORHA were not significant between AD individuals from ADNI and OASIS ($p = 0.1984$).

To discard any potential age and/or sex bias in the index, we analyzed the distribution of NORHA indices within each disease group, separated by each potential confounding factor (Figure 6). Age intervals were determined using the same approach applied by Andrade de Oliveira et al (2015). Empty boxes correspond to intervals with no samples for that specific group ($N = 0$).

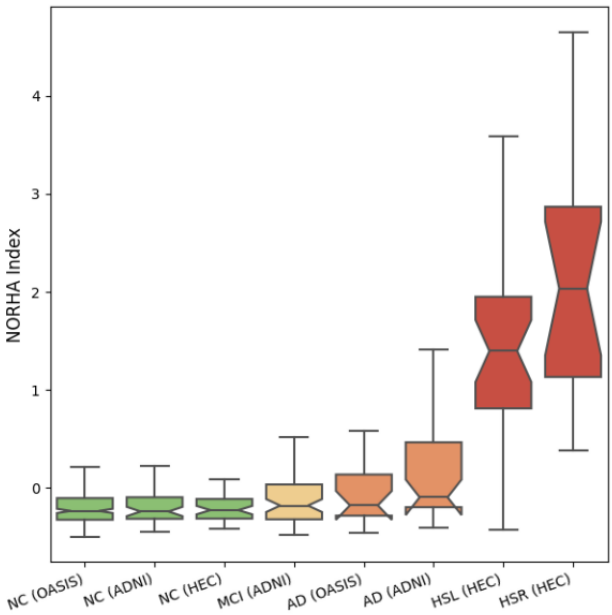

Fig. 4 Box plot depicting the distribution of our normal hippocampi asymmetry index for normal control (NC) subjects (green) and diseased individuals with mild cognitive impairment (MCI, yellow), Alzheimer's disease (AD, orange) and left (HSL) or right (HSR) hippocampal sclerosis (red), as taken from the HEC, OASIS and ADNI test sets.

Boxplots show that the index distributes approximately similar within each individual disease group, regardless of age and sex. Non-parametric hypothesis testing was only possible in sets with $N \geq 16$ (Dwivedi et al, 2017). No statistically significant differences were observed between age and sex subgroups in the NC group ($p > 0.132$). In the AD group, the sample size only allowed for a comparison between the 60-69 and >70 subgroups. Significant differences were only observed between the female subgroups ($p = 0.046$) and not in the male cohort ($p > 0.221$). When comparing males and females within each disease and age subgroup, no statistical differences were observed in most of the compared cases ($p > 0.055$). The only significant differences were seen between males and females with ages ranging 50-59 years, in NC ($p = 0.011$).

Figure 5 presents a qualitative analysis based on four exemplary cases. The first example corresponds to a healthy individual extracted from the HEC test set, to which the model assigned a low NORHA index. Right-side lateralization is observed, which is consistent with previous studies that reported bigger hippocampus volumes on the right hemisphere (Shi et al, 2009). The second example is a subject with HSL taken from the HEC set with a high associated

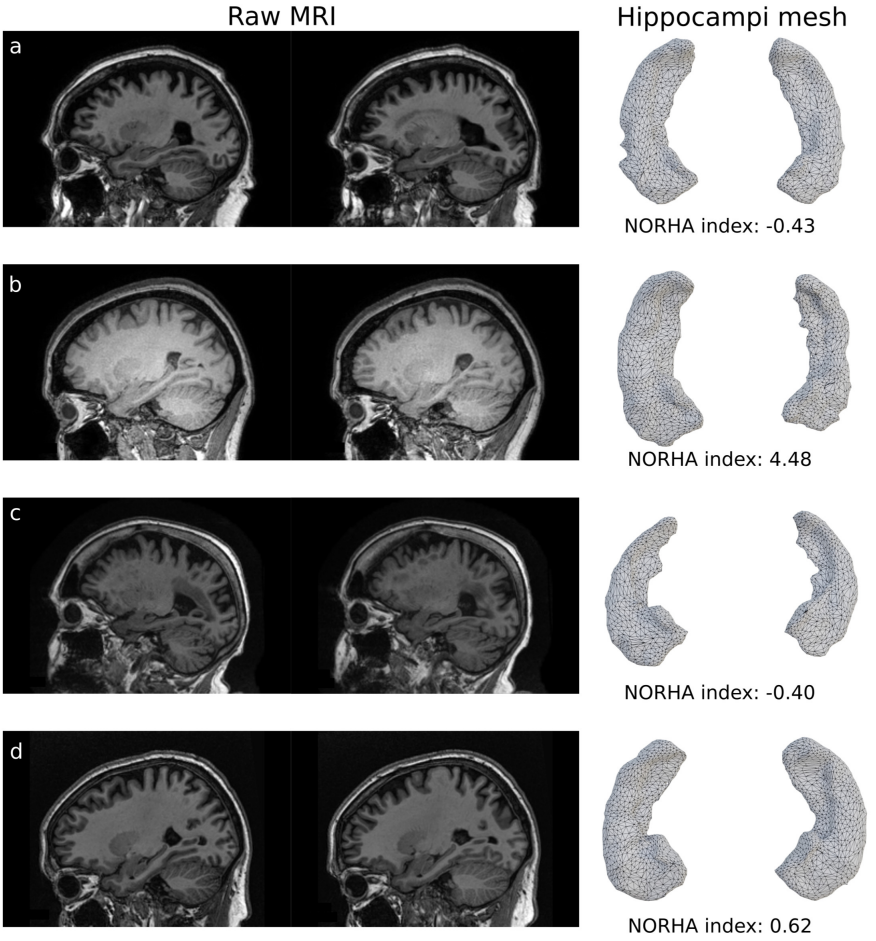

Fig. 5 Qualitative analysis. From left to right: two sagittal slices showing the left and right hippocampi, respectively, and their corresponding segmentation and NORHA indices. From top to bottom: (a) normal case with a low NORHA index showing fairly symmetric hippocampus with slightly higher volume on the right side. (b) HSL case with a high NORHA index and a notorious atrophic left hippocampus. (c) MCI case with a low NORHA index with a normal asymmetry pattern. (d) MCI case with a high NORHA index and an inverted asymmetry pattern, with bigger left-side hippocampus.

NORHA index. The case corresponds to a typical unilateral atrophic appearance of the left hippocampus, considered highly suggestive of hippocampal sclerosis in mesial temporal epilepsy. The third case is a subject suffering from MCI extracted from the ADNI test set. Its associated NORHA index is low, meaning that the model predicts that the morphological differences between the two hippocampi are compatible with a normal scenario. Finally, the last example corresponds also to an MCI case from the ADNI test set, to which

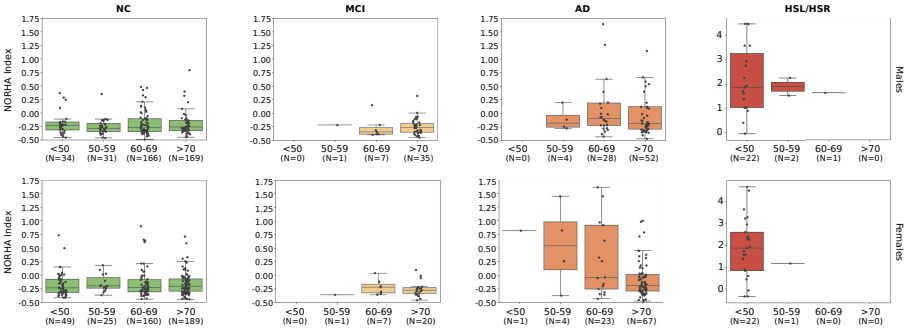

Fig. 6 Analysis of potential age and sex biases in the NORHA index. Boxplots represent index distribution per disease, separating each by age and sex subgroups. N = number of samples within each subgroup. Notice that the HSL/HSR subgroups do not use the same scale for the NORHA index due to high differences in their range with respect to the other subgroups.

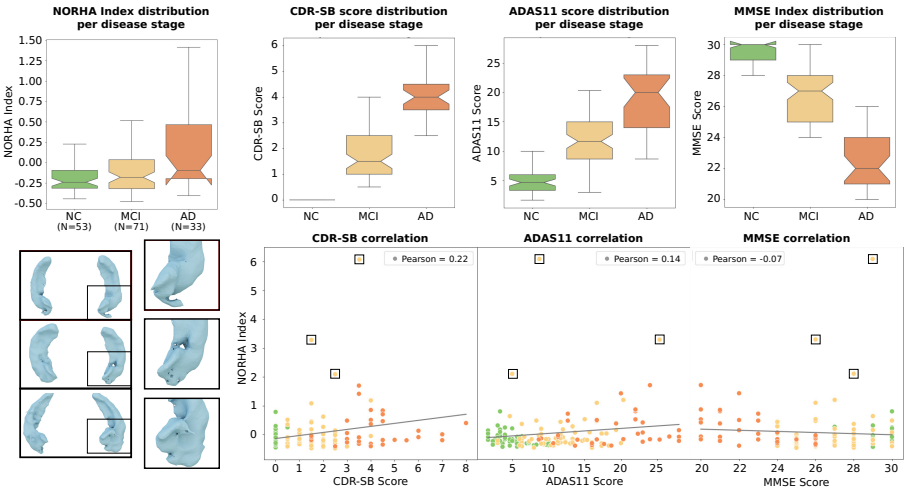

Fig. 7 Top Left: Box plot depicting the distribution of our NORHA index for normal control (NC) subjects (green) and diseased individuals with mild cognitive impairment (MCI, yellow) and Alzheimer's disease (AD, orange) in ADNI test set. Top right: Box plot depicting the distribution of the CDR-SB score, ADAS11 score and MMSE score for the same groups. Bottom: Scatter plots depicting the correlation between our NORHA index and CDR-SB, ADAS11 and MMSE scores. Each sample is color coded according to its ground truth diagnosis (green = NC, yellow = MCI, orange = AD). MCI outliers with potential segmentation mistakes are highlighted with boxes and depicted as 3D meshes.

the model assigned a larger NORHA index with a visually distinctive normal pattern of asymmetry accounting for a slightly bigger left hippocampus.

To understand the relationship between the hippocampal asymmetries measured with our index and the degree of cognitive degradation exhibited by the individuals, we investigated the correlation between our index and different levels of CDR-SB, MMSE and ADAS11 in the ADNI dataset. Figure 7 presents the distribution of the NORHA and cognitive scores for NC, MCI and

AD cases in the test set. Control subjects had CDR-SB values of 0, while those with MCI had values ranging from 0.5 to 4, and subjects with AD had values ranging from 2.5 to 6. A weak positive correlation was found between the cognitive scores and NORHA index, with a correlation coefficient (R) of 0.223 ($p < 0.001$) for CDR-SB, 0.14 ($p < 0.001$) for ADAS11, and -0.07 ($p < 0.001$) for MMSE. Three outliers are observed within the MCI set, which are indicated with boxes in the scatter plots in Figure 7. Their associated meshes are depicted on the left side of the figure. Notice that the three of them show patterns that are inconsistent with the actual shape of a hippocampus, probably due to errors in segmentation. When correlation is analyzed without including these samples, the resulting values are improved in all cases, being $R = 0.27$ ($p < 0.001$) for CDR-SB, $R = 0.30$ ($p < 0.001$) for ADAS11, and $R = -0.32$ ($p < 0.001$) for MMSE.

We quantitatively analyze the effect of the selected descriptive features for characterizing the hippocampal asymmetries by comparing our model with two other alternative OC-SVMs trained using only the 10 morphological features described in Section 2.4.1 and incorporating to these the volume and normalize volume differences described in Section 2.4.2. In particular, we evaluated the three alternatives for discriminating NC individuals from subjects with HSR, HSL, MCI and AD, respectively. Figure 8 depicts their corresponding ROC curves and their associated AUC values.

High AUC values are obtained when discriminating HSL and HSR from NC, regardless of the features used for training. Nevertheless, the model applying the entire set of features achieved statistically better AUC values ($p < 0.031$). When discriminating MCI or AD from NC subjects, much lower AUC values are obtained. In particular, the three alternative NORHA indexes derived from the different sets of features reported similar AUC values for detecting AD both in ADNI and OASIS. Their differences are not statistically significant in ADNI. In OASIS, the model trained using all the features obtained statistically better results than the one trained only with the morphological ones ($p < 0.001$). Finally, notice that the AUC values for detecting MCI almost correspond to a random guess (AUC ≈ 0.5). To understand the discrimination power of NORHA in the context of other alternative features, we performed an experiment comparing NORHA's AUC with those obtained using each of the characteristics described in Section 2.4 as individual abnormality scores. Results are summarized in Table 4. Some features showed better results for specific diseases (e.g. ShapeDNA for MCI, Spherical Disproportion for AD samples in OASIS, and normalized volumetric differences in HSL and HSR). Still, none of them retained this performance over the whole disease spectrum. NORHA, on the other hand, consistently ranks within the best three alternatives among all diseases subgroups.

Finally, we qualitatively study the capabilities of our 14-dimensional feature vectors to distinguish between different diseases and normal subjects by projecting them in a 2D space using t-SNE. Figure 9 depicts with scatter plots

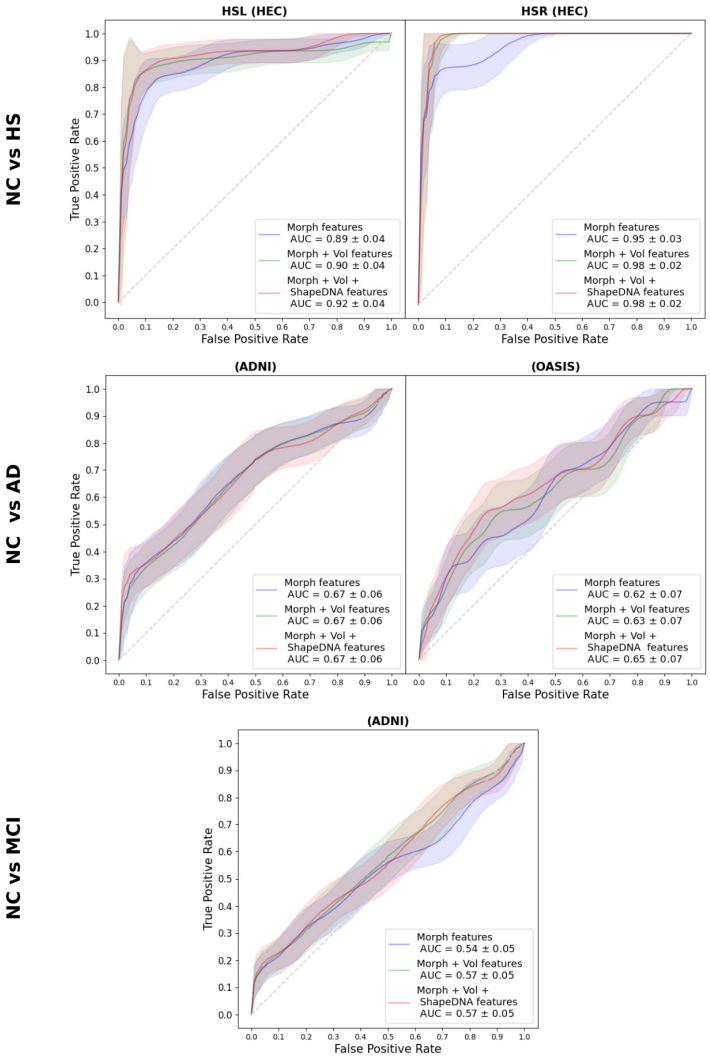

Fig. 8 ROC curves for distinguishing right-sided Hippocampal Sclerosis and Control subjects and left-sided Hippocampal Sclerosis and Control subjects (top). ROC curves for differentiating Alzheimer's Disease and Control subjects (middle). ROC curve for distinguishing Mild Cognitive Impairment and Control subjects (bottom). Each one of the cases was calculated using morphological features (blue line), morphological + volumetric features (green line) and morphological + volumetric + ShapeDNA features (red line).

the resulting projections. Each sample in Figure 9 (a) is color-coded according to its ground truth labeling, using red for diseased cases and green for the healthy ones. In this plot, normal subjects are not grouped within a single cluster but scattered throughout the plot in a non-uniform structure. The majority of the samples corresponding to individuals with MCI and AD are located at the boundaries of the healthy areas, while others are still located

Table 4 Comparison of AUC (95% CI) values obtained by different morphological features as disease scores vs. the NORHA index for discriminating NC from diseased individuals in each test set.

Method	MCI	AD (ADNI)	AD (OASIS)	HSL	HSR
Vol. Diff.	0.54 (0.44-0.64)	0.63 (0.49-0.75)	0.61 (0.46-0.74)	**0.97 (0.89-1.00)**	0.97 (0.93-1.0)
Normalize Vol. Diff.	0.53 (0.44-0.63)	0.64 (0.50-0.77)	0.32 (0.26-0.40)	**0.97 (0.90-1.00)**	**0.98 (0.95-1.0)**
ShapeDNA	**0.60 (0.50-0.70)**	0.53 (0.40-0.66)	0.53 (0.40-0.67)	0.72 (0.61-0.83)	0.27 (0.14-0.41)
Sphericity	0.50 (0.40-0.60)	0.52 (0.38-0.65)	0.36 (0.22-0.51)	0.83 (0.73-0.91)	0.85 (0.71-0.97)
Compactness	0.50 (0.40-0.61)	0.51 (0.37-0.65)	0.36 (0.22-0.50)	0.82 (0.73-0.91)	0.85 (0.71-0.97)
Quadratic Compactness	0.50 (0.40-0.60)	0.51 (0.38-0.65)	0.36 (0.24-0.50)	0.82 (0.72-0.91)	0.85 (0.69-0.97)
Elongation	0.55 (0.45-0.65)	0.61 (0.47-0.73)	0.56 (0.44-0.66)	0.82 (0.72-0.91)	0.80 (0.68-0.92)
Flatness	0.53 (0.42-0.63)	0.46 (0.33-0.58)	0.49 (0.33-0.65)	0.83 (0.72-0.92)	0.82 (0.64-0.95)
Spherical Disproportion	0.50 (0.39-0.60)	0.48 (0.33-0.62)	**0.65 (0.49-0.79)**	0.17 (0.08-0.27)	0.14 (0.02-0.29)
Surface Volume Ratio	0.47 (0.36-0.57)	0.43 (0.29-0.57)	0.57 (0.40-0.73)	0.01 (0.00-0.02)	0.01 (0.00-0.02)
Maximum 2D Diameter	0.47 (0.36-0.57)	0.59 (0.47-0.71)	0.47 (0.36-0.60)	0.72 (0.59-0.84)	0.85 (0.72-0.95)
Maximum 3D Diameter	0.48 (0.38-0.58)	0.53 (0.41-0.66)	0.60 (0.47-0.72)	0.66 (0.54-0.78)	0.76 (0.63-0.87)
Major Axis	0.51 (0.40-0.61)	0.53 (0.40-0.65)	0.56 (0.42-0.70)	0.55 (0.42-0.68)	0.51 (0.33-0.57)
NORHA	**0.57 (0.49-0.66)**	**0.67 (0.57-0.76)**	**0.65 (0.53-0.77)**	**0.92 (0.88-0.96)**	**0.98 (0.95-1.00)**

within this region. On the other hand, individuals suffering from HSR and HSL showing severe deterioration are much better isolated from the healthy point cloud, clustered together at the boundaries of this area. Figure 9 (b) depicts the same t-SNE representation but with samples colored according to its associated NORHA index, as predicted by the proposed method. An alternative representation with their corresponding meshes is shown in Figure 9 (c), including three exemplary cases corresponding to a normal subject, an individual with MCI and one with HSL.

4 Discussion

4.1 Hippocampal asymmetry characterization

The hippocampus is involved in learning, memory function and navigation (Marshall and Born, 2007). Structurally, hippocampi are usually slightly asymmetric, even for healthy individuals. Rightward asymmetry is consistently described in mammals based on ex-vivo examinations and MR images (Pedraza et al, 2004). The anterior part of the right hippocampus is usually larger than the left one in healthy individuals, and this asymmetry increases with age (Lucarelli et al, 2013). On the other hand, some neurodegenerative conditions are known to manifest through changes in the normal asymmetry patterns of the hippocampus. Hence, unilateral hippocampal sclerosis (HS, also called mesial temporal sclerosis) is the most common pathological feature of temporal lobe epilepsy (Thom, 2014). These findings, alongside clinical and neurophysiological data, are of great prognostic value in identifying good candidates for epilepsy surgery. Hippocampal asymmetry has also been associated with mild cognitive impairment and Alzheimer's disease (Shi et al, 2009). Some chronic neurodegenerative diseases also show temporal and hippocampal volume loss, usually associated with variable cognitive decline in cases of MCI or AD (Williams et al, 2013). However, hippocampal atrophy is described bilaterally in both conditions, more prominently in AD and preferentially on the left side of MCI patients, suggesting an inverted pattern compared with the hippocampus of normal subjects (Geroldi et al, 2000). Determining when

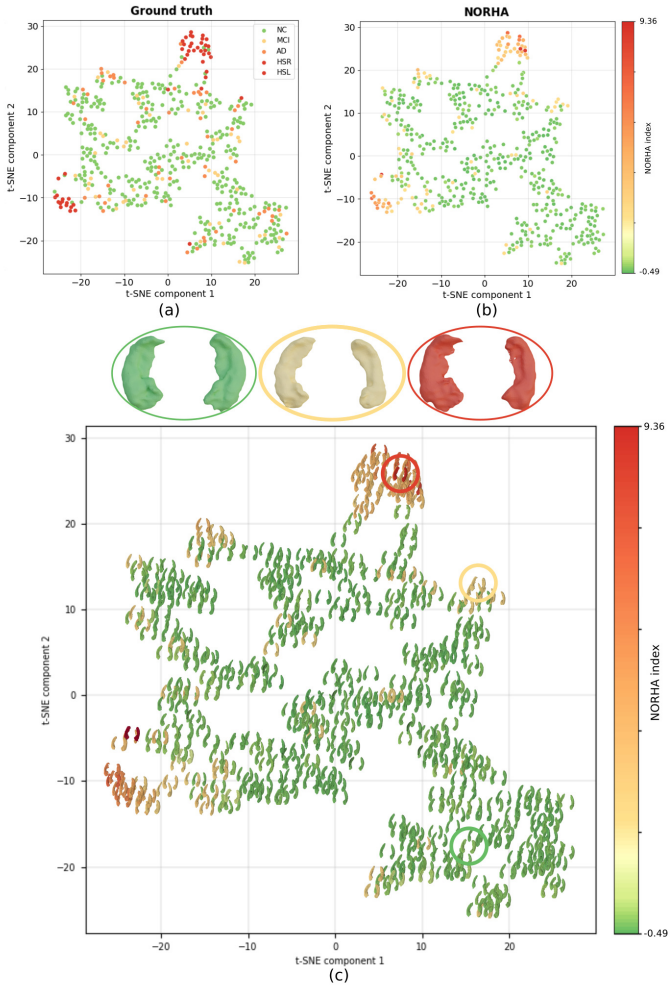

Fig. 9 Distribution of samples in a 2D t-distributed stochastic neighbor embedding (t-SNE) representation of their associated 14-dimensional feature vectors: (a) color-coded by their associated ground truth diagnostic, (b) color-coded by their NORHA index, (c) including the hippocampal meshes color-coded by their NORHA index and three exemplary cases.

differences in asymmetry start to be pathological is challenging by simply visually assessing them in brain MRI. While this may account for high diagnostic accuracy in temporal lobe epilepsy, atypical cases of this disease or other conditions in which the morphometric alterations are too subtle are more difficult to spot (Sun et al, 2017). Moreover, normal anatomical variants in the hippocampal formations, such as incomplete rotation, hippocampal fissure and arachnoid cysts, are usually described and may account for additional diagnostic difficulties (Connor et al, 2004). Furthermore, bilateral conditions affecting both hippocampi may be detected by quantitative analyses when compared with a normal database, although subtle and unilateral modifications are less

identifiable and rarely investigated. It should be noted that in those diseases that manifest with generalized atrophy, although both hippocampi decrease in size, the R>L asymmetry can still remain. (Shi et al, 2009) performed a meta-analysis on hippocampal asymmetry in healthy older adults, people with mild cognitive impairment, and people with Alzheimer's disease and found right > left asymmetry in all groups. However, adults with MCI had more pronounced asymmetry compared to healthy adults and those with AD (Shi et al, 2009). Classic quantitative MRI analysis approaches are focused on assessing the volume of specific structures or subregions, but asymmetry comparisons between hemispheres remain scarce and may be better suited to define some clinical conditions with atypical patterns of hippocampal involvement.

In this paper, we propose an alternative approach based on an automated index that captures the morphological differences between a subject's hippocampi using multiple shape descriptors and then indicates how much the input sample deviates from a normal population. By definition, our NORHA index is able to capture not only pathological asymmetries but also symmetries that are abnormal or that differ from those observed in the training set. This is achieved by means of the OC-SVM model (Schölkopf et al, 1999), a novelty detection algorithm that only requires normal samples to learn the space of common differences between left and right hippocampi during training. This setting distinguishes our method from other alternative proposals such as (Fu et al, 2021), as it does not explicitly model the differences in asymmetry produced by a given condition but the ones that are anatomically expected in a normal scenario. Hence, any sample data can be presented to the model to automatically assess how much this sample differs or not from the healthy cohort, avoiding any biases due to training with a specific disease.

To reduce the burden of manually annotating the hippocampal regions, which is prone to inter-observer variability plus is extremely time-consuming to be done in practice, we applied an automated process based on first preprocessing the input MR scans and then segmenting both hippocampi using a convolutional neural network (Goubran et al, 2020). Therefore, our approach can be used to objectively quantify hippocampal asymmetries without requiring a laborious effort from the operator. Notice, however, that segmentation errors could significantly affect the algorithm, as shown in the outliers depicted in Figure 7, which are captured by NORHA as highly abnormal cases.

The empirical evaluation of our NORHA index indicates that it can effectively characterize normal differences, mapping control subjects to indices that are tightly distributed in low values, even for different datasets (Figure 4). This behavior was expected, as any given healthy sample should be recognized as such by the model. Taking individual differences between morphological features such as volume or more complex characteristics empirically showed to be inadequate to capture the space of normal asymmetries compared to our index (Table 4). Instead, NORHA combines all these features into a common representation space that harnesses all their individual potential to characterize normality into a single, more discriminative value.

When analyzing the index, we observed that it positively correlates with the degree of unilateral deterioration of a hippocampal structure. In particular, the predicted values are significantly increased with respect to the normal cohort in cases with either right or left-sided hippocampal sclerosis. This behavior is also aligned with what is expected for this condition, which is known to be manifested through a high degree of asymmetry between contralateral anatomical structures (Cook et al, 1992). This effectiveness is also reflected in the ROC analysis from Figure 8, where the index reported high AUC values when applied to detect this condition.

Asymmetry estimates obtained from patients with MCI or AD also showed to be higher than normal, although the differences were not statistically significant and resulted in less accurate discrimination (Figures 4 and 8). This could be a consequence of our model not being able to capture bilateral deteriorations, as it is explicitly tailored to quantify unilateral asymmetries. The heterogeneity of the MCI group may result in other possible explanations as well, as some cases may mimic normal samples by showing no changes in asymmetry, while others do present asymmetry changes, as observed in Figure 9. This aligns with previous observations in this particular dataset Nettiksimmons et al (2014), where it was shown that certain subgroups exhibit hippocampal characteristics that contradict the disease trajectory explained by other biomarkers. In Figures 4 and 8 it can also be seen that AD patients from the ADNI and OASIS test sets showed slight deviations in their index distribution. This could be explained either by the differences between datasets or by the fact that ADNI can contain more AD cases with unilateral hippocampal atrophies.

To discard any potential age and sex-related biases in the index, we performed a controlled experiment comparing its distribution between subgroups of different confounding factors. Unfortunately, the low number of samples available for certain groups does not allow us to draw statistical conclusions from all of them. Nevertheless, as shown in Figure 6, our index distributes approximately similarity between age and sex subgroups. This could be thanks to our training set covering a wide range of ages and an approximately similar proportion of males/females. Future work could focus on specializing the index to individual subgroups, e.g. by including these potential confounding factors as input features to the model.

4.2 Potential clinical applications

Our method provides an objective, less observer-dependent way to quantify hippocampal asymmetries that can be used not only to understand differences in normal subjects but to measure anatomical deviations that can be eventually considered abnormal. In this paper, we analyzed common scenarios that are characterized by changes in the hippocampal shape. As an exemplary use case, we analyzed the correlation between our predictions and the CDR-SB, ADAS11 and MMSE cognitive assessments (Figure 7). Interestingly, a weak correlation was found between the measured asymmetries using our

index and the decline in cognitive abilities of the analyzed subjects. This suggests that our index can partially account for the cognitive deterioration observed in the analyzed population. There is a general idea about a border zone of impairment between age-related cognitive decline, mild cognitive impairment and clinically diagnosed dementia (Shi et al, 2009). The search for an imaging biomarker that sheds light on this gray area has been the goal of numerous investigations (McDonald et al, 2000; Sun et al, 2017). The results obtained using NORHA could pave the way towards future large-scale clinical studies applying it to comprehensively analyze its behavior in borderline cases. Nevertheless, this analysis should be carefully considered, taking into account that the cognitive assessments are known to suffer from inter-observer variability (Tractenberg et al, 2001).

We envision that our NORHA index, combined with other noninvasive techniques, may help to diagnose patients with epilepsy and chronic depression (Bernasconi et al, 2003; Mervaala et al, 2000). Furthermore, the design of our NORHA index is sufficiently general to be applied for characterizing deviations in normal asymmetries of other brain regions, which can be helpful, e.g. to detect malformations in brain development due to other epilepsies or to study changes in temporal structures such as the amygdala for autism spectrum disorder (Richards et al, 2020).

4.3 Limitations

Our study should still be read in the context of its own limitations. In this first proof-of-concept analysis, each hippocampus was characterized using a large set of morphological features, including volumetric measurements, ShapeDNA descriptors and other shape properties. Empirical observations demonstrate that utilizing the complete set of characteristics provides better results compared to using individual subsets (as illustrated in Figure 8). However, the method lacks enough sensitivity to be used as an independent tool for diagnosing MCI and AD. This cannot be attributed to the OC-SVM itself, as in Figure 9 we can observe that the selected features are not enough to differentiate these cases from control subjects as notoriously as for HSL/HSR. From a characterization perspective, this could be either a consequence of having some cases in our dataset with bilateral atrophies or our set of features being suboptimal to capture other asymmetry patterns, as previously mentioned in Section 4.1. Automatically learning the shape asymmetries using deep learning methods could potentially alleviate this limitation (LeCun et al, 2015). On the other hand, it should be noted that our approach is focused specifically on differences in asymmetry and not on other structural characteristics of the hippocampus that can be changed by neurodegenerative conditions. For crafting an effective AD diagnostic tool, it would be necessary to integrate complementary shape information beyond left and right differences and also characteristics from other structures also influenced by AD. NORHA could be used and integrated as one of the features in such a system. Nevertheless, notice that this is out of the scope of this paper. Our evaluation of diseased

cases is intended to empirically understand the ability of our index to identify disease-related deviations from normality.

Another influential factor is the quality of the segmentation results. To filter out these cases, we applied a heuristic process that is able to discard clearly wrong meshes. However, we still observed three outliers in the MCI group that were captured by NORHA as potentially abnormal individuals. This further emphasizes the necessity of having accurate brain parcellation models to effectively capture the actual shape of this particular structure.

Another limitation is related to the degree of explainability of our model. In its current form, our method can only provide a value indicating how much the sample deviates from the normal hippocampal asymmetry without indicating which features were the ones that mostly deviated from normality. This is a consequence of choosing an OC-SVM with an RBF kernel, which renders a model in which the coefficients are applied on a transformed, non-invertible feature space. We experimentally studied an OC-SVM with a linear kernel to solve this issue, yet results were significantly less accurate than those obtained when using the RBF. This could also be solved by replacing the hand-crafted features with a deep neural network and then integrating explainability techniques such as GradCAM (Selvaraju et al, 2017).

5 Conclusions

In this study, we present a first proof-of-concept approach for automatically predicting a hippocampal normal asymmetry index from structural brain MRI. Our model can automatically characterize the asymmetries in the hippocampi without requiring manual segmentation. Furthermore, as it is trained using only healthy subjects, it can effectively quantify the normal anatomical differences without being biased towards a specific disease. On the other hand, this approach is general enough to be used with other symmetrical cortical structures. We believe this model can be used in multiple applications, including anatomical studies analyzing normal cohorts–e.g. to understand structural differences in left and right-handed individuals–, or even for discovering potential associations in unilateral anatomical shape deviations and neurodegenerative conditions. Future work will focus on automatically learning the morphological features using convolutional neural networks and introducing qualitative feedback regarding which hippocampus areas are more deviated.

Acknowledgments. This study was partially funded by PIP GI 2021-2023 - 11220200102472CO (CONICET, Argentina) and PICT 2016-0116 (Agencia I+D+i, Argentina). E.I. was supported by an Einstein AI grant from Salesforce company and a PIP 2021-2023 from CONICET. This work was also partially supported by the computing facilities of Extremadura Research Centre for Advanced Technologies (CETA-CIEMAT).

Declarations

Conflict of Interest: The authors declare that they have no conflict of interest.
Patient Data: This articles does not contain patient data.

References

Bernasconi N, Bernasconi A, Caramanos Z, et al (2003) Mesial temporal damage in temporal lobe epilepsy: a volumetric mri study of the hippocampus, amygdala and parahippocampal region. Brain 126(2):462–469

Connor S, Ng V, McDonald C, et al (2004) A study of hippocampal shape anomaly in schizophrenia and in families multiply affected by schizophrenia or bipolar disorder. Neuroradiology 46(7):523–534

Cook M, Fish D, Shorvon S, et al (1992) Hippocampal volumetric and morphometric studies in frontal and temporal lobe epilepsy. Brain 115(4):1001–1015

Dawson-Haggerty et al (2019) trimesh. URL https://trimsh.org/

Dwivedi AK, Mallawaarachchi I, Alvarado LA (2017) Analysis of small sample size studies using nonparametric bootstrap test with pooled resampling method. Statistics in medicine 36(14):2187–2205

Ezzati A, Zammit AR, Habeck C, et al (2020) Detecting biological heterogeneity patterns in adni amnestic mild cognitive impairment based on volumetric mri. Brain imaging and behavior 14:1792–1804

Folstein MF, Folstein SE, McHugh PR (1975) "mini-mental state": a practical method for grading the cognitive state of patients for the clinician. Journal of psychiatric research 12(3):189–198

Fu Z, et al (2021) Altered neuroanatomical asymmetries of subcortical structures in subjective cognitive decline, amnestic mild cognitive impairment, and alzheimer's disease. J Alzheimer's Dis 79(3):1121–1132

Geroldi C, et al (2000) Apolipoprotein e genotype and hippocampal asymmetry in alzheimer's disease: a volumetric mri study. J Neurol Neurosurg Psychiatry 68(1):93–96

Goubran M, et al (2020) Hippocampal segmentation for brains with extensive atrophy using three-dimensional convolutional neural networks. Tech. rep., Wiley Online Library

LaMontagne PJ, et al (2019) Oasis-3: longitudinal neuroimaging, clinical, and cognitive dataset for normal aging and alzheimer disease. MedRxiv

LeCun Y, Bengio Y, Hinton G (2015) Deep learning. nature 521(7553):436–444

Lian Z, et al (2013) A comparison of methods for non-rigid 3d shape retrieval. Pattern Recognit 46(1):449–461

Lister JP, et al (2006) Asymmetry of neuron numbers in the hippocampal formation of prenatally malnourished and normally nourished rats: a stereological investigation. Hippocampus 16(11):946–958

Lucarelli R, et al (2013) Mr imaging of hippocampal asymmetry at 3t in a multiethnic, population-based sample: results from the dallas heart study. AJNR Am J Neuroradiol 34(4):752–757

Lynch C, et al (2006) The clinical dementia rating sum of box score in mild dementia. Dement Geriatr Cogn Disord 21(1):40–43

Van der Maaten L, Hinton G (2008) Visualizing data using t-sne. JMLR 9(11)

Manjón JV, Coupé P (2016) volbrain: an online mri brain volumetry system. Frontiers in neuroinformatics 10:30

Marshall L, Born J (2007) The contribution of sleep to hippocampus-dependent memory consolidation. TiCS 11(10):442–450

McDonald B, et al (2000) Anomalous asymmetry of fusiform and parahippocampal gyrus gray matter in schizophrenia: a postmortem study. Am J Psychiatry 157(1):40–47

Mervaala E, et al (2000) Quantitative mri of the hippocampus and amygdala in severe depression. Psychol 30(1):117–125

Moodley K, Chan D (2014) The hippocampus in neurodegenerative disease. The hippocampus in clinical neuroscience 34:95–108

Mueller SG, et al (2005) The alzheimer's disease neuroimaging initiative. Neuroimaging Clin N Am 15(4):869–877

Nettiksimmons J, DeCarli C, Landau S, et al (2014) Biological heterogeneity in adni amnestic mild cognitive impairment. Alzheimer's & Dementia 10(5):511–521

Andrade de Oliveira A, Carthery-Goulart MT, Oliveira Júnior PPdM, et al (2015) Defining multivariate normative rules for healthy aging using neuroimaging and machine learning: an application to alzheimer's disease. Journal of Alzheimer's Disease 43(1):201–212

O'Bryant SE, et al (2008) Staging dementia using clinical dementia rating scale sum of boxes scores: a texas alzheimer's research consortium study. Arch

Neurol 65(8):1091–1095

Pedraza O, Bowers D, Gilmore R (2004) Asymmetry of the hippocampus and amygdala in mri volumetric measurements of normal adults. J Int Neuropsychol Soc 10(5):664–678

Pedregosa F, et al (2011) Scikit-learn: Machine learning in Python. JMLR 12:2825–2830

Reuter M, Wolter FE, Peinecke N (2006) Laplace–beltrami spectra as 'shape-dna'of surfaces and solids. Comput Aided Des 38(4):342–366

Richards R, et al (2020) Increased hippocampal shape asymmetry and volumetric ventricular asymmetry in autism spectrum disorder. NeuroImage Clin 26:102,207

Rosen WG, Mohs RC, Davis KL (1984) A new rating scale for alzheimer's disease. The American journal of psychiatry

Saad ZS, Reynolds RC, Argall B, et al (2004) Suma: an interface for surface-based intra-and inter-subject analysis with afni. In: 2004 2nd IEEE ISBI: Nano to Macro (IEEE Cat No. 04EX821), IEEE, pp 1510–1513

Schölkopf B, Williamson RC, Smola AJ, et al (1999) Support vector method for novelty detection. In: NIPS, Citeseer, pp 582–588

Selvaraju RR, Cogswell M, Das A, et al (2017) Grad-cam: Visual explanations from deep networks via gradient-based localization. In: Proc. IEEE Int. Conf. Comput. Vis., pp 618–626

Shi F, Liu B, Zhou Y, et al (2009) Hippocampal volume and asymmetry in mild cognitive impairment and alzheimer's disease: Meta-analyses of mri studies

Sun Y, Chen Y, Collinson SL, et al (2017) Reduced hemispheric asymmetry of brain anatomical networks is linked to schizophrenia: a connectome study. Cereb Cortex 27(1):602–615

Team RC, et al (2013) R: A language and environment for statistical computing. The Journal of Open Source Software

Thom M (2014) Hippocampal sclerosis in epilepsy: a neuropathology review. Neuropathol Appl Neurobiol 40(5):520–543

Tractenberg RE, Schafer K, Morris JC (2001) Interobserver disagreements on clinical dementia rating assessment: interpretation and implications for training. Alzheimer Dis Assoc Disord 15(3):155–161

Vergara IA, Norambuena T, Ferrada E, et al (2008) Star: a simple tool for the statistical comparison of roc curves. BMC Bioinform 9(1):1–5

Wachinger C, Salat DH, Weiner M, et al (2016) Whole-brain analysis reveals increased neuroanatomical asymmetries in dementia for hippocampus and amygdala. Brain 139(12):3253–3266

Williams MM, Storandt M, Roe CM, et al (2013) Progression of alzheimer's disease as measured by clinical dementia rating sum of boxes scores. Alzheimers Dement 9(1):S39–S44