Automatic recognition of specific local cortical folding patterns

Léonie Borne, Denis Rivière, Arnaud Cachia, Pauline Roca, Charles Mellerio, Catherine Oppenheim, Jean-François Mangin

NeuroImage 238 (2021) 118208

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

Automatic recognition of specific local cortical folding patterns

Léonie Borne[a,b,*], Denis Rivière[a], Arnaud Cachia[c,d], Pauline Roca[d,e,f], Charles Mellerio[d,e,g], Catherine Oppenheim[d,e], Jean-François Mangin[a]

[a] Université Paris-Saclay, CEA, CNRS, Baobab, Neurospin, Gif-sur-Yvette, France
[b] University of Newcastle, HMRI, Systems Neuroscience Group, NSW, Australia
[c] Université de Paris, LaPsyDÉ, CNRS, Paris, France
[d] Université de Paris, Institut de Psychiatrie et Neurosciences de Paris (IPNP), INSERM, UMR S1266, Paris, France
[e] Groupe Hospitalier Universitaire Paris Psychiatrie et Neurosciences, Sainte-Anne Hospital, Imaging Department, Paris, France
[f] Pixyl, Research and Development Laboratory, Grenoble, France
[g] Centre d'imagerie du Nord, Saint Denis, France

ARTICLE INFO

Keywords:
Cortical sulci
Pattern recognition
Machine learning
Supervised learning
Classification
Convolution neural network

ABSTRACT

The study of local cortical folding patterns showed links with psychiatric illnesses as well as cognitive functions. Despite the tools now available to visualize cortical folds in 3D, manually classifying local sulcal patterns is a time-consuming and tedious task. In fact, 3D visualization of folds helps experts to identify different sulcal patterns but fold variability is so high that the distinction between these patterns sometimes requires the definition of complex criteria, making manual classification difficult and not reliable. However, the assessment of the impact of these patterns on the functional organization of the cortex could benefit from the study of large databases, especially when studying rare patterns. In this paper, several algorithms for the automatic classification of fold patterns are proposed to allow morphological studies to be extended and confirmed on such large databases. Three methods are proposed, the first based on a Support Vector Machine (SVM) classifier, the second on the Scoring by Non-local Image Patch Estimator (SNIPE) approach and the third based on a 3D Convolution Neural Network (CNN). These methods are generic enough to be applicable to a wide range of folding patterns. They are tested on two types of patterns for which there is currently no method to automatically identify them: the Anterior Cingulate Cortex (ACC) patterns and the Power Button Sign (PBS). The two ACC patterns are almost equally present whereas PBS is a particularly rare pattern in the general population. The three models proposed achieve balanced accuracies of approximately 80% for ACC patterns classification and 60% for PBS classification. The CNN-based model is more interesting for the classification of ACC patterns thanks to its rapid execution. However, SVM and SNIPE-based models are more effective in managing unbalanced problems such as PBS recognition.

1. Introduction

The surface of the brain is divided into many convolutions, called gyri, delimited by folds, called sulci. The study of these cortical folds presents two major interesting aspects. First, the morphometry of the cortical folding patterns can be performed with a number of geometric features like depth, length or widening. These features change during the whole life span, for instance during adolescence because of myelin development (Alemán-Gómez et al., 2013), or during aging because of atrophy (Le Guen et al., 2019). Therefore, they are very useful to define biomarkers of pathologies occurring late in life (Bertoux et al., 2019). However, this article does not address this aspect of cortical fold morphometry but focuses on the second aspect which concerns their spatial organisation. Indeed, most of the folding process occurs during the second half of the pregnancy and the spatial pattern of some cortical sulci do not change after birth (Cachia et al., 2016; Im and Grant, 2019). Therefore, they can be used to detect early events at the origin of a specific local brain function. For instance, several studies have reported a correlation between visually recognized sulcal patterns and some cognitive characteristics — e.g. cognitive control efficiency (Cachia et al., 2014; Fornito et al., 2004) or handedness (Sun et al., 2012) - and neuropsychiatric illnesses - e.g. epilepsy (Mellerio et al., 2014) or schizophrenia (Plaze et al., 2015).

The impact of this anatomical characteristic on brain function is still poorly understood today. Indeed, the number of works focusing on the variability of the cortical folding pattern is rather low, probably because

Abbreviations: ACC, Anterior Cingulate Cortex; PBS, Power Button Sign; CNN, Convolutional Neural Network; ROI, Region Of Interest; SNIPE, Scoring by Non-local Image Patch Estimator; SVM, Support Vector Machine.

* Corresponding author at: Hunter Medical Research Institute (HMRI), Lot 1, Kookaburra Cct, New Lambton Heights, NSW, 2305, Australia.
E-mail address: leonie.borne@gmail.com (L. Borne).

https://doi.org/10.1016/j.neuroimage.2021.118208.
Received 1 March 2021; Received in revised form 30 April 2021; Accepted 25 May 2021
Available online 2 June 2021.

the methodological paradigm of the field of neuroimaging, spatial normalization, is not very useful for this kind of exploration. Aligning different folding patterns is not necessarily helping to disentangle them. Therefore, classifying local patterns often requires browsing simultaneously across each individual subject to try to understand how individuals differ. In addition, to discover whether different folding patterns correspond to different functional architecture, a variety of folding configurations have to be matched with a variety of noisy individual functional MRI maps. This difficult task has been carried out in several studies. Tight relationships between folding patterns and functional architecture have been discovered (Amiez et al., 2006; Amiez and Petrides, 2009; 2014; Bodin et al., 2018; Crosson et al., 1999; Derrfuss et al., 2012; Grosbras et al., 1999; Lopez-Persem et al., 2019; Segal and Petrides, 2013; Watson et al., 1993; Weiner et al., 2014; Zlatkina et al., 2016). Such results suggest that each local folding pattern may be the indicator of a specific functional organization. This predictive power would result from the impact of architecture on the dynamics of folding during development (Foubet et al., 2019; Llinares-Benadero and Borrell, 2019; Mangin et al., 2019; Van Essen, 1997). Our goal in this paper is to deliver a tool helping to exploit massively the rare findings of the past, using already annotated databases of patterns. We hope that this will motivate the search for new patterns of interest, and later on the development of new methodologies to automate this quest.

In this article, we propose and compare three different methods allowing the automatic recognition of folding patterns, based on the 3D organization of sulci. The input to our algorithm is a negative print of the cortex entering the folds to their bottom. The methods are tested on two types of sulcal patterns for which there is currently no method to automatically label them: first, the classification of the Anterior Cingulate Cortex (ACC) patterns, and second, the recognition of a specific pre-central pattern called Power Button Sign (PBS).

From a clinical point of view, the choice of these two types of sulcal patterns is justified by previous morphological studies demonstrating their links with brain pathologies and cognition. **Regarding the ACC patterns,** a link has been demonstrated between these patterns and schizophrenia: schizophrenic subjects tend to have symmetric patterns between their right and left hemispheres, unlike controls (Le Provost et al., 2003; Yücel et al., 2003). Later, it was also shown that children with the same lack of asymmetry have poor inhibitory control (Borst et al., 2014; Cachia et al., 2014). Thus this pattern is not only related to a psychiatric disease but also to the healthy brain function. **Regarding the PBS,** this pattern is related to epilepsy. This disease is often due to a focal cortical dysplasia, which corresponds to a disruption of the early stages of embryogenesis leading to abnormal cortical lamination and the presence of dysmorphic neurons. This dysplasia is sometimes linked to abnormal fold patterns (Besson et al., 2008; Mellerio et al., 2014; Régis et al., 2011; Roca et al., 2015), even when the dysplasia is invisible on MRI. Here, we are interested in the automatic detection of one of these patterns: the PBS, present in about 60% of patients with a dysplasia in the vicinity of the motor area, whereas in the general population, this pattern is so rare that we do not know its frequency (Mellerio et al., 2014).

From a computational point of view, the identification of these two types of sulcal patterns is a challenging task because of their complex structure and great variability. While a template pattern can be drawn or inferred from an averaging process, the actual shapes in individual subjects may correspond to complex distortions of this template. **Regarding the sulcal pattern of the ACC,** it can be classified with structural MRI (Cachia et al., 2016; 2014) in two different types (Fig. 1): a “single” type, with only the cingulate sulcus, and the “double parallel” type, with an additional paracingulate sulcus. Automating the recognition of these patterns is particularly difficult because of their high variability (Fig. 2). For this purpose, we have a database of 348 hemispheres, including 207 with the double parallel pattern, of which the cingulate and paracingulate sulci have been manually labeled, by one of the team investigating the meaning of these alternative patterns. Although the size of the

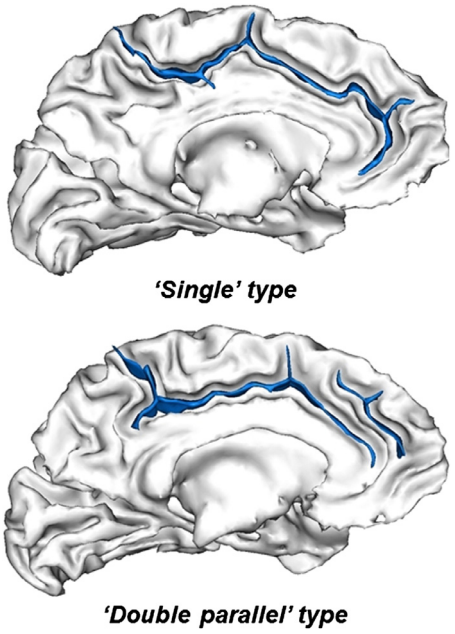

Fig. 1. Two folding patterns identified in the cingulate region. The “single” pattern is characterized by the presence of the cingulate sulcus alone. The “double parallel” pattern has an additional paracingulate sulcus, which must be parallel to the cingulate sulcus and long enough to be identified. Figure reprinted from (Cachia et al., 2014).

paracingulate sulcus is variable, precise criteria on its size were used in order to robustly distinguish the two patterns (i.e. with or without a paracingulate sulcus). **Regarding the PBS,** this pattern is characterized by the interposition of a precentral sulcal segment between the central sulcus and one of its hook-shaped anterior ascending branches (Fig. 3). Automating PBS recognition is particularly challenging because, in addition to the extreme variability of this pattern (Fig. 4), the training database is limited (19 controls and 38 patients with dysplasia, i.e. 114 hemispheres) and the proportion of hemispheres with a PBS is particularly unbalanced (22 subjects, including 1 control and 21 patients with dysplasia, have at least one of the two hemispheres with PBS, i.e. 28 hemispheres with PBS in the database). Thus, these two types of patterns allow us to see the performances of the proposed methods in two fundamentally different cases: for one, the training database is large and balanced, for the other, it is small and unbalanced, which is a recurrent problem when studying rare patterns.

Many other cortical folding patterns exist (Germann et al., 2005; Hotier et al., 2017; Huntgeburth and Petrides, 2012; Kim et al., 2008; Ochiai et al., 2004; Plaze et al., 2015; 2009; Segal and Petrides, 2012; Snyder et al., 2019; Sprung-Much and Petrides, 2018; Zlatkina and Petrides, 2010) and the methods proposed in this article are generic enough to be applied to them. Note that the two chosen patterns are not considered as more interesting than these other patterns but our previous studies had delivered the two annotated datasets used in this article, providing the opportunity to design an automatic method to use the findings on a larger scale. More importantly, these two patterns have been chosen because they are instances of the two very different kinds of situations faced when exploring the cortical folding patterns: 1) the stratification of the whole population using a small number of patterns, two subgroups in the case of ACC; 2) the detection of a rare pattern, usually considered has the signature of a rare developmental event, in the case of the PBS. Nevertheless, it should be noted that while numerous sets of alternative local folding patterns have been described in the literature, few of them have been linked with pathologies as PBS and ACC. The framework proposed in this article to allow the automatic identification of folding patterns that have previously been manually annotated

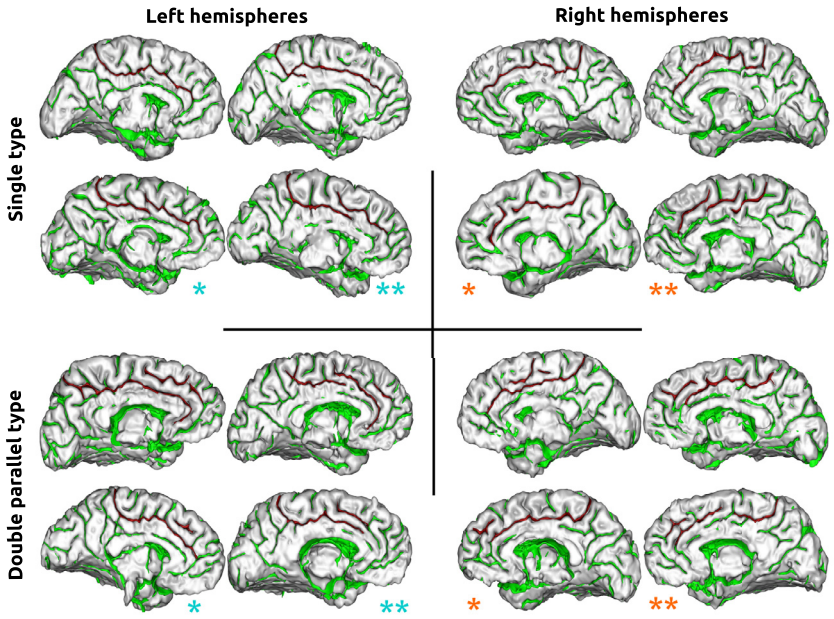

Fig. 2. Illustration of ACC patterns variability. The upper part of the table shows the single patterns and the lower part shows the double parallel patterns. We observe that the position, size and number of components of the paracingulate sulcus are variable. It is sometimes difficult to distinguish with the naked eye between a large paracingulate sulcus and a double cingulate sulcus (see **). Additionally, small paracingulate sulci can sometimes be confused with other small sulci (see *). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

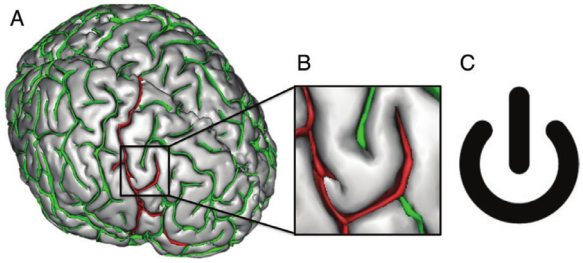

Fig. 3. Typical Power Button Sign (PBS): a fold pattern related to Type 2 Focal Cortical Dysplasia. (A), Sulcal graph on a 3D reconstruction of the brain surface of a 17-year-old male patient shows a typical PBS, with, (B), magnification of the PBS on (A). The PBS, by analogy to, (C), the power symbol, corresponds to interposition of a precentral segment (green) between the central sulcus (red) and a hook-shaped anterior ascending branch (also red). Figure reprinted from (Mellerio et al., 2014). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

on a training set, consists of four main steps: (1) the cortical folds are automatically represented in 3D from structural MRIs, (2) the patterns are manually identified and labeled from this representation in order to extract the features necessary for (3) the training of different machine learning algorithms allowing (4) the automatic classification of the folding patterns. Note that model training requires steps (1), (2) and (3), while model application only requires steps (1) and (4) which are fully-automatic. Three different machine learning approaches are proposed: the first one based on a Support Vector Machine (SVM) classifier, the second one on the Scoring by Non-local Image Patch Estimator (SNIPE) proposed in (Coupé et al., 2012) and the last one on 3D Convolutional Neural Network (CNN).

2. Databases

The Morphologist pipeline of the BrainVISA toolbox, freely available at www.brainvisa.info, was used to segment and provide 3D graphical representation from raw MRI data. This 3D reconstruction of cortical folds was used to manually visualize and label brains. Intra- and inter-rater reproducibility of ACC and PBS pattern labelling have been described previously in (Cachia et al., 2016) and (Mellerio et al., 2014) respectively. The two databases studied are described below.

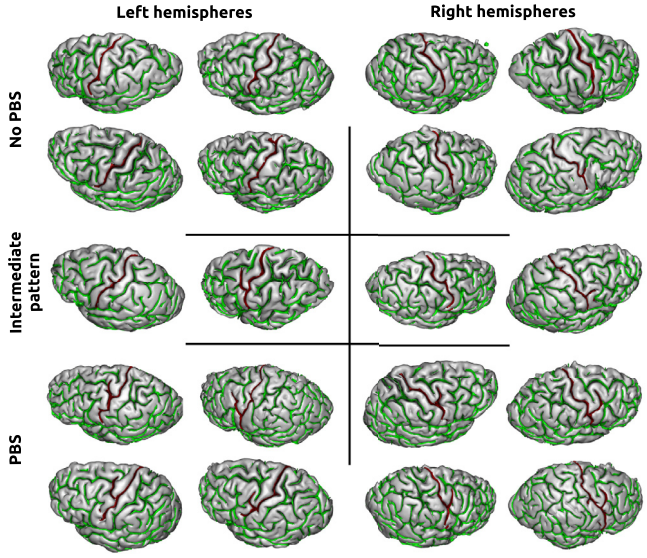

Fig. 4. Illustration of PBS variability. The first two lines represent hemispheres without PBS and the last two lines with PBS. The middle line represents hemispheres with an intermediate pattern. In this study these intermediate patterns are considered as PBS. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2.1. Anterior cingulate cortex (ACC) pattern

The T1w MRIs of 174 subjects (including 71 healthy controls, 59 schizophrenia patients and 44 healthy siblings) were selected from the study detailed in (Chakravarty et al., 2015). The 71 healthy controls (45 men) had a median age of 21 years (minimum 10; maximum 33). The 59 schizophrenia patients (39 men) had a median age of 21 years (minimum 13; maximum 32). The 44 healthy siblings (23 men) had a median age of 20 years (minimum 11; maximum 29). T1w MRI data were acquired with a 1.5-T MR system (Signa 1.5T; GE Healthcare, Milwaukee, Wis), using a 3D spoiled gradient echo with contiguous 1.5 mm slices in the axial plane (echo time, 5 ms; repetition time, 24 ms; flip angle,

45 (DEG); acquisition matrix, 256×192; number of excitations, 1; field of view, 24 cm).

The ACC sulcal pattern was visually identified for each individual hemisphere, following the procedure described in (Cachia et al., 2014) and previously used to label a part of this study's database, as published in (Cachia et al., 2016). The learning base was composed of 348 identified ACC sulcal patterns, including 141 "single" types and 207 "double parallel" types. The cingulate and paracingulate sulci were manually labeled by one of the team investigating the meaning of these patterns.

Further details on this database has been described previously (Cachia et al., 2016; Chakravarty et al., 2015; Giedd et al., 1999). The study was approved by the National Institute of Mental Health review board. Written informed consent was obtained from parents and controls older than 18 years, and written informed assent was obtained from minors.

2.2. Power button sign (PBS)

The T1w MRIs of 57 subjects (including 20 normal controls, 37 epileptic patients) were selected from the study detailed in (Mellerio et al., 2014). The 37 patients (18 men; 28 right handed) had a median age of 24 years (interquartile range, 19–32 years). The 20 control subjects (10 men; 20 right handed) had a median age of 26 years (interquartile range, 25–31 years). T1w MRI data were acquired by using a 1.5-T MR system (Signa 1.5T; GE Healthcare, Milwaukee, Wis) with an inversion-recovery 3D T1-weighted fast spoiled gradient-recalled acquisition (repetition time msec/echo time msec, 10/2; flip angle, 15°; acquisition matrix, 256 × 256; section thickness, 1.2 mm; no gap; in-plane resolution, 0.93×0.93 mm; acquisition time, 6 min 14 s).

The PBS was visually identified for each individual hemisphere according to the criteria described in the same study. In addition to the 15 PBS identified by the experts (including one belonging to a control subject), some hemispheres have been identified as having an intermediate pattern when they show a similar pattern which is not sufficiently pronounced to be classified as such. In order to slightly rebalance the number of samples in each class, intermediate patterns are considered as PBSs in this study. Among the 114 hemispheres in the database, 28 have the PBS (including 13 intermediate patterns, all belonging to epileptic subjects). We manually labeled the central sulcus, its ascending branch and the pre-central sulcus forming the PBS.

Further details on this database has been described previously (Mellerio et al., 2014). This study was found to conform to generally accept scientific principles and ethical standards by the Ethics Review Committee of Ile de France III. Written informed consent was waived for patients and was obtained from control subjects.

3. Method

In this section, the first part will be devoted to the method used to represent cortical folds from structural MRIs. Then, the second part will describe several algorithms that automatically classify the folding patterns from the folds representation described in the previous part. In this study, three different approaches are proposed: the first one based on a Support Vector Machine (SVM) classifier, the second one on the Scoring by Non-local Image Patch Estimator (SNIPE) proposed in (Coupé et al., 2012) and the last one on 3D Convolutional Neural Network (CNN). Finally, the third part will describe how the three obtained models are evaluated and compared.

The code implementing the methods described in this section is available here: https://github.com/brainvisa/morpho-deepsulci.

3.1. Fold representation

The BrainVISA/Morphologist pipeline, used to visualize the sulci in 3D and manually identify the patterns, is also used for data preprocessing (Mangin et al., 2004). This pipeline, applied on all subjects, i.e. 114 hemispheres for the PBS and 348 for the ACC patterns, represents the folds by a set of voxels corresponding to a skeleton of the cerebrospinal fluid filling the fold. The representation of the folds therefore corresponds to a negative mold of the brain. In addition to facilitating manual identification of sulcal patterns, the skeleton representation enables to normalize the data optimally. Indeed, the data are particularly influenced by the type of MRI sequence, the age of the subject (which has a significant impact on the opening of the sulci) or even pathologies. The fold skeletons are affinely registered in the MNI space (Collins et al., 1994). Note that we have preferred an affine registration to a non-linear registration because it could deform the folding patterns enough to make them unrecognisable (which can easily happen given their complex organisation) and that a global registration seems sufficient for the methods presented.

When manually classifying the ACC and PBS patterns, the same morphological criteria were used for the right and left hemispheres. Therefore, for the three approaches presented in this article, the right hemispheres have been flipped in order to be superimposable with the left hemispheres. This significantly increases the size of the training database. However, this manipulation is questionable when applying the methods presented in this article to other folding patterns, as it may actually decrease the performance of the models due to the natural hemispherical asymmetry that can unnecessarily increase the variability of the patterns.

With regard to ACC patterns, the single and double parallel types are respectively characterized by the absence and presence of the paracingulate sulcus. Thus, the objective is to detect the presence of the paracingulate to perform the classification. Similarly, for PBS, the algorithm must detect its presence in the area around the central sulcus. Paracingulate sulcus and PBS will be considered as the "patterns-of-interest". In both databases, the patterns-of-interest were extracted from manual labeling, as described in Fig. 5.

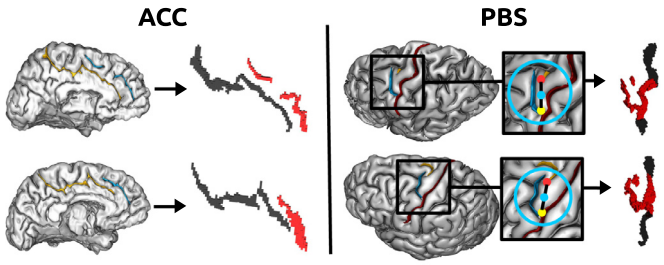

Fig. 5. Extraction of patterns-of-interest from manual labeling. The voxels extracted from the manual labels represent the pattern-of-interest in red and the sulcus to which it is attached in black. **Left:** Paracingulate sulcus, labeled in blue, corresponds to red voxels while cingulate sulcus, labeled in yellow, corresponds to black voxels. **Right:** The central sulcus is manually labeled in red, the ascending branch of the central sulcus in blue and the pre-central sulcus, wedged between the blue and red sulci, in yellow. As these manual labels are insufficient to properly delimit the PBS, the extraction of red voxels requires additional processing. 1. The yellow point corresponds to the intersection between the blue and red sulci. 2. The red point is the point of the yellow sulcus closest to the yellow point. 3. The blue point is in the centre of the segment joining the red and the yellow points. 4. The radius of the blue sphere, with the blue point at its centre, corresponds to $max(d, 20mm)$ with d being the distance between the blue point and the furthest point of the blue sulcus. 5. All the voxels contained in the blue sphere and belonging to the red/yellow/blue sulci define the red voxels. The black voxels correspond to the central sulcus voxels that do not belong to the red voxels. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

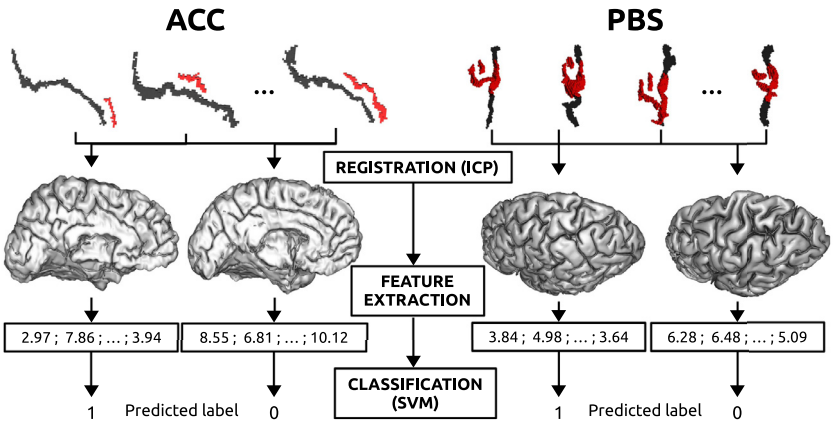

Fig. 6. Illustration of the SVM-based method. For each hemisphere of the training set with a paracingulate sulcus/PBS, the patterns-of-interest (in red) and their attached sulcus (in black) are registered on the fold skeleton of the hemispheres of the test set using the Iterative Closest Point (ICP) algorithm. Once they have been registered, the average distance (from the points of the pattern-of-interest in red to their nearest points in the hemisphere to be classified) is used as a feature for the SVM classifier. Note that only the hemispheres of the training set with the pattern-of-interest are used as features of the SVM classifier. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.2. Classification approaches

3.2.1. Support vector machine (SVM)

The idea of this first approach is to use the distances from the patterns-of-interest of the training database to the image to be classified to train a SVM classifier (Fig. 6).

Registration In order to realign the patterns-of-interest to the hemisphere to be classified, the Iterative Closest Point algorithm (Besl and McKay, 1992), with the robust implementation of Holz et al. (2015), is used. The voxels of the fold skeleton, extracted by the BrainVISA/Morphologist pipeline, are used as point clouds to perform the registration.

In order to constrain the registration of the patterns-of-interest, they are registered with their attached sulcus (i. e. cingulate sulcus for paracingulate sulcus, and central sulcus for PBS), which has been manually labeled. This constraint with large sulci increases the specificity of the pattern and prevents the registration of the patterns-of-interest to be trapped in a spurious local minimum.

Since the Iterative Closest Point algorithm is particularly sensitive to the initial positions of the point clouds to be registered, several initializations are compared in order to avoid local minima. From the initial position of the point clouds, the pattern to be registered and its attached sulcus are translated before applying the Iterative Closest Point algorithm. The amplitude of the translations tested is determined by inner cross-validation. The registration minimizing the distance from the pattern and its attached sulcus to the hemisphere to be classified is selected.

Feature extraction Once registered, the distances between the patterns-of-interest and the hemisphere to be classified are calculated. This distance corresponds to the average Euclidean distance from the voxels of the pattern-of-interest to their nearest voxels in the new hemisphere. The distance is low if the hemisphere to be classified also has the pattern-of-interest.

Classifier The calculated distances are used to train a Support Vector Machine (SVM) algorithm to classify hemispheres with or without the pattern-of-interest. A vector containing the distances to the hemisphere to be classified is used to train the classifier: each feature corresponds to a pattern-of-interest in the training database. Thus, this vector can contain up to 28 features for the PBS and 207 for the ACC. Note that during cross validations, the patterns-of-interest being in the group of subjects for testing are removed from the vector (e.g. for the 10-fold cross validation, the vector therefore contains 25 or 26 features for the PBS and 187 or 188 for the ACC). The Radial Basis Function is used as a kernel. Penalty parameter C of the error term and kernel coefficient γ are determined by inner cross validation. The SVM algorithm used corresponds to the one implemented in the scikit-learn library (Pedregosa et al., 2011). In this implementation, a weight w_i is assigned to each class i in order to calculate the value of the parameter C of the class i such as $C_i = w_i * C$. In order to compensate for the potential imbalance of the classes in the input data, these weights are automatically adjusted to be inversely proportional to class frequencies. Thus, the number of elements n_i belonging to class i among the N elements of the training database is used to calculate the weight w_i such that $w_i = N/(2 * n_i)$.

Hyperparameters The hyperparameters are optimized by 3-fold cross-validated grid-search over a parameter grid. Three hyperparameters are concerned: the translations used to optimize the registration and the γ and C parameters of the SVM classifier corresponding respectively to the kernel coefficient and the penalty parameter of the error term.

Concerning the first hyperparameter, several translations are applied to the patterns to be registered from their initial position. Each of these translations allows us to test a new initialization before applying the Iterative Closest Point algorithm. Three translation amplitudes are considered, $A = [5]$, $A = [10]$ and $A = [20]$, as well as the combinations $A = [5, 10]$ and $A = [10, 20]$. All translations $[t_x, t_y, t_z]$ with $(t_x, t_y, t_z) \in (-A \cup [0] \cup A)^3$ are tested.

Concerning the classifier parameters, the values [0.0001, 0.001, 0.01, 0.1] are tested for the γ parameter and [0.1, 1, 10, 100, 1000] for the C parameter.

3.2.2. Scoring by non-local image patch estimator (SNIPE)

The second approach is based on the Scoring by Non-local Image Patch Estimator (SNIPE) proposed in (Coupé et al., 2012). In order to classify images, SNIPE uses a grading measure based on a nonlocal patch-based framework. This measure estimates the similarity of the patches surrounding the voxels in the image under study with all the patches present in different training populations. In this study, the training database contains two "populations": hemispheres with the pattern-of-interest and those without it (Fig. 7).

Region Of Interest (ROI) delimitation For this approach, the hemispheres to be classified are represented by binary volumes of 2*2*2 mm resolution, containing the fold skeleton. Although the SNIPE method was originally developed to search for matching patches not locally but throughout an entire image, here the patterns-of-interest are specific to local areas of the brain. Thus, it is possible to search for them only on a predefined ROI. This avoids false positives and limits the computing time. Thus, a masking is performed on these volumes in order to extract the ROI. The mask is calculated from the location of the patterns-of-interest in the training database. To avoid omitting a pattern-of-interest whose position would be slightly different from those in the training database, all voxels within 1cm of one of the patterns-of-interest in the training database belong to the ROI. Note that the calculation of the mask is based on the location of the patterns-of-interest in the training database, so it is recalculated for each cross-validation fold, but on average it includes about 2100 voxels of the fold skeleton for the PBS, and 1700 for the ACC patterns.

Optimized PatchMatch algorithm Once the ROI has been extracted, patches with central voxels belonging to the fold skeleton are associated

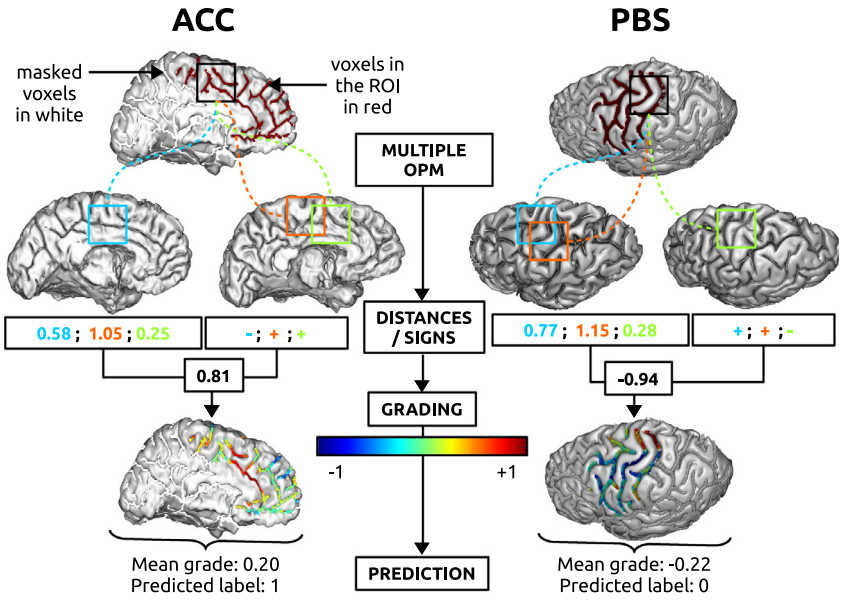

Fig. 7. Illustration of the SNIPE-based method. The column on the left represents the classification of a hemisphere with the pattern-of-interest (ACC study), while the column on the right represents the classification of a hemisphere without the pattern-of-interest (PBS study). For each voxel in the Region Of Interest (ROI) of the hemisphere to be classified, its patch is associated with several similar patches (here 3) in the training database by using the OptimizedPatchMatch (OPM) algorithm. The distance of the similar patches to the patch and their location in a hemisphere with the pattern-of-interest (+) or without it (-), are used to calculate the grade. All grades associated with ROI voxels are averaged to obtain the grade of the hemisphere. A positive grade predicts the presence of the pattern-of-interest and vice versa for a negative grade. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

with similar patches in the training database. To do this, the Optimized Patch Match algorithm (Giraud et al., 2016; Ta et al., 2014) is used. This algorithm searches for similar patches in an atlas library using a cooperative and random strategy resulting in a very low computational burden. Compared to the PatchMatch algorithm (Barnes et al., 2009) from which it is inspired, Optimized PatchMatch is adapted to 3D anatomical atlases by taking into account the rough alignment of images.

Here, the adapted version of the Optimized PatchMatch algorithm implemented in (Borne et al., 2020) is used. This adapted version has been developed specifically for the study of cortical sulci and makes it possible to take advantage of their organisation in the form of meshes. In addition, to manage the unbalanced presence of populations in the training database, the under-represented population is made proportionately more probable to be selected during the initialization by weighting their probabilities. Thus, the probabilities $p(h_i)$ of selecting the hemisphere h_i among the hemispheres $\{h_1, \dots, h_n\}$ of the training set during the initialization is: $p(h_i) = \frac{w(h_i)}{\sum_{j\in\{1,\dots,n\}} w(h_j)}$, with $w(h_k) = \max(1, \frac{n_k}{(n-n_k)})$ and n_k being the number of hemispheres belonging to the class of h_k in the training set.

To obtain several similar patches, the Optimized PatchMatch algorithm is run several times. The patch sizes and the number of similar patches selected are determined by inner cross-validation.

Grading calculation As in (Coupé et al., 2012), in order to estimate the proximity of each voxel in the hemisphere to be classified to both populations, the selected similar patches are used to calculate the degree of proximity to one of the populations. To do so, for each voxel in the hemisphere to be classified, a robust weighted average of the distances of the patches selected to the patch surrounding the voxel is calculated based on the non-local average estimator proposed by Buades et al. (2005). In our case, a negative (respectively positive) classification value indicates that the neighbourhood surrounding the voxel is more specific to hemispheres without the pattern-of-interest than with it. The ROI grades are then averaged to obtain the hemisphere grade.

Hyperparameters The hyperparameters are optimized by 3-fold cross-validated grid-search over a parameter grid. Three hyperparameters are concerned: number of approximate nearest neighbors selected (5, 10, 15, 20, 25, 30), patch sizes (18, 26 or 34mm) and their combinations (all combinations of 1, 2, and 3 different patch sizes are tested). Note that two parameters were set a priori: the size of the search window at 14*14*14 mm and the number of iterations of the propagation and random search steps at 4.

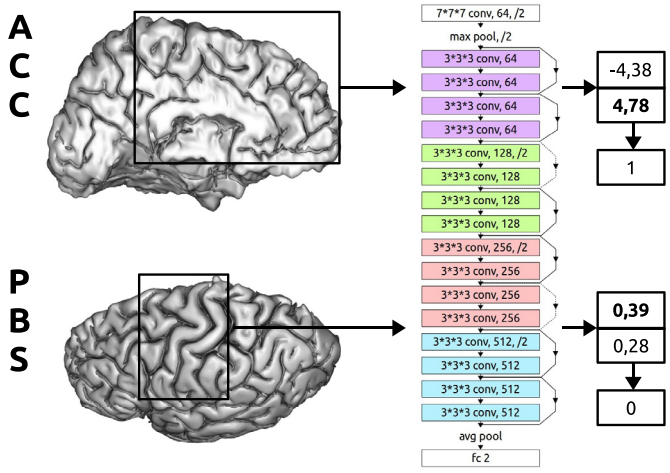

Fig. 8. Illustration of the CNN-based method. The binary volumes of resolution 2*2*2 mm containing the fold skeleton are given as inputs to the neural network. The volume sizes are calculated to contain all the patterns-of-interest of the training database. The neural network used is an 18-layer 3D ResNet. Two values are obtained at the output, one for each class. The highest value determines the predicted class.

In summary, several adaptations have been made compared to (Coupé et al., 2012):

- The search for similar patches is performed using the adaptation of the Optimized PatchMatch algorithm proposed in (Borne et al., 2020) and not using an exhaustive search.
- In order to manage the unbalanced databases, the initialization of the algorithm is weighted to compensate.
- The average of the grades is done on the ROI instead of on a simultaneously segmented region. In fact, since the patterns-of-interest are often missing, it is impossible to define an area to be segmented for the hemispheres without the pattern-of-interest.

3.2.3. Convolutional neural network (CNN)

The last proposed approach uses a 3D Convolutional Neural Network (CNN) (Fig. 8).

Preprocessing The input data of the neural network corresponds to a 3D volume of resolution 2*2*2 mm containing the fold skeleton. Voxels that belong to the fold skeleton are 1; the others are 0. The volume size is calculated to contain all the patterns-of-interest in the training database. Note that unlike the SNIPE-based method where a mask was calculated from the position of the patterns-of-interest to determine the ROI, here it is a parallelepipedic bounding box that is calculated for each fold of the cross validation.

In order to augment the training database, a rotation in a random direction with a random angle (following a Gaussian distribution $\mathcal{N}(0, \frac{\pi}{40}^2)$) and a random translation (each coordinate following a Gaussian distribution $\mathcal{N}(0, 2^2)$) are applied successively to the volumes at each epoch.

Network architecture Concerning the classification of 2D images, spectacular progress has been made in the last decade. This is particularly well illustrated by the ImageNet Large Scale Visual Recognition Challenge, which saw the emergence of several network architectures that are now widely used for image classification. However, for the classification of 3D images, there is currently no such challenge to assess the advantages and disadvantages of several architectures. In the study (Hara et al., 2018), the authors evaluate the performance of adaptations to 3D spatiotemporal images of neural networks used for 2D image classification during the ImageNet Large Scale Visual Recognition Challenge. Based on this study, the architecture of the 3D ResNet was chosen for our problem. The parameters of the architecture used correspond to those of the initially proposed 2D architecture (He et al., 2016), with 3D convolutional layers, a single input channel and 18 layers. In order to be able to manage variable input volume size (since it depends on the folds of the cross-validation), the adaptive average pooling proposed by the Pytorch library (Paszke et al., 2017) is used. Adaptive pooling works like standard pooling, with the stride and kernel size being calculated from the input size ($input\_size$) and the desired output size ($output\_size$). Specifically, the following parameters are used: $stride = floor(input\_size/output\_size)$; $kernel\_size = input\_size - (output\_size - 1) * stride$; $padding = 0$.

The neural network outputs two values, one for each class (i.e. with or without pattern-of-interest). The higher value determines the predicted class.

Training The weight initialization of the convolutional layers follows the method described in (He et al., 2015), using a normal distribution. The stochastic gradient descent with Nesterov momentum (Sutskever et al., 2013) is used for training, with a batch size of 10 samples and a learning rate and momentum determined by an 3-fold inner cross validation. The loss function corresponds to the cross entropy loss, weighted by the inverse of the number of samples of each class in the training database. The learning rate is halved when the loss function has not improved for five consecutive epochs. In accordance with the early stopping strategy, the training is stopped after ten consecutive epochs without improvement. The selected neural network corresponds to the epoch at which the balanced accuracy is the highest.

Ideally, in order to perform early stopping, part of the hemispheres in the training set should be reserved to evaluate the performance of the neural network at each epoch, and train on the others. In this study, in order to maintain a training set of sufficient size during 3-fold inner cross-validation, the validation fold is used to test the model at each epoch and the 2 others folds are used for training.

Hyperparameters In order to choose the hyperparameters (learning rate and momentum), several 3-fold cross-validation loops are performed in turn to adjust the hyperparameters one after the other. First, the momentum is set at 0.9 and the learning rates 1e-2, 1e-3 and 1e-4 are tested in turn. Second, once the learning rates have been tested on all inner cross-validation folds, the learning rate is refined around the value lr that obtained the best balanced accuracy: the values $lr/4$, $lr/2$, $lr * 2$ and $lr * 4$ are tested in turn. Finally, the best learning rate obtained is then tested with momentum 0.8, 0.7 and 0.6 to select the best value. We have chosen not to test values above 0.9 in order to reduce the risk of divergence due to excessive gradient memory.

3.3. Performance evaluation

Cross validation The evaluation of the performance of the trained model was estimated using a double loop of cross validation: one inner 3-fold cross validation loop to select the best hyper-parameters and one outer 10-fold cross validation loop to evaluate classification performances. In order to define the folds of the two cross-validations, special attention was paid to ensure that two hemispheres of the same subject were not separated, one in the training set and the other in the test set. In addition, the hemispheres were distributed so that each fold contained approximately the same proportion of patterns-of-interest.

Classification score Balanced accuracy is used to evaluate the performance of each model (Brodersen et al., 2010). This score is more appropriate than average accuracy in the case of unbalanced databases. It is defined as the average of recall R_c obtained on each class c:

$$R_c = \frac{TP_c}{TP_c + FN_c} \tag{1}$$

with TP_c and FN_c, respectively the number of true positive and false negative samples.

Note that the recall of the class with the pattern-of-interest is called sensitivity, while the recall of the class without the pattern-of-interest is called specificity. These two measures are also reported in the results.

Model comparison In order to compare the models in pairs, a paired T-test was performed between the balanced accuracy lists of each cross validation fold. Note that this comparison method has good repeatability compared to other methods and a modest Type II error but has a high Type I error (Dietterich, 1998). Thus, the results obtained by this test will be questionable if the performances of the models are considered significantly different ($p_{value} < 0.05$), which could correspond to a Type I error, but not if they are considered similar ($p_{value} \geq 0.05$).

In order to determine whether the average balanced accuracy obtained is significant or not, it is compared with chance, i.e. the classification obtained with a Binomial distribution of success probability 0.5 and with a number of samples corresponding to the number of hemispheres to be classified in the database. Since the binomial test used is an exact, two-sided test of the null hypothesis, the p_{value} can be divided by 2 as we test that the accuracy is superior to the chance level.

S score A S score has been defined for each model in order to assess its confidence in the classification of a given hemisphere. For the SVM-based model, the probabilities of presence/absence of a pattern-of-interest is calculated according to the procedure proposed in (Platt, 1999). The proposed S score is the difference between these two probabilities. For the SNIPE-based model, the value of the estimator is directly used as the S score. For the CNN-based model, two scores are obtained at the output of the network, one for each pattern. The S score used corresponds to the difference between these two outputs. In order to determine whether the S score significantly distinguishes fold patterns, the Mann-Whitney test (Mann and Whitney, 1947) is used. In order to compare the effectiveness of the S score in distinguishing between the two classes, we will use it to calculate the Area Under the receiver operating characteristic Curve (AUC).

Looking for rare patterns In order to test the possibility of searching for rare patterns, such as PBS, using the models presented in this study, these were used to search for PBS on an external database: the Human Connectome Project database (Van Essen et al., 2013). This database is composed of healthy volunteers and the T1w structural scans were defaced using the algorithm reported in (Milchenko and Marcus, 2013). The three models proposed were trained on the full training database (i.e. 114 hemispheres, including 28 with a PBS) by fixing the hyperparameters thanks to a 3-fold cross validation. They were then applied to the Human Connectome Project database, containing 1023 subjects, i.e. 2046 hemispheres. In order to see if these models can identify some

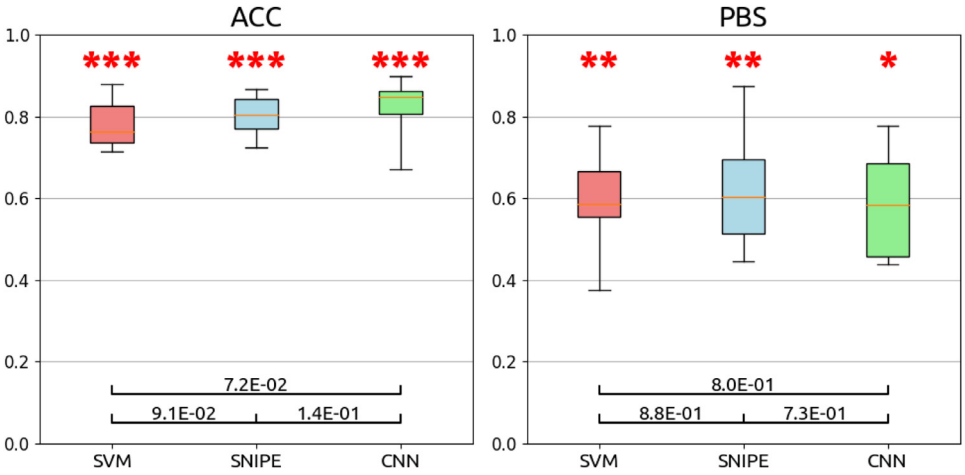

Fig. 9. Balanced accuracies for the 10 folds of the cross-validation by model. The SVM-based model is in red, the SNIPE-based model in blue and the CNN-based model in green. The values of the balanced accuracies obtained for each fold of the cross-validation were compared by paired T-test. The p_{values} obtained are indicated below the compared models. The red stars above the models indicate that the model is better than chance (one star for $p_{value} < 0.05$, two for $p_{value} < 0.02$, three for $p_{value} < 0.001$). The box extends from the lower to upper quartile balanced accuracy values, with a line at the median. The whiskers extend from the box to show the minimum and maximum values.

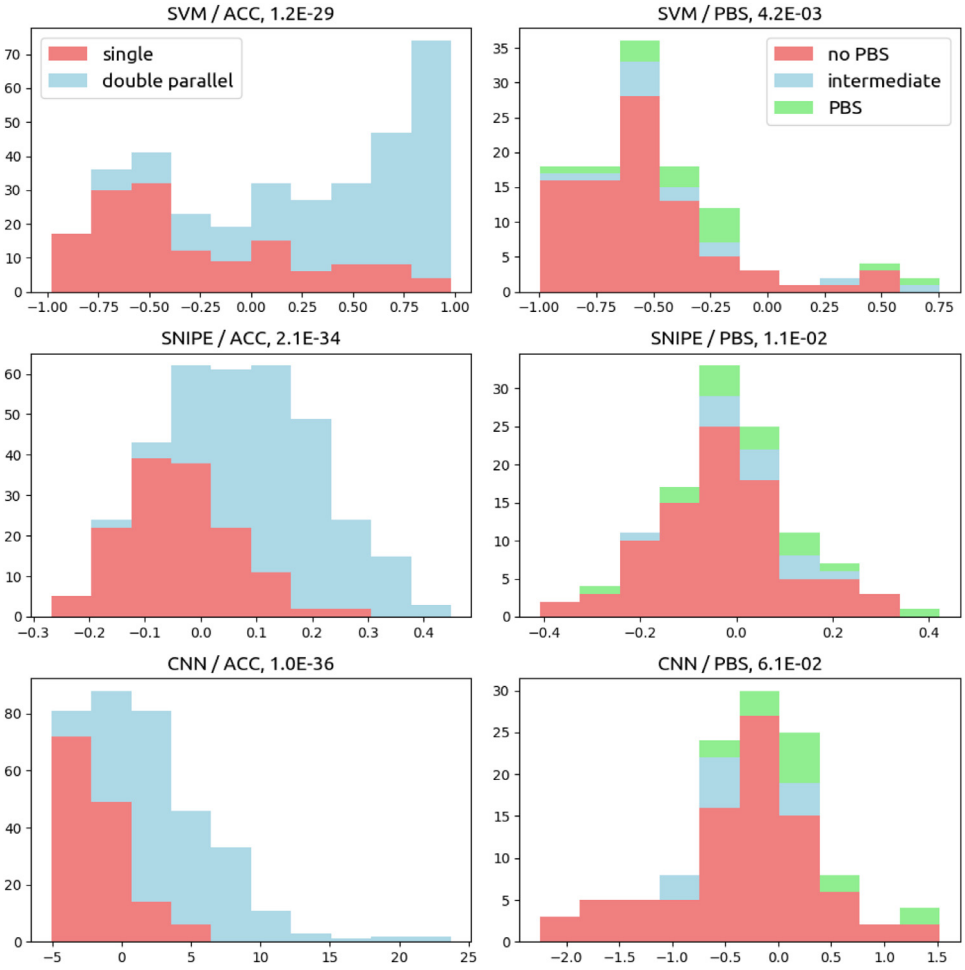

Fig. 10. S score distribution per model. The S scores shown in this figure correspond to the scores obtained in the test folds from the 10-fold cross-validation. Each graph contains a histogram representing the number of samples as a function of the S score. Each color corresponds to a different class, data are stacked on top of each other. The p_{value} of the Mann-Whitney test are indicated in the title of each graph. Note that only the p_{value} of the CNN-based model for PBS recognition is more than 0.05. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

PBSs on this new database, the 9 hemispheres with the highest S score will be examined visually.

4. Results

4.1. Which is the best model?

When analysing the balanced accuracy by model (Fig. 9, Tables 1, 2), we observe that whatever the fold pattern studied, all models are better than chance and that none of them is significantly better than the other

Table 1
Mean balanced accuracy (standard deviation) in % of the 10 folds of the cross-validation for each model presented. The best test scores are highlighted in bold.

	ACC		PBS	
	Train	Test	Train	Test
SVM	78.9 (1.0)	77.9 (5.4)	70.0 (2.4)	60.3 (11.5)
SNIPE	79.7 (1.3)	80.6 (4.5)	66.4 (3.3)	**61.4 (12.3)**
CNN	84.4 (1.9)	**82.7 (6.0)**	70.3 (2.6)	59.0 (12.8)

Table 2
Mean specificity/sensitivity/AUC (standard deviation) in % of the 10 folds of the cross-validation for each model presented. The best score per measure is highlighted in bold.

	ACC			PBS		
	Spec.	Sens.	AUC	Spec.	Sens.	AUC
SVM	75.5(9.9)	80.3(7.1)	85.9(4.0)	69.0(27.5)	51.7(26.3)	61.9(22.0)
SNIPE	71.3(6.6)	**89.8(8.7)**	89.5(3.8)	64.4(13.3)	**58.3(25.0)**	**63.9(13.1)**
CNN	**81.8(7.8)**	83.6(14.0)	**90.6(3.3)**	**71.3(21.5)**	46.7(37.1)	60.5(17.5)

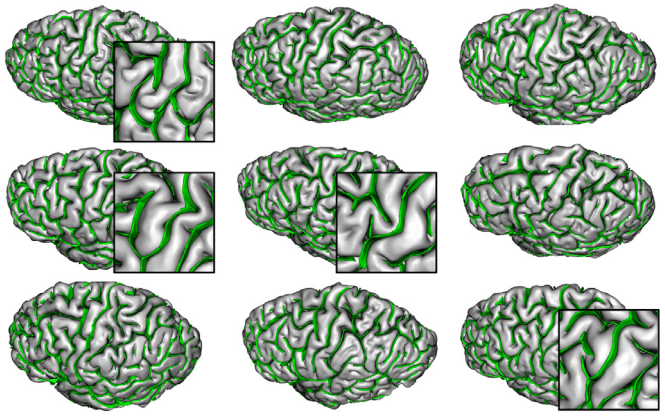

Fig. 11. Looking for PBSs with the SVM-based method. The hemispheres presented are those with the highest S scores on the Human Connectome Project database (score decreasing in the reading direction). Hemispheres with a PBS, or an intermediate pattern, are zoomed. The right hemispheres are flipped for this image.

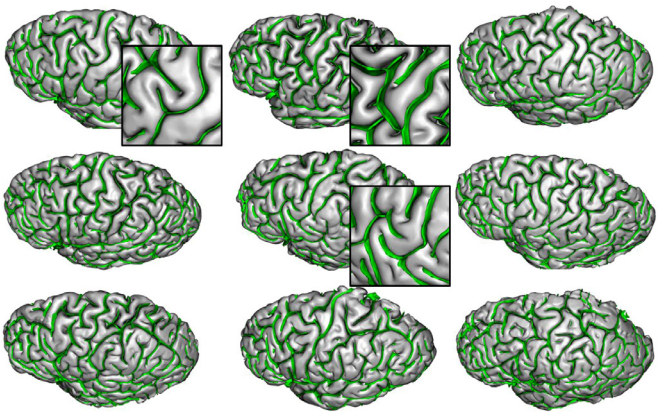

Fig. 12. Looking for PBSs with the SNIPE-based method.

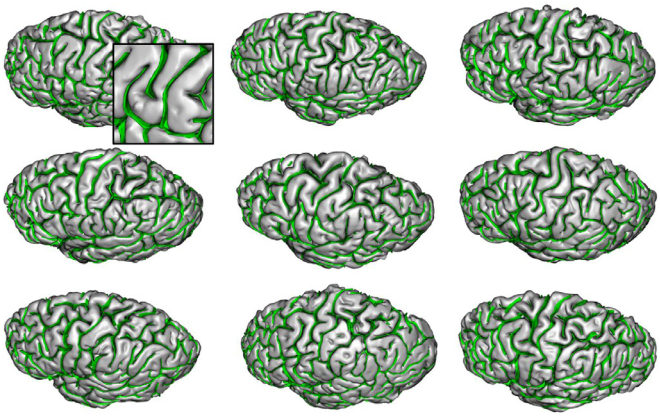

Fig. 13. Looking for PBSs with the CNN-based method.

ones. Thus, the balanced accuracy of the ACC patterns classification is about 80% and that of the PBS about 60%. Models therefore have more difficulty managing rare patterns where the training database is limited and unbalanced. With the current implementation, the CNN-based model requires only a few seconds to determine whether a hemisphere contains a pattern of interest, whereas models based on SNIPE and SVM require at least several minutes (even more without parallelization of the registration of the patterns-of-interest or of the search for similar patches). Concerning training times, these are about several hours (or even days), but they are difficult to compare because they depend heavily on available resources.

Regarding the distribution of the S score defined in section 3.3., the Mann-Whitney test indicates that, except for the CNN-based model, the S scores significantly separate the hemispheres with and without a PBS (Fig. 10). Thus, the CNN-based model does not provide sufficiently reliable scores to assist in the search for PBSs, while the other two models can be used to find PBSs on a new database and enrich the current training database.

4.2. Looking for PBSs

Concerning the search for PBSs on the Human Connectome Project database, the models identified that 30% (for the CNN-based model) to 75% (for the SVM-based model) of the hemispheres presented PBS, which is largely overestimated given the rarity of this pattern in the healthy population. In order to see if these methods can identify some PBSs on this new database, the 9 hemispheres with the highest S score are examined (Figs. 11, 12 and 13). Note that we do not know the proportion of PBSs in the healthy population, so it is difficult to estimate how many PBSs are actually present in the Human Connectome Project database. However, the three models proposed make it possible to find some of them, which is already promising, in particular for SVM and SNIPE-based models that allow us to find more PBSs than the CNN-based model. Indeed, the first two models identify respectively four and three PBSs out of the nine hemispheres examined, while the CNN-based model identifies only one. Note that the PBSs identified with the most confidence by the SVM and SNIPE-based models are different and are not always classified as such by the other model. These methods therefore seem to be complementary for PBSs research. Therefore, by illustrating the usefulness of the proposed models in a particular practical case (the search for a rare pattern in a large database of healthy subjects), these results confirm that the models allow us to meet one of the main motivations of this study.

5. Conclusion

In this study, three methods were proposed for the automatic classification of sulcal patterns. The first method uses the distances of the patterns of the training database to the hemisphere to be classified to train a SVM. The second method is based on the SNIPE patch classification approach proposed by Coupé et al. (2012). The third method uses an 18-layer 3D recurrent neural network.

For the recognition of ACC patterns, the three proposed models have equivalent performance in identifying individual pattern, with a balanced accuracy of about 80%. However, the fastest model to apply is the one based on CNN. For PBS recognition, all three models also have equivalent identification performance, with a balanced accuracy of about 60%. Therefore, it appears that these models perform considerably less efficiently when the database is limited and unbalanced, thus preventing their application for large-scale morphological studies. However, the models provide scores to assess their confidence in the proposed identification, allowing to focus only on the subjects where the models are most confident. Except for the CNN-based model, these scores significantly distinguish between hemispheres with and without PBS. Thus, the CNN-based model is less efficient than the others for PBS identification. This is probably due to the fact that neural networks usually require more training data than classical approaches, which is a problem for the study of rare patterns like the PBS.

In conclusion, none of the proposed models is better than the other two because their performance in identifying individual patterns is equivalent and, although the fastest model is the CNN-based model, it is also the least efficient model for finding rare patterns. It is therefore complicated to advise one model over another because it depends on the pattern-of-interest. However, we believe that these methods perform well enough to later help (1) to identify a link between sulcal shapes and functional organization of the cortex (e.g. by studying the distribution of ACC patterns on larger databases) and (2) to understand their role in neurodevelopmental disease (e.g. by helping to identify rare patterns such as PBS in healthy population). Moreover, these methods are generic enough to be applicable to a wide range of other folding patterns that are the subject of numerous studies (Germann et al., 2005; Hotier et al., 2017; Huntgeburth and Petrides, 2012; Kim et al., 2008; Ochiai et al., 2004; Plaze et al., 2015; 2009; Segal and Petrides, 2012; Snyder et al., 2019; Sprung-Much and Petrides, 2018; Zlatkina and Petrides, 2010).

5.1. How can the current models be improved?

In order to improve the test accuracy of the current proposed models, one possible approach would be to train a 2D neural network from the surface projections of the folds on a 2D image representation. Indeed, 2D networks are easier to train than 3D networks, thanks to their reduced number of parameters to learn. The difficulty of such an approach would be to determine an optimal way to project the brain surface on a 2D grid, without excessively distorting the fold patterns. Such projections have recently been proposed for sulci recognition using a 2D CNN (Parvathaneni et al., 2019). Once this step is completed, it opens the way to many promising neural network architectures that have already proven their worth in 2D image classification, such as for the ImageNet challenge (Russakovsky et al., 2015).

Another possible approach would be to use non-Euclidean neural networks (Bronstein et al., 2017) in order to train the network directly on the surface of the brain and avoid a tricky projection on a 2D grid which necessarily leads to a loss of information.

5.2. How to define fold patterns automatically?

Cortical folds are so variable that manual pattern definition is tricky. Indeed, it is usually the result of long reflections after visualization and study of a large number of subjects by experts. Moreover, the use of a limited database, or of the limited human memory, questions the generalisability of the defined patterns. Thus, many approaches seek to discover folding patterns in an unsupervised manner. These methods define a region of interest and deduce patterns automatically, using clustering algorithms (Duan et al., 2019; Im et al., 2011; Meng et al., 2018; Sun et al., 2009; 2007). However, in this study, we observed that the classification of local patterns is sometimes difficult to carry out because of the continuity between the different configurations. For this reason, in the study of ACC patterns, some subjects were excluded from the study due to a lack of consensus on their identification. Similarly, different levels of PBS presence were used when tagging the training database. Also, other unsupervised approaches prefer to use manifold based techniques to better represent the continuity between two extreme configurations (Sun et al., 2016; 2012). Similarly, it would be interesting to predict a score for the presence of ACC patterns or PBS, rather than attempt to distinguish them into separate groups. Nevertheless, all these unsupervised approaches are based on the prior labeling of the structures of interest, which is sometimes difficult to obtain and may lead to radically different results depending on the region of interest used.

Other approaches are not based on the definition of a region of interest but seek to discover patterns of group related anatomical structure (Takerkart et al., 2017; Toews et al., 2010). In the same vein, some of the proposed methods in this study would be easily adaptable to the search for new patterns that characterize a group. In fact, we believe that fold skeletons could be used to train a model to distinguish patients from controls. If the model achieves significant classification performance, it becomes particularly interesting to look at the patterns that allow it to perform a correct classification. For example, since the SNIPE method is based on the similarity of each patch surrounding the voxels of a fold skeleton to both populations in the learning base, it is then easy to visually identify, for a given skeleton, the sulcal pattern specific to the assigned group. Similarly, by training a CNN to make such classifications, different techniques could help identify areas of the fold skeleton that allow correct classification, such as the use of saliency maps (Simonyan et al., 2014) or occlusions (Zeiler and Fergus, 2014), already used to understand how neural networks work. Such tools would help neuroanatomy experts to identify fold patterns specific to a given group.

Currently, there is no approach to automatically discover folding patterns that gets rid of both a priori definition of region of interest and group of interest. In the future, a fully unsupervised approach would enable the use of the huge databases available today to define patterns generalisable to the whole population, which would greatly facilitate the study and understanding of cortical folds variability.

Credit authorship contribution statement

Léonie Borne: Conceptualization, Methodology, Software, Validation, Resources, Data curation, Writing - original draft, Visualization, Supervision, Project administration. **Denis Rivière:** Conceptualization, Methodology, Software, Validation, Resources, Data curation, Writing - review & editing, Supervision. **Arnaud Cachia:** Conceptualization, Validation, Resources, Data curation, Writing - review & editing. **Pauline Roca:** Conceptualization, Validation, Resources, Data curation, Writing - review & editing. **Charles Mellerio:** Conceptualization, Validation, Resources, Data curation, Writing - review & editing. **Catherine Oppenheim:** Conceptualization, Validation, Resources, Writing - review & editing. **Jean-François Mangin:** Conceptualization, Methodology, Validation, Resources, Data curation, Writing - review & editing, Supervision, Funding acquisition.

Acknowledgments

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No. 785907 & 945539 (HBP SGA2 & SGA3), from the FRMDIC20161236445, from the ANR IFOPASUBA, from the ANR-14-CE30-0014-01 APEX and from the ANR FOLDDICO. Data were provided in part by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University.

References

Alemán-Gómez, Y., Janssen, J., Schnack, H., Balaban, E., Pina-Camacho, L., Alfaro-Almagro, F., Castro-Fornieles, J., Otero, S., Baeza, I., Moreno, D., et al., 2013. The human cerebral cortex flattens during adolescence. Journal of Neuroscience 33 (38), 15004–15010.

Amiez, C., Kostopoulos, P., Champod, A.-S., Petrides, M., 2006. Local morphology predicts functional organization of the dorsal premotor region in the human brain. Journal of Neuroscience 26 (10), 2724–2731.

Amiez, C., Petrides, M., 2009. Anatomical organization of the eye fields in the human and non-human primate frontal cortex. Progress in neurobiology 89 (2), 220–230.

Amiez, C., Petrides, M., 2014. Neuroimaging evidence of the anatomo-functional organization of the human cingulate motor areas. Cerebral cortex 24 (3), 563–578.

Barnes, C., Shechtman, E., Finkelstein, A., Goldman, D.B., 2009. Patchmatch: A randomized correspondence algorithm for structural image editing. In: ACM Transactions on Graphics (ToG), 28. ACM, p. 24.

Bertoux, M., Lagarde, J., Corlier, F., Hamelin, L., Mangin, J.-F., Colliot, O., Chupin, M., Braskie, M.N., Thompson, P.M., Bottlaender, M., et al., 2019. Sulcal morphology in alzheimer's disease: an effective marker of diagnosis and cognition. Neurobiology of aging 84, 41–49.

Besl, P.J., McKay, N.D., 1992. Method for registration of 3-d shapes. In: Sensor Fusion IV: Control Paradigms and Data Structures, 1611. International Society for Optics and Photonics, pp. 586–607.

Besson, P., Andermann, F., Dubeau, F., Bernasconi, A., 2008. Small focal cortical dysplasia lesions are located at the bottom of a deep sulcus. Brain 131 (12), 3246–3255.

Bodin, C., Takerkart, S., Belin, P., Coulon, O., 2018. Anatomo-functional correspondence in the superior temporal sulcus. Brain Structure and Function 223 (1), 221–232.

Borne, L., Rivière, D., Mancip, M., Mangin, J.-F., 2020. Automatic labeling of cortical sulci using patch-or cnn-based segmentation techniques combined with bottom-up geometric constraints. Medical Image Analysis 101651.

Borst, G., Cachia, A., Vidal, J., Simon, G., Fischer, C., Pineau, A., Poirel, N., Mangin, J.-F., Houdé, O., 2014. Folding of the anterior cingulate cortex partially explains inhibitory control during childhood: a longitudinal study. Developmental cognitive neuroscience 9, 126–135.

Brodersen, K.H., Ong, C.S., Stephan, K.E., Buhmann, J.M., 2010. The balanced accuracy and its posterior distribution. In: 2010 20th International Conference on Pattern Recognition. IEEE, pp. 3121–3124.

Bronstein, M.M., Bruna, J., LeCun, Y., Szlam, A., Vandergheynst, P., 2017. Geometric deep learning: going beyond euclidean data. IEEE Signal Processing Magazine 34 (4), 18–42.

Buades, A., Coll, B., Morel, J.-M., 2005. A non-local algorithm for image denoising. In: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), 2. IEEE, pp. 60–65.

Cachia, A., Borst, G., Tissier, C., Fisher, C., Plaze, M., Gay, O., Riviére, D., Gogtay, N., Giedd, J., Mangin, J.-F., Houdé, O., Raznahan, A., 2016. Longitudinal stability of the folding pattern of the anterior cingulate cortex during development. Developmental Cognitive Neuroscience 19, 122–127.

Cachia, A., Borst, G., Vidal, J., Fischer, C., Pineau, A., Mangin, J.-F., Houdé, O., 2014. The shape of the acc contributes to cognitive control efficiency in preschoolers. Journal of cognitive neuroscience 26 (1), 96–106.

Chakravarty, M.M., Rapoport, J.L., Giedd, J.N., Raznahan, A., Shaw, P., Collins, D.L., Lerch, J.P., Gogtay, N., 2015. Striatal shape abnormalities as novel neurodevelopmental endophenotypes in schizophrenia: a longitudinal study. Human brain mapping 36 (4), 1458–1469.

Collins, D.L., Neelin, P., Peters, T.M., Evans, A.C., 1994. Automatic 3d intersubject registration of mr volumetric data in standardized talairach space.. Journal of computer assisted tomography 18 (2), 192–205.

Coupé, P., Eskildsen, S.F., Manjón, J.V., Fonov, V.S., Collins, D.L., 2012. Simultaneous segmentation and grading of anatomical structures for patient's classification: application to alzheimer's disease. NeuroImage 59 (4), 3736–3747.

Crosson, B., Sadek, J.R., Bobholz, J.A., Gökçay, D., Mohr, C.M., Leonard, C.M., Maron, L., Auerbach, E.J., Browd, S.R., Freeman, A.J., et al., 1999. Activity in the paracingulate and cingulate sulci during word generation: an fmri study of functional anatomy. Cerebral cortex 9 (4), 307–316.

Derrfuss, J., Vogt, V.L., Fiebach, C.J., von Cramon, D.Y., Tittgemeyer, M., 2012. Functional organization of the left inferior precentral sulcus: dissociating the inferior frontal eye field and the inferior frontal junction. Neuroimage 59 (4), 3829–3837.

Dietterich, T.G., 1998. Approximate statistical tests for comparing supervised classification learning algorithms. Neural computation 10 (7), 1895–1923.

Duan, D., Xia, S., Rekik, I., Meng, Y., Wu, Z., Wang, L., Lin, W., Gilmore, J.H., Shen, D., Li, G., 2019. Exploring folding patterns of infant cerebral cortex based on multi-view curvature features: Methods and applications. Neuroimage 185, 575–592.

Fornito, A., Yücel, M., Wood, S., Stuart, G.W., Buchanan, J.-A., Proffitt, T., Anderson, V., Velakoulis, D., Pantelis, C., 2004. Individual differences in anterior cingulate/paracingulate morphology are related to executive functions in healthy males. Cerebral cortex 14 (4), 424–431.

Foubet, O., Trejo, M., Toro, R., 2019. Mechanical morphogenesis and the development of neocortical organisation. Cortex 118, 315–326.

Germann, J., Robbins, S., Halsband, U., Petrides, M., 2005. Precentral sulcal complex of the human brain: morphology and statistical probability maps. Journal of Comparative Neurology 493 (3), 334–356.

Giedd, J.N., Blumenthal, J., Jeffries, N.O., Castellanos, F.X., Liu, H., Zijdenbos, A., Paus, T., Evans, A.C., Rapoport, J.L., 1999. Brain development during childhood and adolescence: a longitudinal mri study. Nature neuroscience 2 (10), 861–863.

Giraud, R., Ta, V.-T., Papadakis, N., Manjón, J.V., Collins, D.L., Coupé, P., 2016. An optimized patchmatch for multi-scale and multi-feature label fusion. NeuroImage 124, 770–782.

Grosbras, M.-H., Lobel, E., Van de Moortele, P.-F., LeBihan, D., Berthoz, A., 1999. An anatomical landmark for the supplementary eye fields in human revealed with functional magnetic resonance imaging. Cerebral Cortex 9 (7), 705–711.

Hara, K., Kataoka, H., Satoh, Y., 2018. Can spatiotemporal 3d cnns retrace the history of 2d cnns and imagenet? In: Proceedings of the IEEE conference on Computer Vision and Pattern Recognition, pp. 6546–6555.

He, K., Zhang, X., Ren, S., Sun, J., 2015. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In: Proceedings of the IEEE international conference on computer vision, pp. 1026–1034.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770–778.

Holz, D., Ichim, A.E., Tombari, F., Rusu, R.B., Behnke, S., 2015. Registration with the point cloud library: A modular framework for aligning in 3-d. IEEE Robotics & Automation Magazine 22 (4), 110–124.

Hotier, S., Leroy, F., Boisgontier, J., Laidi, C., Mangin, J.-F., Delorme, R., Bolognani, F., Czech, C., Bouquet, C., Toledano, E., et al., 2017. Social cognition in autism is associated with the neurodevelopment of the posterior superior temporal sulcus. Acta Psychiatrica Scandinavica 136 (5), 517–525.

Huntgeburth, S.C., Petrides, M., 2012. Morphological patterns of the collateral sulcus in the human brain. European Journal of Neuroscience 35 (8), 1295–1311.

Im, K., Grant, P.E., 2019. Sulcal pits and patterns in developing human brains. NeuroImage 185, 881–890.

Im, K., Pienaar, R., Lee, J.-M., Seong, J.-K., Choi, Y.Y., Lee, K.H., Grant, P.E., 2011. Quantitative comparison and analysis of sulcal patterns using sulcal graph matching: a twin study. Neuroimage 57 (3), 1077–1086.

Kim, H., Bernasconi, N., Bernhardt, B., Colliot, O., Bernasconi, A., 2008. Basal temporal sulcal morphology in healthy controls and patients with temporal lobe epilepsy. Neurology 70 (22 Part 2), 2159–2165.

Le Guen, Y., Philippe, C., Riviere, D., Lemaitre, H., Grigis, A., Fischer, C., Dehaene-Lambertz, G., Mangin, J.-F., Frouin, V., 2019. eqtl of kcnk2 regionally influences the brain sulcal widening: evidence from 15,597 uk biobank participants with neuroimaging data. Brain Structure and Function 224 (2), 847–857.

Le Provost, J.-B., Bartrés-Faz, D., Paillère-Martinot, M.-L., Artiges, E., Pappata, S., Recasens, C., Pérez-Gómez, M., Bernardo, M., Baeza, I., Bayle, F., Martinot, J., 2003. Paracingulate sulcus morphology in men with early-onset schizophrenia. The British Journal of Psychiatry 182 (3), 228–232.

Llinares-Benadero, C., Borrell, V., 2019. Deconstructing cortical folding: genetic, cellular and mechanical determinants. Nature Reviews Neuroscience 20 (3), 161–176.

Lopez-Persem, A., Verhagen, L., Amiez, C., Petrides, M., Sallet, J., 2019. The human ventromedial prefrontal cortex: sulcal morphology and its influence on functional organization. Journal of Neuroscience 39 (19), 3627–3639.

Mangin, J.-F., Le Guen, Y., Labra, N., Grigis, A., Frouin, V., Guevara, M., Fischer, C., Rivière, D., Hopkins, W.D., Régis, J., et al., 2019. "plis de passage" deserve a role in models of the cortical folding process. Brain Topography 1–14.

Mangin, J.-F., Riviere, D., Cachia, A., Duchesnay, E., Cointepas, Y., Papadopoulos-Orfanos, D., Collins, D.L., Evans, A.C., Régis, J., 2004. Object-based morphometry of the cerebral cortex. IEEE transactions on medical imaging 23 (8), 968–982.

Mann, H.B., Whitney, D.R., 1947. On a test of whether one of two random variables is stochastically larger than the other. The annals of mathematical statistics 50–60.

Mellerio, C., Roca, P., Chassoux, F., Danière, F., Cachia, A., Lion, S., Naggara, O., Devaux, B., Meder, J.-F., Oppenheim, C., 2014. The power button sign: a newly described central sulcal pattern on surface rendering mr images of type 2 focal cortical dysplasia. Radiology 274 (2), 500–507.

Meng, Y., Li, G., Wang, L., Lin, W., Gilmore, J.H., Shen, D., 2018. Discovering cortical sulcal folding patterns in neonates using large-scale dataset. Human brain mapping 39 (9), 3625–3635.

Milchenko, M., Marcus, D., 2013. Obscuring surface anatomy in volumetric imaging data. Neuroinformatics 11 (1), 65–75.

Ochiai, T., Grimault, S., Scavarda, D., Roch, G., Hori, T., Rivière, D., Mangin, J.F., Régis, J., 2004. Sulcal pattern and morphology of the superior temporal sulcus. Neuroimage 22 (2), 706–719.

Parvathaneni, P., Nath, V., McHugo, M., Huo, Y., Resnick, S.M., Woodward, N.D., Landman, B.A., Lyu, I., 2019. Improving human cortical sulcal curve labeling in large scale cross-sectional mri using deep neural networks. Journal of neuroscience methods 108311.

Paszke, A., Gross, S., Chintala, S., Chanan, G., Yang, E., DeVito, Z., Lin, Z., Desmaison, A., Antiga, L., Lerer, A., 2017. Automatic differentiation in pytorch. NIPS-W.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., 2011. Scikit-learn: Machine learning in Python. Journal of Machine Learning Research 12, 2825–2830.

Platt, J., 1999. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. Advances in large margin classifiers 10 (3), 61–74.

Plaze, M., Mangin, J.-F., Paillere-Martinot, M.-L., Artiges, E., Olie, J.-P., Krebs, M.-O., Gaillard, R., Martinot, J.-L., Cachia, A., 2015. "Who is talking to me?" self-other attribution of auditory hallucinations and sulcation of the right temporoparietal junction. Schizophrenia research 169 (1-3), 95–100.

Plaze, M., Paillére-Martinot, M.-L., Penttilä, J., Januel, D., de Beaurepaire, R., Bellivier, F., Andoh, J., Galinowski, A., Gallarda, T., Artiges, E., Olié, J.-P., Mangin, J.-F., Martinot, J.-L., Cachia, A., 2009. "Where Do Auditory Hallucinations Come From?" A Brain Morphometry Study of Schizophrenia Patients With Inner or Outer Space Hallucinations. Schizophrenia Bulletin 37 (1), 212–221.

Régis, J., Tamura, M., Park, M.C., McGonigal, A., Riviere, D., Coulon, O., Bartolomei, F., Girard, N., Figarella-Branger, D., Chauvel, P., Mangin, J.-F., 2011. Subclinical abnormal gyration pattern, a potential anatomic marker of epileptogenic zone in patients with magnetic resonance imaging–negative frontal lobe epilepsy. Neurosurgery 69 (1), 80–94.

Roca, P., Mellerio, C., Chassoux, F., Rivière, D., Cachia, A., Charron, S., Lion, S., Mangin, J.-F., Devaux, B., Meder, J.-F., Oppenheim, C., 2015. Sulcus-based mr analysis of focal cortical dysplasia located in the central region. PloS one 10 (3), e0122252.

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A.C., Fei-Fei, L., 2015. Imagenet large scale visual recognition challenge. International journal of computer vision 115 (3), 211–252.

Segal, E., Petrides, M., 2012. The morphology and variability of the caudal rami of the superior temporal sulcus. European Journal of Neuroscience 36 (1), 2035–2053.

Segal, E., Petrides, M., 2013. Functional activation during reading in relation to the sulci of the angular gyrus region. European Journal of Neuroscience 38 (5), 2793–2801.

Simonyan, K., Vedaldi, A., Zisserman, A., 2014. Deep inside convolutional networks: Visualising image classification models and saliency maps.

Snyder, W., Patti, M., Troiani, V., 2019. An evaluation of automated tracing for orbitofrontal cortex sulcogyral pattern typing. Journal of neuroscience methods 326, 108386.

Sprung-Much, T., Petrides, M., 2018. Morphological patterns and spatial probability maps of two defining sulci of the posterior ventrolateral frontal cortex of the human brain: the sulcus diagonalis and the anterior ascending ramus of the lateral fissure. Brain Structure and Function 223 (9), 4125–4152.

Sun, Z., Pinel, P., Rivière, D., Moreno, A., Dehaene, S., Mangin, J.-F., 2016. Linking morphological and functional variability in hand movement and silent reading. Brain Structure and Function 221 (7), 3361–3371.

Sun, Z.Y., Klöppel, S., Rivière, D., Perrot, M., Frackowiak, R., Siebner, H., Mangin, J.-F., 2012. The effect of handedness on the shape of the central sulcus. Neuroimage 60 (1), 332–339.

Sun, Z.Y., Perrot, M., Tucholka, A., Rivière, D., Mangin, J.-F., 2009. Constructing a dictionary of human brain folding patterns. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 117–124.

Sun, Z.Y., Rivière, D., Poupon, F., Régis, J., Mangin, J.-F., 2007. Automatic inference of sulcus patterns using 3d moment invariants. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 515–522.

Sutskever, I., Martens, J., Dahl, G., Hinton, G., 2013. On the importance of initialization and momentum in deep learning. In: International conference on machine learning, pp. 1139–1147.

Ta, V.-T., Giraud, R., Collins, D.L., Coupé, P., 2014. Optimized patchmatch for near real time and accurate label fusion. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 105–112.

Takerkart, S., Auzias, G., Brun, L., Coulon, O., 2017. Structural graph-based morphometry: A multiscale searchlight framework based on sulcal pits. Medical image analysis 35, 32–45.

Toews, M., Wells III, W., Collins, D.L., Arbel, T., 2010. Feature-based morphometry: Discovering group-related anatomical patterns. NeuroImage 49 (3), 2318–2327.

Van Essen, D.C., 1997. A tension-based theory of morphogenesis and compact wiring in the central nervous system. Nature 385 (6614), 313–318.

Van Essen, D.C., Smith, S.M., Barch, D.M., Behrens, T.E., Yacoub, E., Ugurbil, K., 2013. The wu-minn human connectome project: an overview. Neuroimage 80, 62–79.

Watson, J.D., Myers, R., Frackowiak, R.S., Hajnal, J.V., Woods, R.P., Mazziotta, J.C., Shipp, S., Zeki, S., 1993. Area v5 of the human brain: evidence from a combined study using positron emission tomography and magnetic resonance imaging. Cerebral cortex 3 (2), 79–94.

Weiner, K.S., Golarai, G., Caspers, J., Chuapoco, M.R., Mohlberg, H., Zilles, K., Amunts, K., Grill-Spector, K., 2014. The mid-fusiform sulcus: a landmark identifying both cytoarchitectonic and functional divisions of human ventral temporal cortex. Neuroimage 84, 453–465.

Yücel, M., Wood, S.J., Phillips, L.J., Stuart, G.W., Smith, D.J., Yung, A., Velakoulis, D., Mcgorry, P.D., Pantelis, C., 2003. Morphology of the anterior cingulate cortex in young men at ultra-high risk of developing a psychotic illness. The British Journal of Psychiatry 182 (6), 518–524.

Zeiler, M.D., Fergus, R., 2014. Visualizing and understanding convolutional networks. In: European conference on computer vision. Springer, pp. 818–833.

Zlatkina, V., Amiez, C., Petrides, M., 2016. The postcentral sulcal complex and the transverse postcentral sulcus and their relation to sensorimotor functional organization. European Journal of Neuroscience 43 (10), 1268–1283.

Zlatkina, V., Petrides, M., 2010. Morphological patterns of the postcentral sulcus in the human brain. Journal of Comparative Neurology 518 (18), 3701–3724.