Ensemble learning for the detection of pli-de-passages in the superior temporal sulcus

Tianqi Song, Clémentine Bodin, Olivier Coulon

Ensemble learning for the detection of pli-de-passages in the superior temporal sulcus

Tianqi Song[a], Clémentine Bodin[b,c] and Olivier Coulon[a,*]

[a]Institut de Neurosciences de la Timone, Aix-Marseille Univ, UMR CNRS 7289, Marseille, France
[b]Center for Research on Brain, Language, and Music, McGill University, Montreal, QC, Canada
[c]Department of Biology, McGill University, Montreal, QC, Canada

ARTICLE INFO

Keywords:
Plis de passage
Cortical folding
Cerebral Cortex
Machine learning
SVM

ABSTRACT

The surface of the cerebral cortex is very convoluted, with a large number of folds, the cortical sulci. These folds are extremely variable from one individual to another, and this large variability is a problem for many applications in neuroscience and brain imaging. In particular, sulcal geometry (shape) and sulcal topology (branches, number of pieces) are very variable. "Plis de passages" (PPs) or "annectant gyri" can explain part of the topological variability, namely why sulci have a variable number of pieces across subjects. The concept of PPs was first introduced by Gratiolet (1854) to describe transverse gyri that interconnect both sides of a sulcus, that are frequently buried in the depth of sulci, and that are sometimes apparent on the cortical surface, hence seemingly interrupting the course of sulci and separating them in several pieces. Nevertheless, the difficulty of identifying PPs and the lack of systematic methods to automatically detect them has limited their use. However, based on a recent characterization of PPs in the superior temporal sulcus, we present here a method to automatically detect PPs in the superior temporal sulcus. Local morphology within the sulcus is characterized using cortical surface profiling, and the three-dimensional PP recognition problem is performed as a two-dimensional image classification problem with class-imbalance. This is solved by using an ensemble support vector machine model (EnsSVM) with a rebalancing strategy. Cross validation and quantitative experimental results on an external dataset show the effectiveness and robustness of our approach.

1. Introduction

The cerebral cortex is a very convoluted surface that folds itself into gyri and sulci, which vary a lot across individuals. As early as the 19th century, anatomists were interested in its organisation and features. They observed that many long primary sulci could be subdivided into several pieces but sometimes appeared as a whole. This topological variability, together with a strong geometrical variability (shape and size of the sulci), leads to difficulties for the systematic study of gyri and sulci across different brains (Ono et al., 1990; Regis et al., 2005) and makes the many sulcal patterns observed on the cortex hard to decipher.

The concept of "Pli de passage" (PP) was introduced by Gratiolet (1854) to describe transverse gyri that interconnect both sides of a sulcus, are frequently buried in the depth of these sulci, and are sometimes apparent on the cortical surface. Later on, Broca specifically reported the presence of three transverse gyri that connect the pre- and post-central gyri along the central sulcus(CS) Broca (1888) : PPs frontal superior, middle and inferior. Such sulcal segmentation was then confirmed by Cunningham's work on cortical morphology and development (Cunningham, 1890a,b, 1897), in which he described how the central sulcus appears during ontogeny from two separate folding seeds merging later on during cortical expansion. The PPs, named "deep annectant gyri" (Cunningham, 1890a,b, 1897), appear early between two isolated parts of sulci, and can eventually get buried deep into the bottom of sulci, with their presence identified as elevations of the sulcal fundus, or interlocking gyri. In the CS, a specific PP has been identified and named "Pli-de-passage fronto-parietal moyen"(PPFM) (Cunningham, 1897; Regis et al., 2005; Cykowski et al., 2008). It has been linked to the location of the primary motor area of the hand (Boling et al., 1999) and sometimes associated with a specific complex U-fiber connectivity (Pron et al., 2021; Catani et al., 2012; Shinohara et al., 2020). Moreover, these potential links with connectivity and function raise the possibility of a clinical relevance, particularly when it comes to localizing specific areas in a context of neurosurgery (Boling et al., 1999). Their potential early

*Corresponding author
ORCID(s):

development (Bodin et al., 2021; Mangin et al., 2019) also points to theit potential use for identifying developmental abnormalities once their normal distribution is known (Leroy et al., 2015; Le Guen et al., 2018; Bodin et al., 2021).

Regis (1994) proposed the first map of PPs over the entire brain. This map gives the approximate locations of PPs, which were deduced from the frequent interruptions of folding patterns within specified brain localization and generally consistent with the initial folding patterns during development (Mangin et al., 2019). Because of the large inter-individual sulcal variability across individuals, this map cannot exactly contain all the PPs and there is no evidence so far that it even contains a stable subset of PPs. However, it provides a framework that explains part of the inter-individual differences. Together with this map, Regis developed the "Sulcal Root" generic model that states a putative spatial organisation of the sulcal units separated by PPs over the entire cortical surface (Regis et al., 2005; Régis et al., 1995). Following the "Sulcal Root" model, Ochiai et al. (2004) proposed a detection of all PPs in the superior temporal sulcus (STS) according to their location and relative depth of the sulcus. Nevertheless, this work hypothesize the existence of a fixed number of PPs in the STS, which is in contradiction with more recent work (Leroy et al., 2015; Le Guen et al., 2018; Bodin et al., 2021). In several recent studies, PPs have been described with various names such as annectant gyrus, gyral bridge, submerged gyrus (Germann et al., 2005), submerged gyral passage (Zlatkina and Petrides, 2010; Sprung-Much and Petrides, 2018) or transition (Huntgeburth and Petrides, 2012). Several authors have pointed out that the PPs provided an explanation for variable sulcal patterns: the presence of a PP, visible or buried deep in the depth of a sulcus, explains the presence of sulci interruptions in some subjects and not others, making PPs relevant morphological landmarks for the understanding of sulcal folding (Regis et al., 2005; Zlatkina and Petrides, 2010; Huntgeburth and Petrides, 2012; Mangin et al., 2019; Bodin et al., 2021).

In the context of understanding gyrification and its variability across individuals, the geometry of the cortical surface has been widely studied. Quantitative descriptions of folding patterns (Zilles et al., 1988; Yu et al., 2007; Toro et al., 2008; Li et al., 2010; Rabiei et al., 2016) and models of the underlying mechanisms (Van Essen, 1997; Toro and Burnod, 2005; Geng et al., 2009; Auzias et al., 2015; Chen et al., 2017) have been proposed . Nevertheless, very few works explicitly studied PPs, and when they did it has been mostly in the central sulcus (CS) and superior temporal sulcus (STS). This is partly due to the lack of precise geometrical definition of the PPs, which are still an ill-defined concept.

Indeed, PPs have been described qualitatively but no precise consensual definition has been proposed, and most attempts have been relying on depth variation in the fundus of sulci. As far as we know, Cykowski et al. (2008) proposed the first observer-independent automatic and quantitative study of depth variations in the CS. In particular, in this work it was shown that a typical depth profile of the CS shows the PPFM as a local minimum of depth. And although not systematic, it is clear that the presence a PP can lead to a local variation (increase) of sulcal depth. This concept has been used to explicitly study PPs in the STS. For instance, Leroy et al. (2015) used the depth profile along the fundus of the STS to detect PPs as depth minima that are blow a depth threshold. An equivalent definition was used by Le Guen et al. (2018) in which depth minima were extracted then filtered by a threshold on their depth value.

The general issue with such approach is that depth variations are useful when their amplitude is above a certain threshold, thus detecting "superficial" or "intermediate" PPs (according to the definition of PPs in (Ochiai et al., 2004)), but they are not able to differentiate small depth variations associated with very deep PPs and depth variations that are caused by noise in the surface geometry or non-PP related variations in general. For instance, inCykowski et al. (2008), it is shown that for 3% of the population there is no depth minimum associated with the PPFM. Moreover, the use of a threshold necessarily leads to false negative that are just below the threshold.

This has led to controversial results such as an asymmetry of the number of PPs in the STS between the left and right hemispheres (Bodin et al., 2021). Indeed, PPs are three-dimensional structures, and their identification based on a two-dimensional depth profile is prone to be deficient (Bodin et al., 2021). In their original 19th century description, it was stated that PPs buried in the depth of sulci were associated with elevations of sulcal fundi but also with interlocking gyri on the walls of sulci (Cunningham, 1890a,b, 1897). The latter has been very little studied until recently (Zlatkina et al., 2016; Bodin et al., 2021). Zlatkina et al. (2016) identified deep PPs which were only 1 mm in depth variation and characterized by an unusual curvature of the sulcus. Such approach based on the geometry of the sulcus itself, together with the concept of the interlocking gyri, led Bodin et al. (2021) to propose a PP characterization based on 3D deformation of sulcal walls (in this case the STS walls). These deformations, named "wall pinches", provide a continuum between superficial PPs for which the pinches on each side are joining to make a clear bridge interrupting the sulcus (Figure 1-a), and deeply buried PPs for which the pinches are facing each other without joining thus showing no or little elevation on the fundus (Figure 1-b). Bodin et al. (2021) also showed that wall pinches are associated with the presence of specific dense u-shape fiber bundles locally joining the two banks of the sulcus. This study provided not

only a new definition of PPs, but confirmed an association with superficial white matter that had been hypothesized before (Regis et al., 2005) or observed at the level of the PPFM in trhe CS (Catani et al., 2012; Pron et al., 2021). Nevertheless, the manual identification of PPs based on this definition is tedious and prone to human errors. This is why we want to implement this definition in order to detect PPs in an automatic and more systematic manner. Starting from the work in Bodin et al. (2021), we propose to automate the detection of pli-de-passages in the superior temporal sulcus. In order to do so using the presence of wall pinches (WPs) on the sulcal banks, we have to keep in mind that on one hand there is no strict quantifiable definition of WPs, and on the other hand we have a dataset on which all WPs and PPs have been delineated by hand (see next section and (Bodin et al., 2021)). Therefore we decided to use supervised learning, using this dataset as a training set. In the next section we present the definition of features that locally characterise the presence of wall pinches, followed by a supervised machine learning-based method that can detect PPs using these features, hence automating the definition provided in (Bodin et al., 2021). After these methodological details, we present experiments on two different datasets to quantify the performances of our method and assess its ability to generalize to different datasets.

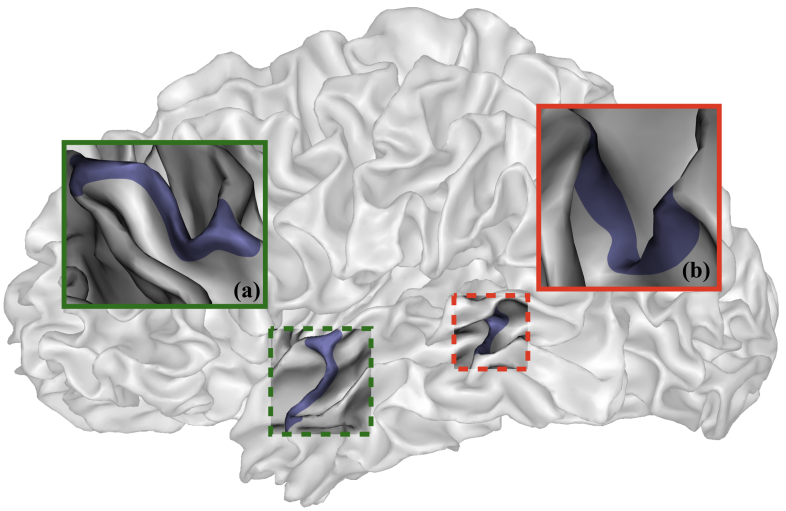


Figure 1: Local morphology of superficial and deep PPs. (a) Superficial PPs; (b) Deep PPs.

2. Materials and methods

2.1. Subjects and image acquisition

In this work, we used two independent datasets, where one is used for training and testing the machine learning models and the other for empirically evaluating the performance of our methods.

2.1.1. Human Connectome Project (HCP) database

The first dataset is composed of structural T1 MR images of 100 subjects from the Human Connectome Project (HCP) database, for which detailed information is available: `https://www.humanconnectome.org/study/hcp-young-adult/document/900-subjects-data-release`. Subjects were randomly having completed the full diffusion and structural acquisitions, being non-twins, right-handed, between 22 and 40 years old.

Data taken from the HCP database were acquired as follows: structural images were acquired using a modified version of Siemens Skyra 3T scanner (Siemens, Erlangen, Germany) with a maximum gradient strength of $100mT/m$, slew rate of $200T/m/s$ (reduced to $91T/m/s$ for diffusion due to peripheral nerve stimulation limits) and a 32-channel head coil. T1-weighted images were acquired using 3D MPRAGE sequence (TR/TE = $2400/2.14ms$, flip angle = 8, FOV = $224 \times 224mm^2$, resolution = $0.7mm$ isotropic). Note that this dataset have been already used in (Bodin et al., 2021).

2.1.2. Temporal Voice Areas (TVA) database

The second dataset is composed of structural T1 MR images of 92 subjects, aged between 17 and 44 years old, from the Temporal Voice Areas (TVA) dataset previously used in (Bodin et al., 2018). The origin dataset contains one-hundred and sixteen healthy subjects that were scanned (n = 116; 53 males, 63 females; mean age 23.7 ± 5.8) as part of published and unpublished experiments of the Voice Neurocognition Laboratory (http://vnl.psy.gla.ac.uk/) of the Institute of Neuroscience and Psychology at University of Glasgow. Participants, drawn mostly from the Glasgow student population, were of various ethnic backgrounds, education and manual lateralization and all provided written informed consent prior to participation. The experiments were approved by the local ethics committee at University of Glasgow. Exclusion of subjects was due to the failure for the depth profile extraction.

All scans were acquired on a 3T Siemens (Erlangen, Germany) Tim Trio scanner at the Centre for Cognitive Neuroimaging (http://www.ccni.gla.ac.uk/), University of Glasgow. High-resolution 3D T1-weighted sagittal scans were acquired for each subject (voxel size 1 mm^3 isotropic; acquisition matrix $256 \times 256 \times 192$).

2.2. Image preprocessing and identification of landmarks

2.2.1. Anatomical images and related maps

All individual T1-images were first segmented using Freesurfer (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki), and then imported into the *Morphologist* pipeline of the BrainVisa (BV) software (http://brainvisa.info) (Mangin et al., 2004) to generate triangular meshes of the grey/white matter interface for both hemispheres of all subjects (see Figure 2-a). These surfaces will be further referred to as ?cortical surface?.Meshes are reconstructed independently for each hemispheres and contain on average 60000 vertices.

Then, we generated several morphological surface maps used for anatomical analysis including sulcal depth, depth potential function (DPF) and curvature map. The depth map for each individual surface was computed as the geodesic distance from the sulcal fundus to the outer cortex (Rettmann et al., 2002) (see Figure 2-b).

The DPF map (Boucher et al., 2009) was computed using an in-house implementation available in BrainVisa, as already done in (Auzias et al., 2015). It is known to provide a regularized estimation of the sulcal depth that considers the information from both convexity and curvature. Importantly, it was also shown independent of brain size and therefore does not require a normalization procedure (Auzias et al., 2015). DPF measure can be either negative or positive depending on whether the vertex is superficial or located in the depth of a sulcus (see Figure 2-c).

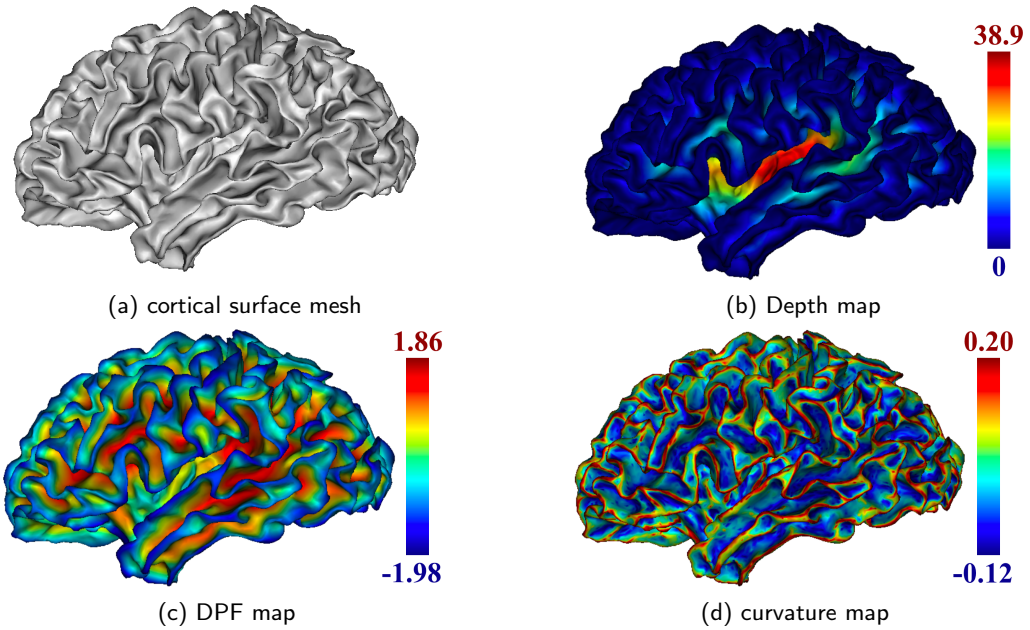
For all cortical surfaces, we generated their mean curvature maps using a finite element method as implemented in BrainVisa. In general, the vertices in the sulcal fundi have low negative curvature value while gyral crests show high positive curvature (see Figure 2-d).

2.2.2. STS identification and drawing

The STS is a fold separating the superior temporal gyrus (STG) from the middle temporal gyrus (MTG) in the temporal lobe (see Figure 3-a). The STS fundi were drawn semi-automatically on the cortical surface for both hemispheres and all subjects of both datasets using the SurfPaint module of the Anatomist visualization software (Le Troter et al., 2011). This was performed by C.B. in the context of (Bodin et al., 2018, 2021). We manually selected the anterior and posterior extremities based on anatomical landmarks identifiable in each subject as described in (Bodin et al., 2018). The anterior extremity was chosen as the tip of the temporal pole, posterior to polar temporal sulcus. In most cases, this sulcus was perpendicular to the STS and a clearly visible PP separates these two folds (Ochiai et al., 2004). The posterior extremity was chosen at the intersection between the STS horizontal main branch and its posterior ascending branches (Segal and Petrides, 2012). The fundus of the STS was drawn automatically between the two extremities as the shortest path that maximizes the DPF between the extremities (Le Troter et al., 2011).

2.2.3. PPs identification

Herein, the PPs were manually labeled on the HCP dataset by C. Bodin in the context of (Bodin et al., 2021) according to the morphological criteria described in (Bodin et al., 2021). Wall pinches were first identified and PPs were then drawn by selecting the two intersections (two vertices) between the extremities of WPs and gyral crests (STG and MTG). A line was then automatically generated between these two vertices, as the shortest path that minimizes the DPF (Bodin et al., 2021). Such PP lines were drawn for the STS on both hemispheres of the 100 HCP subjects (see examples on Figure 3-c). For all points of the STS fundus, those at the intersection with PPs lines were labelled “PP” and the rest were labelled “non-PP” (see Figure 3-d).



(a) cortical surface mesh (b) Depth map

(c) DPF map (d) curvature map

Figure 2: A cortical surface with depth, DPF, and curvature maps.

2.3. Feature extraction and feature images

In order to clearly identify WPs on the gyral walls, Bodin et al. (2021) used DPF and curvature maps on the cortical mesh (Figure 2). These maps were carefully examined at the individual levels, and thresholded to highlight distortions of the cortical surface. Specifically, the variation of DPF values (from blue to red) indicate variations of shape but are more sensitive to depth variations. Curvature maps are more appropriate to detect general shape variations but WPs that are not prominent enough are difficult to observe compared to shape variations at the top of gyri. Overall, both maps make it difficult to detect WPs in a systematic manner with a unique set of parameters across subjects (see Figure 4).

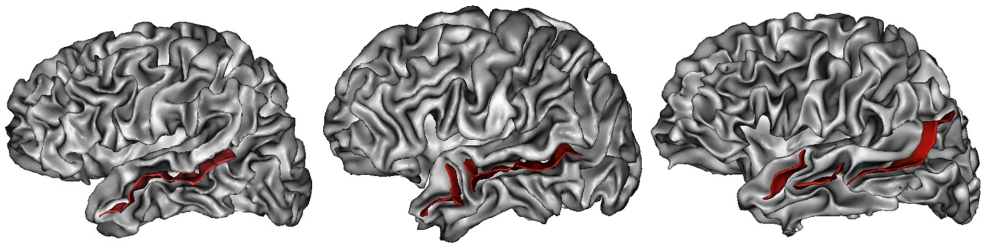
This is why we propose to use different geometrical maps in order to produce features that can reliably characterize wall pinches. In this work, we generated maps using surface profiling (Li et al., 2010) and we used these maps to build feature images for each point of the STS fundus. These feature images will then be used to classify fundus points as PP or non-PP.

2.3.1. Cortical surface profiling

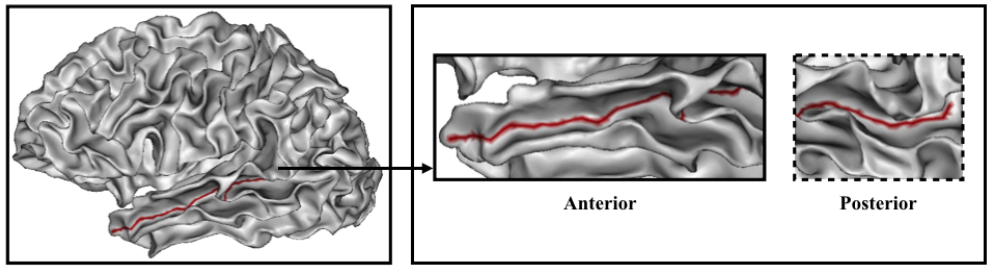
We want to associate each vertex on the STS fundus to a feature vector that characterizes the presence of wall pinches on the two surrounding sulcal walls. The first step is to build a cortical map that can capture the geometry of wall pinches. For this we are using features presented in (Li et al., 2010) and built using cortical surface profiling.

The basic idea of cortical surface profiling is, for each point of the cortical surface, to represent the 3D shape information of a disk-like cortical surface patch around the point, while modeling shape parameters of 2D profiles in all directions within this patch. These 2D profiles capture the shape of the surface in a specific direction around the point of interest. For each vertex, Li et al. (2010) combines a 3D Cartesian coordinate system (Tao et al., 2002) and a 2D polar coordinate system to built a new local 3D coordinate system (see Figure 5-a). For a vertex O on the cortical surface S, we can get its normal direction N and tangent plane P. Consider the normal direction N as the Z-axis in a 3D Cartesian coordinate system and the tangent plane P as the space of a polar coordinate system, with an arbitrary orientation origin R_0. Note that there is no specific given X-axis or Y-axis in this 3D coordinate system in comparison with the 3D Cartesian coordinate system.

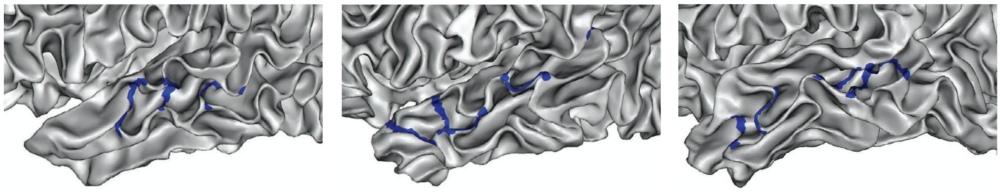
Based on this 3D coordinate system, for a given vertex and its local cortical surface patch, the profiling process is as follow. First we randomly select the starting direction R_0, and let R_α be the radial direction after rotating to an angle α from R_0. The profile of the local surface patch in the direction R_α is recorded as $C(\alpha, x, y)$, where x is the



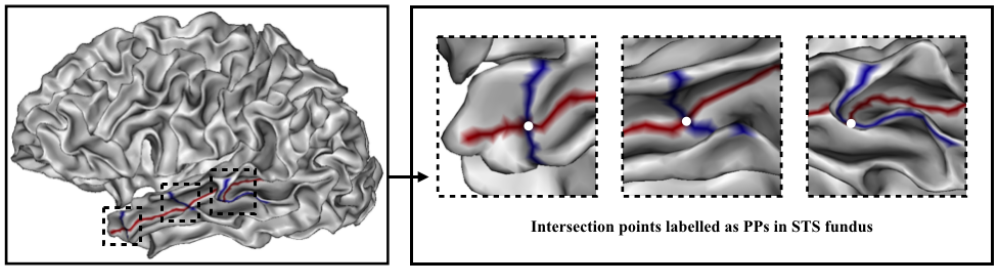
(a) STS



(b) STS fundus



(c) Labeled PPs in STS



(d) Vertices labelled as PPs in STS fundus

Figure 3: Superior Temporal Sulcus (a) Examples of three different STS. The model in red is the STS. Figures adapted from (Lefèvre and Mangin, 2010) (b) Example of STS fundus (in red) generated by the semi-automatic methods. (c) Examples of labeled PPs in STS. Figures adapted from (Bodin et al., 2021) (d) Examples of labelled PPs vertices (in white) in STS fundus.

radial distance of a point on profile to the normal direction N, and y is the normal distance of the point to the plane P The profiling process is conducted with an angular step $\Delta\alpha$ degrees, which will generate $360/\Delta\alpha$ profiles for the local cortical surface patch. Each profile is sampled with a radial step s, and the sampling stops when reached the maximum number of sampling points M. Figure 5-b shows some examples of sampling results. The essence of cortical surface profiling is using a collection of 2D profiles to sample the 3D local surface patch. In the following we used the exact same values than what has been presented in (Li et al., 2010) for the same purpose, i.e. : $\Delta\alpha = 5°$, $s = 0.1$, and $M = 45$.

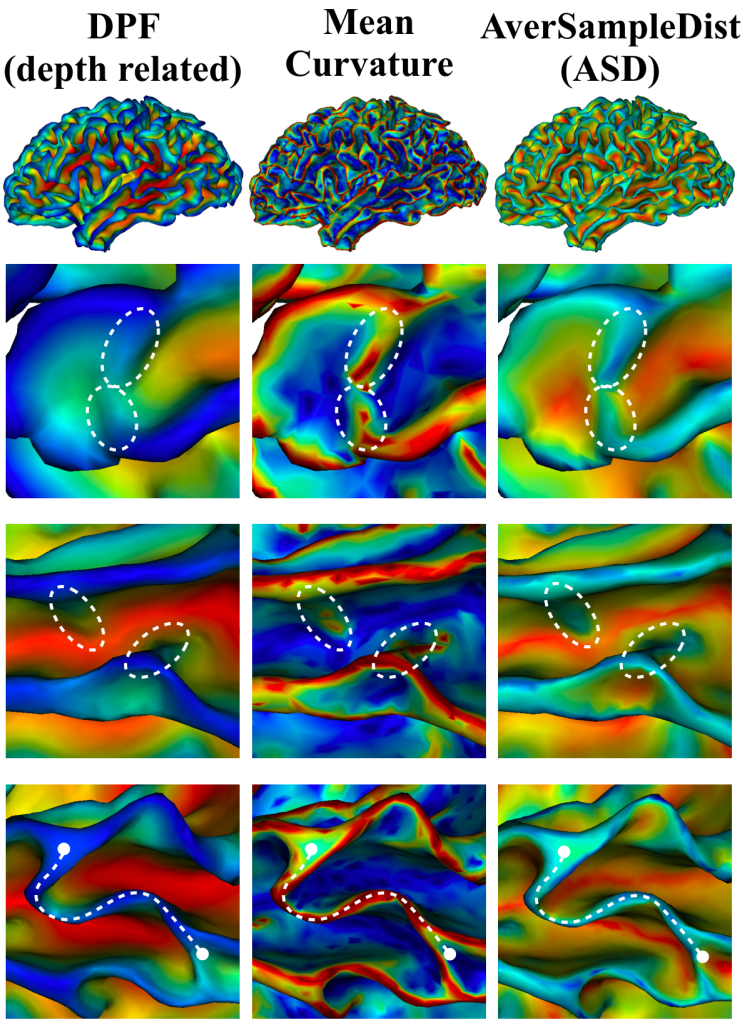


Figure 4: Different feature maps. DPF and curvature maps were used in (Le Guen et al., 2018; Bodin et al., 2021). ASD map was used in (Li et al., 2010). In the second and third rows, the dotted ellipses mark the WPs for the identification of PPs. In the last row, the dotted curve label a superficial PPs.

2.3.2. Computation of morphological feature maps

At every point of the surface, after applying surface profiling, we need to compute a feature that can efficiently characterize the presence of wall pinches on the gyri walls. To do so we propose to use the $AverSampleDis$ (ASD) that was introduced in (Li et al., 2010). ASD values are computed at every point by computing the average of the first order moment of the distance of all points of the profile to the tangent plane (see Figure 5-a), computed as follow:

$$F_{profile} = \frac{1}{M} \sum_{i=1}^{M} y_i \tag{1}$$

where M is the maximum number of points, y_i is the distance of each point on the profile to the tangent plane. The ASD map captures the properties we are interested in and display values on the WP that are equivalent to the values on gyral crests, as illustrated in 4.

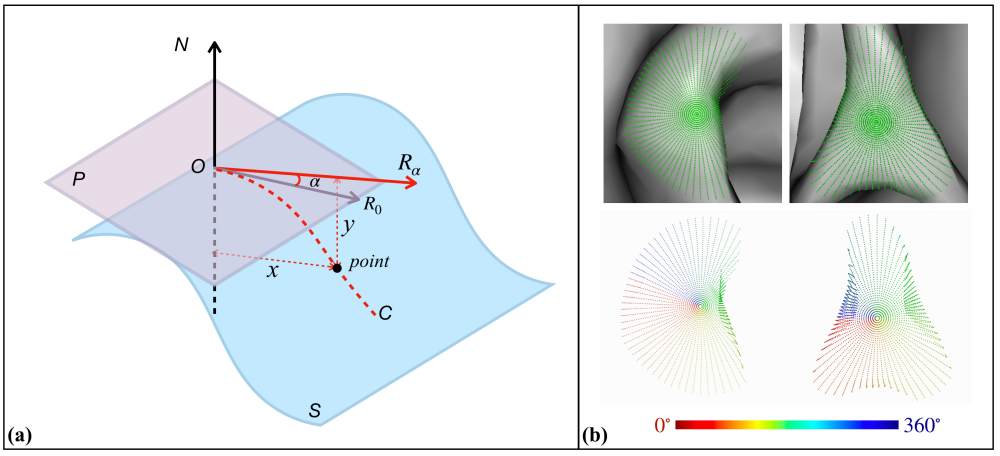


Figure 5: Cortical surface profiling. (a) An illustrative figure to show the coordinate system. O is any vertex on the cortical surface S; P is the tangent plane; N is the normal direction of vertex O; R_0 is the starting direction of sampling; R_α is the sampling direction with α degree away from R_0; C is the sampling profile at direction R_α. (b) Examples of profiles. Sampling parameters: $\Delta\alpha = 5°$, $s = 0.1$, and $M = 45$

2.3.3. Feature images

After building the geometrical map (ASD), we now want to associate a feature vector to each vertex on the fundus of the STS, such that it characterizes the presence of WPs. An illustration of the method is shown in Figure 6. The method is as follow: starting with the ASD map on the mesh of an hemisphere (see Figure 6-a), for each vertex on the STS fundus we use a second cortical surface profiling process to sample ASD values in a disk around the vertex, and build a feature vector. For this second round of surface profiling, we kept the parameter values $\Delta\alpha = 5°$, and $M = 45$, but the radius R of the disk has to be adapted to the depth of the STS, in order to capture enough of the surrounding STS wall geometry without including points that are outside the sulcus (see Figure 6-b). We tested $R = 9, 13.5, 18mm$, by running a simple 10-nearest neighbors (10-NN) classification with a structural similarity metric (SSIM) that provided a larger accuracy for R=13.5mm. For $M = 45$, it corresponds to a value of $s = 0.3mm$. The disk is aligned such that angle 0° corresponds to the direction of the fundus in the antero-posterior direction, which normalizes the orientation of the features across locations and subjects. The profiling results in a circular feature map that is then transformed into a rectangular feature image as shown in (Figure 6-c). This feature image therefore captures the geometry of the surrounding surface, in particular the potential presence of wall pinches (visible in blue on Figure 6-c). On account of the high sampling density, for any given local circular surface patch, the number of sampling points is larger than the number of mesh vertices included in the patch. Therefore, the ASD value at each sampling point is interpolated using barycentric coordinates.

After the feature extraction, each vertex of the STS fundus is associated with a local feature image. We therefore propose to use these feature images to solve the 3D PP recognition problem as a 2D image supervised classification problem.

2.4. Ensemble Support Vector Machines (EnsSVM)

We describe here the supervised learning strategy we use to detect PPs in the STS fundus.

As mentioned before, for all points of the STS fundus, after manual labelling those at the intersection with plis-de-passage were labelled “PP” and the rest were labelled “non-PP”. On the HCP dataset, we got a total of 15703 vertices of STS fundus, where only 865 of them were labelled as PPs, and 14208 were not, leading to an unbalanced data set with a ratio 1 : 16 and PP as the minority class. This could lead to difficulties as most algorithms perform better on balanced datasets.

2.4.1. Augmentation of PPs data

To deal with the imbalanced dataset, we proposed to augment the PPs instances by considering their nearest neighbours. If a point is a PP, then its nearest neighbours on the STS fundus also can be considered as PPs since their surrounding cortical geometry is very similar (see Figure 7-a, b). This increases the number of PPs instances

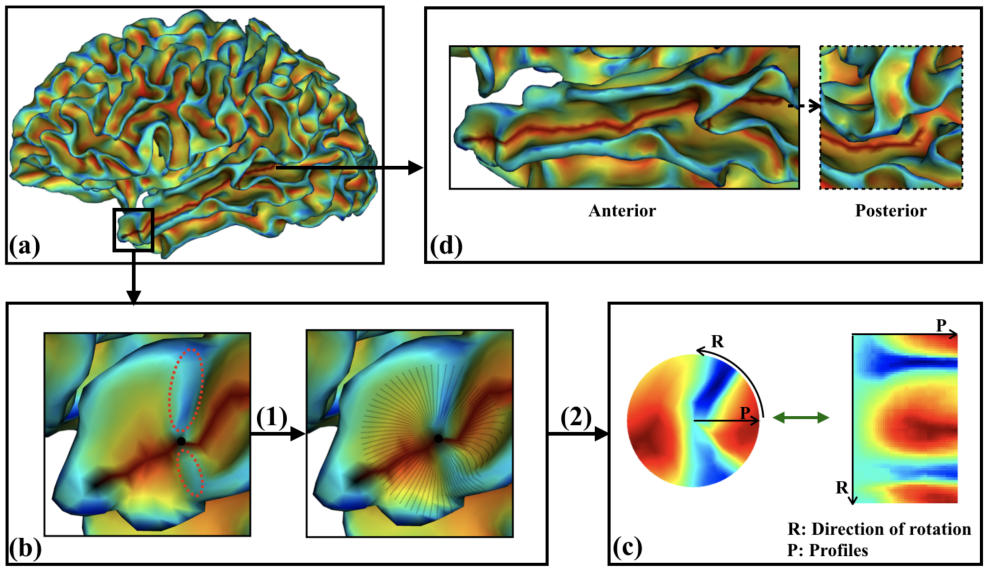


Figure 6: The framework of local feature extraction. (a) Surface with feature map ASD; (b) Surface profiling on a local region. Sampling parameters: $\Delta\alpha = 5°$, $s = 0.3$, and $M = 45$; (c) Feature images generation, from a disk to a rectangular image. (d) STS fundus (dark red).

while reducing the number of Non-PPs. Here we consider the two nearest neighbors, i.e. one on each side of the vertex. Hence by doing so, not only do we rebalance the dataset to some extent, but we also define a PP on the fundus as a small area (called "PP region") around the vertex instead of a single point. Notably, by considering the 2 nearest neighbours, a PP region contains at least 3 points on the STS.

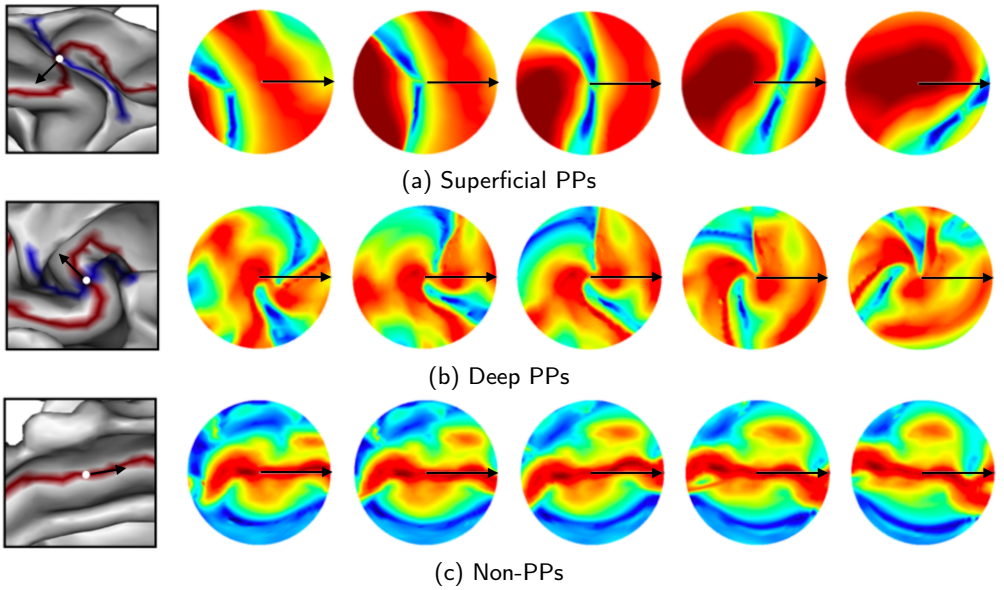


Figure 7: Feature images of neighbouring vertices on STS fundus. On the left is the STS fundus (red) and intersecting PPs (blue). On the right, the feature vector of 5 vertices is showed: the intersection of the fundus and the PP (white point on the left), and two nearest neighbors on each side. The arrows show the direction from the anterior to the posterior parts of the STS fundus.

2.4.2. EnsSVM pipeline

After the PP data augmentation, the unbalanced ratio is reduced to $1:5$. Although improved, it is still an very imbalanced dataset. In order to deal with this aspect and provide an adequate sampling of each category without bias due to the imbalance, we chose a strategy that has been succesful in many ensemble methods (Liu, 2009; Liu et al., 2011). Multiple subsets of the majority class N (non-PPs) are randomly generated $\{N_1, N_2, \ldots, N_T\}$ such that their size is the same as the minority class P, i.e.$|N_i| = |P|$. Then, the union of each pair of N_i and P is used to train a base classifier and a decision is taken by combining all base classifier results with a majority vote.

More precisely, we proposed the following ensemble SVM (EnsSVM) algorithm, illustrated in Figure 8. At first, we apply the augmentation strategy presented in section 2.4.1 and get a total of n_S PPs. We then divide the data into training and test data (note that this step is only needed for the algorithm evaluation, while for a final model applied to detect PPs in new data, we learn from the full dataset). After that, the majority (non-PP) data in the training set is under-sampled n_T times by bootstraping with replacement and the size of each subset is n_S. We then combine each bootstraped subset with the over-sampled positive instance (PPs) to form a perfectly balanced training dataset of size $2 * n_S$ that we use to train a base SVM. Finally, for a test instance, the classification decision is made through a majority voting from all n_T base classifier results.

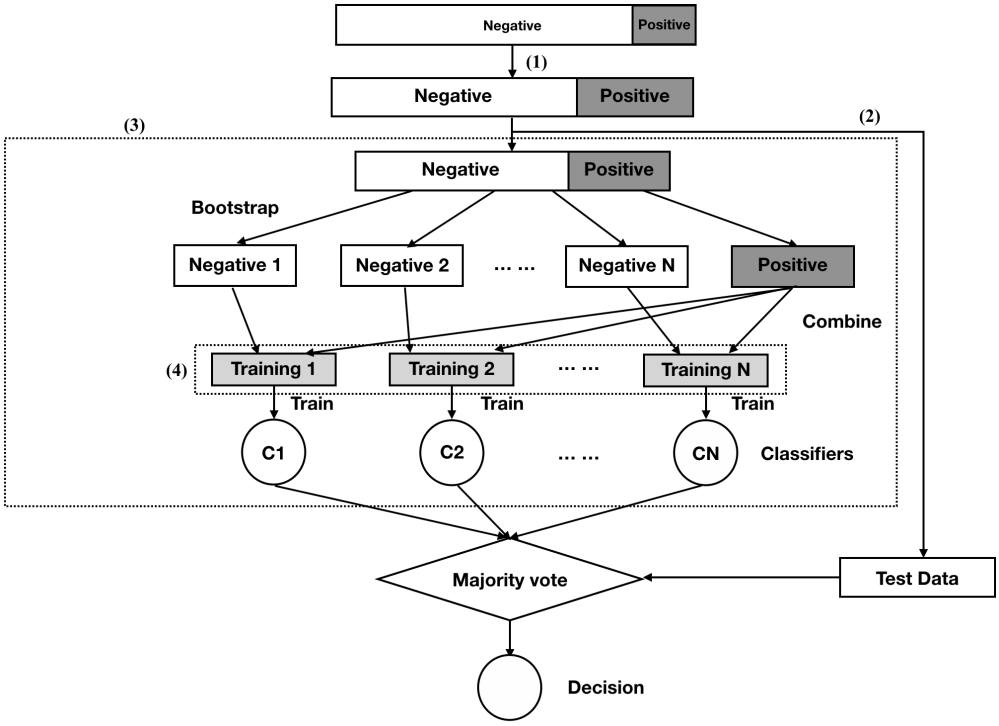


Figure 8: The ensemble SVM algorithm. (1) Augment the PPs instances by considering the two nearest neighbors of PPs as also PPs; (2) Divide the data into training and test data; (3) Build the base classifier; (4) The balanced datasets for training base classifiers.

2.5. Post-processing

Figure 9 gives an example of a result produced by our EnsSVM model. As we can see, instead of detecting single vertices, most of the results include several vertices forming a single connected region. This is expected given the fact that, as mentioned in section 2.4.1, after data augmentation we assume that a PP region on the STS fundus is composed of at least 3 consecutive points. Our hypothesis is that if a point is on a PP, its direct neighbors and possibly second order neighbors should as well, because their surrounding geometry is equivalent, and defining a cortical feature as a single point does not really make sense. On the other hand, a set of more than 5 consecutive points is on average more than 4mm wide and should be considered as containing more than one PP (e.g. when PPs are close enough they will

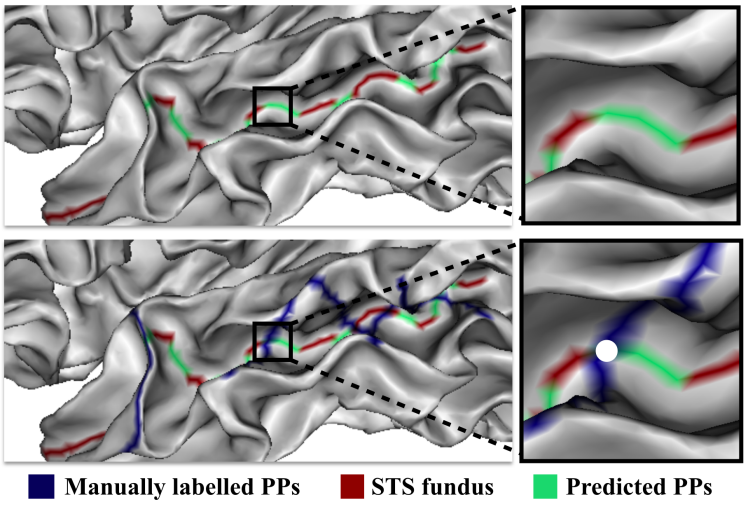


Figure 9: An example of predicted PPs using EnsSVM models.

Table 1
Confusion Matrix

	Ground-truth Positive	Ground-truth Negative
Predicted Positive	TP (True Positive)	FP (False Positive)
Predicted Negative	FN (False Negative)	TN (True Negative)

be detected as a single region). Therefore, once points of the STS fundus have been classified, we apply the following post-processing strategy:

1. For each vertice on STS fundus, get its prediction label and the associated probability from the SVMs.
2. Compute PP regions as connected components of the PP label set.
3. Discard all regions with less than 3 vertices.
4. For regions with 5 vertices or less, select the vertex with maximum probability to represent the PP.
5. For regions with more than 5 vertices,
   a. Split the region at the vertex with minimum PP probability.
   b. If sub-region contains 5 vertices or less, go to step 4. Otherwise, iterate step 5.

2.6. Evaluation measures on the HCP dataset

In this work, our evaluation measures are based on the $ConfusionMatrix$, a two-class problem confusion matrix with positive and negative values is illustrated in Table 1.

The choice of a performance index is critical to evaluate the goodness of a classifier. Traditionally, the overall performance of machine learning algorithm is usually evaluated with the "accuracy":

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

However, it is often meaningless to achieve high accuracy when the data is highly imbalanced. In our case, with a ratio of PPs/Non-PPs $\approx 1/16$, a straightforward method of guessing all instances as non-PPs would achieve an accuracy 93.75%. Although this accuracy seems high, it does not reflect the performance of the strategy used. Therefore a high overall accuracy may misrepresent the performance of classifiers in the context of an imbalanced class problem. Recent work on this specific problem led to the definition of different metrics to propose appropriate performance measures. Our performance measures are therefore expressed as follows:

Table 2
Modified Confusion Matrix

	Ground-truth Negative	Ground-truth Positive
Predicted Positive	TN (True Negative)	FN (False Negative)
Predicted Negative	FP (False Positive)	TP (True Positive)

- The *Recall* measures the accuracy of positive instances.

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

- The *G-mean* is the geometric mean of the accuracy of each class.

$$G\text{-}mean = \sqrt{Specificity \times Recall} \tag{4}$$

where the $Specificity = \frac{TN}{TN+FP}$. Here the sizes of different classes have already been considered so it is a good measurement for class-imbalanced learning.

- The *F-measure*

$$F\text{-}measure = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{5}$$

where the $Precision = \frac{TP}{TP+FP}$. It is designed as the harmonic mean of precision and recall, and can be generalized in F_β

$$F_\beta = (1 + \beta^2) \times \frac{Precision \times Recall}{\beta^2 \times Precision + Recall} \tag{6}$$

where β is a weight for precision and recall. The greater the β, the higher the domination of recall.

- The *Adjusted F-measure* (*AGF*) (Maratea et al., 2014)

$$AGF = \sqrt{F_2 \times InvF_{0.5}} \tag{7}$$

where F_2 is computed by setting $\beta = 2$ in F_β and $InvF_{0.5}$ is computed on the confusion matrix in Table 2:

Compared to the confusion matrix in Table 1, it switches the class labels of each samples (positive samples become negative and vice versa). Using the confusion matrix in Table 2, $InvF_{0.5}$ is computed as the standard $F_{0.5}$

Performances on the HCP dataset used for training and testing will be assessed using these measurements.

2.7. Statistical analysis on TVA data

After assessing results on the HCP dataset, we want to validate its use on an external dataset, the TVA dataset, and quantitatively compare the distribution of PPs in the STS with the manually identified PPs of the HCP data (Bodin et al., 2021), according to different depth ranges and PPs types. We also report the number of PPs per individual and compare the distribution between hemispheres per type of PPs. We also looked at the link between PPS and their position along the STS, as well as with the variations of sulcal depth along the STS fundus. A parametrisation of the STS fundus was performed as in (Cykowski et al., 2008; Bodin et al., 2018), resulting in 101 equally spaced positions along the fundus, with 0 being at the anterior extremity and 100 at the posterior extremity. Depth was computed at each of these positions using the same method as in (Rettmann et al., 2002; Bodin et al., 2018). All comparisons were performed using a non-parametric Wilcoxon signed rank test.

3. Results

In order to assess the performance of our EnsSVM model, we compared it to other machine learning algorithms. Each algorithm was trained and tested on the HCP dataset, with the evaluation measures presented in section 2.6. After this comparison, we present an application of our method to the independent TVA dataset, and discuss a statistical analysis of the prediction results compared to the manual labelling on the HCP dataset.

3.1. Comparison with alternative methods

Two different datasets have been generated in order to compare and evaluate the different algorithms: 1. Balanced dataset, where a subset of the negative dataset was randomly chosen to have the same size as positive dataset; 2. Imbalanced dataset after augmentation of PPs, where the ratio PPs : non-PPs $\approx 1:5$

In this section, we compare the performances of our proposed EnsSVM model with the following methods:

1. single SVM trained on the imbalanced dataset.(Chang and Lin, 2011)
2. single SVM trained on the imbalanced dataset, weighted by the 1 : 5 unbalance ratio. Labels are weighted inversely proportional to class frequencies.
3. single SVM trained on the balanced dataset.
4. bagging SVM trained on the balanced dataset with linear and radial basis function (RBF) kernels respectively.(Breiman, 1996)
5. random forest on the balanced dataset.(Breiman, 2001)
6. extra-trees on the balanced dataset.(Geurts et al., 2006)

Bagging, random forests, and extra-trees are other form of ensemble learning in order to deal with imbalanced data, which is why we chose them for comparison, and random forests are a an extension of Bagging. The main difference between bagging and our ensemble learning is the way the sub-training sets are generated. Bagging is done using the balanced dataset, and sub-training sets are generated randomly from this balanced dataset (bootstrap with replacement), whereas the EnsSVM approach is drawing subsets of the majority class in order to build multiple sub-training sets that are balanced and contain the entire minority class. In our experiments, we trained and tested all the algorithms on the HCP data with a 10-fold cross-validation. For hyper-parameter selection of models, we used a 10×10 nested cross-validation. Here, for each fold of cross-validation, the test data category has the same distribution as the original data and the feature images (or vectors) are generated using the ASD maps. Without specification, we used radial basis function (RBF)(Buhmann, 2003) kernels for all SVMs. In particular, the training of general bagging models and EnsSVM was performed using 21 component classifiers, and the Random Forest and Extra-trees used 500 decision trees (DT), which we determined empirically to be a high enough number to get robust and high performances. All base classifiers were implemented using the scikit-learn packages (Pedregosa et al., 2011).

Table 3 presents the performance for all the methods and the best results are in bold. As visible, EnsSVM achieves the best results in Recall, G-mean and AGF, and the Extra-trees of 500 decision trees performs best in F-measure, whereas the single SVM trained on the imbalanced dataset was the worst. In comparison with all single SVMs, ensemble methods show improved performances. In particular, this shows the importance of a proper sampling for the majority class (negative data) that has a significant impact on prediction results.

We tested the significance of these differences by performing statistical testing (non parametric signed-rank Wilcoxon) over the 10-fold cross-validation on the Recall values. As shown in Table 4, all differences were significant with at least $p<0.05$, except for the difference between SVM on the balanced dataset and SVM with Bagging.

3.2. Prediction results

In order to better understand our results, we focus on the prediction results of our EnsSVM model. Figure 10-a shows the Recall rate across different ranges of depth. Intuitively, the deeper a PP the more buried it is and the more difficult it is to detect. Our model yielded an excellent prediction (mean Recall > 90%) on the most superficial PPs with depth $\in [0, 5.0)mm$. As depth increases, the prediction accuracy decreases, as expected. In particular, for the deepest PPs, with depth above $20mm$, Recall falls below its overall mean (81.1%) but is still high with a score of 78.5%. The mean recall is above 80% for all depth up to $15mm$. In order to compare to previous work on PPs, we used the same

Table 3
The average values of model performance measures

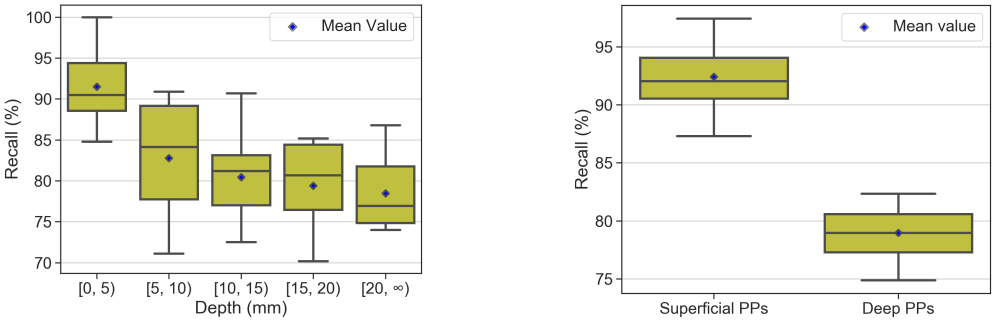
Models	Recall	G-mean	F-measure	AGF
EnsSVM	**0.811**	**0.773**	0.527	**0.774**
Bagging SVM	0.786	0.764	0.520	0.764
Bagging linear SVM	0.715	0.487	0.291	0.540
Random Forest 500DT	0.749	0.762	0.528	0.760
Extra-trees 500DT	0.761	0.766	0.531	0.764
SVM (balanced data)	0.781	0.764	0.522	0.764
SVM (weighted data)	0.704	0.766	**0.547**	0.762
SVM (imbalanced data)	0.298	0.540	0.423	0.550

Table 4
signed-rank Wilcoxon test results between classifiers, for *Recall* on 10-fold cross validation

Models	EnsSVM	BagSVM	BaglinSVM	RF500	ET500	SVMbal	SVMwei	SVMimb
EnsSVM	–	*	**	**	**	**	**	**
BagSVM	*	–	**	**	*	–	**	**
BaglinSVM	**	**	–	**	**	**	**	**
RF500	**	**	**	–	*	*	**	**
ET500	**	*	**	*	–	**	**	**
SVMbal	**	–	**	*	**	–	**	**
SVMwei	**	**	**	**	**	**	–	**
SVMimb	**	**	**	**	**	**	**	–

–: no significant difference, $*$: $p < 0.05$, $**$: $p < 0.01$

DPF-based criterion to separate deep and superficial PPs than the one used in (Bodin et al., 2021; Le Guen et al., 2018). In particular, in (Le Guen et al., 2018) only superficial PPs were detected. We then calculated Recall (Figure 10-b) for both categories. The mean Recall of superficial PPs and Deep PPs is 92.4% and 78.98% respectively. The above results indicate that the EnsSVM has a good prediction accuracy for the detection of PPs, including for the very difficult cases with PPs buried in the depth of the STS, and very good results for the more superficial PPs.



(a) Recall vs. Depth (b) Recall vs. Types of PPs

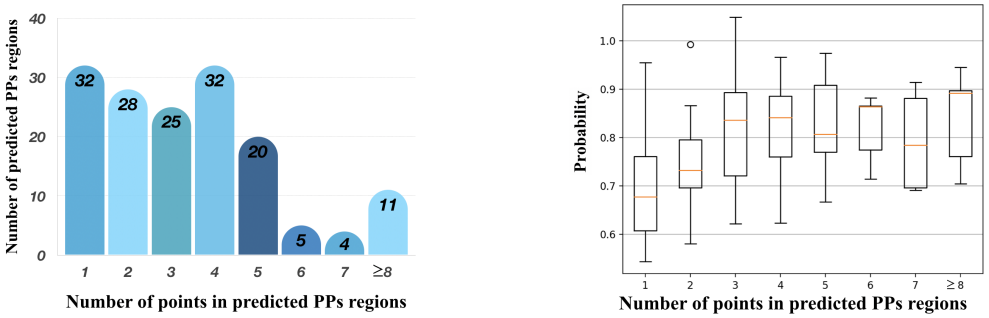
Figure 10: The Recall of EnsSVM across various categories of PPs.

3.3. Post-processing

To further test the post-processing algorithm, we randomly selected 10 subjects with manually labelled PPs (both left and right hemispheres) and used the remaining data to train the EnsSVM model. For each detected PP region, we count the number of points in the region (one region is a set of consecutive points labelled as PP), and show the resulting histogram in Figure 11-a. For each PP region, we computed the associated mean EnsSVM probability. The EnsSVM probability of a PP point is the mean probability of all SVM base classifiers.The probability value as a function of the

number of points in a PP region is shown in Figure 11-(b). As we can see, regions containing one or two points are frequent, with 32 and 28 occurrences respectively. However, they also have the lowest associated probability, with a median probability of 0.67 and 0.73 respectively, below the average probability of all predicted PPs regions containing three or more points (> 0.8).

Interestingly, after comparing with manual labels, the regions containing only one to two vertices are all false positive cases. This therefore confirms the choice we made for our post-processing strategy (section 2.5).



(a) (b)

Figure 11: Statistics of predicted PPs regions using EnsSVM on test dataset of 10 subjects. (a) we count the number of predicted regions with different points; (b) the average probability of each region is calculated by the average of the probabilities of the predicted PPs points in each region.

We also computed the number of regions that were detected as PPs after post-processing, and compared it to the number of PPs manually labelled. As shown in Figure 12, distributions of these numbers across depth intervals are similar between automatic and manual detection.

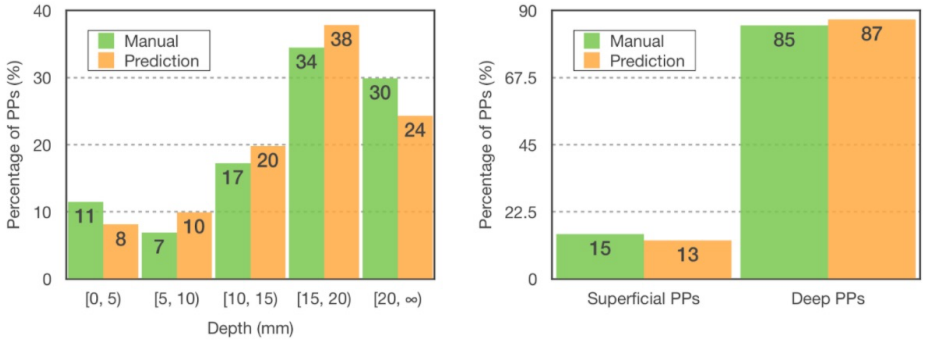


Figure 12: Comparison between the numbers of manually annotated and predicted labels on the testing set of the HCP data across depth ranges.

Beyond the accuracy of the detection, and since we detect regions of 3 to 5 points instead of single vertices, we wanted to test the spatial precision of the detection. We computed the distance between the vertex with highest probability value in each region with the closest manually labelled PP. The average distance is 4.53*mm*, which corresponds on average to 3 vertices (the average distance between consecutive vertices on the fundus of the STS is 1.57*mm* for our dataset). This might be considered slightly high but it is not completely informative because false positives are introducing a strong bias: a false positive having no real counterpart will have a possibly large distance to another PP and will therefore bias the average towards a higher value. On the 10 subjects of the testing set, 87.5% of the PPs are included in a detected region (which is in line with our average Recall score). For these, the mean distance is 1.72*mm*, which on average corresponds to a little more than one vertex away. Therefore we reach a good level of precision for features that we consider being at least 3 vertices wide. For the remaining 12.5%, the mean distance is 9.76mm, which confirms that our overall mean distance is being biased by these.

Table 5
Distances among different max PPs region size

Max region size	3	4	5	6	7	8	9
Distance	4.69	4.73	4.53	4.69	4.57	4.55	4.6

Finally, we investigated our post-processing choice of limiting region size to 5 vertices. We computed the same overall mean distance between detected regions and true PPs, with the maximum size of regions varying from 3 to 9. Results are given in Table 5. It shows that a maximum size of 5 provides the best performance, i.e. the minimum value ($4.53mm$), which also supports our post-processing choice.

3.4. Validation with external data

Using the classifiers trained on the HCP dataset, we conducted a validation on the external TVA dataset. We applied the trained classifiers, and the post-processing of regions, on the TVA dataset and some examples are shown in Figure 13. We then quantitatively compared the distribution of PPs in the STS with the one from the manually identified PPs on the HCP data.

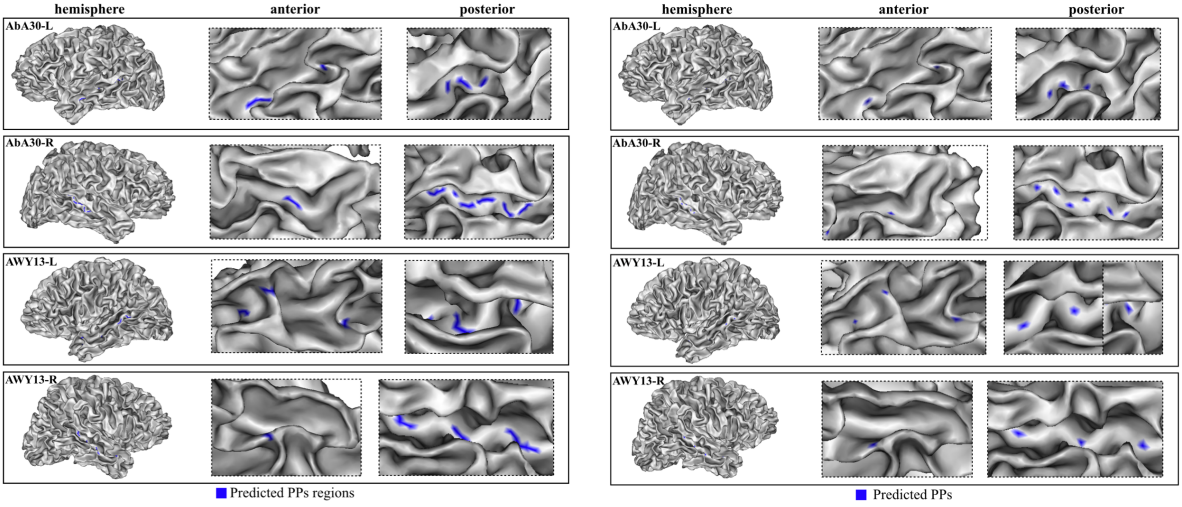
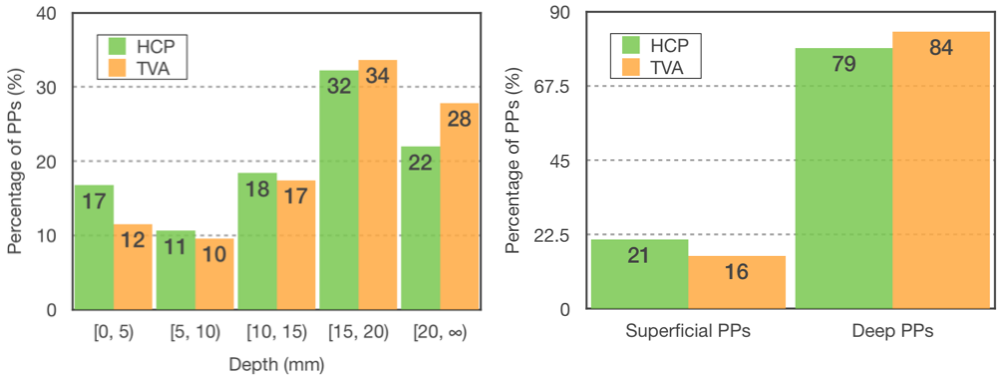


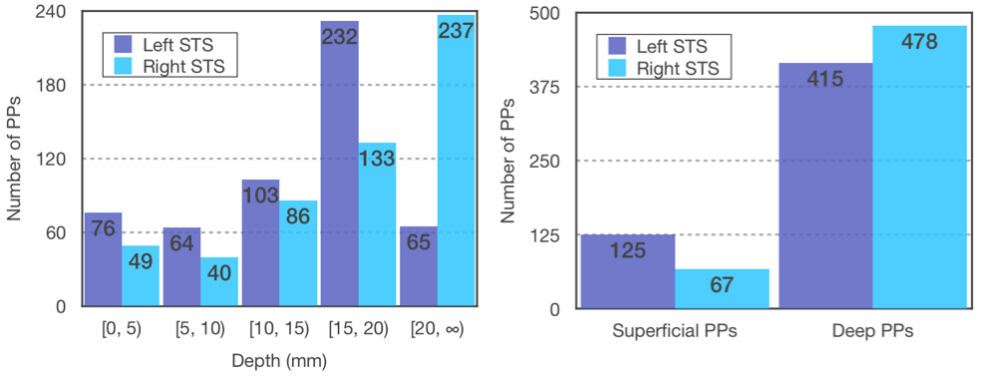
Figure 13: Examples of predicted PPs regions and selected PPs points. The LEFT figures show the PPs regions generated from the EnsSVM. The RIGHT figures show the selected PPs points using post-processsing algorithm.

On both hemispheres of the 92 subjects of the TVA data, 1085 PPs were detected, with 540 PPs in the left STS and 545 PPs in the right hemisphere. We detected 2 to 11 PPs in the left STS ($\text{Mean}_l = 5.86$) against 3 to 10 in the right STS ($\text{Mean}_r = 5.92$), and there was no significant difference between left and right STS across individuals ($p = 0.84$; Wilcoxon signed-rank test), which is identical to the manual results produced on the HCP data in (Bodin et al., 2021). It is noticeable that our automatic classification detected more PPs per hemisphere than the manual identification on the HCP data ($\text{Mean}_l = 4.5$, $\text{Mean}_r = 4.3$). Nevertheless, the distribution of the number of PPs across depth ranges or different types of PPs is similar, as shown in Figure 14-a.

Again, PPs were subdivided into into superficial and deep ones. In the left STS, the number of superficial PPs $N_s = 125$ and the number of deep PPs $N_d = 415$, with a ratio of $1 : 3.32$, while in the right STS, $N_s = 67$ and $N_d = 478$, with a ratio of $1 : 7.13$. The ratios of superficial PPs to deep PPs differed significantly between left and right STS across individuals ($p < 0.05$; Wilcoxon) which is consistent with manual results presented in (Bodin et al., 2021). This difference in distribution across depth ranges between left and right is illustrated in Figure 14(b). It is visible that the PPs in the right STS are located in deeper regions than the left (see examples in Figure 13), which again is consistent with (Bodin et al., 2021) and is in agreement with the fact that the right STS is notoriously deeper than the left STS (Leroy et al., 2015).



(a) HCP vs. TVA



(b) Left STS vs. Right STS

Figure 14: The prediction results of PPs

In order to further study the distribution of PPs along the STS, we counted the number of all PPs according to their coordinate range in the STS, as shown on the histogram in Figure 15. We found that there was no significant difference between left and right hemispheres per STS coordinates ($p = 0.71$; Wilcoxon). As we can see, the number of PPs between left and right hemisphere across ranges of STS coordinates is similar. These results show that the relative positions of PPs between left and right STS are stable at the group level.

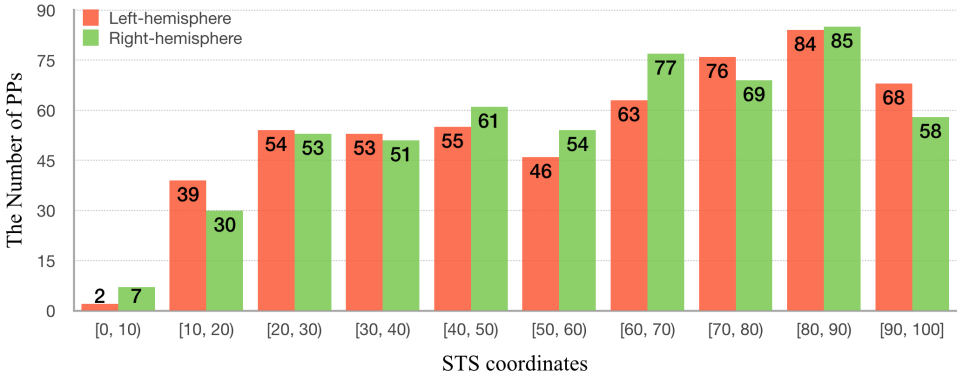


Figure 15: The distribution of PPs based on STS coordinates

We then specifically looked at the superficial PPs (SPs) along the STS at the group level. Figure 16 illustrates both the distribution of superficial PPs and the sulcal depth of the left and right STS. A rightward asymmetry of depth emerges almost simultaneously with a large difference in the number of superficial PPs between left and right: small differences in both depth and number of SPs are visible up to coordinate 48, then, from coordinate 49 to 75, a depth asymmetry becomes progressively larger together with a difference in the the number of SPs between left and right hemispheres. The depth asymmetry in the coordinate interval 49-93 corresponds to the "STAP" (Superior Temporal Asymmetrical Pit) previously described in (Leroy et al., 2015; Bodin et al., 2018). In this region, there is more SPs in the left hemisphere than in the right ($51 > 10$), and the number of SPs differs significantly between left and right hemispheres across individuals ($p < 0.0001$; Wilcoxon). In combination with the distribution of all PPs in this region, we found that the number of PPs was identical between left and right hemispheres in this region, while there were more SPs in the left-hemispheres than the right.

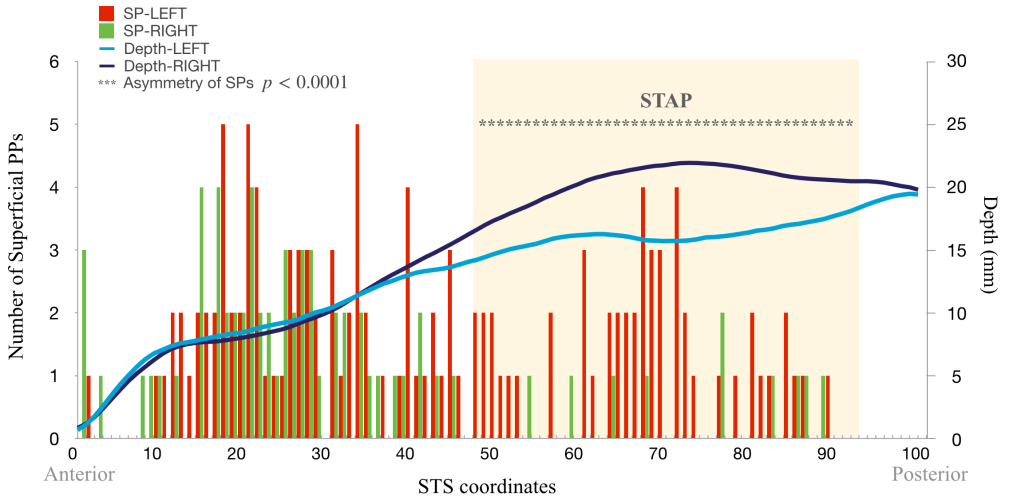


Figure 16: Superficial PPs and the STS depth asymmetry. The histogram shows the number of superficial PPs (SP) between left (in *red*) and right (in *green*) STS. The average depth of left (in *light blue*) and right (in *dark blue*) STS fundus, with depth values (mm) on the right vertical axis. The STAP region defined in (Bodin et al., 2018) is represented by the orange overlap.

4. Discussion

In this work, we were interested in the "annectant gyri" or "plis de passages" (PPs), which have been regarded as potential landmarks to partly explain the variability of cortical foldings across individuals. The PPs describe transverse gyri that interconnect both sides of a sulcus, are frequently buried in the depth of these sulci, and are sometimes apparent on the cortical surface. Based on this definition, the PPs are classified into two categories, namely the "Superficial PPs" and the "Deep PPs". For the superficial PPs, they always appeared with a clear variation of sulcal depth such that some automated algorithms have been implemented (Cykowski et al., 2008; Le Guen et al., 2018) but their use of a strict depth threshold leads to difficulties in distinguishing deep PPs and small fluctuations of the fundus depth (Ochiai et al., 2004). Inspired by the description of interlocking gyri (Cunningham, 1890a,b, 1897), a new general three-dimensional characteristic, named "wall pinches" (WPs), was proposed by (Bodin et al., 2021) that establishes a continuity between deep and superficial PPs. This local geometrical clue laid the foundation for the automatic detection of PPs. In our work, we proposed the first machine learning based method to automatically detect PPs based on these features, trained and tested it on the HCP dataset and validated it on an independent dataset (TVA).

Generally, our method provides a good prediction accuracy for PPs (*Recall* = 81.1%). Results on the independent TVA data also show a similar distribution of PPs across different ranges of depth or categories of PPs as the manual labels of HCP data. Both these results reinforce the robustness of our methods and its ability to identify WPs in order to detect PPs. Meanwhile, the prediction accuracy of PPs across different ranges of depth or categories of PPs also

shows the difficulty in identifying the deep PPs, but with still a Recall of 78.98%. Overall we are able to detect PPs at all depths and without the use of a threshold.

4.1. Selection of feature maps and sampling strategy

One main limitation of the manual labeling of PPs, besides the fact that it is a very tedious task, is the subjective assessment of the presence of WPs. In (Bodin et al., 2021), this is done by using manually tuned, subject dependent, thresholds on both DPF and curvature maps. It took a large amount of training for the main author (C. Bodin) to provide a robust and reproducible detection.

Besides, these maps have drawbacks considering the problem at hand: for the curvature map, very large curvatures are found on gyral crests, leading to much higher values compared to WPs; the DPF mostly shows variations of depth, and since depth represents geodesic distances to the gyral crown, the main direction of variation of the DPF is orthogonal to the WPs. To overcome these limitations, we implemented the *AverSampleDis* (ASD) maps introduced in (Li et al., 2010). Both experimental results and visualisation showed the superiority of new map in PPs identification compared to the DPF and curvature map.

The comparison of several learning algorithms (section 3.1) also stresses the importance of taking into account the imbalanced aspect of the problem. In particular it almost shows the importance of a proper sampling strategy to correct for this. Indeed, learning on the imbalanced dataset performs extremely poorly (SVM on the imbalanced dataset in Table 3) and learning on a balanced dataset does improve performances significantly. Nevertheless it involves subsampling the majority class by a factor 5 in our case, which can randomly result in a bad representation of this group. The efficiency of the ensemble SVM approach that repeats that subsampling in order to provide the best statistical representation of the majority class is visible in Table 3.

4.2. PPs as local landmarks in superior temporal sulcus (STS)

From the prediction results on the TVA dataset, we observe that the number of PPs between left and right STS is similar, which is consistent with the results in (Bodin et al., 2021) but different from the observation in (Le Guen et al., 2018), where they found more PPs in the left than in the right STS. This is caused by the fact that automatic methods applying a threshold on depth, such as the one proposed in (Le Guen et al., 2018), find their limitation when PPs are buried in the depth of sulci, or simply when PPs are visible on the floor of a sulcus that is overall deeper than average. This is the case for the STS, which is notably deeper in the right hemisphere than in the left (Leroy et al., 2015). Such PPs do not pass the thresholds and are therefore rejected as noise (Bodin et al., 2021). Besides, we observed more superficial PPs in the left STS than the right around the STAP region at the group level. It might be due to the fact that the fundus of the left STS is more shallow than on the right one, but it also possibly is the reverse relationship, with more superficial PPs in the left leading to a reduced average depth compared to the right. Nevertheless, the fact that the entire STAP can be detected at the individual level (Leroy et al., 2015) tends to favor the first hypothesis.

We also showed that there is no significant difference between left and right STS in the distribution of PPs across STS coordinates (along the sulcus), reinforcing the fact that PPs are potentially useful landmarks in the STS.

4.3. Wall pinches and their links with connectivity

Our method does not exactly provide the same number of PPs than manual labeling, in both hemispheres. Specifically, we find on the TVA dataset that we detect on average one more deep PP per hemisphere than the manual labelling performed on the HCP dataset. The exact nature of PPs and WPs need to be further explored to understand why we might have such (small) discrepancy and to detect the proper number of PPs, in particular deep ones. One limitation of our work is that we only used the local morphological information (shape features that characterize WPs) to detect PPs. Nevertheless, we showed in a previous study (Bodin et al., 2021) that there is a specific relationship between superficial white matter U-shape fiber connectivity. This can also been noted in other recent studies (Catani et al., 2012; Pron et al., 2021). It is worth also noting that recent work about ”cortical 3-hinges” or “3-hinge Gyri”, which are similar structures to WPs, have reported a stronger structural connectivity for these structures (Li et al., 2010; Ge et al., 2018; Zhang et al., 2020a,b). Finally, a recent Klingler dissection work (Shinohara et al., 2020) has studied cortical features called 3-way junctions, that has a striking resemblance with WPs and 3-hinge gyri. This study showed that 3-way junctions are locations of denser superficial white matter fiber terminations and possibly the hubs of superficial white matter connectivity networks. All these studies converge to suggest that WPs and therefore PPs have a strong relationship with local superficial white matter connectivity, and this information could be used in a automated detection algorithm to provide reliable maps of PPs.

5. Conclusion

In this paper, we present the first machine learning based method to automatically detect plis-de-passage (PP) in the superior temporal sulcus (STS). This difficult problem is addressed using geometrical descriptors and machine learning while taking into account the imbalanced nature of the problem. Our results are very consistent with manual labeling, and show similar distributions of PPs across two different datasets. They further confirm that the number of PPs between left and right STS are similar while there are more superficial PPs in the left hemisphere than the right.

We advocate for the fact that PPs are essential landmarks to understand sulcal variability across individuals (Bodin et al., 2021; Regis et al., 2005), and this work will open the way to their use for understanding cortical variability and organisation.

In order to do so, further work will focus on introducing local white matter connectivity information in our models, and we will study the possibility of generalization of our process to all sulci on the cortical surface.

6. Code availability

All codes for feature computation, classification using EnsSVM, and post-processing, are available on the following public repositery: https://github.com/tianqisong0117/EnsSVMforPPs

7. Acknowledgements

The HCP data were provided by the Human Connectome Project, WUMinn 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. Tianqi Song is supported by a grant from China Scholarship Council(No.201801810039). The project leading to this publication has received funding via the Laennec institute from the Excellence Initiative of Aix-Marseille Universite - A*Midex, a French “Investissements d‘Avenir programme” AMX-21-IET-017. The authors would like to thank Sylvain Takerkart for initial discussions of this work and Pascal Belin for providing the TVA dataset.

References

Auzias, G., Brun, L., Deruelle, C., Coulon, O., 2015. Deep sulcal landmarks: algorithmic and conceptual improvements in the definition and extraction of sulcal pits. Neuroimage 111, 12–25.

Bodin, C., Pron, A., Le Mao, M., Régis, J., Belin, P., Coulon, O., 2021. Plis de passage in the superior temporal sulcus: Morphology and local connectivity. Neuroimage 225, 117513.

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

Boling, W., Olivier, A., Bittar, R.G., Reutens, D., 1999. Localization of hand motor activation in broca’s pli de passage moyen. Journal of neurosurgery 91, 903–910.

Boucher, M., Whitesides, S., Evans, A., 2009. Depth potential function for folding pattern representation, registration and analysis. Medical image analysis 13, 203–214.

Breiman, L., 1996. Bagging predictors. Machine learning 24, 123–140.

Breiman, L., 2001. Random forests. Machine learning 45, 5–32.

Broca, P., 1888. Description elémentaires des circonvolutions cérébrales de l?homme. Mémoires d?anthropologie. Reinwald, Paris 707, 804.

Buhmann, M.D., 2003. Radial basis functions: theory and implementations. volume 12. Cambridge university press.

Catani, M., Dell?Acqua, F., Vergani, F., Malik, F., Hodge, H., Roy, P., Valabregue, R., De Schotten, M.T., 2012. Short frontal lobe connections of the human brain. cortex 48, 273–291.

Chang, C.C., Lin, C.J., 2011. LIBSVM: A library for support vector machines. ACM Transactions on Intelligent Systems and Technology 2, 27:1–27:27. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm.

Chen, H., Li, Y., Ge, F., Li, G., Shen, D., Liu, T., 2017. Gyral net: A new representation of cortical folding organization. Medical image analysis 42, 14–25.

Cunningham, D.J., 1890a. The complete fissures of the human cerebrum, and their significance in connection with the growth of the hemisphere and the appearance of the occipital lobe. Journal of anatomy and physiology 24, 309.

Cunningham, D.J., 1890b. The fissure of rolando. Journal of anatomy and physiology 25, 1.

Cunningham, D.J., 1897. The rolandic and calcarine fissures?a study of the growing cortex of the cerebrum. Journal of anatomy and physiology 31, 586.

Cykowski, M.D., Coulon, O., Kochunov, P.V., Amunts, K., Lancaster, J.L., Laird, A.R., Glahn, D.C., Fox, P.T., 2008. The central sulcus: an observer-independent characterization of sulcal landmarks and depth asymmetry. Cerebral cortex 18, 1999–2009.

Ge, F., Li, X., Razavi, M.J., Chen, H., Zhang, T., Zhang, S., Guo, L., Hu, X., Wang, X., Liu, T., 2018. Denser growing fiber connections induce 3-hinge gyral folding. Cerebral Cortex 28, 1064–1075.

Geng, G., Johnston, L.A., Yan, E., Britto, J.M., Smith, D.W., Walker, D.W., Egan, G.F., 2009. Biomechanisms for modelling cerebral cortical folding. Medical image analysis 13, 920–930.

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, 334–356.

Geurts, P., Ernst, D., Wehenkel, L., 2006. Extremely randomized trees. Machine learning 63, 3–42.

Gratiolet, L.P., 1854. Mémoire sur les Plis cérébraux de l'Homme et des Primatès: Mit Einem Atlas (4 pp. XIV pl.) in fol. 33i. A. Bertrand.

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

Le Guen, Y., Leroy, F., Auzias, G., Riviere, D., Grigis, A., Mangin, J.F., Coulon, O., Dehaene-Lambertz, G., Frouin, V., 2018. The chaotic morphology of the left superior temporal sulcus is genetically constrained. Neuroimage 174, 297–307.

Le Troter, A., Riviere, D., Coulon, O., 2011. An interactive sulcal fundi editor in brainvisa, in: 17th International Conference on Human Brain Mapping, Organization for Human Brain Mapping.

Lefèvre, J., Mangin, J.F., 2010. A reaction-diffusion model of human brain development. PLoS computational biology 6, e1000749.

Leroy, F., Cai, Q., Bogart, S.L., Dubois, J., Coulon, O., Monzalvo, K., Fischer, C., Glasel, H., Van der Haegen, L., Bénézit, A., et al., 2015. New human-specific brain landmark: the depth asymmetry of superior temporal sulcus. Proceedings of the National Academy of Sciences 112, 1208–1213.

Li, K., Guo, L., Li, G., Nie, J., Faraco, C., Cui, G., Zhao, Q., Miller, L.S., Liu, T., 2010. Gyral folding pattern analysis via surface profiling. NeuroImage 52, 1202–1214.

Liu, T.Y., 2009. Easyensemble and feature selection for imbalance data sets, in: 2009 international joint conference on bioinformatics, systems biology and intelligent computing, IEEE. pp. 517–520.

Liu, Y., Yu, X., Huang, J.X., An, A., 2011. Combining integrated sampling with svm ensembles for learning from imbalanced datasets. Information Processing & Management 47, 617–631.

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 32, 1035–1048.

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, 968–982.

Maratea, A., Petrosino, A., Manzo, M., 2014. Adjusted f-measure and kernel scaling for imbalanced data learning. Information Sciences 257, 331–341.

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, 706–719.

Ono, M., Kubik, S., Abernathey, C.D., 1990. Atlas of the cerebral sulci. Thieme Medical Publishers.

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.

Pron, A., Deruelle, C., Coulon, O., 2021. U-shape short-range extrinsic connectivity organisation around the human central sulcus. Brain Structure and Function 226, 179–193.

Rabiei, H., Richard, F., Coulon, O., Lefèvre, J., 2016. Local spectral analysis of the cerebral cortex: new gyrification indices. IEEE transactions on medical imaging 36, 838–848.

Regis, J., 1994. Deep sulcal anatomy and functional mapping of the cerebral cortex (in french).

Régis, J., Mangin, J.F., Frouin, V., Sastre, F., Peragut, J., Samson, Y., 1995. Generic model for the localization of the cerebral cortex and preoperative multimodal integration in epilepsy surgery. Stereotactic and Functional Neurosurgery 65, 72–80.

Regis, J., Mangin, J.F., Ochiai, T., Frouin, V., Rivière, D., Cachia, A., Tamura, M., Samson, Y., 2005. ?sulcal root? generic model: a hypothesis to overcome the variability of the human cortex folding patterns. Neurologia medico-chirurgica 45, 1–17.

Rettmann, M.E., Han, X., Xu, C., Prince, J.L., 2002. Automated sulcal segmentation using watersheds on the cortical surface. NeuroImage 15, 329–344.

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

Shinohara, H., Liu, X., Nakajima, R., Kinoshita, M., Ozaki, N., Hori, O., Nakada, M., 2020. Pyramid-shape crossings and intercrossing fibers are key elements for construction of the neural network in the superficial white matter of the human cerebrum. Cerebral Cortex 30, 5218–5228.

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, 4125–4152.

Tao, X., Prince, J.L., Davatzikos, C., 2002. Using a statistical shape model to extract sulcal curves on the outer cortex of the human brain. IEEE Transactions on Medical Imaging 21, 513–524.

Toro, R., Burnod, Y., 2005. A morphogenetic model for the development of cortical convolutions. Cerebral cortex 15, 1900–1913.

Toro, R., Perron, M., Pike, B., Richer, L., Veillette, S., Pausova, Z., Paus, T., 2008. Brain size and folding of the human cerebral cortex. Cerebral cortex 18, 2352–2357.

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

Yu, P., Grant, P.E., Qi, Y., Han, X., Ségonne, F., Pienaar, R., Busa, E., Pacheco, J., Makris, N., Buckner, R.L., et al., 2007. Cortical surface shape analysis based on spherical wavelets. IEEE transactions on medical imaging 26, 582–597.

Zhang, T., Huang, Y., Zhao, L., He, Z., Jiang, X., Guo, L., Hu, X., Liu, T., 2020a. Identifying cross-individual correspondences of 3-hinge gyri. Medical image analysis 63, 101700.

Zhang, T., Li, X., Jiang, X., Ge, F., Zhang, S., Zhao, L., Liu, H., Huang, Y., Wang, X., Yang, J., et al., 2020b. Cortical 3-hinges could serve as hubs in cortico-cortical connective network. Brain imaging and behavior , 1–18.

Zilles, K., Armstrong, E., Schleicher, A., Kretschmann, H.J., 1988. The human pattern of gyrification in the cerebral cortex. Anatomy and embryology 179, 173–179.

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, 1268–1283.

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