Automatic Detection of Plis De Passage in the Superior Temporal Sulcus using Surface Profiling and Ensemble SVM

Tianqi Song, Clémentine Bodin, Olivier Coulon

AUTOMATIC DETECTION OF PLIS DE PASSAGE IN THE SUPERIOR TEMPORAL SULCUS USING SURFACE PROFILING AND ENSEMBLE SVM

Tianqi Song, Clémentine Bodin, Olivier Coulon

Institut de Neurosciences de la Timone, Aix-Marseille Univ, UMR CNRS 7289, Marseille, France

ABSTRACT

Cortical folding, an essential characteristic of the brain cortex, shows variability across individuals. Plis de passages (PPs), namely annectant gyri buried inside the fold, can explain part of the variability. However, a systematic method of automatically detecting all PPs is still not available. In this paper, we present a method to detect the PPs on the cortex automatically. We first extract the geometry information of the localized areas on the cortex via surface profiling. Then, an ensemble support vector machine (SVM) is developed to identify the PPs. Experimental results show the effectiveness and robustness of our method.

***Index Terms*—** Plis de passage, Cortical folding, Cerebral Cortex, Machine learning, SVM

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, anatomist were interested in its organisation and features. The concept of "Pli de passage" (PPs) 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. Several authors have pointed out that they provided an understanding of the variable interruptions in sulci across individuals, and it was recently demonstrated that PPs can be a useful morphological landmark for models of cortical foldings [1, 2, 3, 4].

PPs have been little studied, mostly in the central sulcus (CS) and superior temporal sulcus (STS). They have been described as protrusions in the fundus of sulci between the two walls, and were therefore generally extracted as local minima of the sulcal depth along the fundus, with a threshold on the depth value [5]. Unfortunately, this definition missed a significant number of deep PPs, and detecting the most buried PPs is a difficult problem since they do not show any depth variations at the fundus. This led to controversial results such as an asymmetry of the number of PPs in the STS between the left and right hemispheres [4]. For this reason, [4] proposed a new morphological characterization of PPs in the STS using the surrounding surface geometry to solve this problem: as shown in Fig.1, the PPs can be characterized by a local deformation of the sulcus walls, named *wall pinches*, which can be observed clearly for superficial PPs but also exist for the deepest buried PPs. This feature establishes a continuum from a superficial apparent transverse gyrus to a completely buried PP with only wall deformations left. It has been shown in [4] that this characterization leads to PPs that are associated with a specific U-shape superficial white matter connectivity. In this paper, we propose to automate the detection of PPs in the STS using machine learning and this new morphological characterization.

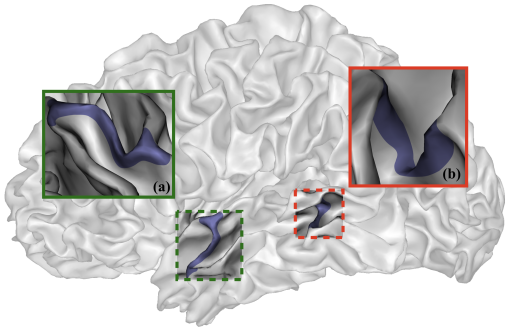


Fig. 1. Local morphology to identify the superficial and deep PPs. (a) Superficial PPs; (b) Deep PPs.

2. MATERIALS AND METHODS

2.1. Subjects and image acquisition

In this work, we used two completely independent datasets, where one is used for training and test the machine learning models and the other for empirically evaluating the performance of our methods. The first dataset is composed of structural T1 MR images of 100 subjects from the the Human Connectome Project (HCP) database, randomly selected, being non-twins, right-handed, between 22 and 40 years old. They have been already used in [4]. The second is composed of structural T1 MR images of 92 subjects, aged between 17 and 44 years old, from the TVA dataset previously used in [6]. Please refer to these two publications for a detailed description of the MR images.

2.2. Preprocessing and identification of PPs

First, we used the Brainvisa software (http://brainvisa.info) on the HCP and TVA data to obtain triangular meshes of the inner cortical surface for both hemispheres of all subjects. The superior temporal sulcus (STS) fundi of each subject were drawn semi-automatically on the cortical meshes (see Fig. 2(d)) using the SurfPaint module of the Anatomist visualization software [7], as presented in [6]. On the HCP subject cortical surfaces, PPs on the STS fundus were manually drawn, as described in [4]. They constitute our reference and learning set. For all points of the STS fundus, those at the intersection with plis-de-passage were labelled PP and the rest were labelled non-PP. We got a total of 15073 vertices, where only 865 of them were labeled as PPs, and 14208 were not. It is a typically unbalanced dataset, with PP as the minority class, which lead to difficulties as many algorithms perform better on a balanced dataset.

2.3. Local feature extraction and feature image

The first step of our method is to capture the local geometrical features of interest at each STS fundus vertex, namely the *wall pinches*, despite their large variability. In order to do so, we used cortical surface profiling [8] to build an appropriate geometrical map on the cortical surface (Fig. 2(a)). In particular, as shown in Fig. 2(a-b), we found that the $AverSampleDis$ maps [8] captured the properties we are interested in and clearly displays the presence of wall pinches.

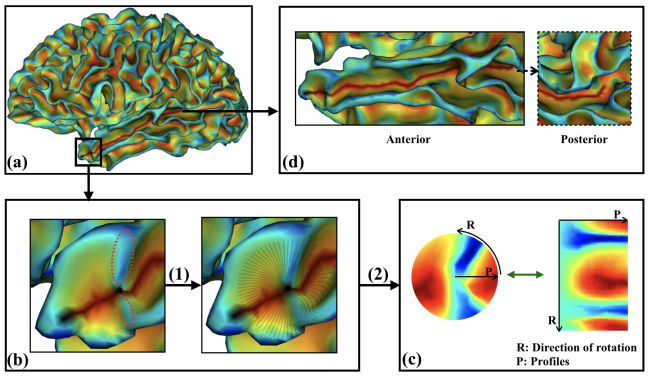


Fig. 2. The framework of local feature extraction.(a) Surface with feature map $AverSampleDis$; (b) Surface profiling on a local region; (c) Feature images; (d) STS fundus (dark red). Steps: (1) Surface profiling; (2) Generate feature images.

Examples of *wall pinches* can be seen on Fig. 2(b) circled with red dashed lines in the case of a quite superficial PP. Once the $AverSampleDis$ computed, at each point of the fundus we build a feature vector as follows. Values around the point are sampled using surface profiling in a disk of radius $r = 13.5mm$ (Fig. 2(b)). 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 resulting circular feature map that is then transformed into a rectangular feature image as shown in (Fig. 2(c)). This feature image capture the geometry of the surrounding surface, in particular the potential presence of wall pinches (visible in blue on Fig.2(c)).

2.4. Ensemble SVM

After the feature image extraction, each vertex of the STS fundus corresponds to local feature image and the three-dimensional (3D) PP recognition problem is converted to a 2D image classification problem. Each feature image has to be classified as PP or non-PP. As mentioned above, the classification of PPs is a learning task from a very unbalanced dataset (ratio $1:16$). To deal with such situation, our method is twofold: first, we applied a data augmentation strategy by considering that if a point is a PP, then its two neighbours on the STS fundus are also considered PPs, and their surrounding cortical geometry is similar. This reduces the unbalance ratio to $1:5$; then, we chose a learning framework dedicated to such situation, the *ensemble SVM* framework (EnSVM) proposed in[9], that consists in repetitively sampling (bootstrapping) the negative dataset in order to train a series of *balanced* SVM classifier. The classification decision is eventually made by majority voting. Fig. 3 illustrates our method.

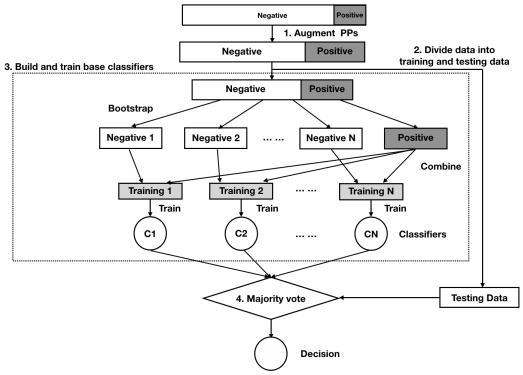


Fig. 3. The ensemble SVM algorithm.

In order to assess this strategy, we compared our approach to other SVM-based methods, namely:

1. a single SVM trained on the augmented dataset.
2. a single SVM trained on the augmented dataset, weighted by the $1:5$ unbalance ratio. Labels are weighted inversely proportional to class frequencies in the original data.
3. a single SVM trained on a balanced version of the augmented dataset, where a subset of the negative dataset was randomly chosen to get a ratio $1:1$.

All SVMs were used with a Gaussian radial basis function kernel, and the training of EnsSVM was performed using 21

component classifiers, which we determined empirically to be a high enough number to get robust and high performances.

2.5. PP regions and post-processing

After our data augmentation strategy presented above, we assume that a PP region on the STS fundus is composed of at least three consecutive points, and a set of more than 5 consecutive points should be considered as containing more than one PP (e.g. when PPs are close enough they will be detected as a single region). Therefore, once the 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.

3. RESULTS AND DISCUSSION

To assess the performance of our method, each algorithm was trained and tested on the HCP dataset, with a number of performance metrics presented below. Then the method was applied to the independent TVA dataset, and results are discussed.

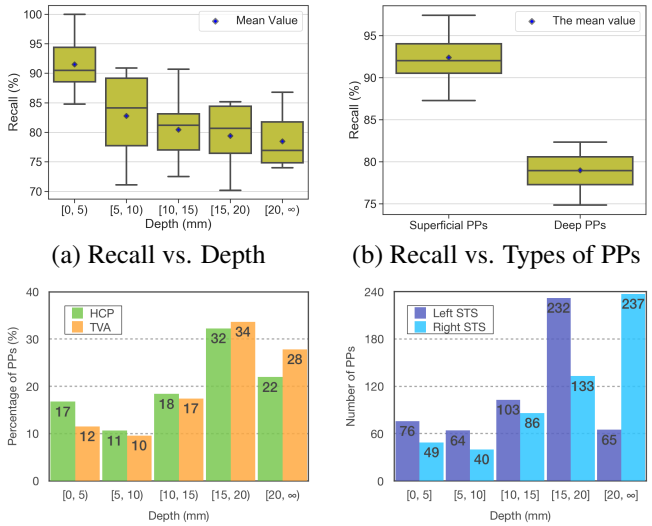
3.1. Experiments and validation framworks

The EnsSVM as well as the 3 other strategies were trained and tested with a 10-fold cross-validation. The evaluation measures used in our experiments were based on the confusion matrix. Generally, the performance of a machine learning algorithm is evaluated with accuracy, but this could be inappropriate for unbalanced data [9]. Hence, we chose several other performance measures: Recall, G-mean, F-measure and Adjusted F-measure (AFG) [9, 10]. Recall measures the accuracy of positive cases, which corresponds to the PPs in our data. G-mean is a metric that can measure the balance between classification performances on both the majority and minority classes. F-measure and AFG combine the recall and precision on the positive class. They measure the overall performance on the minority class. The performances for each method are presented in Table 1. As visible, EnsSVM achieves the best results on all the measures, whereas the single SVM trained on the original data was the worst. In particular, this also demonstrated that a proper way to sampling the negative data had a significant impact on prediction.

Table 1. The Average value of model performance measures

Models	Recall	G-mean	F-measure	AGF
Ensemble SVM	0.811	0.773	0.782	0.775
SVM (balanced data)	0.781	0.765	0.769	0.765
SVM (weighted data)	0.704	0.766	0.753	0.769
SVM	0.298	0.541	0.453	0.649

In order to better understand our results, we focused on the prediction results of our EnsSVM model. Fig.4(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. In particular, for the deepest PPs, whith depth $\in [20, \infty)mm$, 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 below $15mm$. Using the same criteria than [4], we divided PPs in two groups, superficial and deep, and calculated Recall (Fig.4(b)) for both groups. The mean Recall of superficial PPs and Deep PPs is 92.4% and 78.98% respectively. The above results indicate that the ensemble SVM has a relatively good prediction accuracy for the detection of PPs, including for the very difficult cases with PPs buried in the depth of the STS.



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

(c) HCP vs. TVA (d) Left STS vs. Right STS

Fig. 4. The prediction results of PPs

3.2. Validation on TVA Data

In this section, we applied our method on the TVA data and quantitatively compared the distribution of PPs in the STS with the manually identified PPs from the HCP data.

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 (Mean$_l$ $= 5.86$) against 3 to 10 in the right STS (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 [4]. It is noticeable that our automatic classification detected more PPs per hemisphere than the manual identification on the HCP data (Mean$_l$ $= 4.5$, Mean$_r$ $= 4.3$). Nevertheless, the distribution of the number of PPs across depth ranges is similar, as shown in Fig.4(c).

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 [4]. This difference in distribution across depth ranges between left and right is illustrated in Fig.4(d). It is visible that the PPs in the right STS are located in deeper regions than the left, which again is consistent with [4] and is in agreement with the fact that the right STS is notoriously deeper than the left STS [11] .

4. CONCLUSION

In this paper, we present the first method to automatically detect plis-de-passage in the STS. This difficult problem is adressed using geometrical descriptors and machine learning while taking into account the unbalanced nature of the problem (many more points in the STS are not PPs than PPs). Our results are very consistent with manual labeling, and show similar distributions of PPs across two different dataset. As PPs have been shown to be associated with a specific type of superficial white matter, further work will focus on introducing local connectivity information and use multi-feature machine learning models to improve our results.

5. ACKNOWLEDGMENTS

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 authors would like to thank Sylvain Takerkart for initial discussions of this work and Pascal Belin for providing the TVA dataset.

6. REFERENCES

[1] J. Régis, J.-F. Mangin, T. Ochiai, et al. ""sulcal root" generic model: a hypothesis to overcome the variability of the human cortex folding patterns." *Neurologia medico-chirurgica*, volume 45, pp. 1–17, 2005.

[2] V. Zlatkina and M. Petrides. "Morphological patterns of the postcentral sulcus in the human brain." *Journal of Comparative Neurology*, volume 518, pp. 3701–3724, 2010.

[3] J.-F. Mangin, Y. Le Guen, N. Labra, et al. ""plis de passage" deserve a role in models of the cortical folding process." *Brain Topography*, pp. 1–14, 2019.

[4] C. Bodin, A. Pron, M. Le Mao, et al. "Plis de passage in the superior temporal sulcus: Morphology and local connectivity." *bioRxiv*, 2020. URL `https://www.biorxiv.org/content/early/2020/05/28/2020.05.26.116152`.

[5] Y. Le Guen, F. Leroy, G. Auzias, et al. "The chaotic morphology of the left superior temporal sulcus is genetically constrained." *Neuroimage*, volume 174, pp. 297–307, 2018.

[6] C. Bodin, S. Takerkart, P. Belin, et al. "Anatomo-functional correspondence in the superior temporal sulcus." *Brain Structure and Function*, volume 223, pp. 221–232, 2018.

[7] A. Le Troter, D. Riviere, and O. Coulon. "An interactive sulcal fundi editor in brainvisa." In "17th International Conference on Human Brain Mapping, Organization for Human Brain Mapping," , 2011.

[8] K. Li, L. Guo, G. Li, et al. "Gyral folding pattern analysis via surface profiling." *NeuroImage*, volume 52, pp. 1202–1214, 2010.

[9] Y. Liu, X. Yu, J. X. Huang, et al. "Combining integrated sampling with svm ensembles for learning from imbalanced datasets." *Information Processing & Management*, volume 47, pp. 617–631, 2011.

[10] A. Maratea, A. Petrosino, and M. Manzo. "Adjusted f-measure and kernel scaling for imbalanced data learning." *Information Sciences*, volume 257, pp. 331–341, 2014.

[11] F. Leroy, Q. Cai, S. L. Bogart, et al. "New human-specific brain landmark: the depth asymmetry of superior temporal sulcus." *Proceedings of the National Academy of Sciences*, volume 112, pp. 1208–1213, 2015.