Rare cortical folding patterns detection based on an unsupervised deep learning approach

Louise Guillon

université PARIS-SACLAY

Rare cortical folding patterns detection based on an unsupervised deep learning approach

Détection de motifs rares de plissement cortical par une approche de deep learning non supervisée

Thèse de doctorat de l'université Paris-Saclay

École doctorale n°575 Electrical, Optical, Bio-Physics and Engineering (EOBE)
Spécialité de doctorat: Sciences de l'information et de la communication
Graduate School : Sciences de l'ingénierie et des systèmes. Référent : Faculté des sciences d'Orsay

Thèse préparée dans l'unité de recherche **BAOBAB (Université Paris-Saclay, CEA, CNRS)**, sous la direction de **Jean-François MANGIN**, directeur de recherche, le co-encadrement de **Denis RIVIÈRE**, ingénieur-chercheur.

Thèse soutenue à Paris-Saclay, le 07 avril 2023, par

Louise GUILLON

Composition du jury

Membres du jury avec voix délibérative

Josselin HOUENOU Professeur des universités - praticien hospitalier, Univ. Paris Est Créteil	Président
Kiho IM Professeur Associé, Harvard Medical School	Rapporteur & Examinateur
Carole LARTIZIEN Directrice de recherche, CNRS, Inserm, INSA Lyon, Univ. Lyon	Rapporteur & Examinatrice
Guillaume AUZIAS Chargé de recherche, CNRS, INT Marseille	Examinateur
Isabelle BLOCH Professeure, CNRS, Sorbonne Université	Examinatrice
Ninon BURGOS Chargée de recherche, CNRS, Institut du Cerveau Paris	Examinatrice

THESE DE DOCTORAT

NNT : 2023UPAST058

université PARIS-SACLAY

ÉCOLE DOCTORALE
Physique et ingénierie : Electrons, Photons, Sciences du vivant (EOBE)

Titre: Détection de motifs rares de plissement cortical par une approche de deep learning non supervisée
Mots clés: plissement cortical, détection d'anomalies, apprentissage non supervisé

Résumé: Une des caractéristiques visuelles du cerveau est son aspect plissé. Il est en effet constitué de circonvolutions appelées gyri, qui sont délimitées par des sillons appelés sulci. À l'instar des empreintes digitales, les motifs du plissement cortical, c'est-à-dire l'arrangement, les caractéristiques et la forme des sillons sont propres à chaque individu, même s'ils s'inscrivent dans un schéma général propre à chaque espèce. Certains motifs ont été corrélés à des fonctions cognitives comme le contrôle inhibiteur dans la région cingulaire. De plus, des motifs rares ont été corrélés à des troubles neuro-développementaux comme le « Power Button Sign » qui est associé à un certain type d'épilepsie. Les avancées de l'apprentissage profond et en particulier non supervisé constituent une réelle opportunité pour analyser les motifs du plissement cortical, répondant au défi des grandes bases de données et de la variabilité inter-individuelle. Ce travail de thèse a pour but de développer une méthode permettant d'identifier des motifs rares ou anormaux.

À partir d'IRM cérébrales, des squelettes et des cartes de distances correspondant à un moule en négatif du cerveau sont générés. Ils permettent de se concentrer sur la morphologie du plissement. La variabilité inter-individuelle est modélisée par un beta-VAE entraîné uniquement sur des sujets témoins de la cohorte HCP.

À partir de cette représentation, des motifs caractéristiques de la population saine peuvent être identifiés. La première étude est faite dans la région cingulaire. Plus spécifiquement un clustering est appliqué sur l'espace latent. Des motifs représentatifs de chacun des clusters sont générés et permettent de caractériser les différents motifs caractéristiques. Des motifs décrits dans la littérature ont été retrouvés ce qui suggère que l'approche est pertinente.

Une fois des motifs caractéristiques identifiés, les motifs plus rares peuvent être analysés. Dans ce cas, le réseau est appliqué dans la région du sillon central à des cohortes comportant des sujets sains et des sujets avec des motifs rares ou à des données synthétiques présentant des anomalies. Ces dernières permettent de mieux qualifier les déviations identifiées. L'identification de motifs rares se fait dans l'espace latent ou sur la base de l'erreur de reconstruction. Pour évaluer la généralisation, cette approche est appliquée à deux autres régions comportant des motifs anormaux. Les résultats ont montré que l'espace latent et les erreurs de reconstruction apportent des informations complémentaires, adaptées à divers types de déviations. Le pouvoir génératif du beta-VAE permet de mieux comprendre les propriétés encodées dans l'espace latent. La visualisation des erreurs de reconstructions aide à localiser les caractéristiques atypiques identifiées. Enfin, l'application de la méthode aux deux autres cohortes indique une bonne généralisation, malgré un léger effet site.

Ces résultats prometteurs demandent à être répliqués avec davantage de sujets contrôles et de sujets présentant des motifs rares afin d'être confirmés. L'approche que nous proposons pourrait ainsi constituer une première étape vers l'identification de biomarqueurs et vers un modèle systématique à l'échelle du cerveau. Enfin, l'intégration d'autres modalités comme la cytoarchitectonie ou la connectivité enrichirait l'approche proposée.

Title: Rare cortical folding patterns detection based on an unsupervised deep learning approach
Keywords: cortical folding, anomaly detection, unsupervised deep learning

Abstract: One of the visual characteristics of the brain is its folded appearance. It is composed of convolutions called gyri, which are delimited by furrows called sulci. Like fingerprints, the patterns of cortical folding, i.e. the arrangement, characteristics and shape of the sulci, are unique to each individual, although they do fit into a general organization in each species. Some patterns have been correlated with cognitive functions such as inhibitory control in the cingulate region. In addition, rare patterns have been linked with neurodevelopmental disorders such as the "Power Button Sign" which is associated with a certain type of epilepsy. Advances in deep and especially unsupervised learning provide a real opportunity to analyze cortical folding patterns, addressing the challenge of dealing with large databases and high inter-individual variability. The aim of this thesis is to develop a method to identify rare or abnormal patterns.

From brain MRI, skeletons and distance maps corresponding to a negative cast of the brain are generated. They allow the deep network to focus on the folding morphology. Inter-individual variability is modeled by a beta-VAE trained only on control subjects from the HCP cohort.

From this normal representation, typical patterns of the healthy population can be identified. This first study is done in the cingulate region. More specifically, clustering is applied to the latent space. Representative patterns of each of the clusters are generated and allow us to characterize the different typical patterns. Patterns described in the literature were found, which suggests that the approach is relevant.

Once characteristic patterns have been identified, rare patterns can be analyzed. In this case, the network is applied in the central sulcus area to cohorts with healthy subjects and subjects with rare patterns or to synthetic data with anomalies. The latter allows for a better qualification of the identified deviations. The identification of rare patterns is done in the latent space and on the basis of the reconstruction error. To evaluate the generalization, this approach is applied to two other regions with abnormal patterns. The results showed that both the latent space and the reconstruction errors provided complementary information and were more suitable for certain types of deviations. The generative power of the beta-VAE allows a better understanding of the properties encoded in the latent space. The visualization of reconstruction errors helps to localize the identified atypical features. Finally, the application of the method to the two other cohorts indicates a good generalization, despite a slight site effect.

These promising results need to be replicated with more control subjects and subjects with rare patterns in order to be confirmed. The approach we propose could thus constitute a first step towards the identification of biomarkers and towards a systematic model at the brain scale. Finally, the proposed approach would benefit from the integration of other modalities such as cytoarchitecture or connectivity.

Pour Benjamin

Acknowledgments

If today these three years seem to have passed fast and smoothly, some moments were more difficult and I could not have realized this work without the help and support of many people whom I want to thank warmly. These acknowledgements are in both English and French, I apologize in advance for those who do not understand either language.

First of all, I would like to thank my two supervisors Jean-François Mangin and Denis Rivière, for giving me the opportunity to do this thesis within the team. I had contacted Jean-François for another subject that already had a PhD student but I had the chance to work on the subject of this thesis instead. Thank you for your trust and your advice.

I also thank all the members of my jury, for reading my manuscript, their comments and suggestions. We had a very rich discussion which has certainly broadened my vision of my work and the field. I look forward to furthering these conversations if we have the opportunity.

Une des richesses de la recherche repose dans les rencontres que l'on fait au sein du laboratoire et toutes les discussions qui en découlent. À cet égard, les années covid ont été particulièrement difficiles, mais ont néanmoins permis la création de dynamiques. Je pense notamment à la création du groupe de lecture de deep learning qui a été un projet motivant et qui m'a ensuite permis, au fil des séances hebdomadaires, de progresser sur de nombreux points. Zaccharie Ramzi, Benoît Dufumier et Alexandros Popov, merci d'avoir partagé cette initiative. Merci également à tous les membres qui permettent à ce groupe de perdurer, et notamment Corentin Ambroise, qui a repris le lourd fardeau de l'organisation.

Je remercie beaucoup toutes les personnes de NeuroSpin qui m'ont aidée et dont les interactions ont été fructueuses. L'équipe GAIA, les membres de l'équipe Architectures et notre task force "deep folding" pour leur accueil et leur aide.

Plus généralement, ces trois années de thèse ont été ponctuées de très bons moments partagés à Glasgow, autour d'une raclette, d'une bière, d'un café matinal ou d'un plat épicé à Singapour : Cyril Brzenczek, Loïc Dorval, Julie Victor, WenQi Shu-Quartier-Dit-Maire, Pierre Auriau, Corentin Ambroise, Clément Langlet, Robin Louiset, Benoît Dufumier, Chloé Gomez, Vincent Frouin, Ophélie Foubet, Emmanuelle Gourieux, Hugo Boniface; merci beaucoup à vous.

Une caractéristique du milieu de la recherche est que les personnes partent au fur et à mesure de la fin de leur contrat. À mon arrivée à NeuroSpin, j'ai été accueillie avec bienveillance et amitié par l'aquarium, Héloïse de Vareilles, Lisa Perus et Julie Victor. Un immense merci à vous pour votre accueil et tous ces moments

ensemble. Merci en particulier pour votre soutien dans les moments difficiles. Alexandros Popov, Ivy Uszynski, Maëlig Chauvel, Héloïse de Vareilles, notre petite cantine du 14ème a été un réel plaisir et soutien, j'espère qu'elle continuera encore un bon bout de temps ! Je remercie également tout particulièrement Benoît Dufumier, avec qui j'ai partagé à peu près les mêmes temporalités de thèse. Merci pour toutes tes réflexions très pertinentes qui ont grandement contribuer à améliorer mon travail et pour nos discussions plus personnelles qui sont une vraie richesse.

Depuis que j'ai changé de bureau dans NeuroSpin, j'ai eu la chance de mieux connaître Clément Langlet. Nos discussions tant professionnelles que sur des sujets plus larges m'ont beaucoup apporté et ont contribué à m'ouvrir l'esprit, merci pour tout et un immense merci en particulier d'avoir accepté d'être garant technique pour la soutenance, je n'aurais pas pu avoir mieux ! Je tiens également à remercier Ophélie Foubet, que j'ai eu la chance de rencontrer. J'ai beaucoup apprécié nos discussions et partager nos visions. Enfin, un très grand merci à Corentin Ambroise pour sa présence, nos échanges et la relecture de certaines parties de mon manuscrit.

Tout au long de ces trois années, en dehors du laboratoire, j'ai eu la chance d'être soutenue par ma famille et mes amis. Mes parents évidemment, sans qui je ne serais pas là aujourd'hui, leur soutien, leur amour et leur confiance m'ont été très précieux. Delphine a également été importante dans la construction de qui je suis devenue. Je pense aussi très fort à mes deux grands-pères et Mamette qui, même s'ils n'ont pas pu être à mes côtés lors de cette thèse, ont été des exemples très marquants pour moi d'ambition tout en humilité, de courage et de ténacité. J'ai eu la chance d'avoir le soutien de ma deuxième grand-mère, Mami, au cours de ce parcours, toutes nos discussions et en particulier celles sur la place de la femme m'ont beaucoup accompagnées. Isabelle, Adèle, Rose et Lucas, merci d'avoir été là pour moi dans les moments difficiles. Je remercie également chaleureusement la famille Wamo, qui a cru en moi dès le début, malgré les phases de doute. Je suis très honorée de rejoindre votre famille.

Un immense merci également à mes amis pour leur soutien. Chouchou, la vie ne serait pas pareille sans toi et chaque jour est l'occasion de me rappeler la chance que j'ai de te connaître. Margot, notre amitié et tout ce qu'on a partagé m'a beaucoup aidé pour ces trois années, ta bonne humeur et ton humour ont apporté une légèreté bienvenue dans les moments plus sombres. Donatien, après avoir partagé beaucoup de projets pendant notre enfance et notre adolescence, celui-ci était plus personnel mais ta présence depuis nos 3 ans, plus ou moins proche, me porte beaucoup.

Enfin, je n'aurai pas pu faire ce travail sans le soutien inconditionnel de Daniel. Je n'ai pas assez de mots pour exprimer l'étendue de ma gratitude, simplement, merci de croire en moi plus que moi-même et de me pousser à être quelqu'un de meilleur, partager ma vie avec toi est une chance exceptionnelle.

Contents

Introduction

Context

During gestation, the human brain folds resulting in circumvolutions composed of ridges of white matter, the gyri, and furrows, the sulci. The underlying mechanisms are still an area of research but two processes have been highlighted including mechanical and biocellular explanations (Llinares-Benadero and Borrell, 2019). Folding patterns have been intriguing for a long time and their study is not recent. Indeed, the description of the sulci dates back to the 19^{th} century (Cunningham, 1890). In the human population, stability of the folding patterns is observed with an overall similarity of location, shape and arrangements (Ono et al., 1990). This stability is important enough to enable to define a nomenclature of sulci and to develop methods that automate sulci recognition (Rivière et al., 2002; Perrot et al., 2011; Borne et al., 2020). Despite this homogeneity, each brain displays a unique cortical folding, acting as a fingerprint and the inter-individual variability of folding is very important (Wachinger et al., 2015).

The appearance of non-invasive methods like Magnetic Resonance Imaging (MRI) has provided new tools for visualization and studying the link with brain function for example.

On the one hand, some typical folding patterns - whose characteristics can be observed in most people - have been described. For example, the central sulcus is usually composed of one or several knobs. In return, other patterns have been identified as very rare, such as an interruption of the central sulcus which seems to concern less than 1% of the population.

On the other hand, some folding patterns have been linked with function. In particular, the main folds have been proposed as landmarks for the cytoarchitectonics zones (Amunts et al., 2007). Furthermore, some sulcal areas have been correlated to functions such as reading and hand motricity (Yousry et al., 1997; Sun et al., 2016). In return, certain neuro-developmental disorders including focal cortical dysplasia of type 2, a common cause of drug-resistant epilepsy or schizophrenia, were shown to be linked with specific cortical folding configurations (Provost et al., 2003; Mellerio et al., 2014).

The study of folding patterns is thus particularly interesting as it could lead to the discovery of new biomarkers. Another interesting characteristic of folding patterns is that after their formation, during pregnancy, they remain stable throughout life (Cachia et al., 2016). Hence, they can provide a unique insight into the neuro-developmental processes that occurred, acting as pieces of evidence of various events of brain development. Therefore, studying folding patterns can lead to several impacts. First, it can increase our current knowledge of the brain

and potentially enable to correlate functions to structures more easily. In a second step, identifying rare patterns may lead to a better understanding of perturbations that might have occurred during the development and hopefully, to foster the emergence of biomarkers of disorders.

Having large databases to identify typical and rare folding patterns, potentially leading to biomarker discovery is capital. Nevertheless, the high inter-individual variability hinders the identification of these patterns. Initially performed visually, the advances of machine learning and specifically, unsupervised deep learning constitute a real opportunity to address this challenge. Just like AlphaFold (Jumper et al., 2021) has revolutionized the study of protein folding based on deep learning, we hope that deep learning models may stimulate the deciphering of folding patterns. As a matter of fact, such frameworks could help learn a representation of the folding of the population in order to describe both the main typical configurations and the patterns deviating from the modelled norm that human perception cannot offer.

Contributions and thesis organization

This work proposes to use an unsupervised deep learning approach to model the folding inter-individual variability of a healthy population. From this learned representation, we first analyze folding patterns that seem characteristic of a control population before describing what happens at the margins of the modelled norm, looking for rare patterns.

This work has led to several contributions. First, we developed a publicly available preprocessing pipeline to work on specific sulcal regions and focus on the folding characteristics. Then, for the first time, we applied a deep learning framework, a $\beta - VAE$, to model the control folding variability. Based on this normal representation, we first identify typical folding patterns, before trying to detect rare folding patterns. To this end, we generated several datasets of synthetic anomalies. These fake anomalies enable to better describe the abnormal features that our framework can detect, ranging from subtle variations, potentially embedded in the inter-individual variability, to unlikely configurations. Our method has been tested on other regions and disorders. We hope in the near future that our framework can be systematically and easily applied to many other cohorts of patients in order to identify specific rare patterns.

This thesis is organized into four main parts. Part I proposes a state-of-the-art on cortical folding patterns and deep learning frameworks to identify outliers before formulating the questions driving this thesis. Then part II presents the general

material and methods that will be used to model the inter-individual variability of folding. The next two parts (part III and IV) present two applications: the identification of typical folding patterns and the identification of rare or abnormal patterns. Last, we conclude this thesis by drawing the main outputs and presenting some perspectives for future works.

Part I

Background

This first part aims to place this thesis subject in a broader context and how it relates to the state-of-the-art. Specifically, chapter 1 introduces the cortical folding patterns and their characteristics. Then, chapter 2 presents an overview of deep learning, the anomaly detection subfield, the associated challenges and some medical applications. This part ends with the problem formulation, the challenges and our proposed strategy.

Table of Contents

1 - Cortical Folding Patterns

1.1 . From the macroscopic to the microscopic scale: overview of the brain anatomy

After several centuries, the brain is still a fascinating object of study. The brain is composed of three parts: the cerebrum, the cerebellum and the brain stem. The cerebrum is composed of two hemispheres that are connected via the corpus callosum. The hemispheres are divided into four main lobes: frontal, parietal, temporal and occipital. The cerebrum consists in an outer layer, the grey matter, also called the cerebral cortex and an inner layer, the white matter. One striking visual characteristic of the human brain is its convoluted surface that resembles a nut. Indeed, the cerebral cortex is folded forming circumvolutions. It results in gyri, ridges of white matter, that are delimited by furrows, the sulci (Fig.1.1).

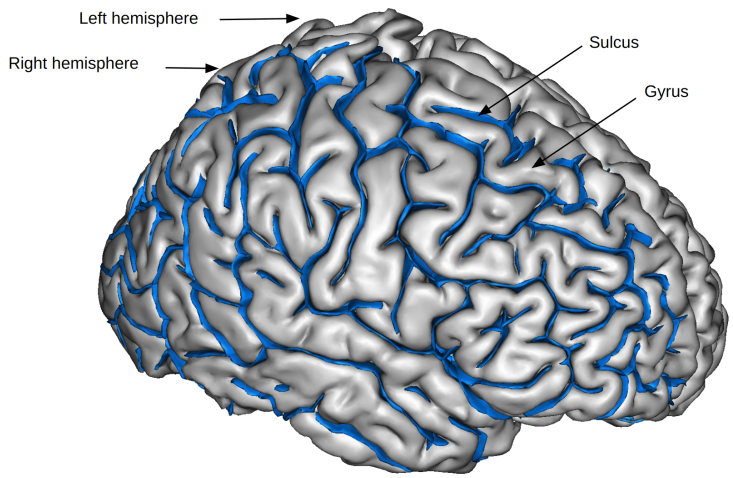


Figure 1.1: *Folded cortex with gyri and sulci.* The blue ribbons are abstract entities representing the sulci which are filled with cerebrospinal fluid.

The cortex is composed of several layers. The different layers are characterized by varying densities of the cells populations. Neuronal cells are located in the cortex. The neurons communicate between each other via the axons, up to the white matter. The white matter is composed of fibers that enable to connect various areas of the cortex.

1.2 . Description and formation of the cerebral sulci

Cortical folding is observed in several species and has been proved to be "an ancestral mammalian trait" (Lewitus et al., 2014). However, the degree of corti-

cal folding (represented by the gyrification index GI) varies across mammals, and some mammalian species, such as the mouse, are lissencephalic, i.e., their brain is not folded. Therefore, folding is not a structural necessity for brain function. In the human brain, studies have managed to describe three successive folding waves that result in primary, secondary and tertiary folds respectively (Chi et al., 1977). The primary folds are the most stable across the population and are located at stereotyped locations (Llinares-Benadero and Borrell, 2019). The existence of these stereotyped locations could be explained by the concept of sulcal roots. Sulcal roots are abstract entities representing the seeds where the first folds form (Régis et al., 2005). Although it is possible to identify some sulcal roots during the development in utero, it is not straightforward to visualize them all, therefore the term *sulcal pits* was introduced. Sulcal pits are the deepest local points along the bottom lines of the sulci (Lohmann et al., 2008; Im et al., 2010; Auzias et al., 2015; Im and Grant, 2019; Mangin et al., 2019). Sulcal roots and sulcal pits are very interesting objects to understand the folding dynamics in particular (Mangin et al., 2019).

The study of the mechanisms involved in the folding process is still an active field of research where historically, two main hypotheses on the folding determinants have been proposed.

The first current gathers several researchers that propose explanations based on mechanical factors. The first theory hypothesized that the cortex *had to fold* due to volume constraints: gyrations enabled to increase the brain surface (Clark, 1945). If this first explanation was quickly discarded, many new mechanistic explanations were formulated. One of the major theories was proposed in 1997 by Van Essen. It introduced the idea that, in the developing brain, the axons could exert a tension between the connected cortical areas that would draw the regions closer together, leading to the formation of the gyri (Van Essen, 1997). This model was however refuted with microdissections showing that if axons were effectively under tension, their orientation was not consistent and could not account for forces that would result in the folded cortex (Xu et al., 2010). Another important theory proposed that cortical folding was caused by differential growths of the cortex layers. An outer shell would have a faster tangential expansion than an inner layer (Richman et al., 1975). This theory was supported by morphogenetic models based on computer simulations that were able to reproduce characteristics of cortical folding (Toro and Burnod, 2005; Tallinen et al., 2014). It was later demonstrated by experiments based on a compound gel matrix which is summarized in Fig.1.2A. Based on a 3D MRI of a fetal brain serving as template, a gel brain was made. Another material, with different properties was added to model the grey matter layer. After immersion in a solvent, the composite gel started to swell at the surface which led to the formation of sulci and gyri. With specific mechanical properties of brain tissues, the outer shell formed very similar folds to those of the human

brain (Tallinen et al., 2016).

The second hypothesis relies on genetic factors. Diverse observations supported this hypothesis. First, several gyrencephalic species, i.e. that have a folded cortex, display similarities in terms of location and shape of the folds. In addition, within a species, the first folds to form are located in stereotyped areas. Such considerations tend to support hypotheses based on a biological and genetic explanation of the folding process (Llinares-Benadero and Borrell, 2019). Studies on the biological origin of the folds mainly pointed out cellular mechanisms such as the neurogenesis and the neuron migration. Specifically, the neurogenesis has been found to be more important in regions that will constitute gyri (Reillo et al., 2011; Lukaszewicz et al., 2006). This contrasted neurogenesis is particularly marked in the inner and outer subventricular zones (ISVZ and OSVZ respectively). The consequence is that there are different densities of immature neurons across the cortex before the folding, with more neurons in the regions of future gyri. The other main event involved in cortical folding is the neurons tangential dispersion during their radial migration. The neurons migrate radially along radial glial fibres (RGFs) but contrary to lissencephalic species whose RGFs are parallel, the RGFs of gyrencephalic species vary across regions and are diverging in future gyri (Fig.1.2B). These two cellular mechanisms are regulated by genetic processes in both space and time. Indeed, spatial patterns of gene expression define a protomap (Rakic, 1988) of the primary folds, leading to the stereotyped locations mentioned earlier. In addition, like the spatial protomap, temporal variations of gene expression in the ventricular zone define a critical period of cells formation that will be crucial for the folding processes (Martínez-Martínez et al., 2016; Borrell, 2018).

These cellular explanations enable to bring light on some parts that the mechanistic model did not explain, however they are not sufficient: two monozygote twins have a more similar folding pattern than two individuals (Lohmann et al., 2008) but they do not present the exact same folding, suggesting that forces other than genetics are involved. Ultimately, and although they have been historically presented as opposites, folding processes appear to be a combination of both mechanical and biological approaches as suggests a recent review (Llinares-Benadero and Borrell, 2019).

1.3 . Perturbations in the cortex development processes may lead to folding anomalies

The previous mechanisms explained the folding formation. However, in the course of mechanical and biological processes, certain disturbances may occur, leading to malformations of the cerebral cortical development and to abnormal or rare folding configurations in particular. The term *malformation of cortical development* was introduced to define a group of disorders in "children with developmental delay and young people with epilepsy" (Barkovich et al., 2012). These

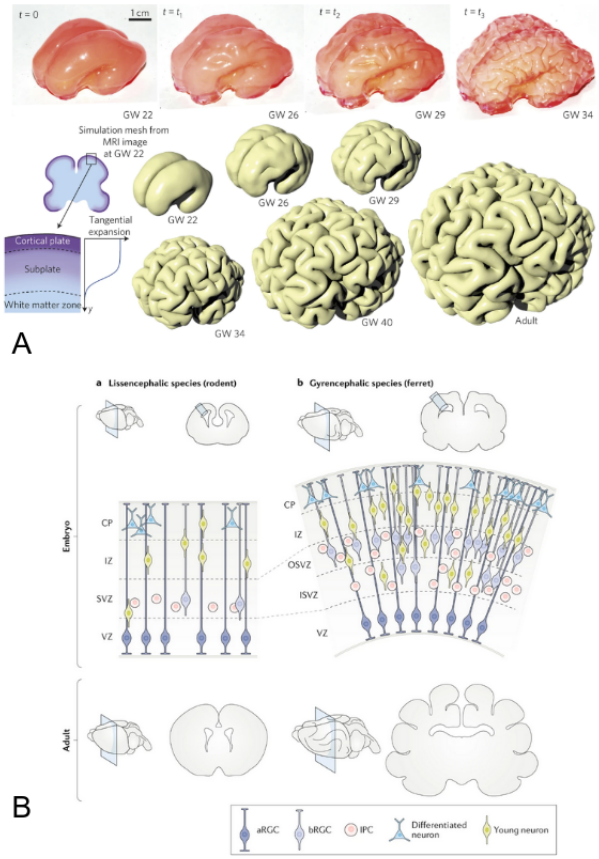


Figure 1.2: *Folding processes theories.* A. Simulation of the folding process with a compound matrix gel (upper row) and with computer modelling. Courtesy of (Tallinen et al., 2016). B. Cellular mechanisms involved in the gyrification. Courtesy of (Llinares-Benadero and Borrell, 2019)

disorders were then categorized according to the developmental steps that were disturbed (Barkovich et al., 1996). Initially proposed in 1996, this classification is constantly evolving due to the progressive increase in the understanding of biological processes and is therefore regularly updated. So far, three classes have been proposed: the malformations resulting from events occurring (1) during neuronal and glial proliferation, (2) during neuronal migration or (3) after the migration of the neurons. The perturbations occurring at these steps may affect the folding patterns in various ways.

The first category gathers malformations resulting from **abnormal neuronal proliferation**. In this case, the proliferation may be reduced, increased or abnormal. For instance, reduced proliferation can lead to microcephaly. In particular, the gene WDR62 seems to play a role in the proliferation and the migration of neuronal precursors and its mutation may result in microcephaly and simplified gyral patterns (Yu et al., 2010). On the contrary, an abnormal proliferation can express

by an increased proliferation which may result in megalencephaly. Specifically, megalencephaly can occur in association with polymicrogyria which presents abnormal folding configurations made of multiple small gyri and shallow sulci. In the case of abnormal proliferation, the development may result in overgrowth normal cells and in the production of abnormal cells in different organs. Specifically, focal cortical dysplasia (FCD) of type 2 is now considered as resulting from abnormal proliferation.

The second category includes malformations due to **abnormal neuronal migration**. It includes heterotopia which is characterized by the presence of a specific type of tissue in a non-physiological area. In such a case, the neurons do not correctly migrate and can remain between the ventricle and the pia. Heterotopia can lead to folding perturbations such as simplified gyral patterns (Sicca et al., 2003; Deleo et al., 2020). Gyral patterns may be also affected by abnormal transmantle migration and result in lissencephaly. On the other hand, when a large number of neurons are located regionally in the deep cerebral white matter, called subcortical heterotopia, the affected area is abnormally small and the cortex is thin and can be microgyric.

The final category describes malformations that may result from **events occurring after the migration**. This is the case of some polymicrogyria and schizencephaly (characterized by abnormal clefts in the hemispheres). It also concerns certain FCDs which could be caused by an injury to the cortex such as severe prematurity, asphyxia or bleeding. These examples also demonstrate folding abnormalities.

To summarize, malformations of cortical development can occur at various stages of the brain development and they seem to result from two main causes: genetic mutations and in utero infections. The neuronal outcomes are extremely variable ranging from no symptoms to severe ones. Exhaustive definitions and brain MR images illustrating each condition are presented in (Oegema et al., 2020; Severino et al., 2020). As seen previously, the folding pattern is a result of the brain development. In case of perturbations due to genetic mutations or exterior events, it is affected and therefore, its study enables to have insights on the neurodevelopment processes.

1.4 . Folding patterns as neuro-developmental markers

The folding process results in approximately 64 sulci per hemisphere. Historically, neuroanatomists have started to study the sulcal shapes and characteristics based on observations made on specimens. In the human population, due to the folding processes detailed previously, a stability of the folding patterns is observed with an overall similarity of location, shape and arrangements (Ono et al., 1990). This stability is important enough to enable to define a nomenclature and to de-

velop methods that automate sulci recognition (Rivière et al., 2002; Perrot et al., 2011; Borne et al., 2020). Despite this homogeneity, each brain displays a unique cortical folding, acting as a fingerprint (Wachinger et al., 2015). In particular, secondary and tertiary folds are highly variable.

Let us define some relevant terms for our study of cortical folding patterns. In our context we have the following definitions:

- ***Shape***: Elementary building block that can be combined to define patterns. For instance a knob or a flat segment (first row of Fig.1.3).
- ***Pattern***: Combination, arrangements of elementary shapes of one or several folds. For instance, several patterns exist in the central region, such as a single knob configuration, a double knob configuration or a rather flat one (see Fig.1.3). It is important to note that in our conception, folding patterns exist in the population as a continuous manifold. Therefore, there is a continuity from the single knob pattern, to the double knob pattern. Similarly, in the actual sulcus representing the flat configuration, we can still see a slight knob.

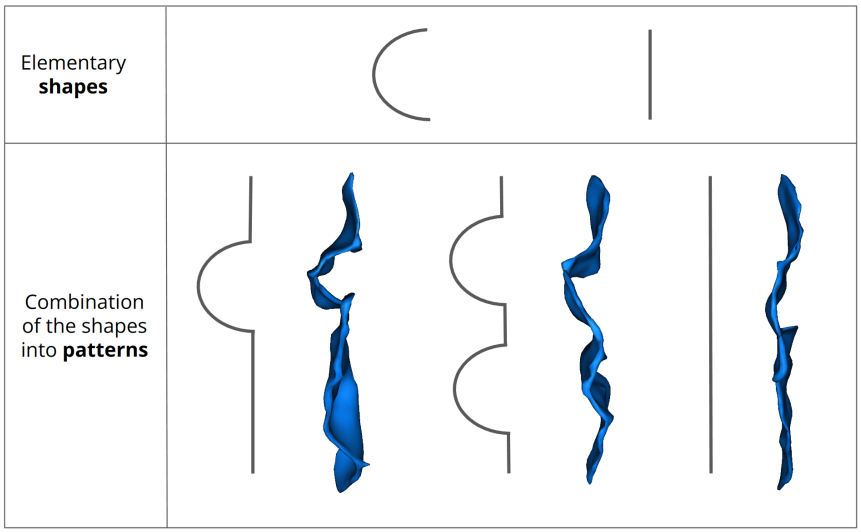


Figure 1.3: *Schematic representation of typical shapes and patterns of the central region.* First row: schematic elementary shapes presenting a knob and a flat segment. Second row: Combination of the schematic shapes into patterns. Each pattern is illustrated by a real example in blue. From left to right: single-knob pattern, double-knob pattern, flat pattern.

The folding variability is so complex that it has long been overlooked. However, thanks to advances in the neuroimaging field, studies have tried to characterize sulci with elementary shapes that amount to building blocks of alternative

patterns. These patterns can have various frequencies of appearance in the population. For instance, the central sulcus is typically composed of one or several knobs (Yousry et al., 1997) (see Fig.1.3). Similarly, the mid-fusiform sulcus usually presents an omega pattern (Weiner et al., 2014). In contrast, some very rare patterns have also been described, such as the interruption of the central sulcus that is present in about only 1% of the population (Mangin et al., 2019).

Folding patterns are especially interesting objects to study as they are related to function. The cerebral cortex is associated with many executive functions such as motor functions or language. At the beginning of the 20th century, Brodmann proposed a map of 52 regions based on the cortex cytoarchitecture. Modern techniques enabled to show that these regions were consistent with functions (Amunts et al., 2007). Certain areas adequately match anatomical landmarks such as gyri and sulci but it is not valid for all areas and using only these landmarks to delineate function areas may result in a lack of precision. Based on functional MRI (fMRI), it has been shown that the central sulcus course corresponds to the border of areas 3 and 4 defined by Broadmann, i.e. the motor area and the sensory area (White et al., 1997). Similarly, the visual cortex is located in the calcarine sulcus (Amunts et al., 2000). However, it is not the case for all areas and even more when it comes to secondary or tertiary folds that display a high inter-individual variability (Amunts et al., 2007). More precisely, based on fMRI, some sulcus parts have been correlated to specific functions. As already mentioned, the central sulcus divides the motor and the sensory areas and distinct parts of the central sulcus have been correlated with the tongue, foot and hand among others (Penfield and Boldrey, 1937; Mangin et al., 2019; Germann et al., 2020). Specifically, the central sulcus main knob has been linked to the hand motricity and is called the "hand knob" (Yousry et al., 1997). In the cingulate region, the cingulate sulcus patterns have been associated to the inhibitory control (Borst et al., 2014). In return, specific patterns were also correlated to neuro-developmental disorders. The Power Button Sign (PBS), a rare configuration of the central sulcus (see Fig.1.4), may be associated with focal cortical dysplasia of type 2 which may cause drug-resistant epilepsy (Mellerio et al., 2014). Patterns in the superior temporal sulcus (STS), central, intraparietal and frontal regions could be related to autism (Levitt et al., 2003; Auzias et al., 2014; Hotier et al., 2017).

A major characteristic of folding patterns is that they constitute "trait features" opposite to "state features" (Cachia et al., 2016). Unlike state features that can evolve during lifespan, trait features remain fixed. For example, sulcal opening is a state feature because it increases with the ageing process (Kochunov et al., 2005; Jin et al., 2018). In return, the pattern of the cingulate sulcus area is a trait feature because it is stable throughout the lifetime after infancy (see Fig.1.5) (Cachia et al., 2016). This difference between trait and state features has also

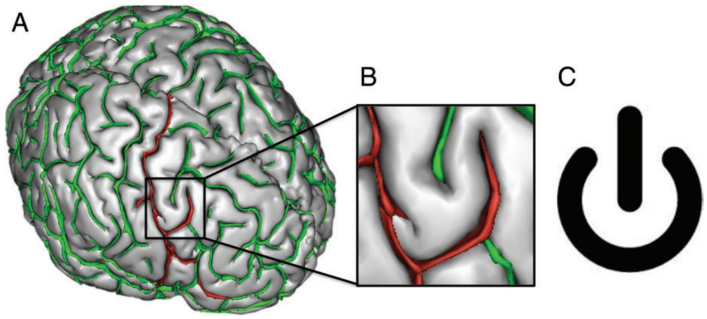


Figure 1.4: *Power Button Sign.* A rare folding pattern associated with FCD2, a common cause of drug-resistant epilepsy. Courtesy of (Mellerio et al., 2014).

been demonstrated in the study of the effects of handedness on the central sulcus shape. For example, forced dextrals show similarities to sinistrals in pattern, but changes in elongation and opening occur when they are constrained to use the right hand for writing (Sun et al., 2012).

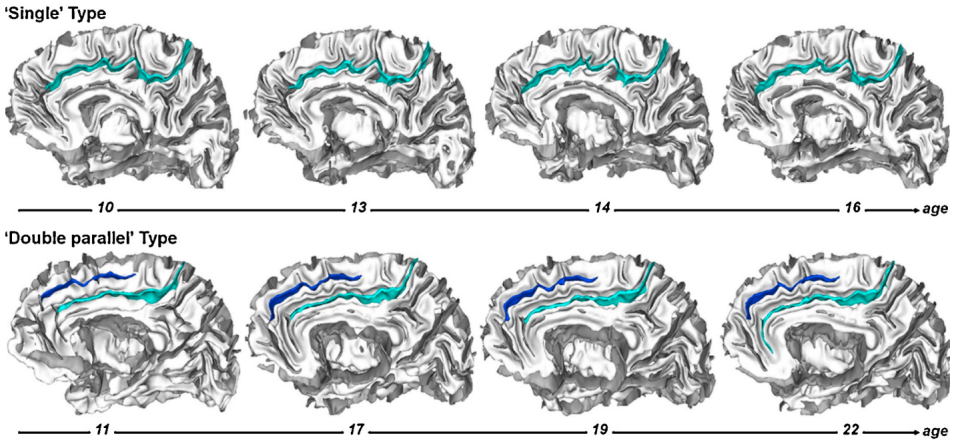


Figure 1.5: *Longitudinal study of folding patterns.* Two patterns have been described in the cingulate region: one long cingulate sulcus (upper row) and one long cingulate sulcus with a smaller parallel sulcus – the paracingulate sulcus– (bottom row). The folding patterns remain stable. Courtesy of (Cachia et al., 2016)

As we have seen in the section 1.3, perturbations can occur during the folding processes and lead to rare or abnormal folding configurations. Therefore, folding patterns consitute a marker of potential disturbances. Since the patterns are trait features, their analysis can provide insight into events that may have happened during the development. Hence, deciphering sulcal complexity and having a better understanding of the underlying shape variability is of great interest as folding patterns could become biomarkers of neuro-developmental disorders.

1.5 . Methods to study cortical folding patterns

Cortical folding can be analysed with diverse approaches depending on the object of the study, whether it is about the folding and sulcal characteristics, or about folding patterns as we defined them in section 1.4.

Many works have studied the global properties of folding and sulcal characteristics. In this case, an approach is to extract from the MR images morphometric features such as the depth, surface curvature or opening of each sulcus. Processing softwares like BrainVISA/Morphologist or FreeSurfer enable to obtain these features. They can then be used to perform statistical analysis in various applications. For instance, they have been used to study the relationship between folding and psychiatric disorders. The sulcation index, depth and length of the sulci have been used to study folding abnormalities in a cohort of patients with Autism Spectrum Disorder (ASD) (Auzias et al., 2014). More recently, the sulcal depth of some secondary and tertiary sulci was linked to some characteristics of ASD patients (Benitez et al., 2022). ASD is not the only psychiatric disorder studied, works have also dealt with schizophrenia for instance (Penttilä et al., 2008). Another feature that can be used to study the sulci is their positioning. It was indeed correlated to ASD (Levitt et al., 2003) and studied for schizophrenic patients (Plaze et al., 2011).

However, these methods cannot completely tackle the study of folding patterns as we define them. To do so, other strategies have been proposed. The most obvious way is to work on folding patterns based on visual descriptions. This can be done directly on the MR images but has the drawback of being in 2d slices. It enabled to identify typical patterns in the mid-fusiform sulcus for example (Weiner et al., 2014). It can also be performed based on the 3D reconstructions of the folds. For instance, the PBS was identified based on such reconstructions (Fig.1.4) (Mellerio et al., 2014).

With the advance of machine learning techniques, other methods have tried to automate the identification of sulcal patterns characteristics. The first step is to choose the way of representing the folds. Different strategies have been developed in this end. A first approach consists in directly working on the folds. This can be performed based on the folds extracted by BrainVISA/Morphologist. In this case, the folds are materialized by a set of voxels. With this input, several methods have been proposed. Each sulcus can be represented as 3D moment invariants, a set of shape descriptors, patterns are then identified thanks to a clustering (Sun et al., 2007). Rather than defining characteristics on each individual sulcus, the pairwise dissimilarity of the sulci after their co-registration can be used to build a dissimilarity matrix that represents the folding variability. From this matrix, it is possible to apply dimension reduction algorithms to capture the main dimensions which represent the principal shape features. This approach was used to study the shape of the central sulcus in the developing brain (de Vareilles et al., 2022), to identify typical folding patterns (Sun et al., 2009) and to characterize the effect of handedness in this region (Sun et al., 2012). This method also enables to compare

folding patterns across species among the primates (Foubet et al., 2022). Another approach is based on sulcal pits. Sulcal pits can be extracted from various MRI preprocessing softwares. Sulcal patterns are then described as a graph of sulcal pits from which similarity matrices are computed. Such approaches were applied to compare the heritability of folding patterns (Im et al., 2011) or identify typical folding patterns (Meng et al., 2018).

Finally, aggregated representations, like averages, may be easier to characterize the folding patterns and the shapes, rather than analyzing and comparing all the subjects. But the selected representation constitutes intrinsically a choice of paradigm. As a matter of fact, on the one hand, we can consider that the set of existing patterns is finite, leading to a certain number of defined patterns (note that this number is still unknown). In such cases, clustering is particularly well suited (Sun et al., 2007, 2009; Meng et al., 2018; Duan et al., 2019). On the other hand, patterns can also be represented in a continuous way. In this case, manifold-based analyses can be used (Sun et al., 2012; de Vareilles et al., 2022; Foubet et al., 2022). To identify typical patterns among the wide inter-individual variability, representative patterns may be used. The representative patterns can be drawn from the subjects' population, analysing the closest subjects to the centroids for instance (Meng et al., 2018), or sulci averages can be used (Sun et al., 2012, 2017; de Vareilles et al., 2022; Foubet et al., 2022).

This first chapter has described cortical folding and the current hypotheses concerning the folding processes. Specifically, we pointed out that several events occurring during neuro-development can affect these processes, potentially resulting in rare or even abnormal folding patterns. Since folding patterns are a trait feature, they remain stable throughout life and constitute *markers* of brain development that may lead to the discovery of biomarkers of neurodevelopmental disorders. However, the inter-individual variability is tremendous making the identification of deviating patterns highly complex. Thanks to its representation power, deep learning can prove to be a very effective way to model inter-individual variability in order to spot rare folding configurations. In particular, since we seek to identify rare or abnormal folding patterns that we do not know about, or only know about in very limited numbers, we address this situation as an anomaly detection problem.

2 - Deep Learning Frameworks for Identifying Outliers

2.1 . Introduction to deep learning

In a few years, deep learning has become one of the major techniques used to model and predict in various everyday life scenarios. The domains of application are very diverse since the inputs, as well as the tasks offer many possibilities. For instance, deep learning can be used to identify objects in images, translate texts from one language to another or forecast events based on time series. Deep learning is principally based on representation-learning methods (LeCun et al., 2015). The latter are capital since the success of machine learning algorithms mainly depends on the learned representations. Unlike some machine learning techniques that require hand-crafted features, the aim of representation learning is to "automatically discover the representation needed for detection or classification" (LeCun et al., 2015). This can be achieved by a combination of modules that learn increasing levels of abstraction. In the case of images, having these successive layers enables to detect edges, which can be then combined to make shapes and objects for instance.

Deep learning, and more globally machine learning can be split into different forms: supervised, unsupervised and reinforcement learning. The most common form is supervised but more and more works now concentrate on the opportunities offered by unsupervised learning.

In a supervised setting, based on input data $X = \{x_1, x_2, ..., x_n\}$ and the associated labels $Y = \{y_1, y_2, ..., y_n\}$ we wish to learn a function f that predicts $\hat{Y} = \{\hat{y_1}, \hat{y_2}, ..., \hat{y_n}\}$. During the learning process, an objective function measures the error between y and $\hat{y}$. The model's parameters Θ are consequently updated to reduce the error. Let us take an example of a classification setting where the aim is to classify x-ray images between controls and patients. The images are fed into a model which outputs a vector $\hat{y}$, that contains the scores for the two categories. The difference between $\hat{y}$ and the true scores y drives the learning process as the model is optimized to have the smallest difference between the predicted and the true scores. Such a framework has been widely used in many domains, including medical applications.

In reinforcement learning, an agent learns from its environment. Given a set of possible actions, the agent chooses one to do and obtains the corresponding reward. The long-term goal of the agent is to maximize its global reward.

Unsupervised learning is inspired by the observations of biological learning: from a lot of information that is not fully labeled, we are able to deduce knowledge and predictions (Barlow, 1989). Thus, unsupervised learning may be used to

extract the underlying causes of a phenomenon. Indeed, it can be seen as a way of modelling the probability density of data (Hinton and Sejnowski, 1999). The concept of unsupervised learning applied to images began with the purposes of dimensionality reduction and compression (Hinton and Sejnowski, 1999). Unlike supervised or reinforcement learning, in unsupervised learning settings, there are no labels or rewards. The labels y are unknown and only the inputs are given to the model. Unsupervised learning can be commonly used to discover patterns in a dataset. The main examples of this type of learning are clustering and dimensionality reduction (Ghahramani, 2004).

In the last decade, deep learning has been widely used. Based on the historical model of the multilayer perceptron (Hornik et al., 1989), many new frameworks have been proposed. The escalation was particularly impressive for images. As a matter of fact, with the introduction of convolutional neural networks (CNN), image analysis, also known as computer vision has become a major field. CNNs are particularly well suited for images. Their creation was inspired by the natural visual perception mechanisms (Fukushima and Miyake, 1982; LeCun et al., 1989; Lecun et al., 1998). Very briefly, based on a succession of convolutional and pooling layers, CNN models are able to detect and extract patterns of increasing complexity. Therefore, they constitute very effective architectures to study images for examples. In the nineties, the first CNN was introduced and applied to handwritten character recognition (Lecun et al., 1998). A major advantage of the CNNs is their ability to generate accurate representations of an image with very little or even no preprocessing step (Gu et al., 2018). Along with the improvement of computational efficiency, more complex architectures have been introduced such as transformers (Jaderberg et al., 2015) or diffusion models (Dhariwal and Nichol, 2021).

2.2 . Overview of anomaly detection

Apart from medical applications which will be discussed in a later section, anomaly detection is widely studied in the deep learning community. The range of domains covered extends from computer vision to time series analysis for example. When analyzing the field, it seems that several words are used interchangeably. Indeed, *outlier* may also be used, which raises the question of their difference. In their review, Chalapathy and Chawla present anomalies and outliers as synonyms: "Anomalies are also referred to as abnormalities, deviants, or outliers in the data mining and statistics literature" (Chalapathy and Chawla, 2019). Hawkins proposed a more precise definition in his monograph on applied probability and statistics: "an outlier is an observation which deviates so much from the other observations as to arouse suspicions that it was generated by a different mechanism" (Hawkins, 1980). Another definition introduces nuances between the two

terms: "'outlier' refers to a data point that could either be considered an abnormality or noise, whereas an 'anomaly' refers to a special kind of outlier that is of interest to an analyst" (Aggarwal, 2017). Fig.2.1 shows the corresponding illustration. According to this vision, data would form a spectrum that ranges from typical normal data to anomalies passing through some noise. At a given threshold of noise, the points are considered as outliers. Consequently, both noise and anomalies would then constitute outliers. In other works, the term "outlier" is used to describe a small part of "normal" data that lies far away (Markou and Singh, 2006). The meaning that will be employed depends on the field of application. For instance, the case of a production line where all samples are standardized will be different from a medical application where each point corresponds to a patient. In the latter, the term "anomaly" could rather be linked to the association with a disorder. Therefore, the use of these two terms could be linked to the notion of inter-individual variability: in a population with very few variations (like the outputs of a production line) the use of "anomaly" may be different from the one in a population that displays a high inter-individual variability.

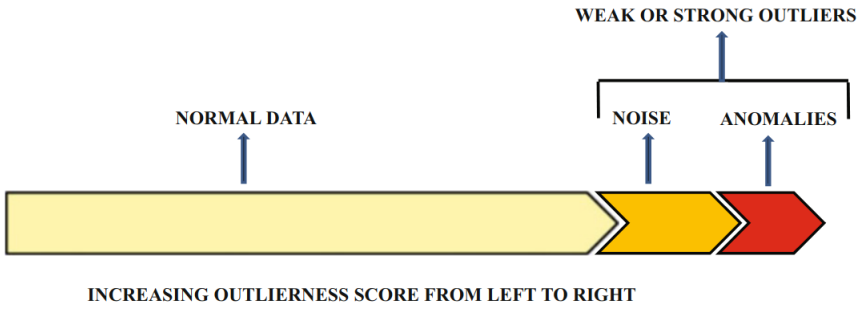


Figure 2.1: *From normal data to outliers.* Schema representing the spectrum of data, from normal to outliers. Courtesy of (Aggarwal, 2017)

Another associated term is "novelty". In a review on the subject, novelty has been defined as "the task of classifying test data that differ in some respect from the data that are available during training" (Pimentel et al., 2014). However, in this review, the use of the term "novelty" is linked to the absence of a category of data during the training process. More specifically, the authors emphasize the abundance of "normal" data which can be used to model a distribution. On the contrary, in these works there is a lack of data describing "abnormalities". Therefore, in this case, novelty detection seems to be a sub-part of anomaly detection in which anomalies are not modelled during training. All in all, currently there is no consensus on a precise definition. Nevertheless, although the three terms might be used interchangeably, the methods to tackle these problems are usually common.

Anomalies can have several natures. Three different types of anomaly have been proposed to encompass most of the applications: *point anomalies*, i.e. sam-

ples that are abnormal with respect to the rest of data; *contextual anomalies*, i.e. samples that are abnormal in a given context; *collective anomalies*, i.e. samples that are abnormal due to their collective occurrence (Chandola et al., 2009).

Anomaly detection remains a challenging task for several reasons. The extensive review of Pang and colleagues presents an overview of these challenges. First, we may not know what are the characteristics of the anomalies. It is therefore harder to detect them. In addition, the anomaly classes may be heterogeneous and by definition, anomalies are scarce which may lead to class imbalance. The interpretability of anomaly detection algorithms is also capital. Lastly, there can be an effect of sampling: is a detected anomaly really an anomaly, or is it due to a sampling effect (Dunning and Friedman, 2014) as illustrated by Fig.2.2?

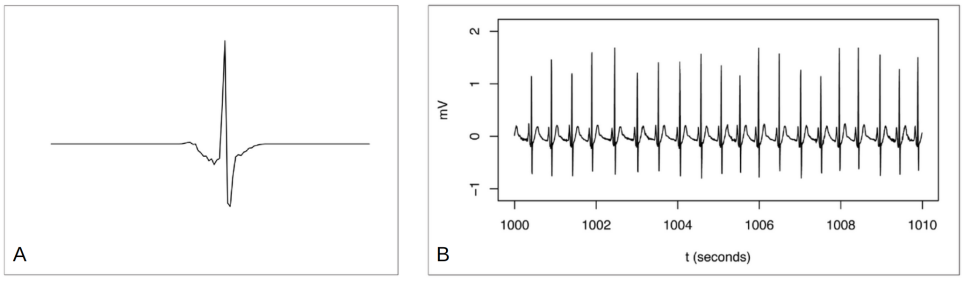


Figure 2.2: *Sampling effect on anomaly detection.* The sampling can have a major impact on the definition of outlier. A. According to this sampling, the peak seems to constitute an anomaly. B. With a larger sample, what seemed to be an anomaly is in fact normal. Courtesy of (Dunning and Friedman, 2014)

These challenges have led to different methods to perform anomaly detection. As a matter of fact, based on the availability or the potential scarcity of abnormal examples, several learning strategies exist. First, if the anomalies are known and are in sufficient number, supervised anomaly detection can be applied. Binary and multi-class classification or segmentation can be used in supervised frameworks. Nevertheless, in most cases we do not know what are the characteristics of the anomalies (Dunning and Friedman, 2014) or we do not have enough abnormal samples. In this case, unsupervised approaches are performed (Fernando et al., 2022). Unsupervised anomaly detection techniques will be described in the next section. Whether the learning is supervised or unsupervised, the output of the algorithm can be either an "outlier score", i.e., a score that expresses the degree of deviation of a point, or a binary label indicating whether the point is an anomaly or not. It is common to obtain the binary labels based on the previous outlier score on which a threshold is applied (Aggarwal, 2017).

2.3 . Unsupervised anomaly detection

Unsupervised anomaly detection can be addressed by various approaches. Prior to advances in deep learning, many "classical", i.e. shallow, machine learning algorithms have been proposed. Different reviews present the diversity of the methods that have been developed to tackle anomaly detection problems (Chandola et al., 2009; Chalapathy and Chawla, 2019; Wang et al., 2019; Smiti, 2020). Very briefly, they can be organized into several categories which include statistical methods, distance-based methods and density-based methods. Statistical approaches consist of detecting whether a point deviates from a distribution (Smiti, 2020). This can be done with boxplots or histograms for instance (Goldstein and Dengel, 2012). The second category is based on distances. Indeed, it is quite intuitive that a sample that is very far from its neighbours can be an outlier. Although distances are difficult to deal with in high dimensions, several methods have been introduced, such as Angle-Based Outlier Detection (ABOD) (Kriegel et al., 2008). ABOD tackles the curse of dimensionality by comparing the angles between difference vectors of the samples. The last category of methods corresponds to density-based methods. One of the main techniques is the *Local Outlier Factor* (LOF) algorithm (Breunig et al., 2000). Instead of a binary label, LOF assigns a degree of "outlierness". The estimation is based on the comparison of the local density of a sample with the local densities of its neighbours. One-class Support Vector Machine (OC-SVM) (Schölkopf et al., 2001) can also be considered as a density-based method. It extends the concept of SVM to the identification of outliers. It is possible to combine the above methods to form ensemble methods, such as the isolation forest (Liu et al., 2008), another widely used algorithm.

Nevertheless, the expansion of deep learning has extended to the anomaly detection field and many *deep anomaly detection methods* have been introduced. In a recent review, deep learning methods for anomaly detection were classified into three paradigms as presented in Fig.2.3: the application of deep learning to extract features that will be inputs to other classical outlier detection methods; the learning of a feature representation of normality and finally, end-to-end anomaly detection frameworks (Pang et al., 2020).

The first framework relies on feature extraction. Specifically, it gathers methods that achieve a dimensionality reduction to obtain relevant representations that are then fed into shallow anomaly detection algorithms.

The second approach aims to detect anomalies by obtaining an accurate representation of normality. Indeed, once normal characteristics have been modelled, it is possible to evaluate whether samples share these characteristics or not. Such frameworks can be driven by several objectives like data reconstruction, generative modelling or self-supervised learning based on an auxiliary task.

One very common model in this framework is based on autoencoders (AE).

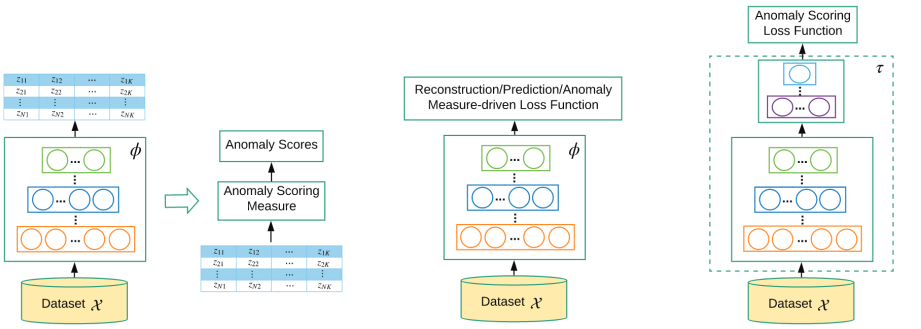


Figure 2.3: *Three frameworks in deep anomaly detection.* Courtesy of (Pang et al., 2020)

These models are composed of two parts: an encoder θ and a decoder ϕ. The encoder learns to project an input x onto a latent space, also called bottleneck, that corresponds to a feature representation comprising far fewer dimensions than the input space. Then, the decoder reconstructs the encoded vectors z. These two parts are optimized to reconstruct the inputs as well as possible. Since the bottleneck contains fewer dimensions and the objective is to minimize the reconstruction error, the model learns to encode the most relevant features of the input. In other words, the model encodes the characteristics that are common and shared by most of the samples. Therefore, in an anomaly detection framework, the training is usually performed on a control population only. The assumption is that only the *normal* features will be encoded. The classic formulation is presented in equations 2.1 and 2.2.

$$z = \theta(x, \Theta_\theta), \hat{x} = \phi(z, \Theta_\phi) \tag{2.1}$$

$$(\Theta_\theta^*, \Theta_\phi^*) = \underset{\Theta_\theta, \Theta_\phi}{\operatorname{argmin}}\, x - \phi(\theta(x, \Theta_\theta), \Theta_\phi) \tag{2.2}$$

This formulation corresponds to a simple autoencoder but many other versions have been proposed since, like denoising AE (Vincent et al., 2008) or variational AE (Kingma and Welling, 2014). Owing to their easy-to-understand approach and to their easy implementation on various kinds of data, AE have been widely used for anomaly detection. The actual detection is usually performed based on the reconstruction error which was initially proposed by (Hawkins et al., 2002). Depending on the type of input data, the layers of the AE can be adapted to form convolutional AE for images or long-term-short-memory (LSTM) AE for time series for instance.

Generative Adversarial Networks (GAN) (Goodfellow et al., 2020) may also be used to model the normality. In the classic framework, the objective is to generate

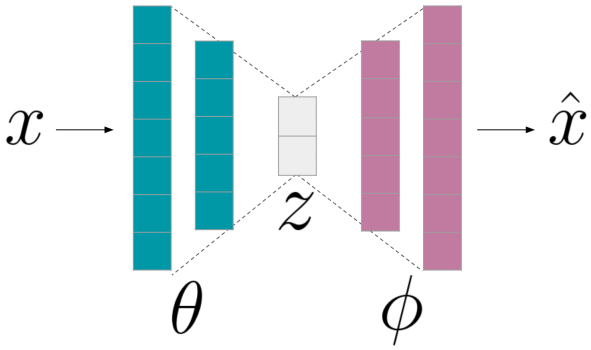


Figure 2.4: *Autoencoder framework.*

realistic new data that have the same properties than the training set, based on two components: a generator and a discriminator. The generator tries to create the most realistic data from the latent space, in order to fool the discriminator that has to determine whether a sample comes from the true data or from the generated data. This model is also adapted in terms of anomaly detection since the generator is usually trained only on control data which results in a latent space that encodes typical features. The discriminator is then used as the anomaly detector. Based on the development of GANs, other approaches have tried to combine the adversarial approach and one-class classification in an end-to-end framework (Sabokrou et al., 2018).

Recently, a new paradigm has emerged, namely self-supervised learning. The assumption of self-supervised learning is to learn relevant feature representations for a downstream task based on pretext tasks. The possible pretext tasks are very diverse. The first proposed was classification of geometric transformations applied to the data (Golan and El-Yaniv, 2018). Many works have also used inpainting as a pretext task (Haselmann et al., 2018; Kang et al., 2021; Li et al., 2021a; Pirnay and Chai, 2021; Xiang et al., 2021; Zavrtanik et al., 2021). In the anomaly detection field, the downstream task is the identification of the anomalies.

However, a drawback of the previous models is that they have not been optimized for anomaly detection tasks. Autoencoders are trained to minimize the reconstruction error and self-supervised frameworks are optimized to perform best on the pretext task. Unlike these models, other interesting approaches have tried to adapt the learning objective to be better suited with the detection of anomalies. Deep support vector data description (SVDD) has been proposed to overcome these limitations and is inspired by one-class classification (Ruff et al., 2018). This model learns to map the data to a space where data points are included in the smallest hypersphere. The projection is optimized to minimize the radius of the hypersphere. Therefore, normal samples fall within the hypersphere whereas anomalous points are projected outside.

To recap, most of the unsupervised frameworks that were presented are based on a training set composed of *normal* data only. Nevertheless, depending on the task and on each specific case, it is possible that the training set used for modeling the normal distribution contains outliers. The presence of these outliers in the modeling of what is *normal* can have consequences on the performances. As a matter of fact, the effects of a contaminated dataset were studied on a GAN (Berg et al., 2019) and on a 3D convolutional AE (Behrendt et al., 2022). In the first setting, they compare the drop in performance between a clean dataset and a dataset containing 2% of anomalies. They deepen this analysis in the second article, gradually increasing the injection from a clean set to 3%, 6% and 12%. The consequences on the performances are major in both cases and starts at a contamination of 3% already.

In addition, in many real life scenarios of anomaly detection it may be expected to have outliers remaining in the train set. Therefore new approaches were introduced to limit the consequences of a contaminated dataset. A possible way is to identify the outliers during training and to remove them (Behrendt et al., 2022). On the contrary, if the outliers are identified, they can also be used to maximise their detection (Berg et al., 2019; Yu et al., 2021; Qiu et al., 2022). Such methods address the limitations deriving from a training objective that is not optimized for anomaly detection and constitute the last framework of Fig.2.3, namely end-to-end anomaly detection. Another framework proposes to combine the principle of GANs and of one-class classifications, gradually injecting outliers to make them more separable (Sabokrou et al., 2018).

Depending on the framework, the actual detection of the anomalies is performed based on the objective function. For example, with autoencoders the anomaly score is usually derived from the reconstruction error (Pang et al., 2020). In GANs, the anomaly score can be a combination of the reconstruction error and of difference between the features (Schlegl et al., 2017; Zenati et al., 2019). GANomaly uses a L1 distance between the encodings of the input and the encodings of the reconstructed inputs (Akcay et al., 2019). In self-supervised settings, the classification scores on the pretext tasks can be used as normality scores from which the anomalies can be deduced (Golan and El-Yaniv, 2018). In the case of SVDD, a threshold is applied on the distance to the center of the hypersphere.

2.4 . Application in medical imaging

In medical imaging, anomaly detection has many applications. First, from a practical point of view, it can help the diagnosis. For instance, anomaly detection can be applied to brain tumor segmentation (Menze et al., 2015), to the identification of lesions (Baur et al., 2019; Simarro Viana et al., 2021) or to disease diagnosis

(Schlegl et al., 2019; Pinaya et al., 2019; Tian et al., 2021; Muñoz-Ramírez et al., 2022; Sato et al., 2022), It can also enable the discovery of new biomarkers (Schlegl et al., 2017). This brief overview of applications highlights the diversity of the types of input that can be used in anomaly detection frameworks which includes X-ray radiography, magnetic resonance imaging or computed tomography scans for example. In addition, these methods can be applied to various organs. Therefore, deep anomaly detection methods constitute a major opportunity for the domain.

Most of the models mentioned previously can be applied to biomedical images. One of the most widely used frameworks is based on autoencoders. Specifically, a comparison of various architectures demonstrated the good performances achieved by these models (Baur et al., 2020). However, medical imaging displays particular characteristics that raise specific challenges. One of the main challenges is the notion of variability. Unlike factory pieces that are stereotyped, each individual has its own characteristics. Even more, individuals can be gathered in wider groups such as age or gender for instance. These properties are particularly challenging as they should not be modelled by an anomaly detection framework. Indeed, they can introduce some noise that is usually not relevant regarding the features of the studied medical condition. In addition, in case of an imbalanced population (for instance much more females than males), they may introduce biases in the learned representations and thus in the predictions. Criteria that are external to the subjects can also impact the data: the machine used to obtain the images can consequently affect the images and is referred as a scanner effect. Another challenge which is present in global anomaly detection but emphasized in medical applications is the dataset size. As a matter of fact, by definition anomalies are rare events but when dealing with pathologies the size of the datasets is even smaller. That is why, unlike anomaly detection in other areas, supervised learning is more used and unsupervised learning is particularly challenging (Fernando et al., 2022). Finally, the characteristics of medical images may differ from classical computer vision: they may be 3D images which leads to a very high number of dimensions for instance.

3 - Problem Formulation

3.1 . Motivation and objective

In the previous chapters we have presented the main characteristics of cortical folding patterns and covered their formation which can be prone to malformations and lead to rare or abnormal folding patterns. As we have seen, since folding patterns are a trait feature that appears during the neuro-development it can be used, at a later time of life, as a testimony, a marker of the underlying processes occurring during brain development and of the potential malformations. In addition, specific cortical folding patterns have been associated with functions as well as disorders. Their study is therefore particularly relevant in understanding brain development and could help find biomarkers for certain neuro-developmental disorders.

The creation of large neuroimaging databases constitutes a first step in the deciphering of the folding patterns. Having large databases enables to get a bigger picture and thus a more accurate analysis of the patterns. Nevertheless, we mentioned the very high inter-individual variability of folding patterns and it is even higher with a larger number of subjects. This makes it difficult to identify what constitutes *typical* folding patterns or on the contrary *rare or even abnormal* patterns. Advances in machine learning and especially deep learning are now opening up new possibilities for studying folding patterns, identifying typical or rare patterns, and hopefully, emerging sulcal biomarkers. With the opportunities offered by representation learning that automatically learns good representations (LeCun et al., 2015) and by the autoencoder framework that enables to encode the most typical features in a latent space, the study of folding patterns may take a turn.

In this thesis, we aim to identify rare cortical folding patterns. Specifically, we adopt an unsupervised deep learning approach to address this task. Since only a few number of rare patterns have already been described, on small cohorts and usually by visual inspection (Mellerio et al., 2014), we leverage the ability of unsupervised frameworks to discover new patterns in the data. Furthermore, we seek to identify *novel* rare folding configurations. Therefore, a supervised learning would not be adequate and unsupervised learning is much better suited.

3.2 . Challenges and proposed strategy

Most of the unsupervised anomaly detection works propose a framework that relies on the modelling of the *normality* (Pang et al., 2020). However, the task difficulty depends on the input, and particularly on the variability observed. In our case, the folding complexity is considerable. Our first challenge is to be able to

effectively model the wide range of typical patterns. Several questions derive from this challenge: how to represent the folding patterns? Which method to use to model the variability?

Once the *normal* variability is encoded, another challenge is to identify the patterns that deviate significantly from the norm and which constitute rare patterns. This raises the question of what level of deviations may be identified. In addition, since few rare patterns have been described, evaluation is challenging and must be addressed.

Finally, in this work, we are interested in local patterns rather than global, hemisphere-wide arrangements of the folds. However, as we have seen in part 1.3, various regions of the cortex may be subject to malformations, therefore, it is interesting to study the whole brain. Nevertheless, combining a local approach with a whole-brain analysis constitutes another challenge.

To address these questions, we propose an unsupervised model that learns the normal inter-individual variability on a control population. Once a representation of the normality is obtained, two applications can be performed:

1. Based on the modelled normal inter-individual variability, we can focus on this representation and **identify typical patterns of the control population**.

2. On the contrary, we can focus on what happens at the margins of the representation of the normal variability to **identify rare patterns**; we can also project patient populations into this representation and analyze their differences, hoping to **discover specific rare patterns**.

These two applications are illustrated on Fig.3.1

3.3 . Contributions

Part II of this work presents the general data and methods used. Both applications rely on a first step which is the modelling of the inter-individual variability. Therefore, they share similarities in the data used and in the general methods applied. In particular, the pre-processing is based on the same representation of folding patterns. An unsupervised deep learning approach based on a $\beta - VAE$ is used in both applications.

Part III concerns the first application: the identification of typical folding patterns. We introduce data and methods that are specific to this application. Specifically, we studied the cingulate region which is of clinical interest. Then, although the deep learning approach is shared with the second application, the methods used to analyze the latent space are different.

Then, we move away from the typical folding characteristics to focus on the rare configurations. Part IV presents our work on the identification of rare folding

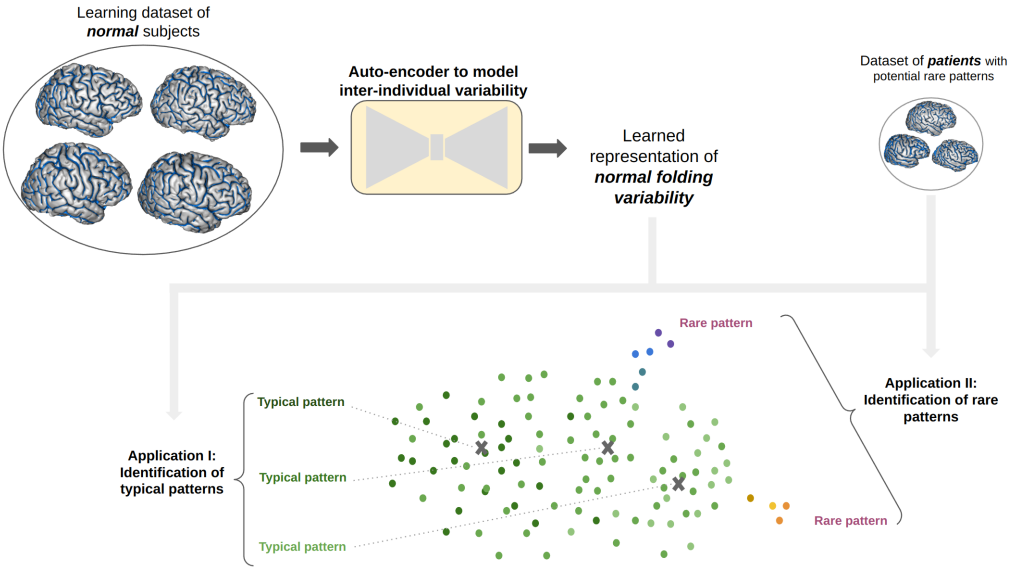


Figure 3.1: *Two possible applications based on the approach.* An autoencoder-based model is trained on a dataset comprising only control subjects to represent the normal folding variability. Once this variability is encoded (colored points), it can be used to identify typical patterns (application I), or to identify rare patterns (application II).

patterns. Several datasets were used. In order to be able to reliably assess our methods, we generate datasets containing synthetic rare configurations. We also evaluate our approach on actual rare patterns. We apply this approach to several areas of the brain.

Finally, part V brings elements for discussing the methods and the results obtained.

Part II

General material and methods to represent the inter-individual variability

In the previous part we have seen that cortical folding patterns are unique to each individual although they follow a global species-specific organization and that unsupervised deep learning could be well adapted to study them. We also presented our approach which consists in modelling the inter-individual variability, and the two possible applications that derive from the learned representations: (1) the identification of typical patterns in a control population, (2) the identification of rare folding patterns.

This second part presents the databases and the general methods that are shared by these two applications. The specific datasets and methods of each application will be presented in parts III and IV. In this section, we first briefly describe the common databases. Then we introduce the representation of cortical folding that we used. Finally, we present the deep learning approaches that were used to model the folding variability.

Table of Contents

4 - Databases

Modelling the inter-individual variability requires databases that have some characteristics. This part introduces only the databases and methods common to the two applications, i.e., that are used to model the normal variability. Therefore, the subjects must be healthy controls and the two databases that will be described in this section include only **control subjects**.

4.1 . The Human Connectome Project dataset

The Human Connectome Project (HCP) dataset (Van Essen et al., 2013) is an open access database. Several releases have made available neuroimaging and behavioral data of 1200 typically developing subjects.

When it comes to modelling the inter-individual variability, HCP has some particularities. First, it comprises only subjects aged between 22 and 36 years old. Compared to many datasets, this means that the subjects are rather young. This has implications in terms of folding patterns. First, as explained in section 1.4, as the subjects are young adults, the folding processes are over and the patterns are already fixed. Moreover, the ageing should not have started to affect the brain and specifically the opening of the sulci. A particularity of this dataset is that it includes twins and their non-twin siblings. This has to be acknowledged since it has been demonstrated that the folding patterns were partly heritable (Im et al., 2011; Ahtam et al., 2021). Another specificity is the homogeneity of the data acquisition method: all MR images were acquired on a single Siemens Skyra Connectom scanner at an isotropic resolution of 0.7mm. This means that compared to most databases, the HCP dataset is homogeneous and very clean, with no site effect for instance.

4.2 . In-house dataset

In addition to the HCP dataset, for both applications we used an in-house dataset that was described in (Borne et al., 2020). It is composed of 62 healthy subjects aggregated from heterogeneous databases. The subjects are mostly men between 25 and 35 years old. The resolution slightly differs between subjects, with the majority of the subjects' MRI having an approximate resolution of 0.93 x 0.93 x 1.20 mm.

5 - Cortical Folding Representation

5.1 . BrainVisa/Morphologist pipeline

For the two databases that have been presented in the previous section, our initial inputs are the T1-weighted MR images. Nevertheless, structural MR images hold numerous pieces of information beyond the morphology of cortical folding. Therefore, we first use the preprocessing software BrainVISA/Morphologist (https://brainvisa.info/) on the raw images. BrainVISA provides tools for various tasks including: Morphologist which performs brain segmentation and sulcal analysis or Anatomist which enables to visualize MR images with advanced functionalities

Morphologist has been developed by the team and is subject to continuous updates (Mangin et al., 1995; Rivière et al., 2002; Mangin et al., 2004a). It is composed of several steps that include skull stripping, bias correction, segmentation of the brain and of the hemispheres, skeletonization of the grey matter and the cerebrospinal fluid (CSF) union (Rivière et al., 2002). The main steps are represented on Fig.5.1A. This leads to so-called skeletons, 3D images representing only the folding and which correspond to a negative cast of the brain.

The skeleton is then segmented into simple surfaces (SS) depending on various parameters such as topological properties and important local sulcal depth variations (Fig.5.1B) (Malandain et al., 1993; Mangin et al., 1995). For example, in Fig.5.1B, small branches (SS2 and SS4) are represented as different simple surfaces from the main ones, SS1 and SS3. The depth variation resulting from the buried gyrus leads to two distinct simple surfaces (SS1 and SS3). Therefore, in this case, the central sulcus is composed of four simple surfaces. The simple surfaces do not comprise junctions. The junctions can be between two SS or between one SS and the hull for instance. Each SS includes voxels which correspond to the folds bottom lines. The skeleton is therefore composed of voxels of different natures: they can be only part of simple surfaces, bottom lines or diverse junctions. These various natures constituting the folds are modelled by specific voxel's values leading to quasi-binary skeletons, composed of several topological values.

The simple surfaces are then aggregated in a graph representation, called folding graph. Each simple surface corresponds to a node of the graph. The nodes bear information such as the manually or automatically attributed label (for instance "S.C._right" for the right central sulcus), the list of the corresponding voxels coordinates in the native space or the size of the simple surface. The bottom line voxels are another attribute of the nodes. The nodes are connected by edges which hold other information like the list of the voxels corresponding to the junction between the two nodes. In addition, the folding graph contains global attributes such

as transformations to referentials. Thanks to BrainVISA/Anatomist, the folding graph can be visualized in 2D or 3D like in Fig.5.1A.5. and 7.

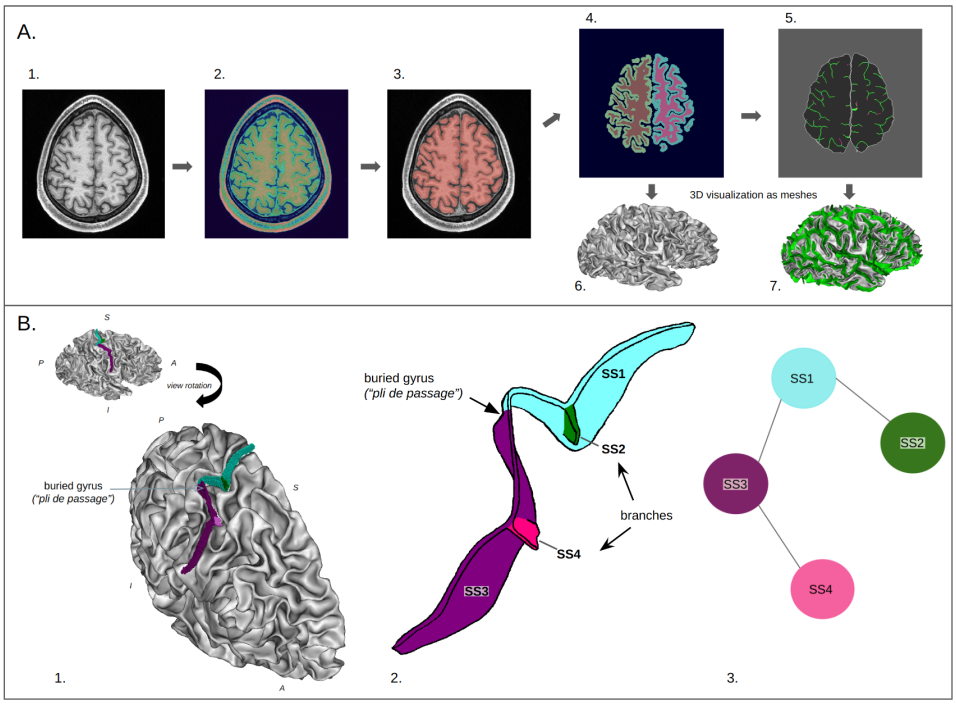


Figure 5.1: *Overview of the BrainVISA/Morphologist pipeline's main steps and of the folds representation.* *A. Main steps of BrainVISA/Morphologist pipeline.* 1. Raw T1-w MRI, 2. Bias-corrected image, 3. Segmentation of the brain, 4. Segmentation of the hemispheres and of the grey and white matter, 5. Skeleton representation of the folding graph, representing a negative cast of A.4. 6. Mesh representation of the white matter of the right hemisphere, 7. Folding graph that represents the folds (in green) as the negative cast of the white matter of the right hemisphere (white mesh). *B. Folds representation.* 1. Example of a central sulcus, which is composed of several elementary entities called simple surfaces (SS). (Orientation: A: Anterior, P: Posterior, S: Superior, I: Inferior). 2. Corresponding schematic representation of the sulcus represented in 1, which is formed by four simple surfaces. Depth variation caused by the buried gyrus and the presence of two branches lead to the division into four different simple surfaces. 3. Corresponding folding graph.

5.2 . Folding Graph Preprocessing

Based on the folding graphs produced by Morphologist, I developed, first on my own and then in collaboration with the team, a preprocessing pipeline to preprocess the folding graphs to make them easier to use in deep learning frameworks. This pipeline is publicly available (https://github.com/neurospin/deep_folding). The main steps and functionalities are presented in Fig.5.2.

5.2.1 . Skeletons, foldlabels and distance maps

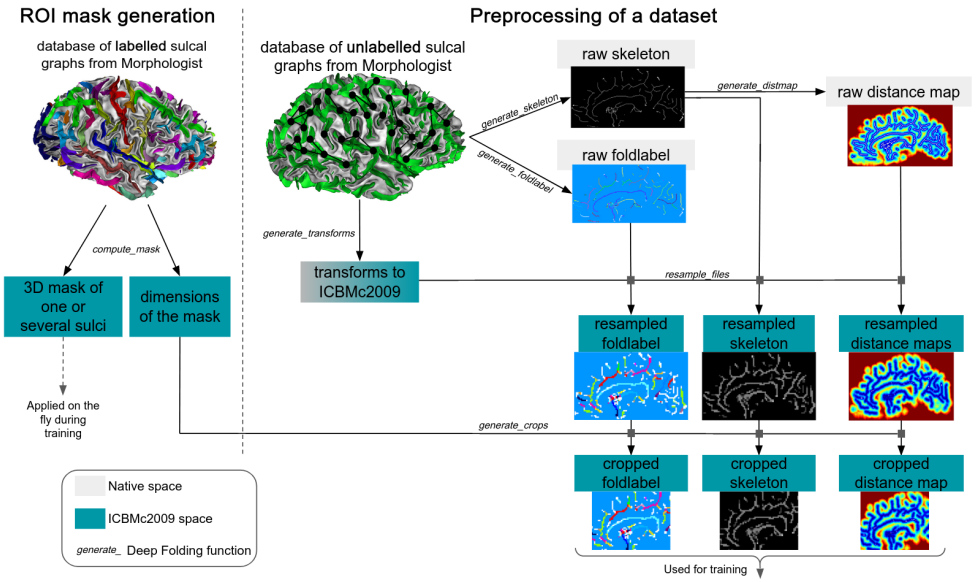


Figure 5.2: *Deep Folding Preprocessing pipeline.* The dashed line divides two functionalities that can be used independently or sequentially. Left: based on a dataset with **labelled sulci**, a 3D mask can be learned. Right: based on a dataset with **unlabelled sulci**, the folding graphs are used to generate raw skeletons and foldlabel images. The skeletons can then be used to generate distance maps. Skeletons, foldlabels and distance maps are resampled to the ICBMc2009 referential. If a region of interest is studied, the dimensions of the adequate mask are used to crop the hemisphere.

Based on the folding graphs of Morphologist, the pipeline enables to generate two types of images: skeletons and "foldlabels" (Fig.5.2).

- **Skeletons**. The skeletons generated by the deep folding pipeline are different from the skeletons of Morphologist. As a matter of fact, since Morphologist's skeletons are the result of the segmentation of the white matter and of the union of grey matter and CSF, the brain hull is included. On the contrary, using the folding graphs to generate the skeletons only translates the SS, the bottom lines and the various junctions. Specifically, for each node and each edge of the graph we write the corresponding voxels with the according topological value. The result is a 3D image of the folds with five different topological values: 30 for the folds bottom lines, 60 for the simple surfaces, 110 for the junctions, 120 for the buried gyri (pli de passage) and 100 for the few remaining voxels. Images of skeletons are presented in Fig.5.2.

- **Foldlabels.** In some cases it can be more relevant to still have the division into elementary SS. The foldlabel images enable to keep the segmentation in SS. Indeed, each node of the graph is sequentially written in the image with a specific number between 1000 and 2000, associated to each SS and which serves to define it. Bottom lines and junctions are written with a

number that corresponds to their associated SS. To ensure to be able to link a bottom line with its associated SS, we add 5000 to the associated SS value. For instance, if a SS is 1551, then its bottom line will be 6551. The result corresponds to "raw foldlabel" in Fig.5.2.

- **Distance maps.** Skeleton-based images have proved to be relevant and were the object of previous studies for fold recognition for instance (Borne et al., 2020, 2021). However, they display some particularities that may make them more difficult to handle. First, skeletons are quasi-binary images representing the folding patterns; hence, they are very sparse images and only very few voxels hold explicit sulcal information in skeletons. On average, sulci represent less than 5% of the voxels. Consequently, it could be interesting to have a representation where the information devoted to folding patterns is more distributed. In addition, skeleton images are not smooth, the voxels that coincide with the sulci, correspond to abrupt changes in intensity. This particularity makes the details of the skeleton harder to represent and reconstruct. When using AE-based models, it is also complex to reconstruct folding patterns in skeletons as they are not continuous images, so there is no notion of the proximity of a voxel to a sulcus. This makes the reconstruction error and the gradient-based learning less efficient. Therefore, we also generate distance maps. In particular, we convert the skeletons into distance maps based on the Chamfer distance, which approximates the Euclidean distance. Sulci are considered objects, and the further away a sulcus is, the larger the value of a voxel is. Unlike skeletons or foldlabel images, distance maps are "continuous" images with smooth variations. Another advantage is that distance maps are built based on the whole hemisphere: the value of a voxel depends on its neighbours. This may be particularly interesting if we work on a region of interest (ROI). Indeed, they give, especially near the border of the ROI, information about objects outside the ROI, which is not the case for ROI made of skeletons. An example of distance map is shown in Fig.5.2.

These three inputs are generated in the native space with the same resolution as the initial T1-W MR images. Then they are affinely registered to the ICBMc2009 space and resampled to an isotropic resolution of either 1 or 2 mm in our applications. Since each individual is in its native space, using such a transformation allows us to compare them. The affine transformations are obtained based on the SPM12 normalization performed during Morphologist processing. It is important to note that our registration is strictly affine. It involves only rotation, translation and scaling towards a template referential. In particular, many methods have proposed nonlinear warping strategies that may erase or distort local folding characteristics. On the contrary, with our registration, the folding patterns remain similar.

5.2.2 . Focusing on a single region: crop definition

As we mentioned in section I, the perturbations that can lead to rare configurations can have an impact on cortical folding at different scales. The study of folding patterns can therefore concern various scales ranging from the whole brain to a local scale, i.e. focused on one or two sulci. In this thesis, we are interested in capturing the ***local*** folding patterns variability, such as the hand knob. Indeed, it is the morphology of the folds that we wish to study. Such characteristics can also be considered at a wider scale; however, as a first step, it seemed easier to focus on smaller regions. Thus, for our two applications, we work on sub-regions, i.e. ROIs, of the brain. It is then necessary to define these ROIs. To study the specific morphology of certain sulci, we chose to define our ROI based on the sulci nomenclature in order to encompass the whole sulcus of interest in our ROI, leading to a ROI centred around one or more sulci.

To define each ROI, we learn a mask of the region over the previously described in-house dataset as the sulci are manually labeled (Borne et al., 2020). The mask is generated directly based on the folding graphs. In the graphs, the sulcus of interest is represented as a set of simple surfaces. We initialize the mask as an empty volume in the ICBMc2009 space. Then, for each subject, the voxels corresponding to the sulcus of interest are registered to the ICBMc2009 space, resampled to the targeted resolution, and increment the mask. The resulting mask is slightly dilated by 5 mm to include potential sulci locations not represented in this manually labelled database. In the end, the mask is a 3D image where the voxels correspond to the frequency of appearance of the sulcus of interest.

Once the mask is defined, we apply it to the dataset we wish to analyze, for instance, the HCP dataset. The mask can be applied to any kind of input: skeletons, foldlabels or distance maps. First, we crop the inputs of the HCP subjects according to the mask bounding box in the ICBMc2009 space (see Fig.5.2). The mask is applied later, on the fly during the training of our network. In the case of a ROI comprising several sulci, we proceed sequentially: we define a mask for each sulcus and we combine them before cropping the subjects.

It is important to note that the sulci labels are only used for obtaining the mask coordinates but are not needed afterwards, once the masks are defined.

5.2.3 . Folds visualization

With our three types of input data, data visualization and shape characterization can be performed directly on 2D slices. This results in visualizations like the ones presented in Fig.5.2. However, having a 3D view of the folds is capital when it comes to analyzing the folding patterns.

It is straightforward to visualize binary images on 3D: we can convert them to meshes with various tools, like one developed by the team: https://github.com/neurospin/dico_toolbox. Therefore, for skeleton and foldlabel data, it is almost direct, we just binarize them. Unlike the skeleton and the foldlabel images,

the distance maps are continuous. Therefore, we binarize our distance maps with an empirically defined threshold of 0.4 and convert them to meshes

5.3 . Requirements on sulci labelling

The labelling of the sulci is not an easy task. Whether it is manual or automatic, several challenges exist. In our effort to model the folding variability and to identify typical and rare patterns, we want to be independent of the sulci labels. Indeed, as we mentioned, the labelling is complex and can be prone to error. Some very efficient tools have been developed and some of them are embedded in BrainVISA/Morphologist pipeline (Borne et al., 2020). However, even if their performance is high, they may still attribute wrong labels, especially in regions which are highly variable and could be of interest. Therefore, it is capital that our approach may be applied to any dataset independently of whether it is labelled or not.

If we recap the previously introduced steps: Morphologist and our folding graph preprocessing do not require labels to generate folding graphs and then skeletons, foldlabels or distance maps. The only step that requires the labels is in the mask definition. Hence, if we use our in-house dataset presented in section 4.2 to define the different masks and thus ROIs, we will not need to have the labels for other datasets.

6 - Deep Learning Approach

Based on the skeletons and the distance maps, we now wish to model the inter-individual folding variability. To do so, we use an approach that relies on deep learning and specifically, on a $\beta - VAE$. Therefore, this section will present the basics that support our deep learning approach.

6.1 . Convolutional Neural Networks

We quickly mentioned convolutional neural networks (CNNs or ConvNets) in part I, we will now go deeper in how they work. CNNs are architectures particularly well-suited to deal with images since they enable to successfully capture spatial dependencies and thus to encode specific image properties. In this part, we will consider a CNN applied to 3D images even though the approach is similar with 2D images. A CNN is usually composed of a succession of three main layers: convolutional, pooling and fully-connected layer. Convolutional layer parameters consist of learnable filters. The filters are smaller than the input. During training, each filter slides across the input. For each location of the filter, a dot product is computed between the filter parameters and the corresponding voxels. Depending on the position of the layer in the network, the learned filters activate for different image patterns. In particular, the first layers of the network focus on the big picture while the deeper layers look at the details: the lower frequencies are learned before the high frequencies. Convolutional layers enable parameter sharing and are especially interesting to deal with images and high-dimensional data as it would be difficult to connect each neuron to all other neurons like in fully connected layers. In convolutional layers, each neuron is connected to only a local region of the input volume. This can be regulated by a hyperparameter called the receiptive field. The size of the output of a convolution layer depends on three other hyperparameters. First, the depth of each output volume: it corresponds to the number of filters associated with a layer. Each of them can focus on something different and thus learn complementary information. The stride enables to indicate the step used for the filters sliding, for instance through each voxel (stride=1) or skipping every other voxel (stride=2). Last, padding adds voxels of a given value (usually 0) around the image.

Between successive convolution layers, pooling layers are usually inserted. They are used to gradually reduce the spatial size of the representation. The most common applied function is MAX. However, it is possible to avoid using pooling layers with a larger stride for example (Springenberg et al., 2015).

Many architectures based on convolutional layers have been introduced. The first successful one was in 1989 (LeCun et al., 1989), later formalized as LeNet (Lecun et al., 1998). Since then, the architectures have been complexified and

have led to major advances in the field. The most famous architectures are AlexNet (Krizhevsky et al., 2012) or more recently ResNet (He et al., 2016) and densenet (Huang et al., 2017).

Now that we introduced the CNNs, let us go back to some of their characteristics that are important. Indeed, each model bears biases: choosing an architecture comes with applying a prior to the space. In particular, several biases are associated with CNN. They introduce an equivariance to translation resulting from the weight sharing and there is a locality bias: pixels that are close, are connected to each other.

In the task of modelling the inter-individual variability of folding patterns, convolutional layers and convolutional networks are well adapted for several reasons. First, thanks to their architecture, they enable to be computationally efficient which is capital as we deal with 3D images. In addition, their ability to explicitly handle the geometry is really interesting when dealing with images. Last, they are known to have a good generalization power.

Convolutional layers can be used in various models, including the autoencoder family. In this case, specific operators can be used. For example, the reconstruction of the latent projection requires increasing the dimensions. To this end, upsampling techniques or transposed convolutions can be used. Upsampling techniques include several methods such as nearest neighbours or linear interpolations. In return, transposed convolutions upsample an input thanks to learnable parameters.

6.2 . Autoencoder framework

6.2.1 . General framework

We briefly introduced the general context of autoencoders in part I. This type of model is widely used for images. One of the most famous autoencoders, the UNet (Ronneberger et al., 2015) which was later adapted to 3d images (Milletari et al., 2016), was designed for segmentation purposes. Such models and their more complex counterparts are also spread in the outlier detection domain. As a reminder, a basic autoencoder is a neural network composed of two parts: the encoder and the decoder. The encoder maps an input x to a vector z in the latent space. The latent vector is then reconstructed by the decoder which outputs x'. The optimization criterion is $||x' - x||$. Depending on the type of the input, several metrics can be used. If dealing with images, L1, L2 or the cross-entropy are usually employed.

The latent space Z, also called information bottleneck, is supposed to encode the main properties of the input space. Based on this assumption, it could be interesting to sample vectors in the latent space and to pass them through the decoder to obtain new samples. However, due to the absence of constraints in the latent space, this usually leads to absurd samples.

6.2.2 . Overview of the VAE

To address this issue, generative models can be used, such as the variational autoencoder (VAE) (Kingma and Welling, 2014). VAE are stochastic networks and in some way, are closer to bayesian models than to AE. Unlike classic AE, they map the input as a distribution p_θ which is parameterized by θ. Let us introduce some notations:

- $\mathbf{x}$ is the input,
- $p_\theta(\mathbf{z})$ is called the prior distribution,
- $p_\theta(\mathbf{x}|\mathbf{z})$ is the likelihood,
- $p_\theta(\mathbf{z}|\mathbf{x})$ is the posterior distribution.

A VAE involves two variables: $\mathbf{x}$ which is observed and $\mathbf{z}$ which is not. $\mathbf{z}$ is sampled from the prior distribution $p_\theta(\mathbf{z})$ while $\mathbf{x}$ can be generated from the conditional distribution $p_\theta(\mathbf{x}|\mathbf{z})$.

However, $p_\theta(\mathbf{z}|\mathbf{x})$ is intractable so an approximation is used: $q_\phi(\mathbf{z}|\mathbf{x})$. These distributions can correspond to an AE framework: $q_\phi(\mathbf{z}|\mathbf{x})$ acts as a probabilistic encoder and $p_\theta(\mathbf{x}|\mathbf{z})$ as a generative model similar to a probabilistic decoder. The approximation $q_\phi(\mathbf{z}|\mathbf{x})$ should be close to the real distribution $p_\theta(\mathbf{z}|\mathbf{x})$. The Kullback-Leibler Divergence D_{KL} (Kullback and Leibler, 1951) is used to quantify the difference between the two distributions. We seek to minimize $D_{KL}(q_\phi(\mathbf{z}|\mathbf{x})||p_\theta(\mathbf{z}|\mathbf{x}))$. If we expand the equation we have:

$$D_{KL}(q_\phi(\mathbf{z}|\mathbf{x})||p_\theta(\mathbf{z}|\mathbf{x})) = \log\, p_\theta(\mathbf{x}) + D_{KL}(q_\phi(\mathbf{z}|\mathbf{x})||p_\theta(\mathbf{z})) - \mathbb{E}_{\mathbf{z}\sim q_\phi(\mathbf{z}|\mathbf{x})} \log\, p_\theta(\mathbf{x}|\mathbf{z}) \tag{6.1}$$

which is equivalent to:

$$\log\, p_\theta(\mathbf{x}) - D_{KL}(q_\phi(\mathbf{z}|\mathbf{x})||p_\theta(\mathbf{z}|\mathbf{x})) = \mathbb{E}_{\mathbf{z}\sim q_\phi(\mathbf{z}|\mathbf{x})} \log\, p_\theta(\mathbf{x}|\mathbf{z}) - D_{KL}(q_\phi(\mathbf{z}|\mathbf{x})||p_\theta(\mathbf{z})) \tag{6.2}$$

The left term of equation 6.2, corresponds to the objective. Indeed, we seek to maximize the log-likelihood of generating real data and that the approximation q_ϕ is the closest to $p_\theta(\mathbf{z}|\mathbf{x})$. This leads to:

$$L_{VAE}(\theta, \phi) = -\log\, p_\theta(\mathbf{x}) + D_{KL}(q_\phi(\mathbf{z}|\mathbf{x})||p_\theta(\mathbf{z}|\mathbf{x})) \tag{6.3}$$

$$L_{VAE}(\theta, \phi) = -\mathbb{E}_{\mathbf{z}\sim q_\phi(\mathbf{z}|\mathbf{x})} \log\, p_\theta(\mathbf{x}|\mathbf{z}) + D_{KL}(q_\phi(\mathbf{z}|\mathbf{x})||p_\theta(\mathbf{z})) \tag{6.4}$$

$$\theta^*, \phi^* = \underset{\theta,\phi}{\operatorname{argmin}}\, L_{VAE} \tag{6.5}$$

In Variational Bayesian methods, this is the variational lower bound, also called evidence lower bound (ELBO) because the KL divergence is never negative. Therefore we have:

$$-L_{VAE} = \log\, p_\theta(\mathbf{x}) - D_{KL}(q_\phi(\mathbf{z}|\mathbf{x})||p_\theta(\mathbf{z}|\mathbf{x})) \leq \log\, p_\theta(\mathbf{x}) \tag{6.6}$$

During the training process, samples are supposed to be drawn according to $\mathbf{z} \sim q_\phi(\mathbf{z}|\mathbf{x})$. However, this sampling would not enable the backpropagation of the gradient. Therefore, the reparameterization trick has been proposed (Kingma and Welling, 2014). Simply put, it expresses $\mathbf{z}$ as a deterministic variable: $\mathbf{z} = T_\phi(\mathbf{x}, \epsilon)$. In this case, ϵ is an auxiliary independent random variable and T_ϕ is a differentiable transformation function that converts ϵ to $\mathbf{z}$. Usually, $q_\phi(\mathbf{z}|\mathbf{x})$ is a Gaussian. Finally, we obtain:

$$\begin{aligned} \mathbf{z} \sim q_\phi(\mathbf{z}|\mathbf{x}) = \mathcal{N}(\mathbf{z}; \mu, \sigma^2\mathbf{I}) \\ \mathbf{z} = \mu + \sigma \odot \epsilon, \text{ where } \epsilon \sim \mathcal{N}(0, \mathbf{I}) \end{aligned} \tag{6.7}$$

where N is a Gaussian of mean μ and standard-deviation σ, and $\mathbf{I}$ is the identity.

One advantage of the VAE is that the latent space, due to the constraint of following a Gaussian distribution, is supposed to be composed of disentangled dimensions. This is especially interesting for interpretability purposes. To emphasize the learning of more disentangled representations, the $\beta - VAE$ was introduced (Higgins et al., 2016). It consists of a modification of the classic VAE to weight D_{KL}:

$$L_{Beta-VAE} = -\mathbb{E}_{\mathbf{z}\sim q_\phi(\mathbf{z}|\mathbf{x})}\, \log\, p_\theta(\mathbf{x}|\mathbf{z}) + \beta D_{KL}(q_\phi(\mathbf{z}|\mathbf{x})||p_\theta(\mathbf{z})) \tag{6.8}$$

Higher values of β would lead to more disentangled representations (Higgins et al., 2016; Burgess et al., 2018). In case of $\beta = 1$, the $\beta - VAE$ is a classic VAE.

6.2.3 . VAE as a generative model

In addition, unlike UNet or VNet which do not have guarantees regarding data generation since there is no constraint on the latent space, VAE models have the advantage of being generative. Indeed, as the latent space is a distribution, it is possible to generate new samples from the latent space. This is particularly interesting for interpretability purposes. As a matter of fact, with such models, it may be possible to understand causal relationships between characteristics in the data, their projection in the latent space and the generated output. This is capital when it comes to attempting to decipher biological processes. For instance, if a model performs classification between patients and healthy controls, we wish to understand the reasons why one subject was predicted as a patient. It can enable to trust the model, i.e., the features involved in the decision have a clinical meaning, or it can allow discovering biomarkers. A recent work presented a bi-directional approach to characterize the impact of variations in specific regions of the inputs

in the latent space, and on the contrary, how perturbations in the latent space could impact specific regions of the output (Liu et al., 2020).

The generation of new samples is performed once the model is trained. We sample a vector from the learned latent distribution which is then passed to the decoder and reconstructed. This method can be used to study the different latent dimensions. For instance, one can sample a latent vector with all the dimensions worth their average and then interpolate along the axis of one specific dimension.

6.3 . Implementation of the beta-VAE

For our two applications, namely the identification of typical folding patterns and of rare ones, I used a $\beta - VAE$ to model the inter-individual variability of folding.

Indeed, in neuroimaging applications, many works have used autoencoder frameworks (Chen and Konukoglu, 2018; Atlason et al., 2019; Baur et al., 2020). More widely, in medical images, GAN have been demonstrated to efficiently work in anomaly detection tasks in optical coherence tomography images (OCT) (Schlegl et al., 2017, 2019). This method was then transposed to brain images (Simarro Viana et al., 2021). In return, VAE-based models have also been widely used (Zimmerer et al., 2019a,b) and a comparison of AE models showed that VAE was one of the most efficient to detect anomalies in brain MR images (Baur et al., 2020).

Therefore, I used a $\beta - VAE$ to model cortical folding characteristics. The input x corresponds to our input data previously described and can be skeletons, foldlabels or distance maps. $\beta > 1$ was preferred in order to favour the disentanglement of the latent dimensions. For both applications, the global architecture was identical. It consisted in fully convolutional encoder and decoder of depth 3 (see Fig.6.1). Several hyperparameters were tuned:

- L: the latent space size,
- β: the weighing coefficient of D_{KL}.
- l_r: the learning rate.

The methods used to tune the hyperparameters depend on the objective and thus will be described later in part III and IV.

In addition to the $\beta - VAE$, I supervised a 6-months internship whose subject was to apply GAN to the identification of rare folding patterns. For the time of the internship, the results were not convincing and I preferred improving our approach with the $\beta - VAE$. This work does not present the results obtained with the GAN but are available at neurospin-projects/2021_adneves_skelegan/internship_report.

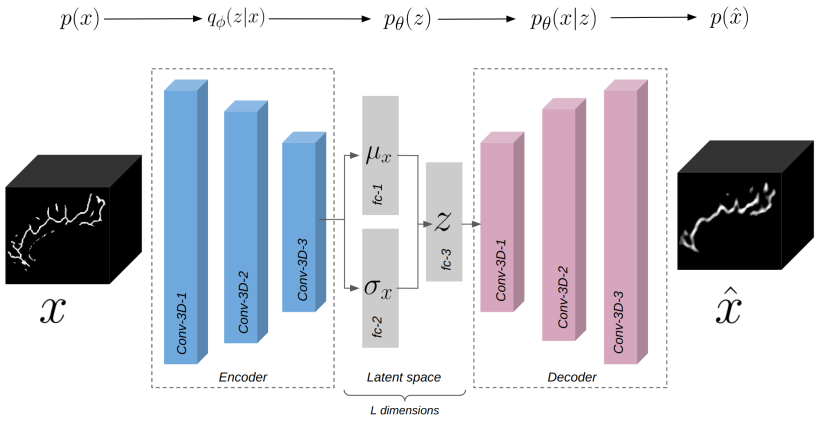


Figure 6.1: $\beta-VAE$ *architecture.* Illustration of the $\beta-VAE$ architecture used. The encoder and the decoder are fully convolutional. The depth is 3.

Part III

Identifying typical patterns

This part concerns the first application presented previously in 3.2. Based on the general methods described in the previous section, we aim to identify typical cortical folding patterns from the learned representation of the normal variability. We will first briefly introduce the context of this application. We will then present the specific material and methods before discussing the results.

Table of Contents

7 - Introduction

7.1 . Motivations

The attempts to describe and classify local folding patterns into categories are numerous and have led to several propositions. In 2004, a research program including the construction of an *"alphabet of the folding patterns"* was described (Mangin et al., 2004b). The use of the term *alphabet* is not coincidental. In this analogy, the sulci would constitute the words of a text. We can easily imagine that several people writing manually the same text would produce very different-looking scripts. Similarly, neurodevelopment results in varying patterns among the population. The different shapes and patterns of each sulcus would be similar to the various ways of manually writing: some words may be divided into several parts for certain authors for instance. Comparing with numbers is more visual, especially with the number 7 which can be associated with two classes, i.e. patterns: with or without a central bar (Fig.7.1). This example highlights another peculiarity: the two classes represent the main diverging characteristics, however, no two examples are strictly identical within a class. Similarly, the components of the folding patterns alphabet would encompass the main divergent features of a region, but the elements associated with each class would still exhibit their own uniqueness such as the digit 7, where on the right side, although both do not have a center bar, one is more rounded at the top.

Creating a dictionary of the typical patterns displayed by the normal population aims to use the patterns as *ideograms* that bring together the diversity of the folding. Adding this layer to analyze the patterns enables simplifying and reducing the folding variability: instead of dealing with as many sulci as individuals, we deal with a finite number of patterns. Having such a dictionary may serve several purposes. First, the patterns could be used for registration in the context of folding studies (Sun et al., 2009). The second purpose is to use the dictionary of *normal patterns* to study neurodevelopmental disorders that could result in rare or abnormal configurations. This goal corresponds to the second application of this thesis which I will discuss in part IV.

7.2 . Challenges and strategy

Identifying the main typical folding patterns is a challenging task. First, the folding variability is massive and thus requires tools powerful enough to account for it. In addition, currently, there is no gold standard in terms of a dictionary of patterns. Indeed, in the 90s, the different patterns were determined visually, based on dozens of specimens (Ono et al., 1990). Since then, several methods have tried to automatically identify the main cortical folding patterns (Sun et al., 2009,

Figure 7.1: *Two* patterns *of the digit 7 from the MNIST dataset.* Left: the digits seven are written with a central bar. Right: Two examples of the digit seven without a central bar. Within the two patterns, each digit exhibits its own specificity: in the left part: the position of the bar and the angle of the main part; in the right, the rounding of the digit or the line's thickness.

2012; Meng et al., 2018; Duan et al., 2019; de Vareilles et al., 2022). However, these methods are applied to different brain areas which makes it more difficult to compare their results for a consensus to emerge. Therefore, the evaluation of our method is complex.

To address this task and its associated challenges we proposed, for the first time, an unsupervised deep learning approach. This work was done in collaboration with Joël Chavas and led to a publication at MICCAI 2022 for which we are co-first authors: Guillon, L., Chavas, J., Pascucci, M., Dufumier, B., Rivière, D., Mangin, J.-F. Unsupervised Representation Learning of Cingulate Cortical Folding Patterns, in: MICCAI 2022, Lecture Notes in Computer Science. Springer Nature Switzerland, Cham, pp. 77–87 (Guillon et al., 2022). The validation strategy was developed later by Aymeric Gaudin, a trainee in the team, and his supervisor Joël Chavas.

8 - Specific Material and Methods

8.1 . The cingulate region

As we already mentioned previously, we are interested in identifying **local** patterns. For this application, we chose to study the cingulate region of the right hemisphere.

The cingulate area is composed of two main sulci, the cingulate sulcus and in some cases the paracingulate sulcus. Fig.8.1 presents the folding morphology of the region. The cingulate is present in each individual although its morphology has been found to be highly variable. Regarding the paracingulate sulcus, a study showed that it was present in almost 70% of cortices but there is an asymmetry as it is more frequent in the left than in the right hemisphere (Yücel et al., 2001).

The presence of these sulci has led to the definition of two patterns as presented in Fig.8.1B: the single type without a paracingulate sulcus and the double parallel type with it (Ono et al., 1990; Cachia et al., 2014). These two sulci are of interest as it has been demonstrated that subjects with poor inhibitory control and schizophrenic patients seemed to have more symmetric patterns between the two hemispheres than controls (Provost et al., 2003; Yücel et al., 2003; Borst et al., 2014; Cachia et al., 2014).

Nevertheless, these two patterns were not the only ones identified in the region. In another study, four different patterns were described (Meng et al., 2018). The patterns they proposed seem to be more fine-grained: they go beyond the presence or the absence of a paracingulate and introduce for instance interruptions of the cingulate. More recently, the same team improved their methodology and proposed five patterns (Duan et al., 2019).

Using this region is particularly relevant for our application and to evaluate our method. First, this region is sufficiently variable to justify the use of our methods. If a region was too stable, one could argue that it would not be appropriate. On the contrary, a region displaying a massive variability could result in many patterns and in a difficult assessment for this first attempt. Moreover, although there is no ground truth regarding the patterns existing in that area, the patterns proposed by the different studies cited previously enable to compare our results. Finally, as we mentioned, the cingulate region has a clinical interest in psychology and psychiatry. In addition, perturbations occurring during neurodevelopment, can affect this region. For example, patients with corpus callosum dysgenesis demonstrate a very different configuration (Bénézit et al., 2015).

In this application aiming at identifying typical patterns, we have a two-fold goal: to summarize the variability into main patterns based on a latent representation, and, within this representation, to be able to detect the presence of the paracingulate sulcus. Indeed, since the presence of the paracingulate constitutes a

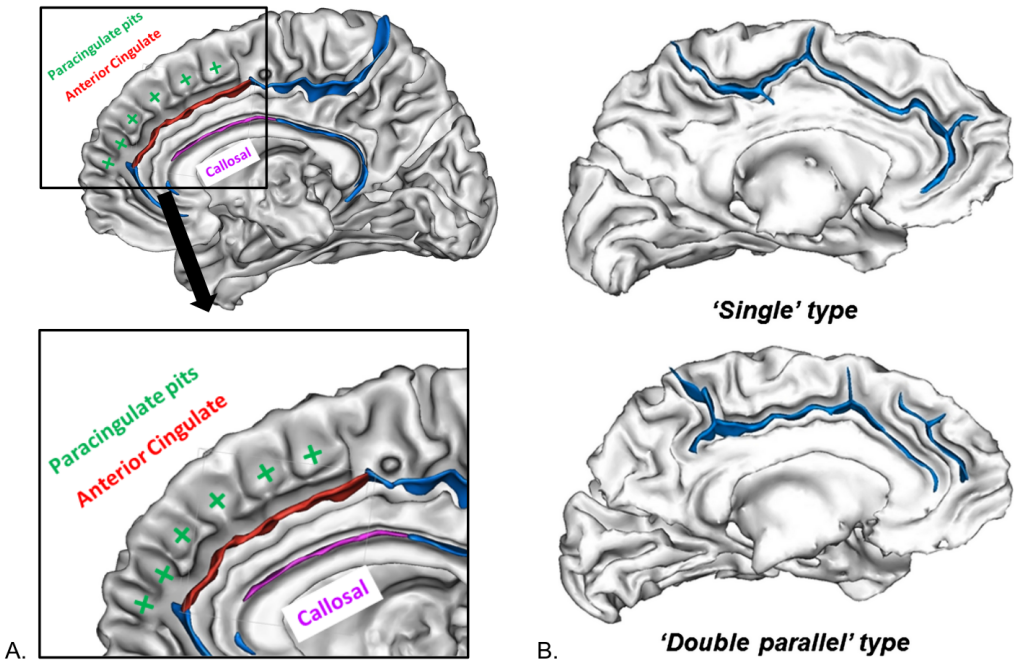


Figure 8.1: *Folding morphology of the Anterior Cingulate Cortex (ACC).* A. Description of typical folding structures in our ROI using the icbm152 average template. The cingulate region is composed of two main sulci: the cingulate sulcus (whose anterior part is in red), which is doubled with paracingulate pits (green crosses) that may be discontinuous or join to form the paracingulate sulcus, leading to numerous combinations. B. Two configurations deriving from the presence of the paracingulate. In the 'single' type, there is no paracingulate, whereas in the 'double parallel' type both the cingulate and the paracingulate are present. Courtesy of (Cachia et al., 2014).

major pattern in the region, assessing whether our latent representation contains features enabling to detect it could be a good proxy to validate our method.

To answer these questions, we used several datasets. First, we learned the folding normality on the HCP dataset which was presented in section 4.1. Then, to be able to assess our method on the detection of paracingulate sulci we need labels of their presence. To this end, we used a specific dataset which consists in an aggregation of several databases which were labelled w.r.t the presence or the absence of a paracingulate sulcus (Rapoport and Gogtay, 2011; Cachia et al., 2014; Chakravarty et al., 2015; Cachia et al., 2016; Tissier et al., 2018; Delalande et al., 2020). We will refer to this dataset as "ACC".

8.2 . Data preprocessing

Since the variability is already very important (see Fig.8.2), we used only the right hemispheres. As presented in part II which detailed the general methods, our

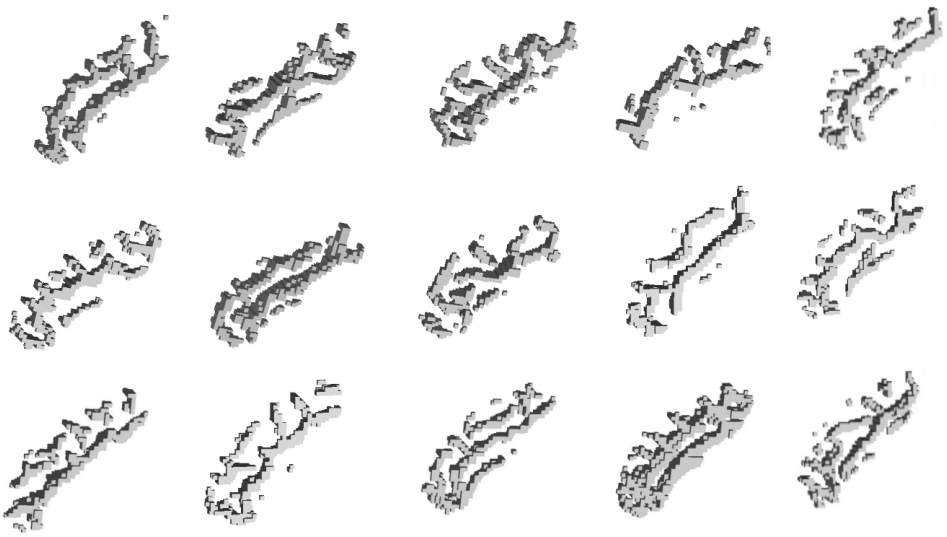


Figure 8.2: *Folding variability of the Anterior Cingulate Cortex (ACC).* 3D crops of the ROI obtained with our Deep Folding preprocessing pipeline. Each crop corresponds to one subject.

input is the skeletons obtained from the folding graphs of BrainVISA/Morphologist. As a reminder, these images are 3D quasi-binary images which represent a negative cast of the brain. We focus on the subregion of the anterior part of the cingulate and the paracingulate. In order to do so, we used our deep folding pipeline and more specifically the part that generates masks of specific sulci (see part 5.2). We ended up with two masks: one for the paracingulate and one for the anterior part of the cingulate sulcus. The two masks allowed us to define our ROI. In this case, the masks were applied prior to training. Our final input is a 2-mm resolution 3D crop of dimension 20x40x40 with integer values that represent local topologies. Fig.8.2 shows examples of our final inputs.

8.3 . Learning cingulate region representations

We compared two unsupervised deep learning models: the $\beta-VAE$ previously introduced and another framework based on contrastive learning which was done by my colleague Joël Chavas. These two models are trained on a population representing the normal variability, i.e. HCP. The aim is to learn to encode the most typical features in the latent space.

- $\beta-VAE$. I will briefly remind the main characteristics of this model, a more in-depth presentation is made in part 6.8. $\beta-VAE$ (Kingma and Welling, 2014; Higgins et al., 2016) is an autoencoder-based model which has the particularity of encoding inputs as distributions. Samples are drawn from this distribution and then reconstructed. The objective function is a combination

of the reconstruction error and the matching of two distributions using the Kullback-Leibler divergence. $\beta - VAE$ is trained to minimize:

$$\mathcal{L}(\theta, \phi; \mathbf{x}, \mathbf{z}, \beta) = -\mathbb{E}_{q_\phi(\mathbf{z}|\mathbf{x})}[\log\ p_\theta(\mathbf{x}|\mathbf{z})] + \beta \mathcal{D}_{KL}(q_\phi(\mathbf{z}|\mathbf{x})||p(\mathbf{z})) \quad (8.1)$$

The first part of the equation can be associated with the reconstruction error. In our case, as we train the $\beta - VAE$ on the binarized skeletons, we used the cross-entropy loss. We chose $p(\mathbf{z})$ to be a reduced centered Gaussian distribution.

- **Sim-CLR**. This model is an instance discrimination contrastive model (Chen et al., 2020). The main hypothesis is that two modified versions (views) of one image should be encoded closer in a latent space than the views of another image. For each sample x of a batch of size N, at each epoch, two views x_i and x_j are generated and fed into the model for which the model outputs are respectively z_i and z_j. The model is trained to bring together views from the same image in the latent space, that is to minimize $\sum_{i=1}^{N} \ell_{i,j=pos(i)} + \sum_{j=1}^{N} \ell_{j,i=pos(i)}$, $\ell_{i,j}$ being the loss function for a positive pair of examples (τ is a temperature parameter) :

$$\ell_{i,j} = -\log \frac{\exp(\text{sim}(z_i, z_j)/\tau)}{\sum_{k=1,k\neq i}^{2N} \exp(\text{sim}(z_i, z_k)/\tau)}, \quad (8.2)$$

In the version of SimCLR implemented by J. Chavas, the view generation is designed specifically for our inputs and leverages their characteristics. Indeed, it uses the discrete topology of the fold skeleton and of the bottom line in particular (described in part 5.1). Fig.8.3 presents visually the bottom lines. Based on this organization, the bottom line can permit to define an augmentation that relies on the topology of the skeleton: the bottom lines (red part in Fig.8.3) remain in all views but the inner part of some folds (orange part in Fig.8.3) are removed. More specifically, the first view combines random rotations between [-10,10]°, over all axes and a 60% rolling cutout with only the bottom lines kept ***inside*** the cutout volume. The second view combines again random rotations between [-10,10]° over all axes followed by a 60% rolling cutout but with the whole skeleton conserved inside the cutout, whereas only bottom values are kept ***outside*** the cutout volume. An example is shown in Fig.8.4. All views are then binarized. This topology-based augmentation forces the model to learn the sheet-based structure of the fold-based skeleton.

For interpretability purposes, we added a decoder to SimCLR. To reconstruct the latent codes, we freezed SimCLR weights and trained a decoder whose input layer is the representation space. The decoder backbone is the one of the $\beta - VAE$ decoder, and the loss is the cross entropy.

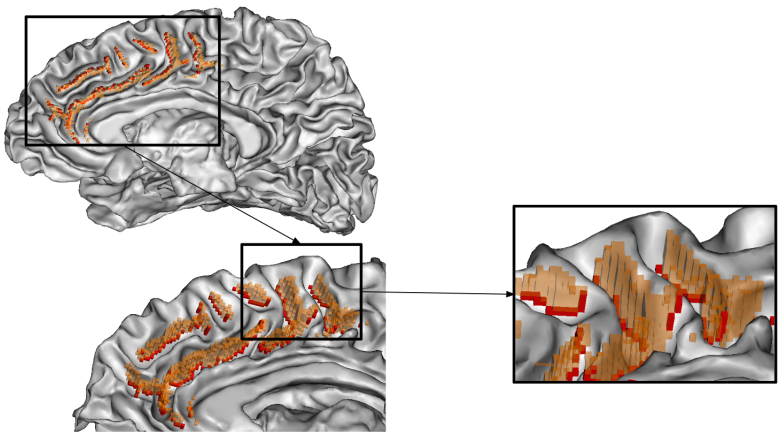


Figure 8.3: *Illustration of folds' bottom lines.* Orange represents the skeleton of the ROI. Red represents the corresponding bottom line. Grey is the white matter. Skeleton is represented here on 1 mm resolution. Note that the skeleton includes its bottom line, we use two colors to facilitate understanding.

Figure 8.4: *Examples of two augmented views.* In the two examples the arrows indicate the cutout where the "walls" of the folds have been deleted but the bottom lines remain.

As indicated in part 6.8, our β-VAE comprises fully convolutional encoder and decoder of symmetrical architectures with three convolutional blocks and two fully connected layers. The backbone of the SimCLR model implemented by J. Chavas is the DenseNet (Huang et al., 2017), followed by two fully connected projection overheads based on (Dufumier et al., 2021). To adapt to our smaller input, the size of the DenseNet network was reduced down to two dense blocks. The latent space corresponds to the representation space of the SimCLR model, which has a better representation quality than output space. Indeed, the contrastive loss could result in a loss of information in the output space useful for the downstream task (Chen et al., 2020).

8.4 . Latent space analysis

Our aim is to identify typical patterns. Therefore, once a condensed normal representation is learned, we have to effectively decipher the main features of this

representation.

8.4.1 . Characterizing folding patterns

To identify folding patterns, once the model is trained, all subjects are encoded to the latent space of both models and reduced to a 2-dimensions space with t-SNE algorithm (van der Maaten and Hinton, 2008). The reduction to two dimensions enables to get more hints of the learned representations and to analyze subjects groups more easily. As we wish to identify a finite number of patterns, we perform a clustering with hierarchical affinity propagation (AP) algorithm (Frey and Dueck, 2007) in the 2-dimensions space. One advantage of AP is that the number of expected clusters does not have to be precised. However, it may output a very large number of clusters, making it difficult to understand from an anatomical point of view. Hence, following the method used in (Meng et al., 2018), we applied the algorithm in an iterative way until a maximum number of five clusters. We stress out that the maximum of five clusters is an arbitrary number and that it has no biological meaning beyond facilitating our understanding.

Once the clustering is performed we have a finite number of clusters where each, hopefully, corresponds to a pattern. The analysis of the main anatomical characteristics of the clusters can be done either on the **latent codes** or on the **input space**. Indeed, for each subject of the clusters we have its skeleton in the input space and its corresponding latent representation.

Based on the latent codes. This first method enables to understand the encoded characteristics in the latent dimensions. More specifically, as the number of dimensions is much smaller than that of the input space, we hypothesized that the latent features enabling to cluster the space condense the variability of the folding characteristics. For the $\beta - VAE$, we leveraged the generative power and sampled the latent representation corresponding to each clusters' centroid and then decoded the images. In order to better understand the characteristics of each cluster, we travelled between them through the latent space to analyse variations across dimensions. For SimCLR, if it has not been originally introduced as a generative model, recent work demonstrates some links between contrastive learning and data-generating processes (Zimmermann et al., 2021). Nevertheless, here, we did not investigate the generative properties of SimCLR and we only reconstructed the latent representation of the nearest subject of each cluster centroid.

Based on the input space. The second method is performed in the input space, based on the skeletons. The clustering enables to obtain labels for all the subjects. We can then compute a local *per*-cluster averaging pattern in the input space (Sun et al., 2012). This method was adapted from the Statistical Probability Anatomical Map (SPAM) strategy. Specifically, all samples are aligned with one chosen as template. They are then summed to obtain a SPAM, also called 'local average'. The SPAMs are thresholded at the same level to be able to visualize in 3D. The threshold is chosen to prevent holes in the generated shapes. All in all,

we use the input skeletons of the subjects constituting each cluster to generate local per-cluster average pattern.

These two methods are both interesting and bring complementary information. As a matter of fact, the reconstructions based on the latent space comprise only *encoded* features. Decoding them enables to get an insight into what our models have retained as essential information to model the normal cingulate variability. In addition, the latent reconstructions are usually simplified representations which makes them easier to interpret. In return, the SPAMs are simplified representations of the *real* inputs. Thus, in a way, they effectively represent normal variability. However, since they are computed based on the input, they may be more sensitive to the diversity of folding and harder to interpret. In addition, they may serve to validate our latent representations depending on whether the patterns are similar or not.

8.4.2 . Ensuring representation generalization

When working on the characterization of typical folding patterns, a major question is whether the patterns found in one population can be recovered in another population. To test the generalization ability of our latent representations, we divided the HCP dataset in halves (HCP_1 and HCP_2) and trained both models (β-VAE and SimCLR) twice, on the two different subsets, leading to two encoders per model. These two encoders are referred as E1 and E2. We then encoded the first part of HCP (HCP_1) with the two encoders leading to two sets of embeddings per model. As mentioned previously, the different embeddings are reduced to a lower dimension space with t-SNE. E1 embeddings are clustered and we report the cluster labels to E2 embeddings. Fig.8.5 illustrates the different steps.

The generalization ability is assessed both visually and quantitatively: visually, we evaluate if the first visualization remains localized in the second visualization, meaning that the learned representations are not dependent on the training data but have captured some general features; quantitatively we measure the adjusted mutual information score (AMI) (Vinh et al., 2010). This metric assesses the similarity between two clusterings. In particular, we compute the AMI based on the cluster labels obtained on HCP_1's projection with E1 and on HCP_1's projection with E2 where we apply another clustering.

8.5 . Quantitative assessment: detection of the paracingulate

In addition to these analyses, we wanted to quantitatively assess our method for the identification of patterns. In particular: do the learned features enable to detect the presence of a paracingulate sulcus? To do so, we used the dataset labelled on the presence of the paracingulate sulcus (ACC). First, we divided this dataset into train and test sets. We trained five different models for both the

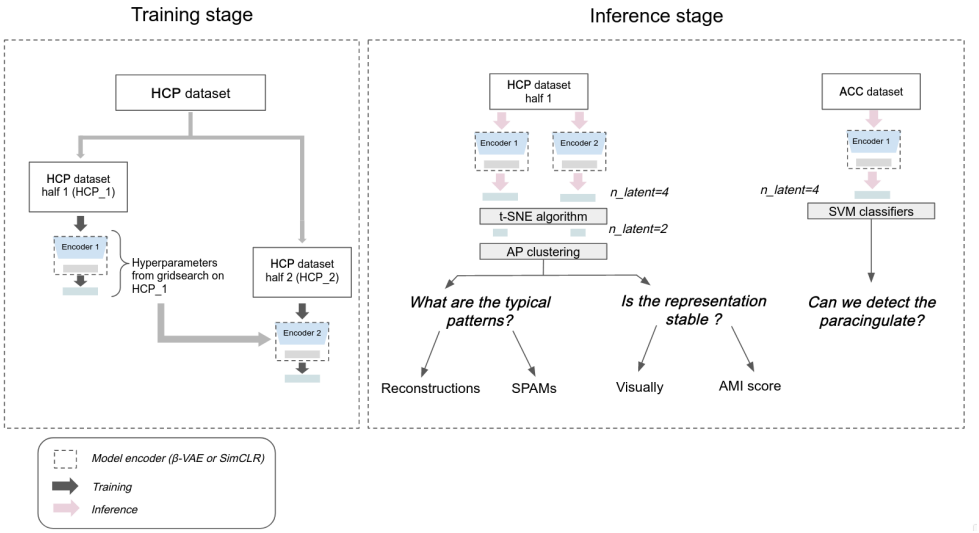


Figure 8.5: *Pipeline to identify typical folding patterns.* HCP dataset is split into two parts HCP_1 and HCP_2. Both models ($\beta - VAE$ and SimCLR) are trained on the two halves HCP_1 and HCP_2. The inference is performed on all four models on HCP_1.

$\beta - VAE$ and SimCLR on HCP_1. Then, we encoded the ACC sets and trained 250 linear Support Vector Machines (SVM) with cross-validation to classify between paracingulate and no paracingulate. The number of SVM (250) has been chosen arbitrarily. The metric used to assess the performances was the area under the receiver operator Curve (AUC).

Note that the development of the quantitative assessment was done by A. Gaudin under the supervision of J. Chavas. I adapted the scripts of the $\beta - VAE$, trained the $\beta - VAE$ and generated the embeddings of the ACC dataset for this model.

8.6 . Training strategy

As presented earlier, the training of the $\beta - VAE$ and of SimCLR was performed on the first half of the HCP dataset (HCP_1), which represents a total of 550 subjects. 80% of the subjects were used for training and the remaining 20% were used for validation (see Fig.8.5).

To find the best hyperparameters (size of the latent space for both models, β value for β-VAE and temperature τ for SimCLR), we performed a gridsearch where the best combination was chosen based on the highest silhouette score on the latent space that corresponded to decent loss value and reconstruction abilities for β-VAE. As a matter of fact, the silhouette score indicates whether the samples are projected in a clustered way or uniformly distributed. The best value is 1 and

the worst -1. 0 indicates the absence of clusters or overlapping clusters. Choosing the hyperparameters leading to a higher silhouette score pushes toward a more clustered representation. We considered the reconstruction quality as an important criterion since we want to interpret the representations based on reconstructions and generations.

We obtained β=2 (tested range 1-8) and τ=0.1 (tested range 0.01-0.3), as well as a latent size of 4 (tested range 2-150) for both models, which enabled to balance between the model performance and the clustering quality. Training of 300 epochs lasted for approximately 1 hour and 2 hours for β-VAE and SimCLR respectively, on an Nvidia Quadro RTX5000 GPU.

9 - Results

9.1 . Latent space structure

Fig. 9.1 (a) and (b) presents the clustering results. The silhouette score with AP on the latent space is 0.17 and 0.42, respectively for β-VAE and SimCLR. It becomes 0.43 and 0.44 when applied to the t-SNE space, indicating a tendency towards a clustered distribution with close clusters. This range of scores is common when dealing with complex data such as neuroimaging modalities (Lebenberg et al., 2019). For both models, four clusters were identified but the organization of the latent space is different: β-VAE latent space seems to distinguish four groups of subjects, separated only by a thin boundary whereas SimCLR latent space is more structured and could be interpreted as a manifold, consistent with the biological reality of folding patterns.

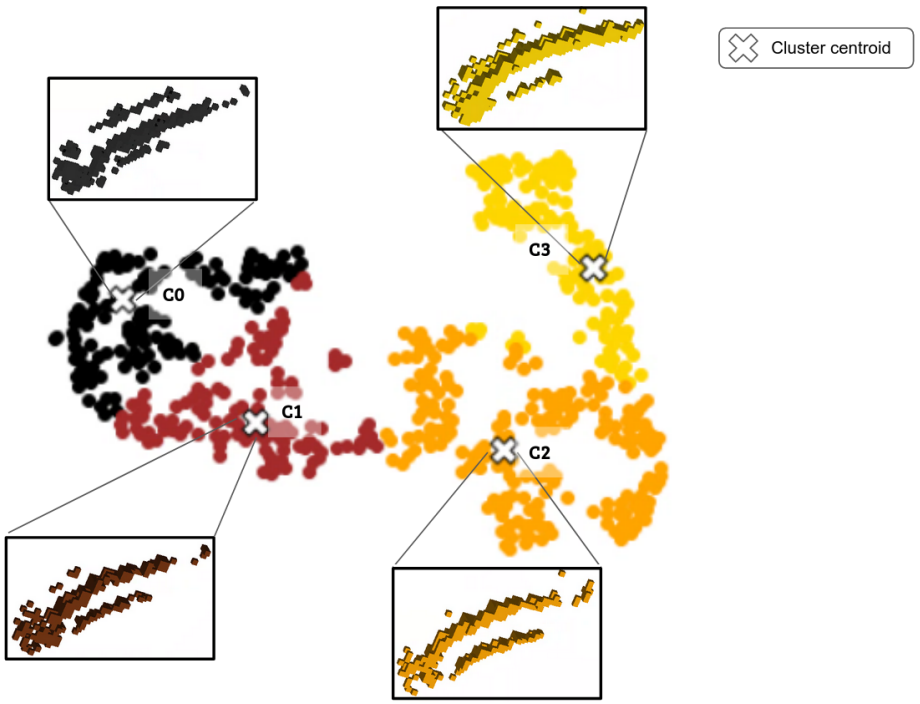
In Fig. 9.2, interpolating from one cluster to another shows that the latent space learned by the $\beta - VAE$ is continuous and regular, and we can progressively see the change of patterns.

9.2 . Deciphering the patterns

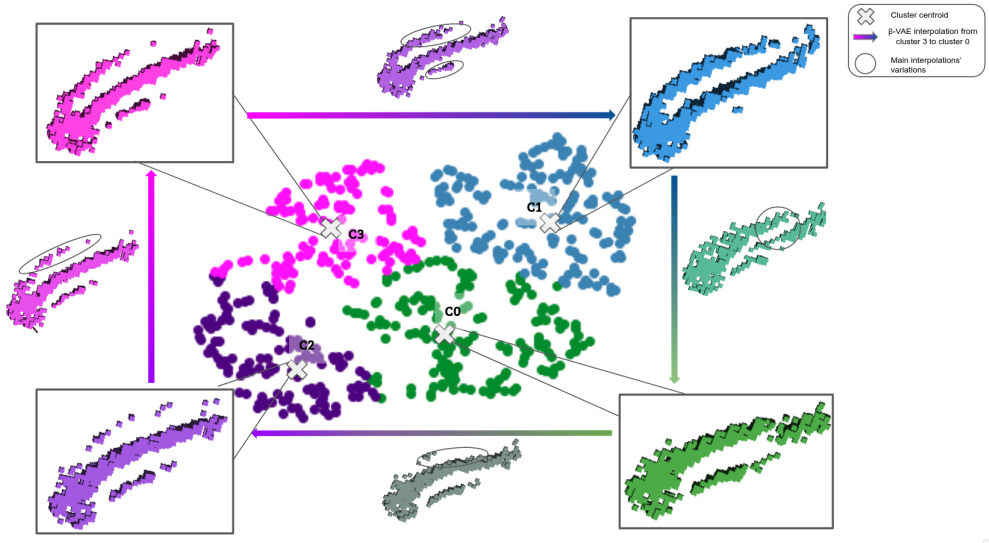
Both models were able to produce reconstructions that are compliant with the inputs, presenting a simplified version of the scene which enables to focus and bring out the most important features (Fig. 9.1 (a) and (b)).

For SimCLR, the black cluster seems to be the only pattern having a continuous paracingulate. The brown and orange clusters could correspond to a long cingulate and a marked callosal sulcus. Similarly, the yellow reconstruction seems to present a callosal and a cingulate sulcus. In addition, on top, there is a paracingulate which is not continuous but rather composed of independent pits as in Fig.8.1A. Based on the latent space organization, similar to a manifold, it is interesting to analyze the reconstructions in terms of evolution. The curvature of the longest sulcus becomes more bent from the black to the yellow cluster. From the black to the orange, we also witness the progressive disappearance of the paracingulate and a more marked presence of the callosal sulcus.

In the case of the $\beta - VAE$, cluster 0 (green) shows another pattern: a split anterior cingulate sulcus. This is even more visible thanks to the interpolations in Fig.9.2. Cluster 1 (blue) presents a cingulate doubled with a long paracingulate, while the pink has a shorter paracingulate and we can discern a piece of callosal sulcus. Lastly, the indigo presents a slight paracingulate split composed of several pits. Linking these visual descriptions to the clustering, it seems that the two



(a) SimCLR representation.



(b) $\beta - VAE$ representation.

Figure 9.1: *SimCLR and β-VAE latent spaces analysis.* (a) t-SNE representation of SimCLR latent space. Insets are decoded latent codes of nearest neighbours for each cluster centroid. (b) t-SNE representation of β-VAE latent space. Insets are decoded latent codes of cluster centroids and of interpolations between cluster centroids.

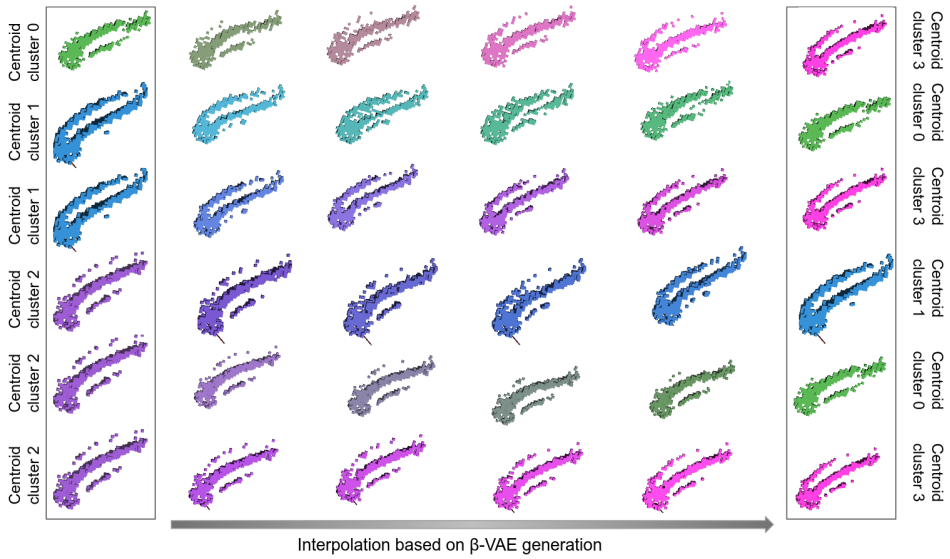


Figure 9.2: *Travelling through β-VAE latent space from one cluster to another.* The extreme left and right columns represent patterns generated from the centroids latent codes, the colours refers to Fig.9.1. Between centroids patterns, new samples obtained travelling through $\beta - VAE$ latent space illustrate the variations from one cluster's particularities to another.

upper clusters display a paracingulate whereas, the other two at the bottom do not.

It is interesting to compare these decoders' outputs with the SPAMs, i.e. cluster averages, based on the input space (figure 9.3). First, for both models, we can notice that the callosal sulcus (lowest sulcus) is similar in all averages. Therefore, it seems that it may not present very different appearances in the population and its variability may not be driving the clustering.

For SimCLR model, the black cluster average could correspond to a simple anterior cingulate. Subjects of the brown cluster could have a sketch of the paracingulate sulcus. The orange average is harder to interpret due to its thickness which can be due to either the presence of two sulci which are aggregated or to the variability of the position of the cingulate among the subjects. Finally, the yellow average includes a sketch of a sulcus parallel to the anterior cingulate, but in the left part of the ROI. Both methods, cluster averages and decoders, represent something different: in the first case, it is the geometrically-aligned average of all subjects in a cluster; in the latter case, it is the reconstruction of *one representative subject* from the latent space. According to our observations, they can converge either to the same (orange and yellow cluster) or to an apparently different (black cluster) representation.

For the β-VAE, we can also identify specificities for all the averages. Blue average seems to be the simplest pattern, a long cingulate without paracingulate or vertical branches. Indigo presents a sulcus parallel to the cingulate divided into

several pieces that may look like a sketch of the paracingulate sulcus. In return, the pink average includes several branches, vertical to the cingulate, that could not merge to create a parallel sulcus. Lastly, the green shows another pattern: a split anterior cingulate sulcus. When comparing to the $\beta - VAE$ reconstructions, we find a similar shape for the green average (cluster 0): a cingulate split in two. Conversely, for the blue cluster, based only on the average pattern, we interpreted a simple cingulate, but in the light of the reconstructions, the swollen anterior part could represent a paracingulate. Cingulate and paracingulate could be merged in the average representation due to positional variations among subjects. The indigo average is also consistent with the reconstruction. For the pink average, it seems to be a little bit different as it presents more vertical branches contrary to the reconstruction which displays a continuous paracingulate.

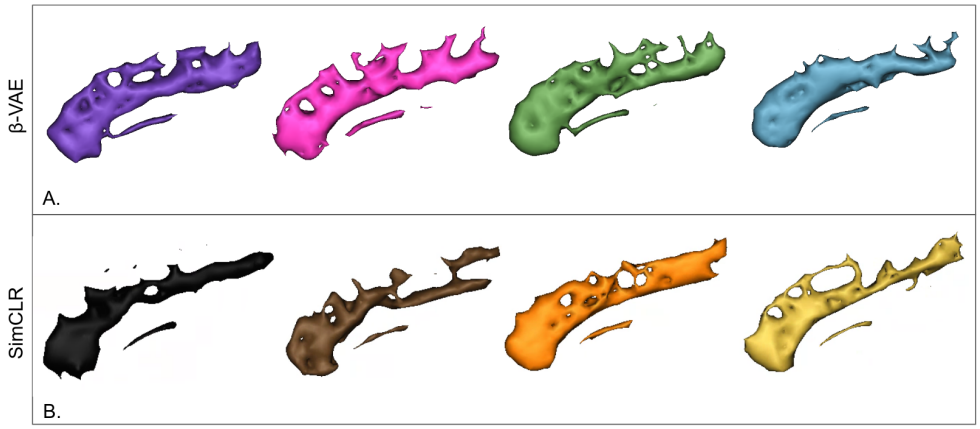


Figure 9.3: *Representative patterns as cluster averages (SPAMs).* Local average sulci obtained for each cluster with β-VAE and SimCLR encodings respectively. Colors match cluster colors of Fig. 9.1.

9.3 . Generalization abilities

Fig. 9.4 presents the visual assessment of the generalization. Part A. corresponds to the clusterings of Fig.9.1: HCP_1 is used to train the models and is then encoded to the latent space before applying the clustering algorithm. Part B. shows the same dataset but projected on the latent space learned on HCP_2. The cluster labels of part A. have been transferred to part B.

The comparison of the two clusterings (Fig. 9.4A. and B.) shows that some clusters remain more stable than others. With β-VAE embeddings, subjects of the blue cluster are still grouped. The pink and indigo clusters remain rather localized. On the contrary, the green cluster spans from the bottom to the left-hand corner through the pink and indigo clusters.

With SimCLR, the subjects of the black and brown clusters stay rather grouped.

It is also the case for the orange cluster even if a part is closer to a yellow area. In return, the yellow cluster is split in two. But overall, clusters' subjects seem to remain close to each other. It is interesting to note that the proximity order of the clusters seems to be kept: ranging from black, brown and orange. It is less obvious for the yellow cluster but in A., it appeared to start from the orange cluster and then to expand towards the black one. The split in half could result from this behaviour.

These observations are confirmed by the AMI score, which is 0.37 for the $\beta - VAE$ and 0.31 for SimCLR.

Both visually and quantitatively, the generalization seems slightly better for the $\beta - VAE$ than for SimCLR.

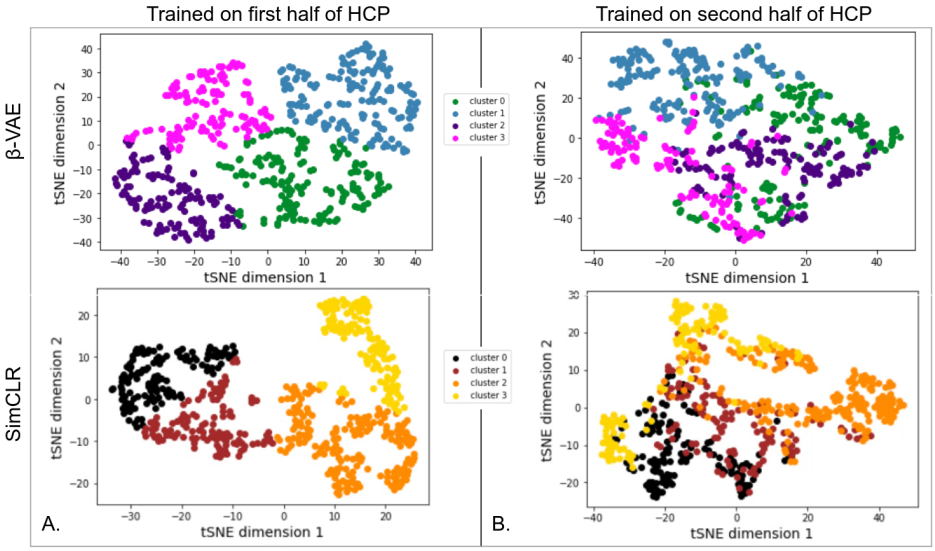


Figure 9.4: *Visual assessment of the generalization ability.* For both models, training is made on HCP_1 (model 1) and HCP_2 (model 2). The inference on HCP_1 leads to two encodings in the latent space obtained respectively with model 1 and model 2. To visualize the stability of the model and of its analysis, the t-SNE representation of model 1 is clustered (A.) and we report the labels on the t-SNE representation of model 2 (B.)

9.4 . Detection of the paracingulate

We now analyze whether the learned latent codes are relevant to detect one pattern described in the region: the 'double parallel type', in other words, the presence of a paracingulate sulcus. For this experiment, the dataset used was the labelled dataset ACC (8.1). Training SVM based on the latent codes to detect the presence of a paracingulate enables to assess whether our models have encoded features relevant for this type of patterns. The results are presented in table 9.1.

Model	Latent space size		
	4	10	30
SimCLR	0.50 (0.44)	0.51 (1.5)	0.46 (1.7)
$\beta - VAE$	0.63 (8.1)	0.69 (3.3)	0.63 (1.5)
PCA	0.54	0.55	0.58
Supervised CNN (Borne et al., 2021)	0.91 (0.03)		

Table 9.1: Classification performances of the presence of a paracingulate sulcus. Mean AUC (standard deviation) in % of the five different trained models. The standard deviation of the PCA was negligible so it is not reported.

We display the results for several sizes of latent space. We also compare the performances with a PCA and the supervised convolutional network proposed in (Borne et al., 2021). The setting of this supervised CNN is different but it establishes a reference. We expect our models to perform better than a PCA. On the contrary, the supervised CNN serves as an upper bound on the performance we could obtain.

First, we observe that in all cases, as we could have expected, the supervised model performs best. Furthermore, surprisingly, the $\beta - VAE$ outperforms massively SimCLR and especially for latent spaces of larger sizes. For example, with 10 latent dimensions, the performance is 0.51 for SimCLR compared to 0.69 for the $\beta - VAE$. In addition, SimCLR learned features lead to equivalent or worse results than a simple PCA.

It is interesting to note that the latent space size has an impact on the performances and that this effect is different depending on the model. With SimCLR, similar performances are obtained for 4 and 10 dimensions but a decrease is observed for 30 dimensions, whereas for the $\beta - VAE$, the AUC is the best for 10 dimensions and there are equivalent results for sizes of 4 and 30.

These results suggest that the latent space of the $\beta - VAE$ could be more suited to detect paracingulate sulci.

10 - Discussion and Conclusion

This part presented the methodology I developed in collaboration with J. Chavas and later A. Gaudin to identify the most typical patterns of the cingulate region. More specifically, based on the preprocessing presented in part 5.2, we compared two unsupervised deep learning models: a $\beta - VAE$ and SimCLR. We proposed a method to learn a representation of the normal variability in a latent space of fewer dimensions than the input space. From this representation, we were able to identify four main patterns in the cingulate region. We also assessed the generalization ability. Finally, thanks to the work of A. Gaudin, we were able to quantitatively evaluate our learned representations on their ability to detect the paracingulate, a pattern described in the literature.

10.1 . Comparison of the two models

Our two models presented different behaviours demonstrating particular properties. Although the clustering gave four clusters for both models, the reduced latent space is very different.

For SimCLR, we obtained four clusters which can be interpreted as different patterns: a configuration with a long cingulate and a paracingulate (black cluster), a long cingulate with pits of paracingulate for the brown cluster, a similar pattern for the yellow cluster but with a longer cingulate sulcus, and a complete absence of paracingulate sulcus for the orange cluster. However, the small differences between the yellow and the brown cluster could question this division into four patterns and we could wonder whether three patterns would not be more accurate. In addition, even if the structure of the latent space can remind a manifold organization, the clustering, especially between the black and the brown clusters, does not seem visually obvious. Therefore, the learned latent space could be not so well adapted for modelling the folding variability. This is also suggested by the classification performances which are worse than a PCA. One possible reason is the backbone used for the SimCLR model which could be too complex for our type of data and the task. This hypothesis is confirmed by some recent works by A. Gaudin and J. Chavas. Using the same encoder architecture as the $\beta - VAE$ achieves at least as good performances as the $\beta - VAE$. Another important element is the fact that, contrary to the $\beta - VAE$ where the reconstructions are the actual centroids, i.e. points that correspond to the average of the subjects of the cluster over all the latent dimensions, for SimCLR it is the reconstruction of **one** subject only, the closest to the centroid. Therefore, they are less representative than the patterns of the $\beta - VAE$. To have more robust representations, a way could be to take inspiration from the SPAMs and generate equivalent representations but on the reconstructions rather than on the input skeletons.

Regarding the patterns identified thanks to the $\beta-VAE$, the four demonstrate their own characteristics and some are similar to patterns described in the literature. In particular, a split of the cingulate sulcus was described as one of the patterns in several works (Meng et al., 2018; Duan et al., 2019). Two long parallel sulci were also identified (Duan et al., 2019). In addition, the generative power of the $\beta-VAE$ is a major advantage as we can sample from the latent space. Indeed, it can be very powerful to help interpret the encoded features and what characterizes each pattern.

Therefore, the reconstructions of the $\beta-VAE$ seem to be better suited for the identification of typical patterns, at least with the current framework of SimCLR.

10.2 . Representations based on the latent and the input space

It is particularly interesting to have visualizations based on the latent space and on the input space. As a matter of fact, as we already mentioned, the local averages are generated based on the *inputs*. Therefore they correspond directly to the normal variability. However, they may be difficult to analyze as all subjects of a cluster are equally weighted in the computation of the average. In case of high variability within a cluster, the average will therefore be blurrier. This can explain the thick appearance of some averages such as the orange or the blue ones. Indeed, having a thick representation may mean that the position is different among subjects: the cingulate may be located lower; it could also mean that some subjects have two parallel sulci but due to variations in the positions, it does not appear clearly. In our work, the local averages are rather difficult to interpret. In order to improve the representation we could weight the subjects depending on their distance to the centroid as it is done in other works (Sun et al., 2007, 2017; de Vareilles et al., 2022; Foubet et al., 2022).

In return, the reconstructions seem to be more precise and easier to interpret. Although in the case of SimCLR, they are just the reconstructions of one subject and thus do not enable to get a shared representation, it is interesting to contrast the $\beta-VAE$ reconstructions with the local averages. With the $\beta-VAE$, the reconstructions are the local averages of the encoded normal variability. We said previously that at least three out of four patterns were consistent across the two visualizations. This means that our model managed to effectively encode the main characteristics of the variability of the inputs.

Therefore, these two visualizations appear to be complementary and the SPAMs enable to validate the learned characteristics of our model.

10.3 . Accurately representing the folding variability

We already presented the two main approaches to study the folding patterns, namely based on a finite number of patterns, i.e. usually based on a clustering or,

on the contrary, based on manifold analyses. In this work, we chose to apply a clustering in order to interpret the patterns more easily. However, we introduced important biases regarding how the latent spaces should be structured (aiming at a higher silhouette score in the gridsearch) and imposing a maximum number of clusters. Indeed, the maximum number of clusters was mainly for practical reasons and does not have a biological meaning. If we had chosen a higher maximum number of clusters, it would have been tricky to analyse but it could have been closer to reality. In addition, according to the distribution of SimCLR latent space, a finer clustering, with a higher granularity could be of interest.

We did not investigate whether the latent space could constitute a manifold. In such a case, other models and another criterion for the gridsearch would be more suited. However, even if we targeted a clustered representation, the interpolations thanks to the $\beta - VAE$ show that the latent space is continuous and that the patterns gradually change from one cluster to another one which is totally compatible with the hypothesis of a manifold.

10.4 . Perspectives

This part proposed several method contributions that can be useful for the community. We introduced topology-based augmentations in the SimCLR setting, which is directly applicable for studies working on skeletons or similar inputs (Harrison et al., 2021; Rao et al., 2021). Moreover, we used for the first time local average folding patterns (Sun et al., 2012) in a deep learning pipeline. Our work also finds a structured latent space for the cingulate region with both models, $\beta - VAE$ and SimCLR and we were able to highlight at least four patterns that are consistent with the literature. The organization obtained with SimCLR seems more conforming with the anatomical reality of folding patterns and can be linked to folding manifolds (Mangin et al., 2016). In return, the generative aspect of the $\beta - VAE$ is a real lever to understand the learned representations and ease the analysis of this complex region.

To encourage a structured and well-separated latent space, we could introduce in the future cluster objectives in the learning phase both for generative models (Danks and Yau, 2021), and for contrastive models (Caron et al., 2018, 2021; Li et al., 2021b).

Finally, we found both cluster averages and decoder outputs to be similar to known cingulate patterns that correlate with executive functions and psychiatric disorders (Cachia et al., 2016). This similarity makes us believe that such latent space structures could correlate with medically relevant parameters. Our study is therefore a first step towards the systematization of the search for main region-specific patterns to then analyze their potential correlations with human cognition and disease.

Part IV

Identifying rare or abnormal patterns

In the previous part, we characterized typical patterns of the cingulate region. This part now concerns our second application where we aim to identify rare or abnormal cortical folding patterns based on the learned representation of the normal variability. First, we will introduce the context of this application before presenting the specific material and methods and discussing the results.

Table of Contents

11 - Introduction

11.1 . Motivations

Neurodevelopment involves many events and processes at different scales. In terms of morphology, they ultimately lead to the uniqueness of each brain and to folding patterns that can be similar to a fingerprint (Wachinger et al., 2015). In some cases, neurodevelopment can be affected by phenomena leading to malformations and rare or abnormal folding patterns. The resulting rare and abnormal patterns may be associated with disorders. Since the folding patterns are a trait feature, i.e. they remain stable throughout life (Cachia et al., 2016), they may be seen as a marker of the different neurodevelopmental processes involved and rare or abnormal folding patterns could act as a signature of certain pathologies. This is all the more interesting as currently, some disorders' diagnoses only rely on survey assessments. In these cases, having a morphologic landmark would be a real improvement. In addition, we usually do not have access to all the steps constituting the chronology of brain development and folding. Recent works have studied extremely preterm newborns and followed longitudinally the development of folding patterns (de Vareilles et al., 2022). This is very helpful in order to better understand brain formation. However, gathering such cohorts is a very complex task. To a lesser extent, studying rare and abnormal patterns in adults and what possibly caused them can help understand both folding formation and potential neurodevelopmental disorders. Therefore, identifying rare or abnormal folding patterns may lead to the discovery of new biomarkers, provide a better understanding of the folding processes and increase our knowledge of the brain.

11.2 . Challenges and strategy

The identification of rare folding patterns has not been deeply investigated so far. Most studies focus on shape descriptors and try to correlate them with psychiatric disorders for instance. However, we are interested in patterns rather than shape descriptors, as these are trait features. Several works have focused on patients suffering from epilepsy and tried to identify specific patterns. In the basal temporal lobe, the different patterns described by Ono (Ono et al., 1990) were used to categorize the patterns and compare their distribution between controls and patients but no new pattern was introduced (Kim et al., 2008). In return, another study brought to light an abnormal configuration in the central region, the PBS (Mellerio et al., 2014). The identification of this pattern was performed visually on a small cohort. The comparison between patients and controls enabled to identify the PBS and to determine whether the pattern was rare and specific to the cohort's patients. However, the major difficulty when doing it visually and

trying to see the difference between controls and patients is that it may result in a kind of overfitting on a small cohort and on a pattern. Indeed, usually, the patterns are not easy to identify which may lead to biases in the identification. In particular, the patterns may be hard to describe and identify due to the high inter-individual variability. In addition, patterns specific to patients of a given cohort may be present in controls of other cohorts. This may be suggested by the results in the automatic detection of the PBS in the HCP dataset (Borne et al., 2021). The CNN-based model proposed in this study identified 30% of PBS in the hemispheres of this all control population. Although some of these hemispheres were wrongly classified, it highlights that (1) the PBS is also present in a healthy population and that (2) its identification is not a trivial problem.

Using an automatic strategy on a wider dataset may constitute an opportunity to address these problems. Indeed, we hope that an automatic model is able to not overfit on a specific pattern. In addition, it has the advantage of being capable to study more subjects. This is capital since the folding is highly variable and a large sample is required to effectively model the normal variability and thus try to avoid generalization shortcomings. Nevertheless, several challenges remain.

First, it is essential to define the types of patterns we are looking for. In particular, in part 2.2, we have seen that the definition of anomaly, outlier or noise is not obvious and may depend on the application. In this thesis, we call a pattern *"rare"* if it deviates from the modelled norm and is present in a control population. In other words, a *rare pattern* is found in a healthy population and is not associated, given the state of the art, with a disorder. In return, an *abnormal pattern* also deviates from the norm but is associated with a disorder. Nevertheless, abnormal patterns may also occur in the healthy population because of resiliency. We stress out that rare patterns are not necessarily abnormal and associated with some disorders. However, whether they have a link with pathologies or not, rare patterns are interesting objects to study as they can constitute traces of neurodevelopmental processes.

A major difficulty in identifying deviating patterns is linked to the ground truth. Indeed, we seek to identify local rare folding patterns but very few have been described and not in all the brain areas. In the absence of a known rare pattern, it is hard to assess our methodology. Moreover, within a rare configuration, there exists a wide diversity of arrangements. This lack of homogeneity makes it hard to evaluate the difficulty of identification of one specific rare pattern for example. In a region where a rare pattern has been described, it is also possible that there are other rare patterns which have not been accounted for. This would hinder the detection performances. Last, as we deal with rare patterns, we do not have many of them, which makes the assessment difficult.

To address these questions, we proposed to design synthetic benchmarks of rare patterns. Having synthetic benchmarks is an effective way to assess our method-

ology more reliably. It also enables to study the *granularity* of deviations that can be detected. We mentioned that rare patterns could deviate for different reasons which could lead to a complex assessment and interpretation. Synthetic benchmarks are a way to precisely control the degree of deviations and therefore to gain insight into what may be identified as deviating. Here, we define granularity as the characteristics and properties of the deviations, such as their size or nature. The analysis of the granularity that can be identified aims to characterize the abnormal features that can be detected and at what level of detail.

In anomaly detection works based on VAE, in most cases, outliers are detected based on the reconstruction error. In this work, we also seek to describe which space is the most relevant to identify deviating patterns: is it based on the reconstruction error, in the input space, that is to say in our case, the *folding space*, or is it the latent space?

Based on the previously described preprocessing we model the inter-individual variability with a $\beta - VAE$. We then investigate the detection power of our methodology both on the latent space and on the *folding space*, using either our synthetic outliers or actual rare patterns. Finally, we assessed the generalization of our approach on other datasets presenting abnormal folding patterns in other regions. Fig.11.1 shows the main steps of the developed framework.

In this part, I first introduce the specific data and methods used, including the synthetic benchmarks generation, training strategy, identification of outliers on both latent and folding space, and generalization. Then I present the results before discussing them.

This part is based on an article submitted to the journal Medical Image Analysis: 'Identification of rare cortical folding patterns with unsupervised deep learning'.

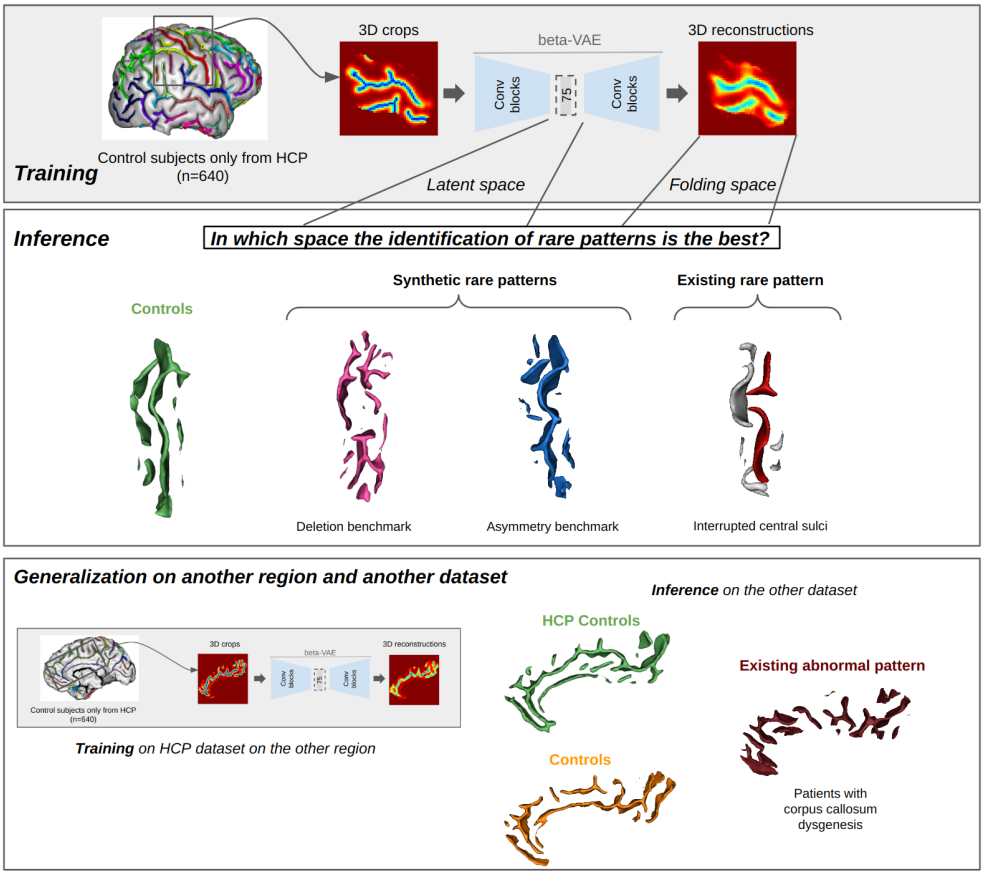


Figure 11.1: *Whole framework for identifying rare folding patterns.* Training is performed on the HCP dataset based on the 3D cropped distance maps. At inference stage, controls and subjects with rare patterns (synthetic benchmarks and interrupted central sulci) are encoded to the latent space and reconstructed. Based on their projection and the reconstruction error, we compare the latent and the folding space. To test the generalization ability, we train another $\beta - VAE$ on a new region (here the cingulate) on the HCP controls with the hyperparameters defined at the training stage of the central sulcus region. Finally, we try to identify existing abnormal patterns (children with CCD in the figure).

12 - Specific Material and Methods

12.1 . The central sulcus region

Folding mechanisms may lead to both global and regional anomalies and these two scales have led to correlations with function disorders (Fernández et al., 2016). In this thesis, we focus on regional patterns rather than on a global representation. Specifically, in this part, we concentrate on the central sulcus (CS) which is a good candidate for our work. Indeed, it is one of the first folds to appear and it is stable enough to be a first step in modelling inter-individual variability. More importantly, usually long and continuous, the central sulcus can be interrupted in very rare cases (less than 1%), making interrupted central sulci relevant patterns to assess our method. Finally, this region is of clinical interest as it is linked to hand motricity and asymmetries have been described (Sun et al., 2012; Bo et al., 2015). For this application, we model the inter-individual variability on the HCP dataset (see section 4.1). Examples of central sulci in HCP are presented in Fig.12.1.

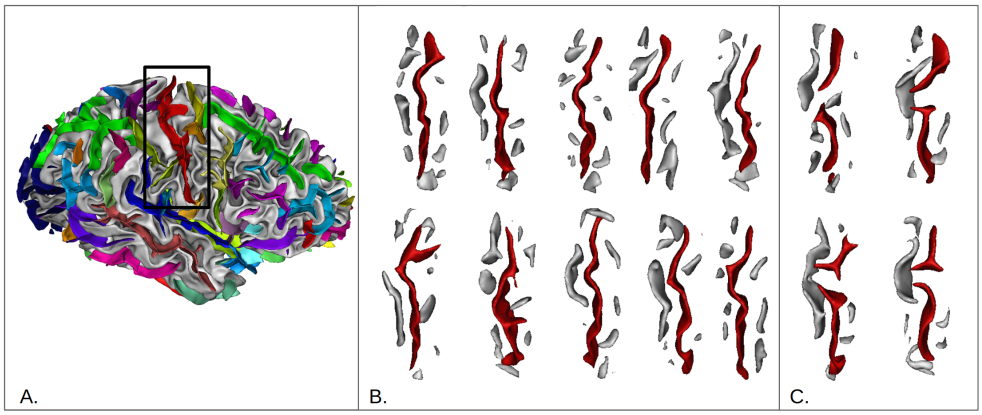


Figure 12.1: *Central sulcus region variability.* A. Localization of the studied region of interest (ROI) on a 3D view of one right hemisphere. The colored ribbons represent sulci, defined as a negative cast of the furrows. The central sulcus is red. B. Examples of non-interrupted central sulci. C. Examples of interrupted central sulci.

The long-term goal of this work is to identify rare folding patterns that have not been characterized yet. However, we first need to assess our method. To do so, we decided to work on a rare pattern already described, the interrupted central sulci. A previous study identified in the HCP dataset seven sulci in the right hemisphere and two in the left (Auzias et al., 2015; Mangin et al., 2019). The identification was based on the depth profiles of the sulcal pit maps, which are defined as the locally deepest points of the cortical surface (Lohmann et al., 2008), and it was then visually confirmed. We chose to work on the right hemisphere rather than on

the left in order to have the highest number of rare patterns. Four examples of interrupted central sulci are shown in Fig.12.1C.

12.2 . Data preprocessing

In this application, we used distance maps as input data. They are obtained thanks to our preprocessing pipeline based on the folding graphs of BrainVISA/Morphologist (see section 5.2 and Fig.5.2). As a reminder, distance maps are continuous images deriving from the skeletons where voxels' value correspond to their distance to the nearest fold. To focus on the central sulcus region, we defined a mask of the right central sulcus. Therefore, the sulci labels are not needed. We then affinely registered the distance maps to the ICMBc2009 space and cropped them according to the mask bounding box. In the end, our input data are 3D crops of the central sulcus region of dimensions 78 x 63 x 88 and isotropic resolution of 1mm. The mask is applied on the fly during the training of the network. Fig.12.1B. and C. show examples of our masked input crops as meshes.

12.3 . Learning central region representations

We used a $\beta - VAE$ to model the inter-individual variability. As a reminder, this model is based on an auto-encoder framework where the input is projected as a distribution into a latent space of fewer dimensions. More details are presented in part 6.8. In this case, as distance maps are continuous images, we used the Mean Squared Error (MSE) as the reconstruction error.

12.3.1 . Preprocessing

For this application, we use distance maps rather than skeletons. Therefore, the input data of the model are the previously defined, just cropped, then masked distance maps. Contrary to the previous application, we do not apply the mask before feeding the data to the model, but rather on the fly during training. For augmentation purposes, random rotations between [-10°, 10°], centered on the mask center, are drawn from a uniform distribution at each epoch and applied to the whole image, before applying the mask that strictly remains at the same position. Such rotations are also sought to limit the edge effects. More precisely, the central sulcus is surrounded by two main folds, the precentral and the postcentral sulci. Parts of these sulci are included in the ROI. Therefore, rotating the distance map *under* the mask enables to capture a wider context and to try to limit their influence.

We observe that skeleton voxels equal 0 in the initial distance maps X and the values increase with the distance to a sulcus, possibly ranging up to 10 mm. Potential reconstruction errors near the sulci would be minor compared to the voxels located far from them at the edge of the mask, whereas we wish the model

to concentrate more on the sulci. Thus, to limit the impact of distance variability far from the sulci, we perform a normalization according to equation 12.1, resulting in values between [0, 1], with the highest values on the folds and a saturation at about 4-5 mm which corresponds to half of the typical width of gyri. An example is shown in appendix A. Finally, we apply a small padding, resulting in samples of dimensions 80 x 80 x 96.

$$X_{norm} = 1 - [2\frac{1}{1+e^{-X}} - 1] \tag{12.1}$$

12.3.2 . Training strategy

Dataset was split into train, validation, and test sets of respectively 640, 161, and 200 subjects. Training is only performed on control data, all identified interrupted central sulci (CS) were added to the test set. The interrupted central sulci were identified based on the detection of the two main sulcal pits of the central sulcus, between which a depth profile was computed to determine the depth of the "pli de passage frontal-moyen" (PPFM), usually a buried gyrus in the sulcus. Subjects with a shallow PPFM were then manually inspected to determine whether the surrounding central sulci were interrupted (Mangin et al., 2019). However, we point out that there may remain some undetected interrupted central sulci in the training set as all subjects were not individually inspected.

To model the normal inter-individual variability, we used the same convolutional $\beta-VAE$ of depth 3 presented in section 6.8. In order to choose the best values for β and latent space dimension L, we performed a gridsearch (β=2-10, L=4-150). To assess each parameter configuration, we used two criteria. Our first criterion is the reconstruction quality. Indeed, we seek to leverage the reconstruction and generative power of the $\beta-VAE$, hence the reconstruction quality must be sufficient. Our second criterion is the detection power on a proxy for the interrupted central sulci. The pre-central and post-central sulci demonstrate some similarities with the central sulcus in terms of orientation, size, and shape. However, they tend to be more interrupted and to present a higher number of ramifications. Therefore, we used the HCP dataset crops of these two other regions as fake outliers. We selected only pre- and post-central sulci which presented some ambiguities with the central sulcus based on the procedure described in appendix B. Finally, our ambiguous set was composed of 28 precentral sulci and 18 postcentral sulci. For each hyperparameter combination, we trained a $\beta-VAE$ on the train set, then a linear SVM was trained to classify between the latent codes of the validation samples and of the pre- or post-central sulci. We kept the hyperparameters that led to the best classification results and good reconstructions (based on reconstruction error and visual inspection). To avoid overfitting we applied an early-stopping strategy.

12.4 . Synthetic rare patterns generation

One of the challenges of our work is the lack of consensual rare patterns to evaluate our methodology. In addition, it would be interesting to be able to quantify the degree of deviation that our model is able to detect. Therefore, several sets of synthetic rare patterns were generated to be used as benchmarks: *Deletion benchmark* and *Asymmetry benchmark*.

12.4.1 . Deletion benchmark

Our first benchmark consists of subjects for whom we have erased one simple surface (SS). As a reminder, a simple surface is an elementary fold segmented based on topological and depth properties (see section 5.1).

Erasing small simple surfaces could be a good proxy to simulate rare patterns because some fold branches may be missing in some people, or a sulcus may be shorter or absent. Large simple surfaces are less likely to be missing but allow us to assess the degree of deviation that can be detected.
Deleting simple surfaces directly on the distance maps would not be interesting as the voxels next to the simple surface indicate the SS position. To tackle this issue, the suppression was done during the generation of the raw skeletons. The distance map is then computed based on the pruned skeletons.

To analyze the granularity of anomaly that can be detected by our method, we generated several benchmarks which vary according to the size of the deleted simple surface (SS). As such, we created four sets where SS size was between 200-500 voxels, 500-700 voxels, 700-1000 voxels, and simple surfaces of more than 1000 voxels. In the following, we name each set with the minimum number of voxels: for instance, *200* corresponds to the benchmark where simple surfaces of size between 200 and 500 were erased. To be deleted, simple surfaces must have a number of voxels included inside the mask corresponding to the range of the different sets. If several simple surfaces meet the criteria, one is randomly chosen to be erased. Otherwise, a subject may not have a simple surface satisfying the requirements. In such cases, the subject is not included in the benchmarks. Finally, from the 200 test subjects, benchmark 200 contains 180 subjects; benchmark 500, 68; benchmark 700, 108 and benchmark 1000, 151 subjects.

To have a better representation of the amount of deleted sulci, Fig.12.2 shows the simple surface sizes distribution in the central sulcus region. The figure shows that our crops contain a large majority of very small simple surfaces (less than 500 voxels) and far fewer large simple surfaces. The smaller simple surfaces are mostly part of the precentral and postcentral sulci, representing more than 85% of the surfaces between 200 and 500 voxels. On the contrary, larger simple surfaces correspond to the central sulcus. Therefore, beyond deleting simple surfaces of varying sizes, the nature of the sulci and thus the location, are also different, especially between the set 200 and the others. The right part of Fig.12.2 shows the number of voxels corresponding to skeletons in our crops. It demonstrates the progressive intensity of anomalies when deleting simple surfaces from 200 voxels to more than 1000 voxels. Indeed, when simple surfaces of more than 1000 voxels

are deleted, it corresponds to a third or a quarter of the skeleton crop. Distance maps are then generated according to 5.2. An example is presented in Fig.12.3.

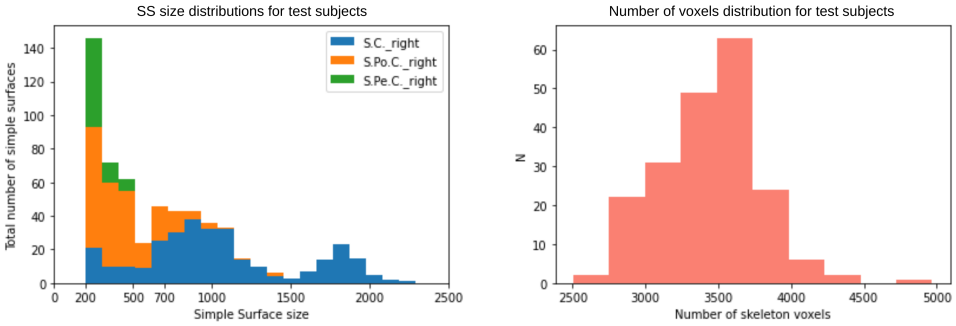


Figure 12.2: *Skeleton's description of the test set.* Left: Stacked histogram representing the distribution of simple surfaces sizes for the test subjects for the three main sulci of our crop, the central sulcus (S.C._right), the precentral sulcus (S.Pe.C._right) and the postcentral sulcus (S.Po.C._right). (Note: The labeling used is automatic and therefore not entirely reliable, but these labels are sufficient to draw conclusions regarding the SS size distribution.) Right: Distribution of the number of skeletons' fold voxels for the test subjects when the mask is applied to the crops.

12.4.2 . Asymmetry benchmark

Our second benchmark leverages the asymmetries described in the central sulcus region which concern several folding features in particular (Sun et al., 2007; de Vareilles et al., 2022; Foubet et al., 2022). Using crops of the left hemisphere as outliers enables to assess whether we can identify shape variations. In practice, this benchmark corresponds to the equivalent crop but in the left hemisphere. Left hemisphere distance maps are generated according to the same methodology as the right. Like our control crops of the right hemisphere, we computed a left central sulcus mask on the labeled dataset. To enforce the exact same crop size, we adapted the mask to match the adequate dimensions by adding or deleting a few voxels. Once the crops were obtained, they were flipped. During training, the right central sulcus mask was applied on the fly. We emphasize that we did not use the interhemispheric plane-symmetric coordinates but a mask specifically designed for the left central sulcus. This is especially important since there is a slight asymmetry in the position of the central sulcus between the two hemispheres (Davatzikos and Bryan, 2002). An example is presented in Fig.12.4.

12.5 . Identifying outliers

Once the model has learned a representation of the inter-individual variability, outliers identification can be performed at two levels. Traditionally, anomaly detection with AE is done based on the reconstruction error and an error map can be obtained comparing the input and the output (Schlegl et al., 2017; Pinaya et al.,

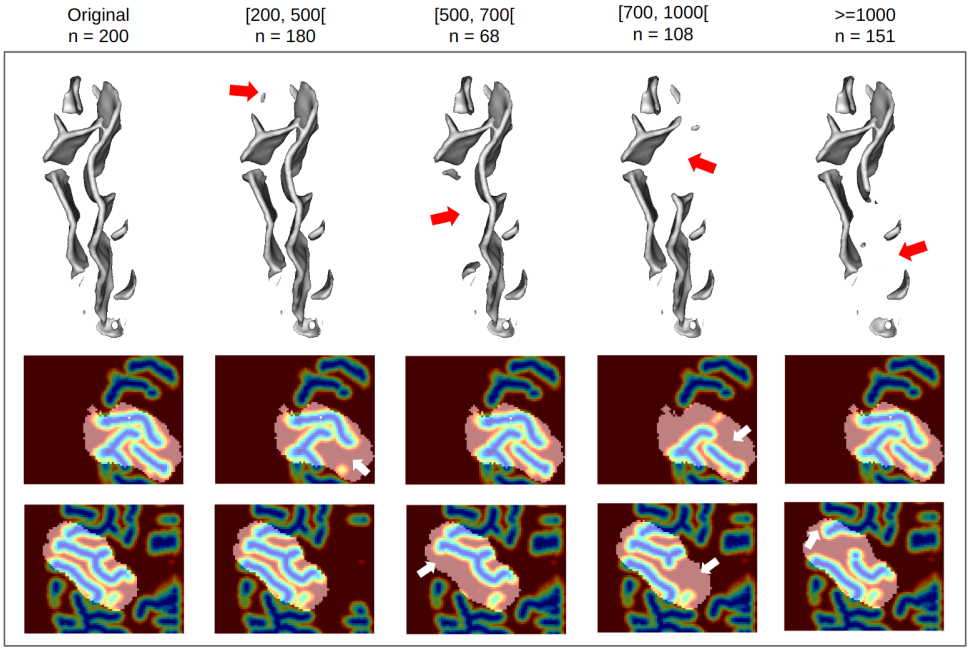


Figure 12.3: *Deletion benchmarks.* Visualization of original sulcal pattern and its altered versions from the four deletion benchmarks showing patterns with increasing simple surface size deleted. Upper row: Mesh visualization. Middle and bottom rows: distance maps on axial view, visualization at depths 15 and 37.

2018; Schlegl et al., 2019). But one can also wonder about the distribution of outliers in the latent space. Are the outliers distributed differently? To answer this question, we investigated the detection power in the outliers' distribution in the latent space and based on the reconstruction errors performed in the input space—which we call *folding space* in our case, as we study folding patterns. For both approaches, control test images and outlier images (deletion benchmarks, asymmetry benchmark and interrupted sulci) are encoded and reconstructed by our trained model.

A specification on data

As mentioned in 12.3.2, our control test set comprises 200 subjects. However, when studying our different outliers sets, data subsets were different since some subjects did not have any SS meeting the benchmark's criteria.

- **Deletion benchmarks**: to avoid any bias, we used only control subjects with a simple surface meeting the benchmark's criteria for each benchmark. Therefore, for benchmark 200, we used 90 controls that have a simple surface between 200 and 500 voxels but that has not been erased, and 90 benchmark subjects, for whom simple surfaces were actually erased. Resulting in $n^{200}_{control} = n^{200}_{deletion} = 90$, $n^{500}_{control} = n^{500}_{deletion} = 34$, $n^{700}_{control} =$

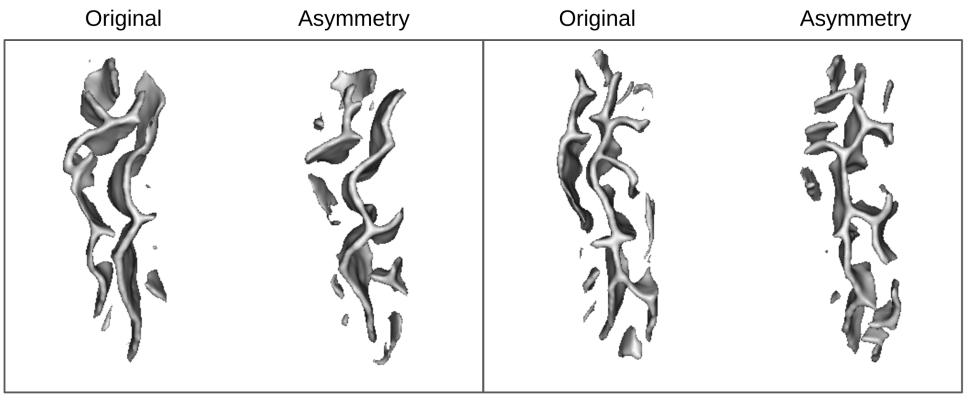


Figure 12.4: *Asymmetry benchmark.* Visualization of the original sulcal pattern and its flipped contralateral version for two subjects.

$n^{700}_{deletion} = 54$ and $n^{1000}_{control} = 75$ and $n^{1000}_{deletion} = 76$.

- **Asymmetry benchmark**: all subjects have their asymmetric counterpart. Hence, 100 subjects were randomly picked among the subjects from the test set for whom we took their asymmetric version. Resulting in $n_{control} = n_{asymmetry} = 100$.

- **Interrupted central sulci**: the whole test set is used as control data, leading to $n_{control} = 200$ and $n_{interrupted} = 7$.

12.5.1 . On the latent space

A hint from the visualization

For both of our benchmarks and the interrupted central sulci, we first sought to have a visualization of data distribution in the latent space. Therefore we projected encoded data into a smaller space of two dimensions with UMAP algorithm (McInnes et al., 2018, 2020). This projection enables us to get a first hint as to how outliers are represented.

Assessing the detection power on the benchmarks

However, the UMAP algorithm drastically reduces dimensions, leading to some information loss. We tried to assess whether relevant information regarding folding patterns was encoded in the latent space. Therefore, we trained linear support-vector machines (SVM) (Pedregosa et al., 2011) on the latent codes with stratified cross-validation to classify between control data and benchmark. Performance is assessed based on the ROC curve.

Quantifying the marginality of interrupted central sulci

As interrupted central sulci are very few, we cannot use classification as we did for the benchmarks. Classic machine learning out-of-distribution (OOD) algorithms are more suited. Therefore, to quantify whether the interrupted sulci are likely to be detected from their location in this reduced space, we applied two classic algorithms, One-Class SVM (OCSVM) (Schölkopf et al., 2001; Pedregosa et al., 2011), and isolation forest (Liu et al., 2008; Pedregosa et al., 2011) based on the data coordinates in the UMAP space. Applying the OOD algorithm in the UMAP space may lead to a loss of information but is easier as there are fewer dimensions and allows us to understand more easily the decision boundary based on the 2D representation. However, interrupted central sulci may not be the rarest pattern, and other folding configurations may be very scarce. Therefore, we also looked at control subjects repeatedly predicted as outliers by these algorithms.

Travelling through the latent space

Finally, to better understand the encoded properties and the learned representations, we leverage the generative power of the $\beta - VAE$. We computed average representations from different sets of data points, taking the mean for each dimension of the latent space. We then reconstructed these vectors. To further analyze the latent space, we traveled through it, going from one point, either the average pattern or a subject, to another point in the latent space, linearly interpolating vectors and reconstructing them.

12.5.2 . On the folding space

Outlier identification in the folding space relies on the model's error. The reconstruction errors' distributions were compared visually and assessed with the Kolmogorov-Smirnov test for the benchmarks and with the Mann-Whitney U-test for interrupted central sulci. For both cases, the null hypothesis was that the two distributions were identical.
The other strength of analyzing this space rather than the latent space is that the model's errors can help understand and locate the rare patterns' characteristics. To localize the errors, we commonly look at the residuals, which are the difference between the input and the reconstruction of the model. This corresponds to what the model has missed or added. To differentiate these two types of errors, we looked at them independently, computing the difference between the input and the output, i.e., the model's omissions, and between the output and the input, i.e., the model's additions. It is particularly interesting in the case of interrupted sulci, as we could expect that the model makes them continuous.

12.6 . Generalization to other regions

To assess the reproducibility in other regions and to ensure that our framework is not limited to the central sulcus, we transposed our methodology to two other regions for which an abnormal pattern has been described.

On the one hand, the isolated corpus callosum dysgenesis (CCD) leads to a cortex anomaly located in the cingulate region. This disorder is a congenital malformation that results in a complete or partial absence of the corpus callosum. The corpus callosum is composed of fibers that connect the two hemispheres.

On the other hand, focal cortical dysplasia of type 2 (FCD2) is a major source of drug-resistant epilepsy and is usually associated with anomalies of gyration and sulcation that may be localized in various areas of the brain. Specifically, in the central region, the Power Button Sign (PBS) has been proposed as a qualitative criterion to diagnose FCD2 (Mellerio et al., 2014).

12.6.1 . Children with CCD

The dataset includes 7 children between 9 and 13 years old presenting an isolated CCD and 7 matched control children (Bénézit et al., 2015). Among the patients, 3 present a complete agenesis, 3 a partial agenesis, and one a hypoplasia, corresponding to "a homogeneous reduction of the callosal size" (Tovar-Moll et al., 2007). In this case, the corpus callosum is completely formed, but abnormally small (Bodensteiner et al., 1994). For all children, the CCD was not associated with other malformations or developmental disorders. As presented before, we used T1-w MR images obtained from a Siemens Tim Trio 3T scanner with an isotropic resolution of 1mm.

The described anatomical anomalies associated with CCD include "sulci radiating on hemisphere medial surface, complete or partial absence of the calloso-marginal sulcus and of the cingulate gyrus" (Bénézit et al., 2015). Therefore, we transposed our method to the cingulate sulcus region. Using the same methodology as presented before, we computed a mask of the cingulate sulcus (gathering the calloso-marginal anterior and posterior fissure in the BrainVISA nomenclature), resulting in crops of dimensions 30 x 128 x 125 and 30 x 130 x 108, which were padded up to 32 x 128 x 128 and 32 x 136 x 112 respectively for the right and left hemispheres. Since the corpus callosum connects the two hemispheres, CCD can be studied equally in both hemispheres. Therefore, we conducted our experiments in the right and in the left hemisphere.

12.6.2 . Patients suffering from FCD2

The dataset includes 19 controls and 29 patients who can have either a positive or a negative MRI (Mellerio et al., 2014). This distinction was proposed by Mellerio et al. (Mellerio et al., 2012): a subject presents a positive MRI if "at least one of the cardinal MR signs of FCD2 (ie, cortical thickening, blurring, cortical and/or subcortical signal changes, transmantle sign) was present" (Mellerio et al., 2014). 12 subjects have a negative MRI and 17 subjects a positive one in the dataset. In addition, the identified lesions may be located either in the right or in the

Characteristics	ctrl	+ / right	+ / left	- / right	- / left
N	19	8	9	7	5
Age	31 (22-53)	21 (11-40)	34 (20-67)	16 (7-42)	18 (7-29)
N 1.5T scanner	19	5	6	7	4

Table 12.1: Characteristics of the dataset of patients suffering from FCD2. + and - indicate whether the MRI is positive or negative. Right and left correspond to the side of the lesion. The age is given as average age (min-max).

left hemisphere, based on histological analyses. Table 12.1 provides the number of subjects of each group. Although we only apply our framework to the right hemisphere as a first step, it is also interesting to analyze patients with the lesion located in the left hemisphere. Indeed, we can expect these patients to behave more like controls. For all subjects, we used the T1-w MR images. The majority was obtained with a GE Healthcare Signa 1,5 T scanner with a resolution of 0.98 x 0.98 x 1.40 mm. The other subjects were scanned with a 3T scanner with a resolution of 1 x 1 x 1.2 (see Table 12.1).

The subjects were included in the original study because their epilepsy was localized in the central region. The consequence of the FCD2 in terms of folding patterns is still an area of research. According to Mellerio et al., patients demonstrate more side branches of the central sulcus and more often a PBS (Mellerio et al., 2014). The PBS is characterized by a precentral branch pointing between the central sulcus and one of its ascending branches (see fig.1.4). Hence, in this case, we worked on a region gathering both the central and the precentral sulcus (in the BrainVISA nomenclature, the precentral sulcus is divided into five entities: the median, marginal, superior, intermediate and inferior precentral sulci). We computed a mask for each of these sulci, merged them and cropped the distance maps, resulting in images of dimensions 78 x 86 x 99 in 1mm resolution.

12.6.3 . Transposition of the method

For these two ROIs, we transposed the method in the same way. Specifically, we trained our $\beta - VAE$ on the HCP dataset to model the inter-individual variability. We used the same data split as before for the central region. Although we trained the model on new regions, we used the hyperparameters obtained with the gridsearch on the central sulcus region for training. Choosing these parameters may lead to sub-optimal performances but enables us to have a first validation of our methodology. Analyses of the latent and the folding spaces are performed following the method described above for the central sulcus.

13 - Results

13.1 . Training results

Each training lasted for approximately 1 hour on an Nvidia Quadro RTX5000 GPU. We obtained with our gridsearch $\beta = 2$ and L = 75.

13.2 . Assessment on synthetic known anomalies

13.2.1 . On the latent space

UMAP latent space visualizations for the four deletion benchmarks are presented in Fig. 13.1. For the benchmark 200, benchmark data are rather homogeneously distributed among control data, suggesting that simple surfaces of sizes between 200 and 500 voxels are too subtle to be encoded differently. Indeed, as shown in Fig. 12.3, small, simple surfaces can correspond to tiny branches that display a high variability in the population. Therefore these synthetic anomalies may be included in the normal variability. The distribution of benchmark 500 seems to be not completely similar to the control's, but the restricted number of subjects makes it hard to conclude. However, the trend becomes more pronounced for benchmarks 700 and 1000 where fake anomalies are gradually gathered and their distributions are different from the controls. These results are confirmed by the ROC curves (Fig.13.1). Even when using all the latent dimensions, classification results are very poor for benchmark 200 (AUC = 0.51), supporting that the deleted branches may be too melted into the inter-individual variability. Classification performances are also very low for benchmark 500 (AUC = 0.70). They start to be slightly better for benchmark 700 (AUC = 0.81) but are very good only for benchmark 1000 (AUC = 0.96).

For the asymmetry benchmark, UMAP visualization demonstrates a good separation between the right and the left hemisphere (Fig.13.2A), which is verified by the classification of the whole latent space (AUC=0.82). These results suggest that specific shape features are encoded among other properties in the latent space.

To better understand the asymmetry characteristics encoded by the model, we leveraged the generative power of our $\beta - VAE$. Fig. 13.2B. and C. show the average patterns for the right (green) and the left hemisphere (blue) as encoded by our model. The hand knob of the right central sulcus seems to be slightly higher and shallower than in the left hemisphere. Moreover, the double-knob configuration appears more prominent in the left hemisphere. To further highlight the main differences between the two hemispheres, we selected the most important dimensions for the classifier, here dimensions 9 and 36. In Fig. 13.3A., control and benchmark data are represented according to these two dimensions. Even if the

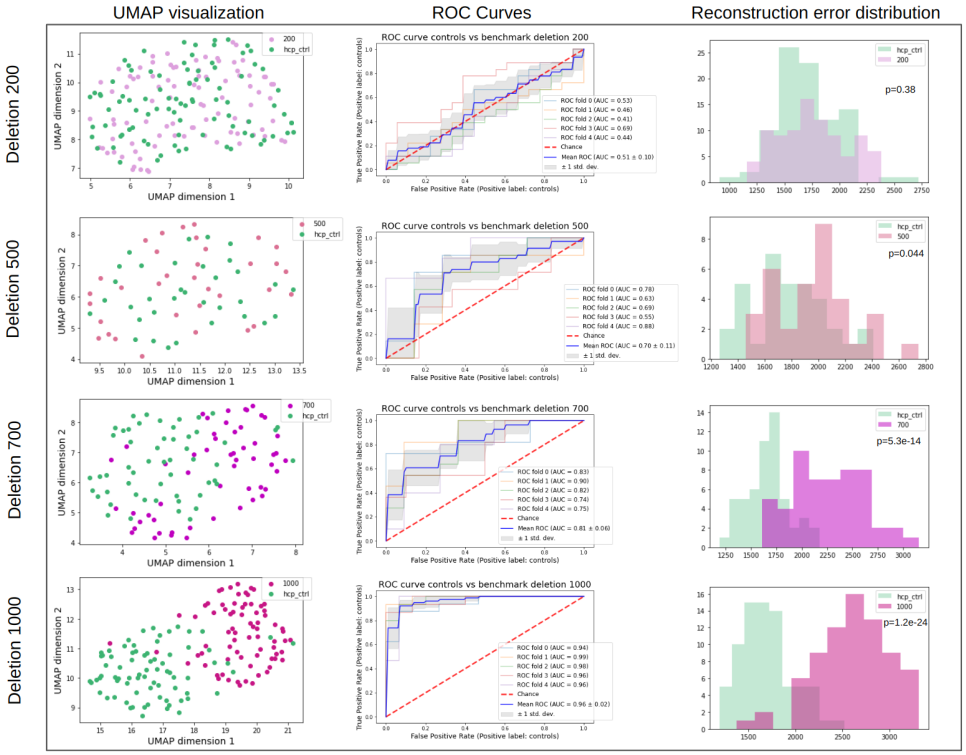
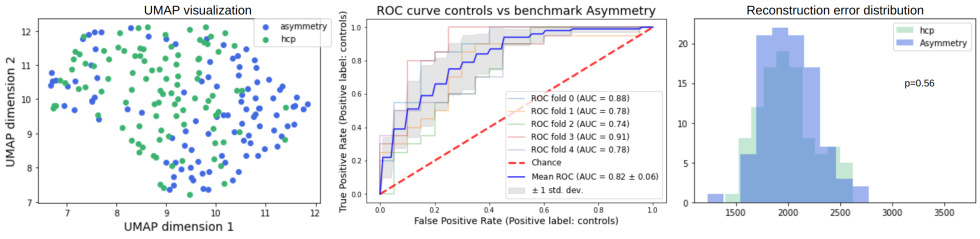


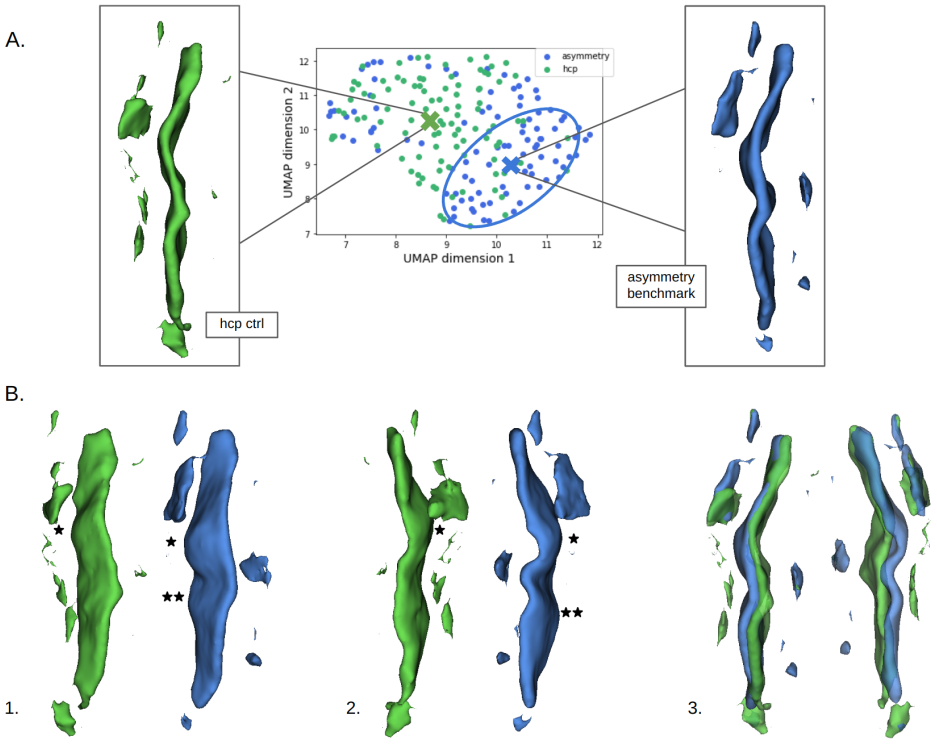
Figure 13.1: *Deletion benchmarks results.* For each row, controls are represented in green and benchmark data in pink. Left column: UMAP projection of benchmark and control data. Middle column: ROC curves of classification of control and benchmark data. Right column: reconstruction error distributions and p-value of the Kolmogorov-Smirnov test with the null hypothesis that the two samples come from the same distribution.

separation is not well marked, we can observe a trend represented by the arrow. We tried to understand the features encoded by the 9th dimension. We took the average for all 75 dimensions of the latent space, and we traveled from the minimum to the maximum of the 9th dimension and reconstructed the resulting vectors. Fig. 13.3B. 1, 2, and 3 represent the reconstructions. These interpolations confirm the trend observed previously. We observe a double-knob configuration in the left hemisphere. The view from underneath and the side view enable visualizing the pli de passage frontal moyen (PPFM). A pli de passage is a gyrus that connects two gyri and which is buried in the depth of some furrows (Mangin et al., 2019). Fig.5.1B.1. and 2. propose a visualization of a "pli de passage" located in the central sulcus, the PPFM. According to the different views from Fig.13.3B. 1, 2 and 3, it seems that the PPFM is smaller in the right hemisphere and located higher in the central sulcus.

13.2.2 . On the folding space



(a) UMAP projection of asymmetry benchmark and control data, ROC curves of classification of control and benchmark data, and reconstruction error distributions.



(b) A. Averages for the control subjects, i.e. right hemispheres (in green), and for the highlighted asymmetry subjects, i.e. left hemispheres (in blue). These averages are also placed on the UMAP dimensions. B. 1. and 2. Respectively side and bottom views of the averages of A. The single star indicates a single-knob configuration, and the two stars indicate the second knob of a double-knob configuration. B. 3. Superposition of the two averages respectively in upper and bottom view.

Figure 13.2: *Asymmetry benchmark results.* Controls are represented in green and benchmark data in blue.

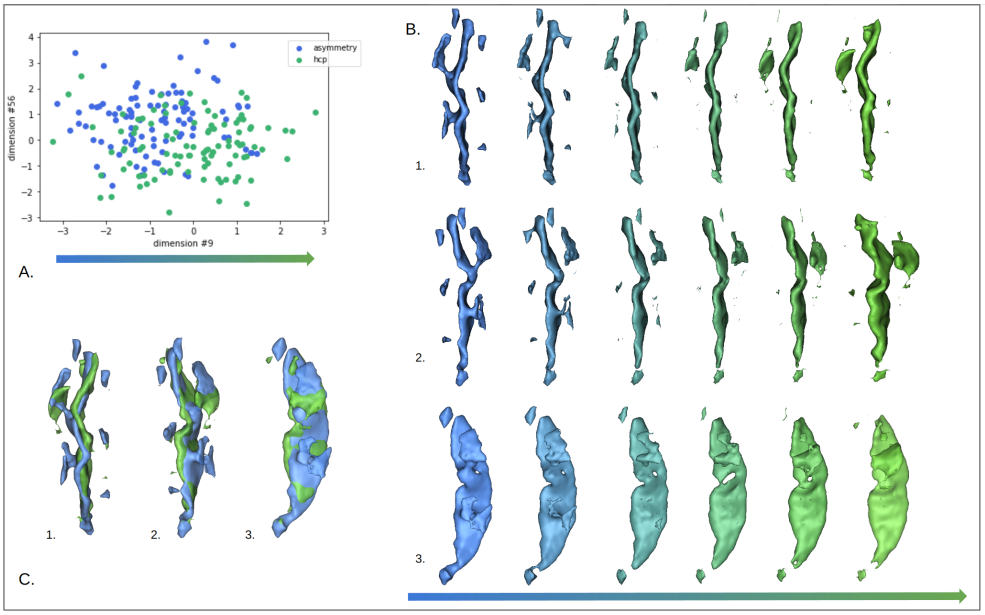


Figure 13.3: *Travelling through the 9^{th} dimension of the latent space.* A. Visualization of controls and asymmetry benchmark according to the most important features of the classifier. B. Interpolations along the 9^{th} dimension. 1, 2, and 3, respectively correspond to the upper, bottom and side view of these interpolations. C. Superposition of extreme interpolations.

We then investigated whether the folding space, i.e, reconstruction errors, was relevant for identifying outliers. For deletion benchmarks, we observe a similar trend as in the latent space. For deletion 200, we cannot see a difference of distributions (p-value = 0.38). However, from deletion 500 we can see a stall with the deletion benchmarks having significantly higher reconstruction errors (p-values of 0.044, 5.3e-14 and 1.2e-24 for benchmarks 500, 700 and 1000 respectively) (Fig. 13.1). On the contrary, for the asymmetry benchmark, there is no significant difference, nor a trend, in the reconstruction error distributions (Fig. 13.2).

13.3 . Application on the case of interrupted central sulcus

13.3.1 . On the latent space

The UMAP projection from the latent space is shown in Fig. 13.4A. On this distribution, we can observe that most interrupted central sulci are at the margin of the point cloud except for one. Thus, it appears that the representation learned by our model enables to project rare patterns at the margin of the population. Interestingly, when we look at the pattern of each one of the interrupted sulci, it seems that a specific pattern, the "T-shape" pattern (Mangin et al., 2019) is specifically located on one side of the representation. Fig. 13.4B. shows the assessment of the marginality of the interrupted sulci based on an OCSVM and

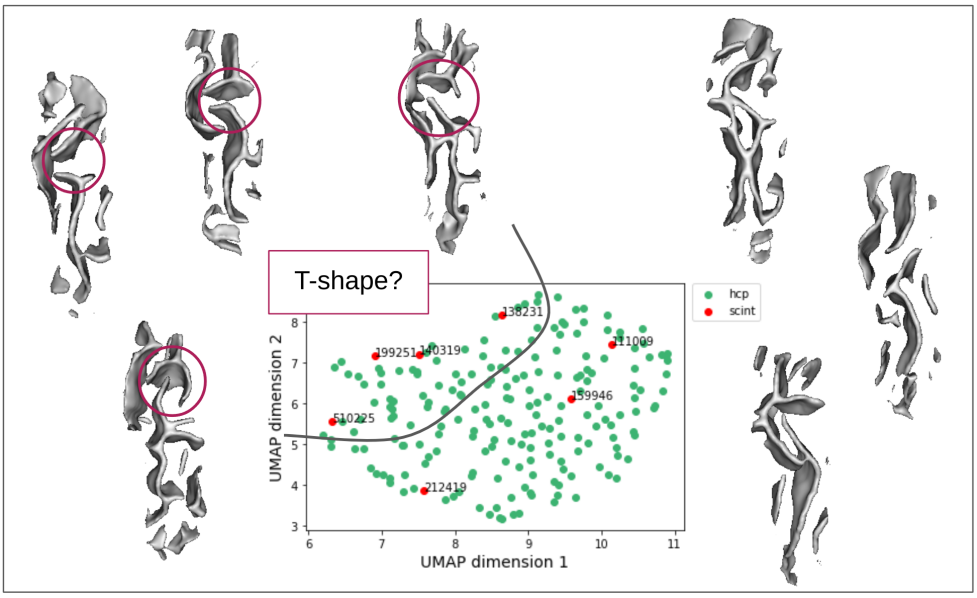
isolation forest. Error margins correspond to various UMAP projections, suggesting that the ability to detect interrupted central sulci (CS) in the UMAP space is very dependent on the UMAP projection. Interrupted CS detection is within the confidence interval, but the curves are close to the superior bound suggesting a tendency. However, interrupted CS positions in the UMAP space are not enough to detect them: detecting 5 interrupted CS out of 7 would lead to more than 40% of false positives. Nevertheless, some other patterns considered as controls and detected as outliers might also be rare.

Fig. 13.5 presents the controls' patterns most often predicted as outliers by the OCSVM. First, we note that the outliers are logically located at the border of the distribution. Moreover, we observe distinct patterns in different regions of the UMAP space. Analyzing the corresponding crops' meshes, we first observe that the subjects share some of the characteristics of the control population (like the presence of a knob for instance). We visually highlighted the subjects of the four regions where we note similarities within the groups. Group B seems to demonstrate a very wide open knob. In addition, the knobs are well defined by the upper and the bottom part of the sulcus. On the contrary, the sulci of group C appear to have larger knobs than usual but they show more continuity with the upper and the bottom parts. The pattern of group D seems to correspond to a rather flat central sulcus with a close, long and continuous postcentral sulcus. The shape characteristics of A are less obvious but the sulci give the impression of having several small knobs, two or even three in the two bottom cases and a small part of the precentral inferior opposite to an upper part of the postcentral sulci. Fig. 13.6 provides a better understanding of these features. For each pattern, we go from the centroid to one of the subjects in each group by interpolating and generating samples. Fig. 13.6A. presents the interpolations from the centroid to the several-knobs pattern. We gradually see the upper part of the hand knob curving and becoming more pronounced until forming a first knob at the top of the sulcus. Another knob in the bottom part appears similarly. Likewise, patterns B, C and D vary progressively until they match the centroid's shape.

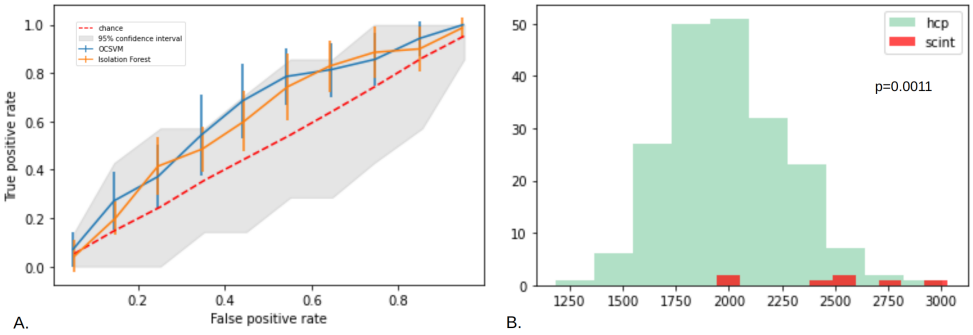
13.3.2 . On the folding space

When analyzing the detection power on interrupted CS in the folding space, we first note that the reconstruction errors' distributions seem to be different between HCP controls and interrupted CS (p-value = 0.0011). This result suggests that our model has more difficulties to reconstruct the input and that reconstruction error could constitute a relevant metric to detect rare or abnormal patterns. However, having only seven subjects strongly limits our conclusions and this should be replicated with more data.

Observing the reconstructions and the residual maps of Fig. 13.7 gives clues into the way our model has encoded the interrupted CS. First, we can note that the reconstruction quality is quite good visually. The model's omissions appear to be quite noisy (blue small fold pieces). The arrow points out an omission



(a) Interrupted central sulci shape distribution in the UMAP space. The 3D folding patterns of the subjects are positioned according to their location in the UMAP space. For instance, the pattern located in the lower left corner corresponds to subject 510225 in the UMAP representation. Subjects with interrupted sulci on the upper left of the UMAP visualization seem to correspond to an interruption with a T-shape pattern.



(b) A. Outlier detection performances using OCSVM and isolation forest on the interrupted CS. B. Controls and interrupted CS reconstruction error distributions.

Figure 13.4: *Interrupted central sulci on UMAP space.* Controls are represented in green and interrupted CS in red.

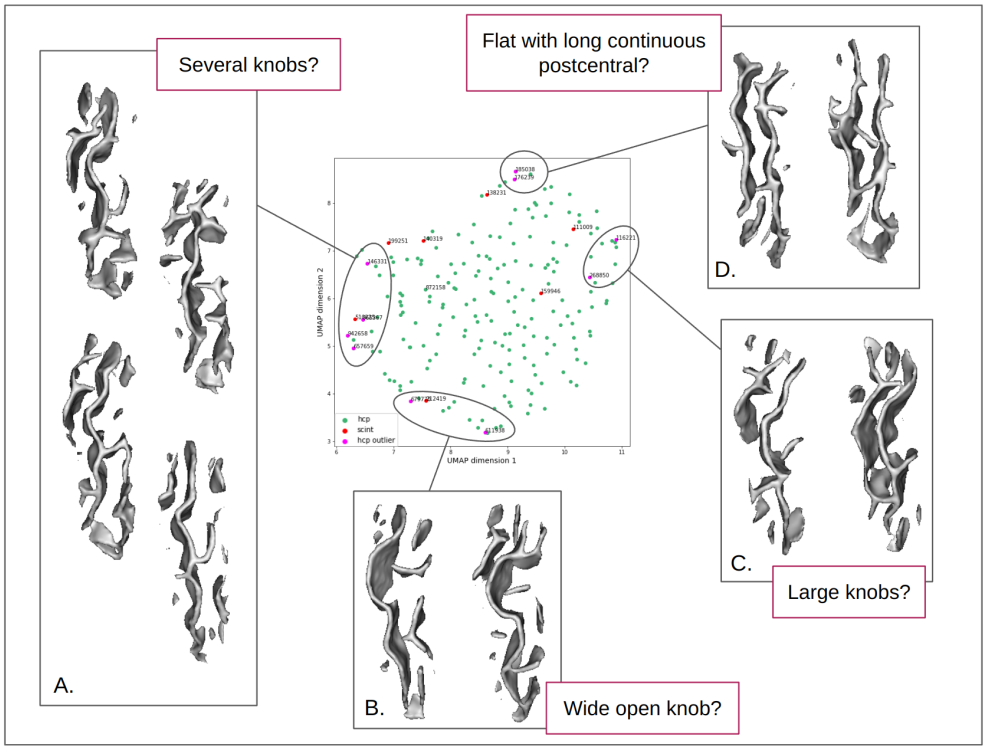


Figure 13.5: *Control subjects identified as outliers.* A, B, C and D correspond to groups of visually similar patterns. The UMAP projection is the same as the one in Fig.13.4. Control subjects identified as outliers are in pink and subjects with interrupted central sulci are still represented in red.

beyond the noise which corresponds to a perpendicular branch pointing toward the frontal cortex. Such a pattern might be an atypical configuration. It is interesting to note that in six out of seven cases, the model transformed interrupted sulci into continuous patterns. This is highlighted by the "output-input" visualizations. Unlike the omissions, the model additions are rather localized. Moreover, the asterisks show where the model has filled the interrupted sulci. Such visualization could be useful to identify rare patterns like interruptions or perpendicular branches.

13.4 . Application to corpus callosum dysgenesis

13.4.1 . On the latent space

We first compare distributions of CCD children (n=7) with control children (n=7) acquired in the same conditions and with HCP adult subjects (n=200). UMAP projections, presented in Fig.13.8A., give different results depending on the hemisphere. For the right hemisphere, it seems that most children controls are included in the distribution of adult controls (hcp_test in green). Five out of the seven subjects having a CCD are located at the margin of the controls, suggesting

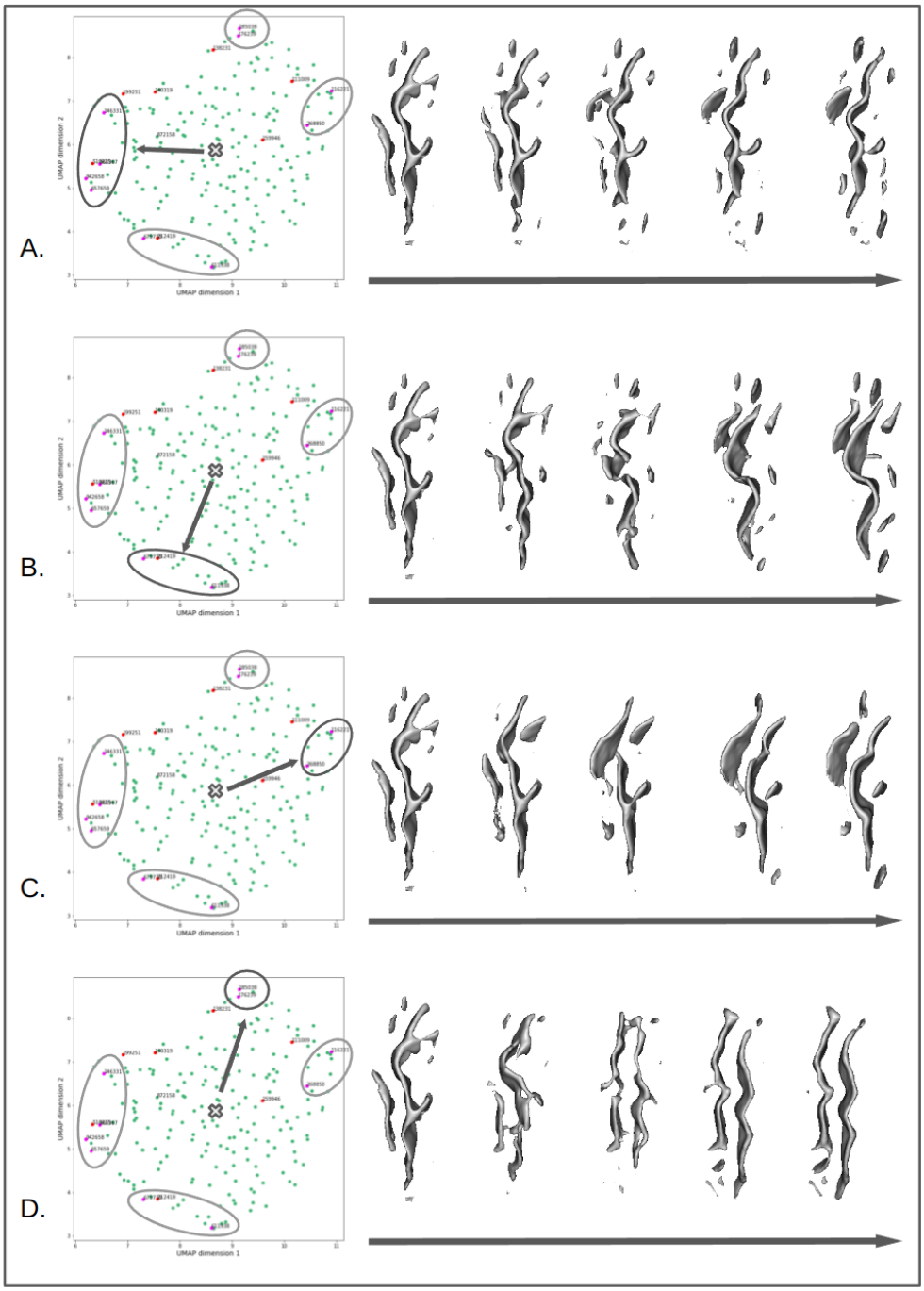


Figure 13.6: *Travelling through the latent space from the centroid to the margin of the UMAP space.* The centroid is the centroid of HCP controls. Then, for each row, interpolations between the centroid and one of the patterns of each group are computed and then reconstructed.

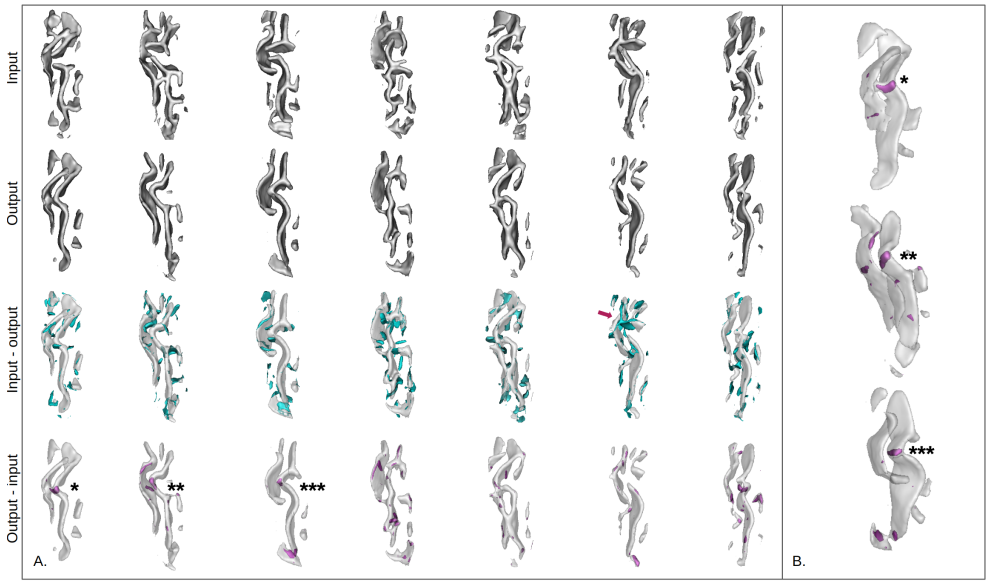


Figure 13.7: *Reconstructions and residuals for all seven interrupted sulci.* A. For all rows, distance maps are converted to meshes for an easier visualization. First row: input data. Second row: reconstruction of the model. Third row: Reconstruction of the model with the difference between the input and the output, i.e. the model's omissions (in blue). The purple arrow highlights an omission corresponding to a perpendicular branch pointing toward the frontal cortex. Last row: Reconstruction of the model with the difference between the output and the input, i.e. the model's additions (in purple). B. Rotated view of the reconstructions represented with asterisks in the last row of A.

that their latent representation differs from the average cingulate sulcus pattern. However, two subjects, one with a complete and one with a partial agenesis, are in the middle of the controls. In the left hemisphere, only three control children are clearly in the control adult distribution. The other four are closer to the CCD subjects but they seem to be still distinct. Indeed, CCD subjects are gathered very close to each other. This could be due to the fact that there may be an age effect between children's and adults' brains or a site effect (different scanners, resolution), which we tried to reduce by using skeleton-based images but which may still remain. Nevertheless, we can still observe a difference in distribution between control children and CCD subjects.

13.4.2 . On the folding space

Regarding reconstruction error distributions (Fig.13.8B.), we observe for both hemispheres that control children seem to have the same distribution as adult controls, which is confirmed by Fig.13.8C. (p-value=0.034 and 0.017 respectively for right and left hemisphere). On the contrary, CCD subjects present higher reconstruction errors that are significantly different from both HCP controls (p-value=3.6e-06 for the two hemispheres) and children controls (p-value=0.0011

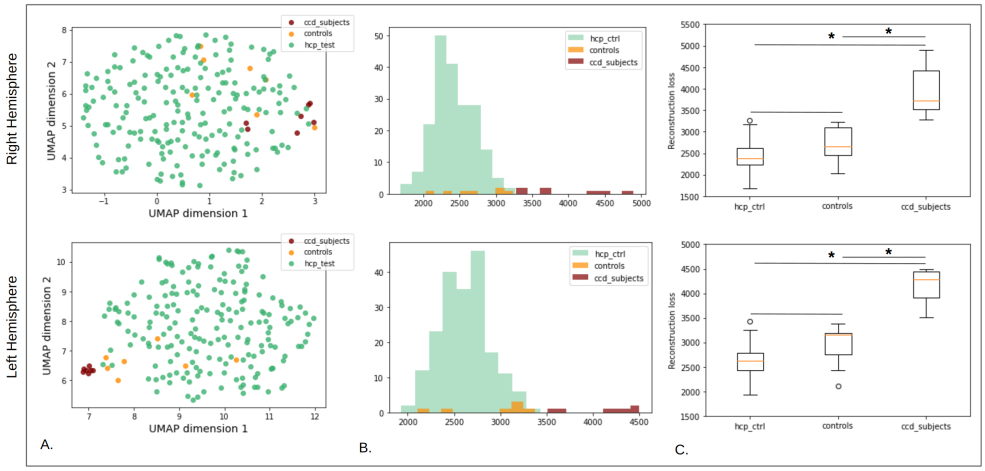


Figure 13.8: *Results on corpus callosum dysgenesis (CCD) subjects.* First row: right hemisphere. Bottom row: left hemisphere. For both rows: A. UMAP projections of CCD subjects, control children and HCP test. B. Reconstruction error distributions for the CCD subjects, control children and HCP test. C. Reconstruction error variations for the CCD subjects, control children and HCP test. Significant differences between populations according to the Mann-Whitney test are indicated with an asterisk.

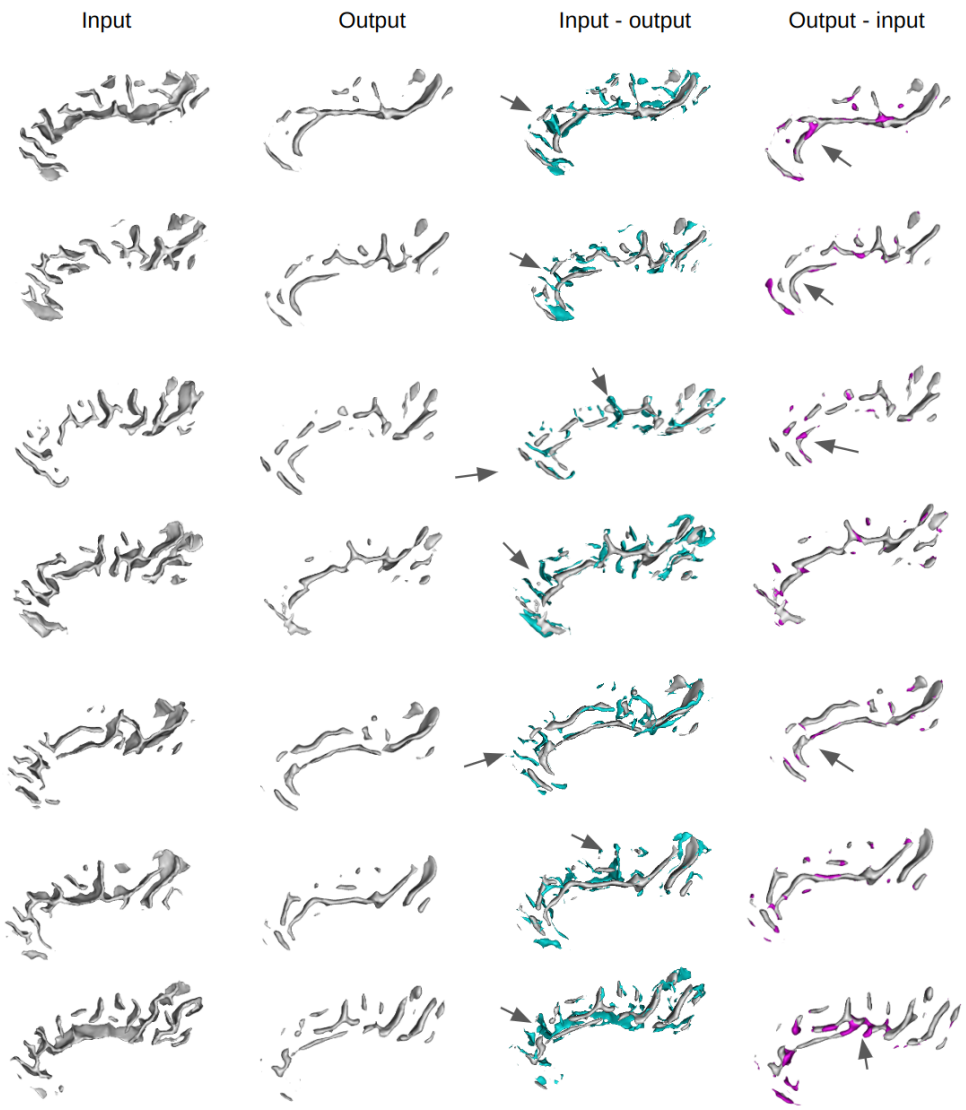
for the two hemispheres). Therefore, it seems that there is a complete individual separability of the CCD patients which is very promising and should be replicated with more data.

The reconstructions presented in Fig.13.9 highlight the singularities of CCD. The model's additions mostly make the cingulate more continuous than initially. The model's omissions are mainly small branches perpendicular to the cingulate sulcus that are radially oriented.

13.5 . Application to patients suffering from FCD2

13.5.1 . On the latent space

Fig.13.10A. presents the latent representation projected to the UMAP space of patients with positive MRI and lesion in the right or left hemisphere (respectively +/right and +/left), patients with a negative MRI and lesion in the right or left hemisphere (respectively -/right and -/left), controls acquired in the same conditions and HCP subjects. We first notice that similar to what we observed with the CCD subjects, the controls do not have the same distribution as HCP controls: they are more located at the margin and seem to represent a transition towards the patients. For patients with positive MRI, those with lesions in the left hemisphere (represented with crosses) do not appear to be distributed like controls. This suggests that they may still have characteristics differing from the controls. 5 out of the 8 patients with positive MRI and right lesion are at the complete margin of the point cloud and beyond the controls of the same database. -/right subjects



(a) CCD subjects.

(b) One control subject from the same cohort.

Figure 13.9: *Right cingulate sulcus reconstructions and residuals for the CCD subjects and one control.* Each row corresponds to a subject. For all columns, distance maps are converted to meshes for easier visualization. First column: input data. Second column: reconstruction of the model. Third column: Reconstructions of the model with the difference between the input and the output, i.e. the model's omissions. Last column: Reconstructions of the model with the difference between the output and the input, i.e. the model's additions. The arrows highlight interesting features added or missed by the model.

seem to be equally projected at the margin or in the HCP distribution. Last, with the exception of one subject, the projection of -/left patients is quite similar to that of controls.

To try to decipher the pattern characteristics of the two groups of patients with the lesion in the right hemisphere, we generated the average patterns. The insets of Fig.13.11 show the three average patterns for the patients having a positive MRI (indigo), the patients with a negative MRI (blue) and the associated controls (orange). The control pattern is composed of the central sulcus (on the left side) and of the precentral sulcus (on the right side). Like the average pattern based on HCP test subjects, the central sulcus of these control subjects is composed of a knob. The precentral sulcus seems rather continuous. The pattern of patients with a positive MRI presents a central sulcus with a wider knob. In addition, we can observe a branch that links the central sulcus to the precentral. Last, the pattern of patients with a negative MRI (blue) shows a central sulcus with a quite small knob. Unlike the control precentral which is quite flat, the precentral of both groups of patients seems curvier, especially for the patients with a negative MRI.

For both groups, we then travelled along the axis from the control centroid to each patient group centroid and beyond (as represented by the arrows). Samples between the insets are linear interpolations which are then reconstructed. The furthest samples are likely to be unrealistic as they were sampled from less represented areas. Going from the control centroid towards the patients with positive MRI, we see the hand knob progressively widening and its bottom part merging with the precentral sulcus. Interestingly, the average pattern of the +/right subjects may even suggest an intermediate PBS pattern. Interpolations beyond are more complex to analyse and seem to present merged central and precentral sulci.

From the control centroid to the patients with a negative MRI, the central sulcus seems to flatten and present a very slight knob in the average pattern. The precentral slightly curves and shows more branches. Beyond the average of -/right patients, the interpolations display unlikely configurations with a CS presenting several interruptions and a long precentral with several branches.

13.5.2 . On the folding space

Regarding the reconstruction error, contrary to CCD subjects, all the samples from the cohort are included in the range of HCP subjects (Fig.13.10B.). This could be due to the lower resolution of most subjects. However, the distribution of both groups of patients seems to be different from that of the controls (Fig.13.10B. and C.). In Fig.13.10B., all patients with negative or positive MRI were considered, whether their lesion is located in the left or in the right hemisphere. Fig.13.10C. presents the details for each group. Both figures suggest that patients with negative MRI have higher reconstruction errors and thus are more difficult to reconstruct. This is confirmed by the p-values reported in table 13.1.

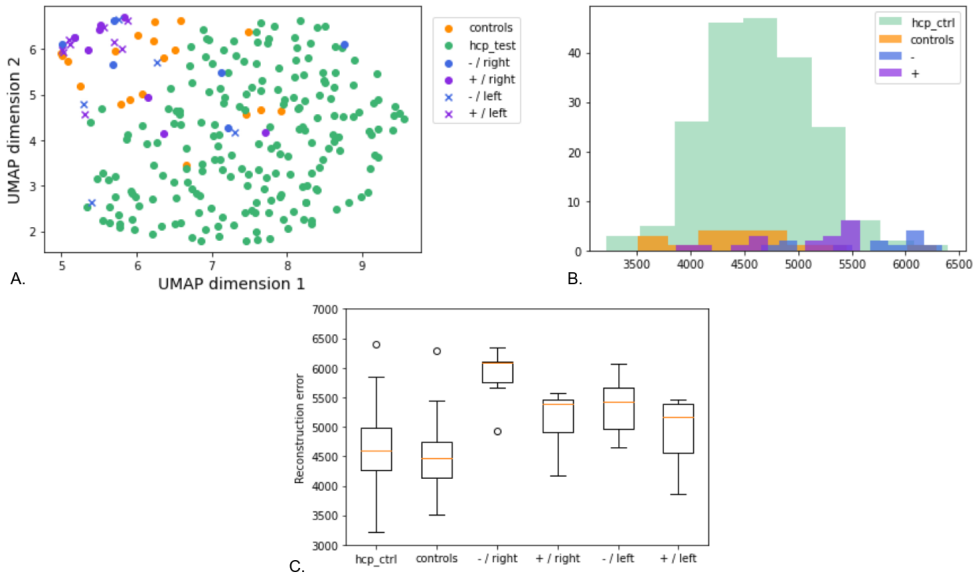


Figure 13.10: *Results on patients suffering from FCD2.* A. UMAP projection of patients with a positive MRI (indigo), patients with a negative MRI (blue), controls of the same dataset (orange) and HCP test subjects (green). The crosses correspond to subjects with a lesion in the left hemisphere. B. Reconstruction error distributions for the four groups (patients with negative or positive MRI, controls and HCP subjects). C. Reconstruction error variations for the different groups. Patients of each group have been separated depending on the lesion's location.

We do not observe a difference in the distributions between HCP subjects and controls (adjusted p-value=0.41). -/right patients are the only group with a distribution significantly different from that of the controls (adjusted p-value=0.0020). For +/right patients, it seems that their reconstruction errors tend to have a different distribution but it is not significant. It is interesting to note that although patients with a lesion located in the left hemisphere present fewer differences with the controls than patients with a lesion in the right hemisphere, they are still not similar to controls. A potential explanation is that the events perturbing the neurodevelopment and resulting in a lesion in the left hemisphere may have had other consequences, less marked, in other areas of the brain. Nevertheless, as we deal with only a small number of patients this should be replicated with more data in order to conclude.

Fig.13.12 presents the reconstructions for the controls and patients with positive and negative MRI (and a lesion located in the right hemisphere). Contrary to the previous results there is no row for the model's addition (output-input) because no error was important enough to appear.

First, we can notice for the controls that despite being simplified versions of the inputs, the reconstructions are quite good. It seems that the main omissions

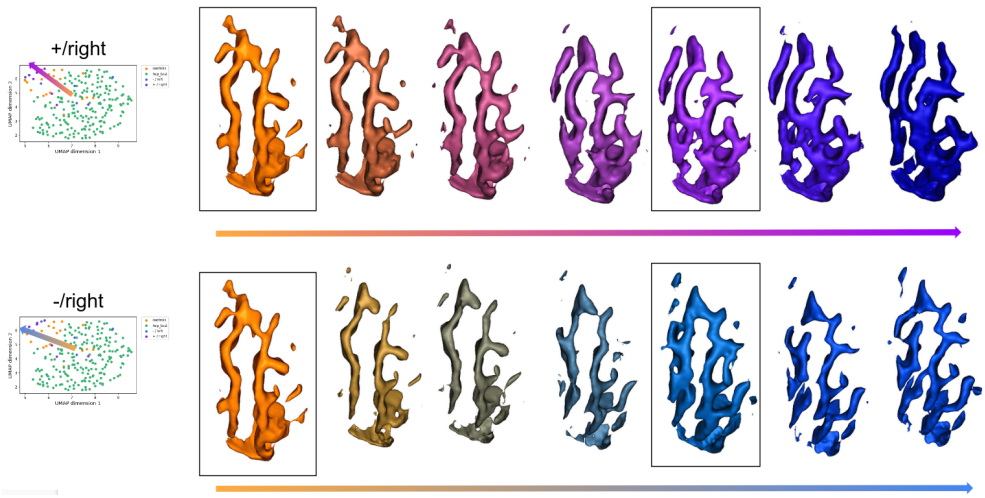
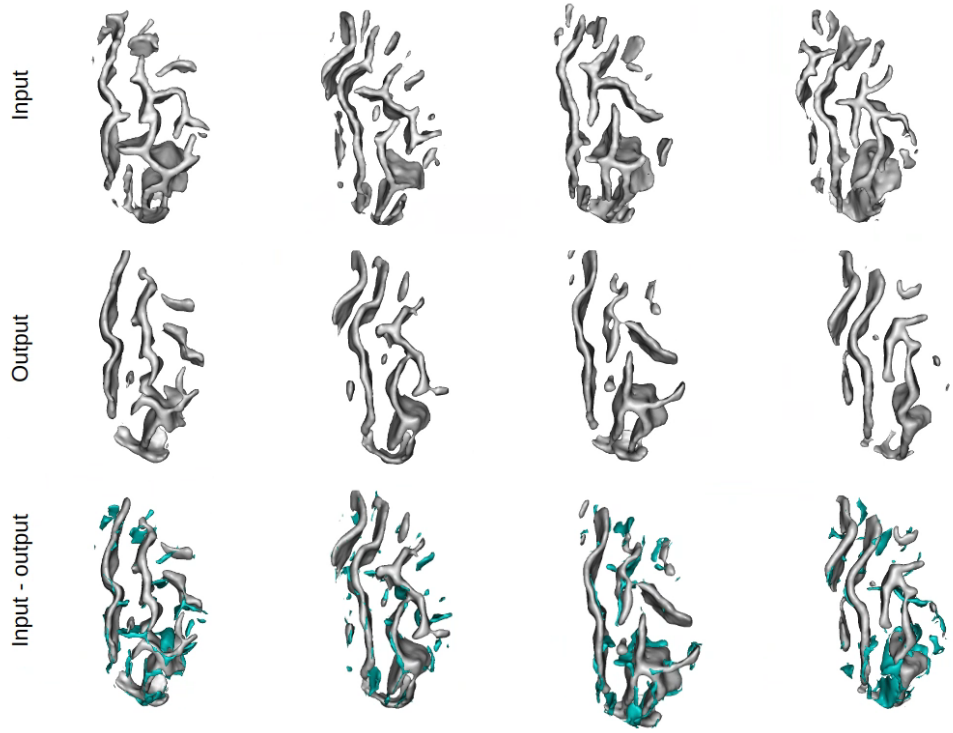


Figure 13.11: *Travelling through the latent space from the control centroid to the centroid of each group of patients and beyond.* The insets are the centroid-generated patterns. We travel through the latent space from the control centroid to each patient group centroid and beyond as illustrated by the arrows. The intermediate patterns between the insets correspond to interpolations along the arrows. First row: From controls to +/right patients. Second row: From controls to -/right patients. Only patients with lesions located in the right hemisphere were used to generate the averages. Colours match Fig.13.10: controls are represented in orange, patients with positive MRI in indigo and patients with negative MRI in blue.

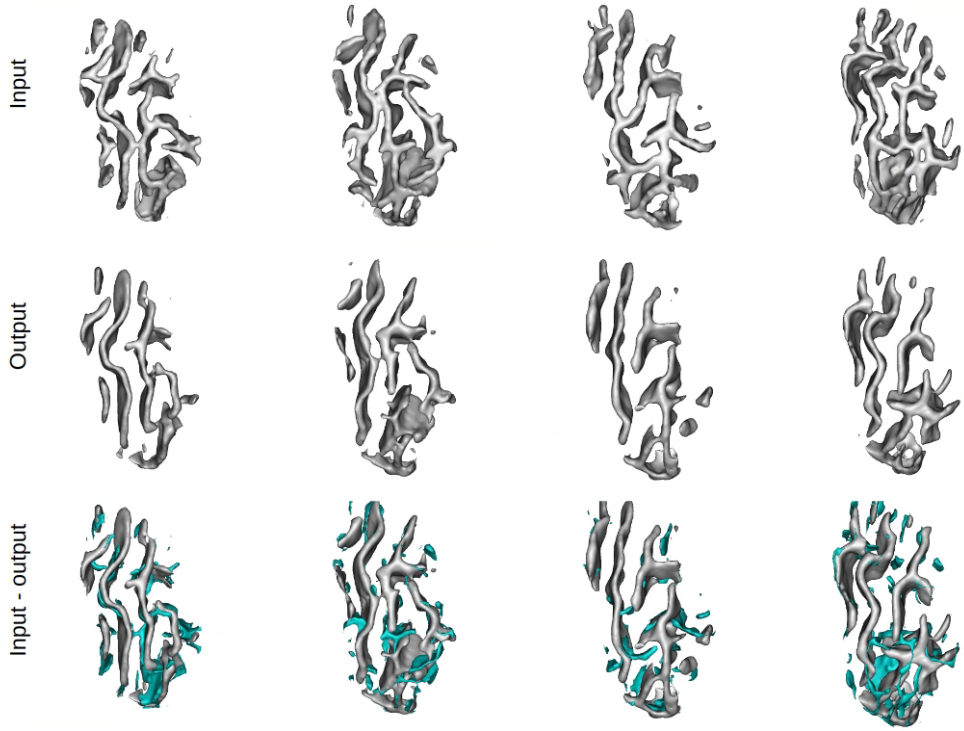
by the model are small branches of the central and precentral sulci.

Fig.13.12 (b) shows the reconstructions of patients with a positive MRI. Similar to the controls, the model's omissions include small branches. However, one important feature is also the connection between the central and the precentral sulci which is partly or entirely missing in all examples. Therefore, a connection between these two sulci may be a rare feature. This could be confirmed by Fig.13.7A. where a perpendicular branch from the CS to the precentral was missing in the reconstruction.

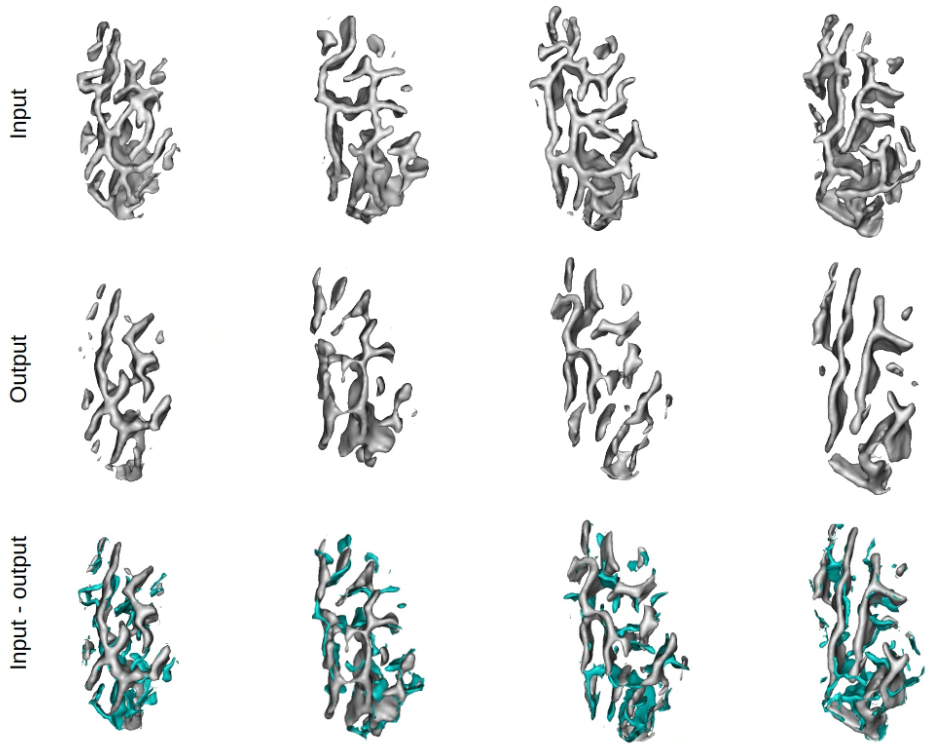
The reconstructions of patients with a negative MRI are shown in Fig. 13.12 (c). First, it visually confirms the results of Fig.13.10: the reconstruction quality is lower. As with patients with positive MRI a connection between the central and the precentral sulci is missing in some cases. In the second column, we can notice that the subject presents a PBS that the model cannot correctly reconstruct.



(a) Control subjects.



(b) Patients with a positive MRI.



(c) Patients with a negative MRI.

Figure 13.12: *Right central region reconstructions and residuals for the patients suffering from FCD2 and controls.* Each column corresponds to a subject. For all rows, distance maps are converted to meshes for easier visualization. First row: input data. Second row: reconstruction of the model. Third row: Reconstructions of the model with the difference between the input and the output, i.e. the model's omissions. Note that all patients represented have the lesion located in the right hemisphere.

Groups	p-value	Adjusted p-value
HCP vs controls	0.083	0.41
controls vs +/right	*0.014*	*0.069*
controls vs -/right	**0.00040**	**0.0020**
controls vs +/left	0.025	0.12
controls vs -/left	*0.0095*	*0.047*

Table 13.1: Significativity of Mann-Whitney U test. The null hypothesis is that two samples have the same distribution. Adjusted p-values were corrected with the Bonferroni method.

14 - Discussion

This part proposed a methodology to study rare folding patterns which was applied to the central sulcus region and to a described rare pattern, interrupted central sulci. Specifically, we represented folding patterns with distance maps and leveraged the generative power of the $\beta - VAE$ to have a better understanding of the learned representations. In addition, we proposed a way to study the granularity of deviations that can be identified and we brought to light several rare patterns in the region. We also compared the identification power of both the latent space and the folding space. Finally, we assessed the generalization of our methodology on developmental anomalies in two other regions.

14.1 . Latent space and folding space, two complementary information

In many anomaly detection works applied to medical images, the detection is performed based on the reconstruction error rather than in the latent space (Schlegl et al., 2019; Baur et al., 2020; Tschuchnig and Gadermayr, 2021; Behrendt et al., 2022). However, both of these spaces have their interest and could bring complementary information. In our work, we studied four types of rare patterns, two synthetic types, deletion and asymmetry benchmarks, and two actual rare patterns. These four categories differ from control data by their own characteristics and thus help to study the granularity detected, that is to say, the typology of rare features that can be identified. For instance, the asymmetry benchmark includes more double-knob configurations. Depending on the size of the deleted simple surface, deletion benchmarks represent different features: benchmarks 200 and 500 represent mainly a missing branch with increasing size, which may represent the normal variability of branches. Benchmark 700 could look like an interrupted sulcus in some cases or in others, like benchmark 1000, an unlikely configuration. Interrupted central sulci present a clear interruption and a rare arrangement of the shapes forming the central sulcus. Last, CCD subjects demonstrate a missing sulcus or missing sulcal parts and branches with different orientations. For patients suffering from FCD2, although PBS has been described, it is not present in all patients and is not specific, so in this case, we do not know whether there is a rare pattern and if so, what the pattern would be.

These different kinds of known deviations from the norm provide clues to the characteristics of rare patterns that can be identified respectively in the latent space or in the folding space of our model. As a matter of fact, the identification performances in the latent and in the folding space vary depending on the kind of patterns. For deletion benchmarks, the folding space, based on the reconstruction error, seems to enable the identification of unusual patterns from smaller modifi-

cations: different distributions are observed from 500 deleted voxels. Whereas in the latent space, the detection requires at least 1000 deleted voxels. Likewise, for the interrupted central sulci, despite the small number of samples, their detection seems to be easier on the basis of reconstruction error than in the latent space. Similar results were obtained on CCD subjects even if the latent representation was encouraging. In return, the error distribution of the asymmetry benchmark is not different from that of the controls, but the benchmark is well detected in the latent space. Therefore, the latent space could be more sensitive to shape arrangements than the folding space. The lack of difference in the error distributions may be due to the fact that the voxel-to-voxel differences between the right and left central sulci are local and subtle and could be embedded in the normal variability. In addition, the reconstruction error is for the entire image. Therefore, in the case of small and very local deviations from the norm, the reconstruction error alone is likely to be insufficient. A way to limit such effects could be to use a more local error, applied to sub-regions or patches for instance.

The difference in the outlier detection performance may also lie in the way our model encodes the outliers. Based on our results, we can consider several cases. First, a rare configuration is represented by several samples present in the training set. This would be the case with the asymmetry benchmark. Indeed, there are more double-knob configurations in the left hemisphere but single and double-knob patterns coexist on both sides. In such a case, the distribution support of the left and right hemispheres are the same, but the densities differ, which could lead to a projection of the outlier at the margin of the latent space but to a good reconstruction. Second, the rare configuration is almost never represented in the training set and the model has not detected and thus encoded its local specificity. Then, the subject would be encoded with a "default" representation and projected in the middle of the other subjects. This would be consistent with the results we described in a previous article (Guillon et al., 2021), where major anomalies (different parts of the brain from the one considered in the train set) were projected in the middle of the point cloud and reconstructed as the average reconstruction. It could be the case of the benchmark deletion 500 and the interrupted central sulcus projected in the point cloud. Last, the outlier configuration is almost never represented in the training set but the model has detected the rare characteristic. The subject is then projected at the margin of the point cloud and the decoder has not learned this part of the latent space leading to a poor reconstruction (interrupted central sulci, CCD subjects).

Nevertheless, in all cases, a strength of the folding space is the possibility to localize the reconstruction errors and, in some cases, the unusual features. If not too noisy, reconstruction errors can be very informative. For example, in the case of interrupted central sulci, looking at the model's addition permits clearly localizing what is atypical in a subject (Fig.13.7). Similarly, in the case of the CCD subjects, the reconstruction errors highlight the presence of radial small branches that are

typical of this brain disorder (Bénézit et al., 2015). The reconstructions of patients suffering from FCD2 may suggest that a connection between the central and the precentral sulci could be a rare pattern. But some noise remains, and it might be interesting to add an additional constraint to represent only errors that correspond to a minimum number of contiguous voxels. This could lead to a good explanation of the abnormality which is of major importance in the field, especially when applied to medical images. Other explanation methods exist, directly on the network such as Grad-CAM (Selvaraju et al., 2020) or on an OC-SVM applied on the learned features (Sohn et al., 2022) for instance; but the use of the reconstruction error is immediate and easy to implement. Hence, the latent space and the folding space, based on reconstruction error, can provide complementary information and both can be used to identify rare patterns.

14.2 . Data size limitations and unknown number of rare patterns

The method should be further qualified because of the low number of our examples of rare patterns. While the study of a known rare pattern is interesting and important, having only seven samples severely limits our conclusions. Similarly, the poor results of benchmarks 500 and 700 in the latent space could be due to their small size, and having larger benchmark datasets could lead to increased performances.

Also, we assessed our method in the CS area on the benchmarks and on an existing rare pattern, but because few rare patterns have been described in this region, there may be other rare configurations in what we consider the control population. For instance, three morphologic variants in the central sulcus region have been introduced, representing 2.9%, 7.0% and 1.8% of the studied population, opposed to 78.2% of "omega" shape, i.e. the central sulcus knob and 10.1% of "epsilon" shape which corresponds to the double-knob configuration (Caulo et al., 2007). This multiplication of rare patterns in the populations would make the identification of interrupted central sulci more difficult.

14.3 . Relevance of synthetic benchmarks

Moreover, we can wonder about the relevance of our synthetic benchmarks. Although synthetic rare patterns are of high interest as they enable to quantify the performances on different degrees of deviations from the norm, few works have been interested in them to our knowledge (Meissen et al., 2022). But the use of fake deviations raises the question: do they constitute adequate rare patterns? Few studies introduced rare folding patterns based on the arrangement of their shapes such as the PBS (Mellerio et al., 2014), an interrupted central sulcus (Mangin et al., 2019) or a flat central sulcus (Sun et al., 2017). Here, we emphasize their

advantage in the study of our understanding of the brain: they are evidence of neurodevelopmental processes and then stable throughout life. But other abnormal sulcal features have been studied and found to be important and correlated with neurodevelopmental disorders, such as the depth, which demonstrated anomalies in autism spectrum disorder (Nordahl et al., 2007; Dierker et al., 2015) or Williams syndrome (Essen et al., 2006) for instance. Despite being another subject of study, a benchmark corresponding to central sulcus depth variations could be interesting to assess whether our framework can be extended to detect such anomalies.

Regarding the current benchmarks we use, we said that small erased SS could remain undetected as this deletion could be embedded in the normal variability. However, there may be several categories of deletion deviations. Some may be minor, as a small SS representing a tiny branch. On the opposite, some small SS, for instance one corresponding to depth change, representing the presence of a pli de passage, and thus leading to an interrupted central sulcus would be expected to be a major feature of the topology. Hence, our criterion, only based on the size of the SS may be insufficient and it could be interesting to add another one, such as topological criteria.

In any case, having an unusual feature (e.g., a missing simple surface or unusual depth) that can be incrementally increased, or comparing several types of features, helps characterize the detection power of a model and the features likely to be detected.

14.4 . Learning relevant representations

When dealing with sulcal patterns and their high complexity, it may be easier to use representations of the folding which attempt to gather several subjects with similar patterns. Local averages of sulci, also called moving averages, enable to concentrate on the main features of the different patterns and are thus very useful to analyze folding patterns (Sun et al., 2012; de Vareilles et al., 2022; Foubet et al., 2022). From a graph-based representation of the sulci, the identification of patterns can be done after computing similarity and applying a clustering (Meng et al., 2018). Our approach proposes another method to learn sulcal representations. From our cropped distance maps, the $\beta-VAE$ learns a mapping to a latent representation which can then be reconstructed. Therefore, rather than explicitly computing pairwise similarity between the subjects, gathering them, and then analyzing the patterns, we hope that our $\beta-VAE$ directly learns shapes that can be combined and arranged in patterns. The representations learned by our model seem to be relevant and consistent with some morphological characteristics of the central sulcus area.
First, the reconstruction of the average representation of the right central sulcus is composed of an upper knob whereas the left average tends more towards a double-knob configuration (respectively green and blue sulci in Fig.13.2B). This is

one of the main known asymmetries in terms of patterns and it appears early in the development. It has been detected in infants of 30 weeks postmenstrual age (de Vareilles et al., 2022) and in adults (Sun et al., 2012).
We also observed differences in terms of curvature of the hand-knob, with a hand-knob more pronounced in the left hemisphere than in the right (Fig.13.2B.). Considering that we study a right-handed population, this could be related to handedness. With the lateralization of the hand motricity, we expect the motor area of the right hand in the left hemisphere and particularly the precentral gyrus to be more developed for right-handed subjects, pushing backward the upper part of the central sulcus which would result in a knob more pronounced. This interpretation is consistent with a study on one-handed subjects that showed that subjects born without a hand had a flatter central sulcus contralateral to the missing hand (Sun et al., 2017).

Another interesting property that was successfully encoded is the PPFM. This pli de passage was first described in 1888 (Broca and Pozzi, 1888) and has been a source of growing interest due to its link with the motor hand area (Boling and Olivier, 2004) and in the context of understanding the formation of the knob regarding evolutionary questions (Hopkins et al., 2014). Our model was able to encode the PPFM in the latent space as well as its asymmetry characteristics. Indeed, we observed that the PPFM is smaller in the right hemisphere which corresponds to central sulcus depth variations described in (Amunts et al., 1996). This is also consistent as the PPFM has been correlated to the hand. Therefore, right-handed subjects tend to have a more developed hand area in the left hemisphere and thus a larger PPFM.
Hence, our latent space has learned relevant *normal* characteristics that are consistent with the region's morphology. It has also enabled to propose four other groups of likely rare patterns (Fig.13.5). The pattern representing a rather flat central sulcus is indeed a non-typical configuration. Less than 2% of the studied subjects were reported to have such a configuration in (Caulo et al., 2007). Moreover, flat central sulci appeared as the most important feature when comparing controls to congenital one-handed subjects who tended to demonstrate flatter central sulci (Sun et al., 2017), confirming that flat central sulci are less frequent patterns. The groups representing large knobs and wide open knobs (Fig.13.5B. and C.) are also an atypical configuration that is present at one extremity of the axis representing the most extreme variations in Human and is closer to configurations we observe in Chimpanzees (Foubet et al., 2022).

14.5 . Generative power of $\beta - VAE$ and comparison with other strategies

Since our proposed framework is able to encode relevant features regarding folding patterns, the generative power of the $\beta - VAE$ can be exploited. Indeed,

reconstructions and interpolations are tools to understand the folding variability. We have just mentioned that the learned patterns were relevant and consistent with those obtained by other methods, but our method has the advantage of being able to reconstruct and interpolate. For instance, interpolations along the main axis of asymmetry variations highlight the evolution from a right to a left hemisphere. It can also be useful to understand the folding process and in particular the formation of interrupted central sulci. As a matter of fact, on Fig.13.6C., an interruption of the central sulcus happens when interpolating from the central subject to one control outlier. When observing the PPFM, we can see that the PPFM increases until reaching the surface of the brain and thus interrupting the central sulcus. Jointly, the inferior sulcal part connects to the precentral sulcus. Such observations may provide additional clues in our understanding of the folding processes.

But other deep learning models could be interesting to study folding patterns. For instance, $\beta - VAE$ reconstructions are known to be blurry contrary to GAN's or more recently, diffusion models. Currently, this shortcoming is limited as we seek to have a simpler representation of folding patterns, still, for more subtle details, another model may be better suited. In addition, in the anomaly detection field, models that add constraints on controls distribution are quite appealing. For instance, deep One-Class Classification and its derivatives have been proposed to push control data into the smallest hypersphere in the latent space (Ruff et al., 2018). This could help increase the detection performance in the latent space. Nevertheless, no matter the architecture or the framework, an important limit to understanding what our model has really encoded is the high number of latent dimensions.

One can also wonder about the representation of the folding patterns. Previously, we mentioned two main strategies: clustering and manifold. Usually, these two approaches are applied to a continuous space. Nevertheless, if we consider sulcal shapes as symbolic entities that can be combined and arranged, we could represent folding patterns based on a discrete space rather than a continuous one. As such, VQ-VAE (van den Oord et al., 2017) seems to be an interesting representation to compare with our present results.

Finally, this framework of outlier detection based on training on control subjects alone may be sensitive to outliers present in the training set. Having a contaminated dataset could severely limit the detection performances, at least in the folding space which is based on the reconstruction error. It has been reported in a brain tumor detection problem that having 3% of outliers in the training set (about 1000 samples) leads to a decrease of 5% of the AUROC and to a 13% decrease if the contamination reaches 12% of the training set (Behrendt et al., 2022). Therefore, one serious shortcoming of our paradigm is that we do not know the outliers we are looking for. Applying our framework to a control population alone in order to bring out rare patterns may limit the different patterns that can be identified. A way to tackle this issue and to increase the patterns detected would

be to exploit the presence of outliers in the training set as proposed in (Qiu et al., 2022). In their technique, the authors introduce an iterative joint training where they assign labels (anomalous or control) to the examples, and then optimize the network's parameters to better identify the anomalies. Such a method could also enable to project the outliers more at the margin of the latent space. The impact of the presence of outliers during training on the latent space has not yet been investigated to our knowledge. If, as we suggested before, outliers present in the training phase are encoded at the margin of the distribution, i.e. in a different area of the latent space, it could be interesting to deepen our analysis, based on clustering for instance.

14.6 . Generalization of the approach: towards an analysis of the whole brain?

This work has shown that our approach had successfully encoded some relevant features of the folding patterns in the central sulcus region but it is attractive to think about the behaviour and results we could obtain in other parts of the brain. Here, we assessed the generalizability of the framework on two other datasets and regions. Our results suggest that our method can well transpose in other brain regions. Specifically, even if we use the hyperparameters (β and L) optimized for another area, the learned representations still enable us to distinguish between control and outlier subjects. This is all the more interesting that in the case of the CCD subjects, the two studied regions are rather different. The central sulcus is one of the first folds to form and is rather stable, contrary to the cingulate region that is more variable (Sun et al., 2009). Therefore, it seems that no matter the folding variability of the zone, our framework can be applied. This encouraging result raises a question regarding the procedure to adopt to extend our analysis to the whole brain. A way could be to define a set of regions, consistent with the cytoarchitecture and function and to train our $\beta - VAE$ on each region. In particular, some areas seem to be interesting from a clinical point of view (Provost et al., 2003; Yücel et al., 2003; Gervais et al., 2004; Borst et al., 2014; Hotier et al., 2017). Our future works may thus focus on proposing an adequate methodology to tackle the whole brain.

On the other hand, when we applied this framework to CCD subjects and patients suffering from FCD2, we also operated a domain shift. Indeed, for the CCD subjects, the dataset to explore included exclusively children while the $\beta - VAE$ was trained on young adults. Despite folding patterns being reported as trait features (Cachia et al., 2016), such an age variation may have an impact. In addition, beyond dealing with children, the site and the scanner are different. Similarly, in the case of patients with FCD2, the subjects were not homogeneous in terms of age and scanners. Such differences have been reported to affect the generalizability and the performances on various targeted tasks. In terms of distributions in the

latent space, despite the fact that the distribution of the controls does not seem to completely overlap the distribution of the HCP controls, the patients still seem to present a different distribution than both controls' populations. Moreover, the domain shift does not seem to have an effect on the folding space where controls reconstruction errors are not significantly different from HCP contrary to CCD subjects that have significantly higher reconstruction errors. The results of patients suffering from FCD2 are more complex to interpret as the consequence, in terms of folding remains unclear. Still, the reconstruction errors of the controls are similar to HCP subjects and patients seem to have higher reconstruction errors. It is particularly interesting for -/right patients as their MRI does not present visual signs of FCD2. However, having only fewer than a dozen of subjects makes it difficult to conclude on the importance of these age and site effects for our task. We will explore these questions in further studies.

14.7 . Perspectives

This part proposed a framework to identify rare and abnormal folding patterns based on the modelling of inter-individual variability. With a new representation of folding patterns, we proposed a model that was able to encode relevant folding characteristics. The use of synthetic rare patterns enlightened the identification power of our model on both the latent space and the folding space. Finally, we successfully generalized our approach to other clinical brain disorders in other regions. Our results open up several avenues of work such as the definition of new synthetic benchmarks that match the characteristics of other known anomalies, the use of other deep learning models that exploit the presence of outliers in the training set, or the use of our framework to better understand the folding process.

Appendices

A - Visualization of an example of input normalization

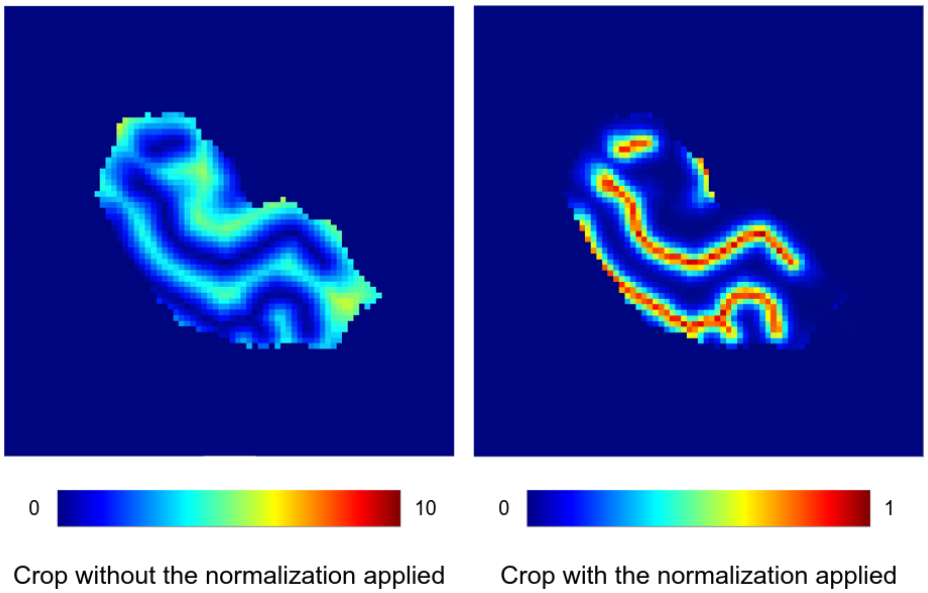


Figure A.1: *Crops without (left) and with (right) the normalization applied.* Example of an input crop without and with the normalization applied and presented in section 12.3.1

B - beta-VAE hyperparameter selection

We seek to choose the best values for the Kullback-Leibler divergence weight (beta) and the number of dimensions of the latent space (L) in the task of identifying interrupted central sulci. To select the best values we used two criteria, the reconstruction quality and the detection power on a proxy for interrupted central sulci. The precentral and postcentral sulci have some similarities with the central sulcus regarding orientation, size and shape. In addition, they are usually more interrupted and have more ramifications which make them a good proxy in the task of interrupted central sulci detection. However, due to the interruptions and the ramifications, the classification between a central sulcus and a pre or postcentral sulcus may be very easy for some subjects. In order to make the classification task non trivial and thus informative, we selected the subjects for whom the task is harder, i.e. that have a pre or a postcentral sulcus that looks like the central sulcus. Therefore, our hyperparameter selection method is based on three steps:

- generating the pre and post central sulci crops
- identifying the precentral and postcentral sulci that present some ambiguities with the central sulcus,
- the actual gridsearch.

These three steps are detailed in the following:

1) Generation of the pre and post central sulci crops. To obtain crops of these two other sulci, like presented in the section 5.2. for the central sulcus, we generated distance maps of the whole hemisphere based on the folding graph of the HCP subjects. Then, like for the central sulcus, we learned masks of the precentral and of the postcentral sulci. In order to have the same image dimensions as the ones for the central sulcus, we adapted the mask to the adequate dimensions by adding or deleting a few voxels. Therefore, we obtain 3D crops of the precentral sulcus and of the postcentral sulcus.

2) "Ambiguous" precentral and postcentral sulci identification. The aim of this second step is to identify "ambiguous" precentral and postcentral sulci. In order to do this, we trained a beta-VAE (L = 150 and beta = 1) on the train set described in section 12.3.2 (only control central sulcus crops). Once the model was trained, we encoded all the training central sulci and all the precentral and postcentral sulci. We then trained two linear SVM to classify the latent codes between the central and precentral sulci, and between central and postcentral sulci. All precentral and postcentral sulci that were wrongly predicted as central sulci were considered as ambiguous. We repeated these steps five times in order to increase the robustness. We found 28 ambiguous precentral sulci and 18 ambiguous

postcentral sulci.

3) Gridsearch. Once we identified ambiguous subjects, we consider them as a proxy for interrupted central sulci, i.e. as the rare patterns to detect. We then performed the gridsearch to find the best hyperparameter configuration to detect them. We trained a beta-VAE on the train set described in 12.3.2 (only control central sulcus crops) for each configuration. Next we encoded the central sulci of the validation set and the ambiguous pre and postcentral sulci. We trained a linear SVM to classify between the latent codes of the validation samples and of the outliers which are composed of the pre and post central sulci.

Part V

General discussion and conclusion

In the previous parts, we have presented our approach which consists in learning the normal folding variability thanks to an unsupervised deep learning model. Based on this representation, we characterized typical patterns of the cingulate region and identified rare or abnormal cortical folding patterns in the central region. This part proposes a summary before discussing several aspects of this thesis.

General discussion and conclusion

Contributions

In this manuscript, we have discussed cortical folding patterns, their formation and their relation with brain functional architecture. In particular, the folding processes can be subject to perturbations which may lead, from a morphological point of view to rare and abnormal patterns that can translate to a wide range of disorders from no symptoms to heavy ones. Moreover, contrary to many markers, folding patterns have the advantage of being *trait features*, i.e., to remain stable throughout life. Therefore, they give a unique insight into the neuro-developmental processes and constitute an opportunity to study the neuro-development and potentially lead to the discovery of biomarkers.

In the meantime, the study of folding patterns has taken a turn with the emergence of large neuroimaging databases. However, to be able to analyse such databases requires specific methods. Unsupervised deep learning constitutes an opportunity to address two challenges: it enables to efficiently study large databases and to bring out representations without any supervision, i.e. without having to know in advance what the typical patterns are.

We presented here a work in three main steps. First, we learned a representation of the folding inter-individual variability. Based on this representation, we proposed two applications: (1) the characterization of typical folding patterns and (2) the identification of rare or abnormal folding patterns. To this end, we used a pre-processing that allows us to focus on the folding information and to work with local patterns. We then trained a $\beta - VAE$ with our particular type of data to learn the folding variability of a control population.

In the first application, we worked on a specific region, the cingulate area which is of clinical interest. We then compared the learned representations of the $\beta - VAE$ and of another unsupervised model, SimCLR. Specifically, in both cases, we performed clustering on the latent space which led to four main patterns, some of which have already been described in the literature. The classification of a specific pattern, the paracingulate, on the basis of the latent representations ensured that the latent codes contain relevant folding information. Therefore, this application showed that our framework could be useful to study folding patterns and could lead to the identification of typical patterns.

The second application aimed to identify rare and abnormal folding patterns. We started by focusing on the central sulcus region where a rare configuration was described. We trained our $\beta - VAE$ to learn a representation of the central

sulcus area. To better qualify the ability to identify rare patterns we generated several benchmarks, each with its own type of deviations. We showed that both the latent space and the folding space, i.e., based on the reconstruction errors, provided information and seemed to be more useful depending on the types of deviations. The learned representations were relevant and we found some characteristics described in the literature. We also found another asymmetry feature that has not yet been reported. In addition, our framework enabled to identify rare patterns within the population. Finally, we evaluated the generalization abilities with two other datasets of patients. In both cases, although it seems that there is a database effect, we are still able to distinguish between patients and controls of the same database based on the folding space. This application has also highlighted the strength of using a generative model to better understand what the model has encoded and to locate the deviations.

Considerations on the size of the latent space

These two applications are complementary: once we have characterized typical folding patterns of a region, it is easier to identify potential deviations. When comparing the two applications, it is interesting to note that although the architecture of the $\beta - VAE$ is the same, the latent space size is very different: 4 for identifying typical patterns, as opposed to 75 for rare patterns. Of course, it has to be noted that the regions considered are different and more importantly, in the works presented here, the resolutions used are different: 2mm when identifying typical patterns versus 1mm when applied to rare patterns. However, in a previous work where we aimed to identify abnormal patterns in the superior temporal sulcus branches, we also worked in 2mm resolution with dimensions of the crop similar to the one we used to study typical patterns of the cingulate region (Guillon et al., 2021). In this case, a gridsearch with latent space sizes between 8 and 100 led to use 100 dimensions in the latent space, suggesting again that a larger number of dimensions in the latent space is better suited when looking for outliers. We should also remind that the gridsearch criteria are different. In the first application, we considered a latent space size leading to a clustered space, based on the silhouette score. Whereas for the second application, we used the latent space size that resulted in the best classification score on a proxy for abnormal patterns. Nonetheless, it could suggest that the number of latent dimensions may depend on the nature of the task and that the identification of typical patterns would require fewer latent dimensions than for outlier detection applied to folding patterns.

The comparison with other works is difficult since the input type is an important element. Indeed, the input dimensions may be closely related to the size of the latent space. Still, we can notice that large latent space sizes have been used in the anomaly detection field. When comparing different AE structures for brain anomaly

detection including the $\beta-VAE$, the latent space size was set at 128 (Baur et al., 2020). More recently, a work studying abnormal shapes of the pancreas with a $\beta-VAE$ compared the performances for different latent dimensions (16, 64, 256, 1024) and suggested that a larger latent space led to better outlier identification (Vétil et al., 2022). In return, when working on the characterization of typical patterns, smaller sizes were chosen: 2 for a $\beta-VAE$ used to characterize brain activity patterns in resting-state functional MRI (Gomez et al., 2022) or 8 in a $\beta-VAE$ used to perform generative modelling of mice brain (Liu et al., 2020).

An explanation for these different sizes could be directly linked to the task, i.e. either identifying typical or rare patterns, and to the role of the latent space. In particular, in an idealistic case, we wish that each dimension of the latent space corresponds to a folding feature. Therefore, when identifying typical patterns, having fewer dimensions would mean that fewer features are encoded and that the variability is expressed only based on certain characteristics. Having fewer features seems to be easier to then perform clustering and find patterns. Indeed, with more features, the variability expressed is increased and identifying what is common between subjects tends to be more complex. In return, in the scope of identifying rare or abnormal patterns, having more features leaves more "degrees of freedom" to encode the variability, which is important in order to encompass the whole range of normal variability. Nevertheless, having a larger number of latent dimensions raises new challenges such as the methods to identify deviations in high dimensionality and the interpretation is harder.

Modelling of cortical folding variability

We mentioned in the background that several methods to study and decipher the folding patterns have been proposed. Specifically, some works introduced manifold-based analyses that correspond to a continuous approach (Sun et al., 2012, 2017; de Vareilles et al., 2022; Foubet et al., 2022). Based on the dimensions of an Isomap, the continuous shape variations are studied along each dimension. This enables to identify tendencies and one or two varying features per dimension. Another approach introduced consists in clustering the subjects and then analyzing average patterns (Meng et al., 2018; Duan et al., 2019). In this work, we used two different approaches. On the one hand, we applied clustering to identify typical patterns and on the other hand, we used a continuous representation to characterize rare patterns. In the latter, clustering does not appear to be particularly well suited to identify outliers. Indeed, as a first hypothesis, we would seek to obtain a representation with only one group of subjects, solely composed of controls, where outliers are positioned at the margin of this representation or farther. Furthermore, we just mentioned that based on our work, it seems that bigger latent spaces are more effective for outlier detection but clustering may be a harder task with more dimensions.

For identifying typical patterns, performing clustering could be more relevant and is a subject of discussion (de Vareilles et al., 2022). Indeed, if we compute the averages of each cluster, either based on the latent representations which are then reconstructed or based on the local averages of the input space, clustering has the advantage of automatically proposing patterns (see Part III). However, in this case, the choice of the number of clusters is important and arbitrary constraints such as the one we used constitute a hard shortcoming. To address this issue we could consider models that perform clustering during training such as Deep cluster (Caron et al., 2018) or SwAV (Caron et al., 2021). In any case, to perform clustering does not necessarily imply that the space is not continuous. On the contrary, we may expect to see continuous variations going from one cluster to another. In particular, the encoded representation could correspond to a distribution where the clusters are regions of higher densities. In terms of patterns, this was shown when we travelled through the latent space (Fig.9.2). Hence, when identifying typical patterns, choosing clustering may not be contradictory with a continuous approach and could propose an easier way to characterize patterns, rather than individual shape characteristics. Therefore, presenting these two approaches as opposites may not be completely accurate. Indeed, as we noted, even when our latent space is optimized to be the most clustered possible, it still shows continuous variations among the clusters, taking advantage of both representations. Therefore, instead of choosing a paradigm between clustering and continuous approaches, another possibility could be to use a method combining both.

Representing cortical folding

The study of folding is complex and can be addressed with many approaches that rely on different objects of interest such as sulci, sulcal pits or gyri for instance. In this work we based our analysis on sulci that we modelled as skeletons or as their continuous counterparts, the distance maps. Even if the distance map is derived from the skeleton, the difference between these two types of data has consequences for the training of the $\beta - VAE$ and the analysis of the latent space: the skeleton is binary, we do not use the same reconstruction loss and the visualization in 3D requires additional steps for distance maps. The main advantages of the distance map are that there is a notion of proximity to a sulcus on the voxels and that we threshold the images to analyse the reconstructions or the generated patterns which enables to have less blurry images (see Appendix C).

Nevertheless, other modelling methods could be relevant and are the subject of several works. In particular, when it comes to modelling shapes, there are different possibilities for representing 3D outputs, including based on voxels like we currently do, points cloud or meshes. A quite recent proposition is based on implicit fields, where a shape is defined by sets of points inside and outside the shape. Such a framework has been applied to unsupervised anomaly detection

in brain images with an auto-decoder architecture (Naval Marimont and Tarroni, 2021) and to supervised and unsupervised classification of complex anatomical shapes (Juhl et al., 2021). Both these applications seem promising and could be a new way to model and study folding patterns.

On the other hand, taking a step back, it is interesting to think that this whole work investigates objects representing the shape of the void. Although we represent the sulci as 3D objects, they are constructions actually filled with cerebrospinal fluid. Nevertheless, these objects are proxies for various brain structures. In particular, the sulci represent the "walls" of the gyri and different patterns, like the presence of an additional sulcus, can correspond to a change in the cytoarchitectonic organization. As an example, in the cingulate area, the presence of a paracingulate sulcus leads to a different distribution of the cytoarchitectonic areas (Amiez et al., 2021). Moreover, around the sulci, there are diverse types of fibres. For instance, five U-shape bundles of the central sulcus have been identified (Pron et al., 2021). The shape of the sulci may also depend on the presence and the characteristics of *pli de passage*, which have been associated with specific U-shape fibres in the Superior Temporal Sulcus (Bodin et al., 2021). Therefore, the sulcal object as we consider it, is at the intersection of diverse modalities and enables to study the effects of multiple aspects at the same time. However, it could be attractive to relate our findings to the cellular organization of the cortex and to connectivity. For the latter, we could consider applying a similar framework to model the inter-individual variability of bundles of fibres around the sulci in order to identify typical or rare patterns.

Impact of the size of the dataset

A potentially important shortcoming of this work is the size of the dataset. For both applications, we used the HCP dataset to learn a representation of the folding variability which resulted in 551 subjects used to identify typical patterns and 640 subjects for rare patterns. Compared to other works, we are lucky to deal with these numbers of subjects but given the inter-individual variability of folding patterns having more subjects would enable to have a more accurate representation. I have not had the time to study the effect of the training set size but this was done in a work following up on what was presented in part III (Chavas et al., 2023). Specifically, they study the impact of the training set size (from n=20 to n=551 which is the configuration of part III) on the classification performances of the presence of a paracingulate sulcus. As we could expect, they demonstrate that better performances are obtained with greater sizes. It is worth noticing that it seems that the curve does not reach the ceiling yet. Recently, a training with 21072 subjects of the UK BioBank dataset demonstrated slightly better performances. Therefore, we would benefit from having more samples. In outlier detection works, the size of the dataset may even be more important. We mentioned the sampling

effect in part IV and the figure 2.2 was particularly expressive. With too little data, the risk is to consider control patterns as outliers due to their undersampling. Several dataset sizes (from 300 subjects to 1200) were compared in outlier detection of the pancreas' shape. The results suggested that in most cases, the more subjects, the better the performances (Vétil et al., 2022).

Last, the number of outlier samples is also capital. In this work, we had only a few subjects: 7 interrupted central sulci, 7 CCD patients and between 4 to 7 patients in the different groups of subjects suffering from FCD2. Having so few outliers severely hinders a reliable conclusion on our results. Therefore, in future works, having more subjects will be capital.

Dataset limitations

We just mentioned the size of the dataset but other considerations regarding our dataset should be acknowledged. First, for training, we used only one cohort, HCP. When our analyses are performed on the same cohort this has the advantage of not being confronted with the site effect. However, when we apply our framework to other cohorts such as the children with CCD or the patients suffering from FCD2 we observe that the distribution of the cohorts' controls is different from the HCP controls. In order to limit this effect, we are currently trying to incorporate controls from other databases in the training.

Moreover, the HCP dataset has some particularities since some subjects are related or even twins. Yet, folding patterns have been shown to be partly heritable (Im et al., 2011; Pizzagalli et al., 2020). In this work, we did not apply a specific methodology to deal with this specificity and we hypothesized that the folding proximity of twins could be embedded in the inter-individual variability. Nevertheless, to validate this hypothesis, we should replicate our work with another cohort without twins, such as UK BioBank.

Towards a new model

In our second application on the identification of rare cortical folding patterns, we used a $\beta - VAE$ which led to promising results. In future works, we wish to deepen another version of this model. Indeed, inspired by the impressive results of self-supervised learning and by some ideas of natural language processing, we started experiments where we added an objective to our $\beta - VAE$. Specifically, we introduced an inpainting objective where a random large simple surface is erased before feeding the input to the $\beta - VAE$. The output of the model is compared to the whole image that comprises the previously erased simple surface. Similar frameworks were used in anomaly detection for natural images (Pirnay and Chai, 2021; Zavrtanik et al., 2021). A comparable approach was introduced as Context-Encoding VAE (CeVAE) and applied to brain anomaly detection (Zimmerer et al.,

2018, 2019b). Such works seemed promising but currently, adding this objective does not lead to better results (see Appendix D). Future works will try to improve this ongoing framework.

Perspectives

This work is promising but calls for exploring many new directions. Indeed, beyond what we mentioned previously, future works could cover several aspects. The first follow-up of this work would be to systematize our method to identify rare folding patterns to the whole brain, covering different regions which are implied in various functions, and to other cohorts of patients. We hope that specific rare patterns could emerge from a cohort of patients.

Regarding the methodology, other unsupervised deep learning models would be interesting to try, such as those offering a clustering during training (Caron et al., 2018, 2021) to identify typical patterns for instance. Similarly, models introducing an objective in the latent space in favour of outlier detection (like deep SVDD (Ruff et al., 2018) for example) could be interesting to gather control samples closer together in the latent space. For both our applications, the VQ-VAE (van den Oord et al., 2017) seems also particularly interesting. The main difference with a classic $\beta - VAE$ is the learned representation which is discrete. If we consider cortical folding best represented by a manifold with continuous variations from one pattern to another, a continuous latent space seems better suited. However, if we consider the patterns as an alphabet, like in the research program presented in 2004 (Mangin et al., 2004b), a discrete latent space could be promising, similar to letters that can be arranged, each one of the components representing a particular shape which are then combined to form the patterns.

Another interesting line of research would be on the outliers detected. In part IV, we visually proposed four groups of outliers based on the latent distribution and tried to draw the characteristics of each group. We are then attracted to automate this approach by applying a clustering for instance. Specifically, in chapter 1.3, we mentioned that several events could occur during brain development and lead to folding anomalies. It would be thus interesting to try to link the groups of outliers with their potential origin.

Last, as we have seen, sulcal objects are particularly relevant to investigate since they are at the interface of several modalities. The study of the link between this work and connectivity, function and cytoarchitecture would be enlightening to increase our knowledge of the brain, its development and the potential disorders.

In conclusion, cortical folding patterns show a very high inter-individual variability that results from the neuro-developmental processes. The characterization of both typical and rare folding patterns can help to decipher the complexity of folding and to better understand brain formation. Unsupervised deep learning methods

appear to be an effective way to tackle this challenging task. This approach is a first step towards a more systemic and systematic framework that would benefit from integrating other modalities such as connectivity and cytoarchitecture.

Publication list

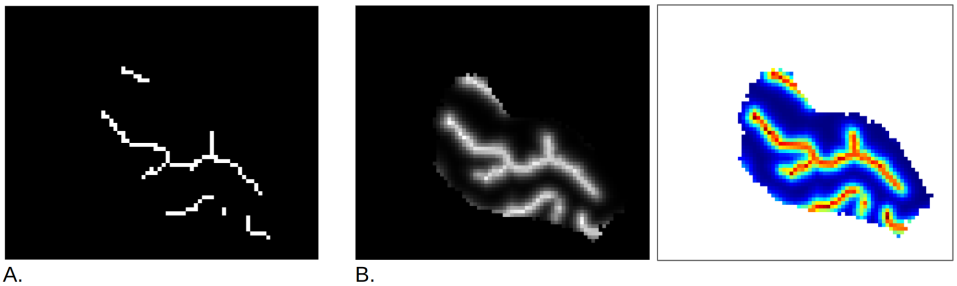
- Guillon, L., Chavas, J., Bénézit, A., Moutard M.-L., Rivière, D., Mangin, J.-F., Identification de motifs de plissement cortical rares dans le cadre de l'épilepsie. Poster à IABM (Colloque Français d'Intelligence Artificielle en Imagerie Biomédicale), 2023.
- Guillon, L., Chavas, J., Bénézit, A., Moutard M.-L., Rivière, D., Mangin, J.-F., Identification of Rare Cortical Folding Patterns using Unsupervised Deep Learning. Article submitted in Medical Image Analysis (MEDIA), 2022.
- Guillon, L., Chavas, J., Pascucci, M., Dufumier, B., Rivière, D., Mangin, J.-F., 2022. Unsupervised Representation Learning of Cingulate Cortical Folding Patterns, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2022, Lecture Notes in Computer Science. Springer Nature Switzerland, Cham, pp. 77–87.
- Guillon, L., Chavas, J., Rivière, D., Makin, T. Mangin, J.-F., Rare Folding Patterns Detection in the Central Sulcus Area with an Unsupervised Deep Learning Method, Poster in Organisation for Human Brain Mapping. Presented at the Organisation for Human Brain Mapping, 2022
- Guillon, L., Cagna, B., Dufumier, B., Chavas, J., Rivière, D., Mangin, J.-F., 2021. Detection of Abnormal Folding Patterns with Unsupervised Deep Generative Models, in: Machine Learning in Clinical Neuroimaging, Lecture Notes in Computer Science. Springer International Publishing, Cham, pp. 63–72.

Appendices

C - Comparison between skeleton and distance map

In this work, we studied cortical folding patterns based on two types of input. First, we used binary skeleton images (Fig. C.1 A.) to identify typical patterns that we then converted to distance maps (Fig. C.1 B.) to characterize rare folding patterns. We proposed distance maps in order to limit some shortcomings of the skeletons (see Chapter 5.2). To assess whether the distance maps are indeed more suited we compared the results obtained using skeletons and distance maps. Similar to Part IV we analyzed both the latent and the folding space on our deletion and asymmetry benchmarks. Note that the deletion benchmarks are slightly different from those presented in Part IV as the method to generate them was later improved (the latest version is used in Part IV).

In the latent space, the fake anomalies seem to be distributed more differently when using distance maps than skeletons for deletion 700 and 1000 (Fig. C.2). For smaller deleted SS, it seems that there is no difference. Regarding the ROC curves, they are slightly better with skeletons for small erased SS (deletion 200 and 500) but the results are equivalent for deletion 700 and a bit better with distance maps for deletion 1000. In the case of benchmark 200, we are not particularly interested in having the small deleted SS identified because they mainly concern the inter-subject variability of small branches instead of actual rare configurations. In addition, these SS generally correspond to the precentral or postcentral sulci, which are not our object of interest. This could be due to the fact that in distance maps, unlike skeletons, some isolated branches located at the edges of the crop



A. B.

Figure C.1: *Skeletons and distance maps.* A. Skeleton. B. Distance maps in levels of grey or in colours for easier visualization. Note that only a slice is represented for visualization but both the inputs are 3D volumes.

may be more related to the rest of the crop due to the information on the distance to the nearest sulcus held by all voxels, particularly thanks to the applied rotations. Note that for benchmark 500, like in Part IV, we have only 34 subjects in each class which limits reliable conclusions. In return, in the folding space, we observe greater differences between the distributions of the reconstruction errors of controls and benchmarks using distance maps than skeletons (Fig.C.4). For benchmark asymmetry, results are roughly similar using both inputs (Fig. C.5a). However, we observe that the averages generated based on the distance maps (Fig. C.5 b) are much more precise. On the contrary, those based on the skeletons are thicker which makes it difficult to analyze properly the differences. To conclude, distance maps seem to lead to equivalent or better results, particularly in the folding space. Furthermore, to visualize the distance maps in 3D a threshold has to be applied, which allows for better-quality images. Future works could improve the distance maps generation process. Indeed, the current method works but could be improved by generating the distance maps on the fly by convolving a Gaussian directly over the skeletons.

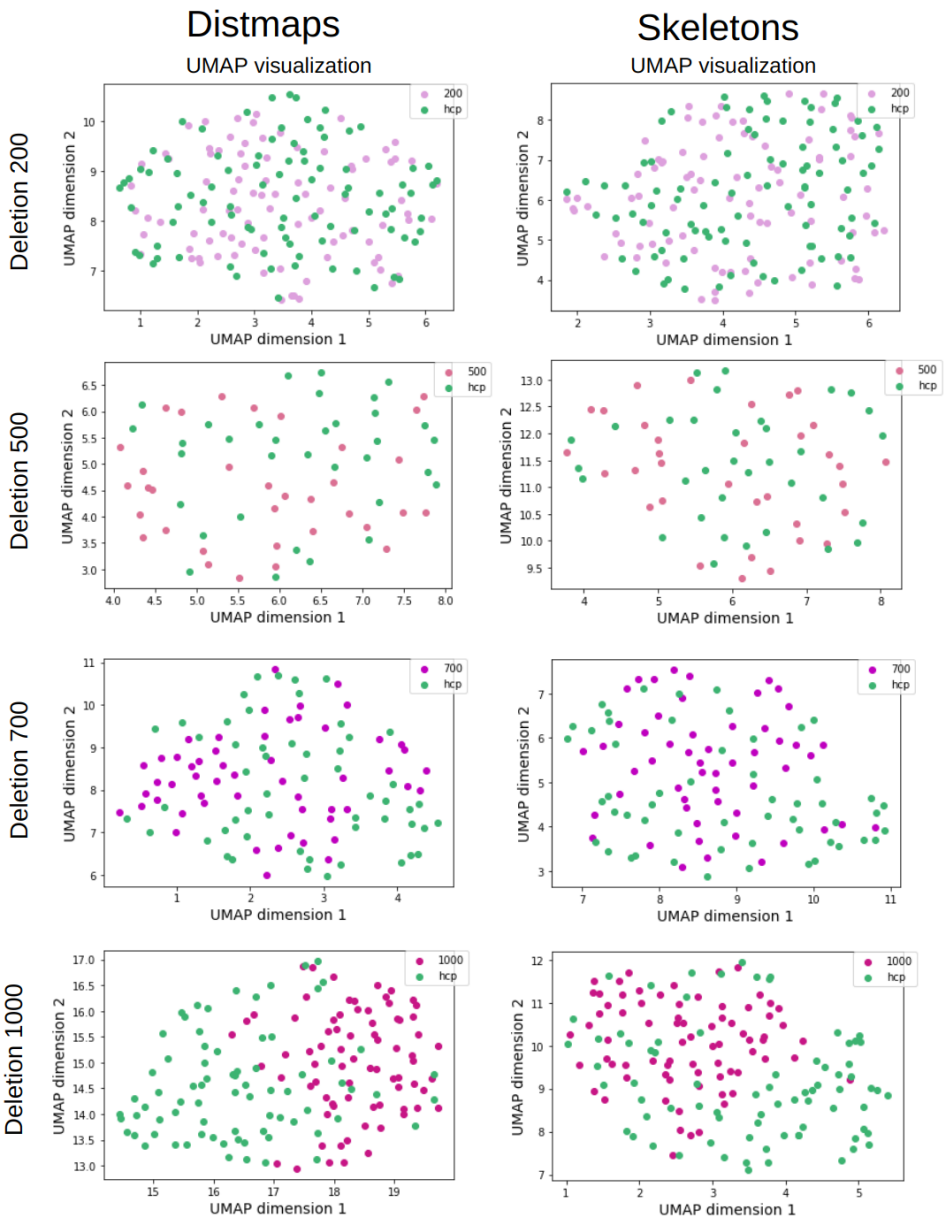


Figure C.2: *Comparison of UMAP projections of deletion benchmarks for skeletons and distance maps.* For both columns, each row corresponds to one benchmark. Left: UMAP projections obtained with distance maps as input. Right: UMAP projections obtained with skeletons as input.

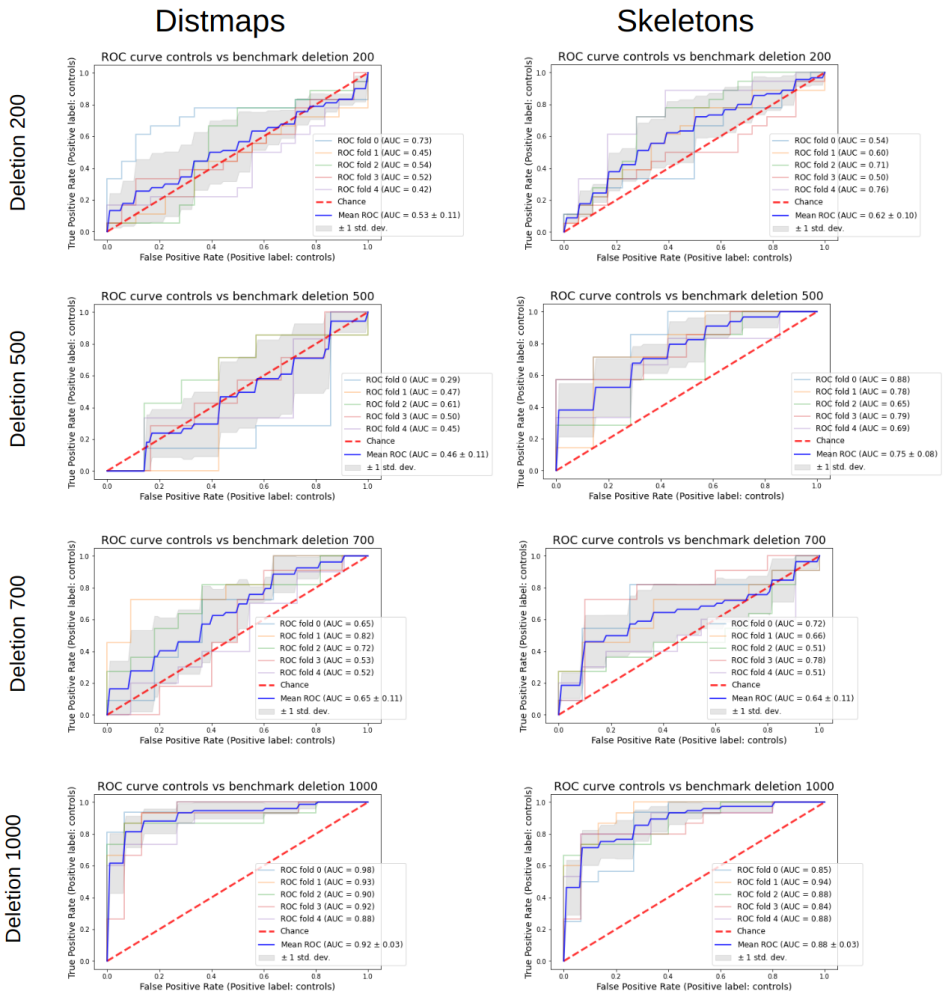


Figure C.3: *Comparison of ROC curves of deletion benchmarks for skeletons and distance maps.* For both columns, each row corresponds to one benchmark. Left: ROC obtained with distance maps as input. Right: ROC obtained with skeletons as input.

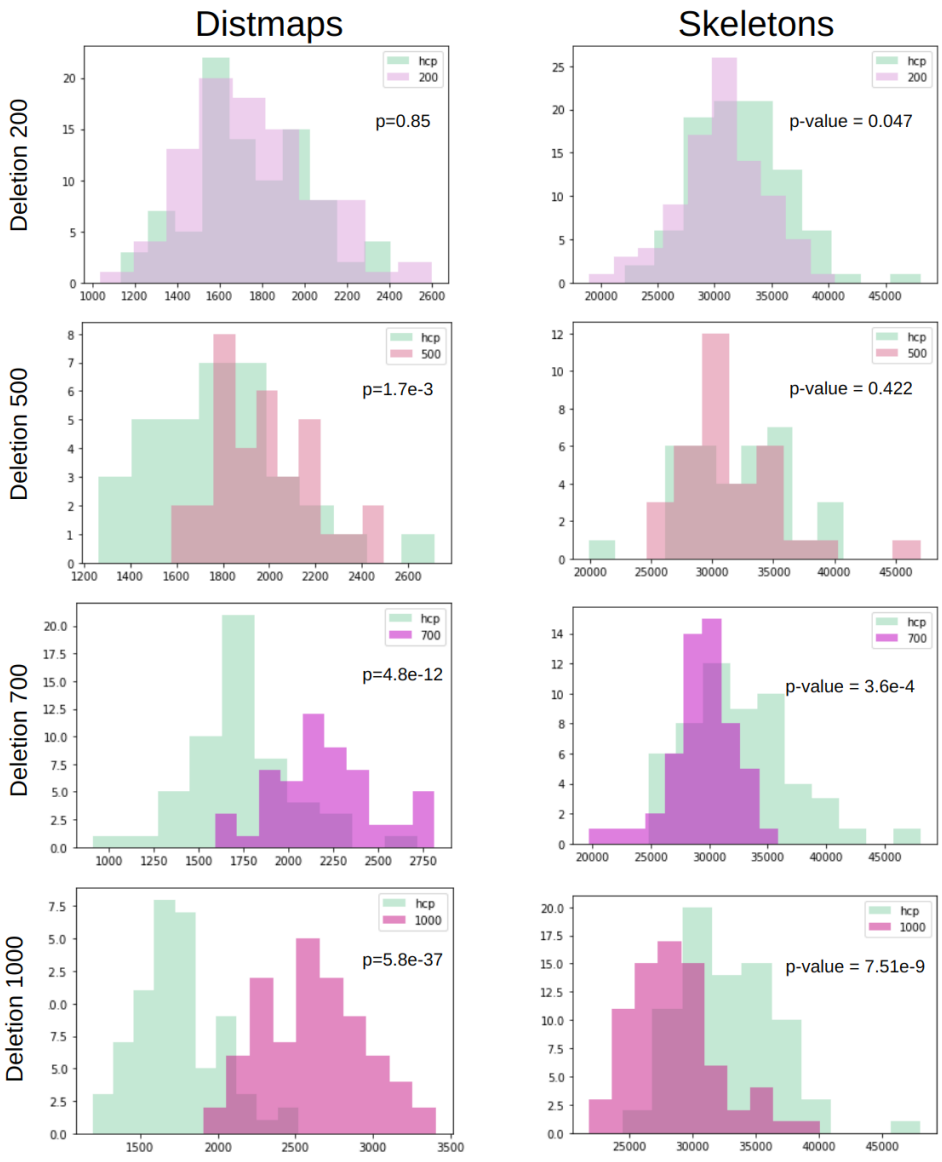
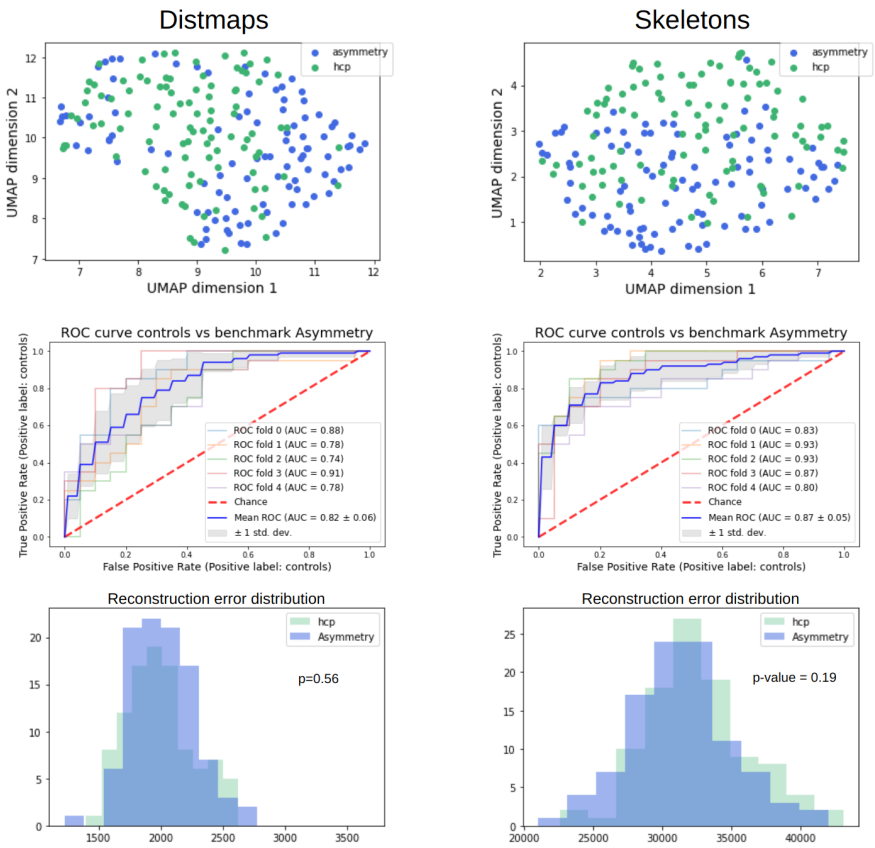
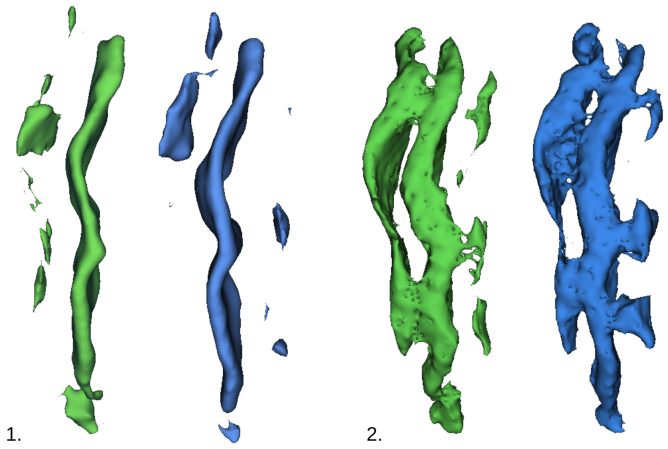


Figure C.4: *Comparison of reconstruction error distributions of deletion benchmarks for skeletons and distance maps.* For both columns, each row corresponds to one benchmark. Left: Reconstruction error distributions obtained with distance maps as input. Right: Reconstruction error distributions obtained with skeletons as input.



(a) UMAP projection of benchmark and control data, ROC curves of classification of control and benchmark data, and reconstruction error distributions.



(b) Averages for control subjects i.e. right hemispheres (in green), and for the asymmetry subjects, i.e. left hemispheres (in blue). 1. Averages obtained using distance maps as inputs and 2. using skeletons as inputs.

Figure C.5: *Comparison of asymmetry benchmark results for distance maps and skeletons.* Controls are represented in green and benchmark data in blue.

D - Current inpainting-based model

Recently, I have started working on an improved version of our framework based on the $\beta - VAE$ presented earlier. However, I have not had the time to fully formalize the approach and perform extensive analyses, so only the preliminary method and results are presented here.

Methods

This new model is inspired by self-supervised learning and some natural language processing applications where a missing word is predicted based on the surrounding words. The hypothesis is that it is possible to predict the shape of a fold based on the neighbouring folds. For instance, if the upper part of the central sulcus, which contains the hand knob, is missing, a model trained to learn the control representation of this area should be able to add this upper part. In the anomaly detection field, inpainting has been used in order to improve the representations learned (Pirnay and Chai, 2021; Zavrtanik et al., 2021). In medical applications, a framework combining a VAE with an inpainting task was introduced to increase the performances of brain anomaly detection (Zimmerer et al., 2018, 2019b).

Here, we propose to add to the $\beta - VAE$ framework presented in II a specific inpainting objective. In particular, rather than masking a random patch, we erase a simple surface (SS) of an input and the model is trained to reconstruct the original image. This additional objective acts as a pretext task, where the downstream task is the learning of a control representation.

To avoid any confusion, we will call our usual masked crops (see Part II) our *region of interest* (ROI). Simply put, the only difference with the previous framework is that instead of the usual ROI, one simple surface of this ROI is masked and the reconstruction is compared with the unmasked ROI. Specifically, during training, for each sample, one random SS is erased with a probability of 80%. To be deleted, SS must comprise at least 300 voxels inside the ROI. At inference, similar to (Zavrtanik et al., 2021), we iteratively mask all the SS of more than 300 voxels in the ROI, encode these altered images and reconstruct them. We retain the maximal reconstruction error to try to distinguish control subjects and rare configurations. Our assumption is that if a SS highly deviates from the modelled norm, the reconstruction error will be higher.

We tested this framework on the same central region crop as before. To assess the method, we used our benchmarks and tested on interrupted central sulci.

Preliminary results

Our preliminary results are presented in Fig.D.1. First, regarding our deletion benchmarks, it seems that deletion 200 and 500 correspond to slightly higher reconstruction errors than in the classic framework. Nevertheless, we observe less difference between the distributions for deletion 700. For deletion 1000 and the asymmetry benchmark, the results are quite similar for the two frameworks. Although it is harder to conclude for the interrupted central sulci as there are only a few of them, it appears that their reconstruction errors are more similar to the controls.

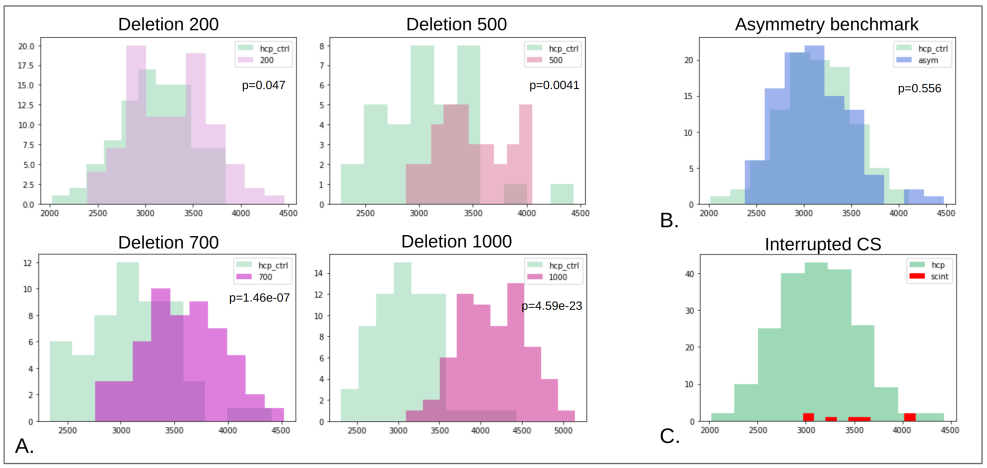


Figure D.1: *Current results obtained with our inpainting-based model.* A. Benchmark deletion results. B. Asymmetry benchmark results. C. Results on interrupted central sulci. In each case, the histograms represent the reconstruction error distribution. The p-value indicated corresponds to the Kolmogorov-Smirnov test with the null hypothesis that the two samples come from the same distribution.

Therefore, regarding deletion benchmarks, it seems that the inpainting-based model is better suited to identify the deviations of small branches that contain less than 700 voxels. However, the difference between the distributions of the reconstruction errors of controls and benchmarks is not massive. In the current state of the approach, it appears that our approach without adding this objective leads to better performances. Additional work is required to conclude.

I did not have the time but several aspects could be further studied. In particular, rather than taking the maximal reconstruction error, it could be interesting to compare the individual distributions of the reconstruction errors according to the size of the deleted SS. Using distributions rather than a maximal error would also have the advantage of *normalizing* the error by the size of the associated SS. Indeed, one could expect that a bigger SS leads to a higher reconstruction error since it is composed of more voxels. Last, locating and analyzing the erased SS

that led to the maximal error could constitute a direct way of locating the deviating pattern.

Résumé étendu en français

Introduction

Une des premières caractéristiques visuelles du cerveau est son aspect plissé, similaire à une noix. Il est en effet constitué de circonvolutions appelées gyri, qui sont délimitées par des sillons, appelés sulci. À l'instar des empreintes digitales, les motifs du plissement cortical, c'est-à-dire l'arrangement, les caractéristiques et la forme des sillons est propre à chaque individu (Ono et al., 1990). Les neuroanatomistes se sont intéressés à la caractérisation des sillons, d'abord à partir de specimens, puis grâce aux avancées de neuro-imagerie. Ces études ont permis de proposer une nomenclature des sillons et plus tard, d'entraîner des modèles à reconnaître automatiquement les différents sillons (Rivière et al., 2002; Perrot et al., 2011; Borne et al., 2020). Cette nomenclature et les performances des modèles suggèrent que même si chaque cerveau possède des motifs de plissement propres, une stabilité est observée.

Le développement de techniques non-invasives telles que l'imagerie par résonnance magnétique (IRM) a permis d'aller plus loin dans l'étude des sillons. Par exemple, des outils de visualisation ont facilité la description des motifs locaux de plissement. En outre, l'IRM fonctionnelle a permis d'étudier le lien entre les formes des sillons et diverses fonctions cognitives. Ainsi, certains motifs de sillons typiques ont pu être corrélés à la dexterité manuelle ou à la lecture par exemple (Yousry et al., 1997; Sun et al., 2016). Au contraire, des motifs rares ont également été décrits, comme l'interruption du sillon central. Certains ont même été corrélés à des troubles du neuro-développement. Par conséquent, l'étude des motifs des sillons corticaux est particulièrement intéressante.

De par la très grande variabilité inter-individuelle observée dans les motifs de plissement, la création de grandes bases de données est absolument nécessaire que ce soit pour identifier des configurations de sillons typiques ou rares. Néanmoins, analyser une multitude de données peut s'avérer compliqué. Les avancées de l'apprentissage automatique et de l'apprentissage profond en particulier, constituent une réelle opportunité pour faire face à un gros flux de données et à la grande variabilité. En effet, de tels modèles pourraient apprendre une représentation du plissement cérébral dans l'objectif de décrire à la fois des configurations typiques ainsi que des des configurations rares, pouvant mener à l'identification de biomarqueurs.

Ce travail de thèse propose d'utiliser une méthode d'apprentissage automatique profond, non-supervisé pour modéliser la variabilité inter-individuelle des motifs de plissement d'une population contrôle. À partir de cette représentation apprise, nous cherchons à identifier des configurations de sillons caractéristiques avant de

nous intéresser aux marges de la représentation pour identifier des configurations rares.

Contexte

Le plissement cérébral n'est pas propre au cerveau humain, d'autres espèces possèdent la caractéristique d'avoir un cerveau plissé qui est un trait ancestral des mammifères (Lewitus et al., 2014). Cependant, l'intensité de plissement varie d'une espèce à l'autre. Chez l'humain, des études ont montré l'existence de trois vagues de plissement successives qui conduisent respectivement aux plis primaires, secondaires et tertiaires (Chi et al., 1977). L'étude des mécanismes impliqués est toujours un champ actif de recherche où historiquement, deux hypothèses principales ont été proposées. La première repose sur des considérations mécaniques et notamment sur des vitesses d'expansion différentes selon les couches du cortex. La deuxième hypothèse est fondée sur des mécanismes biologiques impliquant processus cellulaires et génétiques. Finalement, bien que ces deux propositions aient été opposées, les processus de plissement pourraient s'avérer être une combinaison des approches mécanique et biologique comme le propose une revue récente (Llinares-Benadero and Borrell, 2019).

Cependant, au cours de ces processus, des perturbations peuvent se produire et engendrer divers types de malformations. Ces-dernières se traduisent par un éventail de symptômes de gravité variable. Les perturbations peuvent également conduire à des motifs de plissement rares et anormaux, tels que la lissencéphalie (cerveau plat) ou la dysplasie corticale focale (FCD).

Les processus de plissement aboutissent à la formation d'environ 64 sillons par hémisphère. Au sein de la population humaine, une stabilité des motifs de plissement est observée avec des localisations, formes et arrangements globalement similaires (Ono et al., 1990). Dans le cadre de cette thèse, il est important de définir les termes suivants : ***(1) forme d'un sillon*** : élément de base pouvant être arrangé pour former des motifs. Par exemple un arc de cercle ou un segment (voir figure 1.3); ***(2) motif*** : combinaison, arrangement de formes élémentaires d'un ou plusieurs plis. Par exemple, plusieurs motifs existent dans la région centrale comme une configuration avec une seule bosse, deux bosses, ou un sillon plat.

Grâce aux avancées de neuro-imagerie, des études se sont penchées sur la caractérisation des formes et motifs de plissement. Ainsi le sillon central comporte généralement une ou plusieurs bosses (Yousry et al., 1997). De même le sillon mi-fusiforme présente souvent un motif d'oméga (Weiner et al., 2014). Au contraire, des motifs rares ont été identifiés tels que l'interruption du sillon central. L'étude des sillons est particulièrement intéressante à cause de leur lien avec des fonctions cérébrales. En effet, certains motifs ou formes ont été corrélés à des fonctions comme la bosse du sillon central à la latéralité manuelle. Au contraire, des motifs

de sillons ont aussi été associés à des troubles du neuro-développement comme le "Power Button Sign" (PBS) qui a été corrélé à des cas de dysplasie corticale focale (Mellerio et al., 2014). Par conséquent, l'étude des motifs de plissement cérébral peut être intéressante pour faire émerger des biomarqueurs. Enfin, les motifs de plissement ont l'avantage d'être des propriétés traits, c'est-à-dire qu'ils sont stables au cours de la vie, ce qui en fait des témoins des processus du neuro-développement.

Cependant, l'étude des méthodes de plissement nécessite d'avoir et d'analyser de très grandes bases de données. Depuis quelques années, l'apprentissage profond a révolutionné de nombreux domaines. Spécifiquement, l'apprentissage profond non supervisé est utilisé pour apprendre et faire émerger des représentations dans des données. Ce type d'approche est notamment utilisé dans le domaine de la détection d'anomalie où de nombreux modèles ont été proposés. Une approche très répandue fait appel aux auto-encodeurs (AE). Ces modèles ont la particularité de réduire une image d'entrée à une représentation de plus petite dimension dans un espace appelé "espace latent", avant de la reconstruire en conservant le plus de détails possible. De nombreuses variantes d'AE ont été proposées comme l'auto-encodeur variationnel (VAE) (Kingma and Welling, 2014). Dans ces méthodes, le modèle est entraîné uniquement sur des sujets contrôles. L'hypothèse est que le modèle va apprendre uniquement des caractéristiques communes et ne sera pas capable de reconstruire des anomalies. Les applications de détection d'anomalie sont très étudiées dans le domaine de l'imagerie médicale et permettent notamment d'identifier des tumeurs dans divers organes.

Les processus du neuro-développement peuvent conduire à des motifs de plissement rares ou anormaux, qui en tant que caractère trait, peuvent être utilisés comme des témoins ou des marqueurs des événements sous-jacents impliqués. De plus, ils ont été associés aussi bien à des fonctions cognitives qu'à des pathologies. Leur étude est donc particulièrement pertinente pour aider à comprendre le développement du cerveau et à découvrir des biomarqueurs pour certains troubles neuro-développementaux. Cette thèse a pour objectif d'identifier des motifs de plissement anormaux avec une méthode d'apprentissage profond non-supervisée. Dans un premier temps, nous allons chercher à modéliser la variabilité inter-individuelle du plissement, puis nous analyserons les motifs typiques avant de nous intéresser aux marges de la représentation pour caractériser des motifs rares.

Méthodes générales

La base de données que nous avons utilisée pour modéliser la variabilité de plissement des sujets est la base HCP dont nous utilisons les IRM T1 comme données d'entrée. Cette base est composée d'environ 1000 sujets. Les images IRM T1 sont riches de beaucoup d'information, au-delà de la morphologie des sillons corticaux. Par conséquent, nous avons d'abord traité nos données avec le

logiciel BrainVISA/Morphologist qui permet d'obtenir un graphe de sillons. Nous avons ensuite développé un pré-traitement qui génère deux types d'image 3D à partir des graphes, les squelettes, qui correspondent à un moule en négatif du cerveau avec uniquement les informations concernant les sillons; et les cartes de distances qui sont générées à partir des squelettes. Les cartes de distances sont des images continues où la valeur de chaque voxel correspond à la distance au plus proche sillon. Comme nous nous intéressons à des motifs de plissement locaux, nous avons défini des régions d'intérêt autour de sillons d'intérêt. Pour cela un masque de la zone d'intérêt est appris sur une base de données dont les sillons ont été étiquetés. Le masque est ensuite appliqué sur des données sans étiquettes de sillons. Par conséquent, nous n'avons pas besoin d'avoir des données labellisées pour notre approche.

Pour apprendre la variabilité du plissement nous avons entraîné un $\beta - VAE$ (Higgins et al., 2016). Les VAE (Kingma and Welling, 2014) font partie de la famille des auto-encodeurs mais ont la particularité de projeter les données sous forme de distribution dans l'espace latent. Le $\beta - VAE$ est uniquement entraîné sur des sujets considérés comme contrôles afin d'apprendre une représentation de la diversité des motifs de sillons *normale*. Cette thèse propose deux applications à partir des représentations apprises : la caractérisation de motifs typiques de la population et l'identification de motifs rares.

Identification de motifs typiques

Méthodes spécifiques

Pour cette première application nous nous sommes intéressés à la région du cingulaire. Cette dernière est composée de deux sillons principaux, le cingulaire et, dans certains cas, le paracingulaire. La zone du cingulaire a un intérêt clinique et a été associée à des cas de faible contrôle inhibiteur et de schizophrénie (Provost et al., 2003; Yücel et al., 2003; Borst et al., 2014; Cachia et al., 2014). Dans cette partie, en plus de la base HCP, nous avons utilisé une autre base dans laquelle la présence du sillon paracingulaire est labellisée.

Le pré-traitement décrit précédemment a été appliqué : nos données d'entrée sont des squelettes binaires correspondant à la région cingulaire, la résolution est de 2mm. Les représentations de la région sont apprises par deux modèles, le $\beta-VAE$ et SimCLR (Chen et al., 2020) et la caractérisation des motifs typiques se fait dans l'espace latent. Pour cela, nous avons réduit tout d'abord cet espace à 2 dimensions à l'aide d'un t-SNE, puis nous avons appliqué un clustering. Ensuite, à partir de l'espace latent, les motifs représentatifs des centroïdes de chacun des clusters sont reconstruits et analysés; dans l'espace d'entrée, des sillons "moyens" sont générés à partir des squelettes d'entrée des sujets constituant chacun des clusters. Afin d'avoir une évaluation quantitative de la qualité de nos représentations latentes, un classifieur est entraîné à prédire la présence d'un sillon paracingulaire à partir

des représentations latentes.

Résultats

La taille d'espace latent retenue est de quatre dimensions. Pour les deux modèles, quatre clusters sont obtenus mais l'organisation de l'espace latent est très différente. Le décodage des motifs caractéristiques des clusters a mis en lumière différents motifs. La qualité des motifs obtenus avec SimCLR est moins bonne qu'avec le $\beta - VAE$. Dans les deux cas, un motif représentant le cingulaire et un paracingulaire est observé. Avec le $\beta - VAE$, un autre type de motif décrit dans la littérature a été obtenu : un long cingulaire interrompu. Les motifs représentatifs générés à partir des squelettes d'entrée sont plus difficiles à analyser à cause du rendu moins fin mais on retrouve certains des motifs de l'espace latent.

L'étude de la généralisation suggère que certains clusters sont plus stables que d'autres. Dans l'ensemble les sujets d'un cluster semblent rester proches.

Les performances de classification de la présence d'un paracingulaire sont très différentes entre le $\beta - VAE$ et SimCLR, avec des résultats de 63% et 50% respectivement ce qui suggère que l'espace latent du $\beta - VAE$ est plus pertinent que celui de SimCLR pour détecter le sillon paracingulaire.

Conclusion

Les deux modèles que nous avons implémentés conduisent à un espace latent structuré à partir duquel plusieurs motifs ont émergé. L'espace latent du $\beta - VAE$ semble plus pertinent pour l'identification de sillons typiques dans la région cingulaire et conduit à des motifs décrits dans la littérature. Une explication possible quant aux moins bons résultats de SimCLR est liée à son architecture, qui pourrait être trop complexe pour notre type de données et la tâche. Cette hythèse est confirmée par les travaux récents de l'équipe : utiliser la même architecture d'encodeur que pour le $\beta - VAE$ conduit à des performances quasiment équivalentes à celles du $\beta - VAE$. L'aspect génératif du $\beta - VAE$ est un réel atout pour comprendre les représentations apprises et faciliter l'analyse de régions complexes. Pour encourager l'espace latent à être structuré et à être plus facilement clusterisé, un objectif de clustering pourrait être ajouté au cours de l'entraînement. Les futurs travaux s'efforceront d'analyser de potentielles corrélations entre les motifs identifiés et des mesures de cognition ou de pathologies.

Identification de motifs rares et anormaux

Méthodes spécifiques

Nous avons choisi d'étudier la région centrale pour cette application. En effet, le sillon central est l'un des premiers sillons à se former et il est assez stable pour constituer une première étape dans la modélisation de la variabilité inter-individuelle. En outre, généralement long et continu, ce sillon peut être interrompu dans de rares cas (moins de 1%), ce qui fait des sillons centraux interrompus de

bons candidats pour l'étude des motifs rares.

Les données sont pré-traitées pour obtenir des cartes de distances comme indiqué précédemment. Nos données d'entrée sont donc des images continues 3D où chaque voxel a pour valeur sa distance au plus proche sillon. Dans ce travail, nous nous sommes concentrés uniquement sur l'hémisphère droit.

Pour apprendre une représentation de la région, un $\beta - VAE$ est entraîné sur la cohorte HCP de laquelle les sillons interrompus ont été exclus. Une des difficultés de ce travail est le manque de motifs identifiés comme rares. De plus, il est intéressant de pouvoir quantifier le type de déviations que notre approche permet d'identifier. Pour y répondre, plusieurs ensembles de données synthétiques ont été générés : dans le premier nous avons supprimé des parties de sillons de taille croissante (ensembles suppression), le second correspond à la même région d'intérêt mais dans l'hémisphère gauche (ensemble asymétrie).

L'identification de patterns rares se fait à deux niveaux, dans l'espace latent et dans l'espace des images que nous appelons "espace des sillons", c'est-à-dire à partir de l'erreur de reconstruction. En ce qui concerne l'espace latent, pour avoir une première visualisation de la distribution, nous réduisons cet espace à deux dimensions. Pour quantifier la qualité des représentations de l'espace latent, des classifieurs sont entraînés à distinguer les sujets contrôles des anomalies synthétiques. La génération de nouvelles données échantillonnées à partir de l'espace latent permet de mieux comprendre les propriétés encodées. Dans l'espace des sillons, nous avons comparé la distribution des erreurs de reconstruction. Enfin, pour tester la généralisation de notre approche, nous avons appliqué ce modèle à deux autres régions où des patterns anormaux ont été décrits : des patients sans corps calleux et des patients souffrant de FCD.

Résultats

L'espace latent obtenu est composé de 75 dimensions. Concernant les ensembles de données synthétiques, des résultats différents ont été obtenus selon le type de déviations (suppression de branches ou asymétrie) et selon l'espace. En particulier, pour l'ensemble suppression, dans l'espace latent, plus la taille de la surface supprimée est grande, plus on observe des distributions différentes et de meilleures performances de classification. Similairement, dans l'espace des sillons, les erreurs de reconstructions sont de plus en plus importantes. Pour l'ensemble asymétrie, on observe des distributions différentes dans l'espace latent, en revanche aucune différence n'est observée en comparant les erreurs de reconstruction. La génération de motifs moyens permet d'identifier des caractéristiques typiques des sillons centraux droit et gauche, comme une plus grande fréquence de motifs présentant deux bosses dans l'hémisphère gauche. Concernant les sillons interrompus, ces-derniers semblent avoir tendance à être distribués à la marge des contrôles. Les erreurs de reconstruction semblent également être distribuées différemment mais il est difficile de conclure étant donné le faible nombre de sujets (n=7). La reconstruction et la localisation des erreurs permettent d'indiquer la caractéristique atypique. Par

exemple, dans six cas sur sept le modèle a reconstruit les sillons interrompus en remplissant l'interruption.

Les deux applications aux autres régions montrent que les patients ayant des anomalies de plissement semblent avoir une distribution différente dans l'espace latent et dans l'espace des sillons, suggérant une bonne généralisation de notre approche. Là encore, la localisation des erreurs de reconstruction souligne les caractéristiques atypiques.

Conclusion

Cette partie a proposé une méthode pour étudier les motifs de plissement corticaux rares dans la région centrale et a été appliquée à un motif décrit, le cas de sillons centraux interrompus. Avec la génération de données synthétiques, nous avons proposé une approche pour analyser le type de déviations pouvant être identifiées par notre modèle. Les résultats ont montré que l'espace latent et l'espace des sillons apportaient des informations complémentaires. Enfin, nous avons évalué la généralisation de notre approche grâce à l'étude de deux anomalies situées dans d'autres régions. Cette partie ouvre la voie à de nouveaux travaux qui pourraient s'intéresser à la définition de nouveaux ensembles de données anormales synthétiques ou l'utilisation de méthodes d'apprentissage profond exploitant la présence d'anomalies pendant l'entraînement.

Conclusion générale

J'ai présenté dans cette thèse un travail en trois étapes principales. D'abord une représentation de la variabilité inter-individuelle du plissement est apprise. À partir de cette représentation, deux applications sont proposées, l'identification de motifs caractéristiques et l'identification de motifs rares. Ces deux applications ont permis de montrer la pertinence de l'apprentissage profond non supervisé pour l'étude du plissement cortical.

Ce travail soulève plusieurs points à mentionner.

Tout d'abord il est intéressant de noter les différentes tailles d'espaces latents entre nos deux applications : 4 dans le premier cas pour identifier les motifs typiques, contre 75 pour les motifs rares. Cette différence de taille d'espace latent se retrouve dans d'autres travaux (Baur et al., 2020; Vétil et al., 2022). La détection d'anomalies semble donc nécessiter davantage de dimensions. Une hypothèse explicative concerne l'encodage dans l'espace latent. En particulier, dans le cas idéal, chaque dimension latente correspondrait à une caractéristique. Ainsi, lors de l'identification de motifs typiques, un plus faible nombre de dimensions correspondrait à moins de caractéristiques différentes encodées. La variabilité serait donc exprimée sur un sous-ensemble de caractéristiques seulement. Dans ce cas, le clustering en serait facilité. Au contraire, pour identifier les motifs rares, avoir plus de caractéristiques pour exprimer la variabilité permettrait d'englober tout

l'éventail des particularités de la population plus facilement. Néanmoins, avoir plus de dimensions latentes complique l'analyse de cet espace.

L'étude du plissement est complexe et peut être réalisée avec différentes approches et différents objets d'étude. Dans cette thèse nous avons analysé les sillons en les modélisant avec des images de squelettes ou leur version continue, les cartes de distances. Ces deux types de données semblent mener à de bons résultats comme le montrent nos deux applications mais les cartes de distances présentent certains avantages. En effet, elles permettent d'introduire des notions de proximité au sillon. De plus, la méthode pour les visualiser conduit à des motifs de meilleure qualité. En prenant du recul, il est intrigant de penser que cette thèse analyse des objets représentant la forme du vide. Néanmoins, les objets sulcaux tels que nous les manipulons constituent des proxys pour diverses structures du cerveau. Le sillon représente les murs des gyri. La présence d'un sillon additionel peut conduire à un changement dans l'organisation cytoarchitecturale par exemple (Amiez et al., 2021). Enfin, de nombreuses fibres sont présentes autour des sillons. Par conséquent, les objets que nous considérons sont à l'intersection de plusieurs modalités et permettent l'analyse de multiples aspects. Il serait attractif que de prochains travaux lient nos résultats à l'organisation cellulaire du cortex et aux fibres.

Une limite potentiellement importante de cette thèse réside dans la taille des bases de données utilisées. En effet, des travaux ont montré que les représentations apprises étaient meilleures avec davantage de données (Chavas et al., 2023). Avoir davantage de données représentant des motifs rares est également essentiel pour conclure de façon plus fiable. Dans ce travail nous avions peu de sujets avec des motifs rares ou anormaux, seulement 7 sujets avec des sillons interrompus, 7 patients sans corps calleux et entre 4 et 7 patients souffrant de FCD. D'autres travaux avec davantage de sujets seront nécessaires pour valider notre approche.

De plus, la base de données que nous avons utilisée, HCP, comporte la particularité d'avoir des sujets de la même famille, y compris des jumeaux. Or il a été montré que les motifs de plissement corticaux étaient en partie héritable (Im et al., 2011). Dans cette thèse nous n'avons pas appliqué de méthodologie particulière par rapport à cette caractéristique en faisant l'hypothèse que la proximité des motifs de jumeaux pouvait être inclue dans la variabilité inter-individuelle. Néanmoins, il serait nécessaire de répliquer nos travaux avec d'autres cohortes pour valider cette hypothèse.

Pour conclure, les motifs de plissement cortical présentent une très grande variabilité inter-individuelle qui découle des processus du neuro-développement. La caractérisation de motifs typiques et rares de plissement peut aider à déchiffrer

la complexité du plissement et à mieux comprendre la formation du cerveau. Les méthodes d'apprentissage profond non supervisé semblent être un moyen efficace pour répondre à cette tâche. L'approche que nous proposons est une première étape vers un modèle systématique à l'échelle du cerveau, qui bénéficierait de l'intégration d'autres modalités comme la cytoarchitecture et la connectivité.

Bibliography

Aggarwal, C. C. (2017). An Introduction to Outlier Analysis. In Aggarwal, C. C., editor, *Outlier Analysis*, pages 1–34. Springer International Publishing, Cham.

Ahtam, B., Turesky, T. K., Zöllei, L., Standish, J., Grant, P. E., Gaab, N., and Im, K. (2021). Intergenerational Transmission of Cortical Sulcal Patterns from Mothers to their Children. *Cerebral Cortex*, 31(4):1888–1897.

Akcay, S., Atapour-Abarghouei, A., and Breckon, T. P. (2019). GANomaly: Semi-supervised Anomaly Detection via Adversarial Training. In Jawahar, C. V., Li, H., Mori, G., and Schindler, K., editors, *Computer Vision – ACCV 2018*, Lecture Notes in Computer Science, pages 622–637, Cham. Springer International Publishing.

Amiez, C., Sallet, J., Novek, J., Hadj-Bouziane, F., Giacometti, C., Andersson, J., Hopkins, W. D., and Petrides, M. (2021). Chimpanzee histology and functional brain imaging show that the paracingulate sulcus is not human-specific. *Communications Biology*, 4(1):1–12. Number: 1 Publisher: Nature Publishing Group.

Amunts, K., Malikovic, A., Mohlberg, H., Schormann, T., and Zilles, K. (2000). Brodmann's areas 17 and 18 brought into stereotaxic space-where and how variable? *NeuroImage*, 11(1):66–84.

Amunts, K., Schlaug, G., Schleicher, A., Steinmetz, H., Dabringhaus, A., Roland, P. E., and Zilles, K. (1996). Asymmetry in the Human Motor Cortex and Handedness. *NeuroImage*, 4(3):216–222.

Amunts, K., Schleicher, A., and Zilles, K. (2007). Cytoarchitecture of the cerebral cortex—More than localization. *NeuroImage*, 37(4):1061–1065.

Atlason, H. E., M.d, A. L., Sigurdsson, S., M.d, V. G., and Ellingsen, L. M. (2019). Unsupervised brain lesion segmentation from MRI using a convolutional autoencoder. In *Medical Imaging 2019: Image Processing*, volume 10949, pages 372–378. SPIE.

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

Auzias, G., Viellard, M., Takerkart, S., Villeneuve, N., Poinso, F., Fonséca, D. D., Girard, N., and Deruelle, C. (2014). Atypical sulcal anatomy in young children with autism spectrum disorder. *NeuroImage: Clinical*, 4:593–603.

Barkovich, A. J., Guerrini, R., Kuzniecky, R. I., Jackson, G. D., and Dobyns, W. B. (2012). A developmental and genetic classification for malformations of cortical development: update 2012. *Brain: A Journal of Neurology*, 135(Pt 5):1348–1369.

Barkovich, A. J., Kuzniecky, R. I., Dobyns, W. B., Jackson, G. D., Becker, L. E., and Evrard, P. (1996). A classification scheme for malformations of cortical development. *Neuropediatrics*, 27(2):59–63.

Barlow, H. (1989). Unsupervised Learning. *Neural Computation*, 1(3):295–311.

Baur, C., Denner, S., Wiestler, B., Albarqouni, S., and Navab, N. (2020). Autoencoders for Unsupervised Anomaly Segmentation in Brain MR Images: A Comparative Study. *arXiv:2004.03271 [cs, eess]*. arXiv: 2004.03271.

Baur, C., Wiestler, B., Albarqouni, S., and Navab, N. (2019). Deep Autoencoding Models for Unsupervised Anomaly Segmentation in Brain MR Images. In Crimi, A., Bakas, S., Kuijf, H., Keyvan, F., Reyes, M., and van Walsum, T., editors, *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, Lecture Notes in Computer Science, pages 161–169, Cham. Springer International Publishing.

Behrendt, F., Bengs, M., Rogge, F., Krüger, J., Opfer, R., and Schlaefer, A. (2022). Unsupervised Anomaly Detection in 3D Brain MRI Using Deep Learning with Impured Training Data. In *2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI)*, pages 1–4. ISSN: 1945-8452.

Benitez, J. R., Kannan, S., Hastings, W. L., Parker, B. J., Willbrand, E. H., and Weiner, K. S. (2022). Ventral temporal and posteromedial sulcal morphology in autism spectrum disorder. Pages: 2022.09.01.506213 Section: New Results.

Berg, A., Ahlberg, J., and Felsberg, M. (2019). Unsupervised Learning of Anomaly Detection from Contaminated Image Data using Simultaneous Encoder Training. arXiv:1905.11034 [cs, eess].

Bo, S., Haitao, G., Yuchun, T., Zhongyu, H., Junhai, X., Xiangtao, L., and Shuwei, L. (2015). Asymmetries of the central sulcus in young adults: Effects of gender, age and sulcal pattern. *International journal of developmental neuroscience : the official journal of the International Society for Developmental Neuroscience*, 44. Publisher: Int J Dev Neurosci.

Bodensteiner, J., Schaefer, G., Breeding, L., and Cowan, L. (1994). Hypoplasia of the Corpus Callosum: A Study of 445 Consecutive MRI Scans. *Journal of Child Neurology*, 9(1):47–49. Publisher: SAGE Publications Inc.

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

Boling, W. W. and Olivier, A. (2004). Localization of hand sensory function to the pli de passage moyen of Broca. *Journal of Neurosurgery*, 101(2):278–283.

Borne, L., Rivière, D., Cachia, A., Roca, P., Mellerio, C., Oppenheim, C., and Mangin, J.-F. (2021). Automatic recognition of specific local cortical folding patterns. *NeuroImage*, 238:118208.

Borne, L., Rivière, D., Mancip, M., and Mangin, J.-F. (2020). Automatic labeling of cortical sulci using patch- or CNN-based segmentation techniques combined with bottom-up geometric constraints. *Medical Image Analysis*, 62:101651.

Borrell, V. (2018). How Cells Fold the Cerebral Cortex. *Journal of Neuroscience*, 38(4):776–783. Publisher: Society for Neuroscience Section: Dual Perspectives.

Borst, G., Cachia, A., Vidal, J., Simon, G., Pineau, A., Fischer, C., Poirel, N., Mangin, J.-F., and Houdé, O. (2014). Folding of the anterior cingulate cortex partially explains inhibitory control during childhood: A longitudinal study. *Developmental Cognitive Neuroscience*, 9:126–135. Publisher: Elsevier.

Breunig, M. M., Kriegel, H.-P., Ng, R. T., and Sander, J. (2000). LOF: identifying density-based local outliers. In *Proceedings of the 2000 ACM SIGMOD international conference on Management of data*, SIGMOD '00, pages 93–104, New York, NY, USA. Association for Computing Machinery.

Broca, P. and Pozzi, S. J. (1888). *Mémoires sur le cerveau de l'homme et des primates*. C. Reinwald. Google-Books-ID: d99EAAAAYAAJ.

Burgess, C. P., Higgins, I., Pal, A., Matthey, L., Watters, N., Desjardins, G., and Lerchner, A. (2018). Understanding disentangling in β-VAE. arXiv:1804.03599 [cs, stat].

Bénézit, A., Hertz-Pannier, L., Dehaene-Lambertz, G., Monzalvo, K., Germanaud, D., Duclap, D., Guevara, P., Mangin, J.-F., Poupon, C., Moutard, M.-L., and Dubois, J. (2015). Organising white matter in a brain without corpus callosum fibres. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, 63:155–171.

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

Cachia, A., Borst, G., Vidal, J., Fischer, C., Pineau, A., Mangin, J.-F., and Houdé, O. (2014). The Shape of the ACC Contributes to Cognitive Control Efficiency in Preschoolers. *Journal of Cognitive Neuroscience*, 26(1):96–106.

Caron, M., Bojanowski, P., Joulin, A., and Douze, M. (2018). Deep Clustering for Unsupervised Learning of Visual Features. In Ferrari, V., Hebert, M., Sminchisescu, C., and Weiss, Y., editors, *Computer Vision – ECCV 2018*, Lecture Notes in Computer Science, pages 139–156, Cham. Springer International Publishing.

Caron, M., Misra, I., Mairal, J., Goyal, P., Bojanowski, P., and Joulin, A. (2021). Unsupervised Learning of Visual Features by Contrasting Cluster Assignments. *arXiv:2006.09882 [cs]*. arXiv: 2006.09882.

Caulo, M., Briganti, C., Mattei, P. A., Perfetti, B., Ferretti, A., Romani, G. L., Tartaro, A., and Colosimo, C. (2007). New morphologic variants of the hand motor cortex as seen with MR imaging in a large study population. *AJNR. American journal of neuroradiology*, 28(8):1480–1485.

Chakravarty, M. M., Rapoport, J. L., Giedd, J. N., Raznahan, A., Shaw, P., Collins, D. L., Lerch, J. P., and Gogtay, N. (2015). Striatal shape abnormalities as novel neurodevelopmental endophenotypes in schizophrenia: A longitudinal study. *Human Brain Mapping*, 36(4):1458–1469. Publisher: John Wiley & Sons, Ltd.

Chalapathy, R. and Chawla, S. (2019). Deep Learning for Anomaly Detection: A Survey. *arXiv:1901.03407 [cs, stat]*. arXiv: 1901.03407.

Chandola, V., Banerjee, A., and Kumar, V. (2009). Anomaly detection: A survey. *ACM Computing Surveys*, 41(3):15:1–15:58.

Chavas, J., Gaudin, A., Guillon, L., Fischer, C., Cagna, B., Cachia, A., Rivière, D., and Mangin, J.-F. (2023). Optimizing self-supervised contrastive learning for folding pattern detection. In *Submitted to IPMI*.

Chen, T., Kornblith, S., Norouzi, M., and Hinton, G. (2020). A Simple Framework for Contrastive Learning of Visual Representations. In *Proceedings of the 37th International Conference on Machine Learning*, pages 1597–1607. PMLR. ISSN: 2640-3498.

Chen, X. and Konukoglu, E. (2018). Unsupervised Detection of Lesions in Brain MRI using constrained adversarial auto-encoders. *arXiv:1806.04972 [cs]*. arXiv: 1806.04972.

Chi, J. G., Dooling, E. C., and Gilles, F. H. (1977). Gyral development of the human brain. *Annals of Neurology*, 1(1):86–93. Publisher: John Wiley & Sons, Ltd.

Clark, W. E. L. G. (1945). *Deformation patterns in the cerebral cortex*. Printed at the Oxford University Press by John Johnson.

Cunningham, D. J. (1890). 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(Pt 3):309–345.

Danks, D. and Yau, C. (2021). BasisDeVAE: Interpretable Simultaneous Dimensionality Reduction and Feature-Level Clustering with Derivative-Based Variational Autoencoders. pages 2410–2420. PMLR.

Davatzikos, C. and Bryan, R. N. (2002). Morphometric Analysis of Cortical Sulci Using Parametric Ribbons: A Study of the Central Sulcus:. *Journal of Computer Assisted Tomography*, 26(2):298–307.

de Vareilles, H., Rivière, D., Sun, Z.-Y., Fischer, C., Leroy, F., Neumane, S., Stopar, N., Eijsermans, R., Ballu, M., Tataranno, M.-L., Benders, M., Mangin, J.-F., and Dubois, J. (2022). Shape variability of the central sulcus in the developing brain: A longitudinal descriptive and predictive study in preterm infants. *NeuroImage*, 251:118837.

Delalande, L., Moyon, M., Tissier, C., Dorriere, V., Guillois, B., Mevell, K., Charron, S., Salvia, E., Poirel, N., Vidal, J., Lion, S., Oppenheim, C., Houdé, O., Cachia, A., and Borst, G. (2020). Complex and subtle structural changes in prefrontal cortex induced by inhibitory control training from childhood to adolescence. *Developmental Science*, 23(4):e12898. Publisher: John Wiley & Sons, Ltd.

Deleo, F., Hong, S.-J., Fadaie, F., Caldairou, B., Krystal, S., Bernasconi, N., and Bernasconi, A. (2020). Whole-brain multimodal MRI phenotyping of periventricular nodular heterotopia. *Neurology*, 95(17):e2418–e2426.

Dhariwal, P. and Nichol, A. (2021). Diffusion Models Beat GANs on Image Synthesis. In Ranzato, M., Beygelzimer, A., Dauphin, Y., Liang, P. S., and Vaughan, J. W., editors, *Advances in Neural Information Processing Systems*, volume 34, pages 8780–8794. Curran Associates, Inc.

Dierker, D. L., Feczko, E., Pruett, Jr, J. R., Petersen, S. E., Schlaggar, B. L., Constantino, J. N., Harwell, J. W., Coalson, T. S., and Van Essen, D. C. (2015). Analysis of Cortical Shape in Children with Simplex Autism. *Cerebral Cortex*, 25(4):1042–1051.

Duan, D., Xia, S., Rekik, I., Meng, Y., Wu, Z., Wang, L., Lin, W., Gilmore, J. H., Shen, D., and Li, G. (2019). Exploring folding patterns of infant cerebral cortex

based on multi-view curvature features: Methods and applications. *NeuroImage*, 185:575–592.

Dufumier, B., Gori, P., Battaglia, I., Victor, J., Grigis, A., and Duchesnay, E. (2021). Benchmarking CNN on 3D Anatomical Brain MRI: Architectures, Data Augmentation and Deep Ensemble Learning. *arXiv:2106.01132 [cs, eess]*. arXiv: 2106.01132.

Dunning, T. and Friedman, E. (2014). *Practical Machine Learning: A New Look at Anomaly Detection*. "O'Reilly Media, Inc.". Google-Books-ID: dxZ-IBAAAQBAJ.

Essen, D. C. V., Dierker, D., Snyder, A. Z., Raichle, M. E., Reiss, A. L., and Korenberg, J. (2006). Symmetry of Cortical Folding Abnormalities in Williams Syndrome Revealed by Surface-Based Analyses. *Journal of Neuroscience*, 26(20):5470–5483. Publisher: Society for Neuroscience Section: Articles.

Fernando, T., Gammulle, H., Denman, S., Sridharan, S., and Fookes, C. (2022). Deep Learning for Medical Anomaly Detection – A Survey. *ACM Computing Surveys*, 54(7):1–37.

Fernández, V., Llinares-Benadero, C., and Borrell, V. (2016). Cerebral cortex expansion and folding: what have we learned? *The EMBO Journal*, 35(10):1021–1044. Publisher: John Wiley & Sons, Ltd.

Foubet, O., Sun, Z. Y., Hopkins, W., and Mangin, J.-F. (2022). Comparison of the shape of the Central Sulcus in Hominids. In *Organisation for Human Brain Mapping*, Glasgow, United Kingdom.

Frey, B. J. and Dueck, D. (2007). Clustering by Passing Messages Between Data Points. *Science*, 315(5814):972–976.

Fukushima, K. and Miyake, S. (1982). Neocognitron: A Self-Organizing Neural Network Model for a Mechanism of Visual Pattern Recognition. In Amari, S.-i. and Arbib, M. A., editors, *Competition and Cooperation in Neural Nets*, Lecture Notes in Biomathematics, pages 267–285, Berlin, Heidelberg. Springer.

Germann, J., Chakravarty, M. M., Collins, D. L., and Petrides, M. (2020). Tight Coupling between Morphological Features of the Central Sulcus and Somatomotor Body Representations: A Combined Anatomical and Functional MRI Study. *Cerebral Cortex*, 30(3):1843–1854.

Gervais, H., Belin, P., Boddaert, N., Leboyer, M., Coez, A., Sfaello, I., Barthélémy, C., Brunelle, F., Samson, Y., and Zilbovicius, M. (2004). Abnormal cortical voice processing in autism. *Nature Neuroscience*, 7(8):801–802.

Ghahramani, Z. (2004). Unsupervised Learning. In Bousquet, O., von Luxburg, U., and Rätsch, G., editors, *Advanced Lectures on Machine Learning: ML Summer Schools 2003, Canberra, Australia, February 2 - 14, 2003, Tübingen, Germany, August 4 - 16, 2003, Revised Lectures*, Lecture Notes in Computer Science, pages 72–112. Springer, Berlin, Heidelberg.

Golan, I. and El-Yaniv, R. (2018). Deep Anomaly Detection Using Geometric Transformations. In Bengio, S., Wallach, H., Larochelle, H., Grauman, K., Cesa-Bianchi, N., and Garnett, R., editors, *Advances in Neural Information Processing Systems*, volume 31. Curran Associates, Inc.

Goldstein, M. and Dengel, A. (2012). Histogram-based Outlier Score (HBOS): A fast Unsupervised Anomaly Detection Algorithm. In *KI-2012: poster and demo track, 2012*.

Gomez, C., Grigis, A., Uhrig, L., and Jarraya, B. (2022). Characterization of Brain Activity Patterns Across States of Consciousness Based on Variational Auto-Encoders. In Wang, L., Dou, Q., Fletcher, P. T., Speidel, S., and Li, S., editors, *Medical Image Computing and Computer Assisted Intervention – MICCAI 2022*, Lecture Notes in Computer Science, pages 419–429, Cham. Springer Nature Switzerland.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. (2020). Generative adversarial networks. *Communications of the ACM*, 63(11):139–144.

Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, G., Cai, J., and Chen, T. (2018). Recent advances in convolutional neural networks. *Pattern Recognition*, 77:354–377.

Guillon, L., Cagna, B., Dufumier, B., Chavas, J., Rivière, D., and Mangin, J.-F. (2021). Detection of Abnormal Folding Patterns with Unsupervised Deep Generative Models. In Abdulkadir, A., Kia, S. M., Habes, M., Kumar, V., Rondina, J. M., Tax, C., and Wolfers, T., editors, *Machine Learning in Clinical Neuroimaging*, Lecture Notes in Computer Science, pages 63–72, Cham. Springer International Publishing.

Guillon, L., Chavas, J., Pascucci, M., Dufumier, B., Rivière, D., and Mangin, J.-F. (2022). Unsupervised Representation Learning of Cingulate Cortical Folding Patterns. In Wang, L., Dou, Q., Fletcher, P. T., Speidel, S., and Li, S., editors, *Medical Image Computing and Computer Assisted Intervention – MICCAI 2022*, Lecture Notes in Computer Science, pages 77–87, Cham. Springer Nature Switzerland.

Harrison, J., Lorenzi, M., Legghe, B., Iriart, X., Cochet, H., and Sermesant, M. (2021). Phase-Independent Latent Representation for Cardiac Shape Analysis. In de Bruijne, M., Cattin, P. C., Cotin, S., Padoy, N., Speidel, S., Zheng, Y., and Essert, C., editors, *Medical Image Computing and Computer Assisted Intervention – MICCAI 2021*, Lecture Notes in Computer Science, pages 537–546, Cham. Springer International Publishing.

Haselmann, M., Gruber, D. P., and Tabatabai, P. (2018). Anomaly Detection using Deep Learning based Image Completion. *arXiv:1811.06861 [cs]*. arXiv: 1811.06861.

Hawkins, D. M. (1980). *Identification of Outliers*. Springer Netherlands, Dordrecht.

Hawkins, S., He, H., Williams, G., and Baxter, R. (2002). Outlier Detection Using Replicator Neural Networks. In Kambayashi, Y., Winiwarter, W., and Arikawa, M., editors, *Data Warehousing and Knowledge Discovery*, Lecture Notes in Computer Science, pages 170–180, Berlin, Heidelberg. Springer.

He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep Residual Learning for Image Recognition. pages 770–778.

Higgins, I., Matthey, L., Pal, A., Burgess, C., Glorot, X., Botvinick, M., Mohamed, S., and Lerchner, A. (2016). beta-VAE: Learning Basic Visual Concepts with a Constrained Variational Framework.

Hinton, G. and Sejnowski, T. J. (1999). *Unsupervised Learning: Foundations of Neural Computation*. MIT Press. Google-Books-ID: yj04Y0lje4cC.

Hopkins, W. D., Meguerditchian, A., Coulon, O., Bogart, S., Mangin, J.-F., Sherwood, C. C., Grabowski, M. W., Bennett, A. J., Pierre, P. J., Fears, S., Woods, R., Hof, P. R., and Vauclair, J. (2014). Evolution of the central sulcus morphology in primates. *Brain, Behavior and Evolution*, 84(1):19–30.

Hornik, K., Stinchcombe, M., and White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5):359–366.

Hotier, S., Leroy, F., Boisgontier, J., Laidi, C., Mangin, J.-F., Delorme, R., Bolognani, F., Czech, C., Bouquet, C., Toledano, E., Bouvard, M., Petit, J., Mishchenko, M., d'Albis, M.-A., Gras, D., Gaman, A., Scheid, I., Leboyer, M., Zalla, T., and Houenou, J. (2017). Social cognition in autism is associated with the neurodevelopment of the posterior superior temporal sulcus. *Acta Psychiatrica Scandinavica*, 136(5):517–525.

Huang, G., Liu, Z., van der Maaten, L., and Weinberger, K. Q. (2017). Densely Connected Convolutional Networks. pages 4700–4708.

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

Im, K., Jo, H. J., Mangin, J.-F., Evans, A. C., Kim, S. I., and Lee, J.-M. (2010). Spatial Distribution of Deep Sulcal Landmarks and Hemispherical Asymmetry on the Cortical Surface. *Cerebral Cortex*, 20(3):602–611.

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

Jaderberg, M., Simonyan, K., Zisserman, A., and kavukcuoglu, k. (2015). Spatial Transformer Networks. In Cortes, C., Lawrence, N., Lee, D., Sugiyama, M., and Garnett, R., editors, *Advances in Neural Information Processing Systems*, volume 28. Curran Associates, Inc.

Jin, K., Zhang, T., Shaw, M., Sachdev, P., and Cherbuin, N. (2018). Relationship Between Sulcal Characteristics and Brain Aging. *Frontiers in Aging Neuroscience*, 10.

Juhl, K. A., Morales, X., de Backer, O., Camara, O., and Paulsen, R. R. (2021). Implicit Neural Distance Representation for Unsupervised and Supervised Classification of Complex Anatomies. In de Bruijne, M., Cattin, P. C., Cotin, S., Padoy, N., Speidel, S., Zheng, Y., and Essert, C., editors, *Medical Image Computing and Computer Assisted Intervention – MICCAI 2021*, Lecture Notes in Computer Science, pages 405–415, Cham. Springer International Publishing.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A. W., Kavukcuoglu, K., Kohli, P., and Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873):583–589. Number: 7873 Publisher: Nature Publishing Group.

Kang, S. K., Shin, S. A., Seo, S., Byun, M. S., Lee, D. Y., Kim, Y. K., Lee, D. S., and Lee, J. S. (2021). Deep learning-Based 3D inpainting of brain MR images. *Scientific Reports*, 11(1):1673.

Kim, H., Bernasconi, N., Bernhardt, B., Colliot, O., and Bernasconi, A. (2008). Basal temporal sulcal morphology in healthy controls and patients with temporal lobe epilepsy. *Neurology*, 70(22 Part 2):2159–2165. Publisher: Wolters Kluwer Health, Inc. on behalf of the American Academy of Neurology Section: Articles.

Kingma, D. P. and Welling, M. (2014). Auto-Encoding Variational Bayes. *arXiv:1312.6114 [cs, stat]*. arXiv: 1312.6114.

Kochunov, P., Mangin, J.-F., Coyle, T., Lancaster, J., Thompson, P., Rivière, D., Cointepas, Y., Régis, J., Schlosser, A., Royall, D. R., Zilles, K., Mazziotta, J., Toga, A., and Fox, P. T. (2005). Age-related morphology trends of cortical sulci. *Human Brain Mapping*, 26(3):210–220. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/hbm.20198.

Kriegel, H.-P., Schubert, M., and Zimek, A. (2008). Angle-based outlier detection in high-dimensional data. In *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD '08, pages 444–452, New York, NY, USA. Association for Computing Machinery.

Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. In Pereira, F., Burges, C. J., Bottou, L., and Weinberger, K. Q., editors, *Advances in Neural Information Processing Systems*, volume 25. Curran Associates, Inc.

Kullback, S. and Leibler, R. A. (1951). On Information and Sufficiency. *The Annals of Mathematical Statistics*, 22(1):79–86. Publisher: Institute of Mathematical Statistics.

Lebenberg, J., Mangin, J. F., Thirion, B., Poupon, C., Hertz-Pannier, L., Leroy, F., Adibpour, P., Dehaene-Lambertz, G., and Dubois, J. (2019). Mapping the asynchrony of cortical maturation in the infant brain: A MRI multi-parametric clustering approach. *NeuroImage*, 185:641–653.

LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature*, 521(7553):436–444. Number: 7553 Publisher: Nature Publishing Group.

LeCun, Y., Boser, B., Denker, J., Henderson, D., Howard, R., Hubbard, W., and Jackel, L. (1989). Handwritten Digit Recognition with a Back-Propagation Network. In Touretzky, D., editor, *Advances in Neural Information Processing Systems*, volume 2. Morgan-Kaufmann.

Lecun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324. Conference Name: Proceedings of the IEEE.

Levitt, J. G., Blanton, R. E., Smalley, S., Thompson, P., Guthrie, D., McCracken, J. T., Sadoun, T., Heinichen, L., and Toga, A. W. (2003). Cortical Sulcal Maps in Autism. *Cerebral Cortex*, 13(7):728–735.

Lewitus, E., Kelava, I., Kalinka, A. T., Tomancak, P., and Huttner, W. B. (2014). An Adaptive Threshold in Mammalian Neocortical Evolution. *PLOS Biology*, 12(11):e1002000. Publisher: Public Library of Science.

Li, C.-L., Sohn, K., Yoon, J., and Pfister, T. (2021a). CutPaste: Self-Supervised Learning for Anomaly Detection and Localization. In *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 9659–9669, Nashville, TN, USA. IEEE.

Li, Y., Hu, P., Liu, Z., Peng, D., Zhou, J. T., and Peng, X. (2021b). Contrastive Clustering. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(10):8547–8555. Number: 10.

Liu, F. T., Ting, K. M., and Zhou, Z.-H. (2008). Isolation Forest. In *2008 Eighth IEEE International Conference on Data Mining*, pages 413–422. ISSN: 2374-8486.

Liu, R., Subakan, C., Balwani, A. H., Whitesell, J., Harris, J., Koyejo, S., and Dyer, E. L. (2020). A Generative Modeling Approach for Interpreting Population-Level Variability in Brain Structure. In Martel, A. L., Abolmaesumi, P., Stoyanov, D., Mateus, D., Zuluaga, M. A., Zhou, S. K., Racoceanu, D., and Joskowicz, L., editors, *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, pages 257–266, Cham. Springer International Publishing.

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

Lohmann, G., von Cramon, D. Y., and Colchester, A. C. F. (2008). Deep sulcal landmarks provide an organizing framework for human cortical folding. *Cerebral Cortex (New York, N.Y.: 1991)*, 18(6):1415–1420.

Lukaszewicz, A., Cortay, V., Giroud, P., Berland, M., Smart, I., Kennedy, H., and Dehay, C. (2006). The Concerted Modulation of Proliferation and Migration Contributes to the Specification of the Cytoarchitecture and Dimensions of Cortical Areas. *Cerebral Cortex*, 16(suppl_1):i26–i34.

Malandain, G., Bertrand, G., and Ayache, N. (1993). Topological segmentation of discrete surfaces. *International Journal of Computer Vision*, 10(2):183–197.

Mangin, J.-F., Frouin, V., Bloch, I., Rigis, J., and Lopez-Krahe, J. (1995). From 3D magnetic resonance images to structural representations of the cortex topography using topology preserving deformations. *Journal of Mathematical imaging and Vision*, 5(4):297–318.

Mangin, J.-F., Le Guen, Y., Labra, N., Grigis, A., Frouin, V., Guevara, M., Fischer, C., Rivière, D., Hopkins, W. D., Régis, J., and Sun, Z. Y. (2019). "Plis de passage" Deserve a Role in Models of the Cortical Folding Process. *Brain Topography*, 32(6):1035–1048.

Mangin, J.-F., Lebenberg, J., Lefranc, S., Labra, N., Auzias, G., Labit, M., Guevara, M., Mohlberg, H., Roca, P., Guevara, P., Dubois, J., Leroy, F., Dehaene-Lambertz, G., Cachia, A., Dickscheid, T., Coulon, O., Poupon, C., Rivière, D., Amunts, K., and Sun, Z. (2016). Spatial normalization of brain images and beyond. *Medical Image Analysis*, 33:127–133.

Mangin, J.-F., Riviere, D., Cachia, A., Duchesnay, E., Cointepas, Y., Papadopoulos-Orfanos, D., Collins, D., Evans, A., and Regis, J. (2004a). Object-Based Morphometry of the Cerebral Cortex. *IEEE Transactions on Medical Imaging*, 23(8):968–982.

Mangin, J.-F., Rivière, D., Cachia, A., Duchesnay, E., Cointepas, Y., Papadopoulos-Orfanos, D., Scifo, P., Ochiai, T., Brunelle, F., and Régis, J. (2004b). A framework to study the cortical folding patterns. *NeuroImage*, 23:S129–S138.

Markou, M. and Singh, S. (2006). A Neural Network-Based Novelty Detector for Image Sequence Analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(10):1664–1677. Conference Name: IEEE Transactions on Pattern Analysis and Machine Intelligence.

Martínez-Martínez, M. , De Juan Romero, C., Fernández, V., Cárdenas, A., Götz, M., and Borrell, V. (2016). A restricted period for formation of outer subventricular zone defined by Cdh1 and Trnp1 levels. *Nature Communications*, 7(1):11812. Number: 1 Publisher: Nature Publishing Group.

McInnes, L., Healy, J., and Melville, J. (2020). UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. arXiv:1802.03426 [cs, stat].

McInnes, L., Healy, J., Saul, N., and Grossberger, L. (2018). UMAP: Uniform Manifold Approximation and Projection. *The Journal of Open Source Software*, 3(29):861.

Meissen, F., Wiestler, B., Kaissis, G., and Rueckert, D. (2022). On the Pitfalls of Using the Residual as Anomaly Score. In *Medical Imaging with Deep Learning*.

Mellerio, C., Labeyrie, M.-A., Chassoux, F., Daumas-Duport, C., Landre, E., Turak, B., Roux, F.-X., Meder, J.-F., Devaux, B., and Oppenheim, C. (2012). Optimizing MR Imaging Detection of Type 2 Focal Cortical Dysplasia: Best Criteria for Clinical Practice. *American Journal of Neuroradiology*,

33(10):1932–1938. Publisher: American Journal of Neuroradiology _eprint: http://www.ajnr.org/content/33/10/1932.full.pdf.

Mellerio, C., Roca, P., Chassoux, F., Danière, F., Cachia, A., Lion, S., Naggara, O., Devaux, B., Meder, J.-F., and Oppenheim, C. (2014). The Power Button Sign: A Newly Described Central Sulcal Pattern on Surface Rendering MR Images of Type 2 Focal Cortical Dysplasia. *Radiology*, 274(2):500–507. Publisher: Radiological Society of North America.

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

Menze, B. H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., Burren, Y., Porz, N., Slotboom, J., Wiest, R., Lanczi, L., Gerstner, E., Weber, M.-A., Arbel, T., Avants, B. B., Ayache, N., Buendia, P., Collins, D. L., Cordier, N., Corso, J. J., Criminisi, A., Das, T., Delingette, H., Demiralp, , Durst, C. R., Dojat, M., Doyle, S., Festa, J., Forbes, F., Geremia, E., Glocker, B., Golland, P., Guo, X., Hamamci, A., Iftekharuddin, K. M., Jena, R., John, N. M., Konukoglu, E., Lashkari, D., Mariz, J. A., Meier, R., Pereira, S., Precup, D., Price, S. J., Raviv, T. R., Reza, S. M. S., Ryan, M., Sarikaya, D., Schwartz, L., Shin, H.-C., Shotton, J., Silva, C. A., Sousa, N., Subbanna, N. K., Szekely, G., Taylor, T. J., Thomas, O. M., Tustison, N. J., Unal, G., Vasseur, F., Wintermark, M., Ye, D. H., Zhao, L., Zhao, B., Zikic, D., Prastawa, M., Reyes, M., and Van Leemput, K. (2015). The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). *IEEE Transactions on Medical Imaging*, 34(10):1993–2024. Conference Name: IEEE Transactions on Medical Imaging.

Milletari, F., Navab, N., and Ahmadi, S.-A. (2016). V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. In *2016 Fourth International Conference on 3D Vision (3DV)*, pages 565–571.

Muñoz-Ramírez, V., Kmetzsch, V., Forbes, F., Meoni, S., Moro, E., and Dojat, M. (2022). Subtle anomaly detection: Application to brain MRI analysis of de novo Parkinsonian patients. *Artificial Intelligence in Medicine*, 125:102251.

Naval Marimont, S. and Tarroni, G. (2021). Implicit Field Learning for Unsupervised Anomaly Detection in Medical Images. In de Bruijne, M., Cattin, P. C., Cotin, S., Padoy, N., Speidel, S., Zheng, Y., and Essert, C., editors, *Medical Image Computing and Computer Assisted Intervention – MICCAI 2021*, Lecture Notes in Computer Science, pages 189–198, Cham. Springer International Publishing.

Nordahl, C. W., Dierker, D., Mostafavi, I., Schumann, C. M., Rivera, S. M., Amaral, D. G., and Van Essen, D. C. (2007). Cortical Folding Abnormalities in

Autism Revealed by Surface-Based Morphometry. *The Journal of Neuroscience*, 27(43):11725–11735.

Oegema, R., Barakat, T. S., Wilke, M., Stouffs, K., Amrom, D., Aronica, E., Bahi-Buisson, N., Conti, V., Fry, A. E., Geis, T., Andres, D. G., Parrini, E., Pogledic, I., Said, E., Soler, D., Valor, L. M., Zaki, M. S., Mirzaa, G., Dobyns, W. B., Reiner, O., Guerrini, R., Pilz, D. T., Hehr, U., Leventer, R. J., Jansen, A. C., Mancini, G. M. S., and Di Donato, N. (2020). International consensus recommendations on the diagnostic work-up for malformations of cortical development. *Nature Reviews Neurology*, 16(11):618–635. Number: 11 Publisher: Nature Publishing Group.

Ono, M., Kubik, S., and Abarnathey, C. D. (1990). *Atlas of the Cerebral Sulci*. Thieme-Stratton Corp, Stuttgart : New York, 1er édition edition.

Pang, G., Shen, C., Cao, L., and Hengel, A. v. d. (2020). Deep Learning for Anomaly Detection: A Review. *arXiv:2007.02500 [cs, stat]*. arXiv: 2007.02500.

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., and Duchesnay, (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12(85):2825–2830.

Penfield, W. and Boldrey, E. (1937). Somatic Motor and Sensory Representation in the Cerebral Cortex of Man as Studied by Electrical Stimulation. *Brain*, 60(4):389–443.

Penttilä, J., Paillére-Martinot, M.-L., Martinot, J.-L., Mangin, J.-F., Burke, L., Corrigall, R., Frangou, S., and Cachia, A. (2008). Global and Temporal Cortical Folding in Patients With Early-Onset Schizophrenia. *Journal of the American Academy of Child & Adolescent Psychiatry*, 47(10):1125–1132.

Perrot, M., Rivière, D., and Mangin, J.-F. (2011). Cortical sulci recognition and spatial normalization. *Medical Image Analysis*, 15(4):529–550.

Pimentel, M. A. F., Clifton, D. A., Clifton, L., and Tarassenko, L. (2014). A review of novelty detection. *Signal Processing*, 99:215–249.

Pinaya, W. H. L., Mechelli, A., and Sato, J. R. (2018). Using deep autoencoders to identify abnormal brain structural patterns in neuropsychiatric disorders: A large-scale multi-sample study. *Human Brain Mapping*, 40(3):944–954.

Pinaya, W. H. L., Mechelli, A., and Sato, J. R. (2019). Using deep autoencoders to identify abnormal brain structural patterns in neuropsychiatric disorders: A large-scale multi-sample study. *Human Brain Mapping*, 40(3):944–954.

Pirnay, J. and Chai, K. (2021). Inpainting Transformer for Anomaly Detection. *arXiv:2104.13897 [cs]*. arXiv: 2104.13897.

Pizzagalli, F., Auzias, G., Yang, Q., Mathias, S. R., Faskowitz, J., Boyd, J. D., Amini, A., Rivière, D., McMahon, K. L., de Zubicaray, G. I., Martin, N. G., Mangin, J.-F., Glahn, D. C., Blangero, J., Wright, M. J., Thompson, P. M., Kochunov, P., and Jahanshad, N. (2020). The reliability and heritability of cortical folds and their genetic correlations across hemispheres. *Communications Biology*, 3(1):1–12. Number: 1 Publisher: Nature Publishing Group.

Plaze, M., Paillère-Martinot, M.-L., Penttilä, J., Januel, D., de Beaurepaire, R., Bellivier, F., Andoh, J., Galinowski, A., Gallarda, T., Artiges, E., Olié, J.-P., Mangin, J.-F., Martinot, J.-L., and Cachia, A. (2011). "Where Do Auditory Hallucinations Come From?"—A Brain Morphometry Study of Schizophrenia Patients With Inner or Outer Space Hallucinations. *Schizophrenia Bulletin*, 37(1):212–221.

Pron, A., Deruelle, C., and Coulon, O. (2021). U-shape short-range extrinsic connectivity organisation around the human central sulcus. *Brain Structure and Function*, 226(1):179–193.

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

Qiu, C., Li, A., Kloft, M., Rudolph, M., and Mandt, S. (2022). Latent Outlier Exposure for Anomaly Detection with Contaminated Data. Number: arXiv:2202.08088 arXiv:2202.08088 [cs].

Rakic, P. (1988). Specification of Cerebral Cortical Areas. *Science*, 241(4862):170–176.

Rao, H., Wang, S., Hu, X., Tan, M., Guo, Y., Cheng, J., Liu, X., and Hu, B. (2021). A Self-Supervised Gait Encoding Approach with Locality-Awareness for 3D Skeleton Based Person Re-Identification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 1–1. arXiv: 2009.03671.

Rapoport, J. L. and Gogtay, N. (2011). Childhood onset schizophrenia: support for a progressive neurodevelopmental disorder. *International Journal of Developmental Neuroscience*, 29(3):251–258.

Reillo, I., de Juan Romero, C., García-Cabezas, M. , and Borrell, V. (2011). A role for intermediate radial glia in the tangential expansion of the mammalian cerebral cortex. *Cerebral Cortex (New York, N.Y.: 1991)*, 21(7):1674–1694.

Richman, D. P., Stewart, R. M., Hutchinson, J., and Verne S. Caviness, J. (1975). Mechanical Model of Brain Convolutional Development. *Science*. Publisher: American Association for the Advancement of Science.

Rivière, D., Mangin, J.-F., Papadopoulos-Orfanos, D., Martinez, J.-M., Frouin, V., and Régis, J. (2002). Automatic recognition of cortical sulci of the human brain using a congregation of neural networks. *Medical Image Analysis*, 6(2):77–92.

Ronneberger, O., Fischer, P., and Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. In Navab, N., Hornegger, J., Wells, W. M., and Frangi, A. F., editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, Lecture Notes in Computer Science, pages 234–241, Cham. Springer International Publishing.

Ruff, L., Vandermeulen, R., Goernitz, N., Deecke, L., Siddiqui, S. A., Binder, A., Müller, E., and Kloft, M. (2018). Deep One-Class Classification. In *Proceedings of the 35th International Conference on Machine Learning*, pages 4393–4402. PMLR. ISSN: 2640-3498.

Régis, J., Mangin, J.-F., Ochiai, T., Frouin, V., Riviére, D., Cachia, A., Tamura, M., and Samson, Y. (2005). "Sulcal Root" Generic Model: a Hypothesis to Overcome the Variability of the Human Cortex Folding Patterns. *Neurologia medico-chirurgica*, 45(1):1–17.

Sabokrou, M., Khalooei, M., Fathy, M., and Adeli, E. (2018). Adversarially Learned One-Class Classifier for Novelty Detection. pages 3379–3388.

Sato, J., Suzuki, Y., Wataya, T., Nishigaki, D., Kita, K., Yamagata, K., Tomiyama, N., and Kido, S. (2022). Anatomy-aware Self-supervised Learning for Anomaly Detection in Chest Radiographs. arXiv:2205.04282 [cs].

Schlegl, T., Seeböck, P., Waldstein, S. M., Langs, G., and Schmidt-Erfurth, U. (2019). f-AnoGAN: Fast unsupervised anomaly detection with generative adversarial networks. *Medical Image Analysis*, 54:30–44.

Schlegl, T., Seeböck, P., Waldstein, S. M., Schmidt-Erfurth, U., and Langs, G. (2017). Unsupervised Anomaly Detection with Generative Adversarial Networks to Guide Marker Discovery. *arXiv:1703.05921 [cs]*. arXiv: 1703.05921.

Schölkopf, B., Platt, J. C., Shawe-Taylor, J., Smola, A. J., and Williamson, R. C. (2001). Estimating the support of a high-dimensional distribution. *Neural Computation*, 13(7):1443–1471.

Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., and Batra, D. (2020). Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. *International Journal of Computer Vision*, 128(2):336–359.

Severino, M., Geraldo, A. F., Utz, N., Tortora, D., Pogledic, I., Klonowski, W., Triulzi, F., Arrigoni, F., Mankad, K., Leventer, R. J., Mancini, G. M. S., Barkovich, J. A., Lequin, M. H., and Rossi, A. (2020). Definitions and classification of malformations of cortical development: practical guidelines. *Brain: A Journal of Neurology*, 143(10):2874–2894.

Sicca, F., Silengo, M., Parrini, E., Ferrero, G. B., and Guerrini, R. (2003). Subcortical band heterotopia with simplified gyral pattern and syndactyly. *American Journal of Medical Genetics Part A*, 119A(2):207–210. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/ajmg.a.20111.

Simarro Viana, J., de la Rosa, E., Vande Vyvere, T., Robben, D., Sima, D. M., and Investigators, C.-T. P. a. (2021). Unsupervised 3D Brain Anomaly Detection. In Crimi, A. and Bakas, S., editors, *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, pages 133–142, Cham. Springer International Publishing.

Smiti, A. (2020). A critical overview of outlier detection methods. *Computer Science Review*, 38:100306.

Sohn, K., Li, C.-L., Yoon, J., Jin, M., and Pfister, T. (2022). Learning and Evaluating Representations for Deep One-Class Classification. In *International Conference on Learning Representations*.

Springenberg, J. T., Dosovitskiy, A., Brox, T., and Riedmiller, M. A. (2015). Striving for Simplicity: The All Convolutional Net. In Bengio, Y. and LeCun, Y., editors, *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Workshop Track Proceedings*.

Sun, Z. Y., Cachia, A., Rivière, D., Fischer, C., Makin, T., and Mangin, J.-F. (2017). Congenital unilateral upper limb absence flattens the contralateral hand knob. In *Organisation for Human Brain Mapping*, Vancouver, Canada.

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

Sun, Z. Y., Perrot, M., Tucholka, A., Rivière, D., and Mangin, J.-F. (2009). Constructing a dictionary of human brain folding patterns. *Medical image computing and computer-assisted intervention: MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention*, 12(Pt 2):117–124.

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

Sun, Z. Y., Rivière, D., Poupon, F., Régis, J., and Mangin, J. F. (2007). Automatic Inference of Sulcus Patterns Using 3D Moment Invariants. In Ayache, N., Ourselin, S., and Maeder, A., editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2007*, Lecture Notes in Computer Science, pages 515–522, Berlin, Heidelberg. Springer.

Tallinen, T., Chung, J. Y., Biggins, J. S., and Mahadevan, L. (2014). Gyrification from constrained cortical expansion. *Proceedings of the National Academy of Sciences*, 111(35):12667–12672. Publisher: Proceedings of the National Academy of Sciences.

Tallinen, T., Chung, J. Y., Rousseau, F., Girard, N., Lefèvre, J., and Mahadevan, L. (2016). On the growth and form of cortical convolutions. *Nature Physics*, 12(6):588–593. Number: 6 Publisher: Nature Publishing Group.

Tian, Y., Pang, G., Liu, F., chen, Y., Shin, S. H., Verjans, J. W., Singh, R., and Carneiro, G. (2021). Constrained Contrastive Distribution Learning for Unsupervised Anomaly Detection and Localisation in Medical Images. *arXiv:2103.03423 [cs]*. arXiv: 2103.03423.

Tissier, C., Linzarini, A., Allaire-Duquette, G., Mevel, K., Poirel, N., Dollfus, S., Etard, O., Orliac, F., Peyrin, C., Charron, S., Raznahan, A., Houdé, O., Borst, G., and Cachia, A. (2018). Sulcal Polymorphisms of the IFC and ACC Contribute to Inhibitory Control Variability in Children and Adults. *eNeuro*, 5(1):ENEURO.0197–17.2018.

Toro, R. and Burnod, Y. (2005). A Morphogenetic Model for the Development of Cortical Convolutions. *Cerebral Cortex*, 15(12):1900–1913.

Tovar-Moll, F., Moll, J., de Oliveira-Souza, R., Bramati, I., Andreiuolo, P. A., and Lent, R. (2007). Neuroplasticity in Human Callosal Dysgenesis: A Diffusion Tensor Imaging Study. *Cerebral Cortex*, 17(3):531–541.

Tschuchnig, M. E. and Gadermayr, M. (2021). Anomaly Detection in Medical Imaging – A Mini Review. *arXiv:2108.11986 [cs, eess]*. arXiv: 2108.11986.

van den Oord, A., Vinyals, O., and kavukcuoglu, k. (2017). Neural Discrete Representation Learning. In *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc.

van der Maaten, L. and Hinton, G. (2008). Visualizing High-Dimensional Data Using t-SNE. *Journal of Machine Learning Research*, 9(nov):2579–2605.

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

Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E. J., Yacoub, E., Ugurbil, K., and WU-Minn HCP Consortium (2013). The WU-Minn Human Connectome Project: an overview. *NeuroImage*, 80:62–79.

Vincent, P., Larochelle, H., Bengio, Y., and Manzagol, P.-A. (2008). Extracting and composing robust features with denoising autoencoders. In *Proceedings of the 25th international conference on Machine learning*, ICML '08, pages 1096–1103, New York, NY, USA. Association for Computing Machinery.

Vinh, N. X., Epps, J., and Bailey, J. (2010). Information Theoretic Measures for Clusterings Comparison: Variants, Properties, Normalization and Correction for Chance. *Journal of Machine Learning Research*, 11(95):2837–2854.

Vétil, R., Abi-Nader, C., Bône, A., Vullierme, M.-P., Rohé, M.-M., Gori, P., and Bloch, I. (2022). Learning Shape Distributions from Large Databases of Healthy Organs: Applications to Zero-Shot and Few-Shot Abnormal Pancreas Detection. In Wang, L., Dou, Q., Fletcher, P. T., Speidel, S., and Li, S., editors, *Medical Image Computing and Computer Assisted Intervention – MICCAI 2022*, Lecture Notes in Computer Science, pages 464–473, Cham. Springer Nature Switzerland.

Wachinger, C., Golland, P., Kremen, W., Fischl, B., and Reuter, M. (2015). BrainPrint: A discriminative characterization of brain morphology. *NeuroImage*, 109:232–248.

Wang, H., Bah, M. J., and Hammad, M. (2019). Progress in Outlier Detection Techniques: A Survey. *IEEE Access*, 7:107964–108000. Conference Name: IEEE Access.

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

White, L. E., Andrews, T. J., Hulette, C., Richards, A., Groelle, M., Paydarfar, J., and Purves, D. (1997). Structure of the human sensorimotor system. I: Morphology and cytoarchitecture of the central sulcus. *Cerebral Cortex (New York, N.Y.: 1991)*, 7(1):18–30.

Xiang, T., Lu, Y., Yuille, A. L., Zhang, C., Cai, W., and Zhou, Z. (2021). In-painting Radiography Images for Unsupervised Anomaly Detection. *arXiv:2111.13495 [cs]*. arXiv: 2111.13495.

Xu, G., Knutsen, A. K., Dikranian, K., Kroenke, C. D., Bayly, P. V., and Taber, L. A. (2010). Axons Pull on the Brain, But Tension Does Not Drive Cortical Folding. *Journal of Biomechanical Engineering*, 132(7).

Yousry, T. A., Schmid, U. D., Alkadhi, H., Schmidt, D., Peraud, A., Buettner, A., and Winkler, P. (1997). Localization of the motor hand area to a knob on the precentral gyrus. A new landmark. *Brain*, 120(1):141–157.

Yu, J., Oh, H., Kim, M., and Kim, J. (2021). Normality-Calibrated Autoencoder for Unsupervised Anomaly Detection on Data Contamination.

Yu, T. W., Mochida, G. H., Tischfield, D. J., Sgaier, S. K., Flores-Sarnat, L., Sergi, C. M., Topçu, M., McDonald, M. T., Barry, B. J., Felie, J. M., Sunu, C., Dobyns, W. B., Folkerth, R. D., Barkovich, A. J., and Walsh, C. A. (2010). Mutations in WDR62, encoding a centrosome-associated protein, cause microcephaly with simplified gyri and abnormal cortical architecture. *Nature Genetics*, 42(11):1015–1020. Number: 11 Publisher: Nature Publishing Group.

Yücel, M., Stuart, G. W., Maruff, P., Velakoulis, D., Crowe, S. F., Savage, G., and Pantelis, C. (2001). Hemispheric and Gender-related Differences in the Gross Morphology of the Anterior Cingulate/Paracingulate Cortex in Normal Volunteers: An MRI Morphometric Study. *Cerebral Cortex*, 11(1):17–25.

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

Zavrtanik, V., Kristan, M., and Skočaj, D. (2021). Reconstruction by inpainting for visual anomaly detection. *Pattern Recognition*, 112:107706.

Zenati, H., Foo, C. S., Lecouat, B., Manek, G., and Chandrasekhar, V. R. (2019). Efficient GAN-Based Anomaly Detection. arXiv:1802.06222 [cs, stat].

Zimmerer, D., Isensee, F., Petersen, J., Kohl, S., and Maier-Hein, K. (2019a). Unsupervised Anomaly Localization Using Variational Auto-Encoders. In Shen, D., Liu, T., Peters, T. M., Staib, L. H., Essert, C., Zhou, S., Yap, P.-T., and Khan, A., editors, *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019*, Lecture Notes in Computer Science, pages 289–297, Cham. Springer International Publishing.

Zimmerer, D., Kohl, S., Petersen, J., Isensee, F., and Maier-Hein, K. (2019b). Context-encoding Variational Autoencoder for Unsupervised Anomaly Detection. In *Medical Imaging with Deep Learning - MIDL 2019 - accepted as a short paper*.

Zimmerer, D., Kohl, S. A. A., Petersen, J., Isensee, F., and Maier-Hein, K. H. (2018). Context-encoding Variational Autoencoder for Unsupervised Anomaly Detection. arXiv:1812.05941 [cs, stat].

Zimmermann, R. S., Sharma, Y., Schneider, S., Bethge, M., and Brendel, W. (2021). Contrastive Learning Inverts the Data Generating Process. In *Proceedings of the 38th International Conference on Machine Learning*, pages 12979–12990. PMLR. ISSN: 2640-3498.