Multi-modal brain fingerprinting: A manifold approximation based framework

Kuldeep Kumar [a,c,*], Matthew Toews [a], Laurent Chauvin [a], Olivier Colliot [b,c,d], Christian Desrosiers [a]

[a] Laboratory for Imagery, Vision and Artificial Intelligence, École de technologie supérieure, 1100 Notre-Dame W., Montreal, QC, H3C1K3, Canada
[b] Sorbonne Universités, UPMC Univ Paris 06, Inserm, CNRS, Institut du cerveau et la moelle (ICM) - Hôpital Pitié-Salpêtrière, Boulevard de l'hôpital, F-75013, Paris, France
[c] Inria Paris, Aramis Project-Team, 75013, Paris, France
[d] AP-HP, Departments of Neurology and Neuroradiology, Hôpital Pitié-Salpêtrière, 75013, Paris, France

ABSTRACT

This work presents an efficient framework, based on manifold approximation, for generating brain fingerprints from multi-modal data. The proposed framework represents images as bags of local features which are used to build a subject proximity graph. Compact fingerprints are obtained by projecting this graph in a low-dimensional manifold using spectral embedding. Experiments using the T1/T2-weighted MRI, diffusion MRI, and resting-state fMRI data of 945 Human Connectome Project subjects demonstrate the benefit of combining multiple modalities, with multi-modal fingerprints more discriminative than those generated from individual modalities. Results also highlight the link between fingerprint similarity and genetic proximity, monozygotic twins having more similar fingerprints than dizygotic or non-twin siblings. This link is also reflected in the differences of feature correspondences between twin/sibling pairs, occurring in major brain structures and across hemispheres. The robustness of the proposed framework to factors like image alignment and scan resolution, as well as the reproducibility of results on retest scans, suggest the potential of multi-modal brain fingerprinting for characterizing individuals in a large cohort analysis.

1. Introduction

Despite sharing gross similarities, individual brains show a significant amount of variability (Gordon et al., 2017a) in terms of structure (Mangin et al., 2004), function (Barch et al., 2013; Gordon et al., 2017b; Mueller et al., 2013), and white matter architecture (Bürgel et al., 2006; de Schotten et al., 2011). Recently, various studies have focused on characterizing this variability using brain *fingerprints*, for instance, based on shape (Wachinger et al., 2015), functional connectivity (Finn et al., 2015; Liu et al., 2018), white matter fiber geometry (Kumar et al., 2017), or voxel-wise diffusion density (Yeh et al., 2016a). These studies are motivated by the fact that brain characteristics are largely determined by genetic factors that are often unique to individuals (Thompson et al., 2013). Moreover, various neurological disorders like Parkinson (Geevarghese et al., 2015) and autism (Goldman et al., 2013) have been linked to specific brain abnormalities that are difficult to describe at the population level. With the rapid improvements in MRI acquisition hardware and analysis tools, and thanks to large brain-related initiatives like the Human Connectome Project (HCP) (Van Essen et al., 2013) and UK Biobank (Sudlow et al., 2015), researchers are better poised to study individual subjects in addition to groups (Dubois and Adolphs, 2016; Gordon et al., 2017c), thus taking a step towards precision medicine (Hampel et al., 2017) and precision psychiatry (Finn and Constable, 2016).

The importance of brain fingerprinting is evident from the recent surge in studies on this topic. For example, Yeh et al. (2016a,b) built a local connectome fingerprint using dMRI data and applied this fingerprint to the analysis of genetically-related subjects. Kumar et al. (2017) proposed another dMRI-based fingerprint called *Fiberprint*, which characterizes white matter fiber geometry. Finn et al. (2015) considered the correlation between time courses of atlas-defined nodes to generate a functional connectivity profile and used this profile to identify

---

* Corresponding author. Laboratory for Imagery, Vision and Artificial Intelligence, École de technologie supérieure, 1100 Notre-Dame W., Montreal, QC, H3C1K3, Canada.

E-mail address: kkumar@livia.etsmtl.ca (K. Kumar).

individuals across scan sessions, both for task and rest conditions. Liu et al. (2018) use dynamic brain connectivity patterns for identifying individuals and predicting higher cognitive functions. Moreover, Miranda-Dominguez et al. (2014) proposed a linear model to describe the activity of brain regions in resting-state fMRI as a weighted sum of its functional neighboring regions. Their functional fingerprint, derived from the model's coefficients, has the ability to predict individuals using a limited number of non-sequential frames.

Various morphometry-based fingerprints have also been proposed for structural MRI modalities like T1- or T2-weighted images. For example, Wachinger et al. (2015) quantify the shape of cortical and subcortical structures via the spectrum of the Laplace-Beltrami operator. The resulting representation, called *Brainprint*, is used for subject identification and analyzing potential genetic influences on brain morphology. Toews et al. (2010) represent images as a collection of localized image descriptors and apply scale-space theory to analyze their distribution at the characteristic scale of underlying anatomical structures. This representation is employed to identify distinctive anatomical patterns of genetically-related individuals or subjects with a known brain disease.

So far, fingerprinting studies in the literature have focused on a single modality. However, each modality captures unique properties of the brain and combining multiple modalities can provide a richer, more discriminative information (Calhoun and Sui, 2016; Groves et al., 2012). Hence, the fusion of multiple modalities has been shown superior than single-modality data to identify diseases like schizophrenia, bipolar disorder, major depressive disorder and obsessive-compulsive disorder (Calhoun and Sui, 2016). Multi-modal neuroimaging biomarkers have also been proposed to predict cognitive deficits in schizophrenia (Sui et al., 2015). Similarly, the combination of multiple MRI modalities has led to the improved segmentation of isointense infant brain images (Zhang et al., 2015). Multi-modal imaging data can also be used to predict the brain-age of subjects and detect cognitive impairments (Liem et al., 2017). Detailed reviews on multi-modal methods and investigations for psychopathology can be found in Calhoun and Sui (2016) and Liu et al. (2015a,b).

Due to the challenges of combining multiple modalities in a single framework (Calhoun and Sui, 2016; Liu et al., 2015b), defining a multi-modal brain fingerprinting remains to this day an elusive task. Morphometry-based approaches, such as Brainprint (Wachinger et al., 2015), could potentially be extended to other modalities like dMRI. However, this requires solving non-trivial problems such as the cross-modality alignment of images with different resolutions, the segmentation and correspondence of neuroanatomical structures, etc. Computational efficiency is another important issue when dealing with large-scale, multi-subject and multi-modal datasets like the Human Connectome Project (HCP) (Van Essen et al., 2013) and UK Biobank (Sudlow et al., 2015). In this work, we propose a multi-modal brain fingerprinting that overcomes these challenges using manifold approximation. The key idea is to represent each image as a bag of local features, and derive a subject-level proximity graph using feature correspondences over the entire set of images (Toews et al., 2010). This subject proximity graph provides an approximation of the image appearance subspace (i.e., the manifold), which can be used to obtain a compact fingerprint representation.

Manifold learning has been extensively studied in machine learning (Bengio et al., 2013) with many approaches like Isomap (Tenenbaum et al., 2000), Locally Linear Embedding (LLE) (Roweis and Saul, 2000) and Spectral Embedding (Belkin and Niyogi, 2003)proposed over the years. As detailed in Aljabar et al. (2012), such techniques have also been used for various problems of medical imaging like registration, segmentation, and classification. For example, in Gerber et al. (2010), Gerber et al. use manifold learning to perform a population analysis of brain images. Similarly, a deep learning based approach is explored in Brosch et al. (2013) to learn the manifold of brain MRIs. A key factor in such methods is image representation. For instance, the manifold could be approximated using the Euclidean distance between image pairs,

however, this would not be robust to translation, rotation or scaling, and would suffer from high computational costs.

Representations based on local features, often referred to as bag of features (BoF), offer an efficient alternative for encoding and matching image structures without the stringent requirement of one-to-one correspondence (Lowe, 2004; Tsai, 2012). In brain imaging, BoFs have been shown to automatically identify known structural differences between healthy controls and Alzheimer's subjects in a fully data-driven fashion (Toews et al., 2010). They have also been used successfully to model the development of infant brains (Toews et al., 2012) and align images of different modalities (Toews and Wells, 2013). Despite their numerous advantages, BoFs have thus far not been explored for brain fingerprinting. This is mainly due to their large and variable size, which makes comparing two fingerprints non-trivial.

The key contributions of this work are as follows:

- **Novel framework:** We propose a data-driven approach based on BoFs and manifold approximation that combines the information from multiple imaging modalities into a common fingerprint. By embedding BoFs in a low-dimensional manifold, the proposed approach circumvents the problem of variable representation size, and provides a compact description of brain structure that enables efficient comparisons across subjects. Furthermore, we show how this manifold-based approach can be used to encode non-structural brain characteristics, for instance, modeling functional connectivity profiles from fMRI. To our knowledge, this is the first work to combine structural, diffusion, and functional modalities in a single fingerprint.
- **Large-scale analysis:** We present a comprehensive analysis of the proposed fingerprint using a large-scale dataset from the Human Connectome Project (HCP), where numerous properties/factors are investigated: fingerprint parameters (e.g., manifold dimensionality and proximity graph connectivity), contribution of individual modalities and/or their combination to the fingerprint's discriminativeness, robustness to image alignment and scan resolution, and reproducibility of results with re-test or corrupted scans. Using genetically verified zygosity labels from the HCP twin dataset, we also analyze the proposed fingerprint's ability to identify genetically-related subjects (i.e., monozygotic twins, dizygotic twins and non-twin siblings) from a large cohort, and show our multi-modal fingerprint to outperform single-modality approaches or fingerprints based on raw images. In an analysis of local feature correspondences, we identify for individual modalities the neuroanatomical regions having the most significant differences across groups of genetically-related subjects, between males and females, and across brain hemispheres.

This study extends our preliminary work in Toews and Wells (2016) and Kumar et al. (2018), where BoF representations were used to identify and compare subjects in a population. Here, we show how these variable-length representations can be converted to fixed-sized fingerprints via manifold embedding, and present an out-of-sample strategy to generate fingerprints for new subjects. While our previous work only considered structural and diffusion MRI data, the current study also investigates the benefit of including fMRI-based information, as well as different combinations of sMRI, dMRI, and fMRI data. Additionally, it offers a much deeper analysis where the impact of multiple factors like the inclusion of skull tissue, image alignment, and scan resolution is evaluated. The present study also complements the recent work of Colclough et al. (2017), which analyzes the heritability of functional connectivity profiles from multi-modal data (i.e., fMRI and MEG) using the HCP twin dataset. Unlike this recent work, our study analyzes the relationship between genetic proximity and fingerprint similarity based on a rank retrieval analysis and shows that a higher retrieval accuracy can be obtained when combining structural, diffusion, and functional imaging data.

The rest of this paper is organized as follows. We first present the
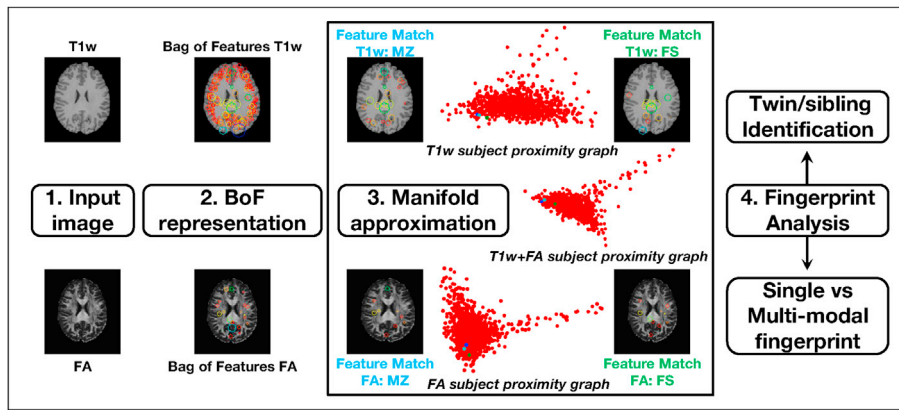
Fig. 1. Pipeline of the proposed framework and analysis. For a given input image, a BoF representation is first obtained by extracting local features. This representation is then converted to a fingerprint by matching features across the entire set of images, and embedding the resulting proximity graph into the manifold. The manifold approximation block shows the 2D embedding coordinates (i.e., fingerprint) of HCP subjects (red dots) obtained with T1w (top), FA (bottom) and combined T1w + FA (middle) images. The fingerprints of a specific subject (blue dot), his/her monozygotic twin (MZ, cyan dot) and full sibling (FS, green dot) are highlighted in each plot. The pairwise feature matches of these two sibling pairs, for T1w and FA images, are shown in the corner images of the block.

proposed multi-modal brain fingerprinting framework, detailing the data pre-processing steps, the BoF representation and proximity graph computation, as well as the manifold embedding of this graph. In Section 3, we then conduct an extensive experimental validation using the T1-weighted, T2-weighted, diffusion-weighted MRI, and resting-state fMRI data of 945 subjects from the HCP dataset. Finally, we conclude with a summary of our contributions and a discussion of possible extensions.

2. Materials and methods

Fig. 1 summarizes the pipeline of the proposed multi-modal brain fingerprint framework, which is comprised of four steps. In the first step, we start with pre-processed structural MRI (sMRI) and diffusion MRI (dMRI) data of 945 subjects from the Human Connectome Project (Van Essen et al., 2012, 2013). Diffusion Tensor Imaging (DTI) and Generalized Q-Ball Imaging (GQI) based Diffusivity measures are obtained from dMRI scans, including: fractional anisotropy (FA), axial diffusivity (AD), mean diffusivity (MD), radial diffusivity (RD) and generalized fractional anisotropy (GFA). The second step then extracts local features from the images of each subject and encodes subjects as a bag of features (BoF). In the third step, the multi-modal fingerprints of subjects are computed using manifold approximation. Towards this goal, a subject-level proximity graph is first constructed by matching the features of each modality across images and identifying pairs of subjects with a high number of correspondences. Fingerprints are then obtained by embedding this graph in a low-dimensional subspace. In the last step, we perform various analyses on the subject fingerprints. The informativeness of individual modalities and their link to genetic proximity are first measured in a twin/sibling identification analysis. This analysis is then extended to multi-modal fingerprints, showing the combined effect and complementarity of multiple modalities. Resting state fMRI network matrices and FreeSurfer derived measures of volume, thickness, and area provided by HCP are also used for fingerprint analysis. Finally, the distribution of feature correspondences between pairs of subjects is used to identify regions showing significant differences across different sibling types. The following subsections describe each of these steps in greater detail.

2.1. Data and pre-processing

We used the pre-processed structural and diffusion MRI data, and the resting state fMRI network matrices of 945 subjects from the HCP1200 release of the Human Connectome Project (Van Essen et al., 2013). The retest data of 42 subjects (out of 45 available) from the same dataset were also considered in our study to evaluate reproducibility. The HCP1200 release provides genetically-verified labels for twins and siblings, and is a rich resource to analyze the importance of environmental and genetic influences for traits, phenotypes, and disorders (Kochunov et al., 2015; Van Essen et al., 2012). Table 1 provides the demographic details of the

Table 1

Demographics. We considered the HCP1200 release subjects with structural MRI, diffusion MRI, and resting-state fMRI netmats data, and for which the HasGT field is true (genetically verified data). The family structure and links are obtained from the output of hcp2blocks.m script listed in data release manual. The sibship sizes are between 1 and 6.

Type	Total	Gender		Age	Handedness	
		F	M	Range (median)	L	R
All	945	501	444	22–36 (29)	84	861
MZ	238	138	100	22–36 (30)	19	219
DZ	126	70	56	22–35 (29)	13	113
NotTwin	581	293	288	22–36 (28)	52	529

subjects used in this study.

Data were acquired on a Siemens Skyra 3T scanner (Sotiropoulos et al., 2013) and pre-processed as described in Glasser et al. (2013). The structural acquisitions include high resolution T1-weighted (T1w) and T2-weighted (T2w) images (0.7 mm isotropic, FOV = 224 mm, matrix = 320, 256 sagittal slices in a single slab), the diffusion acquisition used following parameters: sequence = Spin-echo EPI; repetition time (TR) = 5520 ms; echo time (TE) = 89.5 ms; resolution = 1.25 × 1.25 × 1.25 mm^3 voxels, and the resting-state fMRI acquisition involved four 15 min runs at 2 mm isotropic resolution and a repetition time of 0.72 s (4800 vol per subject). Further details can be obtained from the HCP1200 data release manual.[1] We used the hcp2blocks.m script (described in the HCP1200 release) to generate a FamilyID based matrix, only considering subjects having dMRI, sMRI, and rfMRI netmats data, and for which the HasGT field is true. Using this selection criterion, we obtained a total of 238 monozygotic (MZ) subjects, 126 dizygotic (DZ) subjects, and 581 non-twin subjects. The sibship size ranged between 1 and 6. In a next step, using the mother ID, father ID, family ID and family type information from the output of hcp2blocks.m script, we obtained 119 monozygotic twin pairs, 63 dizygotic twin pairs, 546 full-sibling (FS) pairs, 39 maternal half sibling (MHS) pairs, and 5 paternal half sibling (PHS) pairs. These pairs are used for twin/sibling identification task in the following sections.

For structural MRI, we considered T1-weighted (0.7 mm) and T2-weighted (0.7 mm), with and without skull. The images are in native space and skull stripped, unless explicitly specified. In the case of dMRI data, signal reconstruction was performed with the freely available DSI Studio toolbox (Yeh et al., 2010) using the Diffusion Tensor Imaging (DTI) and Generalized Q-Ball Imaging (GQI) reconstruction options. Four widely used DTI-based measures were extracted to characterize white matter micro-structure: fractional anisotropy (FA), axial diffusivity (AD),

---

[1] https://www.humanconnectome.org/documentation/S1200/.

mean diffusivity (MD) and radial diffusivity (RD). The interpretation of these measures is discussed in Alexander et al. (2007). In addition, to utilize the multi-shell information and high angular resolution of the HCP data, Generalized Q-Ball Imaging (GQI) (Yeh et al., 2010) based measures including generalized fractional anisotropy (GFA) and quantitative anisotropy (QA) were also obtained. For resting state fMRI, we used the connectivity matrices (netmats), provided by the HCP 1200 release, derived using the FSLNets toolbox, either via full correlation or the partial correlation (Smith et al., 2015), the latter being calculated by inverting the covariance matrix. For analyzing the impact of alignment, we also used the MNI space aligned data for T1-weighted (0.7 mm) and T2-weighted (0.7 mm) provided by the HCP 1200 release. In addition, to combine structural modalities with dMRI, and to analyze the impact of scan resolution, we re-sampled T1- and T2-weighted images to a 1.25 mm resolution using the linear registration (FLIRT) tool of FSL (Jenkinson et al., 2012). Finally, our analysis also considered FreeSurfer derived measures of sub-cortical volumes, cortical thickness, and area, as well as T1w/T2w MRI ratio images (0.7 mm, myelin content information).

2.2. Multi-modal brain fingerprint

Generating brain fingerprints of subjects based on their multi-modal data involves multiple steps: extracting local descriptors in images to build a bag of features (BoF) representation of subjects, building a subject proximity graph by comparing their BoF representations, and embedding this graph in a low-dimensional manifold. Additionally, once the manifold has been constructed, an out-of-sample extension strategy is required to compute the fingerprint of new subjects.

2.2.1. Bag of feature (BoF) representation of subjects

In the first step, a set of local descriptors (Lowe, 2004) is obtained from each available image (3D scan). Various local feature extraction and representation approaches (Tuytelaars et al., 2008) can be used, for example, Scale Invariant Feature Transform (SIFT) (Lowe, 1999) or Speeded UP Robust Features (SURF) (Bay et al., 2006). In this work, we use 3D SIFT descriptors as they have been well studied for neuro-image analysis (Toews et al., 2010, 2015; Toews and Wells, 2013) and can be computed efficiently.

3D keypoints are located in the scans of each subject by finding the local extrema (i.e., maxima or minima) of the difference of Gaussians (DoG) occurring at multiple scales. Keypoints with a low contrast or corresponding to edge response are discarded, and remaining ones are encoded into a feature vector (i.e, the descriptor) using the histogram of oriented gradients (HOG) within a small neighborhood. Note that these descriptors are robust to changes in illumination, scale, and rotation, and are thus efficient for comparing images acquired using different scanners or imaging parameters. Each subject is then represented as an orderless bag of features (BoF), containing all the descriptors found in this subject's scans. This representation provides a simple, robust and extensible way of incorporating data from multiple modalities.

2.2.2. Subject proximity graph

Because the BoFs of two subjects may contain different numbers of descriptors, they are difficult to compare directly. To circumvent this problem, we construct an intrinsic manifold of subject appearance using a nearest-neighbor (NN) graph in feature space. In this graph, each descriptor is represented by a node and is connected to its K most similar appearance descriptors based on Euclidean distance. The K-nearest neighbors of each descriptor can be computed in sublinear time, for example, using randomized KD-search trees (Muja and Lowe, 2009). This feature graph is then used to induce a subject proximity graph by considering, for each pair of subjects, the number descriptors in their BoF that are linked in the feature graph.

Let $\mathscr{B}_i^m$ and $\mathscr{B}_j^m$ be the BoFs (i.e., set of descriptors) of subjects i and j for modality $m \in \mathscr{M}$, where $\mathscr{M}$ is the set of available modalities. The

similarity between these subjects is evaluated as

$$s_{ij} = \frac{\sum_{m \in \mathscr{M}} \left| \mathscr{B}_i^m \cap \mathscr{B}_j^m \right|}{\sum_{m \in \mathscr{M}} \left| \mathscr{B}_i^m \cup \mathscr{B}_j^m \right|} = \frac{\sum_{m \in \mathscr{M}} \left| \mathscr{B}_i^m \cap \mathscr{B}_j^m \right|}{\sum_{m \in \mathscr{M}} \left(\left| \mathscr{B}_i^m \right| + \left| \mathscr{B}_j^m \right| - \left| \mathscr{B}_i^m \cap \mathscr{B}_j^m \right| \right)}, \quad (1)$$

where $\left| \mathscr{B}_i^m \cap \mathscr{B}_j^m \right|$ is the number of edges in the feature graph between nodes in $\mathscr{B}_i^m$ and those in $\mathscr{B}_j^m$. When using a single modality, this measure corresponds to the well-known Jaccard similarity. Here, we extend it to a multi-modal setting by comparing the descriptors of each modality separately. We note that the Jaccard distance, defined as one minus the Jaccard similarity, is a *metric* and thus well-suited for constructing the manifold.

When defining the feature graph, K determines the number of nearest-neighbor connections for each descriptor. In manifold learning approaches, this parameter controls the locality of the manifold approximation at each point (Bengio et al., 2013). Its value should be large enough to capture the manifold's local structure, but also restricted so that distances to nearest-neighbors are close to the geodesic. In our experiments, we tested $K \in \{10, 20, 30, 40, 50\}$ and found similar results for these values. In what follows, we report results obtained with $K = 20$.

2.2.3. Manifold embedding

A manifold embedding technique is used to obtain compact brain fingerprints from the subject proximity graph. While various approaches could be employed for this task, for instance, Isomap (Tenenbaum et al., 2000) or locally linear embedding (LLE) (Roweis and Saul, 2000), we performed the embedding using Laplacian Eigenmaps (Belkin and Niyogi, 2003). This technique, which is connected to the well-known Laplace-Beltrami operator, has the advantage of being efficient and allowing out-of-sample extensions.

In Laplacian eigenmaps, each subject i is mapped to a coordinate vector $\mathbf{x}_i \in \mathbb{R}^k$ of the manifold, whose dimension k is a user parameter. The embedding of subjects in the manifold is made such that two subjects i and j with a high similarity s_{ij} will be close to one another. Let $\mathbf{S} \in \mathbb{R}^{n \times n}$ be the adjacency matrix of the subject proximity graph, as defined in Eq. (1), and denote as $\mathbf{L} = \mathbf{D} - \mathbf{S}$ the Laplacian of $\mathbf{S}$, where $\mathbf{D}$ is a diagonal matrix containing the row sums of $\mathbf{S}$. The embedding is accomplished by solving the following problem:

$$arg\,min_{\mathbf{X}} \sum_{i=1}^n \sum_{j=1}^n s_{ij} \left\| \mathbf{x}_i - \mathbf{x}_j \right\|_2^2 = \mathrm{tr}(\mathbf{X}^\mathsf{T} \mathbf{L} \mathbf{X}), \quad \text{s.t. } \mathbf{X}^\mathsf{T} \mathbf{D} \mathbf{X} = \mathbf{I}. \quad (2)$$

The constraint on $\mathbf{X}$ removes an arbitrary scaling factor in the embedding. As described in Belkin and Niyogi (2003), the solution to this problem is given by the leading k eigenvectors of the normalized adjacency matrix $\overline{\mathbf{S}} = \mathbf{D}^{-\frac{1}{2}} \mathbf{S} \mathbf{D}^{-\frac{1}{2}}$, starting from the second one.[2] Once computed, the rows of the matrix $\mathbf{X}$ correspond to the n subject fingerprints of size k.

2.2.4. Out-of-sample extension

The manifold embedding technique described above computes the fingerprint of all subjects at once. If new subjects are added, this process must be repeated over again, which is inefficient and changes the fingerprint of previous subjects. To alleviate these problems, we use an out-of-sample extension of Laplacian eigenmaps, based on the Nystrom method (Bengio et al., 2004; Fowlkes et al., 2004).

Suppose we want to compute the manifold embedding of m new subjects. The first step is to update the nearest-neighbor feature graph with the local descriptors of these new subjects, leaving unchanged the nearest-neighbors of the n base subjects. We then evaluate the pairwise similarities between new subjects and the base ones. Let $\mathbf{P} \in \mathbb{R}^{n \times m}$ be the

---

[2] The first eigenvector contains constant values.

matrix containing these similarities, the adjacency matrix of the extended subject proximity graph $\mathbf{S}_{\text{ext}} \in \mathbb{R}^{(n+m)\times(n+m)}$ is given by

$$\mathbf{S}_{\text{ext}} = \begin{bmatrix} \mathbf{S} & \mathbf{P} \\ \mathbf{P}^{\intercal} & \mathbf{Q} \end{bmatrix}. \tag{3}$$

Using the formula in Belkin and Niyogi (2003), the matrix $\mathbf{Q}$ of similarities between new subjects can be approximated as $\mathbf{P}^{\intercal}\mathbf{S}^{-1}\mathbf{P}$.

To normalize $\mathbf{S}_{\text{ext}}$, we compute the vector of row sums

$$\mathbf{d}_{\text{ext}} = \begin{bmatrix} \mathbf{s}_r + \mathbf{p}_r \\ \mathbf{p}_c + \mathbf{P}^{\intercal}\mathbf{S}^{-1}\mathbf{p}_r \end{bmatrix}, \tag{4}$$

where $\mathbf{s}_r, \mathbf{p}_r \in \mathbb{R}^n$ contain the row sums of $\mathbf{S}$ and $\mathbf{P}$, respectively, and $\mathbf{p}_c \in \mathbb{R}^m$ contains the column sum of $\mathbf{P}$. In the case where m is small compared to n, we have that $\mathbf{s}_r \approx \mathbf{s}_r + \mathbf{p}_r$, and thus $\mathbf{d}_{\text{ext}}$ can be approximated as

$$\mathbf{d}_{\text{ext}} \approx \begin{bmatrix} \mathbf{s}_r \\ \mathbf{p}_c + \mathbf{P}^{\intercal}\mathbf{S}^{-1}\mathbf{p}_r \end{bmatrix}. \tag{5}$$

This strategy, used in O'Donnell and Westin (2007) for white matter fiber segmentation, allows preserving the previous embedding of base subjects. Let $\mathbf{D}_{\text{ext}}$ be the diagonal matrix with entries corresponding to $\mathbf{d}_{\text{ext}}$, the normalized adjacency matrix of the extended subject graph is calculated as $\overline{\mathbf{S}}_{\text{ext}} = \mathbf{D}_{\text{ext}}^{-\frac{1}{2}} \mathbf{S}_{\text{ext}} \mathbf{D}_{\text{ext}}^{-\frac{1}{2}}$. The extended embedding is then obtained following Nystrom's method as

$$\mathbf{X}_{\text{ext}} = \begin{bmatrix} \mathbf{U} \\ \overline{\mathbf{P}}^{\intercal}\mathbf{U}\Lambda^{-1} \end{bmatrix}, \tag{6}$$

where $\mathbf{U}\Lambda\mathbf{U}^{\intercal}$ is the eigen decomposition of $\overline{\mathbf{S}}$, and $\overline{\mathbf{P}}$ is the normalized submatrix in $\overline{\mathbf{S}}_{\text{ext}}$ corresponding to $\mathbf{P}$. Hence, the embedding of base subjects is the same as in Section 2.2.3, and new subjects are embedded as $\overline{\mathbf{P}}^{\intercal}\mathbf{U}\Lambda^{-1}$. Once more, a fingerprint of size k is obtained by considering only the k leading eigenvectors in matrix $\mathbf{U}$, ignoring the constant eigenvector.

2.3. Computational efficiency

Computational and memory requirements are key factors when performing large-scale analyses. In this section, we evaluate these requirements for the main steps of the proposed framework. To highlight the efficiency of encoding images with local descriptors, we also compare our framework to a simple fingerprint using full images as features. Other aspects like scan resolution and image alignment requirements are discussed in Section 3. All experiments were performed on a 3.6 GHz processor with 32 GB RAM.

For the BoF representation of images, we extracted 3D SIFT features using a publicly available tool.[3] Computing these features took about 3 s per image, and approximately 60 min for all 945 images, when processed sequentially. This runtime could, however, be reduced significantly by processing images in parallel. The feature matching routine (Muja and Lowe, 2009), for generating the subject proximity graph from the BoFs of all images, required around 5 min to complete. In comparison, calculating the sum of squared distances (SSD) between full images took 1.7 s on average for a single pair, and 760,000 s for all $(945 \times 944)/2 = 446$, 040 pairs (with parallel computations). In terms of memory, each BoF file is approximately 400 KB in size, compared to 84 MB on average for a NIfTI volume file. In summary, the proposed framework is highly efficient in terms of computational and memory requirements compared to a baseline fingerprint using full images. Moreover, since computing the subject proximity graph has a complexity in $O(N\log N)$ where N is the number of images, and because extending the manifold embedding can be done efficiently using the Nystrom method, our framework is scalable to large datasets.

2.4. Evaluation measures

To measure the link between fingerprint similarity and genetic proximity, we performed a rank retrieval analysis using the sibling information provided in the HCP dataset. In this analysis, we try to identify the twins/siblings of a given subject by comparing its fingerprint with that of all other subjects in the group. Another goal of this analysis is to provide a common platform for the quantitative comparison of individual modalities and their combination. Two standard performance metrics for rank retrieval are used to evaluate the fingerprints: mean recall@k and mean average precision (MAP) (Turpin and Scholer, 2006).

Given a subject i, we rank all other subjects by the similarity (i.e., inverse of Euclidean distance) of their fingerprint to that of subject i. Denote as $\mathscr{T}_i$ the set of target siblings/twins of subject i. For instance, if the target group is non-twin siblings (NT), then $\mathscr{T}_i$ contains the siblings of subject i that are not his/her twin. Moreover, let $\mathscr{S}_i^k$ be the set containing the k subjects with fingerprints most similar to that of i (i.e., the k nearest neighbors of i). For a given value of k, we evaluate the retrieval performance using the measures of recall@k and precision@k:

$$(\text{recall@}k)_i = \frac{|\mathscr{T}_i \cap \mathscr{S}_i^k|}{|\mathscr{T}_i|}, \quad (\text{precision@}k)_i = \frac{|\mathscr{T}_i \cap \mathscr{S}_i^k|}{k}. \tag{7}$$

Mean recall@k, also known as sensitivity, evaluates the proportion of individuals that are genetically related to a given subject, which are within the k individuals most similar to that subject (in terms of fingerprint distance). When analyzing the rank performance for a particular sibling type (i.e., monozygotic twin, dizygotic twin or non-twin sibling), we average values over the set of subjects which have at least one sibling of this type, i.e. the set $\{i, \text{ s.t. } |\mathscr{T}_i| > 0\}$.

We also evaluate performance with the average precision, which extends the above metrics by considering the rank of nearest neighbors:

$$\text{AveP}_i = \frac{1}{|\mathscr{T}_i|} \sum_{k=1}^{n} (\text{precision@}k)_i \times \text{rel}_i(k), \tag{8}$$

where $\text{rel}_i(k)$ is an indicator function with the value equal to 1 if the k-th nearest neighbor of i is relevant (i.e., is in $\mathscr{T}_i$), and zero otherwise. The MAP is obtained by averaging AveP values over all subjects having at least one sibling of the target type.

Finally, we use the d-prime sensitivity index (Gale and Perkel, 2010) to obtain a quantitative measure of separability between the distribution of fingerprint distances corresponding to different sibling types. Let μ_1, μ_2 and σ_1, σ_2 be the means and standard deviations of compared distance distributions (e.g., the distance between monozygotic twins versus between dizygotic twins). The d-prime index is computed as

$$\text{d} - \text{prime} = \frac{\mu_1 - \mu_2}{\sqrt{\frac{1}{2}(\sigma_1^2 + \sigma_2^2)}}. \tag{9}$$

In our experiments, we report absolute values of d-prime, higher values indicating better separability.

3. Experiments and results

A comprehensive set of experiments was conducted to analyze the proposed fingerprint and evaluate its usefulness in various applications. In the first experiment, we analyze the manifold embedding of subjects and measure the impact of manifold dimensionality on the fingerprint's ability to capture genetic proximity. We then perform a detailed rank retrieval analysis, in which fingerprints obtained from a single modality or combinations of multiple modalities are used to identify three types of genetically-related subject: monozygotic twins (MZ), dizygotic twins (DZ) and full siblings (FS). The driving hypothesis of this experiment is

---

[3] http://www.matthewtoews.com/.

that individual modalities capture distinct properties of brain tissues, which can be effectively encoded in the fingerprint, and that combining different modalities can help describe the uniqueness of individual brains. Another goal of this experiment is to measure the relationship between the similarity of fingerprints, for different modality combinations, and genetic proximity.

In another experiment, we assess the impact of the following factors on the proposed fingerprint: image alignment, scan resolution, inclusion of skull, and subject age. This is followed by a reproducibility analysis, performed with the restest scans of 42 subjects, and a comparison with a baseline fingerprint using full images as features. The objective of these experiments is to demonstrate the robustness and performance of the proposed fingerprint, compared to a full image scan-based fingerprint.

We also present applications of the proposed framework for identifying retest scans, duplicate corrupt scans, and incorrectly-reported zygosity labels. In addition, we use the segmentation masks provided with the HCP data to identify cortical and subcortical brain regions where the distribution of feature correspondences between monozygotic twins is significantly different from dizygotic twins. In this analysis, we want to find brain regions which are more influenced by genetic proximity. Finally, we conduct a hemisphere asymmetry analysis using the feature correspondences for different types of siblings.

3.1. Manifold approximation analysis

To analyze the manifold approximation, we generated fingerprints by projecting the subject proximity graph onto a varying number of spectral components (i.e., leading eigenvectors of the normalized adjacency or Laplacian matrix). Fingerprints were normalized by converting each fingerprint to z-scores (centered to have mean 0 and scaled to have standard deviation 1). Fig. 2 (top row) shows a representative 2D spectral embedding of subject proximity graphs obtained using T1w, FA, or both modalities (T1w + FA). As described in Section 2.2.2, modalities are combined by aggregating the feature correspondences in each modality when computing the pairwise subject similarities. In these plots, the position of each red dot corresponds to the 2D fingerprint of a subject. Additionally, in each plot, a single pair of MZ twins is highlighted using blue and cyan dots and their NT sibling highlighted using a green dot.

It can be seen that the distribution of subject embeddings in the manifold varies from T1w to FA, showing that these modalities encode different properties in the fingerprint. Differences between the distributions of FA and T1w + FA fingerprints are in part explained by the fact

that spectral embeddings are defined up to a rotation or axis flipping. Moreover, we observe for all three modality combinations that genetically-related subjects are near to each other in the manifold, and that MZ twins are closer than their non-twin (full) sibling.

In Fig. 2 (bottom row), we measure the impact of manifold dimensionality on the fingerprint obtained with T1w, FA or T1w + FA modalities. The left plot shows the eigenvalues (sorted by decreasing magnitude) of the subject proximity graph's normalized adjacency matrix, which reflect the amount of connectivity information captured by their corresponding eigenvector. This plot indicates that most information is encoded in the first leading eigenvectors and, thus, that a compact fingerprint is possible.

This hypothesis is further confirmed in the middle and right plots of the same row, which evaluate for an increasing number of spectral components (i.e., fingerprint size) how the distributions of distances between MZ fingerprints and between DZ fingerprints differ. The separability between these two distributions of fingerprint distances is measured in terms of d-prime (middle plot) and unpaired t-test p-values (in $-\log_{10}$ scale). In both measures, higher values correspond to a greater separability. For all three modality combinations, a peak separability is observed around 150 eigenvectors, suggesting this value to be suitable for the fingerprint size. The decrease in separability for larger manifold dimensions is due to the fact that the added eigenvectors encode small variations of brain geometry which are not related to genetic proximity. Nevertheless, the difference between fingerprint distances in MZ pairs and in DZ pairs is significant with p-value < 0.01, for all tested manifold sizes and modality combinations.

Comparing the three modality combinations, diffusion-based fingerprints using FA images provide a higher separability than fingerprints generated from T1w, for all manifold sizes. For example, comparing Euclidean distances between MZ pair vs DZ pair fingerprints computed with 150 eigenvectors, we obtain absolute d-prime of 2.398 vs 1.848, and p-value (unpaired t-test, $-\log_{10}$ scale) of 47.676 vs 31.776. However, the separability is increased further when combining both modalities in the fingerprint (absolute d-prime of 2.627, and $-\log_{10}$ p-value of 50.962), in line with our hypothesis that multi-modal fingerprints are more discriminative than those based on a single modality. In addition, in a modality vs modality paired t-test comparing Euclidean distances between MZ pair fingerprints, we find $-\log_{10}$ p-values of 12.057 (T1w vs FA), 21.674 (FA vs T1w + FA), and 59.865 (T1w vs T1w + FA). A similar trend indicating the higher separability of multi-modal fingerprints was also observed for DZ and FS.
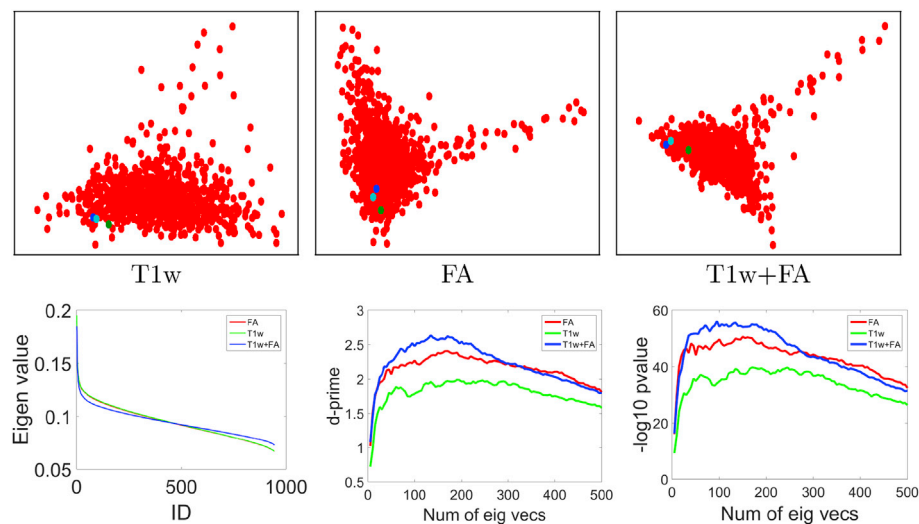
Fig. 2. Compact fingerprint analysis. Top row: representative 2D spectral embedding visualization, blue and cyan dots show one pair of MZ twins and green dot shows their not twin (full) sibling; Bottom row: plots of eigenvalues (excluding the first), absolute d-prime, and $-\log_{10}$ (p-value) (unpaired t-test) for Euclidean distances between MZ pair vs DZ pair fingerprints with an increasing number of eigenvectors.
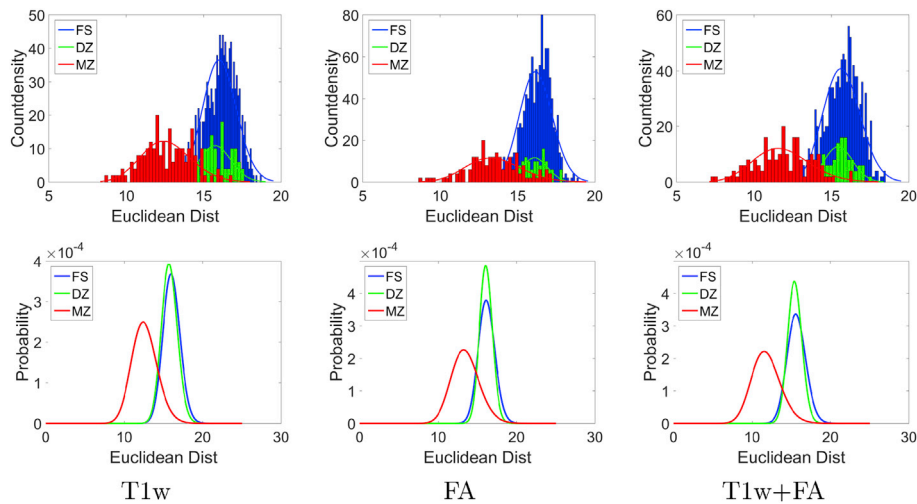
Fig. 3. Compact fingerprint comparison for genetically-related subjects. Count-density histograms (top row) and probability-normalized curves (bottom row; gamma histogram fitting) of Euclidean distances between twin/sibling pair fingerprints using 150 eigenvectors.

Finally, Fig. 3 gives the count histograms and probability density curves (fitted) of Euclidean distances between fingerprints of different sibling types. To generate these results, and in all following experiments, we used a fingerprint of 150 features (i.e., leading eigenvectors of the normalized adjacency matrix). It can be seen that the fingerprints of MZ twins, which share the most genetic material, are significantly more similar than those of DZ twins or full siblings (FS). This follows the results of various twin studies (Peper et al., 2007; Thompson et al., 2013), highlighting the relationship between genetic proximity and anatomical similarity. Comparison of compact fingerprints based on rfMRI (netmat), T1w + rfMRI, and T1w + T2w + FA + rfMRI is reported in Supplement material Fig. 1.

3.2. Identification of genetically-related subjects

In this section, we use genetically verified labels of the HCP dataset to determine whether fingerprints generated using different modality combinations can identify genetically-related individuals within a group of subjects. For combining structural and diffusion modalities, we considered data at 1.25 mm resolution. For resting state fMRI, we utilize the connectivity matrices (netmats) as functional connectivity fingerprints, and obtain the subject proximity graph (manifold approximation) by computing pairwise Pearson correlation. The idea is to closely follow the functional connectivity fingerprint and similarity computation described in Finn et al. (2015) (the parcellation and dataset sizes are different). The multi-modal combinations involving rfMRI are obtained by a linear combination of the rfMRI subject proximity graph with the graph derived from BoFs. Combination weights were determined by grid search, and optimal values of evaluation measures are reported. For measures based on FreeSurfer, we used the unrestricted *csv* file, considering the volume of sub-cortical structures, thickness and area measures for cortical regions. Each of the measures were converted to z-score across subjects and then used as a fingerprint (volume measures are first divided by FS-IntraCranial-Vol). Subject proximity graph is approximated by computing the pairwise Pearson correlation. We refer the reader to Section 2.4 for details on the evaluation protocol and measures.

Table 2 reports the mean average precision (MAP) values (@k = 944) obtained in a rank retrieval of three different siblings types (MZ, DZ and FS), using fingerprints generated from various combinations of the following modalities[4]: FA, MD, GFA, rfMRI netmat (partial

---

[4] DTI = FA + MD + RD + AD; rfMRI netmat = partial correlation and ICA-100.

Table 2
Mean average precision (MAP) obtained with different modality combinations for the identification of genetically-related subjects: monozygotic twins (MZ), dizygotic twins (DZ) and full siblings (FS).

Experiment	Modality	Mean Avg Prec		
		MZ	DZ	FS
sMRI	T1w	0.886	0.160	0.128
	T2w	0.908	0.212	0.111
dMRI	FA	0.964	0.219	0.160
	MD	0.803	0.114	0.086
	GFA	0.968	0.234	0.161
rfMRI	netmat	0.968	0.352	0.205
Modality Combination	T1w + T2w	0.970	0.283	0.183
	T1w + FA	0.977	0.279	0.210
	FA + MD	0.978	0.259	0.198
	T1w + rfMRI	0.990	0.460	0.279
	FA + rfMRI	0.996	0.472	0.301
	T1w + T2w + DTI	0.994	0.392	0.270
	T1w + T2w + FA + rfMRI	**0.997**	**0.546**	**0.371**
Skull Impact	T1w Skull	0.990	0.305	0.230
	T2w Skull	0.980	0.310	0.164
Alignment Impact	T1w MNI	0.852	0.087	0.101
	T2w MNI	0.827	0.147	0.111
Resolution Impact	T1w 1.25 mm	0.831	0.136	0.121
	T2w 1.25 mm	0.879	0.173	0.132
Baseline Comparison	T1w	0.649	0.079	0.052
	T2w	0.520	0.069	0.038
	FA	0.707	0.076	0.049
	Vol + Thck + Area (FreeSurfer)	0.795	0.172	0.106
Retest set	T1w	0.915	0.137	0.130
	T2w	0.917	0.212	0.113
	FA	0.944	0.252	0.158
Random	Rand	0.005	0.005	0.006

correlation, ICA 100), and FreeSurfer volume, thickness and area measures (Vol + Thck + Area). To lighten the presentation, we only report mean average precision (MAP) values, however, mean recall@k results can also be found in Supplement material (Fig. 2 and Table 8). Moreover, detailed results obtained with dMRI based measures (DTI and GQI), rfMRI netmats, and FreeSurfer measures are described in Table 3, 4, and 5 of Supplement material, respectively. The statistical significance of differences between MAP distributions obtained for different modality combinations and sibling types is reported in Table 1 of Supplement

material.

A rich and diverse set of observations can be drawn from Table 2. Comparing modalities, we observe that rfMRI netmat yields the highest MAP among all single-modality fingerprints, these improvements most significant for DZ and FS. For structure-based fingerprints, T1w and T2w provide similar performances across the different sibling types, slightly higher MAP values obtained for MZ and DZ using T2w. Similarly, for diffusion based fingerprints, FA and GFA provide similar performance, both of them outperforming MD. Furthermore, higher MAP values are obtained when combining multiple modalities, the combination of T1w, T2w, FA and rfMRI having the best performance for all sibling types. This applies for combinations within/across structural or diffusion modalities: T1w + T2w outperforms T1w and T2w, FA + MD performs better than FA and MD, T1w + FA outperforms T1w and FA, etc. Similarly, T1w + rfMRI outperforms T1w and rfMRI, and FA + rfMRI performs better than FA and rfMRI.

With respect to the tested sibling types, we observe MAP values between 80.3% and 99.7% when identifying MZ twins, for all modalities and their combinations. This illustrates the high impact of genetic similarity on the structural and diffusion geometry of the brain, as well as on its functional connectivity. Comparing all sibling types, we see higher MAP values for MZ twins compared to DZ twins or full siblings, following the amount of genetic information shared between subjects of these groups (Polderman et al., 2015). In contrast, performances obtained for DZ twins and full siblings are comparable, which reflects the fact that both sibling types have the same genetic proximity. In general, the differences between DZ twins and full siblings were found to be not significant in an unpaired *t*-test for single modalities, with T2w being the exception (Supplement material Table 1). Similar observations can also be drawn from mean recall@k plots and mean recall@10 values (Supplement material Fig. 2 and Table 8), with combined modalities yielding higher recall values than individual ones. In this experiment, FA gives a higher recall than rfMRI for MZ identification, although this difference is not statistically significant. Comparing non-twin siblings, we observe higher MAP values for full sibling identification vs maternal half sibling (MHS) identification (Supplement material Table 6), confirming once again the impact of genetic proximity. However, no clear trend is found for full sibling identification vs paternal half sibling identification (PHS), which is mainly due to the limited sample size (i.e., the dataset contains only 5 PHS pairs).

To quantify the informativeness of one modality versus another, Table 3 reports the relative percentage of MZ and DZ twins identified by both, a single, or none of the modalities.[5] Note that, for a given twin type, each row provides a relative comparison between two modalities, with the sum of the row being 100%. The total number of identification tasks is 238 for MZ and 126 for DZ (the identification of twin 1 for twin 2 and vice-versa are considered two separate tasks). For each task, we consider the $k = 10$ nearest neighbors of a subject in terms of fingerprint distance. The identification is considered a success if the subject's twin is identified within these neighbors. When comparing the relative success rates of single modalities (top half of the table), we observe that FA identifies more twins uniquely than when using T1w or MD. This is particularly noticeable for DZ twins, where 27.78% of DZ pairs were identified by the FA-based fingerprint but not the T1w-based ones. Yet, structural modalities still capture brain tissue properties that are not provided by dMRI, as shown by the 11.90% of all DZ pairs that are identified using T1w but not with FA. Similar observations can be drawn when comparing rfMRI with structural and diffusion modalities. For example, rfMRI identifies 45.24% of DZ pairs that are not identified using T1w within 10 neighbors, while T1w identifies 19.05% unique DZ pairs.

As with the results in Table 2, we see that combining multiple modalities leads to a more discriminative fingerprint. For example, 4.20% of MZ and 60.30% of DZ twins are identified by fingerprints generated from

all modalities (i.e., All MRI = T1w + T2w + FA + rfMRI) but not from fingerprints based only on T1w. Reversely, all MZ twins identified with T1w are also found using T1w + T2w + FA + rfMRI, and only 4.76% of DZ twins are identified uniquely with T1w. This last result could be explained by the fact that subjects can have local similarities due to factors not related to genetics.

3.3. Impact of various factors

Factors like image alignment, scan resolution, skull inclusion and subject age, can be confounds when analyzing the proposed fingerprint. In the following sub-sections, we measure the impact of these factors on the fingerprint's ability to find genetically-related subjects.

3.3.1. Image alignment
Population-level analyses usually require aligning images to a common space or segmenting them into regions of interest, two steps which can be computationally expensive.

Table 2 (sMRI vs alignment impact rows) reports the retrieval performance obtained for fingerprints generated from T1w and T2w images in MNI space (0.7 mm resolution, data provided by the HCP with affine alignment to MNI template). For all sibling types, MNI space-aligned fingerprints (denoted as MNI in the table) obtained lower MAP values than fingerprints using native space data. This observation, which is consistent across T1w/T2w modalities and all sibling types, indicates that image alignment is not required for our fingerprint. Note that similar results were obtained using full images as fingerprints (analyzed in the following section), with a lower MAP for affine-aligned images.

3.3.2. Scan resolution
Scan resolution is another important factor in multi-modal and multi-subject analyses, for example, sMRI data usually offer higher resolutions compared to dMRI.

Table 2 (sMRI vs resolution impact rows) shows that MAP values for the MZ/DZ twin identification task decrease when going from 0.7 mm to 1.25 mm resolution, for both T1w- and T2w-based fingerprints. This is due in part to the reduced number of SIFT features extracted from 1.25 mm resolution images, compared to 0.7 mm resolution ones. However, this is not the case for FS identification tasks, where contrasting trends are observed for T1w and T2w. Moreover, differences in MAP values for the two resolutions are not significant when running an unpaired *t*-test with p-value < 0.01, for any sibling type (see Supplement material Table 2). These results suggest the robustness of our framework to varying scan resolutions.

3.3.3. Inclusion of skull
Since skull size and shape is strongly influenced by genetics, including skull information in fingerprints could increase their discriminative power. In this experiment, we measure the usefulness of skull tissues for identifying pairs of MZ, DZ and FS subjects (facial features are not analyzed).

Table 2 reports the performances of fingerprints based on T1w and T2w image, with or without skull stripping. For both T1w and T2w, as well as all sibling types, including the skull in images improves MAP values. These results are significant with p-value < 0.01 in an unpaired *t*-test (see Table 2 of Supplement material). Hence, skull tissues provides additional feature correspondences which help identify twins and non-twin siblings. However, we should mention that skull stripping is essential to most neuroimaging analyses, and our objective here is only to measure the informativeness of skull tissues on the proposed fingerprint. An extended skull-inclusion analysis, including T1w-by-T2w MRI ratio images (myelin content) and modality combinations are reported in Supplement material Table 10.

3.3.4. Subject age
In twin studies, the age of subjects can be a confound when comparing

---

[5] Results for full siblings are reported in Table 7 of Supplement material.

Table 3
Relative informativeness of fingerprints from different modalities. Comparison between modalities or their combination for the task of identification of a given sibling type. The reported values are **relative percentages** of MZ/DZ twin identification for two modalities, with Mod1 representing successful identifications by modality 1 only. The total number of identification tasks is 238 and 126 for MZ and DZ respectively. Note: identification of twin 1 for twin 2 and vice-versa are considered two separate tasks. The identification is considered a success if the twin is identified within the 10 nearest neighbors of a subject (among 944 subjects).

Experiment	Mod1 vs Mod2	Identification % (MZ)				Identification % (DZ)			
		Both	Mod1	Mod2	None	Both	Mod1	Mod2	None
Single Modality	T1w vs T2w	93.28	2.52	**3.36**	0.84	12.70	13.49	**19.05**	54.76
	T1w vs FA	95.38	0.42	**3.78**	0.42	14.29	11.90	**27.78**	46.03
	T1w vs rfMRI	93.28	2.52	**4.20**	0.00	7.14	19.05	**45.24**	28.57
	FA vs rfMRI	96.64	**2.52**	0.84	0.00	26.19	15.87	**26.19**	31.75
	FA vs MD	88.66	**10.50**	0.84	0.00	10.32	**31.75**	12.70	45.24
Modality Combination	T1w vs All MRI	95.80	0.00	**4.20**	0.00	21.43	4.76	**60.32**	13.49
	T2w vs All MRI	96.64	0.00	**3.36**	0.00	25.40	6.35	**56.35**	11.90
	FA vs All MRI	99.16	0.00	**0.84**	0.00	39.68	2.38	**42.06**	15.87
	rfMRI vs All MRI	97.48	0.00	**2.52**	0.00	49.21	3.17	**32.54**	15.08

Note: All MRI = T1w + T2w + FA + rfMRI.

between different sibling types. For instance, DZ twins and FS siblings share the same amount of genetic material, yet DZ twins could be more similar due to their same age. The HCP data used in this study was acquired in the age range of 22–36, which corresponds to the plateau/saturation in brain and white matter development (Kochunov et al., 2015; Van Essen et al., 2012). Nevertheless, we analyze whether age differences in non-twin siblings is a contributing factor on performance.

Toward this goal, we divided FS sibling pairs in two groups based on the median age difference of 3 years, and measured the MAP in each group for fingerprints generated from T1w, T2w, and FA. Similarly, we also evaluated the impact of absolute age on performance of MZ/DZ. In this case, we divided subjects (not subject pairs) into two groups based on the median subject age of 29 years. As shown in Supplement material Table 9, in general, no significant differences in MAP are observed across these groups. In summary, using the HCP dataset, we found no significant impact of subject age on the proposed fingerprint.

3.4. Comparison to baseline fingerprint

We compared the performance of our fingerprint to a baseline using full images as features. In this baseline, the similarity of two fingerprints is measured as the sum of squared distances (SSD) between intensities of corresponding voxels. Table 2 gives the MAP obtained using this baseline, for T1w, T2w, and FA images in native subject space. For MZ twin identification, the baseline using FA performs better than T1w or T2w, which is consistent with the results of the proposed fingerprint. However, we see that our fingerprint performs consistently better than the baseline, with MAP improvements of 0.237 in T1w, 0.388 in T2w, and 0.257 in FA, for the task of identifying MZ twins. These improvements are significant in a one-sided unpaired t-test with p-value < 0.01 (see Supplement material Table 2). Note that we also tested a similar baseline created from MNI aligned images, however, this led to lower MAP values.

In addition, we used Freesurfer derived measures of sub-cortical volumes, and thickness and area of cortical regions as other baseline fingerprints (see Supplement material Table 5 for detailed analysis on FreeSurfer measures). Higher MAP values are obtained for MZ twin identification using our fingerprint vs Vol + Thck + Area FreeSurfer (0.886 vs 0.795, p-value < 0.01). However, no significant difference is observed for DZ and FS identification.

In summary, while much more compact and efficient (see Section 2.3), our fingerprint based on local features is significantly more informative than a voxel-based representation. It also captures additional information on brain morphology, compared to simple measures of cortical volume, thickness, and area, outperforming this baseline on all identification tasks.

3.5. Results reproducibility

To test the reproducibility of the results, we re-ran the same analysis after replacing the T1w, T2w and FA images of 42 subjects with their retest data. Table 2 gives the MAP values obtained following this process. We observe small differences in MAP, compared to fingerprints using the original data, however, these are not significant (see Supplement material Table 2).

We note that the majority of retest subjects available in the HCP data are MZ twins. Since we do not observe significant differences in identifying this type of twins, it shows that the results are reproducible. The small differences in MAP values for DZ twins and FS siblings could be attributed to slight changes in the ordering of retest subjects' nearest neighbors.

3.6. Applications

In this section, we demonstrate the usefulness of our fingerprint on three different applications: 1) the correction of erroneous zygosity labels, 2) the detection of retest and duplicate scans, 3) the visualization and analysis of local feature correspondences for different modalities, sibling types and neuroanatomical regions.

3.6.1. Zygosity label correction

The Q3 release of the HCP dataset contained self-reported zygosity labels for twins. In the HCP 1200 release, which contains genetically verified zyosity labels, it was found that many self-reported DZ twins were actually MZ twins. In light of this problem, we first evaluate if the proposed framework can be used to identify the twins in large dataset whose self-reported zygosity differs from their true zygosity.

In earlier experiments, we found higher MAP values for MZ twins. Such siblings were always found within the 10 nearest neighbors of a subject (i.e., a mean recall@k of 100% was obtained for $k \leq 10$, Supplement material Table 8), regardless of the modality combination used for the fingerprint. Conversely, a lower percentage of DZ twins could be identified in these lists of nearest neighbors. Based on this idea, we find incorrectly reported MZ candidates as the DZ twins which are within the 10 nearest neighbors of a subject.

Table 4 reports the relative percentage of DZ-to-MZ twins (56 in total) correctly identified by the proposed fingerprint, the baseline using full images, both or none of these methods, for T1w, T2w, and FA modalities. The results show that our fingerprint can identify most incorrectly self-reported MZ twins, with a detection rate of 92.86% for T1w, 100.00% for T2w, and 100.00% for FA. For all modalities, over 32% of cases were identified uniquely by our fingerprint. In contrast, no DZ-to-MZ twins

Table 4

Analysis of self-reported zygosity to genetically verified zygosity detection. The relative percentage of DZ-to-MZ twin identifications by the proposed framework and the full-image baseline. Total number of identification tasks is 56. Identification is considered a success if the twin is identified within the 10 nearest neighbors of a subject.

Modality	Identification %			
	Both	Proposed	Base	None
T1w	60.71	**32.15**	0.00	7.14
T2w	55.36	**44.64**	0.00	0.00
FA	64.29	**35.71**	0.00	0.00

were identified uniquely by the baseline fingerprint. In conclusion, the proposed fingerprint can be used effectively to detect misreported zygosity labels.

3.6.2. Retest and duplicate scan identification

To analyze our fingerprint's ability to detect repeat scans of the same subjects (acquired after a time gap), we used the data of 945 subjects + 42 retest subjects, and considered the task of identifying repeat scan in a rank retrieval analysis.

Following the same evaluation protocol as for identifying MZ/DZ/FS siblings, we obtained a MAP value of 1 for fingerprints generated from T1w, T2w or FA. Thus, in all cases, the single most similar fingerprint to that of a subject corresponded to this subject's retest data. Moreover, when considering the number of local feature correspondences in the subject similarity (i.e., $\sum_{m \in \mathcal{M}} \left| \mathscr{B}_i^m \cap \mathscr{B}_j^m \right|$ in Eq. (1)), we observed more correspondences for the retest data of a subject than for the subject's MZ twin.

Duplicate scans in a dataset, for example resulting from noise corruption, renaming or other manual errors, can introduce bias in analyses. Therefore, we also assessed if our fingerprint could detect duplicate scans of the same subject, corrupted by noise. For this experiment, we introduced duplicate scans for 42 T1w images, to which was added random noise (uniformly distributed random numbers in the $[-a, a]$ range, where $a \in \{20, 60, 100, 150, 200, 400\}$; the mean and stdev of image intensities are respectively 720 and 185). Running a rank retrieval analysis using these duplicates as target, we again obtained an MAP value of 1, for all tested noise levels. As in the retest scan identification task, the number of local feature correspondences was higher with corrupted duplicates than with images of MZ twins. Compared to retest scans, the number of feature correspondences was nearly half for corrupted duplicated, suggesting that noise can reduce correspondences to some extent. Overall, the results of this experiment demonstrate that our fingerprint can preserve brain characteristics over different scans of a subject.

3.6.3. Local feature correspondence analysis

To understand the advantages and limitations of a BoF-based fingerprint compared to voxel-wise or shape-based methods, we perform an in-depth analysis of local feature correspondences between subjects. In order to compare our findings with those of related fingerprint studies like Brainprint (Wachinger et al., 2015), we limit our analysis to genetically-related subjects from HCP and to structural MRI modalities. Other applications of BoF representations for neuro-image analysis have been well studied in the literature (Toews et al., 2010, 2015; Toews and Wells, 2013).

We start with a qualitative visualization of pairwise feature correspondences between subjects of different sibling types. The distribution of correspondences in these modalities is then analyzed using the segmentation maps (WM parcellation) files provided with HCP data. Furthermore, we also report cortical and subcortical regions having significantly different match distributions across sibling types, these regions having a closer relationship to genetic proximity. Finally, we perform a lateral asymmetry analysis in which the distribution of

correspondences in hemispheres are compared. Since fMRI is not required for these analyses, we considered all subjects in the HCP twin dataset having genetically verified labels (only 945 out of 1010 subjects have rfMRI netmats data), giving a total of 139 MZ pairs, 72 DZ pairs, and 1214 full sibling pairs.

3.6.3.1. Scale-space visualization of features correspondences. Analyzing local feature correspondences between sibling pairs provides information in terms of their location as well as scale. In 3D SIFT features, scale coincides with the variance of a Gaussian blur kernel for which the corresponding voxel in the blurred image is a local extrema (Lowe, 1999, 2004). It thus corresponds to a certain degree with the size of structures in which these features are located.

Fig. 4 gives a scale-space visualization of features matched between a subject and his/her MZ twin, as well as the subject's non-twin (full) sibling, for T1w, T2w and FA images (See Supplement material Fig. 3 for DZ and non-twin (full) sibling). The scale information is represented using the radius of circles. Note that circles represent the intersection of 3D spheres with the visible slice and, thus, non-intersecting features are hidden in this 2D visualization.

It can be seen that different image modalities generally result in distinct, complementary feature correspondences throughout the brain. In T1w and T2w images, features are mainly located in the frontal lobe, corpus-callosum, and cerebellum. Smaller-scale features are also visible along various cortical regions, as well as in subcortical structures near the basal ganglia. Moreover, images based on diffusion measures have less correspondences than in structural modalities. These correspondences are located mostly inside or near to white matter: larger-scale features in the corpus-callosum, and smaller-scale ones in the brain stem and along white matter bundles. The distribution of features in prominent brain regions is further analyzed in the next section.

Comparing different sibling types, we see a greater number of correspondences between MZ twins than between DZ twins or full siblings. This observation, which is easier to visualize in T1w and T2w images, is

| T1w | T2w | FA |
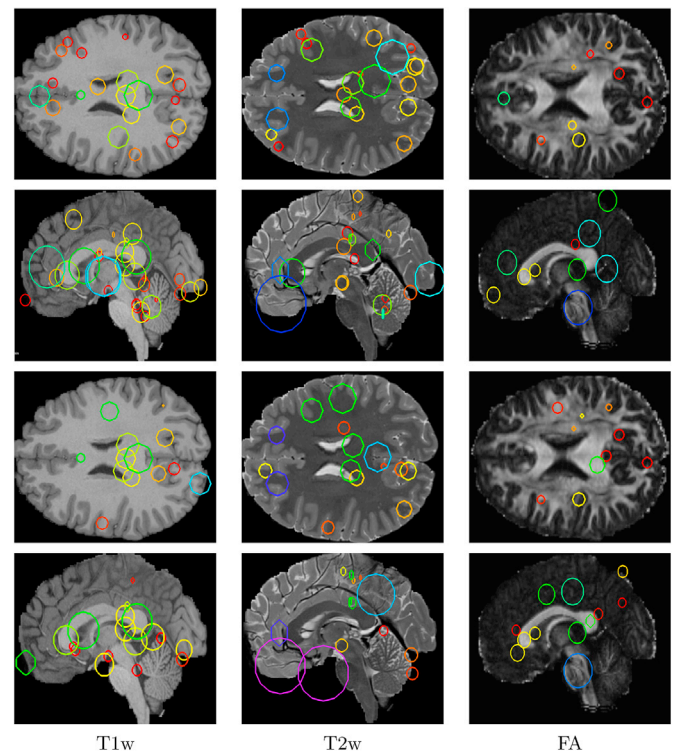
Fig. 4. Example of feature correspondences for a subject and his/her MZ twin (rows 1–2), and the subject's full sibling (FS) (rows 3–4). Scale space is represented using circle radius and color code (for the visible slice).

consistent with other analyses on twin datasets. In terms of feature location and scale, we observe a slightly higher number of correspondences in the frontal cortex for MZ twins, however, no obvious pattern can be drawn from one set of representative plots.

3.6.3.2. Region-wise analysis of feature correspondences. Here, we analyze the distribution of feature correspondences across atlas-defined neuro-anatomical regions, measured over the entire group of subjects. For each scan, segmentation labels were obtained from the Freesurfer-processed data, using LUT table for label descriptions.

Fig. 5 shows the box plot distributions of feature correspondences between pairs of MZ, DZ, and full siblings, for T1w and T2w images. Feature match counts are reported for five broad regions: non-white matter subcortex (s-cort), left/right cortex (crtx-lh/rh) and left/right white matter (wm-lh/rh). Note that mapping local features to a finer cortical parcellation is difficult due to the limited thickness of the cortex. Subcortical regions are further analyzed below.

Comparing across sibling types, we observe a higher number of feature correspondences for MZ pairs across all five regions and both T1w and T2w modalities. This confirms once again that the local features employed in our fingerprint captures brain characteristics related to genetic proximity. Analyzing the region-wise distribution of feature correspondences, all five regions are well represented. Since the number of local features in a region is proportional to its size, it is not surprising that the cortex has the least correspondences. Yet, such features are also produced by intensity variations (i.e., edges), thus explaining why many correspondences are found in the cortex. Finally, when comparing T1w and T2w modalities, we see small differences in the match counts, however these are not statistically significant.

To identify regions showing a strong relationship to genetic proximity, Table 5 gives the p-values ($-\log_{10}$ scale) of an unpaired t-test comparing the mean number of correspondences between subjects of a given sibling type versus another sibling type (e.g., MZ vs DZ). Significance values are provided for the five major regions described above, as well as for 15 prominent subcortical structures matching the analysis by Wachinger et al. (2015). To account for multiple comparisons (i.e., one for each tested region), reported p-values have been corrected using the Holm-Bonferroni procedure (Holm, 1979). Moreover, to account for age and size bias in this analysis, we selected FS pairs with less than 3 years age difference, and matched the number of FS pairs to MZ pairs using a simple bipartite matching based on age.

From Table 5, we observe significant differences between MZ twins and DZ-twins/full-siblings (i.e., corrected p-value < 0.01), for all five major regions and for both T1w and T2w images. In subcortical structures of T1w images, cerebellum white matter and cortex (left and right), lateral ventricles (left and right), left hippocampus and left thalamus
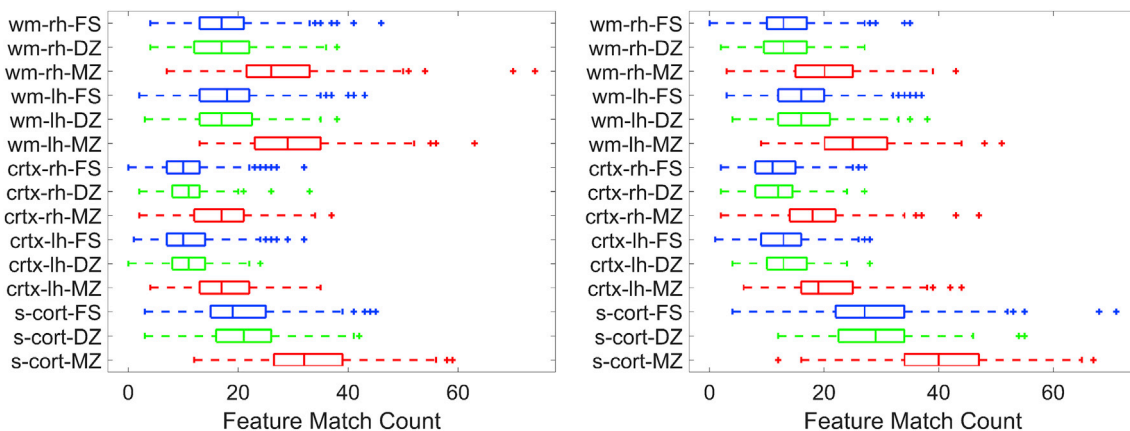
Table 5

Differences in feature match counts obtained for different sibling types in various brain regions, using T1w and T2w. We report HolmBonferroni corrected p-values ($-\log_{10}$ scale) measured using an unpaired t-test. Significant results with corrected p-value < 0.01 are highlighted using bold font.

Label	T1w		T2w	
	MZ vs DZ	MZ vs FS	MZ vs DZ	MZ vs FS
subcortical	**29.31**	**50.31**	**26.06**	**39.41**
Crtx-LH	**22.85**	**35.17**	**23.37**	**38.87**
Crtx-RH	**21.48**	**39.76**	**25.38**	**37.73**
WM-LH	**37.64**	**62.83**	**27.88**	**47.38**
WM-RH	**23.38**	**36.80**	**21.88**	**32.62**
L-Lat-Vent	**5.84**	**11.34**	**5.31**	**7.50**
R-Lat-Vent	**4.21**	**10.72**	**3.99**	**7.57**
L-VentralDC	1.49	**6.06**	**5.16**	2.98
R-VentralDC	0.45	0.60	0.00	0.01
R-Cerebellum-WM	**3.98**	**15.40**	0.00	0.57
L-Cerebellum-WM	**4.82**	**11.11**	2.33	**6.55**
R-Putamen	0.48	0.51	**2.34**	1.24
L-Putamen	0.87	0.35	0.06	0.30
L-Cerebellum-Crtx	**5.74**	**6.26**	**5.58**	**13.81**
L-Thalamus-Proper	**2.71**	**4.03**	0.37	0.01
4th-Ventricle	1.49	**2.24**	1.86	**3.91**
L-Hippocampus	**3.23**	**3.76**	**4.61**	**5.85**
CC-Anterior	1.83	0.51	0.40	0.71
R-Cerebellum-Crtx	**5.96**	**11.93**	**3.38**	**7.57**
3rd-Ventricle	0.45	0.51	0.37	0.43

proper have a significantly different number feature correspondences in MZ twins than in DZ twins or FS subjects. Comparing results obtained with T1w and T2w, the same structures are significant across both modalities, differences in significance reflecting the complimentary of these modalities.

3.6.3.3. Hemisphere asymmetry analysis. In our last experiment, we analyze the symmetry of feature match counts across brain hemispheres, for major structures. Toward this goal, we considered only right-handed (RH) subjects, and limited sibling pairs to subjects with same gender (i.e., a male and his brother, or a female and her sister). For non-twin siblings, we also restricted our analysis to subject pairs with less than 3 years of age difference.

Table 6 gives the results of two-sided unpaired t-tests comparing the feature match counts between cortical or white matter regions (Freesurfer LUT labels) in left- and right- hemispheres. To analyze gender effects, we also report results individually for RH male siblings and RH female siblings. Overall, we observe significant asymmetry in white matter regions (with corrected p-value < 0.01) of MZ twins, the highest significance values obtained for T2w images. No clear pattern is found

Fig. 5. Box plot comparison between MZ, DZ, and FS for pairwise feature correspondence counts for T1w (left) and T2w (right) for major structures. Red, green and blue correspond to MZ, DZ, and FS respectively.

Table 6
Hemisphere asymmetry analysis. For a given modality and twin type, we compare feature match count differences across hemisphere for major structures. Differences are reported as HolmBonferroni corrected p-values (-$\log_{10}$ scale) measured using an unpaired *t*-test, significant results (corrected p-value < 0.01) highlighted using bold font.

Modality	Type	RH Female		RH Male		RH Pairs	
		Crtx	WM	Crtx	WM	Crtx	WM
T1w	MZ	0.95	1.20	0.15	**2.57**	0.87	**2.73**
	DZ	1.05	0.78	0.35	0.99	1.12	0.06
	FS	1.71	0.39	0.84	0.11	1.89	0.09
T2w	MZ	1.95	**9.13**	1.52	**5.77**	**3.00**	**13.97**
	DZ	1.06	**3.60**	1.29	1.23	1.93	**4.06**
	FS	1.04	1.22	1.23	**5.35**	1.90	**5.84**

across sibling types, although hemispherical differences are generally higher in MZ twins than in DZ twins or full siblings. Likewise, no conclusion can be drawn when comparing results for male and female sibling pairs, with significance values varying across different sibling types and modalities.

The asymmetry of function in the brain, for example the hemispheric specializations of language and motor functions, has been extensively studied (Toga and Thompson, 2003). Similarly, studies have analyzed anatomical brain asymmetries based on voxel-based morphometry, sulci and other brain features (Wachinger et al., 2015). The multi-modal and multi-region analysis presented in this work extends previous studies of brain asymmetry in the literature by considering sibling types. Accounting for various confounds, including gender, genetics, handedness and age, this analysis has shown a greater asymmetry in feature correspondences between MZ twins than DZ twins and full siblings, mostly found in white matter regions and T2w images. Moreover, differences in asymmetry appear to be directional.

4. Discussion

In this section, we summarize the findings of this study and emphasize their link to previous investigations. We also highlight its limitations and discuss additional considerations.

4.1. Identification of genetically-related subjects

Our experiments on the task of identifying genetically-related subjects led to various useful observations. We established that the proposed fingerprint, generated from individual modalities or their combination, respects the genetic relationships between siblings, with MZ twins being more similar than DZ twins or full siblings (Peper et al., 2007; Thompson et al., 2013).

3D SIFT features (i.e., keypoints) coincide with the local extrema of a difference of Gaussians function applied in scale space. These features typically lie in high-contrast regions of an image, for instance due to the boundaries between white matter and grey matter (see Fig. 4). More generally, these features represent blob-like structures of varying size and location, which are robust and discriminative for finding correspondences across images. With respect to a voxel-wise full image comparison, the proposed BoF-based fingerprint offers a more compact representation of brain geometry, which is less sensitive to differences in image alignment and contrast. Likewise, compared to standard morphological measures like cortical thickness or sub-cortical region volume/area, our representation may capture a broader range of geometrical brain characteristics, for example distinctive cortical folding patterns only present in a subset of the population.

Analyzing the manifold approximation, we also showed that a discriminative fingerprint could be obtained with only 150 spectral components (i.e., leading eigenvectors of the normalized adjacency matrix of the subject proximity graph). When compared to a baseline using

full images as features, this compact fingerprint yielded significantly better performances, for all modalities and sibling types. This illustrates the high efficiency of our fingerprint and its advantages for comparing large groups of subjects. Moreover, while Laplacian eigenmaps were used to embed the subject proximity graph, the proposed framework is generic and other approaches (e.g., see Bengio et al., 2013) can be employed for this task.

The comparison of fingerprints obtained from structural MRI, diffusion MRI, and resting state fMRI highlighted the informativeness and complementarity of these modalities. Among individual modalities, resting state fMRI based fingerprint performed best for DZ/FS identification and had similar performance to FA/GFA for MZ twin identification. As mentioned in Finn et al. (2015), this could be due to the discriminative power of connectivity profiles, which is a result of integrating imaging data over a relatively long period of time (4800 vol, and 4 runs of 15 min each). The inter-individual variability in the connectivity profile (rfMRI netmat) results in part from the spatial topography (spatial variability in the location of functional regions across individuals) and functional coupling strength. However, assessing the specific contributions of these factors is outside the scope of the current study, and we refer readers to the works of Bijsterbosch et al. (2018) and Gordon et al. (2017c) for further information on this topic. Moreover, while the MAP values for FA/GFA are similar to rfMRI based MZ twin identification, mean recall@10 and relative identification % showed that FA performs slightly better than rfMRI (2.54% unique MZ pair identification as opposed to 0.84% pairs).

Comparing structural and diffusion MRI modalities, we found fingerprints based on FA/GFA to outperform those derived from T1w or T2w. We hypothesize this is caused by the pronounced contrast/magnitude changes in FA maps occurring at the boundary between grey matter and white matter (e.g., endpoints of fiber bundles). As shown qualitatively in Fig. 4, this leads to a more evenly-distributed set of feature matches.

Another interesting observation is the higher MAP values obtained for the identification of DZ twins compared to full siblings, although both sibling types have the same genetic similarity. In Supplemental material Table 1, this difference is found to be significant for several modality combinations (e.g., T1w + T2w + FA + rfMRI, p-value < 0.001). While our experiments accounted for group size and age differences by matching DZ subject pairs with FS pairs having the smallest age differences (one to three years difference), remaining age differences may explain this observation. Further investigation is however required to fully validate this hypothesis.

This work is motivated by the recent increase in multi-modal brain studies. For instance, multi-modal MRI has been shown useful for the analysis of neurodegenerative disorders (Calhoun and Sui, 2016) as well as for identifying subjects with schizophrenia (Sui et al., 2014). Results of this study demonstrate the usefulness of combining multiple modalities in a brain fingerprint. The improvements due to multi-modal combination, for all the twin/sibling types, can be attributed to more comprehensive characterization considering structure, white matter architecture, and functional connectivity. Thus, better performances were obtained with a combined set of modalities than with these modalities alone. Our results are consistent with previous studies underlining the benefit of a multi-modal fusion (Calhoun and Sui, 2016; Groves et al., 2012). As a note, we have focused on major observations only, the comprehensive analysis is open to various other observations including comparison of DTI vs GQI measures, inclusion of T1w-by-T2w MRI ratio images, FreeSurfer measures based identification, etc.

4.2. Advantages and applicability of the proposed fingerprint

Our factor impact analysis demonstrated the robustness of the proposed fingerprint to the non-alignment of images. Since image alignment is key for most population level analysis (Dubois and Adolphs, 2016), by alleviating this requirement, the proposed fingerprint may help save

computational costs and avoids errors introduced during alignment. Experiments have also shown that scan resolution (from 0.7 mm to 1.25 mm) does not have a significant impact on results, although using lower resolution images reduces the number detected features. Data acquired from multiple sites or scanners often need to be brought to the same resolution, introducing small errors during interpolation and re-sampling. The proposed fingerprint may thus be of help for multi-site studies, and pave the way to resolution-independent analyses. Lastly, using retest scans led to no significant changes in results, further validating the robustness of our fingerprint to image acquisition. However, a detailed analysis using multiple single-subject scans with longer between-scan times would be required to fully confirm this claim.

The proposed rank retrieval analysis based on MAP provides a principled approach for comparing different brain fingerprints, which could be utilized in future studies. In this work, we used the proposed fingerprint to find incorrectly reported zygosity labels and identify retest/duplicate scans of the same subjects. Hence, our fingerprint could serve as an efficient and reliable tool for detecting inconsistent information in large cohorts. Another potential application could be to provide physicians with related cases in clinical settings like MCI diagnostic assistance (Gao et al., 2015). For such large-scale content-based image retrieval (CBIR) tasks, computational efficiency and the ability to capture subtle differences between subjects are of high importance (Akgül et al., 2011; Li et al., 2018).

While various twin studies have analyzed genetic influences based on volume, cortical thickness, surface area, and morphometry (Wachinger et al., 2015), this is the first work to use local features and manifold approximation for this problem. Analyzing the distribution of features correspondences across brain regions, in images of different modalities, reveals many interesting insights. Results identify various neuroanatomical regions (e.g., cerebellum, lateral ventricles, ventral diencephalon, hippocampus and thalamus proper) having significantly different match counts in MZ twins than DZ twins or full siblings. These findings relate to those reported in Wachinger et al. (2015), which were obtained on a different dataset (mean subject age of 56 years, compared to a median of 29 years in the HCP dataset). Another key aspect of our analysis is the size of the subject cohort, larger than that of related studies (Peper et al., 2007).

4.3. Limitations and additional considerations

In this work, we used a rank retrieval analysis to evaluate the relation between fingerprint similarity and genetic proximity. Mean recall@k and mean average precision (MAP) were employed to measure sensitivity, specificity, and relative informativeness of fingerprints generated from different modality combinations. However, estimating heritability directly, for instance using the approach described in Ge et al. (2016), would provide a better quantification of genetic influence on fingerprint features. In Elliott et al. (2017), Elliott et al. considered the data of over 8,000 from the UK Bio-bank (Sudlow et al., 2015) to determine the heritability of multi-modal brain imaging phenotypes. Similarly, Colclough et al. report in Colclough et al. (2017) the heritability of multi-modal functional connectivity profiles using 800 HCP subjects. An extensive analysis is required to asses the heritability of the proposed fingerprint and relate our findings to those in these recent studies.

Moreover, when building the subject proximity graph, we assumed the independence of feature correspondences across modalities. However, a deeper analysis could be carried out to investigate false feature correspondences and correlation between features correspondences across modality. As mentioned before, other manifold embedding methods like Locally Linear Embedding (LLE) (Roweis and Saul, 2000) could also be employed for this step.

A limitation of the proposed framework stems from the choice of descriptor for the bag of feature representation. While our data-driven approach based on 3D SIFT descriptors is computationally efficient and provides performances comparable to sophisticated measures like

FreeSurfer derived estimates of volume, thickness, and area, the bag of feature representation is sparse and only captures a subset of relevant information. Our choice for 3D SIFT descriptors was motivated by previous applications in segmentation, classification, registration and group analysis (Toews et al., 2010, 2015; Toews and Wells, 2013), however, other local descriptors could provide improved performance. Also, loss of information occurs when creating the feature graph due to the approximation technique for finding nearest neighbors. This approximation is necessary to reduce computational time.

Another limitation of our framework is that features and consequently results are harder to interpret compared to standard voxel- or region-based analyses. We plan to further explore these questions in our next step, for example, by analyzing the impact of feature correspondence localization on the fingerprint components, or interpreting the contribution of individual modalities in single subject characterization.

In our study, we only used the multi-modal HCP dataset. Further investigations, considering multi-site variations and pre-processing pipeline impact, are needed to establish the generalizability of the proposed approach. Moreover, we observed fingerprints of size 150 eigenvectors to be suitable for twin/sibling identification. Yet, the number of dimensions needed to accurately characterize individuals likely varies depending on the application at hand. Here, we explored separability along the dimension of zygosity, however predicting more subtle dimensions, such as behavior (Shen et al., 2017) or markers of neuropsychiatric illness, may well require additional features.

In this study, we analyzed data from sMRI, dMRI, and rfMRI. However, the proposed framework is generic and could be extended to other modalities like task-fMRI, PET-MRI, and quantitative T1/T2 maps. Finally, an interesting extension of this work would be to assess whether our fingerprint can be used as a biomarker to identify subjects with cognitive or neurological disorders. Publicly available data, for instance from the ADNI dataset (Toews et al., 2010) or Parkinson's Progression Markers Initiative (PPMI) dataset (Marek et al., 2011), could be used for this analysis. Likewise, we plan to investigate how the proposed fingerprints change over time (Yeh et al., 2016b) in a longitudinal setting, observing the relative change in the contribution of individual modalities towards single subject characterization.

5. Conclusion

We presented a brain fingerprint, based on manifold approximation, for the multi-modal analysis of genetically-related subjects. In a rank retrieval analysis, mean recall@k and mean average precision were used to measure the relation between fingerprint similarity and genetic proximity, as well as the contribution/complementarity of information from different MRI modalities. Results indicated that a compact fingerprint of only 150 features could identify genetically-related subjects better than a baseline using full images as features. Our experiments also showed that each modality provides complementary information which can uniquely identify some sibling pairs. Furthermore, we demonstrated the benefit of considering multiple modalities in the fingerprint, combined modalities leading to a better performance than considering these modalities separately. Moreover, our analysis demonstrated the robustness of the proposed fingerprint to various factors, including image alignment, scan resolution, and subject age. The reproducibility of results was also confirmed using retest scans from the HCP dataset, showing our fingerprint to be robust to variability in image acquisition.

The usefulness of our fingerprint was assessed on the tasks of identifying incorrectly reported zygosity and retest/duplicate scans in large dataset. Results of this experiment highlighted the effectiveness of our fingerprint, with MAP values near 100% for all test cases. Moreover, analyzing the distribution of features correspondences across the brain revealed neuroanatomical regions (e.g., cerebellum, lateral ventricles, ventral diencephalon, hippocampus and thalamus proper) with significantly different match counts in MZ twins compared to DZ twins or full siblings. This work could be extended by further investigating the

differences, in terms of feature location and similarity, between dizygotic twins and non-twin siblings. A deeper analysis of aging effects could also be performed, for instance, using longitudinal data. Such analysis would help understand the effect of neuroplasticity on individual brain characteristics.

Acknowledgments

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.neuroimage.2018.08.006.

Code availability

Matlab scripts of the proposed analysis are available at https://github.com/kkumar-iitkgp-livia/Multi_Modal_Brain_Fingerprinting.git.

References

Akgül, C.B., Rubin, D.L., Napel, S., Beaulieu, C.F., Greenspan, H., Acar, B., 2011. Content-based image retrieval in radiology: current status and future directions. J. Digit. Imag. 24 (2), 208–222.

Alexander, A.L., Lee, J.E., Lazar, M., Field, A.S., 2007. Diffusion tensor imaging of the brain. Neurotherapeutics 4 (3), 316–329.

Aljabar, P., Wolz, R., Rueckert, D., 2012. Manifold Learning for Medical Image Registration, Segmentation, and Classification, Machine Learning in Computer-aided Diagnosis: Medical Imaging Intelligence and Analysis, p. 351.

Barch, D.M., Burgess, G.C., Harms, M.P., Petersen, S.E., Schlaggar, B.L., Corbetta, M., Glasser, M.F., Curtiss, S., Dixit, S., Feldt, C., et al., 2013. Function in the human connectome: task-fMRI and individual differences in behavior. Neuroimage 80, 169–189.

Bay, H., Tuytelaars, T., Van Gool, L., 2006. SURF: Speeded up robust features. Comput. Vis.–ECCV 2006 404–417.

Belkin, M., Niyogi, P., 2003. Laplacian eigenmaps for dimensionality reduction and data representation. Neural Comput. 15 (6), 1373–1396.

Bengio, Y., Paiement, J.-f., Vincent, P., Delalleau, O., Roux, N.L., Ouimet, M., 2004. Out-of-sample extensions for lle, isomap, mds, eigenmaps, and spectral clustering. In: Advances in Neural Information Processing Systems, pp. 177–184.

Bengio, Y., Courville, A., Vincent, P., 2013. Representation learning: a review and new perspectives. IEEE Trans. Pattern Anal. Mach. Intell. 35 (8), 1798–1828.

Bijsterbosch, J.D., Woolrich, M.W., Glasser, M.F., Robinson, E.C., Beckmann, C.F., Van Essen, D.C., Harrison, S.J., Smith, S.M., 2018. The relationship between spatial configuration and functional connectivity of brain regions. eLife 7, e32992.

Brosch, T., Tam, R., Initiative, A.D.N., et al., 2013. Manifold learning of brain MRIs by deep learning. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 633–640.

Bürgel, U., Amunts, K., Hoemke, L., Mohlberg, H., Gilsbach, J.M., Zilles, K., 2006. White matter fiber tracts of the human brain: three-dimensional mapping at microscopic resolution, topography and intersubject variability. Neuroimage 29 (4), 1092–1105.

Calhoun, V.D., Sui, J., 2016. Multimodal fusion of brain imaging data: a key to finding the missing link (s) in complex mental illness. Biol. Psychiatr.: Cognit. Neurosci. Neuroimaging 1 (3), 230–244.

Colclough, G.L., Smith, S.M., Nichols, T.E., Winkler, A.M., Sotiropoulos, S.N., Glasser, M.F., Van Essen, D.C., Woolrich, M.W., 2017. The heritability of multi-modal connectivity in human brain activity. Elife 6.

Dubois, J., Adolphs, R., 2016. Building a science of individual differences from fMRI. Trends Cognit. Sci. 20 (6), 425–443.

Elliott, L., Sharp, K., Alfaro-Almagro, F., Douaud, G., Miller, K., Marchini, J., Smith, S., 2017. The genetic basis of human brain structure and function: 1,262 genome-wide associations found from 3,144 GWAS of multimodal brain imaging phenotypes from 9,707 UK biobank participants. BioRxiv 178806.

Finn, E.S., Constable, R.T., 2016. Individual variation in functional brain connectivity: implications for personalized approaches to psychiatric disease. Dialogues Clin. Neurosci. 18 (3), 277.

Finn, E.S., Shen, X., Scheinost, D., Rosenberg, M.D., Huang, J., Chun, M.M., Papademetris, X., Constable, R.T., 2015. Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. Nat. Neurosci.

Fowlkes, C., Belongie, S., Chung, F., Malik, J., 2004. Spectral grouping using the Nystrom method. IEEE Trans. Pattern Anal. Mach. Intell. 26 (2), 214–225.

Gale, S.D., Perkel, D.J., 2010. A basal ganglia pathway drives selective auditory responses in songbird dopaminergic neurons via disinhibition. J. Neurosci. 30 (3), 1027–1037.

Gao, Y., Adeli-M, E., Kim, M., Giannakopoulos, P., Haller, S., Shen, D., 2015. Medical image retrieval using multi-graph learning for MCI diagnostic assistance. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 86–93.

Ge, T., Reuter, M., Winkler, A.M., Holmes, A.J., Lee, P.H., Tirrell, L.S., Roffman, J.L., Buckner, R.L., Smoller, J.W., Sabuncu, M.R., 2016. Multidimensional heritability analysis of neuroanatomical shape. Nat. Commun. 7, 13291.

Geevarghese, R., Lumsden, D.E., Hulse, N., Samuel, M., Ashkan, K., 2015. Subcortical structure volumes and correlation to clinical variables in Parkinson's disease. J. Neuroimaging 25 (2), 275–280.

Gerber, S., Tasdizen, T., Fletcher, P.T., Joshi, S., Whitaker, R., Initiative, A.D.N., et al., 2010. Manifold modeling for brain population analysis. Med. Image Anal. 14 (5), 643–653.

Glasser, M.F., Sotiropoulos, S.N., Wilson, J.A., Coalson, T.S., Fischl, B., Andersson, J.L., Xu, J., Jbabdi, S., Webster, M., Polimeni, J.R., et al., 2013. The minimal preprocessing pipelines for the human connectome project. Neuroimage 80, 105–124.

Goldman, S., O'Brien, L.M., Filipek, P.A., Rapin, I., Herbert, M.R., 2013. Motor stereotypies and volumetric brain alterations in children with autistic disorder. Res. Autism Spectr. Disord. 7 (1), 82–92.

Gordon, E.M., Laumann, T.O., Adeyemo, B., Petersen, S.E., 2017. Individual variability of the system-level organization of the human brain. Cerebr. Cortex 27 (1), 386–399.

Gordon, E.M., Laumann, T.O., Adeyemo, B., Gilmore, A.W., Nelson, S.M., Dosenbach, N.U., Petersen, S.E., 2017. Individual-specific features of brain systems identified with resting state functional correlations. Neuroimage 146, 918–939.

Gordon, E.M., Laumann, T.O., Gilmore, A.W., Newbold, D.J., Greene, D.J., Berg, J.J., Ortega, M., Hoyt-Drazen, C., Gratton, C., Sun, H., et al., 2017. Precision functional mapping of individual human brains. Neuron 95 (4), 791–807.

Groves, A.R., Smith, S.M., Fjell, A.M., Tamnes, C.K., Walhovd, K.B., Douaud, G., Woolrich, M.W., Westlye, L.T., 2012. Benefits of multi-modal fusion analysis on a large-scale dataset: life-span patterns of inter-subject variability in cortical morphometry and white matter microstructure. Neuroimage 63 (1), 365–380.

Hampel, H., OBryant, S., Durrleman, S., Younesi, E., Rojkova, K., Escott-Price, V., Corvol, J.-C., Broich, K., Dubois, B., Lista, S., et al., 2017. A precision medicine initiative for alzheimers disease: the road ahead to biomarker-guided integrative disease modeling. Climacteric 20 (2), 107–118.

Holm, S., 1979. A simple sequentially rejective multiple test procedure. Scand. J. Stat. 65–70.

Jenkinson, M., Beckmann, C.F., Behrens, T.E., Woolrich, M.W., Smith, S.M., 2012. Fsl. Neuroimage 62 (2), 782–790.

Kochunov, P., Jahanshad, N., Marcus, D., Winkler, A., Sprooten, E., Nichols, T.E., Wright, S.N., Hong, L.E., Patel, B., Behrens, T., et al., 2015. Heritability of fractional anisotropy in human white matter: a comparison of human connectome project and ENIGMA-DTI data. Neuroimage 111, 300–311.

Kumar, K., Desrosiers, C., Siddiqi, K., Colliot, O., Toews, M., 2017. Fiberprint: a subject fingerprint based on sparse code pooling for white matter fiber analysis. Neuroimage 158, 242–259. https://doi.org/10.1016/j.neuroimage.2017.06.083. http://www.sciencedirect.com/science/article/pii/S1053811917305669.

Kumar, K., Chauvin, L., Toews, M., Colliot, O., Desrosiers, C., 2018. Multi-modal analysis of genetically-related subjects using sift descriptors in brain MRI. In: Computational Diffusion MRI. Springer, Cham, pp. 219–228.

Li, Z., Zhang, X., Müller, H., Zhang, S., 2018. Large-scale retrieval for medical image analytics: a comprehensive review. Med. Image Anal. 43, 66–84.

Liem, F., Varoquaux, G., Kynast, J., Beyer, F., Masouleh, S.K., Huntenburg, J.M., Lampe, L., Rahim, M., Abraham, A., Craddock, R.C., et al., 2017. Predicting brain-age from multimodal imaging data captures cognitive impairment. Neuroimage 148, 179–188.

Liu, S., Cai, W., Liu, S., Zhang, F., Fulham, M., Feng, D., Pujol, S., Kikinis, R., 2015. Multimodal neuroimaging computing: a review of the applications in neuropsychiatric disorders. Brain Inf. 2 (3), 167–180.

Liu, S., Cai, W., Liu, S., Zhang, F., Fulham, M., Feng, D., Pujol, S., Kikinis, R., 2015. Multimodal neuroimaging computing: the workflows, methods, and platforms. Brain Inf. 2 (3), 181–195.

Liu, J., Liao, X., Xia, M., He, Y., 2018. Chronnectome fingerprinting: identifying individuals and predicting higher cognitive functions using dynamic brain connectivity patterns. Hum. Brain Mapp. 39 (2), 902–915.

Lowe, D.G., 1999. Object recognition from local scale-invariant features. In: Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on, Vol. 2. IEEE, pp. 1150–1157.

Lowe, D.G., 2004. Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vis. 60 (2), 91–110.

Mangin, J.-F., Riviere, D., Cachia, A., Duchesnay, E., Cointepas, Y., Papadopoulos-Orfanos, D., Scifo, P., Ochiai, T., Brunelle, F., Regis, J., 2004. A framework to study the cortical folding patterns. Neuroimage 23, S129–S138.

Marek, K., Jennings, D., Lasch, S., Siderowf, A., Tanner, C., Simuni, T., Coffey, C., Kieburtz, K., Flagg, E., Chowdhury, S., et al., 2011. The Parkinson progression marker initiative (PPMI). Prog. Neurobiol. 95 (4), 629–635.

Miranda-Dominguez, O., Mills, B.D., Carpenter, S.D., Grant, K.A., Kroenke, C.D., Nigg, J.T., Fair, D.A., 2014. Connectotyping: model based fingerprinting of the functional connectome. PLoS One 9 (11), e111048.

Mueller, S., Wang, D., Fox, M.D., Yeo, B.T., Sepulcre, J., Sabuncu, M.R., Shafee, R., Lu, J., Liu, H., 2013. Individual variability in functional connectivity architecture of the human brain. Neuron 77 (3), 586–595.

Muja, M., Lowe, D.G., 2009. Fast approximate nearest neighbors with automatic algorithm configuration. VISAPP 2 (1), 331–340.

O'Donnell, L.J., Westin, C.-F., 2007. Automatic tractography segmentation using a high-dimensional white matter atlas. IEEE Trans. Med. Imag. 26 (11), 1562–1575.

Peper, J.S., Brouwer, R.M., Boomsma, D.I., Kahn, R.S., Pol, H., Hilleke, E., 2007. Genetic influences on human brain structure: a review of brain imaging studies in twins. Hum. Brain Mapp. 28 (6), 464–473.

Polderman, T.J., Benyamin, B., De Leeuw, C.A., Sullivan, P.F., Van Bochoven, A., Visscher, P.M., Posthuma, D., 2015. Meta-analysis of the heritability of human traits based on fifty years of twin studies. Nat. Genet. 47 (7), 702–709.

Roweis, S.T., Saul, L.K., 2000. Nonlinear dimensionality reduction by locally linear embedding. Science 290 (5500), 2323–2326.

de Schotten, M.T., Bizzi, A., Dell'Acqua, F., Allin, M., Walshe, M., Murray, R., Williams, S.C., Murphy, D.G., Catani, M., et al., 2011. Atlasing location, asymmetry and inter-subject variability of white matter tracts in the human brain with MR diffusion tractography. Neuroimage 54 (1), 49–59.

Shen, X., Finn, E.S., Scheinost, D., Rosenberg, M.D., Chun, M.M., Papademetris, X., Constable, R.T., 2017. Using connectome-based predictive modeling to predict individual behavior from brain connectivity. Nat. Protoc. 12 (3), 506.

Smith, S.M., Nichols, T.E., Vidaurre, D., Winkler, A.M., Behrens, T.E., Glasser, M.F., Ugurbil, K., Barch, D.M., Van Essen, D.C., Miller, K.L., 2015. A positive-negative mode of population covariation links brain connectivity, demographics and behavior. Nat. Neurosci. 18 (11), 1565.

Sotiropoulos, S.N., Jbabdi, S., Xu, J., Andersson, J.L., Moeller, S., Auerbach, E.J., Glasser, M.F., Hernandez, M., Sapiro, G., Jenkinson, M., et al., 2013. Advances in diffusion MRI acquisition and processing in the human connectome project. Neuroimage 80, 125–143.

Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., Downey, P., Elliott, P., Green, J., Landray, M., et al., 2015. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med. 12 (3), e1001779.

Sui, J., Castro, E., He, H., Bridwell, D., Du, Y., Pearlson, G.D., Jiang, T., Calhoun, V.D., 2014. Combination of FMRI-SMRI-EEG data improves discrimination of schizophrenia patients by ensemble feature selection. In: Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE. IEEE, pp. 3889–3892.

Sui, J., Pearlson, G.D., Du, Y., Yu, Q., Jones, T.R., Chen, J., Jiang, T., Bustillo, J., Calhoun, V.D., 2015. In search of multimodal neuroimaging biomarkers of cognitive deficits in schizophrenia. Biol. Psychiatr. 78 (11), 794–804 schizophrenia: Glutamatergic Mechanisms of Cognitive Dysfunction and Treatment. https://doi.org/10.1016/j.biopsych.2015.02.017. http://www.sciencedirect.com/science/article/pii/S0006322315001274.

Tenenbaum, J.B., De Silva, V., Langford, J.C., 2000. A global geometric framework for nonlinear dimensionality reduction. Science 290 (5500), 2319–2323.

Thompson, P.M., Ge, T., Glahn, D.C., Jahanshad, N., Nichols, T.E., 2013. Genetics of the connectome. Neuroimage 80, 475–488.

Toews, M., Wells, W.M., 2013. Efficient and robust model-to-image alignment using 3d scale-invariant features. Med. Image Anal. 17 (3), 271–282.

Toews, M., Wells, W.M., 2016. How are siblings similar? how similar are siblings? large-scale imaging genetics using local image features. In: 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI), pp. 847–850. https://doi.org/10.1109/ISBI.2016.7493398.

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

Toews, M., Wells, W.M., Zöllei, L., 2012. A feature-based developmental model of the infant brain in structural MRI. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 204–211.

Toews, M., Wachinger, C., Estepar, R.S.J., Wells III, W.M., 2015. A feature-based approach to big data analysis of medical images. In: International Conference on Information Processing in Medical Imaging. Springer, pp. 339–350.

Toga, A.W., Thompson, P.M., 2003. Mapping brain asymmetry. Nat. Rev. Neurosci. 4 (1), 37.

Tsai, C.-F., 2012. Bag-of-words Representation in Image Annotation: a Review. ISRN Artificial Intelligence.

Turpin, A., Scholer, F., 2006. User performance versus precision measures for simple search tasks. In: Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. ACM, pp. 11–18.

Tuytelaars, T., Mikolajczyk, K., et al., 2008. Local invariant feature detectors: a survey. Found. Trends® Comput. Graph. Vis. 3 (3), 177–280.

Van Essen, D.C., Ugurbil, K., Auerbach, E., Barch, D., Behrens, T., Bucholz, R., Chang, A., Chen, L., Corbetta, M., Curtiss, S.W., et al., 2012. The human connectome project: a data acquisition perspective. Neuroimage 62 (4), 2222–2231.

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

Wachinger, C., Golland, P., Kremen, W., Fischl, B., Reuter, M., Initiative, A.D.N., et al., 2015. Brainprint: a discriminative characterization of brain morphology. Neuroimage 109, 232–248.

Yeh, F.-C., Wedeen, V.J., Tseng, W.-Y.I., 2010. Generalized q-sampling imaging. IEEE Trans. Med. Imag. 29 (9), 1626–1635.

Yeh, F.-C., Badre, D., Verstynen, T., 2016. Connectometry: a statistical approach harnessing the analytical potential of the local connectome. Neuroimage 125, 162–171.

Yeh, F.-C., Vettel, J.M., Singh, A., Poczos, B., Grafton, S.T., Erickson, K.I., Tseng, W.-Y.I., Verstynen, T.D., 2016. Quantifying differences and similarities in whole-brain white matter architecture using local connectome fingerprints. PLoS Comput. Biol. 12 (11), e1005203.

Zhang, W., Li, R., Deng, H., Wang, L., Lin, W., Ji, S., Shen, D., 2015. Deep convolutional neural networks for multi-modality isointense infant brain image segmentation. Neuroimage 108, 214–224.