Efficient Pairwise Neuroimage Analysis using the Soft Jaccard Index and 3D Keypoint Sets

Laurent Chauvin, Kuldeep Kumar, Christian Desrosiers, William Wells III and Matthew Toews

Abstract— We propose a novel pairwise distance measure between image keypoint sets, for the purpose of large-scale medical image indexing. Our measure generalizes the Jaccard index to account for soft set equivalence (SSE) between keypoint elements, via an adaptive kernel framework modeling uncertainty in keypoint appearance and geometry. A new kernel is proposed to quantify the variability of keypoint geometry in location and scale. Our distance measure may be estimated between $O(N^2)$ image pairs in $O(N \log N)$ operations via keypoint indexing. Experiments report the first results for the task of predicting family relationships from medical images, using 1010 T1-weighted MRI brain volumes of 434 families including monozygotic and dizygotic twins, siblings and half-siblings sharing 100%-25% of their polymorphic genes. Soft set equivalence and the keypoint geometry kernel improve upon standard hard set equivalence (HSE) and appearance kernels alone in predicting family relationships. Monozygotic twin identification is near 100%, and three subjects with uncertain genotyping are automatically paired with their self-reported families, the first reported practical application of image-based family identification. Our distance measure can also be used to predict group categories, sex is predicted with an AUC=0.97. Software is provided for efficient fine-grained curation of large, generic image datasets.

Index Terms— neuroimage analysis, invariant keypoints, MRI, family prediction.

I. INTRODUCTION

HEALTH treatment is increasingly personalized, where treatment decisions are conditioned on patient-specific information in addition to knowledge regarding the general population [1]. Modern genetic testing allows us, based on large libraries of human DNA samples, to cheaply predict patient-specific characteristics, including immediate family relationships, and also characteristics shared across the population including racial ancestry, sex, hereditary disease status, etc. [2], [3]. The brain is the center of cognition and a complex organ, tightly coupled to the genetic evolution of animals and in particular that of the human species. To what degree is the human brain phenotype coupled to the underlying genotype? How does the brain image manifold vary locally with genotype, i.e. immediate relatives sharing 25-100% of their polymorphic genes, or between broad groups defined by subtle genetic factors such as racial ancestry or sex?

Large, publicly available databases allow us to investigate these questions from aggregated MRI and genetic data, e.g. 1000+ subjects [4]. The shape of the brain has been modeled as lying on a smooth manifold in high dimensional MRI data space [5], [6], where phenotype can be described as a smooth deformation conditioned on developmental factors including the environment. However the brain is also naturally described as a rich collection of spatially localized neuroanatomical structures, including common structures such as the basal ganglia shared across the population, but also highly specific patterns such as cortical folds that may only be observable in specific individuals or close family members [7].

The keypoint representation is an intuitively appealing means of modeling specific, localized phenomena, i.e. a set of descriptors automatically identified at salient image locations as shown in Figure 1. A keypoint set can be viewed as an element of a high-dimensional manifold, and medical imaging applications such as regression or classification can be formulated in terms of a suitable geodesic distance between sets. As keypoint sets are variable sized, typical metrics based on fixed-length vectors such as L-norms [5], [6] do not readily apply. Distances defined based on set overlap measures as the Jaccard index [8] have proven to be effective in recent studies investigating genetics and brain MRI [9], [10]. For example, by defining set equivalence in terms of nearest neighbor (NN) keypoint correspondences, the Jaccard distance may be used to predict pairwise relationships. Nevertheless, the assumption of binary or hard set equivalence (HSE) between keypoints is a crude approximation given probabilistic uncertainty inherent to natural image structure, and is ill-defined for variable sized datasets where the number of NN correspondences may be variable or unknown a-priori.

Our primary contribution is a new pairwise distance measure for 3D keypoint sets, generalizing the Jaccard index to account for soft set equivalence (SSE) between keypoint elements. SSE is achieved via a kernel framework for keypoint similarity, and we introduce a new kernel that normalizes pairwise keypoint

L. Chauvin, C. Desrosiers and M. Toews are with the Ecole de Technologie Superieure, Montreal, QC H3C 1K3, Canada (email: laurent.chauvin.2@etsmtl.net, christian.desrosiers@etsmtl.ca, matt.toews@gmail.com).

K. Kumar is with CHU Sainte-Justine Research Center, University of Montreal, Montreal, QC H3T 1C5, Canada (email: kuldeep.kumar@umontreal.ca).

W. Wells is with the Harvard Medical School, Boston, MA 02115, USA and the Massachusetts Institute of Technology, Boston, MA 02139, USA (email: sw@bwh.harvard.edu).
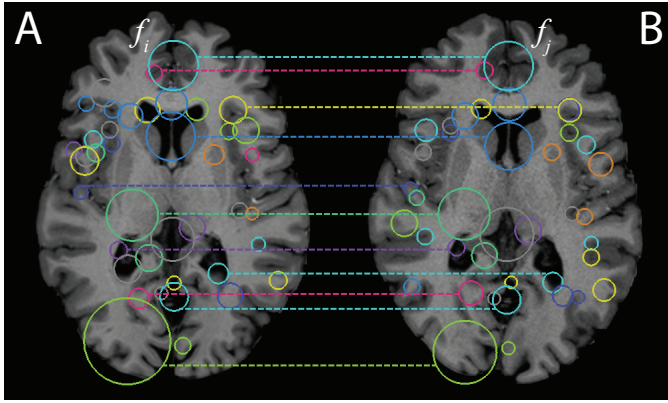
Fig. 1. Illustrating two sets of 3D keypoints $A = \{f_i\}$ and $B = \{f_j\}$ extracted from MRI brain volumes of monozygotic twins. Colored circles represent the 3D locations $\bar{x} \in R^3$ and scales (radii) $\sigma \in R^1$ of keypoints in 2D axial slices, and links are correspondences identified via the SIFT-Rank [11] algorithm, note the degree of similarity between keypoint pairs.

displacement by keypoint scale, accounting for localization uncertainty in scale-space. Our work extends and compares against keypoint-based neuroimaging analysis methods [9], [10], [12], specifically the keypoint signatures approach [10], currently the only method reporting perfect accuracy at identifying scans of the same individual in large public neuroimage MRI datasets, e.g. OASIS [13], ADNI [14] and HCP [15]. Here, following preliminary work [16], we demonstrate that the soft Jaccard distance is specific enough to predict family relationships from medical images, and we present the first results for this new task using 1010 MRIs of 434 families from the HCP [15] dataset. The entire method is based on highly efficient keypoint indexing, and scales to large datasets of volumetric image data.

II. RELATED WORK

Our work is motivated both by the study of the link between genotype and the human brain phenotype from large datasets [17], and by practical applications such as maintaining accurate patient records in hospital Picture Archiving and Communication Systems (PACS). We adopted a memory-based learning model where all data are stored in memory and accessed via highly-specific keypoint queries. Memory-based learning requires no explicit training procedure [18], and approaches Bayes optimal performance as the number of data N becomes large [19]. 3D SIFT keypoints [11] derived from Gaussian scale-space theory [20], [21] are invariant to global similarity transforms and contrast variations, and thus account for misalignment and scanner variability between images. Efficient indexing reduces the quadratic $O(N^2)$ complexity of nearest neighbor (NN) keypoint lookup to $O(N \log N)$, e.g. via $O(\log N)$ KD-tree indexing [22], ensuring that the method scales gracefully to large datasets, e.g. 7500 brain MRIs [10] or 20000 lung CTs [23].

Our method seeks a pairwise distance measure between 3D keypoint sets. As different images generally contain different numbers of elements, a natural choice for pairwise distance between variable sized sets are measures based on

set intersection or overlap, e.g. the Jaccard distance metric [8] first proposed in [24] or the Tanimoto distance [25], [26]. In medical image analysis, the Jaccard distance is typically used to assess pixel-level segmentation accuracy [27], and has proven highly effective for assessing pairwise neuroimage similarity from 3D SIFT keypoint data [9], [12], where hard set equivalence is determined by NN keypoint descriptor correspondences. Set theory is predicated on binary equivalence between set elements, which is difficult to justify in the case of keypoints extracted from image data. Soft set theory was been developed to account for non-binary equivalence [28]–[30], where set operations including intersection and union are defined in terms of non-binary equivalence ranging from [0,1]. Jaccard-like distance was used for retrieving near-duplicate 2D photos [31], where the ratio of soft intersection and union operators was used to account for document frequency. Our work here extends the Jaccard distance proposed in [10] to include keypoint geometry in 3D Euclidean metric space coordinates, and compares to the HSE Jaccard measure used in [9], [12], showing a significant improvement in predicting family relationships.

The keypoint methodology has a number of advantages, such as robustness to occlusion or missing correspondences, invariance to translation, scaling, rotation and intensity contrast variations between images. Our work adopts the 3D SIFT-Rank [11] representation, a robust, general tool used for a variety of imaging tasks, including registration [11], [32], whole body segmentation [33], classification [23], [34], regression [35], without the need for context-specific training procedures or data. As in the original 2D SIFT approach, keypoint geometry is represented as a location $\bar{x}$ and scale σ, these are detected as extrema of a difference-of-Gaussian scale-space [20], as $\text{argmax}_{\bar{x},\sigma} |I(\bar{x}, \sigma) - I(\bar{x}, \kappa\sigma)|$, approximating the Laplacian-of-Gaussian [21], where $I(\bar{x}, \sigma) = I(\bar{x}) * \mathcal{N}(\bar{x}, \sigma^2)$ represents the image convolved with the Gaussian kernel $\mathcal{N}(\bar{x}, \sigma^2)$. The geometry of a keypoint may thus be characterised as an isotropic Gaussian density centered on $\bar{x}$ with variance σ^2 representing spatial extent σ in 3D Euclidean space. Rotationally symmetric Gaussian filtering and uniformly sampled derivative operators lead to scale and rotation invariance. The 3D SIFT-Rank descriptor is a 64-element histogram of rank-ordered oriented gradients (HOG) [11], sampled about a scale-normalized reference frame centered on $\bar{x}$ and quantized uniformly into 2x2x2=8 spatial locations x 8 orientations. Rank transformation provides invariance to arbitrary monotonic variations of descriptor element values [36].

The 3D SIFT-Rank keypoint method we use is a robust and well-established baseline, however our distance measure may generally be used with any keypoint representation. The search for new keypoint detectors and descriptors remains an active research focus, primarily in the context of 2D computer vision. We refer the reader to an extensive literature review of traditional hand-crafted solutions [37] and a recent discussion of 2D SIFT matching technology [38]. Deep learning keypoint methods have been investigated, examples include LIFT [39], DISK [40], LF-Net [41], SuperPoint [42]. Nevertheless, variants of the original SIFT histogram descriptor, including SIFT-Rank [36], DSP-SIFT [43] or RootSIFT [44] remain compet-

itive with descriptors derived from training [45], [46], particularly for non-planar objects [38], while requiring no training and few hyperparameters. Learning-based approaches often rely on hand-crafted keypoint detectors to generate training data [39], [42]. The challenges of developing deep keypoint architectures include training of individual components (i.e. keypoint detector, orientation estimator, descriptor), achieving invariance to rotation or scale changes, the need to retrain for different imaging modalities and body parts, and avoiding bias due to training data procedures [47]. In terms of neuroimage analysis, brain fingerprinting methods have been used to investigate the variability of individuals based on neuroimage-specific preprocessing and data, e.g. brain segmentations from structural MRI [48], [49], neuroanatomic parcellations and functional MRI [50], fiber tracts from diffusion MRI [51]. None of these methods have been sufficiently specific to achieve perfect accuracy at individual identification, and none have been used for the more difficult task of predicting family members. Heritability studies have investigated correlations in MRIs of siblings, including structural [52], [53] and functional MRI [54], however these are not designed for prediction. In contrast, the keypoint methodology we adopt may be applied to generic imaging data with no preprocessing, representing distinctive patterns potentially present only in images of specific individuals or immediate family members. Our work here extends and compares against the keypoint-based neuroimage signatures method [10], the only method reporting perfect accuracy at same-subject image prediction from standard neuroimage datasets (e.g. OASIS [13], ADNI [14], HCP [15]).

Our results here are the first in the literature for the task of family member prediction from volumetric medical image data. Family member prediction has been investigated in the context of 2D face photographs, based on clustering [55], conditional random fields [56], deep neural networks [57], discriminant analysis [58]. These solutions are not directly related to our work, as they generally address aspects and challenges specific to the context of 2D photography, e.g. facial expression and lighting conditions [59], multiple images per person, group photos of relatively few families (e.g. 16) [56]. Accuracy is not generally available for siblings based on genotyping, and near 100% accuracy for monozygotic twin identification from face photographs has not been reported.

III. METHOD

We seek a pairwise distance measure between two sets $A = \{f_i\}$ and $B = \{f_j\}$ that can be used to estimate proximity and thus genetic relationships between subjects from image data. We first describe our method in terms of general set theory, then later include details pertaining to set elements f_i defined as 3D image keypoints. Our measure begins with the Jaccard index or intersection-over-union $J_{HSE}(A, B)$ defined as

$$J_{HSE}(A, B) = \frac{|A \cap B|}{|A| + |B| - |A \cap B|}, \tag{1}$$

where in Equation (1), $J_{HSE}(A, B)$ is defined by binary or hard set equivalence (HSE) between set elements. $A \cap B$ is the intersection operator between sets A and B, and $|A \cap B|$, $|A| = |A \cap A|$, $|B| = |B \cap B|$ represent set cardinality operators

that count the numbers of elements present in each set, where $|A \cap B| \leq \min\{|A|, |B|\}$. The Jaccard index is a similarity measure, and may be used to define various distance measures including the Jaccard distance metric [8] $1 - J_{HSE}(A, B)$ or the Tanimoto distance $-\log J_{HSE}(A, B)$.

In the case of sets of real data samples, for example keypoint descriptors, binary equivalence may be difficult to establish due to noise or uncertainty in the measurement process, and we seek to redefine $J_{HSE}(A, B)$ in Equation (1) such that it more accurately accounts for non-binary equivalence between set elements. This may generally be accomplished by redefining the standard set intersection cardinality operator $|A \cap B|$ in Equation (1) by an analogous soft cardinality operator $\mu(A \cap B)$, leading to the following expression for our proposed Jaccard index based on soft set equivalence $J_{SSE}(A, B)$

$$J_{SSE}(A, B) = \frac{\mu(A \cap B)}{\mu(A) + \mu(B) - \mu(A \cap B)}. \tag{2}$$

In Equation (2), $J_{SSE}(A, B)$ is defined by the soft cardinality operator $\mu(A \cap B)$ as described in the following section, including cardinality operators $\mu(A) = \mu(A \cap A)$, $\mu(B) = \mu(B \cap B)$. In order to ensure that J_{SSE} remains bounded to the range $[0, 1]$, it is important that $\mu(A \cap B)$ be upper bounded by the minimum of the individual soft set cardinalities $\mu(A \cap B) \leq \min\{\mu(A), \mu(B)\}$.

Defining $\mu(A \cap B)$: The cardinality of the soft set intersection $\mu(A \cap B)$ is intended to reflect the uncertainty in equivalence between set elements $f_i \in A$ and $f_j \in B$. We define soft equivalence between elements $f_i \in A$ and $f_j \in B$ as a similarity function or kernel $\mathcal{K}(f_i, f_j)$ ranging from $[0, 1]$, where $\mathcal{K}(f_i, f_j) = 1$ indicates exact equivalence and $\mathcal{K}(f_i, f_j) = 0$ represents the absence of equivalence. We then define a measure $\mu(A \to B)$ over a mapping $A \to B$ from set A to set B as:

$$\mu(A \to B) = \sum_i^{|A|} \max_{f_j \in B} \mathcal{K}(f_i, f_j). \tag{3}$$

The similarity function $\mathcal{K}(f_i, f_j)$ in Equation (3) may generally be defined according to the specific representation of set elements f_i and f_j. The maximum operator $\max_{f_j \in B} \mathcal{K}(f_i, f_j)$ ensures the bound $\mu(A \to B) \leq \mu(A)$ via a partial injective mapping from A to B, embodying the intuition that each element $f_i \in A$ has at most one counterpart $f_j \in B$. Note also that max operator may be computed via computationally efficient NN indexing methods. The soft intersection cardinality $\mu(A \cap B)$ may be defined as a symmetrized version of $\mu(A \to B)$:

$$\mu(A \cap B) = \min\{\mu(A \to B), \mu(B \to A)\},$$

$$= \min\left\{\sum_i^{|A|} \max_{f_j \in B} \mathcal{K}(f_i, f_j), \sum_j^{|B|} \max_{f_i \in A} \mathcal{K}(f_j, f_i)\right\}. \tag{4}$$

Note that the Jaccard distance measure derived using the cardinality operator $\mu(A \cap B)$ as defined by Equations (3) and (4) is symmetric and follows the identity of indiscernibles,

and thus may be considered to be a semi-metric. The triangle inequality holds in the case of binary equivalence [8] but not generally for soft equivalence relationships. Equation (4) may be used to define the cardinality of a single set A as $\mu(A) = \mu(A \cap A) = |A|$ both for hard and soft equivalence since, since for a given element $f_i \in A$, $\max_{f_j \in A} \mathcal{K}(f_i, f_j) = \mathcal{K}(f_i, f_i) = 1$. The standard Jaccard index in Equation (1) is thus a special case of the soft Jaccard index in the case of a binary-valued kernel $\mathcal{K}(f_i, f_j) \in \{0, 1\}$.

Defining $\mathcal{K}(f_i, f_j)$: The soft set intersection cardinality $\mu(A \cap B)$ may be adapted to a specific task by defining $\mathcal{K}(f_i, f_j)$ according to the representation of elements f_i. In the context of this article, a set element f_i is a 3D scale-invariant keypoint $f_i = \{\bar{a}_i, \bar{g}_i\}$ as described in [11], where $\bar{g}_i$ and $\bar{a}_i$ are descriptors of local keypoint geometry and appearance, respectively. Keypoint geometry $\bar{g}_i = \{\bar{x}_i, \sigma_i\}$ consists of 3D location $\bar{x}_i$ and scale σ_i, and appearance $\bar{a}_i$ is a vector of local image information, i.e. a rank-ordered histogram of oriented gradients (HOG) [36].

The kernel $\mathcal{K}(f_i, f_j)$ operates on keypoint elements $f_i = \{\bar{a}_i, \bar{g}_i\} \in A$ and $f_j = \{\bar{a}_j, \bar{g}_j\} \in B$. Here we relax the assumption of hard equivalence using squared exponential kernels with non-zero support, factored into kernels operating separately on local keypoint appearance $\mathcal{K}(\bar{a}_i, \bar{a}_j)$ and geometry $\mathcal{K}(\bar{g}_i, \bar{g}_j)$ variables:

$$\mathcal{K}(f_i, f_j) = \mathcal{K}(\bar{a}_i, \bar{a}_j)\mathcal{K}(\bar{g}_i, \bar{g}_j) \qquad (5)$$

The factorization in Equation (5) is due to the use of descriptors $\bar{a}_i$ that are invariant to 7-parameter similarity transforms of the 3D image coordinate system from which geometry $\bar{g}_i$ is derived. The two kernels in Equation (5) are defined as squared exponential functions as follows.

The appearance kernel $\mathcal{K}(\bar{a}_i, \bar{a}_j)$ is defined by the squared Euclidean distance $\|\bar{a}_i - \bar{a}_j\|_2^2$ between appearance vectors $\bar{a}_i$ and $\bar{a}_j$:

$$\mathcal{K}(\bar{a}_i, \bar{a}_j) = \exp\left(-\frac{\|\bar{a}_i - \bar{a}_j\|_2^2}{\alpha^2}\right) \qquad (6)$$

where in Equation (6), α is a bandwidth parameter that may be estimated adaptively as

$$\alpha = \min_{f_j \in \Omega} \|\bar{a}_i - \bar{a}_j\|_2^2, \ s.t. \ \|\bar{a}_i - \bar{a}_j\|_2^2 > 0, \qquad (7)$$

the minimum Euclidean distance between appearance descriptor $\bar{a}_i \in A$ and the nearest descriptor $\bar{a}_j \in \Omega \setminus A$ within the entire available dataset Ω excluding A. Note this choice of estimator is not strictly symmetric, however it allows the kernel to adapt to arbitrary dataset sizes, shrinking the resolution of prediction as the number of data grows large, and does not affect the symmetry of Equation (4).

The geometry kernel $\mathcal{K}(\bar{g}_i, \bar{g}_j)$ is novel to this work, and is defined as the product of two kernels, one modeling keypoint location conditional on keypoint scale $\mathcal{K}(\bar{x}_i, \bar{x}_j; \sigma_i, \sigma_j)$ and the $\mathcal{K}(\sigma_i, \sigma_j)$ modeling scale alone

$$\mathcal{K}(\bar{g}_i, \bar{g}_j) = \mathcal{K}(\bar{x}_i, \bar{x}_j; \sigma_i, \sigma_j)\mathcal{K}(\sigma_i, \sigma_j) \qquad (8)$$

These kernels are defined as follows

$$\mathcal{K}(\bar{x}_i, \bar{x}_j; \sigma_i, \sigma_j) = \exp\left(-\frac{\|\bar{x}_i - \bar{x}_j\|_2^2}{\sigma_i \sigma_j}\right) \qquad (9)$$

$$\mathcal{K}(\sigma_i, \sigma_j) = \exp\left(-\log^2\left(\frac{\sigma_i}{\sigma_j}\right)\right) \qquad (10)$$

Kernel $\mathcal{K}(\bar{x}_i, \bar{x}_j; \sigma_i, \sigma_j)$ in Equation (9) penalizes the squared distance between keypoint coordinates within a local reference frame, normalized by a variance proportional to the product of keypoint scales $\sigma_i \sigma_j$. This variance embodies uncertainty in keypoint location due to scale, and has a computational form that is reminiscent of mass in Newton's law of gravitation or electric charge magnitude in Coulomb's law. Figure 2 demonstrates how this kernel variance normalizes higher localization error associated with keypoints of larger scale. Kernel $\mathcal{K}(\sigma_i, \sigma_j)$ in Equation (10) penalizes multiplicative difference between keypoint scales (σ_i, σ_j).
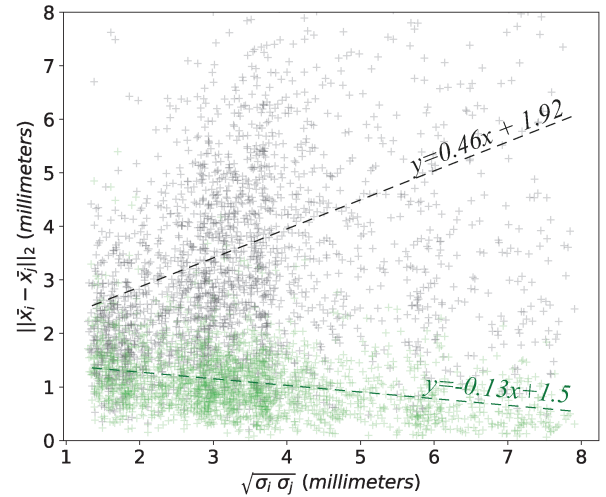

Fig. 2. Graph of spatial localization error $\|\bar{x}_i - \bar{x}_j\|_2$ vs. geometric mean scale $\sqrt{\sigma_i \sigma_j}$. Lines represent the linear regression based on each point set. Note how location error increases with scale (black) due to uncertainty, however this effect is reduced by scale normalization (green) as in Equation (9). This visualization is based on keypoint correspondences between 10 MRI aligned volumes.

Note that while the appearance kernel in Equation (6) is invariant to global similarity transforms due to descriptor invariance, the geometry kernel in Equation (8) measures zero-mean keypoint displacement and thus requires data to be aligned within a common spatial reference frame. Alignment may be established via standard subject-to-atlas registration prior to keypoint extraction or from keypoint correspondences after extraction using feature-based alignment [11]. The following workflow is used to compute pairwise distances for a set of N images:

1) Extract keypoints from each image
2) Align keypoints to an atlas template (optional for pre-aligned images)
3) Identify k-NN correspondences $(\bar{a}_i, \bar{a}_j)$ minimizing the Euclidean distance between descriptors $\|\bar{a}_i - \bar{a}_j\|$
4) Compute pairwise image distances by evaluating kernels from k-NN correspondences

Figure 3 shows the effect of our proposed geometry kernel in regularizing and favoring geometrically plausible correspondences for an example twin sibling pair.
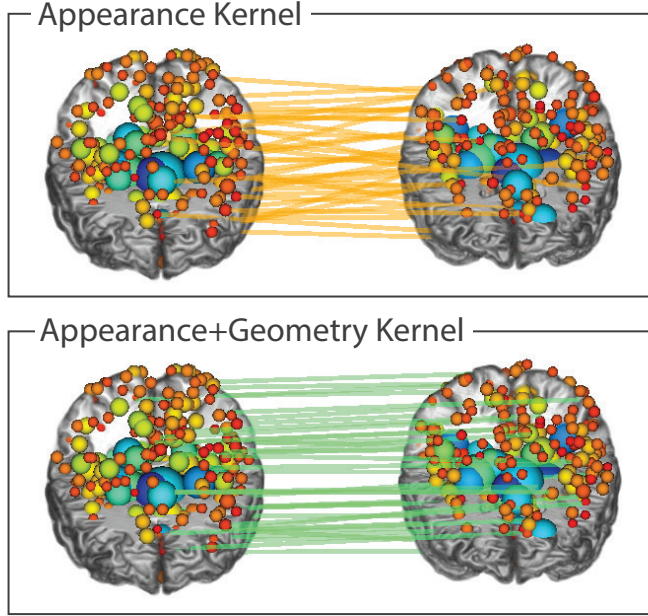

Fig. 3. Visualizing the keypoint correspondences driving classification between a pair of identical twin brains, showing the 100 correspondences with the highest SSE kernel values. The appearance kernel alone (upper, Equation (6)) may include geometrically inconsistent matches (diagonal lines), however these are down weighted when combined with the geometry kernel (lower, Equations (9) and (10)), leading to geometrically consistent matches (horizontal lines) and improved prediction performance.

IV. EXPERIMENTS

Experiments investigate the ability of the Jaccard distance to predict pairwise relationships between whole-brain MRI scans, where relationships include 1) close genetic links between siblings sharing 25-100% of their polymorphic genes and 2) broad genetic links between non-siblings sharing nominal genetic information due to sex and common racial ancestry. We hypothesize that closer genetic proximity will be reflected in higher pairwise similarity and thus lower Jaccard distance. We expect that pairwise distance based on our proposed soft set equivalence (SSE) $d_J(A, B) = -\log J_{SSE}(A, B)$ as defined in Equations (2) and (4) and our geometry kernel as defined by Equations (9) and (10) will lead to improved identification of pairwise relationship labels.

Prediction is based on pairwise Jaccard distance and a leave-one-out protocol, analogous to querying a subject MRI in a hospital PACS. Note that this is a deterministic procedure with no explicit training stage, where the primary hyperparameter is the number of nearest neighbors k per keypoint. Our computational workflow, as previously described, follows three steps 1) 3D SIFT-Rank keypoints are extracted from individual pre-processed images, 2) kNN correspondences are identified between keypoints from all N images, and 3) Jaccard distances are evaluated between all image pairs from kernels and kNN correspondences.

A. Data and Computational Details

Our data set consists of MRI scans of $N = 1010$ unique subjects from the Human Connectome Project (HCP) Q4 release [4], aged 22-36 years (mean 29 years), acquired from a diverse population including 468 males and 542 females, and 434 unique families. The MRI data are provided as T1w volumes at isotropic 0.7mm voxel resolution and preprocessed via a standard neuroimaging pipeline, including rigid subject-to-atlas registration and skull-stripping. Keypoint extraction requires approximately 3 sec. / per image and results in an average of 1,400 keypoints per image for a total of 1,488,065 keypoints. We use a GPU implementation [60] of the 3D SIFT-Rank keypoint algorithm [11] that produces identical keypoints at a $7\times$ speedup. Approximate kNN correspondences between appearance descriptors are identified across the entire database using efficient KD-tree indexing [22], where lookup requires 0.8 sec. / subject for k=200 nearest neighbors on an i7-5600@2.60Ghz machine with 16 GB RAM (1.64 GB used).

Given N=1010 images, there are a total of $N(N-1)/2 = 509545$ pairwise relationships to be evaluated. Each subject pair is assigned by one of five possible relationship labels $L = \{MZ, DZ, FS, HS, UR\}$ for monozygotic twins (MZ), dizygotic twins (DZ), full non-twin siblings (FS), half-siblings (HS) and unrelated non-siblings (UR), for totals of $\{134, 71, 607, 44, 508689\}$ relationships per label. Family relationship labels are based on mother and father identity and zygoticy (for twins), confirmed via genome-wide single nucleotide polymorphism (SNP) genotyping [61]. Note the sparse structure of family relationship labels, where each sibling label (MZ, DZ, FS, HS) is unique to one of 434 families and the vast majority of pairs are UR ($0.998 = 508689/509545$). The nominal probability of randomly guessing the correct family is approximately $0.0023 = 1/434$, and the challenge is thus to efficiently and accurately identify the small number of family relationships. Note that evaluating the similarity of $N(N-1)/2$ pairs via brute force image matching or registration becomes computationally intractable for large N.

Image alignment is fundamental to the functioning of our proposed geometry kernel, which measures zero-mean deviations in keypoint location within a common spatial reference frame. We thus evaluated prediction following two different subject-to-atlas registration methods: the original 3D rigid transform provided by the HCP mapping images to the standard MNI atlas using the FLIRT algorithm [62], [63], and feature-based alignment (FBA) [11] using keypoint correspondences to estimate a 3D similarity transform mapping images to a subject atlas, for 20 different randomly selected atlases. All alignment solutions were very close to the reference MNI alignment (rotation differed by $4.31° \pm 2.84$, translation by $3mm \pm 3.5$), reflecting a degree of inter-subject variability given different atlases. Most notably, however, was that AUC values were virtually identical for all prediction trials, with a maximum standard deviation of $\epsilon = 0.0017$ (see Table I), indicating that prediction was not sensitive to the alignment solution used. The results we report are thus consistent with the reference MNI alignment provided with the HCP data.

B. Close genetic proximity: Siblings

Figure 4 shows distributions of pairwise Jaccard distance conditioned on pairwise relationships. Our proposed kernel based on appearance and geometry (Figure 4, green) generally increases separation between different sibling relationships in comparison to appearance only (Figure 4, orange) as in [10]. A two-tailed Kolmogorov-Smirnoff (KS) test shows all distributions to be significantly different ($p-value < 1e-10$) except those of DZ and FS siblings sharing approximately 50% of their genes ($p-value = 0.0199$). Furthermore, pairwise KS statistics were all lower for appearance+geometry kernels compared to appearance only, indicating that the geometry kernels lead to increased separation of relationship categories. The Jaccard distance may be rapidly evaluated, and pairs which are outliers from their expected distributions may be easily flagged and inspected for irregularities. For example, the highest Jaccard distance outliers indicated noticeable spatial misalignment between a small number of subject pairs (e.g. Figure 4a. The lowest distance outliers for UR pairs may indicate potential sibling relationships. The blue circles near Figure 4b are pairs involving 3 subjects flagged as UR due to genotyping errors (including a MZ twin) [61]. Here, nearest neighbor Jaccard distance was used to correctly predict the self-reported families for all three cases, a result confirmed by the HCP.

As sibling pairs exhibit significantly lower distance than UR pairs, we investigate the degree to which they can be distinguished from unrelated pairs based on a simple distance threshold. Figure 5 shows the Receiver Operating Characteristic (ROC) curves for sibling relationships based on distance, comparing our SSE kernels for appearance and geometry, appearance only and the binary HSE kernel. Table I quantifies the improvement of SSE vs HSE, for various numbers of keypoint nearest neighbors (20, 100, 200), where the highest area-under-curve (AUC) values are obtained for SSE with k=200. The performance of SSE kernels generally increases with the number of NNs, and is always superior for combined appearance and geometry kernels. Conversely, HSE classification performance decreases, even falling below the diagonal in the most challenging case of HS pairs in Figure 5d, as additional hard correspondences appear to accumulate systematically on a subset of typical but unrelated subjects. For completeness, the highest UR prediction result was AUC= 0.951 for SSE, appearance and geometry kernels, k=200.

C. Distant genetic proximity: unrelated subjects

Unrelated (UR) subject pairs share nominal amounts of genetic information, with subtle similarities due to demographic factors such as common sex and ancestral race. We thus expect whole-brain distance to be lowest for pairs of the same race and sex (R,S), highest for different race and sex ($\overline{R},\overline{S}$), and intermediate for either same race (R,$\overline{S}$) or sex ($\overline{R}$,S). While the mean distances for conditional distributions in Figure 6 generally increase with differences in demographic labels, there are many exceptions where pairs with different labels exhibit lower distance than those with the same labels. We note this is consistent with pairwise genetic differences, where

pairs of unrelated individuals from different populations may often exhibit higher genetic similarity than those from the same population [64].

Age difference between subjects is a potential confounding factor in whole brain Jaccard distance, and we plotted the variation of distance vs. age difference in Figure 7. Distance distributions are virtually identical across the HCP subject age range spanning 22-36 years of age, indicating that age difference is not a major confound in this relatively young and healthy HCP cohort where brain morphology is relatively stable. Note that the Jaccard distance was found to increase with age difference in older subjects due to natural aging and neurodegenerative disease in [10].

D. Group Prediction: Sex

While the pairwise Jaccard distance is highly informative regarding sibling relationships, it is insufficient for predicting group labels such as sex, age or disease. A simple modification can be used to predict group labels, by evaluating the distance between a single keypoint set A and supersets formed by the union of group members (excluding subject A), e.g. keypoints for *all Males* $d_J(A, \mathcal{M})$ and *all Females* $d_J(A, \mathcal{F})$. Supersets $\mathcal{M}$ and $\mathcal{F}$ are solely composed of unrelated subjects to avoid potential biases due to family relationships. These distances are combined in a basic linear classifier with a single threshold parameter τ to adjust for differences in the numbers of keypoints per group based on the following equation:

$$Class(A) = \begin{cases} Female & \text{if } d_J(A, \mathcal{F}) - d_J(A, \mathcal{M}) + \tau > 0 \\ Male & \text{otherwise.} \end{cases}$$
(11)

Figure 8 shows ROCs curves for sex prediction obtained by varying τ over the range $(-\infty, \infty)$, comparing Jaccard distance computed with binary, appearance only and combined appearance and geometry kernels. A possible confound is brain size, which is on average slightly larger for males than females [65]. As the appearance kernel is invariant to image scale, the AUC=0.93 reflects prediction accuracy independently of image size. Our proposed combined approach achieves the highest AUC=0.97, again outperforming other options.

V. Discussion

In this paper we propose a novel, highly specific distance measure between volumetric images represented as invariant keypoints. Our measure generalizes the Jaccard index to account for soft equivalence between keypoints, and a novel kernel estimator is proposed to model keypoint geometry in terms of location and scale within normalized image space. The soft Jaccard index is used as a distance measure to predict pairwise relationships between MRI brain scans of 1010 subjects from 434 families, including siblings and twins, and significantly improves upon previous work based on binary equivalence and appearance descriptors [10], [12]. We report the first results for predicting relationships from medical images, a new task, where monozygotic twins can be identified with virtually perfect accuracy. A minor modification allows the Jaccard
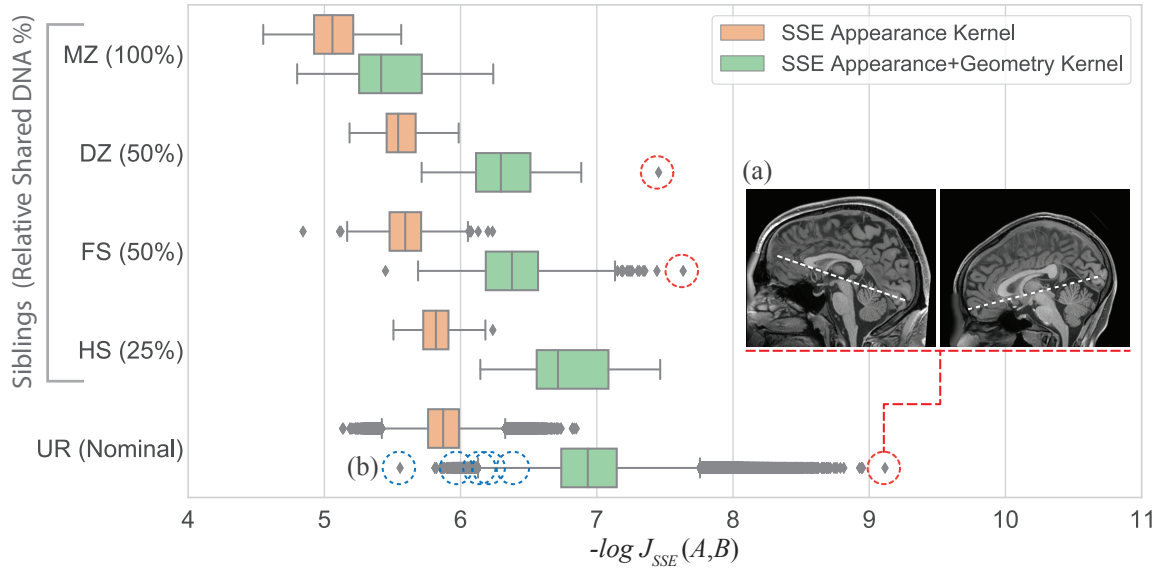
Fig. 4. Distributions of Jaccard distances $-\log J_{SSE}(A, B)$ conditioned on pairwise labels $L = \{MZ, DZ, FS, HS, UR\}$, comparing kernels with and without geometry (green vs. orange). Note that the distance generally increases with genetic separation, and that geometry kernels increase separation between sibling relationship labels. Outliers, i.e. pairs outside of their expected relationship distance distributions, can be easily identified and inspected for irregularity, e.g. (a) unusually high distances may indicate spatial misalignment (red circles) and (b) unusually low UR distances may indicate cases of incorrect family labels (blue circles).
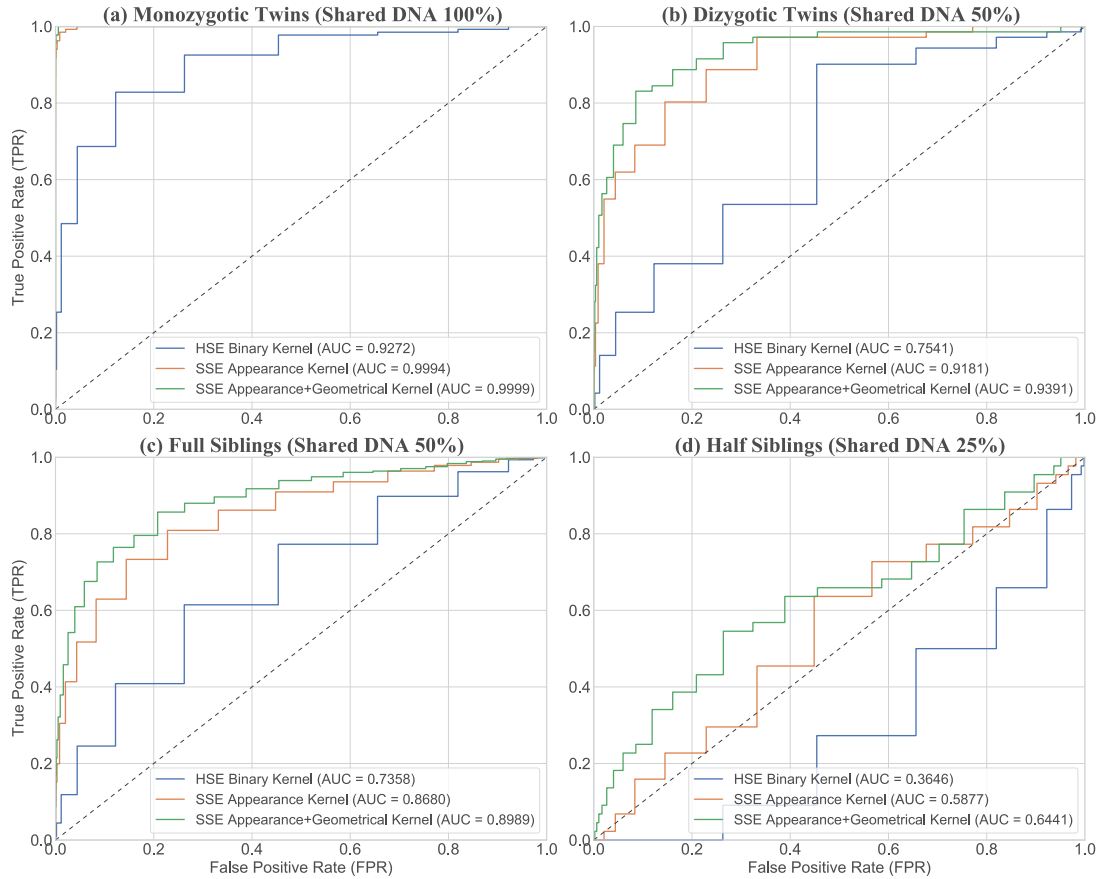

Fig. 5. ROC curves for sibling identification based on Jaccard distance for (a) MZ, (b) DZ, (c) FS and (d) HS pairs, comparing three kernels for k=200 NNs. Note that the prediction AUC is always highest for our proposed kernel combining appearance and geometry (green), and decreases as expected with the degree genetic and developmental difference, e.g. in the order MZ, DZ, FS, HS.

distance to predict group labels such as sex with high accuracy. Our geometry kernel requires spatial normalization, however

trials involving various linear subject-to-atlas alignment solutions including robust keypoint-based alignment [11] show

Kernel / Label	$k = 20$			$k = 100$			$k = 200$		
	HSE [9]	$SSE_{App.}$[10]	$SSE_{App.\,+\,Geo.}$	HSE [9]	$SSE_{App.}$[10]	$SSE_{App.\,+\,Geo.}$	HSE [9]	$SSE_{App.}$[10]	$SSE_{App.\,+\,Geo.}$
Monozygotic	0.9983	0.9993	**0.9996** $\pm\,\epsilon$	0.9544	0.9995	**0.9998** $\pm\,\epsilon$	0.9272	0.9994	**0.9999** $\pm\,\epsilon$
Dizygotic	0.8922	0.8825	**0.9044** $\pm\,\epsilon$	0.8018	0.9034	**0.9250** $\pm\,\epsilon$	0.7541	0.9181	**0.9391** $\pm\,\epsilon$
Full-Sibling	0.8423	0.8433	**0.8753** $\pm\,\epsilon$	0.7611	0.8569	**0.8888** $\pm\,\epsilon$	0.7358	0.8680	**0.8989** $\pm\,\epsilon$

TABLE I

AREA UNDER THE CURVE (AUC) VALUES FOR SIBLING CLASSIFICATION, COMPARING KERNELS AND NUMBERS OF NEAREST NEIGHBORS (K).
$\epsilon = 0.0017$ IS THE MAXIMUM AUC STANDARD DEVIATION OBSERVED FROM 21 DIFFERENT SUBJECT-TO-ATLAS REGISTRATION TRIALS. NOTE
$SSE_{App+Geo}$ KERNELS SIGNIFICANTLY OUTPERFORM SSE_{App} AND HSE FOR THE MORE DIFFICULT DZ AND FS PAIRS.
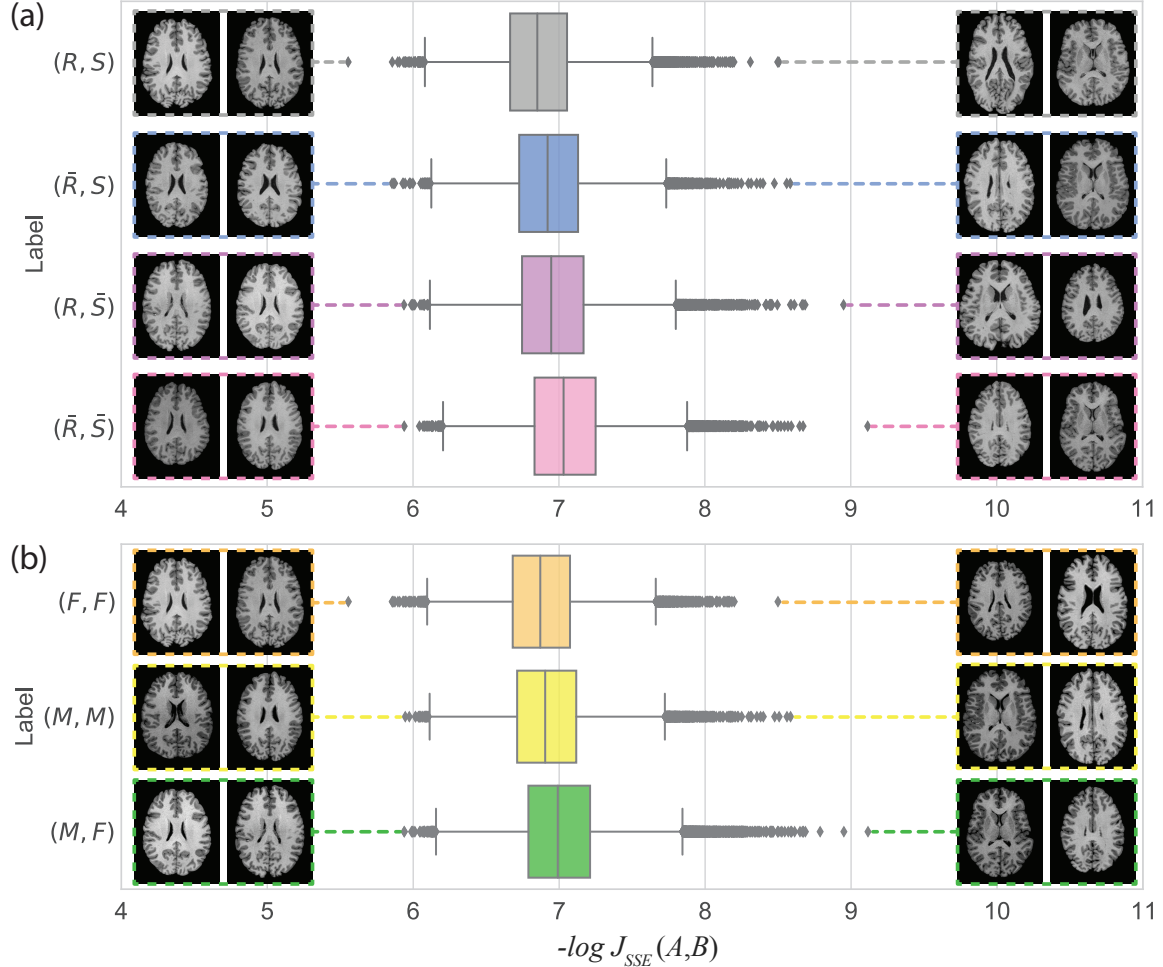

Fig. 6. Jaccard distance $-\log J_{SSE}(A, B)$ distributions between unrelated pairs conditioned on shared demographic information. **(a)** same race, same sex (R, S); different race, same sex $(\bar{R}, S)$; same race, different sex $(R, \bar{S})$, and different race, different sex $(\bar{R}, \bar{S})$ and **(b)** Female-Female (F, F), Male-Male (M, M), Male-Female (M, F). Image pairs corresponding to minimum and maximum distances for each distribution are shown for visualization.

that the prediction performance to be insensitive to the specific alignment solution used. The improved prediction afforded by our geometry kernel thus suggests that family members tend to align in a consistent manner regardless of the solution used.

Our method promises to be a useful tool for curating large medical image datasets for precision medicine and research purposes. A memory-based model using efficient and robust algorithms for 3D keypoint extraction and indexing [11], [66] allows for fine-grained comparisons between $O(N^2)$ image pairs in $O(N \log N)$ computational complexity. The Jaccard distance may thus be used to rapidly validate relationship labels, e.g. unexpectedly low distance may indicate related individuals and unexpectedly high distance may indicate unrelated individuals or spatially misaligned pairs. Previous work identified mislabelled scans of individuals in large neuroimage datasets [10], here our improved method allowed us predict the self-reported families of three subjects labelled as unrelated due to inconclusive genotyping [61], on the basis of nearest neighbor soft Jaccard distance. This exceptional result was confirmed by the Human Connectome Project, and serves as a
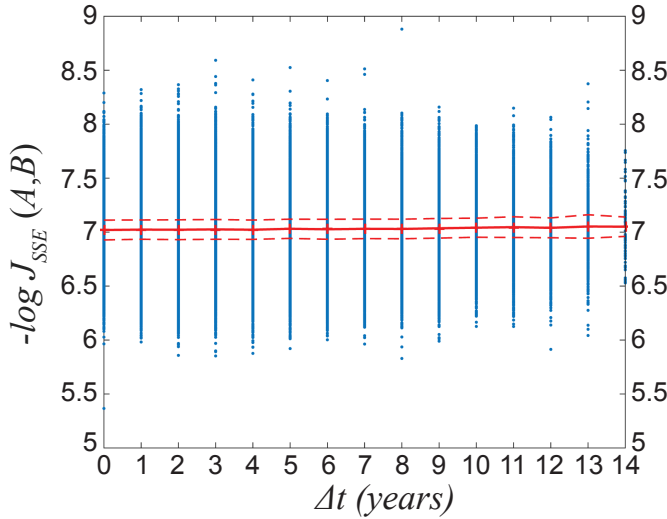
Fig. 7. Distributions of Jaccard distance $-\log J_{SSE}(A,B)$ conditioned on age difference Δt between unrelated pairs (A,B) conditioned on age difference Δt. The mean (solid red line) and standard deviation (dashed red lines) are plotted for each Δt. Age difference has no significant impact on the distance for unrelated healthy young adult brains (age 22-36 years), eliminating a possible confounding factor.
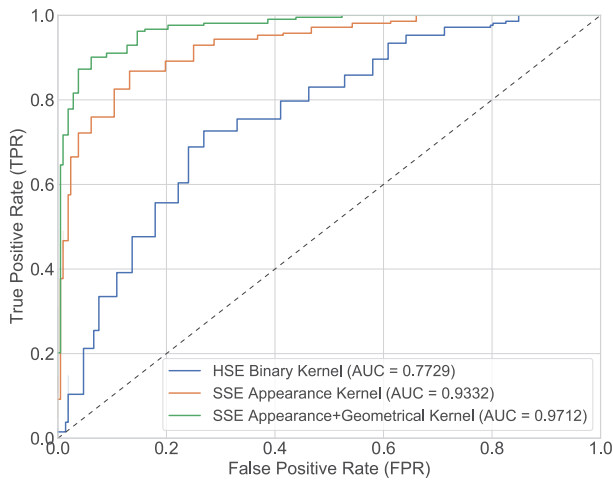

Fig. 8. ROC curves and AUC for sex prediction based on individual-to-group Jaccard distance, comparing binary, appearance only, and combined appearance and geometry kernels.

concrete example as to how our method can be used to validate labels associated with medical imaging data.

The task of pairwise relationship prediction may be viewed as the finest grain of categorization, where the number of unique pairwise relationship labels is linear $O(N)$ in the number of data N, e.g. 434 families from N=1010 images, as opposed to typical classification where all data are associated with a small number of labels (e.g. male, female). Pairwise prediction thus represents a challenge for ubiquitous deep neural network methods, which generally require large numbers of training data per category, and training for specific modalities and body parts. In contrast, generic 3D SIFT keypoints may be used as-is with arbitrary imaging modalities and contexts with no training. In future work, generic keypoint correspondences here could potentially be used to train

domain-specific keypoint models [39], [42]. Auto-encoders used for anomaly detection could potentially be adapted to generate subject-specific codes between pairs of images [67]. Analysis beyond pairs of subjects could be generalized via graph theory to model the clique structure of families [68]. All code required to reproduce our results may be obtained at https://github.com/3dsift-rank.

REFERENCES

[1] M. A. Hamburg and F. S. Collins, "The path to personalized medicine," *New England Journal of Medicine*, vol. 363, no. 4, pp. 301–304, 2010.

[2] G. J. Annas and S. Elias, "23andme and the fda," *New England Journal of Medicine*, vol. 370, no. 11, pp. 985–988, 2014.

[3] A. Auton *et al.*, "A global reference for human genetic variation," *Nature*, vol. 526, no. 7571, p. 68, 2015.

[4] D. Van Essen *et al.*, "The Human Connectome Project: A data acquisition perspective," *NeuroImage*, vol. 62, no. 4, pp. 2222–2231, 2012.

[5] S. Gerber, T. Tasdizen, P. Thomas Fletcher, S. Joshi, R. Whitaker, and Alzheimer's Disease Neuroimaging Initiative (ADNI), "Manifold modeling for brain population analysis." *Medical Image Analysis (MIA)*, vol. 14, no. 5, pp. 643–53, 2010.

[6] T. Brosch, P. Gmbh, I. Technologies, and R. Tam, "Manifold Learning of Brain MRIs by Deep Learning Manifold Learning of Brain MRIs," *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, vol. 16, no. September, 2013.

[7] M. Ono, S. Kubik, and C. D. Abernathey, *Atlas of the cerebral sulci*. Thieme Medical Publishers, 1990.

[8] M. Levandowsky and D. Winter, "Distance between sets," *Nature*, vol. 234, no. 5323, p. 34, 1971.

[9] M. Toews and W. Wells, "How are siblings similar? How similar are siblings? Large-scale imaging genetics using local image features," in *International Symposium on Biomedical Imaging (ISBI)*. IEEE, 4 2016, pp. 847–850.

[10] L. Chauvin *et al.*, "Neuroimage signature from salient keypoints is highly specific to individuals and shared by close relatives," *NeuroImage*, vol. 204, no. 20, 2020.

[11] M. Toews and W. Wells, "Efficient and robust model-to-image alignment using 3D scale-invariant features," *Medical Image Analysis (MIA)*, vol. 17, no. 3, pp. 271–282, 2013.

[12] K. Kumar, M. Toews, L. Chauvin, O. Colliot, and C. Desrosiers, "Multimodal brain fingerprinting: a manifold approximation based framework," *NeuroImage*, 2018.

[13] D. S. Marcus, T. H. Wang, J. Parker, J. G. Csernansky, J. C. Morris, and R. L. Buckner, "Open Access Series of Imaging Studies (OASIS): cross-sectional MRI data in young, middle aged, nondemented, and demented older adults," *Journal of cognitive neuroscience*, vol. 19, no. 9, pp. 1498–1507, 2007.

[14] C. R. Jack Jr *et al.*, "The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods," *Journal of the International Society for Magnetic Resonance in Medicine (ISMRM)*, vol. 27, no. 4, pp. 685–691, 2008.

[15] D. C. Van Essen *et al.*, "The WU-Minn human connectome project: an overview," *NeuroImage*, vol. 80, pp. 62–79, 2013.

[16] L. Chauvin, K. Kumar, C. Desrosiers, J. De Guise, W. Wells, and M. Toews, "Analyzing Brain Morphology on the Bag-of-Features Manifold," in *International Conference on Information Processing in Medical Imaging (IPMI)*, vol. 11492 LNCS. Springer Verlag, 2019, pp. 45–56.

[17] M. R. Sabuncu *et al.*, "Morphometricity as a measure of the neuroanatomical signature of a trait," *Proceedings of the National Academy of Sciences (PNAS)*, vol. 113, no. 39, pp. E5749–E5756, 2016.

[18] O. Boiman, E. Shechtman, and M. Irani, "In defense of nearest-neighbor based image classification," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2008, pp. 1–8.

[19] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Transactions on Information Theory (TIT)*, vol. 13, no. 1, pp. 21–27, 1967.

[20] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision (IJCV)*, vol. 60, no. 2, pp. 91–110, 2004.

[21] T. Lindeberg, "Feature Detection with Automatic Scale Selection," *International Journal of Computer Vision (IJCV)*, vol. 30, no. 2, pp. 79 – 116, 1998.

[22] M. Muja and D. Lowe, "Scalable nearest neighbor algorithms for high dimensional data," *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, pp. 2227–2240, 2014.

[23] M. Toews, C. Wachinger, R. S. J. Estepar, and W. Wells, "A Feature-Based Approach to Big Data Analysis of Medical Images." *International Conference on Information Processing in Medical Imaging (IPMI)*, vol. 24, pp. 339–350, 2015.

[24] E. Marczewski and H. Steinhaus, "On a certain distance of sets and the corresponding distance of functions," in *Colloquium Mathematicum*, vol. 6, no. 1. Institute of Mathematics Polish Academy of Sciences, 1958, pp. 319–327.

[25] D. Bajusz, A. Rácz, and K. Héberger, "Why is Tanimoto index an appropriate choice for fingerprint-based similarity calculations?" *Journal of Cheminformatics*, vol. 7, no. 1, 2015.

[26] D. J. Rogers and T. T. Tanimoto, "A computer program for classifying plants," *Science*, vol. 132, no. 3434, pp. 1115–1118, 1960.

[27] Y. Yuan, M. Chao, and Y. C. Lo, "Automatic Skin Lesion Segmentation Using Deep Fully Convolutional Networks with Jaccard Distance," *IEEE Transactions on Medical Imaging (TMI)*, vol. 36, no. 9, 2017.

[28] D. Molodtsov, "Soft set theory—first results," *Computers & Mathematics with Applications*, vol. 37, no. 4-5, pp. 19–31, 1999.

[29] J. H. Park, O. H. Kim, and Y. C. Kwun, "Some properties of equivalence soft set relations," *Computers & Mathematics with Applications*, vol. 63, no. 6, pp. 1079–1088, 2012.

[30] A. Gardner, J. Kanno, C. A. Duncan, and R. Selmic, "Measuring Distance Between Unordered Sets of Different Sizes," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014, pp. 137–143.

[31] O. Chum, J. Philbin, and A. Zisserman, "Near Duplicate Image Detection: min-Hash and tf-idf Weighting," in *The British Machine Vision Conference (BMVC)*, 2008.

[32] I. Machado *et al.*, "Non-rigid registration of 3D ultrasound for neurosurgery using automatic feature detection and matching," *International Journal of Computer Assisted Radiology and Surgery (IJCARS)*, vol. 13, no. 10, pp. 1525–1538, 2018.

[33] C. Wachinger, M. Toews, G. Langs, W. Wells, and P. Golland, "Keypoint transfer for fast whole-body segmentation," *IEEE Transactions on Medical Imaging (TMI)*, 2018.

[34] M. Toews, W. Wells, D. L. Collins, and T. Arbel, "Feature-based morphometry: Discovering group-related anatomical patterns," *NeuroImage*, vol. 49, no. 3, pp. 2318–2327, 2010.

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

[36] M. Toews and W. Wells, "Sift-rank: Ordinal description for invariant feature correspondence," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2009, pp. 172–177.

[37] D. Mukherjee, Q. M. Jonathan Wu, and G. Wang, "A comparative experimental study of image feature detectors and descriptors," *Machine Vision and Applications*, vol. 26, no. 4, pp. 443–466, 2015.

[38] F. Bellavia and C. Colombo, "Is there anything new to say about sift matching?" *International Journal of Computer Vision*, vol. 128, no. 7, 2020.

[39] K. M. Yi, E. Trulls, V. Lepetit, and P. Fua, "LIFT: Learned invariant feature transform," in *Proceedings of the European Conference on Computer Vision (ECCV)*, vol. 9910 LNCS. Springer Verlag, 2016, pp. 467–483.

[40] M. J. Tyszkiewicz, P. Fua, and E. Trulls, "DISK: Learning local features with policy gradient," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2020.

[41] Y. Ono, E. Trulls, P. Fua, and K. Moo Yi, "LF-Net: Learning Local Features from Images," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2018.

[42] D. Detone, T. Malisiewicz, and A. Rabinovich, "SuperPoint: Self-Supervised Interest Point Detection and Description," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2018.

[43] J. Dong and S. Soatto, "Domain-size pooling in local descriptors: DSP-SIFT," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.

[44] R. Arandjelović and A. Zisserman, "Three things everyone should know to improve object retrieval," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012.

[45] V. Balntas, K. Lenc, A. Vedaldi, and K. Mikolajczyk, "HPatches: A benchmark and evaluation of handcrafted and learned local descriptors," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.

[46] J. L. Schönberger, H. Hardmeier, T. Sattler, and M. Pollefeys, "Comparative Evaluation of Hand-Crafted and Learned Local Features," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.

[47] R. Geirhos, C. Michaelis, F. A. Wichmann, P. Rubisch, M. Bethge, and W. Brendel, "Imagenet-trained CNNs are biased towards texture; increasing shape bias improves accuracy and robustness," in *International Conference on Learning Representations (ICLR)*, 2019.

[48] C. Wachinger *et al.*, "BrainPrint: A discriminative characterization of brain morphology," *NeuroImage*, vol. 109, pp. 232–248, 2015.

[49] S. A. Valizadeh, F. Liem, S. Mérillat, J. Hänggi, and L. Jäncke, "Identification of individual subjects on the basis of their brain anatomical features," *Scientific reports*, vol. 8, no. 1, p. 5611, 2018.

[50] E. S. Finn *et al.*, "Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity." *Nature neuroscience*, vol. 18, no. 11, pp. 1664–71, 2015.

[51] K. Kumar, C. Desrosiers, K. Siddiqi, O. Colliot, and M. Toews, "Fiberprint: A subject fingerprint based on sparse code pooling for white matter fiber analysis," *NeuroImage*, vol. 158, pp. 242–259, 2017.

[52] P. M. Thompson *et al.*, "Genetic influences on brain structure," *Nature neuroscience*, vol. 4, no. 12, p. 1253, 2001.

[53] S. J. van der Lee *et al.*, "Gray matter heritability in family-based and population-based studies using voxel-based morphometry," *Human Brain Mapping (HBM)*, vol. 38, no. 5, 2017.

[54] G. L. Colclough *et al.*, "The heritability of multi-modal connectivity in human brain activity." *eLife*, vol. 6, p. e20178, 2017.

[55] M. R. Kiley and M. S. Hossain, "Who are my family members? a solution based on image processing and machine learning," *International Journal of Image and Graphics*, vol. 20, no. 04, p. 2050033, 2020.

[56] Q. Dai, P. Carr, L. Sigal, and D. Hoiem, "Family member identification from photo collections," in *2015 IEEE Winter Conference on Applications of Computer Vision*. IEEE, 2015, pp. 982–989.

[57] B. Ahmad, M. Usama, J. Lu, W. Xiao, J. Wan, and J. Yang, "Deep convolutional neural network using triplet loss to distinguish the identical twins," in *2019 IEEE Globecom Workshops (GC Wkshps)*. IEEE, 2019, pp. 1–6.

[58] F. Juefei-Xu and M. Savvides, "An augmented linear discriminant analysis approach for identifying identical twins with the aid of facial asymmetry features," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2013.

[59] J. R. Paone *et al.*, "Double trouble: Differentiating identical twins by face recognition," *IEEE Transactions on Information Forensics and Security (TIFS)*, vol. 9, no. 2, 2014.

[60] É. Pepin, J.-B. Carluer, L. Chauvin, M. Toews, and R. Harmouche, "Large-scale unbiased neuroimage indexing via 3d gpu-sift filtering and keypoint masking," in *Machine Learning in Clinical Neuroimaging and Radiogenomics in Neuro-oncology*. Springer, 2020, pp. 108–118.

[61] H. WU-Minn, "1200 subjects data release reference manual," 2017. [Online]. Available: https://www.humanconnectome.org

[62] M. F. Glasser *et al.*, "The minimal preprocessing pipelines for the Human Connectome Project," *NeuroImage*, vol. 80, pp. 105–124, 2013.

[63] M. Jenkinson, P. Bannister, M. Brady, and S. Smith, "Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images," *NeuroImage*, pp. 825–841, 2002.

[64] D. J. Witherspoon *et al.*, "Genetic similarities within and between human populations," *Genetics*, vol. 176, no. 1, pp. 351–359, 2007.

[65] L. Eliot, A. Ahmed, H. Khan, and J. Patel, "Dump the "dimorphism": Comprehensive synthesis of human brain studies reveals few male-female differences beyond size," *Neuroscience & Biobehavioral Reviews*, 2021.

[66] M. Toews, L. Zöllei, and W. Wells, "Feature-based alignment of volumetric multi-modal images," in *International Conference on Information Processing in Medical Imaging (IPMI)*, vol. 23. Springer, 2013, pp. 25–36.

[67] C. Baur, S. Denner, B. Wiestler, N. Navab, and S. Albarqouni, "Autoencoders for unsupervised anomaly segmentation in brain MR images: A comparative study," *Medical Image Analysis (MIA)*, vol. 69, 2021.

[68] J. R. Blair and B. Peyton, "An introduction to chordal graphs and clique trees," in *Graph theory and sparse matrix computation*. Springer, 1993, pp. 1–29.