

Analyzing Brain Morphology on the Bag-of-Features Manifold

Laurent Chauvin¹, Kuldeep Kumar¹, Christian Desrosiers¹,
Jacques De Guise¹, William Wells III^{2,3} and Matthew Toews¹

¹ École de Technologie Supérieure, Montreal, Canada
laurent.chauvin0@gmail.com

² Brigham and Women’s Hospital, Harvard Medical School, Boston, US

³ Computer Science and Artificial Intelligence Laboratory,
Massachusetts Institute of Technology, Boston, US

Abstract. We propose a novel distance measure between variable-sized sets of image features, i.e. the bag-of-features image representation, for quantifying brain morphology similarity based on local neuroanatomical structures. Our measure generalizes the Jaccard distance metric to account for probabilistic or soft set equivalence (SSE), via a novel adaptive kernel density framework accounting for probabilistic uncertainty in both feature appearance and geometry. The method is based on highly efficient keypoint feature indexing and is suitable for identifying pairwise relationships in arbitrarily large data sets. Experiments use the Human Connectome Project (HCP) dataset consisting of 1010 subjects, including pairs of siblings and twins, where neuroanatomy is modeled as a set of scale-invariant keypoints extracted from T1-weighted MRI data. The Jaccard distance based on (SSE) is shown to outperform standard hard set equivalence (HSE) in predicting the immediate family graph structure and genetic links such as racial origin and sex from MRI data, providing a useful tool for data-driven, high-throughput genome wide heritability analysis.

1 Introduction

Health treatment is increasingly personalized, where treatment decisions are increasingly conditioned on patient-specific information in addition to knowledge regarding the general population, i.e. personalized medicine [7]. Modern genetic testing allows us to cheaply predict patient-specific characteristics, including immediate family relationships, and also characteristics shared across the population including racial origin, sex, hereditary disease status, etc. [1], based on a large library of human DNA samples. The brain, the root of cognition, is 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 50-100% of their genes, or between wider sub-populations defined by more subtle genetic factors such as racial origin or sex?

Large, publicly available databases allow us investigate these questions with large samples of coupled MRI and genetic data i.e. 1000+ subjects [23]. Brain shape has been modeled as lying on a smooth manifold in high dimensional MRI data space [6, 2], where phenotype can be described as a smooth deformation conditioned on developmental factors determined by the environment. However the brain is naturally described as a rich collection of spatially localized neuroanatomical structures, including common structures such as the basal ganglia shared across the population, and also intricate cortical folding patterns [15] or pathology only observable in subsets of the population.

An intuitively appealing representation for such phenomena is the bag-of-feature (BoF) model, where the image is described as a set of localized, conditionally independent descriptors identified at image keypoints, i.e. 3D SIFT-Rank features [21]. Bags-of-features can be viewed as existing on a high-dimensional manifold, and medical imaging applications such as regression or classification can be formulated in terms of a suitable geodesic distance metric between BoFs. As BoFs are variable sized sets, typical metrics based on fixed-length vectors such as L-norms [6, 2] do not readily apply. The Jaccard distance metric [10] has proven effective in recent studies investigating genetics and brain MRI [22, 8]. For example, by defining hard set equivalence (HSE) in terms of binary k-Nearest Neighbor (kNN) correspondences between keypoint image descriptors, the Jaccard distance metric can be used to approximate the genetic structure of brain MRIs, predicting family relationships including twins and siblings with high accuracy. The disadvantage of the Jaccard metric is that binary equivalence of feature descriptors is crude approximation given probabilistic uncertainty inherent to natural image structure, and is ill-defined for variable sized datasets where the number of relevant kNN correspondences may be unknown a-priori or variable.

Our primary contribution is to introduce a new Jaccard-like measure between BoFs, modeling probabilistic or soft set equivalence (SSE) [5] between elements. This is achieved via an adaptive kernel density estimator (KDE) [19] accounting for uncertainty in both local feature appearance and geometry, automatically adapting to the quantity of available feature data. A novel kernel is introduced that normalizes variability in pairwise feature displacement by the geometrical mean of feature scales, accounting for localization uncertainty in scale-space. This work extends approaches to analyzing neuroimaging data via local features [22, 8], where the Jaccard distance based on binary or hard set equivalence (HSE) is defined by kNN correspondences.

2 Related Work

The BoF data model was initially used to categorize text documents from locally orderless collections of words [12], then adapted to categorize images [4] following the development of robust image keypoint detectors, e.g. the scale-invariant feature transform (SIFT) [11]. The keypoint representation serves as a highly

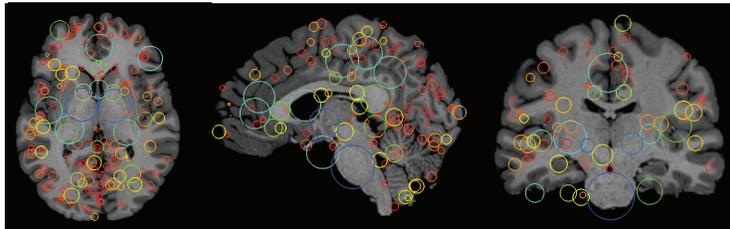


Fig. 1. Illustrating a BoF of 3D SIFT-Rank features [21] extracted from a T1w image from the Human Connectome Project dataset [23]. Circles represent feature location and scale, color also indicates scale.

robust and efficient basis for medical image analysis applications [21, 25], and continues to be a research focus in deep learning [16].

Experimentally, our work is most closely related to neuro fingerprinting methods [24, 22, 9, 8, 3], which can be used to investigate the relationship between genotype and phenotype from magnetic resonance images (MRI) of the human brain from increasingly large, publicly available neuroimaging datasets [23, 18]. The local feature framework is highly robust and requires minimal data preprocessing and scales gracefully to arbitrarily large data sets, as robust, approximate nearest neighbor (NN) correspondences can be identified in $O(\log N)$ complexity given a dataset of N subjects via efficient KD-tree data structures [13]. This allows evaluating all $N(N - 1)/2$ pairwise relationships between subjects in $O(N \log N)$ computational complexity, and has resulted in highest reported accuracy in predicting genetic sibling relationships from structural MRI [8], to our knowledge. Here we compare our SSE Jaccard measure to the standard HSE-based measure used in [22, 8], showing a significant improvement in prediction accuracy.

3 Method

Our method is based on bag-of-features (BoF) representation, where each image is represented as a set of local 3D scale invariant features $\{f_i\}$ automatically extracted from the image as described in [21], primarily for the purpose of highly efficient and robust feature indexing operations. We seek a pairwise distance measure between BoFs $A = \{f_i\}$ and $B = \{f_j\}$ which can be used to estimate proximity and thus genetic relationships between subjects. Previous work has adopted the Jaccard distance metric [10] $d_J(A, B)$ defined as:

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

where $J_{HSE}(A, B)$ is the Jaccard index based on binary equivalence relationships, e.g. hard kNN feature correspondences [22]. We seek a measure based on soft set equivalence in order to more accurately model probabilistic similarity

between local features extracted from natural image data. We do this in a manner similar to [5] by redefining hard set intersection $|A \cap B|$ in Equation (1) by a suitable equivalent $\mu(A \cap B)$, where $0 \leq \mu(A \cap B) \leq |A \cap B| \leq \min\{|A|, |B|\}$, leading to a Jaccard-like measure based on soft set equivalence:

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

Defining $\mu(A \cap B)$: We begin with a generative model of the conditional probability $p(A|B)$ of set A given set B :

$$p(A|B) = \prod_i^{ |A| } p(f_i|B) = \prod_i^{ |A| } \left(\eta + \sum_j^{ |B| } p(f_i|f_j) \right) \quad (3)$$

$$\approx \prod_i^{ |A| } \left(\eta + \max_{f_j \in B} \{p(f_i|f_j)\} \right) \quad (4)$$

where the first equality in Equation (3) is due to the assumption of conditionally IID feature samples $f_i \in A$ and the second from a mixture density model of $p(f_i|B)$ consisting of kernel components $p(f_i|f_j)$ parameterized by $f_j \in B$ and a uniform background component η accounting spurious, noisy features. The approximation in Equation (4) can be used when a distinctive invariant feature f_i typically arises from at most a single mixture component, i.e. a unique image-to-image correspondence. Assuming an exponential kernel function $0 \leq p(f_i|f_j) \leq 1$, the logarithm of Equation (4) is bounded as

$$0 \leq \log p(A|B) = \sum_i^{ |A| } \log \left(\eta + \max_{f_j \in B} \{p(f_i|f_j)\} \right) \leq |A| \log(\eta + 1) \quad (5)$$

We define soft set intersection as follows, and show an upper bound:

$$0 \leq \mu(A \cap B) = \min \left\{ \frac{\log p(A|B)}{\log(\eta + 1)}, \frac{\log p(B|A)}{\log(\eta + 1)} \right\} \leq |A \cap B| \quad (6)$$

Defining $p(f_i|f_j)$: Finally we define the form of the kernel function $p(f_i|f_j)$ used. A scale invariant feature $f_i = \{\mathbf{a}_i, \mathbf{g}_i\}$ is defined by geometry \mathbf{g}_i and an appearance descriptor \mathbf{a}_i . Feature geometry $\mathbf{g}_i = \{\mathbf{x}_i, \sigma_i\}$, consists of 3D location \mathbf{x}_i and scale σ_i . Appearance \mathbf{a}_i is a vector of local image information, i.e. a rank-ordered histogram of oriented gradients (HOG) [21]. The kernel $p(f_i|f_j)$ is factored as

$$p(f_i|f_j) = p(\mathbf{a}_i|\mathbf{a}_j)p(\mathbf{g}_i|\mathbf{g}_j) = p(\mathbf{a}_i|\mathbf{a}_j)p(\mathbf{x}_i|\sigma_i, \mathbf{g}_j)p(\sigma_i|\mathbf{g}_j). \quad (7)$$

In Equation (7), the first equality expresses conditional independence between variables of appearance \mathbf{a}_i and geometry \mathbf{g}_i due to the use of geometrical invariant appearance descriptors. The second equality factors location \mathbf{x}_i and scale σ_i

using the chain rule of probability. Factors $p(\mathbf{a}_i|\mathbf{a}_j)$, $p(\mathbf{x}_i|\sigma_i, \mathbf{g}_j)$ and $p(\sigma_i|\mathbf{g}_j)$ are defined as follow:

$$p(\mathbf{a}_i|\mathbf{a}_j) = \exp\left(-\frac{\|\mathbf{a}_i - \mathbf{a}_j\|_2^2}{\alpha^2}\right) \quad (8)$$

$$p(\mathbf{x}_i|\sigma_i, \mathbf{g}_j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|_2^2}{\sigma_i^2 \sigma_j^2}\right) \quad (9)$$

$$p(\sigma_i|\mathbf{g}_j) = \exp\left(-\log^2\left(\frac{\sigma_i}{\sigma_j}\right)\right) \quad (10)$$

Kernel $p(\mathbf{a}_i|\mathbf{a}_j)$ in Equation (8) penalizes differences in appearance descriptors, where α is an adaptive bandwidth of the KDE, defined as

$$\alpha = \min_{f_j \in \Omega} (d(\mathbf{a}_i, \mathbf{a}_j)), \text{ s.t. } d(\mathbf{a}_i, \mathbf{a}_j) > 0 \quad (11)$$

with $d(\mathbf{a}_i, \mathbf{a}_j)$ is the minimum Euclidean distance between appearance descriptors \mathbf{a}_i and \mathbf{a}_j across the entire dataset. This allows the kernel to adapt to arbitrarily data set sizes, shrinking as the number of data grows large. Kernel $p(\mathbf{x}_i|\sigma_i, \mathbf{g}_j)$ in Equation (9) is a novel formulation penalizing distance between features in normalized image space, where the variance is proportional to the product of feature scales $\sigma_i^2 \sigma_j^2$. This variance embodies uncertainty in feature localization due to scale, and is reminiscent of mass in Newton’s gravitation or electric charge magnitude in Coulomb’s law. Finally, Kernel $p(\sigma_i|\mathbf{g}_j)$ in Equation (10) penalizes multiplicative difference in feature scale in normalized image space.

4 Experiments

The experiments aim to investigate the link between genetics and phenotypical neuroanatomical structure [20] by comparing the proximity graph structure induced by the Jaccard distance measure $d_J(A, B) = -\log J_{SSE}(A, B)$ between pairs of subjects (A, B) from MRI (BoF) observations vs. known genetic relationships. Unless mentioned otherwise, $d_J(A, B)$ will refer to the Jaccard distance with soft set equivalence (SSE).

We compare the set intersection $A \cap B$ defined by standard hard set equivalence (HSE) as in Equation (1) vs. our soft set equivalence (SSE) definition in Equations (2) and (6), with a background distribution η empirically set to 1. We consider the cases of 1) close genetic links between immediate family members sharing 50-100% of their genes and 2) broad genetic links between unrelated non-siblings sharing nominal genetic information due to common racial origin, sex, etc.

Our test set consists of MRI scans of $N = 1010$ unique subjects with genetic ground truth from the Human Connectome Project Q4 release [23], aged 22-36 years (mean 29 years), 468 males and 542 females. There are thus $N(N-1)/2 = 1,020,100$ pairwise relationships, where each pair of subjects is related by one

of four possible relationship labels $L = \{MZ, DZ, FS, UR\}$ for monozygotic twins (MZ), dizygotic twins (DZ), full non-twin siblings (FS) and unrelated non-siblings (UR). Note that evaluating the similarity of $N(N - 1)/2$ pairwise relationships via brute force image registration, e.g. optimizing cross-correlation, quickly becomes computationally intractable as N grows large.

MRI data used consist of skull-stripped T1w images with 0.7mm isotropic voxels, registered to a common reference frame, however the method generally applies to other scalar image modalities (e.g. T2w, FA). Generic 3D SIFT-Rank features are extracted from MRI data, where geometry \mathbf{g}_i is identified as extrema of an difference-of-Gaussian scale-space [11] and local appearance is encoded as a 64-dimensional SIFT-Rank [21] appearance descriptor \mathbf{a}_i . Feature extraction requires approximately 20 sec. / per image and results in an average of 1,400 features per image, for a total of 1,488,065 features. Approximate kNN correspondences between appearance descriptors are identified across the entire database using KD-tree indexing [13], lookup requires 0.8 sec. / subject for k=200 on an i7-5600@2.60Ghz machine with 16 GB RAM (1.64 GB used).

4.1 Close genetic proximity: Siblings

The Jaccard distance quantifies the whole-brain dissimilarity between brain pairs. As the percentage of genes shared is 100% for MZ pairs, on average 50% for DZ and FS pairs, and a nominally low percentage for UR pairs, we hypothesize that the pairwise distances will respect the same ordering on the BoF manifold. Figure 2 shows distributions of all pairwise Jaccard distances conditional on labels, showing that this is indeed the case, with MZ pairs closest, DZ and FS pairs further, and finally UR pairs the farthest.

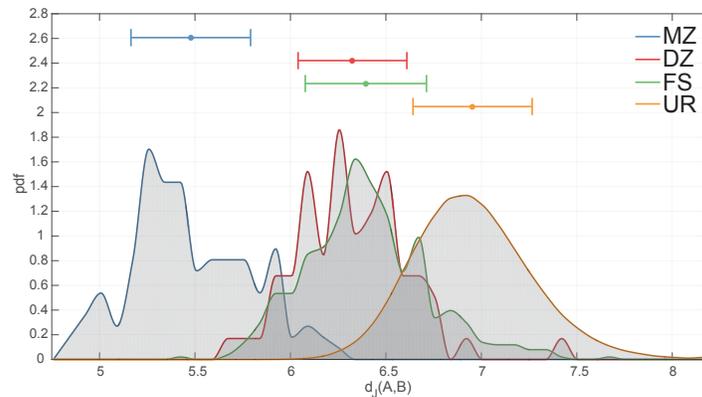


Fig. 2. Distributions of conditional Jaccard distances $p(d_J(A, B)|L)$ conditioned on pairwise labels $L = \{MZ, DZ, FS, UR\}$. $d_J(A, B)$ is evaluated with SSE and NN=200. Multiple peaks in MZ, DZ and FS distributions result from sampling sparsity.

As sibling pairs (MZ,DZ,FS) exhibit significantly lower manifold distance than UR pairs, we investigate the degree to which siblings can be distinguished from unrelated pairs based on distance. Figure (3) shows the Receiver Operating Characteristic (ROC) curves for MZ, DZ and FS relationships based on their proximity on the bag-of-feature manifold, comparing Jaccard distances for hard (HSE) and soft set equivalence (SSE) for various numbers of nearest neighbors (20, 100, 200). Several observations can be made. First, the Jaccard measure we propose for soft set equivalence is always superior to hard set equivalence in terms of classification accuracy. HSE is noticeably sensitive to the number of nearest neighbor (NN) feature correspondences, and the classification performance generally decreases with an increase in the number of NNs. In contrast, the SSE is relatively stable and increases with the number of NNs used. Intuitively, this is because the number NN correspondences above the number of sibling pairs in the data will generally only contribute noise in HSE distance measurements, but these are down-weighted in the SSE based feature proximity. The area under the ROC curve (AUC) values in Table 1 quantify the improvement of SSE vs HSE, where the highest AUC values are obtained for SSE with 200 NN correspondences. The results for HSE are consistent with the work of [8], and our SSE results can thus be considered the state-of-the-art for sibling retrieval from structure MRI.

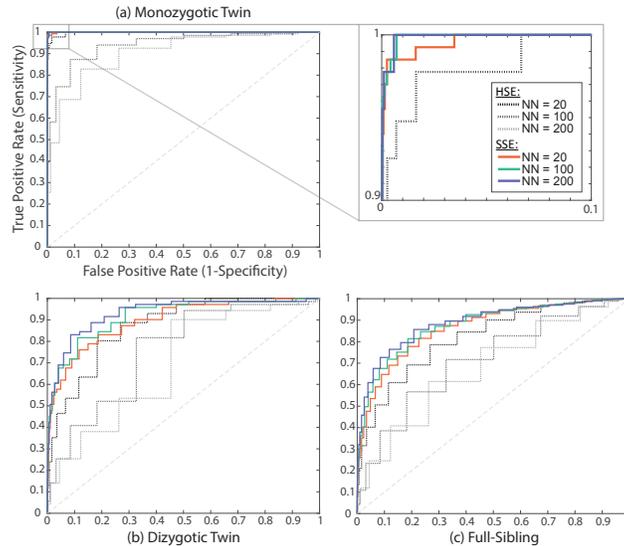


Fig. 3. ROC curves for sibling identification based on Jaccard distance for (a) MZ, (b) DZ and (c) FS pairs. Curves compare hard set equivalence (HSE) vs. the proposed soft set equivalence (SSE) for three values of nearest neighbor (NN) keypoint correspondences (NN=20,100,200).

Label	$NN = 20$		$NN = 100$		$NN = 200$	
	<i>HSE</i>	<i>SSE</i>	<i>HSE</i>	<i>SSE</i>	<i>HSE</i>	<i>SSE</i>
Monozygotic	0.9983	0.9996	0.9544	0.9998	0.9272	0.9999
Dizygotic	0.8922	0.9044	0.8018	0.9250	0.7541	0.9391
Full-Sibling	0.8423	0.8753	0.7611	0.8888	0.7358	0.8989

Table 1. Area Under the Curve (AUC) for monozygotic (MZ), dizygotic (DZ) and full-sibling (FS) classification with HSE and SSE methods for 20, 100 and 200 nearest neighbors (NN). SSE method consistently perform better, with an accuracy increasing with the number of nearest neighbors.

4.2 Distant genetic proximity: unrelated subjects

As compared to siblings, unrelated (UR) subject pairs exhibit a much lower genetic overlap, with subtle similarities due to factors such as common racial origin and sex, and accordingly higher Jaccard distance. We thus expect subtle variations in whole-brain Jaccard distance, with lowest distance for UR pairs of the same race and sex (R,S), medium distance for either same race (\bar{R},\bar{S}) or sex (\bar{R},S), and highest distance for different race and sex (\bar{R},\bar{S}). Figure 4a shows that the conditional distributions of Jaccard distance conditional on race and sex are consistent with this expectation. Note that the mean distance for (\bar{R},S) is lower than (R,\bar{S}) indicating, that sex may be a stronger determinant of whole brain similarity than racial origin.

We investigate more closely the relationship between whole brain distance and pairwise combinations of male/female sex labels in Figure 4 (b). We see that while the Jaccard distance distributions for same sex pairs (F-F, M-M) are highly similar in terms of mean and variance, different sex pairs (F-M) generally exhibit a significantly larger distance, as expected.

A potential confound in whole brain Jaccard distance is age difference. Figure 5 plots the variation in Jaccard distance vs. age difference. 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 HCP cohort where brain morphology is relatively stable. We expect a greater impact age ranges associated with morphological changes, i.e. neurodevelopment in young subjects (e.g. infants) and neurodegeneration in older subjects (e.g. due to natural aging). This will be investigated in future work.

4.3 Sex Prediction

While lower whole-brain Jaccard distance is generally associated with common genetic traits, e.g. same-sex MRI pairs, however it is insufficient for predicting the sex of a brain MRI. We investigate whether the Jaccard distance can be

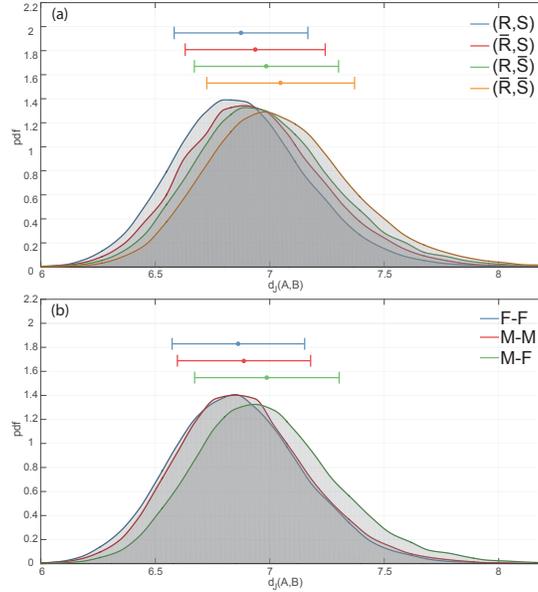


Fig. 4. (a) Distributions of Jaccard distances with SSE (NN=200) for unrelated subjects (UR) and: same race, same sex (blue); different race, same sex (red); same race, different sex (green), and different race, different sex (yellow). Based on a two-sample Kolmogorov-Smirnov test, $p < 1.34e^{-145}$ for all pairs of distributions. (b) Distributions of Jaccard distances with SSE (NN=200) for unrelated subjects (UR) from the same race between: Female-Female (blue), Male-Male (red), Male-Female of Female-Male (green). M-M and F-F distributions are highly similar, and both are significantly different from M-F ($p < 2.83e^{-270}$) based on a two-sample Kolmogorov-Smirnov test.

modified to predict the sex of a brain MRI, by evaluating the distances between a subject BoF A and BoFs compiled *all Males* $d_J(A, Males)$ and *all Females* $d_J(A, Females)$. These distances are combined in a basic linear classifier with a single threshold parameter τ to adjust for differences in the numbers of male and female subjects and features (12):

$$Class(A) = \begin{cases} Female & \text{if } d_J(A, Females) - d_J(A, Males) + \tau > 0 \\ Male & \text{otherwise.} \end{cases} \quad (12)$$

Figure 6 shows ROCs curves for sex classification obtained by varying τ over the range $[-\infty, \infty]$, again comparing Jaccard overlap computed via HSE and SSE. The AUC=0.97 for SSE is significantly higher than AUC=0.91 for HSE. SSE achieves an error accuracy rate of 91%, is slightly higher than 89% obtained using a voxel-wise support vector machine with single and multi-site MRI data reported in a recent work [14]. In comparison, our classification approach is on-the-fly and can be used to classify arbitrary subject groups with no explicit training procedure.

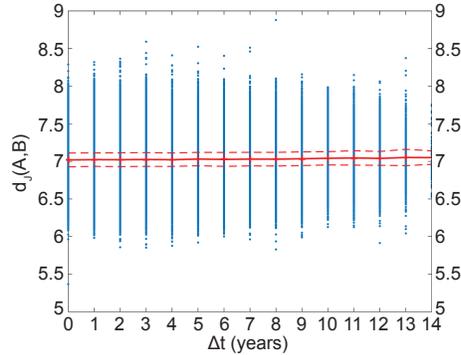


Fig. 5. Conditional distributions of Jaccard distance $p(d_J(A, B)|\Delta t)$ between unrelated subject pairs (A, B) given age difference Δt . The mean (plain red line) and standard deviation (dashed red lines) are plotted for each Δt .

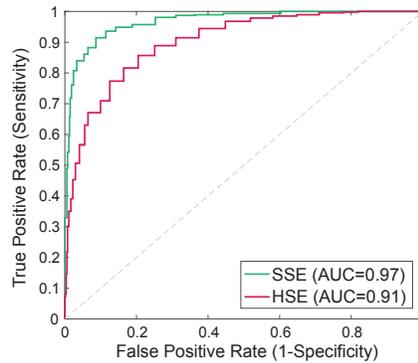


Fig. 6. ROC curves and AUC for sex classification based on Jaccard distances with HSE (red) and SSE (green) methods. AUC is noticeably higher for our SSE method.

5 Discussion

We presented a novel generalization of the Jaccard distance metric to consider probabilistic set equivalence, for the purpose of evaluating distance or dissimilarity between bag-of-feature image representation. We also develop a novel kernel density estimator based on feature geometry. The method is applied to recovering the proximity graph structure between MRI brain scans of 1010 genetically related subjects, including siblings and twins, where both contributions result in significant improvement in recall rates for all sibling types. To our knowledge, these are the highest recall rates for T1w brain scans present in the literature, note that monozygotic twins can be identified with virtually perfect accuracy. Furthermore, a minor modification allows the Jaccard distance to predict subject-specific traits such as sex with high accuracy.

Our method is based on highly efficient algorithms for feature extraction and correspondence, and can be applied to arbitrarily large datasets with minimal preprocessing. It should be noted that the low number of samples per label (i.e. family members) makes it difficult to use recent Convolutional Neural Networks (CNN) approaches, which generally require large numbers of training data. We have found it to be particularly effective at automatically identifying brain scans of the same subject, and we have applied the method to several large medical image datasets to identify previously unknown, incorrectly labeled duplicate subjects. Future work will investigate the method in other image modalities and alternative anatomies, e.g. full body scans. Generic 3D SIFT-Rank features are used here, tuning parameters such as patch size could potentially be optimized for brain MRI data. The relatively frequencies of correspondences across the brain may prove useful in understanding links between genotype and phenotypical variation, potentially in the context of genome wide heritability studies [17]. All code for this work is available upon request.

6 Acknowledgement

This work was supported by NIH grant P41EB015902 (NAC) and a Canadian National Sciences and Research Council (NSERC) Discovery Grant.

References

1. Annas, G.J., Elias, S.: 23andme and the fda. *New England Journal of Medicine* 370(11), 985–988 (2014)
2. Brosch, T., Tam, R., Initiative, A.D.N., et al.: Manifold learning of brain mris by deep learning. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. pp. 633–640. Springer (2013)
3. 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.: The heritability of multi-modal connectivity in human brain activity. *eLife* 6 (jul 2017)
4. Csurka, G., Dance, C., Fan, L., Willamowski, J., Bray, C.: Visual categorization with bags of keypoints. In: *Workshop on statistical learning in computer vision, ECCV*. vol. 1, pp. 1–2. Prague (2004)
5. Gardner, A., Kanno, J., Duncan, C.A., Selmic, R.: Measuring distance between unordered sets of different sizes. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 137–143 (2014)
6. Gerber, S., Tasdizen, T., Fletcher, P.T., Joshi, S., Whitaker, R., Initiative, A.D.N., et al.: Manifold modeling for brain population analysis. *Medical image analysis* 14(5), 643–653 (2010)
7. Hamburg, M.A., Collins, F.S.: The path to personalized medicine. *New England Journal of Medicine* 363(4), 301–304 (2010)
8. Kumar, K., Chauvin, L., Toews, M., Colliot, O., Desrosiers, C.: Multi-modal brain fingerprinting: a manifold approximation based framework. *bioRxiv* (2018)
9. Kumar, K., Desrosiers, C., Siddiqi, K., Colliot, O., Toews, M.: Fiberprint: A subject fingerprint based on sparse code pooling for white matter fiber analysis. *NeuroImage* 158, 242–259 (sep 2017)

10. Levandowsky, M., Winter, D.: Distance between Sets. *Nature* 234(5323), 34–35 (nov 1971)
11. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision* 60(2), 91–110 (nov 2004)
12. McCallum, A., Nigam, K., et al.: A comparison of event models for naive bayes text classification. In: *AAAI-98 workshop on learning for text categorization*. vol. 752, pp. 41–48. Citeseer (1998)
13. Muja, M., Lowe, D.G.: Scalable nearest neighbor algorithms for high dimensional data. *IEEE transactions on pattern analysis and machine intelligence* 36(11), 2227–2240 (2014)
14. Nieuwenhuis, M., Schnack, H.G., van Haren, N.E., Lappin, J., Morgan, C., Reinders, A.A., Gutierrez-Tordesillas, D., Roiz-Santiañez, R., Schaufelberger, M.S., Rosa, P.G., et al.: Multi-center mri prediction models: Predicting sex and illness course in first episode psychosis patients. *NeuroImage* 145, 246–253 (2017)
15. Ono, M., Kubik, S., Abernathy, C.: *Atlas of the cerebral sulci* (1990)
16. Ono, Y., Trulls, E., Fua, P., Yi, K.M.: Lf-net: Learning local features from images. *arXiv preprint arXiv:1805.09662* (2018)
17. Sabuncu, M.R., Ge, T., Holmes, A.J., Smoller, J.W., Buckner, R.L., Fischl, B., Weiner, M.W., Aisen, P., Weiner, M., Petersen, R., et al.: Morphometricity as a measure of the neuroanatomical signature of a trait. *Proceedings of the National Academy of Sciences* 113(39), E5749–E5756 (2016)
18. Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., Downey, P., Elliott, P., Green, J., Landray, M., Liu, B., Matthews, P., Ong, G., Pell, J., Silman, A., Young, A., Sprosen, T., Peakman, T., Collins, R.: UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. *PLOS Medicine* 12(3), e1001779 (mar 2015)
19. Terrell, G.R., Scott, D.W.: Variable kernel density estimation. *The Annals of Statistics* pp. 1236–1265 (1992)
20. Thompson, P.M., Cannon, T.D., Narr, K.L., van Erp, T., Poutanen, V.P., Huttenen, M., Lönqvist, J., Standertskjöld-Nordenstam, C.G., Kaprio, J., Khaledy, M., Dail, R., Zoumalan, C.I., Toga, A.W.: Genetic influences on brain structure. *Nature Neuroscience* 4(12), 1253–1258 (dec 2001)
21. Toews, M., Wells, W.M.: Efficient and robust model-to-image alignment using 3D scale-invariant features. *Medical Image Analysis* 17(3), 271–282 (2013)
22. Toews, M., Wells, W.M.: 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. IEEE (apr 2016)
23. Van Essen, D., Ugurbil, K., Auerbach, E., Barch, D., Behrens, T., Bucholz, R., Chang, A., Chen, L., Corbetta, M., Curtiss, S., Della Penna, S., Feinberg, D., Glasser, M., Harel, N., Heath, A., Larson-Prior, L., Marcus, D., Michalareas, G., Moeller, S., Oostenveld, R., Petersen, S., Prior, F., Schlaggar, B., Smith, S., Snyder, A., Xu, J., Yacoub, E.: The Human Connectome Project: A data acquisition perspective. *NeuroImage* 62(4), 2222–2231 (oct 2012)
24. Wachinger, C., Golland, P., Kremen, W., Fischl, B., Reuter, M.: BrainPrint: A discriminative characterization of brain morphology. *NeuroImage* 109, 232–248 (apr 2015)
25. Wachinger, C., Toews, M., Langs, G., Wells, W., Golland, P.: Keypoint transfer for fast whole-body segmentation. *IEEE transactions on medical imaging* (2018)