Low-Dimensional Non-Rigid Image Registration Using Statistical Deformation Models From Semi-Supervised Training Data

John A. Onofrey*, Xenophon Papademetris, and Lawrence H. Staib

***Abstract*—Accurate and robust image registration is a fundamental task in medical image analysis applications, and requires non-rigid transformations with a large number of degrees of freedom. Statistical deformation models (SDMs) attempt to learn the distribution of non-rigid deformations, and can be used both to reduce the transformation dimensionality and to constrain the registration process. However, high-dimensional SDMs are difficult to train given orders of magnitude fewer training samples. In this paper, we utilize both a small set of annotated imaging data and a large set of unlabeled data to effectively learn an SDM of non-rigid transformations in a semi-supervised training (SST) framework. We demonstrate results applying this framework towards inter-subject registration of skull-stripped, magnetic resonance (MR) brain images. Our approach makes use of 39 labeled MR datasets to create a set of supervised registrations, which we augment with a set of over 1200 unsupervised registrations using unlabeled MRIs. Through leave-one-out cross validation, we show that SST of a non-rigid SDM results in a robust registration algorithm with significantly improved accuracy compared to standard, intensity-based registration, and does so with a 99% reduction in transformation dimensionality.**

***Index Terms*—Brain, deformable registration, dimensionality reduction, medical image analysis.**

I. Introduction

NON-RIGID spatial normalization of different subjects to a common reference space has utility in a variety of medical imaging applications, for instance in population studies of functional brain imaging [1] and in atlas-based brain segmentation [2]. A large number of algorithms have been proposed for non-rigid image registration [3]. However, non-rigid image registration of different subjects is a challenging task made difficult, in part, by highly variable anatomical structure. Accurate anatomical alignment requires non-rigid transformations with a large number of degrees of freedom (DoFs). Inter-subject image registration of such high-dimensionality is further exacerbated by highly non-convex objective functions, which makes optimization ill-posed.

Statistical deformations models (SDMs) have the potential to reduce the dimensionality of the non-rigid transformations by learning the subspace or manifold in which these transformations exist in a population. SDMs attempt to analyze anatomical variation by modeling the non-rigid image deformations from a set of training registrations. SDMs have the potential for more robust registration results by constraining the deformation to be based on previously learned examples from a population. The SDM limits transformation estimates to be realizations from the distribution of deformations even in the presence of imaging noise or artifacts that could otherwise confound standard, unconstrained registration approaches. However, if the SDM is not trained using accurate registrations, the fidelity of the model is questionable. A straightforward method to construct an SDM uses a principal component analysis (PCA) of training deformations [4] to derive a low-dimensional set of deformation basis functions. The performance of PCA-based SDMs can also be limited due to the inability of the PCA analysis to model the covariance of the high-dimensional transformation with orders of magnitude fewer number of training samples; this limitation is an example of the high-dimension-low-sample-size problem. One proposal that addresses this scarcity of samples involves creating synthetic, but physically realistic, deformations of the training data [5].

Alternatively, large medical image databases containing vast amounts of neuroimaging data, such as the 1000 Functional Connectomes Project International Neuroimaging Data-Sharing Initiative [6] (FCP1000), offer an accessible and freely available source of SDM training samples. with which to overcome the limitations of low sample size. In addition, smaller databases like the UCLA Laboratory of Neuroimaging (LONI) Probabilistic Brain Atlas (LPBA40) [7] contain relatively few images, but incorporate richly-annotated information, such as gold-standard manually-segmented volumes of interest (VOIs). Datasets containing such rich information are labor intensive and expensive to create. In this paper, we leverage both the vast quantity of unlabeled anatomical magnetic resonance (MR) brain scans from the FCP1000 dataset and the small, but labeled brain images from the LPBA40 dataset to learn an SDM in a semi-supervised training (SST) framework. The concept of using both labeled and unlabeled training data originates from semi-supervised learning methods, which are widely applicable for classification problems [8], but its use for image registration is novel,

Manuscript received December 15, 2014; revised January 07, 2015; accepted February 02, 2015. Date of publication February 16, 2015; date of current version June 29, 2015. This work was supported by NIH/NIBIB R03 EB012969. *Asterisk indicates corresponding author.*

*J. A. Onofrey is with the Department of Diagnostic Radiology, Yale University, New Haven, CT 06520 USA (e-mail: john.onofrey@yale.edu).

X. Papademetris is with the Departments of Diagnostic Radiology and Biomedical Engineering, Yale University, New Haven, CT 06520 USA (e-mail: xenophon.papademetris@yale.edu).

L. H. Staib is with the Departments of Diagnostic Radiology, Electrical Engineering, and Biomedical Engineering, Yale University, New Haven, CT 06520 USA (e-mail: lawrence.staib@yale.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TMI.2015.2404572

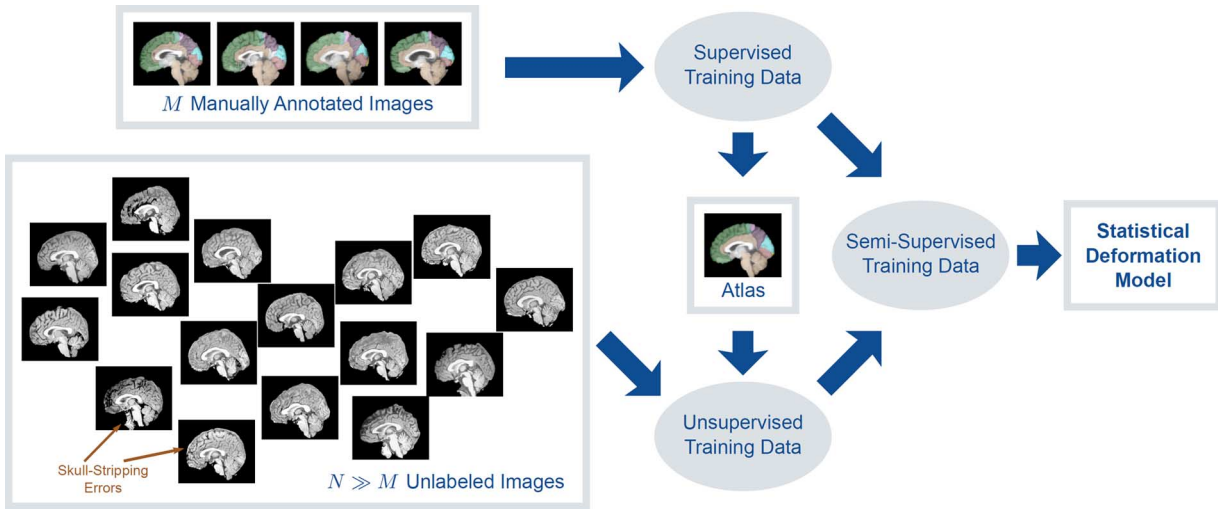

Fig. 1. We train our proposed statistical deformation model using semi-supervised training data that makes use of both a small number of gold-standard, manually-annotated image samples and a large number of unlabeled image samples. Our framework uses the labeled VOIs in the M annotated images to perform non-rigid registration in a supervised manner, from which we create a reference atlas composed of average anatomy and majority-vote label images. We then non-rigidly register the N unlabeled images to the atlas in an unsupervised manner using standard, intensity-only registration. Finally, we perform a principal component analysis of both the supervised and the unsupervised transformations to learn a statistical deformation model.

to the best of our knowledge. As shown in Fig. 1, we use the LPBA40 dataset's labeled MR images to create a set of supervised transformations and augment this supervised sample set with a second set of unsupervised transformations using more than 1200 unlabeled MRIs from the FCP1000. To the best of our knowledge, no prior work has made use of more than 1200 MRIs to learn an SDM for image registration. We demonstrate that SST of PCA-based SDMs is an effective and practical method to derive a set of basis functions for low-dimensional non-rigid registration.

This article expands upon our previous works [9], [10] and we structure the paper as follows: Section II presents an overview of work related to training and using SDMs for non-rigid registration. Section III details our proposed framework for SST of non-rigid SDMs for inter-subject anatomical brain registration. In Section IV, we show through leave-one-out cross validation that SST of our SDM results in a non-rigid registration algorithm with significantly improved accuracy compared to standard, intensity-based registration, and does so with a 99% reduction in transformation dimensionality and a significant computational improvement. Finally, Section V concludes with a discussion of our results and presents future research directions.

II. Background

A. Statistical Deformation Models

Parametric models of shape [11], [12] that statistically model the variation of shape within a particular class of objects were the precursor to SDMs. Cootes *et al.* [12] pioneered active shape models (ASMs) using a principal component analysis (PCA) of landmark point displacements. ASMs rely upon landmark points to describe the object's shape. However, manual identification of corresponding landmarks is expensive and automatic segmentation is itself a challenging task. Furthermore, ASMs do not easily admit modeling of more than one object. SDMs, on the other hand, avoid the explicit segmentation of the object of interest, and instead model how the object or objects deform according to a dense deformation field.

SDMs of dense deformation fields originated from Grenander and Miller's [13] work on computational anatomy. Their work statistically analyzed the dense deformations between corresponding anatomical structures across a sample population. Various approaches exist to statistically analyze the deformation fields. Joshi *et al.* [14] and Gee and Bajcsy [15] both used PCA to provide an orthonormal basis of dense non-rigid deformation, and Cootes *et al.* [16] proposed PCA of diffeomorphic transformations in the groupwise registration setting. Rather than performing PCA on the high-dimensional dense displacement field, Rueckert *et al.* [4] proposed constructing SDMs of the control points of non-rigid B-spline free-form deformations (FFDs) [17]. The FFD control points provide a lower-dimensional parameterization of the dense displacement field, and PCA of the FFD control points is attractive because it has the potential to drastically reduce the deformation dimensionality. Instead of using the deformation field, Vaillant *et al.* [18] performed PCA of geodesics of diffeomorphic flow landmark matching in computational anatomy. Similar to how PCA-based SDMs provide an orthonormal basis of deformations, Amit *et al.* [19] describe using low-frequency basis functions for describing low spatial frequency anatomical variability in the head. Examples of alternatives to PCA-based SDMs include Twining and Marsland's [20] statistical framework on the manifold of diffeomorphic transformations and Singh *et al.*'s [21] use of multivariate partial least squares regression of deformation covariance.

SDMs have found utility both as classifiers and as registration algorithms. While not the focus of this work, SDMs can classify, *i.e.* distinguish, subjects into different population groups based on their deformation characteristics with respect to a reference template image [14], [15], [22], [23], as done in tensor-based morphometry [24]. In contrast, SDMs can be used to drive the registration process itself. Rueckert *et al.* [4] proposed using the

SDM for registration by optimizing the SDM's linear PCA coefficients. While the authors did not implement the method themselves, later works followed this strategy for low-dimensional non-rigid registration. Examples of non-rigid PCA-based SDM registration include: Loeckx *et al.* [25] who registered chest radiographs using the FFD transformation model, Wouters *et al.* [26] who registered brain images using a viscous fluid transformation model, and Kim *et al.* [27] who registered brain images using HAMMER [28]. Instead of performing registration by optimizing the PCA coefficients, Xue *et al.* [29] used their PCA-based SDM as a prior for regularization of high-dimensional non-rigid registrations.

PCA-based SDMs suffer from the curse of dimensionality. In general, the number of training samples available is too small to fully span the high-dimensional space of the non-rigid transformations of an entire image domain. For this reason, SDM-based registration has been limited to small regions of interest within the image domain that focus upon a limited set of anatomy, *e.g.* the corpus callosum within the brain [4]. The PCA SDM's ability to reconstruct high-dimensional transformations is severely limited. Xue *et al.*'s [29] multi-scale wavelet decomposition attempted to address this problem by performing PCA independently at multiple scales.

Yet another recognized problem of SDMs is that they are not very meaningful unless they use accurate or gold-standard training registrations. In the previous works, the accuracy of the training registrations were inherently limited by the behavior of the registration algorithms, which themselves do not guarantee globally optimal results. In contrast, we train the SDM using a set of "gold-standard" deformations that are accurate with respect to annotated volumes of interest (VOIs) relevant to the image analysis task at hand.

B. Manifold Learning-Based Registration

In contrast to PCA, which is a linear method for dimensionality reduction, non-linear dimensionality reduction methods attempt to find a manifold on which the high-dimensional data resides. Within the medical imaging community, a number of authors adopt the Isomap algorithm [30] for learning a manifold of non-rigid image deformation [31]–[34]. These methods use the k-nearest neighbors (k-NN) of the training data to construct the manifold. Jia *et al.* [33] use image intensity differences to determine the nearest neighbors, while both Gerber *et al.* [31] and Hamm *et al.* [32] propose using diffeomorphic deformations as a measure of distance between image pairs. Rather than learning a manifold over the whole image domain as done by the previous methods, Ye *et al.* [34] learn local, regional manifolds within image sub-regions for non-rigid registration. Using the k-NN that define the manifold, manifold-based registration successively registers images that are most similar to each other along the manifold. However, one issue with manifolds using deformations as a distance measure is that non-rigid registration is used to evaluate the k-NN, and these transformations are not necessarily accurate unless they were created using gold-standard registrations. Furthermore, in contrast to linear PCA, it is difficult for manifolds to function as generative models, *e.g.* construct novel transformations on the space defined by the manifold that could be used to register unseen images. Our proposed method could potentially function as an initialization for non-linear manifolds.

C. Basis Functions for Image Registration

In contrast to registration algorithms that build a global non-rigid transformation from a set of local deformations, other authors propose using a set of global non-rigid basis functions to construct a transformation. Friston *et al.* [35] propose non-linear transformation models as a linear combination of displacement basis functions. In this manner, Ashburner and Friston [36] use the low-frequency components of the discrete cosine transform (DCT) basis functions and Christensen [37] uses 3D Fourier series as basis functions, where the number of basis coefficients parameterizes these transformation models. Generic basis functions, however, are not necessarily representative of the actual anatomical deformations found in a population. In this paper, we use PCA to derive a set of deformation basis functions from training data, tailoring them to the problem at hand.

D. Image Registration With Segmented Data

Image registration can potentially be made more accurate by registering salient features within the images rather than using the intensity information by itself. Some registration methods utilize similarity measures that combine both intensity-information and other features of interest. However, the task of automatically segmenting these features and determining their correspondences is itself a particularly challenging problem in medical image analysis. Goshtasby *et al.* [38] alternates between segmentation of regions of interest and registration using these regions. Yezzi *et al.* [39] use active contours that jointly segment and register images. Pohl *et al.* [40] jointly register images to an atlas while simultaneously segmenting the image structures according to the atlas using an Expectation Maximization algorithm.

Because segmentation is itself a challenging problem and the quality of these segmentations can greatly influence the registration quality, the aforementioned joint segmentation-registration methods cannot be considered foolproof. If, on the other hand, a dataset contains gold-standard segmentations that were manually labeled by an expert, registration algorithms could trust these a priori segmentations with impunity. In Section III-A, we introduce a framework that makes use of manually-segmented image data to perform accurate image registration in a supervised manner.

E. Semi-Supervised Learning

The semi-supervised learning (SSL) concept to use both labeled and unlabeled training data is widely applicable for classification problems [8]. Unlike the goal of classification seen in SSL problems, we make use of SSL as a method for inductive learning, in which we learn a model of registration deformations from a set of observed transformations. Extending the concepts of SSL to non-rigid SDM construction is a novel application of SSL, to the best of our knowledge. In our case, the terms "supervised" and "unsupervised" refer to construction of the training data. Our framework makes use of expertly annotated image data to generate *supervised registration* samples, while standard registration algorithms that do not use annotated

information fall into the category of generating *unsupervised registration* samples. In this work, neither the learning algorithm nor the non-rigid registration algorithm use semi-supervised learning techniques. For these reasons, a more appropriate nomenclature for our work is *semi-supervised training*.

Previously published approaches to SDM learning mentioned in Section II-A all fall under the category of unsupervised training. These methods train SDMs with registration estimates of unreliable veracity. However, in this paper, we propose a method of supervised SDM training by leveraging a set of expertly-annotated image data to create accurate registration training transformations. Regardless of training the SDM with labeled or unlabeled data, the aforementioned PCA-based SDMs offer relatively underwhelming registration performance likely due to PCA's inability to model the high-dimensional transformation's distribution with orders of magnitude fewer number of training samples. By including a large number of unsupervised deformation samples with our small set of supervised deformation samples, we propose a method for creating an SDM from semi-supervised training data to counteract the curse of dimensionality.

III. Methods

To train a statistical model of non-rigid deformation using semi-supervised training (SST), our framework requires both supervised and unsupervised deformation training samples as illustrated in Fig. 1. We use a B-spline free-form deformation (FFD) transformation model [17] to perform both types of registrations. The FFD's P B-spline control points displacements provide a low-dimensional parameterization of a dense, non-rigid deformation field $T(\mathbf{x}) \in \mathbb{R}^3$ at all voxels $\mathbf{x} \in \Omega \subset \mathbb{R}^3$, where Ω is the reference image volume in three dimensions.

Rewriting the FFD transformation as a column vector $\mathbf{d} \in \mathbb{R}^{3P}$ of concatenated control point displacements, we denote the set of *supervised* deformation samples $D_s = \{\mathbf{d}_m | m = 1, \ldots, M\}$ and the set of *unsupervised* deformation samples $D_u = \{\mathbf{d}_{M+n} | n = 1, \ldots, N\}$. To create training data in a supervised manner, described in Section III-A, we leverage an image dataset that contains manually annotated VOIs to enforce anatomical correspondence of those particular VOIs during the registration process. To create the unsupervised training data, described in Section III-B, we register images that lack annotated VOIs in an unsupervised manner using image intensity information alone. We register the $M+N$ training sample images to a common reference space with $181 \times 217 \times 181$ volume with $1\ \mathrm{mm}^3$ isotropic resolution (described in Section III-A). Within this volume, an FFD with 5 mm isotropic control point spacing requires $P = 60{,}236$ control points to parameterize a non-rigid deformation with 180,708 DoFs in 3D. In Section III-C, we use both sets D_s and D_u together as semi-supervised training data for constructing an SDM. We then describe how we subsequently use this SDM to non-rigidly register images not included in the training set.

A. Training Data From Supervised Non-Rigid Registration

To generate the set D_s of supervised non-rigid deformations, we leverage a database containing gold-standard, manually-segmented volumes of interest (VOIs) to constrain our deformations with respect to these VOIs [9]. The LPBA40 database [7] contains 40 anatomical, skull-stripped brain MR images I_i, $i = 1, \ldots, 40$, with 56 annotated VOIs (50 cortical structures, 4 subcortical volumes, cerebellum and brainstem). For each subject i, we generate surfaces $S_{i,j}$ for each of the $j = 1, \ldots, 56$ VOIs by extracting an iso-contour of each VOI using Marching Cubes [41]. We smoothed and then sub-sampled each VOI using standard image processing techniques in order to create lower resolution VOI surfaces to reduce the number of surface points, which in turn reduced computational burden. The total number of VOI surface points for each subject averaged across all subjects was 43,809 $\pm$ 3,080 points. Additionally, we use the LPBA40's labeled grey matter, white matter, and CSF tissue segmentations for each subject to normalize the intensities of each subject's anatomical image with respect to subject 1 by performing a least-squares fit of the mean intensity values within each tissue class. We use these normalized images I_i' only for training and do not use them during the actual registration tests performed in Section IV.

We use the MNI Colin 27 image volume as a common reference domain in which to perform our non-rigid registrations. We first register all 40 images to the MNI Colin 27 brain using an affine transformation model that optimizes a robust point matching (RPM) criterion [42] of the brain surfaces, which we generated from each subject's brain mask image included in the LPBA40 dataset. With all images transformed to MNI space, we then select subject 1 to be used as an initial reference template image I_1 for non-rigid registration of the remaining subjects. Using an integrated intensity and point-feature non-rigid registration algorithm [43] and the VOIs, we non-rigidly register the remaining $M = 39$ normalized image I_i' to our single-subject reference template image. This algorithm uses a FFD transformation model with 5 mm isotropic control point spacing and minimizes the sum of squared differences (SSD) similarity measure while penalizing misalignment of the VOI surface points according to the RPM metric [44], [45]. The previously mentioned intensity normalization step motivates our choice of SSD in this case. We denote these transformations $T_{i \rightsquigarrow 1}, i = 2, \ldots, 40$ ($i \rightsquigarrow j$ indicates non-linear registration from space i to j).

Our choice of reference image biases the registrations to that subject's particular anatomy. To remove some of this bias, we compute the mean transformation $\bar{T} = 1/M \sum_{i=2}^{40} T_{i \rightsquigarrow 1}$ and apply the inverse transformation $\bar{T}^{-1}$ to subject 1 to create a deformation bias-corrected (DBC) image $I_{\mathrm{DBC}} = \bar{T}^{-1} \circ I_1$, where $\circ$ is the transformation operator. We then re-register the M subjects to I_{DBC} using the same integrated non-rigid registration as before, and denote these transformations $T_{i \rightsquigarrow \mathrm{DBC}}$. As this set of transformations leverages annotated VOIs and we constrain our registration procedure to align these VOIs, the FFD control point displacements of these transformations constitute our set of supervised registrations $D_s = \{\mathbf{d}_m | m = 1, \ldots, M\}$.

In addition to creating D_s, we also create: *(i)* an atlas composed of an anatomical image $I_{\mathrm{Atlas}} = 1/(M-1) \sum_{i=2}^{40} T_{i \rightsquigarrow \mathrm{DBC}} \circ I_i'$, where I_i' are the normalized images; and *(ii)* a corresponding VOI label image found using a majority-vote of the 39 transformed subject VOIs. I_{Atlas} is a representative template image

for the set of non-rigid FFD transformations with 5 mm control point spacing. We thus use I_{Atlas} as the template reference image for our unsupervised registrations described in the following section.

B. Training Data From Unsupervised Non-Rigid Registration

We use a large number of unsupervised training samples to supplement our small set of supervised deformations. We select $N = 1247$ unlabeled, anatomical brain MR images of healthy, normal subjects from the FCP1000 database [6]. First, we register all images to our common reference space, I_{Atlas}, using an affine transformation $T_{n\rightarrow\text{Atlas}}, n = 1, \ldots, N$ ($i \rightarrow j$ indicates linear registration from space i to j) by maximizing the normalized mutual information (NMI) similarity metric [46]. Following affine registration, we then estimate the non-rigid deformation of all N brains to I_{Atlas}.

Using the same FFD non-rigid transformation model with 5 mm isotropic control point spacing as in Section III-A, we non-rigidly register each subject by maximizing the NMI. However, unlike the expertly skull-stripped LPBA40 images in Section III-A, the FCP1000 images were not manually skull-stripped, and thus artifacts, *e.g.* the optic nerve, remain in some images (as can be seen in Fig. 1). Such artifacts present challenges for intensity-based registration methods. Rather than manually correcting these poor skull-strippings, we perform weighted non-rigid registration [47] that preferentially weights the brain region during the calculation of NMI. For this, we create a brain weight mask image by first dilating I_{Atlas}'s brain mask and then smoothing with a Gaussian kernel. The resulting transformations $T_{n\rightsquigarrow\text{Atlas}}$ and their respective FFD control point displacements comprise our set of unsupervised deformations $D_u = \{\mathbf{d}_{M+n} | n = 1, \ldots, N\}$.

C. Statistical Deformation Model Registration

Using both the supervised and unsupervised training sets, D_s and D_u, respectively, we create a statistical deformation model (SDM) of non-rigid FFD transformations [4]. A principal component analysis (PCA) of the deformations $\mathbf{d}_i, i = 1, \ldots, M + N$ gives a linear approximation of the deformation distribution

$$\mathbf{d} = \bar{\mathbf{d}} + \mathbf{\Phi}\mathbf{w} \tag{1}$$

where $\bar{\mathbf{d}} = 1/M + N \sum_{i=1}^{M+N} \mathbf{d}_i$ is the mean deformation of the $M + N$ training registrations, $\mathbf{\Phi} = (\phi_1 | \ldots | \phi_K) \in \mathbb{R}^{3P\times K}$ is the matrix of orthogonal eigenvectors, and $\mathbf{w} \in \mathbb{R}^K$ is a vector of model coefficients. The number of training samples and the transformation's dimensionality determine the number of eigenvectors such that $K = \min\{M + N, 3P\}$, where typically $M + N \ll 3P$. The k-th eigenvalue λ_k estimates the sample variance along the eigenvector ϕ_k, and we sort them in decreasing order $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_K$. An SDM using K_v parameters accounts for $0 \leq v \leq 100\%$ of the model's cumulative variance $\sum_{k=1}^{K_v} \lambda_k \leq v \sum_{k=1}^{K} \lambda_k$. Thus, using the first K_v coefficients of $\mathbf{w}$ in (1) provides a low-dimensional parameterization (K_v DoFs) of a high-dimensional FFD $\mathbf{d}$, which we denote $T_{\text{SDM}_v}(\mathbf{x}; \mathbf{w})$ for all points $\mathbf{x}$ in the reference image domain Ω.

To non-rigidly register a new image I to our reference image I_{Atlas}, we use the SDM from (1) to optimize the objective function

$$\hat{T}_{\text{SDM}_v} = \arg\max_{\mathbf{w}} J(I_{\text{Atlas}}, T_{\text{SDM}_v}(\cdot; \mathbf{w}) \circ I). \tag{2}$$

Here, J is the NMI similarity metric evaluated throughout the image volume. We solve (2) using conjugate gradient optimization with a hierarchical multi-resolution image pyramid. Unlike other PCA-based SDM registration approaches [4], we do not bound the SDM transformation to have weights within α standard deviations of the mean value, $|w_k| \leq \alpha\sqrt{\lambda_k}$, with $\alpha = 3$ a typical choice, as this can exclude some valid outlier deformations, especially at the finer image scales.

IV. Results

We tested how our semi-supervised training (SST) framework impacts PCA-based statistical deformation model (SDM) construction by performing a series of *leave-one-out* tests. For each supervised deformation $\mathbf{d}_i \in D_s$, we recomputed the SDM in Section III-C by leaving the i-th deformation out of (1). We did not recompute I_{Atlas} as in Section III-A for each leave-one-out test because the single subject's exclusion had negligible effect on the mean intensity image. First, Section IV-A demonstrates how our proposed supervised registration method compares to unsupervised, standard, intensity-based non-rigid registration with respect to creating our atlas I_{Atlas} during training. Section IV-B presents results showing the SDM's ability to reconstruct each of the $i = 1, \ldots, 39$ deformations. Section IV-C then shows how well the SDM registered $\mathbf{d}_i$'s corresponding anatomical image I_i in the LPBA40 dataset to I_{Atlas} using (2). For comparison, we created another SDM using standard, unsupervised-only deformation training samples. We trained this SDM with $M + N$ unsupervised training samples by replacing the LPBA40 supervised registrations in D_s with unsupervised registrations of the same images using our methodology from Section III-B. We also compared our results to standard, intensity-only FFD registration using 5 mm isotropic control point spacing, which we denote FFD_5.

To evaluate registration accuracy, we calculated the mean Dice overlap (MDO) of the $R = 56$ VOIs with respect to the atlas labels [48], [49]

$$\text{MDO}_i = \frac{1}{R} \sum_{r\in R} \frac{2|(T \circ I_i)_r \cap (I_{\text{Atlas}})_r|}{|(T \circ I_i)_r| + |(I_{\text{Atlas}})_r|}$$

where $(A)_r$ denotes the r-th VOI from image A, and T a transformation. We found MDO to be a suitable summary statistic to quantify registration performance rather than analyze each individual VOI Dice overlap separately. Furthermore, since our supervised training framework constrained VOI boundaries to align and a certain amount of registration uncertainty remains at locations away from those boundaries, our use of Dice overlap more appropriately measures accurate alignment of VOI boundaries than a residual sum of square transformation error.

We implemented our code on the GPU using the CUDA parallel programming platform as part of BioImage Suite [50]. The

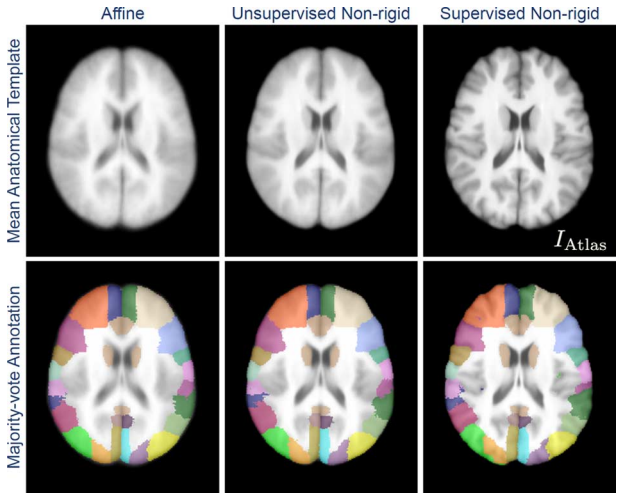

Fig. 2. Mean intensity images of 39 subjects comparing our proposed supervised registration approach using non-rigid intensity/point feature registration (right) compared to standard non-rigid intensity-only (middle) and affine registration methods (left). We overlay the majority-vote atlas labels on top of their corresponding anatomical template images. We used the anatomical template image created by our supervised training approach as the reference image I_{Atlas} for all of our subsequent non-rigid registrations.

eigensystem in (1) may be pre-computed ahead of the registration, and can then be loaded into the algorithm to avoid unnecessarily repetitious PCA computation. We compared the computation time of our approach to standard FFD_5 implemented in BioImage Suite. We ran all algorithms on a workstation featuring an NVIDIA GeForce GTX Titan GPU.

A. Atlas Construction Using Supervised Deformations

We evaluated our supervised deformations by assessing the quality of our atlas I_{Atlas} in comparison to atlases created using standard (unsupervised) intensity-only affine and non-rigid FFD deformations. We created the affine and intensity FFD atlases using the same method described in Section III-A. For affine registration, we used the LPBA40 images transformed to MNI space in Section III-A, and did not require a deformation bias-correction step. For standard, non-rigid FFD, we used the same 5 mm isotropic control point spacing used to create the supervised deformation samples and registered the images to subject 1 by maximizing their NMI, and subsequently re-registered the images after deformation bias-correction.

We qualitatively visualized group anatomical alignment of the 39 subjects using the mean intensity images I_{Atlas}. Fig. 2 shows the I_{Atlas} anatomical template images and majority-vote label images computed using each of the three registration methods. The atlas created using our proposed supervised approach appears much sharper than both atlases generated using standard, intensity-only FFD and affine registrations, which indicates better anatomical correspondences within the group after registration.

We quantitatively assessed registration quality by computing the MDO for all 39 subjects with respect to each registration method's unique deformation bias-corrected reference image I_{DBC}. In this case, we computed MDO with respect to I_{DBC} instead of using I_{Atlas} to provide a more uniform comparison of overlap because the atlas generation process creates

TABLE I
MEAN DICE OVERLAP (MDO) OF THE 39 SUBJECTS IN THE TRAINING SET AFTER OUR PROPOSED SUPERVISED NON-RIGID TRAINING, STANDARD UNSUPERVISED NON-RIGID AND AFFINE TRAINING. REPORTED VALUES ARE Mean $\pm$ std WITH RESPECT TO OUR DEFORMATION BIAS-CORRECTED REFERENCE TEMPLATE IMAGE (SUBJECT 1). ALL RESULTS ARE STATISTICALLY SIGNIFICANT WITH RESPECT TO THE SUPERVISED APPROACH

Method	MDO	Min MDO
Supervised Non-rigid	$\mathbf{75.35 \pm 1.39}$	**72.02**
Unsupervised Non-rigid	71.31 ± 1.58	67.71
Affine	63.36 ± 2.39	56.34

strikingly different atlases with different majority-vote VOI labels for each method, as seen in Fig. 2. As shown in Table I, the atlas generated using supervised deformations aligned the training data significantly better than the unsupervised training methods (two-tailed paired t-test, $p \leq 10^{-25}$). Furthermore, the supervised method had higher worst-case performance (minimum MDO 72.02) than the unsupervised method's mean performance (MDO 67.71), which suggests better VOI alignment across the group.

We also evaluated MDO of the supervised method with respect to I_{Atlas} to measure how well the atlas explained its constituent training data. With respect to I_{Atlas}, we observed MDO 81.76 ± 1.11 (mean and standard deviation) and minimum MDO 78.79 across all subjects. These values represent an upper bound for registration performance using our atlas as these transformations make use of the annotated VOIs to explicitly enforce correspondence of these VOIs in a supervised manner, which are not otherwise available in the more typical unsupervised registration setting that does not have VOI segmentations available.

B. SDM Reconstructions

For each leave-one-out test, we tested the SDM's ability to reconstruct the i-th subject's deformation $\mathbf{d}_i$ by rewriting (1) and solving

$$\hat{\mathbf{d}}_i = \bar{\mathbf{d}}_i + \mathbf{\Phi}_i \mathbf{\Phi}_i^T (\mathbf{d}_i - \bar{\mathbf{d}}_i) \tag{3}$$

where $\hat{\mathbf{d}}_i$ is a least-squares approximation to $\mathbf{d}_i$, and we calculated $\bar{\mathbf{d}}_i$ and $\mathbf{\Phi}_i$ using $\mathbf{d}_j, \forall j \neq i$. We resliced image I_i using $\hat{\mathbf{d}}_i$ and computed the MDO_i. Fig. 3 compares the reconstruction performance of SDMs using different numbers of supervised and unsupervised training samples, $M = 38, 0 \leq N \leq 1247$. We selected the N samples sequentially from D_u, without randomization. Fig. 3 also shows how the SDMs performed when trained using only supervised samples, *i.e.*, $M = 38$, $N = 0$. As to be expected, reconstruction performance increased with the number of samples. However, the inclusion of only a small number of supervised registration samples, $M = 38$ in the case of leave-one-out testing, significantly increased the reconstructive capabilities of the SDM. This observed increase in SDM reconstruction performance is the result of the supervised training samples increasing the PCA model space's variance.

We also experimented using half the number of supervised training deformations, $M = 19$, to see how differences in M affected the SDM. For each leave-one-out test, we randomly

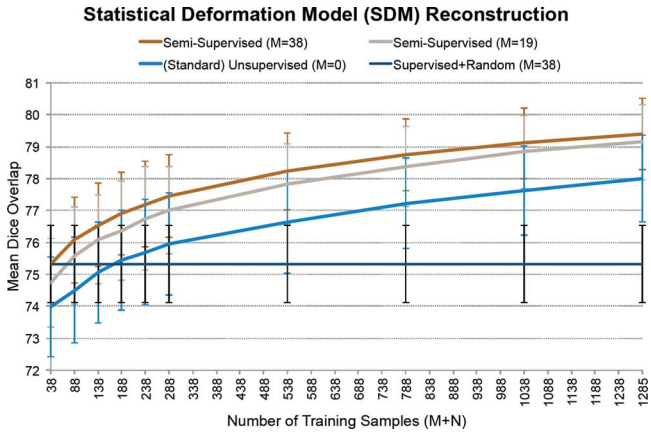

Fig. 3. Mean dice overlap values for leave-one-out transformation reconstructions as a function of the number of M supervised and N unsupervised training samples. The inclusion of a few, $M = 19, 38$, supervised deformation samples significantly increased the SDM's reconstruction abilities compared to using only standard, unsupervised deformation samples, $M = 0$. These results represent a performance ceiling for SDM registration performance. To verify that the unsupervised training samples did have useful correlations for SDM training, we created N synthetic samples of i.i.d. Gaussian noise added to the $M = 38$ supervised training samples (supervised + random training set). For all random training samples $N > 0$, the uncorrelated noise provided negligible changes to the SDM's reconstructive capabilities. The plotted values are the mean with error bars of one standard deviation.

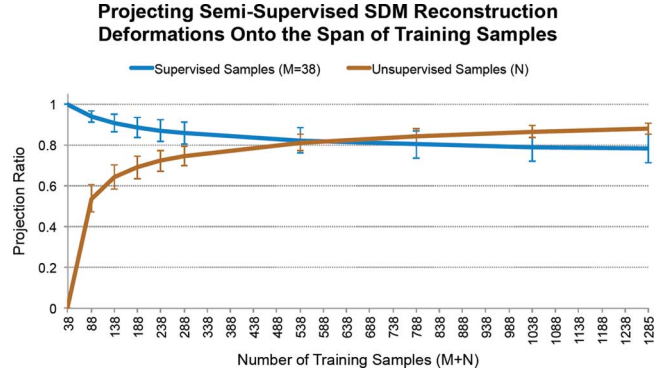

Fig. 4. For an SDM using semi-supervised training, we plot the fraction of the reconstructed deformation $\hat{d}$ that lies within the span of *(i)* $M = 38$ supervised training samples and *(ii)* N unsupervised training samples. Plotted values are means and standard deviations.

selected 19 samples from D_s, and replaced the remaining, unused samples with their corresponding unsupervised deformation. We then selected the remaining unsupervised N samples from D_u as before. Fig. 3 shows that removing half the supervised samples did not result in performance halfway to that found using $M = 0$, and that only a small decrease in SDM reconstruction occurs with large N. We also noted that the MDO curves in Fig. 3 appeared to be continuing to increase with the inclusion of additional training samples.

These reconstruction experiments provided a performance ceiling $(\text{MDO} = 79.40 \pm 1.11)$ for *our SDM's registration capability* as the deformations $\hat{\mathbf{d}}_i$ from (3) were found without the confounds introduced by optimizing an objective function as in (2). Here, $\hat{\mathbf{d}}_i$ represent a least-squares fit transformation to the corresponding supervised deformation $\mathbf{d}_i \in D_s$. With this performance ceiling in mind, Section IV-C presents how well our proposed SDM effectively registered images in practice.

To further demonstrate the utility of the unsupervised registration samples and validate their inclusion into our semi-supervised training framework, we projected the leave-one-out reconstruction results onto both the span of the supervised training samples S_s and the span of the unsupervised training samples S_u. Here, we computed the spans using the Gram-Schmidt process to orthonormalize the respective sets of training samples. In order to see what fraction of the reconstructed deformation $\hat{\mathbf{d}}$ lies on the span of S, we projected $\hat{\mathbf{d}}$ from (3) onto S and compared the magnitude of this projection to the unprojected magnitude $\|\hat{\mathbf{d}}\|$ by calculating the ratio $\|\text{proj}_S(\hat{\mathbf{d}})\|/\|\hat{\mathbf{d}}\|$. Intuitively, if this projection ratio is close to 1 then $\hat{\mathbf{d}}$ lies almost entirely within the space S spanned by those training samples. As seen in Fig. 4, for $N <\sim 550$, the reconstructed deformations fractionally lie more within the space spanned by the supervised training samples, while for $N >\sim 550$, the reconstructions fractionally lie more within the space spanned by the unsupervised training samples. Furthermore, the span of supervised training samples continues to provide a fractionally high contribution to the reconstructed deformation even with a large number of unsupervised training samples.

For a second experiment to verify that the unsupervised training samples did have correlations useful to the construction of the SDM, we created a synthetic set of unsupervised deformations by adding noise to the supervised training samples. We created each of the N synthetic deformation samples by first randomly selecting one of the $M = 38$ supervised deformations from D_s and then adding standard normal i.i.d. Gaussian noise to each component. Including these $M + N$ synthetic samples into the SDM training and recomputing the reconstruction of each leave-one-out sample did not alter the PCA model's performance for any number of added samples $N > 0$ (Fig. 3).

We then used this SDM from supervised training data and synthetic random samples to repeat the reconstruction projection experiments from above by computing the projection ratio of $\|\hat{\mathbf{d}}\|$ onto the span of supervised samples S_s. In this case, the reconstructed deformation lies almost entirely in the span of supervised training samples as the projection ratio using S_s is 1.0 ± 0.0 regardless of the number of synthetic random samples N added. The N random training samples did not significantly change the space of the deformations modeled by the PCA. These results help to show the utility of the unsupervised training samples within our semi-supervised training framework, especially with a large number of included samples.

C. Non-Rigid SDM Registration

Having shown improvements in SDM reconstruction capabilities, we now demonstrate that the SDM from semi-supervised training (SST) samples also works in practice to non-rigidly register images. For each of the 39 leave-one-out test cases, we compared registration using 5 methods: *(i)* SDM from SST $M = 38$, $N = 1247$ training samples; *(ii)* SDM from unsupervised $M = 0$, $N = 1285$ training samples; *(iii)* SDM from supervised $M = 38$, $N = 0$ training samples; *(iv)* SDM from supervised $M = 38$ and $N = 1247$ synthetic random samples

TABLE II
NUMBER OF EIGENVECTORS K_v ACCOUNTING FOR THE FIRST v PERCENT OF SDM CUMULATIVE VARIANCE. M IS THE NUMBER OF SUPERVISED TRAINING SAMPLES AND N IS THE NUMBER OF UNSUPERVISED TRAINING SAMPLES

	SDM Training Method			
	Semi-Supervised	Unsupervised	Supervised	Supervised+Random
v (%)	($M = 38, N = 1247$)	($M = 0, N = 1285$)	($M = 38, N = 0$)	($M = 38, N = 1247$)
25	5	4	-	-
33	-	-	1	-
50	19	19	2	1
75	76	77	11	7
90	228	236	23	16
95	390	405	29	23
99	800	817	35	35
100	1285	1285	38	1285

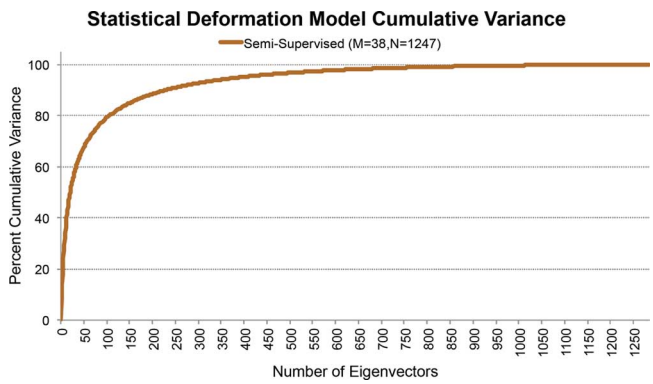

Fig. 5. SDM cumulative variance as function of the number of eigenvectors for SDMs using semi-supervised training with $M = 38$ supervised samples and $N = 1247$ unsupervised samples.

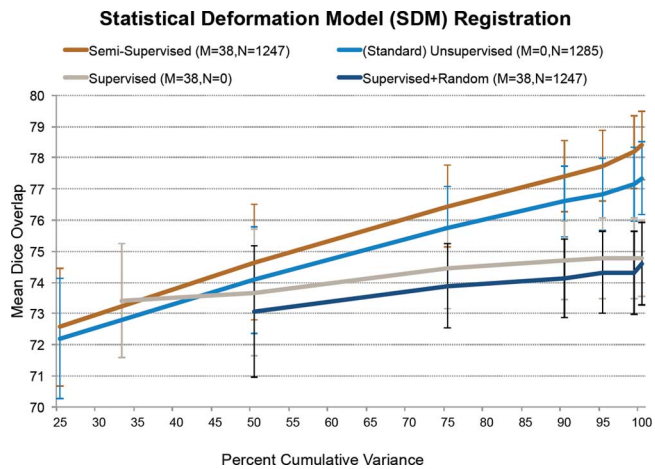

Fig. 6. We evaluate non-rigid registration performance of our SDM registration algorithm using mean dice overlap as a function of SDM cumulative variance. We plot results for SDMs trained using: *(i)* semi-supervised training with $M = 38$ supervised samples and $N = 1247$ unsupervised samples; *(ii)*$N = 1285$ unsupervised training samples; *(iii)* $M = 38$ supervised training samples; and *(iv)* $M = 38$ supervised samples and $N = 1247$ synthetic random samples.

adding i.i.d. Gaussian noise to the supervised samples; and *(v)* standard, intensity FFD_5. For each of the SDM's, we registered the original (unnormalized) image I_i using the first K_v eigenvectors that contained the first $v = 25, 50, 75, 90, 95, 99,$ and 100 percent of cumulative variance and computed MDO_i. For example, Fig. 5 plots the cumulative variance of the SDM using SST. Table II shows the corresponding number of eigenvectors K_v for each percentage of cumulative variance v for each SDM training method. In the cases of SDM training with supervised-only samples and training with supervised and random samples, the first eigenvector accounted for $v = 33\%$ and 50% of the cumulative variance, respectively. The number of eigenvectors defined the dimensionality of the non-rigid deformation.

Fig. 6 shows MDO as a function of v for each of the four SDM types. SDM from SST significantly outperformed SDM from unsupervised training for all values of v (two-tailed paired t-test $p \leq 10^{-7}$). The SDM from supervised training performed worse than both SST and unsupervised training, with the exception of using $v = 0.33$, *i.e.* a single eigenvector. These results were to be expected given the small number of training samples, $M = 38$, and the large percentage of cumulative variance modeled by the first eigenvector. The SDM with supervised and random samples underperformed the SDM using supervised-only training samples for $v < 100\%$, while using all $K_{100\%} = 1285$ eigenvectors produced similar registration results (two-tailed paired t-test $p > 0.05$). These results suggest that the noise in the synthetic random samples adversely affected how our PCA modeled the space of deformations, which in turn led to poor registration performance.

However, we observed that the SDM from SST was unable to fully achieve the reconstructive performance ceiling of the SDM (see Section IV-B). Table III highlights the results of our tests using $v = 100\%$ for each of the registration methods. In terms of MDO, our SDM from SST performed significantly better than FFD_5, and SDMs from unsupervised training and supervised training (two-tailed paired t-test $p < 0.005$). Most impressively, the SDM achieved more accurate registration than FFD_5 while reducing the dimensionality of the non-rigid deformation from 180,708 to 1285, a 99% reduction in DoFs. While we did observe a 63% reduction in mean computation time due to the reduction in DoFs, the reduction was not proportional to the reduction in DoFs. Unlike standard B-spline FFD transformation models that use local displacements of control points, the PCA-based deformation is global in nature, which means

TABLE III

LEAVE-ONE-OUT REGISTRATION RESULTS OF AN SDM FROM SEMI-SUPERVISED TRAINING (SST) USING 100% OF THE VARIANCE, *I.E.* ALL PCA EIGENVECTORS, COMPARED WITH SDMS USING SUPERVISED-ONLY, UNSUPERVISED-ONLY, AND SUPERVISED PLUS RANDOM UNSUPERVISED TRAINING SAMPLES, AS WELL AS STANDARD, INTENSITY FFD REGISTRATION USING 5 MM CONTROL POINT SPACING. M AND N ARE THE NUMBERS OF SUPERVISED AND UNSUPERVISED TRAINING SAMPLES, RESPECTIVELY. MEAN DICE OVERLAP (MDO) VALUES SHOWN ARE Mean $\pm$ std AND MINIMUM FOR 39 SUBJECTS. ALL RESULTS ARE SIGNIFICANTLY DIFFERENT (TWO-TAILED PAIRED T-TEST $p < 0.05$) COMPARED TO SST SDM REGISTRATION

Method	DoFs	MDO	Min MDO	Compute Time (min)
SDM using semi-supervised training ($M = 38, N = 1247$)	1285	$\mathbf{78.40 \pm 1.11}$	**76.07**	27.53 ± 7.56
SDM using unsupervised training ($M = 0, N = 1285$)	1285	77.33 ± 1.17	74.73	27.02 ± 7.88
SDM using supervised training ($M = 38, N = 0$)	38	74.77 ± 1.23	70.83	0.55 ± 0.07
SDM using supervised+random training ($M = 38, N = 1247$)	1285	74.60 ± 1.33	70.92	2.20 ± 0.34
Intensity FFD_5	180,708	78.09 ± 1.26	74.08	73.85 ± 1.39
Affine	12	72.18 ± 2.30	64.11	-

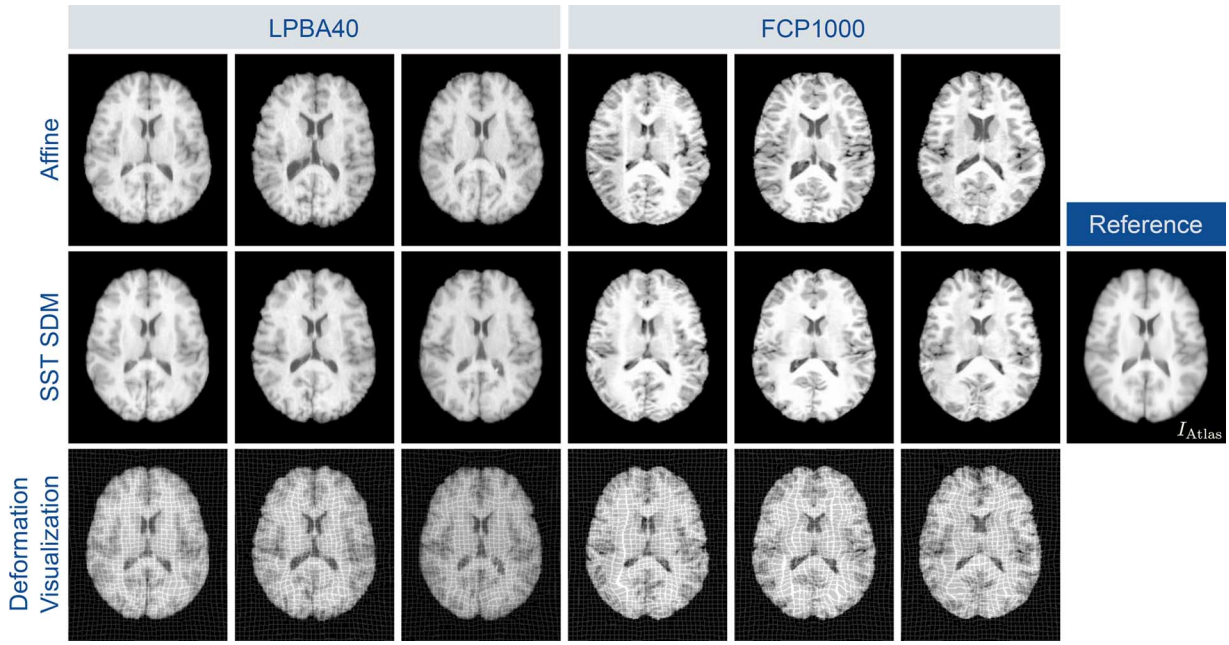

Fig. 7. Example subjects and their respective registrations during leave-one-out tests using our statistical deformation model (SDM) trained using our proposed semi-supervised training (SST). We show example images drawn from both the LPBA40 and FCP1000 training datasets. The first row shows the original images after affine registration to our atlas anatomical template image. The second row shows the same images after SDM registration using 1285 eigenvectors. The third row overlays a 5 mm isotropic grid to visualize the non-rigid deformation.

that changing a single weight parameter results in a global deformation change. This global change means that the image warping procedure used during the optimization of (2) must be performed over the whole image domain, which is computationally slow.

Of particular interest is SST SDM's worst case performance, with minimum MDO of 76.07, being higher than FFD_5's minimum MDO of 74.08. By constraining the registration transformation to the learned space of deformations, the SDM prevented the registration from becoming stuck in worse local minima. Fig. 7 illustrates SST SDM registration results for some example images drawn from both the LPBA40 and the FCP1000 training data.

V. DISCUSSION AND CONCLUSION

We demonstrate the utility of training non-rigid statistical deformation models (SDMs) with both a small set of accurate, supervised deformations and a large set of registrations of unknown quality from a large-scale medical image database. Using an SDM trained in a semi-supervised manner, we show significantly improved registration performance with a 99% reduction in registration DoFs compared to standard registration methods. Constraining a registration's deformation to be from the space of learned deformations in a data-driven manner adds robustness, with our SDM using our proposed semi-supervised training (SST) exhibiting better worst-case performance.

While our use of PCA to learn our SDM is simple, our results show that it is effective given a large enough training set. Both the SDM reconstruction results and the SDM registration results in Sections IV-B and IV-C, respectively, demonstrate that our proposed SST framework provides significantly superior performance compared to standard SDM training using unsupervised training data. Our results indicate some limitations of our approach: *(i)* the SDM reconstructions did not attain our hypothesized upper bound for registration performance found by the supervised registration framework in Section IV-A; and *(ii)* our realized SDM registration performance did not attain the SDM's full reconstructive capabilities. To address the former, Fig. 3 indicates that incorporating a larger set of supervised training samples, greater than our $M = 38$, has the potential to

provide a boost to SDM reconstruction performance; however, our results using $M = 19$ show that this gain is most likely non-linear. Yet another way to improve SDM training might be to incorporate subjects from a more heterogeneous population. For example, our supervised training set contains healthy normal subjects within a relatively narrow age range of 19–40 years [7]. Alternatively, another possible direction is to explore multi-resolution, hierarchical statistical shape models [51] as a way to better model deformations using a limited number of training samples. As for the latter, we require further investigation as to why our realized registration performance lags the SDM's reconstruction capabilities. We hypothesize that potential non-convexity of the objective function prevents our algorithm from attaining the SDM's reconstruction performance bound.

The SDM construction may also benefit from a less biased atlas image. While we did attempt to reduce some of the bias introduced from our use of a single subject as a reference image in Section III-C, use of a groupwise registration registration scheme [52] could further reduce this bias. Ideally, the mean deformation in (1) should be $\bar{\mathbf{d}} = \mathbf{0}$. An unbiased atlas closer to the mean should require less deformation from each of the training images, which has the potential to improve the SDM's capabilities to model correlated deformations.

SDMs might find most utility in the way they were used in this work as an adjunct to the affine registration step, where affine registration provides the initial, global, linear registration, then the SDM registration provides problem-specific, constrained non-rigid registration, and finally, standard non-rigid methods finish the registration process. Such a framework [9], [53] could avoid many of the local minima that confound unconstrained (unsupervised) deformable registration algorithms.

While our objective function in (2) does not bias the deformations towards the mean deformation, it may be of interest to implement the algorithm in a Bayesian framework and compare registration performance and robustness. A limitation of our SDM transformations is that they are not necessarily diffeomorphic. Non-diffeomorphic solutions are possible, particularly for larger weights on the deformation components. In our experience, however, this effect is negligible. We only observed negative Jacobian determinants in 10 of the 39 test cases. In those cases, the percentage of negative voxels within the brain was very small, 0.06%, and the average value was -1.4×10^{-5}. A possible solution could add a soft constraint to the registration objective function penalizing the non-diffeomorphic Jacobians as done for standard FFDs [54].

In addition, we aim to explore methods for non-linear dimensionality reduction, and compare these results to that of linear PCA. We also plan to further validate our algorithm using other databases that contain annotated brain images and to test how the SDM generalizes across a more heterogeneous population. However, the results presented in this paper show that, while straightforward, PCA is indeed effective at modeling the high-dimensional space of non-rigid brain deformations, and that the limiting factors for a PCA-based SDM is the number of training samples used and the quality of those samples. Furthermore, the use of the PCA SDM provides robustness to the image registration process with better worst-case performance than standard, unconstrained registration approaches. Both our theoretical and actual SDM registration results suggest that the space of non-rigid deformations between subjects is of surprisingly low dimensionality.

REFERENCES

[1] A. Gholipour, N. Kehtarnavaz, R. Briggs, M. Devous, and K. Gopinath, "Brain functional localization: A survey of image registration techniques," *IEEE Trans. Med. Imag.*, vol. 26, no. 4, pp. 427–451, Apr. 2007.

[2] P. Aljabar, R. Heckemann, A. Hammers, J. Hajnal, and D. Rueckert, "Multi-atlas based segmentation of brain images: Atlas selection and its effect on accuracy," *NeuroImage*, vol. 46, no. 3, pp. 726–738, 2009.

[3] A. Sotiras, C. Davatzikos, and N. Paragios, "Deformable medical image registration: A survey," *IEEE Trans. Med. Imag.*, vol. 32, no. 7, pp. 1153–1190, Jul. 2013.

[4] D. Rueckert, A. Frangi, and J. Schnabel, "Automatic construction of 3-D statistical deformation models of the brain using nonrigid registration," *IEEE Trans. Med. Imag.*, vol. 22, no. 8, pp. 1014–1025, Aug. 2003.

[5] G. Hamarneh, P. Jassi, and L. Tang, "Simulation of ground-truth validation data via physically- and statistically-based warps," in *Medical Image Computing and Comput.-Assisted Intervention—MICCAI 2008*, ser. Lecture Notes in Computer Science, D. Metaxas, L. Axel, G. Fichtinger, and G. Székely, Eds. Berlin, Germany: Springer, 2008, vol. 5241, pp. 459–467.

[6] B. B. Biswal, "Toward discovery science of human brain function," *Proc. Nat. Acad. Sci.*, vol. 107, no. 10, pp. 4734–4739, 2010.

[7] D. W. Shattuck *et al.*, "Construction of a 3D probabilistic atlas of human cortical structures," *NeuroImage*, vol. 39, no. 3, pp. 1064–1080, 2008.

[8] O. Chapelle, B. Schölkopf, and A. Zien, *Semi-Supervised Learning*. Cambridge, MA: MIT Press, 2006.

[9] J. A. Onofrey, L. H. Staib, and X. Papademetris, "Fast nonrigid image registration using statistical deformation models learned from richly-annotated data," in *Proc. IEEE 10th Int. Symp. Biomed. Imag.*, 2013, pp. 580–583.

[10] J. A. Onofrey, L. H. Staib, and X. Papademetris, "Semi-supervised learning of nonrigid deformations for image registration," in *Medical Comput. Vision. Large Data in Medical Imagingser*, ser. Lecture Notes in Computer Science, B. Menze, G. Langs, A. Montillo, M. Kelm, H. Müller, and Z. Tu, Eds. New York: Springer , 2014, pp. 13–23.

[11] L. Staib and J. Duncan, "Parametrically deformable contour models," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 1989, pp. 98–103.

[12] T. Cootes, C. Taylor, D. Cooper, and J. Graham, "Active shape models-their training and application," *Comput. Vis. Image Understand.*, vol. 61, no. 1, pp. 38–59, 1995.

[13] U. Grenander and M. Miller, "Computational anatomy: An emerging discipline," *Q. Appl. Math.*, vol. 56, no. 4, pp. 617–694, 1998.

[14] S. C. Joshi, M. I. Miller, and U. Grenander, "On the geometry and shape of brain sub-manifolds," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 11, no. 08, pp. 1317–1343, 1997.

[15] J. C. Gee and R. K. Bajcsy, "Elastic matching: Continuum mechanical and probabilistic analysis," in *Brain Warping*, A. W. Toga, Ed. New York: Academic, 1999, pp. 183–197.

[16] T. Cootes, S. Marsland, C. Twining, K. Smith, and C. Taylor, "Groupwise diffeomorphic non-rigid registration for automatic model building," in *Comput. Vision—ECCV 2004*, ser. Lecture Notes in Computer Science, T. Pajdla and J. Matas, Eds. Berlin, Germany: Springer, 2004, vol. 3024, pp. 316–327.

[17] D. Rueckert *et al.*, "Nonrigid registration using free-form deformations: Application to breast MR images," *IEEE Trans. Med. Imag.*, vol. 18, no. 8, pp. 712–721, Aug. 1999.

[18] M. Vaillant, M. Miller, L. Younes, and A. Trouv, "Statistics on diffeomorphisms via tangent space representations," *NeuroImage*, vol. 23, no. 0, pp. S161–S169, 2004.

[19] Y. Amit, U. Grenander, and M. Piccioni, "Structural image restoration through deformable templates," *J. Am. Stat. Assoc.*, vol. 86, no. 414, pp. 376–387, 1991.

[20] C. J. Twining and S. Marsland, "Constructing an atlas for the diffeomorphism group of a compact manifold with boundary, with application to the analysis of image registrations," *J. Comput. Appl. Math.*, vol. 222, no. 2, pp. 411–428, 2008.

[21] N. Singh, P. Fletcher, J. Preston, L. Ha, R. King, J. Marron, M. Wiener, and S. Joshi, "Multivariate statistical analysis of deformation momenta relating anatomical shape to neuropsychological measures," in *Medical Image Computing and Comput.-Assisted Intervention—MICCAI 2010*, ser. Lecture Notes in Computer Science, T. Jiang, N. Navab, J. P. Pluim, and M. A. Viergever, Eds. Berlin, Germany: Springer, 2010, vol. 6363, pp. 529–537.

[22] J. G. Csernansky *et al.*, "Hippocampal morphometry in schizophrenia by high dimensional brain mapping," *Proc. Nat. Acad. Sci.*, vol. 95, no. 19, pp. 11406–11411, 1998.

[23] C. Davatzikos, M. Vaillant, S. M. Resnick, J. L. Prince, S. Letovsky, and R. N. Bryan, "A computerized approach for morphological analysis of the corpus callosum," *J. Comput. Assist. Tomogr.*, vol. 20, no. 1, 1996.

[24] J. Ashburner *et al.*, "Identifying global anatomical differences: Deformation-based morphometry," *Human Brain Mapp.*, vol. 6, no. 5-6, pp. 348–357, 1998.

[25] D. Loeckx, F. Maes, D. Vandermeulen, and P. Suetens, "Temporal subtraction of thorax CR images using a statistical deformation model," *IEEE Trans. Med. Imag.*, vol. 22, no. 11, pp. 1490–1504, Nov. 2003.

[26] J. Wouters, E. D'Agostino, F. Maes, D. Vandermeulen, and P. Suetens, "Non-rigid brain image registration using a statistical deformation model," in *Proc. SPIE*, J. Reinhardt and J. Pluim, Eds., 2006, vol. 6144, pp. 614411–8.

[27] M.-J. Kim, M.-H. Kim, and D. Shen, "Learning-based deformation estimation for fast non-rigid registration," in *IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Workshops*, Jun. 2008, pp. 1–6.

[28] D. Shen and C. Davatzikos, "HAMMER: Hierarchical attribute matching mechanism for elastic registration," *IEEE Trans. Med. Imag.*, vol. 21, no. 11, pp. 1421–1439, Nov. 2002.

[29] Z. Xue, D. Shen, and C. Davatzikos, "Statistical representation of high-dimensional deformation fields with application to statistically constrained 3D warping," *Med. Image Anal.*, vol. 10, no. 5, pp. 740–751, 2006.

[30] J. B. Tenenbaum, V. d. Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science*, vol. 290, no. 5500, pp. 2319–2323, 2000.

[31] S. Gerber, T. Tasdizen, P. T. Fletcher, S. Joshi, and R. Whitaker, "Manifold modeling for brain population analysis," *Med. Image Anal.*, vol. 14, no. 5, pp. 643–653, 2010.

[32] J. Hamm, D. H. Ye, R. Verma, and C. Davatzikos, "Gram: A framework for geodesic registration on anatomical manifolds," *Med. Image Anal.*, vol. 14, no. 5, pp. 633–642, 2010.

[33] H. Jia, G. Wu, Q. Wang, and D. Shen, "Absorb: Atlas building by self-organized registration and bundling," *NeuroImage*, vol. 51, no. 3, pp. 1057–1070, 2010.

[34] D. Ye, J. Hamm, D. Kwon, C. Davatzikos, and K. M. Pohl, "Regional manifold learning for deformable registration of brain MR images," in *Medical Image Computing and Comput.-Assisted Intervention—MICCAI 2012*, ser. Lecture Notes in Computer Science, N. Ayache, H. Delingette, P. Golland, and K. Mori, Eds. Berlin, Germany: Springer, 2012, vol. 7512, pp. 131–138.

[35] K. J. Friston *et al.*, "Spatial registration and normalization of images," *Human Brain Map.*, vol. 3, no. 3, pp. 165–189, 1995.

[36] J. Ashburner and K. J. Friston, "Nonlinear spatial normalization using basis functions," *Human Brain Map.*, vol. 7, no. 4, pp. 254–266, 1999.

[37] G. Christensen, "Consistent linear-elastic transformations for image matching," in *Inform. Processing in Medical Imaging*, ser. Lecture Notes in Computer Science, A. Kuba, M. Amal, and A. Todd-Pokropek, Eds. Berlin, Germany: Springer, 1999, vol. 1613, pp. 224–237.

[38] A. Goshtasby, G. C. Stockman, and C. V. Page, "A region-based approach to digital image registration with subpixel accuracy," *IEEE Trans. Geosci. Remote Sens.*, vol. 24, no. 3, pp. 390–399, 1986.

[39] A. Yezzi, L. Zöllei, and T. Kapur, "A variational framework for joint segmentation and registration," in *IEEE Workshop Math. Methods Biomed. Image Anal.*, 2001, pp. 44–51.

[40] K. M. Pohl, J. Fisher, W. E. L. Grimson, R. Kikinis, and W. M. Wells, "A Bayesian model for joint segmentation and registration," *NeuroImage*, vol. 31, no. 1, pp. 228–239, 2006.

[41] W. E. Lorensen and H. E. Cline, "Marching cubes: A high resolution 3D surface construction algorithm," *SIGGRAPH Comput. Graph.*, vol. 21, no. 4, pp. 163–169, Aug. 1987.

[42] A. Rangarajan *et al.*, "A robust point-matching algorithm for autoradiograph alignment," *Med. Image Anal.*, vol. 1, no. 4, pp. 379–398, 1997.

[43] X. Papademetris, A. P. Jackowski, R. T. Schultz, L. H. Staib, and J. S. Duncan, "Integrated intensity and point-feature nonrigid registration," in *Medical Image Computing and Comput.-Assisted Intervention—MICCAI 2004*, ser. Lecture Notes in Computer Science, C. Barillot, D. R. Haynor, and P. Hellier, Eds. Berlin, Germany: Springer, 2004, vol. 3216, pp. 763–770.

[44] H. Chui and A. Rangarajan, "A new point matching algorithm for non-rigid registration," *Comput. Vis. Image Understand.*, vol. 89, no. 2-3, pp. 114–141, 2003.

[45] X. Papademetris, A. P. Jackowski, R. T. Schultz, L. H. Staib, and J. S. Duncan, "Computing 3d non-rigid brain registration using extended robust point matching for composite multisubject fmri analysis," in *Medical Image Computing and Comput.-Assisted Intervention—MICCAI 2003*, ser. Lecture Notes in Computer Science, R. E. Ellis and T. M. Peters, Eds. Berlin, Germany: Springer, 2003, vol. 2879, pp. 788–795.

[46] C. Studholme, D. Hill, and D. Hawkes, "An overlap invariant entropy measure of 3D medical image alignment," *Pattern Recognition*, vol. 32, no. 1, pp. 71–86, 1999.

[47] J. Suh *et al.*, "Serial nonrigid vascular registration using weighted normalized mutual information," in *IEEE Int. Symp. Biomed. Imag., From Nano to Macro*, 2010, pp. 25–28.

[48] A. Klein *et al.*, "Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration," *NeuroImage*, vol. 46, no. 3, pp. 786–802, 2009.

[49] T. Rohlfing, "Image similarity and tissue overlaps as surrogates for image registration accuracy: Widely used but unreliable," *IEEE Trans. Med. Imag.*, vol. 31, no. 2, pp. 153–163, Feb. 2012.

[50] A. Joshi *et al.*, "Unified framework for development, deployment and robust testing of neuroimaging algorithms," *Neuroinformatics*, vol. 9, pp. 69–84, 2011.

[51] J. Cerrolaza, A. Villanueva, and R. Cabeza, "Hierarchical statistical shape models of multiobject anatomical structures: Application to brain MRI," *IEEE Trans. Med. Imag.*, vol. 31, no. 3, pp. 713–724, Mar. 2012.

[52] S. K. Balci, P. Golland, M. Shenton, and W. M. Wells, "Free-form b-spline deformation model for groupwise registration," in *Medical Image Computing and Comput.-Assisted Intervention: MICCAI . . . Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, 2007, vol. 10, pp. 23–30.

[53] Z. Xue and D. Shen, "Statistically-constrained deformable registration of MR brain images," in *Proc. 4th IEEE Int. Symp. Biomed. Imag., From Nano to Macro*, Apr. 2007, pp. 25–28.

[54] D. Rueckert, P. Aljabar, R. Heckemann, J. Hajnal, and A. Hammers, "Diffeomorphic registration using B-splines," in *Medical Image Computing and Comput.-Assisted Intervention—MICCAI 2006*, ser. Lecture Notes in Computer Science, R. Larsen, M. Nielsen, and J. Sporring, Eds. Berlin, Germany: Springer, 2006, vol. 4191, pp. 702–709.